

Full length article

CEDUP: Using incremental learning modeling to explore Spatio-temporal carbon emission distribution and unearthed patterns at the municipal levelZhiqiang Wu ^{a,b,c}, Renlu Qiao ^{a,b,*}, Xiaochang Liu ^{b,*}, Shuo Gao ^d, Xiang Ao ^d, Zheng He ^b, Li Xia ^e^a Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, 1239 Siping Road, Shanghai, PR China^b College of Architecture and Urban Planning, Tongji University, 1239 Siping Road, Shanghai, PR China^c Department of Mathematics and Theories, Peng Cheng Laboratory, Shenzhen, PR China^d University of Oxford, 11a Mansfield Road, Oxford OX1 3SZ, UK^e School of Management, University of Science and Technology of China, Hefei, PR China

ARTICLE INFO

Keywords:

Carbon dioxide emissions
Spatio-temporal
Incremental learning ensemble model
Municipal
China

ABSTRACT

Carbon emissions reduction has become a world consensus. Cities have an essential role to play in addressing emission reductions. However, previous studies have estimated China's municipal-level carbon emissions based on provincial-level emission patterns, and such a top-down carbon emission accounting approach has led to biased results. Therefore, this study employed an incremental learning ensemble model and a Savitzky-Golay algorithm to measure carbon emission distribution and unearth patterns at the municipal level based on nighttime light (NTL) and regional development characteristics (GDP, population, patents, industry structure). The performance of the proposed method is substantially better than its counterparts in terms of municipal-level estimation (R-square boosted by 20.64%). This research shows significant difference in carbon emission mechanisms between provinces and cities and demonstrates that carbon emissions are time-continuous. It also shows that per-capita carbon emissions are peaking in many China's cities, except in some heavy industrial cities. Our approach provides accurate and dynamic monitoring of municipal emissions in China.

1. Introduction

With the progress of urbanization and industrialization on the global scale, human energy consumption remains at a high level. Mass fossil fuel consumption has led to the continuous rise in carbon emissions (Wang et al., 2016). Many studies demonstrate that this excessive carbon emission has caused tremendous costs and risks, hindering sustainable development (Gao et al., 2021; Wang and Liu, 2017; Wang et al., 2017). Therefore, the world has reached a consensus to control carbon emissions. China is taking an increasing responsibility in the global effort to address climate change (Gao et al., 2022; Shan et al., 2019). China has reached a staggering level of urbanization. Though cities cover less than 3% of the earth's surface, they generate 80% of global greenhouse gas emissions (Fang et al., 2015). Municipal-level carbon accounting can provide detailed information about a city's energy use status, environmental quality, and carbon emissions, helping governments to understand the trends and levels of carbon emissions in cities. Since cities are the grassroots units in China's carbon policy-making, it is especially critical to accurately and dynamically

monitor the level of carbon emissions at the municipal level to achieve the stated goals of China's carbon reduction standards and thereby alleviate global environmental pressure.

The existing literature has established understanding of carbon emission drivers and their influence mechanisms but mainly focuses on the national-level or provincial-level emissions, or on some certain cities in China. Currently, China's government only provides a comprehensive energy consumption yearbook at the provincial level, but that at the municipal level is exceptionally scarce. Creating reliable municipal-level carbon emission inventories is essential for making feasible emission mitigation strategies (Kennedy et al., 2010; Liu et al., 2012; Xi et al., 2011). To address this issue, several methods have been proposed to estimate municipal-level carbon emissions, which fall into two broad categories: namely, bottom-up and top-down methods. The bottom-up method mainly employs the emission point source data, such as cement plants, steel mills, power plants, and spatial data in multiple dimensions, establishing accurate carbon emission grid data (Gurney et al., 2009; Wang et al., 2014). However, this approach is highly limited by the quality of the point source data. A complete census of carbon

* Corresponding authors.

E-mail addresses: RenluQiao@tongji.edu.cn (R. Qiao), 1910222@tongji.edu.cn (X. Liu).

emission sources in China was conducted only once in 2007, after which only incomplete counts (less than 10%) were conducted in specific years (Cai et al., 2018). Therefore, this method only applies to the measurement for fewer years, making it difficult to monitor carbon emissions dynamically from year to year.

Top-down carbon emission estimation is another critical pathway that can efficiently calculate a city's year-by-year carbon emission trend (Chen et al., 2020). Researchers have estimated city-level carbon emissions results based on the provincial-level pattern of carbon emissions (Cazorla and Errandonea, 2014; Chuzhi et al., 2008; Ribeiro et al., 2019; Zhang et al., 2009). With the development of remote sensing technology, nighttime light (NTL) satellite images have become a popular tool for down-sampling from provincial-level to city-level carbon emissions (Lv et al., 2020; Su et al., 2014). Dai et al. (2017) reveals a significant positive correlation between stable NTL data and carbon emission statistics. Nevertheless, the NTL data mainly demonstrates citizens' socioeconomic activity and electricity use, not fossil fuel consumption and associated carbon emissions. Because electricity generation and use are usually in separate places, more significant nighttime lighting does not imply higher carbon emissions, leading to large errors in the calculation of municipal-level carbon emissions (Meng et al., 2014).

China is a vast country containing 333 prefecture-level cities. Since the initiation of economic reforms in 1978, China has witnessed rapid economic growth. However, behind this prosperity, the phenomenon of uneven development has emerged (Cheong and Wu, 2013). This could be due to endogenous and exogenous issues, such as resource endowment and policy regulation, which have contributed to uneven development in China. Given the significant differences in economic development among cities, the underlying carbon emission mechanisms may differ, especially at the provincial scale versus the municipal scale (Shi et al., 2019). The conventional regression method can be inadequate for estimating municipal-level carbon emissions since the carbon emissions comprise highly complicated influence mechanisms (Yang et al., 2020). Machine learning modeling methods can thus be more suitable for establishing the complex relevance of carbon emission (Qiao et al., 2022; Wu et al., 2022a, 2022b).

Overall, the current research on municipal-level carbon emission measurement reveals issues and shortcomings that are required to be addressed. In turn, we launched a project named CEDUP to communicate knowledge with China's government sector and support decision-making processes; CEDUP aims to explore carbon emission distribution and unearth patterns at the municipal level through incremental machine learning modelling that addresses the difference between provincial- and municipal-level emission patterns. CEDUP contains carbon emission patterns for 333 prefecture-level administrative units in China, making it the most comprehensive municipal-level carbon emission data set in China currently. More details on CEDUP can be found at (www.wupen.org).

2. Methods

2.1. Study area

China possesses the world's most significant carbon footprint. Two Chinese administrative levels, provincial and municipal, were selected as the scope of this study as they are the basic units of emission reduction policy-making and implementation in China. Our research include downtown, suburban, and rural areas. Due to missing data, Hong Kong, Macau, and Taiwan were clearly not included in the study. All China's cities and special administrative regions are shown in Fig. A1 and Table A3.

2.2. Estimation model of carbon emissions

As there are considerable differences in development characteristics

at the province and municipal levels in China (Shi et al., 2019), ensemble learning can be a suitable alternative approach to deal with this research question. Extreme gradient boosting (XGBoost) is a modification of the gradient boosting decision tree (GBDT). The algorithm is an ensemble learning model based on the idea of boosting and is trained by using a forward distribution algorithm for a greedy approach. Each decision tree is built by fitting the residuals of the previous decision tree to the target value (Chen and Guestrin, 2016). XGBoost algorithm improves the accuracy and generalization of the model result through second-order Taylor expansion and regularization. Consequently, the study adopted XGBoost to establish the basic carbon emission model.

2.3. Incremental learning of municipal-level carbon emissions

Due to the incomplete energy statistics at the municipal level in China, it is necessary to combine provincial- and municipal-level data to train the carbon emission estimation model together (Chen et al., 2020; Huang et al., 2018; Song et al., 2020; Yang et al., 2020). However, roughly fusing these two levels of the energy consumption data and the characteristics data or only using the province-level data will cause biased estimates of the final municipal-level carbon emission results. Incremental learning can continuously process a continuous flow of information, retaining, integrating, and optimizing old knowledge while absorbing new knowledge. Thus, this study uses an incremental learning algorithm to make the model conform to the municipal-level carbon emission patterns and, at the same time, closer to provincial-level carbon emission patterns. Because of the limited amount of provincial- and municipal-level data, this study chooses joint training (J.T.) and Fine-tuning algorithm to avoid catastrophic forgetting during incremental learning model training (Li, Hoiem, & intelligence, 2017). Specifically, the study uses provincial-level data to train a basic model. Afterward, several new trees are added to the current iterative tree model through a mixture of provincial- and municipal-level data, and the node weights of the base model remain unchanged. We name the model XGBoost-IL.

2.4. Variables of estimation carbon emission model

This study employed a top-down approach provided by IPCC (<https://www.ipcc.ch/>) to estimate the carbon emissions from the energy, which approximates carbon emissions as measured by energy-related statistics. Energy-related carbon emissions account for more than 85% of China's total carbon emissions (Yang et al., 2020). Since the energy consumption data include fuel consumption for local thermal power generation and heating, electricity and heat consumption are excluded from the calculation of final energy consumption to avoid double counting.

$$CE = \sum_i^n E_i \times C_i \times \alpha_i \quad (2-1)$$

CE is the regional carbon emissions; E_i is the energy consumption for energy type i ; C_i is the carbon emission coefficient for energy type i ; α_i is the low calorific value of energy for energy type i . Ultimately, the article calculates energy-related carbon emissions based on the types of fossil energy sources derived from the China Energy Statistical Yearbook, including coal, coke, crude oil, gasoline, kerosene, diesel, fuel oil, liquefied petroleum gas, and natural gas (Yang et al., 2020). The data for calculating energy carbon emissions are from the yearbook statistics of the China Statistics Bureau (<http://www.stats.gov.cn/english/>), with provincial-level energy consumption data (including Beijing, Tianjin, Shanghai, and Chongqing) coming from the China Energy Statistics Yearbook, and municipal-level energy consumption data coming from each municipal statistical bureau. The carbon emission coefficients used in the study are from the General rules for calculating of the comprehensive energy consumption GB/T2589–2020 (<https://openstd.samr.gov.cn>).

Table 1
Variables in the municipal-level carbon emission estimation model.

	Variable	Description
Independent Variables	NTL_PBA	Ground average NTL in the built-up area
	Patent_PC	Number of patents granted per capita
	Population_PBA	Ground average population in the built-up area
	Agriculture_PC	Agricultural product value per capita
	Industry_PC	Industry product value per capita
	Service_PC	Service product value per capita
	Provincial_Emission_PAA	Ground average carbon emissions in the administrative area of the province to which the region belongs
Dependent Variables	Year	Year of data
	Emission_PBA	Ground average carbon emissions in the built-up area

gov.cn/bzgk/gb/). The data span 2005 to 2020, establishing province and municipal levels of carbon emission output dataset for the XGBoost-IL model.

Most previous studies only focus on the impact of NTL on municipal-level carbon emissions in the modeling process without considering the development characteristics of cities, such as industry type and population. To obtain more accurate municipal-level carbon emission estimation, we introduced NTL data, agriculture data, industry data, service data, the number of patents, and population data in the estimation model to characterize the city's industrial structure, technology level, and city volume. The patent data is sourced from the China National Intellectual Property Administration (<https://english.cnipa.gov.cn/>),

while the rest of the data is sourced from the Bureau of Statistics of each municipal and the National Bureau of Statistics of China (<http://www.stats.gov.cn/english/>). The Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS) stable nighttime light data and the Suomi National Polar-orbiting Partnership-Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) nighttime light data are two NTL datasets widely relied upon by researchers. The NTL data of DMSP-OLS and NPP-VIIRS cross-calibrated with each other are used in this study (Chen et al., 2020) to quantify municipal-level carbon emissions in China over a long time series.

However, the area of administrative districts in China's provinces is much larger than that of municipal districts, so the characteristic data mentioned above have huge distribution differences between provinces and municipalities. In Fig. A2, these box diagrams of provincial- and municipal-level total indicators demonstrating the upper and lower quartiles of the provincial data cannot cover that of the municipal data. The median of the total provincial data is 27.94 (Patent) to 12.50 (Population) times higher than that of the total municipal data. As for the mean indicators, the overlap of distribution between the provincial level and municipal level is significantly higher than the mean data. Therefore, the mean data is more suitable for the municipal-level carbon emissions model. The reason is that provincial-level data can include the range of changes in municipal-level data to avoid the carbon emissions model incorrectly classifying municipal data as outliers and to solve the problem of model data overbounds.

The study used eight categories of variables as the input dataset to characterize municipal social development, technology innovation, industry structure, and resource endowment (Table 1). The ground average carbon emissions in the built-up area is treated as the model's output.

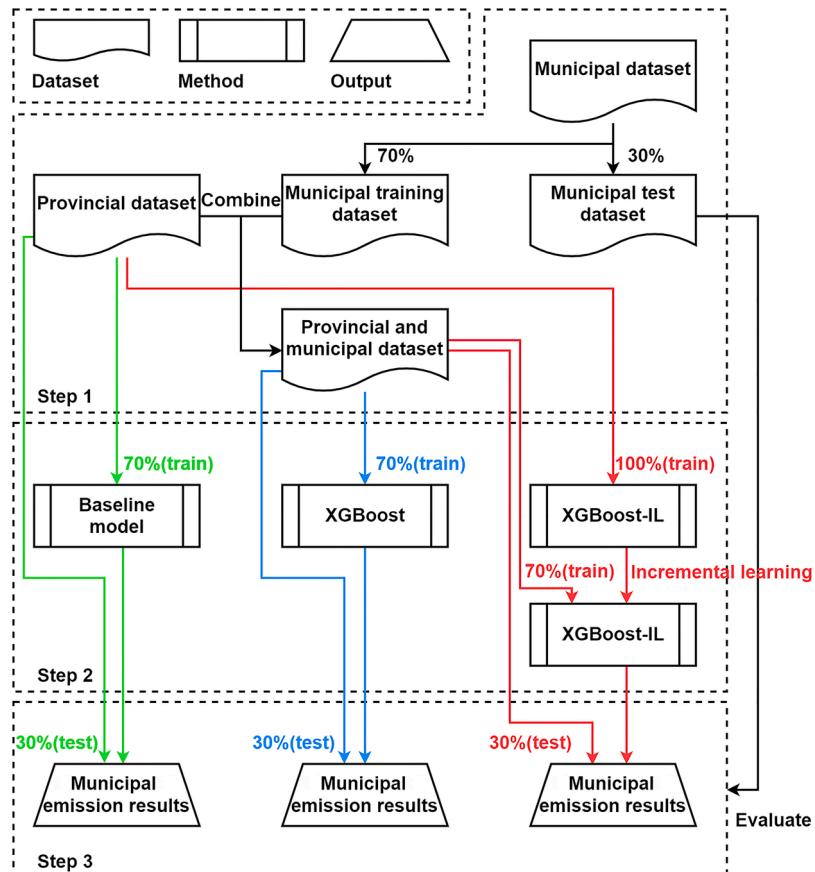


Fig. 1. Evaluation process of carbon emission estimation models.

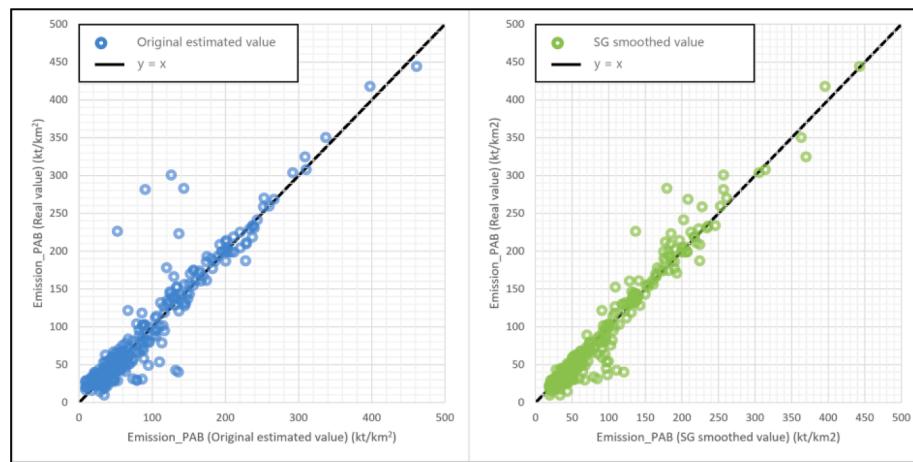


Fig. 2. Comparison of original estimation results and S.G. smoothed results in the municipal test set.

2.5. Evaluating the accuracy of estimation models

Machine learning modeling methods can fit the vast majority of data relationships depending on their complicated structures and nonlinear functions (Wu et al., 2022b). It is difficult to accurately evaluate the performance of the carbon emission estimation model at the municipal level by dividing the dataset straightforwardly. To solve the above problem, this study randomly splits the municipal data into ten copies, three of which were used as the test set for the model, and the remaining seven copies from the training set. The model trains to repeat the above process through cross-validation to cope with the problem. The results of a single test are too one-sided and the training data are insufficient. The specific process is shown in Fig. 1.

To verify the validity of the proposed method, the study examines kinds of carbon emission estimation models, including the traditional baseline mode, the XGBoost algorithm by provincial and municipal data, and the XGBoost-IL algorithm proposed in this study. The baseline model was the dominant approach in previous studies, which fitted carbon emission patterns from provincial data only and then input municipal data to estimate emissions. The carbon emission estimation model is essentially a regression issue, so this study uses the coefficient of determination (R-square) to evaluate the performance of the above algorithms.

2.6. Carbon emissions estimation smoothing

Municipal-level energy consumption data are not complete in China's official datasets, resulting in carbon emission estimates for some cities being calculated based on provincial-level patterns. This can ultimately lead to biases in the municipal carbon emission estimates. However, the annual changes in municipal-level carbon emissions are relatively smooth and do not have excessive fluctuations (Song et al., 2020). Therefore, the study uses Savitzky-Golay filter (S.G.) algorithm to decrease the interference of model anomaly. S.G. is a weighted averaging algorithm with a moving window. But instead of a simple constant window, its weighting coefficients are derived by least-squares fitting of a given higher-order polynomial within a sliding window (Schafer, 2011), which can be defined as follows:

$$Y_{j-SG} = \frac{\sum_{i=-m}^m f(Y_{j+i})}{2m+1} \quad (2-2)$$

Where Y_{j-SG} is the S.G. smoothing results; $2m+1$ is the width of the moving window; $f(Y_{j+i})$ is the function fitted by the least squares method.

2.7. Evaluation of the spatio-temporal characteristics of carbon emissions

The surrounding areas may synergistically influence carbon emissions in different cities. The Moran's I and Getis-Ord G* are commonly used to measure the degree of spatial aggregation, where Moran's I can be expressed as:

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n \omega_{ij}} \times \frac{\sum_{i=1}^n \sum_{j=1}^n \omega_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2-3)$$

In the calculation, y_i and y_j is the value of areas i and j ; \bar{y} is the average value of gross areas; ω_{ij} is the value of the spatial weight between areas i and j . The value of Moran's I is between plus and minus 1. Larger values indicate a more significant spatial correlation.

Getis-Ord G* can be expressed as:

$$G_i = \frac{\sum_{j=1}^n \omega_{ij} y_j - \bar{X} \sum_{j=1}^n \omega_{ij}}{S \sqrt{\left[n \sum_{j=1}^n \omega_{ij}^2 - \left(\sum_{j=1}^n \omega_{ij} \right)^2 \right] / (n-1)}} \quad (2-4)$$

$$\bar{X} = \frac{\sum_{j=1}^n y_j}{n} \quad (2-5)$$

$$S = \sqrt{\frac{\sum_{j=1}^n y_j^2}{n} - (\bar{X})^2} \quad (2-6)$$

In the calculation, the meaning of each variable is the same as that of Moran's I. The Getis-Ord G* is the z-score. For a statistically significant positive z-score, the higher the z-score, the tighter the clustering for high values. For a statistically significant negative z-score, the lower the z-score, the tighter the clustering for the lower values.

3. Experiments and results

3.1. Evaluation results of carbon emission estimation

The above three carbon emission estimation methods can fit the dataset well (see Table A1). At the same time, the R-square exceeds 94.11% on the test set, which means that the generalization of the three models does not appear to be over-fitted. Nonetheless, the model evaluation results on the municipal test set make a huge difference. The XGBoost-IL proposed in this study exhibits an overwhelming advantage, and the R-square is up to 90.63%, while the widely used baseline model's R-square on the municipal test set is barely 24.61%.

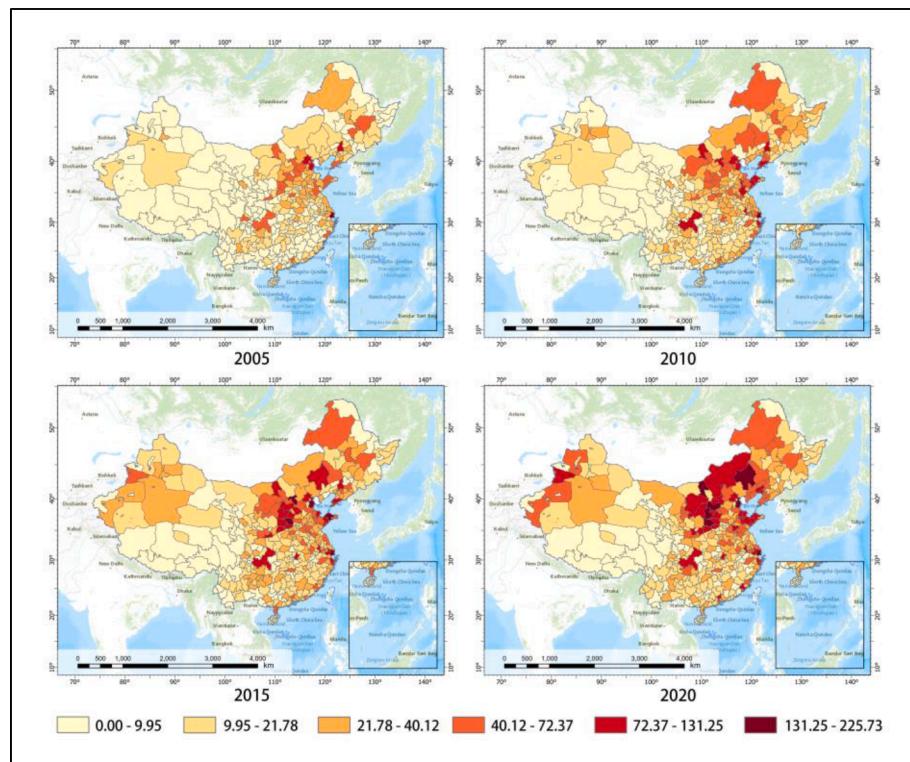


Fig. 3. Municipal-level carbon emission distribution in 2005, 2010, 2015, and 2020 (million tons).

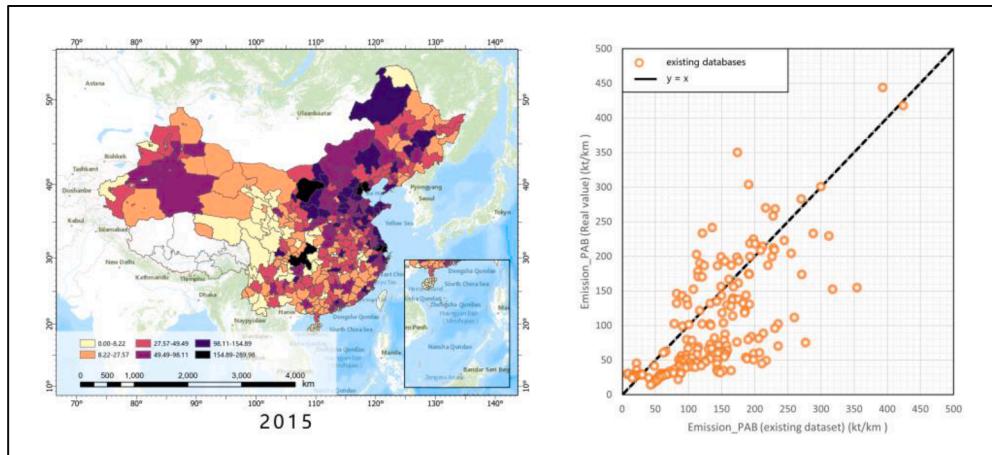


Fig. 4. Municipal-level carbon emission dataset in 2015 from Scientific data (million tons).

3.2. Times series smoothing of municipal-level carbon emission estimation results

Although XGBoost_IL possesses an extremely high accuracy rate on the municipal-level carbon emissions estimation, some of the noisy data can cause the model to generate a few outliers. Based on provincial- and municipal-level data, the study found that the trend in carbon emissions is continuous. Hence, we use the S.G. algorithm to smooth the municipal-level carbon emission estimation results, reducing the interference of anomalous estimates.

Fig. 2 demonstrates errors between the original estimated value and the smoothed value by the S.G. algorithm and the real value. When the point comes close to the black line, the model estimated result is terrific. Unfortunately, the blue scatter chart has few outliers, especially between 200 and 300. However, the green scatter plot shows the S.G.

smoothing process significantly corrects these errors in the blue scatter chart. The study uses error rates (ER) to describe the improvement of estimated results by S.G. smoothing, which can be calculated as follows:

$$ER = \frac{|\hat{y}_i - y_i|}{y_i} \quad (3-1)$$

Where \hat{y}_i is a model-estimated value or a smoothed value; y_i is a sample point real value. For the sake of verifying and quantifying the S.G. algorithm's effectiveness, the study calculates the median and the maximum value of ER in the original estimated and the S.G. smoothed results, respectively (see Table A2).

In the ER aspect, the smoothed results witnessed a substantial improvement. The ER's median of S.Q. smoothed results is 0.07, almost half that of the original estimation results. At the same time, the S.Q. smoothing process dropped the maximum value by 0.91, which is 31%

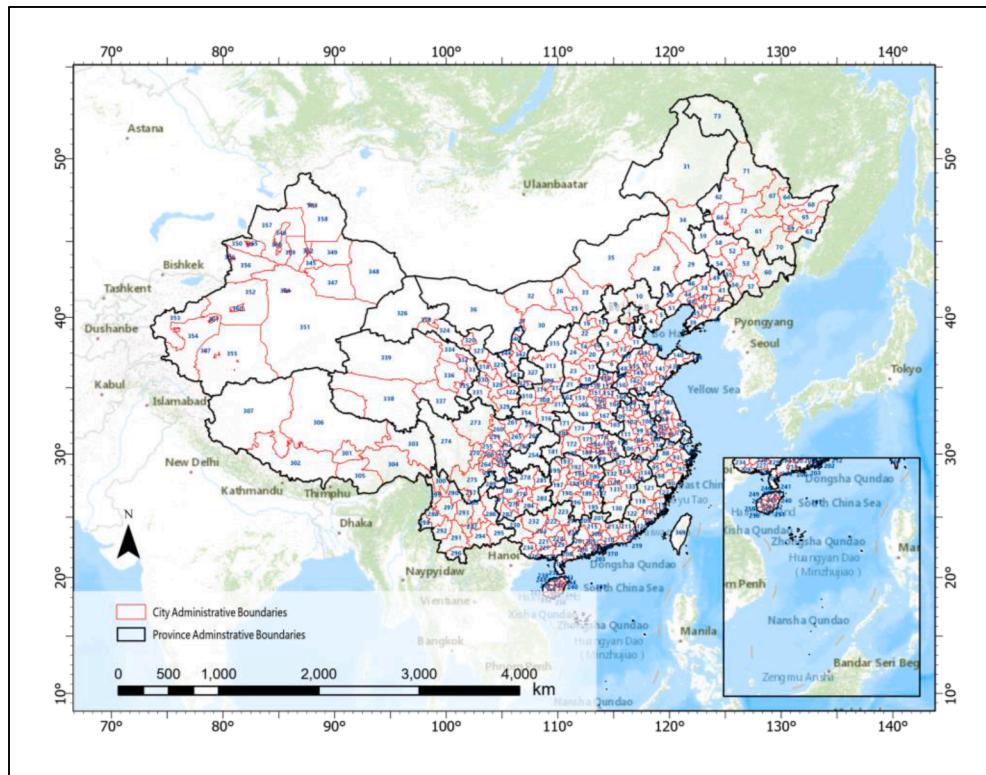


Fig. A1. China's cities and special administrative regions.

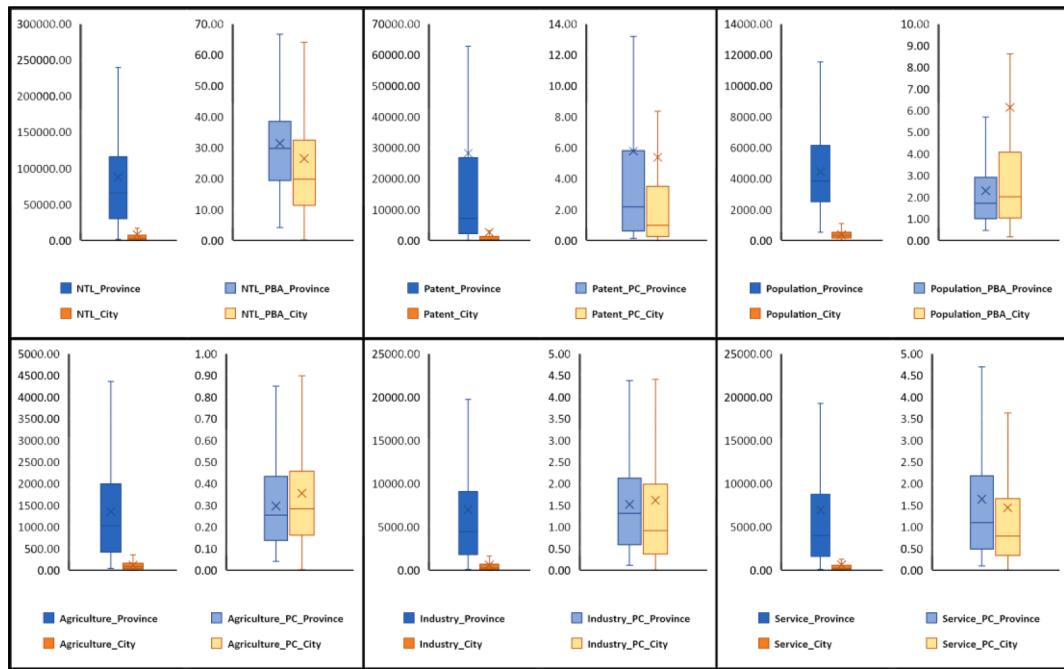


Fig. A2. Differences in data distribution between provincial- and municipal-level mean and total characteristics (P.C.: per capita; PBA: per built-up area).

of the maximum value of the original results' ER.

3.3. Municipal-level carbon emission distribution and evolution in China

Based on the above XGBoost-IL model and S.G. smoothing algorithm, we created CEDUP to obtain 333 Chinese cities' carbon emissions from 2005 to 2020 with enhanced accuracy (see Fig. 3).

Over 15 years, China's carbon emissions have risen year by year, and different cities showed considerable differences. In 2005, Shanghai's carbon emissions exceeded 100 million tons. Likewise, Beijing, Tianjin, and Shenyang are cities with large carbon footprints. By 2010, 6 cities had more than 100 million tons of carbon emissions. Tangshan had emerged as the second largest city in terms of carbon emissions, up to 126.03 million tons, after Shanghai. In 2015, the number of cities with

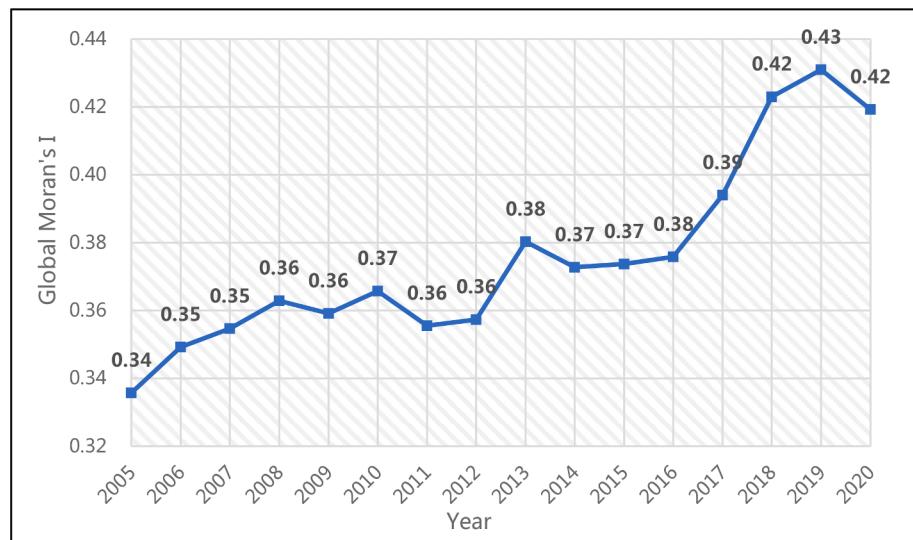


Fig. A3. Trend of global Moran's I from 2005 to 2020.

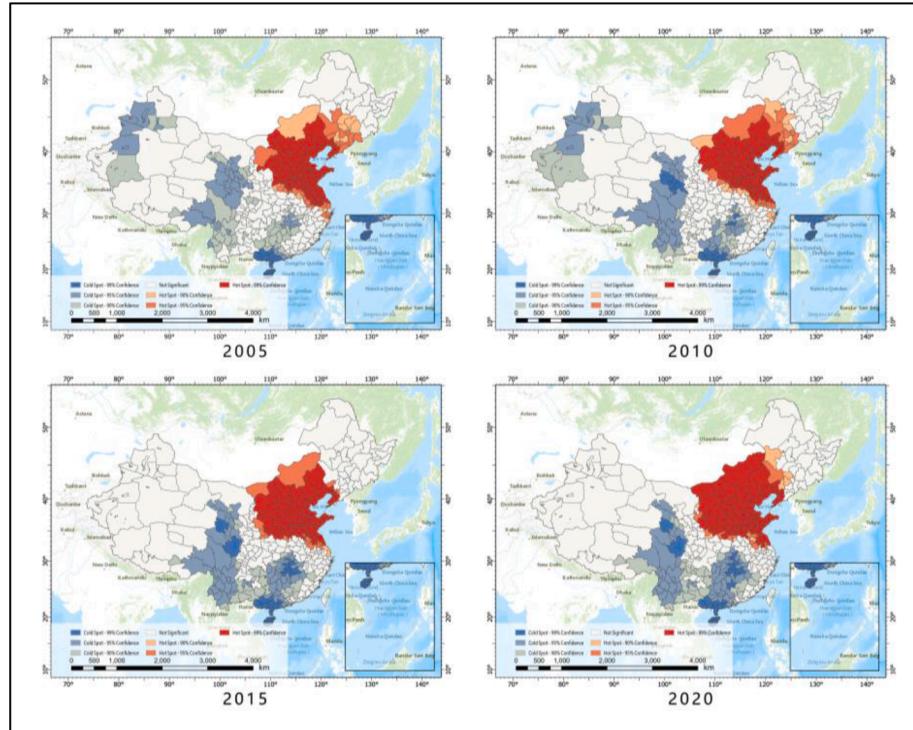


Fig. A4. Agglomeration degree of municipal-level carbon emissions in 2005, 2010, 2015, and 2020.

more than 100 million tons of carbon emissions increased significantly to 14, and 6 of them exceeded 200 million tons. In 2020, Chinese cities' carbon emissions continued to soar. Two cities' exceeded 200 million tons, with Tangshan surpassing Shanghai to become the highest carbon emission city, reaching 225.73 million tons. In addition, Qinghai Golog Tibetan Autonomous Prefecture is the lowest carbon emission city at 5.99 million tons.

3.4. Comparing CEDUP with existing datasets

We compared CEDUP's carbon emissions data with those of the World Bank and CHRED, which are influential carbon accounting organizations worldwide, and highly accurate bottom-up data on China's

carbon emissions, respectively (Cai et al., 2018). Emissions from the three databases for total carbon emissions were close to each other, with a difference of 3.14% between CEDUP and World Bank and 11.32% between CEDUP and CHRED.

Fig. 4 demonstrates a recent dataset recording China's latest county-level carbon emissions data based on top-down methods published in the *Scientific Data* (Chen et al., 2020). We aggregated the emission estimation results at the municipal level to compare our results with them. The distribution of carbon emissions data from *Scientific Data* is highly similar to CEDUP's. Therefore, this study uses the limited data from the statistical bureaus of Chinese cities to calculate the actual municipal-level emissions to assess the estimated our emission results' accuracy, as shown in the scatter plot in Fig. 4. From this, CEDUP can be

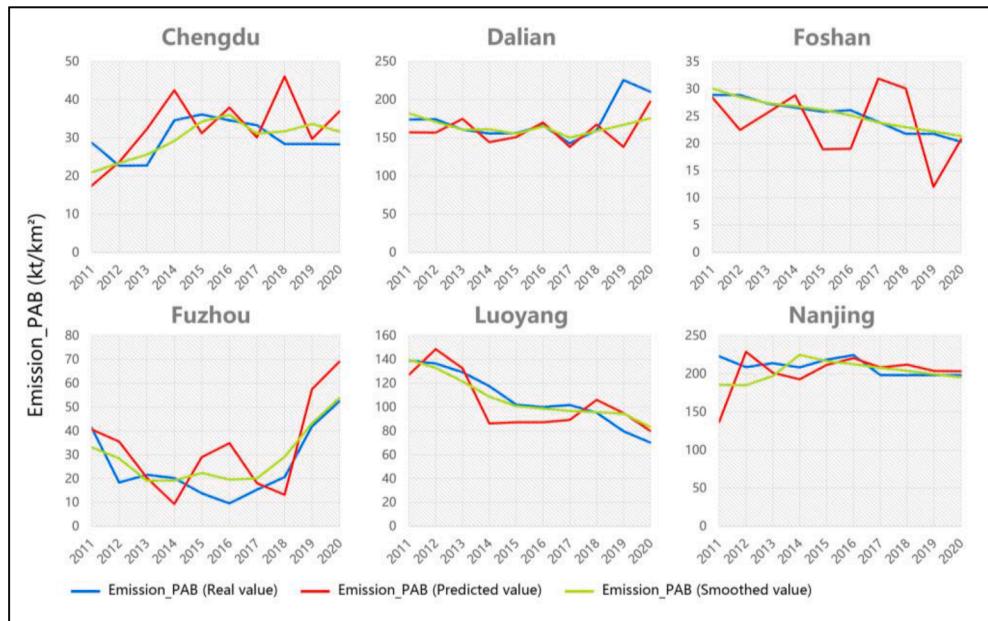


Fig. A5. Carbon emission trends in selected cities from 2011 to 2020.

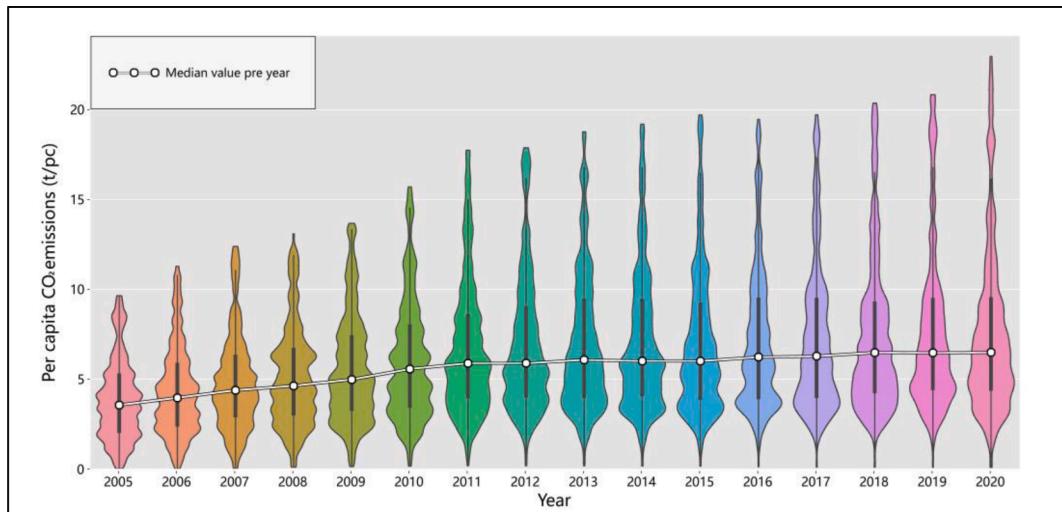


Fig. A6. Carbon emissions per capita in China through time.

Table A1
R-square of methods in three datasets (best results in cross-validation).

Method	R-square (Training set)	R-square (Test set)	R-square (Municipal test set)
Baseline model	99.99%	95.25%	24.61%
XGBoost	99.99%	89.18%	74.07%
XGBoost-IL	99.99%	94.11%	90.63%

Table A2
R-square and ER of results.

Results	R-square	Median	Max
Original estimated	90.63%	0.13	2.94
S.Q. smoothed	94.71%	0.07	2.03

more in line with the real Chinese carbon emission levels than the results of [Chen et al. \(2020\)](#) at the municipal level (see [Fig. 2](#)).

3.5. Spatial aggregation of municipal-level carbon emissions

As seen in [Fig. 3](#), the spatial distribution of Chinese municipal-level carbon emissions possesses certain aggregation characteristics. For quantifying and intuitively describing the carbon emissions' spatial cohesion degree, the study performs a spatial autocorrelation test by global Moran's I test through ArcGIS Pro 2.5.0 platform. Because of the difference in the size of administrative area, we constructed a spatial weight matrix based on Queen's Case method for calculating Moran's I index.

From the data in [Fig. A3](#), there has been a fluctuating upward trend in the global Moran's I index for 15 years, and all results were significant at the 99% confidence level. The index increased from 0.34 in 2005 to 0.43 in 2019, then slightly decreased to 0.42 in 2020, meaning that Chinese cities' carbon emission characteristics were becoming

Table A3

China's cities and special administrative regions.

Province	City/Region	No.
Beijing	Beijing	1
Tianjin	Tianjin	2
Hebei	Shijiazhuang	3
Hebei	Tangshan	4
Hebei	Qinhuangdao	5
Hebei	Handan	6
Hebei	Xingtai	7
Hebei	Baoding	8
Hebei	Zhangjiakou	9
Hebei	Chengde	10
Hebei	Cangzhou	11
Hebei	Langfang	12
Hebei	Hengshui	13
Shanxi	Taiyuan	14
Shanxi	Datong	15
Shanxi	Yangquan	16
Shanxi	Changzhi	17
Shanxi	Jincheng	18
Shanxi	Shuozhou	19
Shanxi	Jinzhou	20
Shanxi	Yuncheng	21
Shanxi	Xinzhou	22
Shanxi	Linfen	23
Shanxi	Lvliang	24
Inner Mongolia Autonomous Region	Huhhot	25
Inner Mongolia Autonomous Region	Baotou	26
Inner Mongolia Autonomous Region	Wuhai	27
Inner Mongolia Autonomous Region	Chifeng	28
Inner Mongolia Autonomous Region	Tongliao	29
Inner Mongolia Autonomous Region	Erdos	30
Inner Mongolia Autonomous Region	Hulunbeier	31
Inner Mongolia Autonomous Region	Bayannur	32
Inner Mongolia Autonomous Region	Ulanqab	33
Inner Mongolia Autonomous Region	Xing'anmeng	34
Inner Mongolia Autonomous Region	Xilinguolemeng	35
Inner Mongolia Autonomous Region	Alashanmeng	36
China and North Korea share	China and North Korea share	37
Liaoning	Shenyang	38
Liaoning	Dalian	39
Liaoning	Anshan	40
Liaoning	Fushun	41
Liaoning	Benxi	42
Liaoning	Dandong	43
Liaoning	Jinzhou	44
Liaoning	Yingkou	45
Liaoning	Fuxin	46
Liaoning	Liaoyang	47
Liaoning	Panjin	48
Liaoning	Tieling	49
Liaoning	Chaoyang	50
Liaoning	Huludao	51
Jilin	Changchun	52
Jilin	Jilin	53
Jilin	Siping	54
Jilin	Liaoyuan	55
Jilin	Tonghua	56
Jilin	Baishan	57
Jilin	Songyuan	58
Jilin	Baicheng	59
Jilin	Yanbian Korean Autonomous Prefecture	60
Heilongjiang	Harbin	61
Heilongjiang	Qiqihar	62

Table A3 (continued)

Province	City/Region	No.
Heilongjiang	Jixi	63
Heilongjiang	Hegang	64
Heilongjiang	Shuangyashan	65
Heilongjiang	Daqing	66
Heilongjiang	Yichun	67
Heilongjiang	Jiamusi	68
Heilongjiang	Qitaihe	69
Heilongjiang	Mudanjiang	70
Heilongjiang	Hehe	71
Heilongjiang	Suihua	72
Heilongjiang	Daxing'anling Area	73
Shanghai	Shanghai	74
Jiangsu	Nanjing	75
Jiangsu	Wuxi	76
Jiangsu	Xuzhou	77
Jiangsu	Changzhou	78
Jiangsu	Suzhou	79
Jiangsu	Nantong	80
Jiangsu	Lianyungang	81
Jiangsu	Huai'an	82
Jiangsu	Yancheng	83
Jiangsu	Yangzhou	84
Jiangsu	Zhenjiang	85
Jiangsu	Taizhou	86
Jiangsu	Suqian	87
Zhejiang	Hangzhou	88
Zhejiang	Ningbo	89
Zhejiang	Wenzhou	90
Zhejiang	Jiaxing	91
Zhejiang	Huzhou	92
Zhejiang	Shaoxing	93
Zhejiang	Jinhua	94
Zhejiang	Quzhou	95
Zhejiang	Zhoushan	96
Zhejiang	Taizhou	97
Anhui	Lishui	98
Anhui	Hefei	99
Anhui	Wuhu	100
Anhui	Bengbu	101
Anhui	Huainan	102
Anhui	Maanshan	103
Anhui	Huaibei	104
Anhui	Tongling	105
Anhui	Anqing	106
Anhui	Huangshan	107
Anhui	Chuzhou	108
Anhui	Fuyang	109
Anhui	Suzhou	110
Anhui	Liuan	111
Anhui	Bozhou	112
Anhui	Chizhou	113
Anhui	Xuancheng	114
Fujian	Fuzhou	115
Fujian	Xiamen	116
Fujian	Putian	117
Fujian	Sanming	118
Fujian	Quanzhou	119
Fujian	Zhangzhou	120
Fujian	Nanping	121
Fujian	Longyan	122
Fujian	Ningde	123
Jiangxi	Nanchang	124
Jiangxi	Jingdezhen	125
Jiangxi	Pingxiang	126
Jiangxi	Jiuzhang	127
Jiangxi	Xinyu	128
Jiangxi	Yingtan	129
Jiangxi	Ganzhou	130
Jiangxi	Ji'an	131
Jiangxi	Yichun	132
Jiangxi	Fuzhou	133
Jiangxi	Shangrao	134
Shandong	Jinan	135
Shandong	Qingdao	136
Shandong	Zibo	137
Shandong	Zaozhuang	138

(continued on next page)

Table A3 (continued)

Province	City/Region	No.
Shandong	Dongying	139
Shandong	Yantai	140
Shandong	Weifang	141
Shandong	Jining	142
Shandong	Taian	143
Shandong	Weihai	144
Shandong	Rizhao	145
Shandong	Linyi	146
Shandong	Dezhou	147
Shandong	Liaocheng	148
Shandong	Binzhou	149
Shandong	Heze	150
Henan	Zhengzhou	151
Henan	Kaifeng	152
Henan	Luoyang	153
Henan	Pingdingshan	154
Henan	Anyang	155
Henan	Hebi	156
Henan	Xinxiang	157
Henan	Jiaozuo	158
Henan	Puyang	159
Henan	Xuchang	160
Henan	Luohe	161
Henan	Sanmenxia	162
Henan	Nanyang	163
Henan	Shangqiu	164
Henan	Xinyang	165
Henan	Zhoukou	166
Henan	Zhumadian	167
Henan	Jiyuan	168
Hubei	Wuhan	169
Hubei	Huangshi	170
Hubei	Shiyan	171
Hubei	Yichang	172
Hubei	Xiangyang	173
Hubei	Ezhou	174
Hubei	Jingmen	175
Hubei	Xiaogan	176
Hubei	Jingzhou	177
Hubei	Huanggang	178
Hubei	Xianning	179
Hubei	Suizhou	180
Hubei	Enshi Tujia and Miao Autonomous Prefecture	181
Hubei	Xiantao	182
Hubei	Qianjiang	183
Hubei	Tianmen	184
Hubei	Shennongjia Forestry District	185
Hunan	Changsha	186
Hunan	Zhuzhou	187
Hunan	Xiangtan	188
Hunan	Hengyang	189
Hunan	Shaoyang	190
Hunan	Yueyang	191
Hunan	Changde	192
Hunan	Zhangjiajie	193
Hunan	Yiyang	194
Hunan	Chenzhou	195
Hunan	Yongzhou	196
Hunan	Huaihua	197
Hunan	Loudi	198
Hunan	Xiangxi Tujia and Miao Autonomous Prefecture	199
Guangdong	Guangzhou	200
Guangdong	Shaoguan	201
Guangdong	Shenzhen	202
Guangdong	Zhuhai	203
Guangdong	Shantou	204
Guangdong	Foshan	205
Guangdong	Jiangmen	206
Guangdong	Zhanjiang	207
Guangdong	Maoming	208
Guangdong	Zhaqing	209
Guangdong	Huizhou	210
Guangdong	Meizhou	211
Guangdong	Shanwei	212

Table A3 (continued)

Province	City/Region	No.
Guangdong	Heyuan	213
Guangdong	Yangjiang	214
Guangdong	Qingyuan	215
Guangdong	Dongguan	216
Guangdong	Zhongshan	217
Guangdong	Chaozhou	218
Guangdong	Jieyang	219
Guangdong	Yunfu	220
Guangxi Zhuang Autonomous Region	Nanning	221
Guangxi Zhuang Autonomous Region	Liuzhou	222
Guangxi Zhuang Autonomous Region	Guilin	223
Guangxi Zhuang Autonomous Region	Wuzhou	224
Guangxi Zhuang Autonomous Region	Beihai	225
Guangxi Zhuang Autonomous Region	Fangchenggang	226
Guangxi Zhuang Autonomous Region	Qinzhou	227
Guangxi Zhuang Autonomous Region	Guigang	228
Guangxi Zhuang Autonomous Region	Yulin	229
Guangxi Zhuang Autonomous Region	Baise	230
Guangxi Zhuang Autonomous Region	Hezhou	231
Guangxi Zhuang Autonomous Region	Hechi	232
Guangxi Zhuang Autonomous Region	Laibin	233
Guangxi Zhuang Autonomous Region	Chongzuo	234
Hainan	Haikou	235
Hainan	Sanya	236
Hainan	Sansha	237
Hainan	Danzhou	238
Hainan	Wuzishan	239
Hainan	Qionghai	240
Hainan	Wenchang	241
Hainan	Wanning	242
Hainan	Dongfang City	243
Hainan	Ding'an County	244
Hainan	Tunchang County	245
Hainan	Chengmai County	246
Hainan	Lingao County	247
Hainan	Baisha Li Autonomous County	248
Hainan	Changjiang Li Autonomous County	249
Hainan	Ledong Li Autonomous County	250
Hainan	Lingshui Li Autonomous County	251
Hainan	Baoting Li Autonomous County	252
Hainan	Qiongzhong Li and Miao Autonomous County	253
Chongqing	Chongqing	254
Sichuan	Chengdu	255
Sichuan	Zigong	256
Sichuan	Panzhihua	257
Sichuan	Luzhou	258
Sichuan	Deyang	259
Sichuan	Mianyang	260
Sichuan	Guangyuan	261
Sichuan	Suining	262
Sichuan	Neijiang	263
Sichuan	Leshan	264
Sichuan	Nanchong	265
Sichuan	Meishan	266
Sichuan	Yibin	267
Sichuan	Guang'an	268
Sichuan	Dazhou	269
Sichuan	Ya'an	270
Sichuan	Bazhong	271
Sichuan	Ziyang	272

(continued on next page)

Table A3 (continued)

Province	City/Region	No.
Sichuan	Aba Tibetan and Qiang Autonomous Prefecture	273
Sichuan	Ganzi Tibetan Autonomous Prefecture	274
Sichuan	Liangshan Yi Autonomous Prefecture	275
Guizhou	Guiyang	276
Guizhou	Liupanshui	277
Guizhou	Zunyi	278
Guizhou	Anshun	279
Guizhou	Bijie	280
Guizhou	Tongren	281
Guizhou	Qianxinan Buyi Miao Autonomous Prefecture	282
Guizhou	Qiandongnan Miao and Dong Autonomous Prefecture	283
Guizhou	Qiannan Buyi Miao Autonomous Prefecture	284
Yunnan	Kunming	285
Yunnan	Qujing	286
Yunnan	Yuxi	287
Yunnan	Baoshan	288
Yunnan	Zhaotong	289
Yunnan	Lijiang	290
Yunnan	Pu'er	291
Yunnan	Lincang	292
Yunnan	Chuxiong Yi Autonomous Prefecture	293
Yunnan	Honghe Hani and Yi Autonomous Prefecture	294
Yunnan	Wenshan Zhuang and Miao Autonomous Prefecture	295
Yunnan	Xishuangbanna Dai Autonomous Prefecture	296
Yunnan	Dali Bai Autonomous Prefecture	297
Yunnan	Dehong Dai Jingpozuzizhiqu	298
Yunnan	Nujiang Lisu Autonomous Prefecture	299
Yunnan	Diqing Tibetan Autonomous Prefecture	300
Tibet Autonomous Region	Lhasa	301
Xizang Autonomous Region	Shigatse	302
Tibet Autonomous Region	Changdu	303
Tibet Autonomous Region	Linzhi	304
Tibet Autonomous Region	Shannan	305
Tibet Autonomous Region	Nagqu	306
Tibet Autonomous Region	Ali	307
Shaanxi	Xi'an	308
Shaanxi	Tongchuan	309
Shaanxi	Baoji	310
Shaanxi	Xianyang	311
Shaanxi	Weinan	312
Shaanxi	Yan'an	313
Shaanxi	Hanzhong	314
Shaanxi	Yulin	315
Shaanxi	Ankang	316
Shaanxi	Shangluo	317
Gansu	Lanzhou	318
Gansu	Jiayuguan	319
Gansu	Jinchang	320
Gansu	Baiyin	321
Gansu	Tianshui	322
Gansu	Wuwei	323
Gansu	Zhangye	324
Gansu	Pingliang	325
Gansu	Jiuquan	326
Gansu	Qingyang	327
Gansu	Dingxi	328
Gansu	Longnan	329
Gansu	Linxia Hui Autonomous Prefecture	330
Gansu	Gannan Tibetan Autonomous Prefecture	331
Qinghai	Xining	332
Qinghai	Haidong	333
Qinghai	Haibei Tibetan Autonomous Prefecture	334
Qinghai	Huangnan Tibetan Autonomous Prefecture	335
Qinghai	Hainan Tibetan Autonomous Prefecture	336
Qinghai	Golog Tibetan Autonomous Prefecture	337
Qinghai	Yushu Tibetan Autonomous Prefecture	338
Qinghai	Haixi Mongol and Tibetan Autonomous Prefecture	339

Table A3 (continued)

Province	City/Region	No.
Ningxia Hui Autonomous Region	Yinchuan	340
Ningxia Hui Autonomous Region	Shizuishan	341
Ningxia Hui Autonomous Region	Wuzhong	342
Ningxia Hui Autonomous Region	Guyuan	343
Ningxia Hui Autonomous Region	Zhongwei	344
Xinjiang Uyghur Autonomous Region	Urumqi	345
Xinjiang Uyghur Autonomous Region	Keramayi	346
Xinjiang Uyghur Autonomous Region	Turpan	347
Xinjiang Uyghur Autonomous Region	Hami	348
Xinjiang Uyghur Autonomous Region	Changji Hui Autonomous Prefecture	349
Xinjiang Uyghur Autonomous Region	Boltara Mongol Autonomous Prefecture	350
Xinjiang Uyghur Autonomous Region	Bayingoleng Mongol Autonomous Prefecture	351
Xinjiang Uyghur Autonomous Region	Aksu Region	352
Xinjiang Uyghur Autonomous Region	Kizilsu Kirghiz Autonomous Prefecture	353
Xinjiang Uyghur Autonomous Region	Kashgar region	354
Xinjiang Uyghur Autonomous Region	Hotan Region	355
Xinjiang Uyghur Autonomous Region	Ili Kazakh Autonomous Prefecture	356
Xinjiang Uyghur Autonomous Region	Tacheng Region	357
Xinjiang Uyghur Autonomous Region	Altay Region	358
Xinjiang Uyghur Autonomous Region	Shihezi	359
Xinjiang Uyghur Autonomous Region	Aral	360
Xinjiang Uyghur Autonomous Region	Tumushuk	361
Xinjiang Uyghur Autonomous Region	Wujiaqu	362
Xinjiang Uyghur Autonomous Region	Beitun	363
Xinjiang Uyghur Autonomous Region	Tielmengguan	364
Xinjiang Uyghur Autonomous Region	Shuanghe	365
Xinjiang Uyghur Autonomous Region	Kekodara	366
Xinjiang Uyghur Autonomous Region	Kunyu	367
Xinjiang Uyghur Autonomous Region	Huyanghe City	368
Taiwan	Taiwan Province	369
Hong Kong Special Administrative Region	Hong Kong Special Administrative Region	370
Macao Special Administrative Region	Macao Special Administrative Region	371

increasingly convergent. This aggregation trend increased sharply, especially from 2016 to 2018.

3.6. Hot spot analysis of municipal-level carbon emissions

To pinpoint the spatial location of the aggregation of high and low values of municipal-level carbon emissions, the study adopts Gi* that combines Moran's I and general G statistics to reveal the spatial characteristics of municipal-level carbon emission distribution. The calculation results are shown in Fig. A4, where red represents the agglomeration degree of high-carbon emission cities and blue represents the agglomeration degree of low-carbon emission cities.

As shown in Fig. A4, high and low-carbon emission cities had different clustering trends during the 15 years. As for high-carbon

emission cities, the overall range of aggregation did not change significantly, even though the degree of its internal aggregation has changed. The agglomeration characteristic of most cities in Inner Mongolia is increasingly significant, especially in Bayannur and Xilingol league. In addition, the degree of high-value convergence in the Yangtze River Delta cluster and some cities in Liaoning had been considerably reduced.

On the other hand, low-value aggregation areas witnessed a crucial migration from 3 to 2 agglomerations over 15 years. In 2005 and 2010, the western region of Xinjiang had a more significant low carbon emission aggregation effect (less than 95% confidence). However, in the following ten years, this aggregation characteristic disappeared.

4. Discussion

4.1. Validity of CEDUP

Previous studies quantified municipal-level carbon emissions based on provincial-level carbon emission patterns through a top-down approach (Chen et al., 2020; Li et al., 2017; Wang et al., 2016). This is because municipal-level statistical data is incomplete. However, the patterns of carbon emissions at the upper level are not the same as those at the lower level. Therefore, the existing database cannot fully support an accurate municipal-level emission analysis for China.

CEDUP focuses on municipal-level carbon emission quantification, whereby a separate test set is divided for municipal-level data to verify the model's validity for municipal-level carbon emission estimation. As for the baseline model, its training set includes only provincial-level data, a common approach in prior studies. However, the model's performance in the test set is abysmal at 24.61%. Like the conclusion reached in Shi et al. (2019) study, the carbon emission patterns at different spatial scales are variable. A simple application of down-sampling modeling approach can cause colossal bias in the municipal-level quantification results.

CEDUP employs the XGBoost-IL approach to solve the above problem of unbalanced data in the provinces and cities. Incremental learning can integrate and optimize existing patterns while constructing new ones (Chefrour, 2019). The evaluation result presents that the estimation of the XGBoost-IL is accurate at the provincial and municipal levels of carbon emissions. The model's generalization is 90.63%, demonstrating CEDUP has high confidence in estimating municipal-level carbon emissions outside the sample cities. This significant improvement is induced by the incremental learning-based model preventing provincial-level patterns from being forgotten while learning about municipal-level patterns. Due to the immense lack of municipal-level data and city's uneven distribution spatially, it can be impossible to comprise carbon emission patterns for all cities in China. When the XGBoost-IL estimates a city's emission that is not part of the learned model of municipal-level emission pattern, the model can calculate an approximate municipal emission result using provincial-level pattern.

On the other hand, the generalization rate of the XGBoost trained simultaneously on the provincial- and municipal-level data is 74.07%, which is a bit away from XGBoost-IL. An explanation for this is that municipal-level datasets is relatively scarce compared to provincial-level datasets, so the XGBoost model shows difficulty in balancing the data at two spatial scales, causing a lack of model-fitting results at both scales. However, the XGBoost model has still acquired part of municipal-level carbon emission rules from the municipal-level data, so there is a significant improvement in the accuracy at the municipal level.

4.2. Temporal autocorrelation of municipal-level carbon emissions

In municipal-level carbon emission research, several studies have shown that carbon emissions at the municipal level possess a spatial autocorrelation (Chen et al., 2020; Shi et al., 2019; Wang et al., 2017; Yang et al., 2020). The analysis of the spatial distribution of municipal-level carbon emissions in this study also proves this (see

Fig. A4). Xiao et al. (2018) used geographically weighted regression models to estimate municipal-level carbon emissions. However, in most studies, although the year is added as an independent variable in the model to characterize the effect of time on carbon emissions, the temporal autocorrelation has not received much attention.

This study selects cities with complete carbon emissions data from 2011 to 2020 to draw the line graphs (see Fig. A5). The blue line shows the actual trend of municipal-level carbon emission intensity over the decade. The red line exhibits the predicted value by the carbon emission model. Because of the systematic bias of machine learning approaches, the carbon emission model will inevitably output a small number of outliers, causing the time-series data to show fluctuations. The green line presents the smoothed estimation results by the S.G. algorithm. The green line fits exceptionally close to the blue line, and the fluctuation of the red line is effectively eliminated, especially in Foshan and Fuzhou. According to these data (see Fig. 2), we can infer that the S.G. algorithm based on time-series window smoothing can successfully improve the accuracy of municipal-level carbon emission quantification results. In general, the long-time series carbon emission results based on CEDUP can improve the abundance of data and the accuracy of the results.

4.3. Policy implications

It is a widely documented fact that China's carbon emissions are upward (Chen et al., 2020; Huang et al., 2018; Shan et al., 2019; Wang and Liu, 2017), and our study confirms this result. Although the total carbon emissions entered a plateau period from 2013 to 2015, they further amplified over the next five years. At the municipal level, carbon emissions in part of China's cities had a dramatic rise, especially in less economically developed areas, such as Tumxuk (rise rate of 60.99% per year) and Haibei Tibetan Autonomous Prefecture (rise rate of 60.82% per year). These cities are usually underdeveloped in the early stages of industrialization and urbanization, adopting the traditional development model characterized by 'pollute first, clean up later' (Wang et al., 2019). On the other hand, for megacities with more than 20 million people, including Shanghai, Beijing, Chongqing, and Chengdu, the growth rates of carbon emissions are more stable. Among them, Chongqing is the fastest-growing city (4.67%), but it is slower than the Chinese average (5.91%). These large cities have urbanized rapidly over the past 20 years, and the associated industrial carbon emissions are approaching their peak and slowing down.

Chinese cities' carbon emissions per capita show a promising trend (see Fig. A6). We employed the violin chart to exhibit the development of Chinese cities' carbon emissions per capita. Fig. A6 demonstrates that the median value of per capita carbon emissions experienced a rapid increase from 3.56 tons per year capita to 5.89 tons per year per capita before 2011, and then the growth trend gradually slowed down until 2020, especially in the last three years, with the per capita carbon emissions leveling off. The result shows that the environmental protection policies implemented by the Chinese government, such as the suppression of fossil energy, a transformation of industrial structure, etc., have yielded significant environmental outcomes. In addition, the total population of China is about to peak between 2025 and 2030, according to the China Seventh National Census (<http://www.stats.gov.cn/english/>). Combining China's per capita carbon emission trends and population trends, we can optimistically predict that China can reach peak carbon emissions before 2030.

4.4. Limitations and future directions

This paper proposes an incremental learning method to estimate municipal-level carbon emissions by combining provincial and municipal development characteristics. However, the industrial structure data used in this paper are more macroscopic and lack more detailed industrial industry divisions, which may reduce the generalization power of the estimation results. In addition, this study finds that carbon emissions

are time-autocorrelated, so introducing a temporal modeling approach in future studies may further improve model accuracy.

5. Conclusions

As one of the world's largest carbon emitters, China's contribution to global climate governance has attracted great attention. In recent years, the Chinese government has taken a series of measures to address climate change and reduce carbon emissions. These measures have not only achieved remarkable results within China but have also had a positive impact on global climate governance. The lack of municipal-level energy statistics makes it difficult to calculate municipal-level carbon emissions. In response to this deficiency, this study provided a new path to quantify spatio-temporal changes in municipal-level carbon emissions between 2005 and 2020. We employ the XGBoost-IL proposed in this paper to establish the nonlinear relationship between area development characteristics and area carbon emissions based on the NTL, GDP, population, number of patents, and industrial output data. The S.G. algorithm then calibrates the model for temporal smoothing of the model quantization results. Finally, we use the reconstructed a municipal test set to effectively evaluate the actual accuracy of the quantification results at the municipal level. Experimental results show that the proposed CEDUP estimates carbon emissions with higher precision than the original machine learning method, and municipal carbon emission quantification accuracy improved by 20.64%. In this context, we find out differences in carbon emission mechanisms between provinces and cities, confirming that the simple adoptions of large-scale regional-level carbon emission patterns to calculate small-scale regional carbon emission can lead to severe deviations. Furthermore, the carbon emissions present time continuity, and the error rate of the quantification results can be significantly reduced by time series calibration.

In addition, we adopt Moran's I, Getis-Ord G*, and violin chart to analyze the spatio-temporal evolution of municipal-level carbon emissions. The analysis reveals that China's high carbon emitting cities are increasingly clustered and shifting to the northwest. The cities in western Xinjiang and central Inner Mongolia have significantly higher levels of carbon emissions. The second major finding was that Chinese cities' per-capita carbon emissions generally tend to peak, but those of some heavy industrial cities are still growing wildly.

Municipal-level carbon emission accounting in China has significant theoretical and practical implications. It can provide a foundation and support for the examination of underlying mechanisms of municipal-level energy consumption and carbon emission, implementation of carbon emission control approaches, and development of urban low-carbon strategies. Additionally, monitoring and evaluation of carbon emission reductions can help determine whether the city is on track to meet its carbon emission reduction goals, allowing for adjustments and improvements to be made in a timely manner.

CRediT authorship contribution statement

Zhiqiang Wu: Conceptualization, Supervision. **Renlu Qiao:** Conceptualization, Writing – original draft, Supervision. **Xiaochang Liu:** Writing – review & editing, Resources. **Shuo Gao:** Writing – review & editing. **Xiang Ao:** Data curation. **Zheng He:** Visualization. **Li Xia:** Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Appendix

See Fig. A1, Fig. A2, Fig. A3, Fig. A4, Fig. A5, Fig. A6 and Table A1, Table A2, Table A3.

References

- Cai, B., Liang, S., Zhou, J., Wang, J., Cao, L., Qu, S., Yang, Z., 2018. China high resolution emission database (CHRED) with point emission sources, gridded emission data, and supplementary socioeconomic data. *Resour. Conserv. Recycl.* 129, 232–239. <https://doi.org/10.1016/j.resconrec.2017.10.036>.
- Cazorla, C., Errandonea, D., 2014. Superionicity and polymorphism in calcium fluoride at high pressure. *Phys. Rev. Lett.* 113 (23), 235902 <https://doi.org/10.1103/PhysRevLett.113.235902>.
- Chefrouir, A., 2019. Incremental supervised learning: algorithms and applications in pattern recognition. *Evol. Intell.* 12 (2), 97–112. <https://doi.org/10.1007/s12065-019-00203-y>.
- Chen, J., Gao, M., Cheng, S., Hou, W., Song, M., Liu, X., Shan, Y., 2020a. County-level CO₂ emissions and sequestration in China during 1997–2017. *Sci. Data* 7 (1), 391. <https://doi.org/10.1038/s41597-020-00736-3>.
- Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 785–794.
- Chen, Z., Yu, B., Yang, C., Zhou, Y., Yao, S., Qian, X., Wu, J., 2020b. An extended time-series (2000–2018) of global NPP-VIIRS-like nighttime light data. <https://doi.org/10.7910/DVN/YGIVCD>. Retrieved from.
- Cheong, T.S., Wu, Y.J.J.o.A.E., 2013. Regional disparity, transitional dynamics and convergence in China. *J. Asian Econ.* 29, 1–14.
- Chuzhi, H., Xianjin, H., 2008. Characteristics of carbon emission in China and analysis on its cause. *Resour. Environ.* 18 (3), 38–42.
- Dai, Z., Hu, Y., Zhao, G.J.S., 2017. The suitability of different nighttime light data for GDP estimation at different spatial scales and regional levels. *Sustainability* 9 (2), 305.
- Fang, C., Wang, S., Li, G., 2015. Changing urban forms and carbon dioxide emissions in China: a case study of 30 provincial capital cities. *Appl. Energy* 158, 519–531. <https://doi.org/10.1016/j.apenergy.2015.08.095>.
- Gao, S., Lim, M.K., Qiao, R., Shen, C., Li, C., Xia, L., 2022. Identifying critical failure factors of green supply chain management in China's SMEs with a hierarchical cause–effect model. *Environ. Dev. Sustain.* 24 (4), 5641–5666. <https://doi.org/10.1007/s10668-021-01675-8>.
- Gao, S., Qiao, R., Lim, M.K., Li, C., Qu, Y., Xia, L., 2021. Integrating corporate website information into qualitative assessment for benchmarking green supply chain management practices for the chemical industry. *J. Clean. Prod.* 311. <https://doi.org/10.1016/j.jclepro.2021.127590>.
- Gurney, K.R., Mendoza, D.L., Zhou, Y., Fischer, M.L., Miller, C.C., Geethakumar, S., 2009. High resolution fossil fuel combustion CO₂ emission fluxes for the United States. *Environ. Sci. Technol.* 43(14), 5535–5541.
- Huang, J., Liu, Q., Cai, Y., Hao, Y., Lei, H., 2018. The effect of technological factors on China's carbon intensity: new evidence from a panel threshold model. *Energy Policy* 115, 32–42. <https://doi.org/10.1016/j.enpol.2017.12.008>.
- Kennedy, C., Steinberger, J., Gasson, B., Hansen, Y., Hillman, T., Havránek, M., Mendez, G.V., 2010. Methodology for inventorying greenhouse gas emissions from global cities. *Energy Policy* 38 (9), 4828–4837. <https://doi.org/10.1016/j.enpol.2009.08.050>.
- Li, Z., Hoiem, D., 2017. Learning without forgetting. *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (12), 2935–2947.
- Liu, Z., Liang, S., Geng, Y., Xue, B., Xi, F., Pan, Y., Fujita, T.J.E., 2012. Features, trajectories and driving forces for energy-related GHG emissions from Chinese mega cities: the case of Beijing, Tianjin, Shanghai and Chongqing. *Energy* 37 (1), 245–254.
- Lv, Q., Liu, H., Wang, J., Liu, H., Shang, Y., 2020. Multiscale analysis on spatiotemporal dynamics of energy consumption CO₂ emissions in China: utilizing the integrated of DMSP-OLS and NPP-VIIRS nighttime light datasets. *Sci. Total Environ.* 703, 134394. <https://doi.org/10.1016/j.scitotenv.2019.134394>.
- Meng, L., Crijns, W., Worrell, E., Huang, B.J.E., 2014. Estimating CO₂ emissions at urban scales by DMSP/OLS nighttime light imagery: methodological challenges and a case study for China. *Energy* 71 (15), 468–478.
- Qiao, R., Li, X., Gao, S., Ma, X., 2022. Improvement of thermal comfort for underground space: data enhancement using variational autoencoder. *Build. Environ.* 207. <https://doi.org/10.1016/j.buildenv.2021.108457>.
- Ribeiro, H.V., Rybski, D., Kropf, J.P., 2019. Effects of changing population or density on urban carbon dioxide emissions. *Nat. Commun.* 10 (1), 3204. <https://doi.org/10.1038/s41467-019-11184-y>.
- Schafer, R., 2011. What is a Savitzky-Golay filter? [lecture notes]. *IEEE Trans. Signal Process.* 28 (4), 111–117.
- Shan, Y., Liu, J., Liu, Z., Shao, S., Guan, D., 2019. An emissions-socioeconomic inventory of Chinese cities. *Sci. Data* 6 (1), 1–10.

- Shi, K., Yu, B., Zhou, Y., Chen, Y., Yang, C., Chen, Z., Wu, J., 2019. Spatiotemporal variations of CO₂ emissions and their impact factors in China: a comparative analysis between the provincial and prefectural levels. *Appl. Energy* 233–234, 170–181. <https://doi.org/10.1016/j.apenergy.2018.10.050>.
- Song, X., Chen, Y., Li, K., 2020. Analyzing spatiotemporal variation modes and industry-driving force research using VIIRS nighttime light in China. *Remote Sens.* 12 (17), 2785.
- Su, Y., Chen, X., Li, Y., Liao, J., Ye, Y., Zhang, H., Kuang, Y., 2014. China's 19-year city-level carbon emissions of energy consumptions, driving forces and regionalized mitigation guidelines. *Renew. Sustain. Energy Rev.* 35, 231–243. <https://doi.org/10.1016/j.rser.2014.04.015>.
- Wang, H., Lu, X., Deng, Y., Sun, Y., Nielsen, C.P., Liu, Y., McElroy, M.B., 2019. China's CO₂ peak before 2030 implied from characteristics and growth of cities. *Nat. Sustain.* 2 (8), 748–754. <https://doi.org/10.1038/s41893-019-0339-6>.
- Wang, J., Cai, B., Zhang, L., Cao, D., Liu, L., Zhou, Y., 2014. High resolution carbon dioxide emission gridded data for China derived from point sources. *Environ. Sci. Technol.* 48 (12), 7085–7093.
- Wang, S., Liu, X., 2017. China's city-level energy-related CO₂ emissions: spatiotemporal patterns and driving forces. *Appl. Energy* 200, 204–214. <https://doi.org/10.1016/j.apenergy.2017.05.085>.
- Wang, S., Liu, X., Zhou, C., Hu, J., Ou, J., 2017. Examining the impacts of socioeconomic factors, urban form, and transportation networks on CO₂ emissions in China's megacities. *Appl. Energy* 185, 189–200. <https://doi.org/10.1016/j.apenergy.2016.10.052>.
- Wang, S., Zhou, C., Li, G., Feng, K., 2016. CO₂, economic growth, and energy consumption in China's provinces: investigating the spatiotemporal and econometric characteristics of China's CO₂ emissions. *Ecol. Indic.* 69, 184–195. <https://doi.org/10.1016/j.ecolind.2016.04.022>.
- Wu, Z., Qiao, R., Zhao, S., Liu, X., Gao, S., Liu, Z., Jiang, Q., 2022a. Nonlinear forces in urban thermal environment using Bayesian optimization-based ensemble learning. *Sci. Total Environ.* 838. <https://doi.org/10.1016/j.scitotenv.2022.156348>.
- Wu, Z., Qiao, R., Zhao, S., Liu, X., Gao, S., Liu, Z., Jiang, Q., 2022b. Nonlinear forces in urban thermal environment using Bayesian optimization-based ensemble learning. *Sci. Total Environ.* 838, 156348 <https://doi.org/10.1016/j.scitotenv.2022.156348>.
- Xi, F., Geng, Y., Chen, X., Zhang, Y., Wang, X., Xue, B., Zhu, Q., 2011. Contributing to local policy making on GHG emission reduction through inventorying and attribution: a case study of Shenyang, China. *Energy Policy* 39 (10), 5999–6010.
- Xiao, H., Ma, Z., Mi, Z., Kelsey, J., Zheng, J., Yin, W., Yan, M., 2018. Spatio-temporal simulation of energy consumption in China's provinces based on satellite night-time light data. *Appl. Energy* 231, 1070–1078. <https://doi.org/10.1016/j.apenergy.2018.09.200>.
- Yang, D., Luan, W., Qiao, L., Pratama, M., 2020. Modeling and spatio-temporal analysis of city-level carbon emissions based on nighttime light satellite imagery. *Appl. Energy* 268, 114696. <https://doi.org/10.1016/j.apenergy.2020.114696>.
- Zhang, M., Mu, H., Ning, Y., 2009. Accounting for energy-related CO₂ emission in China, 1991–2006. *Energy Policy* 37 (3), 767–773. <https://doi.org/10.1016/j.enpol.2008.11.025>.