New Insight Into The Spatial Autoregressive Model of Industrial Wastewater Discharge In The Yellow River Basin

Libo Xia (201514321@stu.ncwu.edu.cn)
North China University of Water Resources and Electric Power https://orcid.org/0000-0002-4646-9547
Zhiliang Wang
North China University of Water Resources and Electric Power
Shuang Du
North China University of Water Resources and Electric Power
Decun Tian
North China University of Water Resources and Electric Power
Feng Chen
North China University of Water Resources and Electric Power

Research Article

Keywords: Yellow River Basin, Industrial Wastewater Discharge, Spatial Autoregressive, Environmental Kuznets Curve, Moran Index, Bayesian

Posted Date: December 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1122988/v1

New Insight into the Spatial Autoregressive Model of Industrial Wastewater Discharge in the Yellow River Basin

Libo Xia[a,*], Zhiliang Wang[b] , Shuang Du[a], Decun Tian[b], Feng Chen[a]

[a] *School of Environmental and Municipal Engineering, North China University of Water Resources and Electric Power, Zhengzhou, 450045, China*

[b] *College of Water Resources, North China University of Water Resources and Electric Power, Zhengzhou, 450045, China*

* Corresponding author, Libo Xia. E-mail address: 201514321@stu.ncwu.edu.cn

Abstract

This article has carried out a statistical analysis of the industrial wastewater discharge (IWD) and gross regional product (GRP) of 79 cities in the Yellow River Basin from 2003 to 2019. By calculating the Moran index of IWD and GRP, the study has found that a certain spatial autoregressive in space. There is an environmental Kuznets curve (EKC) between the environmental pollution and economic development of cities in the Yellow River Basin, and a spatial autoregressive is modelled by a set of random effects that are assigned a conditional autoregressive prior distribution. In the Bayesian environment, Markov chain Monte Carlo (MCMC) is used for inferencing, and the spatial weight matrix is selected to be U-shaped matrix, and the error of the model is minimized. The parameter posterior distribution results of the model showed that the GRP did not show a significant decline. The modified EKC showed that the discharge of industrial wastewater in the entire Yellow River Basin will be reduced. Generally, cities with high pollutant emissions should learn from other cities to reduce emissions, and cities with low GRP need to increase local economic development.

Keywords: Yellow River Basin; Industrial Wastewater Discharge; Spatial Autoregressive; Environmental Kuznets Curve; Moran Index; Bayesian

1 Introduction

As the mother river of China, the Yellow River's ecological environment is of far-reaching significance to the great rejuvenation of the Chinese nation. Through the analysis of the IWD of each city in the Yellow River Basin, suggestions are provided for the next step in the environmental pollution control of the Yellow River Basin. The problems of environmental pollution and economic development have been studied in the 1990s, and it was found that economic growth brought an initial stage of deterioration, followed by an improvement stage (Grossman and Krueger 1995)(Selden and Song 1994). However, from the clean agricultural economy to the polluting industrial economy, some people questioned the existing evidence. The pollution caused by the environment is reduced, and the corresponding economic income level has not been agreed, and may lead to the demise of the classic EKC (Dinda 2004) (Stern 2004), the evidence for the existence of EKC was also questioned (Kijima et al. 2010). After a certain period of research on environmental pollution and economic development, it was found that environmental pollution was not only related to economic levels, but many other factors were slowly being taken into account (Managi 2006). Research on economic development and air pollution was closely related, the regional differences in the relationship between haze pollution and economic growth (Du et al. 2018). A survey of the empirical literature on carbon dioxide emissions found that EKC's estimation of carbon dioxide emissions was essentially uncertain (Shahbaz and Sinha 2019). China's environmental pollution had obvious spatial autoregressive, and economic development and environmental pollution present an inverted U-shaped curve (Cheng et al. 2017). Based on the

Bayesian model averaging method, the influencing factors of air pollution was analyzed, which proved the existence of EKC (Lin and Zhu 2018). Economic globalization will increase the degree of environmental pollution and had an impact on the ecological footprint, and the inverted U-shape was existed (Ulucak and Bilgili 2018). The treatment of environmental pollution should implement different measures according to different spaces and different stages of development (Zhao et al. 2019). And industrial transfer and energy consumption structure were determinants of pollutant emissions in China (Li et al. 2020). In order to accelerate EKC to reach the turning point faster, it was necessary to promote renewable energy consumption (Yao et al. 2019). And emissions and income inequality had a major impact on the EKC turning point (Ridzuan 2019).

When calculating the relationship between environmental pollution and economic development, not just the total statistics and the environmental pollution and economic development of each city, but also the relationship between the spatial location and the calculation of its temporal and spatial characteristics. The Geographically Weighted Regression (GWR) model greatly improves the performance of the model compared to the traditional least square regression model. The GWR model also improves the reliability of the relationship by reducing the spatial autocorrelation (Tu and Xia 2008). Economic development was the main factor in the changes in industrial wastewater discharge in each province during the study period (Geng et al. 2014). The spatial and temporal distribution characteristics of total wastewater discharge in 31 provinces in China from 2002 to 2013 were

analyzed (Chen et al. 2016). According to the results of spatial correlation analysis, the province's nitrogen oxide emissions changes not only affect the province itself, but also the neighboring areas (Diao et al.). A data set composed of 173 countries was used to estimate the EKC that was increased by neighboring per capita income and energy intensity (Zhang et al. 2019). The spatial correlation analysis using global and local Moran I value shown that a significant and positive spatial autocorrelation in environmental, economic and energy aspects, and cities in China was showing a trend of decoupling from economic development (Ma et al. 2020). The study found that encouraging cooperation between neighboring provinces may help reduce the IWD (Zhang et al. 2020). Economic development, investment in environmental governance, industrial structure, and technological innovation all had a certain impact on the amount of IWD, and also shown significant spatial spillover effects (Bu et al. 2021).

In the current research on environmental pollution and economic development, some do not consider spatial factors, but do simple statistical analysis to study the linear relationship between environmental pollution and economic development. The environmental pollution in the early stage increases with economic growth, reaching a certain level. After that, it will decrease with the growth of the economic level, showing an inverted U-shape. Others only consider spatial factors, and had not studied the specific impact of spatial changes on environmental pollution data. This study will use panel data from 79 cities in the Yellow River Basin. In spatial analysis, by calculating different spatial matrices, the spatial influence was calculated

according to the Moran index. Then, in the Bayesian setting, Markov chain Monte Carlo simulation was used to include the spatial influence in the random effects, so that the error of the model was reduced, and the most suitable model for the discharge of industrial wastewater in the Yellow River Basin was found.

2 Methods and data

2.1 Study area

The Yellow River Basin refers to the geographic and ecological area affected by the river from the source to the sea. The relevant area of the Yellow River flowing through the provinces is often referred to as the Yellow River Basin. The Yellow River Basin has a very important position in my country's economic and social development and ecological security. The ecological protection and high-quality development of the Yellow River Basin are of great significance. According to the Yellow River Conservancy Commission's survey of the cities included in the Yellow River Basin, in figure 1, all city in Qinghai Province, Gansu Province, Ningxia Hui Autonomous Region, Inner Mongolia Autonomous Region, Sichuan Province, Shaanxi Province, Shanxi Province, Henan Province, and Shandong Province can be found. According to the availability of data and the convenience of research, 79 cities were selected for EKC analysis of their IWD and GRP.

2.2 Data

The data comes from the *China City Statistical Yearbook* in the National Bureau of Statistics (http://www.stats.gov.cn/), using data from various cities from 2003 to 2019. Discharge of industrial wastewater refers to the total amount of waste water discharged outside the enterprise through all the outlets of the factory area of an industrial enterprise. In some enterprises, the mixed discharge of indirect cooling water and direct cooling water is inseparable and can be included in the statistics. The GRP refers to the final result of production activities of all permanent resident units in

126 a region in a certain period of time calculated at market prices. Therefore, the average
127 value of each city's data from 2003 to 2019 is used, and logarithmic conversion of the
128 volume of industrial waste water discharged (10000 tons) and gross regional product
129 (10000 yuan).

130 *2.3 Model setting*

131 The EKC has different manifestations in different places. It is generally divided
132 into a linear function curve. Environmental pollution has been increasing with the
133 development of the economic level (1). It is also an inverted U-shaped curve. When
134 the economic level reaches a certain level, people pay more attention to
135 environmental pollution and begin to invest in environmental governance (2). The
136 problem of environmental pollution began to decrease. Another is that the pollution
137 problem first increases, after reaching a certain level, there is a certain decrease, and
138 then it starts to increase (3).

139 $$Y_k=\beta_0+\beta_1 X_k+\varepsilon \quad (1)$$

140 $$Y_k=\beta_0+\beta_1 X_k+\beta_2 X_k^2+\varepsilon \quad (2)$$

141 $$Y_k=\beta_0+\beta_1 X_k+\beta_2 X_k^2+\beta_2 X_k^2+\varepsilon \quad (3)$$

142 where, Y_k is logarithm of the IWD, X_k is logarithm of the GRP, $\varepsilon\sim(0, \sigma^2)$.

143 Many problems in spatial data analysis can be explained as image restoration
144 problems. In archaeology and epidemiology, pixel-based Bayesian image analysis is
145 reviewed. The problem of spatial data analysis can be solved by Bayesian method, in
146 which the prior distribution includes known spatial relations (Besag et al. 1991). A
147 new model of spatial dependence, including separate parameters for over-dispersion

148 and spatial dependence intensity. The new dependency structure is incorporated into a
149 general linear mixed model to estimate disease incidence in small geographic areas.
150 The hybrid model allows the local smoothing rate to be adjusted by estimating
151 spatially correlated random effects. Computer simulation studies compare the new
152 model with the built-in autoregressive model, independent model, and non-random
153 effects model (Leroux et al. 2000). Taking the space factor into account, the error of
154 the model is smaller. Spatial autocorrelation is usually used to represent data related to
155 a set of non-overlapping regions. These data appear in a variety of applications such
156 as agriculture, education, epidemiology, and image analysis. Spatial conditional
157 autoregressive models are usually in a hierarchical Bayesian framework, based on
158 MCMC simulation for reasoning. The CARBayes package in R language describes its
159 implementation through MCMC simulation technology, and introduces the use of this
160 model in two working examples in the field of housing price analysis and disease
161 mapping. Disease mapping is the field of epidemiology, which estimates the spatial
162 pattern of disease risk in an extended geographic area in order to identify areas with
163 higher risk levels (Lee 2013). The Bayesian hierarchical model uses a combination of
164 available co-occurrence data and a set of spatial random effects to represent the risk
165 surface. These random effects include those used to simulate any excessive dispersion
166 or spatial correlation in disease data. This is illustrated by using Bayesian analysis to
167 apply conditional autoregressive models to phosphate data sets from archaeological
168 areas. Conditional autoregressive models have been widely used to analyze spatial
169 data in different fields such as population, economy, epidemiology, and geography, as

a model of potential and observed variables (De Oliveira 2012).

$$Y_k|\mu_k \sim f(y_k|\mu_k,\upsilon^2) \quad (4)$$

$E(Y_k)=\mu_k$, K is 79, 79 cities in study area, $\nu^2 \sim (1,0.01)$ inverse gamma prior distribution

$$g(\mu_k)=X_k^T\beta+O_k+\psi_k \quad (5)$$

$$Y_k \sim N(\mu_k,\nu^2) \quad (6)$$

$$\psi_k|\psi_{-k},\ W,\ \tau^2,\ \rho \sim N\left(\frac{\rho\sum_{i=1}^{K} w_{ki}\psi_i}{\rho\sum_{i=1}^{K} w_{ki}+1-\rho}, \frac{\tau^2}{\rho\sum_{i=1}^{K} w_{ki}+1-\rho}\right) \quad (7)$$

In the Bayesian hierarchical function, the connection function is $g(\mu_k)$, the covariate matrix $X = (X_1,\ldots, X_k)$, the offset $O_k= (O_1, \ldots, O_k)$, the spatial structure $\psi_k=(\psi_1,...,\psi_k)$, a multivariate Gaussian prior is assumed for regression parameter β, where $\beta \sim N(\mu_\beta \Sigma_\beta)$, μ_β is the mean of β, diagonal variance matrix Σ_β, $\tau^2 \sim (1,0.01)$ inverse gamma distribution, $\rho \sim (0,1)$ uniform distribution, W is the spatial weight matrix.

3 Results and Discussions

3.1 Data analysis

According to the data of the statistical yearbook, it can be found from the analysis that of the 79 cities studied, Weifang had the highest industrial wastewater emissions, with 2,091 tons, the lowest was Qingyang City, at 246 tons, and the average industrial wastewater emissions across the Yellow River basin was 5,133 tons. In the analysis of the GRP, the GRP of Guyuan City was the least, 1,411,914 yuan, Qingdao was the largest, at 6 3,912,288 yuan, and the average GRP was 13,429,594 yuan. There is a relationship between the GRP and the IWD, and also some spatial differences observe in the EKC. Qingdao City's regional gross domestic product is the largest, industrial wastewater emissions are not Weifang City, the city has begun to pay attention to environmental pollution, take certain measures to reduce pollutant emissions. Weifang City's industrial wastewater emissions, causing a lot of pollution to the environment, at the cost of environmental damage, but did not achieve a certain degree of economic development.

3.2 Model selecting

By comparing the mean square error (MSE) and root mean square error (RMSE) of the above models, the analysis found that the overall EKC in this study area is not much different, and the error of the model is relatively small in table 1. The calculation result is $Y_k=593.94-113.36X_k+7.23X_k^2-0.15\beta_2X_k^3+\varepsilon$, in the figure 4a. In these cities in the Yellow River Basin, as the GRP increases, the discharge of industrial wastewater is still rising, and it is not downward trend.

$$MSE=\frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i-y_i\right| \quad (8)$$

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i-y_i\right)^2} \quad (9)$$

In the Bayesian environment, using MCMC for inference, without calculating the spatial weight, the parameter posterior distribution of X_k, X_k^2, X_k^3 is shown in the following table 2. The result of the model is $Y_k=491.51-94.11X_k+6.03X_k^2-0.15\beta_2X_k^3$, in the figure 4b.

3.3 Spatial analysis

In general, observations from regional units that are close together tend to have similar values. In the spatial analysis, the Moran index (Bivand et al. 2018) of the IWD and the GRP was calculated separately, and it was found that there was a certain degree of spatial autocorrelation, so the parameters of the model cannot be simply determined, and the spatial factors must be considered. As shown in table 3, where z-score is greater than 2.58, or p-value is less than 0.01, it can be explained that it is not spatial randomness, but has a certain spatial correlation. And the spatial distribution characteristics of IWD, the GRP and residuals (ε) are obvious from fig 2a,2b,2c.

$$I=\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}W_{i,j}(X_i-\bar{X})(X_j-\bar{X})}{\left(\sum_{i=1}^{n}\sum_{j=1}^{n}W_{i,j}\right)\sum_{i=1}^{n}(X_i-\bar{X})^2} \quad (10)$$

where, n is 79cities, X_i and X_j represent IWD in area i and j, $\bar{X}$ is the average of IWD in 79 cities, and $W_{i,j}$ is the spatial weight matrix.

By calculating the Moran index of the model's residual, it is found that the error also has a certain spatial agglomeration, so the model needs further parameter determination. The residual autocorrelation reduces the error by using a set of spatially correlated random effects to increase the linear predictor variable as part of

the Bayesian hierarchical model. Spatial random effects are taken into account and are usually expressed in conditional autoregressive models. The model induces spatial autocorrelation through the adjacent structure of regional units. Part of this spatial autocorrelation can be modeled by including known covariate risk factors in the regression model, but after considering these covariate effects, the spatial structure is usually retained in the residuals, which reduces the error of the model.

Then Choosing different spatial weight matrices for spatial analysis, where B is the basic binary code, that is, the spatial weight matrix with adjacent is 1 and non-adjacent is 0. W is the row standardization of B, C is the matrix standardization of B, U is equal to C divided by the number of neighbors, "minmax", the weight is divided by the minimum value of the maximum number of rows and the maximum number of columns of the input weight (Kelejian and Prucha 2010). The models without considering the spatial factors are calculated separately, and the Deviance Information Criteria (DIC) is 160.15. Compared with other weight matrices for calculating the spatial influence, it can be found that the DIC is the largest, so the spatial influence of the model needs to be included. Then, 4 different spatial weights were selected, and found that when the U matrix is selected, the DIC is -15.31, and the model has the smallest error, so the spatial weight matrix is selected as U for the next calculation, as shown in table 4.

We can write the Bayesian deviance (Bakar 2012) as: $D(\theta)=-2L(\theta \mid y)$, $L(\theta \mid y)$ is the log-likelihood of the model.

Now the goodness of fit of a model is obtained as $\bar{D}=E_{\theta \mid y}(D)$ and the model

complexity is written as:$p_D=E_{\theta \mid y}(D)-D\left[E_{\theta \mid y}(\theta)\right]$

$$DIC=E_{\theta \mid y}(D)+p_D=D\left[E_{\theta \mid y}(\theta)\right]+2p_D \quad (11)$$

Watanabe-Akaike information criteria

$$WAIC = \sum_{i=1}^{n} log\, E_{\theta|y}[f(y_i \mid \theta)] + \sum_{i=1}^{n} V_{\theta|y}[log\, p(y_i \mid \theta)] \quad (12)$$

Log pseudo marginal likelihood

$$LPML=\sum_{i=1}^{n} \log\left[f\left(y_i \mid y_{(-i)}\right)\right] \quad (13)$$

3.4 Bayesian inference

After taking the spatial autocorrelation into account, the Bayesian hierarchical model is adopted, and three Markov chains are selected for calculation. The inference of this model is based on three parallel Markov chains. Each chain has run 300,000 samples, and 100000 have been deleted as the burning period. The remaining 200,000 samples were diluted 100 times, resulting in 6,000 samples used to infer 3 Markov chains. The 3 Markov chains is calculated, the posterior distribution of the parameters is shown in the table 5, and the posterior median of the parameters are -91.25, 5.86, and -0.12 respectively, as shown in figure 3. The sample diagram of the simulated parameters is as follows, and the parameters are all convergent.

It can be known through analysis, under the Bayesian hierarchical structure, the spatial error is calculated. After a large number of MCMC simulations, it can be found that the amount of IWD will be reduced in the cities in the Yellow River Basin and the current the comparison. With the development of the economy, there will be more investment in reducing IWD, which will eventually be the curve shown in the figure 4c. The entire Yellow River Basin, as a whole, will have spatial differences in

pollutant emissions, but it will be no spatial differences in the curve of the entire model.

4 Conclusions

This paper studies the discharge of industrial wastewater in prefecture-level cities in the Yellow River Basin, and analyzes the relationship between the discharge of industrial wastewater and the GRP. Results showed that both the amount of IWD and the GRP have a certain spatial autocorrelation, and the error of fitting using the EKC also has a certain spatial autocorrelation. The city with the largest amount of IWD is Weifang City, Qingdao City has the largest GRP, and Jiayuguan City has the largest error in the model. In the Bayesian environment, the inference found that the medians of the posterior parameter distributions of X_k, X_k^2, and X_k^3 are -91.25, 5.86, -0.12. Bayesian inference suggested the model of $Y_k=474.79-91.25X_k+5.86X_k^2-0.12\beta_2X_k^3$, and the DIC of the model is -15.31. The discharge of industrial wastewater was increased with the increasing of GRP, and not showed a certain downward trend. The Bayesian hierarchical structure showed that the spatial error is calculated within the error. After a large number of MCMC simulations, it can be found that the amount of IWD in the cities of the Yellow River Basin should be less than it is now. Results also suggested that the prior space-time effects should be also considered in the spatiotemporal autoregressive model.

Acknowledgments

This paper thanks the National Bureau of Statistics (http://www.stats.gov.cn/) and the package CARBayes (https://cran.r-project.org/web/packages/CARBayes) in the open-source software R language (https://www.r-project.org/).

Declarations

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent to publish Not applicable.

Funding Not applicable.

Competing interests All authors declare no competing interests.

Availability of data and materials All data generated or analyzed in the current study was included in this article.

Author contribution Libo Xia: methodology, software, data analysis, writing—original draft. Zhiliang Wang: writing—review and editing. Shuang Du: writing and editing. Decun Tian: writing and editing. Feng Chen: writing and editing. All authors read and approved the final manuscript.

References

Bakar KS (2012) Bayesian analysis of daily maximum ozone levels

Besag J, York J, Mollié A (1991) A Bayesian image restoration with two applications in spatial statistics Ann Inst Statist Math 43: 1–59. Find this Artic online 43:1–20

Bivand RS, David ·, Wong WS, Wong DWS (2018) Comparing implementations of global and local indicators of spatial association. TEST 27:716–748. https://doi.org/10.1007/s11749-018-0599-x

Bu Y, Wang E, Jiang Z (2021) Evaluating spatial characteristics and influential factors of industrial wastewater discharge in China: A spatial econometric approach. Ecol Indic 121:107219. https://doi.org/10.1016/J.ECOLIND.2020.107219

Chen K, Liu X, Ding L, et al (2016) Spatial characteristics and driving factors of provincial wastewater discharge in China. Int J Environ Res Public Health 13:.

https://doi.org/10.3390/IJERPH13121221

Cheng Z, Li L, Liu J (2017) Identifying the spatial effects and driving factors of urban PM2.5 pollution in China. Ecol Indic 82:61–75. https://doi.org/10.1016/j.ecolind.2017.06.043

De Oliveira V (2012) Bayesian analysis of conditional autoregressive models. Ann Inst Stat Math 64:107–133. https://doi.org/10.1007/s10463-010-0298-1

Diao B, Ding L, Su P, Cheng J The Spatial-Temporal Characteristics and Influential Factors of NOx Emissions in China: A Spatial Econometric Analysis. https://doi.org/10.3390/ijerph15071405

Dinda S (2004) Environmental Kuznets Curve hypothesis: A survey. Ecol Econ 49:431–455. https://doi.org/10.1016/j.ecolecon.2004.02.011

Du G, Liu S, Lei N, Huang Y (2018) A test of environmental Kuznets curve for haze pollution in China: Evidence from the penal data of 27 capital cities. J Clean Prod 205:821–827. https://doi.org/10.1016/j.jclepro.2018.08.330

Geng Y, Wang M, Sarkis J, et al (2014) Spatial-temporal patterns and driving factors for industrial wastewater emission in China. J Clean Prod 76:116–124. https://doi.org/10.1016/j.jclepro.2014.04.047

Grossman GM, Krueger AB (1995) Economic Growth and the Environment*. Q J Econ 110:353–377. https://doi.org/10.2307/2118443

Kelejian HH, Prucha IR (2010) Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. J Econom 157:53–67. https://doi.org/10.1016/J.JECONOM.2009.10.025

Kijima M, Nishide K, Ohyama A (2010) Economic models for the environmental Kuznets curve: A survey. J Econ Dyn Control 34:1187–1201. https://doi.org/10.1016/j.jedc.2010.03.010

Lee D (2013) CARBayes: An R package for Bayesian spatial modeling with conditional autoregressive priors. J Stat Softw 55:1–24. https://doi.org/10.18637/jss.v055.i13

Leroux BG, Lei X, Breslow N (2000) Estimation of Disease Rates in Small Areas: A new Mixed Model for Spatial Dependence. 179–191.

https://doi.org/10.1007/978-1-4612-1284-3_4

Li Z, Song Y, Zhou A, et al (2020) Study on the pollution emission efficiency of China's provincial regions: The perspective of Environmental Kuznets curve. J Clean Prod 263:121497. https://doi.org/10.1016/J.JCLEPRO.2020.121497

Lin B, Zhu J (2018) Changes in urban air quality during urbanization in China. J Clean Prod 188:312–321. https://doi.org/10.1016/J.JCLEPRO.2018.03.293

Ma B, Tian G, Kong L (2020) Spatial-temporal characteristics of China's industrial wastewater discharge at different scales. Environ Sci Pollut Res 27:8103–8118. https://doi.org/10.1007/s11356-019-07488-7

Managi S (2006) Pollution, natural resource and economic growth: An econometric analysis. Int J Glob Environ Issues 6:73–88. https://doi.org/10.1504/IJGENVI.2006.009401

Ridzuan S (2019) Inequality and the environmental Kuznets curve. J Clean Prod 228:1472–1481. https://doi.org/10.1016/J.JCLEPRO.2019.04.284

Selden TM, Song D (1994) Environmental quality and development: Is there a kuznets curve for air pollution emissions? J. Environ. Econ. Manage. 27:147–162

Shahbaz M, Sinha A (2019) Environmental Kuznets curve for CO2 emissions: a literature survey. J Econ Stud 46:106–168. https://doi.org/10.1108/JES-09-2017-0249

Stern DI (2004) The Rise and Fall of the Environmental Kuznets Curve. World Dev 32:1419–1439. https://doi.org/10.1016/j.worlddev.2004.03.004

Tu J, Xia ZG (2008) Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation. Sci Total Environ 407:358–378. https://doi.org/10.1016/j.scitotenv.2008.09.031

Ulucak R, Bilgili F (2018) A reinvestigation of EKC model by ecological footprint measurement for high, middle and low income countries. J Clean Prod 188:144–157. https://doi.org/10.1016/J.JCLEPRO.2018.03.191

Yao S, Zhang S, Zhang X (2019) Renewable energy, carbon emission and economic

growth: A revised environmental Kuznets Curve perspective. J Clean Prod 235:1338–1352. https://doi.org/10.1016/J.JCLEPRO.2019.07.069

Zhang P, Yang D, Zhang Y, et al (2020) Re-examining the drive forces of China's industrial wastewater pollution based on GWR model at provincial level. J Clean Prod 262:121309. https://doi.org/10.1016/j.jclepro.2020.121309

Zhang WW, Sharp B, Xu SC (2019) Does economic growth and energy consumption drive environmental degradation in China's 31 provinces? New evidence from a spatial econometric perspective. Appl Econ 51:4658–4671. https://doi.org/10.1080/00036846.2019.1593943

Zhao X, Zhou W, Han L, Locke D (2019) Spatiotemporal variation in PM2.5 concentrations and their relationship with socioeconomic factors in China's major cities. Environ Int 133:105145. https://doi.org/10.1016/J.ENVINT.2019.105145

TABLES

Table 1 Model comparison table

Model	MSE	RMSE	Adjusted R^2
$Y_k = \beta_0 + \beta_1 X_k + \varepsilon$	0.43	0.6559	0.5698
$Y_k = \beta_0 + \beta_1 X_k + \beta_2 X_k^2 + \varepsilon$	0.427	0.6532	0.5734
$Y_k = \beta_0 + \beta_1 X_k + \beta_2 X_k^2 + \beta_2 X_k^3 + \varepsilon$	0.414	0.6431	0.5864

401 Table 2 The parameter posterior distribution table has no spatial influence

	50%	2.5%	97.5%
Intercept	491.51	-125.75	1093.62
X_k	-94.11	-207.07	22.24
X_k^2	6.03	-1.19	13.06
X_k^3	-0.13	-0.27	0.02

402

403 Table 3 Moran Index Table

	IWD	GRP	ε
I	0.52	0.322754	0.17
z-score	9.220178	5.873645	3.16
p-value	0	0	0.0015

404

405 Table 4 Model selection with different spatial weights

Spatial matrix	weight	DIC	WAIC	LMPL
None		160.15	161.99	-81.13
B		83.82	137.18	-70.82
C		60.24	134.30	-70.07
minmax		25.89	127.59	-60.92
U		-15.31	117.03	-57.31

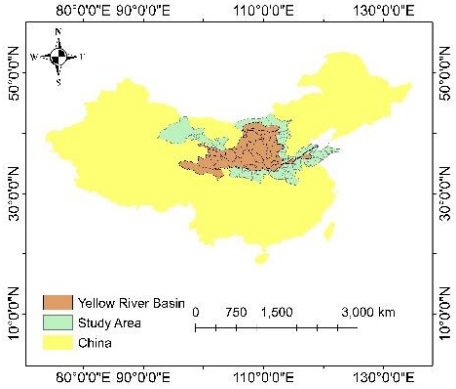
406

407 Table 5 Parameter posterior distribution table

	50%	2.5%	97.5%
Intercept	474.79	-127.81	1073.32
X_k	-91.25	-203.25	21.15
X_k^2	5.86	-1.17	12.81
X_k^3	-0.12	-0.26	0.02

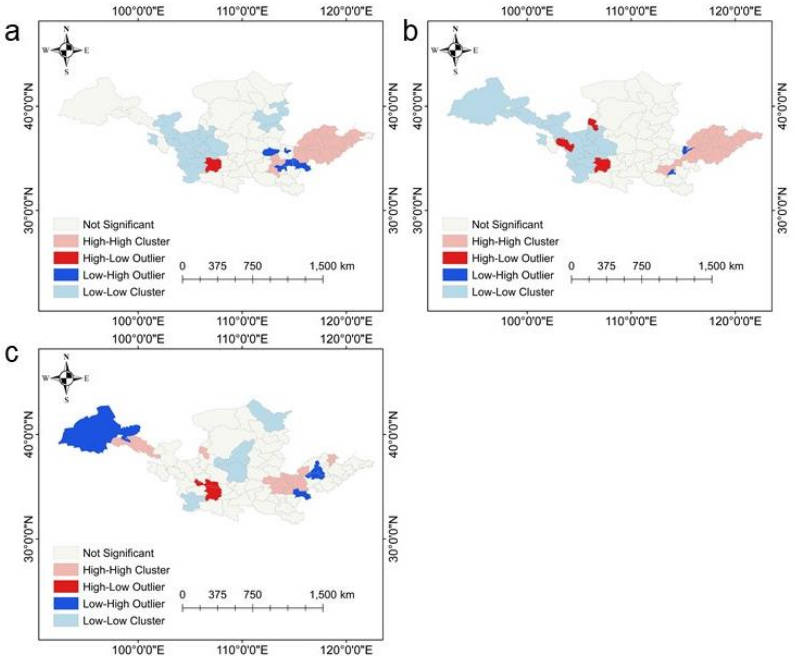
408

409 **FIGURES**



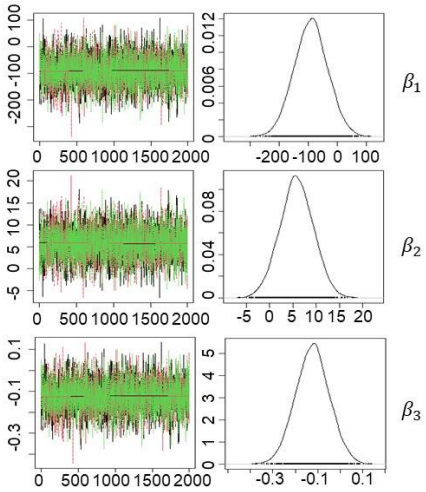
410

411 Fig 1 The map of the study area including 79 cities in the Yellow River Basin.



412

413 Fig 2 Spatial clustering of IWD, GRP and ε in 79 cities.



414

415 Fig 3 Convergence graph of Markov chains with posterior distribution of parameters β_1, β_2, β_3.

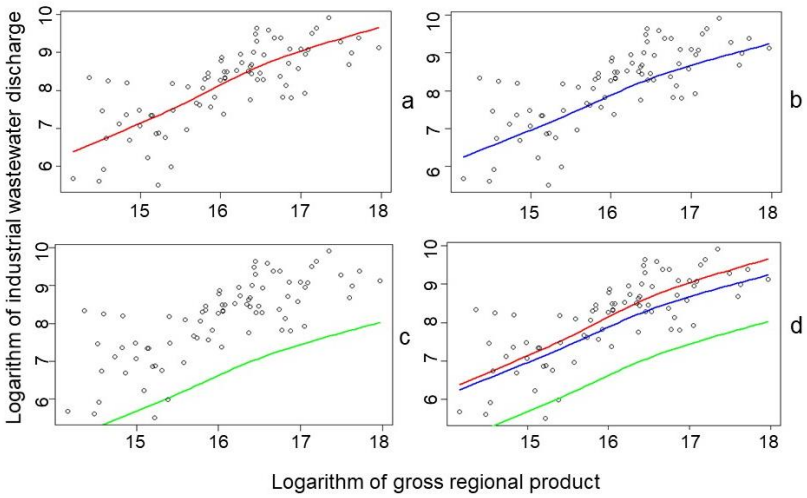


Fig 4 The fitting graphs, a is simple EKC, b is the model of autoregressive priors in Bayesian without spatial, and c is spatial autoregressive model in Bayesian setting, d is comparison among a, b and c.