ISSN: 2688-9323

Robot Application and Environmental Performance from the Perspective of Machine Learning Model: Evidecce from Four Provinces in China

Wenhua Qi*

School of Economics, Nanjing Audit University, Nanjing 211815, China

Abstract

When a new round of industrial revolution is booming, China is also facing pressure on environmental governance. Based on the panel data of the International Federation of Robotics (IFR) and four provinces in China: Jiangsu, Zhejiang, Shanghai and Anhui, this paper uses the random forest model to study the nonlinear relationship between robot application and SO_2 emission on the basis of the traditional econometric model. The results show that the application of robots can significantly reduce the emission of SO_2 emission, and its importance is comparable to the traditional factors that affect pollutant emission. This result provides policy reference for green development of cities.

Keywords

Random Forest; Robot Application; SO_2 Emission; Environmental Performance.

1. Introduction

In recent years, the Chinese government has attached great importance to the development of intelligence, and has issued policy documents such as "Made in China 2025" to support and encourage enterprises to invest in intelligent production. However, in the process of intelligent production represented by industrial robots, the environmental problems caused by it have gradually attracted widespread attention from society. Under the pressure of "carbon peak" goal, the government and academia carefully analyzed the impact mechanism and effect of robot application on carbon emission, but less research was invested in the analysis of other pollutants such as sulfur dioxide. Therefore, this paper selects regional sulfur dioxide emission as an indicator to measure environmental performance, and expands the research on the environmental impact of robot application

2. Literature Review

2.1. Literature on the Industrial Robots

Robots are devices that can perform tasks autonomously or semi-autonomously, and have a wide range of applications and impacts in fields such as industry, service, military, etc. In the new round of industrial revolution driven by artificial intelligence and robotics, the most thought-provoking question is the "machine substitution" problem. Its essence is the relationship between technological progress and labor employment. Along with the development of human society, this problem has always been a hot topic for scholars at home and abroad. Some studies have pointed out that in China, the application of robots will have a certain substitution effect on the labor demand of enterprises, which is especially significant in industries with high market concentration, high external financing dependence and non-state-owned enterprises (Wang and Dong, 2020), and will also have a negative impact on the labor employment level in the following year (Kong et al., 2020). Similarly, there are also similar problems in the labor market of developed countries in Europe and America (Acemoglu and

ISSN: 2688-9323

Restrepo, 2018). This substitution effect is a huge challenge for society: replacing some workers means reducing employment opportunities, which leads to widening income gap, increasing social inequality, and then causing various social problems. However, despite such a big challenge, the fact that robots promote economic development by improving total factor productivity has been confirmed by data from various countries in the world, including China (Yang and Hou, 2020). And promoting robot application can also promote the transformation and upgrading of manufacturing industry by improving production efficiency and optimizing factor allocation structure (Deng and Qu, 2021), which is crucial for the long-term development of the country. Therefore, the development of robots is both a challenge and an opportunity for China. However, it is well known that the first and second industrial revolutions brought about a series of environmental problems such as global warming while promoting human social progress. Will robots as a new round of industrial revolution also aggravate environmental pollution? This issue has attracted widespread attention from the government and academia. Huang Haiyan et al. (2021) empirically concluded that industrial intelligence can effectively reduce carbon emission intensity by using data from China's sub-industries, but still has a significant positive impact on total emissions. Luan et al. (2022) found that robot application would aggravate global warming by increasing energy consumption through analyzing data from 74 countries. Therefore, on the way to achieve "carbon peak" by 2030, the management of robot future development is imminent.

2.2. Literature on the Machine Learning

Machine learning is a method that uses data and algorithms to train models and make predictions and decisions, which has many new impacts on economic research. It can not only help researchers process and generate large-scale, high-dimensional, unstructured data (Huang and Yu, 2018), but also show its prowess in causal identification, prediction and other aspects (Wang et al., 2020). In the study of the relationship between technological progress and environment, many factors are complex and intertwined, and to a large extent they are not theoretically perfect linear relationships. And machine learning models, as an important frontier achievement in the field of artificial intelligence, are very suitable for the study of nonlinear relationship problems (Ang, 2019). Therefore, this paper innovatively introduces this model to study the nonlinear impact relationship between robot application and pollution emission. Another distinction that needs to be made is that machine learning is a small category of artificial intelligence, which is more concrete in algorithm programs; while the robots studied in this paper refer to the many devices invested in industrial production. Both can be "artificial intelligence" in a broad sense, but there are still significant differences in their essence and performance.

3. Data Variables and Research Design

3.1. Data Description

(1) The dependent variable is the industrial SO_2 emission of each city.

(2) The independent variable is the robot installation density (*robot_density*) of each city. Based on the International Federation of Robotics (IFR) database, this paper draws on the construction method of Bartik instrumental variable method widely used by the academic community (Acemoglu and Restrepo, 2020; Wei et al., 2020; Wang and Dong, 2020; Kong et al., 2020), and constructs the robot installation density at the city level. The specific calculation method is as follows:

$$robot_{density_{i,t}} = \sum_{j=1}^{3} \frac{\mathrm{employment_{i,j,t=2008}}}{\mathrm{employment_{i,t=2008}}} * \frac{\mathrm{robot_{j,t}}}{\mathrm{employment_{j,t=2008}}} \tag{1}$$

Where $robot\_density_{i,t}$ represents the robot installation density in city *i* and year *t*; $\frac{employment_{i,j,t=2008}}{employment_{i,t=2008}}$ represents the proportion of employment in the third industry in city *i* and based period(2008); $robot_{j,t}$ represents the number of robots installed in industry *j* and year *t*. $employment_{j,t=2008}$ represents the total employment of industry *j* in 2008.

(3) Control variables mainly include characteristic variables at the city level. Starting from the influencing factors of pollutants, this paper draws on the studies of Ma (2022), Deng (2023) and Wang (2023). The selected control variables mainly include economic development level (GDP, GDP per capita and GDP growth), energy consumption (total energy consumption and coal consumption proportion), industrial structure (industrial structure advanced index), as well as urbanization rate, green rate, capital stock and government expenditure on science and technology. Due to the absence of some data, the final data set used in this paper is the unbalanced panel data with 296 observations from 41 cities in Jiangsu, Zhejiang, Shanghai and Anhui provinces from 2011 to 2019. The main data sources are the statistical almanacs of each city and the International Federation of Robotics (IFR) data set. Table 1 below shows the descriptive statistics of each variable.

Table 1. Descriptive statistics

Variables	N	Min	Max	Average
SO2 emission	296	1,384	240,100	39,220
robot_density	296	2.480	47.310	15.230
Energy	296	155,731	3.279e+07	2.527e+06
Coal_per	296	29.780	94.720	80.200
GDP_per	296	10,090	180,044	63,541
GDP	296	3.725e+06	3.816e+08	3.543e+07
GDP_growth	296	-1.740	16	9.338
Industry_index	296	0.313	2.695	0.905
Rd	296	8,315	3.895e+06	212,262
Urban	296	31.300	89.300	59.010
Capital	296	9.619e+06	4.232e+08	7.515e+07
Green	296	22.260	77.780	41.800

3.2. Research Design

Following the research idea of Wang et al. (2022), the traditional econometric regression model is first used to test the relationship between robot application and SO_2 emission. Fixed effect model is also used to control missing variables that do not change with time. The model is set as follows:

$$SO_2\ emission_{i,t} = \alpha_1 + \alpha_2 robot\ \_density_{i,t} + + \sum_i \alpha_i\ Controls_{i,t} + \lambda_i + \varepsilon_{i,t} \qquad (2)$$

Where, SO_2 emission$_{i,t}$represents the total SO_2 emission in city *i* and year *t*; robot _density$_{i,t}$ represents the installed density of robots in city *i* and year *t*. Controls$_{i,t}$ represents other control variables; λ_i represents fixed effect on cities; $\varepsilon_{i,t}$ represents the residual estimated by the model.

Then, the random forest model, a mainstream machine learning model, is used to study the nonlinear effect of robot application on SO_2 emission. The random forest model is selected for the following two reasons: First, the random forest model can draw the importance of features based on the algorithm of node splitting, which can more intuitively display the importance ratio of each variable. Second, compared with other decision tree-based models, the random forest model can avoid extracting data from the same original data, thus reducing the correlation of data and improving the prediction ability of the model.

ISSN: 2688-9323

4. Result Analysis

Table 2 below reports the regression results of fixed effect model of traditional econometric model. The results show that the application of robots at the city level can significantly reduce SO_2 emission. Then, based on the fact that the dependent variable is a continuous variable, the regression tree is selected as the basic learner and the minimum mean square error (MSE) is taken as the optimization criterion to select the split node. The specific parameter values of the model are shown in Table 3 below. Table 3 also reports the goodness of fit (R_2)of the model. By comparing the goodness of fit of fixed effects in Table 2, it can be found that the random forest model has a better fitting effect on the data. This is mainly due to the data-driven way of machine learning, which can make full use of the variable information with the prediction accuracy as the criterion, so as to get closer to the real complex function form (Wang et al., 2022).

Table 2. Fixed effect model regression results

Variables	SO_2 emission
robot_density	-1,083.508***
	(159.582)
Energy	0.000
	(0.001)
Coal_per	96.463
	(100.654)
GDP_per	-0.034
	(0.107)
GDP	-0.001***
	(0.000)
GDP_growth	-586.508
	(434.758)
Industry_index	14,394.810**
	(6,903.302)
Rd	-0.026***
	(0.010)
Urban	290.982
	(278.892)
Capital	0.000
	(0.000)
Green	21.396
	(226.595)
Constant	58,117.630***
	(17,495.969)
City fixed effect	Yes
Observations	296
Number of cities	41

Notes: * (**, ***) significance at the 10% (5%, 1%) level. Standard errors in parentheses.

Further, in order to investigate the relative importance of robot installation density and factors traditionally considered as pollutants, the importance ranking diagram of characteristic variables as shown in Figure 1 is drawn according to the decreasing degree of residual sum of squares of model variables. It can be clearly seen that the installed density of robots has more influence on sulfur dioxide emissions than GDP. Energy consumption and industrial structure have always been considered as important factors affecting regional pollutant emissions. Thus, the model robustly indicates that there may be a complex nonlinear relationship between robot installation density and environmental performance, and robot installation density is also one of the important factors affecting pollutant emission.

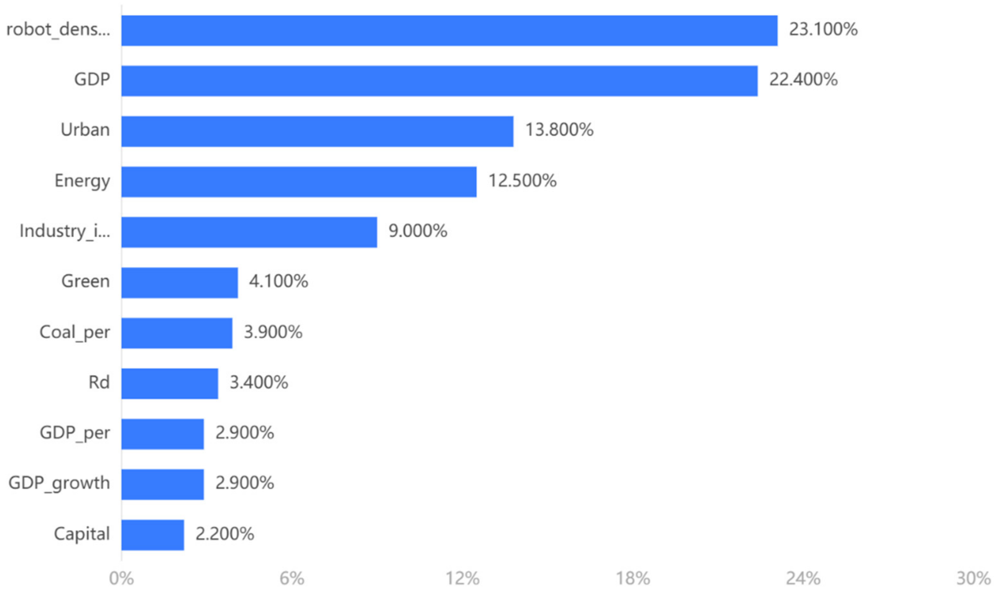

Figure 1. Feature importance

Table 3. The main parameter values and fitting results of random forest model

Parameter	Parameter value
Data segmentation	0.7
Evaluation criteria for node splitting	MSE
Minimum number of samples for internal node splitting	2
The minimum number of samples of leaf nodes	1
The minimum weight of the sample in the leaf node	0
The maximum depth of the tree	10
Maximum number of leaf nodes	50
Number of decision trees	100
There are put back samples	Yes
Goodness of fit R^2	0.931

5. Conclusion

Based on the city-level panel data from 2011 to 2019, this paper adopts the fixed effect model for reference and demonstrates that the application of robots can significantly reduce SO_2 emission. Moreover, from the perspective of machine learning model, it is further found that the importance of the application of robots on the environment is comparable to that of traditional factors. Therefore, this paper suggests that in the process of promoting intelligent development, we should pay attention to enterprises with relatively large intelligent production, establish an environmental supervision system for intelligent production, and enable enterprises to use robots reasonably and arrange production reasonably.

ISSN: 2688-9323

Acknowledgments

This work was supported by the National Undergraduate Innovation and Entrepreneurship Training Programs [grant numbers 202211287029Z]. The author thanks three valuable participants, Wentao Zhang, Yue Wang and Jiaxin Lin, from the Nanjing Audit University. And special thanks to our tutor Fushu Luan for her support and help in this project.

References

[1] Acemoglu D, P. Restrepo. The Race between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment[J]. American Economic Review, 2018,108(6),1488-1542.

[2] Acemoglu D, P. Restrepo. Robots and jobs: Evidence from US labor markets[J]. Journal of Political Economy, 2020, 128(6): 2188-2244.

[3] Ang S., et al. Can Fintech development curb agricultural nonpoint source pollution? [J].International journal of environmental research and public health, 2019, 16(22): 4340.

[4] Luan et al. Industrial robots and air environment: A moderated mediation model of population density and energy consumption[J]. Sustainable Production and Consumption, 2022, 30, 870-888.

[5] Deng Guangyao. Regional variation of carbon emissions from energy consumption and its influencing factors [J]. Statistics and Decision, 2023(06): 56-60

[6] Deng Zhongliang, Qu Xiaobo. The Development of Industrial Robots and the transformation and upgrading of Manufacturing Industry: A Survey based on the use of industrial robots in China [J]. Reform, 2021, No.330(08): 25-37.

[7] Huang Haiyan, Liu Ye, Peng Gang. The impact of industrial intelligentization on carbon emission: An empirical study based on China's industry segmentation [J]. Statistics and Decision, 2021, 37(17): 80-84.

[8] Huang Naijing, Yu Mingzhe. Research progress on the impact of machine learning on economic research [J]. Economic Trends, 2018, No.689(07): 115-129.

[9] Kong Gaowen, Liu Shasha, Kong Dongmin. Robots and employment: An exploratory analysis based on industry and regional heterogeneity [J]. China Industrial Economics, 2020, No 389(08): 80-98.

[10] Ma Luoru, Zhang Yi. Study on the Influence of cross-border E-commerce Policy on urban sulfur dioxide emission [J]. Ecological Economy, 2022, 38(11): 184-189.

[11] Wang Fang, Wang Xuanyi, Chen Shuo. Machine learning in economics research: Review and Prospect [J]. Journal of Quantitative and Technical Economics, 2020, 37(04): 146-164.

[12] Wang Yanqiu, Tao Sijia. Study on the influence and spatial effect of industrial intellectualization on China's industrial carbon emission efficiency [J]. Journal of Technology Economics, 2023, 42(01): 130-140.

[13] Wang Yongqin, Dong Wen. How is the rise of robots affecting China's labor market? Evidence from manufacturing listed companies [J]. Economic Research Journal, 2020, 55(10): 159-175.

[14] Wang Yuanbin, Zhang Yao, Li Jiguang. Digital finance and carbon emission: Research based on micro data and machine learning model [J]. China Population, Resources and Environment, 2022, 32(06): 1-11.

[15] Wei Xihai, Zhang Peikang, Du Yuhong. How robots reshape urban labor market: from the perspective of migrant work tasks [J]. Economic Trends, 2020, No.716(10): 92-109.

[16] Yang Guang, Hou Yu. The use, technology upgrade and economic growth of industrial robots [J]. China Industrial Economics, 2020, No.391(10): 138-156.