Supplement for *GHGPSE-Net: A method towards spaceborne automated extraction of greenhouse-gas point sources using point-object-detection deep neural network*

Yiguo Pang, Denghui Hu, Longfei Tian, Shuang Gao, Guohua Liu

July 11, 2025

S1 Snapshots of syntesized observation

Figure S1: Snapshot of synthesized observations of HGET (3 ppm) across multiple observation footprints.

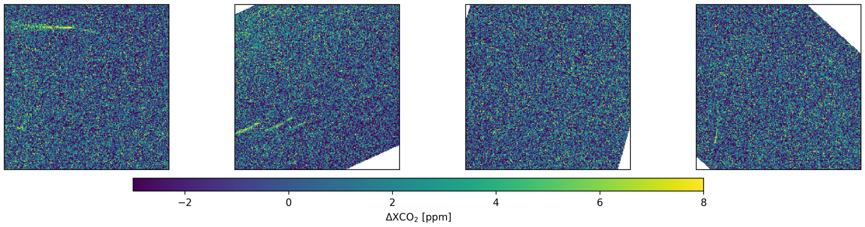


Figure S2: Snapshot of synthesized observations of UCPI (0.7 ppm) across multiple observation footprints.

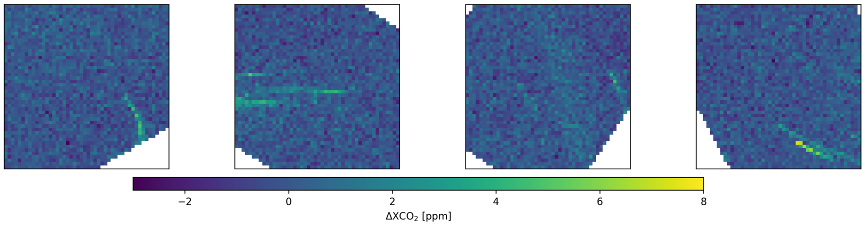


Figure S3: Snapshot of synthesized observations of UCPI (1.5 ppm) across multiple observation footprints.

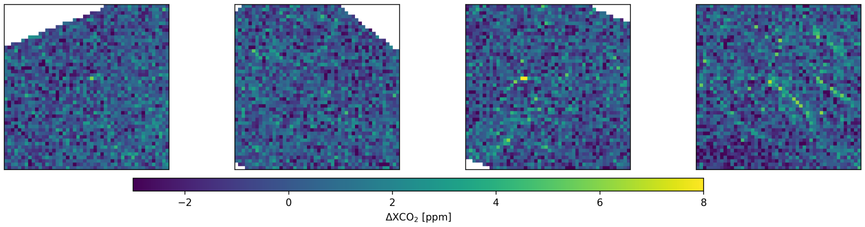


Figure S4: Snapshot of synthesized observations of UCPI (3 ppm) across multiple observation footprints.

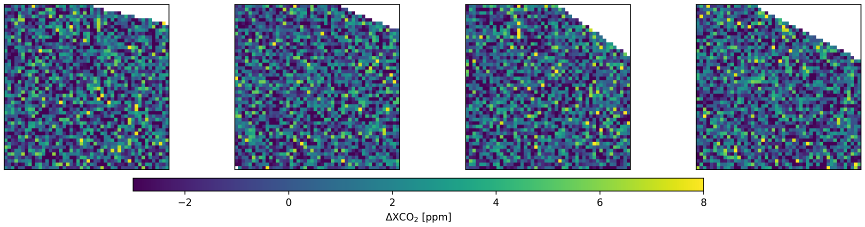
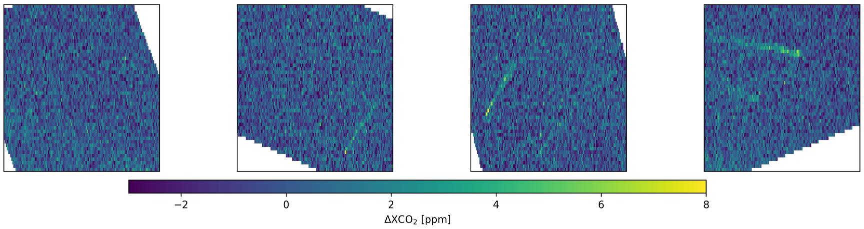


Figure S5: Snapshot of synthesized observations of OCO-3 across multiple observation footprints.



S2 Data augmentation

Figure S6: Distribution of major power density in hotspot area and their CO_2 emission rate.

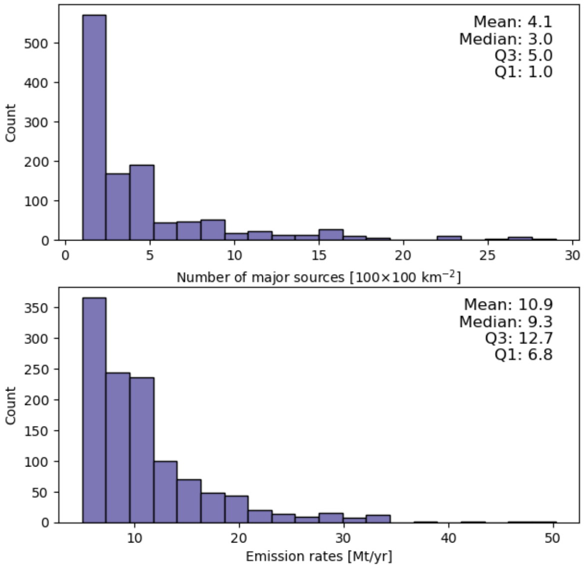


Figure S7: The annual global average XCO_2 under the SSP1-2.6 scenario.

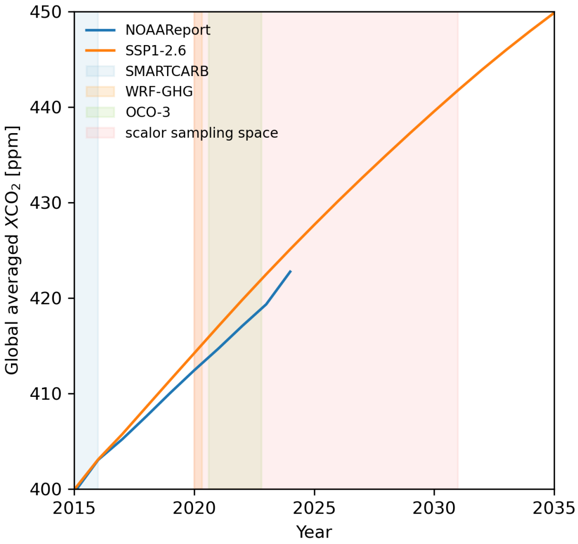


Figure S8: Joint distribution of XCO_2 and CO_2 emission rates.

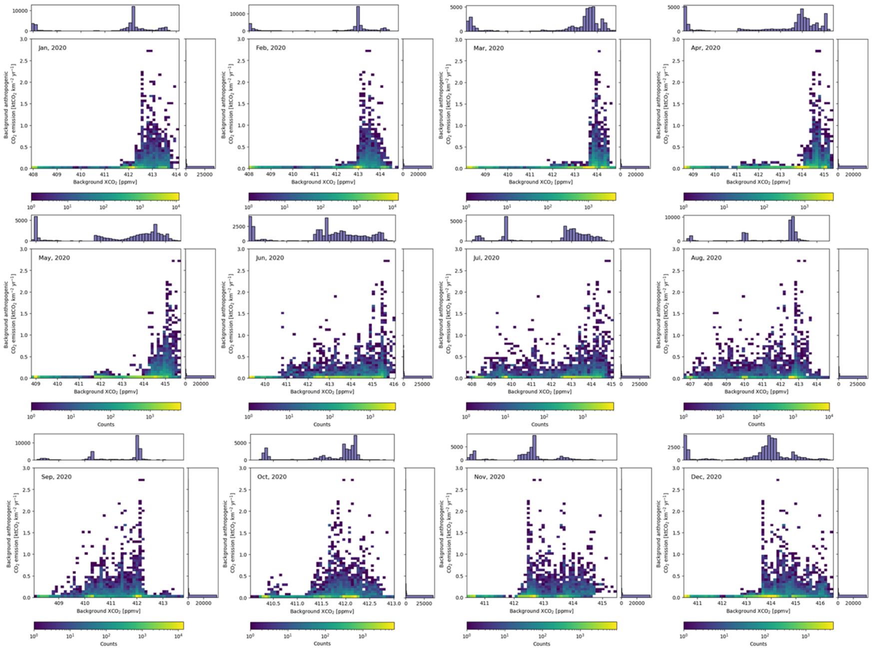


Table S1: Performance comparison of GHGPSE-Nets trained with and without data augmentation.

Data augmentation		Detection			Localization		Quantification				
Train/eval	Test	Precision	Recall	F_1-score	Mean^a	Median^a	Pearson's R	R^2	RMSE^b	MAE^b	MAPE
•	•	0.98	**0.95**	**0.96**	**61.4**	**49.7**	**0.99**	**0.98**	**89.9**	**57.7**	**0.05**
	•	**0.99**	0.09	0.16	310.2	273.5	0.28	-0.68	1072.7	851.3	0.47

^{a} with unit [m]. ^{b} with unit [tCO$_2$ hr^{-1}].

S3 Solving 2D Gaussian-kernel-fitting using constrained least square method

S3.1 Problem defination

Given a 2-D image $\boldsymbol{I}$, which modeled as a combination of N two-dimensional (2-D) Gaussian kernels, we aim to estimate the parameters of these Gaussian kernels. Assuming a shared standard deviation σ, the estimated image $\hat{\boldsymbol{I}}$ is givn by

$$\hat{I}(\boldsymbol{x}) = \sum_{i=1}^{N} a_i G_i(\boldsymbol{x}; \boldsymbol{\mu}_i), \tag{S1}$$

where the pixel location is denoted as $\boldsymbol{x} = [x, y]^T$; a_i denotes the scale of the i-th kernel with $a_i \geq 0$; and $\boldsymbol{\mu}_i = [\mu_{x,i}, \mu_{y,i}]^T$ denotes the center of the i-th kernel and is given by ; G_i is the Gaussian function of the i-th kernel, and is given by

$$G_i(\boldsymbol{x}; \boldsymbol{\mu}_i) = \exp\left(-\frac{1}{2}\frac{(\boldsymbol{x}-\boldsymbol{\mu}_i)^T(\boldsymbol{x}-\boldsymbol{\mu}_i)}{\sigma^2}\right). \tag{S2}$$

Let denote parameter vector as $\boldsymbol{\theta} = [\boldsymbol{a}^T, \boldsymbol{\mu}_x^T, \boldsymbol{\mu}_y^T]^T$, where $\boldsymbol{a} = [a_1, a_2, \ldots, a_N]^T$; $\boldsymbol{\mu}_x = [\mu_{x,1}, \mu_{x,2}, \ldots, \mu_{x,N}]^T$; and $\boldsymbol{\mu}_y = [\mu_{y,1}, \mu_{y,2}, \ldots, \mu_{y,N}]^T$. Let P denote the number of pixels, and let $\boldsymbol{x}_p$ represent the location of the p-th pixel. Then, we can estimate the parameter $\boldsymbol{\theta}$ by fitting the modeled image $\hat{\boldsymbol{I}}$ to observation $\boldsymbol{I}$ using a constrained least square approach. The cost function is formulated as

$$\begin{gathered} \mathcal{L}(\boldsymbol{\theta}) = \textstyle\sum_{p=1}^{P} \left(I(\boldsymbol{x}_p) - \hat{I}(\boldsymbol{x}_p)\right)^2 \\ s.t. \\ a_i \geq 0. \end{gathered} \tag{S3}$$

We can convert the the constrained cost function into a unconstrained function, which are generally easier to be solved, by introducing the penalty methods. Eq.S3 can then be formulated as

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{p=1}^{P} \left(I(\boldsymbol{x}_p) - \hat{I}(\boldsymbol{x}_p)\right)^2 - \lambda \sum_i \log(\text{sigmoid}(a_i)), \tag{S4}$$

where $\lambda > 0$ denotes the penalty coefficient; the sigmoid function is given by

$$\text{sigmoid}(x) = \frac{1}{1+\exp(-x)}. \tag{S5}$$

The parameters $\boldsymbol{\theta}$ can be estimated by minimizing the cost function, and is given by

$$\theta = \arg\min_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}). \tag{S6}$$

S3.2 Solving the minimizing problem using BFGS algorithm

To solve Eq.S6, we adopt the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method, the most popular quasi-Newton algorithm [3]. The BFGS method benefits from not requiring expensive second-order derivative calculations, as it approximates the Hessian to achieve faster convergence. The details of our implementation using BFGS method to solve Eq.S6 are described in Alg.S1.

The gradient functions are given by

$$\begin{aligned} \tfrac{\partial \mathcal{L}}{\partial a_i} &= -2\textstyle\sum_{p=1}^{P}(I_p - \hat{I}_p)G_i(\boldsymbol{x}_p) - \lambda\frac{e^{-a_i}}{1+e^{-a_i}}, \\ \tfrac{\partial \mathcal{L}}{\partial \mu_{x,i}} &= -2a_i\textstyle\sum_{p}(I_p - \hat{I}_p)G_i(\boldsymbol{x}_p)\frac{x_p-\mu_{x,i}}{\sigma^2}, \\ \tfrac{\partial \mathcal{L}}{\partial \mu_{y,i}} &= -2a_i\textstyle\sum_{p}(I_p - \hat{I}_p)G_i(\boldsymbol{x}_p)\frac{y_p-\mu_{y,i}}{\sigma^2}. \end{aligned} \tag{S7}$$

Algorithm S1 BFGS Method

```
Require: Initial parameters $\boldsymbol{\theta}_0$, loss function $\mathcal{L}(\boldsymbol{\theta})$, gradient function $\nabla\mathcal{L}(\boldsymbol{\theta})$,
    maximum iterations $K$, tolerance thresholds $\epsilon_1, \epsilon_2$
Ensure: Optimal parameters $\boldsymbol{\theta}_*$
 1: Initialize $k \leftarrow 0$
 2: while $k < K$ do
 3:     Compute descent direction: $\boldsymbol{d}_k \leftarrow -\boldsymbol{B}_k^{-1}\nabla\mathcal{L}(\boldsymbol{\theta}_k)$
 4:     Set step size according to Wolfe conditions: $\alpha_k \leftarrow \min_{\alpha}\mathcal{L}(\boldsymbol{\theta}_k + \alpha\boldsymbol{d}_k)$
 5:     Compute step: $\boldsymbol{s}_k \leftarrow \alpha_k\boldsymbol{d}_k$
 6:     Update parameters: $\boldsymbol{\theta}_{k+1} \leftarrow \boldsymbol{\theta}_k + \boldsymbol{s}_k$
 7:     if $|\boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k| < \epsilon_1$ or $|\nabla\mathcal{L}(\boldsymbol{\theta}_{k+1})| < \epsilon_2$ then
 8:         break
 9:     end if
10:     Compute gradient variation: $\boldsymbol{y}_k \leftarrow \nabla\mathcal{L}(\boldsymbol{\theta}_{k+1}) - \nabla\mathcal{L}(\boldsymbol{\theta}_k)$
11:     Update Hessian approximation:
            $$\boldsymbol{B}_{k+1} \leftarrow \boldsymbol{B}_k + \frac{\boldsymbol{y}_k\boldsymbol{y}_k^T}{\boldsymbol{y}_k^T\boldsymbol{s}_k} - \frac{\boldsymbol{B}_k\boldsymbol{s}_k\boldsymbol{s}_k^T\boldsymbol{B}_k}{\boldsymbol{s}_k^T\boldsymbol{B}_k\boldsymbol{s}_k}$$
12:     Increment iteration: $k \leftarrow k + 1$
13: end while
```

The step size update in step-4 of Alg.S1 is based on Wolfe conditions, where the step size should follow the Armijo rule and curvature rule, given by

$$\begin{aligned} &i.\mathcal{L}(\theta_k + \alpha\mathbf{d}_k) \leq \mathcal{L}(\theta_k) + c_1\alpha_k\mathbf{d}_k^T\nabla\mathcal{L}(\theta_k), \\ &ii. - \mathbf{d}_k^T\nabla\mathcal{L}(\theta_k + \alpha\mathbf{d}_k) \leq -c_2\mathbf{d}_k^T\nabla\mathcal{L}(\theta_k). \end{aligned} \tag{S8}$$

where $0 < c_1 < c_2 < 1$, and their typical values for quasi-Newton method can be $c_1 = 10^{-4}, c_2 = 0.9$ [3]. For simplicity, we implement this process using the bisectional-backtracking algorithm.

S4 Evaluation for WRF-GHG simulation using IGRA radiosonde data

Figure S9: Temperature residual($T_{\mathrm{WRF-GHG}} - T_{\mathrm{IGRA}}$) profile comparison between WRF-GHG and IGRA radiosonde at Baoshan [1].

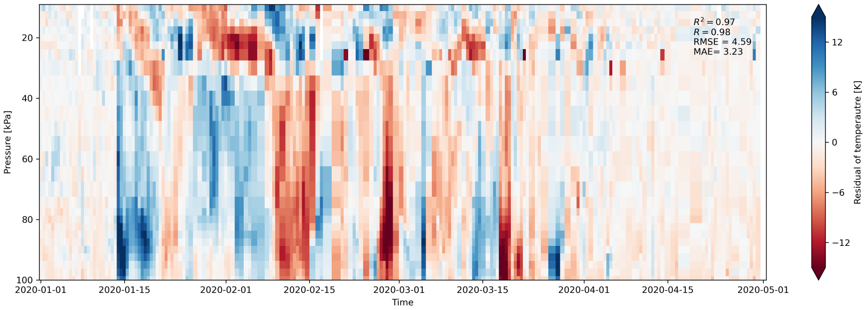


Figure S10: Wind speed residual($U_{\mathrm{WRF-GHG}} - U_{\mathrm{IGRA}}$) profile comparison between WRF-GHG and IGRA radiosonde at Baoshan.

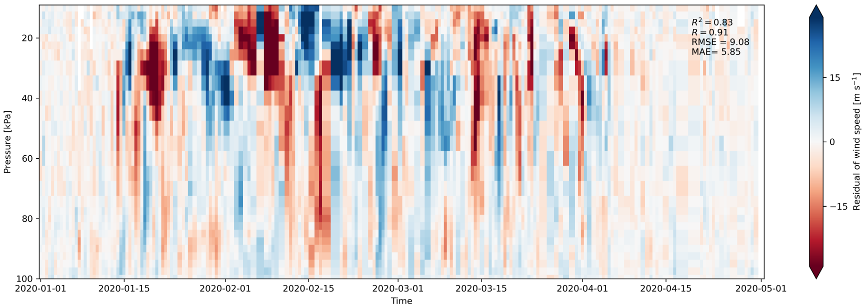
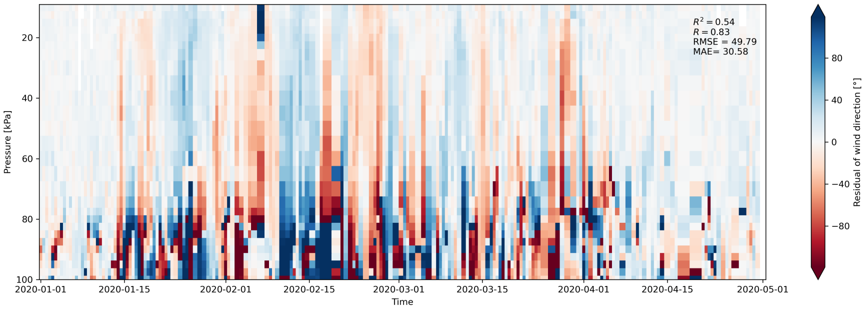


Figure S11: Wind direction residual($\theta_{\mathrm{WRF-GHG}} - \theta_{\mathrm{IGRA}}$) profile comparison between WRF-GHG and IGRA radiosonde at Baoshan.



S5 Evaluation for WRF-GHG simulation using OCO-3 XCO_2 measurements

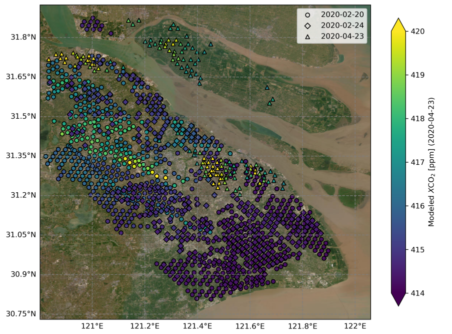


Figure S12: WRF-GHG simulated (a) and OCO-3 observed [4] (b) XCO_2 at OCO-3 observation points. The circle, diamond, and triangular markers represent samples on Feb 20, Feb 24, and Apr 23, 2020, respectively. The basemap is from ESRI.

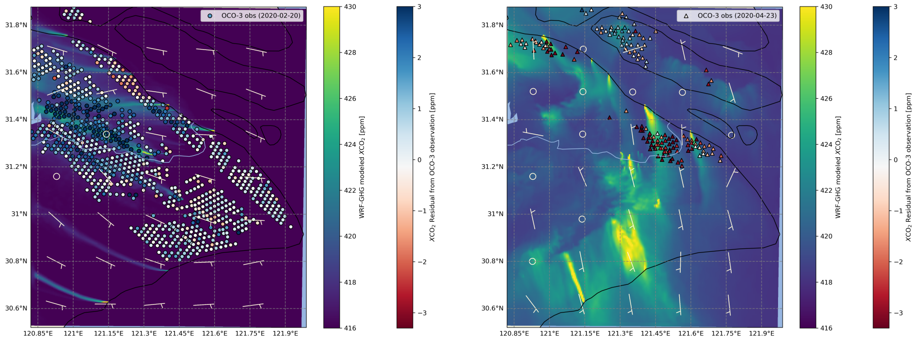
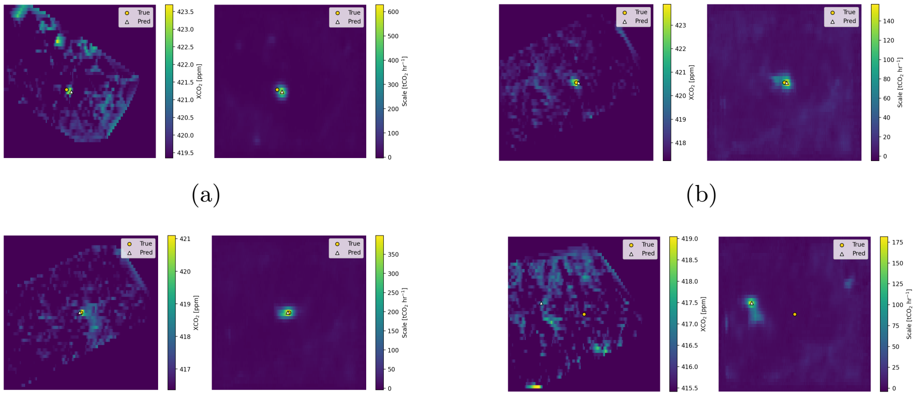


Figure S13: WRF-GHG simulated XCO_2 and residuals between WRF-GHG simulations and OCO-3 observations for (a) Feb 20 and (b) Apr 23, 2020. The major deviations mostly lie downwind of major emission sources, including the metropolitan area in the center and heavy industrial regions to the north near the Yangtze River.

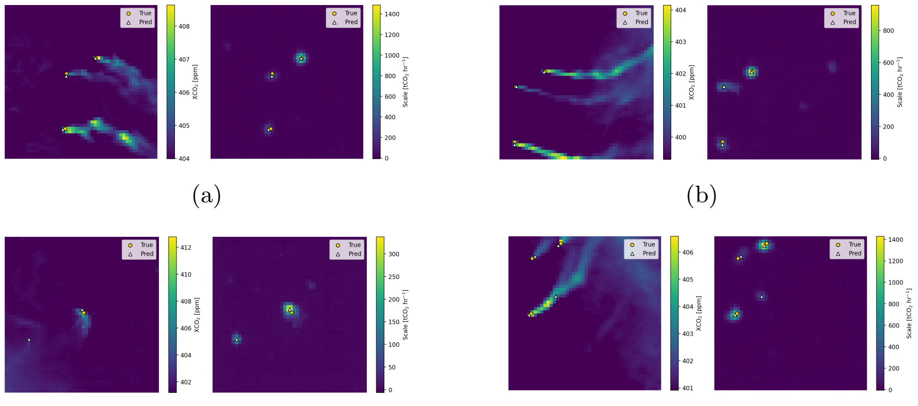
S6 Examples of XCO_2 observations from independent datasets and the corresponding predicted heatmaps.



(a) (b)

(c) (d)

Figure S14: Comparison of OCO-3 XCO_2 observations (left column of each subfigure) and predicted emission heatmaps (right column). Yellow circles indicate true emission sources (power plants), and white triangles indicate detected sources. (a–c) Correct detections with visible plumes near the power plants. (d) False positive case: there is no observable plume near the power plant, and the model detects a plume-shaped structure on the left.



(a)

(b)

(c)

(d)

Figure S15: Comparison of synthetic $X\mathrm{CO}_2$ observations from SMARTCARB datasets [2] (left column of each subfigure) and predicted emission heatmaps (right column). Yellow circles represent true emission sources (power plants), and white triangles represent detected sources. (a–b) Correct detections with visible plumes near the power plants. (c) False positive case: a plume-shaped structure without a true source is detected. (d) Another false positive case, where a companion plume-shaped structure is observable and detected, possibly caused by emissions entering the free troposphere with a wind direction different from that in the planetary boundary layer.

References

[1] Imke Durre, Xungang Yin, Russell S. Vose, Scott Applequist, Jeff Arnfield, Bryant Korzeniewski, and Bruce Hundermark. Integrated Global Radiosonde Archive (IGRA), Version 2, 2016.

[2] Gerrit Kuhlmann, Valentin Clément, Julia Marshall, Oliver Fuhrer, Grégoire Broquet, Christina Schnadt-Poberaj, Armin Löscher, Yasjka Meijer, and Dominik Brunner. SMARTCARB – Use of satellite measurements of auxiliary reactive trace gases for fossil fuel carbon dioxide emission estimation. Technical report, Zenodo, January 2019.

[3] Jorge Nocedal and Stephen J. Wright. Quasi-Newton Methods. In *Numerical Optimization*, pages 135–163. Springer, New York, NY, 2006.

[4] OCO-2/OCO-3 Science Team, Abhishek Chatterjee, and Vivienne Payne. OCO-3 Level 2 bias-corrected XCO2 and other select fields from the full-physics retrieval aggregated as daily files, Retrospective processing v10.4r, 2022.