Engineering Technology
Open Access Journal
ISSN: 2641-8185

Research Article
Volume 6 Issue 5 - May 2026
DOI: 10.19080/ETOAJ.2026.06.555700

Eng Technol Open Acc

Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion

MAK Kai Long[1], KWAN Chun-Kit Wilson[2], MA Ka Yee[2], LAM King-Cheong[2], LOH Wai-Keung Anthony[2]*

[1]*Kent Business School, University of Kent, Canterbury, United Kingdom*

[2]*Division of Science, Engineering and Health Studies, School of Professional Education and Executive Development, The Hong Kong Polytechnic University, Hong Kong, China*

Submission: May 16, 2025; **Published:** May 25, 2026

***Corresponding author:** LOH Wai Keung Anthony, Division of Science, Engineering and Health Studies, School of Professional Education and Executive Development, The Hong Kong Polytechnic University, Hong Kong, China.

Abstract

Air quality in urban settings and environmental disruptions associated with traffic operations significantly impact public well-being and sustainable development of cities. The current research focuses on the interdependence between emissions of particulate matter and characteristics of traffic activity in terms of frequency distributions, especially the interaction patterns observed at different frequency levels and vertical dispersal properties. A set of data related to particulate matter concentrations and frequency-specific indicators were processed via the random forest analysis approach in order to detect non-linear associations between the considered factors including the height of measuring point, the type of pavement, and frequency characteristics. The findings prove that the height of measuring point is the key predictor of the concentrations of PM_1 and $PM_{2.5}$. At the same time, model performance drops when estimating $PM_{4.25}$ and PM_{10} values. Furthermore, the role of frequency-dependent environmental variables in particulate matter variations has been confirmed. However, the contribution of these factors depends on pavement types.

Keywords: Vertical Dispersion; Particulate Matter (PM); Vibration Frequency; Roadside Environment; Random Forest; Infrastructure Interaction

Abbreviations: TRAP: Traffic-Related Air Pollutants; PM: Particulate Matter; RMSE: Root Mean Squared Error

Introduction

It goes without saying that traffic activities are considered one of the key causes of urban particulate matter contamination. The rapid pace of urbanization, high levels of energy use, and a consistent rise in the number of cars all increase traffic emissions in urban areas. It should be noted that land-based transport plays an essential role in densely populated metropolitan areas like Hong Kong [1]. For example, in 2022, there were more than 900,000 vehicles in circulation in Hong Kong, with over 70% of them being private cars. Private vehicle utilization is convenient but often causes traffic jams, high levels of fuel consumption, and a rise in emission rates. To tackle this problem, the government of the Hong Kong Special Administrative Region has developed several solutions. To increase the efficiency of transportation, it introduced smart tools such as HKeToll that reduce the need for idling at tunnels [2]. At the same time, the administration encourages the adoption of sustainable transportation solutions in Hong Kong through the implementation of low-emission policies and promoting environmentally friendly bus services and taxis [3]. Nevertheless, traffic emissions remain a pressing issue that worsens the condition of urban air and increases environmental burden [4].

Extensive literature has been dedicated to the analysis of traffic-related air pollutants (TRAPs) focusing on the origins and dispersion behavior [5-7]. It is common for scientific studies to highlight the detrimental health impact of the prolonged intake of particulates, which manifests itself in respiratory disorders, heart diseases, and elevated mortality rates [8,9]. Particles smaller than one micron (e.g., PM_1) raise particular concerns due to their ability to penetrate deep into the human body, their high surface-to-mass ratio, and the potential for transport of hazardous compounds [10,11]. In addition to airborne contaminants, transportation networks emit numerous physical nuisances, such as noise pollution and ground vibrations. Despite ample documentation of the relationship between noise pollution and cardiovascular risks mediated via sleep disturbance and physiological stress [12], vibration pollution has not received adequate attention in environmental modelling literature. These vibrations emerge from the interaction between tires and roads, the dynamic nature of vehicles, and the state of the roadway infrastructure, potentially propagating throughout neighboring structures and the atmosphere. Notably, mechanical disturbances might interfere with airborne particulates through turbulence enhancement, resuspension, and vertical dispersion phenomena.

Due to the production of particulates and vibrations caused by traffic, researchers are becoming more interested in the possibility that the two variables have some correlation. While past studies have looked into the connection between noise and particulates, little research has been done on the effect of vibration properties, particularly frequencies, on particle behavior. It is important to understand the link to be able to create comprehensive assessment methods that can consider many facets of traffic effects rather than just individual ones. Furthermore, different properties of urban road surfaces can have an effect on both particle emissions and vibrations from them. For example, noise-reducing road surfaces that are common in Hong Kong may change the vibration pattern and consequently the particle dispersion behavior [13,14]. In addition, vertical particle distribution behavior needs to be assessed because of the presence of different urban layers where residential and commercial activities take place. Incorporating vibration data into particulate matter models using machine learning algorithms enables researchers to gain deeper insights into the functioning of urban environments. This research is expected to yield results relevant for developing a more comprehensive approach to the management of particulate matter pollution.

Materials and Methods

Random Forest Model

A random forest model is a supervised machine learning technique developed by [15] and further enhanced by [16]. Random forests have since become one of the most commonly used methods of data-driven modelling because of its robust nature and ability to analyse high dimensional datasets characterised by highly nonlinear interrelations. In general terms, a random forest is made up of decision trees grown on randomly chosen subsamples of the dataset whose outputs are combined to improve predictive performance and avoid overfitting.

In present-day applications, random forest models are widely used in different areas of research. For instance, [17] used a modified random forest algorithm to detect tax compliance risks in the real estate sector, highlighting its applicability in dealing with multivariate financial data. Additionally, [18] used competing risk random forest models to predict fracture risks and mortality in an older population. Therefore, the use of random forest models in research is not limited to any specific application because of their adaptability to various types of problems that require analysis of heterogeneous datasets.

In environmental modelling, the feature that allows random forests to analyse nonlinear interactions between variables proves very useful. Variables such as traffic induced vibration signals, particulate matter concentrations, road surfaces, and geographical features all have nonlinear and dependent relationships. As such, a random forest regression model will be used to analyse the interactions between the different variables in order to predict PM concentrations in relation to vibration frequencies.

Study Design

Study Area

This research was carried out in Hong Kong; an urbanized region located on the southeastern side of the Pearl River Delta. It is a heavily populated region with high-rise buildings, complex road structures, and traffic flows. This makes the region perfect for carrying out studies on the impacts of traffic on the environment. A total of twelve roadside sampling sites were selected based on the diversity of traffic levels, road geometry, and pavements. Different types of sites were identified including those with urban, industrial, and residential settings in order to obtain a comprehensive database. Samples were obtained between May 2023 and November 2023. These sampling sites have been chosen on the basis of traffic level, pavement type, and surrounding construction characteristics in order to understand how each of these variables influences vibration characteristics and the dispersion of particulates. Different pavement types used in this study include PMSMA10, Concrete, SMA20 and PMFC. (Table 1) summarizes these sites and their characteristics:

Vibration Measurement

The vibration features related to traffic have been determined through roadside noise measurements made on a handheld Brüel & Kjær 2245 sound level meter, which measures parameters like LAeq, LCeq, and LZeq [19]. In this research, the frequency-dependent noise level is considered as an indicator for traffic-induced disturbances occurring at the road-vehicle interface. The symbol “L” stands for the noise level; A, C, and Z are weights applied based on different criteria – for instance, auditory

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

response, peak pressure, and no weighting, respectively (Cirrus Research PLC, 2015). As no weighting ("Z" weighting) has been used in this work, the entire frequency range can be preserved. The notation "LZeq" indicates the equivalent noise level during a particular time period; accordingly, it accounts for the entire energy distribution within frequency bands. Data were obtained within one-third octave bands with frequencies ranging from 12.5 Hz to 16 kHz; one-second measurement intervals allowed the analysis of dynamic properties associated with the traffic flow, vehicle movement patterns, and driver behavior.

Table 1: Summary of road pavement data used in this study

Road Type	Road Name (Data Collection Date)	District
PMSMA10	448-458 Kwun Tong Road (15 Jun 23)	Kwun Tong
PMSMA10	85-89 Kwai Fuk Road (23 Jun 23)	Kwai Tsing
PMSMA10	448-458 Kwun Tong Road (04 Jul 23)	Kwun Tong
PMSMA10	11 Yuen Wo Road (12 Oct 24)	Shatin
Concrete	60 Hoi Yuen Road (30 Jun 23)	Kwun Tong
Concrete	60 Hoi Yuen Road (28 Jul 23)	Kwun Tong
Concrete	North Chatham Road, Near Wuhu Street (10 Aug 24)	Kowloon City
SMA20	8 Hung Lok Road (27 Jun 23)	Kowloon City
SMA20	8 Hung Lok Road (11 Jul 23)	Kowloon City
SMA20	8 Hung Lok Road (18 Aug 23)	Kowloon City
PMFC	Whampoa Garden (25 Aug 24)	Kowloon City
PMFC	Lung Cheong Road (09 Nov 24)	Wong Tai Sin

PM Measurement

Concentration of particulate matter was measured using the Trolex AirXD Dust Monitor capable of measuring the concentration of particulates within different diameter ranges of PM_1, $PM_{2.5}$, $PM_{4.25}$, and PM_{10} [20]. This differs from traditional measurement devices that are focused on measuring $PM_{2.5}$ and PM_{10} concentrations [21]. Measurement values obtained are reported in units of micrograms per cubic meter while being in synchronization with the vibration data for time-series analysis. The use of different particle diameters allows examining the impact of vibration on different sizes of particles with unique dispersion and deposition characteristics.

Sampling Setup

Data collection involved sampling at three different levels which were perpendicular to the road's pavement surface as depicted in (Figure 1). The elevation points considered depended on the built environment, but in such a manner that would not inhibit the smooth functioning of the road. Ground level data collection was done closest to the road surface as this is the most exposed area to vehicular pollutants and signals due to its proximity to the sources of the pollution. Sampling at the mid-level was at an elevated point above the ground to determine how particulate matter behaves when subjected to flow dynamics. Sampling at the high level was above mid-level to examine how much dispersion is experienced in order to determine if particulate matter can spread beyond the immediate vicinity. By doing the sampling at three levels as described above, it is possible to gain an overview of the dispersion processes. It will further be possible to investigate the relationship between signal frequencies and particulate matter concentrations in the road pavements. Data was collected during the day time to guarantee safety and to observe traffic movement. Each sampling took at least two hours to collect enough data. The flow chart in (Figure 2) shows how the process was carried out from the planning stages to the analysis stage.

Data Processing

Data Validation

The validity of the collected data was ensured through several processes. All data measurement tools like the sound level meter and Trolex AirXD Dust Monitor were calibrated according to the instructions of the respective manufacturers before starting the measurement process. Compliance testing was done for Trolex AirXD Dust Monitor after each sample period to ensure the accuracy of the results obtained. Furthermore, cross-validation of the collected data was carried out using second-level data from environmental monitoring stations maintained by the Environmental Protection Department. Also, data obtained from the selected spots on different days were used for data stability analysis.

Data Cleaning

Outliers and missing data were detected during initial data analysis. The Z-Score approach was used to detect outliers as it focuses on detecting the presence of values that lie far from the average (Mondal, 2024). Outliers were analyzed and eliminated to avoid influencing the outcome of the model. Interpolation approaches were employed to fill in gaps by using surrounding values.

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

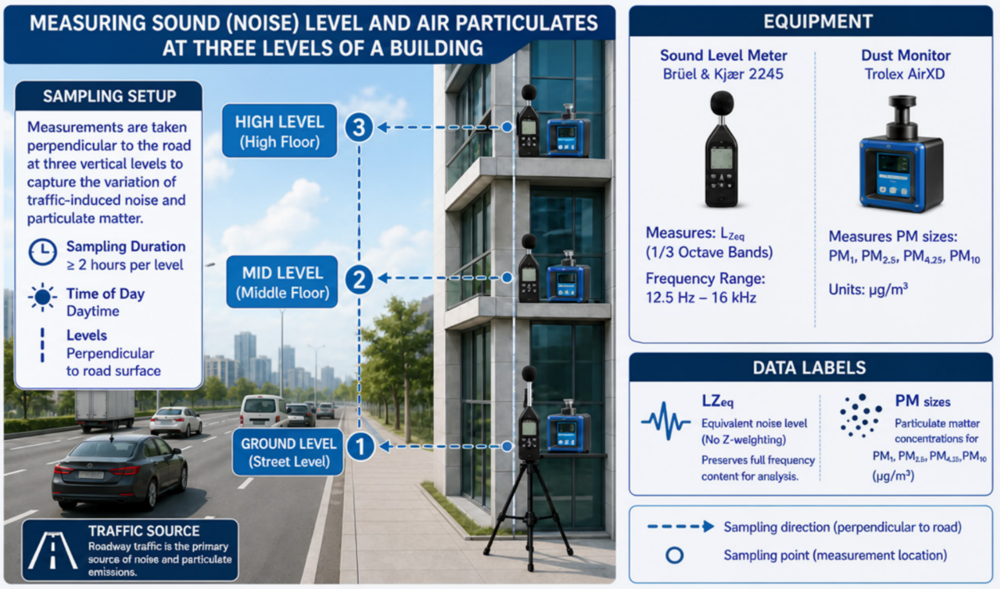

Figure 1: Data collection and sampling setup at three different levels of a building.

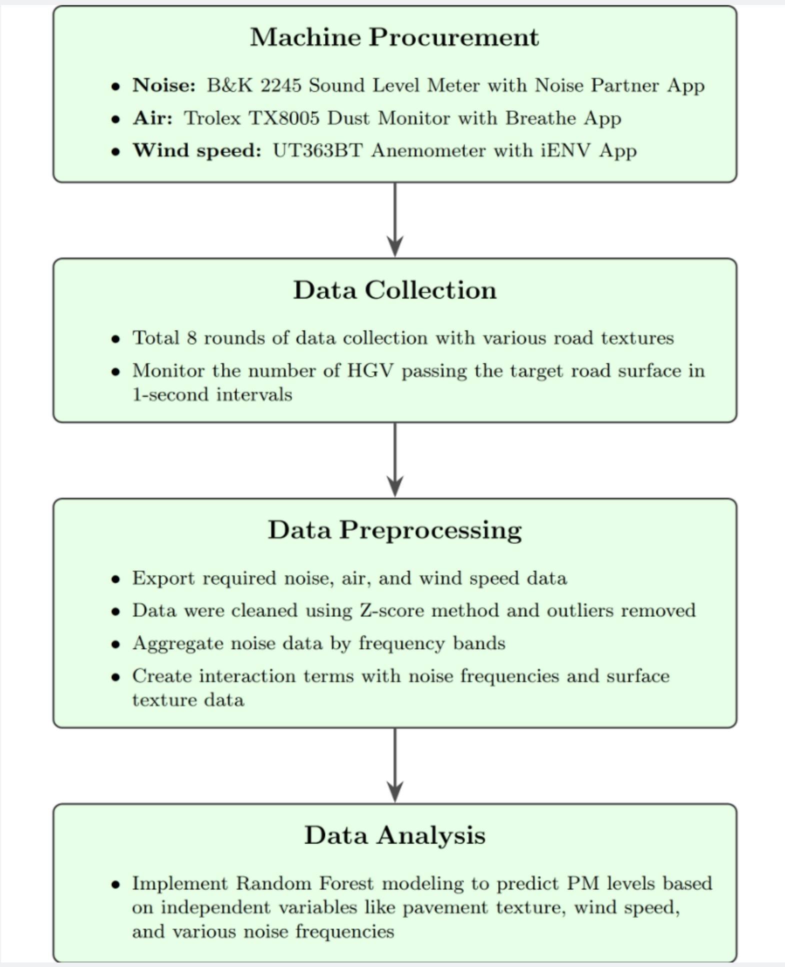

Figure 2: The flow of study

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

Feature Engineering

The following feature engineering process was undertaken to convert raw data to useful features that could be analyzed using machine learning techniques. Categorical features such as road surface features were converted to numeric features so that they could be fed to the random forest algorithm. Temporal features like time of the day and day of the week were computed as they would provide an indication of changes in the traffic situation. The vibration data was reduced to chosen frequency ranges in order to decrease dimensionality without losing any significant information. Interactions between the vibration frequencies and road surfaces were incorporated into the features set.

Data Splitting

Data splitting for training and testing was done using the method of stratified sampling. A ratio of 70:30 was used for training and testing sets, respectively, to ensure consistency in the levels of particulate matter in both sets.

Model Development

In the development of the random forest model, it was necessary to configure the model based on its key parameters such as the number of trees, the maximum depth of the tree, and the minimum number of samples that should be used for the splitting process. The parameters could be optimized through grid searching. Bootstrap sampling was used in the training of decision trees, as well as random selection of features in every split.

Model Evaluation

The root mean squared error and coefficient of determination were used to assess the accuracy of the model based on the testing data. This was done since these measures reflect prediction accuracy and variance explained by the model. To determine the predictors that have significant influence on variability in the particulate matter data, a feature importance analysis was carried out.

Results

Measurement Data

The last data set contained around 4,610,700 data points that were obtained after the cleaning processes from the data collected from twelve different sampling points. These data included vibration-based frequency data, particulate matter concentration, different types of roads used for the sampling process, and the height of measurement. Different sampling points were chosen based on differences in pavement texture. Four samplings were taken from PMSMA10 pavements, three samplings from Concrete pavements, two samplings from SMA20 pavements, and finally three samplings from PMFC pavements. The data was collected in one-second time intervals which ensured the ability to examine the temporal variation of particulate matter concentration and frequency-based environmental signals.

Model Training

Random forest algorithm was set up with some parameters in order to guarantee the satisfactory learning capabilities. First, the parameter (n_estimators) was set to 200 in order to ensure sufficient learning stability and generalization ability of the random forest. The maximum depth of each tree (max_depth) was set with the help of grid search in order to maintain a trade-off between the complexity of the model and its overfitting tendency. Similarly, the min_samples_split parameter was fine-tuned in order to boost the predictive capabilities of the random forest algorithm and maintain computational efficiency. During the training process, each tree in the random forest model was trained based on bootstrap sampling where some random samples of the data were chosen in each iteration. Moreover, a random subset of features is considered during node splitting.

Model Evaluation and Performance Metrics

Once constructed, the model was subjected to the test set of data as a means of assessing its performance. Two major criteria were applied: root mean squared error (RMSE) and the coefficient of determination (R^2). In essence, RMSE serves as a measure of errors committed during the prediction process, whereas R^2 signifies the share of variability explained by the model. In general terms, the model demonstrates fairly good performance concerning the prediction of smaller-sized particles, PM_1 and $PM_{2.5}$; however, its performance is lower when it comes to predicting larger-sized particles, e.g., $PM_{4.25}$ and PM_{10}. This means that the selected factors better explain the variability of ultrafine and fine particles than coarse particles. This claim is also supported by the results of the feature importance assessment, according to which measurement height stands out as one of the key influential factors. At the same time, the influence exerted by frequency-based variables depends on the surface type and particle size.

Feature Importance Analysis

To explore the relationship between PM concentration and frequency-related features, the evaluation of feature importance rankings obtained from the random forest model was conducted across different road textures. The results are summarized in (Tables 2-5). For the PMFC surfaces, the R^2 of the random forest models for PM_1 and $PM_{2.5}$ reached 0.81 and 0.53, respectively, showing satisfactory predictive performance. The important variables include measurement height, whose importance reached approximately 0.80 for PM_1 and 0.50 for $PM_{2.5}$. On the contrary, the prediction performances for $PM_{4.25}$ and PM_{10} became much worse, with the corresponding R^2 being 0.35 and 0.11, respectively. On the PMSMA10 surfaces, the models performed satisfactorily on PM_1 and $PM_{2.5}$ with R^2 reaching 0.88 and 0.71, respectively. Like the results for PMFC surfaces, measurement height played a dominating role for PM prediction, followed by selected frequency-related features. However, the prediction performance for PM_{10} was still relatively poor, with the R^2 being 0.14. For the concrete

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

surfaces, consistent results were found for small particles with R^2 of 0.93 and 0.73 for PM_1 and $PM_{2.5}$, respectively. Similar to previous cases, measurement height maintained a high importance level while decreasing for large particles. As for the SMA20 surfaces, overall prediction performances of the models were relatively lower than other surfaces, especially for large particles. However, it could be seen that measurement height played a dominating role in the models with different levels of importance. Also, some frequency features made contributions to the models. To sum up, in general, it could be noticed that the importance of measurement height is relatively higher when predicting small particle sizes. Frequency-related features are also included but their importance varies according to the surface type.

Table 2: Feature importance table for PMFC type

PM_1		**$PM_{2.5}$**	
R-Square	0.81	R-Square	0.53
Root Mean Squared Error (RMSE)	0.66	Root Mean Squared Error (RMSE)	4.28
Feature Importances		Feature Importances	
Height (m)	0.800022	Height (m)	0.500737
LZeq 315 Hz	0.011566	LZeq 31.5 Hz	0.028658
LZeq 31.5 Hz	0.011442	LZeq 50 Hz	0.027335
LZeq 250 Hz	0.010982	LZeq 40 Hz	0.027264
LZeq 25 Hz	0.010871	LZeq 63 Hz	0.027074
LZeq 40 Hz	0.010724	LZeq 80 Hz	0.026945
LZeq 160 Hz	0.010527	LZeq 20 Hz	0.026576
LZeq 200 Hz	0.01036	LZeq 125 Hz	0.026567
LZeq 50 Hz	0.010194	LZeq 25 Hz	0.025702
LZeq 63 Hz	0.010164	LZeq 200 Hz	0.025589
$PM_{4.25}$		PM_{10}	
R-Square	0.35	R-Square	0.11
Root Mean Squared Error (RMSE)	11.37	Root Mean Squared Error (RMSE)	47.8
Feature Importances		Feature Importances	
Height (m)	0.31885	Height (m)	0.10266
LZeq 31.5 Hz	0.039161	LZeq 31.5 Hz	0.064332
LZeq 40 Hz	0.03887	LZeq 125 Hz	0.058807
LZeq 125 Hz	0.037311	LZeq 160 Hz	0.052566
LZeq 50 Hz	0.037291	LZeq 40 Hz	0.051444
LZeq 25 Hz	0.036483	LZeq 100 Hz	0.05015
LZeq 20 Hz	0.036065	LZeq 50 Hz	0.046955
LZeq 160 Hz	0.035594	LZeq 63 Hz	0.046554
LZeq 63 Hz	0.035521	LZeq 20 Hz	0.046138
LZeq 100 Hz	0.035198	LZeq 80 Hz	0.045115

Table 3: Feature importance table for PMSMA10 type

PM_1		$PM_{2.5}$	
R-Square	0.88	R-Square	0.71
Root Mean Squared Error (RMSE)	0.96	Root Mean Squared Error (RMSE)	2.99
Feature Importances		Feature Importances	
Height (m)	0.812724	Height (m)	0.707655
LZeq 800 Hz	0.06267	LZeq 160 Hz	0.019825

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

LZeq 1000 Hz	0.011981	LZeq 63 Hz	0.019408
LZeq 1600 Hz	0.011759	LZeq 80 Hz	0.019255
LZeq 160 Hz	0.009357	LZeq 50 Hz	0.019163
LZeq 250 Hz	0.008447	LZeq 200 Hz	0.019131
LZeq 200 Hz	0.008417	LZeq 125 Hz	0.0191
LZeq 80 Hz	0.008305	LZeq 100 Hz	0.018599
LZeq 500 Hz	0.008165	LZeq 250 Hz	0.018239
LZeq 1250 Hz	0.007847	LZeq 315 Hz	0.018096
$PM_{4.25}$		PM_{10}	
R-Square	0.48	R-Square	0.14
Root Mean Squared Error (RMSE)	7.6	Root Mean Squared Error (RMSE)	32.16
Feature Importances		Feature Importances	
Height (m)	0.4756	Height (m)	0.164277
LZeq 50 Hz	0.039306	LZeq 50 Hz	0.065775
LZeq 63 Hz	0.037551	LZeq 63 Hz	0.065234
LZeq 80 Hz	0.035884	LZeq 250 Hz	0.059205
LZeq 160 Hz	0.035792	LZeq 80 Hz	0.058275
LZeq 200 Hz	0.034786	LZeq 160 Hz	0.055725
LZeq 125 Hz	0.034358	LZeq 125 Hz	0.055417
LZeq 100 Hz	0.034202	LZeq 200 Hz	0.055238
LZeq 250 Hz	0.033555	LZeq 100 Hz	0.054834
LZeq 315 Hz	0.032475	LZeq 315 Hz	0.053346

Table 4: Feature importance table for Concrete type

PM_1		$PM_{2.5}$	
R-Square	0.79	R-Square	0.5
Root Mean Squared Error (RMSE)	0.0.49	Root Mean Squared Error (RMSE)	4.24
Feature Importances		Feature Importances	
Height (m)	0.705186	Height (m)	0.483591
LZeq 1000 Hz	0.072855	LZeq50 Hz	0.036851
LZeq 315 Hz	0.020451	LZeq 63 Hz	0.035788
LZeq 200 Hz	0.017367	LZeq 80 Hz	0.035338
LZeq 800 Hz	0.015396	LZeq 200 Hz	0.034932
LZeq 250 Hz	0.015032	LZeq 125 Hz	0.034687
LZeq 63 Hz	0.015013	LZeq 100 Hz	0.03437
LZeq 160 Hz	0.014895	LZeq 315 Hz	0.034216
LZeq 400 Hz	0.014758	LZeq 250 Hz	0.034018
LZeq 50 Hz	0.014676	LZeq 160 Hz	0.033473
PM4.25		PM10	
R-Square	0.48	R-Square	0.14
Root Mean Squared Error (RMSE)	7.6	Root Mean Squared Error (RMSE)	32.16
Feature Importances		Feature Importances	
Height (m)	0.305795	Height (m)	0.094766

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

LZeq 50 Hz	0.049775	LZeq 125 Hz	0.068147
LZeq 125 Hz	0.047418	LZeq 50 Hz	0.067499
LZeq 100 Hz	0.047006	LZeq 160 Hz	0.066826
LZeq 63 Hz	0.04667	LZeq 100 Hz	0.062186
LZeq 80 Hz	0.046547	LZeq 63 Hz	0.060797
LZeq 160 Hz	0.046107	LZeq 80 Hz	0.057249
LZeq 200 Hz	0.045942	LZeq 250 Hz	0.056371
LZeq 250 Hz	0.043974	LZeq 1600 Hz	0.055856
LZeq 315 Hz	0.043859	LZeq 200 Hz	0.055339

Table 5: Feature importance table for SMA20 type

PM_1		$PM_{2.5}$	
R-Square	0.38	R-Square	0.21
Root Mean Squared Error (RMSE)	0.03	Root Mean Squared Error (RMSE)	1.84
Feature Importances		Feature Importances	
Height (m)	0.341853	Height (m)	0.19967
LZeq 100 Hz	0.047322	LZeq 200 Hz	0.054499
LZeq 80 Hz	0.044147	LZeq 100 Hz	0.054216
LZeq 125 Hz	0.043618	LZeq 63 Hz	0.053955
LZeq 500 Hz	0.04328	LZeq 125 Hz	0.053535
LZeq 200 Hz	0.043219	LZeq 50 Hz	0.053085
LZeq 50 Hz	0.043153	LZeq 80 Hz	0.052662
LZeq 160 Hz	0.042445	LZeq 250 Hz	0.05162
LZeq 1250 Hz	0.042421	LZeq 160 Hz	0.051426
LZeq 400 Hz	0.04178	LZeq 400 Hz	0.04944
$PM_{4.25}$		PM_{10}	
R-Square	0.09	R-Square	0
Root Mean Squared Error (RMSE)	4.95	Root Mean Squared Error (RMSE)	21.19
Feature Importances		Feature Importances	
Height (m)	0.096283	LZeq 63 Hz	0.078111
LZeq 50 Hz	0.063883	LZeq 200 Hz	0.074688
LZeq 63 Hz	0.062751	LZeq 50 Hz	0.072675
LZeq 160 Hz	0.060526	LZeq 100 Hz	0.066107
LZeq 200 Hz	0.060508	LZeq 80 Hz	0.065562
LZeq 80 Hz	0.059812	LZeq 160 Hz	0.065339
LZeq 125 Hz	0.059646	LZeq 125 Hz	0.063317
LZeq 100 Hz	0.05893	LZeq 800 Hz	0.057031
LZeq 250 Hz	0.057285	LZeq 400 Hz	0.05672
LZeq 315 Hz	0.055734	LZeq 315 Hz	0.056431

Overall Observations

It can be observed that the correlation between the dispersion of particulate matter and the frequency environmental cues is affected by several interactive parameters such as the height at which measurements were taken, the nature of the pavement, and the size of the particles. It is also evident from the model that it performs best when trying to predict ultrafine particles. Larger particles show

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

more variability than what can be explained using the existing set of variables.

Discussion

In addition, this study offers an analysis of the vertical relationship between traffic frequency signals and particulate matter emissions through a random forest regression model. The results from this model show the potential for the management of nonlinear relations within complex data, which involves many features as the random forest includes feature selection in the training process. Firstly, one notable result is that the height at which the measurement was taken is a highly influential factor when it comes to predictions regarding particulate matter concentrations in PM_1, $PM_{2.5}$ for PMFC, PMSMA10 and Concrete pavements. The results also confirm that R^2 for these variables surpasses 0.50, and therefore the predictive abilities of the model are very good. On the other hand, the predictive abilities are fairly poor concerning larger particles such as PM_{10}, since the corresponding R^2 values are generally below 0.20. This indicates that there may be some environmental variables other than those used within the model that affect the prediction of larger particles.

These findings highlight the challenges in predicting coarse particles and point to the necessity of further improvements regarding input variables. Additionally, the ability of the random forest model to predict smaller particle sizes shows the model's effectiveness in managing the complex relation between the numerous environmental variables and particulates in urban settings such as Hong Kong. Low noise road surfacing material is often applied in road construction in order to minimize environmental disturbance. Nevertheless, fossil fuel vehicles along with a complex urban environment create significant pollution and environmental effects. Moreover, the results confirm that the height at which the measurement takes place proves to be the most effective indicator for all sizes of particles especially PM_1 due to their physical properties of being smaller in size, less in mass and more in surface area as compared to other particles. Such an advantage allows them to remain suspended in the environment for long and increases their capacity to absorb harmful chemicals causing serious health problems ranging from heart attack to various types of respiratory problems. The growing trend towards urbanisation and ownership of vehicles in the contemporary world calls for a need to predict the emissions of particulate matter in the environment through advanced models [22].

The results of the current research have wider implications not only for Hong Kong but also other cities facing serious environmental issues with regards to air quality like Los Angeles and Beijing where the policymakers could use the same data-driven approach to understand relationships between traffic and environment in order to adopt an integrated approach towards traffic management considering both PM emissions and physical disturbance caused to the environment [23-28]. One such solution includes the promotion of electric or hybrid cars as well as the adoption of advanced technologies in terms of roads construction. In addition to it, the relationship established between the frequencies of signals and emissions of particulate matter proves that the measures aimed at minimizing environmental disturbances related to traffic could also prove useful in improving air quality. Such measures may include different types of infrastructural changes which, besides addressing the problem of traffic, may affect the pattern of particles emission in the environment. In terms of public health considerations, the results emphasize the necessity of decreasing exposure to ultrafine particulate matter. Campaigns that aim at raising public awareness and encouraging individuals to use alternative means of transportation such as public transit, biking, and walking could reduce the number of privately owned cars and subsequently lower the risks of exposure [29,30]. Some cities where sustainable transport systems have already been implemented, like Amsterdam, may serve as an example in terms of city planning and policy-making.

Another conclusion that can be drawn based on the analysis of feature importance is that RMSE increases along with the growth of particle size. Thus, the model proves less accurate when predicting the behavior of large particles and implies that certain parameters have an impact on their movement that is not present in the current data [31,32]. It may include wind speed, humidity, and atmospheric conditions. Therefore, these variables should be taken into consideration in further studies in order to boost the precision of predictions. As far as the literature review shows, the research in question addresses one of the under-researched topics and uses machine learning algorithms to analyze vertical interactions between particulate matter and frequency characteristics of traffic activity. The results may be helpful in future studies aimed at studying the influence of road surface properties on pollution, designing effective measures to control pollutants, and implementing monitoring systems.

Conclusions

In this study, a detailed examination of the influence of traffic signal frequency properties on particulate matter distribution was carried out. It is found that the most important factor in predicting the concentration of particulate matter, especially for small PM_1, is the measurement height. The model based on a random forest showed high predictability in case of analysis of small PM_1 and $PM_{2.5}$ but low prediction accuracy in case of the prediction of larger $PM_{4.25}$ and PM_{10}. Thus, further studies require more influencing factors to be considered, including meteorological characteristics. Overall, this result proves that taking into account the process of vertical distribution of particulate matter and surface behavior is important for an effective analysis of the particulate matter concentration. From the practical point of view, the results obtained can be helpful in developing strategies aimed at addressing the

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

issues of emission and dispersion of harmful substances. It can be concluded that despite some benefits of the random forest algorithm, such as the ability to analyze nonlinear dependencies and correlations between environmental variables, the processing of large amounts of data can create additional difficulties for prediction due to computational limitations. Thus, future studies are necessary to increase the effectiveness of the model used by expanding the number of considered environmental factors.

Acknowledgements

This research was fully supported by a grant from the Research Grants Council of the Hog Kong Special Administrative Region, China (Project No. UGC/FDS24/E12/21).

References

1. Lee D (2022) To build hong kong into a green, low-traffic city, think beyond technology. The View.
2. HKSAR Government (2022) The Chief Executive's 2022 Policy Address.
3. HKSAR Government (2023) The Chief Executive's 2023 Policy Address.
4. Muresan B, Cesbron J, Lumière L, Gary V, Truong XT (2024) A study of the relationship between rear-of-wheel particle emissions and close-proximity tire/road noise of a passenger car. Science of The Total Environment 918: 170578.
5. Dallmann TR, De Martini SJ, Kirchstetter TW, Herndon SC, Onasch TB (2012) On-road measurement of gas and particle phase pollutant emission factors for individual heavy-duty diesel trucks. Environmental Science & amp Technology 46(15): 8511-8518.
6. Gentner DR, Jathar SH, Gordon TD, Bahreini R, Day DA et al., (2017) Review of Urban Secondary Organic Aerosol Formation from gasoline and diesel motor vehicle emissions. Environmental Science & amp; Technology 51(3): 1074-1093.
7. Zhou L, Salvador CM, Priestley M, Hallquist M, Liu Q (2021) Emissions and secondary formation of air pollutants from modern heavy-duty trucks in real-world traffic - chemical characteristics using on-line mass spectrometry. Environmental Science & Technology 55(21): 14515–14525.
8. Kumar N, Kumar S, Pandey SP (2023) Traffic-related air pollution and associated human health risk. Macromolecular Symposia 407(1): 2100486.
9. Kurt OK, Zhang J, Pinkerton KE (2016) Pulmonary health effects of air pollution. Current Opinion in Pulmonary Medicine, 22(2): 138-143.
10. Lim JM, Lee CY, Lee YJ, Kim H, Lee JH (2022) Characteristics of airborne PM1.0 and associated chemical constituents at a roadside area in Korea. Environmental Engineering Research 28(3): 220089–0.
11. Camfil (2024) PM1 - the new focus to protect human health. Standard & Regulations.
12. Mann S, Singh G (2022) Traffic noise monitoring and modeling — an overview. Environmental Science and Pollution Research, 29(37): 55568–55579.
13. HKSAR Legco (2022) Promoting use of new energy commercial vehicles. Press Release.
14. GovHK (2023) Greener Transportation. Green Living.
15. Ho TK (1995) Random decision forests. Proceedings of 3rd International Conference on Document Analysis and Recognition. Montreal 278-282.
16. Breiman L (2001) Random forests. Machine learning 45: 5-32.
17. Xu C, Kong Y (2024) Random Forest model in tax risk identification of Real Estate Enterprise Income tax PLOS ONE 19(3).
18. Langsetmo L, Schousboe JT, Taylor BC, Cauley JA, Fink HA, et al., (2023a) Advantages and disadvantages of random forest models for prediction of hip fracture risk versus mortality risk in the oldest old. JBMR Plus 7(8).
19. Fallah Shorshani M, Fruin S, Yin X, Mc Connell R, Franklin M (2024) Estimating near-roadway air pollution from multifrequency noise measurements. The Science of the Total Environment 944: 173900.
20. Trolex (2023) Air XD User Manual. In AIR XD_User manual_ P5628.1600_Rev G: Vol. G.
21. IQ Air (2021)
22. Duan L, Song L, Wang W, Jian X, Heijungs R, et al. (2024) Urbanization inequality: evidence from vehicle ownership in Chinese cities. Humanit Soc Sci Commun 11: 703.
23. Hoang LA, Quoc Ho B, Harris N, Drew GH, MI Mohammed (2023) Street-scale dispersion modeling framework of road-traffic derived air pollution in Hanoi, Vietnam Environmental research 233: 116497.
24. Choudhary A, Gokhale S (2016) Urban real-world driving traffic emissions during interruption and congestion. Transportation Research Part D Transport and Environment 43: 59-70.
25. Cirrus Research PLC (2015) A Guide to Noise Measurement Terminology. Cirrus Research plc.
26. Donges N (2024) Random Forest: A complete guide for machine learning. Built In.
27. Gu D, Andreev K, E Dupre M (2021) Major trends in population growth around the world. China CDC Weekly 3(28): 604-613.
28. Hou L, Wang Y, Zheng Y, Zhang A (2022) The impact of vehicle ownership on carbon emissions in the Transportation Sector. Sustainability 14(19): 12657.
29. Lewinson E (2023) A Comprehensive Guide on Interaction Terms in Time Series Forecasting | NVIDIA Technical Blog. NVIDIA Technical Blog.
30. Mondal R (2024) Z Score to identify and remove outliers | Exploratory Data Analysis. Medium.
31. Singh P, Singh N, Singh KK, Singh A (2021) Diagnosing of disease using machine learning. Machine Learning and the internet of Medical Things in Healthcare 89-111.
32. Wallace ML, Mentch L, Wheeler BJ, Tapia AL, Richards M et al., (2023) Use and misuse of random forest variable importance metrics in medicine: Demonstrations through incident stroke prediction. BMC Medical Research Methodology 23(144).

How to cite this article: MAK KL, KWAN CKW, MA KY, LAM KC, LOH WKA. Urban Environmental Interaction Modeling: The Relationship Between Traffic-Induced Vibrations and Particulate Matter Dispersion. Eng Technol Open Acc 2026; 6(4): 555700.
DOI: 10.19080/ETOAJ.2026.06.555700

DOI: 10.19080/ETOAJ.2026.06.555700