Estimation of surface nitrogen dioxide (NO_2) using TEMPO satellite data and machine learning

Neda Kolahi, Costas Armenakis, Mark D Gordon

Department of Earth and Space Science and Engineering, Lassonde School of Engineering
York University, Toronto, Canada - {nedaklh}{armenc}{mgordon}@yorku.ca

Keywords: TEMPO Satellite, Remote Sensing, Random Forest, Surface Nitrogen Dioxide, Air Pollution.

Abstract

Air pollutants such as nitrogen dioxide (NO_2) have detrimental effects on human health and ecosystems. It is therefore very crucial to pinpoint the location of high pollutant concentrations over large areas. Ground-based stations, while offering continuous temporal measurements, cannot provide broader spatial coverage for regions like cities. This study uses Tropospheric Emissions: Monitoring Pollution (TEMPO) satellite observations and a machine learning model to estimate high-resolution surface-level NO_2 concentrations over the Greater Toronto Area (GTA), Ontario, Canada. The random forest regression model was trained with input parameters such as hourly tropospheric NO_2 vertical column density (VCD) values and boundary layer height (BLH), which are the two most effective parameters in feature importance. The model achieved a coefficient of determination (R^2) of 0.84, a root mean square error (RMSE) of 1.703 ppb, and a mean absolute error (MAE) of 0.939 ppb, indicating strong and reliable predictive performance. The findings of this research can support air quality forecasting, public health studies, and urban planning decisions, especially in regions with scarce ground-based pollutant data.

1. Introduction

One of the most important problems that people all around the world are struggling with is air pollution. It is a misconception that air pollution is only a concern for developing countries; in fact, according to the State of Global Air Report 2024, many high-income countries, such as Canada, also experience high levels of pollutants like nitrogen dioxide, sometimes even higher than those in developing nations. Research has consistently shown that long-term exposure to NO_2 is linked to higher risks of death from all causes, as well as from cardiovascular and respiratory diseases. The World Health Organization's (WHO), new air quality standards dramatically tightened the annual mean NO_2 limit from 40 µg/m³ to 10 µg/m³, emphasizing the evidence of its adverse health impacts. In addition, the WHO proposed a 24-hour mean standard of 25 µg/m³ to address not just long-term but short-term NO_2 exposures (World Health Organization, 2021). A positive association between population density, street connectivity, and mortality attributable to NO_2 have been found, indicating that densely built urban environments exacerbate exposure-related health risks (Song et al., 2023).

NO_2 has a huge impact on the environment in addition to its detrimental effects on human health. It contributes significantly to the creation of smog and other harmful pollutants, including particulate matter and ozone, which have an indirect impact on climate change. Although NO_2 is not classified as a direct greenhouse gas, it plays a crucial role as a precursor in the formation of tropospheric ozone. It also contributes to acid rain, reduces visibility by creating haze, and pollutes aquatic ecosystems with nutrients (U.S. Environmental Protection Agency, 2023)

As a result, monitoring accurate NO_2 concentrations is not only crucial for human exposure and air-quality management, but also to help us better understand how NO_2 affects climate indirectly and to support more effective climate mitigation efforts.

Air pollution is typically measured using ground-based monitoring stations, which do not provide sufficient and continuous spatial coverage — a limitation that can be addressed through satellite remote sensing, which has been used in various fields, such as epidemiological research (Wimberly et al., 2021) and environmental justice (Sayyed et al., 2024).

The tropospheric NO_2 vertical column density (VCD) is a parameter that represents the total number of NO_2 molecules per unit area in the troposphere. Hourly values of tropospheric NO_2 are provided by several satellite instruments, such as the Ozone Monitoring Instrument (OMI) onboard NASA's Aura satellite (Levelt et al., 2006), the Tropospheric Monitoring Instrument (TROPOMI) onboard ESA's Sentinel-5P (Reshi et al., 2024), and the Geostationary Environment Monitoring Spectrometer (GEMS) onboard Korea's GEO-KOMPSAT-2B (J. Kim et al., 2020). In this study, we focus on the recently launched Tropospheric Emissions: Monitoring of Pollution (TEMPO) instrument (launched on April 7, 2023), which is a grating spectrometer sensitive to visible and ultraviolet wavelengths developed by NASA that observes atmospheric pollutants over North America during daylight hours. The TEMPO instrument is attached to the Earth-facing side of the geostationary telecommunications satellite Intelsat-40e at an altitude of approximately 35,786 km. Table 1 presents the satellite instruments frequently used in this area, along with their launch years, coverage regions, and spatial and temporal resolutions.

TEMPO data are important for Canada because its vast land area makes it difficult to rely only on ground-based monitoring. One of the main uses of TEMPO data is its assimilation into Environment and Climate Change Canada's (ECCC) air quality forecasting system, helping to improve forecasts across a region that mostly overlaps with TEMPO's observation area. Other important applications include environmental assessments, epidemiological studies, health impact research, and monitoring natural disasters such as forest fires.

To make the most of TEMPO observations, Canada is interested in improving data quality at higher latitudes, where several challenges exist: larger average solar and viewing angles reduce sensitivity of some gases (such as O_3 and NO_2) in the boundary

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-3-2026-157-2026

layer; stratospheric amounts of these gases are higher and more variable, making them harder to correct; TEMPO's pixel sizes increase toward higher latitudes; and retrievals during forest fires are more difficult due to heavy aerosol loading. Snow-covered surfaces also introduce greater uncertainties.

Instrument	Launch Year	Coverage Region	Spatial Resolution (km)	Temporal Resolution
OMI	2004	Global	13 × 24	Daily
TROPOMI	2017	Global	7 × 3.5	Daily
GEMS	2020	East and Southeast Asia	7 × 8	Hourly
TEMPO	2023	North America	2 × 4.75	Hourly

Table 1. Summary of satellite instruments used for tropospheric NO_2 observations, including launch year, spatial and temporal resolutions, and coverage region.

In this section, we focus on previous research that has utilized satellite datasets in combination with machine learning techniques to estimate surface-level NO_2 concentrations.

Griffin et al. showed that using a Random Forest model to estimate surface NO_2 from TROPOMI satellite data significantly outperformed traditional scaling methods based on chemical transport models, achieving higher correlation with ground-based measurements. Meteorological parameters, land cover type, topography, and emission inventories were used as input features. They also concluded that random forest outperformed deep learning models by a significant margin for estimating surface-level NO_2 concentrations over North America (Griffin et al., 2025). Chan et al. presented the estimation of surface NO_2 concentrations over Germany using a neural network model trained with TROPOMI tropospheric NO_2 VCDs and meteorological parameters. The model was validated against ground-based measurements and regional chemical transport model (CTM) simulations, showing good agreement with a Pearson correlation of 0.80 and outperforming CTMs (Chan et al., 2021). Kim et al. used simple linear regression (SLR) and multivariate linear regression (MLR) models to evaluate the relationship between surface monitor NO_2 and TROPOMI column NO_2. The best model achieved a cross-validated R^2 of 0.77 overall, with better performance in suburban and low-traffic areas (Kim et al., 2024). GEMS NO2 VCDs from its level 2 product were employed in different regions of China. This study demonstrated that explicitly incorporating the NO_2 mixing height (NMH) was critical for converting satellite-derived NO_2 vertical column densities (VCDs) into accurate ground-level concentrations. Using a nested XGBoost framework with GEMS measurements, NMH was first predicted from meteorological variables and then used as an input to the main model, increasing the cross-validated R^2 from 0.73 to 0.93 (Ahmad et al., 2024). Another satellite sensor used in this field was the Ozone Monitoring Instrument (OMI), which was integrated with a Geographically Weighted Panel Regression (GWPR) to estimate monthly global ground-level NO_2 concentrations, achieving a Pearson correlation of R = 0.838 and RMSE=7.84 $\mu g/m^3$ (Li & Managi, 2022). All these studies demonstrate the effectiveness of combining satellite data and machine learning models to estimate ground-level NO_2.

2. Materials and Methods

2.1 Study area and Time period

This study focuses on the Greater Toronto Area (GTA), Ontario, Canada, a metropolitan region in the province of Ontario, Canada. All data used in this study were collected during the period from August 2023 to December 2024 in the GTA.

2.2 Dataset

2.2.1 Satellite Tempo data: In this study, hourly tropospheric NO_2 column data are obtained from the TEMPO (Tropospheric Emissions: Monitoring of Pollution) satellite. These Level 3 data have been available since August 1, 2023, and can be accessed through *www.earthdata.nasa.gov*. The data are transformed from NETCDF to CSV format, facilitating visualization and analysis. Additionally, UTC timestamps are converted to local time (EDT) to ensure alignment with ground-based monitoring station data.

A set of filters is implemented to enhance data quality and accuracy. Cloud is one of the most significant variables influencing satellite observation accuracy. The interplay of sunlight with clouds imposes major challenges for satellite remote sensing, in terms of both the spatial complexity of real clouds and the dominance of multiple scattering in radiation transport (Loyola et al., 2018). To minimize this effect, we filtered the TEMPO Level 3 data by applying a cloud fraction threshold of 0.4. This value aligns with NASA's TEMPO Level 3 quality guidelines, which identify 0.4 as the point where cloud-induced scattering begins to stabilize or reduce its influence on retrieval accuracy. In the Greater Toronto Area, where variable cloud cover and lake-effect weather are frequent, this threshold was critical for maintaining a continuous record of diurnal NO2 variations. A more restrictive threshold (e.g., < 0.2) would have led to substantial data gaps. Also, we use the main data quality flag, which indicates the reliability of each data point: 0 (good), 1 (suspect), and 2 (bad). In this study, only data with a flag value of 0 were included. Further filtering criteria included: snow_ice_fraction ≤ 0.5 and solar_zenith_angle ≤ 70°. Only satellite observations that met all these conditions were used in the final analysis.

2.2.2 Surface NO2 Measurements: Hourly NO_2 measurements were obtained from *www.airqualityontario.com*, which provides ground-based air quality data across Ontario. In this study, we focus on seven air quality monitoring stations located within the Greater Toronto Area that measure nitrogen dioxide. These stations are Toronto Downtown, Toronto East, Toronto North, Toronto West, Mississauga, Brampton, and Oakville.

2.2.3 Boundary Layer Height: Another important variable in this study is Boundary Layer Height (BLH). Lower BLH values indicate a smaller vertical space for pollutants to disperse, which leads to higher concentrations near the surface. As a result, adding BLH as a variable significantly improves the accuracy of surface NO_2 concentration predictions (Grzybowski et al., 2023). This research uses BLH data from the climate and weather reanalysis data, which provides hourly estimates at a 0.25° spatial resolution via the Copernicus Climate Data Store. To match the spatial resolution of the Level 3 TEMPO satellite data (0.02° × 0.02°), the BLH data are resampled via linear interpolation.

2.2.4 Meteorological Factors: The zonal wind (U) and the meridional wind (V), which respectively represent the components of wind blowing from west to east and from south to north, along with air temperature (K) at 2 m above the surface, surface pressure (Pa), and precipitation (kg/m²), were extracted from the North American Land Data Assimilation System (NLDAS) dataset, which provides meteorological variables at a spatial resolution of 0.125° × 0.125°.

2.3 Random Forest Model

To estimate surface-level NO_2, we utilize the random forest regression algorithm in Python scikit-learn. Random forest is an ensemble learning method for classification and regression that operates by constructing a multitude of decision trees. The forest averages (for regression) or votes (for classification) among the predictions made by each tree. To minimize overfitting and improve generalization performance, each tree in the forest is trained using a random subset of the training data and a random subset of predictor variables. The performance of Random Forest enhances with an increase in the number of trees (n_estimators in the Python code); however, the incremental improvement diminishes, indicating that a sufficient number of trees is necessary for consistent results, rather than an excessive quantity.

In this study, the input parameters (X) represent the feature data used as predictors in the model. The input parameters (X) include TEMPO vertical column NO_2, snow/ice fraction, meteorological factors such as temperature, relative humidity, wind speed, and boundary layer height, temporal features including day of the week, sine and cosine of the hour, as well as the distance of each ground station from the nearest TEMPO measurement. The target variable (Y) represents the ground-level NO_2 concentrations measured at monitoring stations. During training, X was used to predict Y.

In this research, we use hyperparameter tuning to optimize our results. Using this method, we construct a model for every possible combination of all the given hyperparameter values, assess each model, and choose the design that yields the best outcomes. Hyperparameters were tested and selected for the training of the machine learning algorithm.

We randomly split the dataset using a random state of 42, allocating 80% of the data for training and the remaining 20% for testing. The performances of the models are evaluated by using commonly used statistical indices (Eqs. 1–3), which are the coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_p})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (1)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y_p}\right)^2} \quad (2)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y_p}\right| \quad (3)$$

where y_i and $\hat{y}_p$ are the observed and predicted values, respectively, and $\bar{y}$ is the mean of the observed values. A higher R^2 value and lower RMSE and MAE values indicate better model performance.

3. Results and discussion

The Random Forest model was optimized through hyperparameter tuning using 3-fold cross-validation. As shown in Figure 2, the Random Forest model achieved a high coefficient of determination (R^2 = 0.84), indicating a strong agreement between the observed and predicted surface NO_2 concentrations. In addition, the MAE and RMSE values are 0.94 ppb and 1.70 ppb, respectively, further confirming the model's strong predictive performance and consistency in estimating surface NO_2 concentrations. The Random Forest model was optimized through hyperparameter tuning. The best-performing configuration was obtained with a maximum tree depth (max_depth) of 25 and 400 estimators (n_estimators). These hyperparameters were selected as they maximized the model's predictive performance on the validation data while minimizing overfitting. Table 2 shows the hyperparameters used in the model, including their description, the range of values tested, and the selected values.

Hyperparameter	Description	Range of Values tested	Selected Values
max_depth	Maximum depth of each decision tree	10-60	25
min_samples_leaf	Minimum number of samples required at a leaf node	1-10	1
min_samples_split	Minimum number of samples required to split an internal node	1-10	2
n_estimators	Number of trees in the ensemble	20-500	400

Table 2. Hyperparameters tested and selected for the training of the RF model.

To better interpret the model and identify which variables have the strongest influence on its predictions, we assess the feature importance values provided by the Random Forest algorithm. These scores indicate the relative contribution of each input parameter to the overall predictive accuracy of the model. Table 3 shows the importance of each feature in predicting the target variable.

The feature importance analysis (Fig. 1) reveals that the TEMPO tropospheric NO_2 column is the most significant variable for forecasting surface NO_2 concentrations. This result is consistent with the strong physical relationship between columnar and near-surface NO_2 levels. The second most significant predictor is the boundary layer height (BLH), which is logically anticipated as BLH directly influences pollutant dispersion and vertical mixing. For comparison, we ran the model in a second configuration with an added feature (VCD×BLH) in the model to see if the Random Forest could independently capture the non-linearity between features. Theoretically, this term should aid the model in comprehending the behaviour of vertical column density (VCD)

under varying boundary layer heights (BLH). In the final ranking, this new feature was only more important than rainfall. This result demonstrates that the hierarchical tree structure of the model was already effectively "learning" these non-linear physical dependencies from the unprocessed inputs. As a result, the manual interaction term was unnecessary.

For the final model, the input parameters listed in Table 3 were used. Features such as ice fraction, sine and cosine of the hour, cloud fraction, and solar zenith angle had very low feature importance values (less than 0.005) and were consequently removed to avoid overfitting and reduce the computational burden.

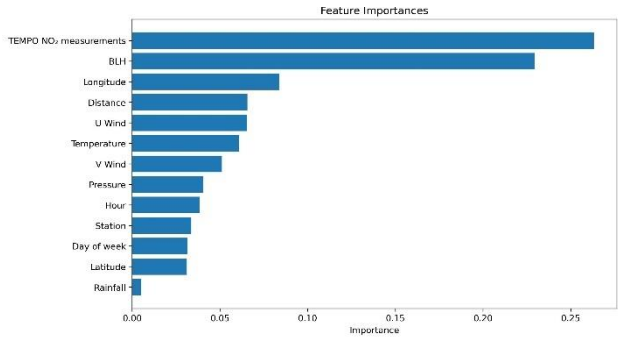

Figure 1. Feature Importance

Feature	Importance
TEMPO NO_2	0.263
BLH	0.229
Longitude	0.084
Distance	0.066
U Wind	0.065
Temperature	0.061
V Wind	0.051
Pressure	0.040
Hour	0.038
Station	0.034
Day of week	0.031
Latitude	0.031
Rainfall	0.005

Table 3. Importance of the input parameters for the random forest predictor.

We compare the predicted NO_2 values with the measured NO_2 values at the 7 ground stations (Fig. 3). Figure 4 shows the MAE and RMSE values for each air quality monitoring station. The Downtown station has the highest errors, which could be due to the heavy traffic, dense urban activity, fast-changing conditions, and possible poor estimation of BLH in the downtown. On the other hand, the Toronto East station shows the lowest error values, meaning the model performs best there. This is likely because the area experiences less variability and fewer complicated emission sources, making the surface NO_2 levels easier for the model to capture.

Figure 5 shows the Mean Error (ME) values for each air quality monitoring station. The mean error values show whether the model tends to overestimate or underestimate NO_2 at each station. Positive values mean the model slightly overpredicts, and negative values mean it underpredicts. Our model slightly overpredicts NO_2 in Brampton, Toronto North, Oakville, and Toronto Downtown, with the strongest overprediction seen in Oakville. On the other hand, the model underpredicts in Mississauga, Toronto West, and Toronto East, with Mississauga showing the largest negative bias.

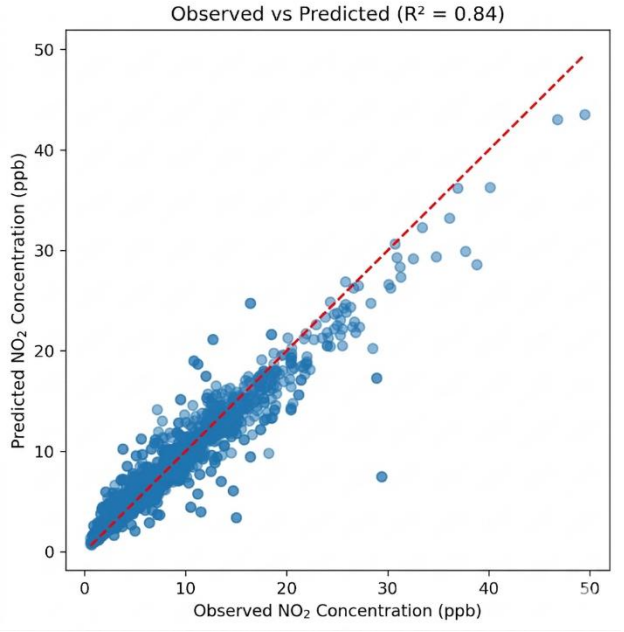

Figure 2. Scatterplot showing the performance of the RF model. The coefficient of determination, $R^2 = 0.84$

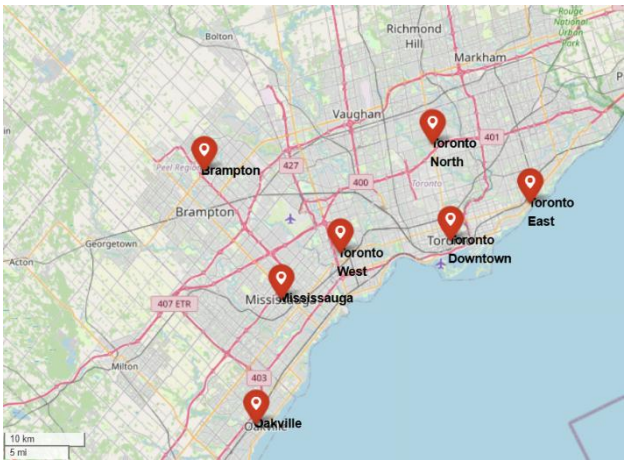

Figure 3. Locations of air quality monitoring ground stations

Additionally, to compare the model’s ability to estimate NO_2 concentrations across different seasons, we assessed its performance separately for each season using the test dataset. As shown in Table 4, the highest RMSE is observed in winter, likely due to more variable atmospheric conditions and the smaller number of observations during this season. This could also be related to a more stable atmosphere in winter and a generally lower BLH. In contrast, the model performs better in spring, summer, and fall, where RMSE values are consistently lower.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-3-2026-157-2026 | © Author(s) 2026. CC BY 4.0 License.

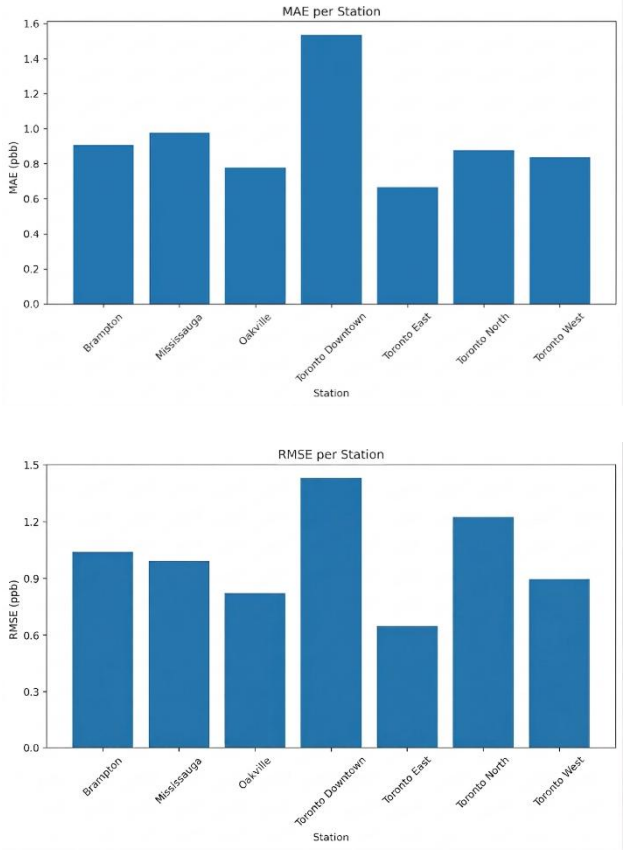

Figure 4. Comparison of MAE (top) and RMSE (bottom) values for each air quality monitoring station

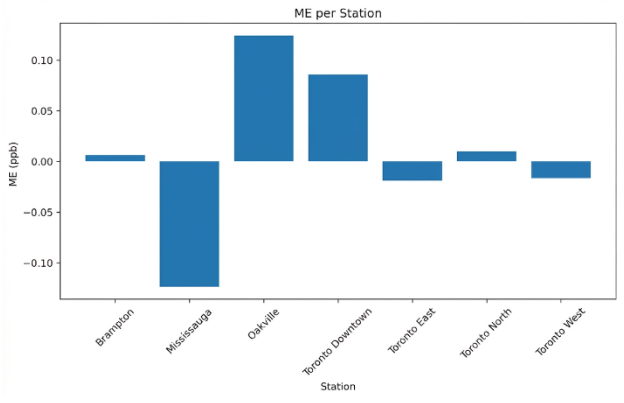

Figure 5. Comparison of ME values for each air quality monitoring station

Season	R^2	RMSE (ppb)
Winter	0.79	3.32
Spring	0.91	1.49
Summer	0.88	1.59
Fall	0.93	1.56

Table 4. Seasonal model performance.

To address the limitations of random data splitting and obtain a more rigorous assessment of the model’s spatial portability, we implemented a Leave-One-Station-Out (LOSO) validation framework (Pandya et al., 2026). In this approach, the model is trained on data from all stations except one, which is withheld entirely to serve as an unseen geographic spot for testing. This process is repeated for every station in the study area. Figure 6 shows that the model’s ability to estimate NO2 at unseen locations is not as high as previously suggested by standard cross-validation. The results of the Leave-One-Station-Out (LOSO) show notable variation in model performance across different locations. The model performs well at stations such as North, Oakville, and Mississauga, with relatively low RMSE values (3.27–3.43 ppb) and MAE values below 2.5 ppb. Moderate performance is observed at the West station, while significantly higher errors are found at Brampton, East, and Downtown, where RMSE values exceed 6 ppb. In particular, the East and Downtown stations show the largest prediction errors, indicating that the model struggles to generalize in these areas.

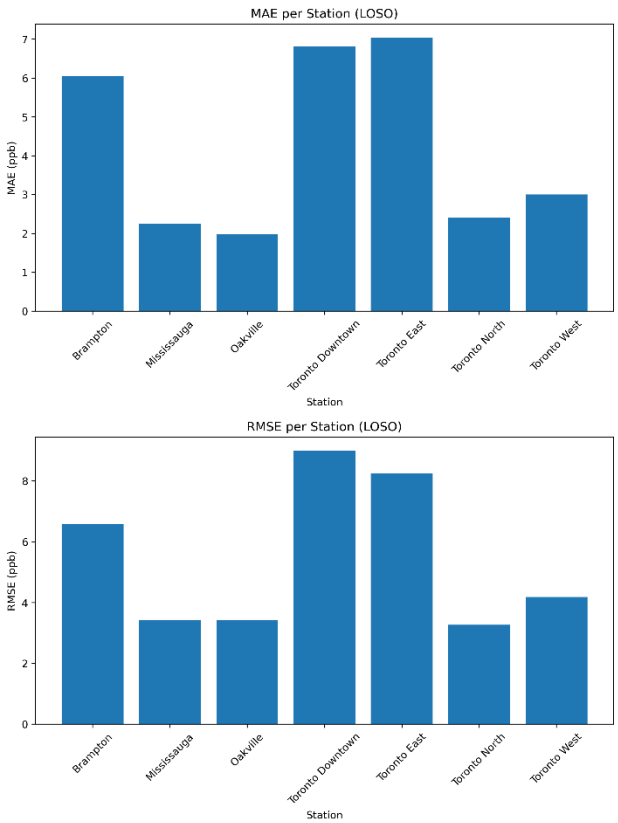

Figure 6. MAE and RMSE at each monitoring station obtained using Leave-One-Station-Out (LOSO) cross-validation, where each station is excluded from training and used as an independent test site

4. Conclusion

In conclusion, our study demonstrates that TEMPO satellite measurements are an effective tool for estimating surface-level NO_2. By using a random forest regression model with key inputs such as tropospheric NO_2 columns and boundary layer height, we achieved an accurate estimation of surface-level NO_2 across the Greater Toronto Area, with an R^2 of 0.84. In our future studies, we plan to investigate the local conditions at each air quality ground station and explore the use of other satellite data, such as

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-3-2026-157-2026 | © Author(s) 2026. CC BY 4.0 License.

those from TROPOMI onboard the Sentinel-5, and investigate the performance of other AI deep learning approaches.

5. Acknowledgement

This work has been financially supported by York University's CIRC research project, Geomatics for analyzing climate change effects on ecosystems and human populations.

References

Ahmad, N., Lin, C., Lau, A.K.H., Kim, J., Zhang, T., Yu, F., Li, C., Li, Y., Fung, J.C.H., & Lao, X.Q., 2024: Estimation of ground-level NO_2 and its spatiotemporal variations in China using GEMS measurements and a nested machine learning model. *Atmospheric Chemistry and Physics*, 24(16), 9645–9665. https://doi.org/10.5194/acp-24-9645-2024

Chan, K.L., Khorsandi, E., Liu, S., Baier, F., & Valks, P., 2021: Estimation of surface NO_2 concentrations over Germany from TROPOMI satellite observations using a machine learning method. *Remote Sensing*, 13(5), 969. https://doi.org/10.3390/rs13050969

Griffin, D., Hempel, C., McLinden, C., Kharol, S.K., Lee, C., Fogal, A., Sioris, C., Shephard, M., & You, Y., 2025: Development and validation of satellite-derived surface NO_2 estimates using machine learning versus traditional approaches in North America. https://doi.org/10.5194/egusphere-2025-1681

Grzybowski, P., Markowicz, K., & Musiał, J., 2023: Estimations of ground-level NO_2 concentrations based on the Sentinel-5P NO_2 tropospheric column number density product. Remote Sensing, 15(2), 378. https://doi.org/10.3390/rs15020378

Kim, E.J., Holloway, T., Kokandakar, A., Harkey, M., Elkins, S., Goldberg, D.L., & Heck, C., 2024: A comparison of regression methods for inferring near-surface NO_2 with satellite data. *Journal of Geophysical Research: Atmospheres*, 129(16), e2024JD040906. https://doi.org/10.1029/2024JD040906

Kim, J., Jeong, U., Ahn, M.H., Kim, J.H., Park, R.J., Lee, H., Song, C.H., Choi, Y.S., Lee, K.H., Yoo, J.M., Jeong, M.J., Park, S.K., Lee, K.M., Song, C.K., Kim, S.W., Kim, Y.J., Kim, S.W., Kim, M., Go, S., Choi, Y., 2020: New era of air quality monitoring from space: Geostationary Environment Monitoring Spectrometer (GEMS). *Bulletin of the American Meteorological Society*, 101(1), E1–E22. https://doi.org/10.1175/BAMS-D-18-0013.1

Levelt, P.F., van den Oord, G.H.J., Dobber, M.R., Mälkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J.O.V., & Saari, H., 2006: The Ozone Monitoring Instrument. *IEEE Transactions on Geoscience and Remote Sensing*, 44(5), 1093–1100. https://doi.org/10.1109/TGRS.2006.872333

Li, C., & Managi, S., 2022: Estimating monthly global ground-level NO_2 concentrations using geographically weighted panel regression. *Remote Sensing of Environment*, 280, 113152. https://doi.org/10.1016/j.rse.2022.113152

Loyola, D.G., García, S.G., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R.J.D., Pedergnana, M., Doicu, A., García, V.M., & Schüssler, O., 2018: The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor. *Atmospheric Measurement Techniques*, 11, 409–427. https://doi.org/10.5194/amt-11-409-2018

Pandya, P. A., Prajapati, G. V., Pradhan, B., Vadalia, D. D., Hirapara, P., Patel, D. J., & Parmar, S. H. (2026). From raw to reliable: machine learning bias correction of reanalysis data for improved drought severity classification. *Journal of Hydrology*, *667*, 134892. https://doi.org/10.1016/j.jhydrol.2025.134892

Reshi, A.R., Pichuka, S., & Tripathi, A., 2024: Applications of Sentinel-5P TROPOMI satellite sensor: A review. *IEEE Sensors Journal*, 24(13), 20312–20321. https://doi.org/10.1109/JSEN.2024.3355714

Sayyed, T.K., Ovienmhada, U., Kashani, M., Vohra, K., Kerr, G.H., O'Donnell, C., Harris, M.H., Gladson, L., Titus, A.R., Adamo, S.B., Fong, K.C., Gargulinski, E.M., Soja, A.J., Anenberg, S., & Kuwayama, Y., 2024: Satellite data for environmental justice: A scoping review of the literature in the United States. *Environmental Research Letters*, 19(3). https://doi.org/10.1088/1748-9326/ad1fa4

Song, J., Wang, Y., Zhang, Q., Qin, W., Pan, R., Yi, W., Xu, Z., Cheng, J., & Su, H., 2023: Premature mortality attributable to NO_2 exposure in cities and the role of built environment: A global analysis. *Science of the Total Environment*, 866, 161395. https://doi.org/10.1016/j.scitotenv.2023.161395

U.S. Environmental Protection Agency (EPA), 2023: Basic Information about NO_2. Available online: https://www.epa.gov/no2-pollution/basic-information-about-no2

Wimberly, M.C., de Beurs, K.M., Loboda, T.V., & Pan, W.K., 2021: Satellite observations and malaria: New opportunities for research and applications. *Trends in Parasitology*, 37(6), 525–537. https://doi.org/10.1016/j.pt.2021.03.003

World Health Organization (WHO), 2021: WHO global air quality guidelines. Particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Executive summary. Geneva: World Health Organization.