https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

Atmospheric Measurement Techniques Discussions

Neural Network Based Estimation of Regional Scale Anthropogenic CO_2 Emissions Using OCO-2 Dataset Over East and West Asia

Farhan Mustafa[1], Lingbing Bu[1], Qin Wang[1], Na Yao[1], Muhammad Shahzaman[2], Muhammad Bilal[3], Rana Waqar Aslam[4], Rashid Iqbal[5]

[1]Collaborative 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
[2]School of Atmospheric Sciences (SAS), Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China
[3]School of Marine Sciences (SMS), Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China
[4]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430079, China
[5]Department of Agronomy, Faculty of Agriculture and Environment, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan

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

Abstract. Atmospheric carbon dioxide (CO_2) is the most significant greenhouse gas and its concentration is continuously increasing mainly as a consequence of anthropogenic activities. Accurate quantification of CO_2 is critical for addressing the global challenge of climate change and designing mitigation strategies aimed at stabilizing the CO_2 emissions. Satellites provide the most effective way to monitor the concentration of CO_2 in the atmosphere. In this study, we utilized the concentration of column-averaged dry-air mole fraction of CO_2 i.e., XCO_2 retrieved from a CO_2 monitoring satellite, the Orbiting Carbon Observatory 2 (OCO-2) to estimate the anthropogenic CO_2 emissions using Generalized Regression Neural Network over East and West Asia. OCO-2 XCO_2 and the Open-Data Inventory for Anthropogenic Carbon dioxide (ODIAC) CO_2 emission datasets for a period of 5 years (2015-2019) were used in this study. The annual XCO_2 anomalies were calculated from the OCO-2 retrievals for each year to remove the larger background CO_2 concentrations and seasonal variabilities. Then the XCO_2 anomaly and ODIAC emission datasets from 2015 to 2018 were used to train the GRNN model, and finally, the anthropogenic CO_2 emissions were estimated for 2019 based on the XCO_2 anomalies derived for the same year. The XCO_2-based estimated and the ODIAC actual CO_2 emissions were compared and the results showed a good agreement in terms of spatial distribution. The CO_2 emissions were estimated separately over East and West Asia. In addition, correlations between the ODIAC emissions and XCO_2 anomalies were also determined separately for East and West Asia, and East Asia exhibited relatively better results. The results showed that satellite-based XCO_2 retrievals can be used to estimate the regional scale anthropogenic CO_2 emissions and the accuracy of the results can be enhanced by further improvement of the GRNN model with the addition of more CO_2 emission and concentration datasets.

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

1 Introduction

Climate change is one of the greatest challenges to the future of Earth arising from global warming, which in turn is accelerated by anthropogenic emissions of greenhouse gases (Lamminpää et al., 2019). The major warming effects are caused by the atmospheric CO_2 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 CO_2 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 CO_2 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 CO_2 emissions. It is becoming increasingly important to find efficient and reliable ways to monitor the CO_2 reduction progresses and evaluation of how well specific CO_2 reduction policies are working.

Satellites provide the most effective way to monitor atmospheric CO_2 with great spatiotemporal resolutions (Mustafa et al., 2021). Several satellites such as GOSAT, GOSAT-2, OCO-2, OCO-3, and TanSAT are orbiting around the Earth and dedicatedly monitoring the atmospheric CO_2 (Crisp, 2015; Liu et al., 2018; Matsunaga et al., 2019; Taylor et al., 2020, p.3). These satellites calculate the average atmospheric CO_2 concentrations in the path of sunlight reflected by the surface through spectrometers carried onboard. OCO-2 measures the CO_2 optical depth with bands centered around 1.6 and 2.0 microns and determines O_2 optical depth with band A, which is centered around 0.76 microns (Crisp et al., 2017, p.2). The information from these bands is combined to calculate the column-averaged dry-air mole fraction of CO_2 (XCO_2) (Crisp et al., 2012; O'Dell et al., 2012). Several studies suggest that XCO_2 can be used to detect the CO_2 concentration induced by anthropogenic activities by removing the background concentration from the satellite XCO_2 retrievals. (Bovensmann et al., 2010; Hakkarainen et al., 2019, p.2; 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 of the United States and China. The XCO_2 retrievals derived from the satellite measurements show a positive correlation with the CO_2 emission inventories (Hakkarainen et al., 2016, p.2; Yang et al., 2019) which implies that these space-based observations can be used to assess the anthropogenic CO_2 emissions by enhancing the anthropogenic XCO_2 concentration.

Asia is the home to the most populous nations with the highest amounts of CO_2 emissions. East Asia, in particular, China significantly contributes to the global carbon budget and has accounted for ~30% of the overall growth in global CO_2 emissions over the past 15 years (EDGAR, 2017). This increment in the CO_2 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 CO_2 emissions per unit GDP by 60–65% relative to 2005 levels, and peak carbon emissions overall, by 2030 (UNFCC, 2015). West Asia is also a region with higher rates of anthropogenic CO_2 emissions (Mustafa et al., 2020) and some of its countries, such as Iran, Saudi Arabia, and Turkey are listed among the 10 largest CO_2 emitting nations in the world. Several studies have been carried out to estimate the CO_2 emissions using various machine learning techniques but most of them do not deal with the spatial

Atmospheric Measurement Techniques
Open Access
EGU
Discussions

distribution. Rao (2021) estimated the CO_2 emissions using Support Vector Machine (SVM). Zhonghan Chen et al. (2018) predicted the CO_2 flux emissions based on published data including latitude, age, potential net primary productivity and mean depth using Back Propagation Neural Network (BPNN) and Generalized Regression Neural Network (GRNN). Yang et al. (2019) estimated the anthropogenic CO_2 emissions using GOSAT XCO_2 retrievals over China and the results showed a good agreement between the estimated and the ODIAC CO_2 emission dataset. In this study, we proposed a method to estimate the regional scale anthropogenic CO_2 emissions using OCO-2 XCO_2 retrievals over East and West Asia. OCO-2 and ODIAC CO_2 datasets were obtained for a period of five years from January 2015 to December 2019. XCO_2 anomalies were calculated from the OCO-2 retrievals for each year, GRNN model was trained using XCO_2 anomalies and ODIAC CO_2 emissions with four years of data from 2015 to 2018 and then anthropogenic CO_2 emissions were estimated for the year 2019 based on 2019 XCO_2 anomalies. The details about the datasets and methods are provided in Section 2. The results including estimated CO_2 emissions and evaluation of these emissions, and correlation between ODIAC CO_2 emissions and XCO_2 anomalies are discussed in Section 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 CO_2 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 microns and the CO_2 bands at 1.61 and 2.06 microns (Hakkarainen et al., 2019). Information from all these bands is combined to calculate the XCO_2. More details about the instrument design, calibration approach, on-orbit performance, and measurement principles are provided in a previous study (Crisp, 2015). A comprehensive study about the validation of OCO-2 XCO_2 retrievals against the Total Carbon Column Observing Network (TCCON) CO_2 dataset reported an absolute median difference of less than 0.4 ppm and the RMS difference less than 1.5 ppm between the two datasets (Wunch et al., 2017). Similar experiments have been carried out for validation of different versions of OCO-2 XCO_2 products and the results showed that the OCO-2 dataset was consistent and reliable for atmospheric CO_2 monitoring (Kiel et al., 2019; O'Dell et al., 2018). In this study, the OCO-2 ACOS/XCO_2 version 10r Level 2 Lite product was used and the dataset was downloaded from the Earthdata platform (https://earthdata.nasa.gov).

2.1.2 ODIAC Dataset

ODIAC is a global fossil-fuel CO_2 ($FFCO_2$) emission dataset with 1 × 1 km, monthly resolution over land and 1×1 degree, annual resolution for international bunkers from the year 2000 onward (Oda et al., 2018). It shares country scale estimates with

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

Atmospheric Measurement Techniques Discussions

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 CO_2 emissions based on the population density while ODIAC incorporates power plant profiles and nighttime light observation for emission distribution (Wang et al., 2020). ODIAC shows a 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). In this study, we used the 2020 version of ODIAC emission dataset (Oda, Tomohiro, 2015).

2.2 Methods

Estimation of anthropogenic CO_2 emissions includes three major steps as shown in Figure 1. The first step includes enhancing the XCO_2 concentration influenced by anthropogenic activities, the second step is about setting up the GRNN model using XCO_2 and ODIAC datasets, and the final step is the validation of estimated CO_2 emissions against the actual ODIAC emission dataset.

CO_2 has a larger background concentration and a longer atmospheric life period compared to other greenhouse gases (Hakkarainen et al., 2019). Because of this, XCO_2 varies by nearly 2% over the seasonal cycle and from pole to pole. In addition, XCO_2 variations influenced by anthropogenic activities are also smaller on the scale of satellite sounding (2–4 km^2). Therefore, high precision is critical for accurate quantification of the XCO_2 anomalies related to anthropogenic activities. To highlight the emission areas, CO_2 seasonal variability and the large background concentrations must be removed.

To highlight the emission areas and extract the information associated with the anthropogenic activities, we used the concept of XCO_2 anomaly suggested by a previous study (Hakkarainen et al., 2019, 2016, p.2), defined as the difference between the individual XCO_2 value measured by satellite and the daily background:

$$XCO_2\ (anomaly) = XCO_2\ (individual) - XCO_2\ (daily\ background) \qquad (1)$$

This equation calculated the XCO_2 anomalies for each observation. Subtraction of daily background concentration removes the seasonal variability. Daily background concentration was obtained by calculating the daily medians. Once the anomalies were calculated, a grid with a spatial resolution of 0.5°×0.5° Longitude/Latitude was defined and the mean against each grid point was calculated for each year. The annual mean of XCO_2 (anomaly) can detrend the seasonal variation (Hakkarainen et al., 2016). The annually-averaged XCO_2 anomalies were resampled at a grid with a spatial resolution of 1°×1° Longitude/Latitude and used along with 1°×1° Longitude/Latitude ODIAC emission dataset to setup the GRNN model.

XCO_2 variations are primarily influenced by anthropogenic activities and terrestrial ecosystems, there are both linear and non-linear mapping between the XCO_2 and the emissions. We adopted the GRNN algorithm to represent the non-linear mapping between the independent variable (XCO_2 anomaly) and dependent variable (CO_2 emission) suggested by a previous study (Yang et al., 2019). The GRNN is a memory-based network that provides estimates of continuous variables and converges to underlying regression. The regression of a dependent variable on an independent variable is the computation of

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

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 Figure 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. Then the standardized values of the independent variable are 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 in the input layer all training data will have the same order of magnitudes (Yang et al., 2019).

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

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 detail about the Holdout Method is 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:

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

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 is denoting the number of training samples.

3 Results and Discussions

3.1 Spatial Distribution of XCO_2 Observations and Anomalies

The satellite-based observations are sensitive to clouds and aerosols, therefore, much of the data is discarded during the preprocessing due to the presence of cloud and aerosol content (Mustafa et al., 2021). Figures 3a and 3b show the quantity of XCO_2 retrievals from 2015 to 2019 on a spatial grid of 0.5°×0.5° Longitude/Latitude over West and East Asia, respectively.

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

Atmospheric Measurement Techniques Discussions

OCO-2 shows a good spatial coverage over East Asia, however, southern parts of the region, in particular, the Tibetan plateau has a relatively lower number of XCO_2 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 XCO_2 retrievals. In the southern parts of West Asia, a very large desert, the Rub' al Kahli is located that stretches across Saudi Arabia, Yemen, Oman, and United Arab Emirates (UAE) and often observes dust storms. The lower number of XCO_2 retrievals in these parts of the region might be due to the ACOS XCO_2 retrieval algorithm that excludes the satellite measurements with high aerosol optical depth and cloud optical thickness (Crisp et al., 2012; O'Dell et al., 2012, p.1).

Figure 3c shows the spatial distribution of five years-averaged XCO_2 anomalies calculated using the method described in section 2.2 over West Asia. The higher concentrations of XCO_2 anomalies were observed over central parts of the region that included Iran, Kuwait, Saudi Arabia, and Iraq. These countries contribute a large fraction of the world's oil. Iran and Saudi Arabia are listed among the top 10 CO_2 emitting nations and produce over 6% of the global CO_2 emissions (Jalil, 2014). In addition, Iran, Saudi Arabia, and Iraq are the major fuel consumers and contribute more than 60% of the region's total fossil fuel CO_2 emissions (Boden et al., 2017). Figure 4d shows the multiyear-averaged XCO_2 anomalies over East Asia. The eastern parts of the region including eastern China, Japan, and South Korea show the highest concentrations of XCO_2 anomalies. China's Beijing-Tianjin-Hebei area, Korea and Japan are the most populated urban regions with the highest amount of anthropogenic emissions in the world (Mustafa et al., 2020).

Figure 3e shows the monthly-averaged XCO_2 over East and West Asia. The monthly-averaged XCO_2 concentrations show seasonal fluctuations. Moreover, the XCO_2 concentrations during each month are higher than those in the same month of the previous year and it reflects that the XCO_2 concentration in the atmosphere is continuously increasing in both regions. The XCO_2 concentration starts increasing from September and reaches its maximum value in April, then it starts decreasing and reaches the minimum value in August. The decrement in its concentration from May to August is due to several reasons, primarily due to the strong photosynthesis and weak respiration rate by the plants, and this process is enhanced during the monsoon or rainy season (Mustafa et al., 2020). The increment in XCO_2 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., 2021).

3.2 Estimated CO_2 Emissions

The annually-averaged XCO_2 anomalies and ODIAC CO_2 emission datasets for a period of four years from 2015-2018 were used as a training dataset for the GRNN model built to estimate the CO_2 emissions using the method described in section 2.2. Then the GRNN model was applied to 2019 annually-averaged XCO_2 anomalies to predict the CO_2 emissions with the same unit as the ODIAC CO_2 emissions. The analyses were carried out separately over East and West Asia. Figures 4a and 4b show the estimated and the actual ODIAC CO_2 emissions over East Asia, respectively. The results show that the estimated and the actual CO_2 emissions exhibit nearly the same spatial distribution pattern. The eastern part of the region shows higher CO_2

emissions and the western and northern parts, in particular, the Tibetan plateau and Mongolia show the minimum CO_2 emissions. The pattern is also similar to XCO_2 anomalies distribution over East Asia (Figure 3d). Figure 4c shows the difference between the estimated and the actual CO_2 emissions over East Asia. The estimated CO_2 emissions show smaller magnitudes compared to the actual emissions over the eastern parts of the region, however, it has larger magnitudes relative to the actual CO_2 emissions over western parts of the region. Figure 4d shows the landcover distribution of East Asia provided by the Copernicus Global Land Services (Buchhorn et al., 2020, p.2). Most of the area where the predicted CO_2 emission is underestimated is covered by agriculture and forest. This underestimation of the CO_2 emissions predicted based on the XCO_2 anomalies indicates the presence of uncertainties in the XCO_2 anomalies that are likely to be produced by the CO_2 uptake of the biosphere which is still remaining in the XCO_2 anomalies. The areas where the predicted CO_2 emission is overestimated are covered by herbaceous vegetation and bare land. This overestimation in the predicted CO_2 emissions might be caused by the nearby CO_2 emitting sources which raise the CO_2 concentration of the nearby areas through atmospheric transport. Previous studies demonstrated that the concentration of atmospheric CO_2 was influenced by atmospheric transport (Cao et al., 2017; Kumar et al., 2014). In addition, the areas where the estimated CO_2 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 CO_2 emissions. The difference between the estimated and the actual CO_2 emissions was ranging from $10^{-3.4}$ tons to 10^{3} tons and the magnitude of difference between $10^{-1.5}$ tons to $10^{1.5}$ tons accounted for 80% of the total grids. Yang et al. (2019) estimated the CO_2 emissions by a machine learning approach using GOSAT XCO_2 retrievals over China and differences between the estimated and the ODIAC actual CO_2 emissions were between -5 Mt to 5 Mt. Our study showed better results and it might be due to the higher resolution of OCO-2 and improvements in the XCO_2 retrieval algorithm

Figures 5a and 5b show the spatial distribution of satellite-based estimated CO_2 emissions and the actual ODIAC CO_2 emissions over West Asia, respectively. The spatial distribution pattern of both the estimated and the original CO_2 emissions is similar with some differences in their magnitudes. Figure 5c shows the difference between the estimated and the actual CO_2 emissions. The satellite-based estimated CO_2 emissions are overestimated compared to the actual ODIAC CO_2 emissions over Iran and Saudi Arabia, whereas, these are underestimated over the rest of the region. Figure 5d shows the landcover distribution of West Asia. It can be seen that the predicted CO_2 emissions are overestimated over the areas that are covered by either urban settlements or bare land. The overestimation of estimated CO_2 over these areas is likely to be caused by atmospheric transportation that influences the spatial distribution of atmospheric CO_2 (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 (Figure 3a). The overestimation of the predicted CO_2 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 XCO_2 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 XCO_2 retrieval algorithm showed uncertainties over deserts (Bie et al., 2018). Similar to East Asia, the predicted CO_2 emissions over West

Atmospheric Measurement Techniques Discussions

Asia are also underestimated over the areas that are covered by agriculture, herbaceous vegetation, or forest and this underestimation might be due to the presence of CO_2 uptake of the biosphere in the XCO_2 anomalies calculated using the satellite-based retrievals. The difference between the estimated and the actual CO_2 emission was ranging from $10^{-3.5}$ tons to 10^2 tons and the magnitude of difference between $10^{-1.5}$ to $10^{1.5}$ tons accounted for 86% of the total grids.

3.3 Correlation Analysis Between OCO-2 XCO_2 Anomalies and ODIAC Emissions

Figure 6 shows the correlation analysis between the ODIAC CO_2 emissions and the XCO_2 anomalies calculated using the OCO-2 retrievals over East and West Asia. A previous study found that the cluster of XCO_2 changes derived from satellite-based observations showed a better and more significant correlation with the CO_2 emissions relative to a single grid of XCO_2 (Yang et al., 2019) and it might be due to the reason that the atmospheric CO_2 measurement is an instantaneous snapshot of the realistic atmosphere (Liu et al., 2015). For that, we segmented the ODIAC emissions which were binned according to every 0.3 tons/year of lgE using mean emissions calculated from annual emissions during 2015–2019, and then the correlation analysis was carried out between the mean of emissions and the mean of the XCO_2 anomalies within the binned regions. The results showed a positive and significant correlation between the two datasets. Figures 6a and 6b show the spatial distribution of segmented ODIAC emissions over East Asia and the scatterplot between the mean of emissions and mean of XCO_2 anomalies, respectively. The two datasets show a positive and significant correlation with the determined coefficient (R^2) of 0.81. The spatial distribution of segmented ODIAC emissions over West Asia and the scatterplot between the mean of emissions and mean of XCO_2 anomalies are shown in Figures 6c and 6d, respectively. The two datasets showed a good correlation with the determined coefficient (R^2) of 0.60. Several studies correlated the satellite-based XCO_2 anomalies with the CO_2 emissions (Fu et al., 2019; Shekhar et al., 2020). Yang et al. (2019) performed a correlation analysis between the GOSAT based XCO_2 anomalies with the ODIAC CO_2 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 values of CO_2 emissions. In our study, the correlation between the CO_2 emissions and XCO_2 anomalies is relatively low for West Asia and it might be due to the uncertainties in the OCO-2 retrievals. A large part of West Asia is covered by deserts and Bie et al. (2018) reported that the ACOS XCO_2 retrieval algorithm showed uncertainties over deserts.

4 Summary and Conclusions

In this study, a neural network-based approach was suggested to estimate the CO_2 emissions using OCO-2 XCO_2 retrievals. The study was carried out using ODIAC and OCO-2 CO_2 products from 2015 to 2019. To remove the CO_2 seasonal variability and the large background concentration from the OCO-2 XCO_2 retrievals, XCO_2 anomalies were calculated for each year. Then a GRNN model was built and XCO_2 anomalies and CO_2 emissions from 2015 to 2018 were used as a training dataset and finally, CO_2 emissions were predicted for 2019 based on XCO_2 anomalies calculated for the same year. The analyses were carried out separately over East and West Asia. The XCO_2-based estimated and ODIAC actual CO_2 emission datasets were compared and both of the datasets showed a good agreement in terms of spatial distribution. The estimated CO_2 emissions showed better results over East Asia compared to West Asia and it might be due to the uncertainties in the XCO_2 retrievals because previous studies reported that the ACOS XCO_2 retrieval algorithm produced uncertainties over deserts. The predicted CO_2 emissions were mostly underestimated over agricultural areas and forests and it was likely to be caused by the uncertainties in the calculated XCO_2 anomalies and these uncertainties were produced due to the presence of CO_2 uptake of the biosphere. The CO_2 emission results were overestimated compared to the actual ODIAC emissions mostly over bare land. These overestimations might be due to the nearby high emission CO_2 sources that raised the XCO_2 concentration due to the effects of atmospheric transportation. Moreover, correlation analysis was also carried out between the ODIAC CO_2 emissions and OCO-2 XCO_2 anomalies and the results were significant with R^2 of 0.81 and 0.60 over East and West Asia, respectively. The results were in agreement with the previous studies.

The results from our study suggested that the CO_2 emissions can be estimated using the observations obtained from the CO_2 monitoring satellites. Currently, several satellites are orbiting around the Earth and dedicatedly monitoring atmospheric CO_2. Joint utilization of the observations obtained from these satellites might reduce the spatiotemporal gaps and uncertainties. In future studies, we intend to improve the GRNN model by the addition of CO_2 uptake datasets and join utilization of multi-sensor data.

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

Competing Interests. All the authors declare that there is not any personal or financial conflict of interest.

Acknowledgements. The authors acknowledge the efforts of NASA to provide the OCO-2 data products. The authors are also thankful to the National Institute of Environmental Studies (NIES) for providing the ODIAC CO_2 emission dataset.

Data Availability Statement. OCO-2 Level 2 XCO_2 product is available at https://earthdata.nasa.gov and ODIAC CO_2 emission dataset is available at http://db.cger.nies.go.jp/dataset/ODIAC/.

Financial Support. This research has been supported by the National Natural Science Foundation of China (NSFC) grant number 41675133.

References

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.

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

Atmospheric Measurement Techniques Discussions

Boden, T., Andres, R., and Marland, G.: Global, regional, and national fossil-fuel co2 emissions (1751-2014)(v. 2017), Environmental System Science Data Infrastructure for a Virtual Ecosystem …, 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 Sensing, 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, 8, 85, https://doi.org/10.3390/atmos8050085, 2017.

Crisp, D.: Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2), in: Proc.SPIE, 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 CO 2 retrieval algorithm - Part II: Global X CO2 data characterization, 5, 687–707, https://doi.org/10.5194/amt-5-687-2012, 2012.

Crisp, D., Pollock, H., Rosenberg, R., Chapsky, L., Lee, R., Oyafuso, F., Frankenberg, C., Dell, C., Bruegge, C., Doran, G., Eldering, A., Fisher, B., Fu, D., Gunson, M., Mandrake, L., Osterman, G., Schwandner, F., Sun, K., Taylor, T., Wennberg, P., and Wunch, D.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, 10, 59–81, https://doi.org/10.5194/amt-10-59-2017, 2017.

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, 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: European Commission. Emission Database for Global Atmospheric Research (EDGAR v4.3.2)., 2017.

Fu, P., Xie, Y., Moore, C. E., Myint, S. W., and Bernacchi, C. J.: A Comparative Analysis of Anthropogenic CO $_2$ Emissions at City Level Using OCO-2 Observations: A Global Perspective, Earth's Future, 7, 1058–1070, https://doi.org/10.1029/2019EF001282, 2019.

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.

Hakkarainen, J., Ialongo, I., and Tamminen, J.: Direct space-based observations of anthropogenic CO2 emission areas from OCO-2, Geophys. Res. Lett., 43, https://doi.org/10.1002/2016GL070885, 2016.

Hakkarainen, J., Ialongo, I., Maksyutov, S., and Crisp, D.: Analysis of Four Years of Global XCO2 Anomalies as Seen by Orbiting Carbon Observatory-2, Remote Sensing, 11, 850, https://doi.org/10.3390/rs11070850, 2019.

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

Atmospheric Measurement Techniques
Discussions

Hoegh-Guldberg, O., Jacob, D., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., and Guiot, J.: Impacts of 1.5 C global warming on natural and human systems, 2018.

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 Strateg Glob Change, 22, 947–972, https://doi.org/10.1007/s11027-016-9709-9, 2017.

Jalil, S. A.: Carbon Dioxide Emission in the Middle East and North African (MENA) Region: A Dynamic Panel Data Study, JEEIR, 2, 5, https://doi.org/10.24191/jeeir.v2i3.9629, 2014.

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.

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, 12, 2241–2259, https://doi.org/10.5194/amt-12-2241-2019, 2019.

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, 10, 1–23, https://doi.org/10.3390/atmos10070354, 2019.

Kumar, K. R., Revadekar, J. V., and Tiwari, Y. K.: AIRS retrieved CO2 and its association with climatic parameters over india during 2004-2011, 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, Remote Sensing, 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., Razlivanov, 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.

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, 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, 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., Ohyama, H., Niwa, Y., Yashiro, H., Kamei, A., Kawazoe, F., Saeki, T., Ishihara, Y., Imasu, R., Teruyuki, N., Nakajima, T. $\sim$Y., Saitoh, N., and Hashimoto, M.: Early Results of GOSAT-2 Level 2 Products, in: AGU Fall Meeting Abstracts, A52H-02, 2019.

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, 12, https://doi.org/10.3390/RS12152498, 2020.

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

Atmospheric Measurement Techniques Discussions

Mustafa, F., Wang, H., Bu, L., Wang, Q., Shahzaman, M., Bilal, M., Zhou, M., Iqbal, R., Aslam, R. W., Ali, Md. A., and Qiu, Z.: Validation of GOSAT and OCO-2 against In Situ Aircraft Measurements and Comparison with CarbonTracker and GEOS-Chem over Qinhuangdao, China, 13, 899, https://doi.org/10.3390/rs13050899, 2021.

Oda, T. and Maksyutov, S.: ODIAC fossil fuel CO2 emissions dataset (version name: ODIAC2016), Center for Global Environmental Research, National Institute for Environmental Studies. 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.

Oda, Tomohiro: ODIAC Fossil Fuel CO2 Emissions Dataset (ODIAC2020, ODIAC2019, ODIAC2018, ODIAC2017, ODIAC2016, ODIAC2015a, ODIAC2013a), https://doi.org/10.17595/20170411.001, 2015.

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., 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, 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.

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, https://doi.org/10.5772/intechopen.97452, 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, 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 Trans. Neural Netw., 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.

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

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, 251, 112032, https://doi.org/10.1016/j.rse.2020.112032, 2020.

UNFCC: Paris Agreement, 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.

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, M. C., 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 Maziere, 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., 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, 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017.

Yang, S., Lei, L., Zeng, Z., He, Z., and Zhong, H.: An Assessment of Anthropogenic CO_2 Emissions by Satellite-Based Observations in China, 19, 1118, https://doi.org/10.3390/s19051118, 2019.

Zhonghan Chen, Xiaoqian Ye, and Ping Huang: Estimating Carbon Dioxide (CO2) Emissions from Reservoirs Using Artificial Neural Networks, Water, 10, 26, https://doi.org/10.3390/w10010026, 2018.

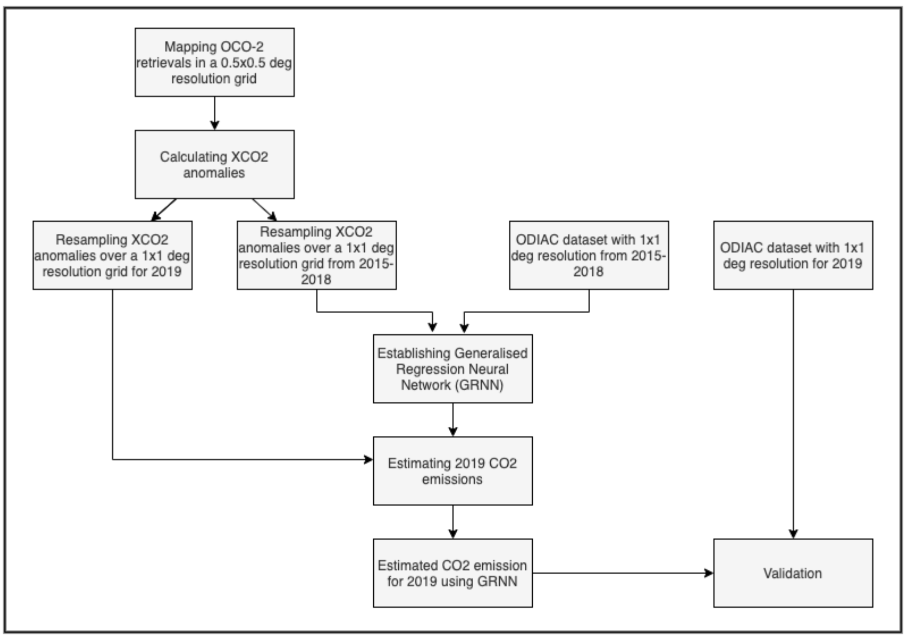


Figure 1: Flowchart explaining steps to estimate the anthropogenic CO_2 emissions using OCO-2 XCO_2 retrievals.

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

Atmospheric Measurement Techniques
Discussions

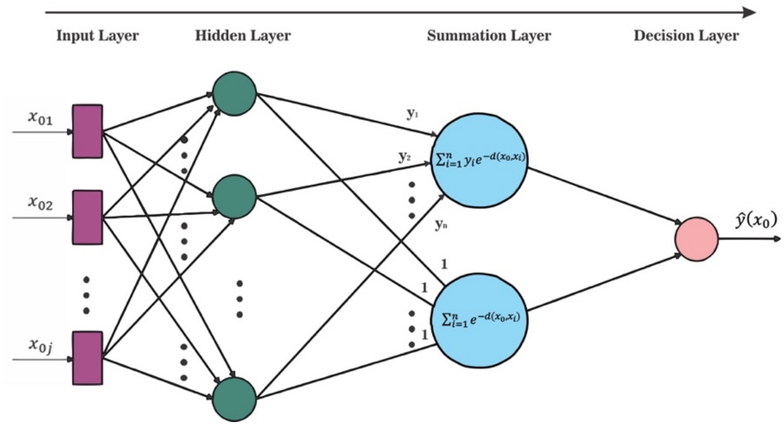


Figure 2: Flowchart explaining steps to estimate the anthropogenic CO_2 emissions using OCO-2 XCO_2 retrievals.

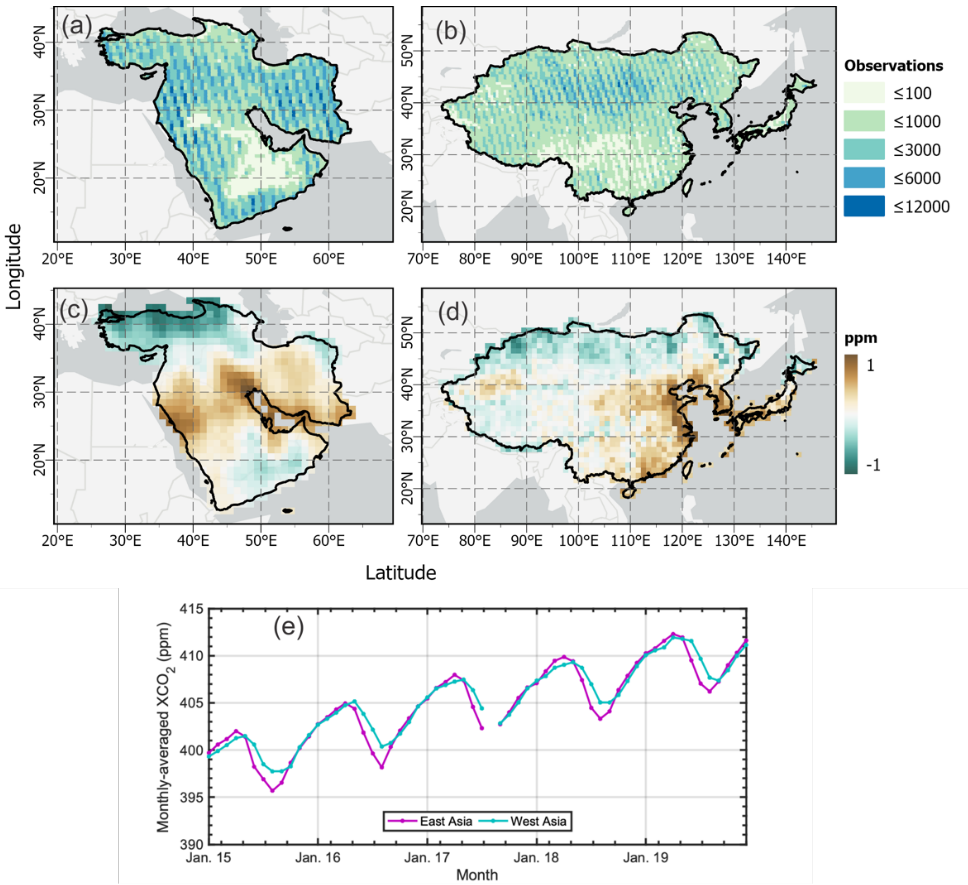
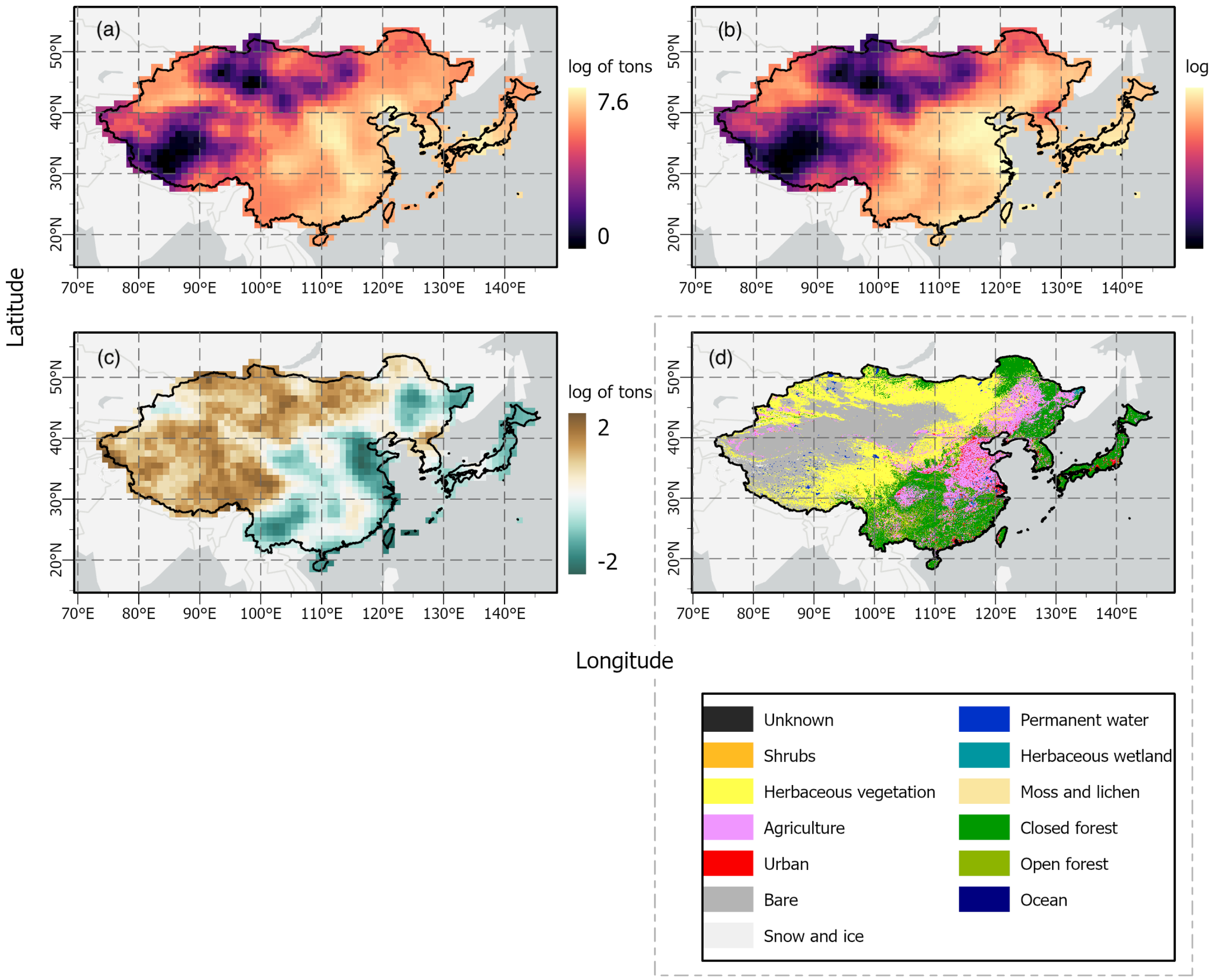


Figure 3: Number of observations in each cell of a 0.5x0.5 deg grid for a period of five years from 2015 to 2019 over (a) West Asia and (b) East Asia; Five years-mean of XCO_2 anomalies calculated using OCO-2 retrieval over (a) West Asia and (b) East Asia; and the monthly-averaged XCO_2 concentration from 2015 to 2019 over East and West Asia. (Basemap credits: OpenStreetMap).

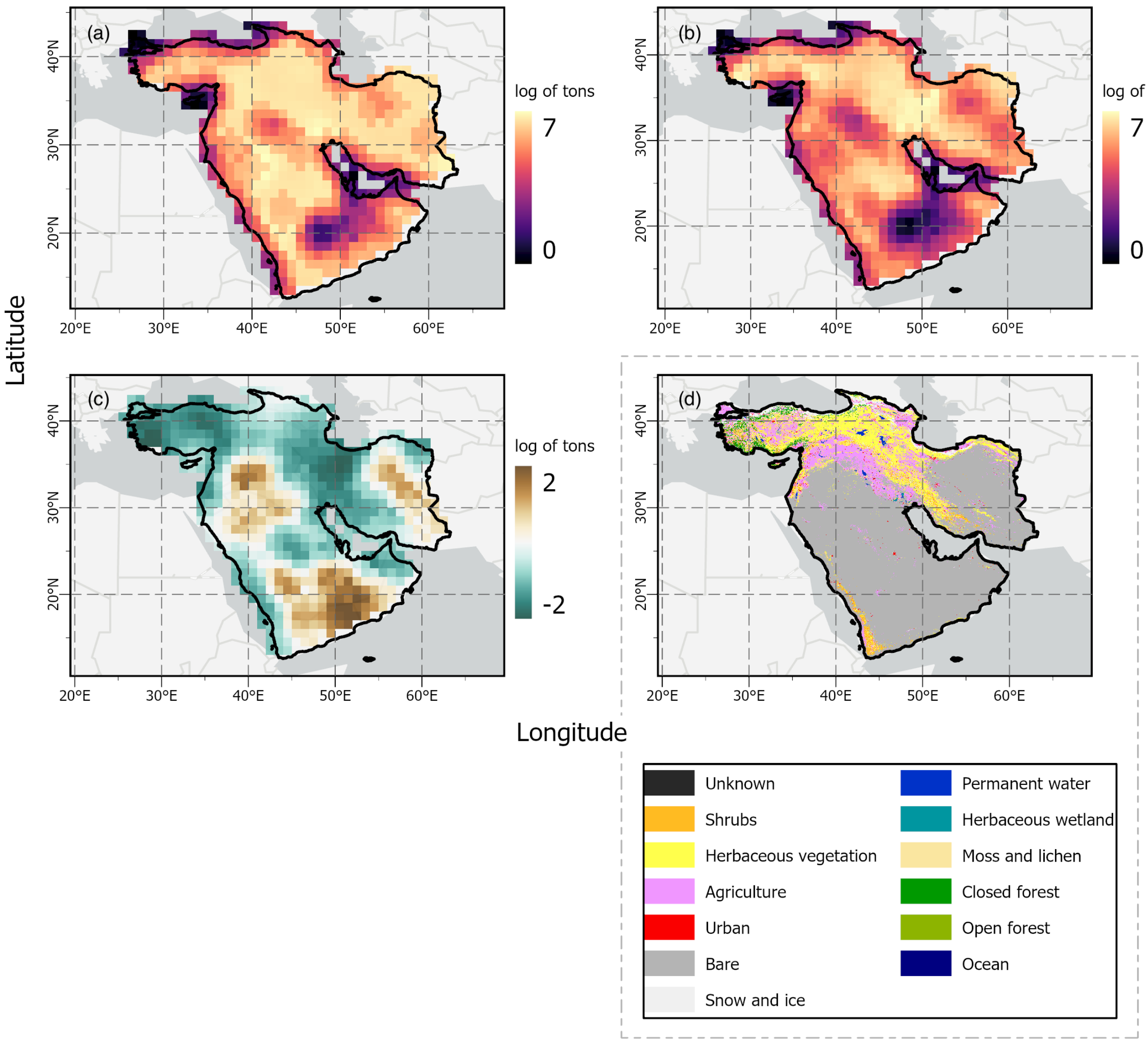
https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021

Atmospheric Measurement Techniques
Discussions



440 **Figure 4: Spatial distribution of (a) OCO-2 XCO_2-based anthropogenic CO_2 emission estimates for 2019 (b) actual ODIAC emissions for 2019, (c) their difference (estimated emission-actual emission), and (d) 100 m resolution landcover distribution provided by Copernicus Global Land Services over East Asia (Basemap credits: OpenStreetMap).**

https://doi.org/10.5194/amt-2021-222
Preprint. Discussion started: 11 August 2021



445 **Figure 5: Spatial distribution of (a) OCO-2 XCO_2-based anthropogenic CO_2 emission estimates for 2019 (b) actual ODIAC emissions for 2019, (c) their difference (estimated emission-actual emission), and (d) 100 m resolution landcover distribution provided by Copernicus Global Land Services over West Asia (Basemap credits: OpenStreetMap).**

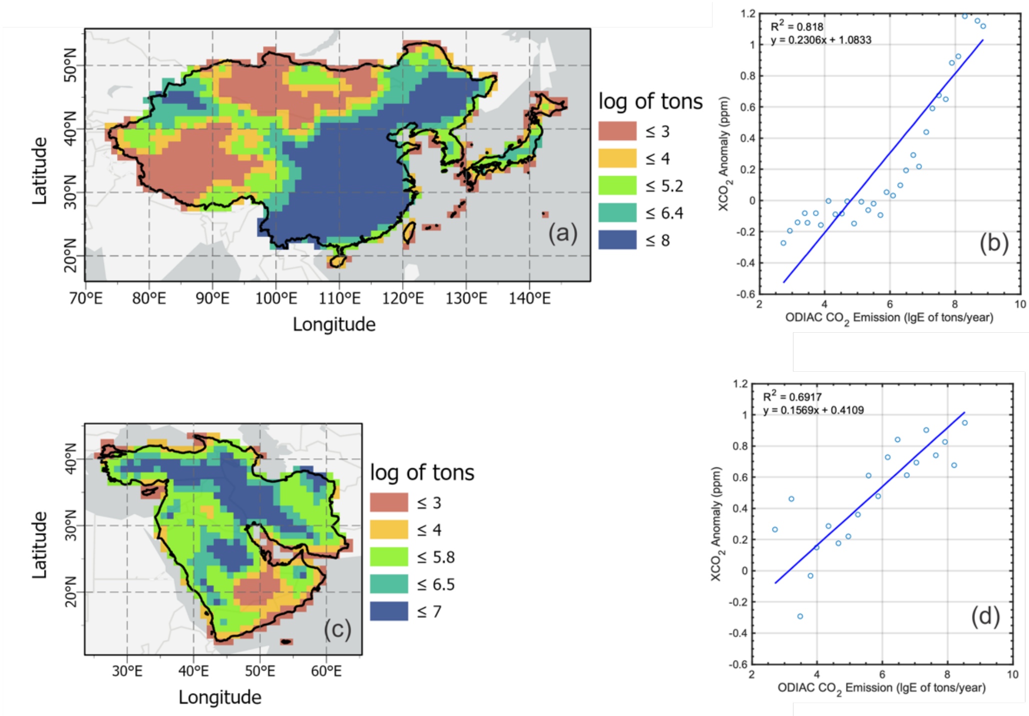


Figure 6: Spatial distribution of segmented ODIAC emissions, where the data are binned by every 0.3 tons/yr of lgE using mean emission calculated from annual emission during 2015–2019 over (a) East Asia and (c) West Asia; the correlation between mean ODIAC CO_2 emissions and mean XCO_2 anomalies calculated from annual XCO_2 during 2015–2018 for (b) East Asia and (d) West Asia (Basemap credits: OpenStreetMap).