Heliyon 10 (2024) e31359

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Coking exhaust contributes to airborne particulate matter in the Beijing–Tianjin–Hebei region

Xiaoming Wan [a,b,*], Weibin Zeng [a,b], Gaoquan Gu [a,b]

[a] Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences, Beijing, 100101, China
[b] University of Chinese Academy of Sciences, Beijing, 100089, China

ARTICLE INFO

Keywords:
Air quality
Land use
Model validation
Prediction
Spatial distribution

ABSTRACT

Coking was regarded as a predominant source of air pollution. Despite the adoption of more environmentally friendly equipment, whether the coking enterprises in the Beijing–Tianjin–Hebei (BTH) region are still causing regional air pollution is worthy of study, which is essential for the control of coking enterprises in this area. To improve the prediction accuracy of large-scale air pollutant distribution, the air particle distribution in the BTH region was simulated via land use regression (LUR) combined with Bayesian maximum entropy (BME); then, the distribution was correlated with the exhaust gas emitted from coking enterprises. Results indicated that the R^2 of the “LUR + BME” method reached 0.95, higher than 0.82 using LUR alone. The air quality distribution presented a pattern of “low in the northern mountains and high in the southern plains”, similar to the distribution of coking enterprises in BTH region. A significant correlation was found between exhaust emissions from coking enterprises and air quality in the BTH region, confirming the contribution of coking emissions to air pollution in this region, and the necessity to continue the strict control on coking enterprises in BTH area.

1. Introduction

Coking plants are prone to producing pollutants such as heavy metals, which introduce pollution risks to the air and health risks to the human body [1,2]. China is the largest coke producer in the world [3]. Coke production in the Beijing–Tianjin–Hebei (BTH) region accounts for about 12 % of the total coke production in China [4]. The coke pollution in the BTH region is regarded as a relatively serious environmental problem in this region [5]. As a result of increasingly stringent environmental policies, coking plants have adopted more environmentally friendly equipment [6]. Whether the coking enterprises in the BTH region are still causing air pollution is worthy of study, which is essential for the air quality improvement in this area.

Particulate matter PM_{10} and $PM_{2.5}$ are the main causes of smoggy weather [7], and they are defined as particles with an aerodynamic diameter of less than 10 and 2.5 μm, respectively [5]. Smog in the BTH area is serious. Air pollution in BTH area has received widespread concern [8]. Industrial activities, traffic, and urban construction activities are considered as the most influential factors of urban air quality [9]. Among them, industrial structure is regarded as the decisive factor affecting $PM_{2.5}$ health risks in BTH area [10]. However, despite the general idea that coking contributes to the air pollution in BTH region, there has been limited study analyzing the correlation between coking exhaust emission and air pollutants from the point of view of spatial distribution.

The simulation of air quality from the regional scale using limited amount of data from monitoring stations is a research focus [7].

* Corresponding author. Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences, Beijing, 100101, China.
E-mail address: wanxm.06s@igsnrr.ac.cn (X. Wan).

https://doi.org/10.1016/j.heliyon.2024.e31359
Received 2 January 2024; Received in revised form 10 May 2024; Accepted 15 May 2024
Available online 16 May 2024

Many relevant exposure assessments involving air pollutants are analyzed directly using the data obtained from spatial monitoring stations. However, the monitoring stations set up in most cities are unevenly distributed and limited in number. Since the monitoring stations can hardly cover all areas, the conclusions drawn are usually incomplete, and they are unable to reflect the actual situation of the region [11]. In BTH region, such varied distribution of monitoring stations is obvious. Using limited point data to simulate the status of the entire region helps promote air pollution control [12,13]. How to accurately reflect the spatial distribution of air pollutants using limited data becomes especially important, and thus becoming a hot research topic. The most used methods for exploring the spatial distribution of $PM_{2.5}$ are spatial interpolation, remote sensing inversion, atmospheric diffusion models, land use models, or a combination of several methods [14]. The land use regression (LUR) model is a popular method used to simulate the spatial variations in urban air pollution. Based on the hypothesis that the underlying surface environment and population density affected the spatial distribution of pollutants, using big data and mathematical statistics, the air pollution was correlated with land use type, meteorological factor and population density. Via such correlation, air pollution exposure can be predicted [9,15,16]. The basic idea of LUR model is to firstly collect the air pollutant concentration of dozens of sampling points as the dependent variable, and then the land use, traffic, population density and other data around the site were obtained as the independent variable. Using these two types of variables, a multiple regression model is to be established, predicting the spatial distribution of air pollutant concentration, and identifying the predominant influencing factors [17]. According to the model, the concentration of pollutants in other regions that are not within the monitoring range can be estimated.

Compared with the traditional air pollutant simulation model, the land use regression model has the following advantages: more factors to consider, a wider prediction range, and higher spatial resolutions [18]. The land use regression model plays an important role in simulating the concentration of pollutants on a large scale and applying it to the hazard assessment of human exposure [19,20].

However, in the traditional LUR model, linear multiple regression equations are usually used to build models. However, the relationship between each variable and the target variable is usually complex and non-linear. Such contradictions indicated that the model cannot accurately express the complex realistic relationship between the dependent factor and the independent factor [21,22]. Integrating LUR with other modeling methods may increase its efficiency [23].

Methods like principal component analysis (PCA) has been combined with LUR to increase the simulation efficiency [24]. Bayesian maximum entropy (BME), as a scientific spatiotemporal data analysis method, can be well combined with LUR [25]. The traditional geostatistical methods focused on actual measured data (hard data), while ignoring some uncertain data that are not directly measured (soft data), including historical data, model fitting data, and the experience and knowledge of experts or existing research results. Different from traditional geostatistics, BME can not only build a spatiotemporal covariance model using known spatial points but also combine hard data and soft data to simulate the spatial distribution of unknown points [26,27]. The integration of LUR and BME can obtain a better prediction of large-scale air pollutant distribution [28].

In this study, the air particulate distribution in the BTH region was first simulated via LUR combined with BME. Aiming to confirm the contribution of the coking industry to air pollution in the BTH region, the simulated air quality was then correlated with the exhaust gas emitted by coking enterprises. Results could provide information for regional air pollution control.

2. Materials and methods

2.1. Study area

The BTH region is located between 36°05' - 42 °40′ north latitude and 113°27' - 119°50′ east longitude, including Beijing, Tianjin, and 11 prefecture-level cities in Hebei Province. It covers an area of 218,000 km^2 and accommodates a population of about 110 million. It is located in the northern part of the North China Plain, bordered by the Yanshan Mountains to the north, the North China

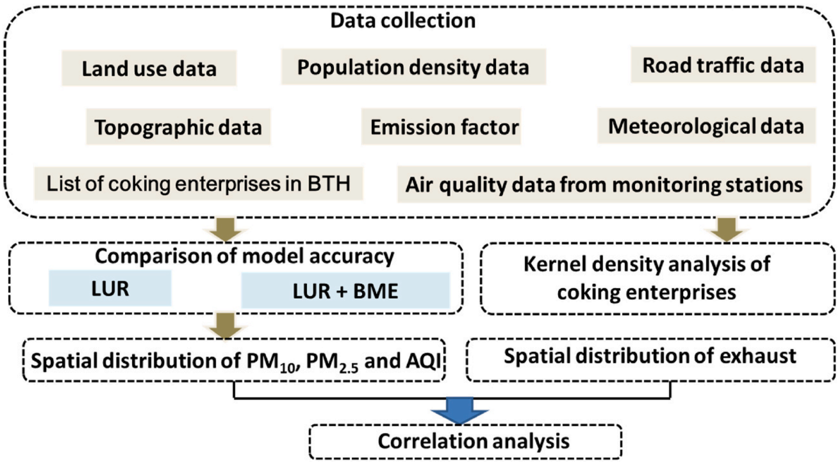

Fig. 1. Research framework.

Plain to the south, the Taihang Mountains to the west, and the Bohai Bay to the east. The terrain is relatively high in the northwest and north areas, and it is relatively flat in the south and east areas. The Yanshan–Taihang Mountain system's structure from the northwest to the southeast gradually transitions into a plain, showing high topographic characteristics in the northwest and low characteristics in the southeast.

The research frame work was described in Fig. 1, mainly including three steps, i.e., data collection, models establishment and the correlation analysis.

2.2. Data collection

The data sources are shown in Table 1. The list of coking enterprises in production was obtained from a list of coking enterprise emission permits and a list of key emission supervision enterprises. Information on the longitude and latitude was extracted according to Baidu map POI. Kernel density analysis is used to calculate the unit density of the measured values of the point and line elements in a specified neighborhood [29]. The kernel density analysis function of ArcGIS is used to analyze the kernel density of coking enterprises in the BTH region, which can directly reflect the distribution of coking enterprises in the BTH region.

The emission factor of coking enterprises is obtained from two databases, i.e., the "Industrial enterprise emission database" and "US-EPA". The pollution emission data of coking enterprises were obtained from the China Industrial Enterprise pollution emission database.

The monthly meteorological data of meteorological stations (102 in BTH area) in the past 30 years, including maximum wind direction (excluding quiet wind) and monthly average wind speed data, were obtained from the National Data Center for Meteorological Sciences. The data on air temperature and precipitation were obtained from the Research Center for Resources and Environmental Data of the Chinese Academy of Sciences. Topographic data were obtained from the United States Geographic Survey.

2.3. Simulate PM_{10} based on LUR

PM_{10} from air monitoring stations in the BTH region were taken as dependent variables, and relevant geographical factors were extracted as independent variables to establish a multiple linear regression model of air quality data in the BTH region.

The LUR model adopted in this study is a multiple linear regression model based on the concentration of air pollutants in the monitoring station and the surrounding variables that may have an impact on air pollutants. The LUR is shown in Eq (1):

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n + \beta \tag{1}$$

where y is the dependent variable, namely the concentration value of PM_{10}; $x_1, x_2, \ldots x_n$ is the independent variable, that is, various influence factors in this study; $\alpha_1, \alpha_2, \ldots \alpha_n$ is the undetermined coefficient; and β is the random variable. Using SPSS 22.0, the variables of the multiple linear regression model were screened via the stepwise regression method, which can effectively avoid multicollinearity.

The independent variables represent the impact of meteorological factors, topographic population, land use types, road traffic, and different meteorological factors on the concentration distribution of pollutants, respectively. With respect to land use type, the variable was expressed using the proportion of land use type in the demarcated buffer zone. With respect to topographic and geomorphic data, it is represented by the DEM data of the location of the enterprise. With respect to the population density data, it is represented by the raster value of the POP data where the enterprise is located. With respect to road traffic data, it is represented by the total length of

Table 1
Data sources.

Data content	Data sources
List of coking enterprises in the Beijing–Tianjin–Hebei region that have obtained production and pollutant discharge permits, the number of enterprises, and the region where the enterprises are located	Tianyan investigation, laboratory research, etc.
Pollution emission data of coking enterprises, such as wastewater, waste gas, smoke, and dust	China Industrial Enterprise pollution emission database
Location of pollution sources: Longitude and latitude	Baidu map POI data crawl
Natural environmental data, including temperature, wind speed, and direction; atmospheric stability; state of the underlying surface	National Meteorological Science Data Center, official websites of local weather bureaus, and statistical yearbooks
Air pollution data, including the AQI, PM2.5, and PM10 data of 80 monitoring stations	National Urban Air Quality Real-time Release Platform
Population spatial distribution data on a 1 km grid	Chinese population spatial distribution kilometer grid dataset. Data registration and publication system of the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (http://www.resdc.cn/DOI), 2017. DOI: 10.12078/2017121101
The distribution of road networks in the region	Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences www.resdc.cn
Land elevation DEM within the study area	United States Geographic Survey (www.usgs.com)
Type of land use in the study area	Resources and Environmental Sciences and Data Center, Chinese Academy of Sciences www.resdc.cn

different types of roads within the delimited buffer. With respect to meteorological element data, the grid values of wind direction, wind speed, relative humidity, and precipitation at the location of the enterprise are expressed. The division of the buffer zone is greater than the safety and health protection distance of the coking plant as the standard; the 2 km scale is selected as the distinction; the 2 km, 4 km, 6 km, and 8 km scales are delimited.

In the process of model building, three rules are followed: (1) a land use variable can have at most one buffer zone in the model to ensure the interpretability of the model; (2) the sign of the variable (±) must be reasonable in order to ensure that the contribution of the variable is meaningful; (3) stepwise regression is used to ensure the simplification of the model while limiting the influence of collinearity among variables. Only variables with significant contributions were retained in the final model (the threshold of variables entering the model was set at $p = 0.05$ and $p = 0.1$ when removed) to prevent overfitting.

2.4. Simulate PM_{10} based on LUR with Bayesian maximum entropy (BME)

The BME model provides a mathematical framework that integrates various kinds of information in a general and unified manner [30]. The simulation of PM_{10} spatial distribution by BME is completed by SEKS-GUI 0.6 package of MATLAB software. The hard data required for BME comprise the PM_{10} data from the air quality monitoring station. According to the results of the LUR model, the

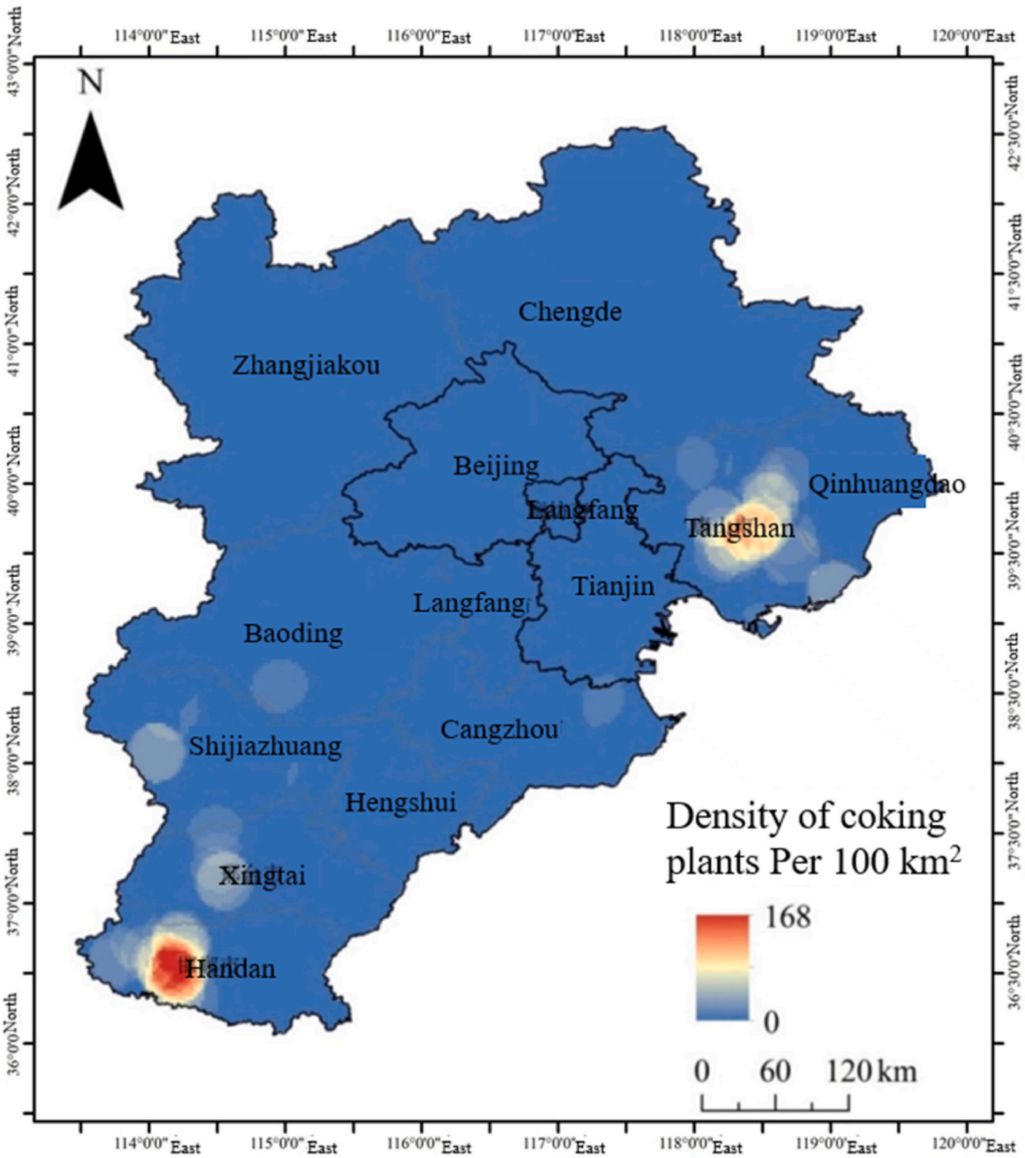

Fig. 2. Kernel density analysis of coking enterprises in the Beijing–Tianjin–Hebei region.

predicted value of particulate matter concentrations is generated, which comprises the soft data required for BME. After importing hard and soft data into SEKS-GUI, data processing is performed. The specific steps include data import, exploratory and covariance analyses of data, and output visual graphics [31].

2.5. Validation of models

SPSS 22.0 was used to evaluate the predictability of the model using the K-fold cross-validation method. Ten-fold cross-validation was performed, which successively left out 10 % of sites from the training dataset, and the model was developed based on the remaining 90 % of sites. The initial sampling was evenly divided into 10 subsamples. A single subsample is retained as the data of the validation model. The other 9 samples were used for training. Cross-validation is repeated 10 times, once for each subsample. The results of the cross-validation were averaged to finally obtain a single estimate. The root mean squared error (RMSE), indicating the absolute difference between measured and predicted concentrations, was used to evaluate the models' prediction ability.

2.6. Correlation between air quality and coking emissions

The “LUR + BME” simulation method was used to calculate the spatial distribution of PM_{10} and $PM_{2.5}$ and the AQI index of air quality in the BTH region. The “value extraction to point” extraction tool of ArcGIS was used to obtain the values of the air quality indicators (AQI, PM_{10}, and $PM_{2.5}$) simulated via the LUR model. The exhaust emission data of coking enterprises were obtained from China Industrial Enterprise's pollution emission database. The correlation between air pollutants and exhaust emissions from coking enterprises was analyzed via the SPSS cross-tabulation analysis method.

3. Results and discussion

3.1. Distribution of coking industries in the BTH region

There are about 90 coking enterprises in the BTH region, which are mainly concentrated in Tangshan and Handan City. The number of key supervision enterprises is 19 in Tangshan, 19 in Handan, and 11 in other cities, totaling 50. The kernel density analysis result can directly reflect the distribution of discrete measured values in a continuous region. From the distribution of coking enterprises (Fig. 2), Handan and Tangshan of Hebei Province are areas with a high density of coking enterprises. In addition, Xingtai and Shijiazhuang of Hebei Province also have coking enterprise distribution densities that are higher than those in the surrounding areas.

3.2. Establishment of the land use regression model

According to the multiple regression results, eight factors were selected based on the significance level and collinearity of the factors, which were cultivated land_8 km, forest land_8 km, grassland_4 km, construction land_2 km, road_8 km, altitude, temperature, and relative humidity. The correlation coefficients of each factor are shown in Table 2.

The final regression equation is as follows:

$$Y(PM_{10}) = 13.89 \times TEM + 4.18 \times HUM - 0.000022 \times LEN_{road\_8km} + 20.27 \times AERA_{construction\_2km} + 19.33 \times AERA_{grass\_4km} * AERA_{arable\_8km}\ AERA - 45.82 + 0.07_{wood\_8km} + 22.60 \times DEM - 317.94$$

where DEM is the elevation of the enterprise's location, m; TEM is the temperature of the location of the enterprise, ^{o}C; HUM is the relative humidity of the location of the enterprise, %; LEN_{road_8km} is the length of the road within the 8 km buffer zone at the location of the business, m; and $AERA_{construction\_2km}$, $AERA_{grass\_4km}$, $AERA_{wood\_8kmarable\_8km}$, and AERA are, respectively, the proportion of construction land, grassland, cultivated land, and forest land within the 2 km, 4 km, 8 km, and 8 km buffer zones where the enterprise is located, %. The regression coefficients are provided in Table 3. As observed in the results, the R^2 of LUR for PM_{10} concentrations in the air reaches 82 %, which is a good simulation result (Table 4).

Potential pollution sources, including traffic and industrial emissions; population density; and factors affecting pollution dispersion, including green space and water bodies, are often utilized to construct LUR models. It has been suggested that industrial activities, traffic pollution, and urban construction activities are the most important factors affecting urban air quality [9]. Similarly to the literature [32], the current study indicated that traffic and land use type are the main variables determining air pollution.

Table 2
Bivariate correlation coefficient.

Impact factors	Arable land_8 km	Woodland_8 km	Grassland_4 km	Construction land_2 km	Altitude	Temperature	Relative humidity	Road_8 km
Relevance	0.45	0.49	0.52	0.34	0.54	0.71	0.74	0.02
p	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 3
Regression coefficients of independent and dependent variables.

Models	Non-standardized coefficient		Standard coefficient	t	Significance
	B	Standard deviation	Beta.		
(Constant)	−318	48.4		−6.57	0.00
Elevation	0.07	0.02	0.42	3.70	0.00
Temperature	13.887	2.59	0.62	5.37	0.00
Relative humidity	4.178	0.72	0.55	5.77	0.00
Road_8 km	−0.000022	0.00	−0.05	−0.85	0.40
Construction land 2 km	20.3	7.83	0.20	2.59	0.01
Grass_4 km	19.3	28.1	0.087	0.69	0.50
Arable land_8 km	22.6	7.45	0.20	3.04	0.00
Woodland_8 km	−45.8	20.0	−0.16	−2.29	0.03

Table 4
Cross validation statistical measures for two estimation methods.

Models	R^2	RMSE
LUR	0.82	1.49
LUR + BME	0.95	1.01

3.3. *The combination of LUR and BME improves fitting*

According to the fitting (Fig. 3) and verification results (Table 4), the prediction of air pollutants via the "LUR + BME" method has been significantly improved. Compared to LUR (Fig. 3a), the R^2 of the "LUR + BME" was apparently higher (Fig. 3b), being 0.95. "LUR + BME" is more suitable for the spatial distribution of air pollutant concentrations and air quality in a large-scale region such as the BTH region.

Air pollution is a worldwide problem. The accurate simulation of air quality from the regional scale has been a research focus [33]. Due to the fitness of LUR to large-scale simulations, it has been used in a considerable number of studies. However, the requirement of a linear relationship between variables limits its application. Therefore, the use of nonlinear regression models, such as geographically weighted regression models, has been adopted to improve the simulation of air quality, which significantly increased R^2 [32,34,35]. The current study indicated that it is feasible to apply "LUR + BME" to the simulation of air pollutant distributions in the BTH region,

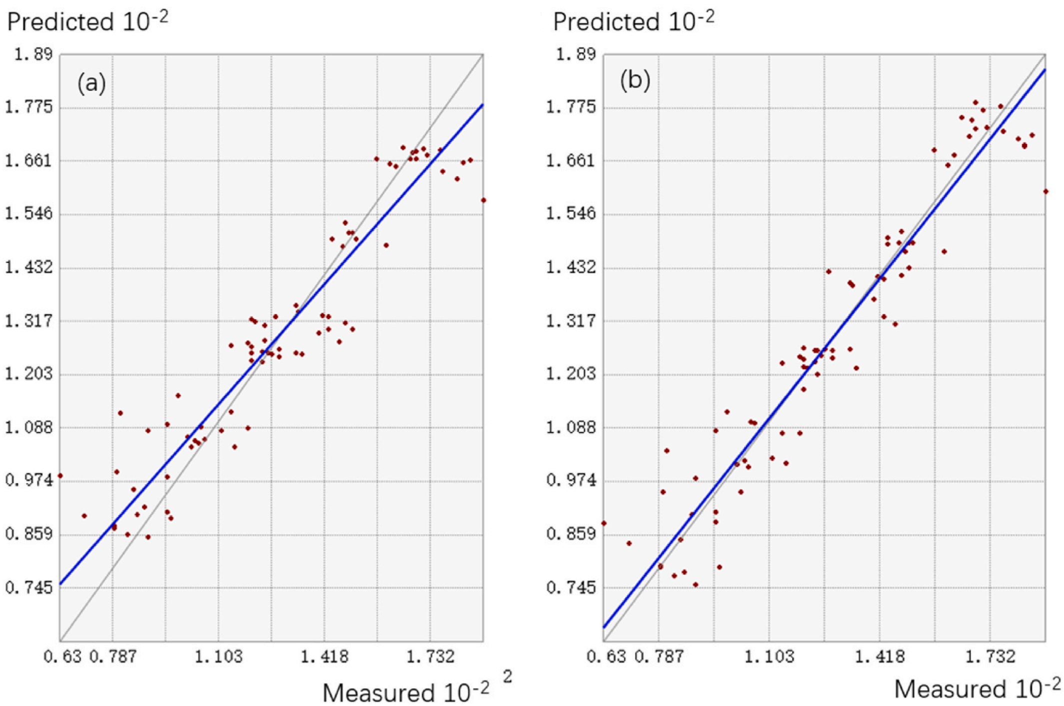

Fig. 3. RMSE evaluation based on LUR (a) and LUR + BME (b).

which obtained a better prediction result than using LUR alone.

Another study has proposed a generalized additive model combining principal component analysis (PCA-GAM) to aid LUR, for the estimation of $PM_{2.5}$ in BTH area, which increased the accuracy of $PM_{2.5}$ estimation [24]. Similar to the current study, they also think than it is necessary to consider both linear and non-linear relationship between $PM_{2.5}$ and variables, which can apparently increase the simulation accuracy.

In our "LUR + BME" mode, LUR is conducted firstly to characterize regional pattern of PM_{10} in BTH. Afterwards, BME is used to extract unexplained variability in the residuals. "LUR + BME" is confirmed to be successful at combining the strengths of both LUR and BME. Similar advantage of "LUR + BME" compared to LUR has been reported [28]. Compared with the previous reports, the R^2 value of the current study is the highest, indicating that the model can be well applied to the BTH region.

3.4. Spatial distribution of air quality

According to the air quality distribution results, the matching degree of different air quality indicators (AQI, PM_{10}, and $PM_{2.5}$) is high. The pollutant concentration shows an increasing trend from northwest to southeast areas; pollutant concentrations are the highest and air quality is the worst in the southern part of the BTH region (Fig. 4). Among them, the areas with poor air quality (Fig. 4a) and high air pollutant concentrations (Fig. 4b and c) are mainly the southern plain region of the BTH region. Baoding, Shijiazhuang, Xingtai, and Handan in the south of the BTH region and Tangshan and Tianjin in the east of the region have relatively poor air quality.

The results were similar to earlier studies made by Yang et al. and Li et al. [24,36], presenting a spatial distribution pattern of "low in the northern mountains and high in the southern plains". Results indicated although the absolute values of $PM_{2.5}$ and PM_{10} have been reduced a lot during the past years with the efforts made to improve the air quality in BTH area, the distribution pattern remained the same. Therefore, it is essential to analyze the spatial distribution of air pollutants and the possible contributing factors of it.

This spatial pattern of air quality was similar to the distribution of coking enterprises in this region (Fig. 2). The coking enterprises with high exhaust emission levels are also located in Tangshan and Tianjin along the coast of Bohai Bay and Handan in Hebei Province, and it can be preliminarily judged that there is a correlation between them.

3.5. Correlation between the emission of coking enterprises and air particulate matter content distribution

Via the correlation analysis of the exhaust emission of coking enterprises and the air quality index, the degree of correlation between the two can be judged. The Sig value of exhaust emissions and the air quality index of coking enterprises is greater than 0.05 (Table 5), indicating that there is no significant difference between exhaust emissions and the air quality index of coking enterprises; therefore, there is a certain correlation between them. By comparing the Phi and V values of the cross-tabular analysis results, it can be observed that the Phi and V values are much higher than 0.1. In the crosstab analysis, Phi and V are indicators that measure the strength of the relationship between the two variables; with their values being greater than 0.1, this indicates that there is a close relationship between the emission of coking enterprises and air particulate matter content distribution.

Coking has been regarded as an important pollution source in the BTH region. Studies have been conducted on the emission characteristics of heavy metals and PAHs from coking enterprises [2], the direct influence of coking exhaust on atmospheric particulates is less studied. Via the current study, it can be suggested that coking still contributes to air pollution in this region, especially in the southern area, although a considerable amount of work has been carried out recently to control pollution from the coking process [37].

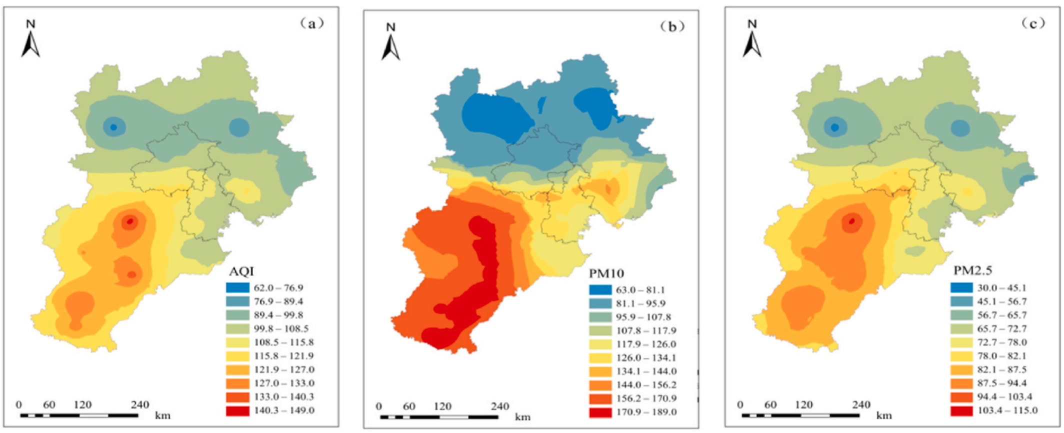

Fig. 4. Air quality distribution map of the Beijing–Tianjin–Hebei region based on "LUR + BME": a) air qualiy (dimensionless); b) PM10 (ug/m^3), c) PM2.5 (ug/m^3).

Table 5
Cross-tabular analysis results of exhaust emissions from coking enterprises and airborne particulate matter content.

Air quality indicators	Exhaust emission		
	Sig.	Phi	V value
AQI	0.28	8.20	0.98
$PM_{2.5}$	0.28	8.20	0.98
PM_{10}	0.28	8.20	0.98

3.6. Environmental implications

Via the analysis of air quality in the BTH region, LUR combined with BME can effectively improve the predictability of air pollutants. In the current study, "LUR + BME" was used to simulate the air quality data in the BTH region with a precision of 95 %. Therefore, the combined use of LUR and BME is feasible for the simulation of regional air quality.

Air pollution in the BTH region is a problem that has received substantial attention [38]. The cross-table analysis showed that the air quality in the BTH region was closely related to the exhaust emission from coking enterprises. The coking enterprises in the BTH region are mainly distributed in Tangshan and Handan, which are located in the southern part of this region. At the same time, air quality was also comparatively poorer in this area than compared to other areas. Coking supervision operations should be intensively implemented in Tangshan and Handan. It is recommended that coking enterprises in BTH area, especially the southern part, should further upgrade their equipment to improve their environmental friendliness. Further, in terms of the urban planning, it is important to consider the environmental conditions like meteorological conditions and population distribution, to further decrease the diffusion of contamination and the potential exposure to vulnerable receptors.

According to the comparison of LUR with LUR + BME, the latter shows better performance than the traditional LUR method, but our model has apparent limitations. It simulated spatial variations in air pollution, but it does not take into account temporal variations. A comprehensive consideration of the spatial and temporal differences of air pollutants in the BTH region may provide more information [39,40]. In addition, our variables focus on land use, traffic, population, and elevation. Including innovative variable forms, such as street patterns, building height, wind direction, and pollution types, may further improve the explanatory power of the model.

4. Conclusions

The predictability of the "LUR + BME" method was significantly improved compared to the LUR method. The root mean square error (RMSE) of the simulation results decreased from 1.49 to 1.01, and the interpretation rate, R^2, of the model increased from 0.82 to 0.95. LUR + BME is feasible for the simulation of air quality on a regional scale. The air quality distribution presented a pattern of "low in the northern mountains and high in the southern plains", similarly to the distribution of coking enterprises. Furthermore, the correlation between exhaust emissions from coking enterprises and air quality in the BTH region indicated a significant correlation.

Data availability

The authors will supply the relevant data in response to reasonable requests.

CRediT authorship contribution statement

Xiaoming Wan: Writing – review & editing, Writing – original draft, Investigation, Funding acquisition, Conceptualization. **Weibin Zeng:** Methodology, Investigation, Formal analysis, Data curation. **Gaoquan Gu:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was funded by the National Natural Science Foundation of China (Grant No. 42077136).

References

[1] J. Luo, X. Feng, H. Han, N. Wang, Identification for discharged characteristics of fine particulate matter from coke chemical industry in northern China, Int. J. Environ. Sci. Technol. 20 (3) (2023) 3351–3360.
[2] L. Mu, L. Peng, X.F. Liu, H.L. Bai, C.F. Song, Y. Wang, Z. Li, Emission characteristics of heavy metals and their behavior during coking processes, Environ. Sci. Technol. 46 (11) (2012) 6425–6430.

[3] X.C. Chen, J.H. Li, M. Jia, S.B. Chen, S.X. Zhang, X. Bo, X. Feng, G.X. Dong, High spatial resolution emission inventory of air pollutants and carbon in China's independent coking industry, Atmosphere 14 (2) (2023).

[4] NBSC, Chinese Statistical Yearbook, Available online:, Chinese Statistics Press, Beijing, China, 2020 http://www.stats.gov.cn/tjsj/ndsj/2020/indexch.htm. (Accessed 10 April 2021).

[5] J. Tian, H.Y. Ni, Y.M. Han, Z.X. Shen, Q.Y. Wang, X. Long, Y. Zhang, J.J. Cao, Primary $PM_{2.5}$ and trace gas emissions from residential coal combustion: assessing semi-coke briquette for emission reduction in the Beijing-Tianjin-Hebei region, China, Atmos. Environ. 191 (2018) 378–386.

[6] Y. Xu, Q.R. Zhong, X. Yun, H.Z. Shen, G.F. Sheng, J.F. Liu, J.M. Ma, J.Y. Hu, Y. Wan, X.J. Wang, C.G. Tian, S. Tao, $PM_{2.5}$-associated health impacts of beehive coke oven ban in China, Environ. Sci. Technol. 53 (19) (2019) 11337–11344.

[7] K.X. Bai, N.B. Chang, W. Gao, Advancing the prediction accuracy of satellite-based $PM_{2.5}$ concentration mapping: a perspective of data mining through in situ $PM_{2.5}$ measurements, Environ. Pollut. 254 (2019) 113047.

[8] H.B. Wang, L.J. Zhao, A joint prevention and control mechanism for air pollution in the Beijing-Tianjin-Hebei region in China based on long-term and massive data mining of pollutant concentration, Atmos. Environ. 174 (2018) 25–42.

[9] C.X. Yan, L.C. Wang, Q. Zhang, Study on coupled relationship between urban air quality and land use in lanzhou, China, Sustainability 13 (14) (2021).

[10] Z.Y. Zhang, C.F. Shao, Y. Guan, C.Y. Xue, Socioeconomic factors and regional differences of PM2.5 health risks in China, 2019, J. Environ. Manag. 251 (2019) 109564.

[11] Y. Jing, L. Pan, Y. Sun, Estimating $PM_{2.5}$ concentrations in a central region of China using a three-stage model, International Journal of Digital Earth 16 (1) (2023) 578–592.

[12] S.Z. Gebreab, D. Vienneau, C. Feigenwinter, H. Bâ, G. Cissé, M.Y. Tsai, Spatial air pollution modelling for a West-African town, Geospatial Health 10 (2) (2015) 205–214.

[13] W.L. Quan, N. Xia, Y.T. Guo, W.Y. Hai, J.M. Song, B.W. Zhang, $PM_{2.5}$ concentration assessment based on geographical and temporal weighted regression model and MCD19A2 from 2015 to 2020 in Xinjiang, China, PLoS One 18 (5) (2023).

[14] B.L. Xu, H. Lin, L.S. Chiu, Y. Hu, J. Zhu, M.Y. Hu, W.N. Cui, Collaborative virtual geographic environments: a case study of air pollution simulation, Inf. Sci. 181 (11) (2011) 2231–2246.

[15] A. Larkin, S. Anenberg, D.L. Goldberg, A. Mohegh, M. Brauer, P. Hystad, A global spatial-temporal land use regression model for nitrogen dioxide air pollution, Front. Environ. Sci. 11 (2023).

[16] S. Pervez, R. Maruyama, A. Riaz, S. Nakai, Development of land use regression model for seasonal variation of nitrogen dioxide (NO_2) in lahore, Pakistan, Sustainability 13 (9) (2021).

[17] G. Hoek, R. Beelen, K. de Hoogh, D. Vienneau, J. Gulliver, P. Fischer, D. Briggs, A review of land-use regression models to assess spatial variation of outdoor air pollution, Atmos. Environ. 42 (33) (2008) 7561–7578.

[18] M. Habermann, M. Billger, M. Haeger-Eugensson, in: Land Use Regression as Method to Model Air Pollution. Previous Results for Gothenburg/Sweden, Symposium on Towards integrated modelling of urban systems, Lyon, France, 2014, pp. 21–28. Lyon, France, Oct 15-15, 2014.

[19] T. Cong-Thanh, N.N. Tuyen, N.T.T. Hang, Iop in Development of Environmental Data for Land Use Regression Models to Assess Fine Particulate Matter Pollution in Ho Chi Minh City, National Conference on GIS Application - GIS and Remote Sensing Applications for Environment and Resource Management, Vietnam Natl Univ Ho Chi Minh City, VIETNAM, 2022. Nov 01-12.

[20] N.O. Lim, J. Hwang, S.J. Lee, Y. Yoo, Y. Choi, S. Jeon, Spatialization and prediction of seasonal NO_2 pollution due to climate change in the Korean capital area through land use regression modeling, Int. J. Environ. Res. Publ. Health 19 (9) (2022) 5111.

[21] A. Molter, S. Lindley, F. de Vocht, A. Simpson, R. Agius, Modelling air pollution for epidemiologic research - Part I: a novel approach combining land use regression and air dispersion, Sci. Total Environ. 408 (23) (2010) 5862–5869.

[22] S. Xu, B. Zou, S. Shafi, T. Sternberg, A hybrid Grey-Markov/LUR model for PM_{10} concentration prediction under future urban scenarios, Atmos. Environ. 187 (2018) 401–409.

[23] B.H.Q. He, M.R. Heal, S. Reis, Land-use regression modelling of intra-urban air pollution variation in China: current status and future needs, Atmosphere 9 (4) (2018).

[24] S. Li, L. Zhai, B. Zou, H.Y. Sang, X.A. Fang, Generalized additive model combining principal component analysis for $PM_{2.5}$ concentration estimation, ISPRS Int. J. Geo-Inf. 6 (8) (2017) 248.

[25] L. Chen, S. Gao, H. Zhang, Y.L. Sun, Z.X. Ma, S. Vedal, J. Mao, Z.P. Bai, Spatiotemporal modeling of $PM_{2.5}$ concentrations at the national scale combining land use regression and Bayesian maximum entropy in China, Environ. Int. 116 (2018) 300–307.

[26] G. Shaddick, H.J. Yan, R. Salway, D. Vienneau, D. Kounali, D. Briggs, Large-scale Bayesian spatial modelling of air pollution for policy support, J. Appl. Stat. 40 (4) (2013) 777–794.

[27] S. Xu, C. Cui, M. Shan, Y. Liu, Z. Qiao, L. Chen, Z. Ma, H. Zhang, S. Gao, Y. Sun, Spatio-temporal prediction of ground-level ozone concentration based on bayesian maximum entropy by combining monitoring and satellite data, Atmosphere 13 (10) (2022).

[28] M.R. Jeanette, L.S. Marc, An LUR/BME framework to estimate PM2.5 explained by on road mobile and stationary sources, Environ. Sci. Technol. 48 (2014) 1736–1744.

[29] Z. Zhang, J. Yang, The density fraction estimation of coarse coal by use of the kernel method and machine vision, Energy Sources, Part A Recovery, Util. Environ. Eff. 37 (2) (2015) 181–191.

[30] L. Chen, S. Liang, X.L. Li, J. Mao, S. Gao, H. Zhang, Y.L. Sun, S. Vedal, Z.P. Bai, Z. X.Haiyu Ma, M. Azzi, A hybrid approach to estimating long-term and short-term exposure levels of ozone at the national scale in China using land use regression and Bayesian maximum entropy, Sci. Total Environ. 752 (2021) 141780.

[31] S. Xu, C. Cui, M. Shan, Y. Liu, Z. Qiao, L. Chen, Z. Ma, H. Zhang, S. Gao, Y. Sun, Spatio-temporal prediction of ground-level ozone concentration based on bayesian maximum entropy by combining monitoring and satellite data, Atmosphere 13 (10) (2022).

[32] L.C. Zhang, X. Tian, Y.H. Zhao, L.L. Liu, Z.W. Li, L.X. Tao, X.N. Wang, X.H. Guo, Y.X. Luo, Application of nonlinear land use regression models for ambient air pollutants and air quality index, Atmos. Pollut. Res. 12 (10) (2021).

[33] G. Hoek, R. Beelen, K. de Hoogh, D. Vienneau, J. Gulliver, P. Fischer, D. Briggs, A review of land-use regression models to assess spatial variation of outdoor air pollution, Atmos. Environ. 42 (33) (2008) 7561–7578.

[34] Y.R. Lyu, K. Kirwa, M. Young, Y. Liu, J. Liu, S.X. Hao, R.K. Li, D.Q. Xu, J.D. Kaufman, A high-resolution computationally-efficient spatiotemporal model for estimating daily $PM_{2.5}$ concentrations in Beijing, China, Atmos. Environ. (2022) 290.

[35] S. Munir, M. Mayfield, D. Coca, L.S. Mihaylova, A nonlinear land use regression approach for modelling NO_2 concentrations in urban areas-using data from low-cost sensors and diffusion tubes, Atmosphere 11 (7) (2020).

[36] X.H. Yang, D.P. Xiao, H.Z. Bai, J.Z. Tang, W. Wang, Spatiotemporal distributions of $PM_{2.5}$ concentrations in the beijing-tianjin-hebei region from 2013 to 2020, Front. Environ. Sci. 10 (2022).

[37] Y.Y. Du, Q. Wan, H.M. Liu, H. Liu, K. Kapsar, J. Peng, How does urbanization influence $PM_{2.5}$ concentrations Perspective of spillover effect of multi-dimensional urbanization impact, J. Clean. Prod. 220 (2019) 974–983.

[38] X.D. Han, H.J. Li, Q. Liu, F.Z. Liu, A. Arif, Analysis of influential factors on air quality from global and local perspectives in China, Environ. Pollut. 248 (2019) 965–979.

[39] Y. Chen, N. Schleicher, M. Fricker, K. Cen, X.L. Liu, U. Kaminski, Y. Yu, X.F. Wu, S. Norra, Long-term variation of black carbon and $PM_{2.5}$ in Beijing, China with respect to meteorological conditions and governmental measures, Environ. Pollut. 212 (2016) 269–278.

[40] C.G. Lorena, C. Daniela, M. Eva, D. Adelaide, M. Annarosa, R.G. Maria, C. Daniele, Inter-annual variability of source contributions to PM_{10}, $PM_{2.5}$, and oxidative potential in an urban background site in the central mediterranean, J. Environ. Manag. 319 (2022) 115752.