Texas A&M Transportation Institute

NOT FOR DISTRIBUTION – NOT FOR REPRODUCTION

INTERIM MEMORANDUM- DRAFT

Air Quality Performance Measures – Investigating the Use of STARS II in Mobile Source Air Quality Modeling

DRAFT FOR REVIEW

Air Quality Performance Measures – Investigating the Use of STARS II in Mobile Source Air Quality Modeling

Air Quality and Conformity Inter-Agency Contract

Subtask 2.3 – FY 2015

Prepared for

Texas Department of Transportation

By

Texas A&M Transportation Institute

August 2015

TECHNICAL MEMORANDUM – DRAFT FOR REVIEW

Inter-Agency Contract (Contract No: 50-4XXIA032)

Subtask 2.3 Air Quality Performance Measures – Investigating the Use of STARS II in Mobile Source Air Quality Modeling

DATE: August 9, 2016

TO: William Knowles, P.E.
Texas Department of Transportation (TxDOT)

COPY TO: Janie Temple, TxDOT
Laura Norton, TxDOT
Justin Malnar, Research Development Office, TTI

FROM: Reza Farzaneh, Ph.D., P.E.
Andrew Birt, Ph.D.
Suriya Vallamsundar, Ph.D.
Joe Zietsman, Ph.D., P.E.
Environment and Air Quality Division,
Texas A&M Transportation Institute

FOR MORE INFORMATION:

Reza Farzaneh
(512) 407-1118
Reza.Farzaneh@tti.tamu.edu

Table of Contents

INVESTIGATING THE USE OF STARS II IN MOBILE SOURCE AIR QUALITY MODELING

1. Project Goals and Outline

This memorandum summarizes work performed by Texas A&M Transportation Institute (TTI) staff in Fiscal Year (FY) 2015 on investigating the use of Texas Department of Transportation's (TxDOT) Statewide Traffic Analysis and Reporting System II (STARS II) for mobile source air quality modeling and related uses. This work was conducted under the TTI-TxDOT Air Quality and Conformity Interagency Contract.

The goal of this study was to investigate a number of potential applications of the STARS II database for transportation air quality analyses and to develop one of these opportunities into a tangible application of the systems' data.

This report proceeds by providing a brief overview of the STARS II system, and a discussion of how traffic data are important for air quality research. The report then presents several big picture ideas for how data from the system could be used for air quality analyses and related uses. The concluding section discusses the ideas generated and the merits of the case study in the light of current air quality analysis trends.

2. Overview of STARS II

STARS II comprises a database of traffic activity and a web-based geographic information system (GIS) interface that can be used to search for and download traffic count data. STARS II is an incremental upgrade of information systems that have been used to monitor and analyze traffic to support decision making within TxDOT and other agencies. The Transportation Planning and Programming Division of TxDOT manages TxDOT's counting program. The program collects traffic count data in the form of:

- Traffic volumes.
- Traffic volumes classified by vehicle type.
- Vehicle speed.
- Vehicle weight.

These traffic data are collected at approximately 362 permanent count locations that are strategically distributed across the state, and from temporary locations that are used within a dynamic sampling plan that covers approximately 75,000 to 80,000 locations annually. The results of these traffic count operations are stored in STARS II where the data are validated, analyzed, and made available for decision support.

Although traffic count data are essential for supporting TxDOT's internal decision making, they also support decision making at a federal level. STARS II addresses a federal requirement that all states implement traffic monitoring system for highways and transportation facilities and equipment (23 CFR 500 part B). The Federal Highway Administration (FHWA) also mandates that states must report standardized traffic information to the Highway Performance Monitoring System (HPMS). The HPMS was initiated in 1978 to collate and distribute data describing the extent, condition, performance, use, and operating characteristics of the nation's highways. The standardized methods for collecting and reporting data to the HPMS makes it the most comprehensive system available for obtaining information about the physical condition of the U.S. highway system. HPMS is used by FHWA to support decisions on the physical condition, safety, service, and efficiency of the national highway system, but is used by organizations such as Environmental Protection Agency (EPA), Metropolitan Planning Organizations (MPOs), and transportation researchers.

HPMS provides comprehensive and standardized data, but it does this at a necessary cost of reducing the detail in some traffic count information. STARS II supports a broader range of traffic measurements, at increased spatial and temporal resolutions than are available within HPMS. It also includes a web-based interface that allows users to rapidly query its data. For these reasons, STARS II has potential to provide high quality traffic information to a range of decision makers.

3. Traffic Data and Air Quality Analysis

Traffic activity data are essential for transportation air quality analyses. This section outlines a number of observations about transportation air quality research methods and trends, and why and how traffic activity data are important to support this research:

Mobile source air quality analysis uses a bottom-up, modeling framework: Transportation air quality analyses most often uses a bottom-up modeling approach to estimate and understand the impacts of transportation on air quality (Figure 1). In this bottom up approach, traffic activity rates are used with emission rates to estimate total mobile source emissions. In some cases, impacts are assessed based on these total emissions, but can also be used as the inputs to dispersion models to estimate the concentration of pollutants through time and space. This bottom up modeling approach is necessary because, although the concentration of pollutants in air can be accurately measured, the rapid dispersion of pollutants in the atmosphere makes it difficult to apportion them to their sources. Although emissions from transportation are an important component of the air quality of a region, they are not the only sources. Other significant pollutant point and area sources include industry, agriculture, and stationary or non-road engines.

The importance of the regulatory process in transportation air quality research: The methods used in transportation air quality analyses have been significantly influenced by regulatory decision making. Although regulatory analyses are an important component of air quality analysis, the majority of air quality analyses actually involve non-regulatory work. Nevertheless, regulatory issues have had a large impact on how non-regulatory air quality analyses are conducted. For example:

a. The regulatory process has driven the development of standardized methods and tools for modeling transportation air quality. Like many other regulatory frameworks, these methods are designed to standardize air quality impact assessment so that results are interpretable, and are easier to use in regulatory decision making.
b. Because traffic activity is an essential component of many analyses, many of the standard methods used to assess air quality impacts are designed in line with the availability of standardized traffic activity data. An example is provided by the widespread use of the EPA Motor Vehicle Emissions Simulator (MOVES) emission rate estimation model, which has been designed to use relatively standard traffic input data (Table 1).
c. National Ambient Air Quality Standards (NAAQS) and non-attainment areas are often the focus of air quality studies. By definition, these areas have the greatest need for air quality work. However, in line with societal concerns about air quality, there is an increasing trend in some areas toward air quality impact

assessments for other pollutants, or focused on more localized or specific air quality issues.

Increasing societal awareness of air quality: This awareness drives a continual need for higher accuracy models and data for air quality analysis, and increased efficiency in assessing air quality impacts. Air quality research must meet the challenge of increasingly stringent air quality goals while at the same time maintaining or improving the efficiency of transportation infrastructure. This requires refinements to analysis methods that allow impact assessments at finer spatial and temporal scales, or that use more explicit representations for how traffic emissions impact air quality (e.g., dispersion model). For these refinements, high quality traffic activity data are especially important.

Traffic activity modeling is a scientific field in its own right: Although traffic activity data underpins much of transportation air quality analysis, understanding and forecasting traffic activity is normally considered an independent field. Most MPOs develop regional travel demand models for planning purposes. These are often used to provide traffic activity data for air quality analyses within the same region. Other research tools used in air quality analysis include traffic simulation models. In practice then, transportation air quality analysis already uses data from TxDOT's traffic data collection program. However, this study provides an opportunity to rethink how traffic activity data can be used in air quality studies, and how increased coordination between individuals involved in traffic monitoring and modeling (including STARS II) can improve air quality analysis.

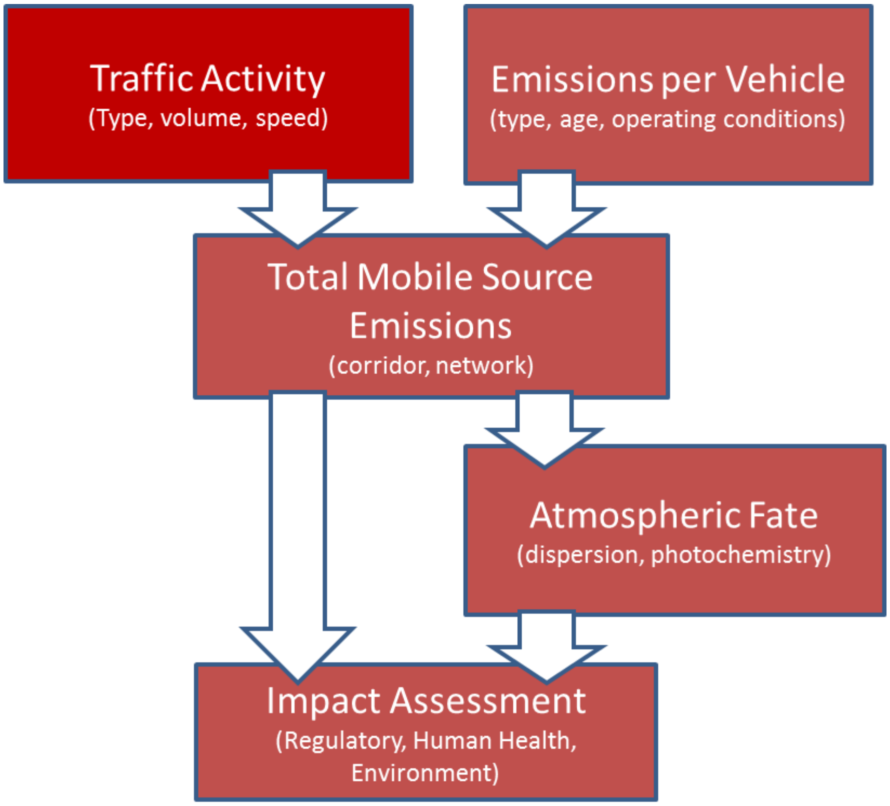

Figure 1 – Overview of transportation air quality analysis illustrating that Traffic Activity data underpins much of Transportation Air Quality research.

Table 1 – Example of the data inputs to MOVES 2014 emission estimation model. Highlighted cells illustrate data types that could be furnished by STARS II.

Data Component	Data Source
External Conditions	
Year of evaluation	User input
Month of evaluation (January, July)	User input
Altitude (high, low)	Meteorological data
Temperature (daily range or ambient temperature for each hour)	Meteorological data
Humidity and solar load	Meteorological data
Vehicle Fleet Characteristics	
Vehicle age distribution	Vehicle registration data from TxDOT
Annual mileage accumulation by vehicle class	Defaults
Diesel sales fractions by vehicle class and model year	Vehicle registration data from TxDOT
Natural gas vehicle fractions	Vehicle registration data from TxDOT
Vehicle Activity	
Vehicle miles traveled (VMT) fraction by vehicle class	VMT mix from classification data
VMT fraction by highway functional system	From link-based model
VMT fraction by hour of day	From automatic traffic recorder (ATR) data
VMT fraction by average speed	Based on link speeds/hour
Source Hours Idling (extended)	Truck VMT and Texas specific studies
Vehicle starts per day by vehicle class and hour	Defaults
Refueling	VMT mix, Vehicle registration
Source Hours Parked	Vehicle registration, VMT mix, speed
Vehicle hot soak time between engine starts	Defaults
Vehicle hot soak time after engine shut down	Defaults
Vehicle diurnal soak time	Defaults
Vehicle trip length (duration) distribution	From travel demand model
Weekend/weekday	User input
Vehicle Gasoline Specifications	
Fuel characteristics (Reid vapor pressure, sulfur content, oxygenate content, fuel volatility, etc.)	Information from TCEQ

4. Qualitative Methods

This section summarizes the results of a qualitative analysis designed to investigate how the STARS II system can be used in novel ways to support Air Quality analysis. The method involved a structured brainstorm session that included participation by 10 transportation analysts. The participants represented a broad range of expertise within the air quality arena: **Joe Zietsman** (air and environmental quality); **Ed Hard** (travel demand modeling); **Brian Bochner** (transportation engineering and air quality policy); **Dennis Perkinson** (emissions estimates for conformity); **Reza Farzaneh** (emission estimates); **Tara Ramani** (air quality and policy); **Suriya Vallamsundar**, (emission estimation, dispersion modeling); **Chaoyi Gu** (emission estimation, MOVES); **Jeremy Johnson** (emissions testing); and **Andrew Birt** (systems modeler, moderator).

The participants were provided with a brief overview of the STARS II system, which outlined the types of data it contains, and the spatial and temporal coverage of these data. The goal of the brainstorm session was to promote the generation of big picture ideas for analysis applications that could be developed now or in the future to support air quality analysis. These broad-scale ideas would later be evaluated based on the feasibility of the idea given a more exhaustive evaluation of the STARS II database, and the perceived value to air quality analyses.

Participants were asked to contribute ideas in the form of the matrix illustrated in Table 2.

Table 2 – Matrix used to record ideas in the brainstorm session.

What is the idea?	**Who is it for?**
Idea 1	Idea 1
Idea 2	Idea 2
…..	…..
Idea n	Idea n
Why is it important?	**What other data is important?**
Idea 1	Idea 1
Idea 2	Idea 2
….	….
Idea n	Idea n

4.1 Qualitative Results

Several different potential analyses and applications were identified through this effort. Some were very specific. Others are more general. Some could be developed with data existing in the STARS II database while others may require at least some additional data (e.g., coverage, duration, frequency). These opportunities are described below.

4.1.1 Periodic, Automatic Emissions Estimates from Key Traffic Count Stations

This opportunity involves selecting key traffic count locations from STARS II together with emission rate estimates to estimate mobile source emissions. As stated, the idea would focus on generating link emissions estimates only. Since traffic counts may not be available for all links, interpolation of volumes, vehicle classifications, speeds, etc. may be needed. The application could use all links within an analysis area or corridor or use a selection of key links to estimate changes or trends occurring in an area. Estimates could be produced automatically on an established frequency or just at times when desired. The outputs could be presented as maps (showing key locations used for the estimates), and/or tables showing *approximate* estimates of emissions for the selected area or corridor. These might resemble the process used to create the annual TxDOT average daily traffic (AADT) maps.

To make this application efficient for repeated use, preparation would need to include the development of the software or human-centric procedures required to perform the computations for sets of links that would be selected by users. The resulting emissions data would not be real-time updates since STARS II database updates are not real time, but could be periodic (i.e., updated in line with temporal updates to the STARS II database). As such they could be used for one-time, periodic, or trends analyses.

4.1.2 Support for Dispersion/Photochemical Modeling

Building upon the preceding idea, this involves estimating on-road emissions for use in estimating dispersion of the pollutants in the atmosphere (i.e., pollutant concentrations). This could be delivered as link sources of pollutants to be interpreted by the end user. An alternative method for general approximations might be to spatially interpolate the traffic data first, calculate emissions based on these interpolations, then use dispersion models to estimate pollutant concentrations. In either case, STARS II offers opportunities (limited to its coverage for each type of data) to provide updated volume, vehicle classification, and speed data for use in on-road emissions estimates as inputs to dispersion/photochemical modeling.

4.1.3 Vehicle Classification

The STARS II database includes a number of stations where network volumes are classified by vehicle-type. Another broad-scale idea is to use these empirical observations to improve, confirm, or validate the VMT by vehicle classifications used for conformity analysis, or other emission estimation studies. Use of STARS II data, depending on the area, could yield more robust samples from which to develop vehicle classification profiles for use in estimating emissions.

4.1.4 Speed Data

Another broad scale idea was to use the speed data from STARS II to corroborate the baseline speed estimates used in conformity or other emission estimation studies. Here the idea centers on the fact that the current method is to estimate link speed using traffic engineering methods (e.g., using volume and highway capacity) rather than using measured speeds. This idea would use empirically measured speeds from STARS II to corroborate these traditional approaches, or to provide an added level of detail into emission estimation. Currently most emission estimations assign a single speed to all vehicles on a road link. This idea would also allow different speeds to be assigned to different vehicle types, providing more refined emission estimates.

4.1.5 Traffic Emissions and Health Impacts – with a Particular Emphasis of Assessing near Road Exposure to Pollutants

This idea modifies some of the previous suggestions by suggesting an application of link-based emission estimates, and dispersion modeling to develop a tool that could be used to study or evaluate the impact of near-road exposure to traffic pollutants on human health. The idea is applicable to both the health of the general public (other road users such as cyclists and pedestrians, and populations that live close to roadways) and populations with high occupational exposure to near road pollutants (e.g., professional drivers, road workers).

4.1.6 Weigh-in-Motion (WIM) Weight Data to Augment MOVES Default Information

The MOVES model currently uses only broadly applicable vehicle weight data to estimate emissions from truck traffic. This idea involves using STARS II WIM data to develop more applicable emission rate estimates for heavy duty vehicles. The central tenet of the idea is that the gross weight of trucks has an important impact on emission rates. As such, empirical real-world data on the weight distribution of trucks using specific segments, or all or portions of the Texas road network could be used to help to refine emission estimates. The idea could apply to emission estimation for regulatory purposes or for specific case studies.

4.1.7 Benchmarking Traffic Counts and Speeds with Other Passive Data

This idea addresses the potential of using STARS II data synergistically with a novel data source, namely passive global positioning system (GPS) and cellular data. Passive data include locational information derived from mobile devices, navigational apps, in-vehicle navigation systems from passenger vehicles and commercial fleets. Passive data are becoming increasingly important for transportation research. They are able to provide traffic data similar to metrics in STARS II such as count estimates by time of day and vehicle speed. However, they are also capable of providing data on the origin and destination of traffic along a corridor or at the count station location. Although the potential of passive data is well realized, there are a number of challenges that must be overcome for it to be used for traffic activity modeling and air quality science. The central tenet of this idea is that passive data provide traffic activity data that are complementary to the data in STARS II (e.g., fine scale temporal and spatial changes in traffic volume), but may require traditional traffic data (e.g., absolute and comprehensive counts by vehicle type) to maximize its full potential. STARS II data could be used in one of a number of related ways:

a. Serve as (and is needed as) a source to expand cell and GPS-based passive data purchased by TxDOT for statewide modeling and/or corridor studies.
b. Corroborate the accuracy and utility of passive data, and vice versa.
c. Help develop algorithms and analyses necessary to realize the full potential of passive data.
d. Assist TxDOT or other stakeholders on the cost-benefit of acquiring passive data at a regional or statewide level.
e. To develop more comprehensive databases or models of vehicle activities based on algorithms that link both datasets (i.e., better spatial and temporal resolution; including additional activity measures; or improving accuracy of existing data).

4.1.8 Emission Estimation for Public Outreach/Educational Tool - Public Information

Some of the previous suggestions could be adapted to provide public outreach or educational materials about air quality conditions. The idea would involve the development of interactive tools (e.g., website) using best estimates of traffic activities, converted to emissions estimates and/or pollutant levels as estimated using dispersion models. These could use estimates of recent emissions or of (periodic) trends. The tools could be delivered to stakeholders to help them understand the relationships between traffic activities, vehicle types, emission estimates, and air borne pollutants. One example could be to highlight areas that most frequently have traffic-induced health alert concerns. Another might be to illustrate how emissions change throughout a year.

4.1.9 On-road Heavy-duty Measurement System (OHMS) - Strategic Positioning of Emission Test Devices

OHMS is a semi-portable, high throughput vehicle emission test device that can be used to rapidly identify high emitting heavy duty diesel engines (trucks, buses, etc.). Installed at a location, it can be used to screen a random sample of the passing truck traffic and notify truck owners of potential maintenance issues. OHMS can be used by MPOs to help identify and eliminate a small number of trucks that disproportionately contribute to local emissions. This idea suggests that STARS II data might be useful for selecting OHMS locations that are optimum for passing truck traffic and that are difficult to be bypassed by truck drivers.

4.1.10 Speed Estimation Based on Statistical Analysis of Other Stations

Speed information is available for only a limited number of locations within STARS II. Furthermore, it is not disaggregated into different vehicle types. Instead STARS II provides a speed distribution of each passing vehicle within a measured time period. However, vehicle speed is an important input into emission models. This idea suggests that it may be possible to disaggregate speed data at certain stations using statistical models that correlate speed distributions with vehicle class distributions from neighboring stations. A similar idea would investigate interpolating speed data across a road network using statistical models that relate locations with speed data to neighboring stations without speed data. In either case, STARS II data could be augmented with other data such as passive GIS.

4.1.11 Houston Galveston Area Council (HGAC) Council - Truck Replacement Study

TTI is currently collaborating with HGAC on a program that provides incentives to replace older trucks operating around the Port of Houston with newer, cleaner models (North Central Council of Governments also has a similar active truck replacement program). A significant output of the HGAC program is GIS data from devices installed on trucks replaced by the program. This high quality data set could be used in conjunction with STARS II to either:

a. Corroborate data in STARS II.
b. Confirm other assumptions used in travel forecast modeling or emissions estimation.
c. Understand the drive cycles of trucks around the Port of Houston (hence emissions) in relation to traffic activity in the area (e.g., diversions, idling in traffic jams).

4.1.12 Speed Limit Enforcement

Some areas of the state have speed limits set to help to reduce mobile source emissions. Vehicles traveling at high speed tend to generate higher emissions per mile than slower moving vehicles. STARS II speed data could be used to identify areas where traffic on high speed roads usually significantly exceeds the posted speed limits. This information could be passed on to enforcement officials for use in prioritizing enforcement. This would have potential benefit for both safety and health purposes.

4.1.13 Confirmation of Emission Estimates for All 254 Texas Counties

This idea suggests that the STARS II traffic data could be used to support current statewide emission inventory estimates. The database could be expanded where needed to provide better coverage or fill in time periods for which data are not current or broad enough.

4.1.14 Corridor Studies - Project Level, Data Extraction Tool/Instructions

TxDOT and other agencies such as the Texas Commission on Environmental Quality (TCEQ) or the Border Environment Cooperation Commission often perform a number of project-level analyses that explore specific air quality issues on corridors or local networks. Here it is often important to obtain rapid, detailed traffic report data that can be used in problem formulation and quantitative analyses. In many cases, STARS II contains high quality data that has potential to be used directly in these studies. STARS II has the potential to provide data with high temporal resolution (hourly and daily counts) or that is classified into specific vehicle types. A significant barrier to using these data is the time and knowledge required to extract these data, and the application of accepted modeling techniques to interpolate or aggregate the data to provide the quantities most useful for a particular case study. This idea suggests the development of a project-level data extraction and modeling tool that can be used and trusted by air quality specialists to improve case study research techniques. At its simplest, this may be a simple instruction manual and or methodology paper that can be followed and referenced in reports. It could also involve the development of bespoke data query tools that could be incorporated into the current STARS II reporting software.

4.1.15 Spatial Interpolation of Traffic Counts to Complement HPMS

This idea expands on number 14 by developing a standard methodology to interpolate traffic information collected at discrete locations. Traffic activity data are an essential precursor to on-road emissions estimation. A significant limitation of the STARS II data to air quality arises because traffic counts are taken at discrete points along the network. In contrast, air quality studies often require emissions generated over a two-dimensional road network (or in the case of dispersion modeling, a three-dimensional volume). The complexity of road networks and traffic presents challenges for interpolating traffic count data from limited data collected at discrete locations. Although HPMS provides network interpolated data for total volumes and percent trucks, data are often required at a higher resolution than the HPMS outputs (e.g., for specific vehicle types, specific time periods). The idea could be implemented through software or an instruction manual demonstrating best practices. In either case, it would produce useful data for air quality impact analysis.

4.1.16 Correlating Abnormal Traffic Activity with Special Events

In many air quality studies, the traffic activity estimates that are used to estimate emissions are based on typical network volumes. However, there is an increasing need for more refined estimates of traffic activity, and their impacts on air quality. The resolution of data in STARS II presents an opportunity to analyze the impact of special events on traffic activity and on air quality. Special events might include football games (e.g., Dallas, Houston, and Austin) and major concerts (e.g., SxSW Austin). Traffic activity data can often be explained using a number of consistent factors including, long-term growth trends, seasonal variation, day of the week variation (weekdays versus weekends, and holiday variation [e.g., Christmas, Thanksgiving). However, traffic data also include a stochastic (unpredictable) component in addition to these reliable trends. This idea would explore whether, once these trends have been incorporated into models of traffic at specific locations, special events could be detected in the traffic patterns, and whether the STARS II data could be used to understand the impacts of these special events. This would work best where sufficient count locations exist on key approach, circulation, and departure routes and be most applicable in areas with air quality problems where single-day or multiple-day events might create air quality concerns.

5. Case Study - An Analysis of Long-Haul Truck Emissions on the I-35 Corridor between Laredo and San Antonio, TX

Using the opportunities identified in the qualitative exercise outlined above, researchers picked a case study designed to illustrate the potential of STARS II to support an analysis of the air quality impacts of long-haul trucks traveling between Laredo and San Antonio, TX. This case study builds upon a contemporary air quality issue that was originally presented to TTI by the TCEQ. Specifically, it involves assessing the air quality impact of a recent policy change (circa 2015) that now allows Mexican long-haul trucks to operate across the border into the interior of the United States.

Previous to the new rule, Mexican trucks were only allowed to operate within a limited region adjacent to the U.S. border called the commercial zone. This meant that freight is currently transported across the border using a drayage system. In the drayage system, freight from the Mexican interior is transported to warehouses at the Mexican border by Mexican domiciled trucks. Then, drayage trucks are used to transport the freight to locations within the commercial zone on the U.S. side of the border. Finally, U.S. domiciled trucks are used to transport the freight to their final destination within the U.S. interior.

In contrast, the new long-haul cross border trucking rules, which were driven by the North American Free Trade Association, allow Mexican domiciled trucks to transport freight from a destination in Mexico, directly across the border to a final destination in the U.S. interior. The new federal law has led to concerns that a difference in emission standards between U.S. and Mexican trucks will impact air quality along the corridor.

This case study was chosen because it represents a real world and contemporary air quality impact assessment that requires a detailed understanding of the current traffic activity along a specific corridor. It requires estimates of long-haul truck activity, rather than the activity of all trucks along the corridor (e.g., single unit trucks). The problem also involves a long-term forecast of truck volume along the corridor. This forecast is necessary because the Mexican government has outlined plans to introduce new truck emission standards by 2018. When implemented, this will result in similar emission rates for U.S. and Mexican trucks (so no emissions impact), but until then Mexican trucks operating on the corridor are expected to have a higher emissions rate than U.S. trucks.

5.1 Case Study Formulation

For the purposes of this document, the case study focuses on the methods used to obtain traffic activity data from STARS II, and on developing heavy duty diesel truck volume models for an emissions impact assessment. The goal is to extract data from the STARS II database, and develop models of long-haul truck activity that could be used in conjunction with MOVES to estimate the long-term impact of the new border crossing rules on emissions along the Laredo-

San Antonio corridor. To fully illustrate the potential of data within STARS II, the database was used to develop a model of long-haul truck activity along the corridor that incorporates the following components:

1. Activity of long-haul trucks (combination trucks rather than all trucks).
2. A forecast of annual growth in truck volume along the corridor.
3. A seasonal component that describes the change in truck activity by month.
4. A daily component that describes the change in traffic activity driven by different days of the week.
5. A stochastic component that describes the unpredictability of long-haul truck activity along the corridor.

5.2 Methods

The STARS II database was queried by location for permanent traffic count stations along the corridor of interest. An additional search term was used to find locations that provided traffic volumes by vehicle classification. The results yielded three stations of primary interest: S219, located on I-35 on the outskirts of Laredo; S210, located on the outskirts of San Antonio; and W531, situated approximately at the midpoint between Laredo and San Antonio. Each station provided data for north- and southbound traffic volumes, by vehicle type, by day between the years 2013 and May 31, 2015. Although each station is classified as a permanent count location, there are significant temporal gaps in the data. Table 3 illustrates the number of days for which classification counts could be obtained. Because the I-35 corridor between Laredo and San Antonio is serviced by a number of arterial roads, an annual summary (for the year 2014) of classified volumes were also collected for key locations along these arterials. Figure 2 illustrates a spatial overview of combination truck volumes on and around the I-35 corridor. This spatial overview of combination truck activities was used to validate a key assumption that most of the long-haul trucks on the corridor between station S219 and S210 must pass through the critical station W531 situated at the midpoint of the corridor.

Table 3 – Temporal coverage of the data at each principal traffic count locations.

Location	Number of Potential Data Points 2013 – 2015 (days)	Number of Data Points Available 2013-2015 (days)	Data Availability (%)
S219 (I-35 near Laredo, TX)	859	605	70.4
W531 (I-35 between Laredo and San Antonio, TX)	859	217	25.3
S210 (I-35 nr San Antonio, TX)	859	651	75.8

Truck crossing data were also obtained for the border port of entry at Laredo, TX. These volumes were collected as an additional qualitative check that the truck activity along the I-35 corridor is similar to the volume of trucks crossing the border. Here, the great majority of trucks crossing the border are engaged in drayage. As such, a 1:1 correspondence is highly unlikely between trucks crossing the border and combination trucks on I-35 because drayage freight may be consolidated in commercial zone warehouses prior to long-haul shipment along I-35. Additionally, some freight may be destined for alternative routes such as US59 to Houston and US83 to Brownsville.

Figure 2 provides a simple, conceptual view of the corridor of interest. It shows that combination truck volumes along the corridor and at the border are broadly similar, and have a similar long term trend. However, as might be expected from Figure 3, station S219 (Laredo) and especially station S210 (San Antonio) have additional combination truck activity resulting from their proximity to major urban areas. In contrast, it is reasonable to assume that combination truck activity at station W531 is most likely the result of trucks that travel the majority of the length of the I-35 corridor.

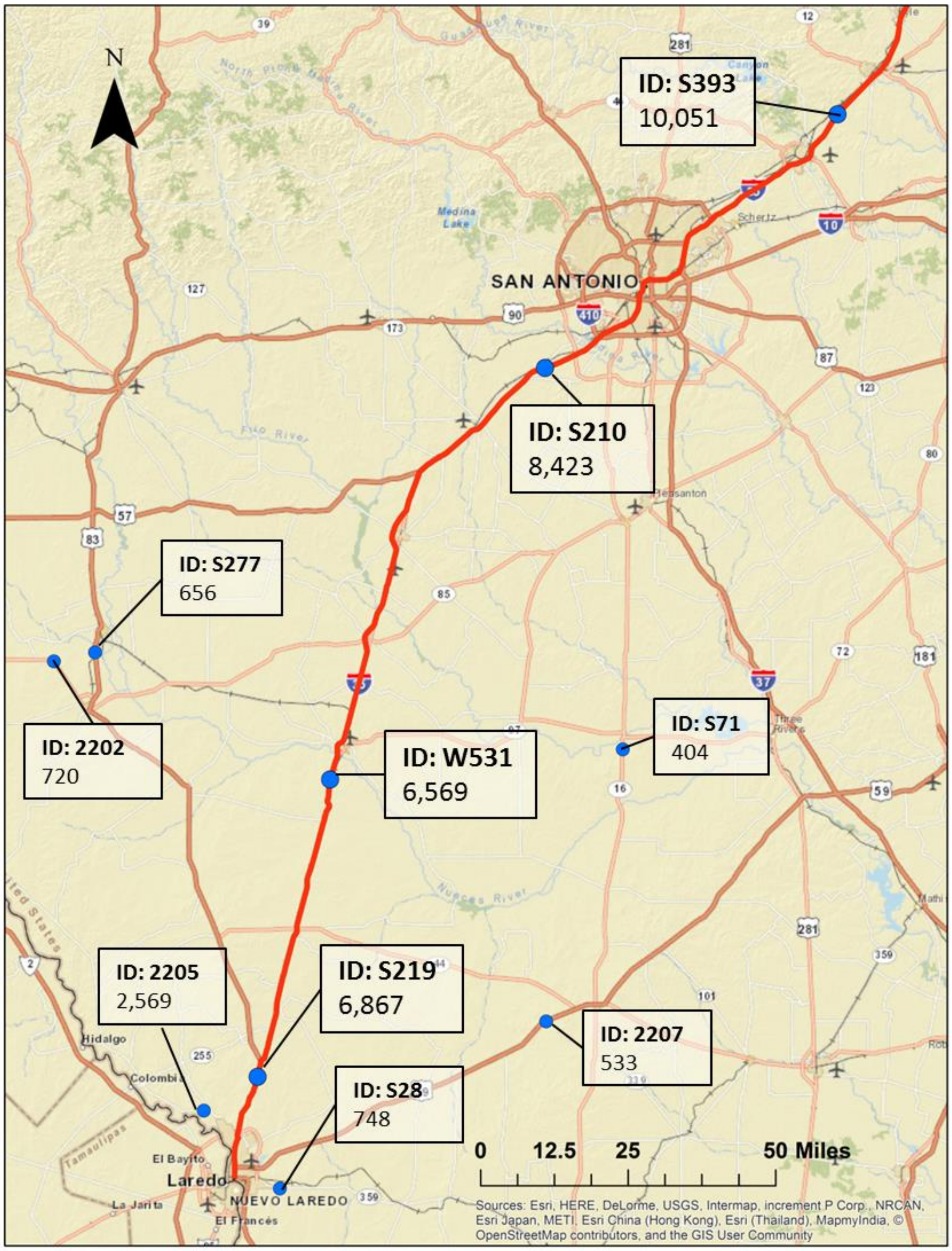

Figure 2 – Overview of study corridor using GIS data on AAD combination truck volumes downloaded from STARS II (data are annual summaries for 2014 directly provided as a result of STARS II analysis rather than researchers' own modeled estimates).

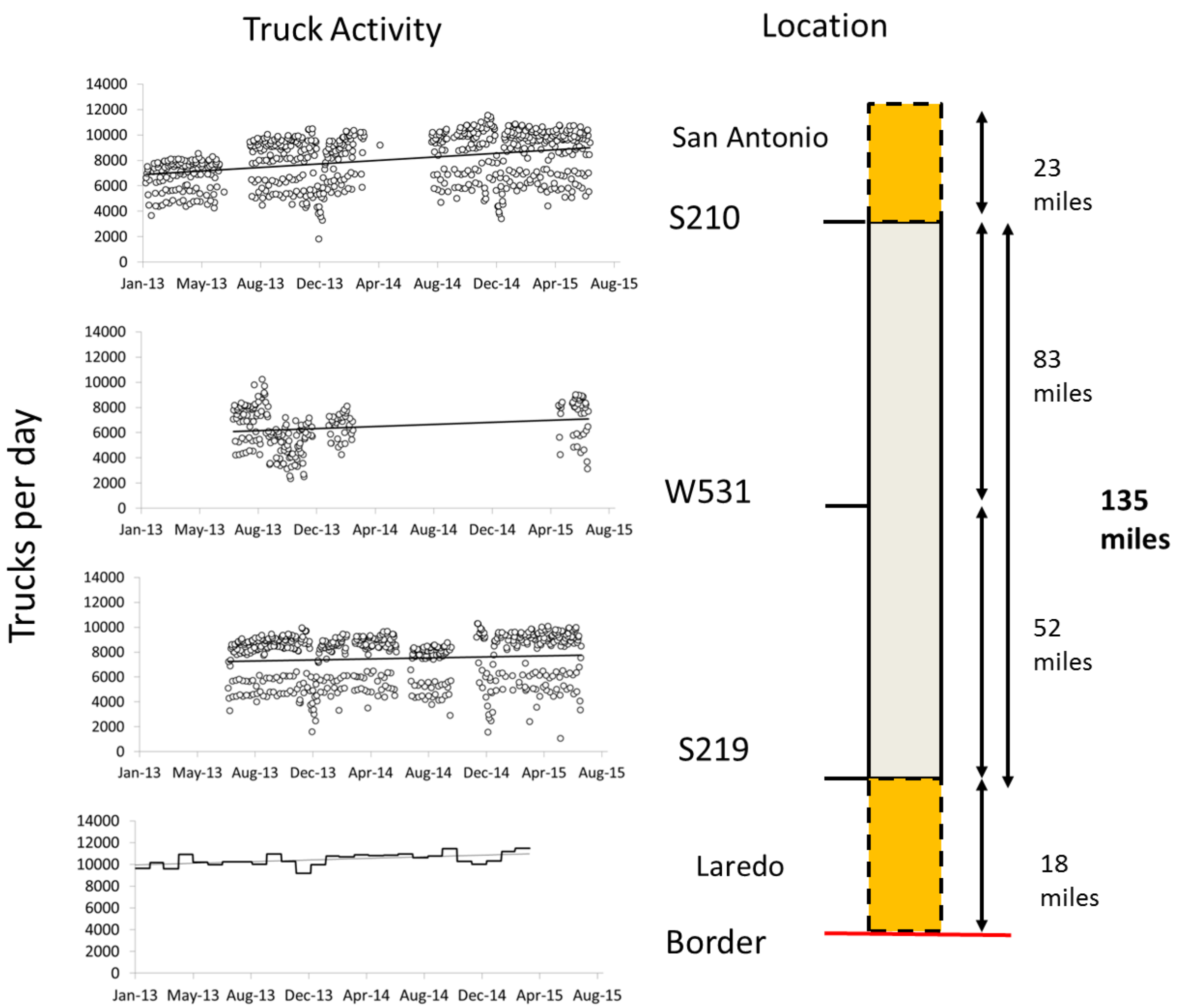

Figure 3 – Simplified representation of truck activity along the I-35 corridor. The graphs on the left hand side of the diagram show detailed temporal traffic counts at the three TxDOT monitoring stations (S210, W531, S219) and truck crossings at the Laredo border. South to north border crossing data were obtained from the US DOT Bureau of Transportation statistics website. Crossings for both directions were estimated as double the south-north estimate based on more limited information on north to south crossings from Texas Center for Border Economic and Enterprise Development. The original border crossing data are provided as monthly crossings. At I-35 stations, truck counts are daily, illustrating a strong daily (weekday) component to truck volumes. The trend lines fitted to each graph illustrate long-term trends in traffic volume.

5.3 Modeling Long-Haul Truck Volumes

To improve understanding of the truck activity at each station, researchers developed a model to interpolate and forecast the observed data:

$$Trucks_{day} = V_{station} + day*(G * M * D * H) + \varepsilon \qquad \text{Equation 1}$$

Where:

$Trucks_{day}$ = Volume of trucks predicted on a given model day.

$V_{station}$ = Baseline volume of trucks on January 1, 2013 (baseline date), at each station.

G = Growth rate in truck volume (Trucks per day).

M = A month specific adjustment factor describing the seasonal trend in truck activity.

D = A day of week adjustment factor describing the day of the week trend in truck activity.

H = A holiday adjustment factor (Thanksgiving and Christmas period Dec. 24 to January 1).

ε = A normally distributed error.

The model in equation was fit to the data obtained from the three stations on I-35 (S210, S219 and W531). Because the traffic at each location is connected, the parameters representing volume growth (G); and monthly (M), day of week (D) and holiday factors (H) were assumed to be the same for each station. In contrast, a different baseline volume ($V_{station}$) was assumed for each station. The models were fitted to the data using a sum of squares regression routine developed using Microsoft Excel. Finally, researchers used the residuals of the model fit to calculate a normally distributed error term. Table 4 shows the fitted model parameters.

5.4 Interpolating and Forecasting Stochastic Daily Truck Activity

The model outlined in Equation 1 and parameters in Table 4 can be used to estimate combination truck volumes to interpolate data missing data for the period 2013–2015 (i.e., the period for which some data exists). It can also be used to forecast future truck volumes (i.e., beyond May 2015). These interpolations or forecasts can be generated deterministically (with no error term), or with an error component that represents the variation in daily combination truck estimates not captured by the model. Figure 4 show deterministic and stochastic forecasts for station W351 for a 2-year time-period and a 15-year time period, respectively. Figure 5 illustrate truck forecasts for station S210.

Table 4 - Parameter values estimated at each traffic count location. The parameter V and the error term are specific to each station. All other parameters were considered to be the same for each station.

Parameter	Station		
	S219	**S210**	**W531**
V	7213.6	6219.2	7691.3
ε (% CV)	N(0.10, 0.17)	N(0.04, 0.16)	N(0.05, 0.18)
H	0.48		
G	1.24		
Month			
January	0.96		
February	1.00		
March	1.00		
April	0.98		
May	0.98		
June	1.00		
July	0.98		
August	1.01		
September	1.00		
October	1.01		
November	1.03		
December	1.05		
Day of Week			
Monday	1.02		
Tuesday	1.09		
Wednesday	1.16		
Thursday	1.18		
Friday	1.14		
Saturday	0.78		
Sunday	0.63		

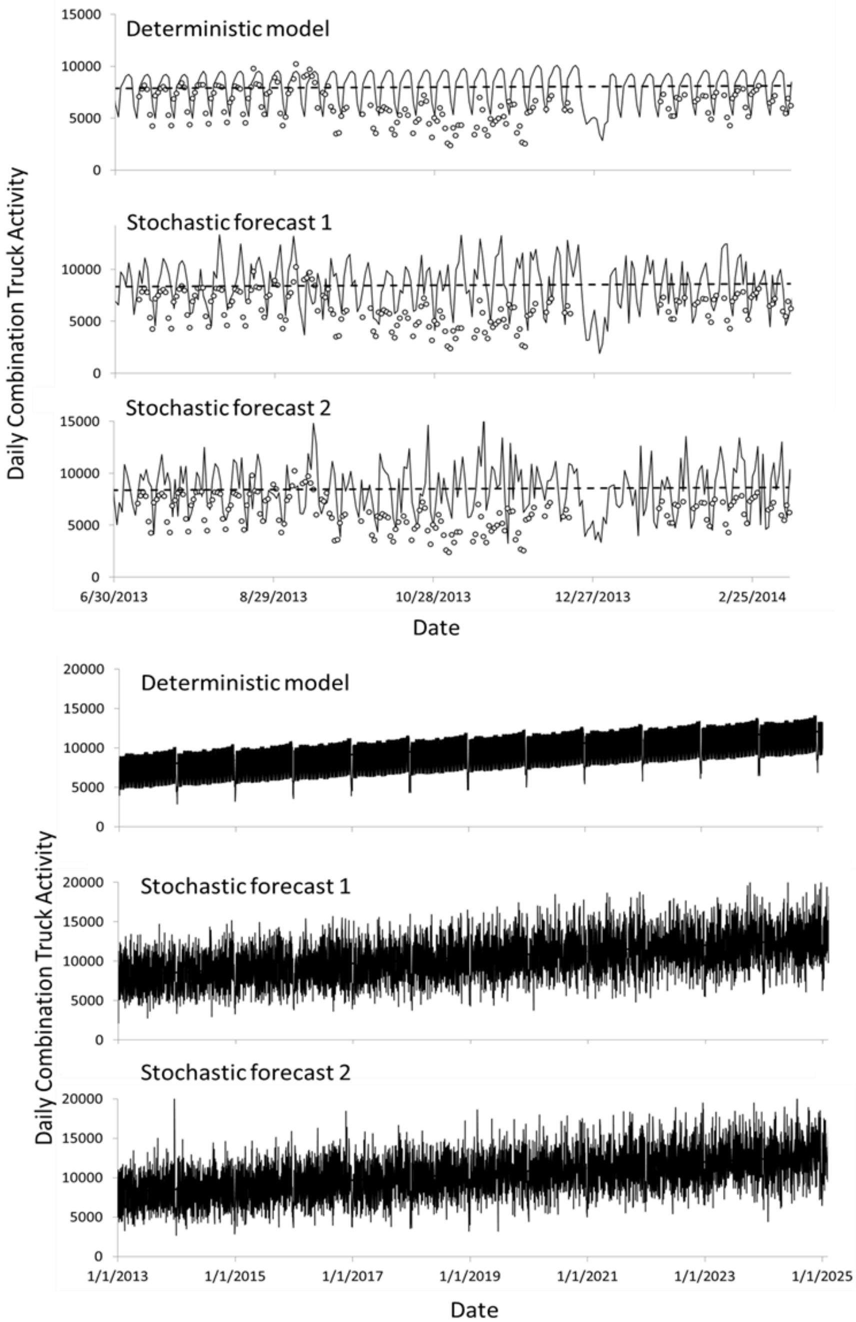

Figure 4 – Forecasts and interpolation of truck volume over two contrasting time periods for station W531). The upper graphs show deterministic and stochastic forecasts for the period January 1, 2013, to Dec. 31, 2014. The forecasts are shown with solid lines. The raw data used to fit the model are shown as open circles (shown on days when data are available). Note the strong daily trends in both the raw data and the models. The lower set of graphs show the same model projected over 2013 and 2025, illustrating the long-term volume growth.

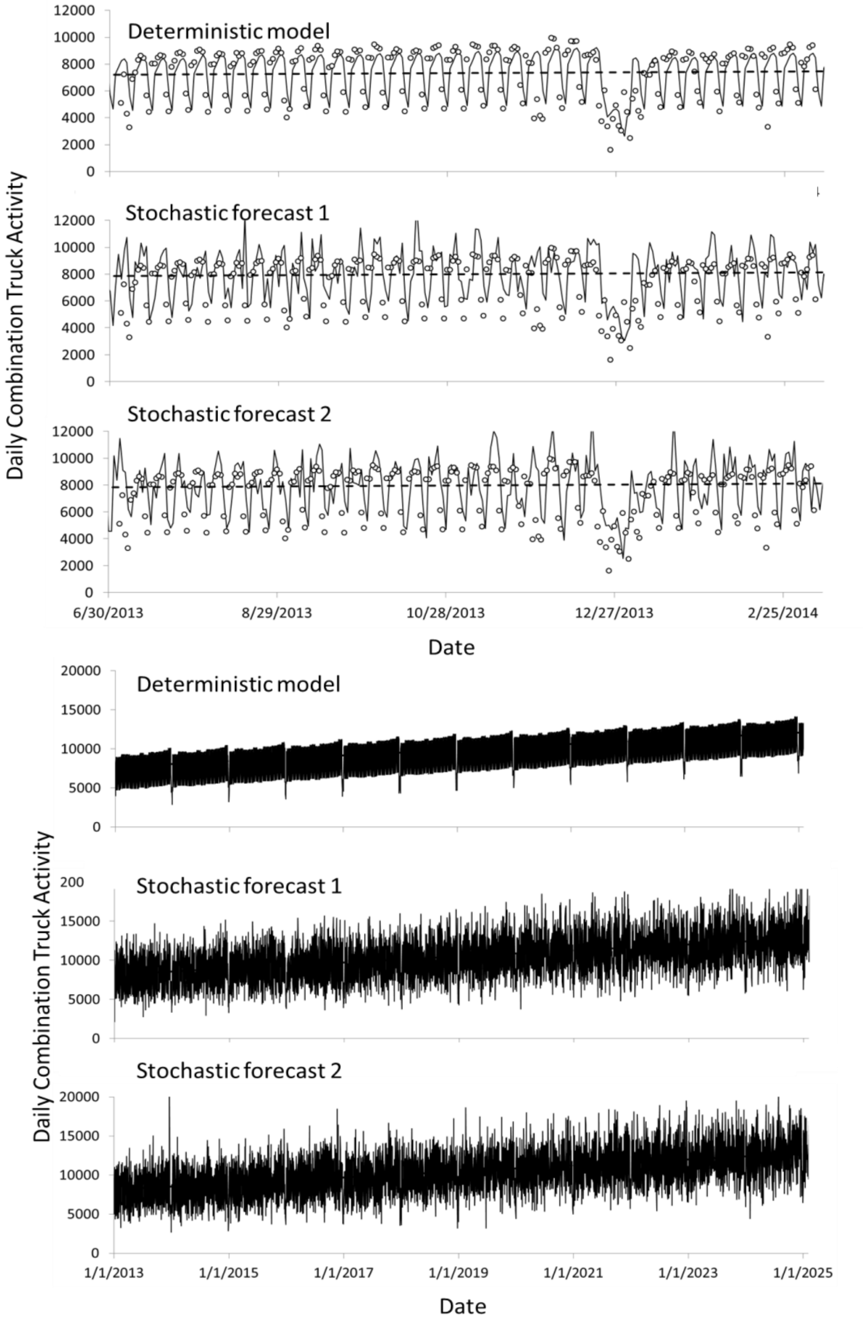

Figure 5 – Forecasts and interpolation of truck volume over two contrasting time periods for station S210. The upper graphs show deterministic and stochastic forecasts for the period January 1, 2013, to Dec. 31, 2014. The forecasts are shown with solid lines. The raw data used to fit the model are shown as open circles (shown on days when data are available). Note the strong daily trends in both the raw data and the models. The lower set of graphs show the same model projected over 2013 and 2025, illustrating the long-term volume growth.

5.5 Application to Emissions Estimation and Air Quality Science

The models outlined here can be used to smooth the raw data, understand its trends and errors, and generally provide an interpreted view of the traffic at a location. In this case study, STARS II most significantly contributes to the case study because its vehicle classification counts can be used to understand the activity of a specific vehicle type (combination long-haul trucks) along a specific corridor. It also provides daily data that are important for correctly interpreting the long-term trend in truck activity along the corridor because truck volumes tend to follow weekday and seasonal trends. If these trends are not accounted for in analysis of the raw data, misinterpretation is possible. Daily data (i.e., replicate counts) are also useful for understanding the unpredictability of volumes along this route (the error term in the model).

Traffic activity data underpins emission estimation. Conceptually, total emissions ($Emissions_{trucks}$) along this corridor can be calculated as:

$$Emissions_{trucks} = Trucks_{day} \times ER_{day} \times CL \qquad \text{Equation 2}$$

Where:

$Trucks_{day}$ = the number of combination trucks on the roadway on a particular day.

ER_{day} = the average emission rates of combination trucks using the roadway (in miles per truck).

CL = the length of the corridor in miles.

In turn, ER_{day} is usually provided by MOVES and requires inputs in the form of the age distribution and emission standards of the fleet, the average speed of the vehicles, and ambient temperature and humidity.

Overall, this case study illustrates that STARS II has considerable potential for providing data for air quality analyses. For this specific air quality problem, researchers were interested in long-haul trucks traveling along a specific corridor. STARS II was used to provide an overview of combination trucks in the region (Figure 2), interpret the long-haul truck activities along the specific corridor (Figure 3), and develop a quantitative model of combination truck volumes along I-35.

The case study as described only requires daily data, but STARS II also has the potential to provide hourly data. Researchers also limited the analysis to modeling combination truck traffic combined for the north- and southbound directions. However, for other applications STARS II can provide data by direction. Finally, this case study required an understanding of the data along only a specific route that researchers assumed carries virtually all of the long-haul activity in the region. Other analyses could involve modeling traffic across a more complex network.

Applied to specific case studies, the models can be used to create deterministic or stochastic forecasts or interpolations of traffic data. The stochastic representations in particular could be used to develop more realistic worst-case scenarios for emission studies. These stochastic models explicitly incorporate all of the uncertainties of traffic at a location. In turn this is useful for understanding how the stochastic variation in traffic activity could affect emissions estimates. This stochastic modeling approach provides a pragmatic approach to modeling traffic data and emissions; in reality, traffic data are provided as limited samples from the real world, and as such do not perfectly represent all traffic activity. The stochastic approach allows researchers to understand how much uncertainty is inherent in the data, and how much this uncertainty is likely to impact the results of an analysis. As such, it provides a way of using the data, even when there may be considerable data gaps or uncertainties in traffic data.

6. CONCLUSIONS

The goal of this study was to investigate potential applications of STARS II database for air quality analysis and to conduct a case study to demonstrate an example for one of those potential applications. Conclusions drawn as a result of work performed in this task are:

- The accuracy of the estimates on which state implementation plans, conformity analyses, and other specialized air quality analyses are based and are directly affected by the availability, amount, and accuracy of the input data—especially traffic data.
- STARS II is a potentially powerful database with applications in many of TxDOT's functions, including air quality analysis. STARS II offers consistency through statewide coverage providing:
  - Traffic volumes.
  - Traffic volumes classified by vehicle type.
  - Speed.
  - Weight.
- The importance of regulatory analyses has driven the development and widespread use of emission rate models and methods that are capable of using detailed traffic count estimates. Traffic activity data (at the most appropriate resolutions) are often the limiting factor in fully using these methods.
- Through a network of permanent count stations on state highways, STARS II has the potential power to provide counts or estimates of traffic volumes (as demonstrated in case study) for any time of season, month, week, day, or hour. However, as described in the case study, time-series models are required to account for temporal and spatial gaps in data.
- Vehicle classification, speed, and WIM data are less widely available in STARS II, but are still available at enough sites to provide samples that may be applicable at the regional analysis levels and some corridors.
- Where sufficiently available, these data could offer opportunities for more detailed fleet mix characteristics for use in emissions modeling than is currently used in most regions. Vehicle classifications are currently applied region-wide, but with enough classification data could be used to understand patterns by region, corridor, or other subarea.
- Several suggestions are offered for using STARS II data for air quality analyses:
  - Periodic, automatic emissions estimates from key traffic count stations.
  - Support for dispersion/photochemical modeling.
  - Vehicle classification.
  - Speed data.
  - Traffic emissions and health impacts, with a particular emphasis of assessing near road exposure to pollutants.

   - WIM weight data to augment MOVES default information.
   - Benchmarking traffic counts and speeds with other passive data.
   - Emission estimation for public outreach/educational tool – public information.
   - SHED – Strategic positioning of emission test devices.
   - Speed estimation based on statistical analysis of other stations.
   - HGAC council – truck replacement study.
   - Speed limit enforcement.
   - Confirmation of emission estimates for all 254 Texas counties.
   - Corridor studies – project level, data extraction tool/instructions.
   - Spatial interpolation of traffic counts to complement HPMS.
   - Correlating abnormal traffic activity with special events
- The extent of use of the STARS II database will depend upon the ease of data extraction and use. The automation of specific data extraction routines for certain air quality analysis types would greatly improve the efficiency of using STARS II data.
- It may be appropriate to selectively expand the STARS II database in areas where air quality analyses can benefit. The additional data would be helpful to produce:
   - More coverage (count stations) in existing count areas and/or in additional areas and/or for more functional classifications.
   - More area coverage, such as in rural areas of urban or partially urban counties.
   - More frequency or duration of counts at existing stations.