

## Review

# Spatial econometric models applied to environmental pollution: a literature review of the socioeconomics drivers.

Manuel Cañaveral <sup>1,\*</sup>, Leonardo Emmendorfer <sup>2</sup>, Debora Spenassato <sup>1</sup> and Ana Azambuja <sup>1</sup>

<sup>1</sup> Institute of Mathematics, Statistics and Physics, Federal University of Rio Grande, Rio Grande, Brazil.

<sup>2</sup> Center for Computational Science, Federal University of Rio Grande, Rio Grande, Brazil.

\* Correspondence: canaveral.mr@gmail.com

**Abstract:** The interest in spatial analysis has been growing in recent years, mainly due to communication technology advance, economic globalization, and the development of new statistical and econometric methods. The main aim of this article is to contribute to the dissemination of spatial econometric applications by presenting some basic theoretical aspects and a literature review of articles that address the socio-economic drivers that lead to environmental pollution. Three spatial regression models are reviewed here: the spatial lag model (SLM), the spatial error model (SEM), and the spatial Durbin model (SDM). A literature search was conducted using specific terms of interest in eight databases, from 1996 to February 2021, where 22 articles were considered for analysis. The results showed that most articles studied environmental problems in China. The most used exploratory spatial analysis model was Moran Index and the most used explanatory spatial analysis models were SDM and SLM.

**Keywords:** spatial econometrics; literature review; socioeconomic drivers; environmental pollution.

## 1. Introduction

Among the major environmental challenges potentially threaten humanity over the forthcoming decades. One should mention climate change, biodiversity loss, unsustainable fishery, emerging pathogens, deforestation, land usage issues, just to name a few. Solving those emerging crises requires us to be able to address increasingly complex problems with large amounts of multidimensional data. Those issues demand sophisticated statistical tools along with the application of complex quantitative models [1].

Under this novel paradigm one should address questions concerning the extent of the impact, the existence of an agglomeration pattern, the occurrence of environmental externalities spilling over geographically to other regions, the factors related to a given phenomenon, among other questions. Given the inherent complexity of the environment-society interrelationships, progress towards the solution of those challenges demands the adoption of increasingly complex statistical tools.

The spatiality of the phenomena under consideration is another important element to be addressed. As Fotheringham [2] point out, it is increasingly recognized that under a multidisciplinary context, most data are spatial. This refers to the fact that most variables have an implicit geographic location, which implies data are two-dimensional (dimension of the variable *per se* and a spatial dimension):

For example, socioeconomic variables such as population density, agricultural production, mortality rate, etc., are obtained through censuses, surveys, and registers. These variables are rarely obtained under an explicit spatial context, but are implicitly associated with a geographic location (municipalities, regions, countries, etc.). In contrast, environmental variables are often obtained from physical measurements with a specific spatial location [3]. For instance, measurements of water and air quality collected at well-identified sampling points.

The inherent need to "integrate" data from various sources creates the need to use spatial econometric techniques. Because by integrating these data will tend to result in spatially dependent as well as spatially heterogeneous observations [3].

It is because of these spatial effects that spatial econometrics has been widely used in recent years in the social sciences in general [4], and in applied research in environmental and natural resources in particular [5].

This paper aims to contribute to the dissemination of spatial econometric applications, presenting some theoretical background and a literature review related to socio-economic drivers related to environmental pollution, focusing on three main spatial regression models.

The remainder of the paper is organized as follows: Section 2 describes the theoretical framework. The spatial regression models presented here are the spatial lag model (SLM), the spatial error model (SEM), and the spatial Durbin model (SDM). The methodology is presented in Section 3. Section 4 provides the results and the conclusions are given in Section 5.

## 2. Spatial econometrics

The increasingly higher interest in spatial analysis in recent years is due to the advances in the communication technologies, the economy globalization [6] and the development of new statistical and econometric tools, among which spatial econometrics (SE) stands out.

The difference between conventional econometrics (CE) and SE is that the former usually neglects interactions that might occur between observations at the geographic level. CE assumes no autocorrelation or independence, i.e., there is no covariance between the residuals of different individuals [7]. This can be formally expressed as:

$$Cov(u_i, u_j | x_i, x_j) = 0 \quad \forall i \neq j \quad (1)$$

Therefore, under this assumption, the perturbation of the individual  $i$  is not related to the perturbation of the individual  $j$  given the values of  $x$ . However, the spatial interactions of the observations violate this independence assumption since, at the geographic level, those might be autocorrelated or spatially dependent.

For example, in environmental sciences, the state of a given ecosystem in one region is not detached from the state of ecosystems in neighboring regions. If the air or water quality is low in a region  $i$ , very likely, in the others regions near  $i$  might be low similarly. This might result either from the inner similarities between regions that share socio-economic patterns that negatively impact ecosystems or due to the negative effects from environmental externalities which spillover to neighboring regions. Spatially dependence among observations is precisely the main interest of SE, unlike CE that majorly focuses on temporal dependence [8].

The choice for a model might also be affected by spatial dependence. Therefore, if spatial data are modeled using ordinary least squares (OLS), which assume sample independence and spatial homogeneity, the results could be affected by significant biases [9]. Therefore, the application of the specialized techniques of spatial econometrics becomes necessary [3].

### 2.1. Spatial heterogeneity

A single general linear model in CE is valid for all sample units:

$$Y = \beta X + \varepsilon \quad (2)$$

In (2) one assumes that  $\beta$  is the same for the entire vector of individuals  $Y$ . In contrast, spatial heterogeneity refers to effect differentiation in space over sample units [10]:

$$Y_i = \beta_i X_i + \epsilon_i \quad (3)$$

Therefore, the spatial heterogeneity model in (3) allows  $\beta_i$  to vary over sample units  $Y_i$ . Models estimated using OLS are not able to properly represent differences between individuals, even if the characteristics of the units of analysis vary, as usually occurs.

For example, environmental pollution (EP) can be modeled as a function of population growth (P), for  $n$  regions, using a simple linear regression logarithmic model, where the results are interpreted as proportions, as:

$$\begin{cases} Y = \beta X + \varepsilon \\ \log EP = \beta(\log P) + \varepsilon \end{cases} \quad (4)$$

Suppose  $\beta = 0.2$ . The OLS interpretation of that model would be as follows: each 1% increase in population leads to a 0.2% increase in the environmental pollution on average, which holds for all regions equally. Besides overestimating pollution, such a homogeneous average increase is quite unrealistic in general.

When modeling heterogeneity, one should ask whether the model parameters would vary spatially. In the example above, it should be questioned whether environmental pollution of all regions react in the same way to population increases. Therefore, a spatial heterogeneity model is applied that basically copes with the differential effects on space.

It is important to note that spatial heterogeneity can also be modeled by advanced CE techniques [11].

## 2.2. Spatial autocorrelation

Spatial autocorrelation is the correlation of a variable with itself and arises when the value of a variable in a location is related to its value in other locations in space [12]. It can be expressed formally as:

$$\text{Corr}[X(s_i); X(s_j)]; \forall i \neq j \quad (5)$$

This means the correlation of a variable  $X$  in the location  $S_i$  in relation to other neighborhood units  $S_j$ . Spatial autocorrelation (also called spatial dependence<sup>1</sup>) can be either positive or negative.

Positive spatial autocorrelation occurs when a variable  $x$  at one location  $i$  exerts a contagion effect on neighboring regions  $j$ , which raises clusters of individuals with geographic proximity.

Negative spatial autocorrelation occurs when a variable  $x$  at one location  $i$  exerts a repulsive effect on neighboring regions  $j$ , which is related to dissimilarity between spatial units that are geographically close to each other.

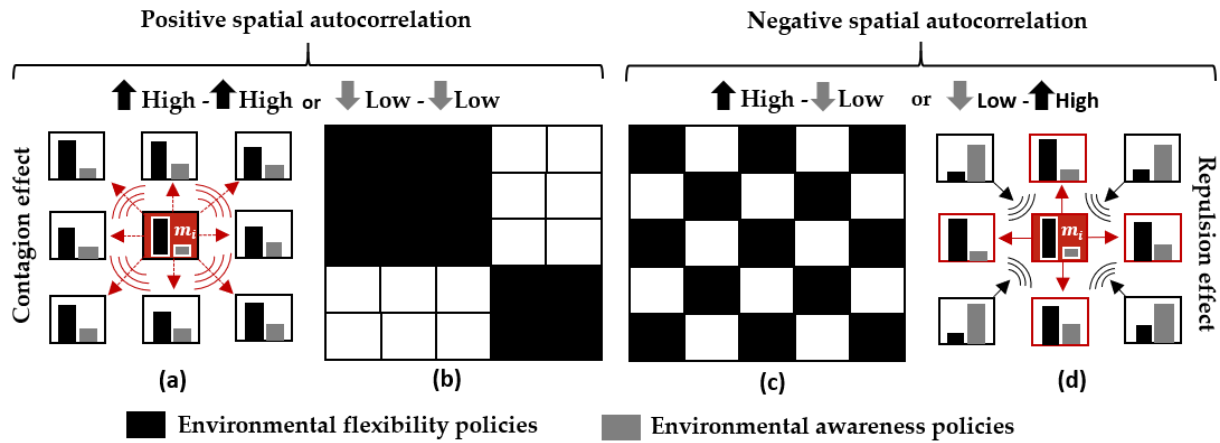
For example, suppose a municipality  $m$  located in a region  $i$  decides to implement environmental flexibility policies that eliminate certain restrictions to the entry of mining investments, which generates economic growth. If other municipalities near  $i$ , motivated by the mining economic benefits, also decide to implement environmental flexibility policies, this configures a positive spatial autocorrelation scenario. On the contrary, if the neighbors observe that mining environmental costs are greater than benefits and decide to restrict mining by implementing environmentally conscious policies, this results a negative spatial autocorrelation scenario.

**Figure 1** illustrates those two contrasting cases. (a) shows a positive spatial autocorrelation scenario where the contagion or spillover effect occurs. Here  $m_i$  presents high values of the environmental flexibilization policies variable and its neighbors also present high values of the same variable, the same happens with the environmental awareness

<sup>1</sup> Spatial autocorrelation is considered in many texts as a synonym of spatial dependence, this is logical because for dependence to exist the values must be autocorrelated. However, in a strict sense, these two terms are not synonymous, since spatial autocorrelation is a weaker expression of spatial dependence, relating only to the first moments of the joint distribution of a variable [12].

Spatial autocorrelation is exploratory, where a contrast statistic is used to detect and quantify the phenomenon. While spatial dependence is an explanatory analysis. That is, it seeks to explain why these dependence relationships occur at the geographical level, based on an underlying theoretical context [13].

policies variable with low-low correlations. (b) shows how on macro level spatial clusters are formed.



**Figure 1.** Positive and negative spatial autocorrelation scenarios. (a) corresponds to the contagion effect at the micro level; (b) corresponds to cluster formation at the macro level; (c) corresponds to a mosaic of dissimilar regions at the macro level; (d) corresponds to the repulsion effect at the micro level.

(d) illustrates how the repulsion effect prevents the cluster formation, therefore, on a macro level, a mosaic-style pattern might arise, as shown in (c). As a result, negative spatial autocorrelations arise from higher values correlated with lower values and vice-versa.

Spatial autocorrelation studies the geographical distribution of the values of a variable for different individuals. Spatial dependence tries to explain how this functional relationship is occurs.

### 2.3. Spatial dependence

When data show spatial dependence, it means that observations from one region  $i$  depend on observations from another region  $j$  [10,14]. This can be formally expressed as:

$$y_i = f(y_j) + \varepsilon_i, \quad \forall i \neq j \text{ with } j = (1, 2, 3, \dots, n) \quad (6)$$

This means that  $y$  in the region  $i$  is explained not only by internal conditions in  $i$ , but also by the values of the same variable  $y$  in neighboring regions  $j$ .

Equation 7 helps to understand the spatial dependence phenomenon, but in practice it is unfeasible, since given a sample  $n$  spatial units, the spatial dependence of each of the  $i$  regions with each of the remaining  $j$  regions must be estimated.

For example, if  $n=5$  one should specify the following system of equations:

$$\begin{aligned} y_1 &= \beta_{21}y_2 + \beta_{31}y_3 + \beta_{41}y_4 + \beta_{51}y_5 + \varepsilon_1 \\ y_2 &= \beta_{12}y_1 + \beta_{32}y_3 + \beta_{42}y_4 + \beta_{52}y_5 + \varepsilon_2 \\ y_3 &= \beta_{13}y_1 + \beta_{23}y_2 + \beta_{43}y_4 + \beta_{53}y_5 + \varepsilon_3 \\ y_4 &= \beta_{14}y_1 + \beta_{24}y_2 + \beta_{34}y_3 + \beta_{54}y_5 + \varepsilon_4 \\ y_5 &= \beta_{15}y_1 + \beta_{25}y_2 + \beta_{35}y_3 + \beta_{45}y_4 + \varepsilon_5 \end{aligned} \quad (7)$$

The first row tells us that in a 5-region system, region 1 depends on regions 2, 3, 4 and 5 plus an error term  $\varepsilon_1$ , and so on.  $\beta_{21}y_2$  means the coefficient of region 2 for region 1.

Therefore, since there are  $(n^2 - n)$  parameters to estimate from just  $n$  observations there might be more parameters than observations available. This prevents the estimation to be performed. The problem arises from spatial dependence multidirectionality. Therefore, restrictions should be imposed on the way observations interact in space, thus reducing the number of parameters to be estimated. This is achieved through the construction

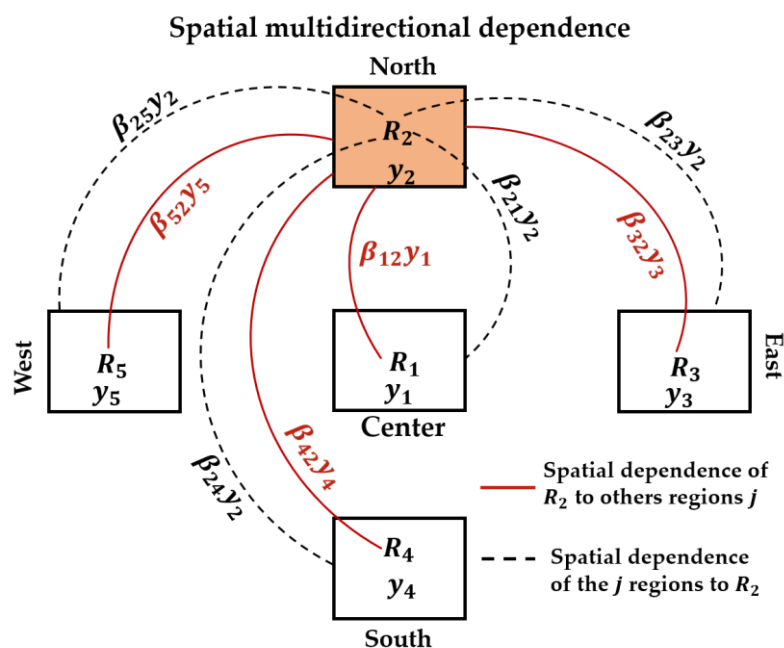
of the spatial weight matrix ( $\mathbf{W}$ ), which controls the multidirectional interdependence relationships.

### 2.3.1. Spatial dependence multidirectionality

Spatial dependence cannot be treated by CE due to the multidirectionality that dominates the interdependence relationships between spatial units [15]. Spatial dependence can occur in several directions across space, while CE usually copes only with single-directional effects in time, where the past explains the present.

Figure 2 presents a hypothetical map of 5 regions with multidirectional interdependencies relationships flows in space.

The equation system (7) showed how a region  $i$  depends spatially on other regions  $j$  that are part of the system. Let us take as reference  $R_2$  located to the north, it depends on  $R_1$  located in the center,  $R_3$  to the east,  $R_4$  to the south and  $R_5$  to the west, the red lines show this multidirectionality with the respective parameters  $\beta$  to be estimated. Reciprocally, regions 1, 3, 4 and 5 also depend on  $R_2$  (dashed lines).



**Figure 2.** Multidirectionality of spatial dependence.

This interdependence of regions (all regions versus all regions) is what generates the multiple parameters to be estimated. The solution is given by the construction of  $\mathbf{W}$ .

### 2.4. Spatial weights matrix

The specification of a spatial weights matrix is one of the most important and conflicting aspects in spatial econometric analysis [16]. This matrix makes it possible to operationalize the relationships in space, incorporating into the general linear model the interdependence relationships of the spatial units. This makes the model cease to be classical and take on the spatial characteristic. The matrix  $\mathbf{W}$  also represents the reduction of interdependencies between regions. The concept is aligned to the Waldo Tobler's first law of geography which states that all things are related to each other, but closer things matter more than distant things [17]:

$$\mathbf{W} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix} \quad (8)$$

$W$  is a positive symmetric  $n \times n$  matrix. The elements  $w_{ij}$  are non-stochastic and exogenous to the model [12,14] and represent the connectivity (adjacency or closeness) between regions  $i$  and  $j$ .

There are several ways to construct the  $W$  matrix, however, the most commonly used approach is based on the binary contiguity of spatial units (first order). This method uses 0 and 1 values to represent the degree of interaction between regions. For instance,  $w_{ij} = 1$  might correspond to spatial units that share a common border, while  $w_{ij} = 0$  might represent noncontiguous areas.

## 2.5. Spatial association indicators

Spatial association indicators are statistics that help identifying whether clustering or dispersion occurs in the data. Spatial autocorrelation includes both global and local measures. Several statistics have been proposed to detect autocorrelation, one of the most widely used is Moran's  $I$  [6].

Global spatial autocorrelation jointly analyzes all the spatial units that make up the sample to determine the occurrence of clusters, or whether data are randomly distributed [18]. Formally expressed as:

$$\text{Global Moran's } I = \frac{N}{\sum_{i,j} w_{ij}} \frac{\sum_{i,j} w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (9)$$

With  $i \neq j$ , where  $x_i$  is the interest variable in region  $i$ ,  $\bar{x}$  is the variable mean value,  $N$  is the number of observations,  $w_{ij}$  is the spatial weight.

The Moran index varies between -1 and 1. A value close to 1 indicates the presence of clusters, and a value close to -1 indicates spatial dispersion in the data. If the value is 0, there is no spatial autocorrelation.

Local spatial autocorrelation is associated to the occurrence of clusters located in specific areas of the territory that concentrate higher or lower values (hot or cold spots) of a variable compared to its mean value [15]. In this case, each spatial unit must be analyzed separately, obtaining a statistic for each region of the sample. It is expressed formally as:

$$\text{Local Moran's } I = \frac{(x_i - \bar{x}) \sum_{j=1}^N w_{ij} (x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (10)$$

The specifications are the same as in the global Moran index.

## 2.6. Spatial regression models

After the spatial autocorrelation analysis confirms the spatial dependence of the geographic elements, it is necessary to establish a spatial econometric model to evaluate the effect and spatial interaction of these elements. In this section, we present spatial autoregressive models, which are used for cross-sectional data.

Spatial dependence can be incorporated into the basic linear regression model through two different approaches: substantive spatial dependence, also known as the spatial lag model, or residual spatial dependence, also known as the spatial error model. Next, we briefly explain both approaches and, additionally, a special case of the SLM model called the Durbin Spatial Model.

### 2.6.1. Spatial lag model

The Spatial Autoregressive Regression Mixed Model or Spatial Lag Model (SLM) is used when an endogenous variable of a regression model is spatially correlated. Dependence is considered as an additional regressor under the form of a spatially lagged dependent variable ( $Wy$ ). This is the most basic model of spatial dependence, which is suitable for expressing situations where the values of a variable depend systematically on the geographic location of that variable. It is formally expressed as:



$$\begin{cases} y = X\beta + \varepsilon \\ y = \rho Wy + X\beta + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \quad (11)$$

Where  $y$  is a vector of  $nx1$  dependent variables.  $X$  is a matrix of  $nxk$  exogenous explanatory variables, and  $\beta$  is a vector of  $kx1$  regression parameters.  $Wy$  is a vector of spatial lags of dimension  $nx1$ , which result from the product of the spatial weights matrix  $W$  and the vector of dependent variables  $y$ , and reflect the weighted average of the values of  $y$  in neighborhood  $i$ .  $\rho$  is a scalar parameter reflecting the strength of spatial dependence. When the scalar parameter  $\rho$  takes the value zero the spatial lag model simplifies to the basic linear OLS regression model. Finally, we assume that the vector  $\varepsilon$  of residuals with dimension  $nx1$  contains independent, normally distributed terms with constant mean and zero variance.

### 2.6.2. Spatial error model

Similarly, spatial autocorrelation could occur only in the disturbance term rising the so-called linear regression model with spatially autoregressive spatial disturbances, also called Spatial Error Model (SEM). It is the most widely adopted specification when the basic linear regression model is ineffective in explaining a spatially autocorrelated phenomenon. The existence of certain factors or variables not explicitly considered in the model transfers to the error terms the configuration of clustering of values (autocorrelation) present in the endogenous variable [15]. The SEM model is expressed formally as:

$$\begin{cases} y = X\beta + \varepsilon \\ \varepsilon = \lambda W\varepsilon + \mu \\ \mu \sim N(0, \sigma^2) \end{cases} \quad (12)$$

Where  $\mu$  is a white noise term and  $\lambda$  the autoregressive parameter reflecting the intensity of interdependencies in the perturbation term.

### 2.6.3. Durbin spatial model

The Spatial Durbin Model (SDM) is a spatial case of the SLM because the SDM model consists of an SLM model augmented by spatially lagged explanatory variables [19]. It is expressed formally as:

$$\begin{cases} y = \rho Wy + \beta X + WX\theta + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \quad (13)$$

Since SDM includes the spatially lagged dependent variable, the explanatory variables matrix and the spatially lagged explanatory variables. That is, SDM depends on the neighbor dependent variables, regional factors proper to the matrix  $WX\theta$ , plus the same factors in the neighboring regions. Therefore, it is more general and has a more effective explanatory power than SLM and SEM.

## 3. Methods

A structured search of articles in eight databases was carried out to identify existing publications on the topic of socioeconomic factors influencing environmental pollution.

The databases consulted were: Taylor & Francis Online – Journals; PMC (PubMed Central); SpringerLink; OneFile (GALE); Social Sciences Citation Index (Web of Science); Directory of Open Access Journals (DOAJ); Scopus (Elsevier); Science Citation Index Expanded (Web of Science).

The search terms were: (socioeconomic drivers) AND (environmental pollution index) AND "spatial econometric", the search was performed by title or keywords or abstract, restricting to the papers in the article. More precisely:

(socioeconomic drivers) was used to find articles that address the influence of socioeconomic factors on environmental pollution.

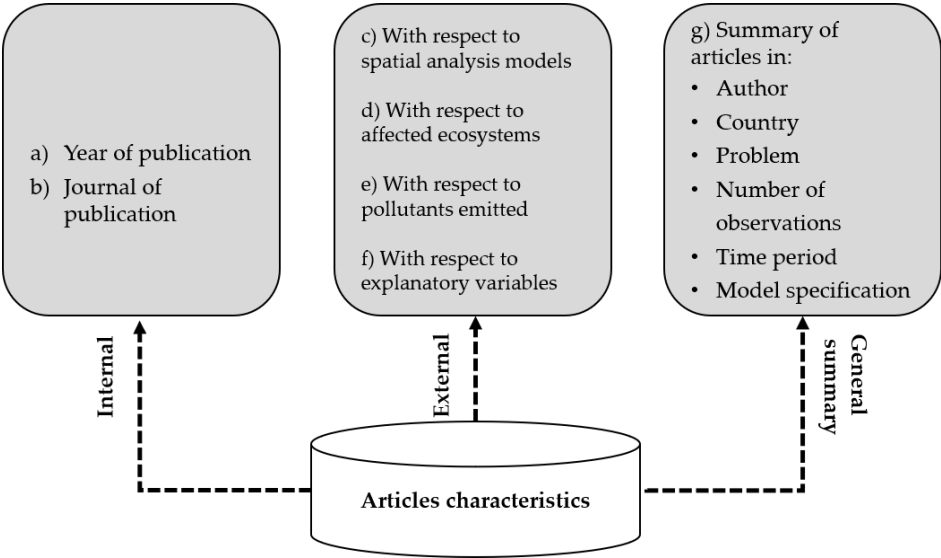
(environmental pollution index) was used to find articles expressing environmental pollution in terms of indexes.

Finally, "spatial econometric" was used to find articles that apply spatial econometric methods to identify the spatial relationship between socioeconomic drivers and environmental pollution.

The time period of the search was from 1996 to February 2021, resulting in 68 articles found. After a review of the subject matter and methods applied, 22 articles were retained. Only articles applying SLM, SEM, and SDM spatial regression models were selected. In addition, the construction of indices to describe environmental pollution was included in their methodological procedures and the specification of econometric models addressed the analysis of the influence of socioeconomic factors on environmental pollution.

4. Results

The literature review results are presented through the descriptive analysis of three sub-themes: internal, external characteristics and general summary of the articles, which are listed in **Figure 3**.

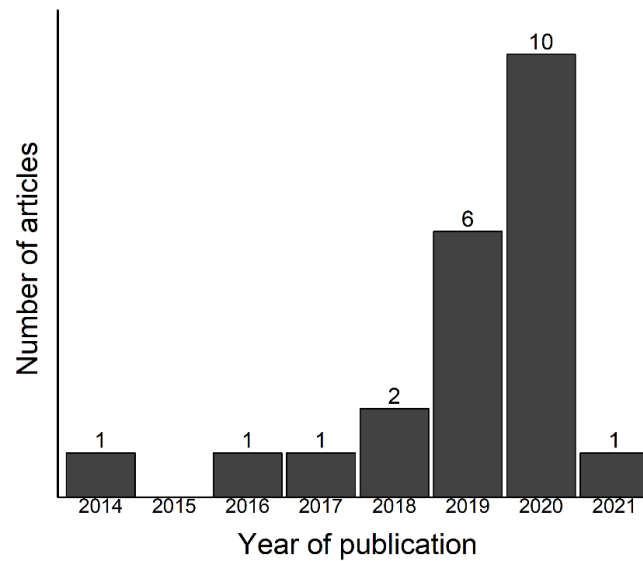


**Figure 3.** Results presentation diagram.

*a) Year of publication*

The publications start from 2014, concentrating in 2019 and 2020 with 6 and 10 articles respectively, as shown in **Figure 4**.





**Figure 4.** Year of publication of the articles.

The intention of researchers to carry out studies on environmental quality taking into account spatial relationships has shown an evident increase in recent years. This may be due to the marked presence of environmental problems and the need to apply statistical inference techniques with a spatial approach.

*b) Journal of publication*

With respect to the journals (**Table 1**), the International Journal of Environmental Research & Public Health was the journal with the highest number of publications and is among the 4 journals with the highest impact factor of the 12 journals.

**Table 1.** Number of publications according to academic journals and impact factor.

Academic Journals	Impact Factor	Frequencies
Atmosphere	2.397 (2019)	1
Carbon Management	1.897 (2019)	1
Complexity	2.591 (2018)	1
Discrete Dynamics in Nature and Society	0.970 (2019)	2
Environment, Development and Sustainability	2.191 (2019)	1
Environmental Economics and Policy Studies	1.430 (2018)	1
Environmental Science and Pollution Research	3.306 (2019)	2
International Journal of Environmental Research & Public Health	3.127 (2019)	8
Journal of Geographical Sciences	3.457 (2019)	1
Public Library of Science	2.740 (2019)	1
Remote Sensing	4.509 (2019)	2
Spatial Economic Analysis	1.231 (2017)	1
Sum		22

*c) Spatial analysis models adopted*

The articles comprised three different explanatory models (*a*) and two exploratory models (*b*). *c* refers to articles that used only one explanatory model, *d* refers to the articles that used all three explanatory models jointly, *e* refers to articles that used only one exploratory model and *f* to those that used all models together. The **Table 2** below shows the distribution of these models.

**Table 2.** Spatial analysis models used by the articles.

Articles	(a)			(b)		(c)			(d)	(e)		(f)
	SLM	SEM	SDM	GMI <sup>1</sup>	LMI <sup>2</sup>	Only SLM	Only SEM	Only SDM	Used SLM-SEM-SDM	Only GMI	Only LMI	All
1			x	x	x			x				
2	x	x	x						x			
3	x	x	x	x					x	x		
4			x					x				
5	x	x		x						x		
6	x	x	x						x			
7	x	x		x						x		
8	x					x						
9			x	x				x		x		
10			x					x				
11	x					x						
12	x		x	x						x		
13	x	x	x	x	x				x			x
14	x	x										
15			x	x				x		x		
16			x	x				x		x		
17			x	x	x			x				
18	x	x	x	x					x	x		
19	x	x	x	x	x				x			x
20	x			x		x				x		
21	x	x	x	x					x	x		
22	x			x	x	x						
<b>Sum</b>	<b>15</b>	<b>10</b>	<b>15</b>	<b>15</b>	<b>5</b>	<b>4</b>		<b>7</b>	<b>7</b>	<b>10</b>	<b>0</b>	<b>2</b>

<sup>1</sup> Global Moran Index; <sup>2</sup> Local Moran Index.

In (a) 15 articles adopted SLM and SDM, and 10 articles used SEM. In (b) GMI was used in 15 articles, which was combined with LMI in 5 of the 15 articles. In 10 articles GMI was used as the single model to explore spatial autocorrelation. In (c) 7 articles opted for the SDM and 4 for the SLM. None study used the SEM as the only explanatory model. In (d) 7 articles used all 3 explanatory models together. In (e) only the GMI was used as the only exploratory model by 10 articles. In (f) only two articles used the 3 explanatory models and the 2 exploratory models together.

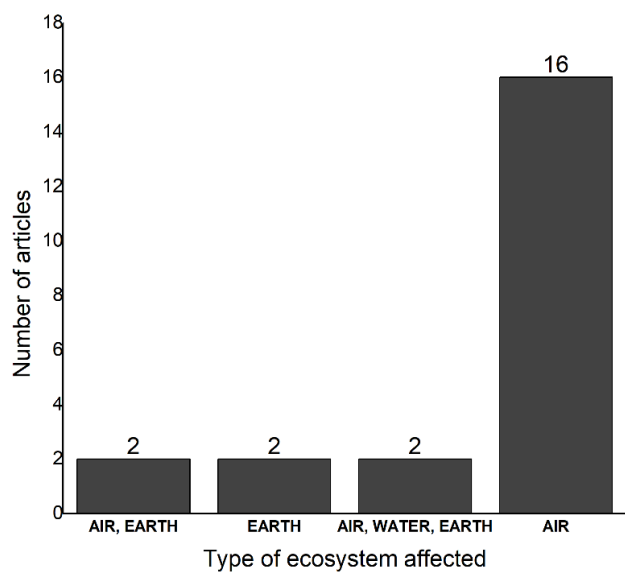
SDM and GMI were the most used models. The reason why SDM was adopted more often is, possibly, due to it is a further complete model that includes spatial lags in the dependent variable, as well as in the independent variables. Therefore, this model contains a greater explanatory richness, in terms of analysis of the spatiality between the variables. GMI was used more often because the global test is usually the most frequent choice among researchers.

In all the articles, the different models were combined. This with the aim of compare the results of the different methods given the same database.

#### *d) Ecosystems affected*

This refers to the main elements that are environmental pollution receptors according to problem of each article.

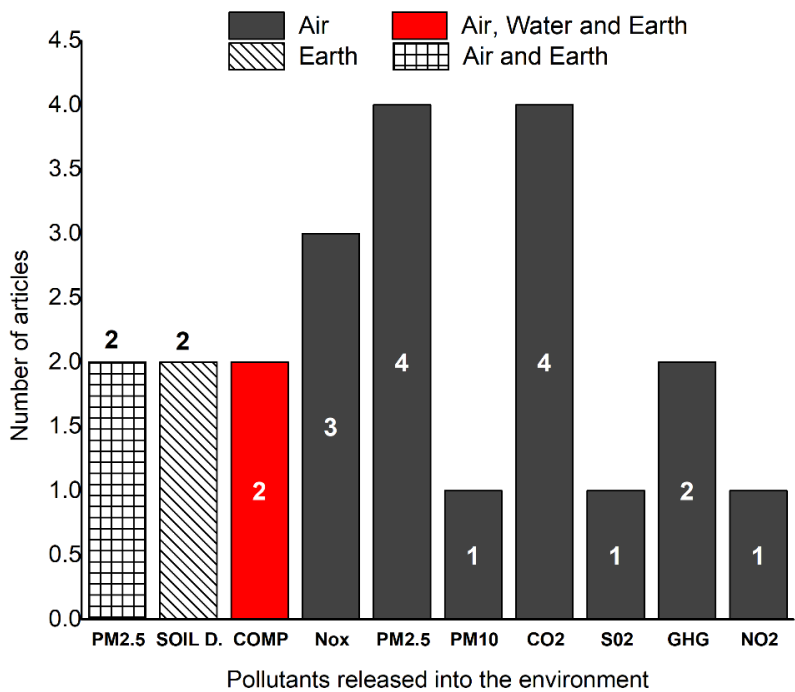
**Figure 5** shows that the main pollution receiver of pollution is with a total of 16 articles. In 4 other articles, air pollution is also a factor intended to be analyzed, along with other affected elements such as water and land. This is because most of the articles come from researchers in China, and air pollution is one of the main environmental problems today in China.



**Figure 5.** Number of articles according to the type of ecosystem affected.

*e) Pollutants emitted*

The main ecosystem affected is the air, according to this literature review. Therefore, the main harmful elements are those that produce atmospheric pollution such as particulate materials of 2.5 (PM<sub>2.5</sub>) and 10 micrometers (PM<sub>10</sub>), nitrogen oxides (NO<sub>x</sub>, NO<sub>2</sub>), carbon (CO<sub>2</sub>) and sulfur dioxides (SO<sub>2</sub>), and greenhouse gases (GHG). **Figure 6** shows pollutants studied by the articles considered.



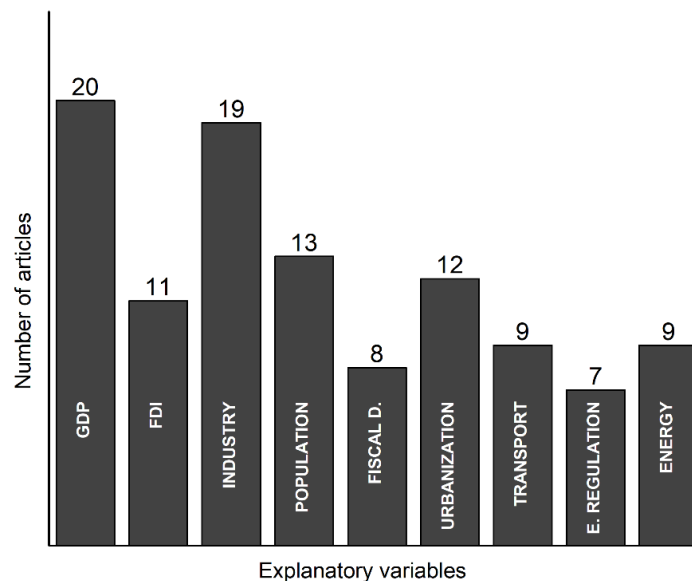
**Figure 6.** Number of articles according to the main pollutants emitted.

The dark bars are 16 articles that investigated air pollution, disaggregated by pollutant. The red bar indicates that there are 2 articles that addressed environmental pollution in a comprehensive manner (COMP), i.e., effects on air, water and land simultaneously. The diagonal striped bar refers to the effects on land, caused by soil degradation (SOIL D.)

with 2 articles. Finally, the bar with squares indicates air and land contamination due to  $PM_{2.5}$ , since land use activities such as construction generate large amounts of this pollutant and this affects soil and air.

#### f) Explanatory variables considered

The econometric models used different and numerous explanatory variables. The **Figure 7** shows the central theme or axes of the variables most used by the models. On the GDP axis, the articles mainly used GDP per capita to represent economic development/growth. In this axis, 20 of the 22 articles, used GDP as one of several socioeconomic drivers explaining environmental pollution.



**Figure 7.** Number of articles according to socioeconomic drivers used as explanatory variables.

Foreign direct investment (FDI) was used to refer specifically to the foreign capital inflow per GDP unit, being used by 11 articles. Industry axis was used to refer mainly to secondary industry output as a GDP proportion, referred to in most articles as industrial structure. It was also used to refer to the number of employees in the secondary industry.

The industrial sector, unlike the primary sector (extractivism, agriculture, fishing, etc.) and the tertiary sector (services), focuses on transformation processes that require the use of fossil fuels and, therefore, have significant environmental impacts.

Industrial sector importance is reflected in the 19 articles that used this variable as an explanatory factor for pollution.

The population axis was used to represent population density (inhabitants/km<sup>2</sup>), population urbanization or urbanization rate (urban population/total population), and population or population growth per se. This axis captures everything that has to do with population, 13 articles used this type of variable, because there is a direct relationship between pollution and population.

The fiscal decentralization axis (Fiscal D.) refers to everything related to public finances. In the 8 articles that used this variable, it was represented as: technical level or technological progress (science and education spending/total public spending), intergovernmental transfers (government transfers/local revenues) and as fiscal decentralization per se, expressed as local spending percentage over total spending.

Fiscal decentralization refers to fact of taking powers away from the central government and assigning them to local governments with their respective financing (i.e., resources transfer to cover the powers). There is no widely accepted consensus on the type of relationship between fiscal decentralization and environmental pollution. However, it is recognized that it has an important impact on the decisions of local governments in terms of socioeconomic and environmental policies.

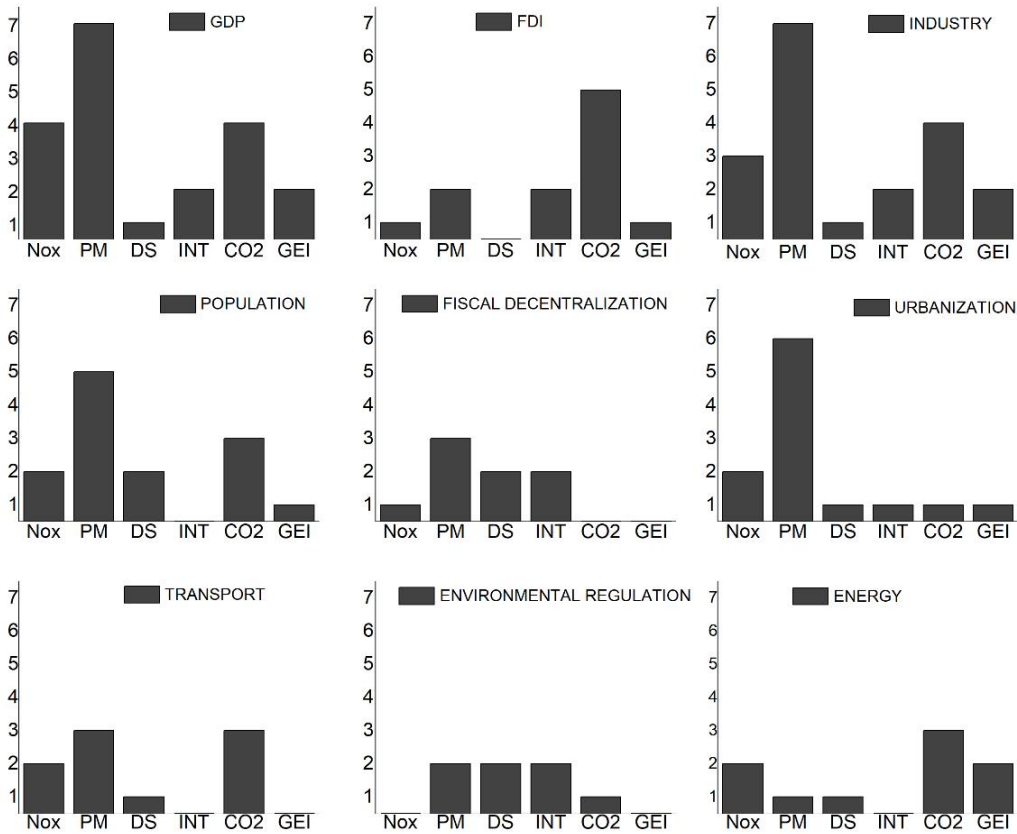
The urbanization axis, unlike the population axis, which focuses on surface areas, was used in 12 articles, usually referred to as territorial urbanization or urbanized area (urban area/total area). Urbanization was also expressed as the average number of employees at work. The use of this variable type is due to the fact that it has a direct relationship with environmental pollution.

The transportation axis comprises everything related to human mobilization means, in the 9 articles it was used as transportation intensity expressed as urban road area per capita to measure automobile exhaust emissions, motor vehicle ownership, rail mileage measured by unit area average rail mileage within a given region, road density and as transportation per se. This variable was mainly used in articles addressing air pollution, given the direct relationship between transportation and air pollutants.

Environmental regulation is everything related to environmental policies and investment in environmental protection. This variable was used in 7 of the 22 articles, which tells us that policies to reduce pollution were not variables widely used in the articles. The relationship between environmental policies and pollution is negative.

With respect to the energy axis, this was represented by energy structure (coal consumption/total consumption of fossil fuels), energy intensity or efficiency (energy consumption/GDP), social energy consumption. These types of variables were used by 9 articles, mainly related to CO<sub>2</sub>, GHG and NO<sub>x</sub> emissions.

The **Figure 8** refers to number of times an article used a given explanatory variable to explain a given pollutant. For example, the GDP variable was used 4 times to explain NO<sub>x</sub> and 7 times to explain PM.



**Figure 8.** Disaggregation of the number of explanatory variables by pollutant.

Variables related to GDP, industry, population and urbanization were most used to explain PM. In summary, the graph above shows how the different socioeconomic drivers used to explain various types of environmental pollution are distributed.

## g) General summary of the articles

**Table 3** presents a summary analysis of each article in the study. The following characteristics are addressed here: the authors, the country under study, the problem, and the specification of the econometric model.

The objectives of each article, in general, are to evaluate the impact of certain socio-economic factors on certain environmental pollutants that are the subject of the problem in the articles.

Most of the articles are studies about environmental problems affecting China. With respect to the problem addressed, they are mostly related to severe levels of air pollution as a product of high rates of economic growth and social development. The model specification is presented in functional form,  $Y = f(X)$ . Certain characteristics are also presented such as T which refers to the time period, N is the sample size and A is the number of years of the study. The exploratory and explanatory methods used in each study are also mentioned here. SLM= Spatial Lag Model, SEM= Spatial Error Model, SDM= Spatial Durbin Model, GMI= Global Moran Index, LMI= Local Moran Index.

**Table 3.** General summary of articles from the literature review.

AUTHOR / COUNTRY / PROBLEM	ECONOMETRIC MODEL SPECIFICATION
<p><b>[20] - CHINA</b></p> <p>In China, accelerated industrialization and urbanization processes, together with impressive economic and social performance, exacerbate natural resources consumption and the emission of pollutants. Thus, deteriorating environmental quality.</p>	<p><math>Y = f(X_1 + X_2 + X_3 + X_4 + X_5)</math>;  T = 2003 al 2012; N=272; A=10; SDM; GMI; LMI  Y =Environmental pollution index,  X<sub>1</sub>= Fiscal Decentralization,  X<sub>2</sub>=Economic Growth (GDP per capita),  X<sub>3</sub>= Foreign Direct Investment (FDI/GDP),  X<sub>4</sub>=Industrial Structure (Industrial GDP/GDP),  X<sub>5</sub>= Population Density (Population//Km<sup>2</sup>).</p>
<p><b>[21] - CHINA</b></p> <p>Increased air pollutants greater acceleration in China, nitrogen oxides (NO<sub>x</sub>), product of rapid economic growth and urbanization.</p>	<p><math>Y = f(X_1 + X_2)</math>;  T = 2010 al 2015; N=30; A=6; SLM; SEM; SDM  Y = Nitrogen oxides (NO<sub>x</sub>), X<sub>1</sub> = Income (GDP per capita)  X<sub>2</sub> = Urbanization (Urban population and Total population).</p>
<p><b>[22] - CHINA</b></p> <p>Atmospheric pollution and poor governance resulting from industrialization and urbanization.</p>	<p><math>Y = f(X_1 + X_2 + \dots + X_6)</math>;  T = 2006 al 2015; N= 31; A= 10; SLM; SEM; SDM; GMI  Y = NO<sub>x</sub> emissions,  X<sub>1</sub> = Economic development,  X<sub>2</sub> = Industrial structure,  X<sub>3</sub>= Energy efficiency, X<sub>4</sub> = Urbanization,  X<sub>5</sub> = Transport, X<sub>6</sub> = Population.</p>
<p><b>[23] - CHINA</b></p> <p>Rapid urbanization in China not only promotes rapid urban population expansion and economic agglomeration, but also aggravates haze pollution.</p>	<p><math>Y = f(X_1 + X_2 + X_3 + X_4)</math>;  T= 2001 al 2016; N= 342; A= 16; SDM  Y = Haze pollution (PM<sub>2.5</sub>),  X<sub>1</sub> = Urban scale (Total population),  X<sub>2</sub> = Urban agglomeration (GDP),  X<sub>3</sub> = GDP per capita,  X<sub>4</sub> = Industrial structure.</p>
<p><b>[24] - CHINA</b></p> <p>Temporal and spatial concentration of PM<sub>2.5</sub> in Hebei as a product of intensive land use.</p>	<p><math>Y = f(X_1 + X_2 + X_3 + X_4)</math>;  T= 2001 al 2016; N= 11; A= 4; SLM; SEM; GMI  Y= Concentrations of PM<sub>2.5</sub>,</p>



	$X_1$ = Intensity of land use, $X_2$ = Land use structure, $X_3$ = Land input level, $X_4$ = Land output benefit.
<p><a href="#">[25]</a> - CHINA</p> <p>Serious haze pollution resulting from the coexistence of industrialization and urbanization in China, with the main components of haze being fine particles and inhalable particles, which can pose a serious threat to human health.</p>	$Y = f(X_1 + X_2 + X_3 + X_4 + \cdots + X_8);$ T= 2008 a 2016; N= 30; A=9; SLM; SEM; SDM $Y$ =Concentrations of $PM_{10}$ , $X_1$ = Industrial transfer, $X_2$ = Population density, $X_3$ =GDP per capita, $X_4$ = Energy (coal consumption/total energy consumption), $X_5$ =Local vehicle usage ( $km^2$ road / $km^2$ city), $X_6$ = Environmental regulation (Industrial pollution investment / industrial GDP), $X_7$ = Industrial structure (Industrial GDP / GDP), $X_8$ = Weather conditions (precipitation and relative humidity).

Table 3 (continued)

AUTHOR / COUNTRY / PROBLEM	ECONOMETRIC MODEL SPECIFICATION
<p><b>[26] - EMERGING ECONOMIES</b></p> <p>Carbon dioxide intensity (<math>CO_2</math>) related to energy.</p>	$Y = f(X_1 + X_2 + \dots + X_5)$ T= 1995 a 2011; N= 39; A=17; SLM; SEM; SDM; GMI; LMI $Y_1$ = GHG incorporated in import trade, $Y_2$ = GHG incorporated in export trade, $X_1$ = Population size, $X_2$ = GDP, $X_3$ = Energy intensity, $X_4$ = Percentage of clean energy/energy consumption, $X_5$ = Percentage of Industrial GDP.
<p><b>[27] - CHINA</b></p> <p>China's rapid industrialization and urbanization coupled with high annual emissions of pollutants has caused the air quality of Chinese cities to deteriorate significantly, threatening the public health and well-being of urban residents.</p>	$Y = f(X_1 + X_2 + X_3 + \dots + X_{10})$ ; T= 2014 al 2017; N= 388; A=4; SLM $Y$ = Concentration of $PM_{2.5}$ , $X_1$ = GDP per capita, $X_2$ = FDI, $X_3$ = Population density, $X_4$ = Urbanized area, $X_5$ = Urban greening rate, $X_6$ = Number of own vehicles, $X_7$ = Percentage of industrial GDP, $X_8$ = Total energy consumption, $X_9$ = Science and technology expenditure to GDP, $X_{10}$ = Soot emissions.
<p><b>[28] - CHINA</b></p> <p>Large-scale urban construction, urban population aggregation and the rapid growth in the number of motor vehicles accompanied by urban development have led to the deterioration of air quality in urban areas in China.</p>	$Y = f(X_1 + X_2 + X_3 + \dots + X_6)$ ; T= 2018 al 2019; N=334; A=2; SLM; GMI; LMI $Y$ = Nitrogen Dioxide ( $NO_2$ ), $X_1$ = Population, $X_2$ = GDP, $X_3$ = Industrial GDP, $X_4$ = Social energy consumption $X_5$ = Total energy consumption, $X_6$ = Ownership of motor vehicles.
<p><b>[29] - OPEC COUNTRIES</b></p> <p>According to the World Bank, energy-related <math>CO_2</math> emissions from developing countries will be 127% higher in the world's most developed economies by 2040. Due to their generally stronger economic growth and continued use of fossil fuels and since oil is the most prominent fossil fuel. OPEC countries are the largest emitters of <math>CO_2</math>.</p>	$Y = f(X_1 + X_2 + X_3 + X_4 + X_5 + X_6)$ ; T= 2006 al 2015; N= 13; A= 13; SDM $Y$ = Carbon dioxide emissions ( $CO_2$ ) $X_1$ = GDP growth (annual %), $X_2$ = Exports as a percentage of GDP $X_3$ = Imports as a percentage of GDP, $X_4$ = Inflation rate (annual percentage) $X_5$ = Employment (percentage of total work), $X_6$ = Governance.
<p><b>[30] - CHINA</b></p> <p>With reform and opening up, the Chinese economy has achieved rapid development; however, the interdependence between economies also involves the interaction of energy consumption and carbon emissions. World Bank data indicated that China has become the largest carbon emitter in the world since 2008.</p>	$Y = f(X_1 + X_2 + X_3 + \dots + X_7)$ ; T= 2004 a 2016; N= 284; A= 13; SLM; SEM; SDM; GMI $Y$ = $CO_2$ emissions, $X_1$ = Economic development (GDP per capita), $X_2$ = Energy efficiency, $X_3$ = Industrial structure (% employed in secondary industry), $X_4$ = Energy structure (% coal consumption), $X_5$ = Population density, $X_6$ = Traffic facilitation ( $km^2$ road/ $km^2$ city), $X_7$ = FDI.

Table 3 (continued)

AUTHOR / COUNTRY / PROBLEM	ECONOMETRIC MODEL SPECIFICATION
<p><b>[31] - SPAIN</b></p> <p>Over the past 20 years, rates of conversion to residential land use have far exceeded population growth rates on the continent [20% vs. 6%]. As a result, the amount of urbanized land consumed per person during this period has more than doubled, leading to the formation of new peripheral cities around traditional urban centers and scattered residential developments located on the urban periphery.</p>	$Y = f(X_1 + X_2 + \dots + X_{25});$ <p>T= 1990 al 2000; N= 3131; A=11; SLM</p> <p><math>X_1</math>= Population growth,  <math>X_2</math>= Vehicles per household,  <math>X_3</math>= % of population between 25 and 45 years old,  <math>X_4</math>= % of population over 65 years old,  <math>X_5</math>= Average number of children per household,  <math>X_6</math>= % of residents with a high degree of education,  <math>X_7</math>= Poverty level approximated by % of resident population without schooling,  <math>X_8</math>= % Immigrants,  <math>X_9</math>= % Employed in manufacturing,  <math>X_{10}</math>= % Retail Employees,  <math>X_{11}</math>= % Employed in other services,  <math>X_{12}</math>= Maximum average temperature,  <math>X_{13}</math>= Minimum average temperature,  <math>X_{14}</math>= Average precipitation,  <math>X_{15}</math>= % Open space,  <math>X_{16}</math>= % Sports and leisure facilities,  <math>X_{17}</math>= Water availability (%),  <math>X_{18}</math>= Inland waters (%),  <math>X_{19}</math>= Terrain roughness index,  <math>X_{20}</math>= Elevation range (km),  <math>X_{21}</math> = Road density, <math>X_{22}</math> = Distance to road (km),  <math>X_{23}</math> = Distance to city center,  <math>X_{24}</math>= Property tax revenues,  <math>X_{25}</math>= Intergovernmental transfers as % of local revenues.</p>
<p><b>[32] - CHINA</b></p> <p>Rapid urbanization in China causing major impacts on land use and scarcity of land resources becoming a constraint to sustainable urban development.</p>	$Y = f(X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8);$ <p>T= 2010 a 2016; N= 287; A= 7; SLM; SEM; GMI</p> <p><math>Y</math> = Urban land use efficiency,  <math>X_1</math> =Economic agglomeration (population density),  <math>X_2</math> = Industrial structure (GDP tertiary industry/GDP secondary industry),  <math>X_3</math> = Economic development (GDP per capita),  <math>X_4</math> = Government intervention (financial expenditure per capita),  <math>X_5</math> = Investment in science and education (science and education expenditure / fiscal expenditure),  <math>X_6</math> = Environmental governance (Gas treatment, wastewater treatment and waste disposal fees),  <math>X_7</math> = Energy consumption intensity (Unit of GDP per energy consumption),  <math>X_8</math> = Marquetization of land</p>
<p><b>[33] - CHINA</b></p> <p>Serious environmental problems in China are the result of economic development and high levels of industrial consumption and emissions. This pollution is accompanied by negative externalities in the form of overflow effects in pollutant discharges.</p>	$Y = f(X_1 + X_2 + X_3 + X_4 + X_5 + X_6);$ <p>T= 2003 al 2017; N= 31; A= 15; SDM; GMI</p> <p><math>Y</math> = Environmental pollution index,  <math>X_1</math>= Fiscal Decentralization,  <math>X_2</math>= Environmental regulation,  <math>X_3</math> = GDP per capita,  <math>X_4</math> = Industrial structure,  <math>X_5</math>= FDI,  <math>X_6</math> = Urbanization.</p>

Table 3 (continued)

AUTHOR / COUNTRY / PROBLEM	ECONOMETRIC MODEL SPECIFICATION
<p><b>[34] - CHINA</b></p> <p>Resource scarcity and increasingly severe environmental pollution in China.</p>	$Y = f(X_1 + X_2 + X_3 + X_4 + X_5);$ <p>T = 2008 a 2017; N= 11; A= 10; SDM; GMI</p> <p>Y = Efficiency of industrial green technology innovation</p> <p>X<sub>1</sub>= Economic development,</p> <p>X<sub>2</sub> = Environmental regulation,</p> <p>X<sub>3</sub> = Government support for technological innovation (government funding of R&amp;D funds),</p> <p>X<sub>4</sub> = FDI,</p> <p>X<sub>5</sub> = Industrial structure.</p>
<p><b>[35] - CHINA</b></p> <p>With the rapid growth of economic development, the imbalance of energy supply and demand pose a critical challenge to our country's energy security. Meanwhile, the increasing and excessive energy consumption leads to greenhouse effect and air pollution, greatly threatening the survival and development of human beings.</p>	$Y = f(X_1 + X_2 + X_3 + X_4);$ <p>T= 1995 a 2016; N= 354; A= 22; SDM; GMI</p> <p>Y = Energy intensity (total energy consumption / GDP),</p> <p>X<sub>1</sub> = Economic growth (GDP),</p> <p>X<sub>2</sub> = Urbanization rate (urban population / total population),</p> <p>X<sub>3</sub> = Industrial structure (industrial GDP / GDP),</p> <p>X<sub>4</sub> = FDI</p>
<p><b>[36] - CHINA</b></p> <p>With sustained economic development, the ecological environment of China is becoming increasingly fragile and the problem of haze pollution is becoming more and more prominent, which has affected the normal life of human beings and the stable development of society.</p>	$Y = f(X_1 + X_2 + X_3 + \dots + X_8);$ <p>T= 1998 a 2016; N= 287; A= 19; SDM; GMI; LMI</p> <p>Y = Haze measurement index (PM<sub>2.5</sub>),</p> <p>X<sub>1</sub> = Economic development (LnGDP),</p> <p>X<sub>2</sub> = Population density,</p> <p>X<sub>3</sub> = Industrial structure,</p> <p>X<sub>4</sub> = Urbanization (average number of employees at work),</p> <p>X<sub>5</sub> = Technical level (% science expenditure in public finances),</p> <p>X<sub>6</sub> = FDI,</p> <p>X<sub>7</sub> = Market scale (total retail sales of social consumer goods),</p> <p>X<sub>8</sub> = Transportation intensity (urban road area per capita).</p>
<p><b>[37] - CHINA</b></p> <p>Over the past four decades, urbanization in China has rapidly increased from 17.92% in 1978 to 58.52% in 2017. This rapid urbanization is accompanied by urban population agglomeration, urban land use and serious industrial emissions leading to high ambient air pollution.</p>	$Y = f(X_1 + X_2 + X_3 + X_4 + X_5);$ <p>T= 2006 a 2016; A=11; SLM</p> <p>Y = PM<sub>2.5</sub>,</p> <p>X<sub>1</sub> = Urbanized population,</p> <p>X<sub>2</sub> = Population density,</p> <p>X<sub>3</sub> = GDP per capita</p> <p>X<sub>4</sub> = Technological progress (% financial expenditure on science and education),</p> <p>X<sub>5</sub> = Industrial structure.</p>

Table 3 (continued)

AUTHOR / COUNTRY / PROBLEM	ECONOMETRIC MODEL SPECIFICATION
<p><b>[38] - CHINA</b></p> <p>Recently, numerous urban areas in China have experienced heavy frequent <math>PM_{2.5}</math> pollution.</p> <p>The Yangtze River Delta [YRD] is a rapidly urbanizing region in China that experienced severe haze pollution in recent decades.</p>	$Y = f(X_1 + X_2 + \dots + X_4);$ T=2002 a 2017; N= 41; A= 16; SLM; SEM; SDM; GMI $Y = PM_{2.5},$ $X_1$ = Population urbanization, $X_2$ = Territorial urbanization (% urban built-up area), $X_3$ = Economic urbanization (regional GDP per capita), $X_4$ = Secondary industry, $X_5$ = Vegetation cover (vegetation area/total area), $X_6$ = Precipitation, $X_7$ = Wind speed.
<p><b>[39] - CHINA</b></p> <p>China has already become the world's largest coal consumer since 2010. Consequently, pollutants such as sulfur dioxide [<math>SO_2</math>] have been emitted enormously for years, which has led China to be the world's largest emitter of pollutants. In particular, over the past two decades, most Chinese cities have still suffered from the severity and extent of <math>SO_2</math> pollution.</p>	$Y = f(X_1 + X_2 + \dots + X_6);$ T= 2008 a 2017; N= 30; A= 10; SLM; SDM; GMI Y= Per capita Sulfur Dioxide Emissions ( $SO_2$ ) $X_1$ = GDP, $X_2$ = FDI, $X_3$ = Secondary industry, $X_4$ = The commerce, $X_5$ = Fossil fuel combustion, $X_6$ = Environmental regulation.
<p><b>[40] - CHINA</b></p> <p>Chinese scale-driven economic growth mode has achieved spectacular success in recent decades. At the same time, rapid economic development has given rise to enormous challenges related to environmental issues, such as rising energy consumption and energy-related <math>CO_2</math> emissions.</p>	$Y = f(X_1 + X_2 + X_3 + X_4);$ T= 1998 a 2015; N=31; A=18; SLM; SEM Y = Carbon Dioxide Emissions ( $CO_2$ ) $X_1$ = Energy structure (% coal consumption/total energy consumption), $X_2$ = Industrial structure (% of industrial GDP/total GDP), $X_3$ = Urbanization level (urban population and total population), $X_4$ = FDI, $X_5$ = Rail mileage (unit area average rail mileage within a given region).
<p><b>[41] - CHINA</b></p> <p>Sustainable development has been seriously challenged by global climate change due to carbon emissions. As a developing country, China pledged to reduce its carbon intensity by 40% to 45% below the 2005 level by 2020. The realization of this target depends not only on substantial transition of the society and economy at the national scale, but also the action and participation of energy saving and emission reduction at the provincial scale.</p>	$Y = f(X_1 + \dots + X_8);$ T= 1997 a 2010; N=30; A= 14; SLM; SEM; SDM; GMI; LMI Y= Carbon intensity, $X_1$ = Total population, $X_2$ = GDP per capita, $X_3$ = Energy intensity (energy consumption/GDP), $X_4$ = Energy structure (% coal consumption/total energy consumption), $X_5$ = Industrial structure, $X_6$ = Urbanization rate (urban population/total population), $X_7$ = Foreign trade openness ((Import - Export) / GDP), $X_8$ = FDI.

## 5. Conclusions

Spatial econometric techniques tend to be increasingly used in empirical research field, providing advances towards the knowledge of the spatial interdependence relationships between socioeconomic drivers that directly and indirectly influence the environment quality.

This literature review presents the basic theoretical aspects of spatial econometrics, as well as an overview of publications in the area, highlighting applications of those techniques in the study of environmental pollution problems and presenting several articles that applied the most widely used methods of spatial econometrics (exploratory and explanatory) to try to answer how socioeconomic drivers impact the environment, discovering how these impacts occur from a spatial perspective. Spatial phenomena are relevant in environmental sciences, as much as in economic-business sciences, health sciences,

among others. Moving towards appropriate methods of spatial estimation improves the quality of the results.

As environmental problems become more acute, as is the case with severe air pollution problems in China, researchers should be able to apply diverse types of methods to promote the search for answers which could help to implement appropriate public policies for pollution mitigation.

**Author Contributions:** M.C. and L.E. designed the study, analyzed the articles, and wrote the manuscript. D. S. and A.A. revised the manuscript. All authors reviewed the manuscript and approved the final version.

**Funding:** This research received no external funding. The APC was funded by L.E.

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

## References

1. Touchon JC, McCoy MW. The mismatch between current statistical practice and doctoral training in ecology. *Ecosphere*. 2016;7(8):e01394.
2. Fotheringham A, Brunsdon C, Charlton M. *Quantitative Geography: Perspectives on Spatial Data Analysis*. SAGE; 2000. 283 p.
3. Anselin L. Spatial Effects in Econometric Practice in Environmental and Resource Economics. *Am J Agric Econ*. 1 de febrero de 2001;83.
4. Goodchild MF, Anselin L, Appelbaum RP, Harthorn BH. Toward spatially integrated social science. *Int Reg Sci Rev*. 2000;23(2):139-59.
5. Bockstael NE. Modeling Economics and Ecology: The Importance of a Spatial Perspective. *Am J Agric Econ*. 1 de diciembre de 1996;78(5):1168-80.
6. Bohórquez IA, Ceballos HV. Algunos conceptos de la econometría espacial y el análisis exploratorio de datos espaciales. *Ecos Econ Lat Am J Appl Econ*. 2008;12(27):9-2.
7. Uriel E. El modelo de regresión simple: estimación y propiedades. *Introd Econ*. 2013;1-49.
8. Pérez Pineda JA. Econometría espacial y ciencia regional. *Investig Económica*. 2006;65(258):129-60.
9. Batarce M, Manterola P. Efectos espaciales en modelos econométricos de demanda de transporte. 2005;
10. LeSage JP. *The theory and practice of spatial econometrics*. Vol. 28. Citeseer; 1999.
11. Yrigoyen CC. Modelos de Heterogeneidad Espacial. 8 de noviembre de 2004 [citado 26 de agosto de 2019]; Disponible en: <https://ideas.repec.org/p/wpa/wuwpem/0411004.html>
12. Vayá E, Moreno R. La Utilidad de la Econometría Espacial en el ámbito de la ciencia regional. *Fundación de Estudios de Economía Aplicada*; 2000.
13. Vilalta CJ. Cómo enseñar autocorrelación espacial. *Econ Soc Territ*. 2005;(18):323-33.
14. Anselin L. *Spatial Econometrics: Methods and Models* [Internet]. Springer Netherlands; 1988 [citado 8 de febrero de 2021]. (Studies in Operational Regional Science). Disponible en: [file:///C:/Users/hp/AppData/Local/Temp/1988\\_Book\\_SpatialEconometricsMethodsAndM.pdf](file:///C:/Users/hp/AppData/Local/Temp/1988_Book_SpatialEconometricsMethodsAndM.pdf)
15. Yrigoyen CC. *Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales*. Dirección General de Economía y Planificación; 2003.
16. Benjanuvatra S. *Essays on Spatial Econometrics* [PhD Thesis]. University of York; 2012.
17. Tobler WR. Cellular geography. En: *Philosophy in geography*. Springer; 1979. p. 379-86.
18. Baronio A, Vianco A, Rabanal C. Una introducción a la econometría espacial. 2012;33.
19. Borrego Sánchez JÁ. Modelos de regresión para datos espaciales. 2018;
20. Liu H, Fang C, Zhang X, Wang Z, Bao C, Li F. The effect of natural and anthropogenic factors on haze pollution in Chinese cities: A spatial econometrics approach. *J Clean Prod*. noviembre de 2017;165:323-33.
21. Ge X, Zhou Z, Zhou Y, Ye X, Liu S. A Spatial Panel Data Analysis of Economic Growth, Urbanization, and NOx Emissions in China. *Int J Environ Res Public Health*. abril de 2018;15(4):725.
22. Diao B, Ding L, Su P, Cheng J. The Spatial-Temporal Characteristics and Influential Factors of NOx Emissions in China: A Spatial Econometric Analysis. *Int J Environ Res Public Health* [Internet]. 4 de julio de 2018 [citado 29 de enero de 2021];15(7). Disponible en: <http://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=29973509&lang=pt-br&site=ehost-live>
23. Fan Q, Yang S, Liu S. Asymmetrically Spatial Effects of Urban Scale and Agglomeration on Haze Pollution in China. *Int J Environ Res Public Health* [Internet]. 5 de diciembre de 2019 [citado 28 de enero de 2021];16(24). Disponible en: <http://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=31817551&lang=pt-br&site=ehost-live>
24. Shao J, Ge J, Feng X, Zhao C. Study on the relationship between PM2.5 concentration and intensive land use in Hebei Province based on a spatial regression model. *PLoS ONE*. 18 de septiembre de 2020;15(9):1-21.
25. Liu Y, Dong F. How Industrial Transfer Processes Impact on Haze Pollution in China: An Analysis from the Perspective of Spatial Effects. *Int J Environ Res Public Health* [Internet]. 1 de febrero de 2019 [citado 28 de enero de 2021];16(3). Disponible en: <http://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=30717159&lang=pt-br&site=ehost-live>



26. Zhong Z, Zhang X, Gao W. Spatiotemporal Evolution of Global Greenhouse Gas Emissions Transferring via Trade: Influencing Factors and Policy Implications. *Int J Environ Res Public Health*. 15 de julio de 2020;17(14):5065.
27. Wang Y, Duan X, Wang L. Spatial-Temporal Evolution of PM 2.5 Concentration and its Socioeconomic Influence Factors in Chinese Cities in 2014-2017. *Int J Environ Res Public Health* [Internet]. 19 de marzo de 2019 [citado 29 de enero de 2021];16(6). Disponible en: <http://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=30893835&lang=pt-br&site=ehost-live>
28. Zheng Z, Yang Z, Wu Z, Marinello F. Spatial Variation of NO<sub>2</sub> and Its Impact Factors in China: An Application of Sentinel-5P Products. *Remote Sens*. enero de 2019;11(16):1939.
29. Ronaghi M, Reed M, Saghaian S. The impact of economic factors and governance on greenhouse gas emission. *Environ Econ Policy Stud*. 1 de abril de 2020;22(2):153-72.
30. Ren Y, Ren X, Hu J. Driving factors of China's city-level carbon emissions from the perspective of spatial spillover effect. *Carbon Manag*. 2 de noviembre de 2019;10(6):551-66.
31. Gómez-Antonio M, Hortas-Rico M, Li L. The Causes of Urban Sprawl in Spanish Urban Areas: A Spatial Approach. *Spat Econ Anal*. 2 de abril de 2016;11(2):219-47.
32. Han X, Zhang A, Cai Y. Spatio-Econometric Analysis of Urban Land Use Efficiency in China from the Perspective of Natural Resources Input and Undesirable Outputs: A Case Study of 287 Cities in China. *Int J Environ Res Public Health*. 2020;17(19):7297.
33. Chen X, Liu J. Fiscal Decentralization and Environmental Pollution: A Spatial Analysis. *Discrete Dyn Nat Soc*. 30 de abril de 2020;NA-NA.
34. Long R, Guo H, Zheng D, Chang R, Na S. Research on the Measurement, Evolution, and Driving Factors of Green Innovation Efficiency in Yangtze River Economic Belt: A Super-SBM and Spatial Durbin Model. *Complexity* [Internet]. 27 de octubre de 2020 [citado 21 de febrero de 2021];2020:e8094247. Disponible en: <https://www.hindawi.com/journals/complexity/2020/8094247/>
35. Yue Y, Tian L, Yue Q, Wang Z. Spatiotemporal Variations in Energy Consumption and Their Influencing Factors in China Based on the Integration of the DMSP-OLS and NPP-VIIRS Nighttime Light Datasets. *Remote Sens*. enero de 2020;12(7):1151.
36. Gan T, Yang H, Liang W, Liao X. Do economic development and population agglomeration inevitably aggravate haze pollution in China? New evidence from spatial econometric analysis. *Environ Sci Pollut Res*. 1 de febrero de 2021;28(5):5063-79.
37. Xie W, Deng H, Chong Z. The Spatial and Heterogeneity Impacts of Population Urbanization on Fine Particulate (PM<sub>2.5</sub>) in the Yangtze River Economic Belt, China. *Int J Environ Res Public Health*. 23 de marzo de 2019;16(6):1058.
38. Cheng L, Zhang T, Chen L, Li L, Wang S, Hu S, et al. Investigating the Impacts of Urbanization on PM<sub>2.5</sub> Pollution in the Yangtze River Delta of China: A Spatial Panel Data Approach. *Atmosphere*. octubre de 2020;11(10):1058.
39. Jiang L, Zhou H, He S. The role of governments in mitigating SO<sub>2</sub> pollution in China: a perspective of fiscal expenditure. *Environ Sci Pollut Res*. 1 de septiembre de 2020;27(27):33951-64.
40. Kangjuan Y, Wang Y. Does regional innovation system efficiency facilitate energy-related carbon dioxide intensity reduction in China? *Environ Dev Sustain*. 2020;1-25.
41. Cheng Y, Wang Z, Ye X, Wei YD. Spatiotemporal dynamics of carbon intensity from energy consumption in China. *J Geogr Sci*. 1 de agosto de 2014;24(4):631-50.