Predicting Carbon Emission at Fine Time Granularity Using Autoformer

Shuyi Wei[1,*], Xin You[2], Yaonan Jiang[3]

[1]College of Computer Science & Technology, Hangzhou Dianzi University, Hangzhou, 310018, China
[2]Zhuoyue Honors College, Hangzhou Dianzi University, Hangzhou, 310018, China
[3] College of Computer Science & Technology, Hangzhou Dianzi University, Hangzhou, 310018, China
eviewei0720@gmail.com, 21051606@hdu.edu.cn, 21052115@hdu.edu.cn

Abstract. Carbon dioxide (CO_2) is a major contributor to global warming. Accurate and high-resolution prediction of CO_2 emissions is critical to achieving carbon neutrality around the world. Previous methods used traditional statistical models or machine learning models, which can only predict the annual CO_2 emissions. With the development of deep learning models, such as GRU, they have been used for prediction. However, there is a lack of predictions for fine time granularity for CO_2 emissions. Thus, in this paper, we propose a carbon emission prediction model at fine time granularity. Autoformer is used to improve the prediction accuracy. A carbon emission dataset that contains three years of China's over 300 cities on a daily basis is used for prediction. The results show that compared with other methods, our model can achieve the highest prediction accuracy. Our model provides high-quality, fine-grained CO_2 emissions data to support global emission monitoring across various urban content.

Keywords: Carbon dioxide prediction, Autoformer, Fine time granularity.

1. Introduction

1.1. Background

Human activities, particularly the burning of fossil fuels for energy, deforestation, and industrial processes, release large amounts of CO_2 into the atmosphere [1]. As one of the major greenhouse gases, CO_2 contributes significantly to global warming, resulting in profound effects on the human environment and poses a substantial threat to the long-term development of the world economy [2,3]. Reducing greenhouse gas emissions and developing a low-carbon economy have become the consensus of all countries in the world [4].

The current CO_2 emission prediction methods can be divided into three types. They are statistical models, traditional machine learning methods, and deep learning models.

The statistical models are not suitable for predicting sequences of CO_2 emissions with complex nonlinear relationships and small sample adaptation.

Traditional machine learning methods require feature selection to establish CO_2 prediction models. Feature selection relies on a significant amount of socioeconomic and energy-related statistical data. The challenge arises when needed data are unavailable, and the feather selection is unpredictable, making it difficult to make accurate predictions [5-8].

With the development of deep learning, models based on Recurrent Neural Networks (RNNs) have been employed for predicting CO_2 emissions. Despite their effectiveness in time-series prediction, RNNs architectures face limitations in capturing long-range dependencies and exhibiting computational inefficiency due to their sequential execution mechanism [9-11].

However, previous studies usually forecasted the emissions on an annual basis based on the statistical yearbook data at the national or province level, which can only provide a rough overview of the changes in carbon emissions. Hence, detailed and spatially explicit estimates of emissions are urgently needed.

By incorporating multi-source data such as electricity generation, regional energy consumption statistics, gridded population density, night lights, urban form, and GDP data, researchers have constructed daily updated CO_2 emission datasets with higher spatial and temporal resolution [12-14]. These datasets offer the opportunity for a more fine-grained spatiotemporal prediction of CO_2 emissions.

1.2. Main Contributions

The contributions of this paper mainly have three aspects. First, we conduct a framework containing each step of our work. Among them, in contrast to previous studies, we use a high-quality city-level CO_2 emission dataset on a daily basis around China, which supports our goal of making fine-grained CO_2 emission prediction. Also, a new methodology is proposed to predict future CO_2 emission.

Second, we bring Autoformer into our model with added time features to make more accurate results. This new model has a better performance than others in terms of the value of evaluation metrics, and also the results shown in plots.

Third, the time series is added to selected deep learning models in order to yield an accurate result with temporal feature. Also, the performance of this method is demonstrated by comparing with other models about the metrics RMSE and MAE, including traditional ARIMA methods, two variants of RNN-based models (RNN and GRU), and two variants of Transformer based models (Transformer and Informer).

The remainder of this paper is organized as follows: Section 2 provides a literature review on CO_2 emission datasets and prediction methods. Section 3 introduces our work ranging from data collection to methodologies we used in detail. In Section 4, the result is accentuated by visualizing and comparing models' performance. The conclusions and future research are summarized in Section 5.

2. Literature Review

This section is divided into two parts: CO_2 emission datasets and prediction models. Each part introduces the main results up to now, which are the foundation of our research.

2.1. CO_2 Emission Datasets

In this part, abundant datasets about CO_2 emissions are searched. The following describes the datasets from coarse-grained to fine-grained with their incongruity with this topic.

The most famous environmental research institution is called World Resources Institute (WRI). In its database, many kinds of datasets about CO_2 emission were found. For instance, 'Global Emissions of CO_2 From Fossil Fuels: 1900-2004' dataset has the long-range data for many years but lack of more detailed geographical features for it only at national level [15].

Also, some data from international organization are discovered. EEA Greenhouse gas data reported by European Union recently shows the greenhouse gas emission including CO_2 in all nations of European Union. This dataset, however, does not have the detailed time statistics based on month [16].

Many datasets have the common flaw, especially in detailed time and space description. In order to look for fine-grained data in terms of geography, it is beneficial to visit national institution.

U.S. Energy Information Administration (EIA) presented 'CO_2 emission aggregates by fuel, state, and sector' dataset. Although it measured carbon emissions in different sectors, it presented a sketchy view of USA CO_2 emission because it was collected in terms of years and provinces [17].

The organize China Emission Accounts and Datasets (CEADs) sorted out the CO_2 emission among Chinese provinces between 1997 and 2021. As same as the USA data introduced above, this dataset also does not have detailed statistics in time and space [18].

Another website also provides CO_2 emission of China, however, it offers more fine-grained data in grid. Unfortunately, just few areas were selected such as Yangtze River Delta economic belt and Pearl River Delta economic belt. This is a hindrance for us to show the whole world's CO_2 emission as much as possible [19].

2.2. CO_2 Emission Prediction Models

In this part, we review the literature on CO_2 emission prediction methods. The methods are classified into three types. They are statistical models, traditional machine learning methods, and deep learning models.

Statistical Models. Statistical models, including regression analysis and time series modeling, are commonly used in CO_2 emission prediction.

Lotfalipour et al. predicted the CO_2 emissions using the grey model and the AutoRegressive Integrated Moving Average (ARIMA) model, and found that the forecasting performance of the grey model is better [20]. Kour et al. used ARIMA model to predict CO_2 emissions in South Africa [21]. Hosseini et al. used two regression models, multiple linear regression and multiple polynomial regression to predict Iran's CO_2 emissions in 2030 [22].

Zhao et al. and Yang et al. used the STIRPAT model to predict CO_2 emissions, and built different scenarios to investigate the trajectories and peak times of CO_2 emissions in the future under different policy impacts [23,24].

Ding et al. proposed a grey multivariable model by making three improvements, including modifying the calculation formula of the background values and optimizing the time response function based on the least-error law, to quantify future Chinese CO_2 emissions from fuel combustion [25]. They further proposed an information-based grey model that integrates the damping accumulative generating operator, data smoothing index, and particle swarm optimization, which can be used to predict the carbon emissions of China's 30 provinces [26].

In general, the regression analysis and ARIMA models are suitable for predicting linear relationships. The STIRPAT model can achieve higher accuracy if the input dataset satisfies preconditions. The grey model is suitable for small sample data for short time series. These statistical models are not suitable for predicting sequences of CO_2 emissions with complex nonlinear relationships.

Traditional Machine Learning Methods. The traditional machine learning methods used models such as Support Vector Regression (SVR), Random Forest (RF), gradient-boosted trees, and shallow neural networks to predict CO_2 emissions, which aims to better predict non-linear relationships within the time series data.

Li et al. predicted city-level CO_2 emissions in China based on a set of open-access data [5]. They first used two feature selection techniques, Recursive Feature Elimination, and Boruta, to extract important features for modeling CO_2 emissions. Then, they trained four machine learning models, namely SVR, RF, Gradient Boosting Machine (GBM), and extreme Gradient Boosting (XGBoost), to predict CO_2 emissions, finding that the XGBoost model achieved the highest estimation accuracy.

Song et al. employed the correlation heatmap and Boruta methods to extract critical variables for modeling [6]. Then, they used the stacking integration framework to integrate several single models, including LightGBM, CatBoost, and NGBoost, to improve the prediction accuracy and model generalization.

Khajavi et al. compared response surface methodology, RF, and SVR model to predict CO_2 emission produced from road transport sector in China [7]. They found that RF model with the Slime Mould Algorithm had the best prediction accuracy.

Traditional machine learning methods require feature selection to establish CO_2 prediction models. Feature selection relies on a significant amount of socioeconomic and energy-related statistical data. Challenge arises when needed data are unavailable and the feather selection is unpredictable, making it difficult to do accurate predictions.

Ren et al. first calculated the carbon emissions and carbon emission intensity of GuangDong province, China from 1995 to 2019 [8]. Then, they constructed a fastlearning network improved by chicken swarm optimization to predict carbon emissions from 2020 to 2060 and explored the peak and neutralization of carbon emissions in nine scenarios. Although this shallow network does not need to select features, its performance is not better than deep learning models.

Deep Learning Models. With the bloom of deep learning, deep neural networks have been proposed to leverage multiple hidden layers to automatically extract intricate data features. Among them, Recurrent Neural Network (RNN) is a type of neural network architecture designed to handle sequential data, which has achieved remarkable success in time series prediction [9]. LSTM (Long Short-Term Memory) and Gated Recurrent Unit (GRU) are two variants of the RNN that address the vanishing gradient problem and capture long-term dependencies in sequential data[10,11].

Huang et al. applied the LSTM to predict CO_2 emissions and then found LSTM had a better prediction performance than back propagation neural network and Gaussian process regression [28].

Zuo et al. proposed an LSTM-STIRPAT model to analyze the CO_2 emission peak of 30 Chinese provinces in China [29]. Initially, the LSTM model is used to predict the CO_2 emissions. According to the prediction results, the provinces are divided into some with peak value and others without peak value. Finally, the empirical analysis of the STIRPAT model was adopted to identify the main drivers of CO_2 emissions.

Wen et al. proposed a hybrid model that integrates ARIMA and LSTM to forecast future carbon emission trends in China [30]. This approach utilizes the strengths of both ARIMA and LSTM to capture linear and nonlinear relationships within the carbon emission data. The final prediction results are obtained by combining the inverse errors of ARIMA and LSTM with the goal of improving prediction accuracy.

3. Methodology

3.1. Framework

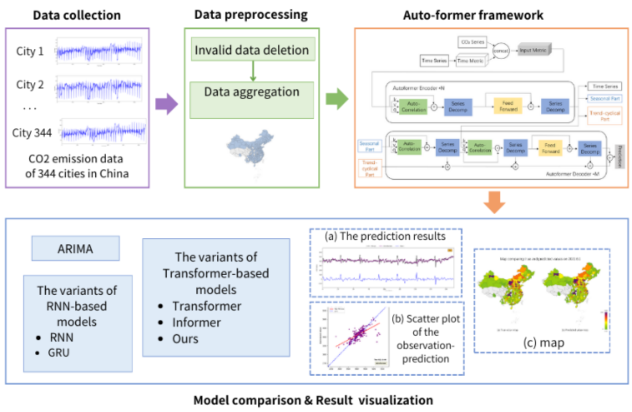

Fig 1. Model comparison & Result visualization

Fig. 1 shows the research framework of the CO2 prediction method. Initially, daily CO2 emission data at the city level in China are collected. Then, the dataset is pre-processed, including invalid data deletion and data aggregation. After that, the Autoformer model is adopted to predict the city-level CO2 emission data on a daily basis. Finally, we compare Autoformer with the ARIMA, the variants of the RNN-based models, and the other variants of the Transformer-based models to show the better performance of it. Moreover, the prediction results are visualized through several graphs to demonstrate an intuitive comparison result.

3.2. Data Collection and Preprocessing

Carbon Monitor is a near-real-time daily CO_2 emission dataset that shows the variations in CO_2 emissions from fossil fuel combustion and cement production at the national level [31]. Daily CO_2 emissions are estimated from a wide range of activity data, including electricity generation data, production data of energy-intensive industrial products, ground transportation mobility data, maritime and air transportation activity data, and fuel consumption data.

GRACED is the near real-time global gridded daily CO_2 emissions dataset from fossil fuel and cement production [32]. Gridded fossil emissions are calculated for different sectors based on the daily national CO_2 emissions from Carbon Monitor, the spatial patterns of the point source emission dataset, the emission database for global atmospheric research, and the spatiotemporal patterns of satellite NO_2 retrievals [31].

Based on the Carbon Monitor and GRACED datasets, Huo et al. constructed the is a near-real-time and city-level CO_2 emission dataset called Carbon Monitor Cities (CM-Cities) [31-33]. CM-Cities provides daily emission estimates from 2019 to 2021 for 1500 cities in 46 countries and disaggregates into five sectors: power generation, residential, industry, ground transportation, and aviation. When delineating city boundaries, CM-Cities establishes two types of spatial scopes: Functional Urban Areas (FUA) and Global Administrative Areas (GADM), which provide information to distinguish community-wide emissions and administrative emissions.

We use the GADM dataset of China from the CM-Cities datasets, which employs the GADM level-2 administrative areas for prefecture-level cities in China [33].

The data needs to be preprocessed. The cities that have empty values in sectors during the study period from 2019 to 2021 are ignored. The missing values are mainly in the power sector during the whole year of 2021. We obtain 300 cities with complete CO_2 emission data, and sum up the values of CO_2 emission in five sectors to obtain the total values of CO_2 emission in one day. Fig. 2 shows the 300 cities on the map.

Fig 2. 300 cities in China

3.3. Our Model

Our model is built on the foundation of Autoformer since its families have shown great modeling ability for long-range dependencies and interactions in sequential data and thus are appealing to time series modeling.

Vaswani et al. proposed the vanilla Transformer with the encoder-decoder structure [34]. Both the encoder and decoder are composed of multiple identical blocks. Each encoder block consists of a multi-head self-attention module and a position-wise feed-forward network while each decoder block inserts cross-attention models between the multi-head self-attention module and the position-wise feed-forward network.

Many variants of Transformer have been proposed to address special challenges in time series modeling, especially the Autoformer model.

Wu et al. studied the long-term forecasting problem of time series and proposed Autoformer, which introduces a decomposition transformer architecture and replaces the attention module with an Auto-Correlation mechanism [35].

Based on Autoformer, we added a time feature. Fig. 3 shows our model's structure. Generally, we concatenate the CO2 emission sequence with the time feature as a combined input metric. However, the time series needs to be processed into a metric. The detailed method is that, the time series is divided into three dimensions: year, month and date, each dimension is encoded into the range between -0.5 to 0.5, putting them together into a metric.

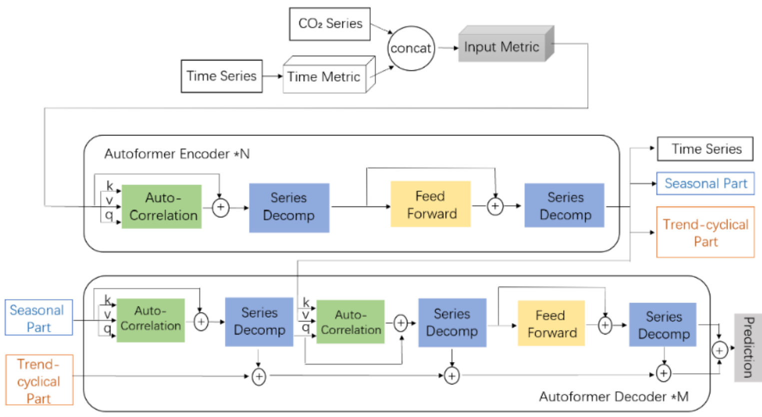

Fig 3. Model's structure

3.4. Evaluation Metrics

The evaluation metrics used in this study are MSE, MAE, and R^2, where n denotes the length of test sets and y_i and $\hat{y}_i$ denote the test set's real and predicted values, respectively. The lower the values of the first two measures, the better the fit.

a) RMSE (Root Mean Square Error): It reflects the mean error of prediction with outlier sensitivity. The calculation of RMSE is given in Eq.1:

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

b) MAE (Mean Absolute Error): It reflects the mean error of prediction with outlier insensitivity. The calculation of MAE is given in Eq.2:

$$MAE = \frac{\sum_{i=1}^{n} \left| y_i - \hat{y}_i \right|}{n} \tag{2}$$

4. Result

4.1. Experimental Details

We have compared Autoformer with other baselines in both traditional machine learning and deep learning fields, including ARIMA, GRU, RNN, Informer and Transformer [9,11,21,34,36]. For GRU, Informer, Transformer, and ours, we choose the same sequence length and prediction length, which is 1 step prediction for every 7 steps sequence. Also, for these models, we divide the datasets into three sections – training, validation and testing – the ratio between them is 7:1:2. While for RNN, we choose the same prediction length (1 step), but change the sequence to 6 steps, which shows a better performance. As for ARIMA and RNN, to meet the needs of the models, we only divide them into two sections – training and testing, and the ratio is 8:2.

4.2. Model Parameter Regulation

ARIMA. To determine the optimal parameters of the ARIMA model, the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) are used to observe the dataset. ACF and PACF can measure the correlation between a time series and its lag. Therefore, the characteristics of the dataset can be observed by analyzing the results of the ACF and PACF of the dataset. Fig 4 shows the ACF and PACF results obtained from the time series of a randomly selected city (Aksu) among all 300 cities. Through analyzing Fig 4, the number of autoregressive terms(p) and moving average terms(q) are determined between 1-5, and the differential order(r) is determined between 0-2. After multiple experiments with different parameters, the optimal parameters of (p, q, r) is (2, 1, 1).

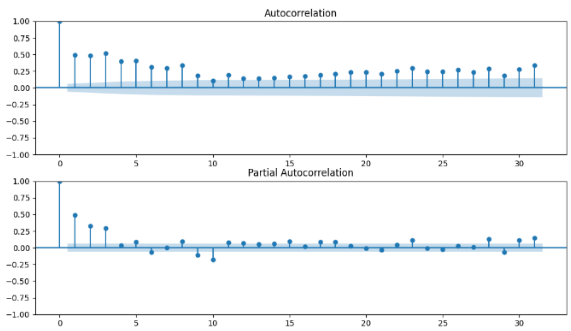

Fig 4. ACF and PACF results

GRU. During the parameter tuning process of GRU, we set the epoch number to 100, and the output results of models with different virous parameters have a common characteristic: the prediction error of the test set will increase significantly after the epoch number is larger than 40, as shown in Fig 5 and Fig 6. Therefore, 40 is the final epoch number we selected. The optimal parameters of GRU are batch_size of 32 and gru_unit of 64.

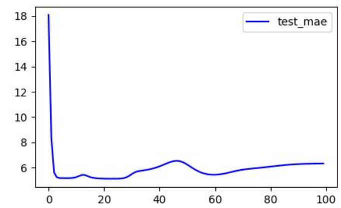

Fig 5. MAE in test set

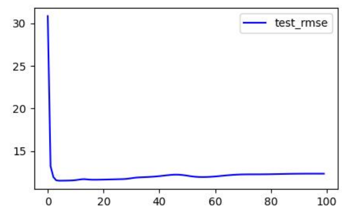

Fig 6. RMSE in test set

Transformer, Informer and Ours. Transformer, Informer and ours use the same model parameter regulation method, comparing the prediction results of models with different label_len. The final optimal label_len of Transformer and Informer is 2, and it is 3 of ours.

4.3. Result Comparison and Visualization

RMSE and MAE results of all models are reported in Table 1, and our model has a relatively accurate performance. To show the results more vividly, we chose Shanghai as a representative to do the visualization, and the visual result is shown in Fig 7, Fig 8, and Fig 9. Within the figures, the first line of the charts shows the prediction results and the real data, the second line of the charts shows the error value of the prediction.

From Fig 7, we can clearly see that ARIMA exhibits a significant error, showing an over-smoothing prediction and having no ability in seasonal and abrupt data prediction. While in Fig 8 and Fig 9 show that the other models can predict the trends, whereas ours performs the best in predicting the peak and valley values and has the best ability to predict the fluctuation data. It can be found by looking at the error line. Additionally, the models from GRU, RNN to Transformer, and Informer display an increasing error in long-term predicting, while our model barely shows. Our model has the smallest error value and the fitted value are quite close to real value at the end of the time line.

In order to best present the error distribution, we draw scatter charts to show the deviation value of the models in Fig 10. Comparing the scatter charts, it is easy to find that our model has the fewest error outliers, with R2 0.44 and RMSE 28.10, which means that it has the best reliability in prediction.

Besides, we create a comparative map (Fig. 11) of 300 cities to show the overall situation of our model's performance, finding that basically all the cities are well forecast.

Table 1. RMSE and MAE in different models

	RMSE	MAE
ARIMA	10.14261225	7.651988617
RNN	6.90614478	4.64482946
GRU	11.507702124	5.187694
Transformer	8.231267646	5.40434041
Informer	8.200191370	5.27341824
Autoformer	6.856545364	4.51108246

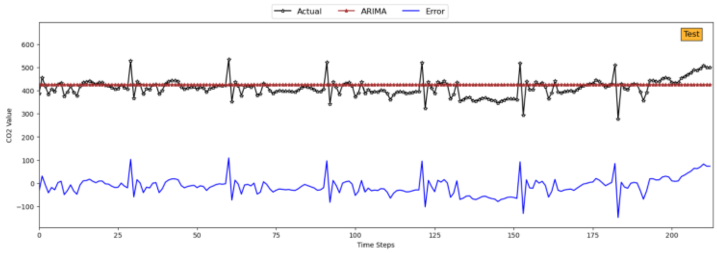

Fig 7. Visualization for ARIMA

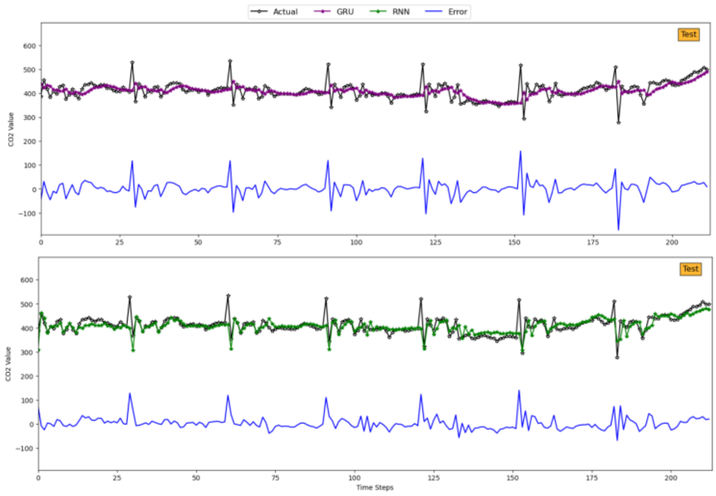

Fig 8. Visualization for GRU and RNN

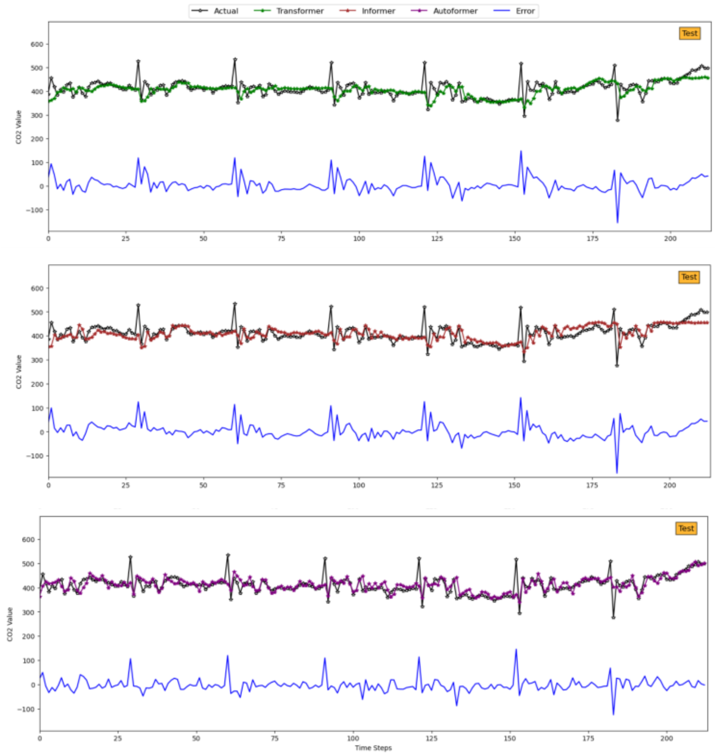

Fig 9. Visualization for Transformer, Informer and Autoformer

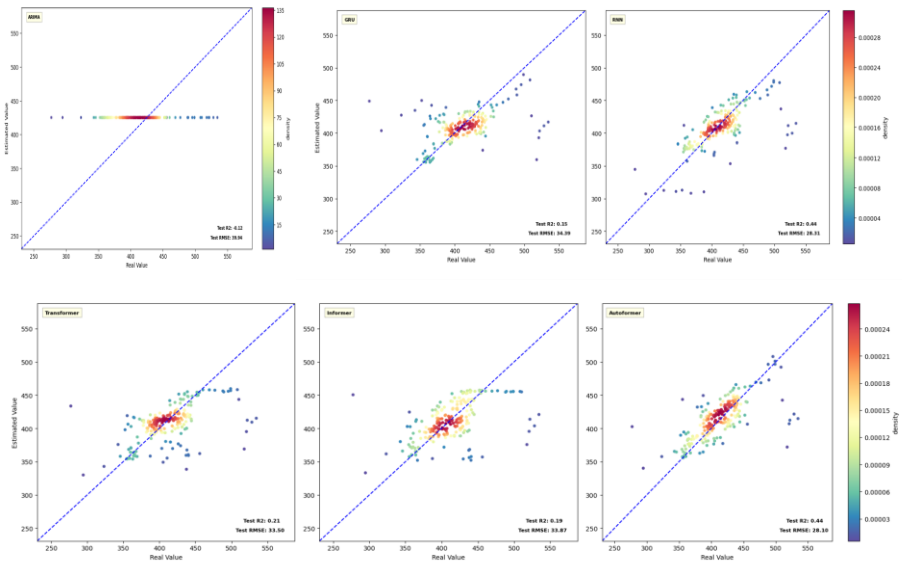

Fig 10. Errors in different models

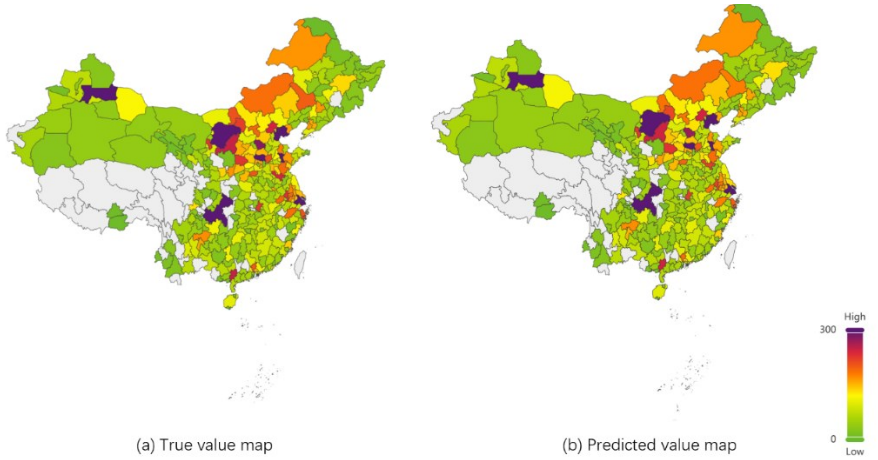

Fig 11. Map comparing true and predicted values on June 1st, 2021

4.4. Synergistic and Integrated

In our experiment, we divide the datasets into three sections, training, validation and testing sections, and the ratio between them is 7:1:2. In order to compare all the models at their best level, we regulate their parameters to ensure each of them is in their best performances, using RMSE and MAE as criterion. Then we do visualization to present our results more understandable and directly. Also, we do the comparison to verify our model is the best either in data or through visualization.

5. Conclusion

Accurate and precise prediction of CO_2 emissions in cities on a daily basis provides the most up-to-date and fine-grained overview of the change in CO_2 emissions. In this paper, we use the Autoformer framework to predict the long-term daily dynamic change process of CO_2 emissions, which has a high prediction accuracy.

A city-level daily CO_2 emission dataset covering three years of CO_2 emission in 300 cities in China is used to demonstrate the effectiveness of our method. The method can be applied to other cities around the world.

Our prediction model focuses on predicting total CO_2 emissions in cities without distinguishing between specific sectors. In the future, we plan to propose new models that support sectoral CO_2 emissions in more detail. Besides, in this study, we construct the prediction model for each city separately. The carbon emission among cities may have potential spatial relationships. In the future, we plan to use the relationship between the spatial distribution characteristics of CO_2 emissions and the economic activities among cities to support the spatiotemporal prediction of CO_2 emissions.

Acknowledgement

Shuyi Wei and Xin You contributed equally to this work and should be considered co-first authors.

Reference

1. Mehmood I, Bari A, Irshad S, et al. Carbon cycle in response to global warming[J]. Environment, climate, plant and vegetation growth, 2020: 1-15.
2. Al-Ghussain L. Global warming: review on driving forces and mitigation[J]. Environmental Progress & Sustainable Energy, 2019, 38(1): 13-21.
3. Dong K, Hochman G, Timilsina G R. Do drivers of CO_2 emission growth alter overtime and by the stage of economic development?[J]. Energy Policy, 2020, 140: 111420.
4. Nyambuu U, Semmler W. Climate change and the transition to a low carbon economy–carbon targets and the carbon budget[J]. Economic Modelling, 2020, 84: 367-376.
5. Li Y, Sun Y. Modeling and predicting city-level CO_2 emissions using open access data and machine learning[J]. Environmental Science and Pollution Research, 2021, 28: 19260-19271.
6. Song Z, Wu Y, Hua C, et al. A Multi-factor Prediction Model for Carbon Productivity Based on Stacking Integration Method[C]//2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI). IEEE, 2022: 83-93.
7. Khajavi H, Rastgoo A. Predicting the carbon dioxide emission caused by road transport using a Random Forest (RF) model combined by Meta-Heuristic Algorithms[J]. Sustainable Cities and Society, 2023, 93: 104503.
8. Ren F, Long D. Carbon emission forecasting and scenario analysis in Guangdong Province based on optimized Fast Learning Network[J]. Journal of Cleaner Production, 2021, 317: 128408.
9. Cho K, Van Merriënboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.
10. Yu Y, Si X, Hu C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural computation, 2019, 31(7): 1235-1270.
11. Cho K, Van Merriënboer B, Bahdanau D, et al. On the properties of neural machine translation: Encoder-decoder approaches[J]. arXiv preprint arXiv:1409.1259, 2014.
12. Liu Z, Ciais P, Deng Z, et al. Carbon Monitor, a near-real-time daily dataset of global CO_2 emission from fossil fuel and cement production[J]. Scientific data, 2020, 7(1): 392.
13. Dou X, Wang Y, Ciais P, et al. Near-real-time global gridded daily CO_2 emissions[J]. The Innovation, 2022, 3(1).
14. Huo D, Huang X, Dou X, et al. Carbon Monitor Cities near-real-time daily estimates of CO_2 emissions from 1500 cities worldwide[J]. Scientific data, 2022, 9(1): 533.
15. "Search Results | World Resources Institute." World Resources Institute, www.wri.org/search?keys=CO2&f%5B0%5D=content_type%3Adata.
16. data.europa.eu. data.europa.eu/data/datasets/data_national-emissions-reported-to-the-unfccc-and-to-the-eu-greenhouse-gas-monitoring-mechanism-14?locale=en.
17. U.S. Energy Information Administration. www.eia.gov/opendata/browser/co2-emissions/co2-emissions-aggregates?frequency=annual&data=value;&sortColumn=period;&sortDirection=desc.
18. China Emission Accounts and Datasets. www.ceads.net/user/search.php?kwtype=0&pagelang=en&searchtype=titlekeyword&typeid=45&q=co2. Accessed 12 May 2024.

19. MEICGreenhouseGases – MEICModel. meicmodel.org.cn/?page_id=2345.
20. Lotfalipour M R, Falahi M A, Bastam M. Prediction of CO2 emissions in Iran using grey and ARIMA models[J]. International Journal of Energy Economics and Policy, 2013, 3(3): 229-237.
21. Kour M. Modelling and forecasting of carbon-dioxide emissions in South Africa by using ARIMA model[J]. International Journal of Environmental Science and Technology, 2023, 20(10): 11267-11274.
22. Hosseini S M, Saifoddin A, Shirmohammadi R, et al. Forecasting of CO2 emissions in Iran based on time series and regression analysis[J]. Energy Reports, 2019, 5: 619-631.
23. Zhao L, Zhao T, Yuan R. Scenario simulations for the peak of provincial household CO2 emissions in China based on the STIRPAT model[J]. Science of The Total Environment, 2022, 809: 151098.
24. Yang F, Shi L, Gao L. Probing CO2 emission in Chengdu based on STRIPAT model and Tapio decoupling[J]. Sustainable Cities and Society, 2023, 89: 104309.
25. Ding S, Dang Y G, Li X M, et al. Forecasting Chinese CO2 emissions from fuel combustion using a novel grey multivariable model[J]. Journal of Cleaner Production, 2017, 162: 1527-1538.
26. Ding S, Zhang H. Forecasting Chinese provincial CO2 emissions: A universal and robust new-information-based grey model[J]. Energy Economics, 2023, 121: 106685.
27. Li Y, Sun Y. Modeling and predicting city-level CO2 emissions using open access data and machine learning[J]. Environmental Science and Pollution Research, 2021, 28: 19260-19271.
28. Huang Y, Shen L, Liu H. Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China[J]. Journal of Cleaner Production, 2019, 209: 415-423.
29. Zuo Z, Guo H, Cheng J. An LSTM-STRIPAT model analysis of China's 2030 CO2 emissions peak[J]. Carbon Management, 2020, 11(6): 577-592.
30. Wen T, Liu Y, he Bai Y, et al. Modeling and forecasting CO2 emissions in China and its regions using a novel ARIMA-LSTM model[J]. Heliyon, 2023, 9(11).
31. Liu Z, Ciais P, Deng Z, et al. Carbon Monitor, a near-real-time daily dataset of global CO2 emission from fossil fuel and cement production[J]. Scientific data, 2020, 7(1): 392.
32. Dou X, Wang Y, Ciais P, et al. Near-real-time global gridded daily CO2 emissions[J]. The Innovation, 2022, 3(1).
33. Huo D, Huang X, Dou X, et al. Carbon Monitor Cities near-real-time daily estimates of CO2 emissions from 1500 cities worldwide[J]. Scientific data, 2022, 9(1): 533.
34. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
35. Wu H, Xu J, Wang J, et al. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting[J]. Advances in Neural Information Processing Systems, 2021, 34: 22419-22430.
36. Zhou H, Zhang S, Peng J, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[C]//Proceedings of the AAAI conference on artificial intelligence. 2021, 35(12): 11106-11115.