Supplement of Atmos. Chem. Phys., 25, 18509–18526, 2025
https://doi.org/10.5194/acp-25-18509-2025-supplement

Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

Constraining urban fossil fuel CO_2 emissions in Seoul using combined ground and satellite observations with Bayesian inverse modelling

Sojung Sim and Sujong Jeong

Correspondence to: Sujong Jeong (sujong@snu.ac.kr)

Sect. S1. Details for WRF configuration

As meteorological initial and lateral boundary conditions for WRF simulations, previous studies related to regional-scale inverse modelling have used data from various atmospheric prediction models, including the North American Regional Reanalysis data (Kunik et al., 2019; Sargent et al., 2018; Zhao et al., 2009), the Mesoscale Forecast Model (Ohyama et al., 2023), the High-Resolution Rapid Refresh model (Fasoli et al., 2018), and FNL Operational Global Analysis data from the National Centers for Environmental Prediction (NCEP) (Ohyama et al., 2023; Wu et al., 2018; Ye et al., 2020), as well as ERA-Interim reanalysis data (Kaminski et al., 2022) and ERA5 (Lian et al., 2022; Nalini et al., 2022) from the European Centre for Medium-Range Weather Forecasts (ECMWF). In this study, we applied the Global Forecast System (GFS) produced by the NCEP and ERA5 as boundary conditions for the WRF simulation. The GFS provides analysis data with a horizontal resolution of 0.5° every 6 h, whereas ERA5 offers atmospheric reanalysis data with a horizontal resolution of 30 km every hour. The meteorological input data for the initial and lateral boundary conditions will be determined by comparing wind speed and direction between observations and model outputs.

To prepare input for the WRF simulation, the WRF Preprocessing System (WPS) is required. WPS defines model domains, interpolates static geographical data onto the grids (geogrid process), extracts meteorological fields from boundary conditions data (ungrib process), and interpolates the meteorological fields horizontally (metgrid process). For the geogrid process, one of the required static geographical datasets is land-use category data. The default land-use category dataset is the IGBP-Modified MODIS 20-category. However, because this dataset was produced for global land-use classification in 2002 (Friedl et al., 2002), it does not reflect the current land-use in Seoul. To address this, we reprocessed the land cover map from the Environmental Geographic Information Service (EGIS) provided by the Ministry of Environment in Korea, making it suitable for use as geographical land-use category data. Significant differences in land cover near Seoul were observed when comparing the IGBP-Modified MODIS 20-category data and the EGIS data (**Fig. S1**). In the IGBP-Modified MODIS 20-category data, both Seoul and its vicinity were classified under the Urban and Built-Up (UB) category, whereas in the EGIS data, area classified as Deciduous Broadleaf Forest (DBF) became distinct near Seoul.

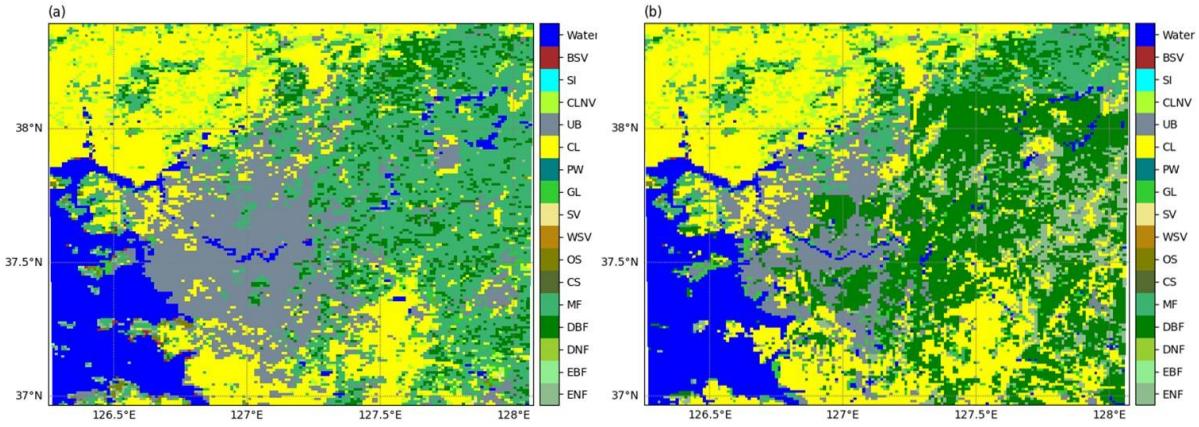

Figure S1: Land-use category maps for (a) IGBP-Modified MODIS 20-category data and (b) EGIS data. BSV, SI, CLNV, UB, CL, PW, GL, SV, WSV, OS, CS, MF, DBF, DNF, EBF, and ENF represent barren or sparsely vegetated regions, snow and ice, cropland/natural vegetation mosaic, croplands, permanent wetlands, grasslands, savannas, woody savannas, open shrub lands, closed shrub lands, mixed forests, deciduous broadleaf forest, deciduous needle leaf forest, evergreen broadleaf forest, and evergreen needle leaf forest, respectively.

For this study, we used a two-way nested configuration featuring a 9 km outer domain (dimensions: 120 × 150 × 40) with two nests: one with a 3 km grid spacing (100 × 100 × 40) and another with a 1 km grid spacing (160 × 160 × 40) (**Fig. S2**). All domains used the WRF Single-Moment 6-Class (WSM6) microphysics scheme (Hong and Lim, 2006), which is a more complex scheme compared to the simple (WSM3) and mixed-phase (WSM5) schemes. For the planetary boundary layer, the Yonsei University (YSU) scheme was applied to all domains (Hong et al., 2006). Additionally, all model domains incorporated the Rapid Radiative Transfer Model for GCMs (RRTMG) shortwave and longwave schemes (Iacono et al., 2008) and the Community Land Model version 4 (CLM4) (Oleson et al., 2010). The Kain-Fritsch scheme for cumulus parameterization (Kain, 2004) was applied only to Domain 1. Grid nudging toward the ERA5 reanalysis data was applied to the wind field (u and v components), temperature, and water vapor mixing ratio across all domains with a nudging coefficient of 3.0×10^{-4} s^{-1}. **Table S1** summarizes the WRF configuration, including model settings and physics options.

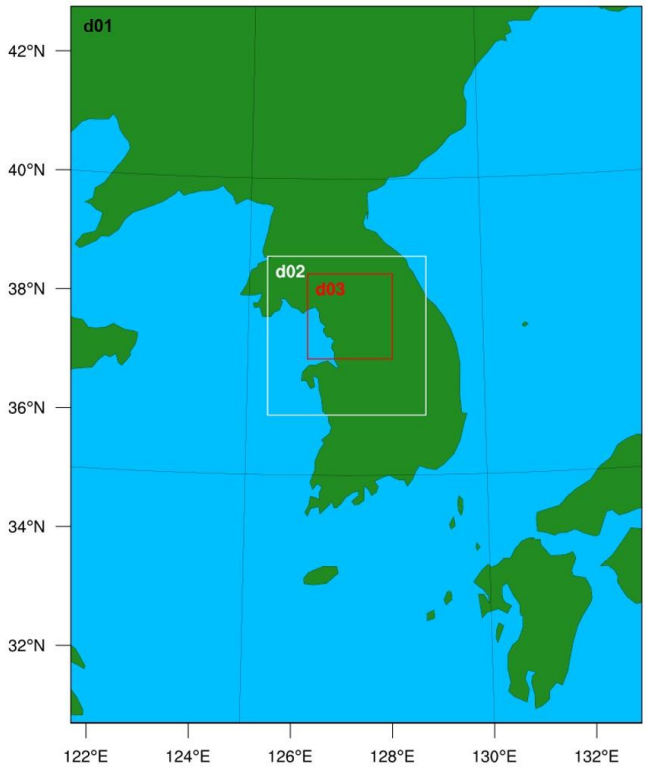

Figure S2: Domain configuration for the WRF simulation.

Table S1: Model settings and physics options for the WRF simulation. Bold text indicates the final data used for the WRF simulation.

Model	**WRF v3.9.1 (WPS v4.0)**		
Meteorological input data	GFS, **ERA5**		
Land-use information	IGBP-Modified MODIS 20-category, **EGIS**		
Domain	Domain 1	Domain 2	Domain 3
Horizontal grid spacing	9 km	3 km	1 km
Dimensions	120×150×40	100×100×40	160×160×40
Vertical layers /Model top	40 sigma layers/50 hPa		
Microphysics	WRF Single–moment 6–class scheme		
Longwave radiation	RRTMG scheme		
Shortwave radiation	RRTMG scheme		
Land surface	CLM4 land-surface model		
Planetary boundary layer	Yonsei University scheme		
Cumulus Parameterization	Kain-Fritsch scheme (only domain 1)		
Grid nudging	Wind field, Temperature, Water vapor mixing ratio		

The wind fields from the WRF simulation output were compared with observational data to evaluate the impact of different input data on the WRF simulation results. For meteorological observations in Seoul, data from the Automated Surface Observing System (ASOS) and Automatic Weather Station (AWS), provided by the Korea Meteorological Administration (https://data.kma.go.kr/), were used. The bias and standard deviation in wind fields between the model and observations were calculated based on the meteorological input data (GFS or ERA5), land-use information (MODIS or EGIS), and grid nudging.

Four cases were compared:

1. When GFS is used as the initial meteorological data in the default configuration (GFS case).
2. When ERA5 is used as the initial meteorological data in the default configuration (ERA5 case).
3. When land-use information is changed from MODIS to EGIS in case 2 (ERA5/EGIS case).
4. When grid nudging is applied in case 3 (ERA5/EGIS+Nudging case).

The differences in wind speed between the model and observations at ASOS in Seoul were 2.38, 2.17, 1.91, and 0.96 m/s for the GFS, ERA5, ERA5/EGIS, and ERA5/EGIS with grid nudging cases, respectively. The wind direction biases for GFS, ERA5, ERA5/EGIS, and ERA5/EGIS with grid nudging were 49.68°, 47.79°, 44.12°, and 40.49°, respectively. **Table S2** shows the mean biases and standard deviations for wind fields between the model and observations at 28 AWS stations in Seoul. The configuration using ERA5 reanalysis data with EGIS land-use data and grid nudging showed the lowest bias in both wind speed and wind direction. Finally, we used the 1 km and 1 h output of the WRF model from the ERA5 with EGIS and grid nudging to simulate optimal meteorological fields.

Table S2: Mean differences and standard deviations in wind speed and direction between model and observational data at 28 AWS stations in Seoul (model minus observation).

Case	Wind speed (m/s)	Wind direction (degree)
GFS	2.95 ± 0.62	51.86 ± 30.78
ERA5	2.65 ± 0.58	49.67 ± 29.38
ERA5/EGIS	2.53 ± 0.55	48.14 ± 29.46
ERA5/EGIS +Nudging	1.46 ± 0.51	44.71 ± 30.62

Sect. S2. Details for footprint calculation from STILT and X-STILT

The STILT model (Fasoli et al., 2018; Lin et al., 2003), coupled with meteorological fields from the WRF model, was used to quantify the sensitivity of ground observations to upwind emission regions. The footprint for each discrete time step $\mathbf{m}$ at coordinates $(\boldsymbol{x_i}, \boldsymbol{y_i})$, transported backward in time from the receptor located at $\boldsymbol{x_r}$ at time $\boldsymbol{t_r}$, can be represented as:

$$\mathbf{f}(\boldsymbol{x_r}, \boldsymbol{t_r} | \boldsymbol{x_i}, \boldsymbol{y_j}, \boldsymbol{t_m}) = \frac{\boldsymbol{m_{air}}}{\boldsymbol{h\bar{\rho}(x_i, y_i, t_m)}} \cdot \frac{\mathbf{1}}{\boldsymbol{N_{tot}}} \sum_{\boldsymbol{p=1}}^{\boldsymbol{N_{tot}}} \boldsymbol{\Delta t_{p,i,j,z \le h}} \quad \text{(Eq. S1)}$$

Where $\boldsymbol{m_{air}}$ is the molar mass of dry air, $\mathbf{h}$ is the atmospheric column height and $\boldsymbol{\bar{\rho}(x_i, y_i, t_m)}$ is the average air density below $\mathbf{h}$ in an element at $(\mathbf{i}, \mathbf{j})$ over time step $\mathbf{m}$. The term $\frac{\mathbf{1}}{\boldsymbol{N_{tot}}} \sum_{\boldsymbol{p=1}}^{\boldsymbol{N_{tot}}} \boldsymbol{\Delta t_{p,i,j,z}}$ represents the time- and volume-integrated influence function, summing the total amount of time each particle $\mathbf{p}$ spends over grid position $(\mathbf{i}, \mathbf{j}, \mathbf{z})$, normalized by the total number of particles $\boldsymbol{N_{tot}}$. A change in CO_2 ($\boldsymbol{\Delta CO_2}$) at the downwind receptor can be obtained by multiplying the footprint ($\mathbf{f}$) by gridded CO_2 flux estimates $\boldsymbol{F(x_i, y_j, t_m)}$:

$$\boldsymbol{\Delta CO_2}(\boldsymbol{x_r}, \boldsymbol{t_r} | \boldsymbol{x_i}, \boldsymbol{y_j}, \boldsymbol{t_m}) = \boldsymbol{F}(\boldsymbol{x_i}, \boldsymbol{y_j}, \boldsymbol{t_m}) \mathbf{f}(\boldsymbol{x_r}, \boldsymbol{t_r} | \boldsymbol{x_i}, \boldsymbol{y_j}, \boldsymbol{t_m}) \quad \text{(Eq. S2)}$$

For satellite measurements, X-STILT (Wu et al., 2018), a modified version of STILT was used to extract XCO_2 signals from OCO-2 and OCO-3 data. X-STILT calculates the weighted column footprint $\boldsymbol{f_w}$, considering column-averaging kernels and pressure weights. The formulation of $\boldsymbol{f_w}$ is as follows:

$$\boldsymbol{f_w}(\boldsymbol{x_{n,r}}, \boldsymbol{t_{n,r}} | \boldsymbol{x_i}, \boldsymbol{y_j}, \boldsymbol{t_m}) = \frac{\boldsymbol{m_{air}}}{\boldsymbol{h\bar{\rho}(x_i, y_i, t_m)}} \cdot \frac{\mathbf{1}}{\boldsymbol{N_{tot}}} \sum_{\boldsymbol{p=1}}^{\boldsymbol{N_{tot}}} \boldsymbol{\Delta t_{p,i,j,z \le h}} \boldsymbol{AK_{norm}}(\mathbf{n}, \mathbf{r}) \mathbf{PW}(\mathbf{n}, \mathbf{r}) \quad \text{(Eq. S3)}$$

Where $\boldsymbol{x_{n,r}}, \boldsymbol{t_{n,r}}$ denotes a column receptor, $\boldsymbol{AK_{norm}(n, r)}$ is the normalized averaging kernel, and $\mathbf{PW}(\mathbf{n}, \mathbf{r})$ is the pressure weighting function. Similar to the footprint from WRF-STILT, a change in XCO_2 ($\boldsymbol{\Delta XCO_2}$) at the downwind column receptor can be yielded by multiplying $\boldsymbol{f_w}$ by gridded CO_2 flux estimates:

$$\boldsymbol{\Delta XCO_2}(\boldsymbol{x_{n,r}}, \boldsymbol{t_{n,r}} | \boldsymbol{x_i}, \boldsymbol{y_j}, \boldsymbol{t_m}) = \boldsymbol{F}(\boldsymbol{x_i}, \boldsymbol{y_j}, \boldsymbol{t_m}) \boldsymbol{f_w}(\boldsymbol{x_{n,r}}, \boldsymbol{t_{n,r}} | \boldsymbol{x_i}, \boldsymbol{y_j}, \boldsymbol{t_m}) \quad \text{(Eq. S4)}$$

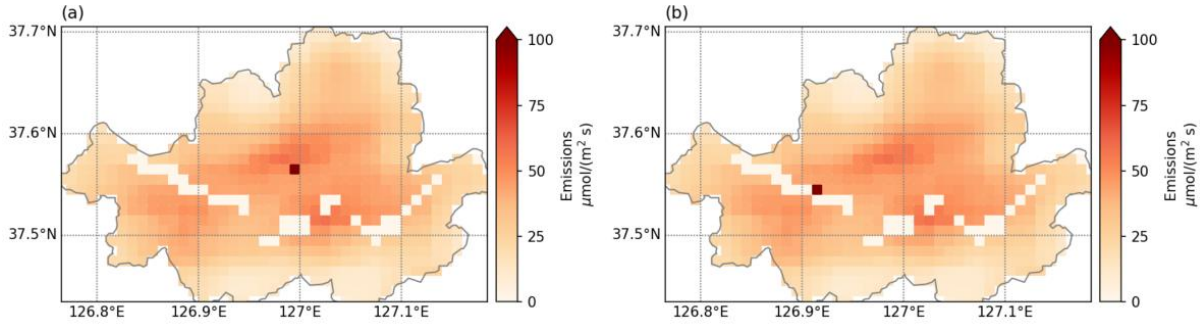

Figure S3: Spatial distribution of CO_2 emissions (ODIAC) in Seoul during December 2021, showing (a) the emissions before and (b) after the relocation of the misaligned point source.

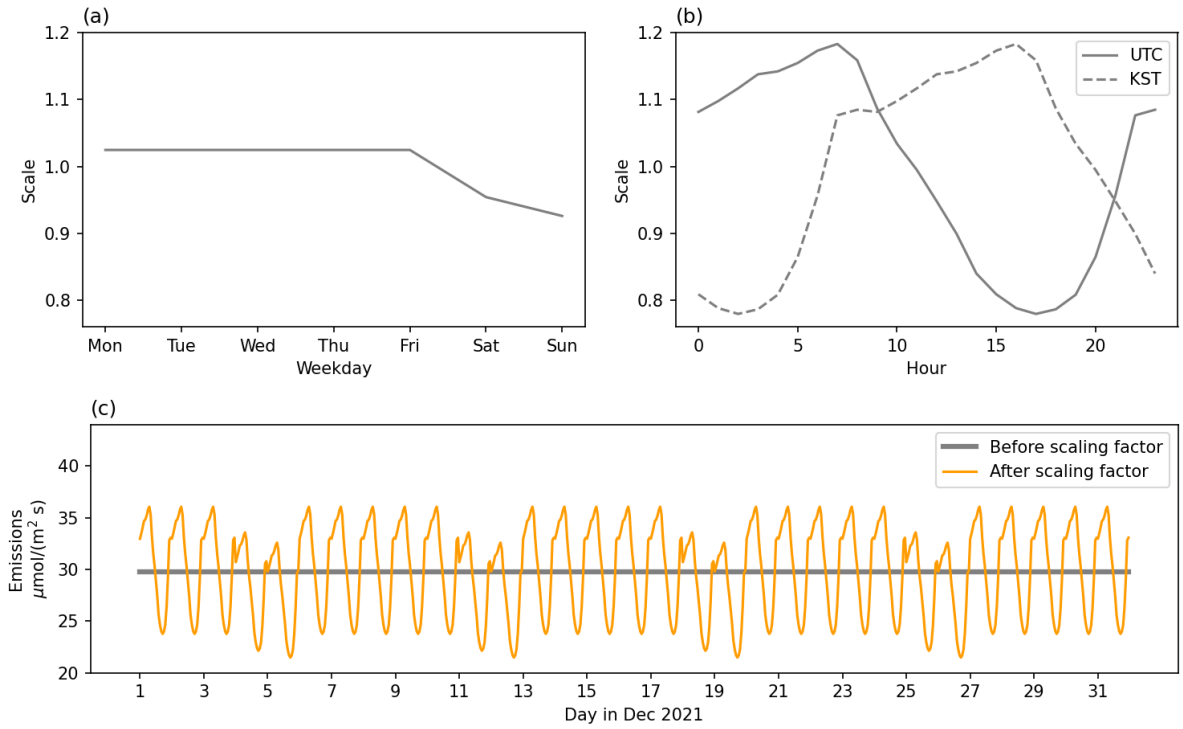

Figure S4: Temporal scaling factors for CO_2 emissions in Seoul: (a) weekly scaling factors, (b) diurnal scaling factors, and (c) a time series of CO_2 emissions before and after applying the temporal scaling factors for December 2021.

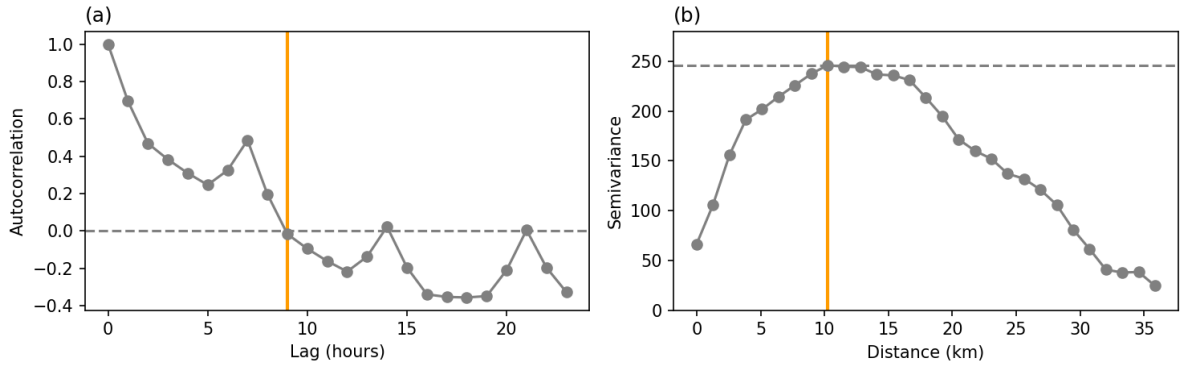

Figure S5: Covariance correlation length parameter analysis. (a) Lagged autocorrelation function of prior emissions uncertainty for daytime (10:00–16:00 KST) in December 2021, averaged over the Seoul domain. The temporal correlation length is approximately 9 h (orange line), where the autocorrelation crosses zero (gray dashed line) for the first time, indicating no correlation. (b) Variogram of prior emissions uncertainty for the Seoul domain, based on daytime (10:00–16:00 KST) data averaged over December 2021 after removing large outliers (above the 95th percentile of uncertainty). The spatial correlation length is approximated as 10 km, rounded down from 10.24 km (orange line), where the empirical variogram reaches the maximum semivariance value (gray dashed line).

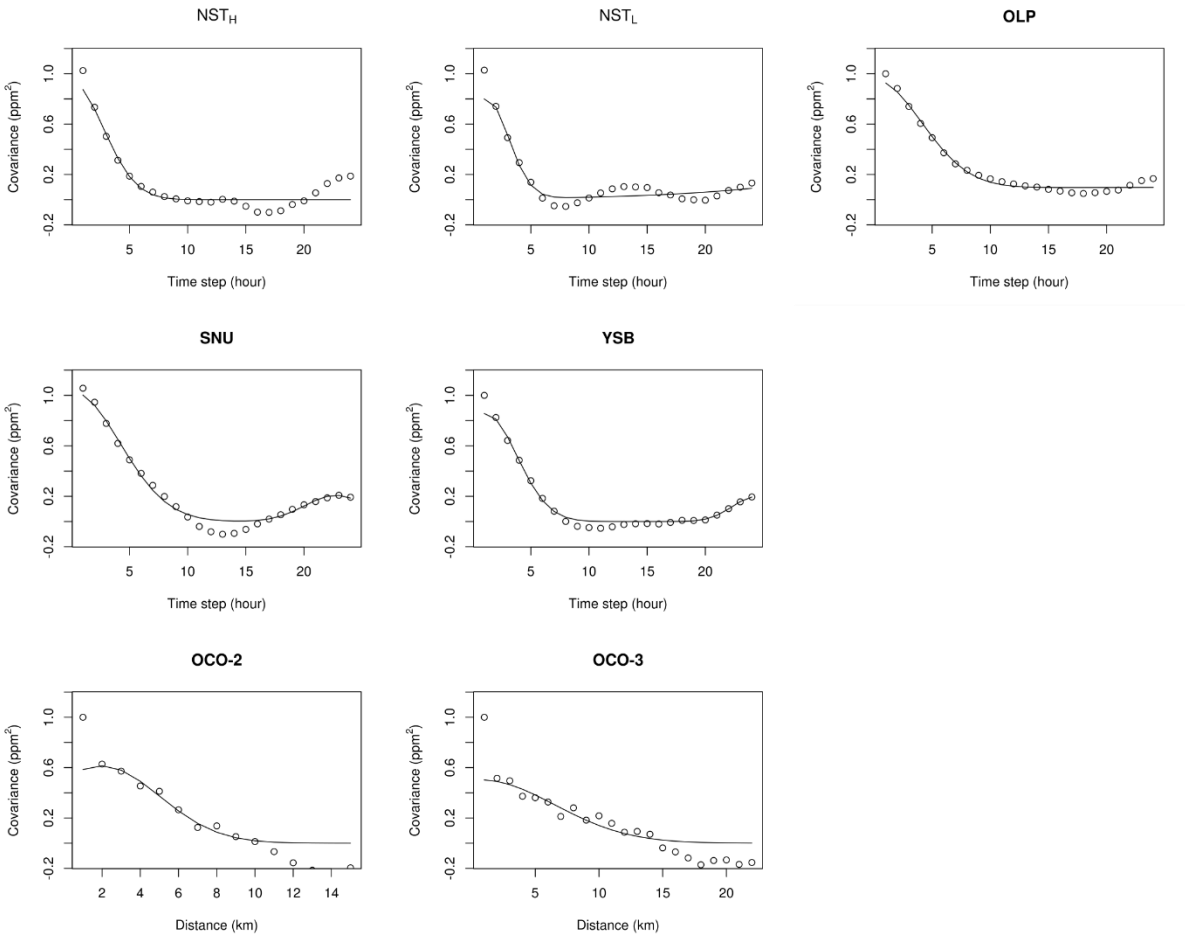

Figure S6: Departure covariances as a function of time steps for ground-based observations (NST_H, NST_L, OLP, SNU, and YSB) and as a function of separation distance for satellite observations (OCO-2 and OCO-3).

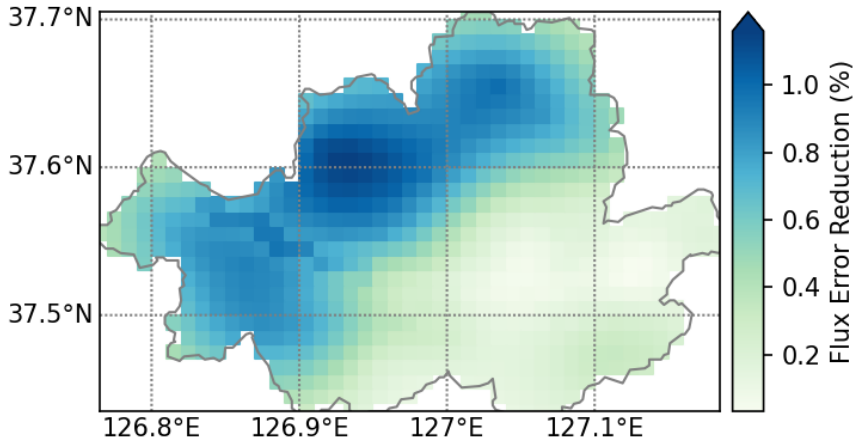

Figure S7: Spatial differences in UR calculated using all observation data compared with using only the five ground sites during December 2021.

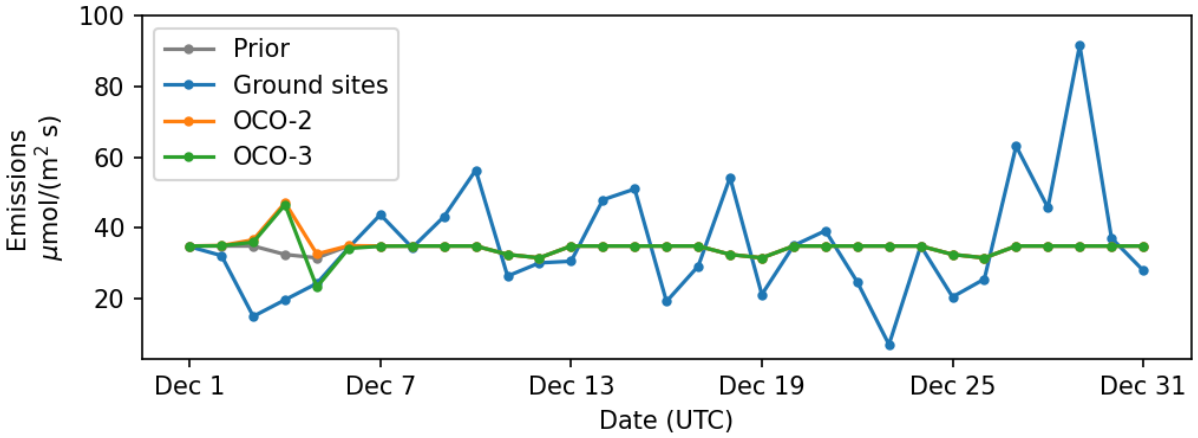

Figure S8: Time series of domain-averaged daily prior and posterior CO_2 emissions derived using different observations (ground sites, OCO-2, and OCO-3).

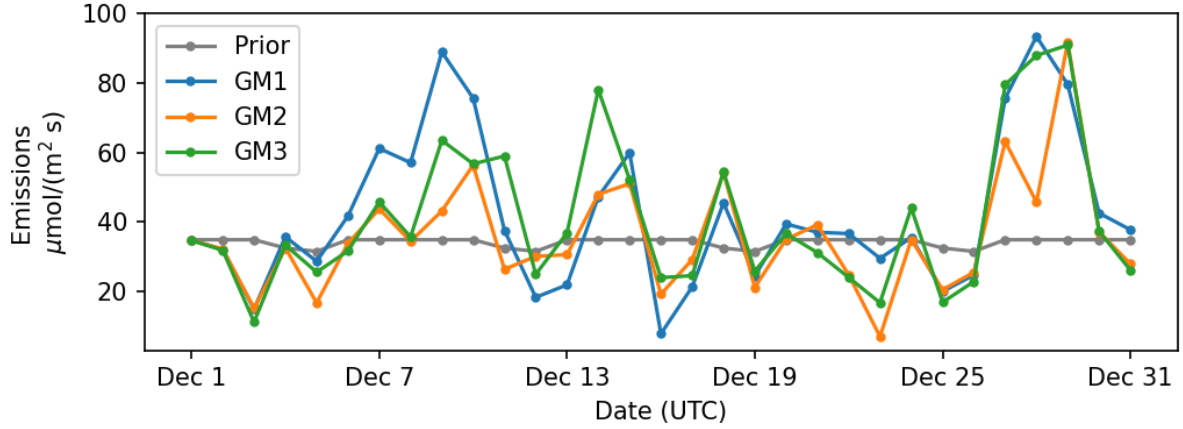

Figure S9: Time series of domain-averaged daily prior and posterior CO_2 emissions derived using different background estimation methods for ground-based observations (GM1, GM2, and GM3).

References

Fasoli, B., Lin, J. C., Bowling, D. R., Mitchell, L., and Mendoza, D.: Simulating atmospheric tracer concentrations for spatially distributed receptors: Updates to the Stochastic Time-Inverted Lagrangian Transport model's R interface (STILT-R version 2), Geosci. Model Dev., 11, 2813–2824, https://doi.org/10.5194/gmd-11-2813-2018, 2018.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C.: Global land cover mapping from MODIS: Algorithms and early results, Remote Sens. Environ., 83, 287–302, https://doi.org/10.1016/S0034-4257(02)00078-0, 2002.

Hong, S. Y. and Lim, J.-O. J.: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6), Asia-Pacific J. Atmos. Sci., 42, 129–151, 2006.

Hong, S. Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Mon. Weather Rev., 134, 2318–2341, https://doi.org/10.1175/MWR3199.1, 2006.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, J. Geophys. Res. Atmos., 113, https://doi.org/https://doi.org/10.1029/2008JD009944, 2008.

Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, J. Appl. Meteorol., 43, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2, 2004.

Kaminski, T., Scholze, M., Rayner, P., Houweling, S., Voßbeck, M., Silver, J., Lama, S., Buchwitz, M., Reuter, M., Knorr, W., Chen, H. W., Kuhlmann, G., Brunner, D., Dellaert, S., Denier van der Gon, H., Super, I., Löscher, A., and Meijer, Y.: Assessing the Impact of Atmospheric CO_2 and NO_2 Measurements From Space on Estimating City-Scale Fossil Fuel CO_2 Emissions in a Data Assimilation System, Front. Remote Sens., 3, 1–21, https://doi.org/10.3389/frsen.2022.887456, 2022.

Kunik, L., Mallia, D. V., Gurney, K. R., Mendoza, D. L., Oda, T., and Lin, J. C.: Bayesian inverse estimation of urban CO_2 emissions: Results from a synthetic data simulation over Salt Lake City, UT, Elem. Sci. Anthr., 7, 1–16, https://doi.org/10.1525/elementa.375, 2019.

Lian, J., Lauvaux, T., Utard, H., Bréon, F. M., Broquet, G., Ramonet, M., Laurent, O., Albarus, I., Cucchi, K., and Ciais, P.: Assessing the Effectiveness of an Urban CO_2 Monitoring Network over the Paris Region through the COVID-19 Lockdown Natural Experiment, Environ. Sci. Technol., 56, 2153–2162, https://doi.org/10.1021/acs.est.1c04973, 2022.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res. Atmos., 108, https://doi.org/https://doi.org/10.1029/2002JD003161, 2003.

Nalini, K., Lauvaux, T., Abdallah, C., Lian, J., Ciais, P., Utard, H., Laurent, O., and Ramonet, M.: High-Resolution Lagrangian Inverse Modeling of CO_2 Emissions Over the Paris Region During the First 2020 Lockdown Period, J. Geophys. Res. Atmos., 127, 1–26, https://doi.org/10.1029/2021JD036032, 2022.

Ohyama, H., Frey, M. M., Morino, I., Shiomi, K., Nishihashi, M., Miyauchi, T., Yamada, H., Saito, M., Wakasa, M., Blumenstock, T., and Hase, F.: Anthropogenic CO_2 emission estimates in the Tokyo Metropolitan Area from ground-based CO_2 column observations, Atmos. Chem. Phys., 23, 15097–15119, https://doi.org/10.5194/acp-23-15097-2023, 2023.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J.-F., Mahowald, N., Niu, G.-Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stockli, R., Wang, A., Yang, Z.-L., Zeng, X., and Zeng, X.: Technical Description of version 4.0 of the Community Land Model (CLM), Univ. Corp. Atmos. Res., 2010.

Sargent, M., Barrera, Y., Nehrkorn, T., Hutyra, L. R., Gately, C. K., Jones, T., Mckain, K., Sweeney, C., Hegarty, J., Hardiman, B., Wang, J. A., and Wofsy, S. C.: Correction: Anthropogenic and biogenic CO_2 fluxes in the Boston urban region, Proc. Natl. Acad. Sci. U. S. A., 115, E9507, https://doi.org/10.1073/pnas.1815348115, 2018.

Wu, D., Lin, J. C., Fasoli, B., Oda, T., Ye, X., Lauvaux, T., Yang, E. G., and Kort, E. A.: A Lagrangian approach towards extracting signals of urban CO_2 emissions from satellite observations of atmospheric column CO_2 (XCO2): X-St, Geosci. Model Dev., 11, 4843–4871, 2018.

Ye, X., Lauvaux, T., Kort, E. A., Oda, T., Feng, S., Lin, J. C., Yang, E. G., and Wu, D.: Constraining Fossil Fuel CO_2 Emissions From Urban Area Using OCO-2 Observations of Total Column CO_2, J. Geophys. Res. Atmos., 125, e2019JD030528, https://doi.org/https://doi.org/10.1029/2019JD030528, 2020.

Zhao, C., Andrews, A. E., Bianco, L., Eluszkiewicz, J., Hirsch, A., MacDonald, C., Nehrkorn, T., and Fischer, M. L.: Atmospheric inverse estimates of methane emissions from Central California, J. Geophys. Res. Atmos., 114, 1–13, https://doi.org/10.1029/2008JD011671, 2009.