Spatial Relationships between Urban Structures and Air Pollution in Korea

Meen Chel Jung 1, Jaewoo Park 2 and Sunghwan Kim 3,*

1 Department of Urban Design and Planning, University of Washington, 3950 University Way NE, Seattle, WA 98105, USA; jmc129@uw.edu
2 Department of Statistics, The Pennsylvania State University, 422 Thomas Building, University Park, PA 16802, USA; jzp191@psu.edu
3 Construction and Economy Research Institute of Korea, 711 Eonju-ro, Seoul 06050, Korea
* Correspondence: shkim@cerik.re.kr; Tel.: +82-2-3441-0734

Received: 15 December 2018; Accepted: 14 January 2019; Published: 17 January 2019

Abstract: Urban structures facilitate human activities and interactions but are also a main source of air pollutants; hence, investigating the relationship between urban structures and air pollution is crucial. The lack of an acceptable general model poses significant challenges to investigations on the underlying mechanisms, and this gap fuels our motivation to analyze the relationships between urban structures and the emissions of four air pollutants, including nitrogen oxides, sulfur oxides, and two types of particulate matter, in Korea. We first conduct exploratory data analysis to detect the global and local spatial dependencies of air pollutants and apply Bayesian spatial regression models to examine the spatial relationship between each air pollutant and urban structure covariates. In particular, we use population, commercial area, industrial area, park area, road length, total land surface, and gross regional domestic product per person as spatial covariates of interest. Except for park area and road length, most covariates have significant positive relationships with air pollutants ranging from 0 to 1, which indicates that urbanization does not result in a one-to-one negative influence on air pollution. Findings suggest that the government should consider the degree of urban structures and air pollutants by region to achieve sustainable development.

Keywords: urban structure; air pollution; Moran’s I; Bayesian spatial model

1. Introduction

Urban structures have become increasingly complicated due to economic factors, such as agglomeration economies and externalities [1–4]. The composition and configuration of land-use types and transportation networks are planned to achieve the best economic growth of regions. For example, Pan et al. [5] found that the distribution of residential and commercial land uses is planned along with network systems in the Chicago Metropolitan Statistical Area. Although each city has a different urban structure based on regional socioeconomic status [6], achieving economic development is a common goal. Given this trend, population concentrations have resulted in 54% of the world’s population living in urban areas [7]. Although large populations and high employment rates indicate that a region has a high level of economic sustainability, many urban regions have begun to experience environmental problems, in particular serious air pollution [8].

Most air pollutants are generated from human activities, including production and consumption, and interactions between the human social system and environment [9,10]. Given that urban structures facilitate human activities and their interactions [11], investigating the relationship between the characteristics of these structures and air pollution is of broad interest to many. In this manuscript, we study this relationship in the context of Korean cities by considering multiple urban structure
covariates and air pollutant emissions simultaneously. We also provide a general theoretical framework for countries entering the terminal stage of urbanization based on this empirical analysis.

The sustainable development paradigm is crucial in terms of the competitiveness of a region [12]; hence, modern governments require different policy formulations and implementations, as opposed to past governments, which were able to implement strong economic growth-oriented policies. For example, Korea achieved rapid economic growth by intensively developing specific areas and focusing on heavy industries. However, this government-led plan was criticized because it neglected all other value except economic growth. Therefore, recent governments have attempted to incorporate environmental value into policy making to attain sustainable development by introducing the smart city concept. Low air quality considerably influences current policy making because residents tend to respond sensitively to changes in air quality [11,13]. Air pollutants can affect public health in direct and indirect ways [14–19]; thus, the need for research on air pollution is strong.

Several studies have been conducted to observe the air quality of various regions [20–24]. Jang [20] found that the atmosphere in Seoul does not satisfy air quality standards due to particle pollution. Fine dust problems, which have become a major social issue in China, affect Korea’s air situation because of the latter’s geographical location. Many urban areas in China show levels of PM2.5 higher than those in other countries [25], and the proportion of particulate matter is higher than the standard suggested by the World Health Organization [26,27]. However, previous studies are limited by the fact that they only observe air pollution in specific regions and do not investigate the relationship between urban characteristics and air quality. Given that the atmosphere over urban areas has become overburdened with various pollutants, analyzing the relationship between urban structures and air pollution is necessary.

Complex urban spaces can be divided into fixed-use areas, such as residential, commercial, and industrial areas, and link-based transportation networks connecting these areas [11]. Since the distributions of complex urban spaces affect air quality [8], attempts to establish the relationship between urban structures and air quality through theoretical models have been made. For example, Borrego et al. [8] conducted simulation studies under different urban planning scenarios to investigate these relationships. Marquez and Smith [11] developed an integrated land use-transport-air quality model to evaluate the effect of urban structure on air quality.

Given the increasing availability of relevant data, a number of empirical studies have analyzed the influence of urban structures on air quality [28–39]. Empirical studies can be classified into two categories: (1) those that conceptualize the types of land use as urban structures [28–32] and, (2) those that regard the degree of sprawl or compactness as the urban structures [33–39]. In this work, we will refer to the first category as the land-use approach and the second category as the sprawl-and-compact development approach. With regard to land-use approaches, Hussein et al. [28] investigated the association between traffic-related variables and sizes of ambient particulate matter. Sánchez-Rodas et al. [30] analyzed the distributions of PM10 in rapidly industrializing regions. Furthermore, levels of air pollutants in green areas have been examined with an increasing preference for easily accessible green spaces in modern cities [31,32]. On the other hand, the emergence of decentralized sprawl and its corresponding concept of compact development have led to attempts to find a relationship with air quality. Cho and Choi [33] conducted panel analyses to evaluate the influence of compact development on air quality in Seoul. The proportion of green areas and number of people within the built-up area variables were used to conceptualize urban compactness. The results revealed no apparent impact of compact development on air pollution levels. The same approach was applied to metropolitan areas or urban regions rather than to an individual city. Using macroscopic structure variables such as centeredness, connectivity, land-use mix, and sprawl index, the researchers found that the amounts of air pollutants increase as the degree of sprawl increases in most metropolitan areas [34–37].

Although previous studies indicate a significant relationship between urban structures and air pollution, both categories of empirical studies have limitations. In the land-use approach, studies
examining comprehensive relationships using multiple urban structure variables are scarce. In addition, direct application of the sprawl-and-compact development approach to Korean cities is challenging because previous theories are based on North American cities. These gaps between the existing hypotheses and empirical analyses fuel our motivation to investigate the administrative area in Korea by using multiple urban land-use covariates and air pollutant variables simultaneously.

Air quality can be measured by estimating the amounts of emissions from pollutant sources or measuring the concentrations of pollutants by monitoring stations [40]. Among the various pollutants, we focus on emissions of nitrogen oxides (NO\textsubscript{X}), sulfur oxides (SO\textsubscript{X}), and two types of particulate matter (PM10 and PM2.5), all of which are meaningful criteria for air pollution. NO\textsubscript{X} are mainly produced in combustion processes, such as in automobile and power plants, while SO\textsubscript{X} primary sources include fossil fuel combustion and natural emissions [41]. PM10 is mainly obtained from the mechanical crushing and abrasion of surfaces while PM2.5 is produced in combustion processes, similar to NO\textsubscript{X} [42]. The phenomenon of high particulate matter concentration repeats every spring and winter has recently been observed in Korea, and studies on the adverse effects of particulate matter have been conducted. In response to findings, the government implemented new policies, such as the “Clean Air Conservation Act” and the “Special Act on the Reduction and Management of Dust,” as regulation measures. Today, NO\textsubscript{X} and SO\textsubscript{X}, which react with other substances in the air, are considered criteria for air pollution in Korea.

Among Organization for Economic Co-operation and Development (OECD) countries, Korea has the highest population density with more than 500 people per km\textsuperscript{2}, which means the country may be particularly vulnerable to air pollution. According to the 2018 Environmental Performance Index, Korea ranked 119th out of 180 countries in air quality; in particular, Korea’s PM2.5-related indicators were at the bottom. However, regional differences in air pollution have been determined due to differences in industrial structures and population densities by region. Therefore, we use spatial regression models to describe the general trend over the entirety of Korea and account for regional characteristics. We first analyze the spatial distributions of air pollutants, namely, NO\textsubscript{X}, SO\textsubscript{X}, PM10, and PM2.5, from pollutant sources in 225 administrative areas in Korea to determine the existence of spatial dependencies of air pollutants. On the basis of these results, we fit the spatial regression models to investigate the relationship between the structural components of cities and air pollutant emissions. In this analysis, four air pollutants and several urban structure variables are used as response variables and covariates, respectively. To the best of our knowledge, no existing approach provides a general spatial model framework for quantifying the relationships between different types of air pollutants and urban structure covariates in Korea. Finally, we suggest policies to achieve sustainable urban areas based on our empirical analyses.

The rest of this paper is organized as follows. In Section 2, we introduce the study area and provide the research background. We also describe the variables and visualize the distributions of four air pollutants. We describe the methodologies applied to perform exploratory spatial data analysis and build the Bayesian spatial linear model in Section 3. In Section 4, we present the main findings of our empirical analysis. We conclude with a discussion and conclusions in Section 5.

2. Data

2.1. Study Area

The spatial extent of this study is Korea (Figure 1). Korea underwent rapid urbanization along with great economic growth over the last 50 years, and, as a result, five major metropolitan areas were formed. To account for spatial uncertainties across the entire region, we set the extent of this study to include the Korean mainland, rather than focusing on a specific region, such as the Seoul metropolitan area.
Korea has a four-level hierarchical administrative structure; in this study, the second-level administrative area, namely, Si (city)-Gun (county)-Gu (district) (SGG), is set as the spatial unit. An SGG corresponds to an administrative county in the United States. However, SGG presents fewer restrictions on the modifiable areal unit problem because the average area of SGGs (438 km$^2$) is smaller than that of US counties (2,584 km$^2$). In this study, the SGG is used as the unit of analysis. Korea consists of 229 SGGs, but four of them (Ongjin, Ulleung, Jeju, and Seogwipo) are excluded from this analysis because they are composed of islands located far from the mainland (Appendix A Figure A1). In addition, these SGGs consistently emit marginal amounts of air pollutants. Therefore, 225 SGGs are used for our analysis.

2.2. Data and Variable

We consider four air pollutants, namely, NOX, SOX, PM10, and PM2.5, as response variables and use population size, area of commercial land use, area of industrial land use, area of park, road length, total land surface, and gross regional domestic product (GRDP) per person as covariates in this study. The number of observations for each variable is 225, and 2015 is set as the base year for all variables. In this section, we provide a detailed description of each variable.

Data on the four air pollutants by SGG are obtained from the National Air Pollutants Emission Service of the National Institute of Environmental Research. This annual dataset includes estimated amounts of air pollutants from 13 major pollutant sources by region (Appendix A Table A1). Except for those on PM2.5, which were collected beginning 2011, data on the three other pollutants were collected from 1999 to 2015. However, analyzing the panel data is challenging because of discontinuities in emission estimates due to annual enhancement of emission factors and addition of new sources or removal of existing sources. Therefore, only the latest dataset (2015) is used for the analysis.

Another dataset on air pollutants is estimated not on the basis of pollutant sources but, rather, on the measurement instruments. Such data are not used in this study due to limitations associated with the control of external factors, including weather conditions. Several studies have found a strong relationship between the concentrations of air pollution and climatic variables, such as wind direction and speed [20,43–47]. In particular, Korea is heavily affected by pollutants from Northeast Asian regions due to its geographical characteristics and westerlies. Therefore, air pollutant data based on measurement instruments can contain air pollutants from both domestic and foreign sources, which can distort the true relationship between the country’s urban structures and air pollutant emissions. To avoid this problem, we use 1-year cross-sectional estimated emission data from domestic pollutant sources.

In this manuscript, we consider multiple covariates to account for complex urban structures. First, we include the population size of SGGs in the models. Population size has been used for scaling...
analysis to reveal diverse urban indicators [48–51]. The population can be used to determine the intensity of various socioeconomic activities causing complex urban structures. Second, we choose areas of commercial and industrial land use as covariates. Although the population can represent the area of residential land use, it cannot adequately cover commercial or industrial areas. For example, if SGGs near major metropolitan areas are bedroom towns, they must have a high proportion of residential area with a low proportion of commercial and industrial areas. By contrast, if SGGs are central business districts (CBDs), they must have a high proportion of commercial space and a low proportion of residential area due to the CBD hollowing-out phenomenon. Therefore, we use areas of commercial and industrial land use to account for the different regional characteristics of urban structures in each SGG. Third, we include the area of green spaces in SGGs in the models. In the context of environmental sustainability, the importance of green spaces to residents has been increasingly emphasized [31,32]. A potential variable for this approach is the area of a green zone defined by the National Land Planning and Utilization Act. However, this method is unsuitable because it includes areas of agricultural lands, such as farmland and production forest land, which are planted with crops focused on economic, rather than environmental, considerations. Hence, we use the area of parks defined as urban parks as a variable. Fourth, the length of all roads in SGGs regardless of pavement type is included in the models. The inclusion of road length in the analysis is reasonable because a considerable volume of air pollutants is directly emitted by the transportation sector [28]. Fifth, to control potential bias due to the different surfaces of each SGG, we include the total land area of each SGG in our models as in [37–39]. Finally, we include per capita real GRDP by SGG for 2015, the base year of which is 2010, to control the economic level on the relationship between urban structure and air pollutants. All urban structure variables are obtained from the Korean Statistical Information Service of Statistics Korea. The descriptive statistics and distributions of the original variables are summarized in Table 1 and Figure 2, respectively. Four air pollutants and seven covariates are transformed into a natural logarithm before descriptive and empirical analyses to improve normality [52–54].

### Table 1. Descriptive statistics of the variables.

<table>
<thead>
<tr>
<th>No. of Observations: 225</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOX (kg/year)</td>
<td>4,696,321</td>
<td>6,633,836</td>
<td>364,356</td>
<td>46,956,502</td>
</tr>
<tr>
<td>SOX (kg/year)</td>
<td>1,487,659</td>
<td>4,943,004</td>
<td>2,526.99</td>
<td>32,832,287</td>
</tr>
<tr>
<td>PM10 (kg/year)</td>
<td>1,004,696</td>
<td>2,520,089</td>
<td>49,270.75</td>
<td>27,237,038</td>
</tr>
<tr>
<td>PM2.5 (kg/year)</td>
<td>420,354.4</td>
<td>1,350,711</td>
<td>19,340.97</td>
<td>14,346,234</td>
</tr>
<tr>
<td><strong>Covariates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population (person)</td>
<td>226,105.9</td>
<td>218,859.7</td>
<td>17,898</td>
<td>1,184,624</td>
</tr>
<tr>
<td>Commercial area (m²)</td>
<td>1,433,670</td>
<td>1,466,489</td>
<td>111,167</td>
<td>10,650,332</td>
</tr>
<tr>
<td>Industrial area (m²)</td>
<td>5,109,502</td>
<td>8,598,202</td>
<td>0</td>
<td>57,804,833</td>
</tr>
<tr>
<td>Park area (m²)</td>
<td>3,318,364</td>
<td>3,801,371</td>
<td>32,000</td>
<td>24,602,000</td>
</tr>
<tr>
<td>Road length (m)</td>
<td>426,654.3</td>
<td>229,819.3</td>
<td>51,497</td>
<td>1,734,727</td>
</tr>
<tr>
<td>Total land surface (m²)</td>
<td>436,449,590</td>
<td>381,005,832</td>
<td>2,826,064</td>
<td>1,819,829,902</td>
</tr>
<tr>
<td>Gross regional domestic product (GRDP) per person (1,000,000 (KRW))</td>
<td>30,559</td>
<td>31.82</td>
<td>6.194</td>
<td>395.476</td>
</tr>
</tbody>
</table>

The Korea Geographic Information Service (GIS) map used to present air pollutant emissions and local spatial autocorrelations is downloaded from the Statistical Geographic Information Service of Statistics Korea. All explanatory spatial data analyses are performed using ArcGIS 10.3 by ESRI.
Figure 2. Regional choropleth maps of (a) Population, (b) Commercial area, (c) Industrial area, (d) Park area, (e) Road length, and (f) GRDP per person.

2.3. Distributions of Air Pollutants

The distributions of the four air pollutants are illustrated in Figure 3. In terms of NO$_X$ and SO$_X$, SGGs show a similar pattern of emission distribution. For example, areas marked by diagonal lines included in the lowest and second-lowest categories (e.g., southern coastal SGGs, southern inland SGGs, eastern coastal SGGs, and northeastern SGGs) emit relatively low amounts of pollutants (Figure 3a,b). By contrast, eastern coastal SGGs (37°–38° north latitude), which are positioned above the diagonal-line regions, southeastern SGGs, and midwest coastal SGGs emit high amounts of NO$_X$ and SO$_X$. These SGGs possess manufacturing-oriented industrial complexes, thermoelectric power plants, and coal-fired power plants that are over 30 years old. PM$_{10}$ emissions are generally higher than PM$_{2.5}$ emissions, but some SGGs are included in the largest category of both types of particulate matter. Although overall patterns are similar among the four air pollutants, this finding does not mean that all individual SGGs have similar levels of emission of each pollutant. In other words, the amounts of pollutants emitted by each SGG may vary. For example, two SGGs (Mungyeong and Yecheon), which feature horizontal stripes in all pollutants, have relatively high emissions of SO$_X$ and PM$_{10}$ (included in the top 50% categories) and low emissions of NO$_X$ and PM$_{2.5}$ (included in the bottom 50% categories); by contrast, other SGGs show the opposite case. These results confirm the necessity of considering regional characteristics when analyzing air pollution.
In summary, we present three findings. (1) Although the four air pollutants are produced by different sources and mechanisms, the overall emissions of SGGs present similar patterns. (2) Some SGGs present different amounts of pollutant emissions due to regional differences in energy consumption and industrial structure. (3) Some adjacent SGGs that produce similar degrees of emissions form spatial clusters. On the basis of these findings, exploratory spatial data analyses are conducted, and spatial linear models are fitted for further analysis.

**Figure 3.** Emissions of (a) \(\text{NO}_X\), (b) \(\text{SO}_X\), (c) PM10, and (d) PM2.5. Regions with diagonal lines have low emissions of \(\text{NO}_X\) and \(\text{SO}_X\). Horizontal stripes in administrative areas (SGGs) show high emissions of \(\text{SO}_X\) and PM10 and low emissions of \(\text{NO}_X\) and PM2.5.
3. Methodology

3.1. Exploratory Spatial Data Analysis

Statistical inference in the presence of spatial autocorrelation is challenging because dependencies among nearby SGGs can lead to unreliable parameter estimates. For example, Dong and Liang [55] noted the presence of global and local spatial dependencies in emissions of various pollutants in China, and Réquia et al. [56] discovered the existence of clusters in various types of pollutant emissions. In this study, spatial autocorrelations of emissions can arise in urban structures, such as in industrial complexes or power plants where emission sources are usually agglomerated due to economic reasons. To determine if spatial autocorrelation issues exist in our study, we use Global Moran’s I, a popular nonparametric statistics of spatial autocorrelation [57]. Moran’s I can detect spatial dependence by calculating the correlation among nearby regions. If Moran’s I is statistically significant, spatial linear regression models should be fitted to account for such dependence among the data.

We calculate Global Moran’s I to check whether or not global spatial autocorrelations occur. Although Global Moran’s I can be a proxy for global spatial autocorrelation, it does not represent local clusters. Local deviations from a global pattern of spatial autocorrelation can occur even if the entire study area has a significant spatial autocorrelation. Thus, we also apply Local Moran’s I, which decomposes Global Moran’s I, to identify local clusters of each pollutant [58]. A positive value of Local Moran’s I means that the emission values for an SGG are similar to those of adjacent SGGs. Then, this SGG becomes part of a high-high or low-low cluster. By contrast, a negative value of Local Moran’s I represents dissimilar emission values among nearby regions. Then, this SGG is part of a high-low or low-high outlier. The details for the spatial weight matrix for Moran’s I are provided in Section 3.2.

3.2. Spatial Linear Regression Model

In the presence of spatial autocorrelations, adding a set of spatially correlated random effects to the linear predictor is the most common approach to account for such dependencies. These random effects are typically modeled with a conditional autoregressive (CAR) prior [59], which can incorporate spatial dependence among neighboring areal units. Lee [60] compared several variants of CAR models and identified the model proposed by Leroux et al. [61] as the most general and practical approach; therefore, we use this model in our analysis. The hierarchical structure of spatial linear models renders the Bayesian approach convenient to use. For example, Bayesian spatial regression models have been widely used to analyze the relationship between air pollution and health in Scotland [62] and estimate the effects of air pollution on respiratory hospital admissions in London [63]. Liu et al. [64] predicted air quality in Xiamen, China, by using a Bayesian hierarchical model.

In this context, Bayesian spatial linear regression models are fitted to examine the relationship between the covariates and response variables, namely, NO\textsubscript{X}, SO\textsubscript{X}, PM10, and PM2.5 models. We have n = 225 number of observations across the spatial domain and p = 7 number of predictors. We let Y be a 225 \times 1 response variable vector and X be a 225 \times 7 matrix of values of covariates. We note that Y is log-transformed air pollutants and X is log-transformed urban structure and control variables. Our models can be written as follows:

\[ Y = X\beta + \varphi, \]

where \( \beta \) is a regression coefficient and \( \varphi \) is a spatial random effect. \( \varphi \) is modeled by a CAR [59,61] prior distribution to incorporate spatial dependence among observations as follows:

\[ \varphi_i | \varphi_{-i} , W, \tau^2, \rho \sim N\left( \frac{\rho \sum_{j=1}^{n} W_{ij} \varphi_j}{\rho \sum_{j=1}^{n} W_{ij} + 1 - \rho}, \frac{\tau^2}{\rho \sum_{j=1}^{n} W_{ij} + 1 - \rho} \right), \]
where $\rho$ is a spatial autocorrelation parameter and $\tau^2$ is a variance parameter. The covariance structure of $\phi$ is determined by selecting spatial weight matrix $W$ that summarizes the spatial associations among observations. The priors for our parameters of interest are:

$$
\beta \sim N(0, 100^2), \tau^2 \sim \text{Inverse Gamma}(1, 0.01), \rho \sim \text{Uniform}(0, 1).
$$  \hspace{1cm} (3)

We infer the model parameters $(\beta, \tau^2, \rho)$ via a Bayesian approach by using a Markov chain Monte Carlo (MCMC) algorithm. The code for statistical analysis is implemented by using the CARBayes [65] package in R [66].

For spatial weight matrix $W$, we use k-nearest neighbor weights based on the Euclidean distance of the centroid of areal units. An SGG is unlikely to have more than eight neighbors; therefore, we set the range of $k$ from 3 to 8. We use the Bayesian information criterion (BIC) to select $k$ for each model. The implementation proceeds as follows. (1) We fit spatial regression models for each of the response variables ($\text{NO}_X, \text{SO}_X, \text{PM10}, \text{PM2.5}$) with different $k$ values. (2) BIC values are calculated, and a small BIC indicates a good model. We then select models with the smallest BIC for each of the response variables. (3) To validate our analysis, we calculate the Global Moran’s $I$ for each model’s residuals. When the Global Moran’s $I$ value is significantly small, the models successfully account for spatial autocorrelation.

4. Results

The Global Moran’s $I$ values for air pollutants are summarized in Table 2. All values are significantly positive, which means that pollutants are clustered together. In particular, the Global Moran’s $I$ value for PM2.5 is higher than that for other pollutants. Hence, PM2.5 is more clustered than the other pollutants are. Our result reveals the existence of spatial autocorrelation in emissions and thus provides theoretical justifications for using spatial linear regression models.

<table>
<thead>
<tr>
<th>NO\textsubscript{X}</th>
<th>SO\textsubscript{X}</th>
<th>PM10</th>
<th>PM2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Moran’s $I$</td>
<td>0.177 *</td>
<td>0.191 *</td>
<td>0.192 *</td>
</tr>
<tr>
<td>($&lt;0.001$)</td>
<td>($&lt;0.001$)</td>
<td>($&lt;0.001$)</td>
<td>($&lt;0.001$)</td>
</tr>
</tbody>
</table>

* indicate statistically significant estimates.

We also examine the local clusters and outliers of each pollutant (Figure 4). Each air pollutant has a different local spatial pattern. For example, most of the high-high clusters of NO\textsubscript{X} are located in western coastal areas, and many of the high-high clusters of SO\textsubscript{X} are at 37º north latitude, especially in either eastern or western coastal regions. With regard to two types of particulate matter, most of the high-high clusters of PM10 and PM2.5 are at 37º north latitude or in eastern coastal areas at 36º north latitude. In terms of low-low clusters, NO\textsubscript{X} have more clusters than other pollutants do, and the clusters are distributed all over the study area. Meanwhile, SO\textsubscript{X} and PM2.5 have distinctly few low-low clusters indicating that an SGG with low emission of either SO\textsubscript{X} or PM2.5 does not cluster together.

Several implications for further analysis are deduced on the basis of the spatial pattern of clusters. (1) The distributions of statistically calculated clusters are in discord with those of descriptive clusters in Figure 3 to some degree. For example, not all diagonal-line areas of NO\textsubscript{X} and SO\textsubscript{X} indicate low emissions form low-low clusters. (2) Several SGGs are included in the high-high clusters of four pollutants, indicating that common urban characteristics can occur among these SGGs. In other words, several urban structure factors can influence the emissions regardless of air pollutant type. For example, several western coastal SGGs where national industrial complexes are located are included in each high-high cluster of four pollutants. (3) By contrast, the different distribution of clusters indicates that the relationship might differ by air pollutants. The degree of influence may vary by pollutant even if urban structure variables affect emissions. The results of Global and Local Moran’s $I$ suggest potential...
spatial dependencies of air pollutants. Therefore, we use a spatial regression model as a rigorous approach to investigate the relationship between urban structures and each air pollutant.

*Figure 4.* Spatial clusters and outliers of (a) NO$_X$, (b) SO$_X$, (c) PM10, and (d) PM2.5.

We select $k$ for each spatial regression model based on the BIC values summarized in Table 3. On the basis of the results in Table 3, we choose $k = 8$ for the NO$_X$ model, $k = 7$ for the SO$_X$ model, and $k = 6$ for the PM10 and PM2.5 models. Figure 5 shows the extent of spatial dependencies among SGGs for different $k$ values.

**Table 3.** Bayesian information criterion (BIC) for each model.

<table>
<thead>
<tr>
<th></th>
<th>NO$_X$ Model</th>
<th>SO$_X$ Model</th>
<th>PM10 Model</th>
<th>PM2.5 Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 3$</td>
<td>$-407.249$</td>
<td>$-323.998$</td>
<td>$330.579$</td>
<td>$390.741$</td>
</tr>
<tr>
<td>$K = 4$</td>
<td>$-505.341$</td>
<td>$-513.312$</td>
<td>$330.661$</td>
<td>$337.124$</td>
</tr>
<tr>
<td>$K = 5$</td>
<td>$-524.871$</td>
<td>$-411.576$</td>
<td>$307.652$</td>
<td>$-461.686$</td>
</tr>
<tr>
<td>$K = 6$</td>
<td>$-456.650$</td>
<td>$-456.619$</td>
<td>$-214.554$</td>
<td>$-484.580$</td>
</tr>
<tr>
<td>$K = 7$</td>
<td>$-537.169$</td>
<td>$-555.418$</td>
<td>$327.363$</td>
<td>$-455.380$</td>
</tr>
<tr>
<td>$K = 8$</td>
<td>$-561.229$</td>
<td>$828.030$</td>
<td>$318.267$</td>
<td>$387.368$</td>
</tr>
</tbody>
</table>

1 Bold indicates the smallest BIC for each model.
positively estimated. The coefficients for industrial area and GRDP per person are positively significant. Increments in the area of industrial land use commonly show increments in emissions. An industrial area primarily is filled with the positive influence of commercial area on the emission of SO\(_X\). Meanwhile, an increase in the area of commercial land use is significant only on SO\(_X\) (0.529%). Given that the commercial area considered here includes traditional and distribution commercial districts, the positive influence of commercial area on the emission of SO\(_X\) is sensible. Increments in the area of industrial land use commonly show increments in emissions of air pollutants by less than 1%. These results show that the negative effects of urbanization can be mitigated in the concept of efficiency.

In general, significant covariates show positive values, and the magnitude of the values is smaller than that of the one-to-one correspondence in the relationships. In other words, a 1% increase in urban structure variables results in increments in emissions of air pollutants by less than 1%. These results show that the negative effects of urbanization can be mitigated in the concept of efficiency.

First, when the population increases by 1%, the emissions of NO\(_X\), PM10, and PM2.5 increase by 0.565%, 0.456%, and 0.379%, respectively. These positive effects are reasonable because population size determines the size of the residential area and the degree of socioeconomic activities in the regions. Meanwhile, an increase in the area of commercial land use is significant only on SO\(_X\) (0.529%). Given that the commercial area considered here includes traditional and distribution commercial districts, the positive influence of commercial area on the emission of SO\(_X\) is sensible. Increments in the area of industrial land use commonly show increments in emissions. An industrial area primarily is filled with secondary industries that are one of the primary sources of air pollutants, especially SO\(_X\). However,
a park area, which is assumed to have negative impacts on emissions, is not significant in all models. This is because although the park area exists in the region, we cannot determine the amounts, kinds, and spatial patterns of plants within the park that are directly related to the reduction of air pollutants. The road length variable is not statistically significant to the emissions of all air pollutants, which may seem counterintuitive, because a potential confounding issue exists in this variable. For example, a two-lane back road and an eight-lane main road can have the same length. However, the emissions from the two roads are completely different because of the difference in traffic volumes, road widths, and average vehicle speeds. Considering this additional information for each road is challenging in practice. The total land surface is positively significant to the emissions of air pollutants, except for SO\textsubscript{X}. In particular, the emissions of the two types of particulate matter increase more than the emission of NO\textsubscript{X}. This result indicates that large areas usually have many particulate matter pollutant sources. Moreover, increases in GRDP per person have positively significant effects on pollutant emissions, but their degrees are relatively small. We conclude that the impact of regional economic level on air pollution is lower than that of the urban structure in Korea.

Table 5 shows the Global Moran’s I values for each model’s residuals. The absolute values of all pollutants, except for the SO\textsubscript{X} model, are insignificant. Thus, the Bayesian models (NO\textsubscript{X}, PM10, and PM2.5) effectively account for the spatial dependencies. Although the SO\textsubscript{X} model’s Global Moran’s I value is significant, the value decreases from 0.191 to 0.103, which means the model accounts for spatial autocorrelation.

<table>
<thead>
<tr>
<th></th>
<th>NO\textsubscript{X} Model</th>
<th>SO\textsubscript{X} Model</th>
<th>PM10 Model</th>
<th>PM2.5 Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Moran’s I</td>
<td>−0.004 (0.999)</td>
<td>0.103 * (&lt;0.001)</td>
<td>−0.059 (0.071)</td>
<td>−0.042 (0.223)</td>
</tr>
</tbody>
</table>

* indicate statistically significant estimates.

5. Discussion and Conclusions

Korea has pushed ahead with urbanization through government-initiated planning. Hence, its urban structures have become complicated. At the same time, air pollution as a serious environmental issue has recently elicited attention in Korea. Spatial models for underlying mechanisms are critical for estimating emissions of air pollutants from urban structures and for accounting for regional characteristics. Against this background, we study the spatial relationship between urban structures and emissions of air pollutants via exploratory data analyses and Bayesian spatial regression models. Our approach provides a general framework for analyzing spatial relationships between different air pollutants and urban structure covariates. The results show a smaller increase than the direct proportion in the emissions of air pollutants with increasing urban structure variables. This finding is crucial because it suggests that urbanization progress does not result in a one-to-one negative influence on air pollution. Previous studies have focused on investigating the relationship between few urban structure variables and the emission of a particular air pollutant. By contrast, we use multiple regression models to address the uncertainty in the underlying mechanisms. Population size is significant, except in the SO\textsubscript{X} model, and it conforms to the results of other studies [33,37]. Furthermore, population coefficients are estimated to be less than one, which indicates that the innovation effect occurs in urban areas in the context of scaling analysis [48–51]. The positively significant regression coefficients for commercial and industrial areas in the SO\textsubscript{X} model are similar to the results of Cárdenas Rodríguez et al. [38].

The insignificant park area in all the models supports the study of McCarty and Kaza [37]. However, the spatial pattern of the forest variable considered in their research has a positive influence on the atmospheric environment. The traffic-related variable, which was assumed to be the major cause of low air quality in the study of Hussein at al. [28], is not significant in our study. This result can be due
to the confounding of the road length variable as we pointed out in Section 4. Furthermore, a variable related to local economic levels can be used to identify unique relationships between air pollution and economic growth in Korea. On the basis of the positive coefficients for GRDP per person, we find that economic growth negatively affects air conditions in Korea. This finding is in contrast with the results of McCarty [37] and Cárdenas Rodríguez et al. [38], who found that economic level covariates show negative relations with air pollutant variables in the United States and Europe, respectively. Our findings align with the results of Kim and Oh [67], who found positive relationships between economic growth and air pollution in Korea. This fact demonstrates the importance of implementing Korea-specific air-related policies.

Policies for urban development should be planned according to the degree of urban structure and the pollutant level by region. In this process, scenario planning can be introduced to achieve sustainable development by considering ecosystem functions [68,69]. One of the applicable scenarios is that most SGGs are in the urbanization terminal stage. This scenario indicates that many regions have a complex urban structure with a large population that is difficult to change. In this situation, rapidly changing existing areas that are closely related to the emission of air pollutants to other uses is difficult because of numerous economic and social limitations. Hence, the government should recommend a means of maintaining current uses and reducing emissions without manipulating the existing ecosystem. The approach is linked to a smart city, which is a concept for the future in Korea. The smart city plan aims to respond to urban environmental problems, including air pollution, and ultimately block pollution in advance by introducing new technologies in the existing urban structure [70,71]. Thus, the government can reduce air pollution without radical urban transformations.

Furthermore, regional policies should be applied differently depending on the pollutant level because emissions by pollutants vary from region to region. Cárdenas Rodríguez et al. [34] found that positively significant correlations occur between nitrogen dioxide and PM10 and between sulfur dioxide and PM10. Thus, a region with a high level of a specific pollutant may also have a high level of other air pollutants. However, regional differences remain in air pollution. The local government should examine and understand the condition of regional air pollution and apply air pollution reduction policies in consideration of its urban structure characteristics. An air pollution policy, the “Clean Air Conservation Act,” suggests local government action toward regional pollution, but the policy still targets high administrative units. Furthermore, if a region and its neighboring areas have a similar level of a specific air pollutant, then a macro approach that considers these areas’ structures together is required.

In this study, we use multiple urban structure covariates by considering composition factors, such as size and area, in investigating the underlying mechanisms. We note that the configuration of the urban structures, which can affect the emission of air pollutants as well, must be considered. For example, the widely distributed industrial complex and agglomerated industrial complex may have a different influence on emissions even if both complexes have the same area. The remote-sensing method might provide raster data for configuration information. However, in the interest of clear model interpretation and the limited availability of configuration data on urban structures, our urban structure covariates appear to be a good choice. The other open question is how to conceptualize the green area and transportation factors in the urban area. We use park area and road length as practical proxies, but they have potential confounding issues as we pointed out in the previous sections. However, finding appropriate proxies poses significant challenges in practice.

In summary, we investigate the relationship between urban structures and emissions of air pollutants in Korea. The descriptive distribution of emissions of four air pollutants presents the degree of local emission by pollutants. SGGs with agglomerated industry complexes or large populations usually emit large amounts of air pollutants. We use Bayesian spatial linear regression models to account for existing spatial dependencies and to study the underlying mechanisms. We utilize population size, commercial area, industrial area, park area, road length, total land surface, and GRDP per person as spatial covariates. The results provide a theoretical implication for countries that are at
the end of the acceleration stage or at the beginning of the terminal stage of urbanization. The models
indicate that urbanization does not result in a one-to-one negative influence on air pollution. However,
the government should consider sustainable development because a positive relationship exists
between urbanization and air pollutant emissions. In this context, the smart city concept and scenario
planning can be considered in the policy making process to achieve a sustainable environment.

Author Contributions: M.C.J. conceived the design of this study, performed an explanatory data analysis, and
wrote the manuscript. J.P. conducted the statistical analysis and wrote the method section. S.K. conceived the
design of this study and performed the visualization of GIS maps. M.C.J., J.P., and S.K. participated in revising
the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to the anonymous reviewers and Karen Dyson for providing
constructive feedback on the manuscript. This research is based on the result of “Seoul Research Competition
2017”, which is the outstanding paper award.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Figure A1. Locations of four excluded SGGs.

Table A1. Classification of pollutant sources.

<table>
<thead>
<tr>
<th>Pollutant Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Combustion in energy and transformation industry</td>
</tr>
<tr>
<td>2. Non-industrial combustion plants</td>
</tr>
<tr>
<td>3. Combustion in manufacturing industry</td>
</tr>
<tr>
<td>4. Production processes</td>
</tr>
<tr>
<td>5. Extraction and distribution of fossil fuels</td>
</tr>
<tr>
<td>6. Solvent and other product use</td>
</tr>
<tr>
<td>7. Road transport</td>
</tr>
<tr>
<td>8. Other mobile sources and machinery</td>
</tr>
<tr>
<td>9. Waste treatment and disposal</td>
</tr>
<tr>
<td>10. Agriculture</td>
</tr>
<tr>
<td>11. Other sources and skins</td>
</tr>
<tr>
<td>12. Fugitive dust</td>
</tr>
<tr>
<td>13. Biomass combustion</td>
</tr>
</tbody>
</table>

References
2. Fujita, M.; Ogawa, H. Multiple equilibria and structural transition of non-monocentric urban configurations.
3. Garcia-López, M.-À.; Muñiz, I. Urban spatial structure, agglomeration economies, and economic growth in


35. Bereitschaft, B.; Debbage, K. Urban Form, Air Pollution, and CO2 Emissions in Large U.S. Metropolitan Areas. *Prof. Geogr.* 2013, 65, 612–635. [CrossRef]


60. Lee, D. A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spat. Spatio-Temporal Epidemiol.* 2011, 2, 79–89. [CrossRef]


64. Liu, Y.; Guo, H.; Mao, G.; Yang, P. A Bayesian hierarchical model for urban air quality prediction under uncertainty. *Atmos. Environ.* 2008, 42, 8464–8469. [CrossRef]


70. Jin, J.; Gubbi, J.; Marusic, S.; Palaniswami, M. An Information Framework for Creating a Smart City through Internet of Things. *IEEE Internet Things J.* 2014, 1, 112–121. [CrossRef]