

## Article

# XCO<sub>2</sub> Data Full-Coverage Mapping in China Based on Random Forest Models

Ruizhi Chen <sup>1,2</sup>, Zhongting Wang <sup>2</sup>, Chunyan Zhou <sup>2</sup>, Ruijie Zhang <sup>1,2</sup>, Huizhen Xie <sup>2</sup> and Huayou Li <sup>2,\*</sup><sup>1</sup> Chinese Research Academy of Environmental Sciences, Beijing 100012, China<sup>2</sup> Satellite Application Center for Ecology and Environment, Ministry of Ecology and Environment, Beijing 100094, China

\* Correspondence: lihuayou@secmep.cn

**Abstract:** Carbon dioxide (CO<sub>2</sub>) is a key driver of global climate change. Since the Industrial Revolution, the rapid rise in atmospheric CO<sub>2</sub> levels has significantly intensified global warming and climate-related issues. To accurately and promptly monitor changes in CO<sub>2</sub> concentrations and to support the development of climate policies, this study proposes a method based on random forest models to generate a continuous monthly dataset of CO<sub>2</sub> column concentration (XCO<sub>2</sub>) across the entire Chinese region from 2004 to 2023. The study integrates XCO<sub>2</sub> satellite observations from SCIAMACHY, GOSAT, OCO-2, and GF-5B, alongside nighttime light remote sensing data, meteorological parameters, vegetation indices, and CO<sub>2</sub> profile data. Using the random forest algorithm, a complex relationship model was established between XCO<sub>2</sub> concentrations and various environmental variables. The goal of this model is to provide XCO<sub>2</sub> estimates with enhanced spatial coverage and accuracy. The XCO<sub>2</sub> concentrations predicted by the model show a high level of consistency with satellite observations, achieving a correlation coefficient (R-value) of 0.9959 and a root mean square error (RMSE) of 1.1631 ppm. This indicates that the model offers strong predictive accuracy and generalization ability. Additionally, ground-based validation further confirmed the model's effectiveness, with a correlation coefficient (R-value) of 0.956 when compared with TCCON site observation data.

**Keywords:** carbon dioxide; random forest; atmospheric remote sensing; Chinese region; GF-5B



Academic Editor: Stephan Havemann

Received: 6 November 2024

Revised: 21 December 2024

Accepted: 23 December 2024

Published: 27 December 2024

**Citation:** Chen, R.; Wang, Z.; Zhou, C.; Zhang, R.; Xie, H.; Li, H. XCO<sub>2</sub> Data Full-Coverage Mapping in China Based on Random Forest Models. *Remote Sens.* **2025**, *17*, 48. <https://doi.org/10.3390/rs17010048>

**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Greenhouse gases can absorb and emit radiation within the thermal infrared range, leading to the greenhouse effect, a critical driver of global climate change [1,2]. Carbon dioxide (CO<sub>2</sub>) is the most significant greenhouse gas produced by human activities, and its atmospheric concentration has been rising at an alarming rate since the Industrial Revolution, greatly contributing to global warming and climate change [3,4]. Accurately monitoring CO<sub>2</sub> concentrations is essential for addressing climate change and achieving carbon reduction targets [5]. Traditional ground-based monitoring methods face limitations due to sparse station distribution, making it challenging to obtain large-scale, continuous observations. In contrast, satellite-based monitoring methods, with their high spatial resolution, have been widely adopted for acquiring global and regional CO<sub>2</sub> data [6,7].

Over the past two decades, numerous satellites have been launched to monitor atmospheric carbon dioxide, including SCIAMACHY, GOSAT, OCO-2 [8–10], and China's TANSAT, GF-5, and DQ-1 [11,12]. These satellites use spectroscopic analysis technology to obtain XCO<sub>2</sub> data, which is widely used in scientific research and policymaking [3,13].

However, factors such as cloud cover, sensor limitations, and orbital constraints can lead to data gaps in satellite observations [12]. These gaps hinder the comprehensive analysis of XCO<sub>2</sub>'s spatial and temporal variations, particularly at regional scales.

To address these challenges, various data fusion and interpolation methods have been developed to generate continuous datasets [14,15]. However, traditional interpolation techniques, such as kriging, often overlook key factors influencing CO<sub>2</sub> concentrations, resulting in reduced spatial resolution and less accurate quantification of temporal changes [16,17]. In contrast, machine learning (ML) techniques have shown promising results in regards to data fusion and prediction, as they can model complex relationships between input and target variables without explicitly defining mathematical models. By integrating multiple auxiliary data sources, including emission inventories, meteorological parameters, and vegetation indices, ML models can provide more accurate XCO<sub>2</sub> estimates and reduce uncertainties associated with satellite observations [18].

In addition, although numerous datasets have been generated from single-satellite observations, they often have certain limitations. For instance, datasets derived from SCIAMACHY are constrained by the early technological capabilities, resulting in lower detection accuracy. Similarly, data from the GOSAT satellite suffer from insufficient coverage due to limitations in its detection principles. While OCO-2 satellite data excel in both detection accuracy and coverage, its relatively late launch limits the availability of long-term time-series datasets. Consequently, one of the most significant advantages of multi-source fused datasets is their superior spatiotemporal coverage compared to that of single-satellite datasets, as well as their ability to correct earlier data.

As the world's largest emitter of carbon dioxide, China has made significant efforts in recent years to control emissions and improve air quality [19–22]. Understanding the spatial and temporal variations of XCO<sub>2</sub> across China is crucial for assessing the effectiveness of these measures and guiding future policy decisions. However, existing satellite XCO<sub>2</sub> datasets contain data gaps and uncertainties, particularly in regions with limited observations. Therefore, there is an urgent need to develop advanced methods to generate continuous, high-resolution XCO<sub>2</sub> datasets for China.

In this study, we developed a method based on random forest methods to generate a continuous monthly XCO<sub>2</sub> dataset for China from 2004 to 2023. By integrating XCO<sub>2</sub> satellite observations from SCIAMACHY, GOSAT, OCO-2, and GF-5B with nighttime remote sensing data, meteorological data, vegetation indices, and CO<sub>2</sub> profile data, our aim is to produce XCO<sub>2</sub> estimates with enhanced spatial coverage and accuracy. This paper provides a comprehensive analysis of the temporal and spatial variations of XCO<sub>2</sub> in China and evaluates the model's performance through overall validation, representative region validation, and ground-based validation.

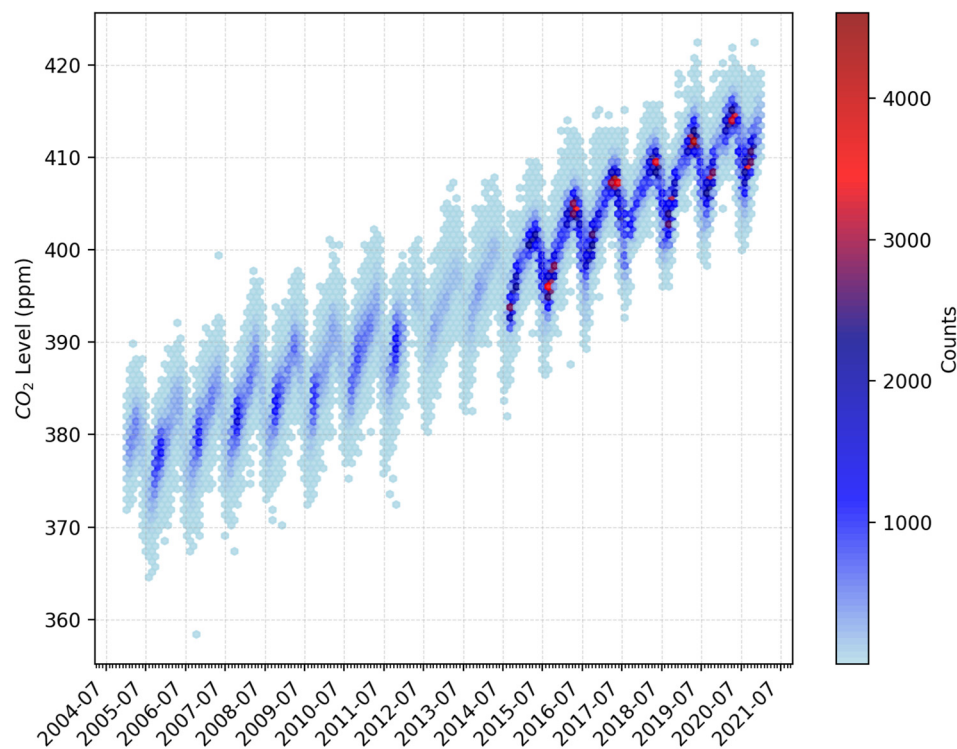
Our study stands out by incorporating GF-5B satellite data as a key input for model training. As of now, the application of GF-5B data is still in its early stages, and its full potential and technical advantages have yet to be widely recognized and validated. Through this research, we aim to leverage GF-5B's high spectral resolution to generate high-quality remote sensing datasets and explore its applications in the field of hyperspectral remote sensing. By systematically demonstrating its superiority and reliability in practical applications, our study not only seeks to expand the application scope of GF-5B data but also aims to enhance the global influence of China's hyperspectral satellite technology. This contributes to showcasing China's technical expertise in hyperspectral remote sensing while providing a "Chinese solution" and "Chinese perspective" to global sustainable development challenges.

## 2. Materials and Methods

### 2.1. Data Sources and Data Preprocessing

#### 2.1.1. Satellite Data

This study utilizes satellite data from SCIAMACHY, GOSAT, OCO-2, GF-5B, and OMI, including both XCO<sub>2</sub> and NO<sub>2</sub> data. The XCO<sub>2</sub> data from these satellites is presented in Figure 1, and the specific parameters of each satellite are listed in Table 1. Below is a brief introduction to the various satellite datasets and their preprocessing methods.



**Figure 1.** Display of original XCO<sub>2</sub> data from multi-source carbon satellites.

**Table 1.** Summary of satellite product information.

Satellite	SCIAMACHY	GOSAT	OCO-2	GF-5B	OMI
Time Coverage	2003.01–2012.03	2009.04–2016.12	2014.09–2023.12	2023.01–2023.12	2004.10–2023.12
Date Version	V02.01.02	9r	11r	-	V3 (OMNO2d)
Monitoring Indicators	CO <sub>2</sub>	CO <sub>2</sub>	CO <sub>2</sub>	CO <sub>2</sub>	NO <sub>2</sub>
Observation Time	10:00	13:00	13:36	13:30	13:45
Width of Coverage	960 km	790 km	10.6 km	865 km	2600 km
Spatial Resolution	30 × 60 km	10.5 km	2.25 × 1.5 km	10.3 km	13 × 24 km
Data Precision	~14 ppm	~1 ppm	~1 ppm	1~4 ppm	-

It is worth noting that the representation of CO<sub>2</sub> obtained through different methods is different, and CO<sub>2</sub> gas is uniformly mixed in the atmosphere; its total vertical column often varies due to factors such as terrain and pressure. In order to render CO<sub>2</sub> gas comparable under different conditions, XCO<sub>2</sub> (the mixed ratio of CO<sub>2</sub> in the entire air column after normalization with O<sub>2</sub>) is often used for representation. The reason for using O<sub>2</sub> as a representative is that its content in the atmosphere is known and has little variation. The CO<sub>2</sub> concentration data used in this paper are dry air mole fraction data, which will be represented by XCO<sub>2</sub> from now on without further explanation.

SCIAMACHY is a satellite launched by the European Space Agency (ESA) to measure CO<sub>2</sub> concentrations. It employs three different inversion algorithms: DOAS, WFM-DOAS, and BESD. Previous studies have shown that the BESD algorithm significantly reduces scattering-induced errors and excels at retrieving CO<sub>2</sub> and CH<sub>4</sub> concentrations. The accuracy of the BESD algorithm is within 3 ppm for individual observation points, with a regional bias of approximately 0.5 ppm. Therefore, this study utilizes the BESD V02.01.02 product, covering the period from January 2003 to March 2012. GOSAT, launched by Japan in 2009, is the world's first satellite dedicated to greenhouse gas monitoring. This study uses the NASA-released GOSAT-ACOS retrieval product, version 9r, which spans from April 2009 to December 2016, with a retrieval accuracy of approximately 1 ppm. OCO-2, launched by NASA in 2014, monitors atmospheric CO<sub>2</sub> concentrations. Its data covers the period from September 2014 to December 2023, with a retrieval accuracy of about 0.5 to 1 ppm. The GF-5B high-spectral-resolution observation satellite is an important member of China's high-spectral-resolution observation satellites, specializing in fine-resolution remote sensing observations of the atmosphere and surface, which was successfully launched on 7 September 2021. It is equipped with the Greenhouse Gas Monitoring Instrument (GMI), which exhibits the world's leading high-spectral-resolution and sensitivity, and is mainly used to monitor the concentration of carbon dioxide (CO<sub>2</sub>) and methane (CH<sub>4</sub>) in the atmosphere. Its inversion uses the optimal estimation physical method, and the CO<sub>2</sub> inversion accuracy is 0.67%, better than the 1% design target. This paper uses data covering January to December 2023. The NO<sub>2</sub> data is sourced from the Aura satellite, launched by NASA in 2004, which carries the Ozone Mapping and Profiler Suite (OMI) instrument. This study selects the level 3 daily product, covering the period from January 2004 to December 2023.

Based on the satellite-related parameters and product technical documentation, this paper processes the aforementioned satellite data as follows: First, data range selection and quality screening are performed, focusing on observation data within China. Subsequently, data quality control is implemented. For XCO<sub>2</sub> data, to mitigate the effects of cloud and snow/ice cover, this study utilizes data flagged as "good" quality. For GOSAT and OCO-2 satellite data, only those with uncertainties below 0.5 ppm are selected. For NO<sub>2</sub> data, we use tropospheric NO<sub>2</sub> concentrations measured in molecules per square centimeter, selecting data with cloud cover percentages below 30%. Finally, outlier removal is conducted on all filtered data using a four-standard-deviation filter to eliminate abnormal data points.

Considering the original resolution of both the satellite data and auxiliary data, and given that this paper aims to generate a monthly grid dataset with a resolution of 0.25°, we utilized Python 3.11 software to rasterize the satellite data. To generate raster images at the target resolution, the satellite data processed with a four-standard-deviation filter are rasterized. Specifically, empty rasters are created for each month, and the value for each grid cell is calculated. For a given grid cell, if no satellite observations fall within it, the cell is assigned a "NoData" value. If a single observation falls within the cell, the observed value is directly assigned. In cases where multiple observations fall within the same grid cell, the cell value is computed as a weighted average, with weights inversely proportional to the uncertainty of the observations. This process ultimately produces a satellite-derived raster dataset with a resolution of 0.25°.

### 2.1.2. Auxiliary Data

Studies have indicated [23] that XCO<sub>2</sub> concentrations are primarily influenced by anthropogenic emissions, natural emissions, and uptake, as well as meteorological factors. Therefore, this paper selects DMSP/OLS and NPP/VIIRS nighttime light remote sensing data to simulate the impact of anthropogenic emissions; uses MODIS NDVI and EVI data

products to simulate the impact of carbon emissions and absorption by vegetation; and employs the ERA5 reanalysis global climate dataset to analyze the influence of meteorological factors on XCO<sub>2</sub> concentrations, specifically including sea level pressure, 2 m air temperature, boundary layer height, solar radiation flux, surface pressure, total column water vapor concentration, total precipitation, and wind speed and direction at 100 m. Taking into account the influence of other factors, we also incorporate digital elevation model (DEM) and CO<sub>2</sub> vertical profile data from CarbonTracker as auxiliary variables into the model. Given that the spatiotemporal resolutions of the data used in this study vary, and the data quality is uneven, preprocessing of various types of data is required to construct a dataset with a unified resolution for model training. Below is a brief introduction to the processing flow of the auxiliary data.

Studies have indicated [24] strong correlation between nighttime light remote sensing data and anthropogenic CO<sub>2</sub> emissions. However, DMSP/OLS and NPP/VIIRS data face challenges related to comparability across different sensors and years, with NPP/VIIRS data containing a significant amount of low-value noise. In this paper, we adopt the method proposed by Zhang et al. [25] to process the data, generating a monthly nighttime light remote sensing grid dataset with a resolution of 0.25° for the period from 2004 to 2023. Other auxiliary data also need to be unified to a 0.25° grid resolution. To increase the resolution, this paper employs bilinear interpolation, while first-order conservative interpolation is used for decreasing the resolution. Aside from selecting specific XCO<sub>2</sub> data based on satellite overpass times, the processing of CarbonTracker data follows the same approach as that for the other datasets. The specific parameters of the auxiliary data are presented in Table 2.

**Table 2.** Summary of auxiliary information.

Type	Variable	Temporal Resolution	Space Resolution	Data Source
Light	Light Brightness	Monthly	30 km × 60 km	DMSP/OLS
		Monthly	0.74 km	NPP/VIIRS
Vegetation	EVI, NDVI	14 d	0.05° × 0.05°	MODIS
Meteorology	AP, AT, BLH, SP, TCW, TP, WEV, WN, WE	Monthly	0.25° × 0.25°	ERA5
CT Model	CO <sub>2</sub> profile	3 h	3° × 2°	Carbon Tracker

The raster datasets, encompassing both satellite data and auxiliary data, were bundled and structured into data blocks for machine learning purposes. Ultimately, around 600,000 data entries were generated for model training, with 89,751 valid observations from the SCIAMACHY satellite, 43,686 valid observations from the GOSAT satellite, 461,935 valid observations from the OCO-2 satellite, and 11,935 valid observations from the GF-5B satellite. In this study, the dataset was divided into training sets and test sets at a ratio of 4:1 [26].

## 2.2. Training and Evaluation of Machine Learning Models

### 2.2.1. Random Forest Model

Random forest (RF) is an ensemble learning algorithm primarily used for classification and regression tasks. It improves the accuracy and robustness of the model by combining the prediction results of multiple decision trees, with the result obtained through voting or weighted average calculation of all the trees' predictions [27]. This study uses the random forest algorithm to establish the correlation between XCO<sub>2</sub> satellite observations and environmental variables. The algorithm model can be represented as follows:

$$P_{CO_2} \propto P_{Light} + P_{NO_2} + P_{EVI} + P_{NDVI} + P_{Me} + P_{DEM} + P_{CT}$$

$P_{CO_2}$  represents the actual satellite observation value;  $P_{Light}$ ,  $P_{NO_2}$ ,  $P_{EVI}$ ,  $P_{NDVI}$ ,  $P_{Me}$ ,  $P_{DEM}$ , and  $P_{CT}$  represent the brightness value, NO<sub>2</sub> concentration, EVI, NDVI, ERA5 meteorological reanalysis data, DEM, and CarbonTracker’s XCO<sub>2</sub> concentration value, respectively, at the corresponding location of the satellite observation.

### 2.2.2. Model Validation Metrics

Metrics such as RMSE, MAE, mean bias, linear correlation coefficient, and coefficient of determination were used to evaluate model performance, as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

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

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i|$$

where  $\bar{y}$  is the mean value of the XCO<sub>2</sub> measurements,  $\hat{y}_i$  is the value of the predicted XCO<sub>2</sub>, and N is the number of data samples in the dataset.

### 2.2.3. Workflow

As shown in Figure 2, based on the random forest model, the workflow of this paper includes the following three steps:

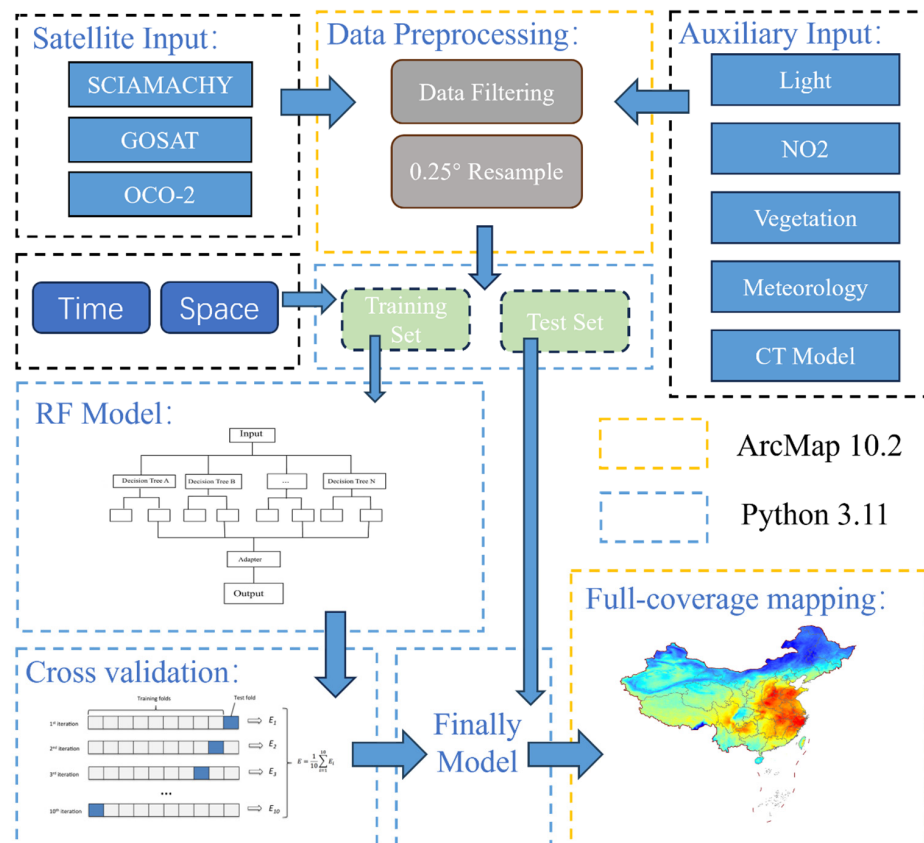


Figure 2. The workflow of XCO<sub>2</sub> full-coverage mapping.

Data preprocessing: Firstly, it involves the uniform inversion condition processing of satellite data to ensure that the data are inverted under the same conditions, improving accuracy and comparability. Subsequently, the processed satellite data, along with other auxiliary data for analysis, are converted to a uniform spatial resolution of  $0.25^\circ$ .

Selection of effective datasets for model training and validation, and determination of the dataset to be predicted: In the model training and validation stage, the effective dataset is further divided into a training set and a validation set, which are used for model training and performance evaluation, respectively. By continuously adjusting the model parameters, the model is better fitted to the training set data, and the validation set is used to assess the model's predictive ability. If necessary, the model is adjusted and retrained.

Prediction and data analysis: In this stage, after the model converges, the entire dataset to be predicted is input into the model for prediction, resulting in a spatially and temporally continuous CO<sub>2</sub> dataset. Subsequently, a post-evaluation of the dataset is conducted to check its accuracy and reliability. The dataset is then deeply analyzed to extract useful information, providing support for subsequent decisions or research.

#### 2.2.4. Hyperparameter Optimization

In the random forest model, adjusting the structure and complexity of the trees is crucial for enhancing predictive performance and ensuring stability. This process involves carefully selecting hyperparameters, known as hyperparameter tuning. Among these, five parameters—`n_estimators`, `max_depth`, `min_samples_split`, `min_samples_leaf`, and `max_features`—significantly impact model performance.

The `n_estimators` parameter specifies the number of decision trees in a random forest model, while `max_depth` sets the maximum depth of a single decision tree, controlling its growth. Deeper trees can create more complex models, which may affect accuracy. Generally, higher values for these two parameters lead to better model fitting but significantly increase computational costs. Therefore, selecting moderate values is crucial for efficient model operation.

The `min_samples_split` parameter determines the minimum number of samples required for a node to split, limiting conditions for further subtree division. Increasing this value when the sample size is large can help improve model accuracy. The `min_samples_leaf` parameter sets the minimum number of samples required in a leaf node, establishing the minimum sample size for leaf nodes. Finally, the `max_features` parameter determines the number of features considered for each node split, influencing the diversity of each tree.

Cross-validation is a commonly used method for evaluating model performance and selecting optimal hyperparameters. It allows for a more accurate assessment of the impact of different hyperparameter combinations on model performance. Grid search is a method for finding the best combination by iterating through a specified range of hyperparameters. First, the range and step size for the hyperparameters are defined, and all possible combinations are generated. For each combination, 10-fold cross-validation is used to evaluate model performance, and the combination with the best results is selected. In this paper, we choose the grid search method for hyperparameter optimization to identify the optimal parameters. Accuracy is selected as the evaluation metric to comprehensively assess the final model's performance.

In order to solve the inherent risk of overfitting in the random forest model, it is necessary to manually adjust the search space of the parameters and repeat the above grid search process. To prevent the model from becoming too complex, we limited the maximum tree depth (`max_depth`) to 20 and set the minimum number of samples required to split the internal nodes (`min_samples_split`) to 5. These parameters were optimized by

10-fold cross-validation, balancing model accuracy and generalization. The parameters of the specific hyperparameters and the initial search space are shown in Table 3.

**Table 3.** Hyperparameter model.

Hyperparameter	Hyperparameter Search Space	Final Hyperparameter
n_estimators	[100, 200, 300, 500, 1000, 1500, 2000]	1400
max_depth	[10, 20, 50, None]	20
min_samples_split	[2, 5, 10]	5
min_samples_leaf	[1, 2, 4]	2
max_features	['auto', 'sqrt', 'log2']	'sqrt'

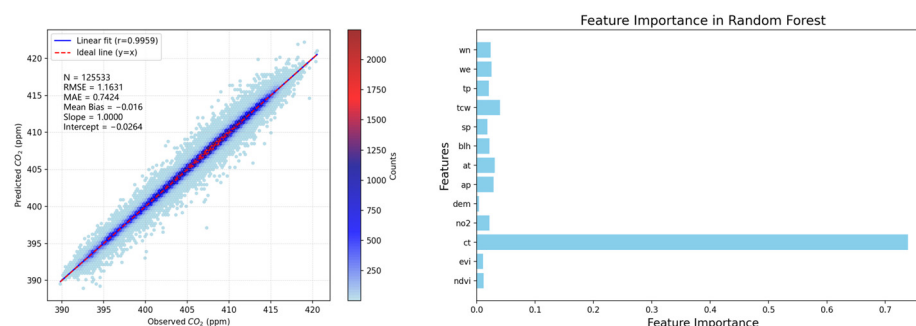
### 3. Results

#### 3.1. Overall Model Performance and the Importance of Variables

To evaluate model performance, we assessed the completed model from multiple perspectives, including overall model performance, performance in typical regions, and performance across different years.

Through comparison with the test dataset (accounting for 20% of the total data), we found that the model's prediction results exhibit excellent agreement with the satellite observation data, with an R-value greater than 0.99; the root mean square error (RMSE) is 1.1631 ppmv; the mean absolute error (MAE) is 0.7424 ppmv; the mean bias is  $-0.016$  ppmv; the slope of the fitted line is 1.0000, and the intercept is 0.0264. The overall fitting results indicate a strong correlation between the predicted XCO<sub>2</sub> and the observed XCO<sub>2</sub> within the Chinese region.

A sensitivity analysis was also performed to assess the impact of various input variables on the deep learning approach. The result is shown in Figure 3. The CarbonTracker results are the most critical variable, with an importance of 0.7. This is followed by various meteorological data. This result occurs mainly because these variables are related to terrestrial activities, including carbon uptake from photosynthesis driven by vegetation, solar radiation, and temperature, as well as carbon emissions from biosphere respiration influenced by temperature changes.



**Figure 3.** Test set overall results in China from 2004 to 2020.

After validating the model's accuracy, this study examines its predictive performance in three representative regions: Shandong, the Sichuan Basin, and the Pearl River Delta, each chosen for its unique emissions profile and regional characteristics. The specific location distribution is shown in Figure 4. Shandong Province, located along the eastern coast of China, boasts a mild climate with distinct seasons. As a major hub for industry and energy consumption, Shandong's industrial structure is significantly skewed towards heavy and chemical industries, resulting in persistently high carbon emissions. The Sichuan Basin is included in the study due to its relatively enclosed terrain and wet, rainy climate. However, frequent cloudy weather in this region poses challenges in obtaining high-quality



satellite data. The Pearl River Delta, an economic hub along the southern coast of China, features a hot and humid climate and is dually driven by urbanization and industrialization, with carbon emissions that cannot be overlooked. These three regions are not only economically developed and densely populated but also exhibit significant total carbon emissions. Coupled with relatively accurate energy consumption data, they provide ideal conditions for model validation. Therefore, this paper selects these three representative regions to comprehensively assess the model’s strengths and limitations in regards to carbon emission prediction and further validate the model’s predictive capabilities under different geographical and climatic conditions through practical application.

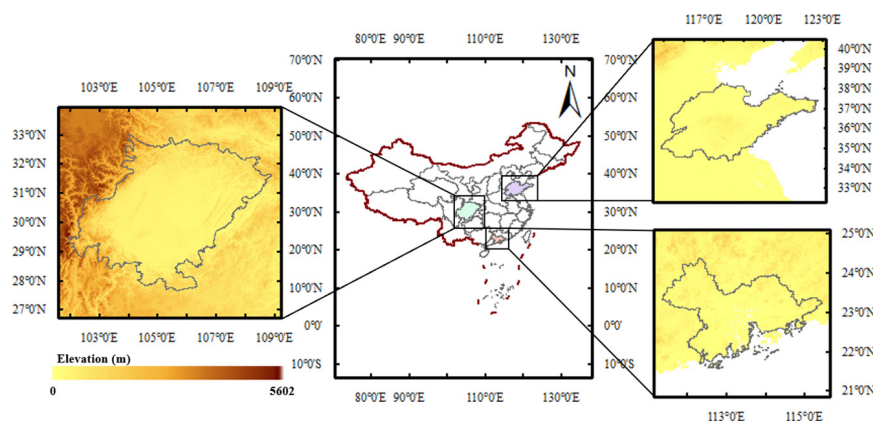


Figure 4. Representative regions used in this study.

As shown in Figure 5, the R-values of the model in China’s Shandong Province, Sichuan Basin, and Pearl River Delta regions are 0.9950, 0.9875, and 0.9821, respectively; the root mean square errors (RMSE) are 0.5748 ppmv, 0.8408 ppmv, and 0.9611 ppmv, respectively. These results indicate that the model exhibits good fitting performance in the typical regions and can accurately predict the regional column concentrations of carbon dioxide. Among the three typical regions selected in this paper, Shandong Province yields the best model fitting results with a slope and intercept of 0.9980 and 0.8086, respectively, and the fitting line is close to the 1:1 line. The R-values of the model in the Pearl River Delta and Sichuan Basin are slightly lower than those in Shandong, with intercepts of  $-4.8848$  and  $1.5278$  for the fitting lines, respectively. The poorer results in these regions can be attributed to the increased instability of the model due to the limited number of effective satellite observations. Overall, however, the predictive capability of the model is consistent at both regional and overall scales, indicating good generalization ability of the model.

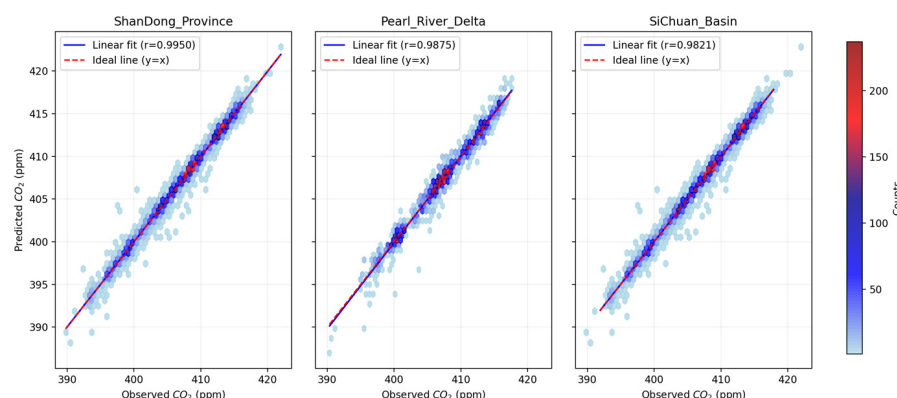


Figure 5. Test set overall results from 2004 to 2020 in the representative regions.

As shown in Table 4, we validated the model's prediction performance across different years. The results revealed that the model's prediction accuracy improved annually. During the period from 2004 to 2013, primarily relying on SCIAMACHY satellite data and constrained by data volume and uncertainty, the model's fitting performance was not ideal. From 2015 to 2016, we incorporated GOSAT and OCO-2 satellite data. These two satellites offered higher observation quality, and the increased training data volume significantly enhanced the model's fitting performance. Between 2017 and 2020, the model was trained solely on OCO-2 satellite data, achieving the best fitting results. Because the original GOSAT dataset used in this article only contains data from 2009–2016, data after 2016 does not include GOSAT. The  $R^2$  values for each year exceeded 0.92. This outstanding performance can be primarily attributed to the high quality and greater volume of effective data provided by the OCO-2 satellite.

**Table 4.** Test results of RF in China for each year from 2004 to 2020.

Year	Size N	RMSE	Accuracy	
			MAE	$R^2$
2004–2008	14,525	1.5778	1.2258	0.8466
2008–2013	15,392	1.5500	1.1996	0.8614
2014	8593	0.8983	0.4063	0.9174
2015	16,092	0.7353	0.4673	0.9291
2016	14,469	0.6885	0.4550	0.9393
2017	11,658	0.6805	0.4425	0.9284
2018	15,408	0.6748	0.4367	0.9462
2019	14,440	0.6828	0.4465	0.9395
2020	14,956	0.5003	0.3538	0.9665
2021	14,241	0.5712	0.4073	0.9571
2022	15,204	0.6521	0.4367	0.9477
2023	20,541	0.7941	0.4367	0.9324
All	170,519	1.1231	0.7124	0.9844

### 3.2. Ground-Based Station Validation

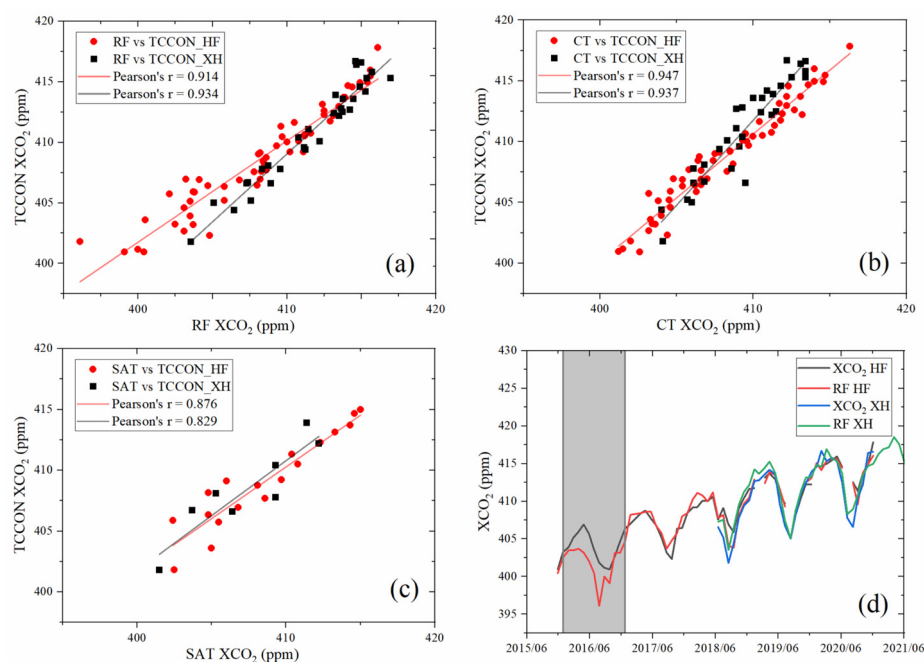
Ground-based monitoring is a crucial method for understanding changes in atmospheric  $CO_2$  concentrations. Its high precision and temporal resolution make it widely used in model accuracy assessments. TCCON is a network of ground-based Fourier transform spectrometers that record direct solar spectra in the near-infrared spectral region, from which accurate and precise column-averaged abundances of  $CO_2$ ,  $CH_4$ ,  $N_2O$ , HF, CO,  $H_2O$ , and HDO are retrieved and reported. The Hefei station is part of a high-resolution Fourier transform spectrometer (FTS) observation platform established by the Key Laboratory of Environmental Optics and Technology (Anhui Institute of Optics and Fine Mechanics, AIOFM) at the Chinese Academy of Sciences. In November 2018, the station successfully obtained data quality certification from the Total Carbon Column Observing Network (TCCON), making it China's first TCCON standard and benchmark station. The data provided by the Hefei station is indispensable for optimizing satellite algorithms, validating models, and advancing carbon neutrality research. The Xianghe station, located in Xianghe County, Langfang City, Hebei Province, serves as an important field observation and research base for the Institute of Atmospheric Physics at the Chinese Academy of Sciences. On 3 September 2021, after a rigorous evaluation by the TCCON Science Steering Committee, the Xianghe station was officially accepted as a TCCON standard station. Since June 2018, the station has continuously acquired high-quality spectral data and retrieval products for greenhouse gases, providing solid support for key research areas such as carbon peaking and carbon neutrality.

Although the comparison between different remote sensing instruments (such as TCCON and OCO-2) needs to take the differing sensitivity of  $XCO_2$  into account by

applying the average kernel, previous studies have shown that the difference between the corrected satellite data and the original data is around 0.2 ppm [28]. Compared with the differences between satellite and FTS data, the effect of XCO<sub>2</sub> comparison, with or without the application of prior profiles and mean cores, is small. Therefore, this paper directly compares satellite and FTS data, without considering the effects of different prior profiles and mean cores.

In this paper, data from the TCCON stations in Hefei and Xianghe are selected for model validation and accuracy assessment. The observed data are averaged on a monthly scale. The effective observation data from the Hefei station spans from October 2015 to December 2020, totaling 58 data points, while the Xianghe station covers the period from June 2018 to December 2020, with a total of 31 data points.

The validation results are shown in Figure 6a. These results indicate that the XCO<sub>2</sub> simulations performed by the random forest model are excellent, with mean absolute errors (MAEs) of 1.0616 for Hefei and 1.1267 for Xianghe. The root mean square errors (RMSE) are 1.5144 and 1.3357, respectively, with linear correlation coefficients of 0.914 and 0.934. Figure 6b compared the CT model data with the results of TCCON. The linear correlation coefficients of the HF station and the XH station were 0.947 and 0.937, respectively. The accuracy of the results from the HF station was slightly better than that of RF model, while the results from the XH station were similar to those of the RF model. Figure 6c compares the accuracy of the original satellite observation data with that from the TCCON stations. Figure 6d compares the monthly average XCO<sub>2</sub> concentration from the TCCON sites with the monthly XCO<sub>2</sub> concentration data obtained by the RF model. The concentrations from 2017 to 2020 are generally consistent, while the data from 2016 (indicated by the gray band area in the figure) show significant discrepancies, with a maximum difference of 5.8 ppm in August 2016. The poor results in 2016 may be attributed to the inadequate intercalibration of GOSAT and OCO-2 satellite data in overlapping years, resulting in substantial errors at the same observation point and poor model performance. Overall, the model results show significant consistency with station monitoring results, indicating that the model effectively estimates XCO<sub>2</sub> concentrations.



**Figure 6.** (a) Data pairs compared with TCCON for RF; (b) data pairs compared with TCCON for CT; (c) data pairs compared with TCCON for SAT; (d) monthly TCCON XCO<sub>2</sub> compared with RF, The shaded areas show the comparison results for 2016.

### 3.3. Comparison of Fitting Results of Different Models

To further evaluate the superiority of the model, we also compared the RF model to other commonly used models, such as ERT, XGBoost, and ANN. Considering the problem of computational cost, this paper selects the training set and validation set of data for three years from 2017 to 2019. The model is evaluated from the perspective of the whole model and the representative region. In previous experiments, it was found that the model performance of the three typical regions was similar, and the observation data were more effective in Shandong Province. Therefore, Shandong Province was used as the representative region for this area. The evaluation indexes of each model are shown in Table 1. The verification results show that the  $R^2$  values of RF and ERT are 0.952 and 0.929, respectively, while the  $R^2$  values of XGBoost and ANN are 0.941 and 0.912, respectively. XGBoost came the closest to RF in the validation results. Taking into account the model evaluation metrics and combining insights from relevant studies, based on our training dataset, we concluded that RF was the best choice. The verification results of different models are shown in Table 5.

**Table 5.** Comparison of evaluation metrics between the RF and commonly used models based on 10-fold cross-validation results.

Model	Overall Model Performance			Representative Regions Performance		
	$R^2$	MAE (ppm)	RMSE (ppm)	$R^2$	MAE (ppm)	RMSE (ppm)
RF	0.952	0.8424	1.0649	0.943	0.8414	1.2651
ERT	0.929	1.0624	1.2341	0.891	1.2654	1.5213
XGBoost	0.941	0.7366	1.2100	0.945	0.7996	1.1367
ANN	0.912	1.0023	1.5214	0.902	1.3177	1.7246

## 4. Discussion

### 4.1. Data Coverage Rate

To discuss the spatial coverage of the generated dataset, we compare the monthly average spatial coverage of the original satellite dataset with that of the RF model-generated dataset. The dataset of a single satellite is the dataset generated after the original data are rasterized. The multi-source carbon satellites raster dataset is a raster dataset generated in 2.1.1. The RF-model dataset is a dataset generated using the random forest algorithm. The spatial resolution of the above datasets is  $0.25^\circ$ . The spatial coverage values are shown in Table 6.

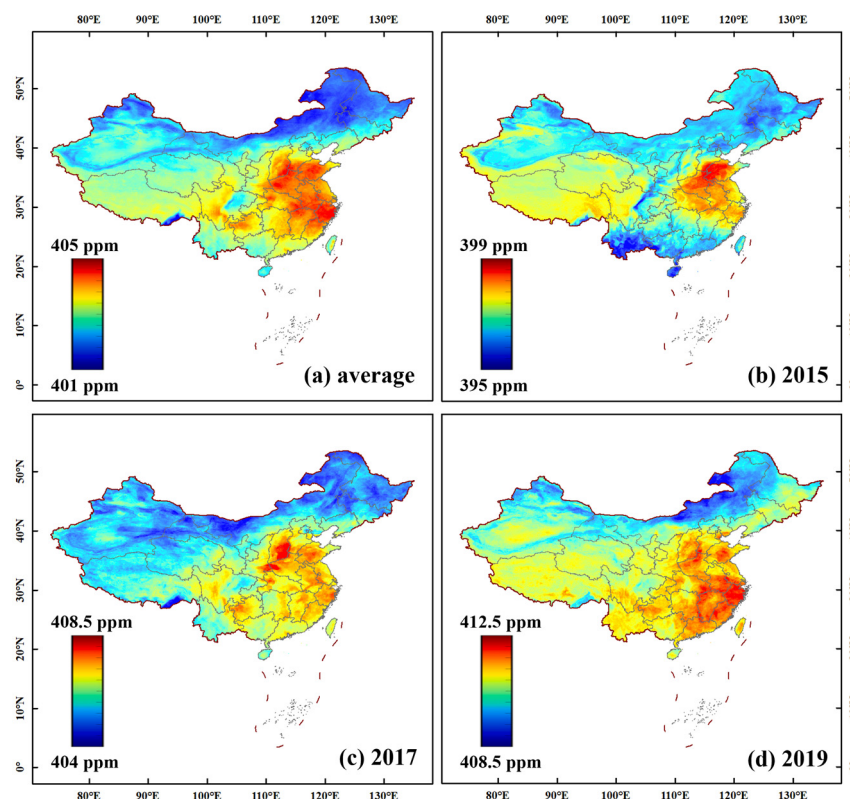
**Table 6.** Spatial coverage rate of different satellite datasets.

Datasets	Data Coverage Year	Monthly Data Coverage Rate
SCIAMACHY	2004–2012	2.8%
GOSAT	2019–2016	0.7%
OCO-2	2015–2023	5.2%
GF-5B	2023	0.7%
Multi-source carbon satellites raster dataset	2004–2023	6.1%
RF-model dataset	2004–2020	100%

### 4.2. Spatial Distribution Characteristics of Multi-Year Average $XCO_2$

As shown in Figure 7a, during the period from 2015 to 2020, the average concentration of  $XCO_2$  in China was approximately 403 ppmv. In terms of geographical distribution, the central and eastern regions displayed significantly higher concentrations, especially in the North China Plain, Zhejiang Province. These areas exhibited relatively high  $XCO_2$  concentrations due to frequent human activities. In contrast, the concentrations in the western Qinghai-Tibet Plateau and northeastern regions were lower, which may be related to lower emissions or higher natural carbon absorption in these areas. This study can

clearly reveal regions with higher XCO<sub>2</sub> concentrations, such as Shandong Province and Zhejiang Province, demonstrating a deeper understanding and analysis of these areas. These findings are of great significance for further understanding and addressing climate change issues.



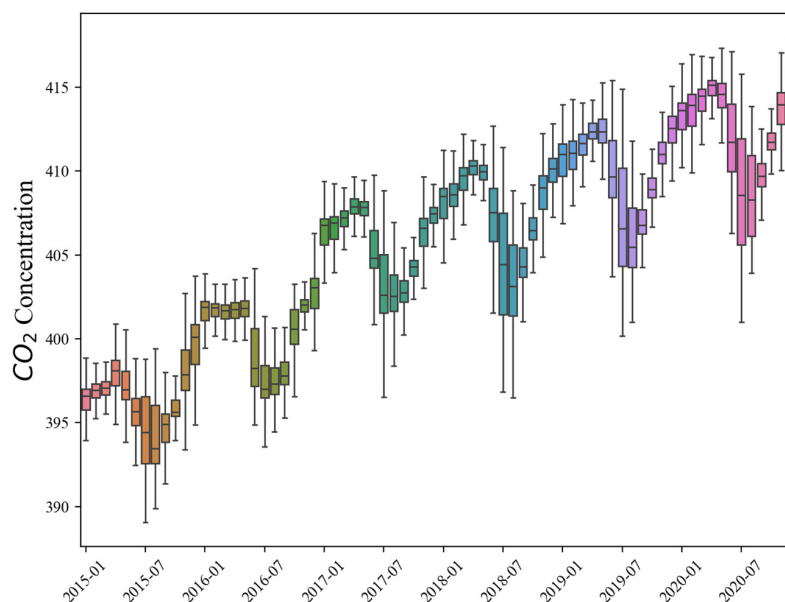
**Figure 7.** Spatial distribution pattern of XCO<sub>2</sub>: (a) average from 2015 to 2020; (b) result for 2015; (c) result for 2017; (d) result for 2019.

#### 4.3. Temporal Variation Characteristics of XCO<sub>2</sub>

Figure 7b–d displays the spatial distribution of XCO<sub>2</sub> in southeast China from 2015, 2017, and 2019, with the concentration of XCO<sub>2</sub> increasing at a rate of approximately 2 ppm per year. From 2015 to 2020, the overall concentration of XCO<sub>2</sub> in southeast China showed an upward trend, and the coverage area of the high-value regions, indicated in red, expanded annually. In 2015, the high-value regions of XCO<sub>2</sub> were mainly concentrated in the North China Plain, and gradually spread towards central China as the years progressed. By 2018, the high-value coverage areas had almost encompassed the entire southeast China area, and from 2018 to 2020, these high-value regions gradually stabilized. These changes may be related to factors such as the industrialization process, increased energy consumption, and climate change in southeast China.

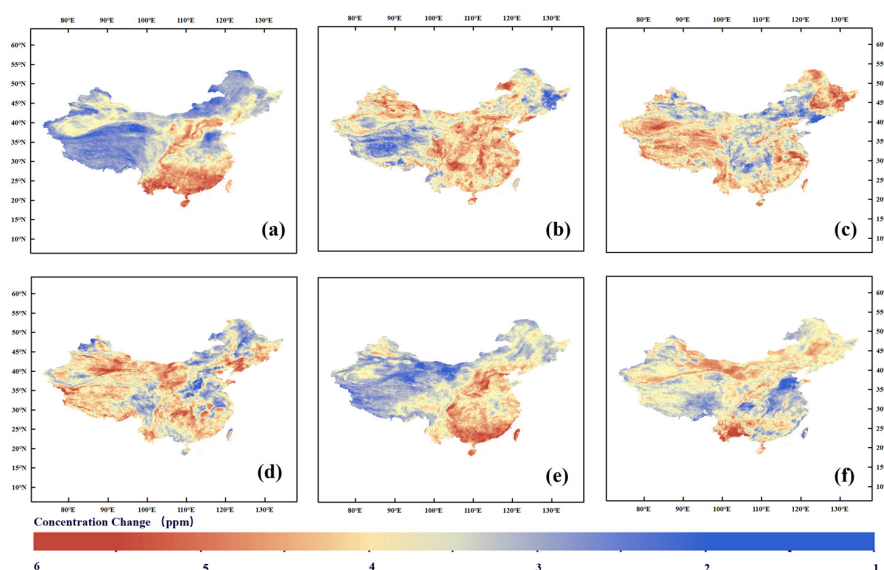
Data from 2015 to 2020 were selected from the generated continuous monthly CO<sub>2</sub> column concentration (XCO<sub>2</sub>) dataset, and 10,000 data points were randomly selected from the monthly raster images to describe the XCO<sub>2</sub> concentration in this month, as shown in Figure 8. The figure reveals pronounced seasonal fluctuations in CO<sub>2</sub> concentrations, with higher levels in winter and relatively lower levels in summer. This pattern may be attributed to increased heating demands during winter and enhanced CO<sub>2</sub> absorption capacity due to lush vegetation in summer. These findings indicate that the model has effectively captured the spatiotemporal variations in CO<sub>2</sub>. The difference between the monthly changes in 2015–2016 and 2017–2020 may be due to differences in the payloads

and inversion algorithms of the GOSAT and OCO-2 satellites, resulting in inferior mapping results for overlapping years than for years using single-satellite data.



**Figure 8.** Interannual variation in XCO<sub>2</sub> from 2015 to 2020.

By subtracting the XCO<sub>2</sub> mapping results from 2015 to 2020 year by year, we obtain the results shown in Figure 9, which is used to describe the change in XCO<sub>2</sub> distribution. The results in Figure 9a are calculated as follows: Subtract the 2020 result from the 2015 data and divide by five. This illustrates the average change in XCO<sub>2</sub> over the years, revealing that the regions with the most rapid growth in XCO<sub>2</sub> are primarily located in the Pearl River Delta and Hebei Province. Figure 9b–f depicts the distribution of XCO<sub>2</sub> changes across different years. While the interannual variation patterns are not evident, the XCO<sub>2</sub> concentrations in the Pearl River Delta (PRD) region exhibit a consistent trend of rapid increase. It is suggested that the PRD should focus on the energy sector and the industrial sector to optimize the energy structure and improve energy efficiency.



**Figure 9.** Interannual variation distribution of XCO<sub>2</sub>: (a) average from 2015 to 2020; (b) variation from 2019 to 2020; (c) variation from 2018 to 2019; (d) variation from 2017 to 2018; (e) variation from 2016 to 2017; (f) variation from 2015 to 2016.

## 5. Conclusions

This paper focuses on elucidating the important scientific topic of evaluating the spatiotemporal distribution patterns of atmospheric carbon dioxide in China. By applying machine learning methods, it achieves efficient integration of multiple satellites and various data types, thereby constructing a high-resolution dataset for spatially and temporally continuous XCO<sub>2</sub> (carbon dioxide column concentration) in China's atmosphere. This dataset not only provides a solid foundation for in-depth research but also emphasizes the crucial role of system validation and evaluation, aiming to ensure data accuracy and reliability, thereby enhancing the credibility and practical value of scientific research.

Based on the constructed dataset, this paper further delves into the spatiotemporal distribution characteristics of China's atmosphere and the underlying influencing factors from different dimensions. Through detailed analysis, we reveal how key driving factors affect the distribution and variation of atmospheric XCO<sub>2</sub> and accordingly, propose specific scientific insights and conclusions. This provides important scientific evidence and a reference for understanding climate change in China, as well as globally.

Although the machine learning model developed in this study demonstrates satisfactory performance, its architectural design lacks sufficient spatiotemporal adaptability, limiting its ability to accurately capture key features within both temporal and spatial neighborhoods. In addition, the model showed a decline in performance in years when data from different satellites overlapped. Due to the differences in payload parameters and inversion algorithms among different satellite systems, the data processing method of directly weighted averages in the same grid may be one of the reasons for the performance degradation. While random forest methods have achieved excellent performance, their tendency to overfit auxiliary data must be addressed, especially in large-scale distributed predictions. Combining additional regularization methods or exploring alternative algorithms, such as gradient enhancement or neural networks, can further improve the reliability of the model. Future studies should further explore fusion methods for diverse satellite data to address these challenges. Neural network models with spatiotemporal attention mechanisms may present a more effective solution. Therefore, future research should prioritize optimizing and innovating machine learning model designs. Specifically, integrating advanced carbon-monitoring satellite data, such as from GOSAT-2 and OCO-3, could facilitate the development of a globally continuous spatiotemporal CO<sub>2</sub> dataset.

## 6. Proclamation of AI-Assisted Generative Writing and AI-Supported Technologies

The writers employed ChatGPT-4o during the writing process to enhance readability and improve language. Following their use of this tool/service, the writers examined and made any necessary edits to the text, and they assume full responsibility for the publication's content.

**Author Contributions:** Conceptualization, R.C. and Z.W.; data curation, R.Z. and H.X.; formal analysis, R.C.; methodology, R.C. and Z.W.; supervision, Z.W., C.Z. and H.L.; validation, R.C., Z.W. and R.Z.; writing—original draft, R.C.; writing—review and editing, R.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the National Key Research and Development Program of China (2022YFE0209100) and the National Natural Science Foundation of China (Grant No. 41971324).

**Data Availability Statement:** The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

**Acknowledgments:** We thank ECMWF for providing the ERA5 dataset, available from <https://cds.climate.copernicus.eu/datasets> (accessed on 24 December 2024). We thank the NASA JPL Laboratory

for providing the GOSAT and OCO-2 dataset, which can be downloaded from <https://search.earthdata.nasa.gov/search> (accessed on 24 December 2024). CarbonTracker data can be downloaded at <https://gml.noaa.gov/aftp/products/carbontracker> (accessed on December 24, 2024). We thank the TCCON community for providing data used in this study, and the data can be downloaded at <https://tccodata.org> (accessed on 24 December 2024). The MOD13A3 dataset was downloaded at <https://search.earthdata.nasa.gov/search> (accessed on 24 December 2024). We thank all those institutions and their associates.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. He, C.; Ji, M.; Grieneisen, M.L.; Zhan, Y. A review of datasets and methods for deriving spatiotemporal distributions of atmospheric CO<sub>2</sub>. *J. Environ. Manag.* **2022**, *322*, 116101. [[CrossRef](#)]
2. Hu, K.; Feng, X.; Zhang, Q.; Shao, P.; Liu, Z.; Xu, Y.; Wang, S.; Wang, Y.; Wang, H.; Di, L.; et al. Review of Satellite Remote Sensing of Carbon Dioxide Inversion and Assimilation. *Remote Sens.* **2024**, *16*, 3394. [[CrossRef](#)]
3. Zhao, S.; Liu, M.; Tao, M.; Zhou, W.; Lu, X.; Xiong, Y.; Li, F.; Wang, Q. The role of satellite remote sensing in mitigating and adapting to global climate change. *Sci. Total Environ.* **2023**, *904*, 166820. [[CrossRef](#)] [[PubMed](#)]
4. Fernández-Martínez, M.; Sardans, J.; Chevallier, F.; Ciais, P.; Obersteiner, M.; Vicca, S.; Canadell, J.G.; Bastos, A.; Friedlingstein, P.; Sitch, S.; et al. Global trends in carbon sinks and their relationships with CO<sub>2</sub> and temperature. *Nat. Clim. Change* **2019**, *9*, 73–81. [[CrossRef](#)]
5. Jeong, K.; Hong, T.; Kim, J. Development of a CO<sub>2</sub> emission benchmark for achieving the national CO<sub>2</sub> emission reduction target by 2030. *Energy Build.* **2018**, *158*, 86–94. [[CrossRef](#)]
6. Li, C.; Wang, X.; Ye, H.; Wu, S.; Shi, H.; An, Y.; Sun, E. Assessment of thermal power plant CO<sub>2</sub> emissions quantification performance and uncertainty of measurements by ground-based remote sensing. *Environ. Pollut.* **2024**, *361*, 124886. [[CrossRef](#)]
7. Xie, F.; Ren, T.; Zhao, C.; Wen, Y.; Gu, Y.; Zhou, M.; Wang, P.; Shiomi, K.; Morino, I. Fast retrieval of XCO<sub>2</sub> over east Asia based on Orbiting Carbon Observatory-2 (OCO-2) spectral measurements. *Atmos. Meas. Tech.* **2024**, *17*, 3949–3967. [[CrossRef](#)]
8. Crisp, D.; Pollock, H.R.; Rosenberg, R.; Chapsky, L.; Lee, R.A.; Oyafuso, F.A.; Frankenberg, C.; O'Dell, C.W.; Bruegge, C.J.; Doran, G.B.; et al. The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products. *Atmos. Meas. Tech.* **2017**, *10*, 59–81. [[CrossRef](#)]
9. O'dell, C.W.; Eldering, A.; Wennberg, P.O.; Crisp, D.; Gunson, M.R.; Fisher, B.; Frankenberg, C.; Kiel, M.; Lindqvist, H.; Mandrake, L.; et al. Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm. *Atmos. Meas. Tech.* **2018**, *11*, 6539–6576. [[CrossRef](#)]
10. Wang, J.; Feng, L.; Palmer, P.I.; Liu, Y.; Fang, S.; Bösch, H.; O'Dell, C.W.; Tang, X.; Yang, D.; Liu, L.; et al. Large Chinese land carbon sink estimated from atmospheric carbon dioxide data. *Nature* **2020**, *586*, 720–732. [[CrossRef](#)]
11. Wunch, D.; Wennberg, P.O.; Osterman, G.; Fisher, B.; Naylor, B.; Roehl, C.M.; O'Dell, C.; Mandrake, L.; Viatte, C.; Kiel, M.; et al. Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO<sub>2</sub> measurements with TCCON. *Atmos. Meas. Tech.* **2017**, *10*, 2209–2238. [[CrossRef](#)]
12. Crisp, D. Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2). *Proc. SPIE Earth Obs. Syst.* **2015**, *9607*, 960702.
13. Schimel, D.S.; Carroll, D. Carbon Cycle-Climate Feedbacks in the Post-Paris World. *Annu. Rev. Earth Planet. Sci.* **2024**, *52*, 467–493. [[CrossRef](#)]
14. Li, J.; Jia, K.; Wei, X.; Xia, M.; Chen, Z.; Yao, Y.; Zhang, X.; Jiang, H.; Yuan, B.; Tao, G.; et al. High-spatiotemporal resolution mapping of spatiotemporally continuous atmospheric CO<sub>2</sub> concentrations over the global continent. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *108*, 102743. [[CrossRef](#)]
15. Sheng, M.; Lei, L.; Zeng, Z.C.; Rao, W.; Song, H.; Wu, C. Global land 1° mapping dataset of XCO<sub>2</sub> from satellite observations of GOSAT and OCO-2 from 2009 to 2020. *Big Earth Data* **2023**, *7*, 180–200. [[CrossRef](#)]
16. Grosz, B.; Horváth, L.; Gyöngyösi, A.Z.; Weidinger, T.; Pintér, K.; Nagy, Z.; André, K. Use of WRF result as meteorological input to DNDC model for greenhouse gas flux simulation. *Atmos. Environ.* **2015**, *122*, 230–235. [[CrossRef](#)]
17. He, Z.; Lei, L.; Zhang, Y.; Sheng, M.; Wu, C.; Li, L.; Zeng, Z.C.; Welp, L.R. Spatio-Temporal Mapping of Multi-Satellite Observed Column Atmospheric CO<sub>2</sub> Using Precision-Weighted Kriging Method. *Remote Sens.* **2020**, *12*, 576. [[CrossRef](#)]
18. Zhang, M.; Liu, G. Mapping contiguous XCO<sub>2</sub> by machine learning and analyzing the spatio-temporal variation in China from 2003 to 2019. *Sci. Total Environ.* **2023**, *858*, 159588. [[CrossRef](#)]
19. Liu, Z.; Deng, Z.; He, G.; Wang, H.; Zhang, X.; Lin, J.; Qi, Y.; Liang, X. Challenges and opportunities for carbon neutrality in China. *Nat. Rev. Earth Environ.* **2022**, *3*, 141–155. [[CrossRef](#)]



20. Shi, Q.; Zheng, B.; Zheng, Y.; Tong, D.; Liu, Y.; Ma, H.; Hong, C.; Geng, G.; Guan, D.; He, K.; et al. Co-benefits of CO<sub>2</sub> emission reduction from China's clean air actions between 2013–2020. *Nat. Commun.* **2022**, *13*, 5061. [[CrossRef](#)] [[PubMed](#)]
21. Yuan, B.; Li, C.; Yin, H.; Zeng, M. Green innovation and China's CO<sub>2</sub> emissions—The moderating effect of institutional quality. *J. Environ. Plan. Manag.* **2022**, *65*, 877–906. [[CrossRef](#)]
22. Zheng, X.; Lu, Y.; Yuan, J.; Baninla, Y.; Zhang, S.; Stenseth, N.C.; Hessen, D.O.; Tian, H.; Obersteiner, M.; Chen, D. Drivers of change in China's energy-related CO<sub>2</sub> emissions. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 29–36. [[CrossRef](#)]
23. Cui, Y.; Zha, H.; Jiang, L.; Zhang, M.; Shi, K. Luojia 1-01 Data Outperform Suomi-NPP VIIRS Data in Estimating CO<sub>2</sub> Emissions in the Service, Industrial, and Urban Residential Sectors. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 3000905. [[CrossRef](#)]
24. Shi, K.; Shen, J.; Wu, Y.; Liu, S.; Li, L. Carbon dioxide (CO<sub>2</sub>) emissions from the service industry, traffic, and secondary industry as revealed by the remotely sensed nighttime light data. *Int. J. Digit. Earth* **2021**, *14*, 1514–1527. [[CrossRef](#)]
25. Zhang, B.; Li, J.; Wang, M.; Duan, P. Mutual Correction of DMSP/OLS and NPP/VIIRS in Mainland China. *Remote Sens. Inf.* **2021**, *36*, 99–107.
26. Talekar, B.; Agrawal, S. A Detailed Review on Decision Tree and Random Forest. *Biosci. Biotechnol. Res. Commun.* **2020**, *13*, 245–248. [[CrossRef](#)]
27. Yu, R.; Zhao, G.; Chang, C.; Yuan, X.; Wang, Z. Random Forest Classifier in Remote Sensing Information Extraction: A Review of Applications and Future Development. *Remote Sens. Inf.* **2019**, *34*, 8–14.
28. Wang, W.; Tian, Y.; Liu, C.; Sun, Y.; Liu, W.; Xie, P.; Liu, J.; Xu, J.; Morino, I.; Velazco, V.A.; et al. Investigating the performance of a greenhouse gas observatory in Hefei, China. *Atmos. Meas. Tech.* **2017**, *10*, 2627–2643. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.