Temporal and Spatial Distribution Characteristics of NO_x Emissions of City Buses on Real Road Based on Spatial Autocorrelation

Qikai Peng^{1,2}, Jiaqiang Li^{1,2}, Yanyan Wang^{1,2}, Longqing Zhao^{1,2}, Jianwei Tan³,
Chao He^{1,2*}

¹ School of Mechanical and Transportation Engineering, Southwest Forestry University, Kunming 650224, China

² Key Lab of Vehicle Emission and Safety on Plateau Mountain, Yunnan Provincial Department of Education, Kunming 650224, China

³ School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China

ABSTRACT

To characterize the spatial and temporal distribution of NO_x exhausted by urban buses, we measured real-world on-road NO_x emissions from these vehicles in the city of Kunming, China, using an onboard monitoring platform. To fill the data gaps and produce a complete data set, we combined Bayesian network modeling and probabilistic inference. The complete data set was then used to generate an NO_x emission heat map, and spatial autocorrelation was applied to evaluate the distribution characteristics. The results show that our method for filling in the missing data provides highly accurate values, with spatial autocorrelation indices of 0.648, 0.836, 0.935, and 0.798 for the morning, midday, afternoon, and evening, respectively. The NO_x emissions showed spatial correlation during all four periods, whereas the pollutive emissions showed spatial aggregation. According to the heat map, the NO_x concentrations peaked during the midday and the afternoon. Furthermore, regardless of the period, the largest emissions accumulated in Road Sections 1–3 and 6–9, and the highest as well as the fastest-growing emission intensity occurred in Road Sections 5–9.

Keywords: Buses, Spatial autocorrelation, NO_x emissions, Temporal and spatial distribution characteristics

OPEN ACCESS

Received: July 12, 2020

Revised: December 20, 2020

Accepted: January 29, 2021

* Corresponding Author:

hcsmile@163.com

Publisher:

Taiwan Association for Aerosol
Research

ISSN: 1680-8584 print

ISSN: 2071-1409 online

Copyright: The Author's institution. This is an open access article distributed under the terms of the [Creative Commons Attribution License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are cited.

1 INTRODUCTION

Air pollution is an important global risk factor for disease and death (Lim *et al.*, 2012; Apte *et al.*, 2015). Measurement of air pollution is critical for epidemiology and air quality management, especially for ground-based urban air pollution detection (Brauer *et al.*, 2015; Carvalho, 2016). In recent years, with the rapid development of the urban economy, the number of vehicles has sharply increased, and traffic emissions have become an important source of urban air pollution (Zhang and Batterman, 2013; Su *et al.*, 2015). Due to the rapid development of urbanization in China, urban air pollution has shifted from local single-point pollution to complex regional pollution sources (Chai *et al.*, 2006). One of the important sources of composite air pollution in China is vehicle exhaust emissions, which are especially important for pollutants such as fine particulate matter (PM_{2.5}), carbon monoxide (CO), nitrogen oxides (NO_x), and volatile organic compounds (VOCs) in urban air (Fan *et al.*, 2015). Urban diesel buses are the main contributors to urban NO_x emissions. Through photochemical reactions NO_x can produce photochemical smog, which is extremely harmful to the human body (Wang *et al.*, 2011; Li *et al.*, 2012; Liu *et al.*, 2013). Strengthening the control of NO_x from urban public transport is important to reduce urban air pollution and optimize quality of life. Therefore, it is important to study the spatial and temporal distribution characteristics

of urban public transport NO_x emissions and to analyze its correlations to improve urban air quality and strengthen air quality management (Zhang *et al.*, 2017; Zhang *et al.*, 2019).

Due to the particularities of the urban bus operating environment, which consist of frequent starting, braking, idling, and low-speed driving conditions, the engine is mostly running at low-speed, high-torque, and idle-speed conditions. To control motor vehicle exhaust emissions more effectively, researchers analyzed the spatial and temporal distribution characteristics and laws of vehicle exhaust emissions based on different models and methods. A statistics-based Monte Carlo simulation method was used to analyze the uncertainty of emissions estimates by simulating a large number of vehicle emissions data points (Lv *et al.*, 2019). The Corinair method was used to estimate the emission reductions brought about by urban bus modernization and to estimate the contributions of particulate matter and gaseous pollutants emitted by city buses in the Krakow region (Bogacki and Bzdziuch, 2019). Hao *et al.* (2017) and Carslaw *et al.* (2019) summarized the results of remote-sensing measurements of vehicle emissions in the United Kingdom, providing detailed information on NO₂ and total NO_x (NO₂ + NO) emissions and analyses of NO_x emissions trends in recent years. Liu *et al.* (2018) used a Global Positioning System (GPS) receiver to collect per-second activity data, used 2-second speed and acceleration to determine the bus operating mode, and applied the vehicle-specific power (VSP) to estimate the corresponding emissions. Chen *et al.* (2018) used a chassis dynamometer to explore the impact of various factors on the emissions of heavy-duty diesel vehicles and to test and analyze Euro VI standard heavy-duty diesel vehicles under different loads, fuels, driving environments, and resistance conditions. He *et al.* (2016) used a large number of methods to obtain real-time dynamic traffic flow information and used the Computer Programme to Calculate Emissions from Road Transport (COPERT) model to analyze the spatiotemporal characteristics of Beijing vehicle emissions. Wang *et al.* (2012) simulated the exhaust emissions of diesel vehicles under different working conditions and analyzed their emission characteristics. Chen *et al.* (2015) investigated and analyzed the various conditions of Chengdu light-duty gasoline passenger vehicles and used the International Vehicle Emission (IVE) model to analyze their exhaust emission characteristics and to establish a local emissions inventory. Kan *et al.* (2018) carried out an autocorrelation analysis of vehicle exhaust emissions in Nanjing through the portable emissions measurement system (PEMS). Different methods and ideas can be used to obtain the desired results and create a database and inventory of local vehicle pollutant emissions characteristics (Wang and Li, 2016).

Liu *et al.* (2018) built a high-resolution temporal-spatial vehicle emission inventory of the city of Foshan based on detailed hourly traffic data and a COPERT emissions model. Kan *et al.* (2018) proposed a method that estimated fuel consumption and emissions at a fine-grained level based on vehicle GPS trajectories and microscopic model Comprehensive Modal Emission Model (CMEMs). Alam *et al.* (2017) estimated transit bus emissions by using Motor Vehicle Emission Simulator (MOVES) embedded drive cycles with local data.

In these studies, real traffic conditions were included to characterize temporal and spatial distribution; however, the emission data were calculated by emission models rather than from actual road emissions. PEMS have been widely used to record various types of vehicle emissions data under real driving conditions, but cannot reflect the long-term emission characteristics of transit buses (Apte *et al.*, 2017).

This paper is based on a study of the spatial and temporal distribution characteristics of NO_x from urban buses on real roads, using a data-filling method combining a Bayesian network and probabilistic reasoning to fill in the missing entries in the data stored in the tensor model. Then the data set was generated again for storage, and a tensor model was used to facilitate data storage and extraction. Spatial autocorrelation analysis combined with the ArcGIS online platform was used to determine the spatial and temporal distribution characteristics of pollutant emissions, which was helpful for understanding the spatial and temporal distribution characteristics of urban bus NO_x emissions.

2 DATA SOURCES AND RESEARCH METHODS

2.1 Data Sources

The No. 80 bus in Kunming was selected as the research object. The data used in this study

came from an online vehicle exhaust monitoring platform that was established to obtain dynamic real-time vehicle exhaust data for specific routes throughout the day. A terminal device was installed on the bus, which includes an NO_x sensor, a GPS receiver, an onboard diagnosis system (OBD) data reader and a wireless data transmission unit. The NO_x sensor is an electrochemical sensor, which can directly measure the bus NO_x emission concentration second by second. The GPS system can provide navigation and positioning information, such as real-time longitude, latitude, and altitude, and positioning function to obtain the coordinates of each point. The collected data are transmitted to the corresponding platform by General Packet Radio Service (GPRS). The time for data collection and experimentation in this study was 20 weeks, extending from 2017 January 2 to 2017 May 22. The data collection hours were approximately 06:00–22:30, covering both working and non-working days and including the actual operation of the whole route throughout the day.

2.2 Research Methods

The real-time NO_x emissions obtained from the online monitoring platform were extracted, integrated, and stored in the tensor model. The missing data points were then filled in using a Bayesian network and probabilistic reasoning to produce a complete data set. Finally, spatial autocorrelation coefficients generated by the ArcGIS software platform were used to study the spatial and temporal distribution characteristics of NO_x.

2.3 Instantaneous Emission Calculation

We extracted real-time NO_x emissions quality data from the online monitoring platform and calculated the instantaneous emission rates, which were used to calculate the corresponding emission factor. Therefore, the complete working condition data and the corresponding NO_x transient emission rates (NO_x instantaneous wet-based concentrations) were obtained. The NO_x transient emission rate is obtained according to Eq. (1):

$$\text{NO}_x = 0.001587 \times \text{NO}_{x\text{conc}} \times G_{\text{exh}} / 3600 \quad (1)$$

In the formula, NO_x is the instantaneous emission mass of gaseous pollutant (g s⁻¹), NO_{xconc} is the instantaneous wet-based concentration of each gaseous pollutant in the original exhaust gas (ppm), and G_{exh} is the instantaneous exhaust flow rate (kg h⁻¹).

The instantaneous pollutant emissions of the corresponding road sections were accumulated to obtain the total pollutant emissions of each section.

2.4 Introduction to Tensor Models

Tensors, also known as *N-mode matrices*, are higher-order generalizations of scalars, vectors, and matrices (Mislevy, 1991; Ali and Foroosh, 2016; Günther *et al.*, 2017). In the sense of isomorphism, a zero-order tensor ($r = 0$) is a scalar, a first-order tensor ($r = 1$) is a vector, and a second-order tensor ($r = 2$) is a matrix. $A \in \mathbb{R}^{l_1, l_2, \dots, l_n}$ can be thought of as a tensor of order N .

The NO_x data used in this experiment included 17 platforms and 16 regional sections of the whole route, starting from Kunming station and ending at the western passenger station. The total length of the study route was about 10.8 km, and the total time was 20 weeks of real-time diesel bus emissions data. According to the real-time emission characteristics, a five-dimensional tensor form of spatial/time (59,400) × day (7) × emission (2) × position (16) × week (20) was constructed. There were 59,400 dimensions per day, representing the NO_x emission data read second by second on that day; there were 7 dimensions in a day, representing the data of the week, and there were 20 dimensions in a week, representing the tensor model of a 20-week pattern. There were 16 location dimensions, which meant that there were 17 bus stations and 16 regional sections, and there were 2 emission dimensions, namely emission concentration and instantaneous emission rate.

In this study, the tensor structure was selected to store data because the data conversion process was convenient. For a large number of data sets, the tensor model can more easily extract, process, store, and transform data, which is conducive to saving time and improving efficiency. For example, Tensor (960, 4, 1, 3, 12) represents the concentration of NO_x emissions in the database

on the fourth day of Week 12 at 06:16 in Position 3, as shown in Fig. 1. From the above description, the NO_x emission concentration and instantaneous emission rate at each corresponding point in time and space can be accurately obtained. This study used the tensor model to store NO_x emission information. The required data set can be constructed quickly, and the corresponding relationships of NO_x emission characteristics in space and time can be revealed.

2.5 Missing Data Processing

Due to various reasons such as damage to experimental equipment, irregular data entry, or limited data acquisition equipment capacity, the missing data phenomenon often occurred in real databases (Sidiropoulos *et al.*, 2017). Based on the analysis of urban bus exhaust emissions by big data, this paper abandoned the traditional missing data-filling algorithm (Acar *et al.*, 2010) to use another based on the Bayesian network and probabilistic reasoning, which is suitable for a large number of data sets. Unlike the commonly used Bayesian network construction algorithm, this study started from the mining of attribute correlations to build the network for the specific application premise of missing-value filling and filled in the data on the basis of Map-Reduce.

A Bayesian network is a graphical probabilistic network based on probabilistic reasoning, and the Bayesian formula is the basis of the probabilistic network. A Bayesian network is represented by a direct acyclic graph, where each node represents a variable and each directed edge between two points represents the casual relationship between two variables. Using the various relations between the Bayesian network variables, the joint probability distribution of these variables is decomposed into several independent conditional probability distributions. The formula is shown as Eq. (2):

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i / \pi(x_i)), \quad (2)$$

where $\pi(x_i)$ is the parent node set of the variable, and when $\pi(x_i) = \emptyset$, $p[x_i / \pi(x_i)] = p(x_i)$. The independence of this condition in the Bayesian network reduces the complexity of the model and brings convenience to uncertainty reasoning with high-dimensional data.

In the data used in this study, the attributes were highly correlated. When the Bayesian network was constructed, the choice was made to explore from the perspective of attribute correlation rather than from the casual relationship between variables, which provided a large amount of available effective inference information and greatly improved the credibility of the reasoning probability. For the application type in this study, it was only necessary to infer the missing data according to the known relevant information, and only the related variables that have an impact on the missing variables need to be noted. The experimental data came from the actual traffic data sets *Date1* and *Date2* of the test center. To test the performance of the

Fig. 1. A tensor model for storing emission information.

algorithm in a big-data environment, *Date1* and *Date2* were synthesized according to the correlation coefficient, and Bayesian networks with 19 edges and 12 nodes and with 25 edges and 17 nodes were generated respectively to generate test data. Table 1 gives a partial raw data set in which data with the * symbol represent data to be extracted by generating a missing data element, and Table 2 represents the corresponding attribute of the actual missing data.

Probabilistic reasoning is the process of deducing unknown information from information with probabilistic properties. Probabilistic reasoning in Bayesian networks deals with two types of problems: the posterior probability and the maximum-posterior hypothesis problem. The former is used to calculate the posterior probability distribution of the unknown variables, whereas the latter is used to solve the combination of the values of the unknown variables to maximize the posterior probability. It is formally defined as a given evidence variable $E = e$, and the value of the query variable q^* satisfying Eq. (3) can be found as follows:

$$q^* = \operatorname{argmax} P(Q = q / E = e). \quad (3)$$

Q is successively equal to each missing variable, and its possible value can be inferred from known information.

When using large amounts of real-time data, serial processing is very slow, and therefore parallel computation is often used to calculate the inference probabilities of the data (Jin, 2013; Jin *et al.*, 2013). The whole probabilistic inference process was divided into three parts: *Para Collect*, *Para Estimate*, and *Impute*, which served to collect parameters, to estimate parameters, and to fill in missing values, respectively, and which were executed sequentially.

After this, the corresponding probabilistic inference was made in the real missing data set. The three attributes of *concentration*, *exhaust*, and *quality* were selected, the Batch Normalization (BN) algorithm was executed on a parallel computer, and the obtained experimental results were compared (Farhangfar *et al.*, 2007; Jin, 2013). Table 3 shows the test results. The data set used in this study was very large and would have been inconvenient to display and operate; therefore, small pieces of real data were used for comparison, but the algorithm was executed according to the original data set. The large real-time data of pollutants leads to a large number of nodes, but it cannot offset the advantage of running time obtained by the parent node and the child node in the Bayesian network, so as to improve the computing efficiency.

Next, the NO_x correlation coefficients were analyzed by the data-filling methods, which were also used to fill in data for concentration, exhaust flow, and quality in the calculation process. In Table 3, *pt/pf* represents the average probability of “true” and “false” values of the three variables. Table 3 shows that the algorithm can effectively distinguish the “true” value from other low-probability values and obtain a probability closer to the actual truth value; *Top* represents

Table 1. Partial raw data sets.

0 Velocity	1 Direction	2 Speed	3 Fuel	4 Load	5 Concentration	6 Exhaust	7 Quality
1.6	58	648	13	19	563	41.605	0.037137*
3	56	650	13	19	576*	41.605	0.037995
2.5	58	651	13	19	720	43.175*	0.049286
1.6	58	651	13	19	790	40.82*	0.051128
3	57	651	13	18	837	39.25	0.052086
3	56	648	14	19	837	42.39	0.056253*
2	55	650	14	19	836*	43.96	0.058267

Table 2. Experimental data set.

Data set	Record (10,000 ⁻¹)	Attr	Size (GB)
Date1	429.4	14	0.624
Date2	427	14	0.569
Test1	96,482.6	12	112
Test2	84,567.9	17	96

Table 3. Experimental results.

Data set		pt/pf	Top (%)
Date1	Concentration	0.94/0.32	96
	Exhaust	0.89/0.27	90
	Quality	0.92/0.22	93
Date2	Concentration	0.86/0.14	92
	Exhaust	0.76/0.09	85
	Quality	0.79/0.11	84

the proportion of the maximum probability of the truth value among all the alternative values among all the inferred filling values obtained by calculation. This discussion makes it clear that the reliability of this algorithm is represented by the probability value obtained by probabilistic reasoning. The higher the probability value, the closer the truth value is to the actual value, and the greater is the accuracy of the data filling.

2.6 Spatial Autocorrelation

Spatial autocorrelation refers to the potential interdependence between observed data points of some variables within the same distribution region (Martin, 1996). The spatial autocorrelation coefficient is used to measure the spatial distribution characteristics of physical or ecological variables and their influence on the field. It can also be used to characterize the spatial distribution characteristics of attributes. Urban bus exhaust emissions have obvious spatial distribution characteristics, and therefore the spatial autocorrelation analysis method is very suitable for analyzing the spatial distribution characteristics of these emissions. This study used the ArcGIS software to perform spatial autocorrelation analysis of pollutant properties.

Global spatial autocorrelation provides a description of the spatial characteristics of the observed values of variables over the whole study area. This analysis was performed to test whether the dependence degree of independent variables over the whole region has an aggregation effect. Global spatial autocorrelation statistical methods include Moran's *I*, Geary's *C*, and the Getis-Ord *G*. Moran's *I* index is the most commonly used and is calculated as in Eq. (4):

$$I = \frac{n}{\sum_{i=1}^n (y_i - \bar{y})^2} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n \sum_{j=1}^n \omega_{ij}} = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n \omega_{ij}} \frac{Y^T W Y}{Y^T Y} \quad (4)$$

where *I* is the global spatial autocorrelation index; *y* represents the observed value of a particular spatial unit; $\bar{y}$ is the average value of this attribute value; *n* is the number of regional units; and ω_{ij} is the spatial weighting coefficient matrix, which represents the neighborhood relationship of each spatial unit.

Local spatial autocorrelation refers to whether there is a correlation between the observed values in each spatial unit in the study area and its surrounding units. If there is no such property, the local spatial autocorrelation position that may be hidden will be discussed. If such a property exists, the spatial heterogeneity will be analyzed. The global Moran's *I* coefficient can reveal the aggregation degree of similar spatial attributes, thus indicating the degree of spatial correlation; however, it cannot identify the exact spatial location of an aggregation area. Local spatial autocorrelation can compensate for this deficiency (Zhao *et al.*, 2009). The calculation of local spatial autocorrelation includes G statistics and LISA; in this study, local Moran's *I* (LISA) was used and can be calculated as follows:

$$I_i = \left(\frac{y_i - \bar{y}}{\sigma} \right) \sum_{j=1}^n \omega_{ij} \left(\frac{y_j - \bar{y}}{\sigma} \right) \quad (5)$$

where *I_i* is the local autocorrelation coefficient of region *i* and ω_{ij} is the spatial weighting matrix.

The variation range of the global Moran's I index is between $(-1, 1)$. When Moran's $I > 0$, the study area presents positive spatial autocorrelation, and the attribute values present a clustered spatial distribution. As the Moran's I index becomes closer to 1, the positive correlation becomes stronger. When Moran's $I < 0$, the study area presents negative spatial autocorrelation, and the attribute values present a dispersed distribution pattern. As the Moran's I index becomes closer to -1 , the negative correlation becomes stronger. A high value of local Moran's I indicates that regional units with similar variable values are clustered in space, whereas a low value indicates that dissimilar regional units are adjacent in space. When Moran's index is close to 0, there is no spatial autocorrelation between attribute values, and they are randomly distributed in space (Wang *et al.*, 2014).

3 RESULTS AND DISCUSSION

3.1 Temporal Distribution Characteristics

Figs. 2 and 3 show the NO_x emissions of the No. 80 diesel bus in Kunming during the course of one week. Fig. 2 shows the all-day emissions characteristics of NO_x on weekdays and Fig. 3 the characteristics on non-weekdays. Figs. 2 and 3 illustrate that there is no obvious regular emission characteristic discriminating between weekdays and non-weekdays. The pollutants show similar changes over time, that is, there are obvious distribution of high and low peaks in a given period. The peak hours were mainly around 7, 11, 17, and 21 and corresponded to the peak hours of morning, afternoon, and evening traffic flow. The overall emission levels on weekdays were higher than on non-weekdays, and this significant difference was consistent with overall traffic flow in cities on different days and the daily travel of pedestrians.

The time emission characteristics of NO_x have obvious high and low peaks, which are similar to the conclusions of Tang *et al.* (2018) on the time distribution characteristics of pollutants emitted by motor vehicles in Hangzhou based on real emission data from the motor vehicle emissions management database. Because this article described the daily time variation characteristics from Monday to Sunday, compared with Tang *et al.*'s emission inventory based on annual and monthly changes, it can describe the emission characteristics of pollutants at different times of the day in more detail. In the course of the day, the changes in high and low peaks were more obvious in each time period, and the pollutant emission amounts in each time period were significantly different from those in other time periods. Therefore, this paper better reflects the temporal emission characteristics of pollutants.

3.2 Trends in Emissions during Working Days and Non-working Days

As shown in Fig. 4, NO_x emissions on weekdays were higher than on non-weekdays as a whole, with the average bus speed on weekdays being approximately 26 km h^{-1} and that on non-weekdays being approximately 30 km h^{-1} . In the four peak areas, the rate of change of NO_x on weekdays was higher than on non-weekdays; the reason for this lies in the fact that passenger flow is generally higher in the peak period on weekdays, and to increase passenger volume, ease traffic, and facilitate pedestrians, the departure frequency increase markedly in these time slots, leading to a significant increase in pollutant discharge in this period. Because there was no significant increase in human and vehicle flow on non-working days compared with that on working days, the dispatching frequency was relatively uniform. During the time periods of 16:00 and 20:00, the non-working-day emissions exceeded the working-day emissions; the former is due to the increase of NO_x emission due to the increase of bus departure frequency, and the latter is to meet the increase of non-working-day pedestrian travel volume, in line with the pedestrian travel law.

There is an obvious "weekend effect" in the concentration of NO_x in Kunming. The concentration of NO_x on weekends is lower than that on weekdays. The research results are similar to the observation results of Huang (2011) in Shanghai and similar to the study results of Wu (2013) on the temporal-spatial distribution of NO_x in Xi'an. NO_x emissions in Xi'an were similar to those reported in this paper, and the peak alternation was obvious in different time periods according to the daily variation. The weekend effect of NO_x in Kunming was closely related to the difference in human work and rest arrangements between weekends and weekdays (Tang *et al.*, 2010).

Fig. 2. Emission characteristics of NO_x pollutants from 06:00 to 22:00 on weekdays of urban diesel buses: (a) Monday, (b) Tuesday, (c) Wednesday, (d) Thursday, (e) Friday.

Fig. 3. Emission characteristics of non-workday pollutant NO_x from 06:00 to 22:00 on urban diesel buses (a. Saturday, b. Sunday).

Fig. 4. Comparison of NO_x emissions of diesel buses in working days and non-working days.

Table 4. Spatial autocorrelation analysis.

	Moran / index	Expected index ($\times 10^{-4}$)	Variance ($\times 10^{-5}$)	Z score	Distance threshold ($\times 10^{-4}$)
Morning	0.648	2.86	0.2	322.614	2.19
Midday	0.836	2.62	1.2	532.456	2.16
Afternoon	0.935	2.46	1.2	613.883	2.16
Evening	0.798	2.86	1.2	478.174	3.00

3.3 Spatial Emission Characteristics in Different Time Periods

The NO_x autocorrelation coefficient for real-time operation of the Kunming No. 80 bus was calculated by the ArcGIS software, and its spatial aggregation and heterogeneity were also analyzed. The results were obtained from multiple random trials with the ArcGIS software, and the mean value was obtained, as shown in Table 4. The spatial autocorrelation coefficients of the four time periods were all positive and high, indicating that they had obvious aggregation characteristics in space; in other words, adjacent exhaust emissions concentrations were the same, showing an obvious spatial high-low aggregation pattern.

GIS spatial autocorrelation analysis showed spatial clustering or heterogeneity of points with higher or lower exhaust emissions, showing positive/negative global spatial autocorrelation, that is, the results of a comparison between the emission values of a spatial area and its surrounding areas. By combining these two analysis methods, the heterogeneity of NO_x spatial distribution in four time periods (the morning hours are from 6 to 8, the noon hour is from 11 to 13, the afternoon time is from 16 to 18, and the evening session is from 20 to 22) located along the whole operation route was determined. Then the emission changes for different sections in the same time period were compared, and a brief analysis was performed on the influencing factors of emission changes for different sections in different time periods along the whole route. Fig. 5 shows a thermal diagram of the whole route at different time periods.

The platform, intersection, and traffic-light positions along the whole route are marked in Fig. 5(a). The spatiotemporal emission characteristics of the whole route were comprehensively analyzed by taking into account the main factors affecting pollutant emissions. Figs. 5(a), 5(b), 5(c), and 5(d) present heat maps of NO_x emission concentrations and show that there were significant differences in NO_x emissions at different time periods along the whole bus route during one day, according to the color intensity of the movement routes in the four time periods. The overall emission-level variation in the different time periods was obvious, which conformed to normal urban traffic flow and the actual traffic situation. In addition, different sections of the whole route during the same time period indicated an obvious emission color, and the whole driving route was affected by platforms, traffic lights, and intersections as shown in Fig. 5(a), resulting in obvious changes in emission levels along different sections of the whole route at the same time period.

According to the combined results of the four time periods in Fig. 5 and the spatial autocorrelation analysis in Table 4, the exhaust emissions at different time periods were significantly different, and the exhaust emissions at different sections during the same time period were obviously different, indicating an obvious time/space variation in bus exhaust emissions. According to the analysis results in Table 4, Moran's *I* coefficients were all positive, which indicated that the urban bus exhaust emissions presented clear spatial distribution characteristics in their particular road sections. Moran's *I* coefficient for NO_x was lower in the morning and at night, indicating that unit spaces of dissimilar regions were adjacent. In other words, the distribution of sample points was dispersed and had some spatial heterogeneity, and the emissions concentration was low. Moran's *I* coefficient for NO_x was higher at noon and in the afternoon, indicating that its sample point distribution was relatively clustered and its spatial clustering is relatively obvious, especially in the afternoon. The local spatial autocorrelation coefficient was as high as 0.935, indicating that regional units with similar variables were spatially clustered. The main reason for this situation was that the No. 80 bus runs at idle speed and low speed for a long time during the evening rush hour. The starting point is located at the railway station and the passenger station is near the terminal. At rush hour in these areas, the flow of people and cars peaks for the day, and the emissions reach their daily maximum.

The results of pollutant autocorrelation analysis are consistent with the spatial distribution of pollutants obtained by Li and Zhang (2011), which had obvious spatial clustering characteristics. The results of the thermal map analysis are similar to the results for the thermal map of the Beijing motor vehicle emissions network made by Guo *et al.* (2017). High emission intensity and high emissions occur at road sections (platforms, traffic lights, and intersections) where abnormal values of traffic flow appear, and the emissions change significantly in each section.

3.4 Emission Intensity of Each Section

Fig. 6 shows the distribution characteristics of NO_x on the corresponding sections of the entire bus route at different times. Fig. 6 shows that NO_x emissions are significantly different in different sections at different times. The emissions from Sections 1–3 and 5–8 at different periods were significantly higher than those of other sections at the same time because the emissions of Sections 1–3, including the starting-point railway station, a section with all-day traffic congestion, and sections with multiple intersections and traffic lights, were significantly higher than those from other sections. The four Sections 5–8 pass through a dense residential area, although there were no heavily congested intersections or intersections with traffic lights. Due to this environment, driving conditions became more complicated, the bus was always moving at low speed and idle speed, and frequent deceleration to yield to pedestrians and frequent acceleration to start and

Fig. 5. Heat map of urban diesel buses in the morning, noon, afternoon and evening.

Fig. 6. NO_x emission changes of No. 80 diesel bus in different time periods and different sections of the route.

Fig. 7. Diagram of NO_x emission intensity of No. 80 diesel bus in different time periods and different sections of the route.

leave from a stopped condition increased significantly. Because of these conditions, the emissions from this section of road were high throughout the day, which was consistent with the actual traffic conditions, and therefore it was a high-emission section.

Exhaust emission intensity reveals the impact of pollutants on surrounding pedestrians and residents. The greater the exhaust emission intensity, the greater is the harm that it will do to surrounding residents and pedestrians. Fig. 7 shows the time-space variation characteristics of NO_x emission intensity at different sections during different time periods. NO_x emissions on roads containing major intersections and adjoining dense residential areas were 1.8–3.2 times higher than average, and the NO_x emission intensity of the road section containing the main intersections was 0.5–1.2 times the average value, whereas the NO_x emission intensity of the road section containing dense residential areas was 2–3.8 times the average. During the peak period, NO_x emissions were 1.5–2.8 times the average, and the NO_x emissions intensity was 1.1–1.8 times the average. Combining the effects of NO_x emission intensity, emissions, and time, it was found that NO_x emissions in dense residential areas and those containing major intersections increase the hazard to pedestrians by 3–8% (dense residential areas) and by 15–33% (road intersections).

Fig. 7 shows that the highest intensity of NO_x emissions during the day occurred in the afternoon, followed by noon, then evening, and finally morning, which was consistent with the spatiotemporal emission characteristics of the whole road section as described above. In Fig. 7, the rates of change of emission intensity in Sections 6–8 at each period were 2.56, 5.31, and 3.04 kg km⁻², respectively, representing high emission intensity. Because these five roads are located in complex and dense residential areas, platforms have been built relatively close together to facilitate travel by surrounding residents, and as a result, the length of each road section is relatively short, being 390, 301, and 296 m respectively. Therefore, the emission intensities of Road Sections 6–8 are 1.1, 2.7, and 1.9 times higher than average, respectively, mainly because short road sections have been set up for the convenience of surrounding residents. NO_x emissions in each time period were similar to those reported by Li *et al.* (2017) in their study on vehicle emissions inventory in Foshan. In the morning, at noon, and in the afternoon, the pollutant discharge rate was high, and the pollutants reached their daily peak around 17:00. However, the difference in this study was the identification of an evening pollutant discharge peak instead of a slow decline from 17:00 onwards.

4 CONCLUSIONS

This study investigated the spatial and temporal distribution of NO_x emissions exhausted by diesel buses by applying spatial autocorrelation to emission data—specifically, heat maps based on the temporal variation, intensity, and quantity of the emissions—for the Kunming No. 80 bus. The following conclusions can be drawn:

- (1) In this paper, Bayesian network and probabilistic reasoning are used to fill the data of concentration, exhaust and quality. The results show that the true probability of Date1 is 0.94, 0.89, and 0.92, respectively, and the true probability of Date2 is 0.86, 0.76, and 0.79, respectively. By selecting true values with high probabilities—96%, 90%, and 93% for the first set and 92%, 85%, and 84% for the second—we reduced errors in filling the data, thereby drastically enhancing the accuracy of our calculations.
- (2) The concentrations of the NO_x emissions exhibited obvious peaks and troughs throughout the day, and the quantity of discharged pollutants clearly varied between high- and low-emission periods.
- (3) The heat maps for the afternoon and evening revealed higher pollutive emissions, and during these periods, all of the nodes and roads displayed a high-emission trend. However, during the morning and evening, the pollutants displayed a lighter color and sparser distribution, and the emissions were lower and less varied over the entire route. The spatial autocorrelation indices for the morning, midday, afternoon, and evening—0.648, 0.836, 0.935, and 0.798, respectively—indicated spatial correlation between the four periods as well as a degree of heterogeneity during the morning and evening for the NO_x, although the aggregation of this species was difficult to observe owing to the time constraints.
- (4) Sections 1–3 and 6–9 produced large emissions during every period and contributed to the majority of the pollutants measured along the entire route. The highest and fastest-growing emission intensity was found on the road segment comprising Sections 5–9, though.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from National Natural Science Foundation of China (51968065), Applied Basic Research Project of Yunnan Province [2017FG001(–010), 2017FG001(–093)] and Major Science and Technology Projects in Yunnan Province (2018ZE001).

DISCLAIMER

Reference to any companies or specific commercial products does not constitute an endorsement, recommendation or favoring by the company.

REFERENCES

- Acar, E., Dunlavy, D.M., Kolda, T.G. (2009). Link Prediction on Evolving Data Using Matrix and Tensor Factorizations. 2009 IEEE International Conference on Data Mining Workshops, Presented at the 2009 IEEE International Conference on Data Mining Workshops (ICDMW), IEEE, Miami, FL, USA, pp. 262–269. <https://doi.org/10.1109/ICDMW.2009.54>
- Alam, A., Hatzopoulou, M. (2017). Modeling transit bus emissions using MOVES : Comparison of default distributions and embedded drive cycles with local data. J. Transp. Eng. Part A. Syst. 143, 04017049. <https://doi.org/10.1061/JTEPBS.0000067>
- Ali, M., Foroosh, H. (2016). Character recognition in natural scene images using rank-1 tensor decomposition. IEEE Intern Confer Image Proc. <https://doi.org/10.1109/ICIP.2016.7532888>
- Apte, J.S., Marshall, J.D., Cohen, A.J., Brauer, M. (2015). Addressing global mortality from ambient PM_{2.5}. Environ. Sci. Technol. 49, 8057–8066. <https://doi.org/10.1021/acs.est.5b01236>
- Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., Hamburg, S.P. (2017). High-resolution air pollution mapping with google street view cars: Exploiting big data. Environ. Sci. Technol. 51, 6999–7008. <https://doi.org/10.1021/acs.est.7b00891>
- Bogacki, M., Bódziuch, P. (2019). Urban bus emission trends in the Krakow metropolitan area (Poland) from 2010 to 2015. Transp. Res. Part D 67, 33–50. <https://doi.org/10.1016/j.trd.2018.11.003>
- Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R.V., Dentener, F., Dingenen, R. van, Estep, K., Amini, H., Apte, J.S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P.K., Knibbs, L.D., Kokubo, Y., Liu, Y., Ma, S., et al. (2016). Ambient air pollution exposure estimation for the global burden of disease 2013. Environ. Sci. Technol. 50, 79–88. <https://doi.org/10.1021/acs.est.5b03709>
- Carslaw, D.C., Farren, N.J., Vaughan, A.R., Drysdale, W.S., Young, S., Lee, J.D. (2019). The diminishing importance of nitrogen dioxide emissions from road vehicle exhaust. Atmos Environ. 1, 100002. <https://doi.org/10.1016/j.aeaqa.2018.100002>
- Carvalho, H. (2016). The air we breathe: Differentials in global air quality monitoring. Lancet Respir. Med. 4, 603–605. [https://doi.org/10.1016/S2213-2600\(16\)30180-1](https://doi.org/10.1016/S2213-2600(16)30180-1)
- Chai, F.H., Chen, Y.Z., Wen, Y., Duan, L., Xie, S.D., Xue, Z.G., Cao, D., Liu, J., Song, G.J., Wang, S.L. (2006). Regional total air pollutant control technology and demonstration research. Environ. Sci. Res. 19, 163–171. (in Chinese with English Abstract)
- Chen, J.H., Fan, W.B., Qian, J., Li, Y., Zhao, W.Y. (2015). Establishment of the light-duty gasoline vehicle emission inventory in Chengdu by the International Vehicle Emission model. Acta Sci. Circum. 35, 2016–2024. (in Chinese with English Abstract) <https://doi.org/10.13671/j.hjkxxb.2014.1075>
- Chen, L., Wang, Z., Liu, S., Qu, L. (2018). Using a chassis dynamometer to determine the influencing factors for the emissions of Euro VI vehicles. Transp. Res. Part D 65, 564–73. <https://doi.org/10.1016/j.trd.2018.09.022>
- Fan, S.B., Tian, L.D., Zhang, D.X., Qu, S. (2015). Beijing motor vehicle emission characteristics based on actual road traffic flow information. Environ. Sci. 36, 2750–2757. (in Chinese with English Abstract) <https://doi.org/10.13227/j.hjxx.2015.08.004>
- Farhangfar, A., Kurgan, L.A., Pedrycz, W. (2007). A novel framework for imputation of missing values in databases. IEEE Trans. Syst. Man Cybern. Part A Syst. Humans 37, 692–709. <https://doi.org/10.1109/TSMCA.2007.902631>
- Günther, R., Wenzel, T., Wegner, M., Rettig, R. (2017). Big data driven dynamic driving cycle development for busses in urban public transportation. Transp. Res. Part D 51, 276–89. <https://doi.org/10.1016/j.trd.2017.01.009>
- Guo, J.J., Fan, S.B., Wang, J.L., Tian, L.D., Zhang, D.X., Sun, G.H. (2017). Spatial Clustering Characteristics of Vehicle Emissions in Beijing. Environ. Sci. Technol. 40, 141–146. (in Chinese with English Abstract)
- Hao, L., Chen, W., Li, L., Tan, J., Wang, X., Yin, H., Ding, Y., Ge, Y. (2017). Modeling and predicting low-speed vehicle emissions as a function of driving kinematics. J. Environ. Sci. 55, 109–117. <https://doi.org/10.1016/j.jes.2016.05.037>

- He, J., Wu, L., Mao, H., Liu, H., Jing, B., Yu, Y., Ren, P., Feng, C., Liu, X. (2016). Development of a vehicle emission inventory with high temporal–spatial resolution based on NRT traffic data and its impact on air pollution in Beijing – Part 2: Impact of vehicle emission on urban air quality. *Atmos. Chem. Phys.* 16, 3171–3184. <https://doi.org/10.5194/acp-16-3171-2016>
- Huang, Z.X. (2011). Observation of spring ozone and its precursors in Shanghai suburbs. *Environ. Sci. Technol.* 34, 87–89+99. (in Chinese with English Abstract) <https://doi.org/10.3969/j.issn.1003-6504.2011.05.020>
- Jin, L. (2013). Key Technology Research on Missing Value Filling in Incomplete Data. Master's Degree, Harbin Instit Technol. <https://doi.org/CNKI:CDMD:2.1014.001204>
- Jin, L., Wang, H.Z., Huang, S.B., Gao, H. (2013). Big data missing value filling algorithm based on Map-Reduce. *J. Comput. Res. Dev.* 50, 312–321. (in Chinese)
- Kan, Z., Tang, L., Kwan, M.P., Zhang, X. (2018). Estimating vehicle fuel consumption and emissions using GPS big data. *Int. J. Environ. Res. Public Health* 15, 566. <https://doi.org/10.3390/ijerph15040566>
- Li, C.X., Zhang, Z.H. (2011). Spatial statistics analysis of regional environmental pollution In China. *Energy Procedia* 5, 163–168. <https://doi.org/10.1016/j.egypro.2011.03.029>
- Li, L.J., Wang, Y., Li, J.X., Xin, L.Z., Jin, J. (2012). Study on atmospheric heavy pollution in Beijing from 2000 to 2010. *China Environ. Sci.* 32, 23–30. (in Chinese) <https://doi.org/10.3969/j.issn.1000-6923.2012.01.004>
- Lim, S.S., Vos, T., Flaxman, A.D. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380, 2224–2260. [https://doi.org/10.1016/S0140-6736\(12\)61766-8](https://doi.org/10.1016/S0140-6736(12)61766-8)
- Liu, H., Wu, B., Liu, S., Shao, P., Liu, X., Zhu, C., Wang, Y., Wu, Y., Xue, Y., Gao, J., Hao, Y., Tian, H. (2018). A regional high-resolution emission inventory of primary air pollutants in 2012 for Beijing and the surrounding five provinces of North China. *Atmos. Environ.* 181, 20–33. <https://doi.org/10.1016/j.atmosenv.2018.03.013>
- Liu, J., Ge, Y., Wang, X., Hao, L., Tan, J., Peng, Z., Zhang, C., Gong, H., Huang, Y. (2017). On-board measurement of particle numbers and their size distribution from a light-duty diesel vehicle: Influences of VSP and altitude. *J. Environ. Sci.* 57, 238–248. <https://doi.org/10.1016/j.jes.2016.11.023>
- Liu, Y.H., Ma, J.L., Li, L., Lin, X.F., Xu, W.J., Ding, H. (2018). A high temporal-spatial vehicle emission inventory based on detailed hourly traffic data in a medium-sized city of China. *Environ. Pollut.* 236, 324–333. <https://doi.org/10.1016/j.envpol.2018.01.068>
- Lv, W., Hu, Y., Li, E., Liu, H., Pan, H., Ji, S., Hayat, T., Alsaedi, A., Ahmad, B. (2019). Evaluation of vehicle emission in Yunnan province from 2003 to 2015. *J. Cleaner Prod.* 207, 814–825. <https://doi.org/10.1016/j.jclepro.2018.09.227>
- Martin, D. (1996). An assessment of surface and zonal models of population. *Int. J. Geog. Inf. Syst.* 10, 973–989. <https://doi.org/10.1080/02693799608902120>
- Mislevy, R.J. (1991). Review of statistical analysis with missing data. *J. Educ. Stat.* 16, 150–155. <https://doi.org/10.2307/1165119>
- Sidiropoulos, N.D., Lathauwer, L.D., Fu, X., Huang, K., Papalexakis, E.E., Faloutsos, C. (2017). Tensor Decomposition for Signal Processing and Machine Learning. *IEEE Trans. Signal Process.* 65, 3551–3582. <https://doi.org/10.1109/TSP.2017.2690524>
- Su, J.G., Apte, J.S., Lipsitt, J., Garcia-Gonzales, D.A., Beckerman, B.S., de Nazelle, A., Texcalac-Sangrador, J.L., Jerrett, M. (2015). Populations potentially exposed to traffic-related air pollution in seven world cities. *Environ. Int.* 78, 82–89. <https://doi.org/10.1016/j.envint.2014.12.007>
- Tang, G.Q., Li, X., Wang, X.K., Xin, J.Y., Hu, B., Wang, L.L., Ren, Y.F., Wang, Y.S. (2010). Effects of synoptic type on surface ozone pollution in Beijing. *Huan Jing Ke Xue* 31, 573–578. (in Chinese with English Abstract)
- Tang, W., Yang, Q., Huang, C., Lu, B., Xia, Y., Jing, B. L., Lu, Q., Lu, J. (2018). Study on characteristics of pollutant emission from motor vehicles in Hangzhou based on large data analysis and IVE model. *Acta Sci. Circum.* 38, 71–78. (in Chinese with English Abstract) <https://doi.org/10.13671/j.hjkxxb.2017.0241>
- Wang, J.Q., Yang, J.J., Gao, J.D., Ma, J., Liu, S.X., Gao, H.Y. (2011). Progress in diesel vehicle exhaust

- emission control technology. *Sci. Technol. Guide* 11, 67–75. (in Chinese)
- Wang, L., Li, T.Z. (2016). Spatial autocorrelation analysis of urban bus emission. *J. Transp. Eng. Inf.* 14, 95–100. (in Chinese with English Abstract) <https://doi.org/10.3969/j.issn.1672-4747.2016.01.016>
- Wang, W.X., Kuang, Y.Q., Huang, L.S. (2014). Spatial autocorrelation analysis on energy-related carbon emission in Guangdong. *Environ. Sci. Technol.* 37, 180–87+193. (in Chinese with English Abstract)
- Wang, X.L., Liu, H.Y., Li, T.H., Deng, K. (2012). The emission model of diesel bus based on CMEM in China. *J. Harbin Inst. Technol.* 44, 78–81. (in Chinese with English Abstract) <https://doi.org/10.11918/j.issn.0367-6234.2012.06.018>
- Wu, S.G. (2013). Study on NO_x Pollution Characteristics and Exposure Dose Evaluation of Typical Urban Traffic Micro-environment in Xi'an City. Northwest University.
- Zhang, J.W., Xian, Y., Chen, B.Z. (2017). Study on the emission inventory from mobile source in Nanchang. *Acta Sci. Circum.* 37, 2449–2458. (in Chinese with English Abstract) <https://doi.org/10.13671/j.hjkxxb.2017.0022>
- Zhang, K., Batterman, S. (2013). Air pollution and health risks due to vehicle traffic. *Sci. Total Environ.* 450–451, 307–316. <https://doi.org/10.1016/j.scitotenv.2013.01.074>
- Zhang, L., Hu, X., Qiu, R., Lin, J. (2019). Comparison of real-world emissions of LDGVs of different vehicle emission standards on both mountainous and level roads in China. *Transp. Res. Part D* 69, 24–39. <https://doi.org/10.1016/j.trd.2019.01.020>
- Zhao, X.F., Huang, X.J., Zhang, X.Y., Zhu, D.M., Lai, L., Zhong, T.Y. (2009). Application of spatial autocorrelation analysis to the COD, SO₂ and TSP emission in Jiangsu province. *Chin. J. Environ. Sci.* 30, 1580–1587. (in Chinese with English Abstract) <https://doi.org/10.3321/j.issn:0250-3301.2009.06.004>