



Neural-network-based estimation of regional-scale anthropogenic CO₂ emissions using an Orbiting Carbon Observatory-2 (OCO-2) dataset over East and West Asia

Farhan Mustafa¹, Lingbing Bu¹, Qin Wang¹, Na Yao¹, Muhammad Shahzaman², Muhammad Bilal³, Rana Waqar Aslam⁴, and Rashid Iqbal⁵

¹Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, Key Laboratory of Meteorological Disasters, Ministry of Education, Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China

²School of Atmospheric Sciences (SAS), Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China

³School of Marine Sciences (SMS), Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China

⁴State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430079, China

⁵Department of Agronomy, Faculty of Agriculture and Environment, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan

Correspondence: Lingbing Bu (lingbingbu@nuist.edu.cn)

Received: 22 July 2021 – Discussion started: 11 August 2021

Revised: 14 October 2021 – Accepted: 15 October 2021 – Published:

Abstract. Atmospheric carbon dioxide (CO₂) is the most significant greenhouse gas, and its concentration is continuously increasing, mainly as a consequence of anthropogenic activities. Accurate quantification of CO₂ is critical for addressing the global challenge of climate change and for designing mitigation strategies aimed at stabilizing CO₂ emissions. Satellites provide the most effective way to monitor the concentration of CO₂ in the atmosphere. In this study, we utilized the concentration of the column-averaged dry-air mole fraction of CO₂, i.e., XCO₂ retrieved from a CO₂ monitoring satellite, the Orbiting Carbon Observatory-2 (OCO-2), and the net primary productivity (NPP) provided by the Moderate Resolution Imaging Spectroradiometer (MODIS) to estimate the anthropogenic CO₂ emissions using the Generalized Regression Neural Network (GRNN) over East and West Asia. OCO-2 XCO₂, MODIS NPP, and the Open-Data Inventory for Anthropogenic Carbon dioxide (ODIAC) CO₂ emission datasets for a period of 5 years (2015–2019) were used in this study. The annual XCO₂ anomalies were calculated from the OCO-2 retrievals for each year to remove the larger background CO₂ concentrations and seasonal variability.

The XCO₂ anomaly, NPP, and ODIAC emission datasets from 2015 to 2018 were then used to train the GRNN model, and, finally, the anthropogenic CO₂ emissions were estimated for 2019 based on the NPP and XCO₂ anomalies derived for the same year. The estimated and the ODIAC CO₂ emissions were compared, and the results showed good agreement in terms of spatial distribution. The CO₂ emissions were estimated separately over East and West Asia. In addition, correlations between the ODIAC emissions and XCO₂ anomalies were also determined separately for East and West Asia, and East Asia exhibited relatively better results. The results showed that satellite-based XCO₂ retrievals can be used to estimate the regional-scale anthropogenic CO₂ emissions, and the accuracy of the results can be enhanced by further improvement of the GRNN model with the addition of more CO₂ emission and concentration datasets.

1 Introduction

Climate change is one of the greatest challenges to the future of Earth, and it stems from global warming, which is accelerated by anthropogenic emissions of greenhouse gases (Lammimpää et al., 2019). The major warming effects are caused by atmospheric CO₂ emissions, and significant amounts of these emissions are contributed by fossil fuel combustion and some industrial activities, such as the calcination of limestone during cement production (Hutchins et al., 2017). The levels of atmospheric CO₂ are continuously increasing (Mustafa et al., 2020), and if these levels continue to increase at the same rate, 1.5 °C of global warming will be reached between 2030 and 2052, which will cause more climate extremes (Hoegh-Guldberg et al., 2018).

Estimates of CO₂ emissions at national, regional, and global levels are now widely reported and have become an important element of public policy and mitigation strategies. Many countries are making efforts to reduce CO₂ emissions. Over the past few decades, significant work has been carried out to compile the regional and the global inventories of CO₂ emissions from anthropogenic activities (Olivier et al., 2005; Janssens-Maenhout et al., 2015; Gurney et al., 2009; Oda and Maksyutov, 2015). Most of the emission inventories employ bottom-up methods using available human activity data, emission factors, and corresponding technologies. The bottom-up methods incorporate energy consumption datasets along with other information, such as fuel purity and efficiency. However, it is known that such information can be subject to errors and biases, leading to considerable discrepancies and uncertainties in emission estimates, especially in the case of rapidly growing developing economies such as China and India (Guan et al., 2012; Korsbakken et al., 2016). These discrepancies can result in ~40 % to ~100 % uncertainty in emission estimations at the country and the local scales, respectively (Peylin et al., 2013; Wang et al., 2013). Moreover, defining the uncertainty in the inventory datasets is also a challenging task, and the intercomparisons of various inventories do not necessarily reveal all of the uncertainties, as different inventories sometimes use common sources of information (Konovalov et al., 2016)^{CE2}. It is becoming increasingly important to find efficient and reliable ways of monitoring CO₂ reduction progress and to evaluate how well specific CO₂ reduction policies are working.

Satellites provide the most effective way of monitoring atmospheric CO₂ with great spatiotemporal resolution. Several satellites such as the Greenhouse Gases Observing Satellite (GOSAT), GOSAT-2, the Orbiting Carbon Observatory-2 (OCO-2), OCO-3, and TanSAT^{CE3} are orbiting the Earth and are dedicated to monitoring atmospheric CO₂ (Crisp, 2015; Liu et al., 2018; Matsunaga et al., 2019; Taylor et al., 2020; Bao et al., 2020; Hong et al., 2021; Yang et al., 2018). These satellites calculate the average atmospheric CO₂ concentration in the path of sunlight reflected by the surface using spectrometers carried onboard. OCO-2 measures the CO₂

optical depth with bands centered around 1.6 and 2.0 μm^{TS1} and determines the O₂ optical depth using the A-band, which is centered around 0.76 μm (Crisp et al., 2017; O'dell et al., 2012). The information from these bands is combined to calculate the column-averaged dry-air mole fraction of CO₂ (XCO₂) (Crisp et al., 2012). Several studies suggest that XCO₂ can be used to detect the CO₂ concentration induced by anthropogenic activities by removing the background concentration from the satellite XCO₂ retrievals (Bovensmann et al., 2010; Hakkarainen et al., 2019; Keppel-Aleks et al., 2013). The results from these studies have reported an enhancement of nearly 2 ppm over megacities and high-density urban regions in the US and China. The XCO₂ retrievals derived from the satellite measurements show a positive correlation with the CO₂ emission inventories (Hakkarainen et al., 2016; Yang et al., 2019) which implies that these space-based observations can be used to assess the anthropogenic CO₂ emissions by enhancing the anthropogenic XCO₂ concentration.

Asia is home to the world's most populous nations with the highest CO₂ emissions. East Asia, in particular China, significantly contributes to the global carbon budget and has accounted for ~30 % of the overall growth in global CO₂ emissions over the past 15 years (EDGAR^{TS2}, 2017). This increment in the CO₂ levels is mainly due to the rapid economic growth and anthropogenic activities (Shan et al., 1997). China has pledged to make aggressive efforts to reduce the CO₂ emissions per unit gross domestic product (GDP) by 60 %–65 % relative to 2005 levels, and peak carbon emissions overall, by 2030 (UNFCCC^{TS3}, 2015). West Asia is also a region with higher rates of anthropogenic CO₂ emissions (Mustafa et al., 2020), and some of its countries, such as Iran, Saudi Arabia, and Turkey, are listed among the 10 largest CO₂ emitting nations in the world. Several studies have been carried out to estimate the CO₂ emissions using various machine learning techniques, but most of them do not deal with the spatial distribution. Rao (2021) estimated the CO₂ emissions using Support Vector Machine (SVM). Zhonghan et al. (2018) predicted the CO₂ flux emissions based on published data including latitude, age, potential net primary productivity (NPP), and mean depth using the Back Propagation Neural Network (BPNN) and Generalized Regression Neural Network (GRNN) models. Yang et al. (2019) estimated the anthropogenic CO₂ emissions using GOSAT XCO₂ retrievals over China, and the results showed good agreement between the estimated values and the ODIAC CO₂ emission dataset. In this study, we have improved the model initially developed by Yang et al. (2019) to estimate the regional-scale anthropogenic CO₂ emissions using OCO-2 XCO₂ retrievals over East and West Asia. MODIS NPP, OCO-2, and ODIAC CO₂ datasets were obtained for a period of 5 years from January 2015 to December 2019. XCO₂ anomalies were calculated from the OCO-2 retrievals for each year; the GRNN model was trained using XCO₂ anomalies, MODIS NPP, and ODIAC CO₂ emissions with 4 years

of data from 2015 to 2018; and then anthropogenic CO₂ emissions were estimated for the year 2019 based on 2019 NPP and XCO₂ anomalies. Atmospheric CO₂ monitoring satellites can detect and analyze the anthropogenic CO₂ signatures, and the satellite-based estimation of anthropogenic CO₂ emissions can be helpful in investigating the carbon emissions as a data-driven method, which is different from the conventional method of calculating an emission inventory. Although the estimation of anthropogenic CO₂ emissions using satellite datasets is a challenging task, as some other factors such as the atmospheric transport and the terrestrial ecosystem play notable roles in controlling the spatial distribution of atmospheric CO₂ (Cao et al., 2017), this data-driven method can still provide meaningful help with respect to quantifying anthropogenic CO₂ emissions that will be important for evaluating the effects of anthropogenic CO₂ emission reduction at regional as well as global scales.

The remainder of this paper is structured as follows: the details of the datasets and methods are provided in Sect. 2, and the results, including the estimated CO₂ emissions, an evaluation of these emissions, and the correlation between ODIAC CO₂ emissions and XCO₂ anomalies are discussed in Sect. 3.

2 Materials and methods

2.1 Datasets

2.1.1 OCO-2 dataset

The Orbiting Carbon Observatory-2 (OCO-2) was launched by the National Aeronautics and Space Administration (NASA) on 2 July 2014 to monitor the concentration of atmospheric CO₂ at regional and global levels (Crisp, 2015). It carries a three-channel imaging grating spectrometer that collects high-resolution, bore-sighted spectra of reflected sunlight. Spectra are collected in the molecular oxygen A-band at 0.765 μm and the CO₂ bands at 1.61 and 2.06 μm (Hakkarainen et al., 2019). Information from all of these bands is combined to calculate the XCO₂. The spatial resolution of OCO-2 is 2.25 km × 1.29 km. More details about the instrument design, calibration approach, in-orbit performance, and measurement principles are provided in a previous study (Crisp, 2015). In this study, we used the OCO-2 Atmospheric Carbon Observations from Space (ACOS)/XCO₂ version 10r product that was generated using the ACOS Level 2 Full Physics (L2FP) retrieval algorithm, which used a Bayesian optimal estimation framework to derive estimates of XCO₂ from spectral measurements of reflected solar radiation (O'dell et al., 2012; Crisp et al., 2012). A comprehensive study on the validation of OCO-2 XCO₂ retrievals against the Total Carbon Column Observing Network (TCCON) CO₂ dataset reported an absolute median difference of less than 0.4 ppm and a root-mean-square (RMS) differ-

ence of less than 1.5 ppm between the two datasets (Wunch et al., 2017). Similar experiments have been carried out for the validation of different versions of OCO-2 XCO₂ products, and the results have shown that the OCO-2 dataset was consistent and reliable for atmospheric CO₂ monitoring (Kiel et al., 2019; O'dell et al., 2018). The quality and the quantity of the XCO₂ product have been improved with the developments in the ACOS FP retrieval algorithm. The latest OCO-2 XCO₂ product has single sounding precision of ~0.8 ppm over land and ~0.5 ppm over water, and RMS biases of 0.5–0.7 ppm over both land and water (O'Dell et al., 2021). The evolution of the ACOS L2FP retrieval algorithm from v7 to v10 is summarized in Table 1.

No major changes were made in the ACOS v9 L2FP retrieval algorithm relative to v8 except for the sampling of the meteorological prior. The trace gas absorption coefficient tables (ABSCO) were updated in various versions of the ACOS L2FP retrieval algorithms. The source of the prior meteorology was changed from the European Center for Medium-Range Weather Forecasts (ECMWF) in ACOS v7 to the NASA Goddard Modeling and Assimilation Office (GMAO) Goddard Earth Observing System (GEOS) Forward Processing – Instrument Team (FP-IT) products for v8 and v9. The aerosol prior source was changed from the GMAO Modern-Era Retrospective analysis for Research and Applications (MERRA) product in v7–9 to Goddard Earth Observing System 5 (GEOS5) FP-IT in v10. Moreover, an additional stratospheric aerosol layer was introduced in ACOS v8–10. The prior value of aerosol optical depth (AOD) for each retrieved aerosol type was lowered from 0.0375 in v7 to 0.0125 in v8–10. The CO₂ prior developed by the Total Carbon Column Observing Network (TCCON) team using the ggg2014 algorithm remained same in v7, v8, and v9 of the algorithm. Another major change was switching the land surface model from a purely Lambertian land surface model to a bidirectional reflectance distribution function (BRDF) model (Taylor et al., 2021).

2.1.2 ODIAC dataset

ODIAC is a global fossil fuel CO₂ (FFCO₂) emission dataset with a 1 × 1 km monthly resolution over land and a 1 × 1° annual resolution for international bunkers from the year 2000 onward (Oda et al., 2018). It shares country-scale estimates with the Carbon Dioxide Information Analysis Center (CDIAC) but distributes the emissions differently within the countries and includes gridded international bunker emissions (Oda and Maksyutov, 2015). CDIAC distributes the CO₂ emissions based on the population density, whereas ODIAC incorporates power plant profiles and nighttime light observations for emission distribution (Wang et al., 2020). ODIAC shows better agreement with the US bottom-up inventory (Gurney et al., 2009) than CDIAC, and it is commonly used in flux inversions (Crowell et al., 2019; Lauvaux et al., 2016; Maksyutov et al., 2013; Takagi et al., 2011).

Table 1. Evolution of the Atmospheric Carbon Observations from Space (ACOS) Level 2 Full Physics (L2FP) retrieval algorithm (Taylor et al., 2021).

	ACOS v7	ACOS v8/9	ACOS v10
1 Spectroscopy	ABSCO v4.2	ABSCO v5.0	ABSCO v5.1
2 Meteorology prior source	ECMWF	GEOS5 FP-IT	No changes
3 Aerosol prior source	MERRA monthly climatology	No changes	GEOS5 FP-IT with tightened prior uncertainty
4 Retrieved aerosol types	Water, ice, and two MERRA types	With stratospheric aerosol	No changes
5 AOD prior value (per type)	0.0375	0.0125	No changes
6 CO ₂ prior source	TCCON ggg2014	No changes	TCCON ggg2020
7 Land surface model	Lambertian	BRDF	No changes

In this study, we used the 2020 version of ODIAC emission dataset that is freely available and can be downloaded from <http://db.cger.nies.go.jp/dataset/ODIAC/TS6>.

2.2 Methods

The estimation of anthropogenic CO₂ emissions includes three major steps, as shown in Fig. 1: the first step includes enhancing the XCO₂ concentration influenced by anthropogenic activities; the second step involves setting up the GRNN model using the XCO₂, NPP, and ODIAC datasets; and the final step is the validation of estimated CO₂ emissions against the actual ODIAC emission dataset.

The OCO-2 XCO₂ dataset was downloaded from the EARTHDATA platform (<https://earthdata.nasa.gov/TS7>); to ensure the reliability of the data, screening and filtering of the dataset was carried out following the instructions given in the OCO-2 Data User Guide (DUG). Each sounding that is processed using the ACOS L2FP retrieval algorithm is assigned either a “good” (0) or “bad” (1) quality flag based on screening criteria derived from comparisons with TCCON and modeled CO₂ fields. It is generally advised that users should use the good-quality soundings for regional- and local-scale studies because the soundings flagged as bad-quality might include biases that compromise their utility for the application. In this study, the OCO-2 XCO₂ retrievals were included if (i) they were flagged good (flag of 0) and (ii) the standard deviation of the good soundings for the day was less than 2 ppm. CO₂ has a larger background concentration and a longer atmospheric lifetime than other greenhouse gases (Hakkarainen et al., 2019). Hence, XCO₂ varies by nearly 2 % over the seasonal cycle and from pole to pole. In addition, XCO₂ variations influenced by anthropogenic activities are also smaller on the scale of satellite soundings (2–4 km²). Therefore, high precision is critical for the accurate quantification of the XCO₂ anomalies related to anthropogenic activities. To highlight the emission areas, CO₂ seasonal variability and the large background concentrations must be removed.

To highlight the areas associated with the anthropogenic CO₂ emission, XCO₂ anomalies were calculated by subtracting the daily XCO₂ median (daily background) from the individual XCO₂ observation – a method suggested by previous

studies (Hakkarainen et al., 2019, 2016):

$$\begin{aligned} XCO_2 (\text{anomaly}) = XCO_2 (\text{individual}) \\ - XCO_2 (\text{daily background}). \end{aligned} \quad (1)$$

This equation calculated the XCO₂ anomalies for each observation. Subtraction of the daily background concentration removes the seasonal variability. The space-based soundings are irregularly distributed and have spatiotemporal gaps because a large amount of the satellite observations is removed after screening for clouds and other artifacts. To deal with the spatiotemporal gaps, kriging interpolation was used, and a mapping dataset was generated with a spatial resolution of 0.5° × 0.5° (latitude × longitude) and a temporal resolution of 16 d. Finally, the mean against each grid cell was calculated for each year from 2015 to 2019. The annual mean of XCO₂ (anomaly) can detrend the seasonal variation (Hakkarainen et al., 2016). The annually averaged XCO₂ anomalies were resampled at a grid with a spatial resolution of 1° × 1° (latitude × longitude) and used along with 1° × 1° (latitude × longitude) ODIAC emission dataset to set up the GRNN model.

During the process of photosynthesis, living plants convert CO₂ into sugar molecules that they use for food. In the process of making food, they also release the oxygen we breathe. Plant productivity plays a crucial role in the global carbon cycle by absorbing the CO₂ released by anthropogenic activities. The net primary productivity (NPP) shows how much CO₂ is absorbed by plants during photosynthesis minus how much CO₂ is released during respiration. A negative NPP value means that CO₂ is released into the atmosphere, and a positive value represents the absorption of atmospheric CO₂. To improve the model results, an NPP dataset (MOD17A3HGF) provided by MODIS has also been used in this study. It provides information about annual NPP and is distributed by NASA’s Land Processes Distributed Active Archive Center (LP DAAC). The NPP dataset with a spatial resolution of 500 m was downloaded from the LP DAAC website (<https://lpdaac.usgs.gov/products/mod17a3hgf/006/TS8>). The annual NPP is derived from the sum of all 8 d Net Photosynthesis (PSN) products (MOD17A2H) from the given year. The MODIS NPP dataset was reprojected and resampled to the spatial res-

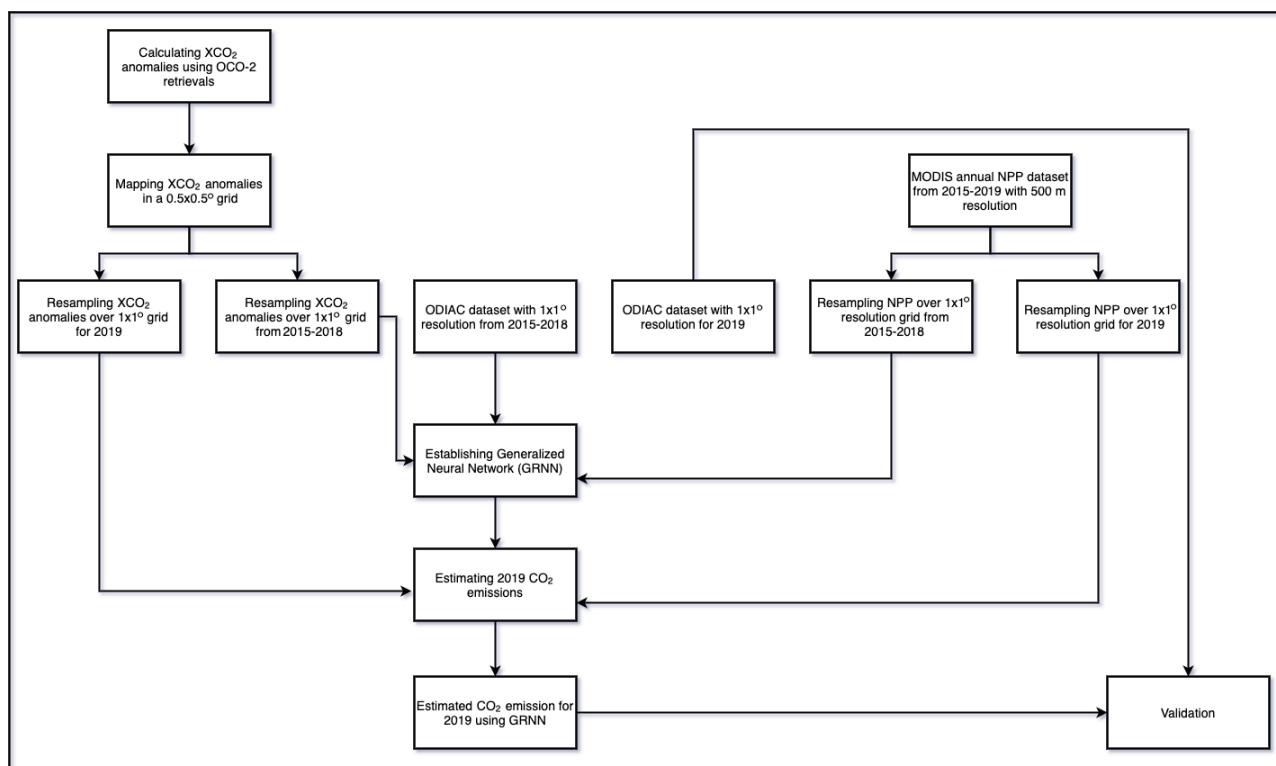


Figure 1. Flowchart explaining the steps involved in estimating the anthropogenic CO₂ emissions using MODIS NPP and OCO-2 XCO₂ retrievals.

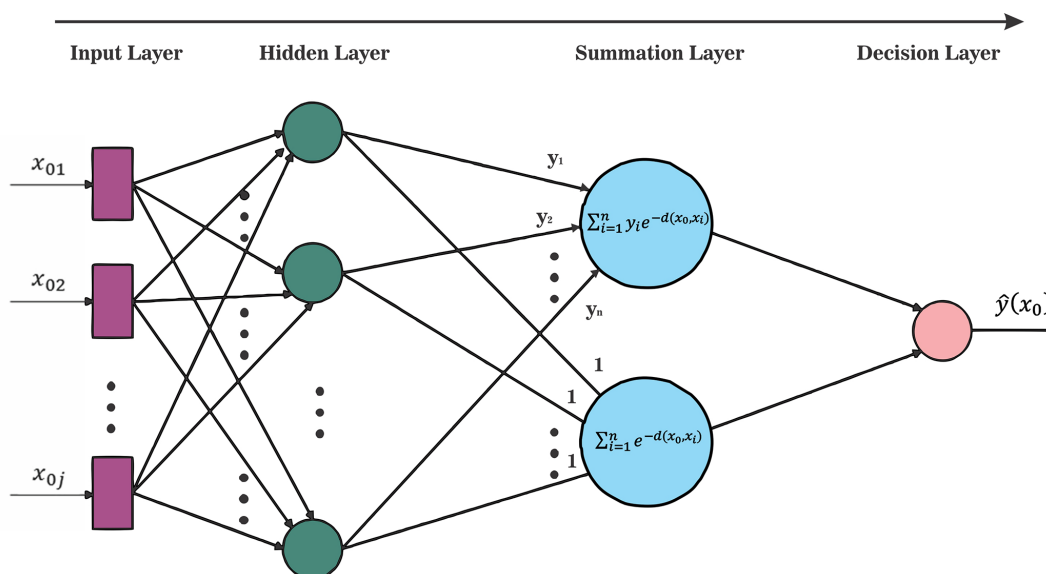


Figure 2. Flowchart explaining the steps involved in estimating the anthropogenic CO₂ emissions using OCO-2 XCO₂ retrievals (Yang et al., 2019).

olution of $1^\circ \times 1^\circ$ (latitude \times longitude) for each year and used along with the ODIAC and OCO-2 datasets to train the GRNN model and as well predict the CO₂ emissions.

XCO₂ variations are primarily influenced by anthropogenic activities and terrestrial ecosystems, and there is both linear and nonlinear mapping between the XCO₂ and the emissions. We adopted the GRNN algorithm to represent the nonlinear mapping between the independent variables (XCO₂ anomaly and NPP) and the dependent variable (CO₂ emissions). The GRNN is a memory-based network that provides estimates of continuous variables and converges to an underlying regression. The regression of a dependent variable on an independent variable is the computation of the most probable value of the dependent variable for each value of the independent variable based on a finite number of possibly noisy measurements of the independent variable and the associated values of the dependent variable. The dependent and the independent variables are usually vectors (Rooki, 2016). The architecture of GRNN is shown in Fig. 2. It consists of four layers including an input layer, a hidden layer, a summation layer, and a decision layer. In the input layer, each neuron corresponds to the independent variable that is expressed as a mathematical function, and the independent variable values are standardized. The standardized values of the independent variable are then transferred to the neurons in the hidden layer. In this layer, each neuron stores the values of the dependent and independent variables and calculates a scalar function. The third layer, known as the summation layer, contains two neurons: the denominator summation unit, which sums the weight values being received from the hidden layer, and the numerator summation unit, which sums the weight values multiplied by the actual target-dependent variable value for each hidden neuron. Finally, the target-dependent value is obtained in the decision layer by dividing the value accumulated in the numerator summation unit by the value in the denominator summation unit. To develop a neural network, the dependent and the independent training variables must be standardized so that all training data will have the same order of magnitude in the input layer (Yang et al., 2019) CE9.

$$d(x_0 - x_i) = \sum_{j=1}^p \left[\frac{x_{0j} - x_{ij}}{\sigma} \right]^2, \quad (2)$$

where p is the dimension of the variable vector x_i , σ is the spread parameter, and an optimal spread parameter value is obtained after several runs following the mean squared error of the estimated values, which must be kept at a minimum (Rooki, 2016). In this study, values of spread parameters were optimized using the “Holdout Method”. More details about the Holdout Method are provided in a previous study (Specht, 1991). The weight of the denominator neuron was set to 1.0. The predicted target dependent variable was

defined by the following equation:

$$\hat{y}(x_0) = \frac{\sum_{i=1}^n y_i e^{-d(x_0, x_i)}}{\sum_{i=1}^n e^{-d(x_0, x_i)}}, \quad (3)$$

where the values calculated with the scalar function in a hidden neuron i are weighted with the corresponding values of the training samples y_i . n denotes the number of training samples.

3 Results and discussion

3.1 Spatial distribution of XCO₂ observations and anomalies

The satellite-based observations are sensitive to clouds and aerosols; therefore, many of the data are discarded during preprocessing due to the presence of clouds and aerosols (Mustafa et al., 2021b). Figure 3a and b show the quantity of XCO₂ retrievals from 2015 to 2019 on a spatial grid of $0.5^\circ \times 0.5^\circ$ (latitude \times longitude) over West and East Asia, respectively. OCO-2 shows good spatial coverage over East Asia; however, the southern parts of the region, in particular the Tibetan Plateau, have a relatively lower number of XCO₂ retrievals. The Tibetan Plateau is the most extensively elevated surface on Earth, and satellite measurements show larger uncertainties over this region (Yang et al., 2019). In the case of West Asia, the southern parts of the region have a lower number of XCO₂ retrievals. A very large desert, the Rub’ al Kahli, is located in this area; it stretches across Saudi Arabia, Yemen, Oman, and the United Arab Emirates (UAE) and often observes dust storms. The lower number of XCO₂ retrievals in these parts of the region might be due to the ACOS XCO₂ retrieval algorithm that excludes satellite measurements with a high aerosol optical depth and cloud optical thickness (Crisp et al., 2012; O’dell et al., 2012).

Figure 3c shows the spatial distribution of the 5-year averaged XCO₂ anomalies calculated using the method described in Sect. 2.2 over West Asia. The higher concentrations of XCO₂ anomalies were observed over the central parts of the region that included Iran, Kuwait, Saudi Arabia, and Iraq. Iran and Saudi Arabia are listed among the top 10 CO₂ emitting nations and produce over 6 % of the global CO₂ emissions (Jalil, 2014). In addition, Iran, Saudi Arabia, and Iraq are the major fuel consumers of the region and contribute more than 60 % of the region’s total fossil fuel CO₂ emissions (Boden et al., 2017). Figure 3d shows the multiyear averaged XCO₂ anomalies over East Asia. The eastern parts of the region including eastern China, Japan, and South Korea show the highest concentrations of XCO₂ anomalies. China’s Beijing–Tianjin–Hebei area, Korea, and Japan are the most populated urban regions with high amounts of anthropogenic emissions in the world (Mustafa et al., 2020).

Figure 3e shows the monthly averaged XCO₂ over East and West Asia. The monthly averaged XCO₂ concentrations

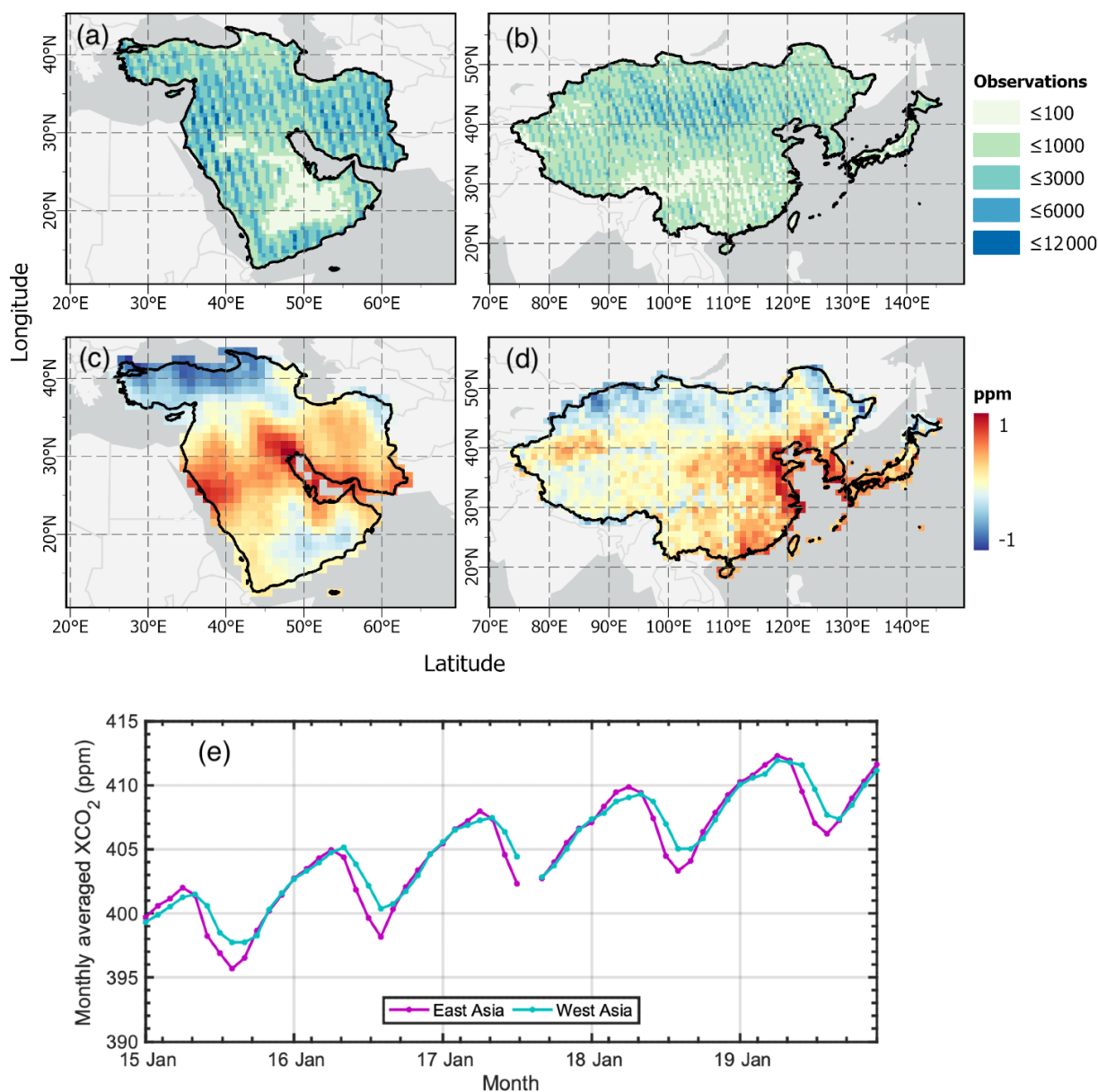


Figure 3. Number of observations in each cell of a $0.5 \times 0.5^\circ$ grid for a period of 5 years from 2015 to 2019 over (a) West Asia and (b) East Asia; the 5-year mean of XCO₂ anomalies calculated using OCO-2 retrievals over (c) West Asia and (d) East Asia; and (e) the monthly averaged XCO₂ concentration from 2015 to 2019 over East and West Asia. (The base map was sourced from OpenStreetMap.) ^{TS10}

show seasonal fluctuations. Moreover, the XCO₂ concentrations during each month are higher than those in the same month of the previous year, which reflects that the XCO₂ concentration in the atmosphere is continuously increasing in both regions. The XCO₂ concentration starts increasing from September and reaches its maximum value in April; it then starts decreasing and reaches its minimum value in August. The decrement in its concentration from May to August is due to several reasons; however, it is primarily owing to the strong photosynthesis and weak respiration rate of plants, which is enhanced during the monsoon or rainy

season (Mustafa et al., 2020). The increment in the XCO₂ concentration from September to April is likely to be caused by weak photosynthesis and strong respiration, the use of heating systems in winter, and strong microbial activity (Cao et al., 2017; Mustafa et al., 2021a).

3.2 Estimated CO₂ emissions

The annually averaged XCO₂ anomalies, MODIS NPP, and ODIAC CO₂ emission datasets for a period of 4 years from 2015 to 2018 were used as a training dataset for the GRNN model built to estimate the CO₂ emissions using the method

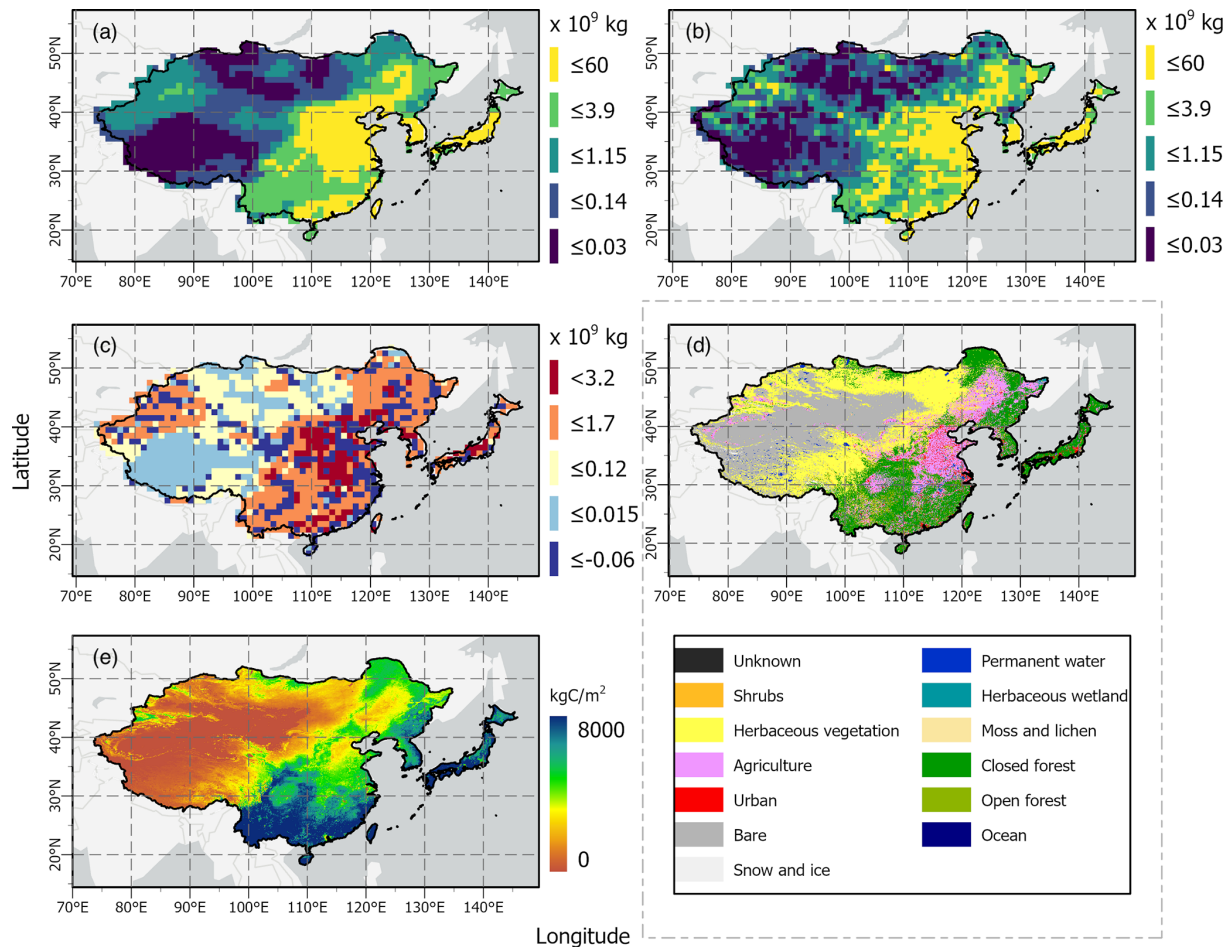


Figure 4. Spatial distribution of (a) OCO-2 XCO₂-based anthropogenic CO₂ emission estimates for 2019, (b) actual ODIAC emissions for 2019, (c) their difference (estimated emission minus actual emission), (d) 100 m resolution land cover distribution provided by the Copernicus Global Land Service over East Asia, and (e) the spatial distribution of NPP. (The base map was sourced from OpenStreetMap.)

described in Sect. 2.2. The GRNN model was then applied to 2019 annually averaged XCO₂ anomalies and NPP datasets to predict the CO₂ emissions with the same unit as the ODIAC CO₂ emissions. The analyses were carried out separately over East and West Asia. Figure 4a and b show the estimated values and the ODIAC CO₂ emissions over East Asia, respectively. The results show that the estimated values and the inventory CO₂ emissions exhibit nearly the same spatial distribution pattern. The eastern part of the region shows higher CO₂ emissions, and the western and northern parts, in particular the Tibetan Plateau and Mongolia, show the minimum CO₂ emissions. The pattern is also similar to the XCO₂ anomalies distribution over East Asia (Fig. 3d). The estimated CO₂ emissions have a relatively smoother distribution pattern compared with the ODIAC CO₂ emissions, which might be due to the interpolation of the OCO-2 dataset. Figure 4c shows the difference between the estimated and the inventory CO₂ emissions over East Asia. The estimated CO₂ emissions are generally overestimated rela-

tive to the ODIAC CO₂ emissions; however, the emissions are underestimated over some parts of the region as well. Figure 4d shows the land cover distribution of East Asia provided by the Copernicus Global Land Service (Buchhorn et al., 2020). The predicted CO₂ emissions are overestimated over most of the regional parts; however, this overestimation is more significant over agricultural areas that are located near high-density regions, e.g., eastern China. Eastern China, Japan, and Korea are known to be among the regions with the highest CO₂ emissions, and this underestimation over the agricultural areas might be caused by the nearby CO₂ emission sources which raise the CO₂ concentration of the nearby areas through atmospheric transport. Previous studies have demonstrated that the concentration of atmospheric CO₂ is influenced by atmospheric transport (Cao et al., 2017; Kumar et al., 2014). The areas where the predicted CO₂ emissions are underestimated are covered by agriculture, forest, and vegetation. This underestimation of the predicted CO₂ emissions over these areas indicates the presence of uncertainties

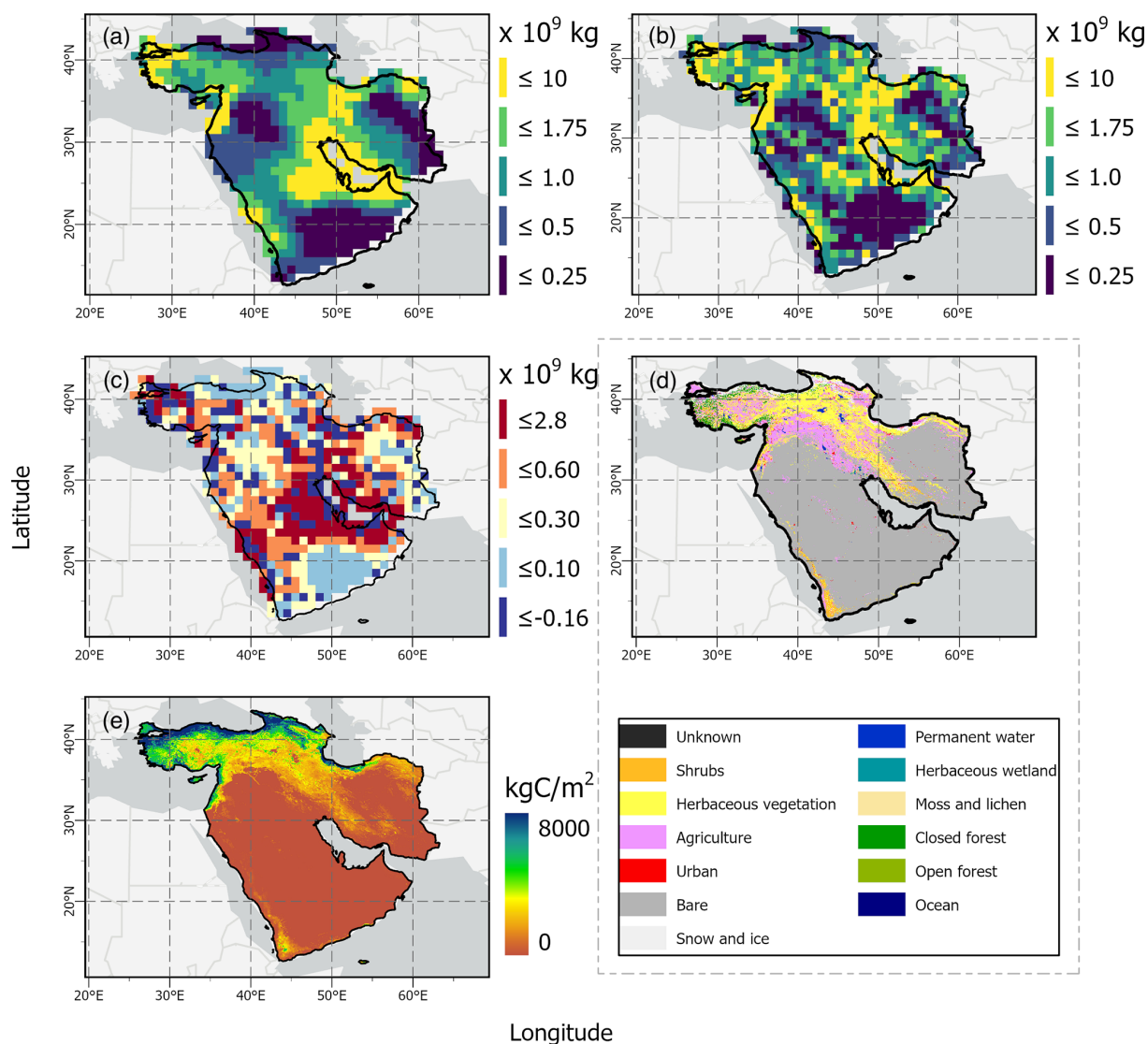


Figure 5. Spatial distribution of (a) OCO-2 XCO₂-based anthropogenic CO₂ emission estimates for 2019, (b) actual ODIAC emissions for 2019, (c) their difference (estimated emission minus actual emission), (d) 100 m resolution land cover distribution provided by the Copernicus Global Land Service over West Asia, and (e) the spatial distribution of NPP. (The base map was sourced from OpenStreetMap.)

in the XCO₂ anomalies that are likely to be produced by the CO₂ uptake of the biosphere which still remains in the XCO₂ anomalies. In addition, the areas where the estimated CO₂ emissions are overestimated have higher elevations. OCO-2 observations show larger uncertainties over elevated and mountainous areas, especially the Tibetan Plateau where the OCO-2 retrievals are significantly overestimated (Kong et al., 2019; Mustafa et al., 2020), and this might also have a contribution to the overestimation of estimated CO₂ emissions. The difference between the estimated and the ODIAC CO₂ emissions ranged from -0.06×10^9 to 3.2×10^9 kg, and the magnitude of difference between -1×10^9 and 1×10^9 kg accounted for 84 % of the total number of grid cells. Yang et al. (2019) estimated the CO₂ emissions using a similar

machine learning approach with GOSAT XCO₂ retrievals over China, and the differences between the estimated values and the ODIAC CO₂ emissions were between -5×10^9 and 5×10^9 kg. Moreover, the predicted results from the above-mentioned study exhibited less CO₂ emissions overall relative to the ODIAC emissions, contradicting our results. Our study showed better results, which may be due to the fact that (i) we improved the predictive model with the addition of an NPP dataset (Fig. 4e), (ii) we utilized the higher-resolution XCO₂ retrievals provided by OCO-2, and (iii) we incorporated the OCO-2 XCO₂ retrievals processed using the latest version of the retrieval algorithm. The newer version of the ACOS L2FP retrieval algorithm has improved the quan-

tity and the quality of the satellite-based observations (Taylor et al., 2021).

Figure 5a and b show [TS14](#) the spatial distribution of satellite-based estimated CO₂ emissions and the actual ODIAC CO₂ emissions over West Asia, respectively. The spatial distribution pattern of both the estimated and the original CO₂ emissions is similar with some differences in their magnitudes. CO₂ emissions in the eastern parts are relatively larger compared with other parts of the region. Figure 5c shows the difference between the estimated values and the ODIAC CO₂ emissions. The satellite-based estimated CO₂ emissions are generally overestimated compared with the actual ODIAC CO₂ emissions. The estimated CO₂ emissions are notably larger over Iran and Saudi Arabia. Figure 5d shows the land cover distribution of West Asia. It can be seen that the predicted CO₂ emissions are overestimated over the areas that are covered by either urban settlements or bare land. The overestimation of estimated CO₂ over these areas is likely to be caused by atmospheric transportation that influences the spatial distribution of atmospheric CO₂ (Cao et al., 2017). Moreover, a large part of West Asia is covered by deserts, and these deserts observe a notably lower number of OCO-2 retrievals (Fig. 3a). The overestimation of the predicted CO₂ emissions over the largest desert of the region, the Rub' al Kahli, located in southern parts is likely to be caused by the uncertainties in the satellite-based XCO₂ anomalies, and these uncertainties are likely to be produced due to a lower number of OCO-2 retrievals. In addition, a previous study also indicated that the ACOS XCO₂ retrieval algorithm showed uncertainties over deserts (Bie et al., 2018). Similar to East Asia, the predicted CO₂ emissions over West Asia are also underestimated over areas that are covered by agriculture or vegetation, and this underestimation might be due to the presence of CO₂ uptake by the biosphere in the XCO₂ anomalies calculated using the satellite-based retrievals. The difference between the estimated values and the ODIAC CO₂ emissions ranged from -0.16×10^9 to 2.8×10^9 kg, and the magnitude of the difference between -1×10^9 and 1×10^9 kg accounted for 88 % of the total number of grid cell. [CE12](#)

3.3 Correlation analysis between OCO-2 XCO₂ anomalies and ODIAC emissions

Figure 6 shows the correlation analysis between the ODIAC CO₂ emissions and the XCO₂ anomalies calculated using the OCO-2 retrievals over East and West Asia. Yang et al. (2019) found that the cluster of XCO₂ changes derived from satellite-based observations showed a better and more significant correlation with the CO₂ emissions relative to a single grid of XCO₂, which might have been due to the fact that the atmospheric CO₂ measurement is an instantaneous snapshot of the realistic atmosphere (Liu et al., 2015) [CE13](#). For the correlation analysis, we segmented the ODIAC emissions, which were binned every 0.3 ty^{-1} of lgE using mean

emissions calculated from annual emissions during 2015–2019, and then carried out an analysis between the mean of the emissions and the mean of the XCO₂ anomalies within the binned regions. The results showed a positive and significant correlation between the two datasets. Figure 6a and b show [TS15](#) the spatial distribution of segmented ODIAC emissions over East Asia and the scatterplot between the mean of the emissions and the mean of the XCO₂ anomalies, respectively. The two datasets show a positive and significant correlation with a determined coefficient (R^2) of 0.81. The spatial distribution of segmented ODIAC emissions over West Asia and the scatterplot between the mean of the emissions and the mean of the XCO₂ anomalies for this region are shown in Fig. 6c and d, respectively. The two datasets showed a good correlation with a determined coefficient (R^2) of 0.60. Several studies have correlated satellite-based XCO₂ anomalies with CO₂ emissions (Fu et al., 2019; Shekhar et al., 2020). Yang et al. (2019) performed a correlation analysis between the GOSAT-based XCO₂ anomalies and the ODIAC CO₂ emissions over China and found a significant correlation with a determined coefficient (R^2) of 0.82 which increased up to 0.95 if the analysis was carried out with higher CO₂ emission values. In our study, the correlation between the CO₂ emissions and XCO₂ anomalies is relatively low for West Asia, which might be due to the uncertainties in the OCO-2 retrievals. A large part of West Asia is covered by deserts, and, as previously stated, Bie et al. (2018) reported that the ACOS XCO₂ retrieval algorithm showed uncertainties over deserts.

4 Summary and conclusions

In this study, anthropogenic CO₂ emissions were estimated using satellite datasets and employing a neural-network-based method. The study was carried out using ODIAC CO₂ emissions, OCO-2 XCO₂, and MODIS NPP datasets from 2015 to 2019. To remove the CO₂ seasonal variability and the large background concentration from the OCO-2 XCO₂ retrievals, XCO₂ anomalies were calculated for each year. A GRNN model was then built; XCO₂ anomalies, NPP, and CO₂ emissions from 2015 to 2018 were used as a training dataset; and, finally, CO₂ emissions were predicted for 2019 based on the NPP and XCO₂ anomalies calculated for the same year. The analyses were carried out separately over East and West Asia. The satellite-based estimated values and the ODIAC CO₂ emission datasets were compared, and both of the datasets showed good agreement in terms of spatial distribution. The estimated CO₂ emissions showed better results over East Asia compared with West Asia, which might be due to the uncertainties in the XCO₂ retrievals: previous studies have reported that the ACOS XCO₂ retrieval algorithm produced uncertainties over deserts. The predicted CO₂ emissions were generally overestimated, and this overestimation was larger over the areas that were closer to the high-density urban regions. The overestimations might be

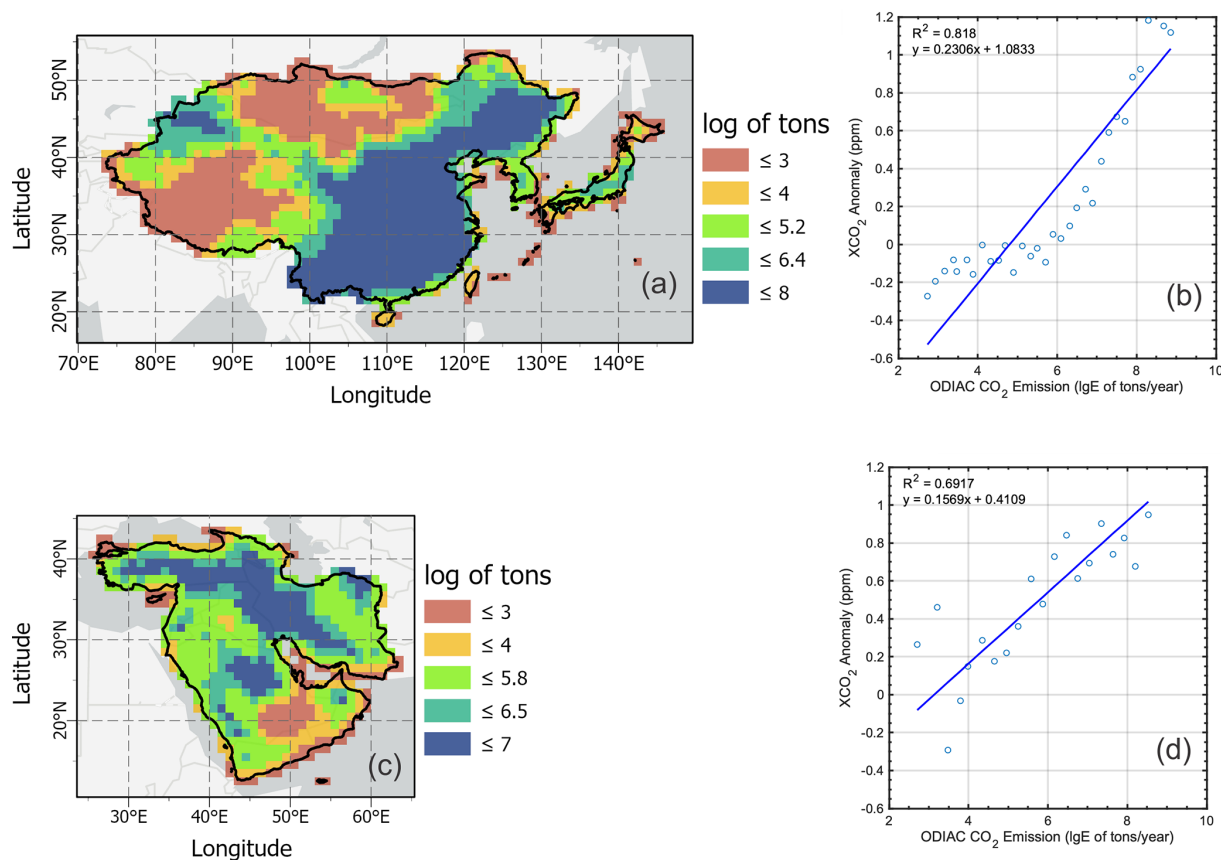


Figure 6. The spatial distribution of segmented ODIAC emissions, where the data are binned in 0.3 t yr^{-1} [TS13](#) of lgE [CE11](#) bins using the mean emission calculated from the annual emissions from 2015 to 2019 over (a) East Asia and (c) West Asia. The correlation between mean ODIAC CO₂ emissions and mean XCO₂ anomalies calculated from annual XCO₂ from 2015 to 2018 for (b) East Asia and (d) West Asia. (The base map was sourced from OpenStreetMap.)

due to the nearby high-emission CO₂ sources that raised the XCO₂ concentration due to the effects of atmospheric transport. The satellite-based estimated CO₂ emissions were underestimated over some parts of the regions, mostly areas covered by agricultural land and vegetation; this was likely caused by the uncertainties in the calculated XCO₂ anomalies, and these uncertainties were produced due to the presence of the CO₂ uptake of the biosphere. We compared our results with a previous study carried out using a similar predictive model incorporating GOSAT XCO₂ retrievals [CE14](#). The referenced study generally underestimated the predicted CO₂ emissions, with larger differences relative to ODIAC CO₂ emissions, contradicting our results. Our study showed relatively better results, which might be due to several reasons: (i) we improved the predictive model with the addition of an NPP dataset, (ii) we incorporated OCO-2 XCO₂ retrievals that have a higher spatial resolution compared with the GOSAT XCO₂ retrievals, and (iii) we used a XCO₂ product processed using the latest version of the ACOS L2FP retrieval algorithm. The newer version of the algorithm has improved the quantity and the quality of the XCO₂ retrievals.

Moreover, correlation analysis was also carried out between the ODIAC CO₂ emissions and the OCO-2 XCO₂ anomalies, and the results were significant with R^2 values of 0.81 and 0.60 over East and West Asia, respectively. These results were in agreement with the previous studies.

The results from our study suggest that CO₂ emissions can be estimated using observations obtained from CO₂ monitoring satellites. Currently, several satellites are orbiting the Earth and are dedicated to monitoring atmospheric CO₂. Joint utilization of the observations from the old and the latest satellites, such as OCO-3, GOSAT-2, and TanSAT, might reduce the spatiotemporal gaps and uncertainties. In future studies, we intend to improve the GRNN model via the addition of CO₂ uptake datasets and the joint utilization of multi-sensor data.

Data availability. The OCO-2 Level 2 XCO₂ product is available from <https://earthdata.nasa.gov/TS16>, and the ODIAC CO₂ emission dataset is available from <http://db.cger.nies.go.jp/dataset/ODIAC/> (CGER, 2021) [TS17](#).

Author contributions. FM carried out the analysis under the supervision of LB, with input and support from QW, NY, MS, MB, RWA, and RI. FM wrote the original article with feedback from all the co-authors.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The authors acknowledge the efforts of NASA with respect to providing the OCO-2 data products. The authors are also thankful to the National Institute of Environmental Studies (NIES) for providing the ODIAC CO₂ emission dataset. The lead author (Farhan Mustafa) is thankful to Thomas E. Taylor (Colorado State University, USA) for providing help with summarizing the evolution of the ACOS L2FP retrieval algorithm.

Financial support. This research has been supported by the National Natural Science Foundation of China (grant no. 41675133). [TS18](#)

Review statement. This paper was edited by Dmitry Efremenko and reviewed by two anonymous referees.

References

- [TS19](#) Bao, Z., Zhang, X., Yue, T., Zhang, L., Wang, Z., Jiao, Y., Bai, W., and Meng, X.: Retrieval and Validation of XCO₂ from TanSat Target Mode Observations in Beijing, *Remote Sens.-Basel*, 12, [TS20](#), <https://doi.org/10.3390/rs12183063>, 2020.
- Bie, N., Lei, L., Zeng, Z., Cai, B., Yang, S., He, Z., Wu, C., and Nassar, R.: Regional uncertainty of GOSAT XCO₂ retrievals in China: quantification and attribution, *Atmos. Meas. Tech.*, 11, 1251–1272, <https://doi.org/10.5194/amt-11-1251-2018>, 2018.
- Boden, T. A., Andres, R. J., and Marland, G.: Global, regional, and national fossil-fuel CO₂ emissions (1751–2014) (v. 2017), Environmental System Science Data Infrastructure for a Virtual Ecosystem (ESS-DIVE) (United States), Carbon Dioxide Information Analysis Center (CDIAC), Oak Ridge National Laboratory (ORNL), Oak Ridge, TN, USA [TS21](#), 2017.
- Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO₂ emissions from space and related applications, *Atmos. Meas. Tech.*, 3, 781–811, <https://doi.org/10.5194/amt-3-781-2010>, 2010.
- Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., and Smets, B.: Copernicus Global Land Cover Layers—Collection 2, *Remote Sens.-Basel*, 12, 1044, <https://doi.org/10.3390/rs12061044>, 2020.
- Cao, L., Chen, X., Zhang, C., Kurban, A., Yuan, X., Pan, T., and de Maeyer, P.: The Temporal and Spatial Distributions of the Near-Surface CO₂ Concentrations in Central Asia and Analysis of Their Controlling Factors, *Atmosphere*, 8, 85, <https://doi.org/10.3390/atmos8050085>, 2017.
- Center for Global Environmental Research (CGER): ODIAC Fossil Fuel Emission Dataset, CGER [data set], available at: <http://db.cger.nies.go.jp/dataset/ODIAC/>, last access: 2021. [TS22](#)
- Chen, Z., Ye, X., and Huang, P. [TS23](#): Estimating Carbon Dioxide (CO₂) Emissions from Reservoirs Using Artificial Neural Networks, *Water*, 10, 26, <https://doi.org/10.3390/w10010026>, 2018.
- Crisp, D.: Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2), in: Proc. SPIE, 8 September 2015 [TS24](#), <https://doi.org/10.1117/12.2187291>, 2015.
- Crisp, D., Fisher, B. M., O'Dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Mandrake, L., McDuffie, J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O'Brien, D. M., Oyafuso, F., Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T. E., Thompson, D. R., Wennberg, P. O., Wunch, D., and Yung, Y. L.: The ACOS CO₂ retrieval algorithm – Part II: Global XCO₂ data characterization, *Atmos. Meas. Tech.*, 5, 687–707, <https://doi.org/10.5194/amt-5-687-2012>, 2012.
- Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A. M., Oyafuso, F. A., Frankenberg, C., O'Dell, C. W., Bruegge, C. J., Doran, G. B., Eldering, A., Fisher, B. M., Fu, D., Gunson, M. R., Mandrake, L., Osterman, G. B., Schwandner, F. M., Sun, K., Taylor, T. E., Wennberg, P. O., and Wunch, D.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, *Atmos. Meas. Tech.*, 10, 59–81, <https://doi.org/10.5194/amt-10-59-2017>, 2017.
- Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevalier, F., Liu, J., Deng, F., Feng, L., McKain, K., Chatterjee, A., Miller, J. B., Stephens, B. B., Eldering, A., Crisp, D., Schimel, D., Nassar, R., O'Dell, C. W., Oda, T., Sweeney, C., Palmer, P. I., and Jones, D. B. A.: The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network, *Atmos. Chem. Phys.*, 19, 9797–9831, <https://doi.org/10.5194/acp-19-9797-2019>, 2019.
- EDGAR [TS25](#): Emission Database for Global Atmospheric Research (EDGAR v4.3.2), European Commission, 2017.
- Fu, P., Xie, Y., Moore, C. E., Myint, S. W., and Bernacchi, C. J.: A Comparative Analysis of Anthropogenic CO₂ Emissions at City Level Using OCO-2 Observations: A Global Perspective, *Earths Future*, 7, 1058–1070, <https://doi.org/10.1029/2019EF001282>, 2019.
- Guan, D., Liu, Z., Geng, Y., Lindner, S., and Hubacek, K.: The gigatonne gap in China's carbon dioxide inventories, *Nat. Clim. Change*, 2, 672–675, <https://doi.org/10.1038/nclimate1560>, 2012.
- Gurney, K. R., Mendoza, D. L., Zhou, Y., Fischer, M. L., Miller, C. C., Geethakumar, S., and de la Rue du Can, S.: High Resolution Fossil Fuel Combustion CO₂ Emission Fluxes for the United States, *Environ. Sci. Technol.*, 43, 5535–5541, <https://doi.org/10.1021/es900806c>, 2009.

- Hakkaraianen, J., Ialongo, I., and Tamminen, J.: Direct space-based observations of anthropogenic CO₂ emission areas from OCO-2, *Geophys. Res. Lett.*, 43, [TS26](https://doi.org/10.1002/2016GL070885), <https://doi.org/10.1002/2016GL070885>, 2016.
- 5 Hakkaraianen, J., Ialongo, I., Maksyutov, S., and Crisp, D.: Analysis of Four Years of Global XCO₂ Anomalies as Seen by Orbiting Carbon Observatory-2, *Remote Sens.-Basel*, 11, 850, <https://doi.org/10.3390/rs11070850>, 2019.
- Hoegh-Guldberg, O., Jacob, D., Bindi, M., Brown, S., Camilloni, I.,
10 Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., and Guiot, J.: Impacts of 1.5 °C global warming on natural and human systems, *Global warming of 1.5 °C. An IPCC Special Report*, [TS27](https://doi.org/10.1017/97810161172061), 2018.
- Hong, X., Zhang, P., Bi, Y., Liu, C., Sun, Y., Wang, W., Chen, Z., Yin, H., Zhang, C., Tian, Y., and Liu, J.:
15 Retrieval of Global Carbon Dioxide From TanSat Satellite and Comprehensive Validation With TCCON Measurements and Satellite Observations, *IEEE T. Geosci. Remote*, 1–16, <https://doi.org/10.1109/TGRS.2021.3066623>, 2021.
- Hutchins, M. G., Colby, J. D., Marland, G., and Marland, E.: A comparison of five high-resolution spatially-explicit, fossil-fuel, carbon dioxide emission inventories for the United States, *Mitig. Adapt. Strat. Gl.*, 22, 947–972, <https://doi.org/10.1007/s11027-016-9709-9>, 2017.
- Jalil, S. A.: Carbon Dioxide Emission in the Middle East and
25 North African (MENA) Region: A Dynamic Panel Data Study, *Journal of Emerging Economies and Islamic Research*, 2, 5, <https://doi.org/10.24191/jeeir.v2i3.9629>, 2014.
- Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, *Atmos. Chem. Phys.*, 15, 11411–11432,
30 <https://doi.org/10.5194/acp-15-11411-2015>, 2015.
- Keppel-Aleks, G., Wennberg, P. O., O'Dell, C. W., and Wunch, D.: Towards constraints on fossil fuel emissions from total column carbon dioxide, *Atmos. Chem. Phys.*, 13, 4349–4357, <https://doi.org/10.5194/acp-13-4349-2013>, 2013.
- 40 Kiel, M., O'Dell, C. W., Fisher, B., Eldering, A., Nassar, R., MacDonald, C. G., and Wennberg, P. O.: How bias correction goes wrong: measurement of XCO₂ affected by erroneous surface pressure estimates, *Atmos. Meas. Tech.*, 12, 2241–2259, <https://doi.org/10.5194/amt-12-2241-2019>, 2019.
- 45 Kong, Y., Chen, B., and Measho, S.: Spatio-temporal consistency evaluation of XCO₂ retrievals from GOSAT and OCO-2 based on TCCON and model data for joint utilization in carbon cycle research, *Atmosphere*, 10, 1–23, <https://doi.org/10.3390/atmos10070354>, 2019.
- 50 Kononov, I. B., Berezin, E. V., Ciais, P., Broquet, G., Zhuravlev, R. V., and Janssens-Maenhout, G.: Estimation of fossil-fuel CO₂ emissions using satellite measurements of “proxy” species, *Atmos. Chem. Phys.*, 16, 13509–13540, <https://doi.org/10.5194/acp-16-13509-2016>, 2016.
- 55 Korsbakken, J. I., Peters, G. P., and Andrew, R. M.: Uncertainties around reductions in China’s coal use and CO₂ emissions, *Nat. Clim. Change*, 6, 687–690, <https://doi.org/10.1038/nclimate2963>, 2016.
- Kumar, K. R., Revadekar, J. V., and Tiwari, Y. K.: AIRS retrieved CO₂ and its association with climatic parameters over
60 india during 2004–2011, *Sci. Total Environ.*, 476–477, 79–89, <https://doi.org/10.1016/j.scitotenv.2013.12.118>, 2014.
- Lamminpää, O., Hobbs, J., Brynjarsdóttir, J., Laine, M., Braverman, A., Lindqvist, H., and Tamminen, J.: Accelerated MCMC for Satellite-Based Measurements of Atmospheric CO₂,
65 *Remote Sens.-Basel*, 11, 2061, <https://doi.org/10.3390/rs11172061>, 2019.
- Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O’Keefe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Razli-
70 vanov, I., Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban CO₂ emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), *J. Geophys. Res.-Atmos.*, 121, 5213–5236, <https://doi.org/10.1002/2015JD024473>, 2016.
- 75 Liu, D., Lei, L., Guo, L., and Zeng, Z.-C.: A Cluster of CO₂ Change Characteristics with GOSAT Observations for Viewing the Spatial Pattern of CO₂ Emission and Absorption, *Atmosphere*, 6, 1695–1713, <https://doi.org/10.3390/atmos6111695>, 2015.
- Liu, Y., Wang, J., Yao, L., Chen, X., Cai, Z., Yang, D., Yin, Z., Gu, S., Tian, L., Lu, N., and Lyu, D.: The TanSat mission: preliminary global observations, *Sci. Bull.*, 63, 1200–1207, <https://doi.org/10.1016/j.scib.2018.08.004>, 2018.
- Maksyutov, S., Takagi, H., Valsala, V. K., Saito, M., Oda, T., Saeki, T., Belikov, D. A., Saito, R., Ito, A., Yoshida, Y., Morino, I., Uchino, O., Andres, R. J., and Yokota, T.: Regional CO₂ flux estimates for 2009–2010 based on GOSAT and ground-based CO₂ observations, *Atmos. Chem. Phys.*, 13, 9351–9373, <https://doi.org/10.5194/acp-13-9351-2013>, 2013.
- Matsunaga, T., Morino, I., Yoshida, Y., Saito, M., Noda, H.,
90 Ohyama, H., Niwa, Y., Yashiro, H., Kamei, A., Kawazoe, F., Saeki, T., Ishihara, Y., Imasu, R., Teruyuki, N., Nakajima, T. S. Y., Saitoh, N., and Hashimoto, M.: Early Results of GOSAT-2 Level 2 Products, in: *AGU Fall Meeting Abstracts*, December 2019, abstract #A52H-02 [TS28](https://doi.org/10.1016/j.scib.2018.08.004), 2019.
- 95 Mustafa, F., Bu, L., Wang, Q., Ali, M. A., Bilal, M., Shahzaman, M., and Qiu, Z.: Multi-year comparison of CO₂ concentration from NOAA carbon tracker reanalysis model with data from GOSAT and OCO-2 over Asia, *Remote Sens.-Basel*, 12, [TS29](https://doi.org/10.3390/RS12152498), <https://doi.org/10.3390/RS12152498>, 2020.
- 100 Mustafa, F., Bu, L., Wang, Q., Yao, N., Shahzaman, M., Bilal, M., Aslam, R. W., and Iqbal, R.: Neural Network Based Estimation of Regional Scale Anthropogenic CO₂ Emissions Using OCO-2 Dataset Over East and West Asia, *Atmos. Meas. Tech. Discuss.* [preprint], <https://doi.org/10.5194/amt-2021-222>, in review, 2021a.
- Mustafa, F., Wang, H., Bu, L., Wang, Q., Shahzaman, M., Bilal, M., Zhou, M., Iqbal, R., Aslam, R. W., Ali, M. A., and Qiu, Z.: Validation of GOSAT and OCO-2 against In Situ Aircraft Measurements and Comparison with CarbonTracker and GEOS-
110 Chem over Qinhuangdao, China, *Remote Sens.-Basel*, 13, 899, <https://doi.org/10.3390/rs13050899>, 2021b.
- O’Dell, C. W., Connor, B., Bösch, H., O’Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I.,
115 Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO₂ retrieval algorithm – Part 1: Description and

- validation against synthetic observations, *Atmos. Meas. Tech.*, 5, 99–121, <https://doi.org/10.5194/amt-5-99-2012>, 2012.
- O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, *Atmos. Meas. Tech.*, 11, 6539–6576, <https://doi.org/10.5194/amt-11-6539-2018>, 2018.
- Oda, T.: ODIAC Fossil Fuel CO₂ Emissions Dataset, [TS30](https://doi.org/10.5194/amt-11-6539-2018), 2015. [TS31](https://doi.org/10.5194/amt-11-6539-2018)
- Oda, T. and Maksyutov, S.: ODIAC fossil fuel CO₂ emissions dataset (version name: ODIAC2016), Center for Global Environmental Research, National Institute for Environmental Studies [TS32](https://doi.org/10.17595/20170411.001), <https://doi.org/10.17595/20170411.001>, 2015.
- Oda, T., Maksyutov, S., and Andres, R. J.: The Open-source Data Inventory for Anthropogenic CO₂, version 2016 (ODIAC2016): a global monthly fossil fuel CO₂ gridded emissions data product for tracer transport simulations and surface flux inversions, *Earth Syst. Sci. Data*, 10, 87–107, <https://doi.org/10.5194/essd-10-87-2018>, 2018.
- Olivier, J. G. J., Van Aardenne, J. A., Dentener, F. J., Pagliari, V., Ganzeveld, L. N., and Peters, J. A. H. W.: Recent trends in global greenhouse gas emissions: regional trends 1970–2000 and spatial distribution of key sources in 2000, *Environm. Sci.*, 2, 81–99, <https://doi.org/10.1080/15693430500400345>, 2005.
- Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO₂ inversions, *Biogeosciences*, 10, 6699–6720, <https://doi.org/10.5194/bg-10-6699-2013>, 2013.
- Rao, M.: Machine Learning in Estimating CO₂ Emissions from Electricity Generation, in: Uncertainty Management in Engineering – Topics in Pollution Prevention and Controls [Working Title], IntechOpen [TS33](https://doi.org/10.5194/bg-10-6699-2013), 2021.
- Rooki, R.: Application of general regression neural network (GRNN) for indirect measuring pressure loss of Herschel-Bulkley drilling fluids in oil drilling, *Measurement*, 85, 184–191, <https://doi.org/10.1016/j.measurement.2016.02.037>, 2016.
- Shan, Y., Guan, D., Zheng, H., Ou, J., Li, Y., Meng, J., Mi, Z., Liu, Z., and Zhang, Q.: China CO₂ emission accounts 1997–2015, *Scientific Data*, [TS34](https://doi.org/10.1016/j.measurement.2016.02.037), 1–14, 1997.
- Shekhar, A., Chen, J., Paetzold, J. C., Dietrich, F., Zhao, X., Bhattacharjee, S., Ruisinger, V., and Wofsy, S. C.: Anthropogenic CO₂ emissions assessment of Nile Delta using XCO₂ and SIF data from OCO-2 satellite, *Environ. Res. Lett.*, 15, 095010, <https://doi.org/10.1088/1748-9326/ab9cfe>, 2020.
- Specht, D. F.: A general regression neural network, *IEEE T. Neural Networ.*, 2, 568–576, <https://doi.org/10.1109/72.97934>, 1991.
- Takagi, H., Saeki, T., Oda, T., Saito, M., Valsala, V., Belikov, D., Saito, R., Yoshida, Y., Morino, I., Uchino, O., Andres, R. J., Yokota, T., and Maksyutov, S.: On the Benefit of GOSAT Observations to the Estimation of Regional CO₂ Fluxes, *SOLA*, 7, 161–164, <https://doi.org/10.2151/sola.2011-041>, 2011.
- Taylor, T. E., Eldering, A., Merrelli, A., Kiel, M., Somkuti, P., Cheng, C., Rosenberg, R., Fisher, B., Crisp, D., Basilio, R., Bennett, M., Cervantes, D., Chang, A., Dang, L., Frankenberg, C., Haemmerle, V. R., Keller, G. R., Kurosu, T., Laughner, J. L., Lee, R., Marchetti, Y., Nelson, R. R., O'Dell, C. W., Osterman, G., Pavlick, R., Roehl, C., Schneider, R., Spiers, G., To, C., Wells, C., Wennberg, P. O., Yelamanchili, A., and Yu, S.: OCO-3 early mission operations and initial (vEarly) XCO₂ and SIF retrievals, *Remote Sen. Environ.*, 251, 112032, <https://doi.org/10.1016/j.rse.2020.112032>, 2020.
- Taylor, T. E., O'Dell, C. W., Crisp, D., Kuze, A., Lindqvist, H., Wennberg, P. O., Chatterjee, A., Gunson, M., Eldering, A., Fisher, B., Kiel, M., Nelson, R. R., Merrelli, A., Osterman, G., Chevallier, F., Palmer, P. I., Feng, L., Deutscher, N. M., Dubey, M. K., Feist, D. G., Garcia, O. E., Griffith, D., Hase, F., Iraci, L. T., Kivi, R., Liu, C., De Mazière, M., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Pollard, D. F., Rettinger, M., Roehl, C. M., Schneider, M., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Tê, Y., Velazco, V. A., Vrekoussis, M., Warneke, T., and Wunch, D.: An eleven year record of XCO₂ estimates derived from GOSAT measurements using the NASA ACOS version 9 retrieval algorithm, *Earth Syst. Sci. Data Discuss.* [preprint], <https://doi.org/10.5194/essd-2021-247>, in review, 2021.
- UNFCCC [TS35](https://doi.org/10.5194/essd-2021-247): Paris Agreement, [TS36](https://doi.org/10.5194/essd-2021-247), 2015.
- Wang, J. S., Oda, T., Kawa, S. R., Strode, S. A., Baker, D. F., Ott, L. E., and Pawson, S.: The impacts of fossil fuel emission uncertainties and accounting for 3-D chemical CO₂ production on inverse natural carbon flux estimates from satellite and in situ data, *Environ. Res. Lett.*, 15, 085002, <https://doi.org/10.1088/1748-9326/ab9795>, 2020.
- Wang, R., Tao, S., Ciais, P., Shen, H. Z., Huang, Y., Chen, H., Shen, G. F., Wang, B., Li, W., Zhang, Y. Y., Lu, Y., Zhu, D., Chen, Y. C., Liu, X. P., Wang, W. T., Wang, X. L., Liu, W. X., Li, B. G., and Piao, S. L.: High-resolution mapping of combustion processes and implications for CO₂ emissions, *Atmos. Chem. Phys.*, 13, 5189–5203, <https://doi.org/10.5194/acp-13-5189-2013>, 2013.
- Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Notholt, J., Warneke, T., Petri, C., De Mazière, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO₂ measurements with TCCON, *Atmos. Meas. Tech.*, 10, 2209–2238, <https://doi.org/10.5194/amt-10-2209-2017>, 2017.
- Yang, D., Liu, Y., Cai, Z., Chen, X., Yao, L., and Lu, D.: First Global Carbon Dioxide Maps Produced from TanSat Measurements, *Adv. Atmos. Sci.*, 35, 621–623, <https://doi.org/10.1007/s00376-018-7312-6>, 2018.
- Yang, S., Lei, L., Zeng, Z., He, Z., and Zhong, H.: An Assessment of Anthropogenic CO₂ Emissions by Satellite-Based Observations in China, *Sensors-Basel*, 19, 1118, <https://doi.org/10.3390/s19051118>, 2019.

Remarks from the language copy-editor

- CE1** Please note that this manuscript has undergone copy-editing according to the standards of American English.
- CE2** Please check that the meaning of your sentence is intact.
- CE3** Please note that, according to our standards, we always define abbreviations/acronyms at the first instance of use (in the abstract as well as in the rest of the text). Please check the other edits made in such cases throughout the paper and let me know if further changes are required. Thank you.
- CE4** Please confirm the change.
- CE5** Do you mean “single-sounding precision”?
- CE6** Do you mean “1 km × 1 km”?
- CE7** Do you mean “1° × 1°”?
- CE8** Do you mean “the mean of each grid cell”?
- CE9** Please check that the meaning of your sentence is intact.
- CE10** Please confirm the change here and again in the caption of Fig. 5.
- CE11** What does “lgE” refer to? This is currently unclear. Please advise.
- CE12** This seems to be repeated information from page 8. Is it required again here? Please check.
- CE13** Please check that the meaning of your sentence is intact.
- CE14** Please add a citation for the study that you are referring to here.

Remarks from the typesetter

- TS1** Please check.
- TS2** Please check and confirm the change.
- TS3** Please check and confirm the change.
- TS4** Please confirm or provide a different short running title.
- TS5** Reference missing from reference list.
- TS6** Please provide date of last access.
- TS7** Please provide date of last access.
- TS8** Please provide date of last access.
- TS9** Please confirm.
- TS10** The composition of Fig. 3 has been adjusted to our standards. Please also note the language edits to Figs. 1 and 3.
- TS11** Please check.
- TS12** Please check.
- TS13** Please note that units have been changed to exponential format throughout the text. Please check all instances.
- TS14** Please check.
- TS15** Please check.
- TS16** Please provide a direct link to the dataset and, if possible, a DOI instead of a URL. In any case, please provide a reference list entry including creators, title, and date of last access.
- TS17** Please confirm citation.
- TS18** Please note that the funding information has been added to this paper. Please check if it is correct. Please also double-check your acknowledgements to see whether repeated information can be removed or changed accordingly. Thanks.
- TS19** Please ensure that any datasets and software codes used in this work are properly cited in the text and included in this reference list. Thereby, please keep our reference style in mind, including creators, titles, publisher/repository, persistent identifier, and publication year. Regarding the publisher/repository, please add “[dataset]” or “[code]” to the entry (e.g. Zenodo [code]).
- TS20** Please provide article number or page range.
- TS21** Please check and confirm addition of all necessary information taken from the original document.
- TS22** Please provide date of last access and confirm the year. Please also confirm the entire reference list entry.
- TS23** Please check and confirm the change of all author names and initials based on the original document/DOI.
- TS24** Please add the title and venue of the event.
- TS25** Please check and confirm the change.
- TS26** Please provide article number or page range.
- TS27** Please add the name of publisher/publishing institution and place of publication.
- TS28** Please check and confirm the change/addition

- TS29** Please provide article number or page range.
- TS30** Please add the name of publisher/publishing institution and place of publication.
- TS31** Reference not cited in text.
- TS32** Please add the place of publication.
- TS33** Please add additional data to adequately locate the source.
- TS34** Please add the name of publisher/publishing institution and place of publication.
- TS35** Please check and confirm the change.
- TS36** Please add the name of publisher/publishing institution and place of publication.