Research article

Spatially differentiated effects of socioeconomic factors on China's NO\textsubscript{x} generation from energy consumption: Implications for mitigation policy

Abstract

Nitrogen oxides (NO\textsubscript{x}) has become the priority of China's air pollution control, but the regional socio-economic factors responsible for NO\textsubscript{x} generation are embedded with spatial disparities, which leads to different effects of air quality policy at the local level. This study applied a geographically weighted regression (GWR) model to investigate the drivers of NO\textsubscript{x} generation from energy consumption (NGEC) in China's 30 provinces, to explore nonstationary spatial effects of NGEC. The results showed that population size has always been the dominant factor in spatial NGEC across all regions of China, although there is a minor north-south difference. However, the
effect of per capita GDP and energy intensity leads to a significant north-south difference when they are influencing NGEC, which shows a minor west-east difference from thermal power generation (TE). We also found that in Northern and Northeast China, the transition towards cleaner energy structure based on natural gas has started correlating significantly with NO\textsubscript{x} generation through a weakly negative effect in 2015. Our findings show alternative strategies on NO\textsubscript{x} reduction, which include the spatially differentiated effect of regional socioeconomic factors on energy consumption.

Keywords: NO\textsubscript{x} generation from energy consumption; Driving factors; Geographically weighted regression; China

1 Introduction

China's severe air pollution has been mainly induced by massive energy consumption based on fossil fuels, which has been driven by rapid industrialisation and urbanization over the past decades. The NO\textsubscript{x} emissions from China's fossil fuel consumption could account for more than 90% of the total emissions in China (Cui et al., 2013). Moreover, the severe haze and smog pollution in highly industrialized and populated areas, such as Beijing-Tianjin-Hebei and surrounding areas, the Yangtze River Delta, and the Weihe-Fenhe plain (Cai et al., 2018; Yang et al., 2018), have become an important issue (Guo et al., 2014; Huang et al., 2018). Among all kinds of air pollutants, nitrogen oxides (NO\textsubscript{x}) have attracted much more attention from the scientific community, as NO\textsubscript{x} is a necessary precursor to cause fine particles, ozone, and other regional pollutants (Huang et al., 2014). Therefore, it would be crucial to control NO\textsubscript{x} emissions from energy consumption for effectively reducing particulate matter (like PM\textsubscript{2.5}) concentrations.

In response to such concerns, a series of policies were issued by the Chinese government to mitigate NO\textsubscript{x} emissions, such as "the 13th Five Year Plan of Energy Saving and Emission Reduction", which aims to reduce national NO\textsubscript{x} emissions by 15% in 2020. The Blue-Sky Plan has also outlined provincial NO\textsubscript{x} reduction targets based on regional air quality goals. However, the top-down implementation of those policies is difficult due to the regional heterogeneity of socioeconomic factors and the geographical imbalances of industrialisation and urbanization (Kanada et al., 2013; Liang et al., 2016). Therefore, it is imperative to investigate the driving mechanisms behind regional NO\textsubscript{x} emissions for developing a spatially differentiated strategy on the NO\textsubscript{x} reduction targets.

For such an investigation, it is crucial to estimate NO\textsubscript{x} generation related to energy consumption. Previous studies mostly focused on generating the NO\textsubscript{x} emissions inventory using the bottom-up emissions inventory (Hao et al., 2002; Huang et al., 2011) and satellite remote sensing (Cho et al., 2017; Jiang et al., 2016). The statistical data on NO\textsubscript{x} emissions have provided a critical foundation for assessing the characteristics of NO\textsubscript{x} spatial distribution using quantitative analysis. For example, Wang (2013) described the spatial characteristics of NO\textsubscript{x} emissions intensity in China based on exploratory spatial data analysis (ESDA), which showed that provincial NO\textsubscript{x} emissions intensity had spatial autocorrelation and agglomeration.
However, the current studies exploring the relations between spatial socioeconomic factors and NO\textsubscript{x} emissions have shown a significant research gap. Spatial econometric models that involve spatial autocorrelation and heterogeneity are usually employed in the research on the correlation between socioeconomic factors and air pollution (Hao and Liu, 2016; Kang et al., 2016; Zhou et al., 2017). Especially, Diao et al. (2018) have ever surveyed the relationship between NO\textsubscript{x} emissions and its determinants, using the spatial lag model (SLM), the spatial error model (SEM), and the spatial Durbin model (SDM), which all employed global models. However, these studies only assume the spatial stationarity as a prerequisite in the SEM, SLM and SDM, which all considered proximity effects. So they could not generate a separate parameter for each observation and thereby reveal the different spatial links of every object being studied (Griffith and Paelinck, 2018).

The emissions of air pollutants usually have spatial heterogeneity, thanks to the spatial difference in economic development (Wang et al., 2019). Specific research methods thus have to be applied to capture such spatial variability and non-stationarity. As a varying coefficient method, the Geographical Weighted Regression (GWR) model has recently gained more focus with the purpose to explore the location-specific impacts from various drivers on environmental pollution. Wang and Fang (2016) investigated the determinants of urban PM\textsubscript{2.5} concentrations in the Bohai Economic Rim. Similarly, Xu and Lin (2017) and Xu et al. (2017) used the GWR model to evaluate the mechanisms of CO\textsubscript{2} emissions from China's manufacturing and agricultural sectors. Fan et al. (2018) studied the impacts of urban form on air pollutant emissions in China, including NO\textsubscript{x} emissions.

Those studies compensated the research gap mentioned above but still have the following limits. First, previous research has mostly focused on the effect of end-of-pipe reduction technologies on NO\textsubscript{x} emissions, but less attention on how local socioeconomic aspects and energy consumption factor cause the spatial distribution of NO\textsubscript{x} generation at the source. Second, although some studies have considered spatial effects, they did not explain the heterogeneity between the driving forces of air pollutants because of the limitation of standardized coefficient regression methods (Mashhoodi, 2018). Third, previous studies focus on actual NO\textsubscript{x} emissions, which are mostly affected by end-of-pipe measures and socioeconomic factors together. China's strict environmental regulations of air pollution have made NO\textsubscript{x} emissions dropped dramatically (Wang et al., 2018). Thus we need to understand the role of NO\textsubscript{x} generation, which is mainly affected by socioeconomic factors instead of end-of-pipe measures.

As a result, this study aims to fill those gaps by answering the following questions: (a) how different are the impacts of energy consumption and socioeconomic context upon generation from NO\textsubscript{x} energy consumption (NGEC) across provincial regions in China? (b) what kind of impact mechanism based on spatial heterogeneity could work on NO\textsubscript{x} generation in China? (c) how the coal-to-gas policy and energy intensity improve the NO\textsubscript{x} generation across all regions?

Accordingly, this study would proceed with the following steps. First, through quantifying the provincial NGEC in China, we identify the characteristics of its spatial distribution and aggregation. Second, we use the GWR model to investigate location-specific effects of the driving factors on NGEC, where the estimated parameters outputted by the GWR model vary across provinces. Finally, this paper proposes an alternative
strategy to manage the relationship between China's energy consumption, economic development, and NO\textsubscript{x} reduction.

The following structure would be in this study: Section 2 presents the framework and methods used in this study and explains the sources of the data used. Section 3 outlines and presents the results. Section 4 offers a thorough discussion of spatial influences and policy implications. Finally, the main conclusions are summarized and are outlined in Section 5.

2 Methodology

In this study, all statistical analysis would be conducted according to Fig. 1. We introduced Exploratory Spatial Data Analysis (ESDA) to examine univariate spatial autocorrelation in NGEC. Also, we would do a multivariate analysis by applying Ordinary Least Squares (OLS) regressions to examine initial relationships without spatial dependence, and test all possible combinations of variables. We chose the best OLS model based on AIC scores and examined variables retained for collinearity using Variance Inflation Factors (VIF). Statistically significant Koenker (BP) statistics could provide analysis variation in the relationships between variables and NGEC. Geographically Weighted Regression (GWR) could provide new insights for local differentiated NO\textsubscript{x} control.
2.1 Estimation of NO\textsubscript{x} generation from energy consumption

Based on the China Energy Statistical Yearbook (2006–2016), NO\textsubscript{x} generation from energy consumption would be estimated for China’s 30 provinces from 2005 to 2015, using the bottom-up emission methods (Huang et al., 2011; Wang et al., 2018):

\[
E_{tij} = \sum_{ij} EF_{ijf} \times Q_{ijf0}
\]  

(1)

Where \(E_{tij}\) is the amount of NGEC at year \(t\); the subscripts \(i, j, \) and \(f\) represent the province, sector, and fuel type in terms of energy consumption, respectively; \(EF\) is the NO\textsubscript{x} generation factor, and \(Q\) represents the quantity of fuel consumption for each sector. Fuel types in this study covered coal, diesel oil, coke, gasoline, fuel oil, crude oil, coke oven gas, kerosene, natural gas, liquefied petroleum gas, other gas, and refinery gas. The factors of NO\textsubscript{x} generation for each fuel (Appendix A) were obtained from Kato and Akimoto (1992) and...
Hao et al. (2002), which are widely applied in China-related environmental studies (Tian et al., 2001; Gao et al., 2006; Jiang et al., 2016). The generation factors for the heating and agricultural sectors refer to industry and wholesale, respectively. The accounting details for NGEC are given in Appendix B.

2.2 Exploratory spatial data analysis

Spatial autocorrelation models are popular to characterize spatial distribution patterns (Diao et al., 2018; Xu and Lin, 2017, 2018). Moran’s I which measures spatial autocorrelation (Moran, 1948; Geary, 1954) can be further classified into Global Moran’s I and Local Moran’s I (Anselin and Griffith, 1988). In this study, we use Global Moran’s I to estimate the degree of spatial dependence and heterogeneity of NGEC among 30 provinces in China from the years 2005 to 2015, applying Open Geoda 1.2. The formula is as follows:

\[
Moran' I = \frac{n}{\sum_{i,j=1}^{n} W_{ij}} \frac{\sum_{i=1}^{n} \sum_{j \neq i}^{n} W_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}
\]

(2)

where \(x\) is a variable measured in each of the \(i = 1, 2, \ldots, n\) locations, and \(W_{ij}\) is the element in row \(i\) and column \(j\) of the spatial weights matrix. The Z-score is calculated using

\[
Z = \frac{I - E(I)}{\sqrt{Var(I)}},
\]

where \(E(I)\) and \(Var(I)\) are the expectation and variance, respectively.

The local Moran statistic is used to analyse spatial clustering and can provide more detailed insights into the location-specific nature of spatial dependence. The specific formula is as follows:

\[
I_i = \frac{z_i}{\sum_{i=1}^{n} z_i^2} \times z_i^2
\]

(3)

where \(z_i\) expresses the observation for region \(I\) for a variable as a deviation from the mean, and the \(z_i^2\) is the spatial lag for location \(I\), obtained as:

\[
z_i^2 = \sum_{j=1}^{n} a_{ij} z_j
\]

(4)

In the local spatial autocorrelation implementation, each observation can be placed into one of four types: HH indicates that both the province itself and the neighbouring provinces have higher values; LL denotes that both the province itself and the neighbouring provinces have low NO\(_x\) emissions; LH indicates low values surrounded by high values; finally, HL indicates high values surrounded by low values. Additionally, Local Indicators of Spatial Association (LISA) aggregation map were used to present the spatial distribution of results.
2.3 STIRPAT model

According to the Environmental Kuznets Curve (EKC) hypothesis (He and Wang, 2012), the STIRPAT model is widely used in the energy-related field (Poumanyvong et al., 2012; Liu et al., 2015; Wang et al., 2017a; Shafiei and Salim, 2014; Xu and Lin, 2017):

\[ I = aP^bA^cT^d e \]

where \( a \) is the constant, \( b, c, d \) are the exponential terms of \( P \) (Population), \( A \) (Affluence), and \( T \) (Technology), respectively, and \( e \) is the error term.

The STIRPAT model preserves the multiple correlations between human driving forces of the IPAT model and considers human driving forces such as population, affluence, technology as primary factors influencing environmental pressure changes. The original model is often improved to suit the various purposes and needs of different empirical studies. In this study, we took the logarithmic of all variables to eliminate possible heteroscedasticity. At the same time, we also standardized variables \(^1\) as the variables have different meanings and units:

\[ \ln I = \ln a + b (\ln P) + c (\ln A) + d (\ln T) + \ln e \]

Combining the STIRPAT model with the existing literature, we chose eight influencing factors to explore the impacts of socioeconomic factors on NGEC using the stepwise regression. We obtained the following model (Eq. (7)):

\[ \ln y_{ij} = \beta_0 + \beta_{ps} \ln P_{ij} + \beta_{PD} \ln PD_{ij} + \beta_{FDP} \ln GDP_{ij} + \beta_{FDI} \ln FDI_{ij} + \beta_{TE} \ln TP_{ij} + \beta_{EI} \ln EI_{ij} + \beta_{TE} \ln TE_{ij} \]

where \( y_{ij} \) represents NO\(_x\) generation from energy consumption in the province \( I \) in year \( j \), \( \beta_0 \) is a regression constant, \( \beta_i \) denotes the partial regression coefficients of the \( i \)th explanatory variable, and \( e \) is the error.

Economic growth is an important economic factor on air pollution emissions, which has always been a key concern of many studies (Xu and Lin, 2017; Hao et al., 2015; West et al., 2013; De Foy et al., 2016). As an increasingly important air pollutant, NGEC had been closely linked to the rapid growth of China's economy over the past years. So we selected include three variables in this study to show the \( A \) (Affluence) factors: GDP per capita (PGDP), foreign direct investment (FDI) and the proportion of tertiary industry (TP). Moreover, considering the effects of technological progress on NGEC, we applied three factors, including the energy intensity (EI), thermal power generation (TE) and natural gas consumption ratio (NGR) as the \( T \) (Technology) factors in this study. For example, if NGEC decreases with the decline of EI, then it is implied that technological progress plays a positive role (Wu et al., 2016; Xu and Lin, 2016; Xu et al., 2016; Diao et
The $P$ (Population) factors were measured by two variables in this study. One is population size (PS), which could have a strong impact on energy consumption and air pollution emissions (Lamsal et al., 2013; Lyu et al., 2016); the other is urban density (UD), which is closely related to industrial layout, urban planning and population policy. The detailed descriptions of input indicators are given in Table 1.

![Table 1](image)

The presentation of Tables and the formatting of text in the online proof do not match the final output, though the data is the same. To preview the actual presentation, view the Proof.

Summary of all variables in the modelling analysis.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Definition</th>
<th>Units of measurement</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGEC</td>
<td>NO$_x$ generation from energy consumption</td>
<td>tonne</td>
<td>1,154,666</td>
<td>796,482</td>
</tr>
<tr>
<td><strong>Independent variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PS</td>
<td>Population size</td>
<td>$10^4$ people</td>
<td>4421.01</td>
<td>2654.75</td>
</tr>
<tr>
<td>UD</td>
<td>Urban density</td>
<td>People/km$^2$</td>
<td>2370.43</td>
<td>1386.30</td>
</tr>
<tr>
<td>PGDP</td>
<td>Per capita GDP</td>
<td>Yuan</td>
<td>23,896</td>
<td>15,443</td>
</tr>
<tr>
<td>FDI</td>
<td>Foreign direct investment</td>
<td>$10^4$ USD</td>
<td>2,799,081</td>
<td>3,214,462</td>
</tr>
<tr>
<td>TP</td>
<td>The proportion of tertiary industry</td>
<td>Per cent</td>
<td>46.82%</td>
<td>8.55%</td>
</tr>
<tr>
<td><strong>Technology factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EI</td>
<td>Energy intensity</td>
<td>$10^4$ tonne/billion yuan</td>
<td>1.6360</td>
<td>0.9653</td>
</tr>
<tr>
<td>TE</td>
<td>Thermal power generation</td>
<td>$10^8$ kW·h</td>
<td>1087.11</td>
<td>946.76</td>
</tr>
<tr>
<td>NGR</td>
<td>Natural gas consumption ratio</td>
<td>Per cent</td>
<td>7.28%</td>
<td>7.07%</td>
</tr>
</tbody>
</table>

2.4 Geographical Weighted Regression

It may be more realistic to assume that human activities are heterogeneous in different regions (Tenerelli et al., 2016); thereby, a GWR model could be adopted to solve this problem. Two essential prerequisites are needed when we start applying the GWR model (Wheeler and Páez, 2010). One is that the samples of socioeconomic
phenomena must have spatial autocorrelation; The other is that there should be spatial non-stationarity among the variables. The GWR model could be correctly applied only after the global and local spatial autocorrelation analysis. The form of the traditional linear regression model of GWR is:

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_i X_i + \ldots + \beta_n X_n + \epsilon \tag{8} \]

where \( \beta_0 \) represents a constant, \( \beta_i \) represents the regression coefficient, which is estimated using the ordinary least squares (OLS) method, and \( \epsilon \) represents a random disturbance term, which satisfies the spherical disturbance hypothesis.

By allowing a local weight based on a spatial location matrix, we could describe the distance between the observed location and the estimated point location. So the GWR model can be re-expressed as below:

\[ y_i = \beta_0 (u_i, v_i) + \sum_{k}^{n} \beta_k (u_i, v_i) x_{ik} + \epsilon_i \tag{9} \]

In this equation, \( y_i \) is the dependent variable of NO\textsubscript{x} generation from energy consumption in province \( i \), \( \beta_0 (u_i, v_i) \) is the intercept coefficient of province \( i \), \( \beta_k (u_i, v_i) \) is the location regression coefficient, \( \{u_i, v_i\} \) denotes the coordinates of the province \( i \), \( x_{ik} \) is the value of the \( k \)th independent variable, and \( \epsilon_i \) is the random location-specific error term of \( i \)th province. The location estimates are obtained by weighting the instances around province \( i \) according to Eq. (10):

\[ \hat{\beta} (u_i, v_i) = (X^T W (u_i, v_i) X)^{-1} X^T W (u_i, v_i) y \tag{10} \]

\[ W_{ij} = \exp \left( - \left( \frac{d_{ij}}{b} \right)^2 \right), \text{ if } d_{ij} < b \]

\[ W_{ij} = 0 \quad \text{otherwise} \tag{11} \]

where \( \hat{\beta} (u_i, v_i) \) is the estimate of the parameter in \( \{u_i, v_i\} \), \( X \) indicates a vector of independent variables, \( y \) represents a vector of dependent the variable, \( W (u_i, v_i) \) is a spatial weight matrix using the fixed Gaussian function, which refers to the weight of instance observed for province \( i \) for estimating the coefficient for province \( j \). In Eq. (10), \( d_{ij} \) is the distance between \( i \) and \( j \), and \( b \) is referred to as the bandwidth. In this study,
using ArcGIS 10.2, the fixed bandwidth was determined by the corrected Akaike Information Criterion (AIC) of the GWR model.

2.5 Data sources

The data set consists of the cross-sectional data for the 30 provinces of mainland China; the Tibet Autonomous Region was not included due to incomplete data. Based on the China Statistical Yearbook (2006–2016), China Energy Statistical Yearbook (2006–2016), and China City Statistical Yearbook (2006–2016), we collected cross-sectional data for the years 2005–2015, including PS, urban area, PGDP, FDI, industry structure, EI, and NGR. UD was equal to urban population divided by the urban area. In order to cut the effect of inflation, PGDP was converted into constant prices based on 1995. EI was equal to energy use from energy consumption divided by real GDP. TE was obtained from energy balance sheets of each province in the China Energy Statistical Yearbook (CESE), under the item of the output of thermal power in the transformation.

3 Results

3.1 Temporal and spatial distribution features of NGEC

As shown in Fig. 2a, we observed that national NGEC grew rapidly from 2005 to 2012, then peaked in 2012, and then slowly declined after 2012. This is because China has listed NO\textsubscript{x} emissions as one of the controlled indicators since 2011. At the provincial level (Fig. 2b), the provinces with red boxes in Fig. 2b showed a similar trend as the national level, but the provinces with blue boxes showed a different trend. It is worth noting that the declining trend of national NO\textsubscript{x} emissions is slow after 2012, which might indicate that the reduction of China's NGEC in future still need more actions from improving the industrial structure and energy consumption structure although end-of-pipe reduction policies could play a role in reducing NGEC. More importantly, we could not ignore the trend of NGEC in various provinces, and it is worth to explore how NGEC changes in different provinces were affected by different socio-economic and energy factors (see Fig. 3).
National and provincial characteristics of NGEC from 2005 to 2015. (a) national level; (b) provincial level.

Fig. 3

Figure Replacement Requested
We chose the years of 2005, 2010, 2015, to represent the spatial distribution trend of NGEC over ten years, which are mainly based on China's Five-Year Plan\(^2\). We observed that NGEC had shown spatial heterogeneity in China from 2005 to 2015 (see Fig. 3). The high NGEC accumulation areas were mostly concentrated in Inner Mongolia, Guangzhou, the Yangtze River Delta region, Beijing-Tianjin-Hebei region and its surrounding areas. Comparably, low NGEC accumulation areas were located in the western less-developed areas. So, the reason might be that the spatial difference of rapid economic growth and urbanization has led to an increasing spatial difference in energy consumption size since 2000. We would apply the spatial correlation method and the GWR model to explore what and how the spatial socioeconomic drivers cause the spatial aggregation of \(\text{NO}_x\) emissions in China.

### 3.2 Spatial correlation analysis on NGEC

The results in Table 2 shows the global Moran's I values for the spatial correlation of NGEC from 2005 to 2015. All the p-values in every single year from 2005 to 2015 had a significance level of 1%, which means that the null hypothesis can be rejected. Additionally, the Moran's I value was more than zero for each year, indicating that there was a spatial autocorrelation in NGEC. Furthermore, the Moran's I index declined from 2005 to 2015, which further reveals that the spatial agglomeration weakened over the past ten years. In terms of Z-score, the index value for each year was over 1.65, suggesting that there was a positive spatial autocorrelation of HH and LL in terms of NGEC.

<table>
<thead>
<tr>
<th>Year</th>
<th>Moran's I</th>
<th>Expected Index</th>
<th>Z-score</th>
<th>p-value</th>
</tr>
</thead>
</table>

The presentation of Tables and the formatting of text in the online proof do not match the final output, though the data is the same. To preview the actual presentation, view the Proof.
The three maps of local indicators of spatial association (LISA) agglomeration (Fig. 4) depict the results of the year 2005, 2010 and 2015. Based on Moran's I, the spatial autocorrelation could be classified into four types including HH, LL, HL and LH. The red and blue provinces indicate HH and LL spatial clusters of NGEC, respectively. The HH depicts that Shandong and Jiangsu province are two cluster centre of remarkable high NGEC in China, and Qinghai province is a cluster centre of low NGEC. Additionally, HH and LL region was much stable in the past ten years. So, we would further explore how the spatially socioeconomic drivers cause high NGEC in Shandong, Jiangsu, and their surrounding regions.
3.3 Spatial correlations between driving factors and NGEC

According to the prerequisites of the GWR model, we should first test the OLS model of NGEC, which is presented in Table 3. The empirical results indicate that all of the independent variables were statistically significant at a 95% confidence level, and Variance Inflation Factors (VIFs) was relatively low. The Adjusted-$R^2$ values indicate that more than 96% of the variation in NGEC can be explained using this model. In general, the factors selected in the model were comprehensive and representative, and the fit of the model was high.

We have also observed that the Koenker statistic (BP) was significant, which indicates there was a spatial instability between the model-dependent variable and the independent variables. Moreover, this spatial instability reduces the fit of the model. Therefore, we set up a GWR model that could accommodate the spatial

<table>
<thead>
<tr>
<th>Variables/Year</th>
<th>2005</th>
<th>p-value</th>
<th>2010</th>
<th>p-value</th>
<th>2015</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnPS</td>
<td>0.6260</td>
<td>0.000000*</td>
<td>0.6462</td>
<td>0.000000*</td>
<td>0.5716</td>
<td>0.000000*</td>
</tr>
<tr>
<td>lnUD</td>
<td>−0.0728</td>
<td>0.042311*</td>
<td>−0.0580</td>
<td>0.040405*</td>
<td>−0.0095</td>
<td>0.004391*</td>
</tr>
<tr>
<td>lnPGDP</td>
<td>0.4412</td>
<td>0.000037*</td>
<td>0.3662</td>
<td>0.000117*</td>
<td>0.3617</td>
<td>0.00009*</td>
</tr>
<tr>
<td>lnFDI</td>
<td>−0.1081</td>
<td>0.033541*</td>
<td>−0.0131</td>
<td>0.036560*</td>
<td>−0.0247</td>
<td>0.030002*</td>
</tr>
<tr>
<td>lnTP</td>
<td>0.0164</td>
<td>0.157231</td>
<td>0.0115</td>
<td>0.25606</td>
<td>−0.0311</td>
<td>0.027016*</td>
</tr>
<tr>
<td>lnEI</td>
<td>0.2357</td>
<td>0.000785*</td>
<td>0.2614</td>
<td>0.000243*</td>
<td>0.3813</td>
<td>0.000021*</td>
</tr>
<tr>
<td>lnTE</td>
<td>0.4321</td>
<td>0.000000*</td>
<td>0.4370</td>
<td>0.000000*</td>
<td>0.2947</td>
<td>0.000160*</td>
</tr>
<tr>
<td>lnNGR</td>
<td>−0.0672</td>
<td>0.114338</td>
<td>−0.0516</td>
<td>0.274890</td>
<td>−0.1148</td>
<td>0.018649*</td>
</tr>
<tr>
<td>Adj-$R^2$</td>
<td>0.9729</td>
<td>0.9665</td>
<td>0.9707</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AICc</td>
<td>18.3831</td>
<td>−8.8646</td>
<td>−0.0264</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-B</td>
<td>0.9190</td>
<td>0.9004</td>
<td>0.4098</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K(BP)</td>
<td>0.0335</td>
<td>0.0476</td>
<td>0.0426</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTES: ‘*’ indicates a p-value of less than 0.05. Adj-$R^2$, AIC, J-B, K(BP) are Adjusted R-squared, corrected Akaike Information Criterion, the p-value for Jarque-Bera statistic, and p-value for Koenker statistic, respectively.
instability. We found that the corrected Akaike information criterion (AICc) in the GWR model is lower than in the OLS model, which means the GWR model performed better than the OLS regression model (Table 4).

According to GWR results, the coefficients which characterised to identify the temporally and spatially varying relationships are all collected in Fig. 5. Moreover, each box could reflect the spatial distribution of the relationships between a specific variable and provincial NGEC, also the positive or negative correlations of drivers. We found that, in all provinces of China, PS, PGDP, EI, and TE were positively correlated with NGEC, and NGP and UD were negatively correlated with NGEC. Specifically, TP only negatively correlated with NGEC in all regions. The top three rankings for elasticity in 2005 and 2010 are PS, PGDP, and TE, while EI replaced TE and ranked the third place in 2015, which means that the effect of EI on NGEC has been strong since 2015.
4 Discussion

4.1 The spatial impact of energy factors on NGEC

It is known that energy consumption activities will directly influence the growth of NGEC. Policy-makers usually regulate the energy consumption activities with two approaches: energy efficiency and energy structure (Han et al., 2007). Energy intensity (EI) is the main indicator to measure the energy efficiency of a given economy, while the energy structure represents the input structure of primary energy in the economy.

We found that EI, of all factors, had a positive and significant correlation with spatial NGEC across all provinces from 2005 to 2015 (Fig. 6a), which featured the highest growth with time (increasing from 0.17–0.29 to 0.33–0.42). Also, as indicated in Fig. 7, Northern China (0.21–0.25, 0.30–0.33) featured a stronger correlation than Central (0.18–0.21, 0.27–0.29) and Western (0.17–0.26, 0.22–0.34) China in 2005 and 2010. However, Southern China had a similar effect as Northern China only in 2015. We also observed a trend of increasing spatial convergence of EI from 2005 to 2015 (Fig. 5).
Accordingly, these findings are meaningful to understand the impact mechanism of EI, and can lead to policy implications for future NO\textsubscript{x} reduction. Energy intensity is closely linked to energy policy and industrial structure. In the Northern and Eastern China, the energy-saving technologies and standards are more frequently updated thanks to the stringent regulation (Zhang et al., 2019a; Wu et al., 2019). Such context of policy-making could mobilise more financial investment on improving EI, and also lead to more economic.
benefits for the reduction of NO$_x$ generations in those regions. On the other hand, China's trend towards green development is leading to a transition from high energy-intensity to low energy-intensity, which could take a prohibiting effect on NO$_x$ generation. Consequently, these factors caused a stronger correlation with NO$_x$ generation in Northern and Eastern China than the rest of the region.

We found that thermal power plants have a decreasing impact (dropping from 0.40–0.46 to 0.27–0.35) on NGEC (Fig. 6b). There is an impact gap of TE between Western China (0.31–0.39) and Eastern China (0.27–0.31) in 2005 (Fig. 8). China has made ultra-low NO$_x$ emission standards for thermal power plants, but the implementation of those standards is stricter in Eastern China than in Western China.

Another energy factor NGR, however, has shown a regional differentiation between Northern China (0.13–0.16) and Southern (0.08–0.10) China. We found that the spatial distribution of NGR significantly correlated the spatial distribution of NO$_x$ generated significantly in 2015 (Fig. 9), which is significantly different from the statistically non-significant result in 2005 and 2010 (Fig. 6c) although its effect in 2015 is still weaker than other factors. We further observed that Northern China clearly featured a stronger correlation with NGEC than other regions.
4.2 The spatial effects of economic factors on NGEC

As economic factors in our model, both PGDP and FDI can represent regional economic development (Jiang et al., 2018; Hille et al., 2019). Based on the OLS results, PGDP has been identified as one of the crucial factors resulting in NGEC increase (Amri, 2017; Ding et al., 2017), while FDI shows a minor negative effect on national NGEC, which is similar to the conclusion of Jiang et al. (2018).

However, we observed that GDP growth has always dominantly affected the change of NGEC in China over the past ten years (Fig. 6d), although such effect is slowly dropping (from 0.37–0.49 to 0.29–0.41). Meanwhile, there is a significant spatial correlation effect from PGDP, where a stronger correlation happens in Northern and Eastern China (Fig. 10).
Comparably, FDI only shows a weak effect (−0.17–0.05) on NGEC (Fig. 6e), which worked positively in Western China and negatively in Eastern, Northern, Northeast China in 2015 (Fig. 11). Such finding is compliant with the “Pollution Haven Hypothesis” (Hao et al., 2018; Shahbaz et al., 2015) when applied to Western China (e.g., Xinjiang, Qinghai, Gansu, Sichuan, Yunnan, Chongqing, and Guizhou). There is, however, lack of evidence that FDI can lead to more air pollution across all regions. In general, FDI can either contribute to the introduction of environmentally-friendly or clean energy technologies in some cases, be associated with high pollution industries and cause more energy consumption and NGEC. As a result, it is crucial to set up the regionalised environmental regulations towards FDI entry strategies in Western China, which shall consider the environmental capacity for the region and the categories of investing industries.
4.3 The spatial influence of population factors on NGEC

The population factor is the most correlative (0.59–0.67) factor on NGEC, i.e., a bigger population size correlated to more NGEC (Diao et al., 2018; Wang et al., 2018; Wang et al., 2017b). We found that PS's effect rapidly decreases in China over the past decades (Fig. 6f). We also observe a shift from the east-west difference (2005) to the north-south difference (2015) in the relationship between PS and NGEC (Fig. 12).

Actually, these results bring some meaningful signals to policy-makers. First, the mitigation policy should give more attention on the changing relationship in the Southern China, where the increasing NGEC was drifting away from the impact of local population size over time, i.e., more people in the region no longer means a worse environment. For example, in Guangdong province, the contribution of tertiary industry surpassed the secondary industry in 2013. Such upgrading industrial structure has driven the change of the employment structure, which means more labour forces shifted from agriculture industry and manufacture industry to the service industry. There is a different case in the Beijing-Tianjin-Hebei regions, where PS is still closely linked with high EI manufacturing, thus has driven more NGEC from 2005 to 2015. These findings call for further research on the nexus of population-urbanization-employment and its impact on industrial structure and NO\textsubscript{x} generation.

4.4 Identifying the spatial differentiation of the impact mechanism

In order to obtain a comprehensive overview of impactors, this section further classified China into six sub-national control zones of NO\textsubscript{x} (Table 5) by comparing driving forces on provincial NGEC. We defined the
concept of “average regional influence” based the average value of the regression coefficients of these provinces in the same region (Xu and Lin, 2018). Based on that, we ranked all the driving factors by their impacts on NGEC and took the top three as the dominant factors in every region (Table 5), which should be given a priority concern in each NO\textsubscript{x} control zone.

As mentioned earlier, PS has been always the dominant factor correlating with NGEC over the past ten years across all provinces in China. Additionally, we found that there was an evolving spatial-homogeneity of the impact mechanism on the NGEC. In 2005, there no apparent spatial difference across Central China, Eastern China, Northern China, and Northeast China, where the PGDP led NGEC were more significant than the impact of TE, but in Western China and Southern China, TE takes the place of PGDP. In 2010, TE became the second important factor instead of PGDP across most regions, other than Northeast China. However, in 2015, EI and PGDP ranked second and third across most regions, except for Northeast China. These results clearly show the change of dominating driving forces on NGEC over the past years.

Because the impact mechanism varied around China and is very different from the OLS model results, which only focus on the homogeneity level. Our research conveys important messages in managing China's NO\textsubscript{x} emission. We know that EI has gradually grown into the dominant factor across all provinces, so the progress of EI would become increasingly essential for the control of NO\textsubscript{x}. Policy-makers should balance well PGDP

### Table 5

<table>
<thead>
<tr>
<th>Regions</th>
<th>Covered provinces</th>
<th>Dominating factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2005</td>
</tr>
<tr>
<td>National scope</td>
<td>All</td>
<td>PGT</td>
</tr>
<tr>
<td>Northeast China</td>
<td>Heilongjiang, Jilin, Liaoning</td>
<td>PGT</td>
</tr>
<tr>
<td>Northern China</td>
<td>Beijing, Tianjin, Hebei, Shanxi, Shandong, Henan</td>
<td>PGT</td>
</tr>
<tr>
<td>Central China</td>
<td>Hubei, Jiangxi, Hunan</td>
<td>PGT</td>
</tr>
<tr>
<td>Eastern China</td>
<td>Jiangsu, Anhui, Shanghai</td>
<td>PTG</td>
</tr>
<tr>
<td>Southern China</td>
<td>Fujian, Guangdong, Hainan</td>
<td>PTG</td>
</tr>
<tr>
<td>Western China</td>
<td>Eleven other provinces</td>
<td>PTG</td>
</tr>
</tbody>
</table>

*: P represents population scale, G represents GDP per capita, T represents thermal power generation, and E represents energy intensity.
and energy consumption in the future. Especially, Northeast China should get more attention from policies because its energy-intensive industry, such as petrochemical, energy and metallurgical, heavily relies on TE.

4.5 Policy implications

Our findings not only fill in the knowledge gap of the socioeconomic context of NO\textsubscript{x} generations but also generate insights for designing and implementing air quality policies in China.

Firstly, we have observed that China's continuous energy policy of improving energy intensity has a growing effect on reducing NO\textsubscript{x} generation, regardless of regional variety. We thus suggest that it is important to carry on stringent regulation on energy intensity due to its synergistic impact on reducing both energy consumption and air pollution. These findings are particularly useful to Northern China and Southern China, where energy intensity has become the most dominant factor to influence NO\textsubscript{x} generation.

Secondly, our findings call for a reassessment on the policy to increase natural gas consumption ratio all over China. Winter heating and manufacturing industries in Northern China have become a burden for local air quality management. In 2015, China initiated a drastic, statewide coal-to-gas initiative, which aimed to significantly increase the numbers and scale of natural gas-based power plants and heating facilities, in order to reshape energy structure. In July 2019, the initiative was called off and is currently under evaluation. Our research contributes to such evaluation because we found that, before this initiative, i.e., during 2005–2015, natural gas consumption rate has a strong prohibiting effect over NO\textsubscript{x} generation only in Northern and Northeast China. Such an effect is, however, not strong enough to replace the impact of energy intensity from 2010 to 2015. Thus, there is no evidence that the coal-to-gas initiative, if carrying on, has a positive effect of reducing NO\textsubscript{x} generation.

Thirdly, our findings are linking the regional economic transition context towards NO\textsubscript{x} generation. We found that Northern China is experiencing a much faster decouple process between GDP growth and NGEC comparing with Southern China, but such impact gap has become smaller recently. The significance of such a decoupling effect in Northern China is due to the dominant share of energy-intensive industries in those regions. In contrast, Guangdong province as part of Southern China, for example, has not been dependent on energy-intensive industries in the past decade, and thus shows a different decoupling curve with NGEC (Wu et al., 2017) than Northern China. Such location-specific and spatially varying finds are different from previous studies which mainly implicitly presumed that economic development for all provinces has an unvarying impact on NO\textsubscript{x} (Diao et al., 2018; Ge et al., 2018; Zhang et al., 2019b).

4.6 Uncertainty analysis

To examine the reliability of our results and the uncertainty of data, we conducted a preliminary uncertainty analysis outlined below. In this study, the uncertainty is mainly associated with the estimated NGEC, socioeconomic determinants and the modelling. First, NGEC was calculated by the bottom-up emission methods (Eq. (1)): count various types of energy consumption in various industries, and this could avoid double-counting about non-fuel use of energy, for example, NO\textsubscript{x} generated during the production process. This estimating method has been accepted by the academic and widely used in the previous study. Moreover,
NO\textsubscript{x} generation factor, distinguished by different economic sectors and energy types, was obtained from the previous studies (Kato and Akimoto, 1992; Hao et al., 2002) and it has been widely applied in many studies (Tian et al., 2001; Gao et al., 2006; Jiang et al., 2016), which could prove it reliable and scientific. Meantime, NO\textsubscript{x} generation factors can cause the uncertainty of the estimated NGEC. NO\textsubscript{x} generation factor was measured by actual measurement, which may cause additional uncertainty due to differences in the accuracy of the actual measurements. Second, uncertainties and errors associated with socioeconomic determinants (Table 1) may stem from officially published statistical yearbooks in China (Bai et al., 2018), for example, statistics might exist deviation in the statistical process. Third, as we know, models are simplified representations of real-world systems; they typically do not always mimic actual conditions. Variables used for modelling would introduce potential uncertainties. For example, uncertainty in the variability of urban density results from spatial data accuracy. Meanwhile, for the uncertainties of modelling, although models with high explanatory power and the significances of most variables reached the level of <0.05, Residual Sum of Squares still exists, and thus models carried some uncertainties.

5 Conclusions

The control on NO\textsubscript{x} generation is centring now in China's air pollution policy system because of its increasing concern about air quality. Considering the limitation of the OLS model on estimating the spatial agglomeration of NO\textsubscript{x}, this study applied the geographically weighted regression (GWR) model to analyse nonstationary spatial effects, including the spatial difference and spatial agglomeration of EI, clean energy structure, FDI, economic growth, and the effect of PS on NGEC. We observed that NGEC presents an increasing spatial heterogeneity varying from Northeast China to Southern China. The spatial aggregation with the highest NO\textsubscript{x} generation clearly concentrated in Shandong, Jiangsu, and its surrounding areas were shown.

We found that energy intensity has always shown a strong and positive correlation with spatial difference and agglomeration of NGEC over the past years, especially in Northern and Northeast China. Thermal power plants as an important contributor have a decreasing impact on NGEC, but there still keep the impact gap between Western and Eastern China. Differently, the spatial distribution of another energy factor, nature gas shares, freshly correlated the spatial NO\textsubscript{x} generation only in 2015 and featured by a south-north gap. Usually, PGDP of economic development often plays a crucial role in NGEC. However, we found that it did not always work this way because of its west-east difference of spatial correlation. However, the GWR model results show that another economic factor FDI correlated NGEC only positively in Western China and but negatively in Eastern, Northern, Northeast China in 2015. Moreover, the size of the PS in Northern and Northeast China shows positively correlated to high NGEC, although its effect was dropping across the last ten years. However, it did not happen in Southern China, including Guangdong province, which resulted mostly from the shift of labour force across the industry sector, which could affect the NGEC generation downward.

Conflicts of interest

None.

Uncited reference
Acknowledgements

This work was supported by the Tianjin Natural Science Foundation [grant number 18JCZDJC39900], the National Natural Science Foundation of China [grant numbers 71373134], the Special Foundation to build universities of Tianjin [grant number C0291760], the China Scholarship Council [grant number 201806200003], the Tianjin Natural Science Foundation [grant number 18JCZDJC39900] and the Fundamental Research Funds for the Central Universities.

References


Huang, R.J., Zhang, Y., Bozzetti, C., Ho, K.F., Cao, J.J., Han, Y., Zotter, P., 2014. High secondary aerosol contribution to particulate pollution during haze events in China. ON Nat. 514 (7521), 218.


Xu, B., Lin, B., 2018. Do we really understand the development of China’s new energy industry? Energy Econ. 74, 733–745.


**Footnotes**

**Text Footnotes**

[1] Standardized variables are variables that have been standardized to have a mean of 0 and a standard deviation of 1. Standardizing makes it easier to compare, even if variables were measured on different scales.

[2] China's Five-Year Plans are a series of social and economic development initiatives, which can reflect the socioeconomic development trend of China's NGEC.
Highlights

- We examined spatially differentiated effects of NO\textsubscript{x} generation using GWR.
- Economic growth and energy intensity showed the largest south-north spatial effects.
- PS always dominantly affected spatial NGEC with a minor north-south difference.
- NGR correlated significantly with NGEC through a weakly negative effect in 2015.
- Spatially differentiated reduction policies need region-wise redesigning.

Queries and Answers

**Query:** Your article is registered as a regular item and is being processed for inclusion in a regular issue of the journal. If this is NOT correct and your article belongs to a Special Issue/Collection please contact p.sivakumar@elsevier.com immediately prior to returning your corrections.

**Answer:** Yes

**Query:** Please confirm that the provided emails “jfwangnk@126.com, lingxuan.liu@lancaster.ac.uk” are the correct address for official communication, else provide an alternate e-mail address to replace the existing one, because private e-mail addresses should not be used in articles as the address for communication.

**Answer:** Yes, the provided emails are correct.

**Query:** Please confirm that given names and surnames have been identified correctly and are presented in the desired order and please carefully verify the spelling of all authors' names.

**Answer:** Yes

**Query:** The citations “Wang et al. (2013); Hao et al., 2016; Wang et al. (2016); Anselin, 1988; Wheeler, 2010” have been changed to match the author name in the reference list. Please check here and in subsequent occurrences.

**Answer:** Yes.

**Query:** For figure(s) 1,10,11,2,3,4,5,6,7,8,9, More than one set of versions have been provided. Please confirm that we have used the correct one.

**Answer:** Yes, we have checked all the figures and used the correct one.

**Query:** Please note that ‘Fig. 3’ was not cited in the text. Please check that the citation(s) suggested by the copyeditor are in the appropriate place, and correct if necessary.
**Answer:** We have checked and deleted the citation suggested by the copyeditor, and added "Fig. 3" in the appropriate place. Please see the details in the paper.

**Query:** The citations “Wang et al., 2018a” has been changed to match the date in the reference list. Please check here and in subsequent occurrences.

**Answer:** Yes.

**Query:** Uncited references: This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or, alternatively, delete it. Any reference not dealt with will be retained in this section. Thank you.

**Answer:** We have checked the "Pasquier and André, 2016" and deleted it in the reference.

**Query:** Have we correctly interpreted the following funding source(s) and country names you cited in your article: China Scholarship Council, China; National Natural Science Foundation of China, China; Fundamental Research Funds for the Central Universities; Tianjin Natural Science Foundation, China?

**Answer:** Yes

**Query:** Please provide the volume number or issue number or page range or article number for the bibliography in Ref(s). Hille et al., 2019.

**Answer:** Yes, we have added the the volume number and the page range.