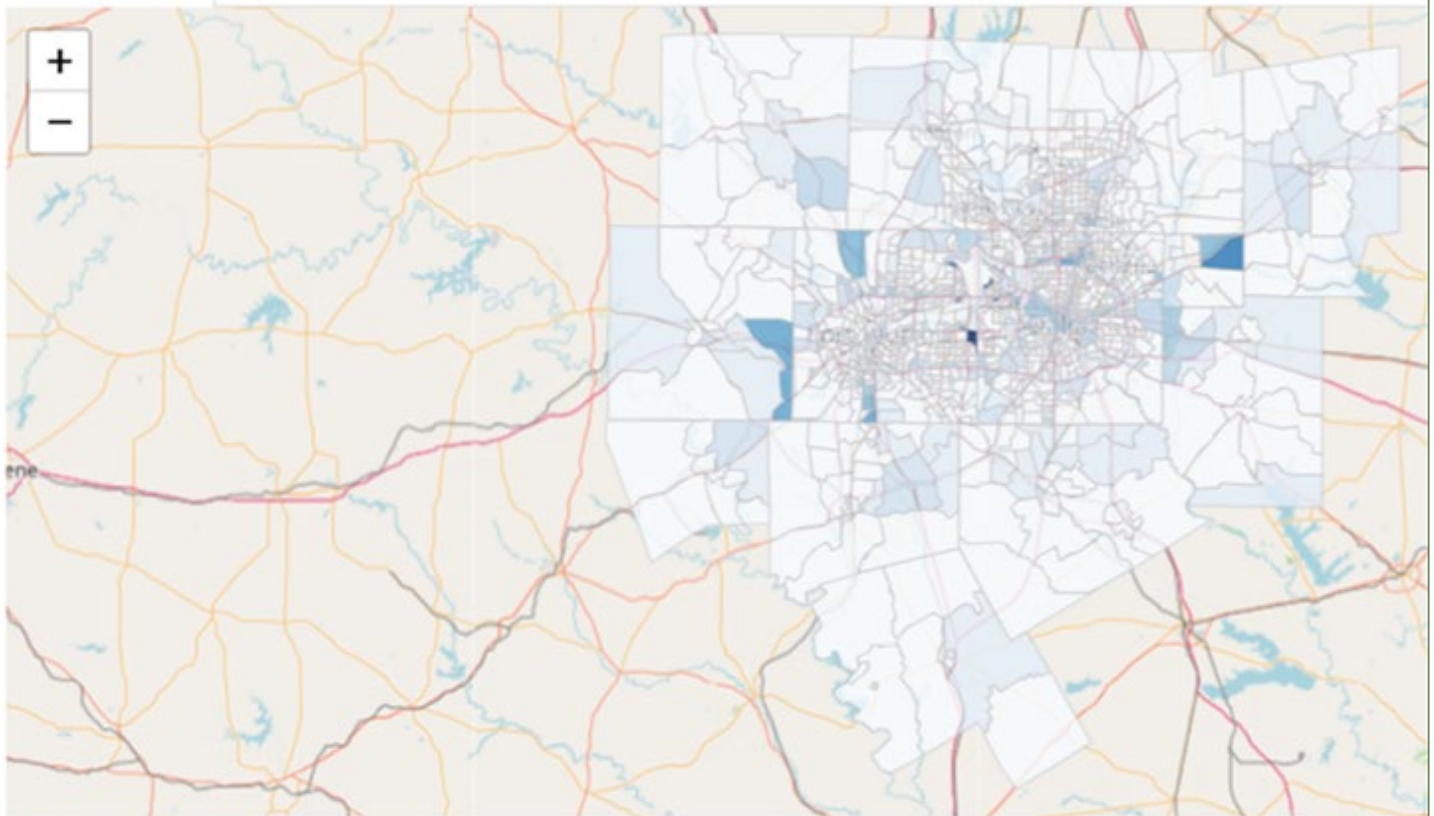# DEVELOPMENT OF EMISSIONS ESTIMATION AND DISPERSION MODELING FRAMEWORK FOR THE HEALTH IMPACT ASSESSMENT

Total Premature Deaths avoided: 52

June 13, 2024

Disclaimer

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Development of Emissions Estimation and Dispersion Modeling Framework for the Health Impact Assessment		5. Report Date June 13, 2024	
		6. Performing Organization Code	
7. Author(s) Rohit Jaikumar, Ph.D. Tao Li, Ph.D. Guo Quan Lim, Ph.D. Tara Ramani, P.E., Ph.D. Madhusudhan Venugopal, P.E.		8. Performing Organization Report No. 06-60-TTI	
9. Performing Organization Name and Address: Texas A&M Transportation Institute 3135 TAMU College Station, Texas 77843		10. Work Unit No.	
		11. Contract or Grant No. 69A3551747128	
12. Sponsoring Agency Name and Address Office of the Secretary of Transportation (OST) U.S. Department of Transportation (USDOT)		13. Type of Report and Period Final November 2022 to June 2023	
		14. Sponsoring Agency Code	
15. Supplementary Notes This project was funded by the Center for Advancing Research in Transportation Emissions, Energy, and Health University Transportation Center, a grant from the U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program.			

16. Abstract

A health impact assessment (HIA) is an essential tool for evaluating the potential health effects of environmental changes on populations. Recognizing the significance of an HIA in the context of air pollution, this project introduced a comprehensive framework for assessing the health impacts of air quality variations. Implemented as an R Shiny application, this framework streamlines the HIA process, making it user-friendly and adaptable to diverse data sources. The application encompasses the entire HIA process, from data preprocessing to results visualization. A case study using the North Central Texas Council of Governments travel demand model network demonstrated the framework's efficacy, revealing a notable reduction in the concentrations of particulate matter that is 2.5 micrometers or smaller and associated health risks in 2026 compared to 2019. This innovative tool offers a transparent and accessible approach for researchers, policymakers, and the public to assess and address air pollution-related health impacts.

17. Key Words Integrated modeling, health impact assessment, AERMOD, dispersion modeling, emissions, traffic-related air pollution		18. Distribution Statement No restrictions. This document is available to the public through the CARTEEH UTC website. http://carteeh.org	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 24	22. Price $0.00

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

Executive Summary

Background and introduction: A health impact assessment (HIA) is a systematic process used to evaluate the potential health effects of a policy, program, or project on a population. It provides evidence-based recommendations to decision-makers to protect and promote public health. Given the increasing concerns about environmental health, especially related to air pollution, the HIA has become a crucial tool for policymakers. This project built upon previous work by Texas A&M Transportation Institute researchers that assessed the health impacts of automated vehicles and the use of high spatial resolution data for health impact assessment of traffic-related air pollution.

Objective: The primary goal was to create a comprehensive framework, implemented as an R Shiny application, to assess the health impacts of air quality changes. This framework aims to simplify the complex processes involved in HIAs, making them accessible to a broader audience, including researchers, policymakers, and the public.

Framework features: Notable features of the framework include the following:

- **Holistic integration:** Covers the entire HIA process.
- **User-centric design:** Provides an intuitive interface for ease of use.
- **Flexibility:** Handles diverse data sources and formats.
- **Transparency:** Provides a clear understanding of underlying assumptions and calculations.

Implementation: The framework was developed using the R programming language as an R Shiny application. It offers a step-by-step process, from data input to results visualization, allowing users to conduct a full-chain health impact assessment seamlessly.

Application structure: The application is divided into five main tabs:

1. Preprocessing.
2. Input.
3. Run AERMOD.
4. Visualize results.
5. Health impact analysis.

Each tab serves a specific purpose, guiding users through the modeling chain, from preprocessing emissions and traffic data to visualizing results and comparing different scenarios.

Results—Case study: The North Central Texas Council of Governments travel demand model network was used as a case study to test the R Shiny application. The study involved various stages, from preprocessing traffic data to modeling emissions and dispersions to assessing health impacts. The results showed a significant reduction in the concentrations of particulate matter that is 2.5 microns or smaller in 2026 compared to 2019, leading to a decrease in premature deaths.

Table of Contents

List of Figures

1   Background and Introduction

This project leveraged previous work conducted by Texas A&M Transportation Institute researchers related to a full-chain health impact assessment of automated vehicles (AV) through changes in air quality and the use of high spatial resolution data for a health impact assessment of traffic-related air pollution in a large area. The first study focused on evaluating AV through a health lens while the second study further contributed to the traffic-related air pollution HIA exercise by focusing on methodological issues, specifically studying the potential impacts and applications of using datasets of higher spatial resolution as part of the process.

The activity used in the study was based on the old Dallas-Fort Worth travel demand model (TDM) for the year 2018 received from the North Central Texas Council of Governments (NCTCOG) through an email correspondence, and emission rates were based on the older Motor Vehicle Emission Simulator (MOVES) model (USEPA, 2022). MOVES3[1] is the Environmental Protection Agency's (EPA's) latest mobile source emission model, which estimates less exhaust particulate matter (PM) with a diameter smaller than 2.5 microns ($PM_{2.5}$) compared to previous versions of the model due to decreased extended idling activity and lower heavy-duty emission rates (USEPA, 2023). Also, MOVES3 has a better representation of PM emissions resulting from brake and tire wear. Moreover, the previous models did not consider resuspended dust, which is a significant source of $PM_{2.5}$ emissions.

A complex modeling chain, involving the following components, is required to simulate the dispersion of transportation-related emissions from the tailpipe of the car to the lungs of an individual:

- **Traffic modeling**: Estimates the traffic volumes, speeds, and driving patterns.
- **Emission modeling**: Estimates the emissions associated with the traffic patterns.
- **Dispersion modeling**: Estimates the transport of the emissions into the atmosphere based on the physical forces and chemical reactions.
- **Exposure modeling**: Estimates the amount of pollution an individual is exposed to based on the strength and duration of exposure.
- **HIA**: Estimates the mortality burden or risk of disease based on exposure to a particular pollutant.

Estimating accurate transportation emissions is a vital task in an HIA. The EPA recommends using local activity and emission rates for developing transportation emissions. The emissions generated at the link level were used as the input for dispersion modeling to estimate the concentration at finer scales, such as a traffic analysis zone, neighborhood zip code, or hot spot like an intersection. The refined concentration and exposure findings can then be used to estimate the health risk by employing the Centers for Disease Control and Prevention (CDC) and EPA datasets.

The error or inaccuracy in the HIA modeling chain can arise from any of the steps identified previously and may result in an underestimation or overestimation of the local risk. The proposed study focused on developing a framework to improve the emissions and dispersion modeling estimates by using the latest available network and local data to develop the activity and emission rates that are needed to predict health risks.

2   Problem Statement

HIA modeling is a systematic process that evaluates the potential health effects of a policy, project, or program. At its core, the HIA modeling chain begins with the collection of data on the sources and amounts of pollutants being

---

[1] MOVES3 was the latest emission model available when the project started in the late 2022 calendar year.

released into the environment, known as the emissions inventory. This foundational step is crucial as inaccuracies or omissions at this stage can lead to critical errors in subsequent steps.

Following the emission inventory, the process delves into air quality modeling. Here, mathematical models are employed to predict how pollutants will disperse in the atmosphere both spatially and temporally. The accuracy of these predictions hinges on the quality of the model, the input data, and a comprehensive understanding of atmospheric processes.

Once the dispersion of pollutants is understood, the next step is exposure assessment. This step determines the extent and duration of human exposure to these pollutants, taking into account factors like population density, activity patterns, and pollutant concentrations. However, it is often challenging to account for individual variations in exposure, resulting from factors such as time spent indoors versus outdoors or activity levels among individuals.

Building upon the exposure assessment, health impact functions are then applied. These functions establish relationships between exposure levels and specific health outcomes, such as respiratory diseases or premature deaths. Establishing clear cause-and-effect relationships between pollutants and health outcomes can be difficult, especially when considering long-term effects or the interactions between multiple pollutants.

The final step in the HIA modeling chain is risk characterization. This step synthesizes the data from the previous stages to estimate the overall health impact in terms of morbidity, mortality, or other relevant metrics. This stage requires the integration of diverse data sources and often involves making assumptions or estimations, which introduces potential uncertainties.

While the HIA modeling chain is a powerful tool, it is not without its challenges. The accuracy of the entire modeling process is only as good as the data provided; incomplete or outdated emission inventories can skew results. Moreover, no model can perfectly replicate the complexities in the real world; simplifications and assumptions are inevitable, leading to potential inaccuracies. Additionally, external factors like climate change, socioeconomic shifts, or changes in healthcare infrastructure can influence health outcomes but might not always be considered in the HIA.

3   Approach

The evolution of HIA methodologies has paved the way for the development of our proposed framework. This framework is not just a tool, but an integrated system designed to bridge the gaps often found in traditional HIA processes. It seeks to provide a comprehensive solution for assessing the health impacts of air quality changes, ensuring that every step, from data input to results interpretation, is seamless, accurate, and user-centric.

The primary objective of this framework is to ease the way HIAs are conducted. It aims to simplify the complex processes involved, making them more accessible to a wider range of stakeholders, from researchers and policymakers to the public. Our vision was to create a platform where data-driven insights can be generated with precision, and results can be visualized intuitively, fostering informed decision-making.

At the core of the framework is its integrated workflow. Users initiate the process by uploading their datasets, including the TDM, emission inventories, meteorological data, population demographics, and health metrics. The system was designed to handle diverse data sources to ensure its flexibility and adaptability.

Once the data are uploaded, the framework begins a series of computational processes. It starts with the preparation of emissions data for different roadway links in a manner suitable for air quality modeling. Users can compare different scenarios at this preprocessing step to ensure the accuracy of the input data. The next step involves preparing the input files required for air quality modeling using AERMOD, a regulatory model developed

by the American Meteorological Society and EPA to predict the dispersion and concentration of pollutants in the atmosphere. Once the AERMOD input files are prepared, the framework helps users run AERMOD without leaving the user interface. In the next step, the tool has built-in post-processing scripts to extract the data from the AERMOD outputs and visualize the data on a map or as a time series. Also, users can compare the results from different scenarios.

Following this activity, the exposure assessment is performed. Here, the framework calculates the extent of human exposure to the predicted pollutant concentrations. It considers various factors, such as population density, activity patterns, and the geographical distribution of pollutants, to determine a population-weighted average concentration of the pollutant at the census tract level. This granularity ensures that the exposure metrics generated are not just averages but are representative of actual population exposures. Building on the exposure data, the framework then applies health impact functions. These functions, derived from epidemiological studies, correlate specific pollutant exposure levels with overall mortality.

One of the standout features of the proposed framework is its emphasis on visualization. Recognizing that raw data or tables can often be overwhelming, the platform offers interactive visual tools. Users can explore the health impacts spatially through detailed maps, observing how impacts vary across regions. Time series charts allow users to see how impacts change over time, providing insights into both short-term and long-term effects.

Additionally, the framework offers real-time feedback. As users adjust parameters or input new data, the visualizations update instantly, allowing for more dynamic exploration and hypothesis testing.

3.1   Advancements Over Previous Models

Several factors set this framework apart from its predecessors:

- **Holistic integration:** By encompassing the entire HIA process, from data input to results visualization, the framework eliminates the disjointedness often found in traditional methods.
- **User-centric design:** The platform was designed with usability in mind. Its intuitive interface ensures that even those without a deep understanding of HIAs can navigate and derive insights.
- **Flexibility:** The system's ability to handle diverse data sources and formats ensures it can be applied in various scenarios, regions, or research contexts.
- **Transparency:** Every step in the process is transparent, allowing users to understand the underlying assumptions, models, and calculations. This transparency fosters trust and facilitates stakeholder engagement.

In essence, the proposed framework is more than just a tool; it is a comprehensive solution that seeks to make health impact assessments more accurate, understandable, and actionable.

4   Implementation

The framework was developed using the R language in the form of an R Shiny application. R Shiny is a powerful tool for creating interactive web applications directly from R scripts. It allows for real-time data processing and visualization, making it an ideal platform for our health impact assessment framework.

The developed R Shiny application serves as an interactive platform for users to conduct a full-chain health impact assessment. The application was designed hierarchically, where the output from one tab feeds into the next, allowing users to progress through the modeling chain seamlessly.

The application boasts a clean and intuitive design, ensuring ease of use, as shown in Figure 1. The sidebar layout is consistently used across all tabs, with input elements on the left and output visualizations on the right. The design also incorporates conditional panels, which dynamically adjust based on user actions, such as adding a second scenario for comparison.
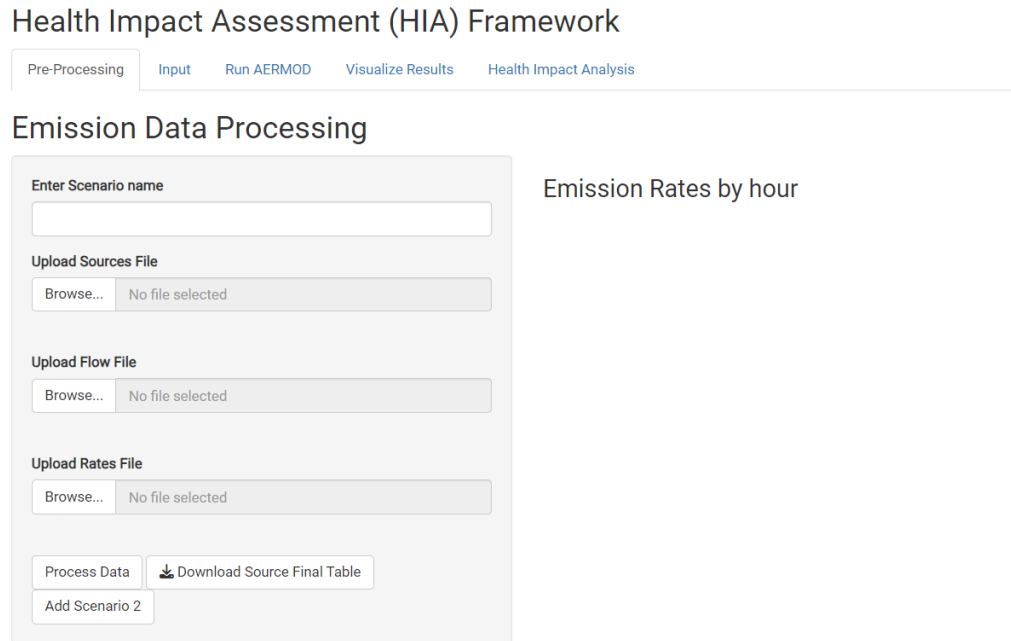

Figure 1. User interface of the health impact assessment framework.

The application is divided into five main tabs:

1. Preprocessing.
2. Input.
3. Run AERMOD.
4. Visualize results.
5. Health impact analysis.

The tabs were designed in a hierarchical order starting from left to right with the output of the tab on the left feeding into the next tab. The hierarchical design ensures a logical flow, guiding users step-by-step through the modeling chain. Action buttons are strategically placed to prompt users to move to the next step, and download buttons allow users to save their processed data at any stage. The application also provides real-time visual feedback, such as plots and maps, enhancing user engagement and understanding.

Sections 4.1–4.5 describe each of the tab's functionalities in detail.

4.1   Preprocessing Tab

The preprocessing tab (Figure 2) is the initial step where users can preprocess the emissions and traffic data, estimate emissions, create geographical projections, and have the inputs converted into a format suitable for the AERMOD. Here, users can input distinct scenario names to add more clarity to their projects. Uploading the source, flow, and rate files becomes a seamless task, ensuring that all necessary data are at their fingertips. One of the standout features is the ability to visualize emission rates on an hourly basis, providing a granular view of the data.

If users wish to draw comparisons, they can effortlessly add a second scenario, making comparative analysis straightforward.

Figure 2. User interface and functionality of the emission data processing tab.

The key points to keep in mind when using the preprocessing tab include the following:

- The sources file should contain details about each emission source, including its location (start and end points), speed, number of lanes, and area type. The identification code (ID) is crucial as it links data across different files.
- The flow file provides flow data for different hours and directions for each source. The ID in this file should match the source file.
- The rates file contains emission rates for different pollutants based on the road description and average speed.
- Once the data are uploaded and processed, the framework will generate a box plot visualizing the hourly emission rates. This visualization helps in understanding the distribution of emission rates and comparing them between different scenarios.

- Users can download the processed data for further analysis or recordkeeping.

4.2   Input Tab

The input tab (Figure 3) has a user-friendly interface designed to streamline the processing of source, receptor, and template data. By simply uploading their respective data files, users can rely on the system to automatically generate the necessary AERMOD input format. This seamless integration is further enhanced by the platform's visualization capabilities. Here, users can actively view the spatial distribution of their sources and receptors on a map, providing a more interactive and comprehensive understanding of their data.
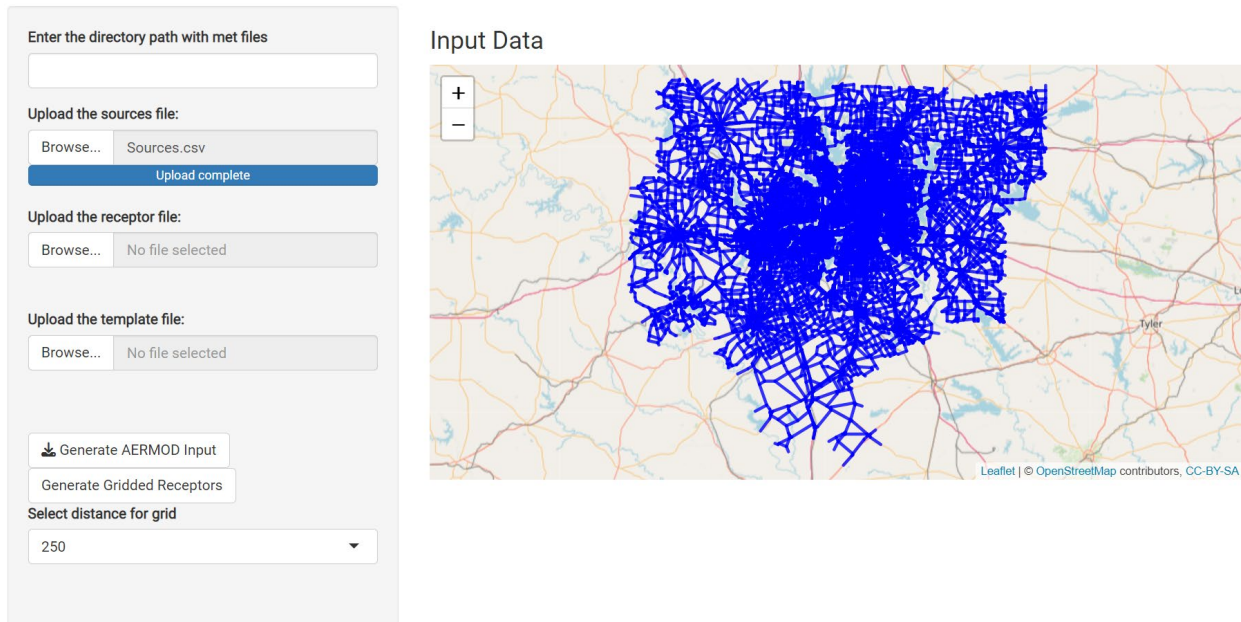

Figure 3. User interface and functionality of the input tab.

For the data upload and reading process, the platform is equipped to handle various data types. Users can upload a comma-separated values (CSV) file containing their source data (generated from the output of the previous tab), which the system efficiently reads using R's *read.csv* function. Similarly, the receptor data, also in CSV format, is converted into the receptor data frame table. Furthermore, the option exists to upload a template file. This file acts as the foundational structure for generating the AERMOD input. It is filled with placeholders that the system intelligently replaces with the actual data. A sample template file is provided with the R code for easy editing.

When users click on the *Generate AERMOD Input* button, the framework edits the template file based on the source and receptor tables. For receptors, the system identifies a specific placeholder line in the template and then formats the receptor data into the AERMOD format using the *sprintf* function. This formatted data then takes the place of the placeholder in the template. The source data, on the other hand, undergoes a three-step processing sequence. Initially, the system formats the source location data, which includes start and end coordinates. Subsequently, emission factors for each source are formatted, followed by the hourly emission data for each source. Each of these formatted sections finds its place in the template, replacing its respective placeholder.

The tool automatically generates visualizations as users upload the data. As soon as the source data are uploaded, a map is generated, showcasing the spatial distribution of the sources as line segments. This map not only provides

a visual representation but also helps to quality check the data spatially. With the upload of the receptor data, the map evolves to display receptor locations, marked distinctly as points.

As a final step of the input tab, users are presented with an option to download this data in the AERMOD input format. With a click, a file named *AERMOD_Input.txt* is generated, containing all the processed data.

Also, recognizing the importance of feedback during data-intensive operations, a modal dialog appears with the message *Processing data... Please wait*. This simple yet effective feature ensures users are always in the loop, and aware that their data are undergoing processing.

4.3   Run AERMOD Tab

As shown in Figure 4, the run AERMOD tab serves as a dedicated platform for executing the AERMOD for air dispersion. Users can specify both the input file and the directory containing the AERMOD executable. The system was designed to seamlessly run the AERMOD and the console output will be readily available for users to review.
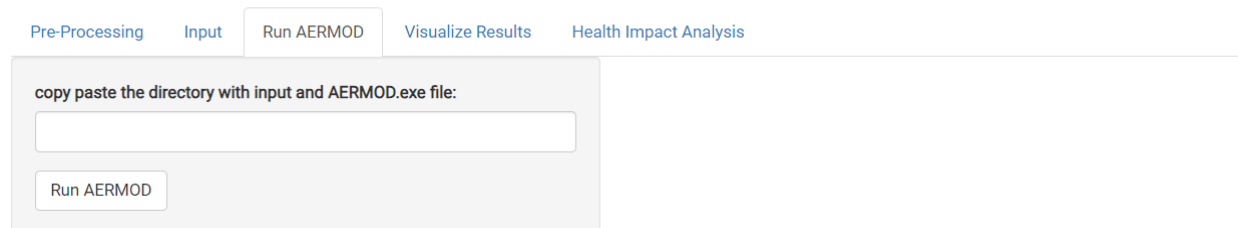

Figure 4. User interface and functionality of the run AERMOD tab.

The AERMOD execution function, termed *run_aermod*, was designed to accept various input parameters, including the path to the AERMOD input file, the desired name for the output file, the directory housing the AERMOD executable, and a Boolean flag to determine the return type of the modeling results. First, the function verifies the presence of the *aermod.exe* file in the designated directory. The function prepares an input file for AERMOD based on the input provided, which can be either a text string, file path, or data frame. It then builds a shell command that directs to the directory containing the *aermod.exe* file and subsequently executes it. After the AERMOD run, the output file is saved in the working directory, which can be renamed as per the user's preference. If the silent flag is deactivated, the function will also read and return the modeling results.

Through the *capture_console_output* function, the system ensures that as AERMOD runs, the console output is captured. This real-time feedback mechanism allows users to monitor progress and instantly view any messages or errors that may be generated during the AERMOD run.

By clicking the *Run AERMOD* button, a series of systematic steps are initiated. First, a modal dialog window informing users with a *Running AERMOD... Please wait* message emerges. The system then activates the console output to capture the functions. It then verifies if users have selected a directory and then constructs a path to the AERMOD input file within the selected directory. With everything in place, the *run_aermod* function is called, which executes AERMOD with the specified parameters. Upon completion, the modal dialog window disappears.

4.4   Visualize Results Tab

As shown in Figure 5, the visualize results tab helps users extract the modeling results from the AERMOD output file and visualize the data. This tab helps interpret the outcomes of the AERMOD results. Also, the tab helps users do a comparative analysis between two distinct scenarios.
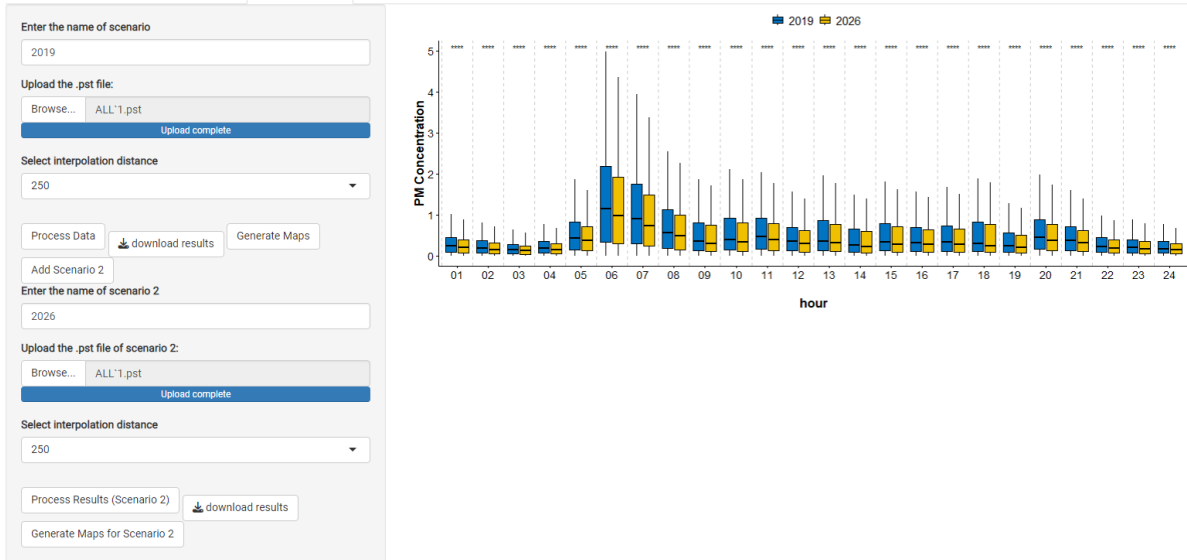
Figure 5. User interface and functionality of the visualize results tab.

The system employs a *read_pos* function to handle the post-processed data. This function reads the entire post file generated by the AERMOD run, filters relevant data lines based on specific patterns, and formats these lines into a singular text block. The data are then stored in a data table.

Next, the *process_pst_file* function reads the post-processed data and splits the date and hour for clarity. To expedite the processing, parallel processing was leveraged to harness the power of multiple cores. For every unique date-hour combination, the function transforms data to align with specific coordinate systems, defines a bounding box by analyzing the geographical extent of all the receptors in the data, and crafts a grid based on the user's input ranging from 250 m to 1,000 m. The inverse distance weighting method is then used for spatial interpolation of the results and the data are visualized on a map. For each map, the visuals of each date-hour combination are saved as an image file in the same folder as the R Shiny application. Once all the processes are completed, the parallel processing setup shuts down.

The system was designed so that the user's interaction with it would be streamlined and intuitive. When users opt to generate maps for individual scenarios via the *Generate Maps* button, a modal dialog window signaling data processing will appear. The *process_pst_file* function is then called to process and visualize data for the chosen scenario. A console message confirms the map generation, and the modal dialog recedes.

For users seeking to compare the analysis between two scenarios, they can click on the *Add Scenario 2* button, which will tell the system to execute the *process_pst_file* function twice, once for each scenario. Once they upload results from both scenarios, a box plot appears comparing the two scenarios temporally by the hour. A t-test for each hour will be performed and the results will be represented on top of each hour. The plot helps users compare the effect of different interventions undergone for a particular scenario and determine whether those interventions have brought statistically significant changes in pollutant concentrations.

4.5   Health Impact Assessment Tab

As shown in Figure 6, the health impact assessment tab helps users understand the intricate relationship between air quality changes and health impacts. Designed with precision, this tab delves deep into the data of baseline and control scenarios, merging them with population metrics to deduce the number of premature deaths that could be averted due to shifts in $PM_{2.5}$ concentrations. The results from the analysis are visualized on an interactive map, and users can download the results in a CSV format for further analysis.
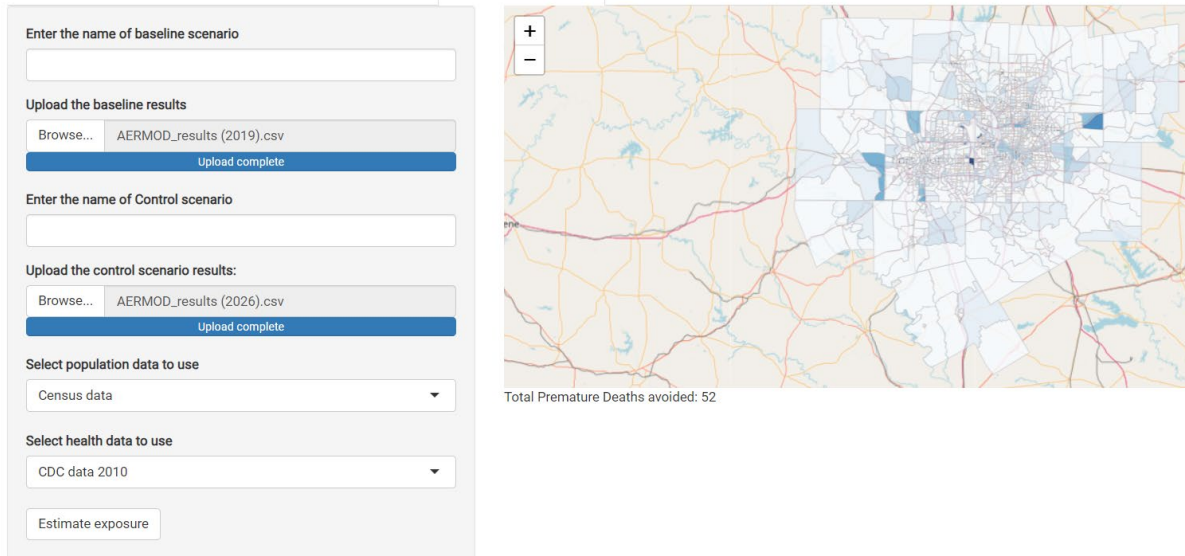
Figure 6. User interface and functionality of the health impact assessment tab.

The system employs reactive functions to read and process the baseline and control data. For the baseline scenario, the system reads the user-provided CSV file (generated from the previous tab) and distributes the data to each receptor to calculate the mean concentration. A similar approach is repeated for the control scenario, ensuring that the data are summarized effectively.

The core of the HIA lies in the processing and visualization of health data. As users initiate the process, a modal dialog window signaling the commencement of data processing emerges. The system then reads the population data; users can select the source of this data, such as Census data. Then, the system performs spatial processing of the population weighted average $PM_{2.5}$ concentration levels for both baseline and control scenarios. The system then calculates the variance in concentrations between the baseline and control scenarios. This variance becomes the foundation for computing the relative risk (RR) and the population attributable fraction (PAF). Finally, the system calculates premature deaths attributable to $PM_{2.5}$ based on the PAF, population count, and mortality rate. Mortality rates are defaulted to the CDC Places dataset, which reports at the county level. The system distributes the county-level mortality rates to individual census tracts based on the population of each tract. Alternatively, users can also select the United States Small-Area Life Expectancy Estimates Project dataset, which presents finer-resolution mortality rates at the census tract level.

Based on the user's selection of which census data and mortality data to use, the map on the main tab updates to represent the number of premature deaths averted. The result is an interactive map, with polygons symbolizing census tracts. Each polygon's hue reflects the premature deaths averted, and a popup adds to the user experience by providing tract-specific details.

The aggregated value of the number of premature deaths averted across all tracts is calculated and displayed at the bottom of the map. Users can also download the health impact results in a CSV format. This CSV file contains data for each census tract, from both the baseline and control concentrations, regarding RRs, PAFs, and the premature deaths attributable to $PM_{2.5}$.

5   Case Study Results

The project team chose to use the North Central Texas Council of Governments (NCTCOG) TDM network as a case study to test the framework. This section explains the different components of the modeling process, which

involved TDM network preprocessing up to visualizing results and comparing different scenarios using the application.

5.1 Traffic Data

The NCTCOG TDM network contains information on the flow and speed of vehicles by time period (morning peak—AM, evening peak—PM, and off peak—OP). The vehicle flow is estimated at different hours using hourly activity factors for the respective county for the 2019 base year and 2026 future year. Vehicle speeds were constant for all hours inside the same time periods for both analysis years. The data tables were inputted into the preprocessing tab of the framework to compare the vehicle flow (Figure 7) and speed (Figure 8) at different hours for all links for both analysis years. As seen in Figure 7, the vehicle flow has two peaks: one at 8–9 AM and another at 5–6 PM for both analysis years. For all time periods, the vehicle flows in 2026 were significantly higher compared to the vehicle flows from 2019.
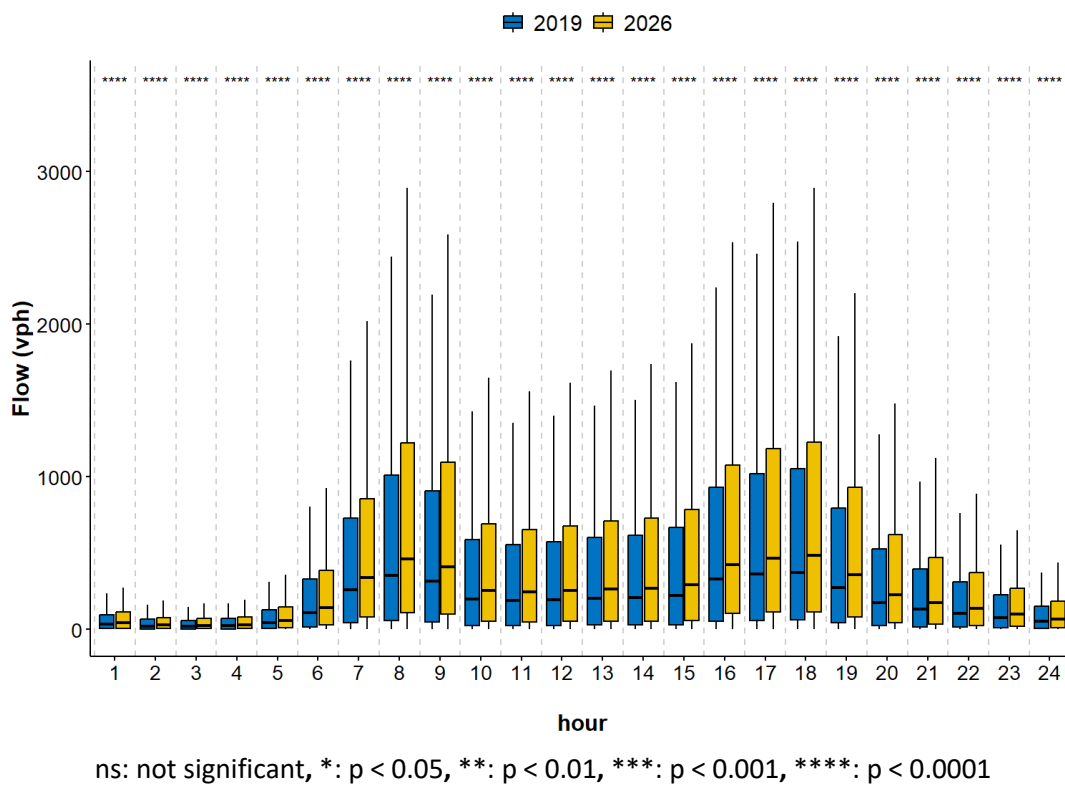

ns: not significant, *: p < 0.05, **: p < 0.01, ***: p < 0.001, ****: p < 0.0001

Figure 7. Comparison of flow (vph) for the analysis years of 2019 and 2026 during weekdays.

ns: not significant, *: p < 0.05, **: p < 0.01, ***: p < 0.001, ****: p < 0.0001
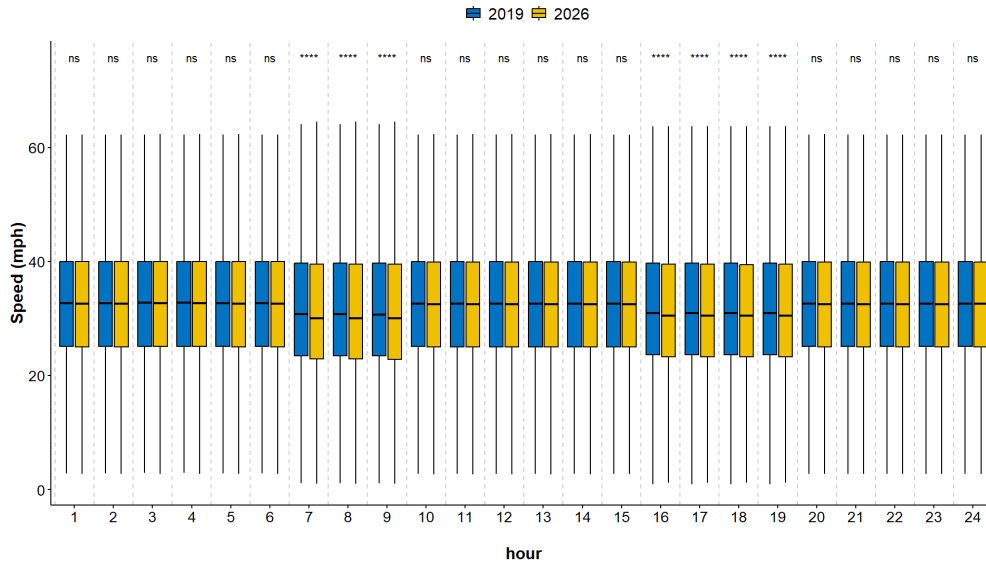
Figure 8. Comparison of speed (mph) for the analysis years of 2019 and 2026.

5.2   Emission Model

The emission rates for both analysis years were derived from the MOVES Emission Rate Lookup Table generated from the MOVES3 model. As shown in Figure 9, the rates for both analysis years were inputted into the preprocessing tab and comparison charts were generated for both analysis years. The emission rates for PM_{10} (selected for the case study) decreased as speeds increased and the 2019 rates were slightly higher than the 2026 rates for all road types. The rates were then used to estimate the emission rates for the AERMOD model (in g/m^2-s), which were combined using the flows, speeds, and rates tables inputted into the preprocessing tab.
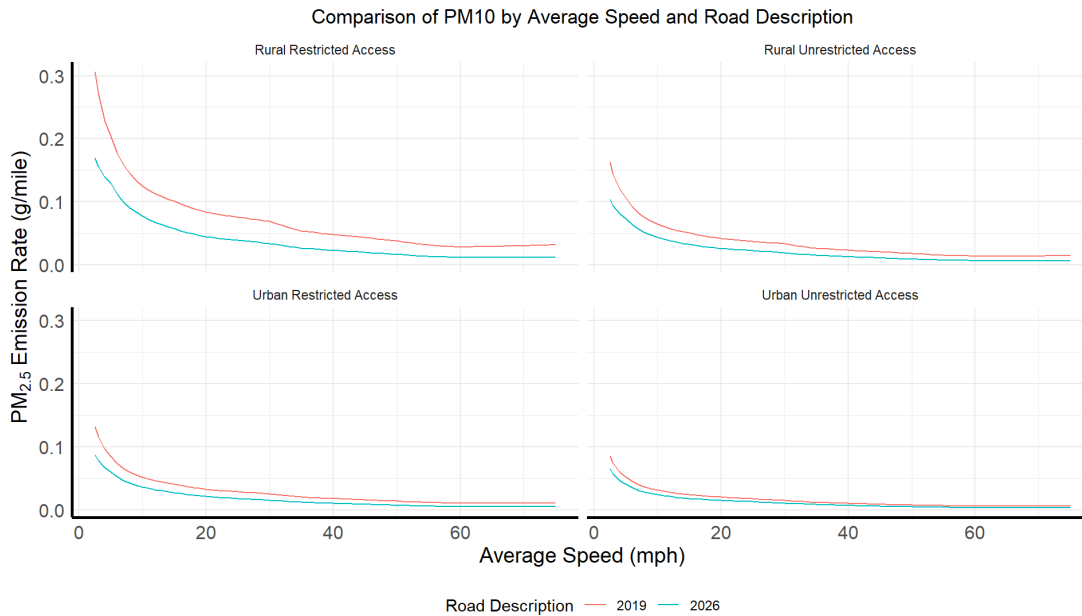

Figure 9. Comparison of emission rates (g/veh-mi) for the analysis years of 2019 and 2026 by road type and average speed.

As shown in Figure 10, the emission rates for 2019 were significantly higher for all time periods compared to 2026. Although vehicle flow significantly increased in 2026 compared to 2019, the vehicles in 2026 were cleaner compared to 2019 based on pollutant emission rates. Thus, the overall emissions were lower in 2026 when compared to 2019.
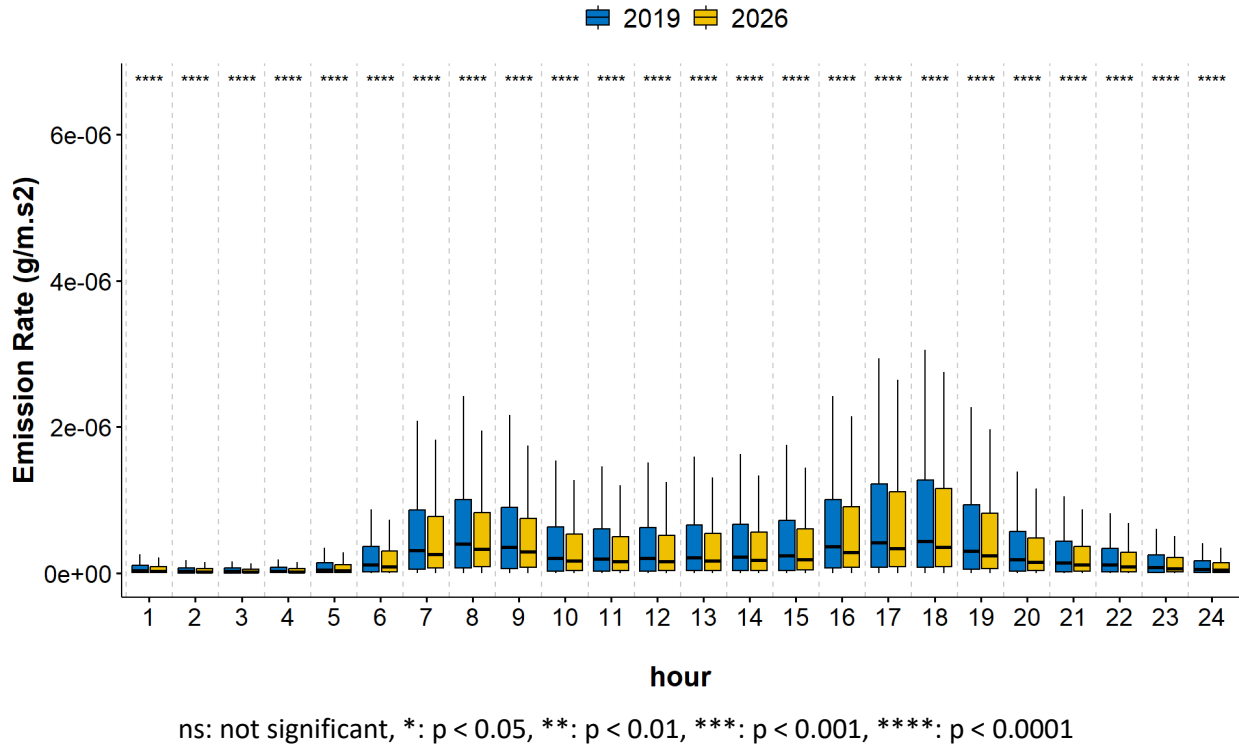

ns: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$

Figure 10. Comparison of emission rates (g/m²-s) for the analysis years of 2019 and 2026 by hour of day.

5.3   Dispersion Model

The hourly emissions and the geometry generated by the preprocessing tab were inputted into the input tab. Receptors were generated at 50 m, 100 m, and 150 m around the centroid of the sources using the functions described in Chapter 4. The meteorological data was obtained from the Texas Commission on Environmental Quality website, which provides preprocessed metadata files for each county that were ready to use for the AERMOD. The AERMOD input, as shown in Figure 11, was generated using the run AERMOD tab. The run produced an AERMOD output and a *.pst* file containing the hourly results. The *.pst* file was processed to extract the results, which are discussed in the next section.

```
CO STARTING
CO TITLEONE  NO TITLE SPECIFIED
CO MODELOPT  DFAULT  CONC  NODRYDPLT  NOWETDPLT
CO RUNORNOT  RUN
CO AVERTIME  1
CO POLLUTID  OTHER
CO SAVEFILE  TMP.FIL
CO FINISHED

SO STARTING
SO ELEVUNIT  METERS
SO LOCATION  2249    LINE    697998.3750  3648145.5000  698417.8125  3648150
SO LOCATION  2256    LINE    697515.9375  3648137.0000  697998.3750  3648145
SO LOCATION  2257    LINE    699031.5000  3648142.5000  699233.5000  3648141
SO LOCATION  2258    LINE    698417.8125  3648150.0000  699031.5000  3648142
SO LOCATION  2259    LINE    697212.2500  3648132.0000  697515.9375  3648137
SO LOCATION  2262    LINE    696799.2500  3648126.5000  697212.2500  3648132
SO LOCATION  2263    LINE    696536.2500  3648125.7500  696799.2500  3648126
SO LOCATION  2264    LINE    695670.5625  3648110.2500  695856.1875  3648111
SO LOCATION  2371    LINE    697998.3750  3648145.5000  698007.7500  3647802
SO LOCATION  2373    LINE    695227.3125  3647616.7500  695501.8750  3647925
SO LOCATION  2380    LINE    697520.9375  3647684.7500  698007.7500  3647802
SO LOCATION  2381    LINE    697515.9375  3648137.0000  697520.9375  3647684
SO LOCATION  2402    LINE    697219.7500  3647622.2500  697520.9375  3647684
SO LOCATION  2403    LINE    697212.2500  3648132.0000  697214.1250  3647924
SO LOCATION  2406    LINE    696926.1250  3647619.0000  697219.7500  3647622
SO LOCATION  2407    LINE    696813.8125  3647617.5000  696926.1250  3647619
SO LOCATION  2408    LINE    696799.2500  3648126.5000  696813.8125  3647617
SO LOCATION  2409    LINE    696548.3750  3647617.2500  696813.8125  3647617
SO LOCATION  2457    LINE    695897.8750  3647459.5000  696548.3750  3647617
SO LOCATION  2479    LINE    699031.5000  3648142.5000  699026.2500  3647787
SO LOCATION  2482    LINE    698018.8125  3647328.5000  698937.3750  3647349
SO LOCATION  2483    LINE    698007.7500  3647802.5000  698018.8125  3647328
```

Figure 11. Snippet of the AERMOD input file generated by the application.

Figure 12 shows the spatial variation of PM$_{2.5}$ concentrations (in µg/m^3) in the NCTCOG area as estimated by the AERMOD model. The concentrations were higher in city centers and by major highways for both years. The PM$_{2.5}$ concentrations appeared to be higher in 2019 when compared to 2026, due to the lower emission rates in 2026.
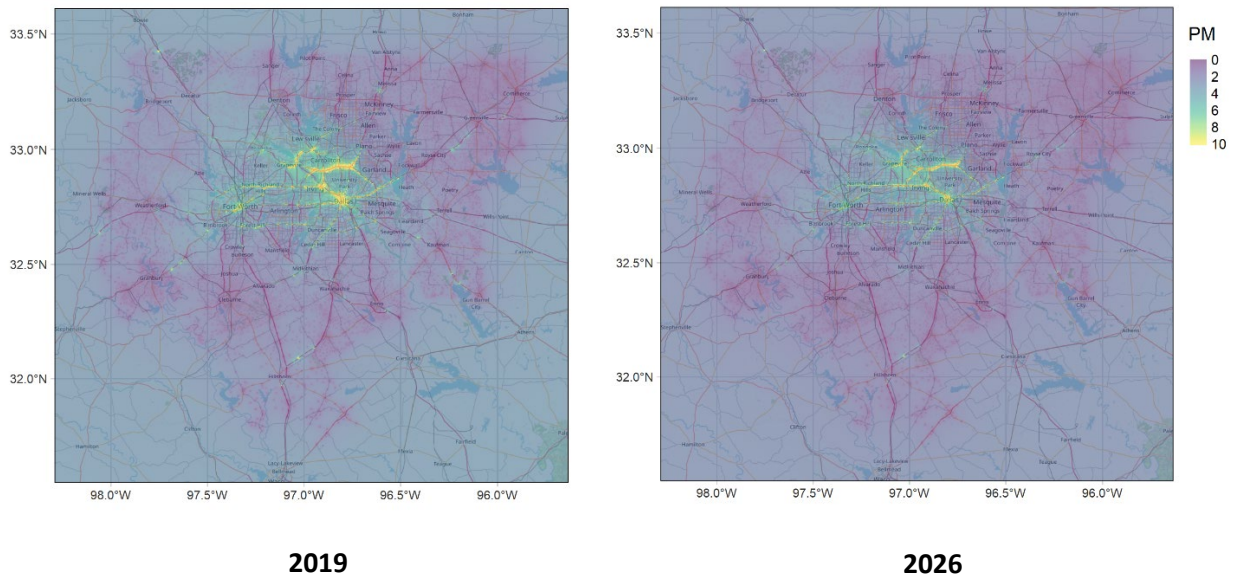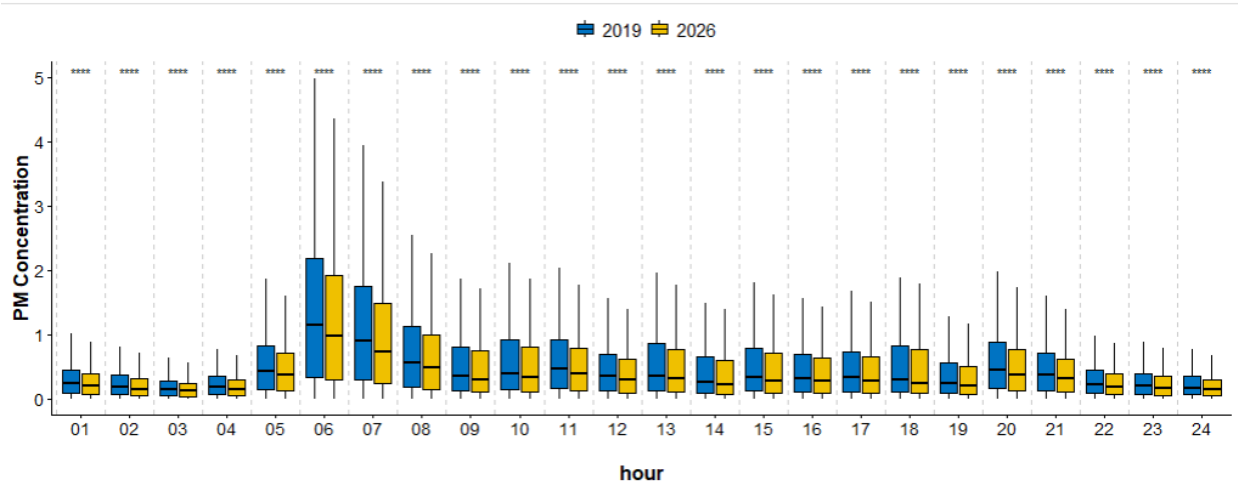

2019   **2026**

Figure 12. Comparison of PM$_{2.5}$ concentrations (µg/m^3) in NCTCOG area for the analysis years of 2019 and 2026.

Figure 13 shows the diurnal variation of the $PM_{2.5}$ concentrations for both analysis years. The concentration was lower for 2026 when compared to 2019, mainly due to the lower emission rates for 2026. As shown in Figure 14, the $PM_{2.5}$ concentrations were lower at 8–9 AM on the July 1 analysis day due to the increase in wind speeds (although the emissions were higher during that time period) as the day progressed, which helped disperse the PM concentrations. Receptors were placed at increasing distances (50 m, 150 m, and 300 m) from the roadways to determine the exposure of $PM_{2.5}$ to the populations living close to roadways. Figure 15 shows the diurnal variation of $PM_{2.5}$ concentrations based on the distance of receptors from the sources. The $PM_{2.5}$ concentrations decreased as the distance from the sources increased.

ns: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$

Figure 13. Comparison of $PM_{2.5}$ concentrations for the analysis years of 2019 and 2026.
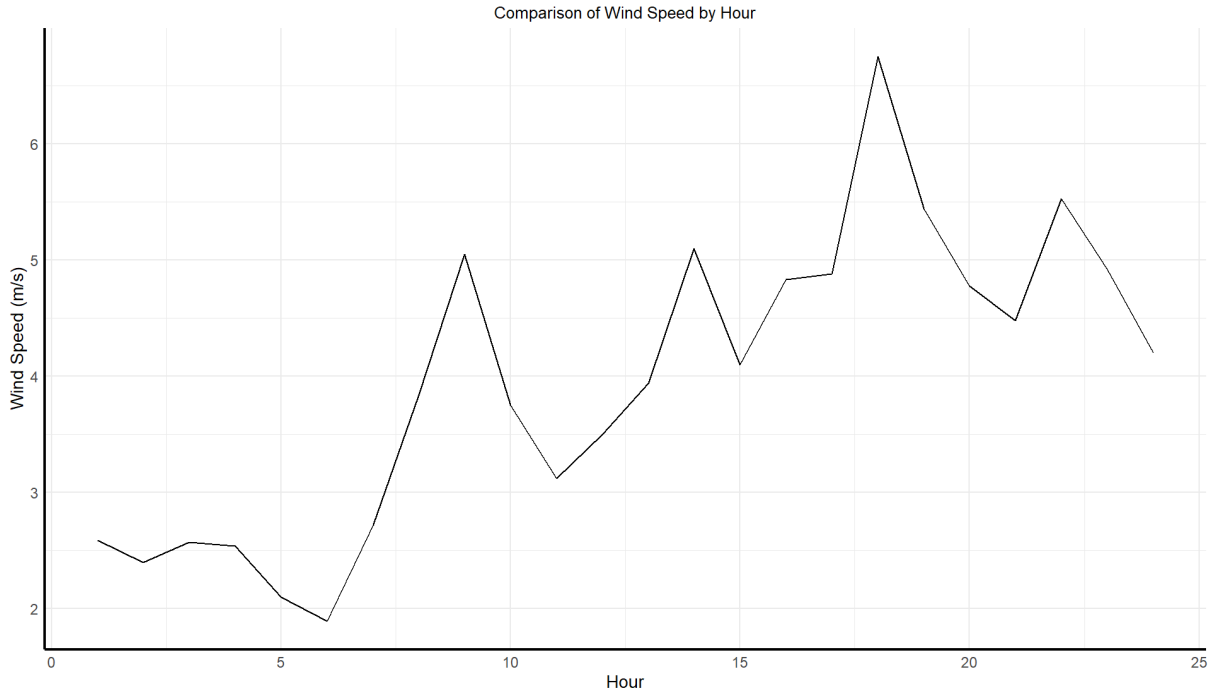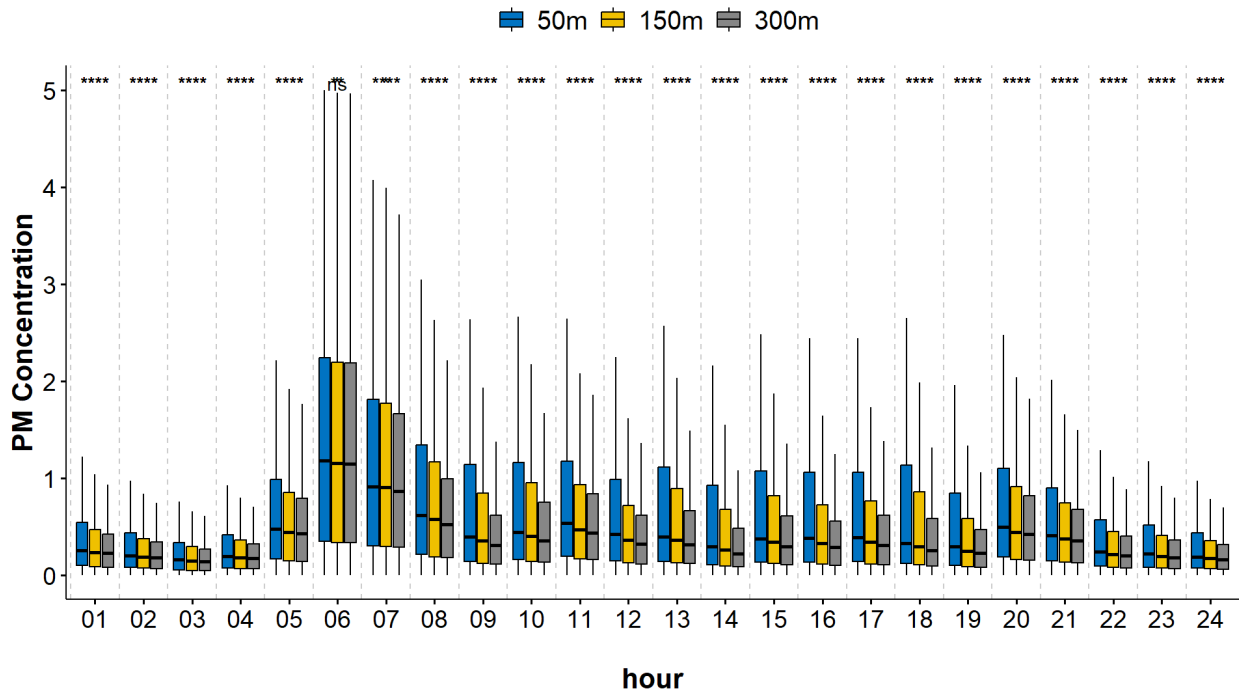
Figure 14. Diurnal variation of wind speed for the modeled day of July 1st.

ns: not significant, *: p < 0.05, **: p < 0.01, ***: p < 0.001, ****: p < 0.0001

Figure 15. Diurnal variation of PM$_{2.5}$ concentrations by the distance of receptors from the source.

Once the results from the dispersion model were obtained, an HIA was performed to evaluate the number of premature deaths that could be avoided by the reduction of PM$_{2.5}$ in 2026 when compared to 2019. Figure 16 shows the geographical representation of the analysis results.
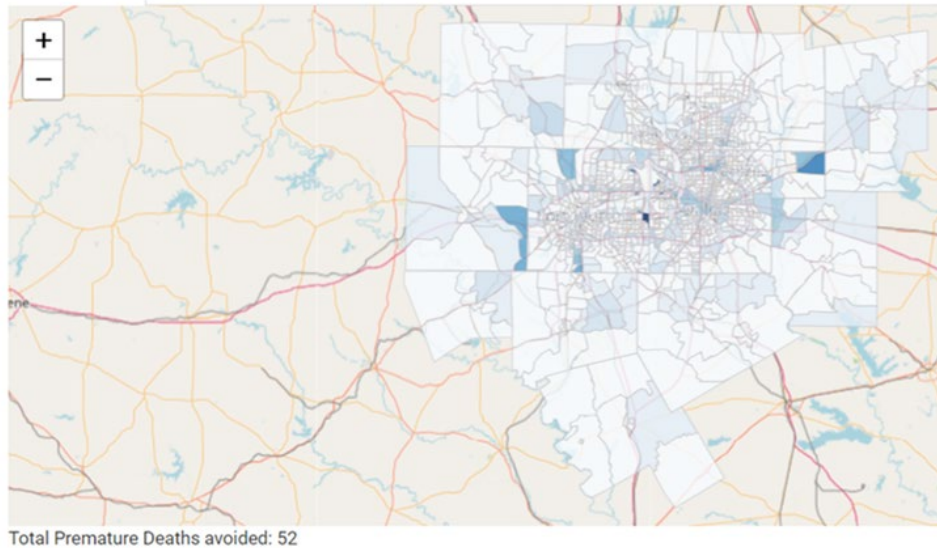
Figure 16. Premature deaths avoided in 2026 by reduction of PM$_{2.5}$ when compared to 2019.

6   Conclusions and Recommendations

The HIA framework documented in this report is a comprehensive tool that integrates various stages of the HIA modeling chain, from data input to results visualization. The framework's design is user-centric, ensuring that even those without a deep understanding of HIA can navigate and derive insights. The framework's emphasis on visualization ensures that users can easily interpret the results. Interactive maps, time series charts, and comparative analyses provide a holistic view of the health impacts of air quality changes. The platform's ability to work with different sources and formats of data enhances its applicability in various scenarios, regions, or research contexts.

The case study using the NCTCOG TDM network showcased the framework's capabilities. The results from the case study highlighted considerable reductions in PM$_{2.5}$ concentrations in 2026 compared to 2019, leading to fewer premature deaths. This reduction was attributed to cleaner vehicles and improved emission rates in the 2026 fleet.

Future work could focus on incorporating more advanced dispersion and photochemical modeling techniques (e.g., CAL3QHC, CAMx, etc.), integrating real-time data sources, and expanding the framework's applicability to other environmental health impact assessments.

7   Outputs, Outcomes, and Impacts

Outputs from this study include the following:

- **Comprehensive HIA framework:** A user-centric tool that integrates various stages of the HIA modeling chain, from data input to results visualization.
- **Interactive visual tools:** Detailed maps, time series charts, and comparative analyses that allow users to interpret the health impacts of air quality changes.
- **Integrated workflow:** An integrated workflow that allows users to seamlessly upload datasets ranging from TDM, emission inventories, and meteorological data, to population demographics and health metrics.
- **R Shiny application:** An interactive platform developed using the R language that allows users to conduct a full-chain health impact assessment.

Outcomes of this study include the following:

- **Enhanced decision-making:** Policymakers, researchers, and the public can make informed decisions based on the insights derived from the HIA framework.
- **Improved air quality modeling:** The framework provides a more accurate representation of how pollutants disperse and concentrate in the atmosphere.
- **Increased awareness:** Stakeholders gain a better understanding of the health ramifications of air quality shifts.
- **Policy recommendations:** The results from the HIA can guide regulatory, legislative, or policy changes to improve air quality and public health.
- **Standardized HIA process:** The framework offers a standardized approach to conducting HIAs, ensuring consistency and reliability across different studies.

Impacts of this study include the following:

- **Reduced health risks:** A decrease in premature deaths and other health risks due to improved air quality.
- **Economic benefits:** A decrease in healthcare costs associated with air pollution-related diseases and conditions.
- **Enhanced public awareness:** A more informed public that can advocate for better air quality measures and make personal decisions that contribute to cleaner air.
- **Environmental benefits:** A decrease in air pollution leading to a healthier ecosystem that benefits flora, fauna, and water sources.

In summary, the outputs of this project provide tangible products and tools, the outcomes lead to changes in the transportation system and its framework, and the impacts reflect the broader effects on the transportation system and society at large.

8 Research Outputs, Outcomes, and Impacts

8.1 Technology Transfer Outputs, Outcomes, and Impacts

R Shiny Application: An interactive platform developed using the R language that allows users to conduct a full-chain health impact assessment.

8.2 Education and Workforce Development Outputs, Outcomes, and Impacts

N/A

9 References

U.S. EPA. (2022). Latest Version of Motor Vehicle Emission Simulator (MOVES). MOVES 2014b:Latest Version of Motor Vehicle Emission Simulator. U.S. Environmental Protection Agency (U.S.EPA). Retrieved from https://www.epa.gov/moves/latest-version-motor-vehicle-emission-simulatormoves

USEPA. (2023). Motor Vehicle Emission Simulator: MOVES3 (Version 3.0) [Computer software]. https://www.epa.gov/moves

10 Appendix

Appendix A. Template for AERMOD Input Generator

Appendix B. Framework R Codes on Shiny Application