Open Access

Journal of Energy and Power Technology

LIDSEN Publishing Inc.

Original Research

Traffic NO_x Pollution Prediction and Health Cost Estimation Using Machine Learning: A Case Study of Toronto, Canada

Hamidreza Shamsi [1, †], Ehsan Haghi [2, ‡, †], Manh-Kien Tran [2, †], Sean Walker [3, †], Kaamran Raahemifar [4, †], Michael Fowler [2, †, *]

1. Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada; E-Mail: hshamsi@uwaterloo.ca
2. Department of Chemical Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada; E-Mails: ehaghi@uwaterloo.ca; kmtran@uwaterloo.ca; mfowler@uwaterloo.ca
3. Department of Chemical and Biomolecular Engineering, University of South Alabama, Mobile, AL 36688, USA; E-Mail: seanwalker@southalabama.edu
4. College of Information Sciences and Technology (IST) Data Science and Artificial Intelligence Program, Penn State University, State College, Pennsylvania, PA 16801, USA; E-Mail: kvr5517@psu.edu

‡ Current Affiliation: Musqueam Indian Band, 6735 Salish Drive, Vancouver, BC, V6N 4C4

† These authors contributed equally to this work.

* **Correspondence:** Michael Fowler; E-Mail: mfowler@uwaterloo.ca

Academic Editors: Francesco Gabriele Galizia and Marco Bortolini

Special Issue: Energy Efficiency in Flexible and Reconfigurable Manufacturing: Emerging Trends, Models and Applications in the Industry 4.0 Era

Journal of Energy and Power Technology
2022, volume 4, issue 4
doi:10.21926/jept.2204043

Received: March 11, 2022
Accepted: December 15, 2022
Published: December 25, 2022

Abstract

Road traffic is a significant source of air pollution that has a harmful impact on human health. To reduce the health and environmental impacts of fossil fuel consumption in the transportation sector, many countries have implemented policies to promote the deployment of electric vehicles (EVs). A vital factor to consider when designing policies to support EV use

is the monetized health impacts of fossil fuel consumption. This research aims to investigate the health benefit of replacing internal combustion engine vehicles (ICEVs) with zero-emission vehicles in the city of Toronto, Canada. A long short-term memory (LSTM) model is developed in this work to predict future NO_x concentrations considering the effect of the traffic volume, weather, time of day, and historical NO_x concentrations. The developed model is then used to predict long-term NO_x concentrations and annual average NO_x reduction from zero-emission vehicle deployment in four different scenarios in Toronto. Additionally, interpolation methods are used to predict the pollution reduction in all Dissemination Areas (DA) of Toronto, and a health cost assessment is conducted to estimate the health benefit in all the scenarios. The results of the modeling in this work show that the western areas of Toronto experience higher NO_x concentration reduction in all scenarios. These reductions are the result of the higher correlation between traffic volume and pollution in those areas. The results also show that with a 10% reduction in ICEV traffic volume, 70 premature deaths can be prevented annually, equivalent to 560 million CAD in health benefits per year.

Keywords
Air pollution; machine learning; ZEV; health cost; long short-term memory

1. Introduction

The transportation sector is one of the most energy-consuming economic sectors in the world. In 2018, about 50% of the world's oil consumption was used for road transportation [1]. Additionally, transportation is an essential contributor to economic development and social growth and this sector is expected to grow with population and economic growth. The utilization of this sector has risen over the years, with the number of passenger-kilometers increasing by 260% in the Organisation for Economic Co-operation and Development (OECD) countries between 1970 and 2008 [2]. However, since most automobiles use fossil fuels such as gasoline and diesel, the transportation sector negatively affects society through air pollution and greenhouse gas (GHG) emissions. In Canada, the transportation sector accounts for 24% of total annual emissions, with road transport accounting for 82.5% of the transportation sector [3]. GHG emissions from the transportation sector are from fossil fuel combustion in internal combustion engine vehicles (ICEVs) that provide transportation services. The GHG emissions and contribution to climate change are not the only externality of using fossil fuels in the transportation sector. Air pollution is the largest environmental cause of diseases and premature death in the world with an estimated 9 million premature deaths caused by pollution in 2015 accounting for 16% of all deaths worldwide [4]. In some countries, air pollution can be the cause of more than 25% of all deaths. In Canada, it was estimated that there were on average, about 21,000 premature deaths caused by air pollution during the 2000s [5]. Some of the most common pollutants released from ICEVs include nitrogen oxides (NO_x), sulfur oxides (SO_x), and particulate matters ($PM_{2.5}$ or PM_{10}), all of which have negative effects on human health [6]. NO_x, SO_x, and particulate matter ($PM_{2.5}$ or PM_{10}) air pollutants have been found to be a factor in causing serious diseases such as chronic obstructive pulmonary disease, chronic ischemic heart disease, diabetes, strokes, and tracheal, bronchial, and lung cancers [7].

Due to the negative effects on human health and high GHG emissions from traffic-related fossil fuel consumption, governments around the world have started to focus on cleaner transportation solutions [8]. Electric vehicles (EVs) are a promising alternative to conventional ICEVs. Depending on the electricity source used to charge the vehicle, EVs can be zero-emission vehicles (ZEVs) and not release any harmful pollutants during the use phase [9]. EVs can use batteries [10, 11] or hydrogen fuel cells [12] as energy storage systems although these types of EVs are often considered to be distinct technologies. Batteries are better suited for light-duty and passenger vehicles while hydrogen fuel cells are more fitted to be used in medium and heavy-duty trucks as well as trains and industrial vehicles. Since heavy-duty trucks release more air pollutants than passenger vehicles, the use of hydrogen fuel cells in heavy-duty trucks is a more promising pathway for the reduction of traffic-related air pollution [13].

Despite the advantages of EV deployment such as reducing air pollution and GHG emissions from the transportation sector, the transition from conventional vehicles to alternative options has been facing major challenges. The cost of infrastructure including generation capacity, transmission lines, and charging/refueling infrastructure to support EVs, for instance, is a major roadblock to their mass rollout. As a result, the transition to EV utilization in the transportation sector, as in any socio-technical system, must occur gradually and logically to ensure optimal outcomes in terms of socioeconomic benefits [14]. In recent years, many researchers have been working on investigating the benefits of EVs and their effect on the reduction of traffic-related air pollution and its negative health impact. Schneidemesser et al. [15], conducted a study to quantify the air pollution from vehicle use in urban areas. The authors considered the effect of environment, density traffic, and vehicle type (buses, trucks, personal cars) on the particle concentrations. The focus of the analysis done was to assess the cyclists' exposure to pollutants by analyzing tracks with accompanying video footage. Ventura et al. [16], assessed the effect of vehicle inspection and maintenance programs on air pollution. The authors analyzed the available data from 2014 to 2017 in Rio de Janeiro state, Brazil. The conclusion of the analysis showed that lack of maintenance leads to the increase of CO and hydrocarbon emission from vehicle use up to 5 times compared to national limits. Ke et al. [17] developed a model to estimate energy consumption, GHG emissions, and pollutant emissions from different light-duty passenger vehicles. Each vehicle type emission was estimated using the assumed vehicle emission factors of various vehicle technologies. The emission factors for VOCs, NO_x, and $PM_{2.5}$ for: multiport fuel injection (MPFI) vehicles, gasoline direct injection (GDI) vehicles, and hybrid electric vehicles (HEVs). The results of the analysis performed by Ke et al. showed that while battery electric vehicles (BEVs) have a high potential in reducing pollution, the reduction of NO_x from vehicle use, however, depends on the source of electricity and will drop if the share of non-fossil electricity in imported power reaches 30%. Some researchers have used empirical data collection to assess the pollutant emissions from different types of vehicles. Kebede et al. [18], for instance, used random roadside testing of different public transport vehicles to analyze the standard compatibility of on-road vehicles. The analysis was done by collecting data from random roadside inspections of diesel-fueled vehicles in Addis Ababa, Ethiopia. The authors tested 358 vehicles manufactured between 1960 and 2017 including minibuses, mid-sized buses, and large buses.

The limits of traditional methods have made machine learning a popular tool in recent years for air pollution modeling. Machine learning is used to refer to a wide range of techniques that use available data to gain knowledge about the correlation of different parameters and enable forecasting. Bougoudis et al. [19], developed a machine learning model to forecast the air pollutant

concentration in the Attica area, Greece. The model developed used clustered datasets that shared similar characteristics including pollutant concentration, day, hour, month, temperature, and relative humidity. In this way, the model gained knowledge on the correlation of the different factors on pollutant concentration. The model then used the knowledge to forecast pollutant concentration based on contributing factors. Lautenschlager et al. [20] developed a machine learning for air pollution modeling that could work based on openly available data source OpenStreetMap. Sinnott and Guan [21] assessed the potential of linear regression models, artificial neural network (ANN), and long short-term memory (LSTM) models in $PM_{2.5}$ pollution prediction. The results of the analysis done in showed LSTM forecasted the $PM_{2.5}$ concentration with the highest accuracy. Both linear regression and ANN models did not perform well in forecasting high $PM_{2.5}$ concentration values.

Air pollution modeling is used in the literature to quantify the reduction of air pollution by mass EV rollout. A common approach to model the spatial distribution and concentration of traffic-related air pollution is the land-use regression model [22]. In recent years, machine learning methods have been utilized to obtain more accurate predictions of air pollution released by vehicles [23]. Wen et al. [24], for instance, developed a novel spatiotemporal convolutional LSTM neural network to predict air pollution concentration. The model inputs were the delay between $PM_{2.5}$ concentrations from monitoring stations as well as some meteorological and aerosol data such as humidity, temperature, wind speed, planetary boundary layer height, and aerosol optical depth (AOD) near the stations. $PM_{2.5}$ concentration data from over 1000 air quality monitoring stations in China were used to validate the model [16]. The use of meteorological and aerosol data was found to improve the accuracy of the proposed model significantly. Tong et al. [25], proposed a deep learning spatiotemporal model combining LSTM and recurrent neural network to predict the daily concentration of $PM_{2.5}$. Due to the lack of data, only three features were used, which were longitude, latitude, and time. The model was validated using ground $PM_{2.5}$ data from the US Environmental Protection Agency (EPA)'s Air Quality System (AQS) and was found to have acceptable accuracy. Adding more features such as AOD, land use, roads, emissions, elevation, and weather conditions could improve the accuracy of the model, as it was determined that the temporal correlation was superior to the spatial correlation. Huang et al. [26], predicted the $PM_{2.5}$ concentration in the air in Beijing and Shanghai using a deep neural network model which was the combination of LSTM and convolutional neural network. The input features used in the model included NO_x concentration, accumulated wind speed, and accumulated hours of rain. Results showed that the overall accuracy of the model was verified with low average mean absolute error (MAE) and root mean square error (RMSE) from the model outputs. Fan et al. [27], presented a spatiotemporal prediction framework using a deep recurrent neural network consisting of LSTM layers and fully connected layers. This model was able to handle missing data in the time series. The model was trained using real-world air quality and meteorological datasets in the Jingjinji area of China, and it was found to have a high accuracy in predicting sudden heavy pollution events and average patterns. Qin et al. [28], were able to forecast the $PM_{2.5}$ concentrations in some regions of China as a time series. A convolutional neural network was used as the base layer to extract input features, which were meteorological data and pollutant concentrations in their air pollution prediction model. An LSTM network was also used to extract the time-series features for the input data. This approach performed well in predicting $PM_{2.5}$ concentrations, but it could be improved by adding more factors such as geomorphic conditions. Šimić et al. [29], compared the performance of five different machine

learning regressors in regards to predicting PM_{10} and NO_2 concentrations in the city of Zagreb, Croatia, and found that Lasso regression was the best performing algorithm. Additionally, it was shown that seasonal weather conditions and traffic locations affected the concentrations significantly. Li et al. [30], used machine learning algorithms to determine hourly street-level $PM_{2.5}$ and NO_x concentrations. Random forest was determined to the best-performing algorithms out of the six that were utilized and evaluated. It was also found that non-emission factors, like non-local pollution and temperature, accounted for a significant amount of pollution concentration predictions, while the rest came from direct emission contributors like vehicles. Other studies, such as Lana et al. [31] have shown that meteorological conditions and local temperature pressure and humidity conditions can have a greater control over pollution levels than traffic levels. When concentrations of NO_x are lower, traffic levels have a greater impact on pollution levels [32, 33].

The review of the air pollution modeling literature shows that the research is mainly focused on analyzing $PM_{2.5}$ emissions. However, NO_x emission modeling has not been thoroughly investigated in the literature, although NO_x is a harmful pollutant that is released in large quantities by diesel vehicles which are very common in the transportation sector. NO_x is the cause of many health concerns, including serious respiratory diseases [34]. About 10,000 premature deaths in Europe in the year 2013 can be attributed to high NO_x emissions from light-duty diesel vehicles [35]. In that sense, it is crucial to use machine learning methods to investigate and predict NO_x concentration from fossil fuel consumption and to quantify the health benefits of NO_x reduction via transport electrification.

In this work, NO_x concentration in Toronto is predicted using an LSTM model that is trained based on previous timestep data. The Keras module in Python is utilized to develop the model. LSTM is a recurrent neural network model that uses a sequence of data to predict the outcomes in the future. It is widely used to predict energy consumption, weather forecast, traffic forecast, or air pollution concentration [28, 36-38]. The neural network undergoes a sensitivity analysis to determine the best network parameters. A long-term prediction is then presented using the developed model. Finally, the health and economic benefits of ZEVs are estimated using the results from the machine learning air pollution prediction model.

Air pollution concentration in an area is a function of pollutant sources and weather conditions. Traffic-related air pollution, HVAC (heating, ventilation, and air conditioning) system air pollution, and out-of-Toronto air pollution are critical sources of pollutant emissions in the city of Toronto. Weather parameters including temperature, wind, humidity, precipitation, and solar radiation are also factors that can affect pollution concentrations in different areas.

This research contributes to the literature by:

- Using a machine learning model to predict air pollution in the city of Toronto;
- Using traffic data, in addition to air pollution data and weather data, as an input to the LSTM model that is used to predict air pollution centration in different areas of Toronto; and
- Estimating monetized health impacts from pollution emission based on the concentration predicted by the model.

2. Materials and Methods

This section provides an overview of the methodology and the assumptions used for modeling NO_x emission in this work. Figure 1 shows the methodology used in this work to estimate the health cost caused by traffic-related air pollution in Toronto.

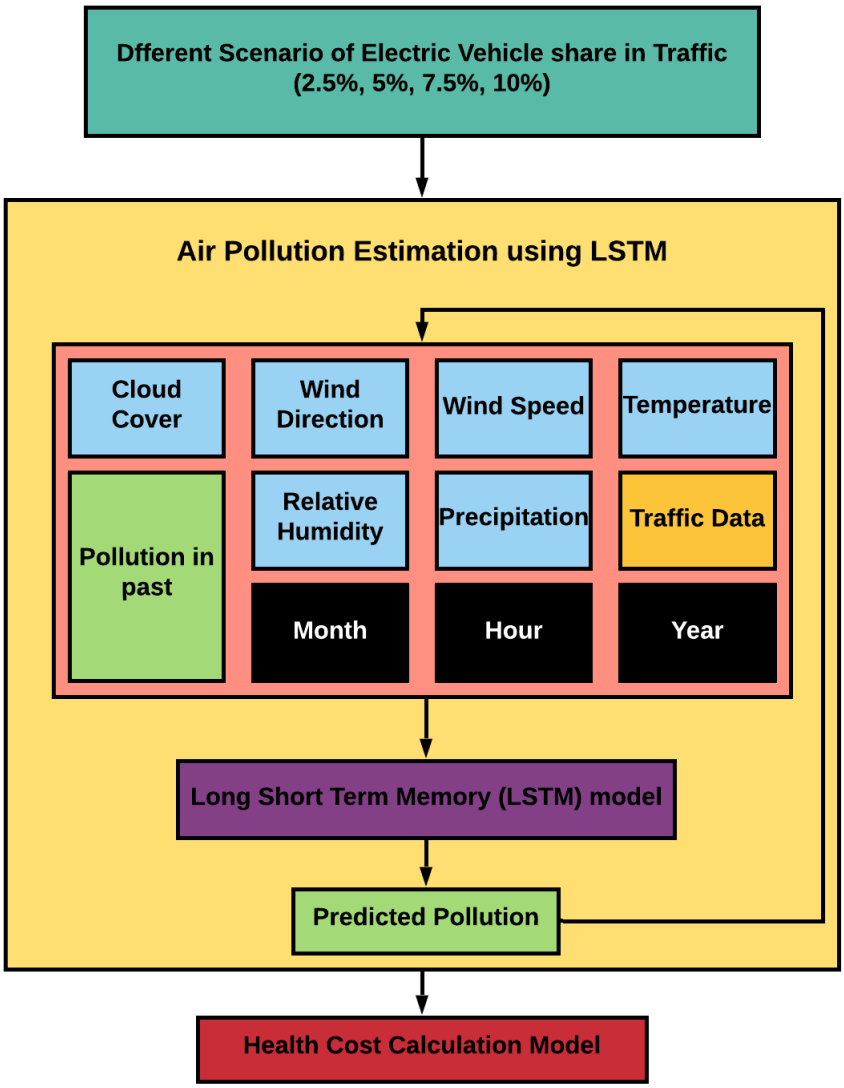

Figure 1 Methodology used to estimate the health benefit of increasing EV share in the city of Toronto.

In the model developed in this work, an LSTM model is first developed using features including weather data and time of the day as shown in Figure 1. The objective of the LSTM model is to predict the NO_x pollution based on the weather, time, traffic count, and NO_x pollution from previous timesteps. These features are important factors affecting the pollution quantity in the air [39, 40]. The developed model is then used to estimate the annual average NO_x concentration in four different locations in the city of Toronto in four different scenarios representing four traffic compositions. The annual average NO_x concentration in Toronto's Dissemination Areas (DA) is estimated using nearest, linear, and cubic interpolation methods. Based on the estimated pollution reduction and population data, the decrease in mortality rates is calculated using Hazard Ratio (HR).

Finally, the prevented deaths in all scenarios are converted to monetary values using Value of Statistic Life (VSL) [41].

Figure 2 shows the features used to predict the NO_x concentration in Toronto. There are four air pollution monitoring stations that report the hourly concentration of different pollutants in Toronto: Toronto Downtown, Toronto East, Toronto North, and Toronto West [42]. Additionally, the hourly traffic count data in forty stations were used to input the LSTM model [43]. The other input to the LSTM model is the weather data as it has a significant effect on pollution dispersion. Time data, including year, month, day of the week, and hour of the day are used to consider the temporal effects on NO_x concentration.

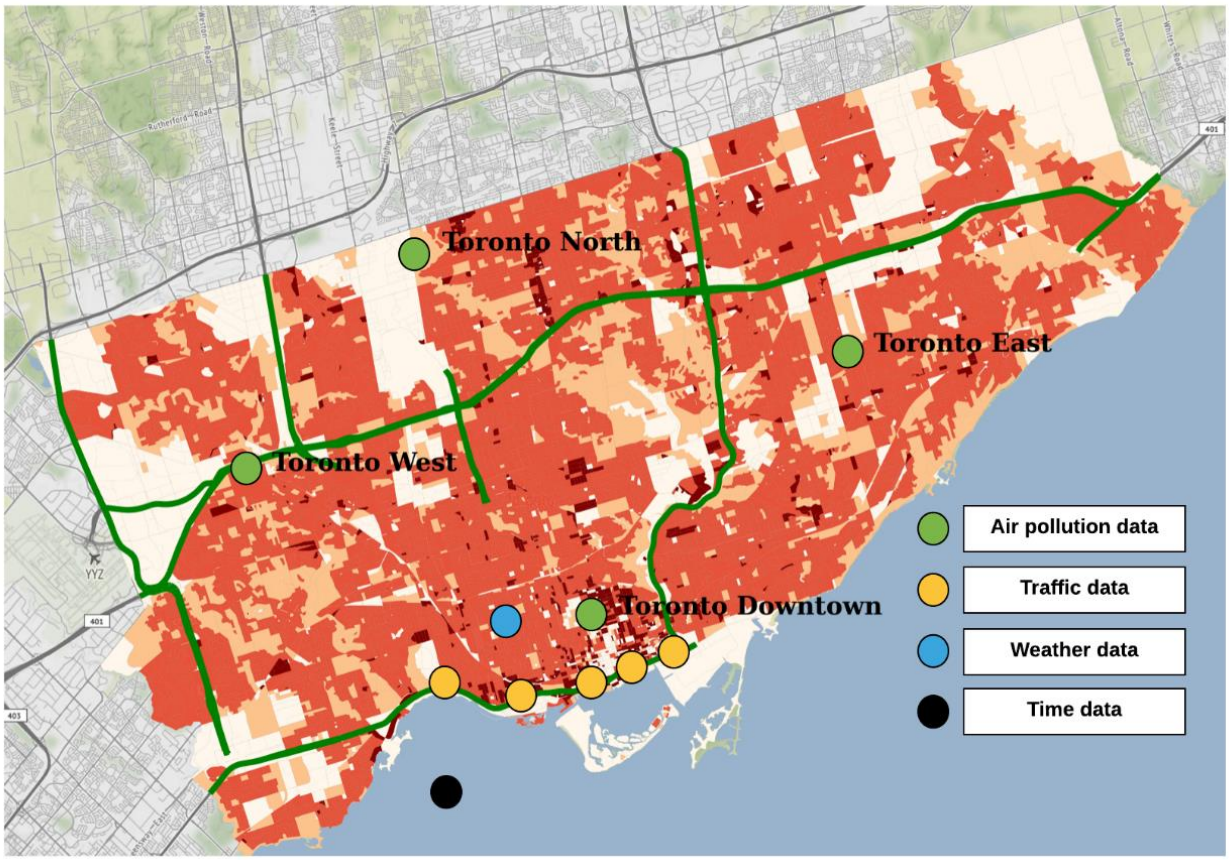

Figure 2 Schematic of the features used in the LSTM model.

2.1 Weather Data

The hourly weather data in the Toronto City Centre station is extracted for the 2015-2017 interval [44]. It includes temperature, relative humidity, wind direction, and wind speed. To shed light on the weather properties in this study, a brief explanation is represented below.

2.1.1 Temperature

Figure 3 shows the hourly average temperature for each month in the city of Toronto. As can be seen in Figure 3 the temperature drops to a minimum in the early morning and has a maximum in

the afternoon in all months. A comparison of Figure 3 and Figure 8 shows the idea that an increase in the temperature on the ground has a significant impact on the amount of NO_x concentration.

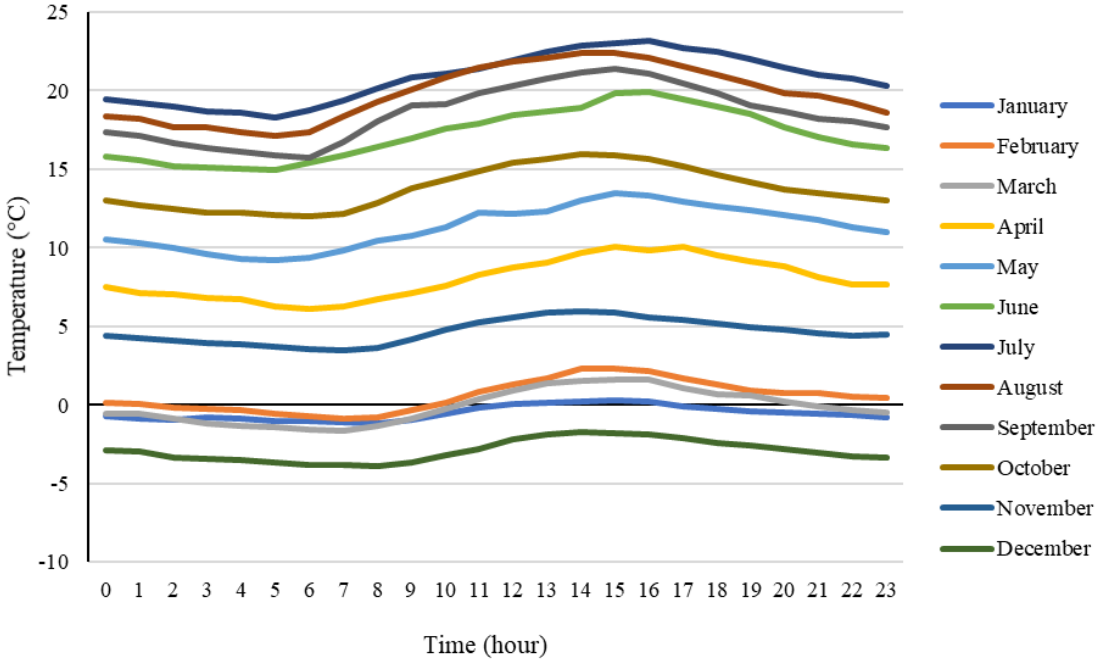

Figure 3 Monthly average of Temperature during a day.

2.1.2 Wind Speed

Figure 4 shows the hourly average wind speed changes during the day for each month. Except for December and January, the wind speed profile shows a vast difference between morning and afternoon wind speed. Also, it can be seen that the wind speed is higher in colder months. Wind speed not only affects the dispersion of the pollutants but also can influence the amount of pollutants coming from other regions.

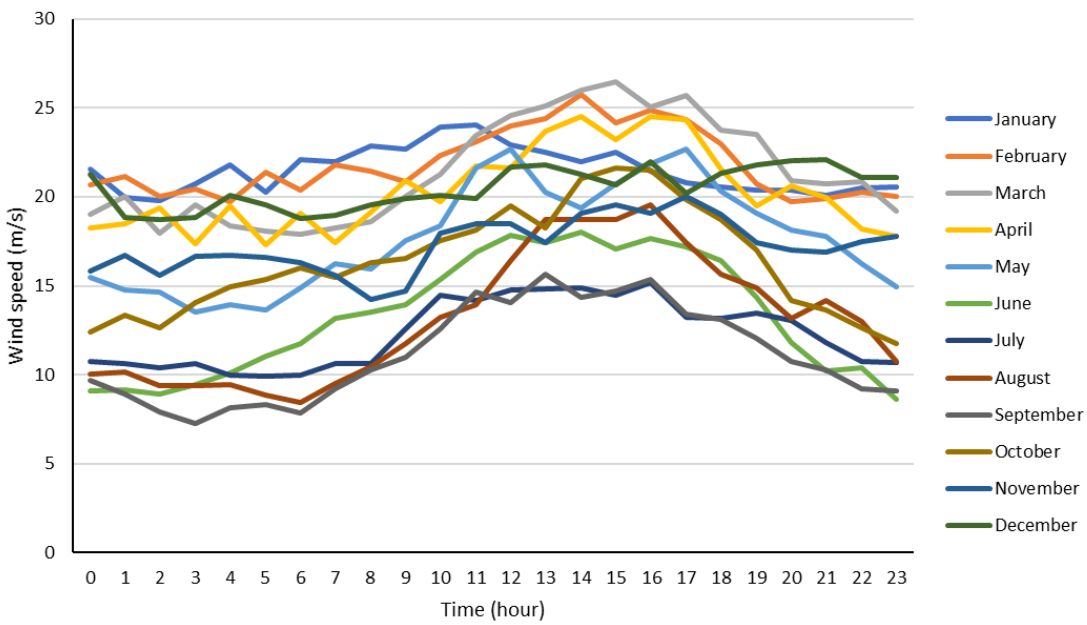

Figure 4 Monthly average of Wind Speed during a day.

2.1.3 Precipitation

The hourly data for precipitation could not be found on the web; however, the number of days with precipitation depicted in the figure below can show the importance of such a dataset. Figure 5 shows the number of days with precipitation (rain or snow) in each hour during a year. Instead of precipitation quantity, the weather condition is used as a feature to train the machine learning model. In this regard, three weather conditions are considered: No precipitation, Rainy, and Snowy.

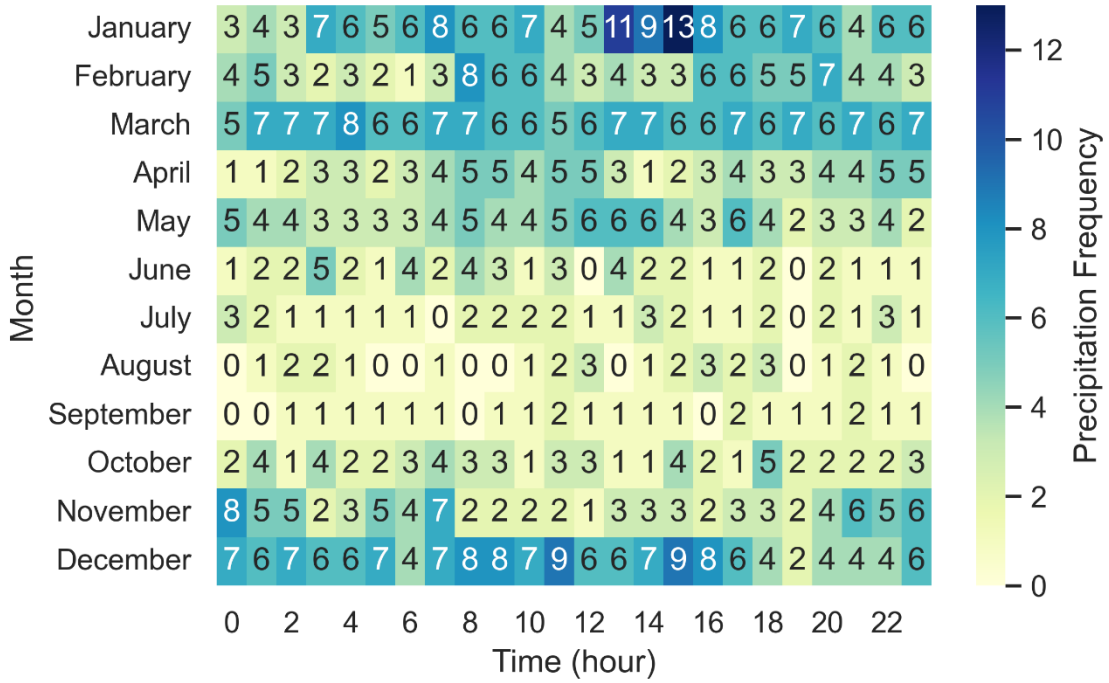

Figure 5 Number of days with precipitation in every single hour.

2.2 Traffic Count Data

Figure 6 and Figure 7 show the traffic count distribution over a year for a single location (Lake Shore Blvd, East Bound, West of Oarsman Dr) in the city of Toronto is shown [43]. The number of vehicles on weekdays is higher than on weekends. Also, the traffic pattern is different for weekdays and weekends because of work commuters during weekdays. As a result, the day of the week is chosen to be an input feature to the LSTM model for predicting air pollution.

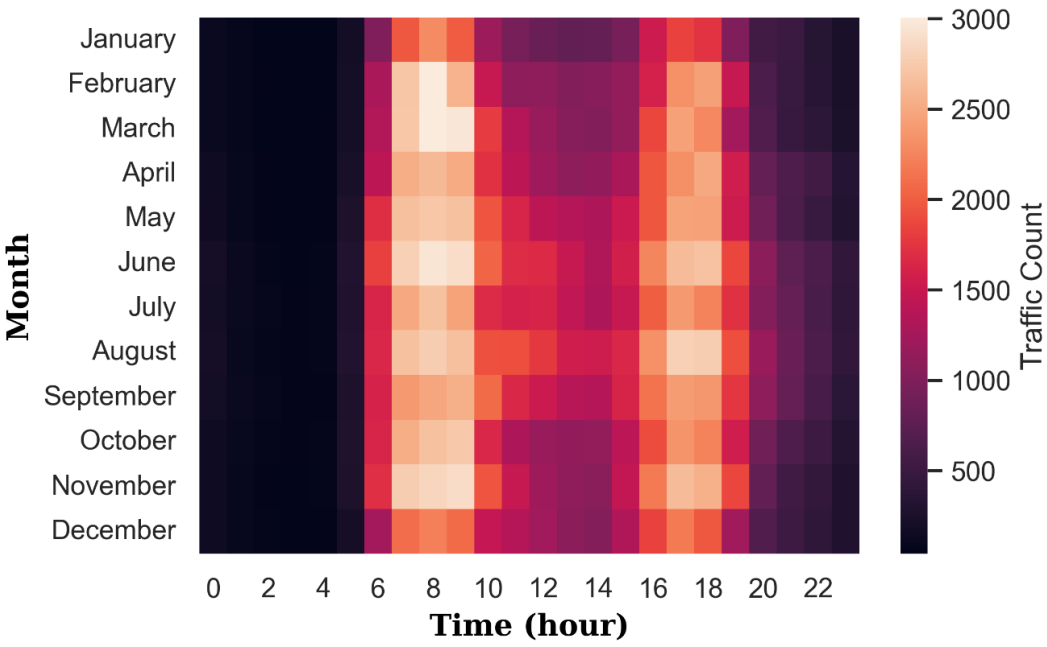

Figure 6 Monthly average traffic count data (weekdays).

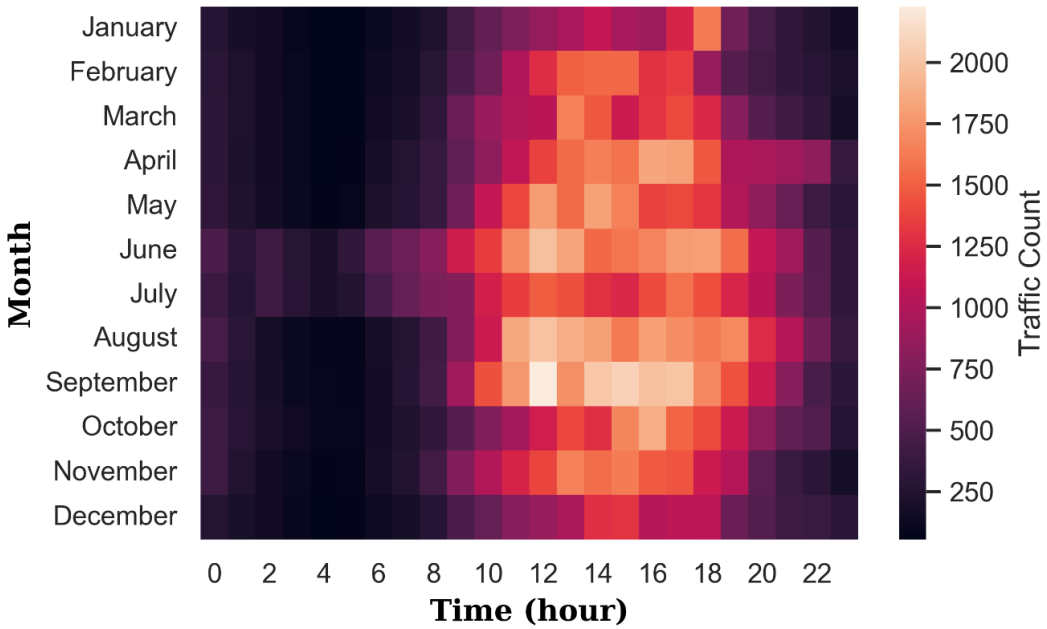

Figure 7 Monthly average traffic count data (weekends).

2.3 Air Pollution Data

Figure 8 shows the monthly average NO_x concentration at the Toronto Downtown air pollution measurement station. The minimum amount of air pollution, at this station, happens in the months of May through July, during the Spring and early Summer. The main reasons for this pattern are the higher use of heating systems and the variation of sunlight in Winter and Summer.

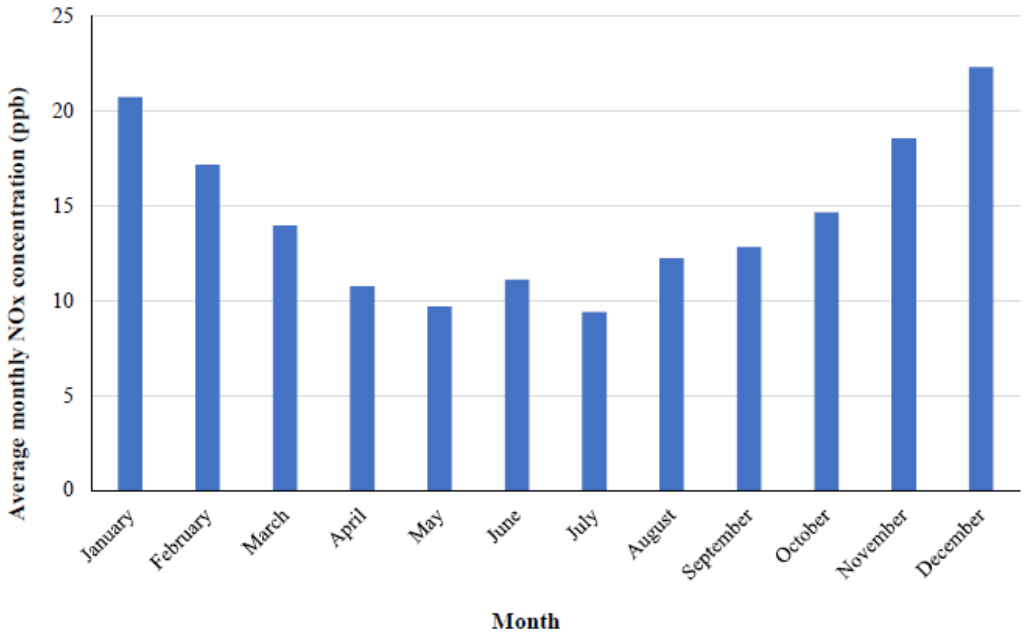

Figure 8 Monthly average NO_x - Toronto Downtown station.

Figure 9 shows the hourly average NO_x concentration during a day in 3 different months. As can be seen, NO_x concentration has a consistent pattern during the day for all three months. It reaches a maximum in the early morning as many people commute to work, so the number of vehicles is a key factor in the morning. Also, after sunset, the pollution starts to increase again. This happens because the sun heats the ground during the day, which results in more air movement during the day.

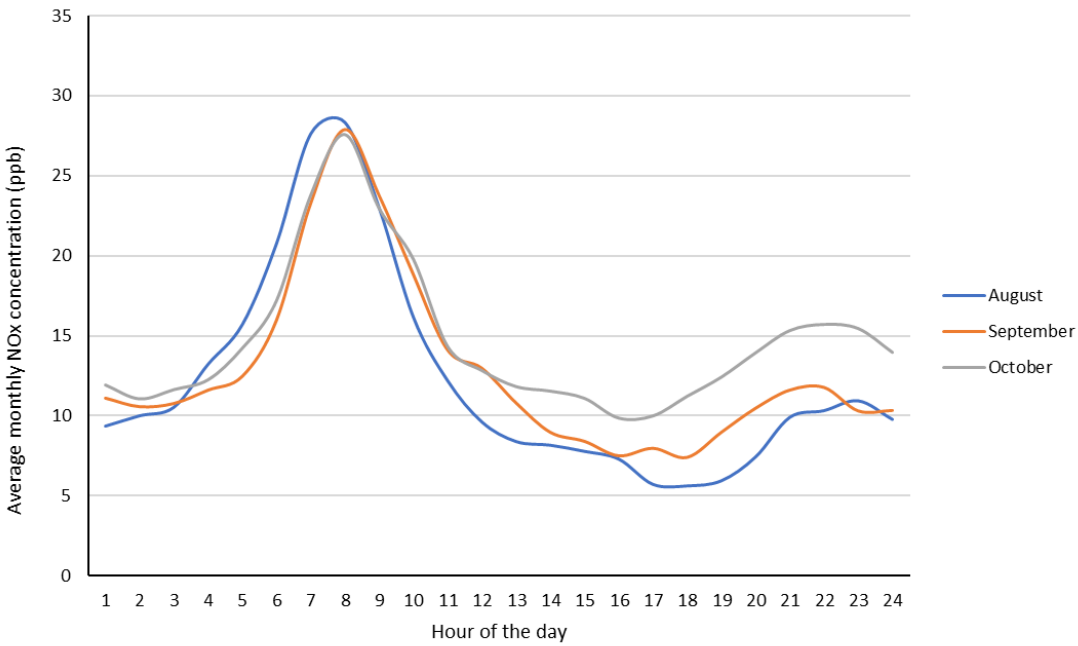

Figure 9 Monthly average NO_x concentration during a day in August, September, and October.

2.4 Air Pollution Estimation Model

To estimate the air pollution at four locations in the city of Toronto, an LSTM model is used as a predictive tool. LSTM models are a type of recurrent neural network for learning sequence prediction problems. Keras library is being used in this problem to build the deep learning model. Keras is an open-source library that acts as an interface for the TensorFlow library which is a machine learning and artificial intelligence library. Figure 10 shows the features used to learn the LSTM model. The features include weather, traffic count, time, and pollution from past timesteps. To prepare the data, it was found that there are missing values, especially among the 40 traffic count station data. To increase the number of data points, the traffic count stations with many missing values are disregarded. This decreases the number of stations used in the analysis from more than 40 stations to 8 stations. Also, to overcome the complexity of the model, the number of traffic count stations at the input to the model is reduced from 40 to 3. In other words, the 3 traffic count locations that have the largest impact on air pollution are kept, and the other 37 locations' data is disregarded. MinMaxScaler function is used to scale all the features into the range [0,1]. This will help to find the importance of each feature while disregarding its order of magnitude. The maximum epoch is considered high enough to make sure that the model is learned completely. Also, an early stopping function is added to the model to avoid overfitting. In this regard, if the amount of validation loss does not get better in 50 epochs, the learning process will stop, and the weights are extracted from the best validation loss epoch.

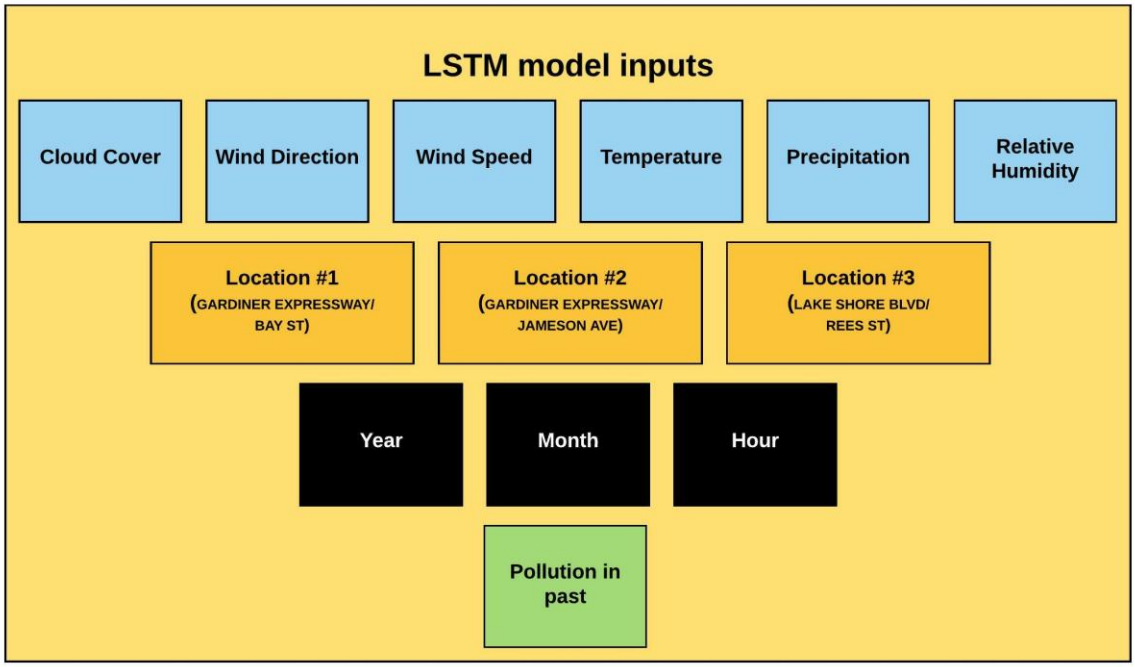

Figure 10 Input features for LSTM NO_x concentration prediction model.

A sensitivity analysis is conducted to find the best model parameters such as number of timesteps, number of layers, number of nodes in each layer, loss function, optimizer. The final structure of the neural network in TensorFlow is shown in Figure 11. Also, Table 1 shows the learning model parameters in TensorFlow for this work.

Table 1 LSTM parameters.

Parameter	Value
Loss function	Mean squared logarithmic error
Optimizer	adam
Batch Size	32
Early Stopping Patience	50
Number of hidden layers	2
Number of timesteps	2
Scale function	MinMaxScaler function
K-Fold cross-validation	K = 5

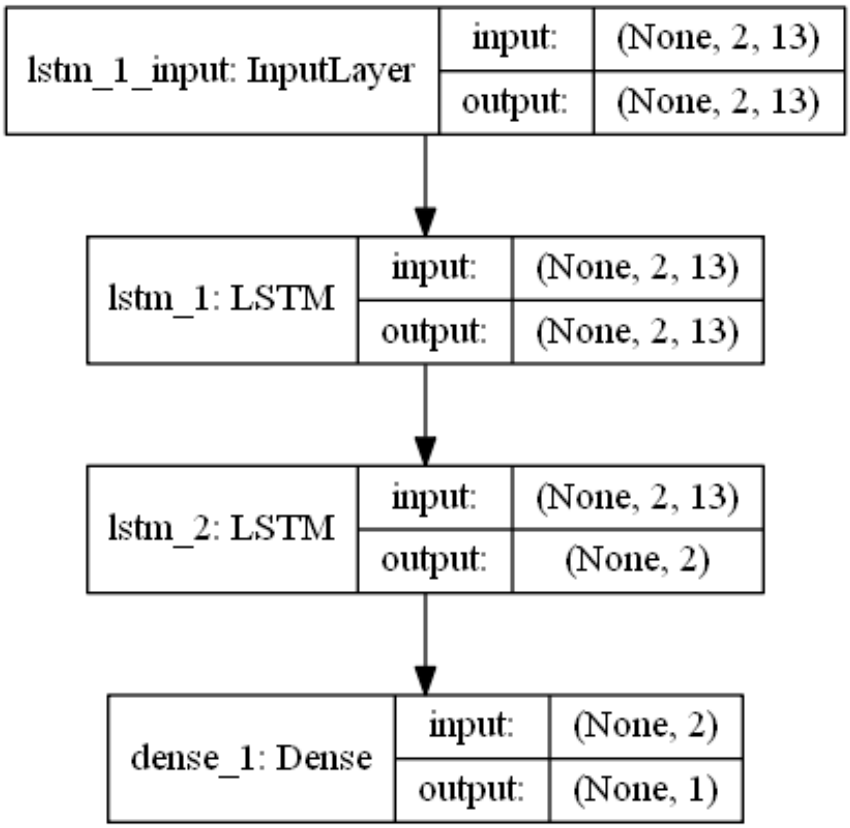

Figure 11 LSTM neural network structure.

5-fold cross-validation is utilized to find the effect of timestep numbers in predicting NO_x concentration. K-fold cross-validation is a method for finding the best evaluation of a machine learning model. In this method, the dataset is randomly partitioned into k subsamples. The model error, then, is calculated by utilizing the k-1 subsample as training data, and the last subsample as the validation data. The mean-squared error is calculated for different timesteps and different air monitoring stations. The results are shown in Table 2. As can be seen, except for one station, the mean-squared error is lower when assuming two timesteps to predict the NO_x concentration.

Table 2 Mean-squared error for different timesteps and different locations.

Number of timesteps	Toronto Downtown	Toronto East	Toronto West	Toronto North
1	5.5176	6.9964	9.364	6.1298
2	5.4924	6.812	9.3108	6.0052
4	5.5172	6.8026	9.3168	6.006

One of the challenges of forecasting air pollution using LSTM is long-term prediction is not possible. In other words, it is less challenging when the objective of the problem is to predict the pollution concentration in time t using data from time t-1, t-2, t-3, …, t-n; however, it is more difficult to forecast more steps in the future.

As the aim of this study is to calculate the impact of changing the share of EVs on NO_x emissions, it is assumed that all features except traffic count data stay unchanged between each scenario. Then, by changing the traffic counts based on different scenarios, the effect of EVs and ICEVs on pollutant concentration can be estimated. The challenging part is that NO_x concentration at time t heavily depends on concentration at times t-1, t-2, ..., t-n, where n is the number of timesteps in the LSTM model. In other words, NO_x concentration at previous timesteps not only affects the concentration at time t but also has an impact on concentration at future timesteps. To overcome this problem, the predicted values of NO_x concentration are used as an input to the model. In other words, assuming timestep = 2, the NO_x concentration at time t = 3 is predicted using the actual data at time t = 1 and t = 2; however, the NO_x concentration at time t = 4 is predicted using actual concentration at time = 2 and predicted concentration at t = 3. Also, pollution at t = 5 is predicted using predicted values of pollution at t = 3 and t = 4. It is worth mentioning that the actual values of other inputs such as weather and time data are used in all timesteps. Following this method, air pollution data is provided to the model just for the first two steps. As a result, a higher prediction error is seen due to error propagation. The modifications to the classical method in this LSTM model predict NO_x concentration in the long term, which makes the model capable of finding the effect of traffic count on long-term NO_x concentration.

2.5 Health Cost Calculation

Figure 12 shows the model used to calculate the health benefit from increasing EV market share. First, traffic count from different scenarios is imported to the learned LSTM model. The LSTM model, which is previously learned using current data, can predict hourly NO_x concentration in different scenarios. An annual average of NO_x concentration is then calculated in four air monitoring stations. To acquire the annual average NO_x concentration in all dissemination areas, three different interpolation methods are used: linear, cubic, and nearest method. Using Hazard Ratios (HRs), the risk of mortality in all scenarios is calculated. Having population, current average mortality rate, and decreased risk of mortality, the annual prevented deaths in all DAs are calculated. Finally, the annual deaths prevented are converted into monetary values using the Value of Statistical Life (VSL) in Canada.

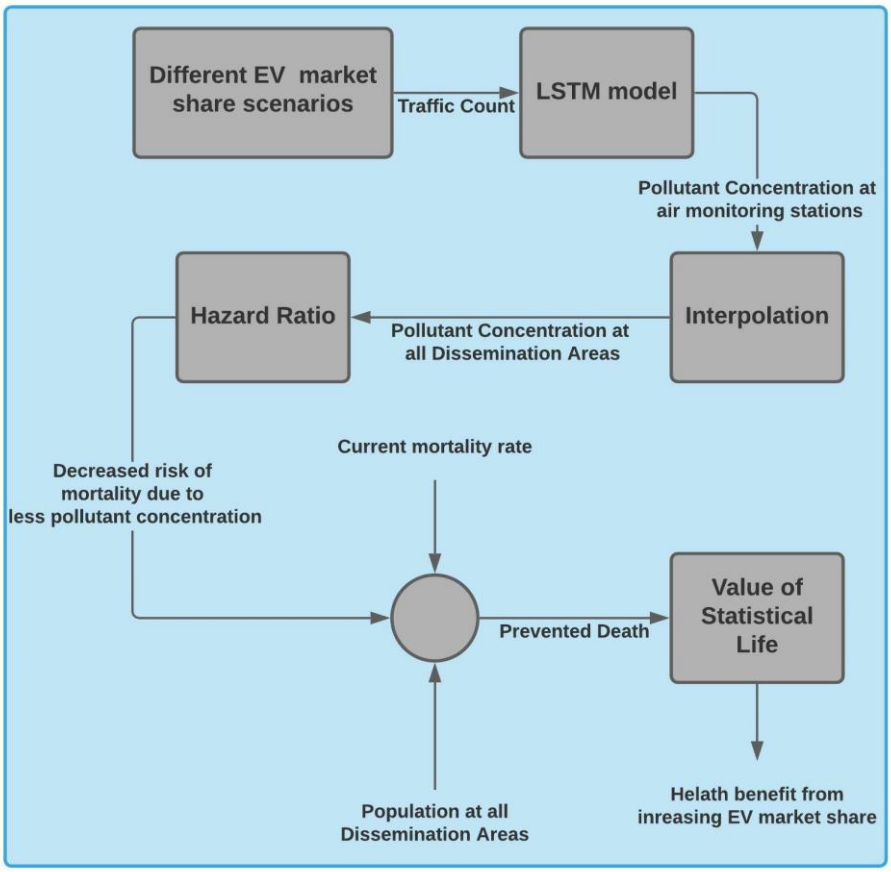

Figure 12 Calculation Process of health benefit due to increased market share of EV.

Table 3 Different scenarios are considered to investigate the performance of the model on different inputs. Also, to avoid major errors, a significant EV market share increase is not considered, because it might negatively affect the accuracy of air pollution prediction. Table 3 shows the different reductions in traffic count in different scenarios.

Table 3 Reduction in traffic count in different scenarios.

Scenario	Reduction in Traffic
1	2.5%
2	5%
3	7.5%
4	10%

3. Results

Figure 13 shows the comparison between the predicted and actual amount of NO_x concentration in Downtown Toronto station using different methods of implementing the LSTM model. In the original LSTM model, during the whole time series, the actual pollution is used to predict future pollution. However, in the alternative method of implementing the LSTM model, a prediction is made based on the previously predicted data. As can be seen in Figure 13, the error in the original

method is less than the error in the modified method. The mean-squared error in the original method and the modified method are 5.31 ppb and 8.82 ppb, respectively. The annual average of the predicted NO_x concentration in the Toronto Downton air pollution monitoring station is 15.83 ppb and 13.92 ppb using the original and modified method, respectively, where the actual annual average of NO_x concentration in this station is 16.36 ppb. The results show that although the error in the modified method is higher than the original method, the model can still make a reasonable prediction. The reason is that although pollution is predicted based on previously predicted pollutions, actual data are used for other features such as weather, time, and traffic count data. Also, as noted before, using the modified version provides the ability to make long-term predictions, which is the necessity of the current research.

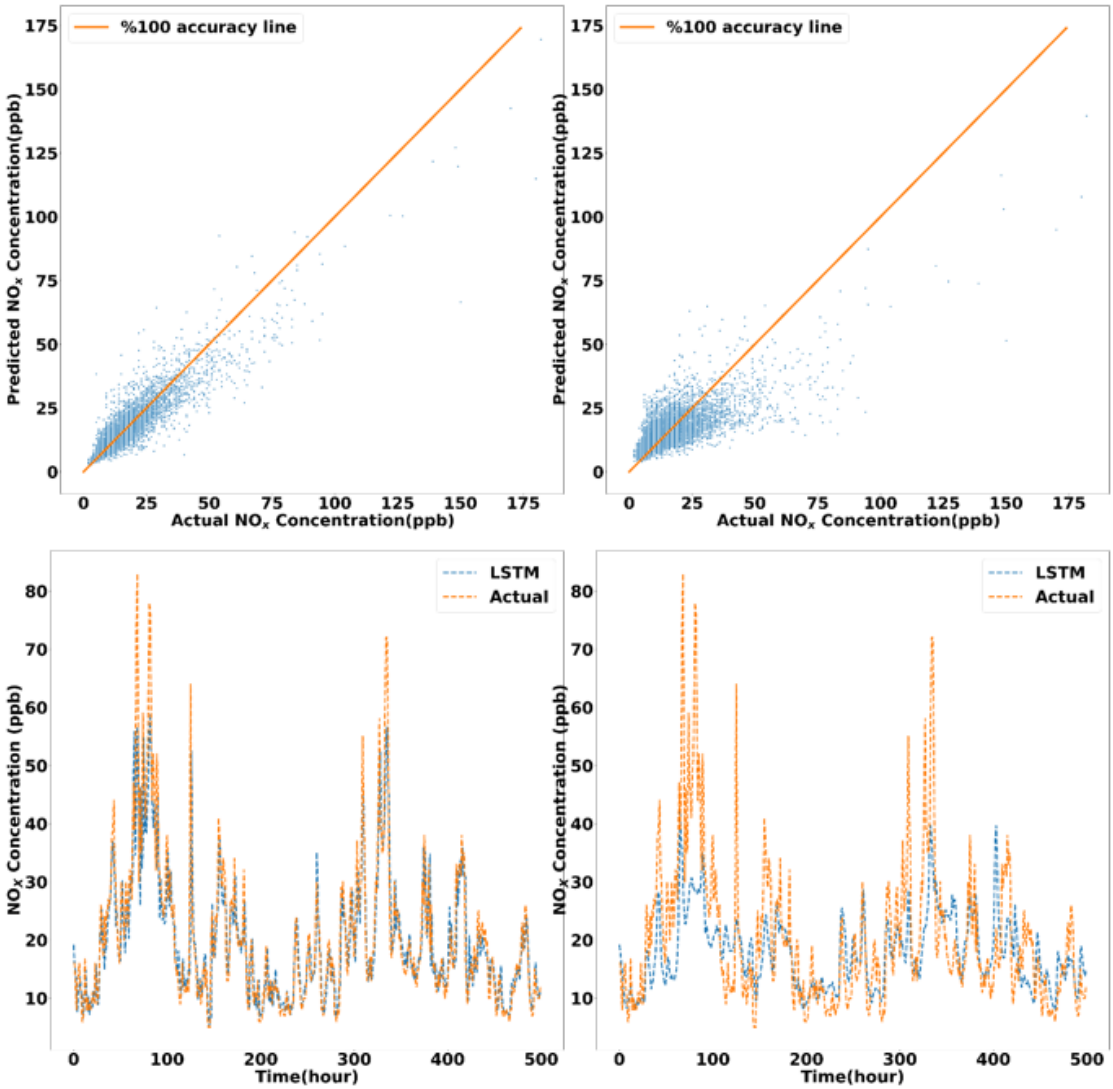

Figure 13 Toronto Downtown NO_x concentration prediction using LSTM and modified LSTM model. (Top left) Comparison between predicted and actual NO_x concentration using LSTM model. (Top right) Comparison between predicted and actual NO_x concentration using modified LSTM model. (Bottom left) First 500 hours of predicted and actual NO_x concentration using LSTM model. (Bottom right) First 500 hours of predicted and actual NO_x concentration using modified LSTM model.

Due to the stochastic nature of the algorithm or evaluation procedure, or differences in numerical precision, the output of the model may be different every time the learning procedure is conducted. In order to overcome this issue, the training process is applied five times for each location. Then, the mean-square error (MSE) and the annual average error are calculated as an average of the different learned models. The results are shown in Table 4.

Table 4 LSTM model error.

Parameter	Toronto Downtown	Toronto East	Toronto North	Toronto West
Actual NO_x concentration annual average (ppb)	17.3281	18.57291	16.25966	24.64313
Predicted NO_x concentration annual average (ppb)	16.35	18.03422	15.96259	24.6445
Annual average prediction error (%)	5.98%	2.99%	1.86%	0.01%
Mean-Squared Error	8.690333	13.72669	11.54073	18.61467

As can be seen in Table 4, the model is highly accurate for all locations. The MSE is lower at the Toronto Downtown station. The reason could be due to the proximity of the Toronto Downtown monitoring station and traffic count locations. Due to the lack of data, almost all traffic stations are in downtown Toronto. This could result in higher MSE in other locations. Nevertheless, the model shows a high accuracy in predicting annual average NO_x concentration in all stations, which is required for health cost calculations. The higher prediction accuracy in Toronto West station can be because of the higher influence of traffic count and the less influence of the residential heating system, as the area is less populated and close to Highway 401. In other words, the features included in this research have a higher influence on air pollution concentration in Toronto West station.

The models are next utilized to calculate NO_x concentration in different scenarios as mentioned in Table 3. Again, to overcome the stochastic nature of the model, calculations are made five times for each location and each scenario. The results of these analyses are shown in Table 5.

Table 5 Annual average NO_x reduction in different scenarios (ppb).

Scenario	Toronto Downtown	Toronto East	Toronto North	Toronto West
Scenario 1 (2.5% reduction in traffic count)	0.17	0.27	0.31	0.34
Scenario 2 (5% reduction in traffic count)	0.34	0.55	0.60	0.67
Scenario 3 (7.5% reduction in traffic count)	0.51	0.84	0.88	1.01
Scenario 4 (10% reduction in traffic count)	0.68	1.11	1.14	1.34

As can be seen in Table 5, the highest NO_x reduction can be achieved at the Toronto West station. There are two explanations for this possible reduction First, Toronto West station is in the vicinity of Highway 401, which is one of the busiest highways in North America, thus, the pollution at the Toronto West station is more dependent upon traffic volume. Secondly, the average NO_x concentration is higher in the western regions of Toronto, so the change in traffic volume can have a higher impact on NO_x reduction. The lowest reduction happens in the Toronto Downtown station that is due to low average concentrations and its location near Lake Ontario.

To find the annual average NO_x concentration reduction in all locations in Toronto, three different methods of interpolation are used: linear interpolation, cubic interpolation, and nearest interpolation. For the regions outside the quadrilaterals in between four air monitoring stations, the nearest method is used. The results of interpolation in different scenarios and different interpolation methods are shown in Figure 14.

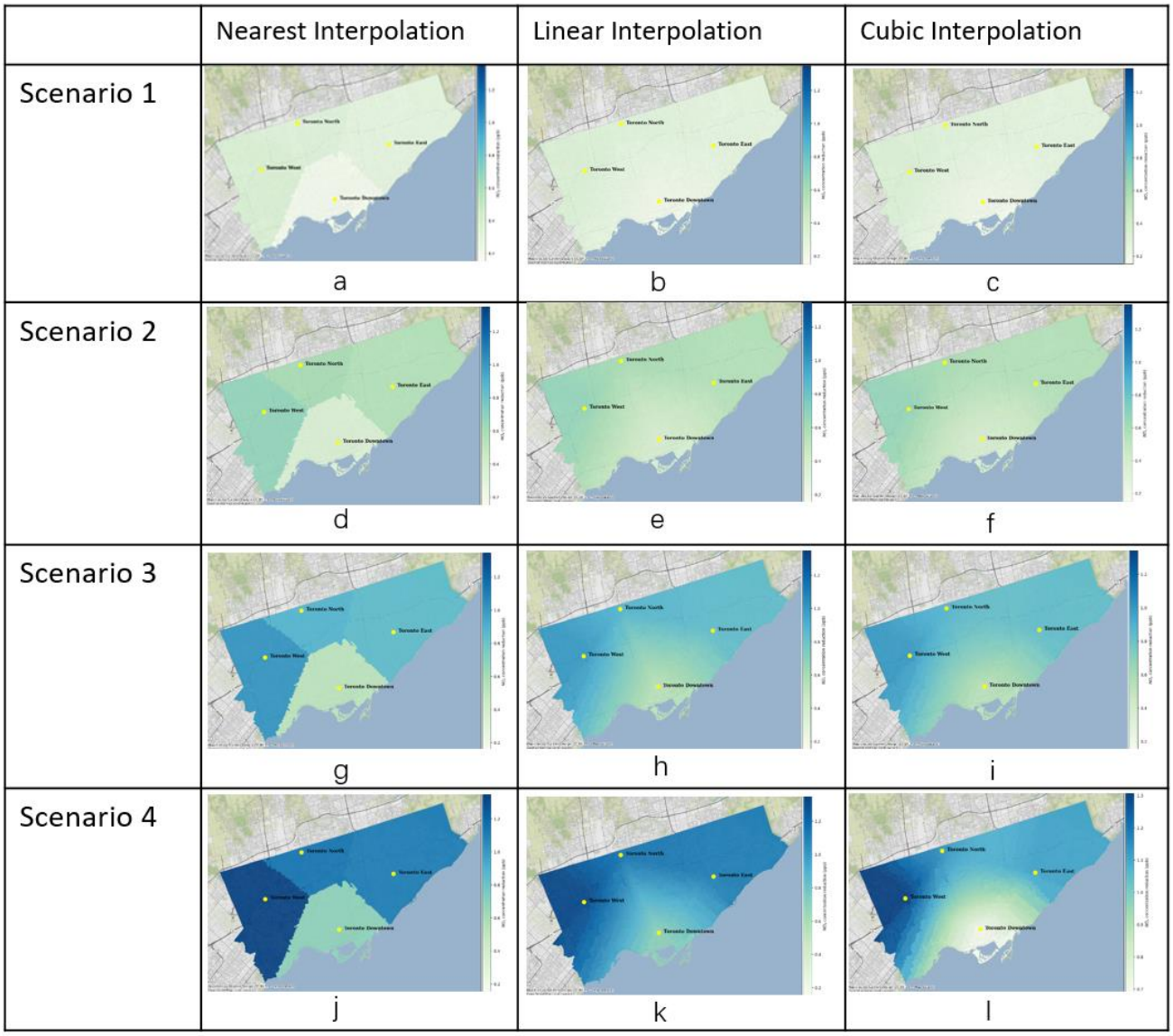

Figure 14 NO_x concentration reduction in various scenarios and interpolation methods.

As can be seen in Figure 14, linear interpolation and cubic interpolation show a very slight difference. As stated before, the highest NO_x reduction occurs in the western regions of Toronto, and the lowest reduction occurs in southern regions. Scenario 4, which is reducing fossil fuel vehicles by 10%, has the highest pollution reduction as expected. It is worth mentioning that interpolation causes possible inaccuracy in NO_x reduction estimation; however, with more air monitoring locations, the model will become more accurate. In other words, the higher number of pollution monitoring stations, the higher accuracy can be achieved. Although the small number of pollution monitoring stations causes estimation errors, especially near highways, it will still give a good estimation of pollution reduction.

Figure 15 shows the reduced mortality in Scenario 4 using cubic interpolation. As can be seen, the reduction in mortality depends on the population and NO_x concentration reduction in different regions. For instance, although the pollution reduction is significant in Pearson International airport because of the low population density in that area, the mortality reduction is not high.

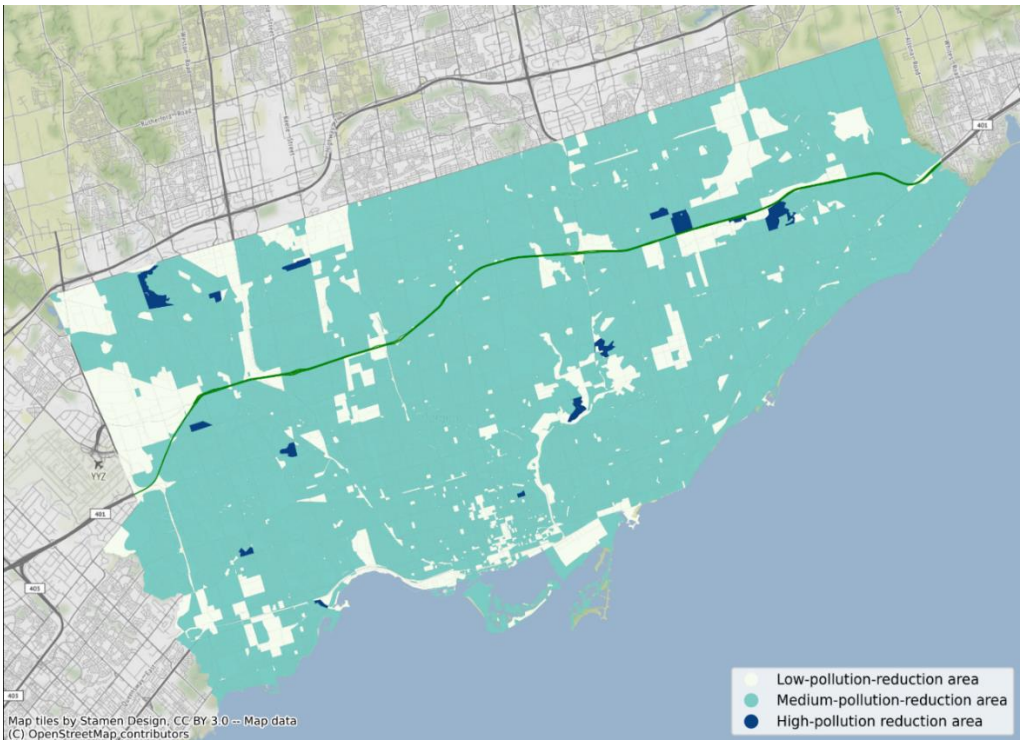

Figure 15 Mortality decrease by area.

In Figure 15, If the reduced number of premature death due to pollution is larger than 1, it is flagged as high-reduction, if it is between 0.1-1 it is considered medium reduction, and if it is less than 0.1, it is assumed low-reduction. In other words, if the reduced probability of mortality in one region multiplied by the population in that region is higher than 1, it is assumed a high pollution reduction; if it is between 0.1-1 it is assumed medium, and if it is less than 0.1, it is assumed low.

Calculating the mortality reduction in different scenarios using different interpolation methods, it was found that the interpolation method does not impact the total annual death prevented each year. The total prevented death in different scenarios is shown in Table 6. As can be seen, more than Canadian dollars (CAD) 500 million per year can be saved in terms of health benefits from reducing NO_x concentration in Scenario 4. The results show a linear reduction, which is mostly because of the linear output of the LSTM model.

Table 6 Prevented mortality and health benefit of different scenarios.

Scenario	**Annual Prevented Mortality**	**Health benefit (CAD million/year)**
Scenario 1 (2.5% reduction in traffic count)	18	144
Scenario 2 (5% reduction in traffic count)	35	280
Scenario 3 (7.5% reduction in traffic count)	53	424
Scenario 4 (10% reduction in traffic count)	70	560

4. Conclusion

Air pollution has a significant effect on the human health condition. Different pollutants such as $PM_{.25}$, NO_x, CO, O_3 cause lots of premature death and diseases. The aim of this study was to estimate the health cost of the fossil-fuel-based transportation system. Having the current transportation system's health cost, the benefit of replacing fossil-fuel vehicles with EVs can be investigated. To forecast air pollution, a machine learning method called long short-term memory (LSTM) is used to predict air pollution based on traffic data, weather conditions, and previous air pollution data. The developed model, then, is used to estimate the NO_x concentration in different scenarios of EV penetration in the market. Using hazard ratio, the impact of reducing the number of fossil-fuel vehicles is investigated. Finally, having the population concentration in different neighborhoods, the total number of prevented death in each scenario is evaluated. Finally, using Value of Statistical Life (VSL), the health benefit of different EV penetration scenarios is calculated. The results show that the highest NO_x reduction occurs in Toronto West. Also, the NO_x concentration reduces linearly as the share of fossil-fuel vehicles decreases. The reduction in mortality heavily depends on air pollution reduction as well as population concentration. Finally, the 10% reduction in fossil-fuel vehicles can save 70 people's lives just due to a reduction in NO_x. In terms of monetary, this can benefit the economy by saving CAD 560 million per year.

Acknowledgment

This work was supported by the Department of Chemical Engineering at the University of Waterloo, Canada Research Chair Tier I—Zero-Emission Vehicles and Hydrogen Energy Systems Grant number: 950-232215, and The Natural Sciences and Engineering Research Council of Canada (NSERC), Discovery Grants Program, RGPIN-2020-04149.

Contributions

Hamidreza Shamsi contributed to the model development and analysis, Ehsan Haghi contributed to the analysis, Manh-Kien Tran contributed to the methodology and analysis Sean Walker contributed to the literature review and preliminary analysis. Kaamran Raahemifar contributed to the analysis, Michael Fowler contributed to the methodology.

Funding

This article was funded by the Natural Sciences and Engineering Research Council, Grant RGPIN-2020-04149.

Competing Interests

The authors have declared that no competing interests exist.

References

1. IEA. Final consumption–Key World Energy Statistics 2020–Analysis [Internet]. Paris: IEA; 2020 [cited date 2021 November 29]. Available from: https://www.iea.org/reports/key-world-energy-statistics-2020/final-consumption.

2. Jakubiak-Lasocka J, Lasocki J, Siekmeier R, Chłopek Z. Impact of traffic-related air pollution on health. In: Environ exposure to pollutants. Advances in experimental medicine and biology. Cham: Springer; 2014. pp. 21-29.
3. Perera P, Hewage K, Sadiq R. Are we ready for alternative fuel transportation systems in Canada: A regional vignette. J Clean Prod. 2017; 166: 717-731.
4. Landrigan PJ, Fuller R, Acosta NJ, Adeyi O, Arnold R, Baldé AB, et al. The Lancet Commission on pollution and health. Lancet. 2018; 391: 462-512.
5. Brauer M, Reynolds C, Hystad P. Traffic-related air pollution and health in Canada. CMAJ. 2013; 185: 1557-1558.
6. Badyda AJ, Dabrowiecki P, Lubinski W, Czechowski PO, Majewski G. Exposure to traffic-related air pollutants as a risk of airway obstruction. In: Respiratory regulation-clinical advances. Dordrecht: Springer; 2013. pp. 35-45.
7. Forouzanfar MH, Afshin A, Alexander LT, Anderson HR, Bhutta ZA, Biryukov S, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016; 388: 1659-1724.
8. Shamsi H, Haghi E, Raahemifar K, Fowler M. Five-year technology selection optimization to achieve specific CO_2 emission reduction targets. Int J Hydrog Energy. 2019; 44: 3966-3984.
9. Tran MK, Sherman S, Samadani E, Vrolyk R, Wong D, Lowery M, et al. Environmental and economic benefits of a battery electric vehicle powertrain with a zinc–Air range extender in the transition to electric vehicles. Vehicles. 2020; 2: 398-412.
10. Tran MK, Fowler M. A review of lithium-ion battery fault diagnostic algorithms: Current progress and future challenges. Algorithms. 2020; 13: 62.
11. Tran MK, Akinsanya M, Panchal S, Fraser R, Fowler M. Design of a hybrid electric vehicle powertrain for performance optimization considering various powertrain components and configurations. Vehicles. 2020; 3: 20-32.
12. Kast J, Vijayagopal R, Gangloff Jr JJ, Marcinkoski J. Clean commercial transportation: Medium and heavy duty fuel cell electric trucks. Int J Hydrog Energy. 2017; 42: 4508-4517.
13. Shamsi H, Tran MK, Akbarpour S, Maroufmashat A, Fowler M. Macro-Level optimization of hydrogen infrastructure and supply chain for zero-emission vehicles on a canadian corridor. J Clean Prod. 2021; 289: 125163.
14. Kopackova H, Libalova P. Smart city concept as socio-technical system. Proceedings of the 2017 International Conference on Information and Digital Technologies (IDT); 2017 July 5th; Zilina, Slovakia. New York: IEEE. pp. 198-205.
15. von Schneidemesser E, Steinmar K, Weatherhead EC, Bonn B, Gerwig H, Quedenau J. Air pollution at human scales in an urban environment: Impact of local environment and vehicles on particle number concentrations. Sci Total Environ. 2019; 688: 691-700.
16. Ventura LM, de Oliveira Pinto F, Gioda A, Márcio de Almeida DA. Inspection and maintenance programs for in-service vehicles: An important air pollution control tool. Sustain Cities Soc. 2020; 53: 101956.
17. Ke W, Zhang S, He X, Wu Y, Hao J. Well-to-wheels energy consumption and emissions of electric vehicles: Mid-term implications from real-world features and air pollution control progress. Appl Energy. 2017; 188: 367-377.

18. Kebede L, Tulu GS, Lisinge RT. Diesel-fueled public transport vehicles and air pollution in Addis Ababa, Ethiopia: Effects of vehicle size, age and kilometers travelled. Atmos Environ X. 2022; 13: 100144.
19. Bougoudis I, Demertzis K, Iliadis L. HISYCOL a hybrid computational intelligence system for combined machine learning: The case of air pollution modeling in Athens. Neural Comput Appl. 2016; 27: 1191-1206.
20. Lautenschlager F, Becker M, Kobs K, Steininger M, Davidson P, Krause A, et al. OpenLUR: Off-the-shelf air pollution modeling with open features and machine learning. Atmospheric Environ. 2020; 233: 117535.
21. Sinnott RO, Guan Z. Prediction of air pollution through machine learning approaches on the cloud. Proceedings of the 2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT); 2018 Dec 17th; Zurich, Switzerland. New York: IEEE. pp. 51-60.
22. Lee HJ, Chatfield RB, Strawa AW. Enhancing the applicability of satellite remote sensing for $PM_{2.5}$ estimation using MODIS deep blue AOD and land use regression in California, United States. Environ Sci Technol. 2016; 50: 6546-6555.
23. Wang A, Xu J, Tu R, Saleh M, Hatzopoulou M. Potential of machine learning for prediction of traffic related air pollution. Transp Res D Transp Environ. 2020; 88: 102599.
24. Wen C, Liu S, Yao X, Peng L, Li X, Hu Y, et al. A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. Sci Total Environ. 2019; 654: 1091-1099.
25. Tong W, Li L, Zhou X, Hamilton A, Zhang K. Deep learning $PM_{2.5}$ concentrations with bidirectional LSTM RNN. Air Qual Atmos Health. 2019; 12: 411-423.
26. Huang CJ, Kuo PH. A deep CNN-LSTM model for particulate matter ($PM_{2.5}$) forecasting in smart cities. Sensors. 2018; 18: 2220.
27. Fan J, Li Q, Hou J, Feng X, Karimian H, Lin S. A spatiotemporal prediction framework for air pollution based on deep RNN. ISPRS Ann Photogramm Remote Sens Spat Inf Sci. 2017; 4: 15-22.
28. Qin D, Yu J, Zou G, Yong R, Zhao Q, Zhang B. A novel combined prediction scheme based on CNN and LSTM for urban $PM_{2.5}$ concentration. IEEE Access. 2019; 7: 20050-20059.
29. Šimić I, Lovrić M, Godec R, Kröll M, Bešlić I. Applying machine learning methods to better understand, model and estimate mass concentrations of traffic-related pollutants at a typical street canyon. Environ Pollut. 2020; 263: 114587.
30. Li Z, Yim SH, Ho KF. High temporal resolution prediction of street-level $PM_{2.5}$ and NO_x concentrations using machine learning approach. J Clean Prod. 2020; 268: 121975.
31. Lana I, Del Ser J, Padro A, Velez M, Casanova-Mateo C. The role of local urban traffic and meteorological conditions in air pollution: A data-based case study in Madrid, Spain. Atm Env. 2016; 145: 424-438
32. Kaminska J. A random forest partition model for predicting NO_2 concentrations from traffic flow and meteorological conditions. Sci of Tot Env. 2019; 651: 475-483.
33. Sayegh A, Tate J, Ropkins K. Understanding how roadside concentrations of NOx are influenced by the background levels, traffic density, and meteorological conditions using Boosted Regression Trees. Atm Env. 2016; 127: 163-175.
34. Boningari T, Smirniotis PG. Impact of nitrogen oxides on the environment and human health: Mn-based materials for the NO_x abatement. Curr Opin Chem Eng. 2016; 13: 133-141.

35. Jonson JE, Borken-Kleefeld J, Simpson D, Nyíri A, Posch M, Heyes C. Impact of excess NO_x emissions from diesel cars on air quality, public health and eutrophication in Europe. Enviro Res Lett. 2017; 12: 094017.
36. Fu R, Zhang Z, Li L. Using LSTM and GRU neural network methods for traffic flow prediction. Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC); 2016 Nov 11th; Wuhan, China. New York: IEEE. pp. 324-328.
37. Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy. 2018; 148: 461-468.
38. Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In: Advances in neural information processing systems. La Jolla: Neural Information Processing Systems Foundation, Inc.; 2015. pp. 802-810.
39. Catalano M, Galatioto F. Enhanced transport-related air pollution prediction through a novel metamodel approach. Transp Res D Transp Environ. 2017; 55: 262-276.
40. Orun A, Elizondo D, Goodyer E, Paluszczyszyn D. Use of Bayesian inference method to model vehicular air pollution in local urban areas. Transp Res D Transp Environ. 2018; 63: 236-243.
41. Kniesner TJ, Viscusi WK. The value of a statistical life. Van Law Res Pap. 2019. No. 19-15. doi: 10.2139/ssrn.3379967.
42. Air Quality Ontario. Homepage [Internet]. Toronto: Air Quality Ontario; 2022 [cited date 2021 November 29]. Available from: http://www.airqualityontario.com/.
43. CityofToronto/vz_challenge. vz_challenge/transportation/traffic_volumes [Internet]. San Francisco: GitHub, Inc.; 2018 [cited 2021 December 19]. Available from: https://github.com/CityofToronto/vz_challenge/tree/master/transportation/traffic_volumes.
44. Weather Dashboard for Toronto. Homepage [Internet]. Toronto: Weather Dashboard for Toronto [cited date 2021 December 15]. Available from: https://toronto.weatherstats.ca/.