Applied Energy Symposium 2021: Low carbon cities and urban energy systems
September 4-8, 2021, Matsue, Japan
Paper ID: 11

Historical and prospective trajectories of operational carbon in China's commercial buildings: an assessment via LASSO-GWO approach

Xiwang Xiang[1], Xin Ma[2], Zhili Ma[3], Minda Ma[4*]

1 School of Management Science and Real Estate, Chongqing University, Chongqing, 400045, PR China (xiwang2718@gmail.com)

2 School of Science, Southwest University of Science and Technology, Mianyang, 621010, PR China (cauchy7203@gmail.com)
3 School of Management Science and Real Estate, Chongqing University, Chongqing, 400045, PR China (mzlmx@cqu.edu.cn)
4 Department of Earth System Science, Tsinghua University, Beijing, 100084, PR China (Corresponding Author, maminda@tsinghua.edu.cn)

ABSTRACT

The rapid growth of energy consumption in commercial building operations hinders the pace of carbon emission reduction in China's building sector, thus bringing great challenges to the successful realization of low-carbon development in China. This study uses historical data on carbon emissions from China's commercial building operation to establish the STIRPAT model. The model parameters are estimated by LASSO regression, and the Grey Wolf Optimizer (GWO) is used to optimize the nonlinear coefficients of the LASSO regression model. The proposed model is used to evaluate historical carbon emission reduction levels and estimate the peak value of future carbon emissions in China. Findings show that: (1) The main driver forces of carbon dioxide emissions from the commercial building sector in China are population size, GDP per capita, and energy intensity of carbon emissions, and their elastic coefficients are 0.5097, 0.2870, and 0.2006, respectively. (2) The peak emissions of the commercial building sector are 1269.42 $MtCO_2$, and the peak year is estimated to be 2029. Overall, this study analyzes the historical emission reduction levels and prospective peaks of carbon emissions in China's commercial building sector from a new perspective. The research results can help governments and decision-makers formulate effective emission reduction policies and can also provide references for the low-carbon development of other countries and cities.

Keywords: carbon-dioxide mitigation, commercial building operations, STIRPAT model, LASSO regression, carbon emissions peak;

NONMENCLATURE

Abbreviations	
LASSO	Least Absolute Shrinkage and Selection Operator
GWO	Grey Wolf Optimizer
STIRPAT	Stochastic Impacts by Regression on Population, Affluence and Technology
$MtCO_2$	Million tons of carbon dioxide
Symbols	
C	Carbon emissions of commercial buildings
P	Population size
G	Gross Domestic Product (GDP)
Gs	Gross Domestic Product of service industry
F	Floor space of commercial buildings
E	Energy consumption of commercial buildings

1. INTRODUCTION

The building sector is one of the three main drivers of global energy consumption and carbon emissions [1]. According to the report of the 2019 United Nations Climate Conference: Carbon emissions released by the global building sector account for 40% of total emissions [2]. Carbon emissions in the building sector have been highly valued worldwide. For the present situation of China, the growing tertiary industry economy and the rising energy demand for commercial buildings have brought severe challenges to China's low-carbon development [3]. Nevertheless, an army of studies has suggested that the building sector is a critical breakthrough in mitigating future global climate change [4-6]. Therefore, to further advance the process of China's energy conservation and emission reduction, especially the carbon emissions caused by the commercial building sector. Efficiently and reasonably assessing historical emission reduction levels and predicting the roadmap for peaking carbon emissions in the future is exceedingly important and should not be delayed.

Currently, the stochastic impacts by regression on population, affluence, and technology (STIRPAT) model is a popular tool used to analyze the influencing factors of carbon emissions [7-8], which is essentially a multiple regression model. Due to the multicollinearity between influencing factors, ordinary least squares estimation cannot accurately estimate the result. Ridge regression is often regarded as an effective means to handle this issue, which sacrifices the accuracy of the model in exchange for a more reasonable result [9]. In other words, all the influencing factors considered can be estimated by ridge regression. In fact, researchers are more interested in figuring out the major contributors to carbon emissions, not merely for analyzing too many variables brings high control costs, but more importantly, the coupling relationship between influencing factors also makes it difficult for emission reduction policies to be effectively implemented.

On the other hand, statistical methods [10], artificial intelligence models [11], and scenario analysis [12-13] are often used to predict future carbon emissions behavior. Scenario analysis mainly simulates and predicts the pathway of carbon emissions by setting changes in variables. However, to the best of the authors' knowledge, the current scenario analysis results are mainly realized via the exponential form of the STIRPAT model (KAYA identity). Obviously, the contribution of different driving forces and the multicollinearity between them are ignored.

For that, this paper proposes an extended STIRPAT model for the carbon emissions of commercial buildings in China. To the best of the authors' knowledge, this study first utilizes the LASSO regression method to estimate the results of the proposed STIRPAT model. In addition, the Grey Wolf Optimizer (GWO) is used to determine the nonlinear parameters of the LASSO regression model. Specifically, this work is based on historical data to assess China's carbon emission reduction levels during the period 2003-2018. Thereafter, further used the model to predict the low-carbon path and peak scenarios of China's commercial building sector from 2019 to 2060.

The most important contribution of this work is to propose a parameter estimation method based on LASSO regression for the extended STIRPAT model, which not only eliminates the collinearity between variables, but also gives full play to the advantages of LASSO regression that is good at feature selection. In addition, the popular intelligent optimization algorithm GWO has also been creatively introduced to optimize the nonlinear parameters of LASSO regression. The optimized model can avoid over-fitting, and can also simulate historical carbon emissions more intelligently and accurately. Overall, this study established the extended STIRPAT model with LASSO-GWO from a new perspective. The new approach evaluates the historical emission reduction level by simulating the historical emission data of commercial buildings and predicts its future emission reduction path.

2. METHOD

2.1 STIRPAT model

The stochastic impacts by regression on population, affluence, and technology model, also named as STIRPAT model, was initially put forward by Dietz and Rosa in 1971 [14] based on the IPAT model to meet the needs of statistical testing, which is an import tool often used to reveal the impacts of human behavior on environmental. After 50 years of development, the STIRPAT model has been widely recognized and applied in many disciplines such as energy economics, environmental economics, and climate change economics by virtue of its unique advantages. In addition, more and more studies have shown that the STIRPAT model plays an important role in the study of factors affecting carbon emissions.

In generally, the classic STIRPAT model is written as follows:

$$I = aP^b A^c T^d e \tag{1}$$

where a represents the constant term; b, c and d represents the parameters to be estimated; e represents the stochastic error term. I represents the environmental pressure caused by influencing factors; P , A and T represent the total population, affluence, and technical level, respectively. In empirical analysis, Eq. 1 is often transformed into a linear logarithmic form, written as follows:

$$\ln(I) = \ln a + b\ln(P) + c\ln(A) + d\ln(T) + \ln e \tag{2}$$

The STIRPAT model can expanded by considering different influencing factors, in order to explain the impact of these driving factors on environmental changes. Recently, the research of Ma et al. [15] showed that the carbon dioxide emissions of China's commercial buildings are closely related to six factors such as per capita GDP and energy consumption intensity. In view of this, the extended STIRPAT model used in this paper to study the carbon dioxide of the commercial buildings in China is defined as follows:

$$\ln C = \alpha_p \ln P + \alpha_g \ln g + \alpha_s \ln s + \alpha_{EA} \ln EA + \alpha_{e_c} \ln e_c + \alpha_{K_c} \ln K_c \tag{3}$$

The specific explanation of the symbols mentioned in extend STIRPAT model is tabulated in Table 1.

Tab 1 Definition and description of variables in the model

Symbol	Definition	Description
g	$g = \frac{G}{P}$	GDP per capita
s	$s = \frac{G_s}{G}$	Industrial structure
EA	$EA = \frac{F_c}{G_s}$	Industrial efficiency
e_c	$e_c = \frac{E_c}{F_c}$	Energy intensity
K_c	$K_c = \frac{C_c}{E_c}$	Total emission factor

2.2 LASSO regression with Grey Wolf Optimizer (GWO)

In most of the past studies on the analysis of factors affecting carbon emissions, ridge regression has been regarded as a mainstream tool for eliminating multicollinearity. The LASSO [16] is a shrinkage method like ridge, with subtle but important differences, and it can be defined as follows:

$$\min_{\beta} \left\{ \| Y - X\beta \|_2^2 + \alpha \| \beta \|_1 \right\} \tag{4}$$

Combined with the extended STIRPAT model proposed in this paper, Y can be written as $\ln C$, and X can be written as $(\ln P, \ln g, \ln s, \ln EA, \ln e_c, \ln K_c)$, β can be written as $\left(\alpha_P, \alpha_g, \alpha_s, \alpha_{EA}, \alpha_{e_c}, \alpha_{K_c}\right)^{\mathrm{T}}$. α is an adjustable parameter, which is often determined by cross-validation. Therefore, an optimization problem for non-linear parameter α can be formatted mathematically as follows:

$$\min_{\alpha} J(\alpha) = \frac{1}{n} \sum_{k=1}^{n} \left| \frac{\ln C(k) - \ln C(k)}{\ln C(k)} \right| \times 100\% \tag{5}$$

$$st. \begin{cases} \ln C = \alpha_p \ln P + \alpha_g \ln g + \alpha_s \ln s + \alpha_{EA} \ln EA + \alpha_{e_e} \ln e_c + \alpha_{K_c} \ln K_c \\ \left(\alpha_p, \alpha_g, \alpha_s, \alpha_{EA}, \alpha_{e_e}, \alpha_{K_c}\right)^{\mathrm{T}} = \theta^{\mathrm{T}} \\ \theta^{\mathrm{T}} = \min_{\beta \in \mathbb{R}^v} \left\{ \frac{1}{N} \| y - X\beta \|_2^2 + \alpha \| \beta \|_1 \right\} \end{cases}$$

It is quite difficult to solve Eq. 5 with traditional mathematical methods. To handle this issue, the Grey Wolf Optimizer (GWO) [17] was employed to solve the optimal nonlinear parameter α, which is a popular heuristic swarm intelligence optimization algorithm with simple operation and efficient convergence ability.

2.3 Data source

In order to study the driving factors affecting carbon dioxide emissions in China's building sector, all the driving factors were calculated according to the time-series cross-section data of China from 2003 to 2018, and the calculation method was listed in Tab 1. The data on population size (P) and economy level (G_S) are accessed from the Statistical Yearbook of China, which is directly cited in this paper. Data on the gross floor space (F_c), energy consumption (E_c) and carbon emissions (C) of commercial building sector were collected from the CBEED (Researchgate.net/project/China-Building-Energy-and-Emission-Database-CBEED).

3. RESULTS AND DISCUSSION

3.1 Multicollinearity test

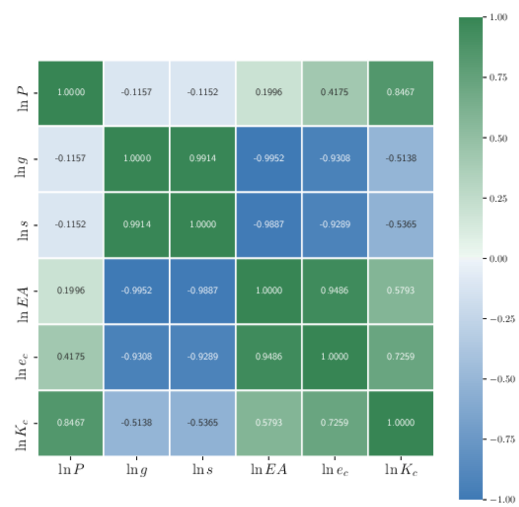

Fig 1 Multicollinearity test results of driver forces.

Multicollinearity refers to the highly correlated linear relationship between two or more variables in the regression model, which directly leads to the increase of the standard deviation of the parameter estimation and reduces the accuracy of the parameter estimation. It also means that any slight disturbance in the data can result in a significant change in the estimated results. In this subsection, the Pearson correlation coefficient is applied to examine the correlation of driver forces, and the test results are illustrated in Fig 1. Obviously, the correlation coefficient of most driver forces is greater than 0.75, which shows that there is serious multicollinearity between driver forces. In other words, the regression coefficient of the regression model has low reliability and cannot effectively explain the influencing factors of carbon emissions in the commercial building sector.

3.2 Major Driver Forces

In this paper, the historical emissions data from 2003 to 2018 is used to construct an extended STIRPAT model, and the LASSO regression with GWO (LASSO-GWO) is employed to estimate the regression coefficient of the model. Furthermore, the data from 2003 to 2016 was used as the training set to estimate the linear parameters of the STIRPAT model (coefficients of driver forces) according to Eq. 2, and the data from 2017 to 2018 were used as the test set to optimize the nonlinear parameters according to Eq. 3. The convergence curve of the GWO was drawn in Fig 2 a, and the actual data of carbon emissions and the model results were plotted in Fig2 b.

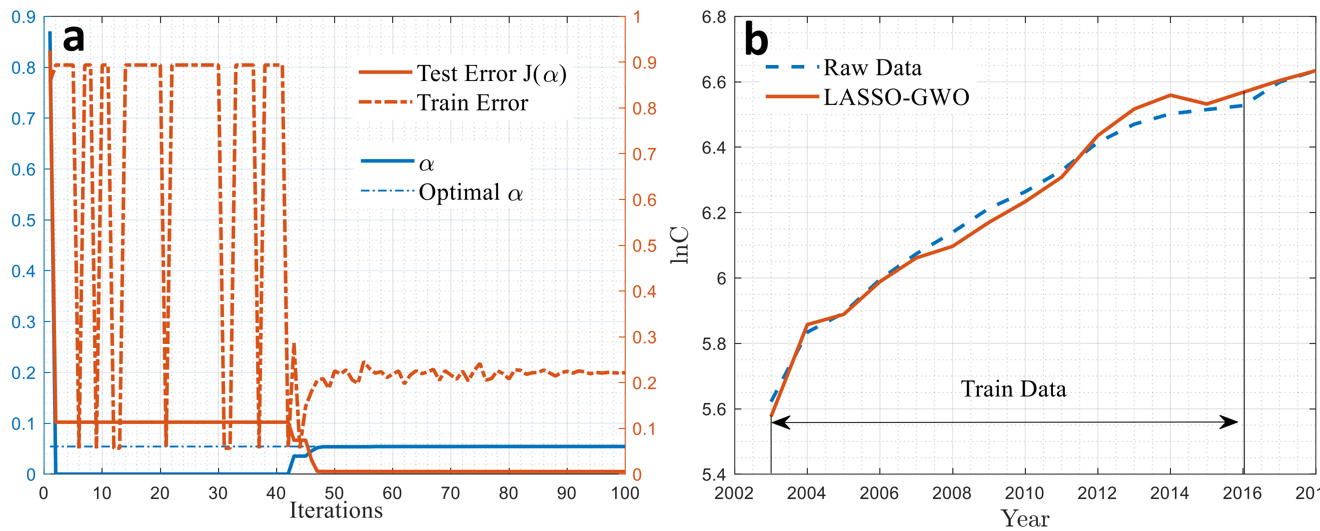

Fig 2 (a) Convergence curve by GWO for LASSO regression and (b) Comparison between the raw data and the calculated results of the proposed LASSO-GWO model

Fig. 2 a shows that GWO has converged to the global optimum after about 47 iterations, the optimal value of the nonlinear parameter is 0.5306, the training set and the test set eventually converge to 0.2261 and 0.0543, respectively. With the iteration of the algorithm, it can be clearly seen that the error of the training set shows a trend of first decreasing and then increasing, and finally stabilizes, and the error of the test set has been decreasing. The main reason for this phenomenon is that LASSO-GWO shows satisfactory generalization ability to avoid falling into over-fitting. This should not be surprising because the LASSO regression model with the L1 regularization term reduces the complexity of the model by constraining the linear parameters of the model, thereby improving the generalization ability of the model. In addition, we should also notice that the GWO fell into the local optimum during the early period of the algorithm iteration. At this moment, the nonlinear parameter is 0, and the LASSO regression model degenerates back to the ordinary least square multiplicative regression, and the linear parameters are estimated by the linear least-squares rule. As mentioned earlier, the estimation results of the parameters can be extremely unstable due to the interference of multicollinearity, which can also be observed from the convergence curve. On the contrary, when the parameter $\alpha = 0.5306$, the Lasso regression established in this work shows more stable results. Then, further solve Eq. 2 to obtain the extended STIRPAT model proposed in the paper as follows:

$$\ln C = 0.5097 \ln P + 0.2870 \ln g + 0.2006 \ln e_c \qquad (3)$$

Fig. 2 b shows the comparison between the results of the expanded STIRPAT model and the actual carbon emissions of the commercial building sector. The overall fit is very good. Therefore, the STIRPAT model proposed in this article can effectively explain carbon emissions from the commercial building sector in China. In the past period of time, the increase in carbon dioxide emissions from the commercial building sector is mainly due to changes in population size (P), national wealth level (g), and energy consumption intensity (e_c) in the public building sector. The size of the population has a significant promoting effect on carbon dioxide emissions. Its elasticity coefficient is 0.5097, which means that for every 1% increase in the population of China, carbon dioxide emissions will increase by 0.5097%. This is because the increase in the size of the population will directly lead to the squeezing of the space for human activities in commercial buildings, thereby increasing environmental pressure. On the other hand, the increase in carbon dioxide emissions from the commercial building sector is also importantly related to economic development. For every 1% increase in per capita GDP, carbon dioxide emissions will increase by 0.2806%. As the basic prerequisite for economic development, energy consumption is also the main force in the production of carbon dioxide emissions. And the results suggested for every 1% increase in per energy consumption intensity, carbon dioxide emissions will increase by 0.2006%.

3.3 Low-carbon pathway of carbon emission

The LASSO-GWO model proposed in this paper is also used to predict the peak roadmap of China's commercial building sector. Similarly, the data from 2003 to 2016 was applied to establish the model, and the data from 2017 to 2018 was applied to optimize the nonlinear parameter. Besides, this work further predicts the trajectory of carbon emissions from 2019 to 2060 based on the scenario settings in the literature [18].

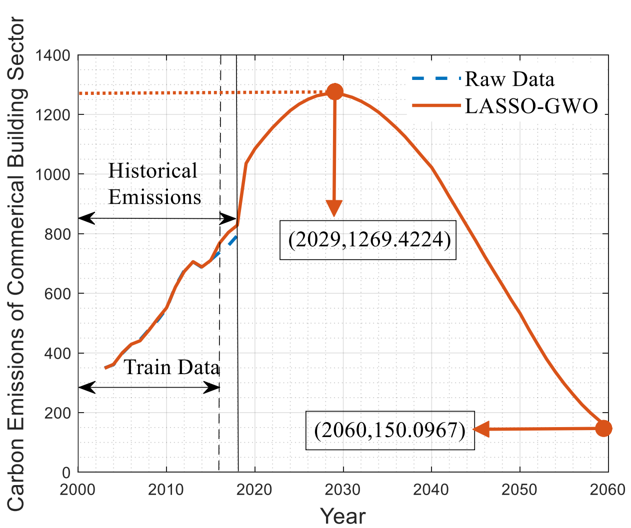

Fig 3 Historical carbon emissions and future peak road map of China's commercial building sector

The modelling and prediction results based on the LASSO model are illustrated in Fig 3. The modelling results coincide with historical emissions of China's commercial building sector, which checks the reliability of the model established. The prediction results show that the carbon emissions from the commercial building sector in China peak at 1269.4224 $MtCO_2$ in 2029, and then decrease at a rate of 36.1703 $MtCO_2$ per year to 150.0967 $MtCO_2$ by 2060.

4. DISCUSSION

In this paper, the extended STIRPAT was established, and the LASSO-GWO was first proposed to estimate the model results. The new extended STIRAPAT model with LASSO-GWO not only effectively eliminates multicollinearity between driver factors but also boosts the generalization ability of the model. Although this study has successfully performed a historical assessment of carbon emissions and predicted its peak path, several gaps that exist in this study can be addressed in the future. First, the existing parameters in the model can be further expanded. A reasonable parameter setting can more reliably dig out the major driver factors in historical carbon emissions. Moreover, the analysis of the pathway on low-carbon is mainly based on the scenario analysis results, which ignored the negative impact of the uncertainty in the carbon emission reduction process of the building sector, especially for the commercial buildings. For that, upcoming studies should spend considerable effort to combine the extended model with dynamic scenario analysis.

5. CONCLUSION

This study proposed an extended STIRPAT model, and the LASSO regression is first applied to estimate the regression coefficients. Also, the GWO is applied to determine the optimal nonlinear parameters of the LASSO. Then, the extended STIRPAT model with LASSO-GWO is utilized to make historical assessments of carbon emissions from 2001 to 2018 and to predict China's future low-carbon path from 2019 to 2060. Overall, the core findings of this research can be summarized as follows:

The major driver forces for carbon emission in commercial building operations are population size, GDP per capita, and energy intensity of carbon emissions, and their elastic coefficients are 0.5097, 0.2870, and 0.2006, respectively. The size of the population has a vital role in promoting carbon dioxide emissions in China's commercial building sector, which is closely related to the accelerated urbanization process in my country in recent years. In order to alleviate this pressure, it's necessary for China to control the size of the population and raise the people's low-carbon awareness. Also, Per capita GDP and energy consumption intensity are also crucial factors leading to the increase in carbon emissions. The adjustment of the industrial structure and the optimization of the energy structure are the leading solutions, such as vigorously develop low-carbon industries and fully develop and utilize renewable energy

Carbon emissions from the commercial building sector will peak in year 2029, with a peak of 1269.42 $MtCO_2$. Based on historical data and scenario assumptions, the LASSO-GWO model was established to predict the future development path of carbon emissions. The results show that China's commercial building sector will reach its peak in 2029, with a peak of 1269.4224 $MtCO_2$. After reaching the peak, the carbon emissions of the commercial building sector will decline at a rate of 36.1703 $MtCO_2$ per year until drop to 150.09 $MtCO_2$ by 2060.

ACKNOWLEDGEMENT

This study was supported by the China Postdoctoral Science Foundation (2020M680020), the Shuimu Tsinghua Scholar Program of Tsinghua University (2019SM139), the National Planning Office of Philosophy and Social Science Foundation of China (18BJL034), and the Fundamental Research Funds for the Central Universities of China (2020CDJSK03YJ07).

REFERENCE

[1] Ma M, Yan R, Du Y, et al. A methodology to assess China's building energy savings at the national level: an IPAT–LMDI model approach[J]. Journal of cleaner production, 2017, 143: 784-793.

[2] Ma M, Ma X, Cai W, et al. Low carbon roadmap of residential building sector in China: Historical mitigation and prospective peak[J]. Applied Energy, 2020, 273: 115247.

[3] Guillén-Lambea S, Rodríguez-Soria B, Marín J M. Comfort settings and energy demand for residential nZEB in warm climates[J]. Applied Energy, 2017, 202: 471-486.

[4] Ma M, Cai W. What drives the carbon mitigation in Chinese commercial building sector? Evidence from decomposing an extended Kaya identity[J]. Science of The Total Environment, 2018, 634: 884-899.

[5] Schäuble D, Marian A, Cremonese L. Conditions for a cost-effective application of smart thermostat systems in residential buildings[J]. Applied Energy, 2020, 262: 114526.

[6] McNeil M A, Feng W, du Can S R, et al. Energy efficiency outlook in China's urban buildings sector through 2030[J]. Energy Policy, 2016, 97: 532-539.

[7] Wang M, Che Y, Yang K, et al. A local-scale low-carbon plan based on the STIRPAT model and the scenario method: The case of Minhang District, Shanghai, China[J]. Energy Policy, 2011, 39(11): 6981-6990.

[8] Shahbaz M, Loganathan N, Muzaffar A T, et al. How urbanization affects CO2 emissions in Malaysia? The application of STIRPAT model[J]. Renewable and Sustainable Energy Reviews, 2016, 57: 83-93.

[9] Wang M, Che Y, Yang K, et al. A local-scale low-carbon plan based on the STIRPAT model and the scenario method: The case of Minhang District, Shanghai, China[J]. Energy Policy, 2011, 39(11): 6981-6990.

[10] Shamsuzzaman M, Shamsuzzoha A, Maged A, et al. Effective monitoring of carbon emissions from industrial sector using statistical process control[J]. Applied Energy, 2021, 300: 117352.

[11] Kumar S, Muhuri P K. A novel GDP prediction technique based on transfer learning using CO2 emission dataset[J]. Applied Energy, 2019, 253: 113476.

[12] Fang K, Tang Y, Zhang Q, et al. Will China peak its energy-related carbon emissions by 2030? Lessons from 30 Chinese provinces[J]. Applied Energy, 2019, 255: 113852.

[13] Liu D, Xiao B. Can China achieve its carbon emission peaking? A scenario analysis based on STIRPAT and system dynamics model[J]. Ecological indicators, 2018, 93: 647-657.

[14] Ehrlich P R, Holdren J P. Impact of population growth[J]. Science, 1971, 171(3977): 1212-1217.

[15] Ma M, Ma X, Cai W, et al. Carbon-dioxide mitigation in the residential building sector: a household scale-based assessment[J]. Energy Conversion and Management, 2019, 198: 111915.

[16] He Y, Qin Y, Wang S, et al. Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network[J]. Applied energy, 2019, 233: 565-575.

[17] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. Advances in engineering software, 2014, 69: 46-61.

[18] Ma M, Cai W. Towards low carbon pathway of the commercial building sector for 2070: Lessons from China[C]//Appl Energy Symp. 2020: 10-17.