Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-07/0-4946-2		

4. Title and Subtitle	5. Report Date
DEVELOPMENT OF A PROTOTYPE DYNAMIC CONGESTION AND INCIDENT PREDICTION SYSTEM	October 2006 Published: March 2007
	6. Performing Organization Code
7. Author(s) Kevin Balke, Praprut Songchitruksa, Nadeem Chaudhary, Chi-Lueng Chu, Sangita Sunkari, Paul Nelson, Shamanth Kuchangi, Vipin Tyagi, and Dvahg Swaroop	8. Performing Organization Report No. Report 0-4946-2
9. Performing Organization Name and Address Texas Transportation Institute The Texas A&M University System College Station, Texas 77843-3135	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. Project 0-4946
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, Texas 78763-5080	13. Type of Report and Period Covered Technical Report: September 2005-August 2006
	14. Sponsoring Agency Code

15. Supplementary Notes

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.
Project Title: Dynamic Traffic Flow Modeling for Incident Detection and Short-Term Congestion Prediction
URL: http://tti.tamu.edu/documents/0-4946-2.pdf

16. Abstract

This report is a summary of the processes and procedures used to develop the Dynamic Congestion and Incident Prediction System (DCIPS). The DCIPS is a prototype tool designed to illustrate to the Texas Department of Transportation (TxDOT) the potential of having a tool to predict when and where traffic and environmental conditions might lead to the formation of incidents and congestion. In addition to the introduction, this report contains five major sections. The "System Design and Architecture" section describes the major components, architecture, and data flows associated with developing the DCIPS prototype tool. The "User's Guide" section provides instructions and procedures for installing, operating, and interpreting the results of the prediction models contained in the DCIPS. The "Proof-of-Concept Testing" section describes how researchers used hardware-in-the-loop simulation to test the functionality and operation of the DCIPS prototype tool. The "Issues Affecting Implementation" section documents some of the data quality issues researchers encountered as they attempted to deploy the prototype system. The "Summary and Lessons Learned" section highlights some of the major findings from this research project as well as next steps for future research activities.

17. Key Words	18. Distribution Statement
Incident Prediction, Traffic Modeling, Short-Term Congestion Prediction, Traffic Forecast	No restrictions. This document is available to the public through NTIS: National Technical Information Service Springfield, Virginia 22161 http://www.ntis.gov

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	130	

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

DEVELOPMENT OF A PROTOTYPE DYNAMIC CONGESTION AND INCIDENT PREDICTION SYSTEM

by

Kevin Balke, Ph.D., P.E.
Center Director, TransLink®
Texas Transportation Institute

Praprut Songchitruksa, Ph.D.
Assistant Research Scientist
Texas Transportation Institute

Nadeem Chaudhary, Ph.D., P.E.
Senior Research Engineer
Texas Transportation Institute

Chi-Lueng Chu
Ph.D. Candidate
Department of Computer Science
Texas A&M University

Sangita Sunkari
Programmer/Analyst II
Texas Transportation Institute

Paul Nelson, Ph.D.
Research Associate
Texas Engineering Experiment Station

Shamanth Kuchangi
Graduate Research Assistant
Department of Civil Engineering
Texas A&M University

Vipin Tyagi
Ph.D. Candidate
Department of Mechanical Engineering
Texas A&M University

and

Dvahg Swaroop, Ph.D.
Associate Professor
Department of Mechanical Engineering
Texas A&M University

Report 0-4946-2
Project 0-4946
Project Title: Dynamic Traffic Flow Modeling for Incident Detection and Short-Term Congestion Prediction

Performed in cooperation with the
Texas Department of Transportation
and the
Federal Highway Administration

October 2006
Published: March 2007

TEXAS TRANSPORTATION INSTITUTE
The Texas A&M University System
College Station, Texas 77843-3135

DISCLAIMER

This research was performed in cooperation with the Texas Department of Transportation (TxDOT) and the Federal Highway Administration (FHWA). The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the FHWA or TxDOT. This report does not constitute a standard, specification, or regulation.

This report is not intended for construction, bidding, or permit purposes. The engineer in charge of the project was Kevin N. Balke, P.E. #66529.

The United States Government and the State of Texas do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

ACKNOWLEDGMENTS

This project was conducted in cooperation with TxDOT and FHWA. The authors would like to express their appreciation to Mr. Al Kosik, P.E., and Mr. Fabian Kalapach, P.E., with the TxDOT Traffic Operations Division for serving as project coordinator and project director for this project. The authors would also like to acknowledge Mr. Brian Burk, P.E. (TxDOT Austin District), Mr. Ron Fuessel (TxDOT Traffic Operations Division), and Mr. Henry Wickes, P.E. (TxDOT Traffic Operations Division), for serving as project advisors on this project. Without their insight, knowledge, and assistance, the authors would not have been able to complete this project. Furthermore, the authors would also like to acknowledge Mr. Wade Odell, P.E., and Ms. Sandra Kaderka of the TxDOT Research and Technology Implementation Office for their assistance in administering this research project.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

INTRODUCTION

BACKGROUND

Historically, freeway traffic management software has been designed to allow operators to react to incidents and congestion after they have already occurred. While reacting to unexpected events will always remain a critical part of freeway operations, freeway operators need to proactively manage traffic on the freeway to minimize the impact of events or even possibly prevent them from occurring in the first place. The purpose of this research project was to combine several existing research disciplines (such as traffic flow prediction and modeling) with real-time measurements of freeway performance to develop a tool that the Texas Department of Transportation (TxDOT) can use to proactively manage traffic operations on the freeway. As part of this research, we examined techniques and technologies that TxDOT can use, in conjunction with real-time and archived loop detector data, to forecast if and when traffic conditions are likely to produce incidents. We also examined methods TxDOT can use to predict traffic flow parameters (such as speed, volume, and occupancy) 1, 5, 10, and 15 minutes into the future. We developed and tested a prototype tool that TxDOT can implement in their control centers that will combine these two models to predict potential problem locations in real time, before they occur.

The overall goal of this research project was to produce a tool that TxDOT can implement in freeway management centers that will allow them to use traffic detector information currently being generated in their freeway management systems to make real-time, short-term predictions of when and where incidents and congestion are likely to occur on the freeway network. The idea was to combine roadway network modeling, traffic flow simulation, statistical regression and prediction methodologies, and archived and real-time traffic sensor information to forecast when and where: 1) traffic conditions will exist that are likely to produce an incident and 2) platoons of traffic will merge together to create congestion on the freeway. To accomplish this goal, we identified four objectives as part of this research effort:

1. Develop a methodology for identifying and predicting when and where incidents are likely to occur on the freeway system by comparing traffic detector data from around known incident conditions.

2. Develop a model to predict traffic flow parameters 15 to 30 minutes into the future based on current and historical traffic flow conditions.
3. Develop a prototype tool that can be implemented by TxDOT in their freeway management centers that combines the ability to predict potential incident conditions and short-term congestion.
4. Conduct a demonstration of the prototype tool.

We conducted this research over a two-year period. In the first year, we focused on the first two objectives: developing the traffic and incident prediction models. In the second year of this project, we focused on the development of a prototype tool. This report documents the development of the prototype tool.

SUMMARY OF YEAR 1 ACTIVITIES

In the first year of the research project, we focused on identifying techniques and strategies for predicting and forecasting traffic and incident conditions using real-time traffic sensor and other information. In conducting this research, the research team split into several independent groups, each focusing on different aspects of the problem. One group focused on using weather and traffic flow conditions as predictors of incident conditions. Other groups focused on developing models for producing short-term forecasts of potential congestion, using current measured traffic conditions. These research activities became the foundation for the work we performed in the second year of this project. The results of these activities are summarized in Report 0-4946-1, *Dynamic Traffic Flow Modeling for Incident Detection and Short-Term Congestion Prediction: Year 1 Progress Report (1).*

ORGANIZATION OF REPORT

This report is a summary of the processes and procedures used to develop the Dynamic Congestion and Incident Prediction System (DCIPS). The DCIPS is a prototype tool designed to illustrate to TxDOT the potential of having a tool to predict when and where traffic and environmental conditions might lead to the formation of incidents and congestion. In addition to this section, this report contains five major sections. The "System Design and Architecture" section describes the major components, architecture, and data flows associated with developing the DCIPS prototype tool. The "User's Guide" section provides instructions and procedures for installing, operating, and interpreting the results of the prediction models contained in the DCIPS. The "Proof-of-Concept Testing" section describes how we used hardware-in-the-loop simulation to test the functionality and operation of the DCIPS prototype tool. The "Issues Affecting Implementation" section documents some of the data quality issues we encountered as we attempted to deploy the prototype system. The "Summary and Lessons Learned" section highlights some of the major findings from this research project as well as next steps for future research activities. At the end of the report, we have included four appendices that describe the specifics of the traffic condition and incident prediction model developed as part of this research.

SYSTEM DESIGN AND ARCHITECTURE

The Texas Transportation Institute's (TTI) Dynamic Congestion and Incident Prediction System is a prototype tool that has been designed and developed for TxDOT to be implemented in freeway management centers. The DCIPS allows TxDOT to use traffic detector information currently being generated in their freeway management systems to make real-time, short-term predictions of when and where incidents and congestion are likely to occur on the freeway network.

The system combines roadway network modeling, statistical regression and prediction methodologies, and real-time traffic sensor information to forecast when and where: 1) traffic conditions will exist that are likely to produce an incident and 2) platoons of traffic will merge together to create congestion on the freeway. To accomplish this, TTI researchers have incorporated four prediction and forecasting models:

- *Cumulative Flow Forecast Model*—This model forecasts the flow-in-process level of service and delay-in-flow level of service over a span of 15 minutes for all the stations in the designated target section of the freeway.
- *Speed-Density Forecast Model*—This model forecasts speed and density for a given station in the designated target section of the freeway.
- *Incident Prediction Model*—This model uses weather and traffic flow conditions as predictors of incident conditions. It forecasts the incident probability, collision probability, and hazard score for a given station in the designated target section of the freeway.
- *Crash Potential Model*—This model forecasts the crash potential for all the stations in the designated target section of the freeway.

Figure 1 shows the high-level system architecture of the DCIPS. Each of the major components of the DCIPS is discussed below.

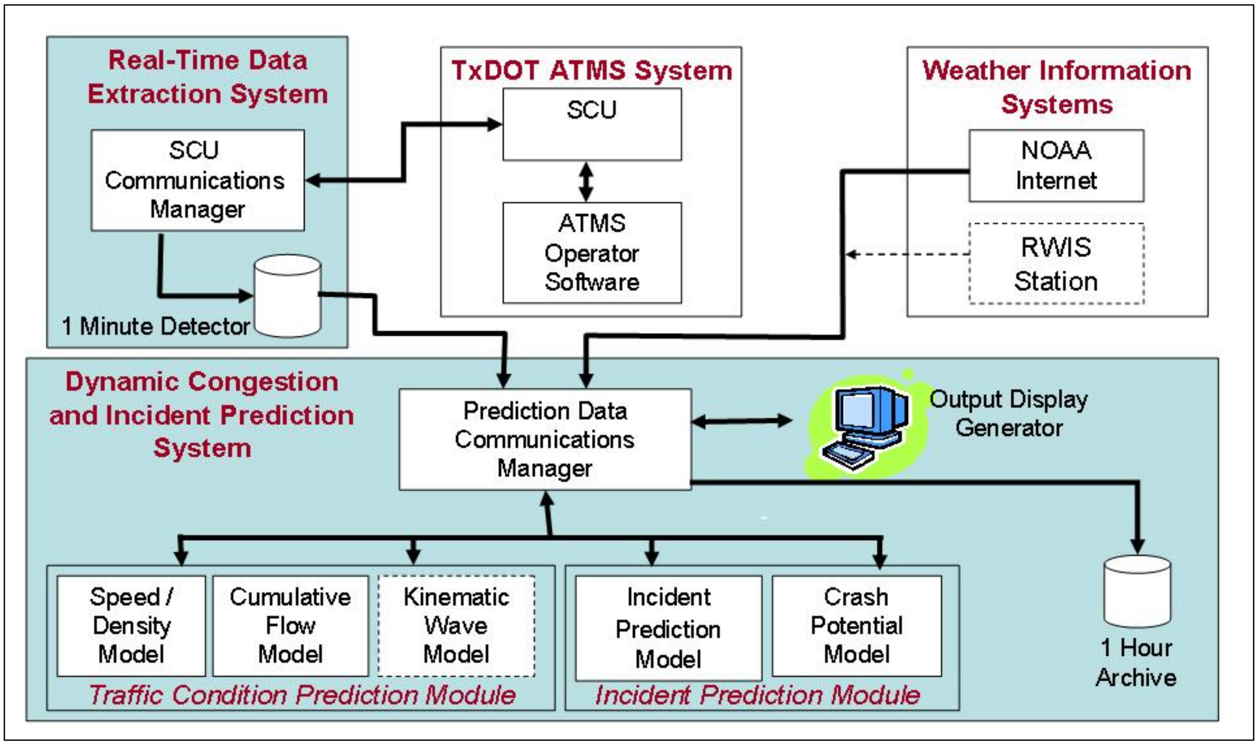


Figure 1. High-Level System Architecture Diagram of the Dynamic Congestion and Incident Prediction System.

SYSTEM INPUT DATA

The DCIPS was designed to operate in real-time using a live data source obtained from the field. The system uses two types of operational data: traffic operations data obtained from the TxDOT System Control Unit (SCU) and weather data from the National Oceanographic and Atmospheric Administration (NOAA) or from Road Weather Information System (RWIS). Below is a discussion of the data inputs into the DCIPS.

Traffic Operations Data

The DCIPS obtains its traffic operations data from TxDOT's SCU. The SCU collects traffic operations data from loop detectors installed on the freeway. The traffic operations data are updated once every minute. The following are types of traffic operations data used by the DCIPS:

- *Volume*—This value represents a count of the number of vehicles passing over the detector station in a minute.

- *Occupancy*—This value represents the percent of time that a detector indicated the presence of a vehicle during that minute.
- *Speed*—This value represents average speed of a vehicle passing through a trap-loop detection station during that minute.
- *Percent Trucks*—This value represents the percent of vehicles crossing a trap detector exceeding a user-defined length threshold. If the measured vehicle length exceeds the user-defined length threshold, the vehicle is counted as a truck. At the end of the 1-minute interval, the number of trucks is converted to the percent of trucks.

For the purposes of this project, we assumed that each lane on the freeway as well as each lane on exit and entrance ramps was its own detection station and that detector data were *not* aggregated across lanes.

Weather Data

The system was also designed to take two sources of weather data: information directly from Road Weather Information Systems that might be installed adjacent to the freeway as well as real-time weather data from the National Oceanographic and Atmospheric Administration weather stations. The system was designed to take weather information directly from a NOAA weather website, like that shown in Figure 2. This website contains current readings of the wind speed and direction, visibility and sky conditions, current temperature, dew point, relative humidity, and barometric pressures.

The system also uses the latitude and longitude of the city to estimate sunrise and sunset times. Sunrise and sunset were determined to be significant factors in predicting incident likelihood from the detector data (*1*).

SYSTEM COMPONENTS

Real-Time Data Extraction Subsystem

While not a component of the DCIPS, the Real-Time Data Extraction System (RTDES) is an integral part of the Dynamic Congestion and Incident Prediction System. The Real-Time Data Extraction System consists of two components: the System Control Unit Communications Manager and the 1-Minute Detector Data database. The SCU Communications Manager is a software program that extracts the 1-minute volume data from the SCU and places them in a "storage bin" from which other applications can extract the loop detector data. The SCU Communications Manager requests an update of the detector information in the SCU at a rate of 10 updates per second. Upon receiving the information from the SCU, the SCU Communications Manager compares the most recently extracted information to information stored in a database to determine if the recent information represents new detector information. If the SCU Communications Manager determines that the data have been updated, it then stores the most recent data in the 1-Minute Detector Data database. If the detector data have not been updated, the SCU Communications Manager continues requesting new data from the SCU until a new update has been received.

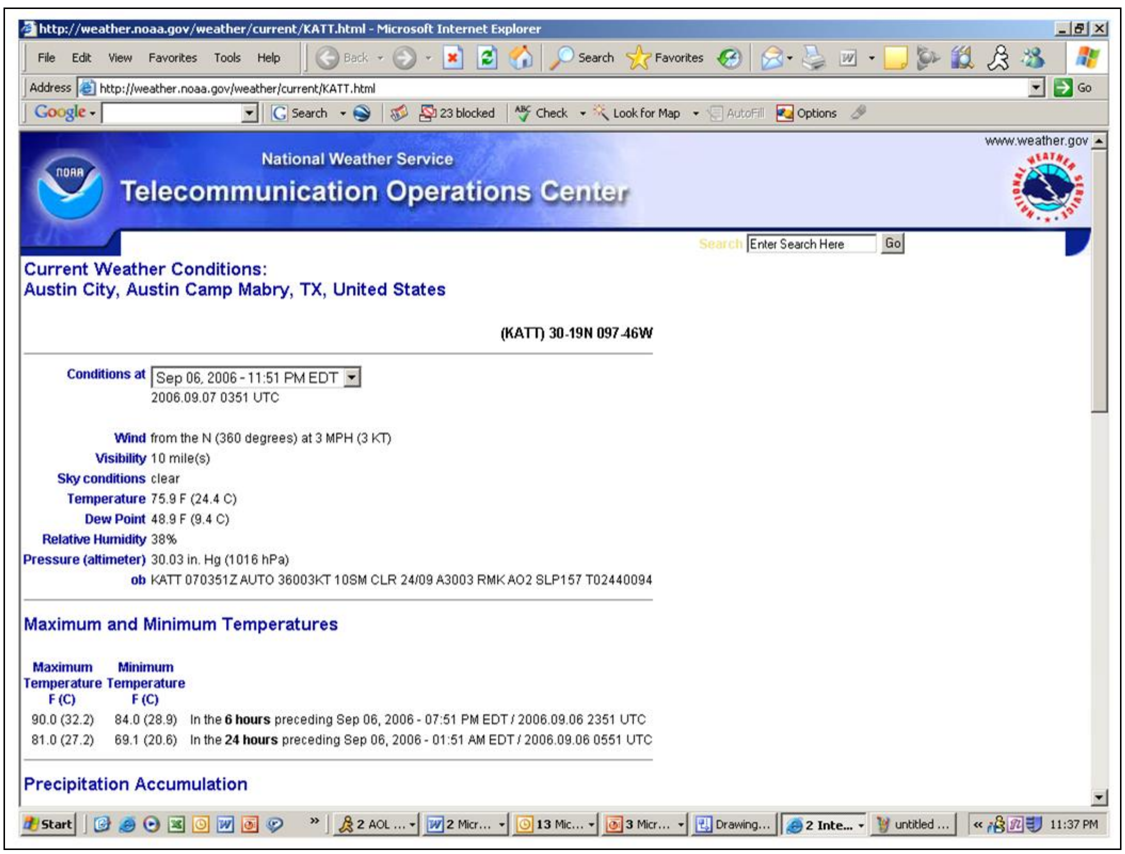


Figure 2. Screen Capture of National Weather Service (NWS) Website from Which the DCIPS Extracts Real-Time Weather Condition Information.

Dynamic Congestion and Incident Prediction Subsystem

As shown in Figure 1, the DCIPS subsystem consists of four primary components: the Prediction Data Communications Manager, the Traffic Condition Prediction Module, the Incident Prediction Module, and the Output Display Generator.

Prediction Data Communications Manager

The primary function of the Prediction Data Communications Manager is to extract the current traffic condition data from the RTDES at least once every minute and distribute them to each of the various traffic and incident prediction modules. The Prediction Data Communications Manager is also responsible for obtaining the current weather information from the NWS website and parsing the data stream to extract the relevant pieces of weather information used in the Incident Prediction Model (i.e., visibility, sky conditions, sunrise time, and sunset time). The Prediction Data Communications Manager is also responsible for saving all the input data for the different modules in a database up to 1 hour for evaluation purposes. After an hour, the data from the database are appended into a text file, and the database is cleared. The Prediction Data Communications Manager communicates with the two prediction modules through an interface that has controls that allow the modules to connect to the databases, read live data, and display their predictions. Each module processes the data, calculates predictions, and saves the output in the database every 1 minute. The structures of the databases where these data are stored are shown in Appendix A.

Traffic Condition Prediction Module

The Traffic Condition Prediction Module is the second major component of the DCIPS. Currently, the Traffic Condition Prediction Module contains two traffic prediction models—the Cumulative Flow Model and the Speed-Density Forecast Model—but has been designed to support the addition of other prediction models as they are developed. The following provides a brief description of the two models currently contained in the Traffic Condition Prediction Model. A more detailed description of these models can be found in Appendices B and C and in Report 0-4946-1 *(1)*.

Cumulative Flow Model. The Cumulative Flow Model was derived from work performed by Newell *(2)* and Cassidy and Windover *(3)*. The Cumulative Flow Model is based on the premise that under normal conditions, traffic detected at an upstream detector station can be expected to arrive at a downstream station at a time equivalent to the travel time of traffic. Flow that does not arrive at the expected travel time could be an indication of a potential operational problem. Figure 3 illustrates the concept of the Cumulative Flow Model. A complete description of the development of this model can be found in Appendix B.

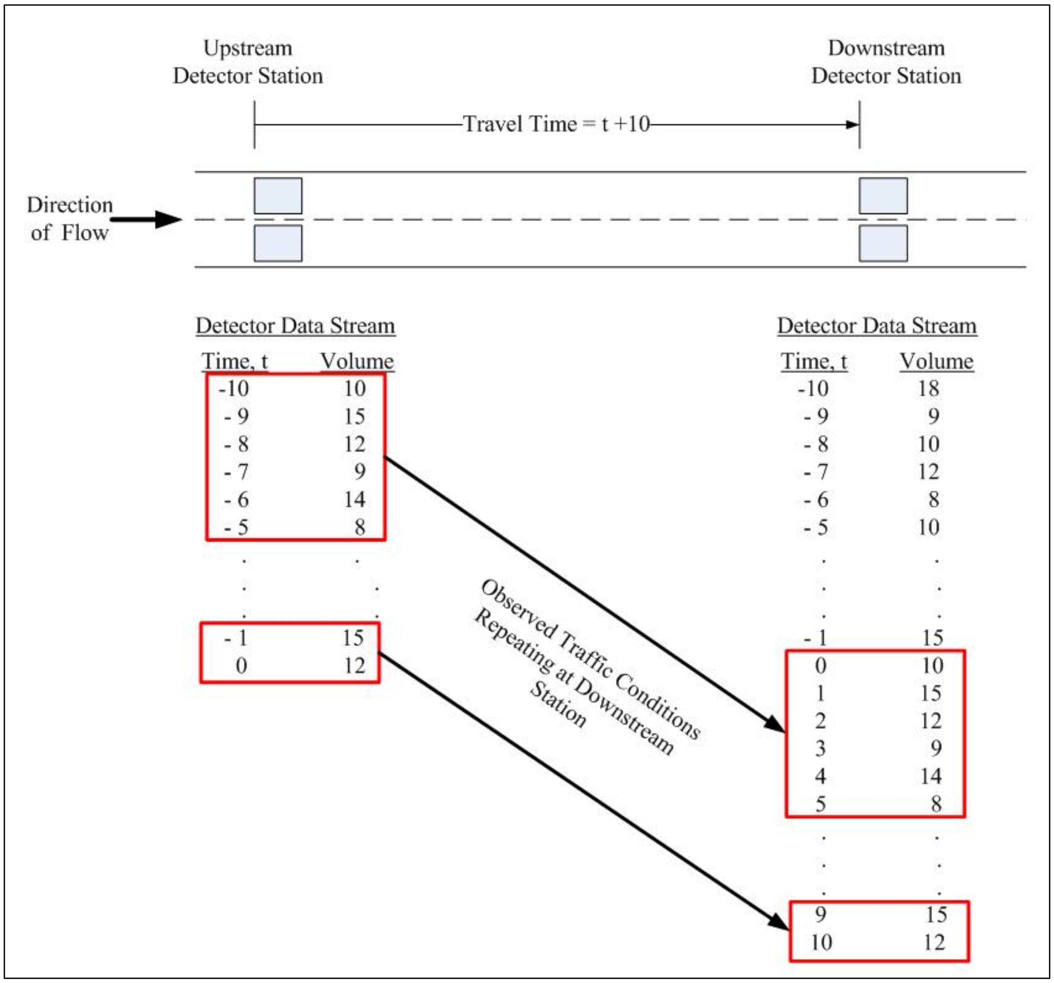


Figure 3. Illustration of the Concept Underlying the Cumulative Flow Model.

In this application, the Cumulative Flow Model produces two predictions—the flow-in-process and the delayed-flow. The flow-in-process measure is the number of vehicles traveling in a freeway section in a particular time and can be viewed as a proxy to density. As such, flow-in-process is translated into the number of passenger cars per mile per lane and the level of service. Thresholds proposed in Chapter 23 of the *Highway Capacity Manual* (*4*) are used for classifying the level of service of the flow-in-process.

The model also produces a delayed-flow ratio measure, which can be used as an indicator for assessing the level of congestion of a particular freeway section. The delayed-flow ratio is the ratio of the number of passenger cars being delayed to the number of passenger cars traveling in that section (or flow-in-process). A high delayed-flow ratio implies that a high percentage of vehicles are delayed and the operator should be alerted to a potential developing situation.

If a freeway section is able to maintain a good quality of service, the delayed-flow ratio will not provide extra information on the operation status of that freeway section. To avoid this false alarm, the delayed-flow ratio is only calculated and shown for freeway sections with levels of service E and F.

Speed-Density Forecast Model. The other traffic prediction model we included in the Traffic Condition Prediction Module is the Speed-Density Forecast Model. This model used a 5-minute moving average to predict expected traffic conditions the next minute into the future. The same approach was used to forecast traffic volume conditions multiple time steps into the future. Long-term forecasts into the future (more than 5 minutes into the future) were based solely on forecasted conditions. We used this approach to develop 1-minute, 5-minute, 10-minute, and 15-minute estimates of speed and density. Figure 4 illustrates the moving-average concept used to develop predictions of speed and density at individual detector stations.

Incident Prediction Module

In addition to providing predictions of future traffic conditions, we also developed a module to provide predictions of when and where traffic and environmental conditions were "favorable" for incident conditions to form. We included two different models that produce forecasts of when incident conditions are likely to occur. We called these models the Incident Prediction Model and the Crash Potential Model. A brief overview of these two models is provided below. A more in-depth discussion of the methodology used to develop these models can be found in Appendices C and D and in Report 0-4946-1 (*1*).

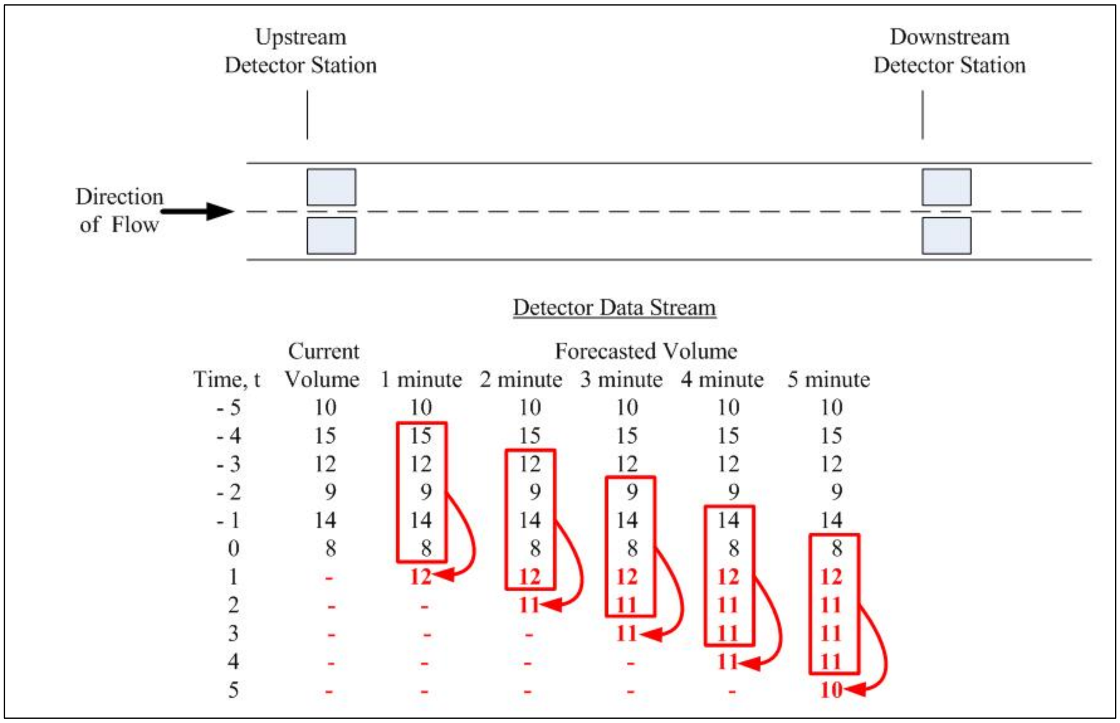


Figure 4. Illustration of Moving-Average Concept Used in the Speed-Density Forecast Model.

Incident Prediction Model. The Incident Prediction Model utilizes real-time weather and traffic data and a probability-based statistical model to produce a likelihood forecast of where incidents are likely to occur along freeway segments. To provide traffic-related inputs, 1-minute traffic data—volume, speed, and occupancy—as observed through inductive loop detectors were aggregated at 5-minute intervals. Current weather conditions were retrieved from the Internet on an hourly basis to provide weather-related inputs such as visibility and sky conditions. A statistical technique known as logit models was used to calibrate the model by examining historical relationships between incident and weather/traffic data.

Crash Potential Model. The final model included in the Incident Prediction Module was the Crash Potential Model. This model applies some of the concepts and procedures used in predicting high-accident locations and safety analyses (5 - 9) to estimate the likelihood that a crash will occur at a particular location given the forecast of the expected traffic volume and existing geometric conditions. The approach uses statistical procedures to identify real-time accident "precursors" in the traffic stream. The precursors are identified by determining what the traffic conditions were like just prior to an accident occurring and are developed from historical information about traffic conditions, geometric conditions, and accident experience at a particular location. In this research, we attempted to use both density and coefficient of variation of speed as real-time precursors to where accidents were likely to occur. Coefficient of variation of speed (CVS) is a measure of the amount of fluctuation (or variation) and is computed as the standard deviation in speed divided by the average speed. A high value of CVS implies that a large amount of fluctuation in speed exists in the traffic stream, which has been shown to cause accidents *(9)*.

Output Display Generator

The last major component of the DCIPS is the Output Display Generator. Each model processes the data, calculates its predictions, and saves the output in the database every 1 minute. The Output Display Generator plots graphs based on the predictions by the respective models. Examples of the output plots generated for each model are presented and discussed below.

In addition to generating the output display, the Output Display Generator is responsible for generating an archive of the model results. Output from each model is kept in the database that is updated every minute. Model outputs are retained in this database for 1 hour. After every hour, the data are moved to a daily text file (similar to what is done with the Austin Advanced Transportation Management System [ATMS] loop detector data). The weather data, which are updated every minute, are also saved into text files every hour per day.

Cumulative Flow Model Outputs. A sample of the type of output produced by the Cumulative Flow Model is provided in Figure 5. The Cumulative Flow Model provides a forecast of the flow-in-process and the ratio between the delayed-flow and flow-in-process to forecast the anticipated level of service. The user has the option to designate a targeted section of the freeway. The level-of-service thresholds provided in Chapter 23 of the *Highway Capacity Manual* *(4)* are used for classifying the level of service of the flow-in-process. The user has the option of predicting level of service based on the projected traffic conditions 1 minute, 5 minutes, 10 minutes, and 15 minutes into the future. If a freeway section is able to maintain a good quality of service, the delayed-flow ratio does not provide extra information on the operation status of that freeway section. To avoid false projections of poor performance, the delayed-flow ratio is only calculated and shown for freeway sections with levels of service E and F. Higher values of delayed-flow ratio indicate greater levels of predicted congestion.

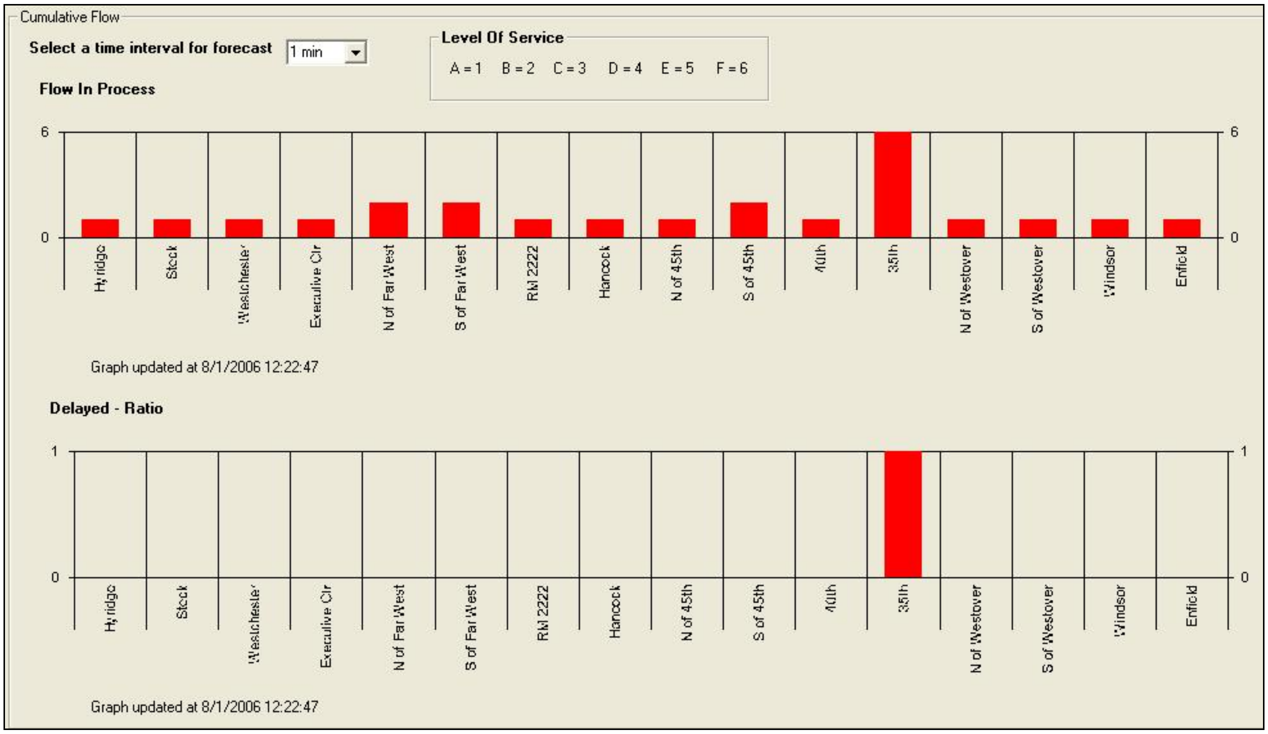


Figure 5. Sample Output of Cumulative Flow Model.

Speed-Density Forecast Model Outputs. This model forecasts the predicted speed and density for a given station in the designated target section of the freeway. The output manager speed is plotted in miles per hour (mph), while density is plotted in vehicles per hour per lane (vphpl). We have provided the user with the capability to select the time interval for which to display the speed and density forecasts—1, 5, 10, or 15 minutes into the future. Forecasted values are displayed for each station in a particular direction of travel.

Figure 6 shows a sample output of the graphs produced by the Speed-Density Forecast Model. Density values are displayed in green, while speed values are displayed in red.

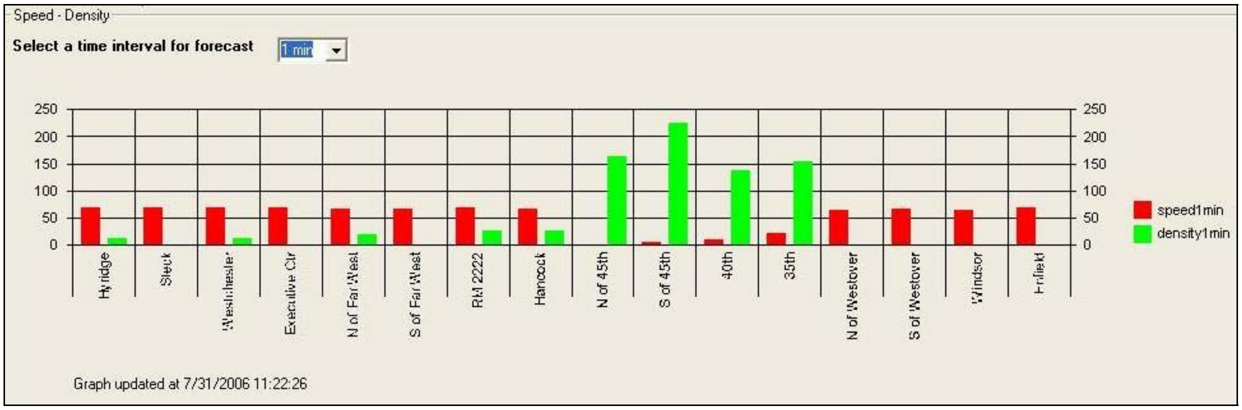


Figure 6. Sample Output of Speed-Density Forecast Model.

Incident Prediction Model Outputs. Figure 7 shows an example of the output generated for the Incident Prediction Model.

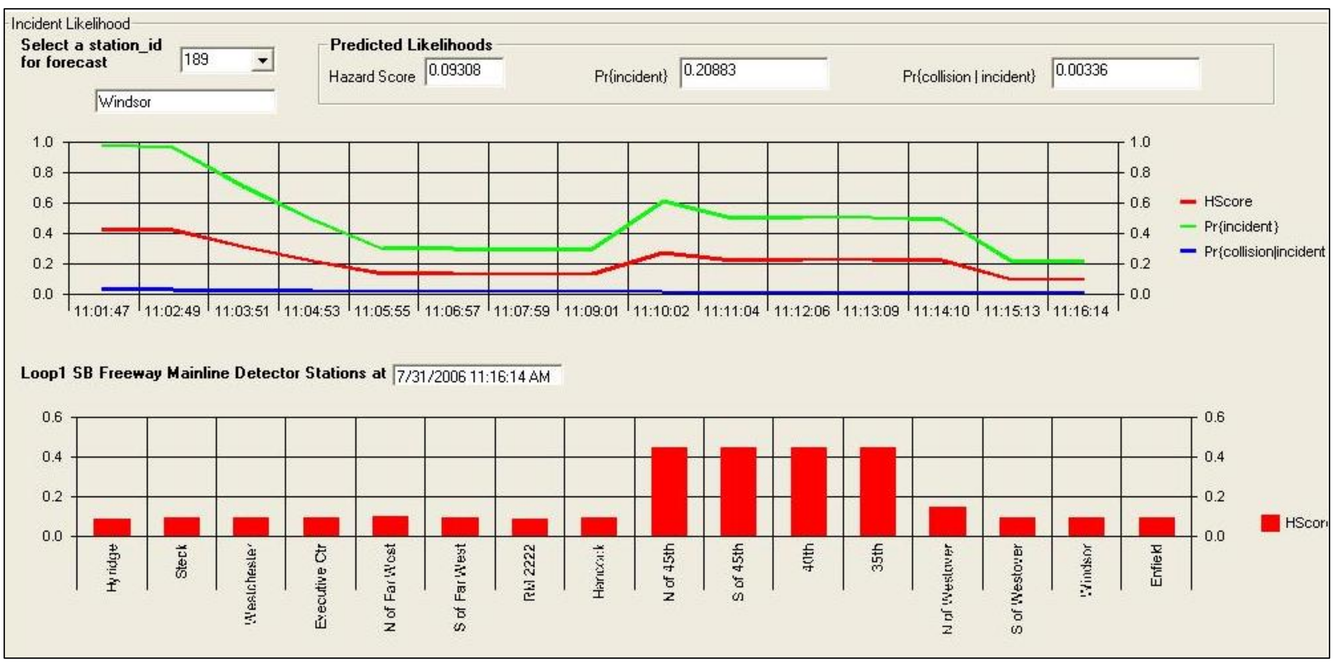


Figure 7. Sample Output of Incident Prediction Model.

The graphical display consists of two graphs updated in real time. The top graph displays the profiles of predicted probabilities and hazard scores over the last 15 minutes at the selected detector station. The user can specify the detector station of interest in this interface. Three predicted components are displayed in this graph: 1) probability of an in-lane incident,

2) probability of an incident being a collision given that an incident has occurred, and 3) hazard score. Using this graph, control center operators can monitor the trends to identify how long the predicted outputs at the selected location have been in a particular state. A persistent check can also be performed using this display. For instance, control center operators may require hazard scores to exceed a particular threshold for at least 3 minutes consecutively before any appropriate actions are warranted.

The bottom graph displays the most recent hazard scores across all the detector stations along the freeway mainline of the test bed. Control center operators can use this graph to quickly review traffic situations throughout the entire test bed and determine where the traffic and weather conditions are critical for freeway operations. Higher hazard scores would imply more hazardous conditions at particular detector stations.

Dual thresholds for hazard scores can be used to designate locations that require attention from control center operators. The lower threshold, once exceeded, can be used to put operators on alert conditions, while the higher threshold can be used to identify the locations that require immediate attention. The hazard score thresholds of 0.16 and 0.28 are recommended based on historical observations of Loop 1 and US 183 in Austin, Texas. These thresholds are considered preliminary and should be used only if there is no better information available. These thresholds are expected to be updated over time to reflect the current freeway conditions once the system has been deployed. To reduce the rate of false alarms, a persistence check can be applied before an alert can be issued.

Crash Potential Model Outputs. In our implementation, the Crash Potential Model uses real-time and forecasted measures of speed and density to predict when and where incidents are likely to occur. Speed and density measures are extracted from the loop detectors and fed into the model to produce an estimate of the probability that an accident will occur given the current or predicted conditions. The system then converts this likelihood estimate into the crash potential. Crash potential is defined as the expected number of crashes over a certain period of exposure. Exposure is denominated as vehicle miles of travel. The Crash Potential Model reports crash potential for that particular minute it is evaluating by looking at the precursors over the current time and preceding 4 minutes. A sample of an output screen for the Crash Potential Model is shown in Figure 8.

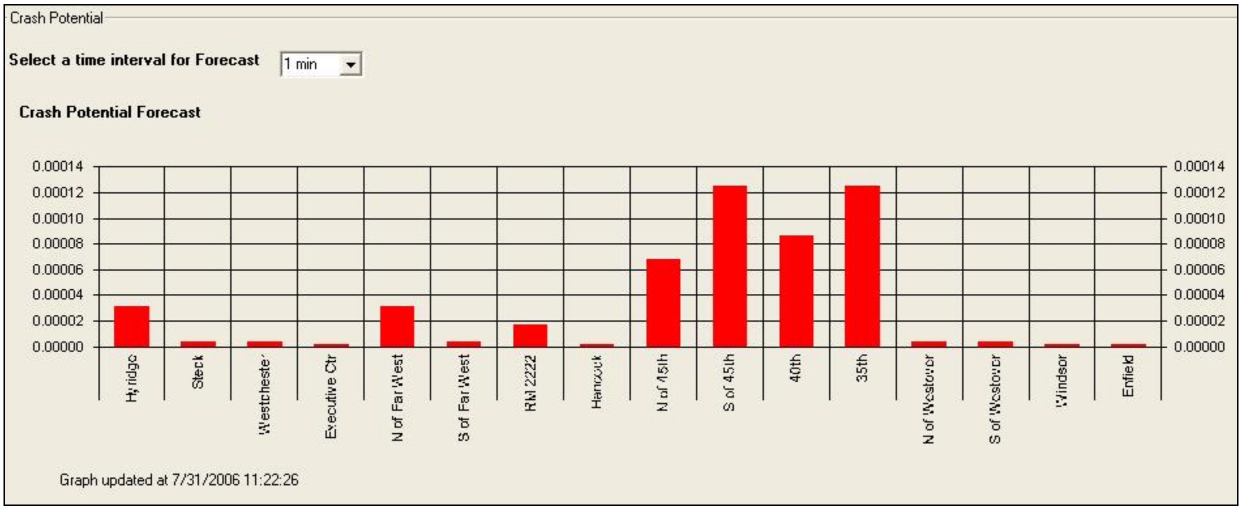


Figure 8. Sample Output of Crash Potential Model.

DATA FLOWS

Figure 9 is a data flow diagram for the DCIPS prototype. The diagram shows what data elements are used by each of the model components and how the data move between these components. The DCIPS uses 1-minute speed, volume, and occupancy information collected from traffic monitoring systems (i.e., loop detectors) from the TxDOT ATMS. The DCIPS also uses weather information that is extracted from a website produced by the National Weather Service. The SCU Communications Manager is responsible for extracting traffic condition data from the TxDOT ATMS. This information is stored in a database that is part of the SCU Communications Manager subsystem. The Prediction Data Communications Manager is responsible for extracting the 1-minute traffic condition data from the database and distributing the appropriate traffic condition information to each of the prediction models. The Prediction Data Communications Manager is also responsible for securing the current weather data from the NWS website. The Output Display Generator uses the output from each of the prediction models to generate the appropriate model displays described above. Also, the Output Display Generator is responsible for time-stamping and placing the model output results into an archival database that can be used to evaluate the accuracy of the prediction models.

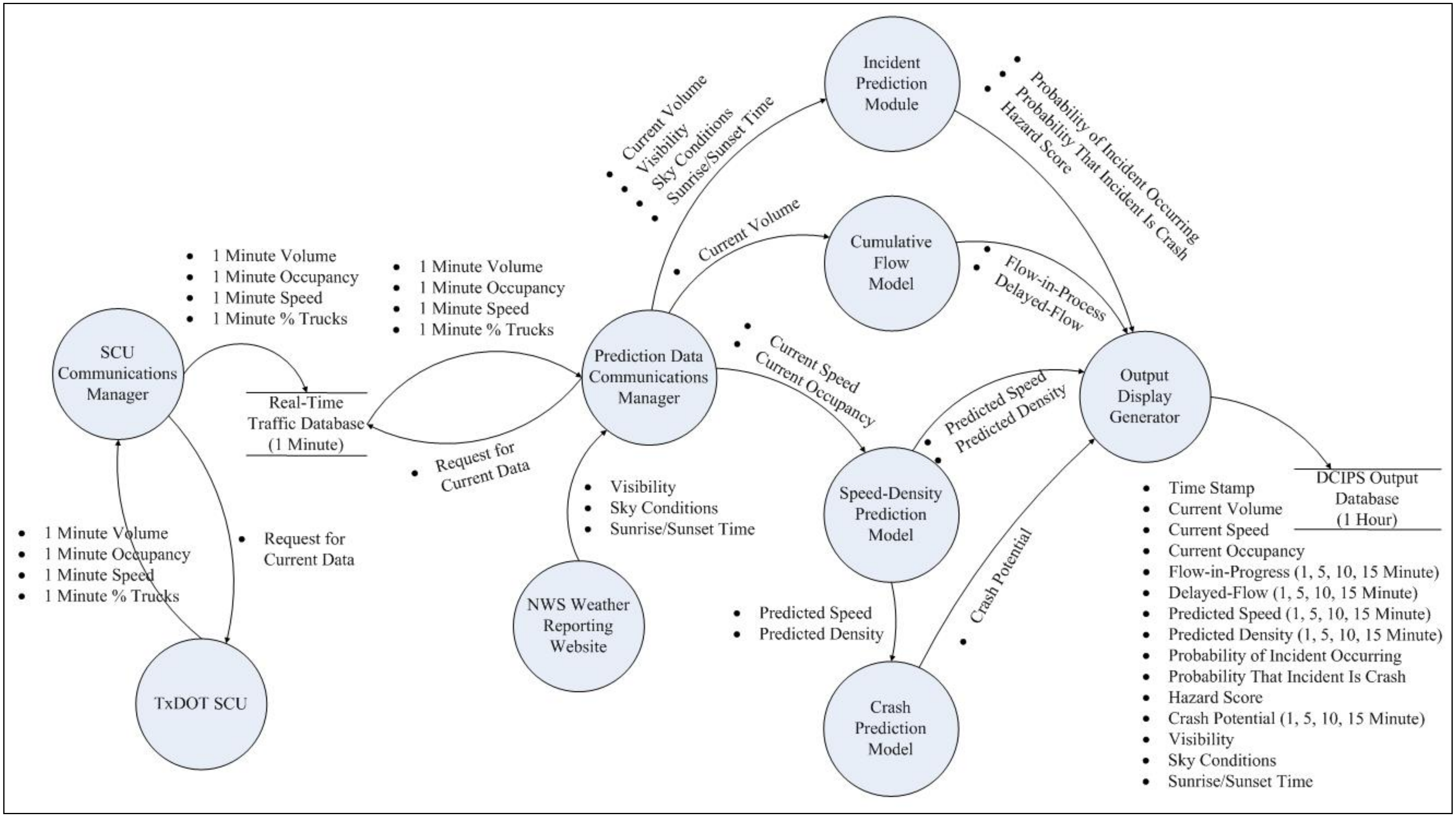


Figure 9. Data Flows in the Dynamic Congestion and Incident Prediction System.

USER’S GUIDE

In this section of the report, we provide a user’s guide on how to install, operate, and interpret the model results from the DCIPS prototype tools. First, we will describe the software and hardware requirements needed to operate the DCIPS tools. We will then describe the processes and procedures for installing, operating, and removing the software program on a personal computer (PC). The system has been designed to operate as a stand-alone application using real-time traffic and weather condition information.

SOFTWARE AND HARDWARE REQUIREMENTS

The SCU Communications Manager software is a stand-alone computer program designed to query and extract traffic detector data directly from TxDOT’s System Control Unit. Written in Visual Basic, the SCU Communications Manager requests the most recent update of traffic information from the SCU, compares this information to previously stored data to determine if conditions have changed, and then requires new information from a database. It connects to the TxDOT SCU through one of the standard output ports on the SCU.

The DCIPS software is a compilation of several individual programs—all written in Visual Basic. A setup program has been prepared to allow all of the components to be installed automatically. Ideally, the SCU Communications Manager and the DCIPS software need to be installed on the same computer in the control center. The folders for DCIPS on the machine will be located at C:\Program Files\DCIPS. The DCIPS folder will also contain the folders for the day files consisting of input traffic and weather data after you execute the application.

Both the SCU Communications Manager and the DCIPS have been designed to run on a PC. We recommend that the computer have a minimum of 1 gigahertz central processing unit (CPU) and 512 megabytes (MB) of random access memory (RAM), and runs one of Microsoft’s operating systems: Windows NT®, Windows 2000®, or Windows XP®. Both the SCU Communications Manager and the DCIPS use a Microsoft Access® database for retaining necessary input data. Both the SCU Communications Manager software and the DCIPS software need to be installed on the computer before the user can run the DCIPS.

INSTALLING THE DCIPS SOFTWARE

We have provided TxDOT with an installation compact disc (CD) that contains the DCIPS software. Included on this CD is a program to automatically install the DCIPS programs. After placing the CD in the appropriate drive, the user should double-click the *Setup* file located in the disc directory. This should start the installation process and cause the installation program to display the welcome screen shown in Figure 10. To continue with the installation process, the user should click the *OK* button located at the bottom of the welcome screen. If the user does not wish to install the software, he or she can click the *Exit Setup* button on the bottom right of the welcome screen.

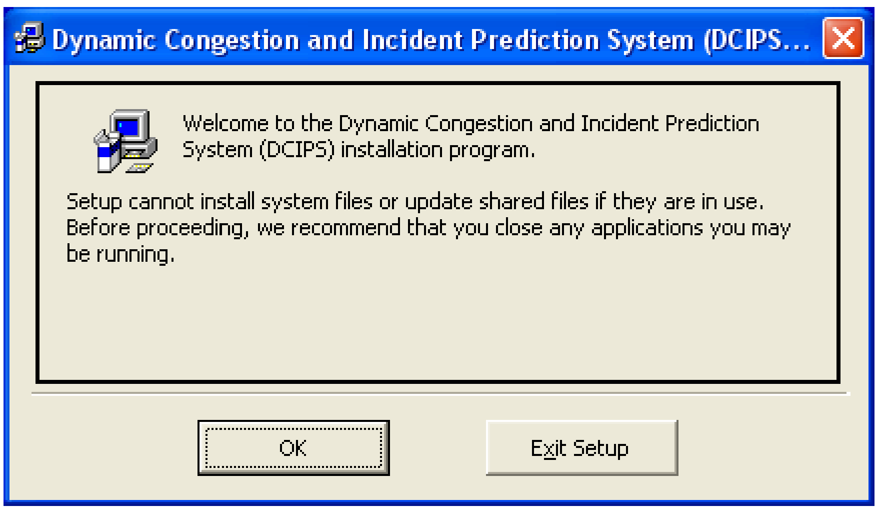


Figure 10. Welcome Screen of the DCIPS Installation Program.

When the user hits the *OK* button in Figure 10, it will cause the *DCIPS Setup* screen to appear. This screen is shown in Figure 11.

The user should then click the install button (which looks like a computer and software, as show in Figure 11) on this screen to continue with the installation program. This will cause the setup program to create a new directory—the Predictor Data Communications Manager—in the C:\Program Files directory and to install the DCIPS in that directory.

If the user would like to install the DCIPS program in a subdirectory other than the one specified, he or she can change directories by clicking the *Change Directory* button. This will allow the user to enter the path name to the directory where he or she would like the program installed. After changing the directory, the user should then click the install button to continue the installation process.

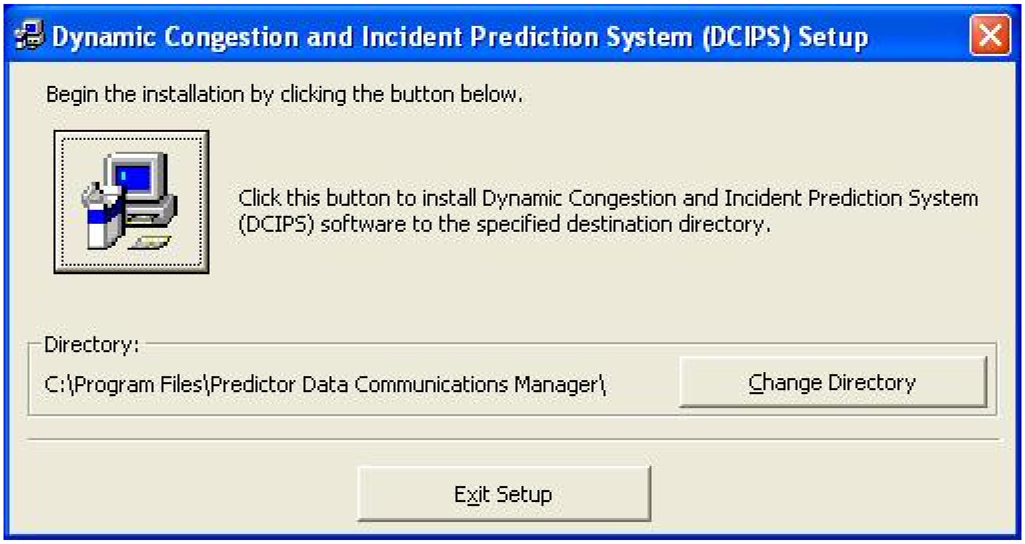


Figure 11. The Initial Setup Screen for Launching the DCIPS Installation Program.

The user can quit the installation program by clicking the *Exit Setup* button at the bottom of the screen.

After hitting the install button, the screen shown in Figure 12 should appear. As a default, the installation program will install the DCIPS in the *Dynamic Congestion and Incident Prediction System* Program Group. If the user would like to use a different program group, he or she can enter a new group name directly or select one from the *Existing Groups* listed on the screen.

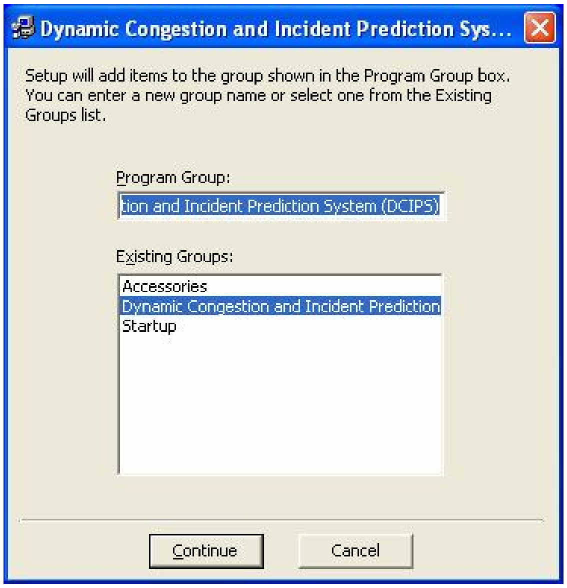


Figure 12. Input Screen Where the User Can Specify the Location for Installing the DCIPS.

Once the user has entered the desired program group name, he or she should hit the *Continue* button. This will cause the setup program to install the DCIPS program on the computer. The user can terminate the installation process by clicking the *Cancel* button.

If the installation process is successful, the user will see the screen shown in Figure 13. To clear the screen, the user should then click the *OK* button.

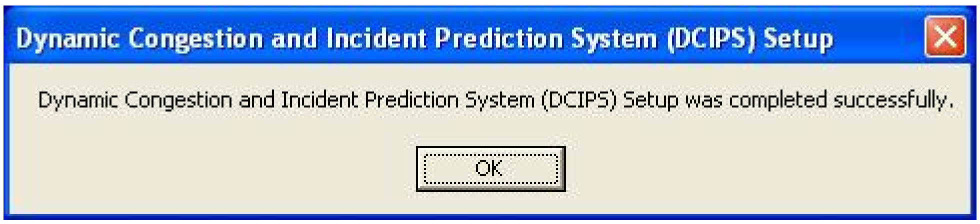


Figure 13. Screen Indicating That the DCIPS Has Been Successfully Installed.

CONNECTING TO SCU COMMUNICATIONS MANAGER'S DATABASE

After installing the DCIPS on the computer, the user still needs to establish a connection between the DCIPS and the SCU Communications Manager's database. This can be done using the Microsoft Windows® Administrative Tools program found on the *Control Panel* screen (see Figure 14). To start the process of connecting the DCIPS to the SCU database, the user should double-click the *Administrative Tools* icon.

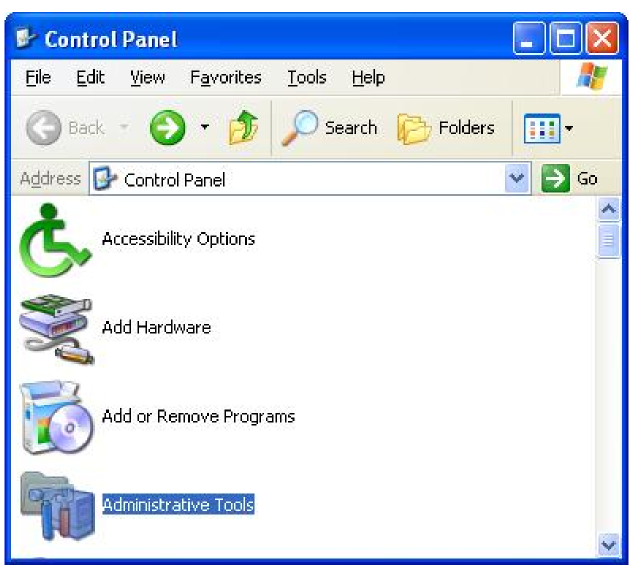


Figure 14. The *Administrative Tools* Icon Located in the Control Panel.

Once the user double-clicks the *Administrative Tools* icon, the computer should display the screen shown in Figure 15. The user then double-clicks the *Data Sources (ODBC)* shortcut on this screen.

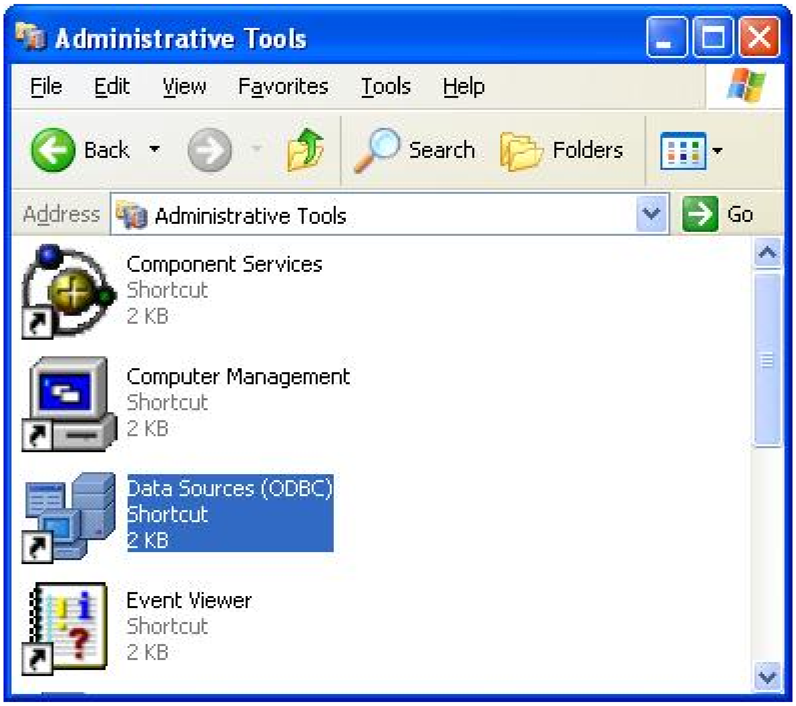


Figure 15. Shortcut for Configuring the Data Source for the DCIPS.

Double-clicking the *Data Sources (ODBC)* shortcut will cause the computer to open the *ODBC Data Source Administrator* screen. The user should then select the *System DSN* tab located near the top of the screen. After clicking this tab, the user should see the input screen shown in Figure 16. From this screen the user would click the *Add* button on the right side of the screen. This will cause the computer to display a screen similar to that shown in Figure 17. From this screen, the user should scroll down the list of available data source drivers and select *Microsoft Access Driver (*.mdb)* from the list. If the *Microsoft Access Driver (*.mdb)* option is not shown in the list, the user will need to install Microsoft Access®. The user can either cancel out of the *ODBC Data Source Administrator* screen and install Microsoft Access®, or contact his or her system administrator for assistance. The user would then need to repeat the steps for connecting the database after the software has been installed.

Once the appropriate database driver name has been highlighted, the user can click the *Finish* button at the bottom of the screen. This should open a new screen similar to that shown in Figure 18.

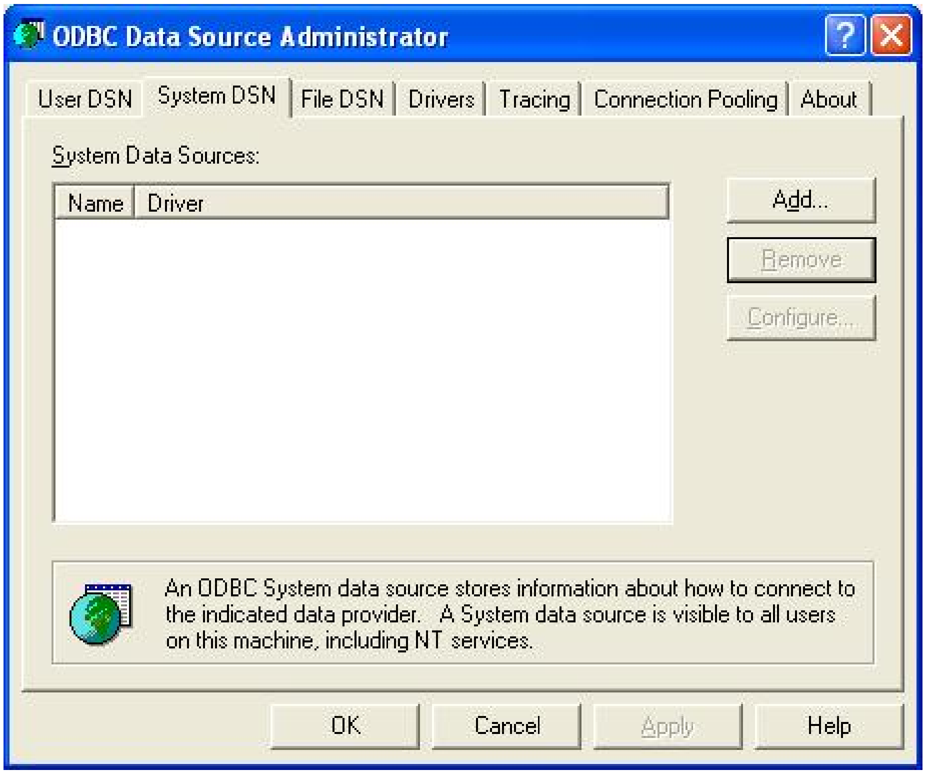


Figure 16. ***ODBC Data Source Administrator*** **Screen.**

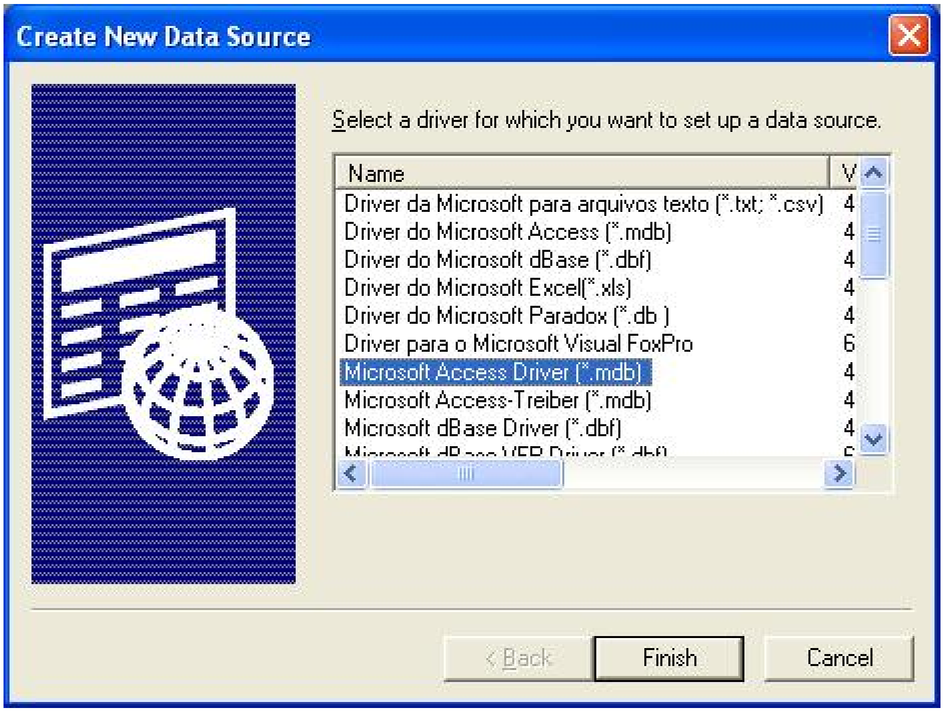


Figure 17. Screen for Selecting Data Source Driver.

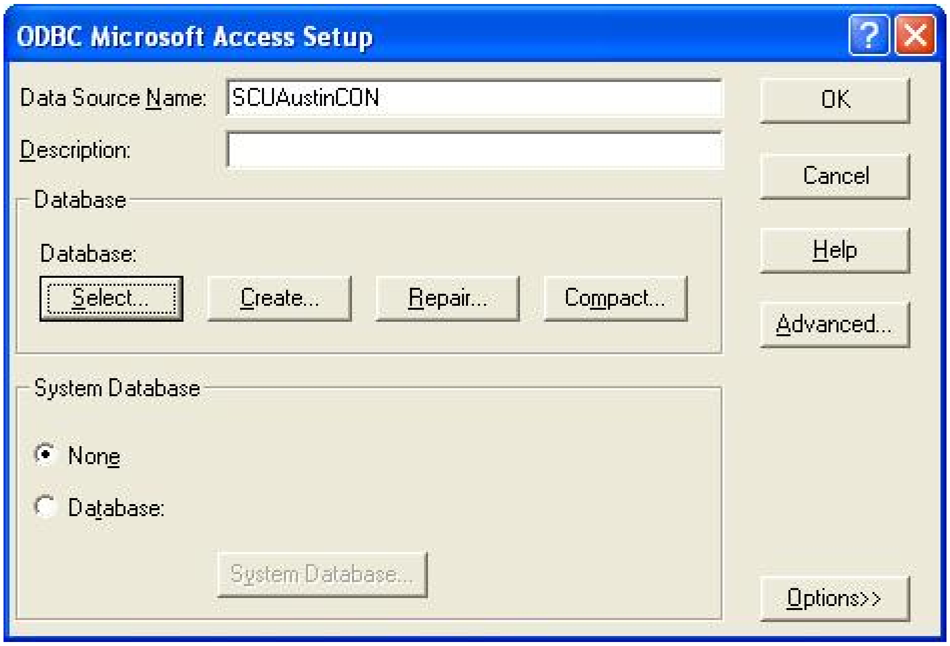


Figure 18. ***ODBC Microsoft Access Setup*** **Screen.**

In the space provided next to the *Data Source Name*, the user should enter the name of the database to which he or she is connecting. In our example, we have called this data source name *SCUAustinCON*. After providing a data source name, the user then needs to provide a name and path where this data source is located. To do this, the user should click the *Select* button in the *Database* section of the input screen. This will open a screen similar to that shown in Figure 19. From this screen, the user should enter the path name where the *SCU_Austin_LiveDatabase*—the database used by the SCU Communications Manager—is located. After entering the path name, the user should click the *OK* button on the right-hand side of the screen. That will return the user to the *ODBC Microsoft Access Setup* screen, except this time the path name will be displayed in the *Database* section (see Figure 20). The user can then click the *OK* button on the right-hand side of the screen to complete the connection process. The DCIPS will be connected to the SCU Communications Manager database, and the DCIPS will be able to extract the traffic condition information from the SCU.

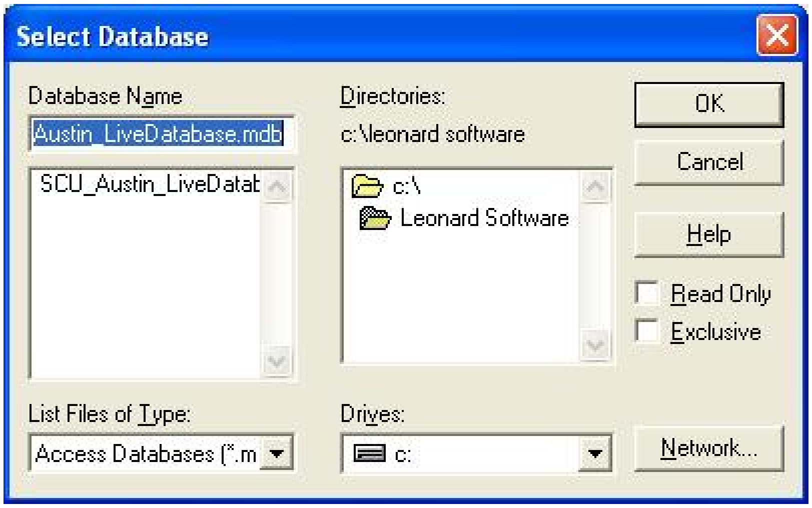


Figure 19. Input Screen for Entering Path to SCU Real-Time Database.

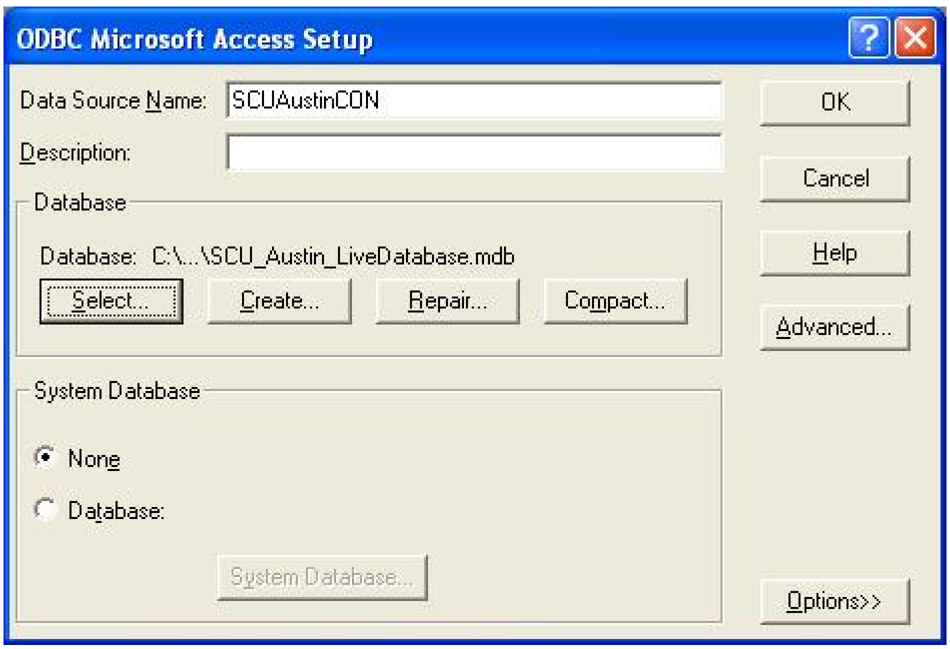


Figure 20. *ODBC Microsoft Access Setup* Screen with Path to SCU Real-Time Database.

OPERATING THE DCIPS

Once the software programs have been installed and the connection has been made between the SCU Communications Manager and the DCIPS, the user is ready to begin operating the prototype tool. The installation program places the DCIPS program in the Program Manager. To execute the DCIPS, the user should click *Start* and then *All Programs*. The user should then navigate through the programs listed there until he or she finds the *Dynamic Congestion and*

Incident Prediction System (DCIPS) shortcut. Double-clicking the shortcut will cause the DCIPS program to launch. Figure 21 shows the initial DCIPS screen after the user launches the program.

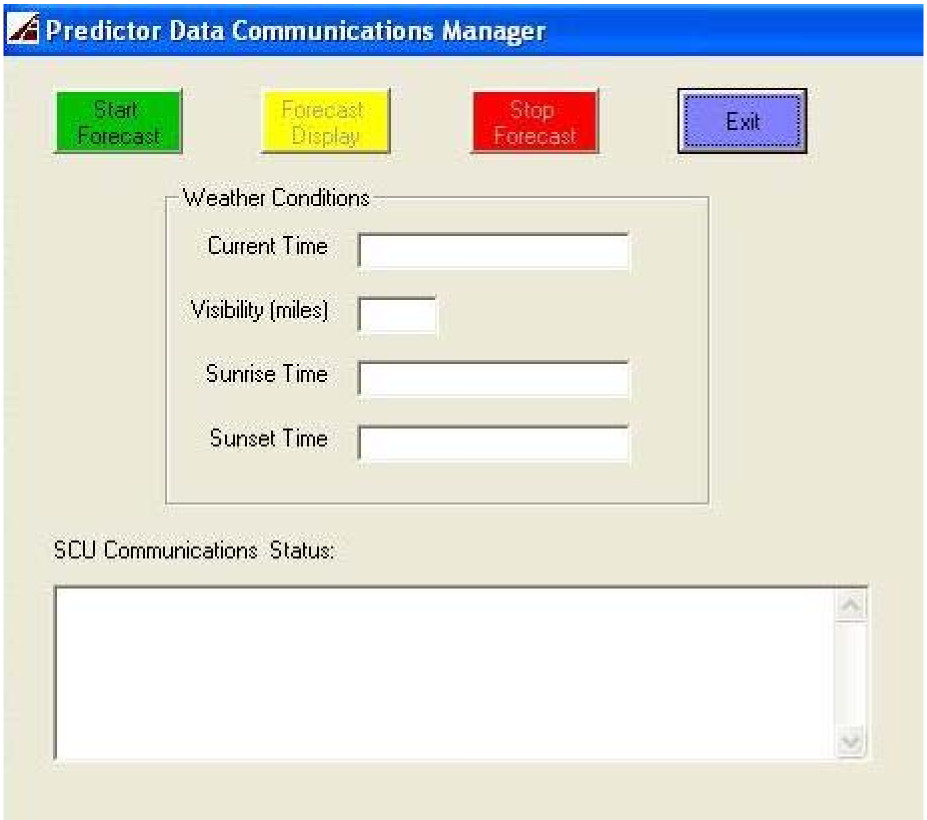


Figure 21. Initial Start Screen of the DCIPS Prototype Tool.

The user should then click the green *Start Forecast* button in the upper left-hand corner of the initial screen to start the DCIPS producing forecasts. The DCIPS will go out to the identified NWS website location and download the current visibility and sunrise and sunset times, as well as establish a connection to the SCU Communications Manager.

If a successful connection to these subsystems is made, the user should see a screen similar to that shown in Figure 22. The screen will update to provide two pieces of status information: the current weather conditions downloaded from the NWS website and the status of the communications to the SCU. *SCU Communications Status* shows the status of the DCIPS with the SCU Communications Manager. If the connection is good, it displays a message *Connection successful* along with a time stamp. If the connection is lost, it displays a message *No real time data available from SCU* along with a time stamp.

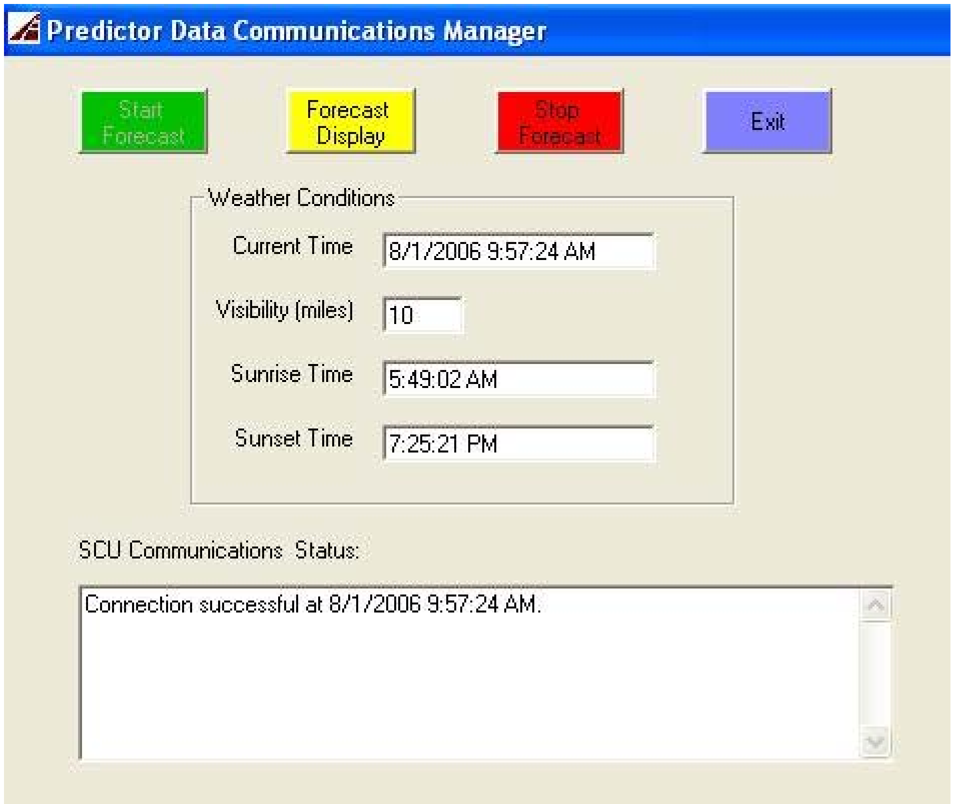


Figure 22. Initial Screen after User Clicks *Start Forecast* Button.

After clicking the green *Start Forecast* button, three other buttons will be activated: a yellow *Forecast Display* button, a red *Stop Forecast* button, and a blue *Exit* button. The red *Stop Forecast* button will cause the DCIPS to stop forecasting traffic and incident conditions. This is not the same as exiting the program. The operator might use the *Stop Forecast* button if he or she wants to essentially "pause" the application (i.e., stop producing forecasts) without causing it to terminate permanently. To fully terminate the program, the user needs to click on the blue *Exit* button to exit from initial startup screen. Note, clicking on the red X in the top right-hand corner of the startup screen does not exit from the application. When the user tries to exit that way, he or she gets the warning message shown in Figure 23. The only way the user can exit from the DCIPS is by clicking the *Exit* button.

Figure 23. Error Screen If the User Tries to Exit the DCIPS Incorrectly.

To begin receiving traffic conditions and incident forecast information, the user should click on the yellow *Forecast Display* button (see Figure 22). This will cause the DCIPS to display a screen similar to that shown in Figure 24. From this screen, the user can see the status of the communications with the SCU Communications Manager (in the *SCU Communication Status* window on the upper left-hand side of the screen). The screen also contains a *Bad Data Status* window (upper right-hand side of the screen) that shows the user an error warning if the DCIPS has detected and corrected any bad data from the SCU.

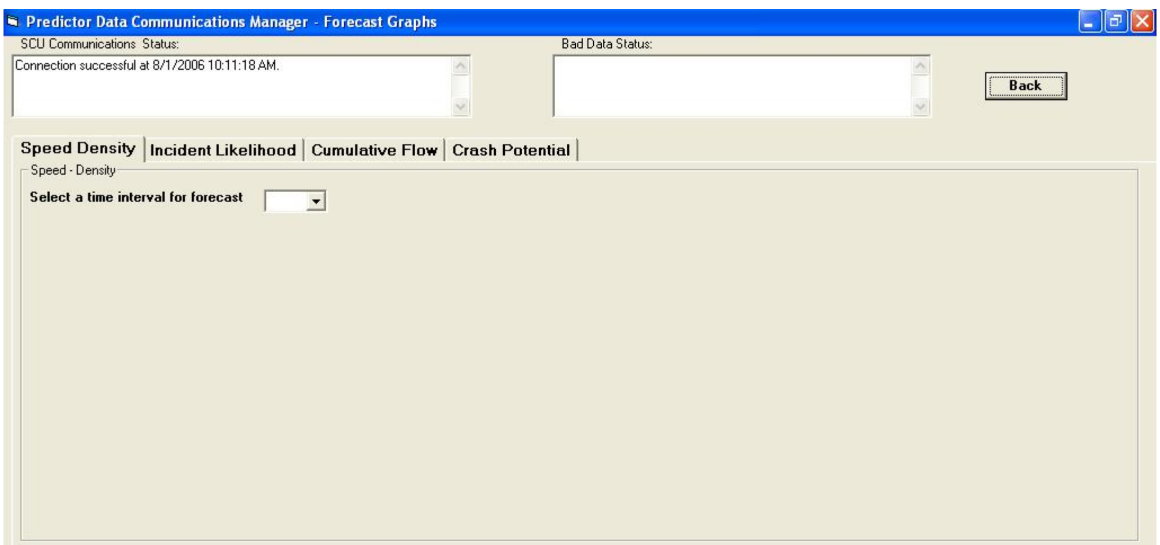


Figure 24. Initial Status Screen for Displaying Traffic Condition and Incident Predictions.

This screen also contains four tabs, each one corresponding to the traffic conditions and incident prediction models: the Speed-Density Model, the Incident Prediction Model, the Cumulative Flow Model, and the Crash Potential Model.

REMOVING THE DCIPS

The user can remove the DCIPS program from the computer using the *Add or Remove Programs* feature in the Windows® software. To access this feature, the user should access the

Control Panel through the *Start* menu on the desktop. From the *Control Panel* screen, the user should select the *Add or Remove Programs* icon. After clicking the *Add or Remove Programs* icon, the user sees a listing of all of the programs installed on the PC. The user should scroll down the list and highlight *Dynamic Congestion and Incident Prediction System (DCIPS)* from the list. The user should then click the *Change/Remove* button displayed on the screen. Clicking that button should cause the screen shown in Figure 25 to be displayed to the user. If the user wishes to continue with the removal process, he or she clicks the *Yes* button on the left-hand side of the screen. Clicking the *No* button will cause the removal process to terminate without removing the program.

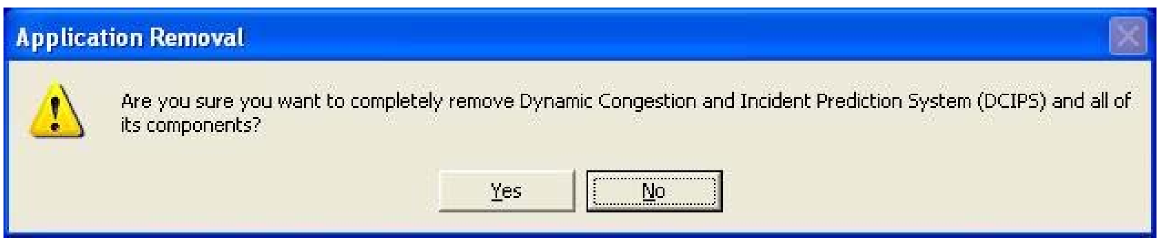


Figure 25. Confirmation Screen for Removing the DCIPS.

Once Windows® has successfully removed the program; the user will receive the message shown in Figure 26 confirming that the program has been removed from the computer.

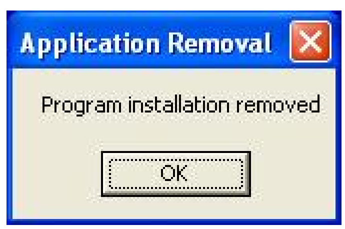


Figure 26. Confirmation Screen That the DCIPS Has Been Successfully Removed from the Computer.

PROOF-OF-CONCEPT TESTING

In this section, we first describe how the system was set up for integration and Real-Time Incident Prediction System (RIPS) tool testing. Then, we ran the example test scenario to illustrate how the RIPS would work in the field.

TEST SETUP

We tested the developed RIPS in the laboratory setting using the hardware-in-the-loop simulation concept. VISSIM was selected as a simulation test bed where a 7-mile segment of Loop 1 was modeled and loop detectors were placed along the simulated network to generate detector observations. We used the Austin detector database to ensure that the detector placement in the simulated network corresponded to actual locations. The testing concept can be depicted as shown in Figure 27.

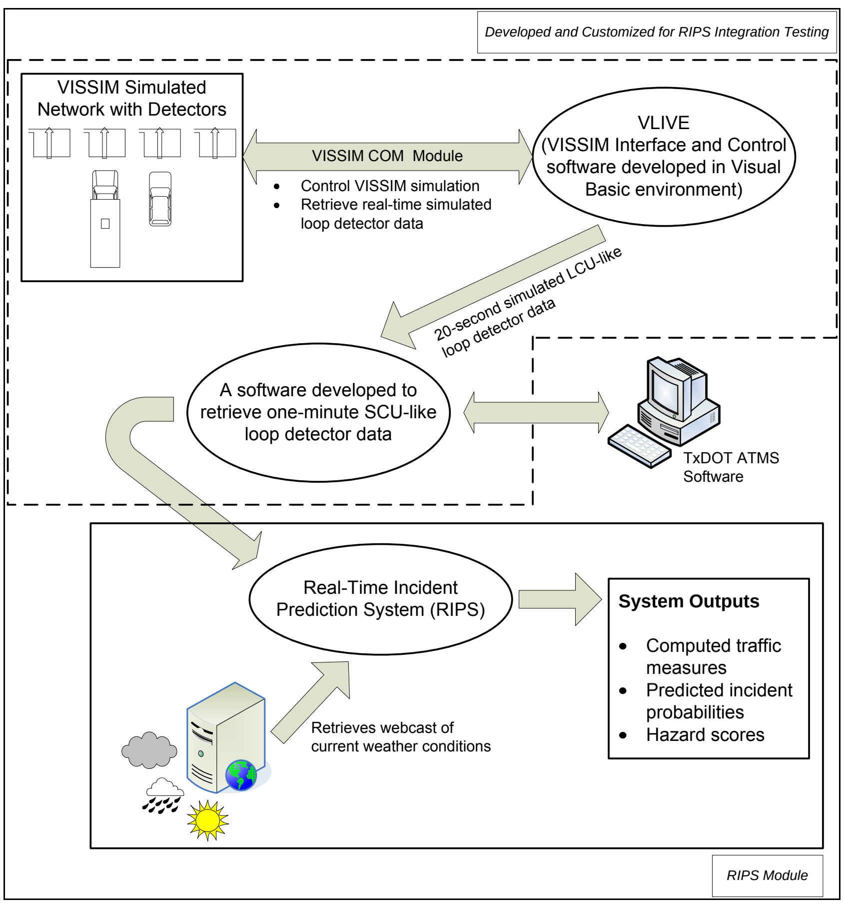


Figure 27. Diagram of RIPS Integration Testing.

In a simulated network, a single-lane-blocked incident was coded in the simulated network using vehicle actuated programming (VAP) in order to illustrate the effects of the incident on traffic flow parameters. Incident location, starting time, and duration can be specified and modified by users.

From the diagram, VLIVE software (see Figure 28) was developed in a Visual Basic (VB) environment integrating the VISSIM COM capability through VB's graphical user interface designed for users to control and observe changes in traffic parameters in run-time. VLIVE retrieves the simulated loop detector data from VISSIM and aggregates them into a 20-second data format similar to Local Control Units (LCU). These 20-second data from VLIVE are then combined in 1-minute data at the SCU level. Simple validation checks such as threshold checking are performed at the SCU level. A program was developed to retrieve 1-minute SCU data directly from the ATMS and then store them into a database.

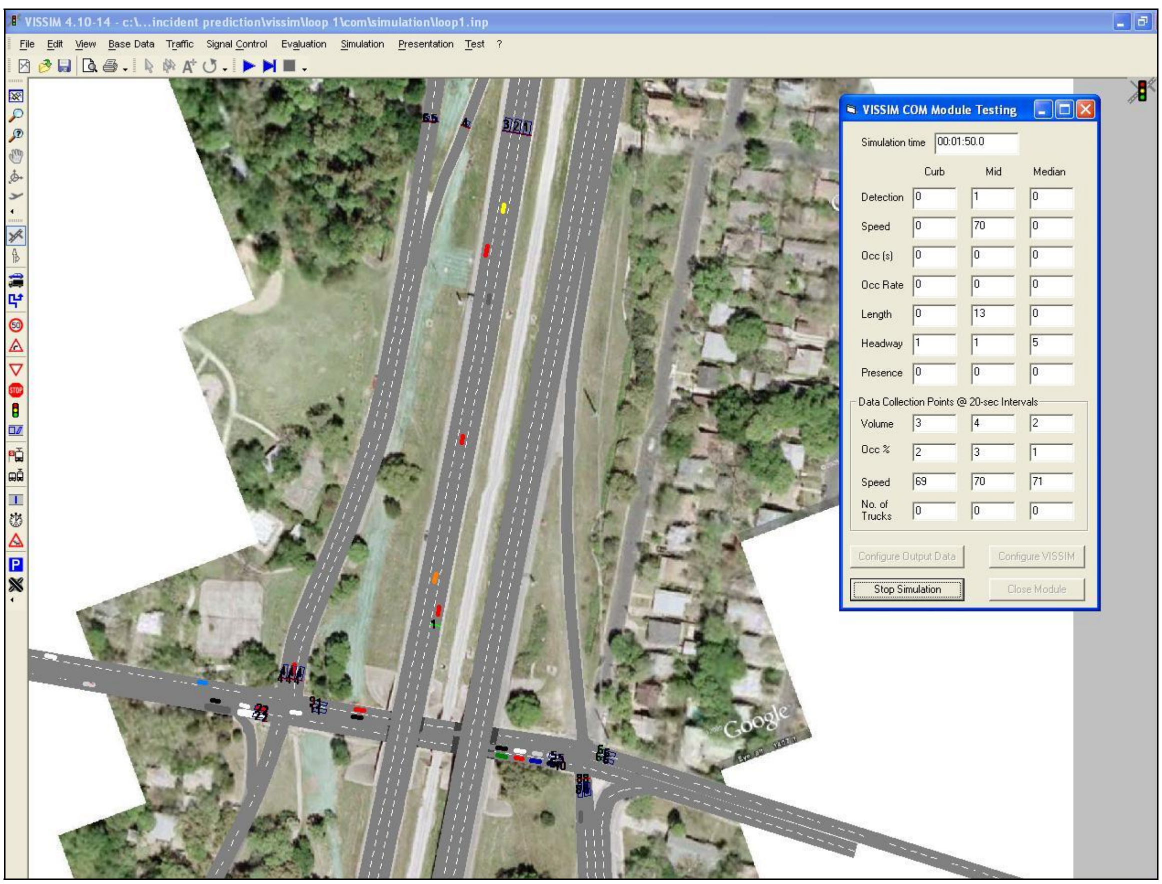


Figure 28. VLIVE Software and the Simulated VISSIM Network.

SIMULATED NETWORK

A 7-mile freeway segment of Loop 1 located in the west of Austin, Texas, from US 183 to Lake Austin Boulevard was selected as a simulation test bed. The test bed consisted of a total of 69 individual inductive loop detectors on mainline and ramp sections. The web-based weather conditions from the Camp Mabry station located near the middle of the test bed were retrieved to provide inputs into the incident prediction models.

A one-lane-blocked incident was programmed to occur on the test bed in order to examine the effects of an incident on traffic flow parameters as observed through inductive loop detectors. The incident location, duration, and occurrence time can be modified by users through VISSIM VAP.

EXAMPLE SCENARIOS AND RESULTS

A test incident scenario was a one-lane-blocked incident that lasted for 10 minutes near the south terminus of the test bed. Figure 29 through Figure 34 depict the simulated traffic conditions over the period of incident timeline and the corresponding prediction results. Three scenarios are presented herein: non-incident condition, during-incident condition, and post-incident condition.

For non-incident traffic and normal weather conditions (see Figure 29), the model outputs shown in Figure 30 show hazard scores in the range of 0.10 to 0.16, which are below the suggested threshold of 0.18, indicating that the current freeway traffic and weather conditions are unlikely to produce incidents. The top graph in Figure 30 displays the most recent hazard scores across all the detector stations along the freeway mainline of the test bed. Control center operators can use this graph to quickly review traffic situations throughout the entire test bed and determine where the traffic and weather conditions are critical for freeway operations. The bottom graph in Figure 30 shows the predicted outputs over the past 15 minutes at the selected detector station specified by users. Using this graph, control center operators can monitor the trends to identify how long the predicted outputs at a selected location have been in a particular state. A persistent check can also be performed using this graph. For instance, control center operators may require hazard scores to exceed the threshold for at least 3 minutes continuously before any appropriate actions will be put in place.

Similarly, Figure 31 and Figure 32 illustrate simulated traffic conditions during an incident and the corresponding model outputs respectively. The corresponding displays for a post-incident scenario are presented in Figure 33 and Figure 34. In this case, it should be noted that there is a delay period for hazard scores and predicted probabilities to decrease once an incident has been cleared. This corresponds with a notion that a risk of a secondary incident can be above normal right after the incident has been cleared and before traffic completely returns to pre-incident conditions.

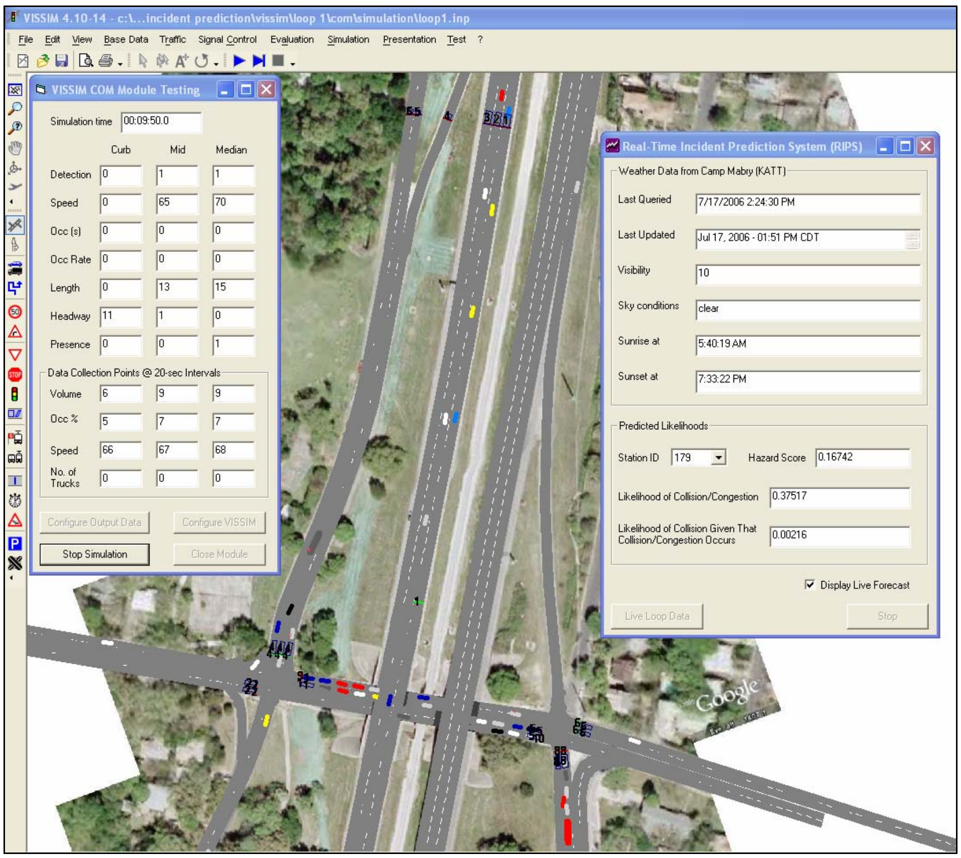


Figure 29. Non-Incident Traffic Conditions.

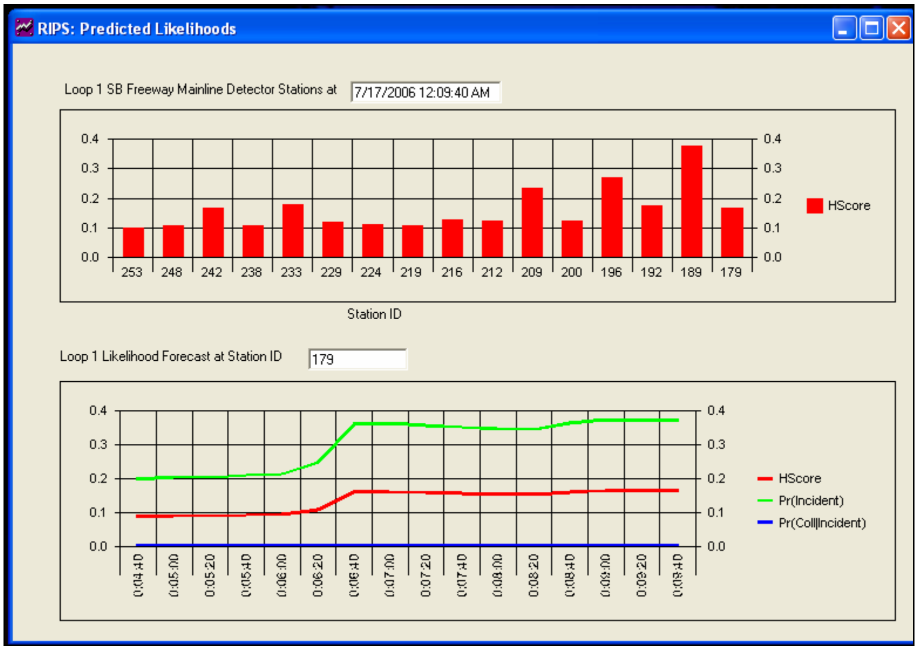


Figure 30. Example Model Outputs for Non-Incident Traffic Conditions.

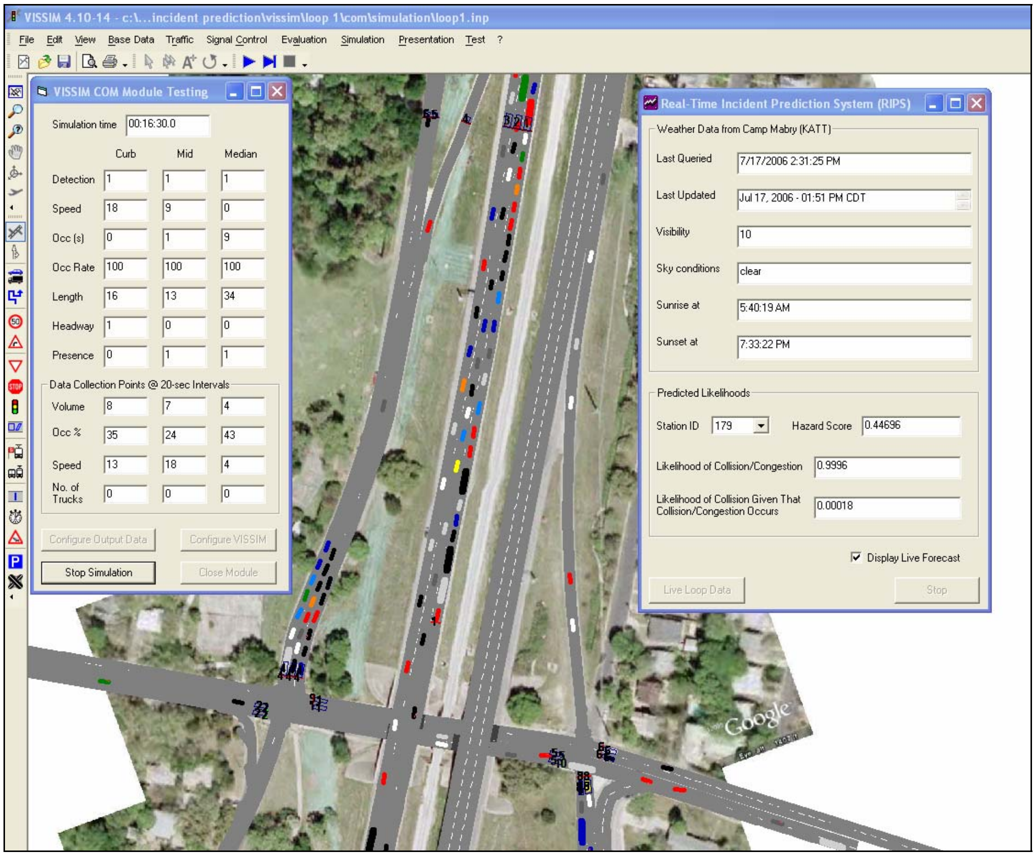


Figure 31. During-Incident Traffic Conditions.

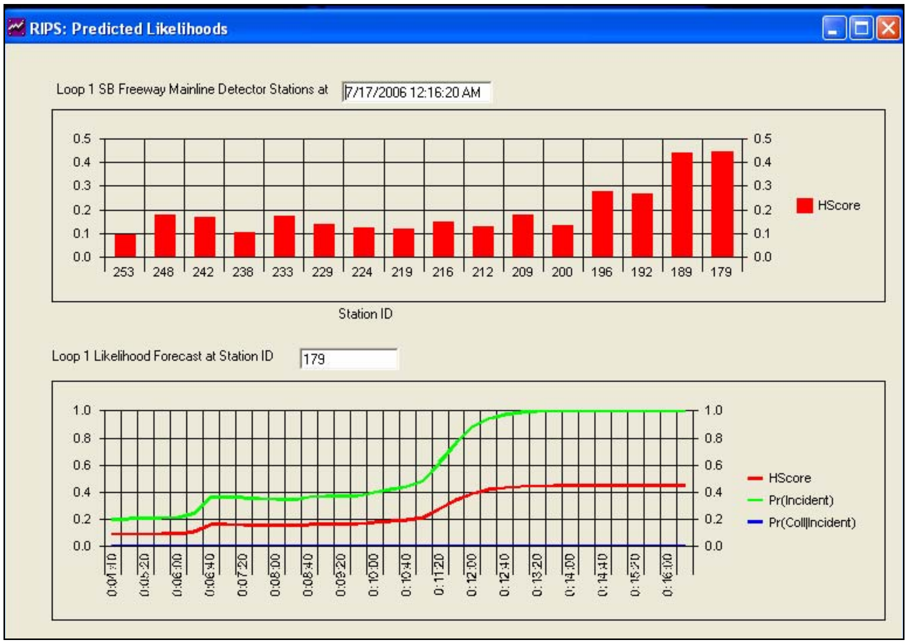


Figure 32. Example Model Outputs During-Incident Traffic Conditions.

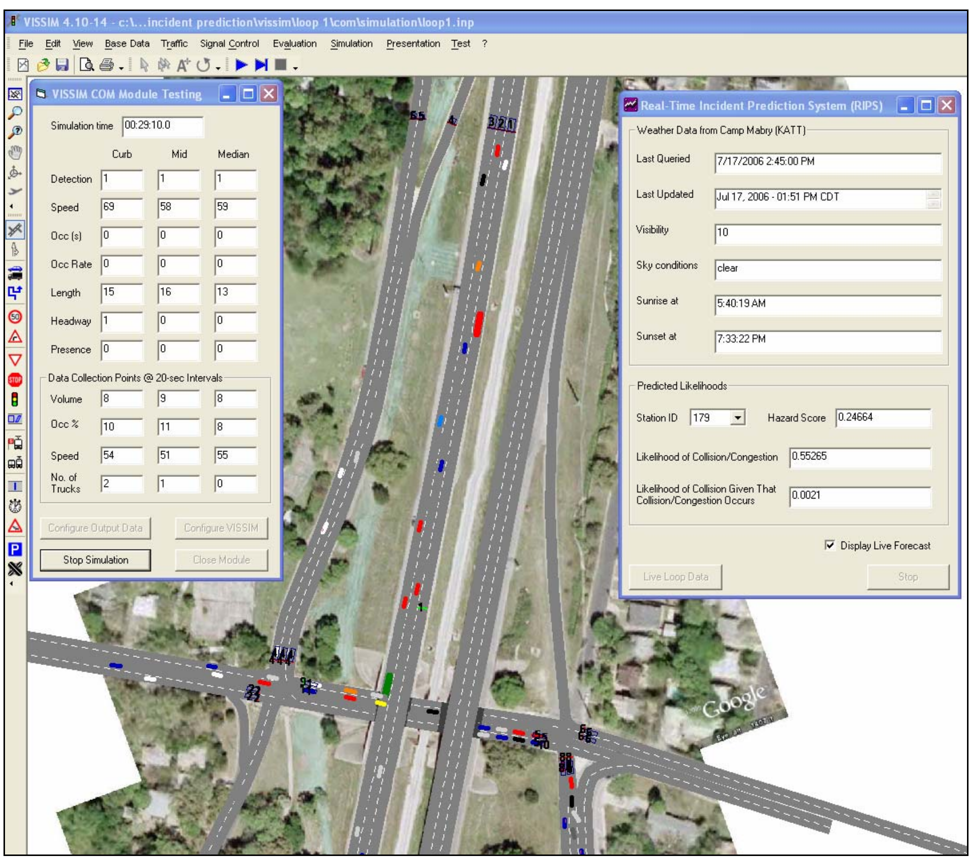


Figure 33. Post-Incident Traffic Conditions.

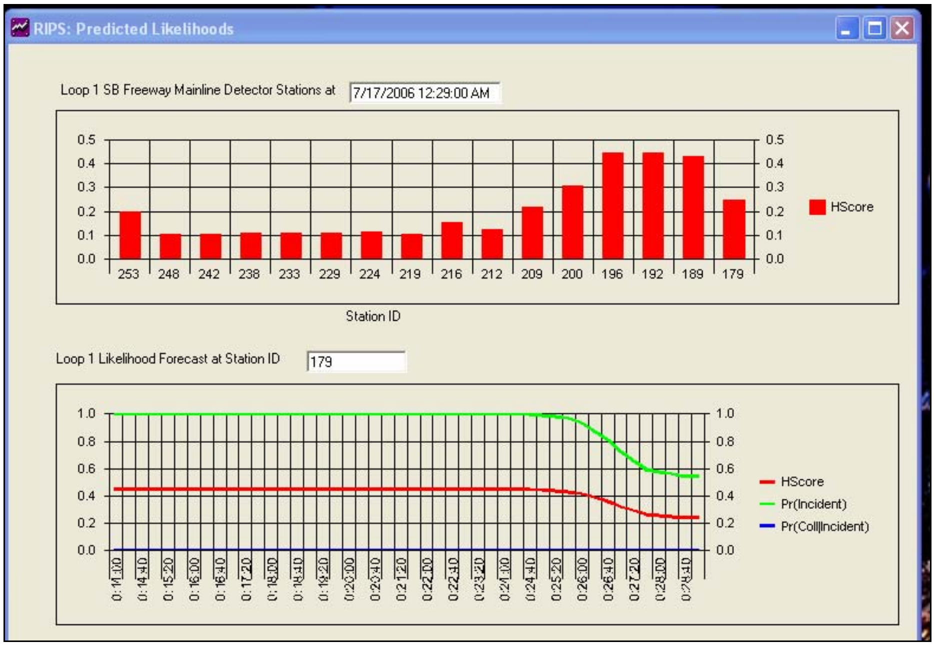


Figure 34. Example Model Outputs for Post-Incident Traffic Conditions.

ISSUES AFFECTING IMPLEMENTATION

Unlike simulation data, real data are seldom perfect. To develop and implement an effective model for deployment in traffic management systems, it is crucial to have reasonably accurate data on volume, occupancy, and speed.

After analyzing Austin data, we have concerns about two problems, namely missing data and erroneous data. Instances of missing data include negative entries (-1) during certain 1-minute periods and entries of zero (0) for several consecutive time periods. Erroneous data arise when the detector registers wrong data. These two types of problems found in the Austin data are discussed below.

MISSING DATA

Missing data are a prevalent problem for Austin freeway systems. When examining the data, we found that a significant number of detectors are inoperative, as of November 14, 2005. Numerous detector stations on US 183 and Loop 1 register values of 0 or -1 all the time. For example, on the southbound direction of Loop 1, all three detectors of the Westover Road station report -1 for all the fields (Table 1), while the detectors of the ISD station (the next detector station downstream) report 0 for all the fields all the time (Table 2).

Table 1. Detectors of Westover Road Station on Southbound Loop 1.

Time Stamp		Detector ID	V	O	S	T	Detector ID	V	O	S	T	Detector ID	V	O	S	T	
180024	…	6002821	-1	-1	-1	-1	6002822	-1	-1	-1	-1	6002823	-1	-1	-1	-1	…
180124	…	6002821	-1	-1	-1	-1	6002822	-1	-1	-1	-1	6002823	-1	-1	-1	-1	…
180224	…	6002821	-1	-1	-1	-1	6002822	-1	-1	-1	-1	6002823	-1	-1	-1	-1	…
180325	…	6002821	-1	-1	-1	-1	6002822	-1	-1	-1	-1	6002823	-1	-1	-1	-1	…
180424	…	6002821	-1	-1	-1	-1	6002822	-1	-1	-1	-1	6002823	-1	-1	-1	-1	…
180524	…	6002821	-1	-1	-1	-1	6002822	-1	-1	-1	-1	6002823	-1	-1	-1	-1	…

Note: V=Volume, O=Occupancy, S=Speed, T= Percent Trucks

Table 2. Detectors of ISD Station on Southbound of Loop 1.

Time Stamp		Detector ID	V	O	S	T	Detector ID	V	O	S	T	Detector ID	V	O	S	T	
180024	…	6003811	0	0	0	0	6003812	0	0	0	0	6003813	0	0	0	0	…
180124	…	6003811	0	0	0	0	6003812	0	0	0	0	6003813	0	0	0	0	…
180224	…	6003811	0	0	0	0	6003812	0	0	0	0	6003813	0	0	0	0	…
180325	…	6003811	0	0	0	0	6003812	0	0	0	0	6003813	0	0	0	0	…
180024	…	6003811	0	0	0	0	6003812	0	0	0	0	6003813	0	0	0	0	…
180124	…	6003811	0	0	0	0	6003812	0	0	0	0	6003813	0	0	0	0	…

Note: V=Volume; O=Occupancy, S=Speed, and T=Percent Trucks

Further examination of data in the southbound direction of Loop 1 from November 14, 2005, to November 16, 2005, suggested that detectors on Loop 1 may have an upper limit on occupancy for error checking because the reported occupancies are never more than 25 percent. Since peak period occupancy values on Loop 1 regularly exceed 25 percent, up to half of all observations during the peak period are missing. Examples are the detectors on the freeway section from Westchester to RM 2222 from around 7:30 a.m. to 9:00 a.m., and detectors on the southbound Loop 1 freeway section from RM 2222 to south of 45th Street.

ERRONEOUS DATA

As defined previously, erroneous data are incorrect data registered by the detector, LCU or SCU. Some types of erroneous data are easy to identify. For example, erroneous data arise when the detector reports positive occupancy and zero volume. Some detectors may occasionally report this type of erroneous data. But if a detector consistently outputs positive occupancy and zero volume, the detector must be examined. We found that detector 2011611 on Thunder Creek reported a significant number of cases of this type of erroneous data during 5:00 a.m. to 10:00 p.m. of the week of November 7, 2005 (ranging from 36 to 95 cases). The number of erroneous data reported by this detector warrants further examination.

Another type of erroneous data arises when a detector consistently reports constrained speed values. This type of erroneous data can be easily identified by simple time-of-day checks. For example, Figure 35 shows that the detectors on the Lazy Lane and Ohlen Road stations on northbound US 183 consistently report extremely slow speeds (less than 35 mph).

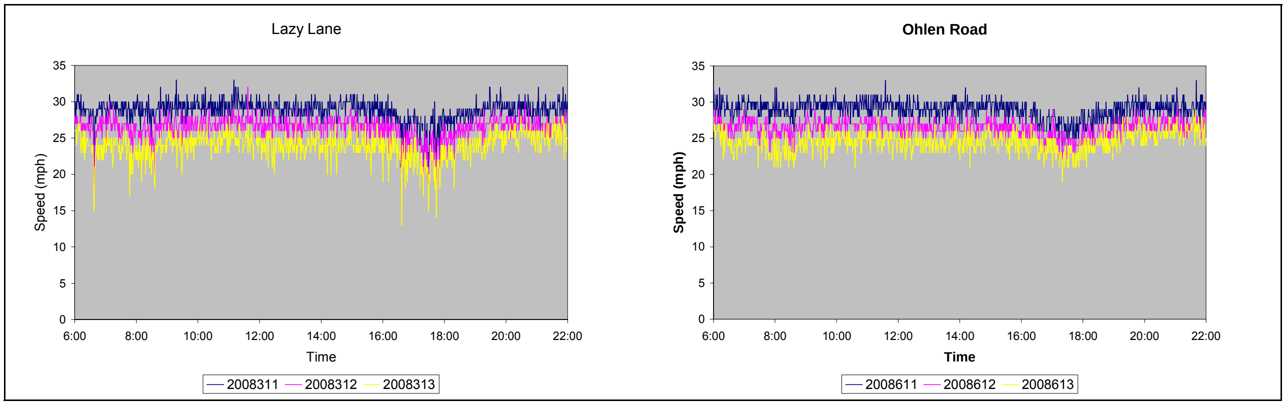


Figure 35. Constrained Speed Values.

However, in general, erroneous data are difficult to locate. For example, volume data collected from 2:00 p.m. to 6:00 p.m. on June 4, 2002 (a Tuesday), for the Guadalupe Street, Lamar Boulevard, and Lazy Lane stations on northbound US 183 did not reveal any abnormality. This site is illustrated in Figure 36.

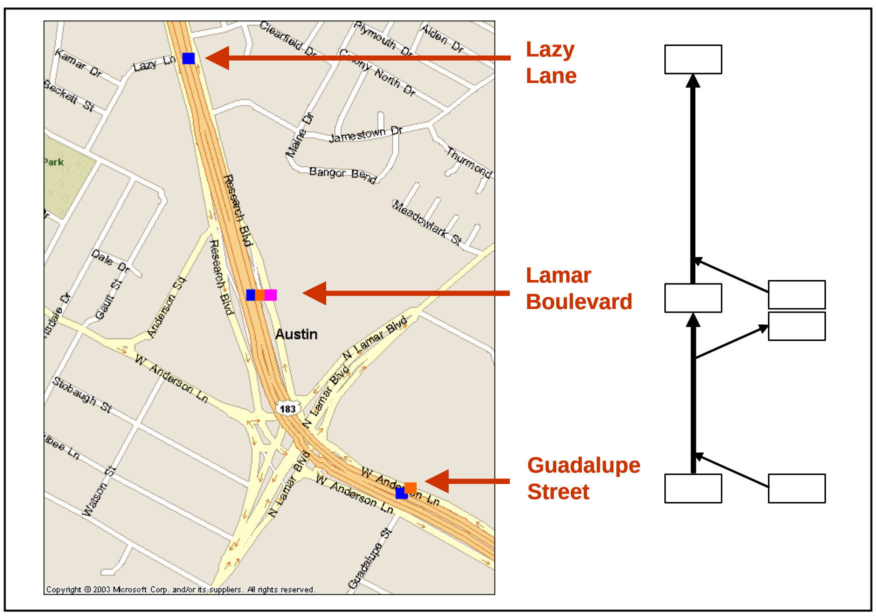


Figure 36. Study Site on US 183 in Austin, Texas.

When we applied the input-output model to the collected data, we were alarmed when examining the flow-in-process of the site. Figure 37 presents the flow-in-process and delayed-flow plots for this site. In the section from Lamar Boulevard to Lazy Lane, both flow-in-process and delayed-flow increased steadily at an approximate rate of 100 vehicles per hour from 2:00 p.m. to 5:00 p.m., and jumped to around 500 vehicles from 5:00 p.m. to 6:00 p.m. Recall that flow-in-process is defined as the number of vehicles that are traveling in the section at any particular time; it is improbable to have 500 vehicles in a half-mile three-lane section simultaneously.

For the upstream section from Guadalupe Street to Lamar Boulevard, the flow-in-process and delayed-flow demonstrate a horizontal trend from 2:00 p.m. to 4:00 p.m. However, from around 4:15 p.m., both flow-in-process and delayed-flow dropped significantly, resulting in negative values. In other words, more vehicles are leaving the section than vehicles entering the section, a logically impossible scenario.

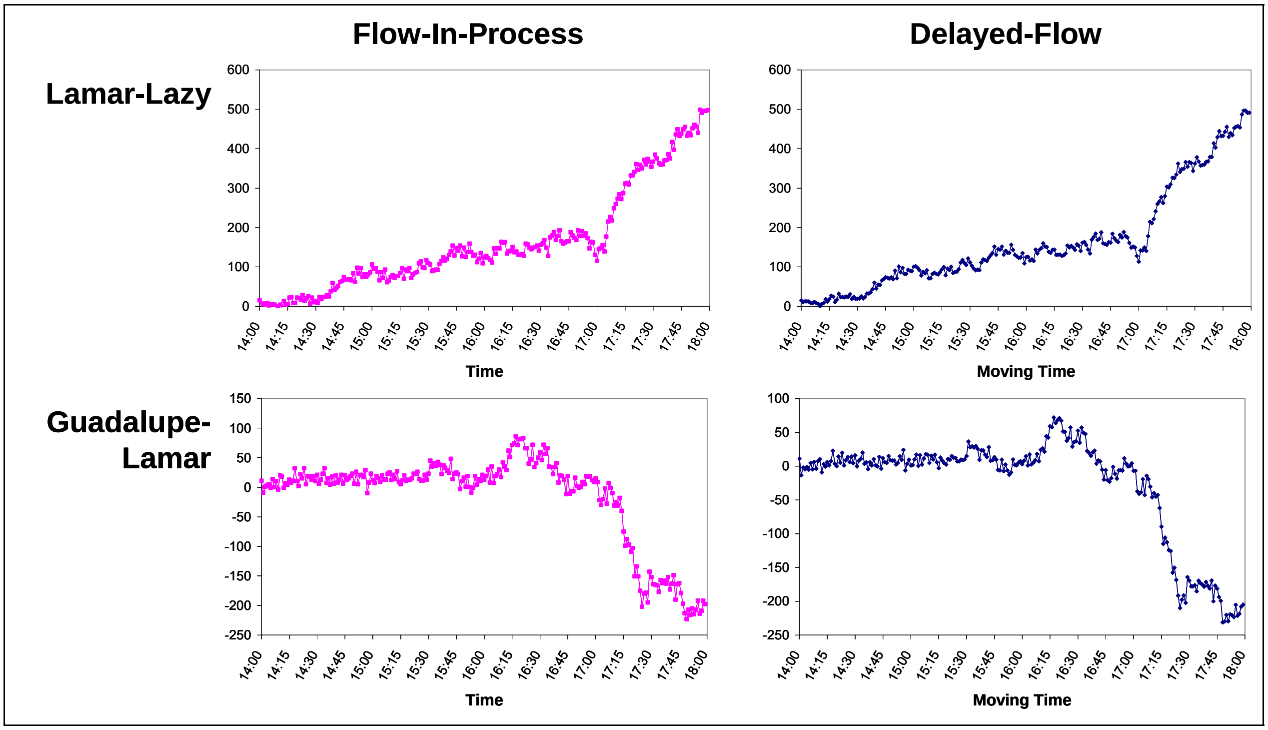


Figure 37. Flow-in-Process and Delayed-Flow of Guadalupe-Lamar-Lazy Site.

The above analysis revealed that the detector station at Lamar Boulevard recorded more vehicles than those at the Guadalupe Street station and Lazy Lane station. It is perhaps due to faulty detectors on Lamar Boulevard only. Thus, we ignored the data from the Lamar Boulevard station and plotted the flow-in-process and delayed-flow against time for the section Guadalupe-Lazy in Figure 38.

In Figure 38, both flow-in-process and delayed-flow show an upward trend, which indicates an incident that lasted a couple of hours. However, there is no incident logged during that period. Therefore, we suspect that the detectors may not provide accurate volume counts.

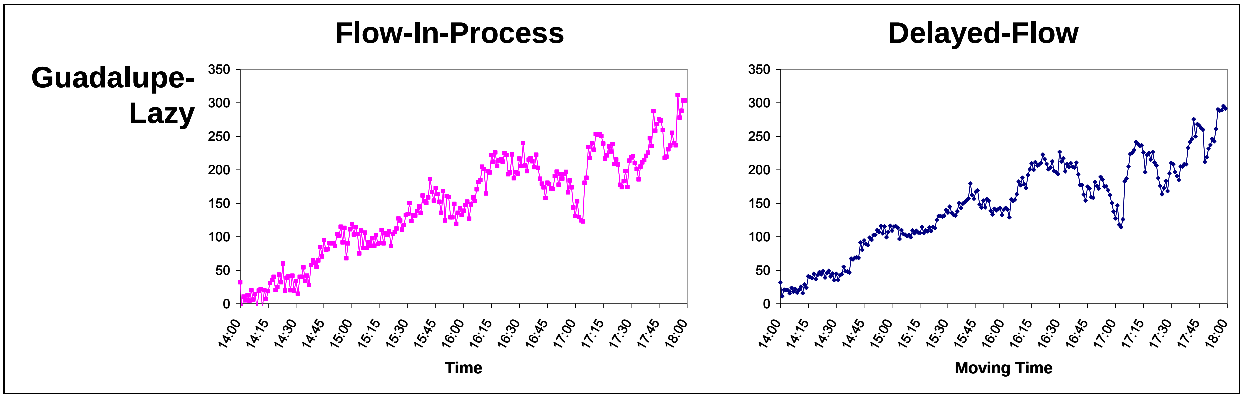


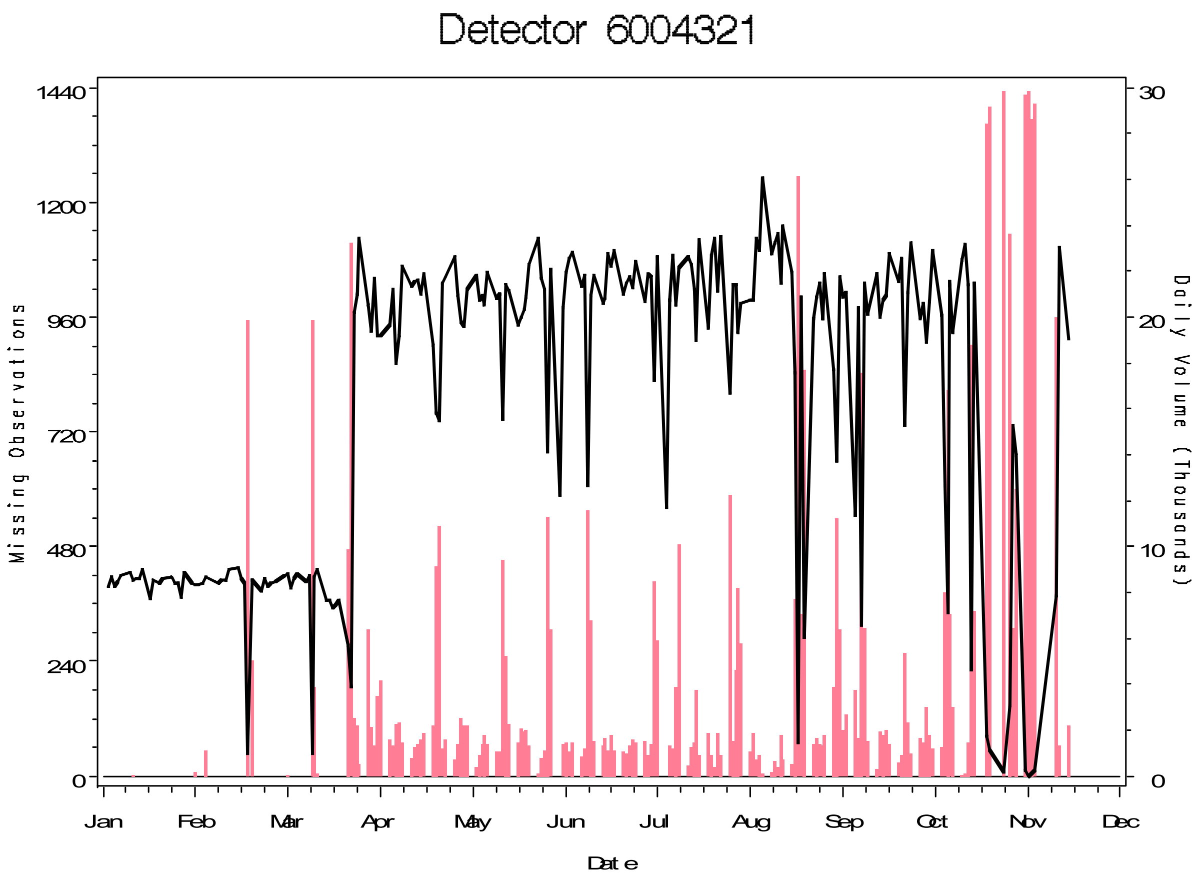
Figure 38. Flow-in-Process and Delayed-Flow of Guadalupe-Lazy Site.

Also, we observed that many detectors on Loop 1 started to have different patterns in total daily volumes in March 2005, which may be the result of a SCU software update or problem solving by TxDOT. Data collected by detector 6004321 at the 45th Street station on southbound Loop 1 exemplify this pattern shift as shown in Figure 39. Because of this pattern shift, we must be careful when using the historic data.

Since erroneous data appear to be valid and difficult to identify in general, erroneous data may have far more significant implications than missing data on both model development and implementation.

CONCLUDING REMARKS

Accurate data are a crucial key to the success of developing and implementing a congestion and incident prediction system. On one hand, a model that is developed based on inaccurate past data may not reflect the real environment and may render the results useless. On the other hand, a valid model may provide invalid results if the real-time data are faulty. Therefore, it is crucial that further investigation be carried out to identify the causes of these errors and that steps be taken to correct these problems.



Daily Traffic Volume   Missing Observations

Figure 39. Daily Volumes and Missing Observations.

SUMMARY AND LESSONS LEARNED

SUMMARY

Historical data can be a valuable resource for traffic management center operators. This project showed that historical traffic condition information, coupled with information from other sources, such as weather information, can be used to generate models that operators can use to produce short-term forecasts of traffic conditions. By tying these models to real-time traffic and weather condition information, operators can use these models to identify where incidents and congestion have the potential to occur based on the current travel conditions.

In this research project, we examined several different strategies and techniques for developing short-term forecasts (i.e., up to 15 minutes into the future) of where traffic congestion and incidents were likely to form on a freeway network using real-time traffic and weather information. As part of this research, we developed and incorporated four different prediction models for forecasting when and where incidents and traffic congestion were likely to occur. In the first model, the Cumulative Flow Forecast Model, we used traffic condition data from two adjacent detector stations to forecast the flow-in-process level of service and delay-in-flow level of service over a span of 15 minutes in the designated target section of the freeway. We used the Speed-Density Forecast Model, which uses standard moving-average techniques to forecast traffic conditions, to develop 15-minute forecasts of traffic conditions at a single detector station based on observed historical trends. In the Incident Prediction Model, we developed a model to predict the probability that an incident would occur, the probability that the incident would be a collision, and a hazard score for a given station in the designated target section of the freeway, based on current weather and traffic flow conditions as predictors. With the Crash Potential Model, we applied standard accident rate forecasting techniques to predict the likelihood that current and forecasted traffic conditions might result in a collision.

All four models were then integrated in a prototype tool, called the Dynamic Congestion and Incident Prediction System. The DCIPS was designed to be installed in a TxDOT traffic management center (TMC) connected in real time to the System Control Unit, which provides 1-minute volume, speed, and occupancy information collected from detector stations on the freeway. The DCIPS system uses this information along with current visibility and weather data to develop forecasts of locations where incidents and congestion are likely to occur. The prototype tool was tested in a hardware-in-the-loop simulation environment in the TTI TransLink® Research Center laboratory.

LESSONS LEARNED

The following lists several of the lessons that we learned as part of this research project:

- This research showed that using historical information to develop forecasts of potential traffic conditions was possible. This research activity, however, was hampered by not having good quality data. TxDOT needs to develop and employ procedures for ensuring the quality of the data that are stored in their databases. Good quality data are not only essential to producing forecasting models, but they are also critical to supporting planning activities for operations. The quality of the data feeding the process has an impact on the quality of the results of the forecasting tools.
- This research also showed that it was possible to "tap into" the TxDOT ATMS data stream and extract data in real time without slowing down or interfering with the core ATMS software. By tapping into the SCU data stream, TxDOT can develop multiple applications that utilize the same data stream without interfering with the operator's ability to assess traffic performance and implement control strategies.

FUTURE RESEARCH ACTIVITIES

The following is a list of future research and implementation activities that need to be performed if TxDOT wants to continue development of prediction and forecasting tools:

- TxDOT currently is working on techniques to identify erroneous and inconsistent traffic data before they get stored in databases. This should significantly improve the quality of the data. After this occurs, we recommend to TxDOT that they recalibrate the models developed as part of this research activity to ensure that they can accurately represent traffic conditions.
- TxDOT needs to evaluate how their operators might use and react to forecasted information. For example, TxDOT needs to determine what action(s) is(are) appropriate for an operator to take if he or she is alerted to the potential of an accident or congestion occurring at a location. Furthermore, TxDOT needs to develop tools and techniques for evaluating the impacts of an operator taking action based on

predicted information to determine if the potential action would have a positive or negative impact on the predicted conditions.

- Much research is currently underway at many locations throughout the United States and elsewhere to develop traffic forecasting models and tools—some of which incorporate real-time measures of traffic conditions into the modeling process. Models such as DYNAMIT, DYNASMART, and others have the potential to provide forecasted traffic conditions based on a dynamic traffic assignment model. TxDOT should continue to monitor the development of these tools and begin integrating and incorporating them as a traffic management tool in their control centers.

REFERENCES

1. K. Balke, N. Chaudhary, C. Chu, S. Kuchangi, P. Nelson, P. Songchitruksa, D. Swaroop, and V. Tyagi. *Dynamic Traffic Flow Modeling for Incident Detection and Short-Term Congestion Prediction: Year 1 Progress Report*. Report No. FHWA/TX-06/0-4946-1, Texas Transportation Institute, The Texas A&M University System, College Station, Texas, September 2005.

2. G. Newell. "A Simplified Theory of Kinematic Waves in Highway Traffic, Part II: Queueing at Freeway Bottlenecks." *Transportation Research,* Part B, Vol. 27, 1993, pp. 289-303.

3. M. J. Cassidy and J. R. Windover. "Methodology for Assessing Dynamics of Freeway Traffic Flow." *Transportation Research Record 1484*, 1995, pp. 73-79.

4. *Highway Capacity Manual*. Transportation Research Board, National Research Council, Washington, D.C., 2000.

5. C. Oh, J. S. Oh, S. G. Ritchie, and M. Chang. *Real-Time Estimation of Freeway Accident Likelihood*. Technical Report, Institute of Transportation Studies, University of California, Irvine, December 2000.

6. C. Lee, F. Saccomanno, and B. Hellinga. "Analysis of Crash Precursors on Instrumented Freeways." *Transportation Research Record 1784*, 2002, pp. 1-8.

7. C. Lee, F. Saccomanno, and B. Hellinga. "Real-Time Crash Prediction Model for Application to Crash Prevention in Freeway Traffic." *Transportation Research Record 1840*, 2003, pp. 67-77.

8. M. Abdel-Aty, A. Pande, N. Uddin, H. Al-Deek, and E. Radwan. *Linking Crash Patterns to ITS-Related Archived Data*. Final Report, University of Central Florida, Orlando, 2004.

9. K. M. Kockelman and J. Ma. "Freeway Speeds and Speed Variations Preceding Crashes, within and across Lanes." Presented at ITS America 2004, 14th Annual Meeting and Exposition, San Antonio, Texas, 2004.

APPENDIX A: STRUCTURE OF DATABASE TABLES IN DCIPS

The following items documents the structure of the database tables used in the development of the DCIPS.

1. **TrafficData**—Contains real-time input data provided by the SCU Communications Manager.

Field Name	Data Type	Field Size	Description
det_id	Text	8	Unique detector identification number
volume	Number	Integer	1-minute volume
occupancy	Number	Integer	Occupancy, 0-100%
speed	Number	Integer	Average speed, mph
percenttrucks	Number	Integer	Average truck percentage, 0-100%
ftime	Date/Time		Time stamp of last data record read
VolBadCount	Number	Integer	Keeps track of the number of times bad volume data were encountered and corrected with the previous value for each respective detector
SpdBadCount	Number	Integer	Keeps track of the number of times bad speed data were encountered and corrected with the previous value for each respective detector
OccBadCount	Number	Integer	Keeps track of the number of times bad occupancy data were encountered and corrected with the previous value for each respective detector
TruckBadCount	Number	Integer	Keeps track of the number of times bad percent trucks data were encountered and corrected with the previous value for each respective detector

2. **Station**—Contains the relationship between each station and detectors, i.e., the list of detectors for each station.

Field Name	Data Type	Field Size	Description
station_id	Text	3	Unique station identification number
det_id	Text	7	Unique detector identification number

3. **StationInfo**—Contains detailed information for each station.

Field Name	Data Type	Field Size	Description
station_id	Text	3	Unique station identification number
station_desc	Text	15	Name of station
hwy	Text	5	Name of highway
direction	Text	1	Direction of traffic flow—N for north, S for south, E for east, W for west
seq	Number	Integer	Sequential order of station based on its location
type	Text	3	Type of station—ENT for entrance ramp, EXT for exit ramp, FWY for freeway

4. **Roadway**—Contains roadway information for each station.

Field Name	Data Type	Field Size	Description
station_id	Text	3	Unique station identification number
roadway_type	Number	Integer	Type of roadway—1 for curved/ramp influence or 0 for straight/no ramp influence

5. **Weather**—Contains weather data read from a web-based weather forecaster.

Field Name	Data Type	Field Size	Description
udate	Text	15	Date of update
utime	Text	15	Time of update
visibility	Text	25	The distance (in miles) that a person would be able to clearly identify an object.
skyconditions	Text	50	A description of the sky conditions (e.g., “clear”, “cloudy”, “partly cloudy” etc.)
sunrisetime	Text	10	Time of sunrise
sunsettime	Text	10	Time of sunset

6. **IncidentLikelihoodForecast**—Contains the forecasted output data from Incident Likelihood Forecaster (ILF).

Field Name	Data Type	Field Size	Description
station_id	Number	Integer	Unique station identification number
incidentprobability	Number	Double	Probability for an incident to occur
collisionprobability	Number	Double	Probability for a collision to occur
HScore	Number	Double	Hazard score
ftime	Date/Time		Time stamp of forecast

7. **CumulativeFlowForecast**—Contains the forecasted output data from Cumulative Flow Forecaster (CFF).

Field Name	Data Type	Field Size	Description
station_id	Text	3	Unique station identification number
FipLos	Text	4	Flow-in-process level of service for 1 minute
DfLos	Number	3	Delay-in-flow level of service for 1 minute
FipLos5min	Text	4	Flow-in-process level of service for 5 minutes
DfLos5min	Number	3	Delay-in-flow level of service for 5 minutes
FipLos10min	Text	4	Flow-in-process level of service for 10 minutes
DfLos10min	Number	3	Delay-in-flow level of service for 10 minutes
FipLos15min	Text	4	Flow-in-process level of service for 15 minutes
DfLos15min	Number	3	Delay-in-flow level of service for 15 minutes
ftime	Date/Time		Time stamp of forecast

8. **CrashPotentialForecast**—Contains the forecasted output data from Crash Rate Forecaster (CPF).

Field Name	Data Type	Field Size	Description
station_id	Text	3	Unique station identification number
crashpotential1min	Number	Double	Crash rate for 1 minute
crashpotential5min	Number	Double	Crash rate for 5 minutes
crashpotential10min	Number	Double	Crash rate for 10 minutes
crashpotential15min	Number	Double	Crash rate for 15 minutes
ftime	Date/Time		Time stamp of forecast

9. **SpeedDensityForecast**—Contains the forecasted output data from Speed Density Forecaster (SDF).

Field Name	Data Type	Field Size	Description
station_id	Text	3	Unique station identification number
density1min	Number	Integer	Density for 1 minute
speed1min	Number	Double	Speed for 1 minute
density5min	Number	Integer	Density for 5 minutes
speed5min	Number	Double	Speed for 5 minutes
density10min	Number	Integer	Density for 10 minutes
speed10min	Number	Double	Speed for 10 minutes
density15min	Number	Integer	Density for 15 minutes
speed15min	Number	Double	Speed for 15 minutes
ftime	Date/Time		Time stamp of forecast

Figure 40 below shows the entity relationship diagram (ERD) of the tables described above.

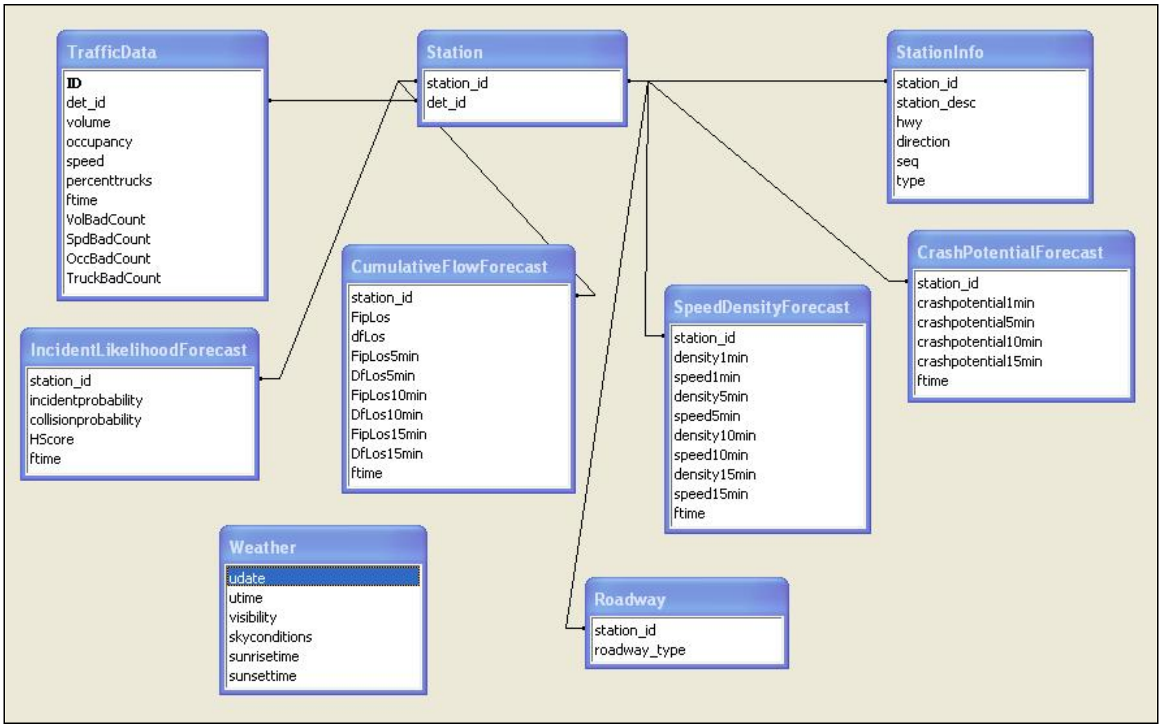


Figure 40. Entity Relationship Diagram between the Internal Databases in the DCIPS.

APPENDIX B: DEVELOPMENT OF INPUT-OUTPUT ANALYSIS OF CUMULATIVE FLOW MODEL

This section presents a simple methodology for evaluating the operating states of a freeway using cumulative flow data from pairs of adjacent detectors along the freeway. The methodology is based on the works of Newell (*1*) and Cassidy and Windover (*2*). By examining the system from downstream to upstream, the proposed methodology provides a means of evaluating the operating states of the freeway system.

BACKGROUND

We first provide an overview of cumulative flow and the moving-time coordinates (MTC) system, which are the foundations of the proposed methodology described in the next section. Readers interested in further details are referred to Newell (*1*) and Cassidy and Windover (*2*).

Cumulative Flow

Cumulative flow is the measure of the total number of vehicles passing over a detector from some referenced time (i.e., since 8:00 a.m.). Define:

$A(x,t)$ = cumulative number of vehicles that have passed detector location x by time t,

x_u = an upstream detector location,

x_d = a downstream detector location, and

y = freeway section of interest bounded by x_u and x_d.

For any section y, we can construct an input-output diagram by drawing the cumulative flow curves, $A(x_u, t)$ and $A(x_d, t)$, with respect to time as shown in Figure 41. This input-output diagram is a very useful tool for analyzing the characteristics of freeway traffic flow in section y. More specifically, the vertical distance between the two curves at some time t_j is the total number of vehicles in section y at that time. In this document, we define this quantity as flow-in-process of section y. Furthermore, the horizontal distance between the curves at height j represents the travel time of the jth vehicle from x_u to x_d.

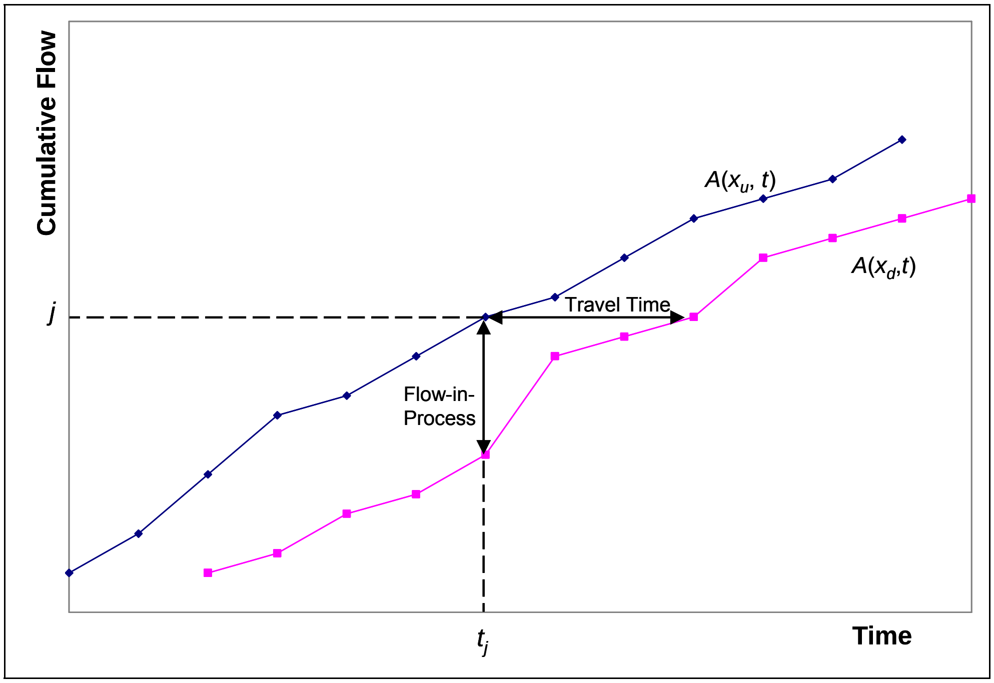


Figure 41. Input-Output Diagram.

Moving-Time Coordinates System

Even though the input-output diagram provides us with information on flow-in-process and travel time, additional information can be obtained by using an MTC system.

Define moving time, $t'(x,t)$, as follows:

$$t'(x,t) = t + v(x_d - x) \tag{B-1}$$

where t is the actual data collection time (i.e., 8:00 a.m., 8:01 a.m., 8:02 a.m., etc.) at x_u and v is the average free-flow travel time per unit distance (i.e., per foot, per meter, etc.) from x_u to x_d. We can now define new cumulative flow curves as follows:

$$A'(x_u,t') = A(x_u, t'\text{-v}(x_d\text{-}x_u)) \tag{B-2}$$

and

$$A'(x_d,t') = A(x_d, t'). \tag{B-3}$$

This transformation is equivalent to shifting the upstream curve $A(x_u,t)$ (in Figure 41) to the right by an amount equal to the free-flow travel time from x_u to x_d. In this MTC system, a vehicle traveling at v takes zero time to travel from x_u to x_d in the absence of delay inside section y. Figure 42 illustrates the cumulative curves of Figure 41 in an MTC system. As shown in

Figure 42, the vertical distance between the two curves at any time t_j represents current delay to vehicles in section y. In this study, we refer to this quantity as delayed-flow.

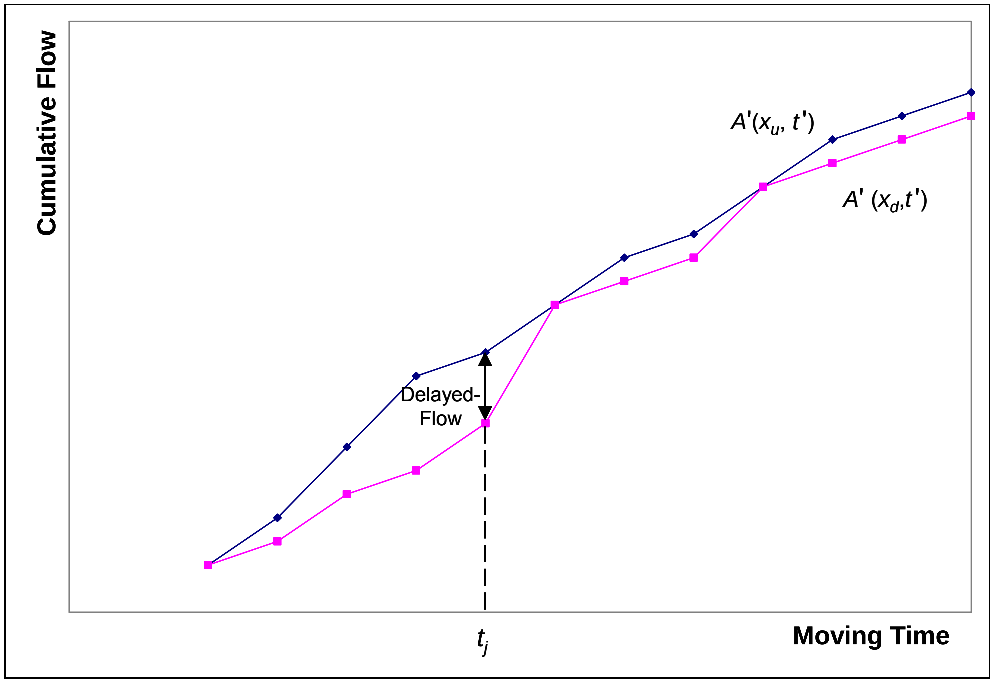


Figure 42. Input-Output Diagram Using Moving-Time Coordinates System.

PROPOSED METHODOLOGY

Characteristics of Flow-in-Process and Delayed-Flow

In this section, we discuss characteristics of flow-in-process and delayed-flow under different traffic conditions on a freeway section.

First, consider a perfect case where:

1. the vehicles arrive uniformly at x_u,
2. the vehicles' arrival rate at x_u is constant, and
3. the vehicles travel from x_u to x_d at the average free-flow speed.

In this case, the two cumulative curves will be parallel (Figure 43a), and the upstream cumulative curve will superimpose the downstream cumulative curve in the MTC system (Figure 43b). As a result, flow-in-process will be constant, and delayed-flow will be zero.

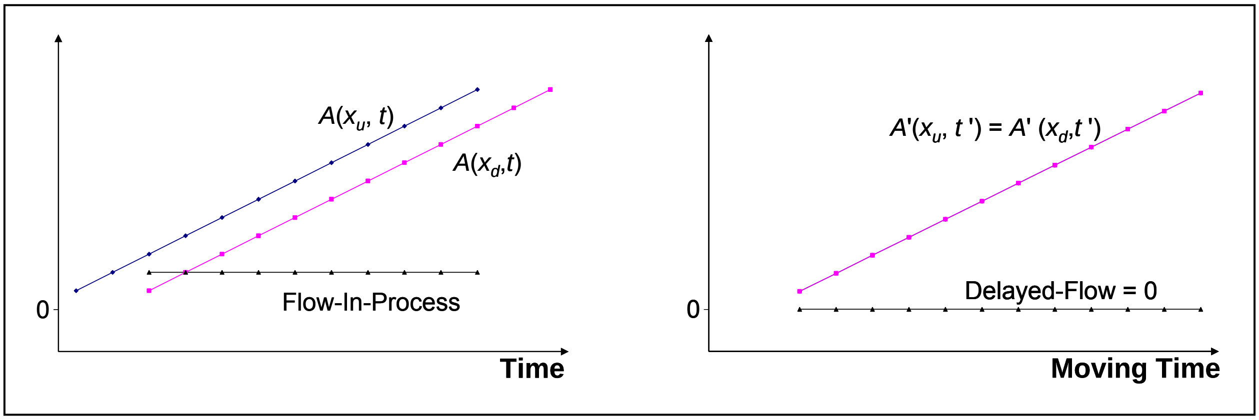


Figure 43. Constant Vehicle Flow.

Suppose there is an increase in traffic flow at time *t* while the travel time is not affected. In this case, there will be more vehicles traveling in the section starting at time *t*. As a result, there will be a jump in flow-in-process while delayed-flow will remain zero (Figure 44). Similarly, if traffic flow decreases at time *t* without affecting travel time, there will be a decrease in flow-in-process while delayed-flow will remain unchanged (Figure 45).

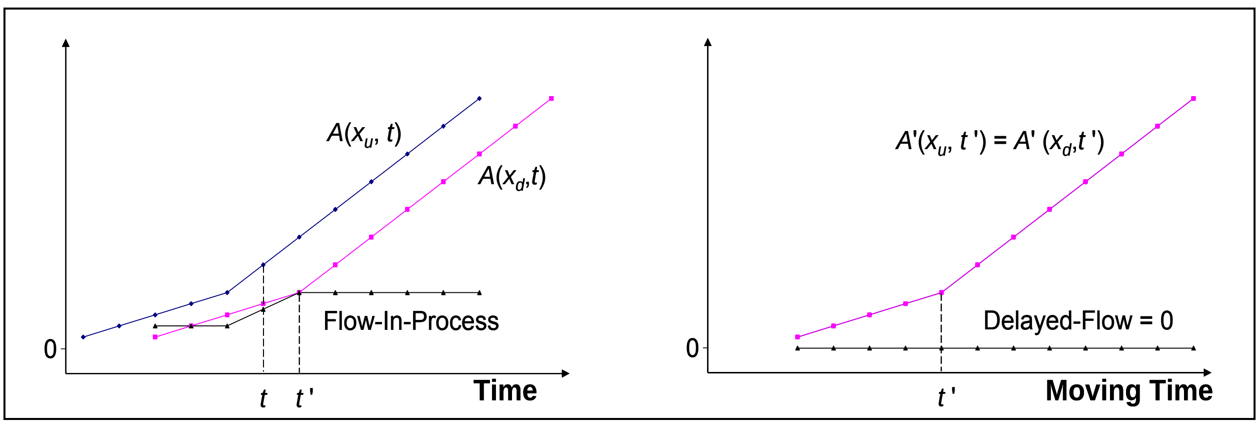


Figure 44. Increased Vehicle Flow.

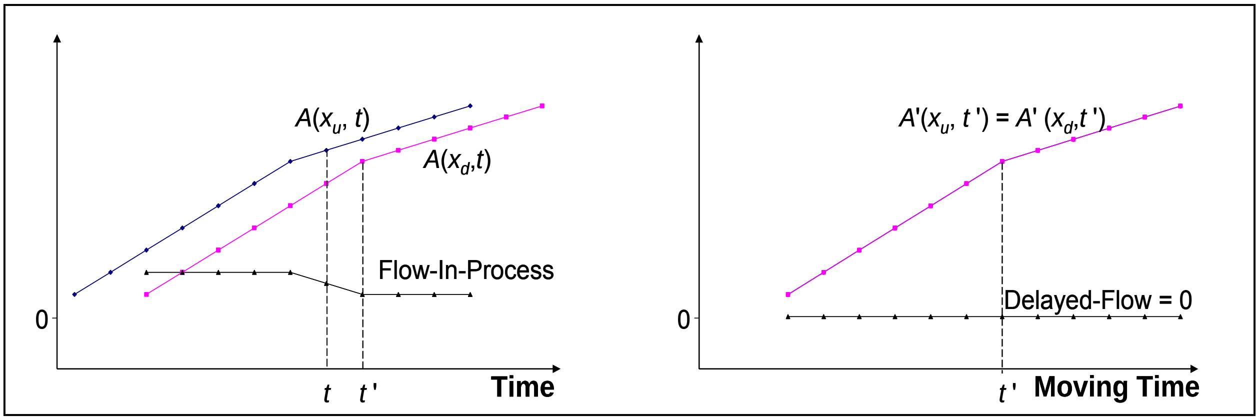


Figure 45. Decreased Vehicle Flow.

Now, suppose that an incident occurs in the section at time t_1' resulting in increased travel time between the two detector locations. In the MTC system, this situation will cause the cumulative flow from x_d to be lower than that from x_u, resulting in a positive delayed-flow. When the incident clears at some time $t_2' > t_1'$, the number of vehicles leaving the downstream detector location will increase, resulting in a decrease in flow-in-process and delayed-flow. Examples of the input-output diagrams and the corresponding flow-in-process and delayed-flow for the case described above are illustrated in Figure 46.

If the shockwave resulting from the incident reaches x_u, the operation of the adjacent upstream freeway will be compromised. This spillback of congestion into the upstream freeway section can be identified by doing a similar analysis for that section.

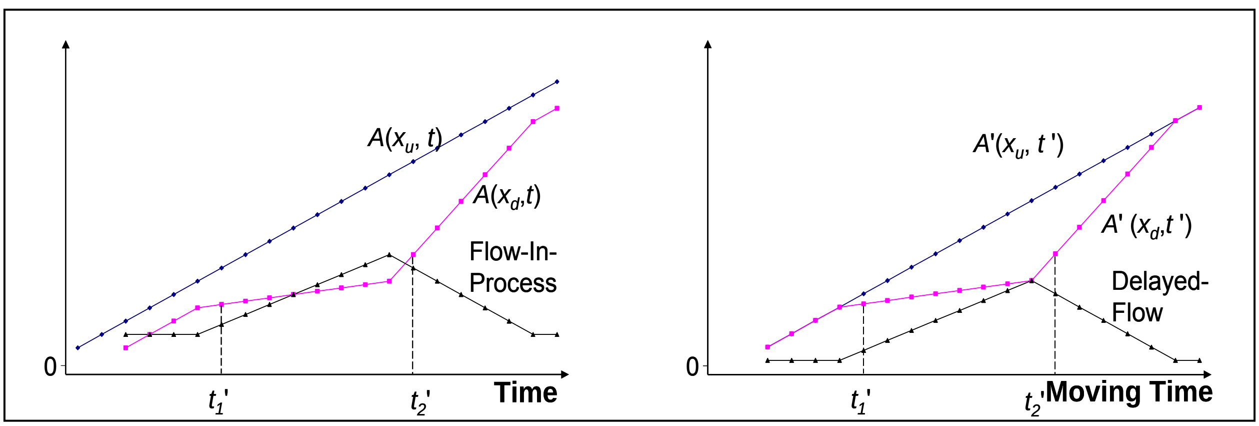


Figure 46. Incident at Freeway Section without Spillback.

Figure 47 illustrates the effects of spillback on the affected freeway sections. Figure 47a presents the input-output diagrams for section y_d (section where the incident occurs). Figure 47b shows the input-output diagrams for the upstream section y_u. In this example, the incident occurs at t_1', and the shockwave spills back to the upstream section y_u at t_1. The figure also shows that the incident clears (that is, flow-in-process at section y_d starts to decrease) at t_2', while the effects of the incident in the upstream section start to subside at $t_2 > t_2$'.

Input-output diagrams of section y_u (Figure 47b) are similar to those in Figure 46, indicating the effect of the incident does not reach the section further upstream. Should the shockwave reach the upstream section of y_u, Figure 47b will be similar to Figure 47a.

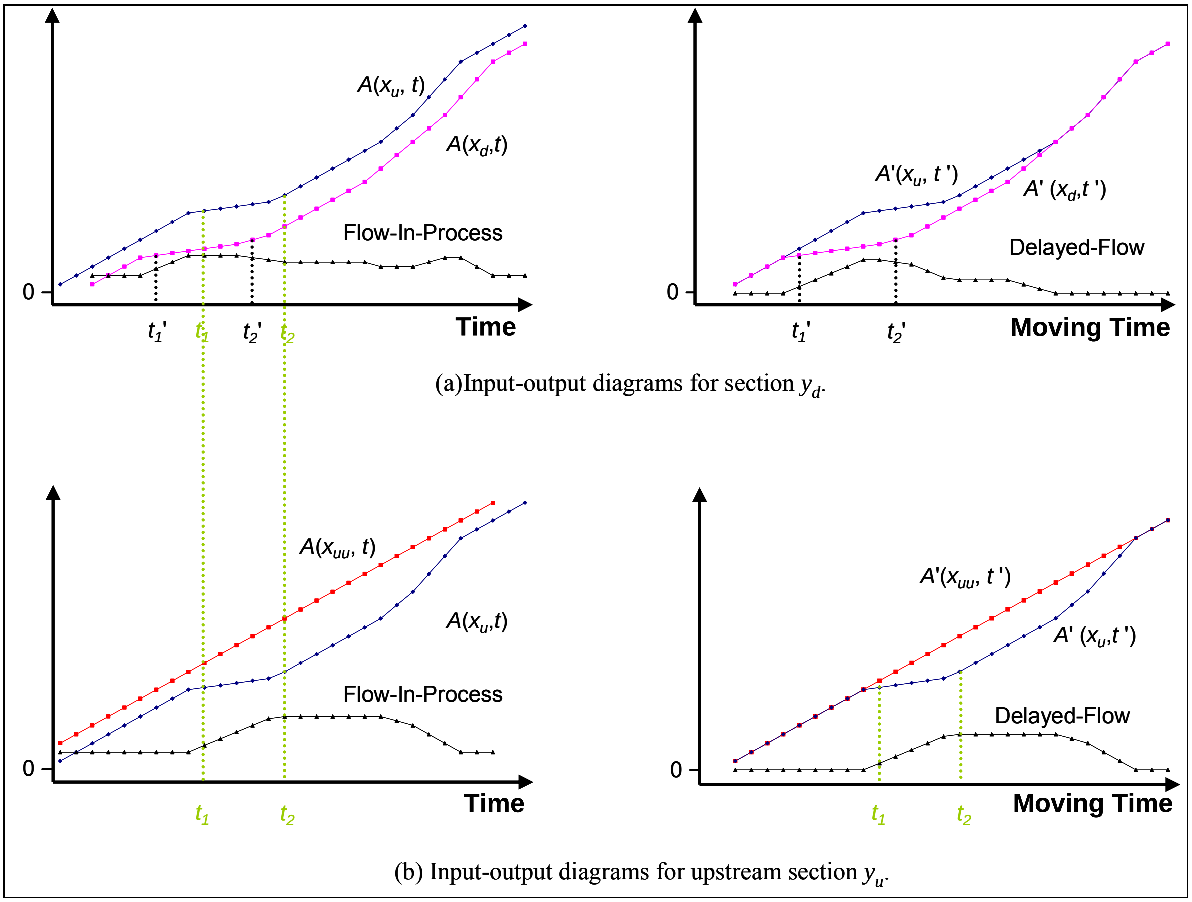


Figure 47. Incident at Freeway Section with Spillback.

Methodology

Based on the analysis, we proposed evaluating the current operating states of the freeway system by monitoring the flow-in-process and delayed-flow of each adjacent pair of detectors along the freeway starting from the farthest downstream section y_n and then back to the farthest upstream freeway section y_1 (Figure 48).

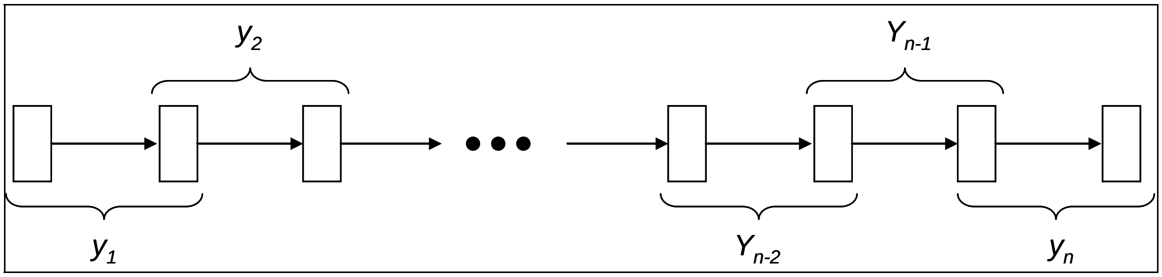


Figure 48. Freeway Detector System.

As discussed previously, the flow-in-process and delayed-flow can identify whenever there is a volume change on any freeway section. Therefore, these indicators can identify the start and end of peak-flow periods as well as incidents. In addition, analysis of areas under the elevated sections of a delayed-flow curve may quantify the amount of delay at each freeway section. However, this issue needs further investigation. For instance, if an incident occurs at freeway section y_i, monitoring delayed-flow for that section will identify the incident. Also, delayed-flow of upstream section y_{i-1} will identify when the shockwave from section y_i arrives at section y_{i-1}. Furthermore, the effects of the shockwave reaching farther upstream (i.e., section y_{i-2}) can be identified using data from that section. For this reason, our proposed methodology evaluates the system from the farthest downstream freeway section to the farthest upstream freeway section.

There are three primary advantages of this proposed methodology:

- It is scalable and can be applied to freeways with different sizes.
- Using the principle of flow conservation, it can be applied to freeway sections with different configurations (Figure 49).
- It depends on a limited amount of information, i.e., volume counts and free-flow speed.

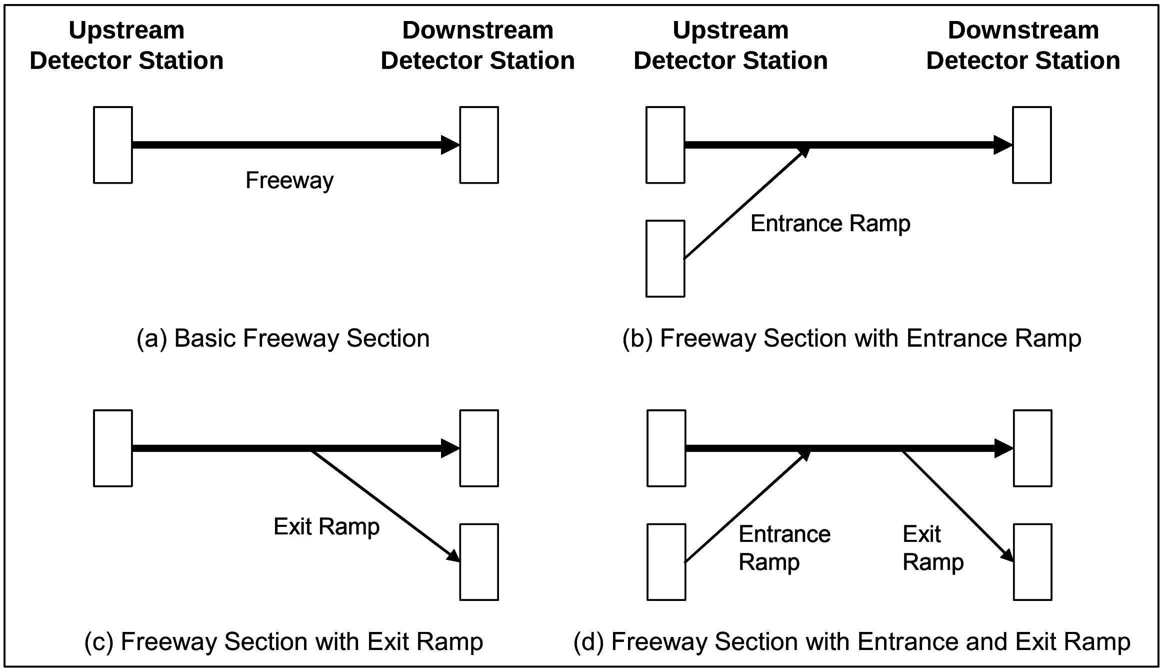


Figure 49. Different Freeway Section Configurations.

ILLUSTRATION USING SIMULATION

In this section, we illustrate the proposed methodology by applying it to count data obtained from computer simulation using VISSIM 4.00 (*3*). The simulated system is a three-lane freeway with four basic freeway sections in the study area (Figure 50). In addition, the freeway speed was uniformly distributed in the range of 55 to 75 vehicles per hour.

The simulation was divided into three time periods. Table 3 provides the durations and arrival rates for these periods. The arrival rate increases significantly after the first 15 minutes of simulation. The period of high demand lasts for 30 minutes, at which point the arrival rate drops to the initial rate. For the study described here, we conducted one 1-hour simulation without any incident and one 1-hour simulation with a single 10-minute incident, which started 1800 seconds into the simulation and blocked the middle lane only. As illustrated in Figure 50, the incident occurred at section y_3 immediately downstream of detector station x_3.

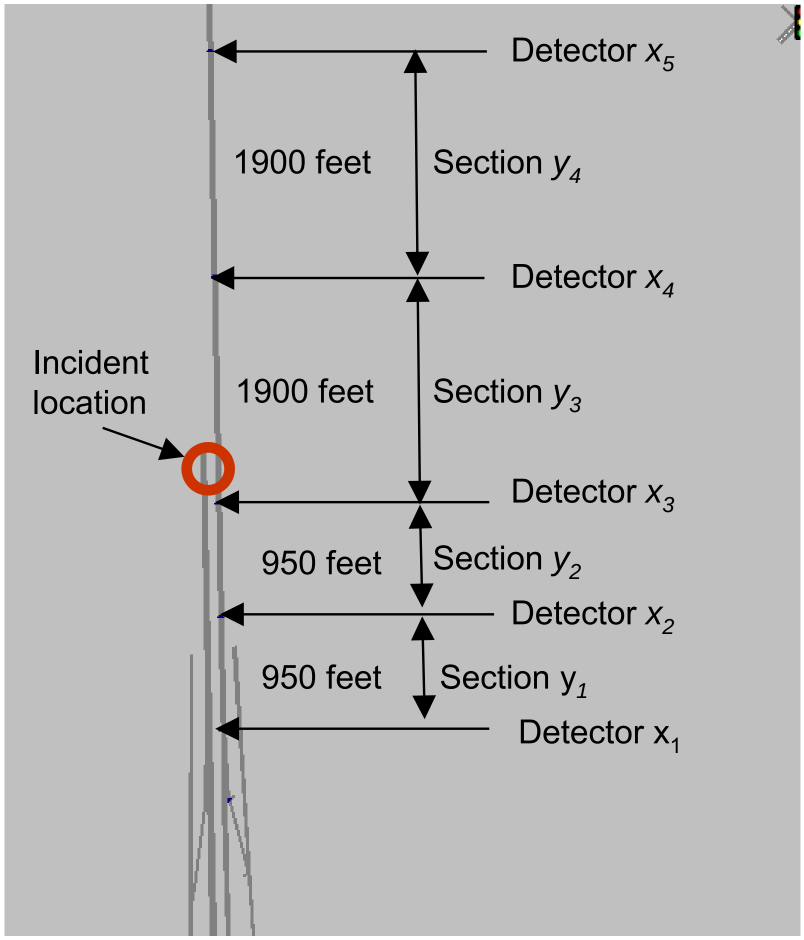


Figure 50. Sample Freeway System.

Figure 51 shows the flow-in-process and delayed-flow plots for the no-incident scenario. Here, the curves of flow-in-process and delayed-flow with respect to time are not piecewise linear as in the previous section. This is because vehicle arrivals are random as opposed to the uniform arrivals assumed for previous illustrations. Furthermore, individual vehicles travel at different speeds and not the average free-flow speed as assumed earlier. These characteristics are similar to that observed in the real world. Nevertheless, the curves of flow-in-process and delayed-flow can be approximated by piecewise linear curves to produce curves similar to those presented in the previous section.

In Figure 51 the flow-in-process for each section increased sharply some time after 900 seconds of simulation and decreased sharply some time after 2700 seconds of simulation time. In general, this result agrees with the arrival (demand) data presented in Table 3. The elevated portion of the flow-in-process curves does not start at 900 seconds but is slightly shifted to the right. This shift is equal to the travel time from the vehicle entry point into the system to the detector locations and shows when the increased traffic started impacting each section. Also note that the flow-in-process of sections y_3 and y_4 is larger than that of sections y_1 and y_2 during the period with increased traffic flow. In general, longer sections have larger flow-in-process, which is defined as vehicles traveling in the section at a particular time. Since section y_3 and y_4 are longer than sections y_1 and y_2, it is expected that the flow-in-process of sections y_3 and y_4 is larger than that of sections y_1 and y_2. Despite a significant increase in demand, however, delayed-flow remains approximately the same at all freeway sections during the entire simulation period. This is because no disruptions occurred on any section.

Table 3. Arrival Rate Distribution.

From (Seconds)	To (Seconds)	Arrival Rate (vph)
0	900	4000
900	2700	7500
2700	3600	4000

Figure 52 presents the flow-in-process and delayed-flow plots for the scenario with incident. The behaviors of flow-in-process and delayed-flow plots for each section are similar to those in Figure 51 until 1800 seconds. In other words, these plots detect an increase in traffic and no increase in delay. At around 1800 seconds, the delayed-flow of section y_3 increases significantly, which indicates an incident. Later, the delayed-flow and flow-in-process of upstream sections y_2 and y_1 also increase, while the delayed-flow of section y_3 does not drop. It serves as a signal that the shockwave created by the incident has reached the upstream sections.

More detailed information can be obtained by magnifying (Figure 53) the scale of plots for sections y_1, y_2, and y_3. Figure 53a shows how time lags in peaks of delayed-flow at adjacent sections identify the progression of the shockwave. Note that the delayed-flow of downstream section y_4 is not affected, while the flow-in-process decreases during the incident. This happens because fewer vehicles enter section y_4 during the incident.

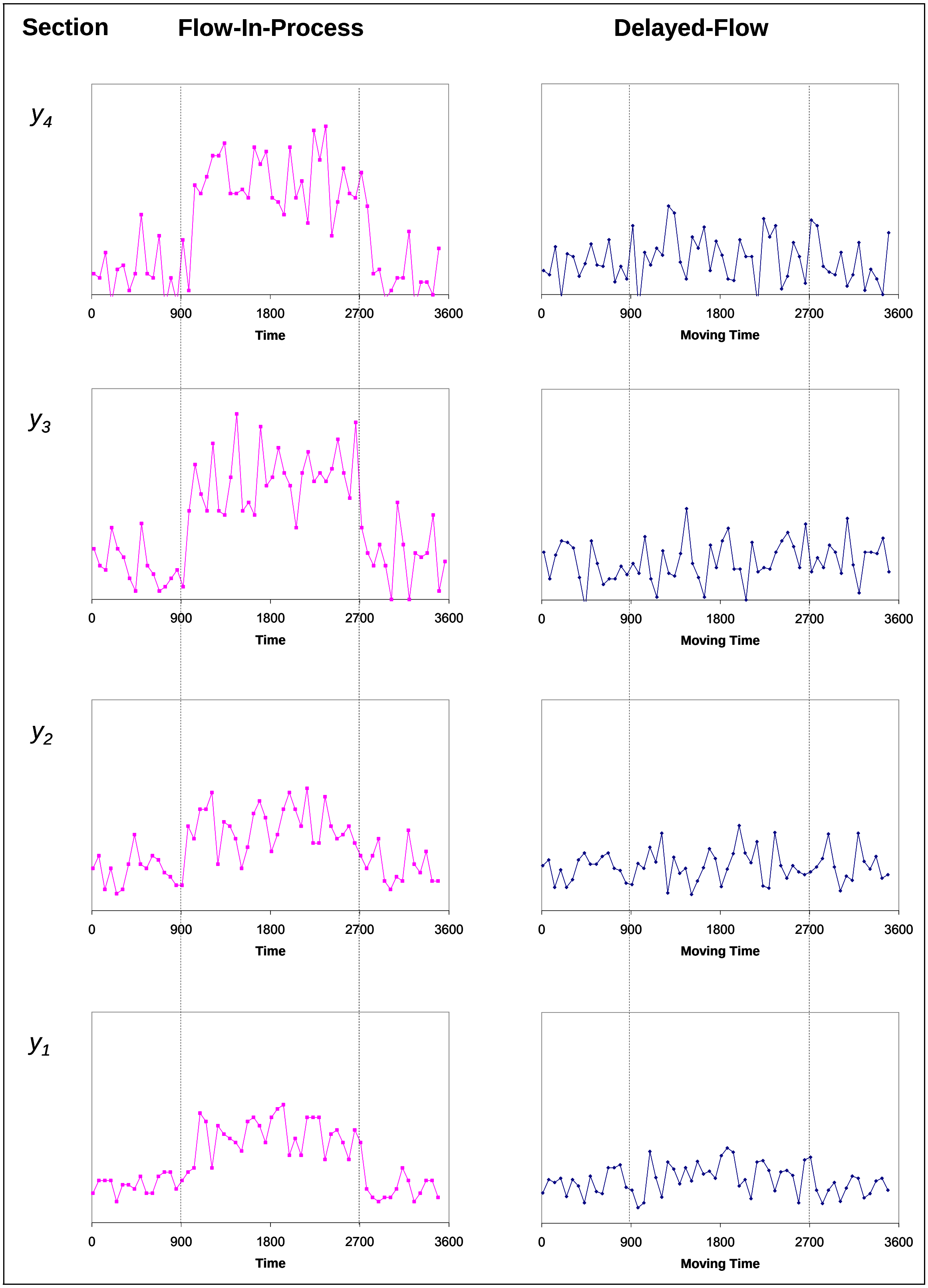


Figure 51. Scenario without Incident.

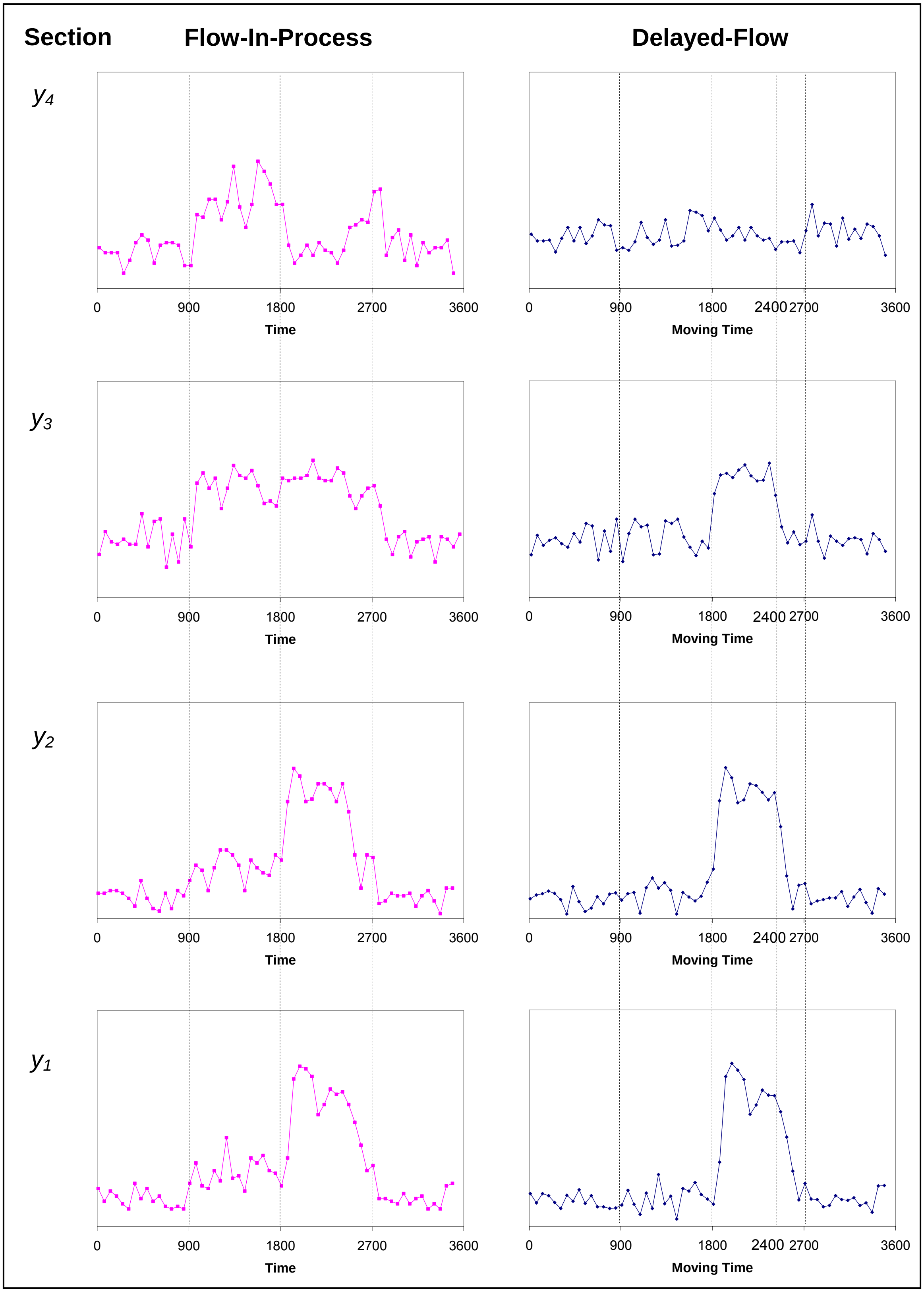


Figure 52. Scenario with Incident.

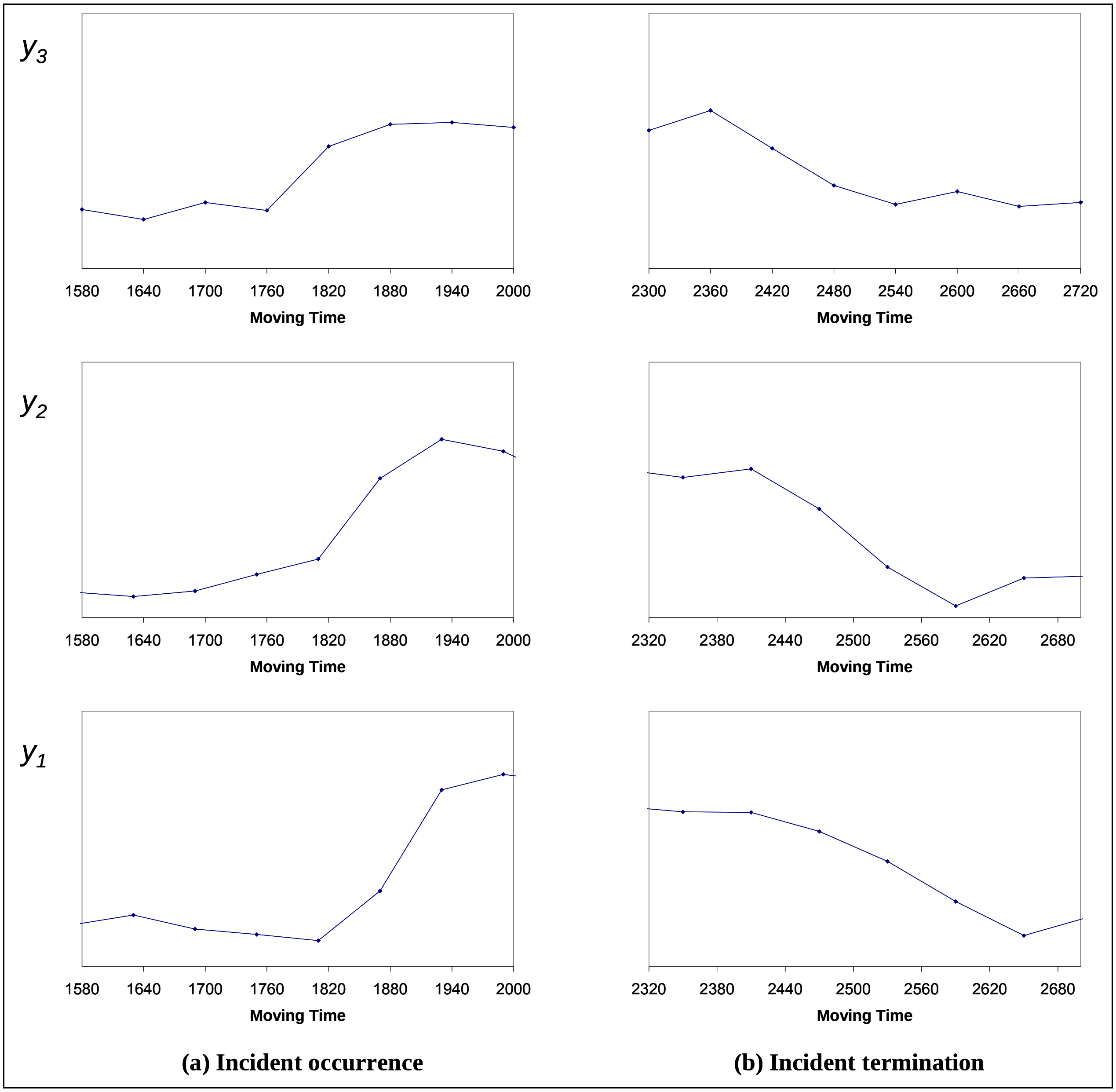


Figure 53. Delayed-Flow during Beginning and End of the Incident.

Similarly, Figure 53b shows time lags between the times at which congestion starts to clear at adjacent sections. Since the incident lasted for 600 seconds, one would expect the delayed-flow for the section with the incident to quickly drop to the level prior to the incident. This is the case for section y_3 at time 2400 seconds. The delayed-flow of section y_1 and y_2 does not drop to the expected levels until several minutes after the incident cleared. This is not a surprising result.

FREEWAY OPERATION STATUS PREDICTION

While the proposed methodology evaluates the current operating state of a freeway system, it serves the traffic control operator if a prediction mechanism is added to the methodology. As such, a simple moving average-based model is employed for predicting flow-in-process and delayed-flow from 1 minute to 15 minutes in advance.

Modified Moving-Average Model

Define:

m = number of moving-average periods,

$o(t)$ = observed value (flow-in-process or delayed-flow) at time t, and

$p(t, u)$ = predicted value (flow-in-process or delayed-flow) of time $t+u$ at time t.

Then,

$$p(t,u) = \begin{cases} \dfrac{\sum_{v=0}^{m-u} o(t-v)}{m}, & \text{if } u = 1; \\ \dfrac{\left(\sum_{v=0}^{m-u} o(t-v) + \sum_{v=1}^{u} p(t,v)\right)}{m}, & \text{if } u > 1. \end{cases} \qquad \text{(B-4)}$$

As shown in Equation B-4, the m-period moving-average model is used for forecasting. For one-period forecast, the predicted value is simply the m-period average of the past m observed value. For two-period forecast, $p(t,2)$, since the $o(t+1)$ is not available at time t, the predicted value $p(t,1)$ is used instead. As such, $p(t,u)$ will be calculated using predicted values only if $u>m$.

This model is tested using the results of the simulation illustrated in the previous section. Based on experiment results, the number of the moving average is chosen to be five periods since it generally provides the smallest average mean absolute errors for both predicted flow-in-process and delayed-flow.

For example, the 5-minute forecasts against time at section y_2 for the two scenarios (with and without incident) are shown in Figure 54 and Figure 55.

When there is no incident, the modified moving-average model is acceptable since the flow-in-process and delayed-flow are stable as in Figure 55. However, if an incident happens at time t, there is no indicator from the observed data at time t-5. Therefore, the pattern of the predicted flow-in-process and delayed-flow has an obvious 5-minute lag from the actual observed data.

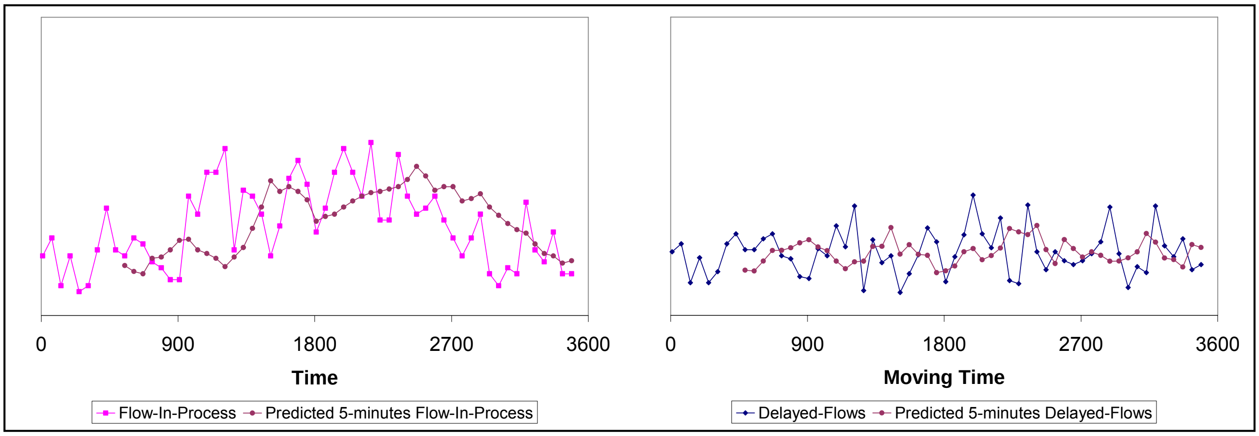


Figure 54. Five-Minute Forecasts at Section y_2 for Scenario without Incident.

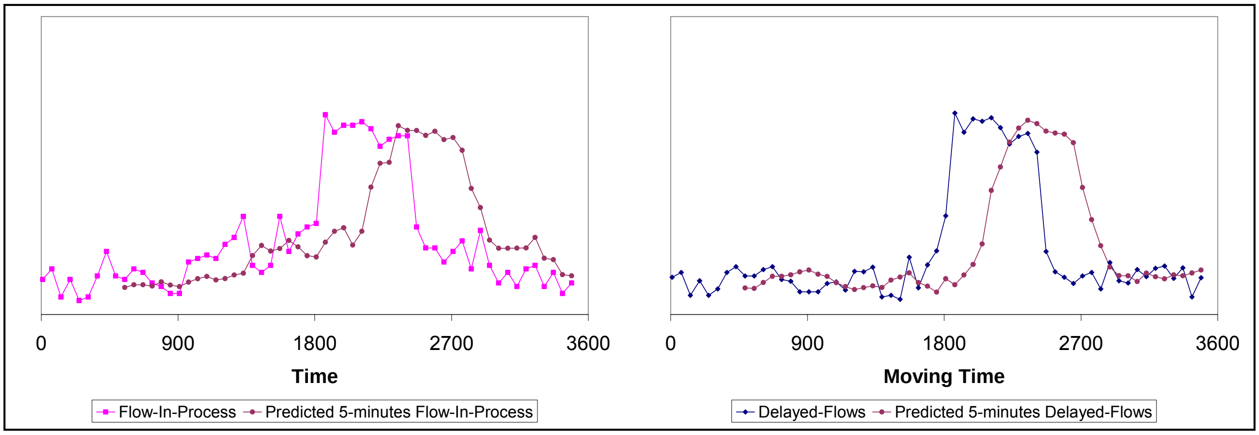


Figure 55. Five-Minute Forecasts at Section y_2 for Scenario with Incident.

Ideas for Developing Better Prediction Model

As shown in the previous subsection, the simple modified moving-average model is not effective in forecasting flow-in-process and delayed-flow when an incident happens. This result is partly because the forecast is based solely on the information of a single detector station and valuable information is ignored using this localized approach. For example, since the current traffic volume entering the farthest upstream freeway section y_1 (Figure 48) and all the entrance ramps are known, the entering and exiting volume of each downstream section can be estimated by projecting flow with approximate travel time. Another example is that if the delayed-flow of section y_3 increases significantly, it may be a signal of an incident. Therefore, the operations of the upstream section may be affected in a later time. As such, the forecast for flow-in-process and delayed-flow of the upstream section must take into consideration the possible shockwave.

We believe that a better prediction model can be developed by taking a system-wide approach for projecting flow from the upstream to downstream freeway section and then adjusting the predicted flow-in-process and delayed-flow from the downstream to upstream section.

INTERPRETATION OF INPUT-OUTPUT MODEL RESULTS

As discussed in the previous section, flow-in-process and delayed-flow are useful in detecting the operation state of the freeway. However, one may still need to perform a simple analysis to obtain knowledge about the operation status of the freeway given the nominal data on flow-in-process and delayed-flow. To free the operator from doing further analysis given the performance measures, the level of service of flow-in-process and the ratio of delayed-flow and flow-in-process (delayed ratio) are derived and shown instead of the nominal data in the user interface. The advantages of using these alternate measures are that they are easy to understand and that they immediately paint the general picture of the operation status of the freeway system.

Level of Services of Flow-in-Process

Recall that flow-in-process is the number of vehicles traveling in a freeway section in a particular time. Therefore, it can be viewed as a proxy for density. As such, flow-in-process is translated into the number of passenger cars per mile per lane, and the level-of-service thresholds proposed in Chapter 23 of the *Highway Capacity Manual* (*4*) are used for classifying the level of service of the flow-in-process.

Level-of-service thresholds shown in Chapter 23 of the *Highway Capacity Manual* (*4*)are reproduced in Table 4.

Table 4. Level-of-Service Thresholds from the *Highway Capacity Manual*.

Level of Service	Density Range (Flow-in-Process Range) in Passenger Car per Mile per Lane
A	0-11
B	>10-18
C	>18-26
D	>26-35
E	>35-45
F	>45

The advantage of using level of service is that it is easy to understand. However, subtle information is lost due to the level-of-service classification. For example, when traffic volume is increasing, the freeway section may show the same level of service for a couple of minutes before showing an inferior level of service. As such, a delay in detecting the changes in volume may result.

Delayed-Flow Ratio

The main purpose of delayed-flow is as an indicator for identifying the level of congestion of a particular freeway section. Therefore, a good indicator for delayed-flow is the ratio of the number of passenger cars being delayed to the number of passenger cars traveling in that section. If this ratio is high, a high percentage of vehicles are delayed and the operator should be notified. This indicator can be approximated by:

Delayed-Flow Ratio = 100 percent (Delayed-Flow/Flow-in-Process).

If a freeway section has a good quality of service based on flow-in-process, the delayed-flow will not provide extra information about the operation status of that freeway section. For example, there may be a couple of passenger cars traveling on a freeway at midnight. If these vehicles are traveling at a speed of 5 miles below the expected speed, the delayed-flow may be large compared to the flow-in-process. To avoid this false alarm, the delayed ratio is only calculated and shown for a freeway section with levels of service E and F.

Example

To explain the results of the input-output model, we presented a sample output as shown in Figure 56. Here, the level of service of flow-in-process and the delayed-flow ratio are plotted in bar-chart formats with the station name printed underneath. The stations are arranged in order from upstream to downstream of the freeway.

In the bar chart, the performance measures correspond to the immediate downstream section of the corresponding station. For example, the level of service shown under the station name RM2222 refers to the level of service of the section between RM2222 and Hancock.

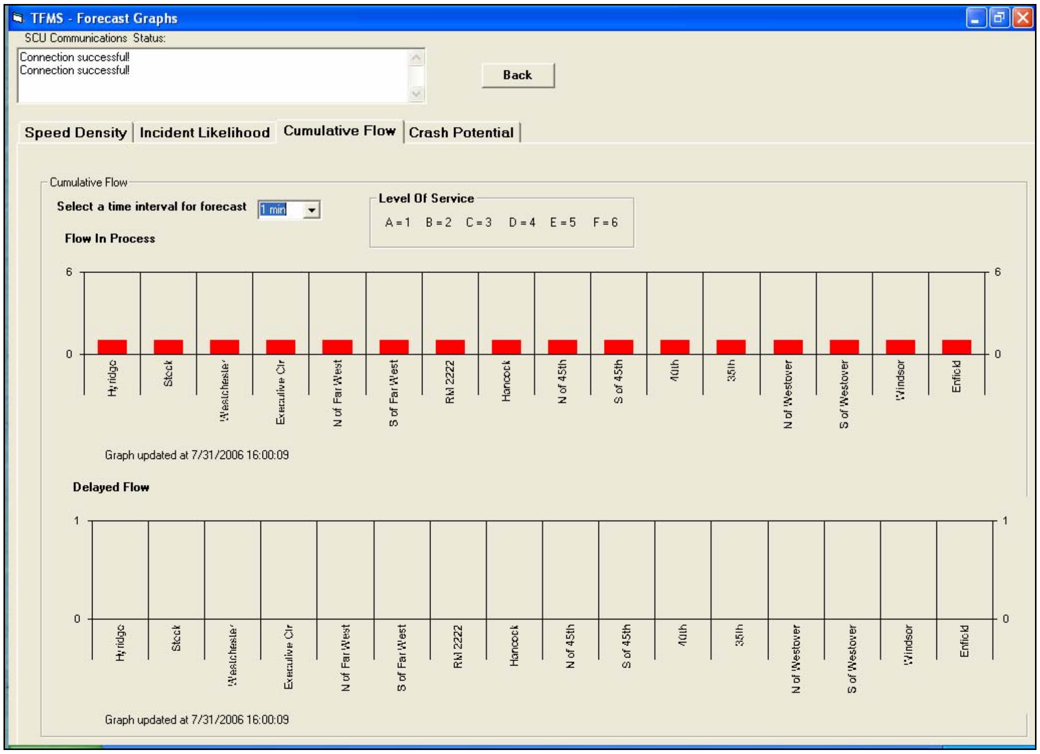


Figure 56. Sample Output of the Input-Output Model.

Figure 57 shows a sample output when a major incident occurs downstream of 35th Street where all lanes are blocked for 5 minutes. In this case, the section level of service is F, and the delayed ratio is calculated and shown. Since all the lanes are blocked, no vehicle can enter the downstream section; thus, the delayed ratio is 1. In this case, we expect that the operation of the upstream section will be affected soon. This shockwave effect is demonstrated in Figure 58, which is obtained 3 minutes after the incident.

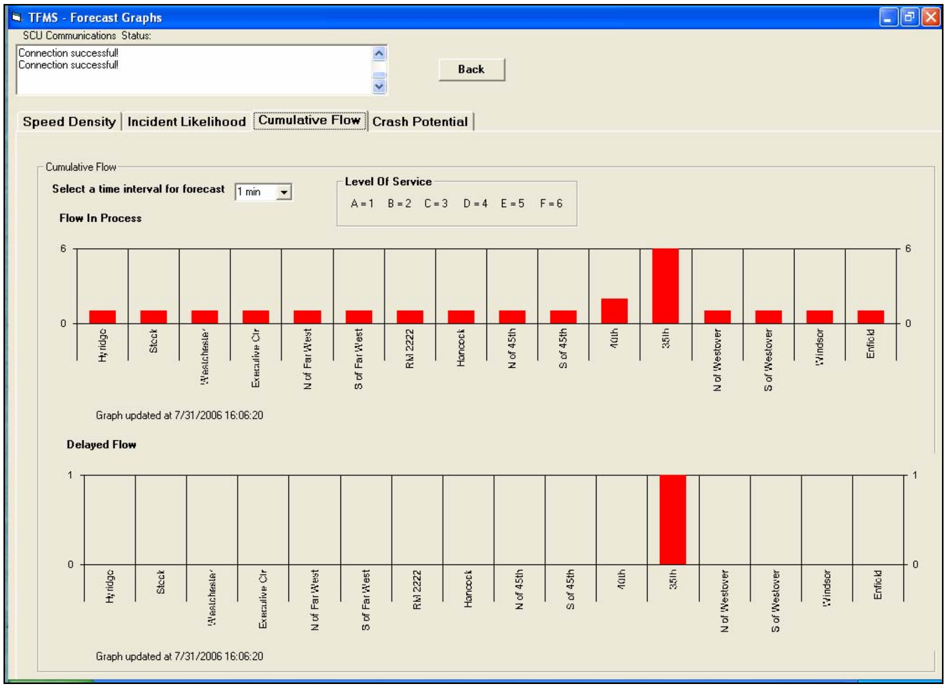


Figure 57. Major Incident Downstream of 35th Street.

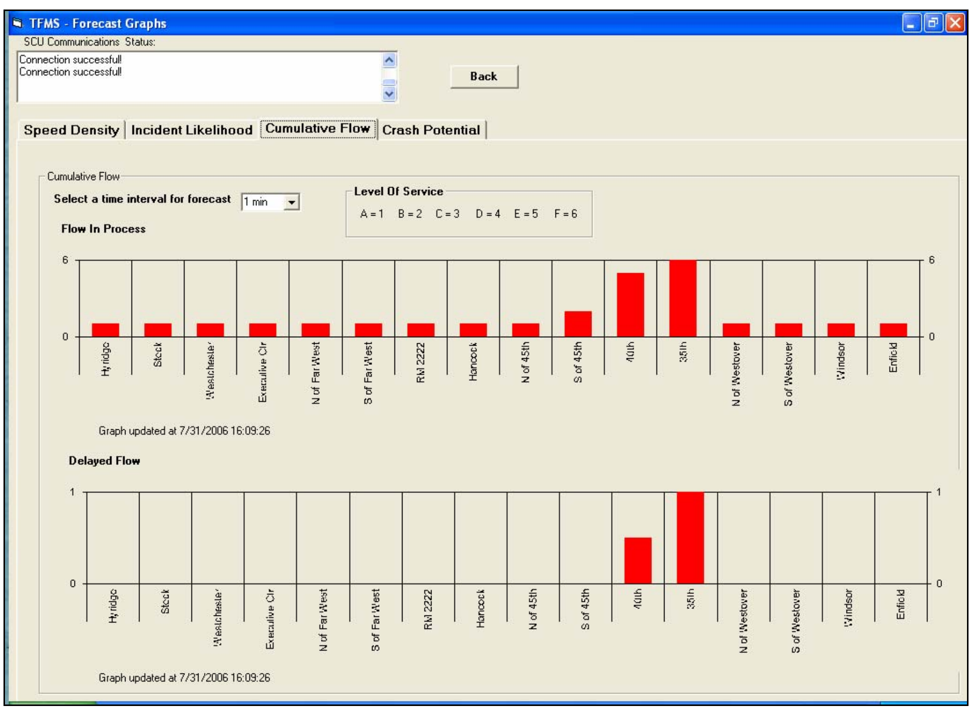


Figure 58. Freeway Operation 3 Minutes after Major Incident.

REFERENCES

1. G. Newell. "A Simplified Theory of Kinematic Waves in Highway Traffic, Part II: Queueing at Freeway Bottlenecks." *Transportation Research,* Part B, Vol. 27, 1993, pp. 289-303.

2. M. J. Cassidy and J. R. Windover. "Methodology for Assessing Dynamics of Freeway Traffic Flow." *Transportation Research Record 1484*, 1995, pp. 73-79.

3. *VISSIM 4.00 User Manual*. PTV, Stumpfstraβe 1, D-76131 Karlsruhe, Germany, 2004.

4. *Highway Capacity Manual*. Transportation Research Board, National Research Council, Washington, D.C., 2000.

APPENDIX C: DEVELOPMENT OF INCIDENT PREDICTION MODEL

Reliable prediction, fast detection, and timely response are the keys to a successful freeway incident management system. This study was initiated as a part of a project sponsored by TxDOT to develop a prototype tool that TxDOT can integrate and implement with their current Advanced Transportation Management System software for use in their freeway management centers. This prototype tool will allow TxDOT operators to better detect when and where conditions on the freeway are likely to produce an incident or congestion is likely to occur. By having a better idea about when and where these conditions are likely to occur, TxDOT operators can take proactive steps to reduce demand and better manage freeways.

The objectives of this study are twofold:

- Develop a method to predict the likelihood of incident occurrence on freeways based on real-time loop detector, weather, and environmental data.
- Implement and test the developed method through a prototype real-time incident prediction system.

Past studies indicated that several traffic measures are potential precursors of freeway accidents (*1, 2, 3*). Further, a recent study by Songchitruksa and Balke (*4*) concluded that integrating weather and environmental data into traffic data could provide significant benefits for the performance of real-time incident prediction. Factors such as visibility and daylight conditions can enhance the incident prediction performance when compared to the models with only traffic data.

Most past studies did not discuss if and how the models can be implemented in a real-time environment. In this study, we first discuss the development of a real-time incident prediction model. Then, we focus our discussion on the design and development of a prototype real-time incident prediction tool that can be integrated with TxDOT ATMS. We set up a simulation test bed using VISSIM to generate the same real-time detector data as those obtained from the System Control Unit. This test bed was designed to test the real-time interaction capability of the prediction module with incoming traffic and weather data streams while avoiding the need to actually deploy the system in the field. This setup allowed us to correct any problems that may arise unexpectedly and fine-tune the system to meet the user's need.

MODEL DEVELOPMENT

We developed an integrated framework utilizing the weather conditions, environment, and loop data to enhance the ability to predict freeway incident occurrences. We explored several modeling approaches for RIPS. First, we calibrated the multinomial logit model considering three outcomes—no incident, congestion, and collision. Then, separate data sources were used to calibrate the multinomial logit (MNL) models in the following order: weather/environment, loop, and both loop and weather/environment. The estimation results revealed that the performance of the MNL models was unsatisfactory because: 1) the estimated parameters are unstable and not robust to different subsets of explanatory variables and 2) the parameter estimates cannot be easily interpreted because they were confounded by the bi-level nest structure.

To address this issue, we recalibrated the model using sequential binary logit model estimation. The estimation results become consistently stable and amenable for logical interpretation. We proposed a method to compute incident occurrence potential or hazard score based on the predicted likelihoods. Specification of critical hazard score requires the use of the prediction performance curve. The procedure to establish the prediction performance curve is provided in this study. Using this curve, the critical threshold for hazard score can be specified for any given desirable detection and false alarm rates.

In this section, we described the data used in this study. Then, the model development and selection process as well as the result interpretation are explained. The procedure to compute hazard score and specify critical threshold is discussed subsequently.

Data

In this study, we selected the freeway sections on Loop 1 and US 183 in Austin, Texas, for our analysis. These sections have area-wide installation of loop detectors. We obtained the weather data from the Camp Mabry weather station, which is the nearest weather station in the vicinity of the studied freeway sections. There are three sources of data that we used to develop incident prediction models in this study: 1) weather and environmental data, 2) loop detector data, and 3) incident data.

Weather and Environmental Data

Weather information was obtained from climatological data provided by the National Climatic Data Center (NCDC) and can be accessed online. Weather records are usually archived on an hourly basis. The data are archived at more frequent intervals (every 10 to 20 minutes) during special weather events such as heavy fog and thunderstorms.

Each weather record can contain both qualitative and quantitative fields. Some fields are ordinal-qualitative. For example, a weather type "TS" would represent moderate thunderstorm while "TS+" and "TS-" indicate heavy and light thunderstorm, respectively. In addition, there can be a combination of weather types. For instance, if a light thunderstorm and a haze are occurring at the same time, the weather type record is "TS- HZ." In order to perform the analysis, these text formats must be recoded as indicator variables, and a combination of weather types must be split into a set of single indicator variables. Weather data were recoded and stored in comma-delimited files. The data in these files are summarized in Table 5.

Table 5. Detailed Records of Processed Weather Data.

Acronym	Description	Acronym	Description
sta	Station: AUST, CAMP, GEOR	valGust	Gust value
mo	Month	staPr	Station pressure
yr	Year	prTend	Pressure tendency
day	Day	seaLvlPr	Sea level pressure
rtime	Recorded time	repType	Report type
staType	Station type	precipTt	Precipitation total
maint	Maintenance indicator	vis	Visibility (numeric)
skyCond	Sky condition	skyCLR	Clear sky indicator
visibil	Visibility	skyFEW	Few clouds indicator
wthrType	Weather types	skySCT	Scattered clouds indicator
dryBulbF	Dry bulb temperature (F)	skyBKN	Broken clouds indicator
dryBulbC	Dry bulb temperature (C)	skyOVC	Overcast indicator
wetBulbF	Wet bulb temperature (F)	wDZ	Drizzle indicator
wetBulbC	Wet bulb temperature (C)	wRA	Rain indicator
dewPtF	Dew point (F)	wTS	Thunderstorm indicator
dewPtC	Dew point (C)	wFG	Fog indicator
relHumid	Relative humidity	wFZ	Freeze indicator
windSpd	Wind speed (mph)	wSQ	Squall indicator
windDir	Wind direction	wHZ	Haze indicator
windGust	Wind gust indicator	wSN	Snow indicator

Incident Data

The incident logs were imported from the Excel® format and then manually reviewed and edited for errors in the records. Incident logs were recoded and stored in the same format as weather records. A summary of records contained in the incident logs is shown in Table 6. One caveat about incident records is that the reported date and time at the beginning and at the end of an incident may be inaccurate due to a possibility of delay in the reporting process. In addition, a number of minor freeway incidents may have never been reported. However, this lagged time is associated with the type of incident. For instance, the mean time to reporting of fatal or injury accidents is more likely to be shorter than property damage only (PDO) accidents.

Table 6. Detailed Records of Processed Incident Data.

Acronym	Description	Acronym	Description
inIndex	Incident index	logEvent	Logged date and time
inNumber	Incident number	clEvent	Cleared date and time
dir	Direction	comment	Additional comments
roadway	Roadway	affected	Affected structure (freeway/ramps/frontage/interchange)
crossSt	Cross street	singleLn	Affected lanes (specific)
detector	Detector description	logDate	Logged date
entrRamp	Entrance ramp	logTime	Logged time
exitRamp	Exit ramp	lane1	Affected lane 1 indicator
befAfter	Before/After/At	lane2	Affected lane 2 indicator
block	Block number	lane3	Affected lane 3 indicator
intype	Incident type	lane4	Affected lane 4 indicator
notified	Notifier	leftSh	Affected left shoulder indicator
		rightSh	Affected right shoulder indicator

Loop Detector Data

Each record of loop data contains a time stamp and a series of detector ID, volume, occupancy, speed, and percent truck. Loop data are currently archived on an hourly basis. Hourly data were combined into a single file for each day. Each record of loop data is referenced by a time stamp and detector ID. File names were created using the following format: "NAME YYYYMMDD.txt." The "NAME" represents a roadway name, which is LP0001 or US0183. For example, the loop data from all detectors along US 183 on June 14, 2003, were stored in the file name "US0183 20030614.txt."

The inventory of loop detectors in Austin contains the information about each detector including detector ID (detID), station ID (staID), roadway (rdway), direction (dir), lane number (lane), type (freeway or ramp), and detector description (detdesc).

Computed Traffic Measures

There is a large catalog of measures that can be computed from a stream of 1-minute observations of loop detectors. In this study, average volume, average speed, average occupancy, and CVS were evaluated in the model development. The computation procedure requires a specification of window size for moving averages. First, the calculation is carried out for each individual lane detector. Then, station averaging is applied to a set of detectors that belong to the same station.

For each individual lane detector, the average volume, average speed, average occupancy, and in-lane variation of speed were computed. Moving-average window size can be specified in the calculation. In this analysis, 3-, 5-, and 8-minute moving averages were tested.

The average volume per minute is calculated as

$$\overline{q} = \frac{\sum_{i=1}^{N} q_i}{N} \tag{C-1}$$

where

q_i = 1-minute volume count of i^{th} interval and

N = number of 1-minute intervals in a specified averaging window.

The average occupancy is calculated as

$$\overline{o} = \frac{\sum_{i=1}^{N} o_i}{N} \tag{C-2}$$

where

o_i = 1-minute average percent occupancy.

Also, occupancy is a proportional indicator of density.

The weighted average speed is calculated as

$$\overline{v} = \frac{\sum_{i=1}^{N} q_i v_i}{\sum_{i=1}^{N} q_i} \tag{C-3}$$

where

v_i = 1-minute weighted average speed of i^{th} interval.

The weighted average speed has an advantage over the arithmetic mean in that zero-count intervals are not used in a calculation, which avoids underestimation of mean values. The weighted average speed better describes the true fluctuation of vehicles' speed over time, particularly during nighttime when there is a preponderance of zero-count intervals.

CVS is a measure of the amount of fluctuation in traveling speeds. Past studies indicate that a breakdown in traffic flow will significantly increase the CVS and thus the likelihood of accidents (*1*,*2*,*4*).

Because a speed observation can be zero when there is no vehicle, the computation of CVS can be done in many variations. To illustrate, assume that we are considering CVS over a 5-minute interval. The first case is to compute the CVS using the 1-minute average speeds over 5-minute intervals while each interval is weighted equally. In this manner, zero-count intervals, which are typically the case at night, will increase the value of CVS. In other words, CVS may be large because zero-count intervals can cause abrupt changes in speed values. The second case is to compute the CVS as in the first case, but each interval is weighted by the volume counts. Mathematically, this can be expressed as

$$CVS = \frac{\sigma_{v_i}}{\bar{v}} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N} q_i \left(v_i - \bar{v}\right)^2}}{\bar{v}}. \tag{C-4}$$

The CVS values are generally sensitive to differences in speeds in low-volume conditions. The CVS calculation using Equation C-4 can be modified such that the moving weighted average speeds are used instead of 1-minute average speed. As a result, CVS can be computed as

$$CVS_{\bar{v}} = \frac{\sigma_{\bar{v}_i}}{\bar{\bar{v}}} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(\bar{v}_i - \bar{\bar{v}}\right)^2}}{\bar{\bar{v}}} \tag{C-5}$$

where

$\bar{\bar{v}} = \frac{1}{N}\sum_{i=1}^{N} \bar{v}_i$ and

$CVS_{\bar{v}}$ = the fluctuation of moving-average speeds over N intervals.

Equation C-5 was used for the CVS computation in this analysis to mitigate the effect of changes in speed values in low-volume conditions. The CVS parameter also helps decrease the false alarm rate. Other measures such as standard deviations of volume and occupancy were not analyzed since an earlier study of these indicators did not find results encouraging *(5)*.

Data from a group of lane detectors at the same location are referred to as "station data." The computed lane measures are averaged across lanes to obtain station average measures. Four measures are computed as in the case of lane data. For each 1-minute interval, missing or invalid measures in each lane detector can either be omitted or specially treated. In this study, we treat any intervals that have invalid computed lane measures as invalid intervals.

The station average volume is defined by

$$\bar{q}_s = \frac{1}{\ell}\sum_{j=1}^{\ell}\bar{q}_j. \tag{C-6}$$

where

ℓ = the number of lanes at the station.

The station average occupancy, the station average speed, and the station average CVS are defined by Equations C-7, C-8, and C-9, respectively.

$$\bar{o}_s = \frac{1}{\ell}\sum_{j=1}^{\ell}\bar{o}_j. \tag{C-7}$$

$$\bar{v}_s = \frac{1}{\ell}\sum_{j=1}^{\ell}\bar{v}_j. \tag{C-8}$$

$$\overline{CVS} = \frac{1}{\ell}\sum_{j=1}^{\ell}CVS_j. \tag{C-9}$$

The station average measures were computed for every minute of valid lane detector data. These data are further matched with both incident-affected and incident-free conditions to produce a data set for model development.

Data Integration

Incident logs were merged with weather records using incident logged date and time and weather record reported time. Only incident logs with weather records within 60 minutes of incident logged time were considered in the analysis. The merged data file "iwlimdep04.txt" consists of incident logs matched with weather records on IH 35, US 183, and Loop 1 from year 2002 to 2004. The fields in the data file are inIndex, dir, inType, lane1, lane2, lane3, lane4, leftSh, rightSh, mo, yr, day, logTime, rtime, daytime, twi, dryBulbF, wetBulbF, precipTt, vis, skyCLR, skyFEW, skySCT, skyBKN, skyOVC, wDZ, wRA, wTS, wFG, wFZ, wSQ, wHZ, wSN, and rdway.

For modeling purposes, incident logs must be matched with corresponding loop detector data. Only collision incidents on US 183 and Loop 1 from 2003 to 2004 were considered in the previous analysis because loop installations are somewhat limited on other roadways.

The files "i.staID.03.txt" and "i.staID.04.txt" contain matched incident logs and station IDs for all types of incidents that occurred on freeways in 2003 and 2004, respectively. Cross street, roadway, and direction descriptions in incident logs were used to manually identify detector stations in the vicinity of an incident. Station IDs can be located for 762 and 3662 incidents in 2003 and 2004, respectively. Note that these figures represent all types of incidents where the congestion incidents represent a significant portion of all incidents in 2004 due to changes in occupancy thresholds.

From the analysis of loop data, the file "mdat.txt" contains computed loop-related measures for collision incidents in 2003 and 2004 on US 183 and Loop 1 that can be paired with specific detector station IDs. Three moving-average windows and eight values of incident detection times were used to compute the loop-related measures for a total of 117 collision records. The values of 3, 5, and 8 minutes were used for a moving-average window. Incident detection times were varied from 0 to 35 minutes at 5-minute intervals. Each record contains the following fields: incident detection time (min.bef), incident index (inIndex), moving-average window (interval), time stamp of detector data (tstamp), mean volume (vol.mean), mean occupancy (occ.mean), mean speed (spd.mean), and mean CVS (cvs.mean).

The sample for model calibration must consist of incident-affected and incident-free data. For incident-affected data, each record contains incident logs, weather records, loop data, and environment data. Incident-free data represent traffic conditions free from incident impacts. First, incident logs were paired with 5-minute moving-average measures computed from loop data at 15 minutes before reported incident occurrence times. Congestion, collision, and stall incidents in 2003 and 2004 from US 183 and Loop 1 were considered in the analysis, resulting in a total of 4187 incidents. After extracting loop data for each incident, the number of incidents valid for the analysis was reduced to 3822 incidents due to missing or erroneous loop data. When the data were further matched with weather records, the number of incidents valid for the analysis reduced to 3808. For incident-affected conditions, the file "inc.iwloop.txt" contains incident data matched with weather records and loop data for model calibration.

Incident-free data were selected randomly using the following set of criteria:

- Only data during weekdays between 6 a.m. and 12 a.m. (midnight) from 2003 to 2004 were considered due to limited monitoring outside TMC operating hours.
- Only days without reported incidents were used in random sampling.
- Five detector stations were randomly selected for each day.
- Ten time stamps were randomly selected for each station ID for the computation of measures from loop data.

A total of 2540 records of incident-free loop data were randomly sampled. Then, each record of computed loop data was matched with weather records at the Camp Mabry station. Only 20 records could not be matched with historical weather data, thus leaving 2520 records representative of incident-free data.

The subsequent modeling analysis used a combination of 5-minute moving-average and 15-minute duration before incident for calculation of traffic measures since it was found to be the most satisfactory specification for the incident prediction task *(4)*.

Preliminary Model Estimation

First, three types of incidents were considered for model calibrations—collision, congestion, and stall. Nested and non-nested MNL models were used to assess and identify key determinants of incident types. Preliminary results indicated that stall incidents are difficult to predict due to their irregular pattern of occurrence. Stall incidents often involve mechanical breakdowns of vehicles and are fairly random. Factors such as seasonality and ambient temperature were tested during the model development; however, none of the factors were found to be significant enough to justify further consideration of stall incidents in the model calibration.

Therefore, stall incidents were excluded from further calibration attempts. Fortunately, stall incidents represent only a minute proportion of all incidents and take place mostly on freeway shoulders, thus causing only minimal disruption to mainline traffic flow.

A comparison of incident records between 2003 and 2004 revealed a sudden increase in the number of congestion incidents on US 183. In order to avoid the bias in the sample, approximately 20 percent of congestion incidents in 2004 were randomly selected for the sample data set.

There are a total of 3719 observations in the final data set, which can be broken down into 2520 non-incident records, 117 collision incidents, and 1082 congestion incidents. The nest structure as shown in Figure 59 was used for the calibration of nested MNL models.

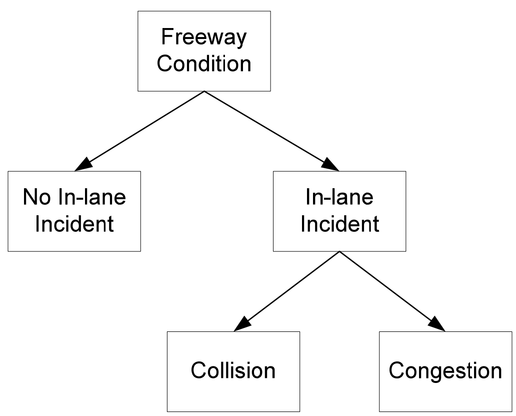


Figure 59. Nest Structure for Nested MNL Models.

Models were calibrated for three different scenarios: 1) both loop and weather data are available, 2) only loop data are available, and 3) only weather data are available. Model estimation results before probability adjustments for each corresponding scenario are shown in Table 7 through Table 9, respectively. The results indicated that selected explanatory variables are statistically significant regardless of the available data set.

Estimation results were corrected for the biased sampling. In the data sampling procedure, we followed a so-called choice-based sampling, which is essentially the case where certain outcomes are intentionally oversampled in order to understand the decision-making process. As opposed to random sampling, the estimation results in this scenario must be properly corrected using the true proportion of each outcome in the population.

Table 7. Estimated Nested MNL Model Using Loop and Weather Data.

Variables	Descriptions	Estimated Coefficients	t-ratio	P-value
Attributes of utility functions of incident types of the in-lane incident branch				
Estimated coefficients of the utility function of congestion incidents				
JC	Constant for the congestion incident alternative	3.418	7.253	0.000
JVOL	Five-minute station average of traffic volume (veh/min/lane)	0.105	5.470	0.000
JOCC	Five-minute station average of occupancy (%)	0.037	1.790	0.073
JDIR	Direction indicator (1 if NB; 0 if SB)	-0.808	-4.906	0.000
JPEAKHR	Peak hour indicator (1 if during 7:30am to 9:00am or 4:00pm to 6:00pm; 0 if otherwise)	-1.702	-5.903	0.000
JDAY	Daytime indicator (1 if during the day; 0 if otherwise)	-2.255	-5.167	0.000
Estimated coefficients of the utility function of collision incidents				
CCVS	Five-minute station average of variation in speeds	3.616	2.250	0.024
CTS	Thunderstorm indicator (1 if thunderstorm is present; 0 if otherwise)	1.185	2.396	0.017
Attributes of utility function of the branch choice				
Estimated coefficients of the utility function of the in-lane incident branch				
YC	Constant for the in-lane incident branch	-3.244	-6.240	0.000
YDAY	Daytime indicator (1 if during the day; 0 if otherwise)	1.485	3.788	0.000
YOCC	Five-minute station average of occupancy (%)	0.060	3.066	0.002
YCVS	Five-minute station average of CVS (%)	11.135	6.556	0.000
YTS	Thunderstorm indicator (1 if thunderstorm is present; 0 if otherwise)	1.061	2.793	0.005
YLVIS	Natural log of visibity in miles	-0.466	-4.661	0.000
YWIND	Wind speed (mph)	-0.128	-7.625	0.000
Estimated inclusive value parameters				
Φ(none)	Estimated inclusive value parameter of the non-incident branch		….. Fixed at 1.0 …..	
Φ(in-lane)	Estimated inclusive value parameter of the in-lane incident branch	0.781	5.868	0.000
	Log likelihood at convergence	-1635.104		
	Restricted log likelihood	-2929.240		
	ρ^2 statistic	0.442		
	Number of observations	3295		

Table 8. Estimated Nested MNL Model Using Only Loop Detector Data.

Variables	Descriptions	Estimated Coefficients	t-ratio	P-value
Attributes of utility functions of incident types of the in-lane incident branch				
Estimated coefficients of the utility function of congestion incidents				
JC	Constant for the congestion incident alternative	3.723	7.754	0.000
JVOL	Five-minute station average of traffic volume (veh/min/lane)	0.087	4.810	0.000
JOCC	Five-minute station average of occupancy (%)	0.054	2.546	0.011
JDIR	Direction indicator (1 if NB; 0 if SB)	-0.761	-4.950	0.000
JPEAKHR	Peak hour indicator (1 if during 7:30am to 9:00am or 4:00pm to 6:00pm; 0 if otherwise)	-1.897	-6.070	0.000
JDAY	Daytime indicator (1 if during the day; 0 if otherwise)	-2.265	-5.173	0.000
Estimated coefficients of the utility function of collision incidents				
CCVS	Five-minute station average of variation in speeds	4.132	2.621	0.009
Attributes of utility function of the branch choice				
Estimated coefficients of the utility function of the in-lane incident branch				
YC	Constant for the in-lane incident branch	-5.044	-9.807	0.000
YDAY	Daytime indicator (1 if during the day; 0 if otherwise)	1.563	3.591	0.000
YOCC	Five-minute station average of occupancy (%)	0.056	2.594	0.009
YCVS	Five-minute station average of CVS (%)	11.165	6.573	0.000
Estimated inclusive value parameters				
Φ(none)	Estimated inclusive value parameter of the non-incident branch		….. Fixed at 1.0 …..	
Φ(in-lane)	Estimated inclusive value parameter of the in-lane incident branch	0.861	5.810	0.000
	Log likelihood at convergence	-1709.441		
	Restricted log likelihood	-2959.738		
	ρ^2 statistic	0.422		
	Number of observations	3335		

Table 9. Estimated Nested MNL Model Using Only Weather Data.

Variables	Descriptions	Estimated Coefficients	t-ratio	P-value
Attributes of utility functions of incident types of the in-lane incident branch				
Estimated coefficients of the utility function of congestion incidents				
JC	Constant for the congestion incident alternative	3.861	9.544	0.000
JDIR	Direction indicator (1 if NB; 0 if SB)	-0.696	-5.285	0.000
JPEAKHR	Peak hour indicator (1 if during 7:30am to 9:00am or 4:00pm to 6:00pm; 0 if otherwise)	-1.062	-5.545	0.000
JDAY	Daytime indicator (1 if during the day; 0 if otherwise)	-0.710	-2.071	0.038
Estimated coefficients of the utility function of collision incidents				
CTS	Thunderstorm indicator (1 if thunderstorm is present; 0 if otherwise)	0.951	2.107	0.035
Attributes of utility function of the branch choice				
Estimated coefficients of the utility function of the in-lane incident branch				
YC	Constant for the in-lane incident branch	-4.245	-4.893	0.000
YDAY	Daytime indicator (1 if during the day; 0 if otherwise)	1.994	3.985	0.000
YTS	Thunderstorm indicator (1 if thunderstorm is present; 0 if otherwise)	0.862	2.373	0.018
YLVIS	Natural log of visibity in miles	-0.445	-5.679	0.000
YWIND	Wind speed (mph)	-0.132	-9.599	0.000
Estimated inclusive value parameters				
Φ(none)	Estimated inclusive value parameter of the non-incident branch	….. Fixed at 1.0 …..		
Φ(in-lane)	Estimated inclusive value parameter of the in-lane incident branch	1.525	5.458	0.000
	Log likelihood at convergence	-2249.225		
	Restricted log likelihood	-3378.399		
	ρ^2 statistic	0.334		
	Number of observations	3719		

From Table 7, using both loop and weather data, the following factors were found to increase the likelihood of incident occurrence on freeways:

- daytime,
- increase in average occupancy,
- increase in average CVS,
- presence of thunderstorm,
- degradation of visibility condition, and
- decrease in wind speed.

Given that an in-lane incident has occurred, the following factors were found to be significant determinants for differentiating a collision from congestion:

- increase in average CVS,
- presence of thunderstorm,
- decrease in traffic flow,
- decrease in occupancy,
- northbound direction,
- peak hours, and
- daytime.

The models were re-calibrated using separate loop and weather/environment data to evaluate if the coefficient estimates are robust. The estimation results are provided in Table 8 and Table 9. From both tables, while the overall model goodness-of-fit decreases, the signs of estimated coefficients and the statistical significance of explanatory variables are consistent with the previous case where both data sources were used for calibration. The implication here is that the combination of loop detector data and weather/environment data can provide an incident prediction model that is more robust than using one data source alone.

There are disadvantages, however, to using MNL models for real-time incident prediction. First, the interpretation of the influence of coefficient estimates on the probabilities of the outcomes is complicated because it is confounded by the nest structure. For example, an increase in CVS can increase the likelihood of in-lane incident occurrence as well as the likelihood of collision given that an in-lane incident has occurred. According to the nest structure, the probabilities of the branch choice and the alternatives in the branch are both affected by the same variable. Therefore, the structure of MNL models requires the exact calculation of the final probability to quantify the impact of changes in explanatory variables on the probabilities of the outcomes. Second, the estimation results of MNL models are relatively less robust to the choice of independent variables when compared to binary logit models. This characteristic of MNL models implies that the addition or removal of independent variables in the models may change the statistical significance of other existing variables. This model behavior is undesirable and can be avoided through the use of binary logit models as discussed in subsequent sections.

Selected Incident Prediction Models

The previous calibrations using MNL models showed that while the model choice is appropriate theoretically, the results and inherent disadvantages make them less desirable. To resolve this issue, we calibrated two binary logit models using both loop and weather/environment data with each model representing the same branch outcomes as in the previous MNL estimations. This approach is referred to as sequential binary logit estimation.

Sequential Binary Logit Estimation

Two binary logit models were estimated. The first model has two possible outcomes: an in-lane incident versus no incident. Given that an in-lane incident has occurred, the second prediction model also has two outcomes: collision versus congestion. Note that this sequential structure is similar to the MNL model except that this requires two separate model calibrations.

The first model predicts the likelihood of having an in-lane incident within the next 15 minutes in the vicinity of the detector station. The second model predicts the type of incident given that an incident has occurred, i.e., the likelihood of having a collision versus congestion. The binary logit estimation results of the first and second models are shown in Table 10 and Table 11, respectively.

Table 10. Estimated Binary Logit Model for In-Lane Incident versus No Incident.

Data set: Loop and weather data (US-183 and Loop 1 from 2003 to 2004)

Before constant adjustments

Variables	Descriptions	Estimated Coefficients	t-ratio	P-value
Attributes of utility functions of the in-lane incident alternative				
ONE	Constant for the in-lane incident alternative	2.796	2.433	0.0150
OCCMEAN	Five-minute station average of occupancy (%)	0.114	4.187	0.0000
SPDMEAN	Five-minute station average of weighted average mean speed	-0.034	-2.351	0.0187
CVSMEAN	Five-minute station average of coefficient of variation in speed	9.244	1.834	0.0667
LOGVIS	Natural log of visibity in miles	-0.644	-3.432	0.0006
DAYTIME	Daytime indicator (1 if during the day; 0 if otherwise)	0.924	2.886	0.0039
PEAKHR	Peak hour indicator (1 if during 7:30am to 9:00am or 4:00pm to 6:00pm; 0 if otherwise)	-1.758	-5.551	0.0000
SUMMER	Summer indicator (1 if summer; 0 if otherwise)	0.936	3.381	0.0007
DIR	Direction indicator (1 if NB; 0 if SB)	-0.751	-3.105	0.0019
	Log likelihood at convergence	-251.222		
	Restricted log likelihood	-390.198		
	ρ^2 statistic	0.356		
	Number of observations	563		

To interpret the results in Table 10, an explanatory variable with a positive coefficient estimate implies that freeway conditions with a presence of such factors are more likely to produce an in-lane incident versus no incident and vice versa for negative coefficient estimates. The following factors were found to be significant precursors of in-lane incident occurrence:

- increase in average occupancy,
- decrease in average speed,
- increase in average CVS,
- decrease in visibility condition,
- daytime,
- non-peak periods,
- summer season, and
- southbound traveling direction.

Table 11. Estimated Binary Logit Model for Congestion versus Collision Incident.

Data set: Loop and weather data (US-183 and Loop 1 from 2003 to 2004)

Before constant adjustments

Variables	Descriptions	Estimated Coefficients	t-ratio	P-value
Attributes of utility function of the congestion alternative				
ONE	Constant for the congestion alternative	1.593	1.692	0.0906
OCCMEAN	Five-minute station average of occupancy (%)	0.142	4.713	0.0000
SPDMEAN	Five-minute station average of weighted average mean speed	0.030	2.369	0.0179
LOGVIS	Natural log of visibity in miles	-0.805	-3.316	0.0009
DAYTIME	Daytime indicator (1 if during the day; 0 if otherwise)	-1.887	-3.023	0.0025
PEAKHR	Peak hour indicator (1 if during 7:30am to 9:00am or 4:00pm to 6:00pm; 0 if otherwise)	-2.540	-5.510	0.0000
SUMMER	Summer indicator (1 if summer; 0 if otherwise)	1.704	3.887	0.0001
DIR	Direction indicator (1 if NB; 0 if SB)	-0.891	-2.057	0.0397
SKYCLR	Clear sky indicator (1 if yes; 0 if otherwise)	1.454	3.866	0.0001
	Log likelihood at convergence	-107.890		
	Restricted log likelihood	-179.593		
	ρ^2 statistic	0.399		
	Number of observations	284		

The estimation results are quite intuitive. The results imply that incidents are more likely in congested traffic conditions and/or poor visibility conditions. In-lane incidents are also more likely to occur during daytime, summer, non-peak period, and southbound direction.

The estimation results of the second binary logit model are presented in Table 11. This model predicts the type of incident given that an incident has occurred. Two possible outcomes were considered in this model: congestion versus collision. Oftentimes, the incident prediction model cannot differentiate whether the observed freeway conditions are more likely to produce a collision or normal congestion. This model aims to distinguish between these two types of events based on the available weather and traffic information.

Positive coefficient estimates in Table 11 signify the increase in likelihood of having congestion rather than a collision and vice versa for negative coefficient estimates. The following conditions were found to be more likely to produce a collision:

- decrease in average occupancy,
- decrease in average speed,

- decrease in visibility condition,
- peak periods,
- non-summer seasons,
- northbound traveling direction, and
- sky conditions other than clear sky.

From both models, it can be concluded that, in addition to traffic measures, visibility condition, sky condition, and seasonal indicator can be useful predictors of incident occurrence and incident type.

Outcome Probability Adjustments

Since the collision outcomes are overrepresented in the sample, an estimation correction must be made. To correct the constant estimates, each constant must have the following subtracted from it:

$$\ln\left(\frac{SF_i}{PF_i}\right), \tag{C-10}$$

where

SF_i = the fraction of observations having outcome *i* in the sample and

PF_i = the fraction of observations having outcome *i* in the total population.

To adjust the estimation results in Table 10, the number of incident-affected and incident-free 15-minute intervals must be estimated. Based on historical incident data, there were 87 incidents on average per station per year. A study by Drakopoulos reported that the average time from a crash occurrence until it was cleared was 52.2 minutes. Therefore, it was logical to assume that each crash occurrence affects the traffic conditions for three intervals on average. This gives $PF_{incident}$ = (87×3)/(52×5×19) = 0.05 and $PF_{non\text{-}incident}$ = 0.95. Based on the sample used in the model calibration, the sample fractions can be calculated straightforwardly: $SF_{incident}$ = 278/563 = 0.494 and $SF_{non\text{-}incident} = 1 - SF_{incident}$ = 0.506.

Similar adjustments apply to the results in Table 11. Let us consider all the incident records as population data. The data set used in the model calibrations was randomly sampled from this population. Therefore, the following adjustments can be estimated: $SF_{congestion}$ = 191/284, $PF_{congestion}$ = 4000/4200, $SF_{collision}$ = 93/284, and $PF_{collision}$ = 200/4200.

Hazard Score

The final models consist of two sequential binary logit models where the first one predicts the likelihood of an in-lane incident versus no incident and the second one predicts the likelihood of an incident being a collision versus congestion. Control center operators may find these two probabilities difficult to interpret in actual implementation since they can be overwhelmed by multiple sources of information simultaneously.

Two predicted components shown in Table 10 and Table 11 after probability adjustments can be expressed as shown in Equations C-11 through C-18.

The first predicted component can be estimated by computing two utility functions as follows:

$$\begin{aligned} U_{inc} &= 0.505 + 0.114(OCCMEAN) - 0.034(SPDMEAN) + 9.244(CVSMEAN) \\ &-0.644 \times Log(VIS) + 0.924(DAYTIME) - 1.758(PEAKHR) + 0.936(SUMMER) \\ &-0.751(DIR) \end{aligned} \tag{C-11}$$

$$U_{none} = 0.630 \tag{C-12}$$

where

U_{inc} and U_{none}= the utility functions of in-lane incident and no incident alternatives, respectively.

The probabilities of having an in-lane incident and no incident can be calculated as:

$$\Pr(\text{In-Lane Incident}) = \frac{e^{U_{inc}}}{e^{U_{inc}} + e^{U_{none}}} \tag{C-13}$$

$$\Pr(\text{No Incident}) = 1 - \Pr(\text{In-Lane Incident}) \tag{C-14}$$

The second predicted component determines the likely incident type based on two utility functions:

$$\begin{aligned} U_{cong} &= 1.941 + 0.142(OCCMEAN) + 0.030(SPDMEAN) - 0.805 \times \log(VIS) \\ &-1.887(DAYTIME) - 2.540(PEAKHR) + 1.704(SUMMER) - 0.891(DIR) \\ &+1.454(SKYCLR) \end{aligned} \tag{C-15}$$

$$U_{coll} = -1.928 \tag{C-16}$$

where

U_{cong} and U_{coll} = the utility functions of congestion and collision outcomes, respectively. The corresponding probabilities can be estimated as:

$$\Pr(\text{congestion} \mid \text{in-lane incident has occurred}) = \frac{e^{U_{cong}}}{e^{U_{cong}} + e^{U_{coll}}} \tag{C-17}$$

$$\Pr(\text{collision}|\text{in-lane incident}) = 1 - \Pr(\text{congestion} \mid \text{in-lane incident}) \tag{C-18}$$

To simplify the interpretation process, we considered three techniques to combine two predicted probabilities in an efficient and meaningful manner: 1) linear combination, 2) higher order function, and 3) Euclidean distance.

A linear combination is essentially a linear function of predicted probabilities:

$$H_t = \sum_{\forall i} \beta_i \Pr(i) \tag{C-19}$$

where

H_t = hazard score at time t and

β_i = a coefficient of predicted probability for outcome i.

A higher order function or a nonlinear combination is similar to a linear combination except that the power of Pr*(i)* can be greater than 1.

Euclidean distance is a normalized form of a quadratic function, which can be calculated as:

$$H_t = \left[\sum_{\forall i} \beta_i \Pr(i)^2 \right]^{1/2} \tag{C-20}$$

where

$\sum_{\forall i} \beta_i = 1$.

The Euclidean distance approach was selected in this study since it has the following desirable properties:

- The scores are scaled and normalized from 0 to 1.
- The score calculation is scalable if additional probability outcomes should be added.

- The higher score implies a more hazardous or incident-prone condition, and the lower score implies a less hazardous or incident-prone condition.
- The normalized score can be visualized as a scaled distance from the origin in the n-dimension where n is the number of possible outcomes. The further distance implies a more critical condition.
- All three techniques considered require appropriate assignment of weighting values or β_i.

We utilized the normalized expected cost approach for this task. This approach views β_i as a normalized cost associated with each outcome *i*. The sequential binary estimation approach yields three possible outcomes with the following probabilities: Pr(No Incident), Pr(Congestion), and Pr(Collision). The probability of no incident can be estimated from Equation C-14. The probabilities of congestion and collision outcomes can be determined as follows:

$$\Pr(\text{Congestion}) = \Pr(\text{Congestion}|\text{In-lane Incident}) \cdot \Pr(\text{In-lane Incident}) \qquad \text{(C-21)}$$

$$\Pr(\text{Collision}) = \Pr(\text{Collision}|\text{In-lane Incident}) \cdot \Pr(\text{In-lane Incident}) \qquad \text{(C-22)}$$

Also, it should be noted that the summation of all outcome probabilities is equal to 1. Now, let us define c_i as the associated cost of outcome *i* where *i* can be either congestion, collision, or nothing. Costs of a collision and congestion are the equivalent monetary amounts if they do occur or the equivalent monetary savings if they have been avoided through appropriate proactive strategies. In this case, it is obvious that $c_{nothing} = 0$. Since the accurate valuation of $c_{congestion}$ and $c_{collision}$ is beyond the scope of this study, we assumed $c_{congestion}$ = \$50,000 and $c_{collision}$ = \$200,000 to illustrate how to translate the expected cost to the values for β_i.

Since $\sum_{\forall i} \beta_i = 1$, we can rescale the cost components to satisfy this property as follows:

$$\beta_i = \frac{c_i}{\sum_{\forall i} c_i} \qquad \text{(C-23)}$$

In this manner, β_i can also be viewed as the normalized cost of an outcome *i*. Therefore, using the Euclidean approach, the hazard score at time *t* which combines all the predicted probabilities can be expressed as:

$$H_t = \left[0 \cdot \Pr(\text{No Incident})_t^2 + (0.2) \cdot \Pr(\text{Congestion})_t^2 + (0.8) \cdot \Pr(\text{Collision})_t^2 \right]^{1/2} \quad \text{(C-24)}$$

Denote that $H_t \in (0,1)$. The system implementation as described in subsequent sections utilizes Equation C-24 for the calculation of the hazard score.

Threshold Selection Procedure

We considered two procedures to specify the critical threshold for the hazard score: 1) percentile based and 2) performance based.

In the percentile-based approach, thresholds are specified based on the percent of the time the score should be allowed to exceed the thresholds. There can be more than one threshold. For example, TMC managers can specify, based on historical data, two thresholds at 80th percentile and 90th percentile, respectively. The lower threshold may be designated as an alert condition, while the higher threshold would then require immediate attention by control center operators.

To mitigate the number of false alarms as well as excessive hours of unnecessary monitoring, a simple filter control policy could be applied. One common filter policy is the persistence check. For example, the TMC can require the hazard score to exceed the specified thresholds for an extended period of time, say at least 2 minutes, before any follow-up actions would take place.

A performance-based approach employs performance characteristics of prediction models to define the thresholds based on the tradeoff between two criteria: missing prediction rate and false alarm rate. Hypothetical relationships of false alarm and missing prediction rates versus a range of thresholds are illustrated in Figure 60.

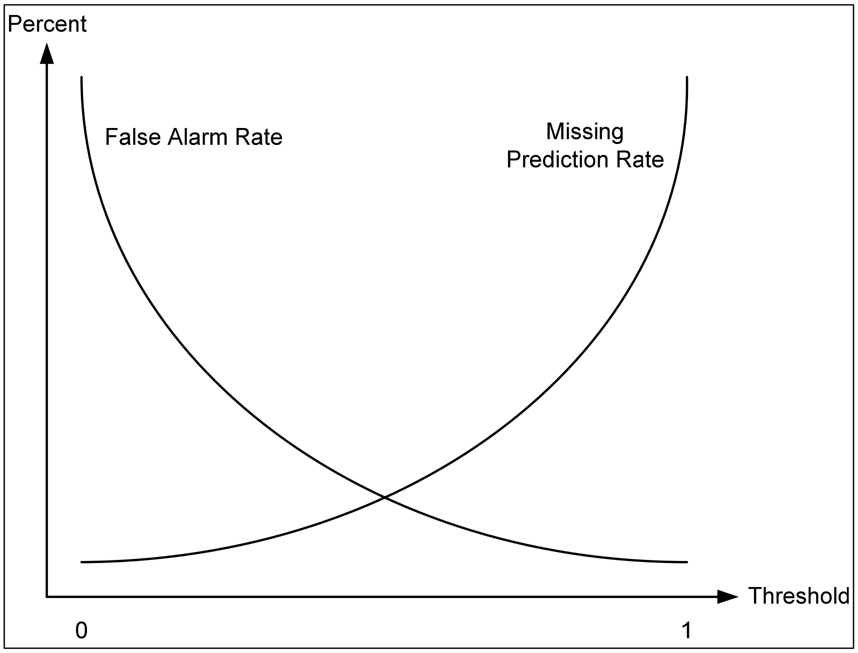


Figure 60. Missing Prediction and False Alarm Rates versus Score Thresholds.

The false alarm rate is defined as a ratio of the number of false predictions to the number of actual incidents. The missing prediction rate is defined as a ratio of the number of incidents that cannot be predicted in advance to the number of actual incidents.

Based on Figure 60, TMC managers can specify the critical threshold with respect to their incident management objectives. For example, if the objective is to balance the false alarm and missing prediction rates (in other words, both rates are considered equally important), the critical threshold should be specified where these two curves intersect. If control centers are constrained by manpower resources for monitoring, the threshold could be moved further to the right to decrease the false alarm rates. On the other hand, if the objective is to place more importance upon the ability to predict incidents in advance, the threshold should be shifted toward the left.

Although the performance-based approach would provide a well-established and defendable means to specify a critical threshold, it does require extensive data collection and observation in the field for threshold calibration. Therefore, the percentile-based approach was suggested for the specification of critical thresholds at the current stage of this system.

PROTOTYPE DEVELOPMENT

First, data management is discussed. Then, system design and development including the real-time implementation of the models are explained. Design considerations for real-time system implementation learned from this study are provided.

System Design and Development

The system was developed in a Visual Basic environment to take advantage of its database capability. The design consists of a main module that controls and interacts with three submodules: weather/environmental, database, and computation. The system architecture is depicted in Figure 61.

The weather/environmental module accesses the weather data via the Internet and extracts required inputs for the model. It also determines lighting conditions by using latitude and longitude to estimate sunrise and sunset times. The outputs from this module are directly passed onto the computational module.

The database module provides connectivity to three data sources. In the field, the system will retrieve the traffic data from the ATMS software. However, in order to test the functionality of the system without interrupting the actual control center operation, we designed the system to have an option for retrieving simulated traffic data. We selected VISSIM as our software of choice in this study to generate freeway traffic data in a hardware-in-the-loop (HITL) manner. The HITL design and implementation are discussed in subsequent sections. This module also interacts with the database to collect intermediate outputs internally and to store final outputs into the central database.

The computational module calculates loop-related traffic measures in real time. Both traffic and weather/environmental inputs are used to predict the probabilities of incidents following the models described above. The hazard score is also calculated using the Euclidean distance approach.

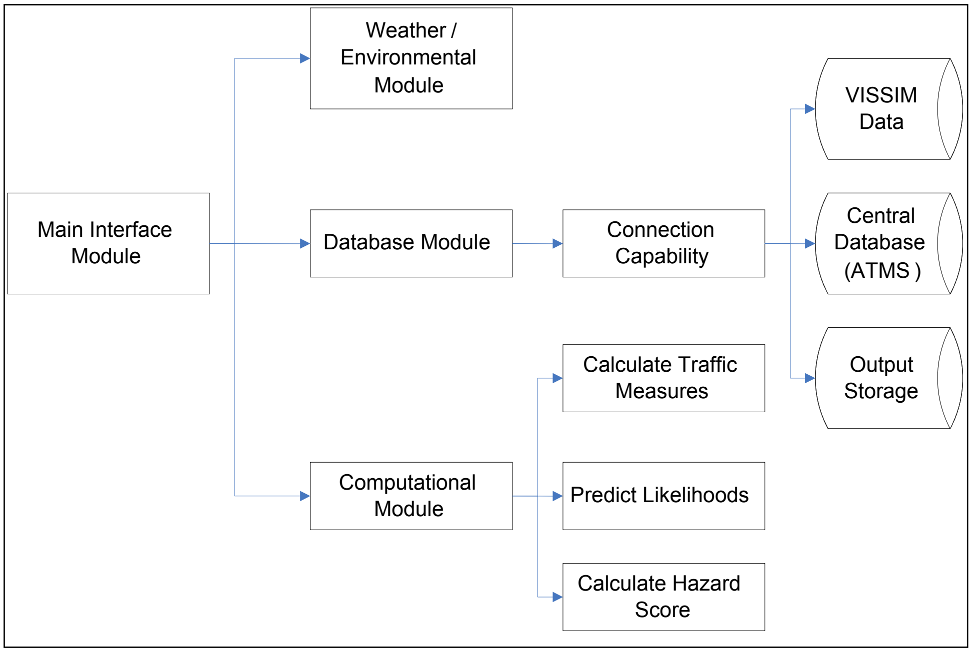


Figure 61. Prototype System Architecture.

The prototype was implemented in Visual Basic environment with ActiveX Data Object (ADO) capability. The main graphical user interface (GUI) is shown in Figure 62. This interface provides basic information about current weather information and the last time that the weather and predicted values were updated. It also shows predicted probabilities corresponding to the detector station ID selected by the users.

From this main GUI, the users can turn on the graphical display (see Figure 63) of real-time predicted likelihoods and hazard scores by checking the *Display Live Forecast* checkbox. The users can also view the real-time traffic measures (see Figure 64) derived from live loop detector data by clicking the *Live Loop Data* command button.

The graphical display shown in Figure 63 consists of two graphs updated in real time. The top one displays the calculated hazard scores across all the detector stations at the most recent time. Control center operators can use this information to monitor area-wide freeway networks and quickly determine where the current traffic and weather conditions are likely to produce incidents. The bottom graph displays the profiles of predicted probabilities and hazard scores over the last 15 minutes at the selected detector station. The user can select the detector station of interest from the main GUI. Three predicted components are displayed in this graph: 1) probability of an in-lane incident, 2) probability of an incident being a collision given that it has occurred, and 3) combined hazard score.

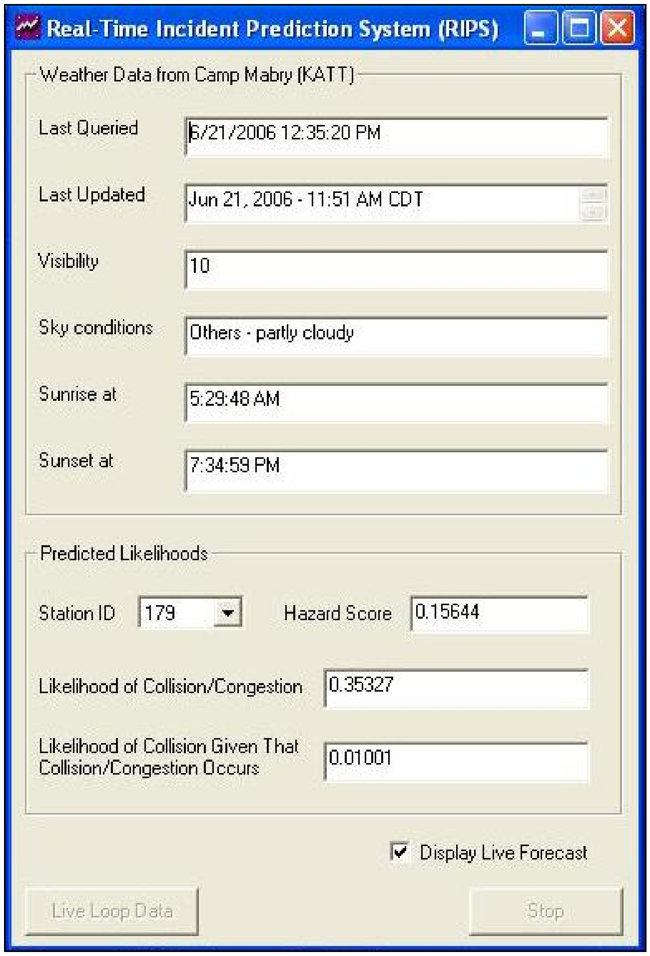


Figure 62. RIPS—Main User Interface.

The real-time computed traffic measures shown in Figure 64 allow the users to view the actual average flow, occupancy, speed, and coefficient of variation in speed at all the detector stations. Control center operators can use this real-time traffic data together with the forecast values to support their decisions concerning appropriate freeway management strategies.

The system checks at a fixed interval if the traffic data have been updated. The system is designed to execute, update, and store the results only when there are new incoming data; otherwise, the system will stay in a rest mode until the next checking period.

At the code level, the program was developed using a concept of classes to allow easy code maintenance and facilitate the integration process with different modules.

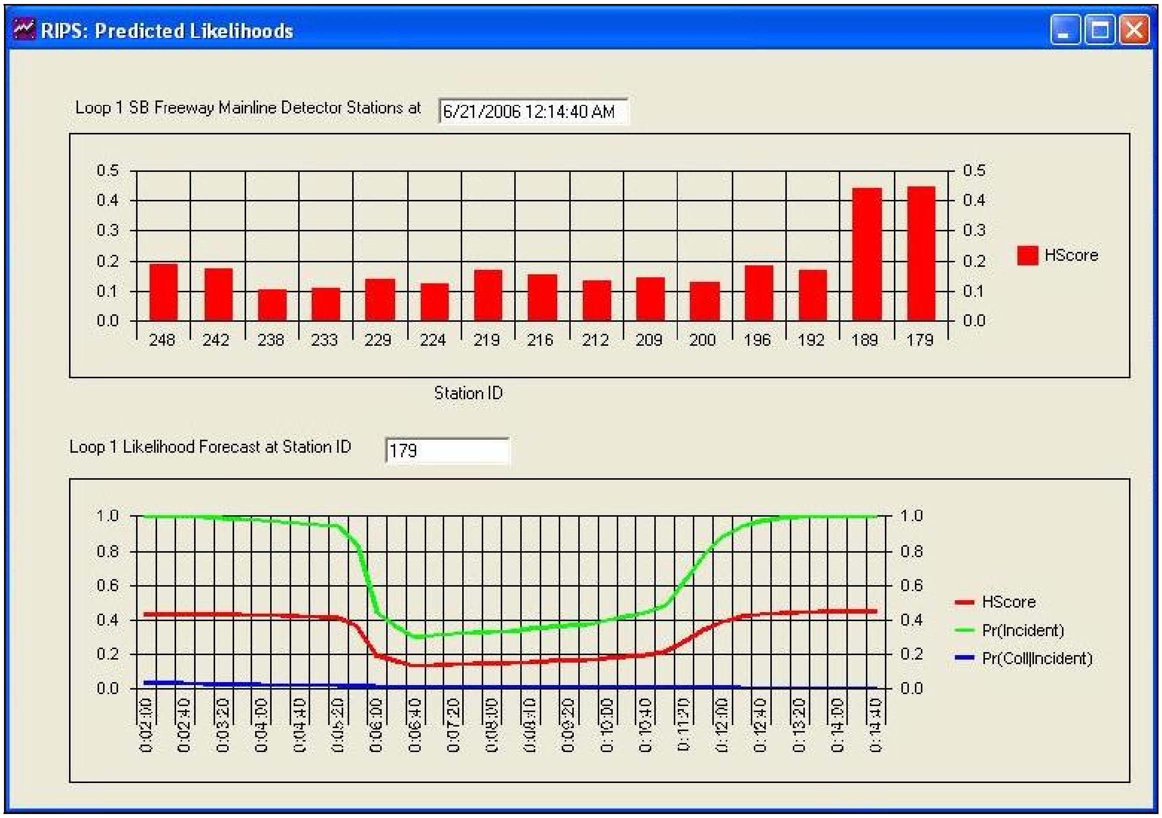
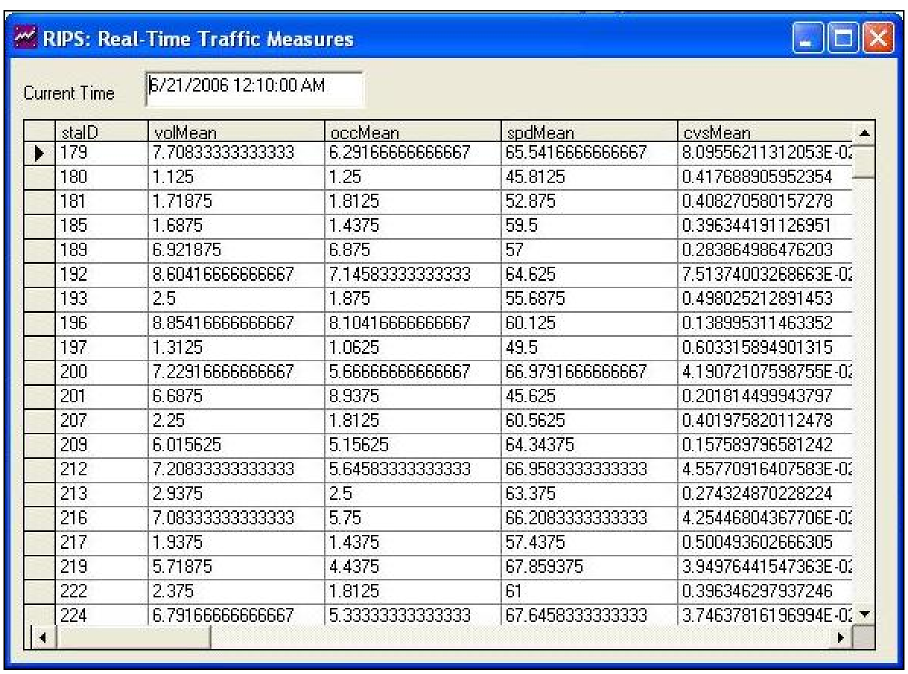


Figure 63. RIPS—Predicted Likelihoods.



RIPS: Real-Time Traffic Measures

Current Time 6/21/2006 12:10:00 AM

staID	volMean	occMean	spdMean	cvsMean
179	7.70833333333333	6.29166666666667	65.5416666666667	8.09556211312053E-0
180	1.125	1.25	45.8125	0.417688905952354
181	1.71875	1.8125	52.875	0.408270580157278
185	1.6875	1.4375	59.5	0.396344191126951
189	6.921875	6.875	57	0.283864986476203
192	8.60416666666667	7.14583333333333	64.625	7.51374003268663E-0
193	2.5	1.875	55.6875	0.498025212891453
196	8.85416666666667	8.10416666666667	60.125	0.138995311463352
197	1.3125	1.0625	49.5	0.603315894901315
200	7.22916666666667	5.66666666666667	66.9791666666667	4.19072107598755E-0
201	6.6875	8.9375	45.625	0.201814499943797
207	2.25	1.8125	60.5625	0.401975820112478
209	6.015625	5.15625	64.34375	0.157589796581242
212	7.20833333333333	5.64583333333333	66.9583333333333	4.55770916407583E-0
213	2.9375	2.5	63.375	0.274324870228224
216	7.08333333333333	5.75	66.2083333333333	4.25446804367706E-0
217	1.9375	1.4375	57.4375	0.500493602666305
219	5.71875	4.4375	67.859375	3.94976441547363E-0
222	2.375	1.8125	61	0.396346297937246
224	6.79166666666667	5.33333333333333	67.6458333333333	3.74637816196994E-0

Figure 64. RIPS—Computed Traffic Measures.

Output Interpretation

First, control center operators must define critical thresholds for hazard scores. A percentile-based threshold selection approach was used to define thresholds as discussed earlier. To illustrate this procedure, hazard scores were calculated using randomized historical traffic, weather, and environment data. To ensure appropriate representation of field conditions, 5 percent of incident-affected and 95 percent of incident-free data were randomly sampled from historical observations. This proportion is consistent with the ratios used to adjust outcome probabilities.

Utilizing the randomized historical data, two predicted probabilities were calculated using the calibrated binary logit models. The hazard scores were then calculated using Equation C-24. The 90^{th}, 95^{th}, and 99^{th} percentiles of hazard scores were found to be 0.10, 0.16, and 0.28, respectively. If control center operators decided to specify the lower and upper thresholds of the hazard score to be 0.16 and 0.28, respectively, this would imply that the RIPS is likely to produce forecast values that may require the following to immediately assess the current conditions and possibly initiate appropriate freeway management strategies: 1) 5 percent of the time for control center operators to be alerted to potential incidents once the lower threshold is exceeded and 2) 1 percent of the time for control center operators once the upper threshold is exceeded.

To reduce the rate of false alarms, a persistence check can be applied before an alert can be issued. For instance, the control center operators may require the lower and upper thresholds to be exceeded continuously for more than 3 and 2 minutes, respectively.

The hazard score thresholds of 0.16 and 0.28 are recommended based on historical observations of Loop 1 and US 183 in Austin, Texas. These thresholds are considered preliminary and should be used only if there is no better information available. These thresholds are expected to be updated over time to reflect the current freeway conditions once the system has been deployed.

Deployment Considerations

Primary concerns related to deployment of the RIPS have been identified during the prototype development of RIPS, which are as follows:

- data integrity,

- data management, and
- communications (data flows).

Preliminary evaluation of Austin loop detector data indicated two prevailing issues related to data integrity: 1) missing data and 2) erroneous data.

Common causes of missing data are loop failure and malfunctioned data communication. Missing data can be detected by examining the archived loop data for a series of consecutive values of -1 or 0 for an extended period of time. In Austin, missing data were reported when the observed occupancy values exceeded the maximum threshold specified for the detector (e.g., 25 percent). This partly explains why missing data frequently appear during the congested period of certain freeway segments in Austin.

Erroneous data occur when observed data are different from actual values. There exist numerous patterns of erroneous data, and only some can be easily identified. Several scenarios of erroneous data are quite subtle and therefore difficult to detect programmatically. In addition, it is fairly complicated to enumerate all the check scenarios for erroneous data. Instances of erroneous data include:

- invalid combinations of volume, occupancy, and speed values;
- constrained speed observations;
- undercounting and overcounting of traffic flows; and
- repetitive data patterns.

Invalid combinations of volume, occupancy, and speed values can be determined based upon general knowledge of traffic flow properties; for example, observed data can be flagged as erroneous if a positive speed value is reported while the occupancy is zero. Another scenario of erroneous data is constrained speed observations where the speed profile appears to be normal but the observed speed values are significantly below the anticipated segment speed. For instance, the detectors at Lazy Lane on US 183 observe average speeds below 35 mph throughout the entire day (see Figure 65).

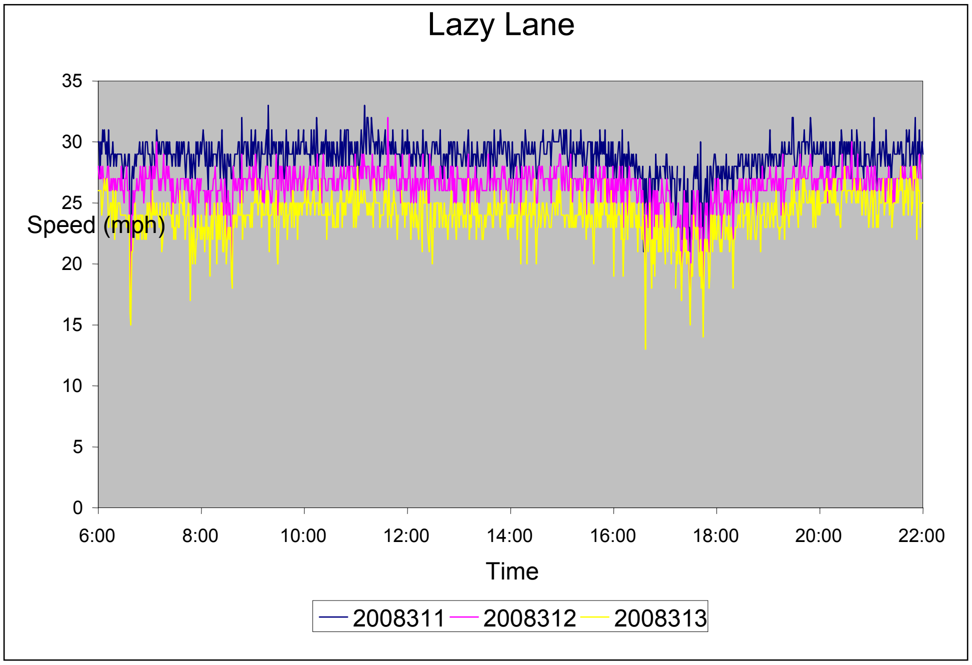


Figure 65. Speed Profiles of Detectors at Lazy Lane.

Some erroneous data checks, such as detector undercounting and overcounting, require the examination of the vehicle conservation property at multiple successive detectors. Let us denote a flow-in-process as the difference between total in-flow and out-flow at two consecutive detector stations. One would expect to observe the flow-in-process to oscillate around zero over time if the detectors are functioning properly. However, this is not always the case as shown in Figure 66 where the flow-in-process between Lamar Boulevard and Guadalupe Street on US 183 in Austin was drifting over time.

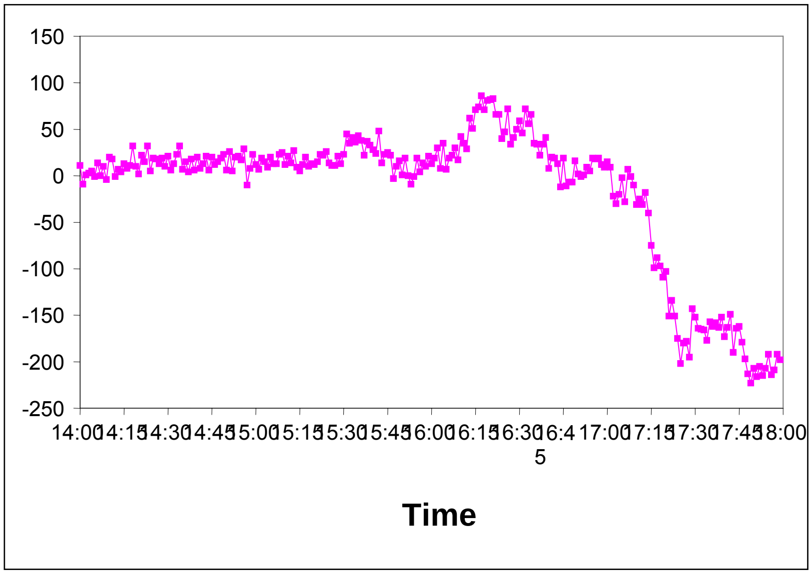


Figure 66. Example of Drifting Flow-in-Process.

Data management must be properly designed to handle growing flows of incoming data such as loop and weather data as well as system outputs. Issues to be addressed include how long the incoming data should be archived for quality control, how long the prediction outputs should be kept for measuring and fine-tuning system performance, and what database management system (DBMS) is cost-effective and well suited with the current ATMS architecture and future data requirements.

Real-time data communications require the consideration of the following issues: 1) time synchronization, 2) time lag, and 3) processing speed.

RIPS interacts with multiple data sources, and each has its own registered time stamps and archived intervals. In this case, there are three sources of time stamps—loop detector data, weather records, and environment data. Common time reference must be used to flag all the incoming data as well as prediction outputs.

Time lag is defined as a delay period from the moment that the actual event has occurred to the moment that the system registers the event. The event could be any data sources such as vehicle speed or weather condition.

Processing speed is inversely proportional to the size of the network handled by the RIPS. In order to keep the RIPS running in real time, additional computing resources must be sought as the number of detector stations grows. Dedicated computing resources for each freeway segment might be another viable solution to this problem as well.

The above issues have not been effectively addressed in the current implementation of this prototype.

REFERENCES

1. C. Lee, F. Saccomanno, and B. Hellinga. "Analysis of Crash Precursors on Instrumented Freeways." *Transportation Research Record 1784*, 2002, pp. 1-8.

2. C. Lee, F. Saccomanno, and B. Hellinga. "Real-Time Crash Prediction Model for Application to Crash Prevention in Freeway Traffic." *Transportation Research Record 1840*, 2003, pp. 67-77.

3. M. Abdel-Aty, A. Pande, N. Uddin, H. Al-Deek, and E. Radwan. *Linking Crash Patterns to ITS-Related Archived Data*. Final Report, University of Central Florida, Orlando, 2004.

4. P. Songchitruksa and K. Balke. *Assessing Weather, Environment, and Loop Data for Real-Time Freeway Incident Prediction.* Submitted for possible publication to the Transportation Research Board, 2006.

5. C. Oh, J. S. Oh, S. G. Ritchie, and M. Chang. *Real-Time Estimation of Freeway Accident Likelihood*. Technical Report, Institute of Transportation Studies, University of California, Irvine, December 2000.

6. A. Drakopoulos, M. Shrestha, and E. Ornek. "Freeway Crash Timeline Characteristics and Uses," *Transportation Research Record 1748*, 2001.

APPENDIX D: DEVELOPMENT OF CRASH POTENTIAL MODEL

The purpose of this model is to correlate some traffic precursors with incident occurrence so that estimation of incidents for various traffic conditions is possible. The model is classified in statistical terms as a categorical log-linear model. A detailed description of the model components, its development, and real-time working procedure is given in this section.

MODEL DESCRIPTION

Selection of Crash Prediction Precursors

A precursor is a variable that is derived from traffic stream data and whose variations can indicate or point to a desirable pattern in traffic flow behavior. Recent research in incident prediction has widely used the concept of precursors in models for predictions. Several researchers have worked with various precursors and tested the potential of those precursors for incident prediction.

The first task in this research was to identify the appropriate incident precursor. A detailed review of the literature on incident prediction/precursors was carried out. In this review several traffic flow variables have been considered for their potential as incident precursors, such as density, speed, CVS, flow, volume headways, etc. Every study in this process was reviewed for the extent of predictability and was accordingly classified as positive ("+") if the result showed satisfactory correlation between incident precursors and occurrence of incidents, negative ("-") if no correlation was found, and zero ("0") for undeterminable cases. It was found that a greater amount of research was done with CVS as a precursor than any other precursors. Some studies even distinguish between CVS along the lane and across the lanes. But most studies refer to CVS along the lanes.

Traffic volume has traditionally been a precursor of interest to many researchers for statistically relating to crash frequency. This precursor has been statistically quite significant, and research using models with volume is capable of describing 60 percent of the incidents *(1)*. However, the longer aggregation time involved in deducing traffic volumes has been a restriction of using volume as a precursor in real-time prediction systems. Volume has been predominantly useful in highway or intersection safety-oriented studies.

Hourly flow, which is a shorter time aggregation of volume, is another precursor that has been used by some researchers to predict accident rate. The results of models involving hourly flow have indicated a definitive correlation between hourly flow and accident rate; for example, in the work of Hiselius *(2)*, an increasing rate of accidents with hourly flow is indicated. It has further been investigated by segregating vehicle types. In the case of cars, there is a constant increase in accident rate with hourly flow, but in the case of trucks there is a decreasing rate with hourly flow. Another study *(3)* also affirms that hourly flow provides a better understanding of interactions like incidents. However, there has not been elaborate work on hourly flow as a precursor and a convenient prediction model for real-time applications.

Time headway has been tried as a causal precursor, and it has been shown that shorter headways have been reasons for collisions *(4)*. But again, there has been no convincingly explanatory model for use of this precursor in real-time incident prediction systems.

A precursor that has been widely used and found to be sensitive in accident prediction is the coefficient of speed variation along the lane *(5-9)*. Most of the models using CVS as a precursor have a very satisfactory correlation with accidents. Abdel-Aty et al. *(8)* have documented that their model's crash prediction level was around 62 percent, and similarly convincing results have been reported in most of the other models. In all the studies involving CVS, the provision to aggregate the precursor values over an optimally small time period has been an advantage in sensing and predicting the variation of traffic behavior. However, a study *(9)* has taken a totally opposite stand with regard to the ability of CVS to serve as an incident precursor. Several reasons can be attributed to such a difference in conclusions, such as method of obtaining speed data, aggregation interval, and statistical methodology. But in comparison to other precursors, CVS is the most likely choice of precursor for further investigation for use in real-time prediction models.

Traffic density is another parameter that has a good correlation in explaining incidents; traffic density is usually used in conjunction with CVS *(7, 8)*. After a careful review, density and CVS along the lanes have been chosen as the potential candidates for further investigation and use in the model for this project. A table summarizing the review of candidate incident precursor is given in Table 12.

Table 12. Review of Precursors Used for Incident Prediction.

Precursors	Number of Studies Reviewed	Positive Results	Negative Results	Neutral/Weak Results
Speed variation along the lane	8	6	2	-
Speed variation across the lanes	1	-	-	1
Occupancy or density	2	2	-	-
Volume	2	2	-	-
Hourly flow	2	2	-	-
Headway	1	1	-	-

Hence, for this module, CVS and density will be used as the primary precursors. Apart from these two primary precursors, two indicators, namely peak-hour indicator and roadway type indicator, will be used. These precursors will be defined according to Lee et al. *(6)*.

Defining the Precursors

One-minute aggregated data that were archived by TxDOT were used in deriving the precursors for this module. With the speed data available from the detectors for every minute, a moving-average coefficient of variation of speed was calculated as given by Equation D-1. The CVS was calculated over a 5-minute period, starting from the time interval against which the precursors are reported and including the preceding four intervals.

$$CVS = \frac{\sqrt{\dfrac{n\sum_{1}^{n}(S_i)^2 - \left(\sum_{1}^{n} S_i\right)^2}{n(n-1)}}}{\dfrac{\sum_{1}^{n} S_i}{n}} \qquad \text{(D-1)}$$

where

S_i = speed in miles per hour at time t and

n = number of time intervals.

Average occupancy, which is another precursor, is simply calculated as a moving average of the 1-minute occupancy available in the data set. The average occupancy was calculated over a 5-minute period consistent with the CVS aggregation interval.

The precursors obtained were tagged with either "peak" or "non-peak" for time-of-day indication. Precursors occurring anytime between 6:30 a.m. to 9:30 a.m. and 4:00 p.m. to 7:00 p.m. were considered peak, and precursors obtained at other times were considered non-peak. The roadway-type factor does not vary with each and every recording of data but is fixed for every detector location. The roadway type was classified as "straight" or "other" for each detector considered in this study. This classification was based on the horizontal alignment of the roadway section and the presence of ramps near the detector station. A detector station on a straight alignment and far from the influence area of the ramps is tagged as "straight"; otherwise, it is tagged as "other."

Model Formulation

A categorical log-linear model was chosen to predict the likelihood of crash rate using the selected precursors. As indicated earlier, the proposed model very closely follows that suggested by Lee et al. *(7)*, except that one of the additional precursors used by Lee et al. *(7)*—the speed difference between the two adjacent detectors along a lane (i.e., lateral CVS)—has not been considered in this model. The model correlates the expected number of crashes on any section of a freeway to the combined effect of the categorical precursors that prevail 5 minutes before the time of prediction. The functional form of the model is in Equation D-2.

$$N / EXP^{\beta} = f\left(C * \lambda_{CVS(i)} * \lambda_{Occ(j)} * \lambda_{R(k)} * \lambda_{P(l)}\right) \qquad \text{(D-2)}$$

where

N = the expected number of crashes over the analysis time frame,

EXP = the exposure in vehicle-kilometers of travel,

C = constant,

$\lambda_{CVS(i)}$ = effect of the crash precursor variable *CVS* having *i* levels,

$\lambda_{Occ(j)}$ = effect of the crash precursor variable *Occ* having *j* levels,

$\lambda_{R(k)}$ = effect of road geometry (control factor) having *k* levels,

$\lambda_{P(l)}$ = effect of time of day (control factor) having *l* levels, and

β = coefficient for exposure.

MODEL CALIBRATION

Lee et al. *(7)* found that a proportion of 50:30:20—low, medium, and high values of precursors, respectively—gave the best fit for the categorical model that was developed in their study. For each of the precursors, CVS, and occupancy, the boundary values for the lowest 50 percent of the precursor values and next 30 percent of the precursor values will be determined from the processed detector data. The boundary values used to determine the categories are given in Table 13.

Table 13. Boundary Values for the Precursors.

Category	CVS	Occ (%)
L (50) or 1	<= 0.043	<= 3.6
M (30) or 2	> 0.043 & <= 0.227	> 3.6 & <= 5.8
H (20) or 3	> 0.227	> 5.8

This model is designed so that CVS has categories 1, 2, and 3, wherein 1 represents the low range of speed variation and 3 represents the high range of speed variation. Similarly, occupancy has three categories from 1 to 3, wherein 1 represents the low value of occupancy and 3 represents the high value of occupancy. Roadway type and peak-hour factor have two categories each, as explained previously. Therefore, all 36 different categories are defined by taking a combination of different levels of precursors. For all 36 categories, the number of incidents in the study section during the 2-year study period was extracted from the incident logs, which form the input for the model calibration. The model was calibrated using the maximum likelihood methodology *(10)*. The results of the calibration are given in Table 14.

Table 14. Results of Parameter Estimation for Crash Potential Model.

Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
C	2.693	0.832	3.237	0.001	1.062	4.324
$\lambda_{CVS=1}$	-1.395	0.566	-2.466	0.014	-2.504	-0.286
$\lambda_{CVS=2}$	-0.373	0.357	-1.045	0.296	-1.071	0.326
$\lambda_{CVS=3}$	0(a)	.	.	.	.	.
$\lambda_{Occ=1}$	-2.059	0.299	-6.884	0.000	-2.646	-1.473
$\lambda_{Occ=2}$	-1.632	0.361	-4.522	0.000	-2.339	-0.924
$\lambda_{Occ=3}$	0(a)	.	.	.	.	.
$\lambda_{P=0}$	-0.615	0.301	-2.047	0.041	-1.205	-0.026
$\lambda_{P=1}$	0(a)	.	.	.	.	.
$\lambda_{R=0}$	-0.462	0.173	-2.670	0.008	-0.801	-0.123
$\lambda_{R=1}$	0(a)	.	.	.	.	.
β	0.043	0.099	0.437	0.662	-0.151	0.237

Discussion of Parameter Estimates

The estimated parameters (λ) for the model are shown in Table 14. CVS and Occ parameters with subscript 1 indicate the lowest level of that precursor, 2 indicate medium level, and 3 indicate highest or most severe level, while subscript 0 for peak-time factor (λ_P) indicate a non-peak hour and 1 indicate peak hour. Similarly, subscript 0 for roadway type (λ_R) indicates a straight section of road without any on-ramps or off-ramps. β is a parameter exposure. C is a constant, which in the present model setup can be interpreted as the maximum risk that an incident is predicted. Information obtained in Table 14 is useful in analyzing two different aspects of the model parameter estimates. Firstly, the effect of different categories in a given precursor can be analyzed, and the effect of different precursors to the extent they can influence incident prediction can be analyzed. Secondly, the statistical significance of each of the parameters can be assessed. Both of these kinds of analyses are presented in the following paragraphs of this section.

Physical Interpretation of Estimated Parameters

Any traffic model is verified when it can reflect that which can be observed in real time or when it can explain some observed phenomenon. This kind of verification is called physical interpretation of the model. One category in each precursor is set to zero value. This category is referred to as the aliased cell, and all other estimates will be in reference to the respective precursor's aliased cell. Let us examine the physical significance of each precursor. Parameters for CVS show that the medium and low categories have a negative sign with decreasing value (real scale) as we move from the high to low category. The negative parameter means that when the traffic state transitions from the high category of CVS to the medium category, the risk of incident occurrence decreases by an amount. A further transition from the medium state to low CVS results in a further decrease in accident likelihood. The results agree with the general observations one can make on a highway. Many other studies, too, have shown that as CVS increases, there is a higher probability of incidents than under lower CVS. An approximate quantitative feel for the decrease in the risk is indicated by the numerical values of the estimated parameters. The model also indicates that the reduction in risk of accident is less from high to medium CVS when compared to the reduction in risk from the medium to low CVS categories.

Occupancy, too, shows a trend similar to that for CVS. As we get from higher to lower occupancy, the risk of crash occurrence decreases. However, the calibrated model shows that the reduction in risk of crash occurrence is more significant as we move from a higher occupancy state to a medium occupancy state. And the same is true when there is a transition from a medium occupancy state to a low occupancy state. The results are aligned with practical observations. Accidents on highways are more likely during a high occupancy level than during low occupancy since high occupancy requires high attention from drivers so as not to become involved in accidents. Our model reflects this observation. But, comparing the values of CVS categories with occupancy categories, occupancy is more sensitive in responding to incident occurrence than CVS. The range of estimated values between high and low levels is much larger for occupancy than for CVS, which means that the risk of an accident is reduced even if traffic is operated smoothly by reducing variations in the speed of the traffic stream somehow. In contrast, by reducing the occupancy on a section, there is a high likelihood that risk in number of crashes can be considerably reduced, more than what we could achieve by controlling speed variation.

From the results, we see that the peak-hour factor and roadway-type indicator have a negative value for the estimated parameters with subscript 0. This means that non-peak hours have less chance of incidents to occur than peak hours. Also, we can state from the results that straight sections of roadway, without on-ramps and off-ramps, have less likelihood of incident occurrence than curved sections or sections near ramps. The overall results show that the influences of the roadway-type indicator and peak-hour indicator are less sensitive parameters when compared to CVS and occupancy in influencing the risk of accidents. However, peak-hour factor and roadway type are still significant precursors in incident forecasting.

Statistical Significance of Estimated Parameters

Another way of looking at the significance of the model is by analyzing the statistical results of the parameter estimation. The columns "Z" and "Sig." in Table 14 are the primary indicators to measure the significance of the estimated parameters. Looking at the "Z" values, everything except $\lambda_{CVS=2}$ and β are statistically significant with "Z" values being greater than 1.96 at 95 percent confidence level. Significance values of less than 0.05 are expected to declare any parameter used in the model to be statistically significant for use in the model.

In the above context $\lambda_{CVS=2}$ and β are obviously less significant to be included in the model from a statistical standpoint. However, $\lambda_{CVS=2}$ is still included in the model for two reasons. Reduction in the number of CVS categories would mean a lesser number of traffic states to represent traffic flow, and removal of $\lambda_{CVS=2}$ would reduce the total possible states to two-thirds of the current possible states. Reducing the total number of possible states would make the model less sensitive to the change in traffic state and make it harder to predict the likelihood of accidents, which is not desirable. Justification for retaining the exposure parameter β is made because of the well-established relationship between the number of accidents and the exposure. The number of accidents alone is insufficient in making logical decisions about accident risks. Looking at two roads with a similar number of accidents, if the data for the first road were reported over a 2-year period and the data for the second road were reported over a 1-year period, there is a difference in terms of risk. In the example, the second road has a higher risk or likelihood of incident than the first road. Hence, exposure is an important factor in explaining the risk component in incident reporting. So, though β is statistically very insignificant, it has a greater role for the physical interpretation of the overall model.

Working Procedure in Real Time

A schematic of the working procedure for the incident prediction model for the given minute and given section of roadway is shown in Figure 68. Input requirements for the incident prediction model are the predicted speed and density at a particular section. Average speed and density over all the lanes on a particular section of roadway are considered. The predicted speed and density are obtained from a traffic model. Other information that is required is time of day and the roadway-type indicator. The archived data of projected speed and density for the preceding 4 minutes for that particular section are extracted from the database. The precursors are calculated with this 5-minute traffic data.

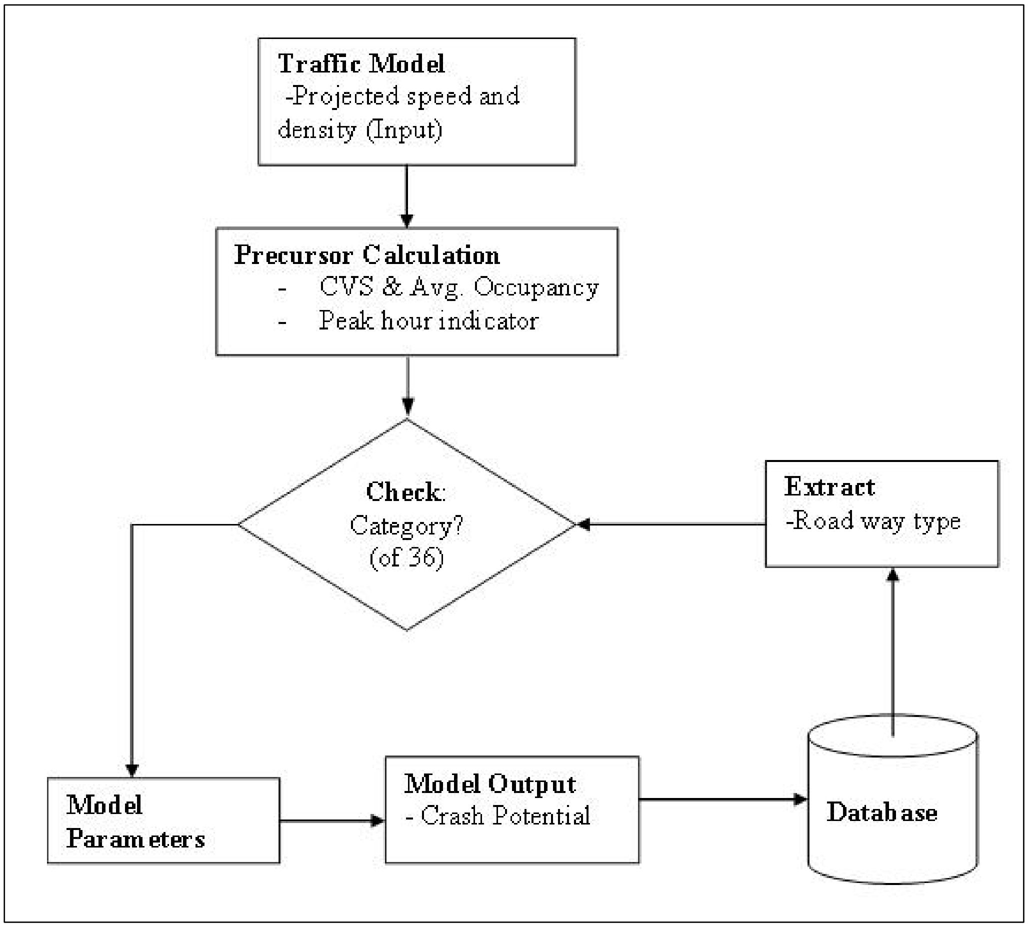


Figure 67. Schematic for Working Procedure of Crash Potential Model.

Once the precursors are determined, a check is made based on the predetermined boundary values to classify the current traffic condition among one of the 36 possible categories (the categories and boundary values were defined earlier). Based on the categories, the respective parameters for the model are considered and crash potential is calculated.

This procedure is illustrated for a single station and for a given minute. The same cycle is carried out for all the detector stations that will be considered for the study and will run for a continuous period of 1-minute intervals. However, in the beginning of the model run, the output will not be generated for the first 4 minutes due to insufficient data for precursor calculation.

OUTPUT INTERPRETATION FOR CRASH POTENTIAL MODEL

Crash potential is defined as the expected number of crashes over a certain period of exposure. Exposure is denominated as vehicle miles of travel. The incident likelihood model as a stand-alone reports crash potential for the particular minute it is evaluating by looking at the precursors over the current time and preceding 4 minutes.

The advantage of this model in predicting future crash potential can be realized by using this model with a traffic prediction model. The process in which the incident likelihood model works in the current setup is depicted in Figure 68.

The following example demonstrates how the crash potential is calculated from the model. Consider that CVS and occupancy are calculated from a 1-minute predicted speed and density on a section of a roadway that is just before an on-ramp during a morning peak hour with an average annual daily traffic (AADT) of 100,000 vehicles. We assume that both CVS and occupancy fall in the lower range as per the respective boundary values defined for this model. Hence, in terms of our model, the next 1-minute traffic condition falls into a category defined as CVS=1, Occ=1, P=1, and R=1.

Equation D-3 shows our formulation of this categorical situation:

$$CP = \frac{e^{(C+\lambda_{\mathrm{CVS(i)}}+\lambda_{Occ(i)}+\lambda_{R(k)}+\lambda_{\mathrm{P(l)}}+\beta\ln(\mathrm{EXP}))}}{\mathrm{AADT}} \tag{D-3}$$

Therefore, the crash potential for the next 1 minute, represented as $CP1$, is shown in Equation D-4:

$$CP1 = \frac{e^{(2.693-1.395-2.059+0+0+0.43\ln(100000))}}{10^5}$$

$$CP1 = 0.00066$$

Similarly, depending on the category in which the predicted traffic condition exists, the model will estimate the corresponding crash potential. However, the threshold in determining if the crash potential should be flagged as high risk, medium risk, or low risk can be determined by doing further field studies and by engineering judgment.

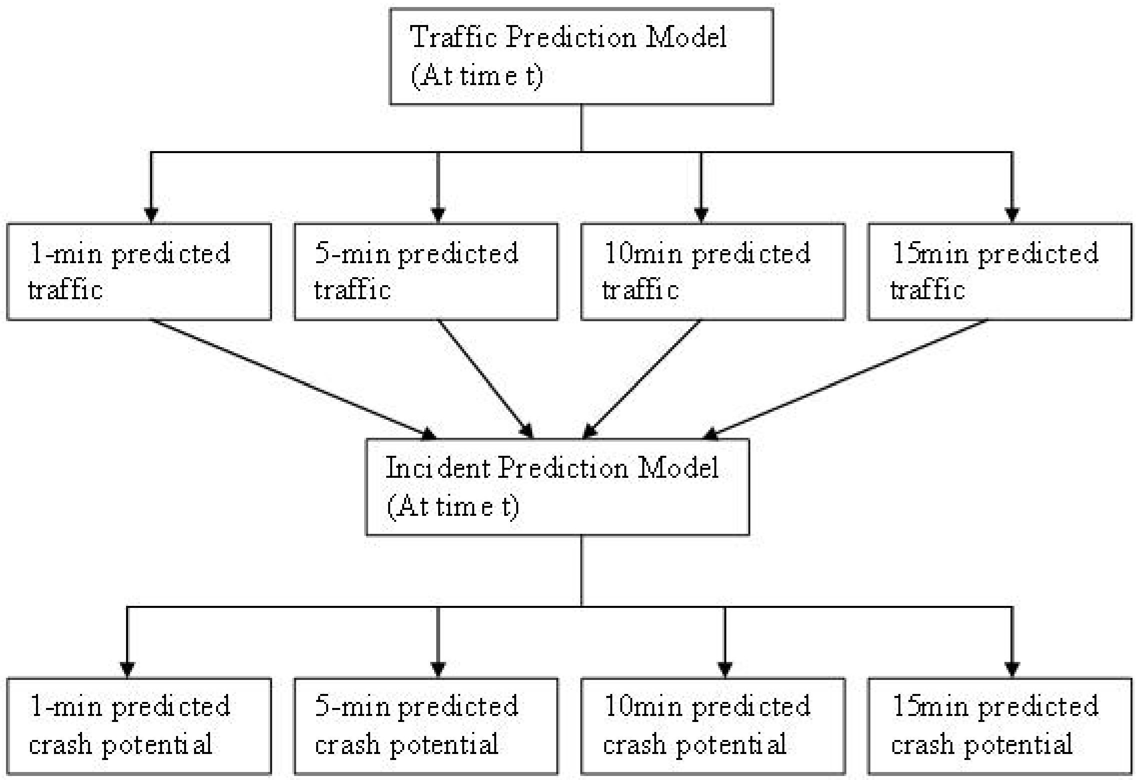


Figure 68. Model Output of Crash Potential Model.

REFERENCES

1. P. Gribe. "Accident Prediction Models for Urban Roads." *Accident Analysis & Prevention*, Vol. 35, 2003, pp. 273-285.

2. L. W. Hiselius. "Estimating the Relationship between Accident Frequency and Homogeneous and Inhomogeneous Traffic Flows." *Accident Analysis & Prevention*, Vol. 36, 2004, pp. 985-99.

3. A. Cedar and M. Livneh. "Relationships between Road Accidents and Hourly Traffic Flow—I." *Accident Analysis & Prevention*, Vol. 14, 1982, pp.19-34.

4. G. A. Davis. "Collective Responsibility in Freeway Rear-End Collisions. An Application of Casual Models." Presented at Transportation Research Board Annual Meeting, Washington, D.C., 2005.

5. C. Oh, J. S. Oh, S. G. Ritchie, and M. Chang. *Real-Time Estimation of Freeway Accident Likelihood*. Technical Report, Institute of Transportation Studies, University of California, Irvine, December 2000.

6. C. Lee, F. Saccomanno, and B. Hellinga. "Analysis of Crash Precursors on Instrumented Freeways." *Transportation Research Record 1784*, 2002, pp. 1-8.

7. C. Lee, F. Saccomanno, and B. Hellinga. "Real-Time Crash Prediction Model for Application to Crash Prevention in Freeway Traffic." *Transportation Research Record 1840*, 2003, pp. 67-77.

8. M. Abdel-Aty, A. Pande, N. Uddin, H. Al-Deek, and E. Radwan. *Linking Crash Patterns to ITS-Related Archived Data*. Final Report, University of Central Florida, Orlando, 2004.

9. K. M. Kockelman and J. Ma. "Freeway Speeds and Speed Variations Preceding Crashes, within and across Lanes." Presented at ITS America 2004, 14th Annual Meeting and Exposition, San Antonio, Texas, 2004.

10. SPSS, Inc. *SPSS 13.0*. Chicago, 2004.