Supplementary Material

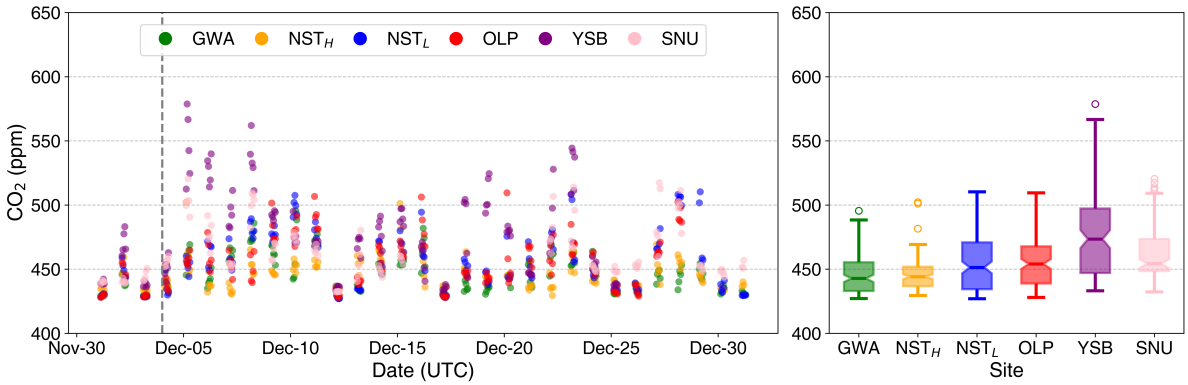


Figure S1. Time series of CO_2 concentrations and their distribution used for inversions over Seoul in December 2021 (adapted from Sim and Jeong, 2025). Only data from 3-7 UTC (12-16 KST) are shown. The period before the gray dashed line (December 1-3) represents the spin-up of the atmospheric transport model and was excluded from the inversions.

Text S1. Details for WRF configuration

We followed the WRF configuration described in previous study Sim and Jeong (2025) with additional enhancements to improve the accuracy of the WRF simulation. We used sea surface temperature (SST) data from the UK Met Office's Operational Sea surface Temperature and Sea-Ice Analysis (OSTIA) during the WRF Preprocessing System (WPS) and applied observation nudging.

Because South Korea is surrounded by three seas, this geographical feature requires careful consideration of sea surface temperature in meteorological inputs. The global European Center for Medium-Range Weather Forecasts Re-Analysis data (ERA5), which was applied as the boundary conditions for the WRF simulations, provides SST data at 0.25° resolution based on OSTIA. As shown in Fig. S2, we incorporated higher-spatial resolution OSTIA SST data at 5 km as an intermediate file to improve the spatial information of SST in our simulations.

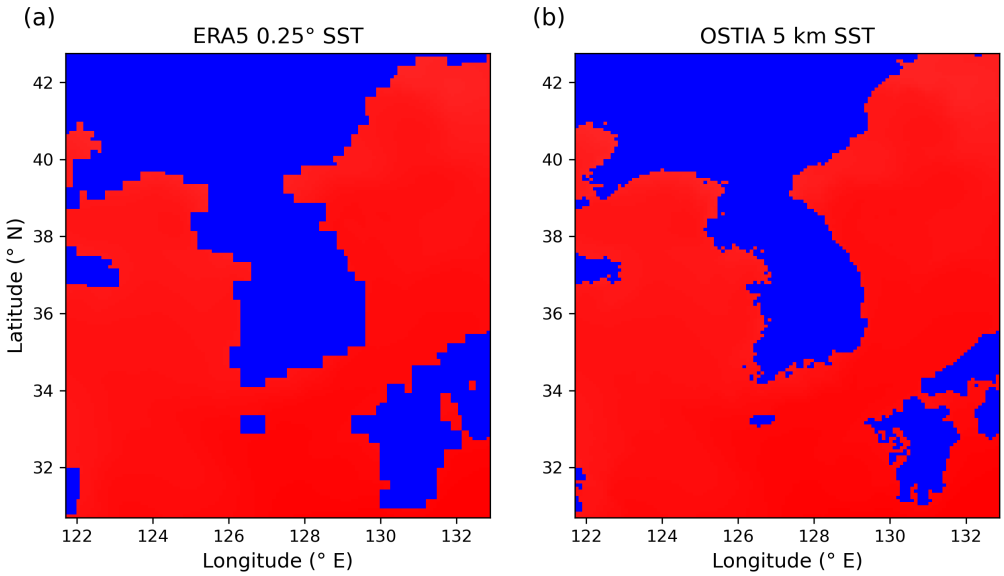


Figure S2. SST input maps for (a) default SST data from ERA5 and (b) high resolution (5 km) SST data from OSTIA.

For observation nudging, we assimilated observed wind direction, wind speed, temperature, and relative humidity from four Korea Meteorological Administration Automated Synoptic Observation System (ASOS) sites located in domain 3 (Table S1). Figure. S3 shows the monthly averaged Temperature (TT) field from the WRF preprocessing system (metgrid file) for domain 3, confirming that observation nudging was successfully applied at these ASOS stations.

Table S1. Locations and details of four ASOS sites used for observation nudging in WRF domain 3.

Site name	Instrument type	Latitude (°N)	Longitude (°E)	Elevation (m)
Gangwha Island	ASOS	37.7074	126.4463	47.84
Baekryeong Island	ASOS	37.9740	124.7124	36.00
Seoul	ASOS	37.5714	126.9658	85.67

Incheon	ASOS	37.4777	126.6249	68.99

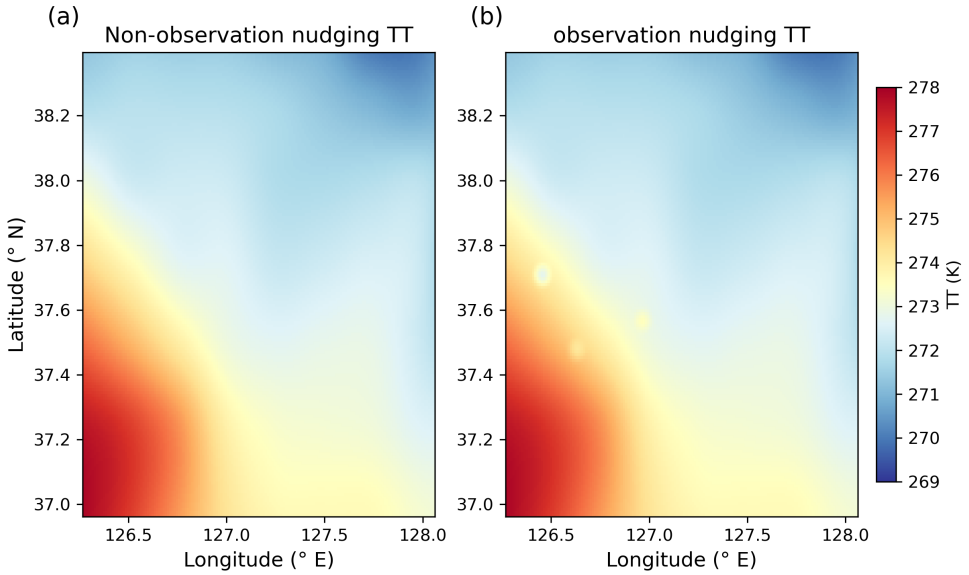


Figure S3. Monthly averaged Temperature (TT) field from WRF preprocessing system (metgrid output) for domain 3 (a) without observation nudging and (b) with observation nudging.

The WRF configuration is shown in Table S2 and the domain configuration for the WRF simulation is described in Fig. S4. For the domain configuration, we used a two-way nesting technique with a 9 km upper domain (domain 1, 120×150) containing two nested inner domains, domain 2 with 3 km resolution (100×100) and domain 3 with 1 km resolution (160×160), all with 40 vertical levels. Initial and boundary conditions were obtained from ERA5 with a horizontal resolution of 30 km every hour. Land use data were primarily obtained from the Environmental Geographic Information Service (EGIS) provided by the Ministry of Climate, Energy and Environment. For areas not covered by the EGIS dataset, the IGBP-Modified MODIS 20-category dataset was used as the default dataset. WRF Single-moment 6-class (WSM6) is used as microphysics scheme (Hong and Lim, 2006), the Rapid Radiative Transfer Model for General Circulation Models (RRTMG) is applied for shortwave and longwave schemes (Iacono et al., 2008; Morcrette et al., 2008). The planetary boundary layer scheme used is the Yonsei University (YSU) scheme (Hong et al., 2006). We incorporated the Revised MM5 Monin-Obukhov scheme for surface layer scheme (Jiménez et al., 2012) and the Community Land Model version 4 (CLM4) for land surface processes (Oleson et al., 2013). The Kain-Fritsch cumulus parameterization was applied only to upper domain 1 (Kain and Fritsch, 1990; Kain, 2004). For grid nudging, we assimilated temperature, wind field, and water vapor mixing ratio from the ERA5 reanalysis. For observation nudging, we set the horizontal radius of influence to 10 km, and the vertical radius of influence to 0.1 in the sigma vertical coordination system, following the WRF-LES-FDDA system developed by the national institute of meteorological sciences (NIMS) (NIMS, 2023). The detailed configurations applied to the WRF simulations is described in Table S2.

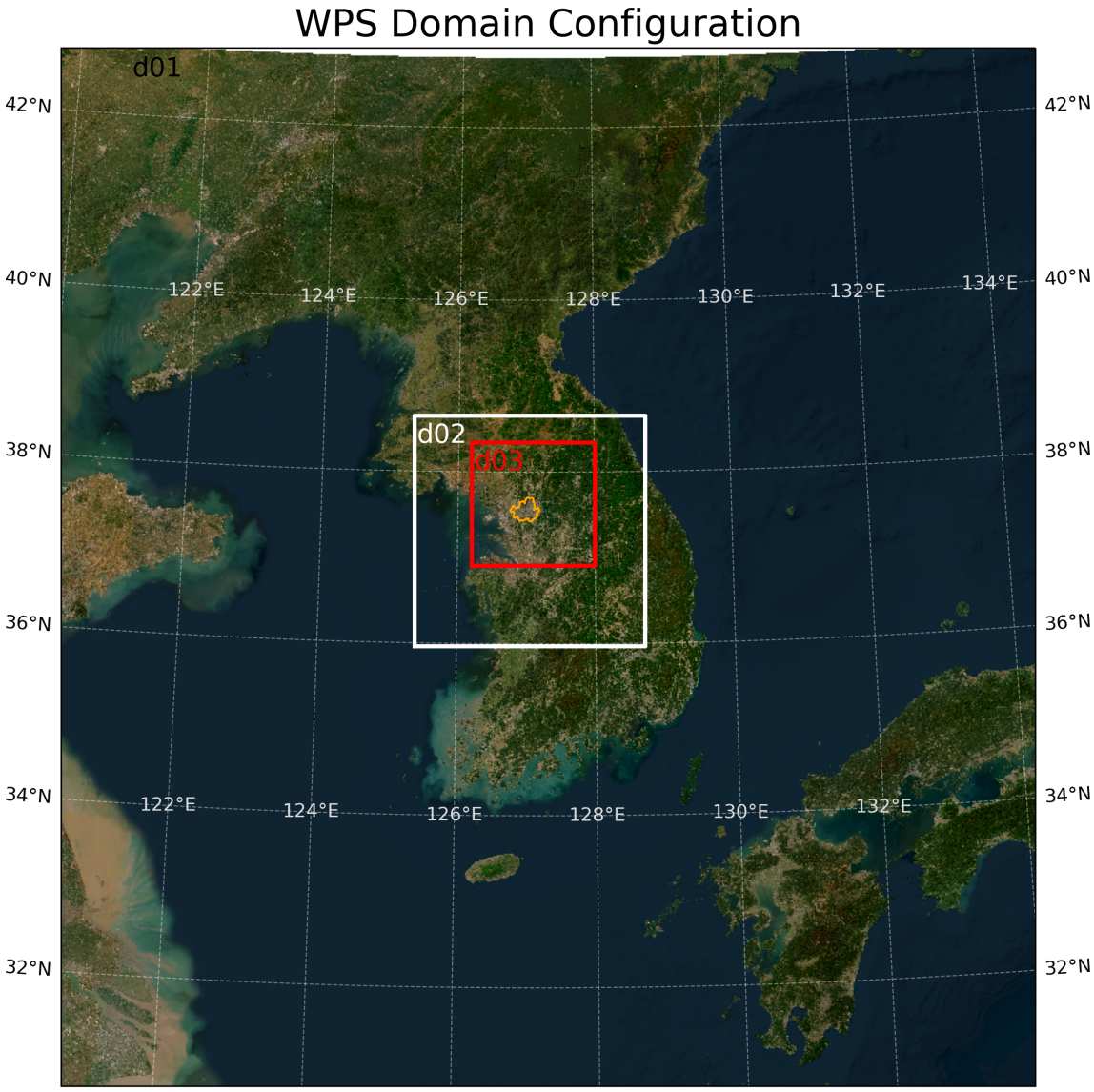


Figure S4. A domain configuration for the WRF simulations (Orange line indicates the boundary of Seoul).

Table S2. Model configurations for the WRF simulations.

WRF v3.9.1 (WPS v4.0)				
Meteorological input data		ERA5 (30 km, hourly)		
Land use information		IGBP-Modified MODIS 20-category, EGIS		
Domain (Number of grids)		Domain 1 (120×150)	Domain 2 (100×100)	Domain 3 (160×160)
Horizontal grid spacing		9 km	3 km	1 km
Vertical layers /Model top		40 sigma layers/ 50 hPa		
Physical scheme	**Microphysics**	WRF Single-moment 6-class		
	Longwave radiation	Rapid radiative transfer model for General Circulation Models		

	Shortwave radiation	Rapid radiative transfer model for General Circulation Models		
	Surface Layer /Land surface	Revised MM5 Monin-Obukhov Scheme/CLM4 land-surface model		
	Planetary boundary layer	Yonsei University Scheme		
	Cumulus Parameterization	Kain-Fritsch	-	-
	Grid Nudging	Wind field, Temperature, Water vapor mixing ratio		
	Observation Nudging	-	-	Wind field, Temperature, Relative Humidity

We validated the WRF model performance by comparing the simulated meteorological field against observations from the ASOS station in Seoul. To assess the model performance of wind speed, wind direction, temperature, and relative humidity, the mean absolute error (MAE), index of agreement (IOA), mean bias error (MBE), and Pearson correlation coefficient (r) were calculated as summarized in Table S3. The correlation coefficients between modeled and observed values exceed 0.7 for wind speed, wind direction, temperature, and relative humidity, and the IOA values also exceed 0.7, suggesting that the 1 km high resolution meteorological fields from WRF simulations adequately reproduce the meteorological conditions over Seoul during this period.

Table S3. The performance of WRF simulation between ASOS observation in Seoul.

WRF variable	MAE (±1 σ)	IOA	MBE	r
Wind Speed (m/s)	1.31 (± 1.32)	0.74	0.87	0.81
Wind Direction (θ)	0.83 (± 1.14)	0.86	0.29	0.74
Temperature (K)	1.18 (± 0.95)	0.98	-0.74	0.98
Relative Humidity (%)	6.86 (± 5.91)	0.93	4.37	0.90

Text S2. Detail for Lagrangian transport model STILT

We used Stochastic Time-Inverted Linear Transport (STILT) coupled with meteorological field from WRF simulations to derive footprints. The footprints represent the sensitivity of receptor concentration changes to surface fluxes. STILT is a Lagrangian particle dispersion model that releases an ensemble of virtual particles backward in time from the receptor location and calculates this sensitivity based on their trajectories. Following Lin et al. (2003) and Fasoli et al. (2018), the receptor concentration change is calculated stochastically from the number of particles residing within the domain and their residence times as described in Eq. (S1):

$$\Delta C_{m,i,j}(X_r, t_r) = f(X_r, t_r | x_i, y_i, t_m) F(x_i, y_i, t_m) \quad \text{(S1)}$$

Here, $F(x_i, y_i, t_m)$ is the surface flux at location (x, y) and time t, and $f(X_r, t_r | x_i, y_i, t_m)$ is the footprint, which converts surface fluxes to concentration change at the receptor location (X_r) at time (t_r). Footprints (f) is defined as follows (Lin et al., 2003; Fasoli et al., 2018):

$$f(X_r, t_r | x_i, y_i, t_m) = \frac{m_{air}}{h\bar{\rho}(x_i, y_i, t_m)} \frac{1}{N_{tot}} \sum_{p=1}^{N} \Delta t_{p,m,i,j,k} \quad \text{(S2)}$$

The term $\frac{1}{N_{tot}} \sum_{p=1}^{N} \Delta t_{p,m,i,j,k}$ denoted as the residence time that a given particles spends within a spatial grid cell during a specified time interval. $\frac{m_{air}}{h\bar{\rho}(x_i, y_i, t_m)}$ converts surface fluxes into atmospheric concentrations. The molar mass of air (m_{air}) is defined as 28.97 g mol^{-1} in this study, and $\bar{\rho}(x_i, y_i, t_m)$ represents the mean air density at location (x, y) and time t within the atmospheric boundary layer, calculated following Eq. (S3) as described in Fasoli et al. (2018).

$$h'_k(p, t) = z_r + \int_0^{t_k} \sigma_\omega \sqrt{2T_L \left(t + T_L \left(e^{-\frac{t}{T_L}} - 1 \right) \right) dt} \quad \text{(S3)}$$

t_k represents the time step component and σ_ω is standard deviation of vertical velocities encountered by particles along the STILT backward trajectory during the integration time step. T_L is the Lagrangian decorrelation time scale. According to Eq. (S3), the effective mixing depth $h'_k(p, t)$ in STILT gradually converges from the receptor height above ground level (z_r) to the planetary boundary layer (PBL) height from the meteorological fields.

In this study, we derived footprints using meteorological fields from WRF for December 2021, releasing 1000 particles and tracking them backward in time window for 24 hours. The STILT receptors were configured at the latitude, longitude and inlet height of the six observation sites used in the inverse modeling. Fig. S5 and S6 show the December 2021 daytime averaged footprints for all six observation sites and each individual site, respectively. The sensitivity of footprint increases with proximity to the receptor locations, with higher sensitivities observed at sites with lower observation inlet heights (e.g., OLP, YSB, and SNU). In addition, all six sites exhibited relatively high footprint values to extend toward northwest. GWA and NST_H, which have relatively high inlet heights above ground level (AGL), exhibited weaker footprints, with patterns more concentrated toward the northwest. In contrast, OLP, SNU, and YSB, which have relatively low AGL inlet heights, indicated stronger footprints and broader spatial distribution extending beyond the northwestern direction.

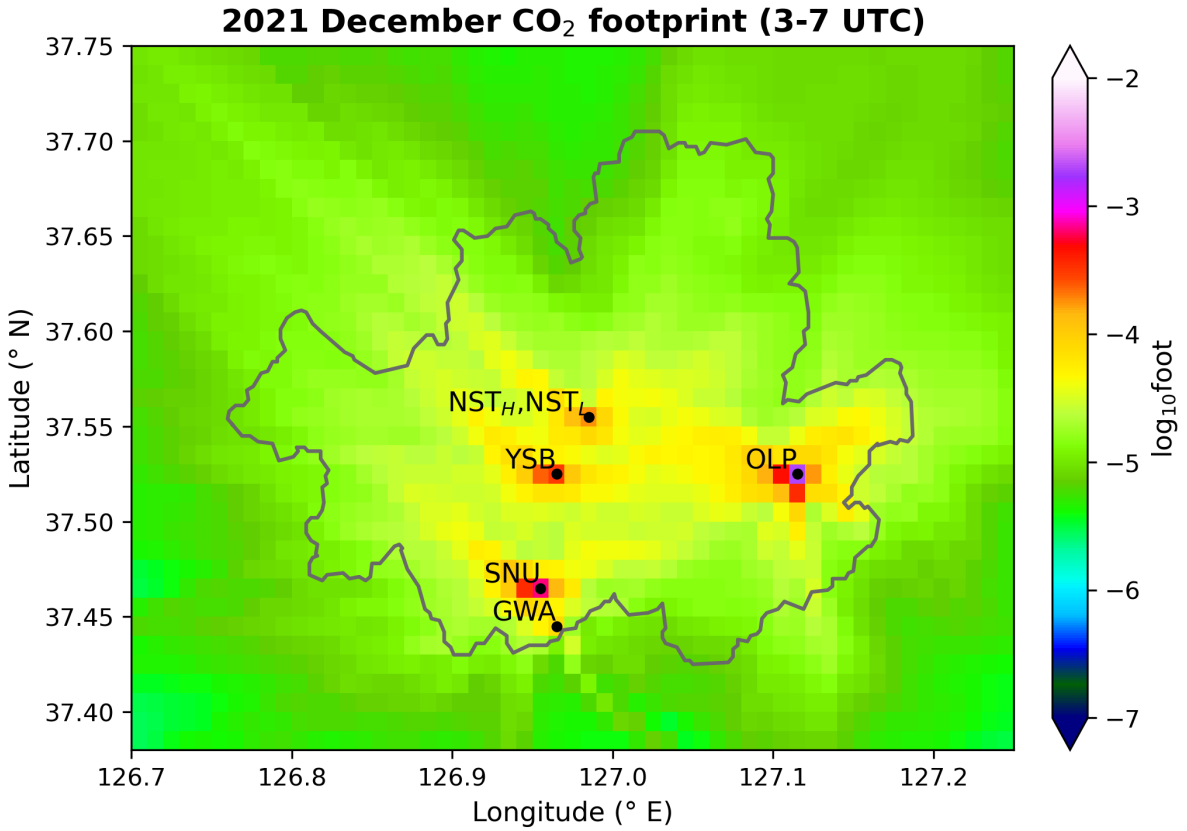


Figure S5. A spatial distribution of the December 2021 mean daytime (3-7 UTC) footprints for all six observation sites. The unit of footprint is ppm μmol^{-1} m^2 s.

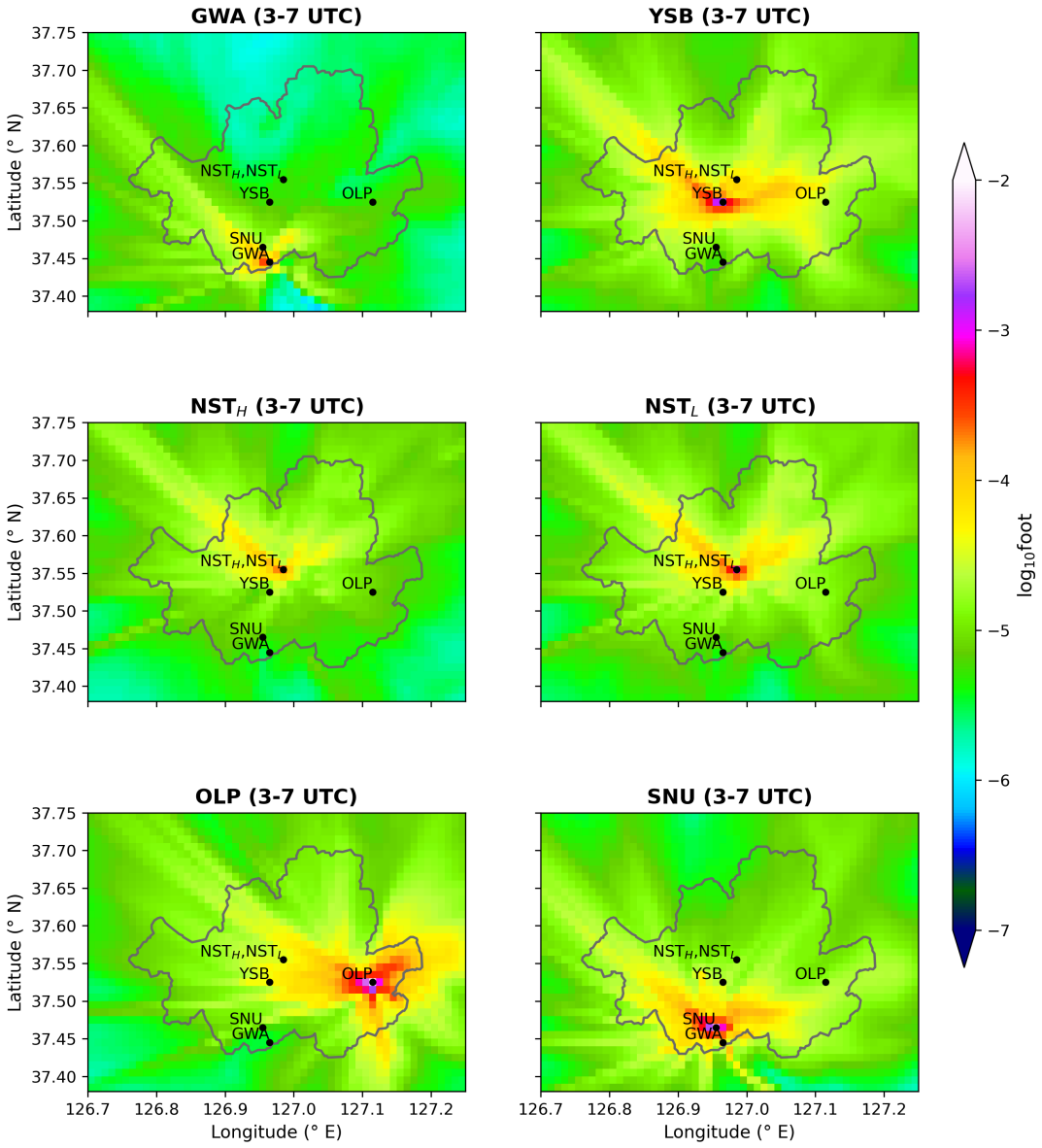


Figure S6. December 2021 daytime averaged footprints for each of the six observation sites used in this study. The unit of footprint is ppm μmol^{-1} m^2 s.

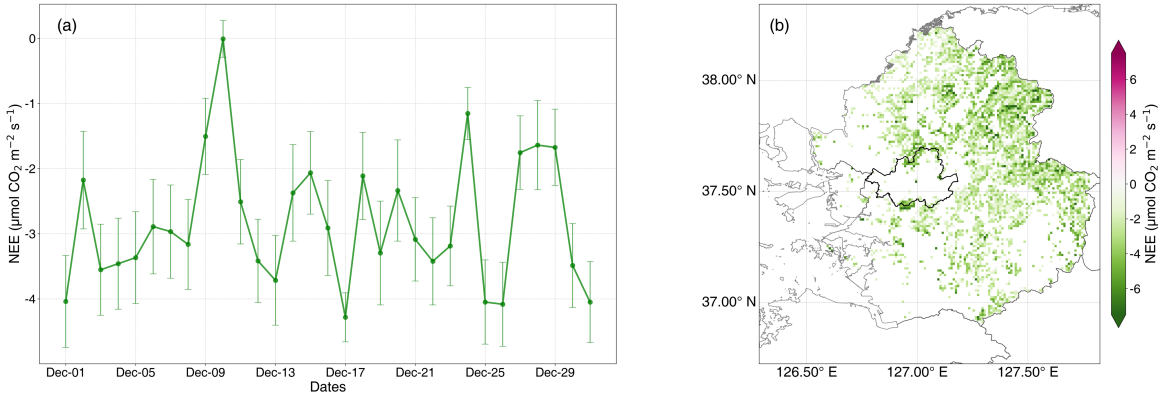


Figure S7. A graph of daytime (3-7 UTC) biospheric CO_2 flux (Net Ecosystem Exchange, NEE) for December 2021. (a) daily mean timeseries with error bars representing ± 1 σ and (b) monthly mean spatial distribution.

Text S3. Sensitivity test for background concentrations

We compared 12 different methods for estimating background CO_2 concentration to select the more appropriate approaches for this study. Following previous studies, we classified these methods into five categories: (1) using background air masses from remote sites, (2) utilizing statistically processed data from high-altitude observation sites, (3) statistically filtered data from individual site, (4) applying statistical methods to data from all observation sites, and (5) estimating upwind regions based on wind direction and using high-altitude sites in those upwind regions.

For background air mass methods, we selected Minamitorishima (MNM) as the global background and Anmyeondo (AMY) as the regional background. Both sites are greenhouse gas concentration observation sites as the part of the World Meteorological Organization Global Air Watch (WMO GAW) program. The MNM site was installed on Japan island for background CO_2 concentration (Liu et al., 2021; Wada et al., 2007) and AMY site has been operational since 1998 to measure background atmospheric conditions unaffected by greenhouse gas emissions from South Korea (Cho et al., 2007; Lim et al., 2012). Data from both sites were obtained from the World Data Centre for Greenhouse Gases. For high-altitude urban background estimates, we used NST_H and GWA, which have high observation heights above ground level (336 m and 231 m, respectively). For these high-altitude sites, we derived background concentrations using either daily minimum values (Fasoli et al., 2018; Lian et al., 2023; Zhao et al., 2009) or 2–day moving average (Mallia et al., 2023; McKain et al., 2012). Additionally, for GWA, which has long-term measurement recording since 2009, we also calculated background concentration by filtering data using curve-fitting methods (NOAA, 2021). For methods using individual observation sites, we applied daily minimum values or the 24–hour moving 5^{th} percentile (Chandra et al., 2016; Sim and Jeong, 2025). We also used the 5^{th} percentile across all observation sites collectively (He et al., 2024). For wind direction-based methods, we dynamically selected upwind sites based on hourly wind direction simulated by WRF at each observation site, following the approach of Lauvaux et al. (2016). When the wind direction was between 0° and 90° or 270° and 360° (northern sector), we assumed the upwind region was from the north and selected NST_H (located in norther Seoul at a high inlet height) as the background site. Conversely, when the wind direction was between 90° and 180° or 180° and 270° (southern sector), we assumed the upwind region was from the south and selected GWA (located in southern Seoul at a high inlet height) as the background site. For the selected upwind site, we applied either the 5^{th} percentile within a 24–hour moving window or the daily minimum values as the background concentrations. These background concentration methods are summarized in Table S4.

Table S4. 12 methods for background concentrations

Classification		**Methods**	Abbreviation
Background air mass	**Global background air**	Hourly mixing ratio at MNM	mnm_h
	Regional background air	Hourly mixing ratio at AMY	amy_h
High-altitude sites		NST_H daily minimum	nsh_dm

	NST_H minimum after 2-day moving average	nsh_2dma
	GWA daily minimum	gwa_dm
	GWA minimum after 2-day moving average	gwa_2dma
	GWA curve fitting (2009-2022)	gwa_ccg
Individual observation sites	Daily minimum at each site	each_dm (*_dm)
	5th percentile within a 24 hours window per site	each_24hm5p (*_24hm5p)
All observation sites	5th percentile across all sites	all_5p
Wind direction based upwind high-altitude sites	Northern sector (0~90°, 270~360°): 5th percentile within a 24 hours window at NST_H	dir_24hm5p
	Southern sector (90~180°, 180~270°): 5th percentile within a 24 hours window at GWA	
	Northern sector (0~90°, 270~360°): daily minimum at NST_H	dir_dm
	Southern sector (90~180°, 180~270°): daily minimum at GWA	

We statistically evaluated the differences between observed ΔCO_2 ($\Delta CO_{2,obs}$) derived using the 12 background concentration methods and modeled $\Delta CO_{2,mod}$ from two different priors (Hs_{ODIAC}, Hs_{KCP}). We considered median value of $\Delta CO_{2,obs}$ - $\Delta CO_{2,mod}$, MBE, Pearson correlation coefficient (r), root mean squared error (RMSE), IOA, and mean percent error (MPE) to identify the three most appropriate background concentration methods for this study. The statistical metrics are summarized in Table S5 and Table S6.

According to this analysis, we selected three methods as most appropriate: (1) each_24hm5p, (2) nsh_dm, (3) gwa_2dma. The each_24hm5p method consider site-specific background concentrations, with median differences close to zero and reasonable MBE and RMSE values. However, this method showed lower IOA values, suggesting potentially reduced consistency for outliers. The nsh_dm method exhibited relatively high correlation and IOA values, suggesting great agreement between modeled and observed values. The gwa_2dma method showed significantly low MBE and RMSE values, indicating high accuracy when comparing $\Delta CO_{2,obs}$ and $\Delta CO_{2,mod}$. The MPE result revealed that systematic biases in each method. nsh_dm revealed that $\Delta CO_{2,mod}$ were higher than $\Delta CO_{2,obs}$, while both gwa_2dma and each_24hm5p showed the opposite tendency, with $\Delta CO_{2,mod}$ lower than $\Delta CO_{2,obs}$.

Table S5. Comparison of observed and modeled ΔCO_2 using ODIAC as prior emissions across 12 background concentration methods.

Abbreviation	Median	Coeff r	MBE	RMSE	IOA	MPE (%)
mnm_h	24.32	0.45	26.26	32.04	0.53	220923
amy_h	8.88	0.32	9.95	21.31	0.56	79504
nsh_dm	5.96	0.46	8.91	20.23	0.63	-22649

nsh_2dma	-7.00	0.40	-5.87	18.21	0.59	-127381
gwa_dm	12.82	0.46	15.08	23.98	0.60	101818
gwa_2dma	2.51	0.40	3.85	18.25	0.61	7383
gwa_ccg	13.44	0.44	16.44	25.22	0.59	74055
each_dm (*_dm)	2.63	0.24	4.84	18.76	0.53	101824
each_24hm5p (*_24hm5p)	1.93	0.19	3.03	18.47	0.49	90189
all_5p	12.71	0.44	15.34	24.61	0.59	48751
dir_24hm5p	6.75	0.56	13.95	29.71	0.59	-22236
dir_dm	8.07	0.54	15.07	31.54	0.58	-17200

Table S6. Comparison of observed and modeled ΔCO_2 using KCP as prior emissions across 12 background concentration methods.

Abbreviation	Median	r	MBE	RMSE	IOA	MPE (%)
mnm_h	22.73	0.45	24.28	30.97	0.55	920076
amy_h	7.76	0.34	7.97	21.34	0.58	313457
nsh_dm	4.67	0.48	6.93	19.95	0.66	-108223
nsh_2dma	-8.10	0.42	-7.85	19.73	0.60	-552238
gwa_dm	10.83	0.47	13.10	23.21	0.63	400688
gwa_2dma	1.24	0.42	1.87	18.66	0.63	16542
gwa_ccg	12.82	0.46	14.46	24.38	0.62	319879
each_dm (*_dm)	1.80	0.27	32.86	19.50	0.54	400489
each_24hm5p (*_24hm5p)	1.18	0.21	1.05	19.54	0.50	349147
all_5p	11.45	0.45	13.36	23.87	0.62	212052
dir_24hm5p	5.12	0.52	11.91	29.22	0.61	-497681
dir_dm	6.84	0.51	13.57	30.96	0.60	-29374

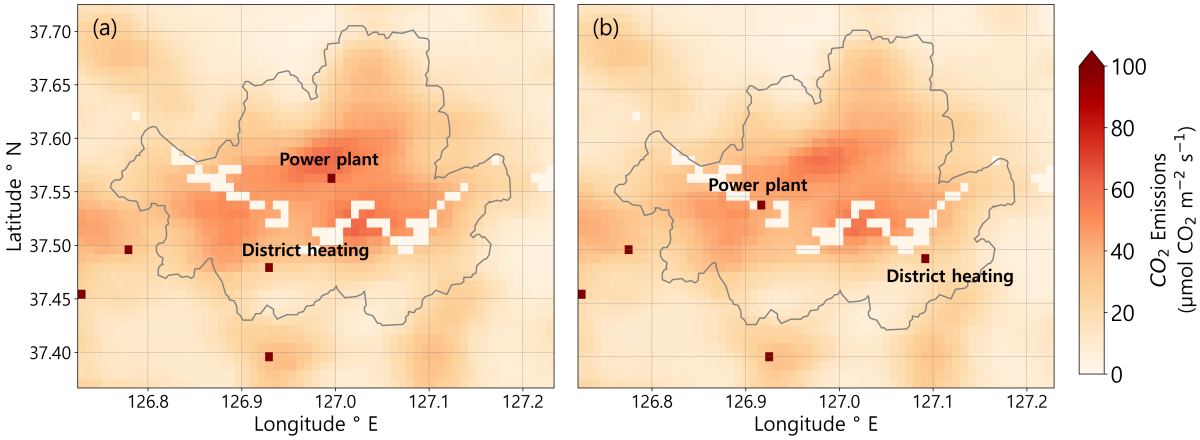


Figure S8. A spatial distribution of ODIAC emissions over Seoul for December 2021, showing (a) original emissions with misallocated point sources and (b) corrected emissions after relocating point sources.

Text S4. Prior error covariance

Text S4.1 Uncertainties of Prior Emissions

Figure. S9 shows the uncertainties (σ) of prior emissions KCP and ODIAC. KCP uncertainties are approximately 14.87 % of the mean emissions, with values of 5.5 (± 4.0) μmol CO_2 m^{-2} s^{-1} and a total monthly daytime uncertainty of 92.9 Gg CO_2. In contrast, the ODIAC daytime emission uncertainty (±1 σ) represents 81.43 % of the mean daytime emissions, with values of 28.5 (± 14.2) μmol CO_2 m^{-2} s^{-1}, corresponding to a total December 2021 daytime uncertainty of 480.62 Gg CO_2.

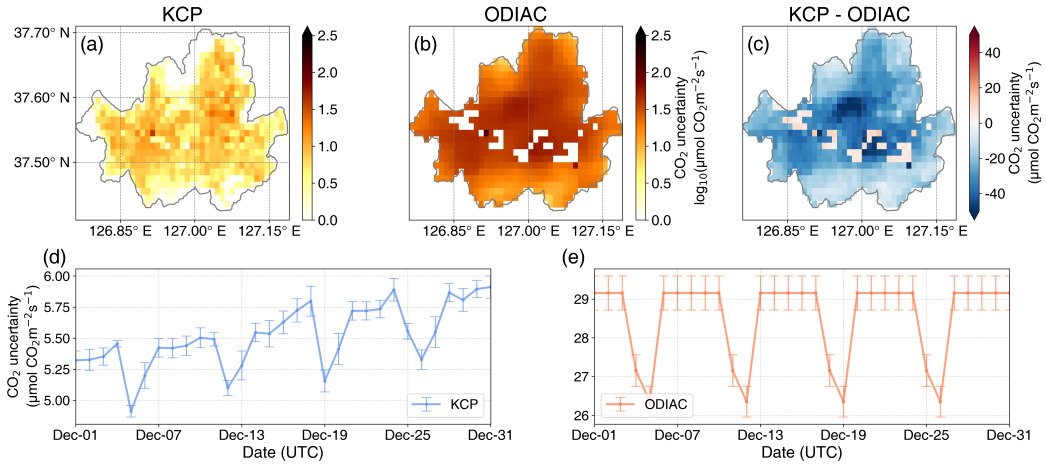


Figure S9. Daytime (3-7 UTC) emission uncertainties for December 2021. Spatial distributions of (a) KCP uncertainties, (b) ODIAC uncertainties, (c) their difference (KCP-ODIAC), and daily mean uncertainty time series for (d) KCP and (e) ODIAC with error bars representing ± 1 σ.

KCP uncertainties are proportional with emission magnitudes across most regions. The spatial distribution of ODIAC uncertainties shows very low values (1 μmol CO_2 m^{-2} s^{-1}) over the Han River, while else where the uncertainties are proportional to emission levels. As shown in Fig. S9c, KCP uncertainties are higher than ODIAC uncertainties in grid cells classified as rivers in the ODIAC dataset, while KCP shows significantly lower than ODIAC elsewhere.

Figure. S9d and e exhibit the time series of emission uncertainties. KCP uncertainties vary between 4.75 to 6 μmol CO_2 m^{-2} s^{-1}, while ODIAC uncertainties fluctuate between 26 to 30 CO_2 m^{-2} s^{-1}. KCP emission uncertainties shows an increasing trend, whereas ODIAC maintains relatively constant uncertainty levels throughout the entire month. The day-of-week patterns in emission uncertainties are consistent with those in emissions, resulting in similar temporal variations in uncertainties between the two datasets.

Text S4.2 Temporal and spatial correlation length scales

In this study, we derived the temporal correlation length (l_τ) through autocorrelation analysis and the spatial correlation length (l_s) through variogram analysis of each emission inventory. Autocorrelation analysis is a technique for analyzing temporal correlation in time series data by considering all correlations between different time points as covariances to assess the degree of temporal correlation (Hartmann, 2016). The minimum time at which the autocorrelation converges to zero is considered the l_τ, representing the time scale over which correlation exist (Kunik et al., 2019; Mallia et al., 2020).

Spatial correlation was assessed using variogram analysis, a widely applied geostatistical method for characterizing spatial dependence (Yoo et al., 2008). The spatial correlation length (l_s) was defined as the minimum lag distance at which the semivariance reaches a sil, indicating the distance beyond which spatial correlation becomes negligible (Kunik et al., 2019; Mallia et al., 2020).

To determine the most appropriate variogram model for the KCP and ODIAC uncertainty datasets, several theoretical models were fitted to the experimental variogram. Model performance was evaluated using the coefficient of R^2 calculated between the experimental semivariances and the model-fitted semivariances at the corresponding lag distances. As summarized in Table S7, the J-bessel variogram model achieved the highest R^2 values for both inventories and was therefore selected.

Table S7. A Selection of the best-fitting variogram model based on R^2

Model type	R^2 with ODIAC	R^2 with KCP
Gaussian	0.948	0.375
Exponential	0.884	-
Matern	0.948	0.529
Stable	0.948	-
Rational	0.947	-
Circular	-	-
Spherical	-	-
Super Spherical	-	-
JBessel	0.997	0.965

The result indicates that KCP exhibits a temporal correlation length (l_τ) of 5 hours and the spatial correlation length (l_s) of 7 km. For ODIAC, l_τ is 9 hours and l_s is 8 km. The l_s values were rounded toward the converged values. Fig. S10 exhibits the variogram and autocorrelation analysis for ODIAC, while Fig. S11 shows the corresponding results for KCP.

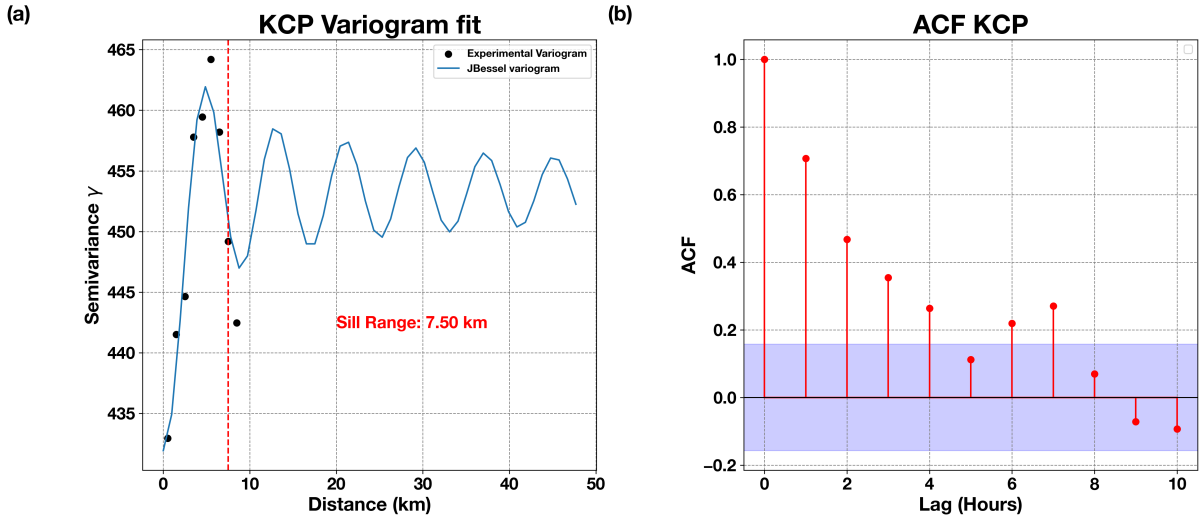


Figure S10. Covariance correlation length scales. (a) A variogram of daytime (3-7 UTC) KCP uncertainty over Seoul domain for December 2021. (b) A lagged autocorrelation function of KCP uncertainty for daytime (3-7 UTC) in December 2021 over the Seoul domain.

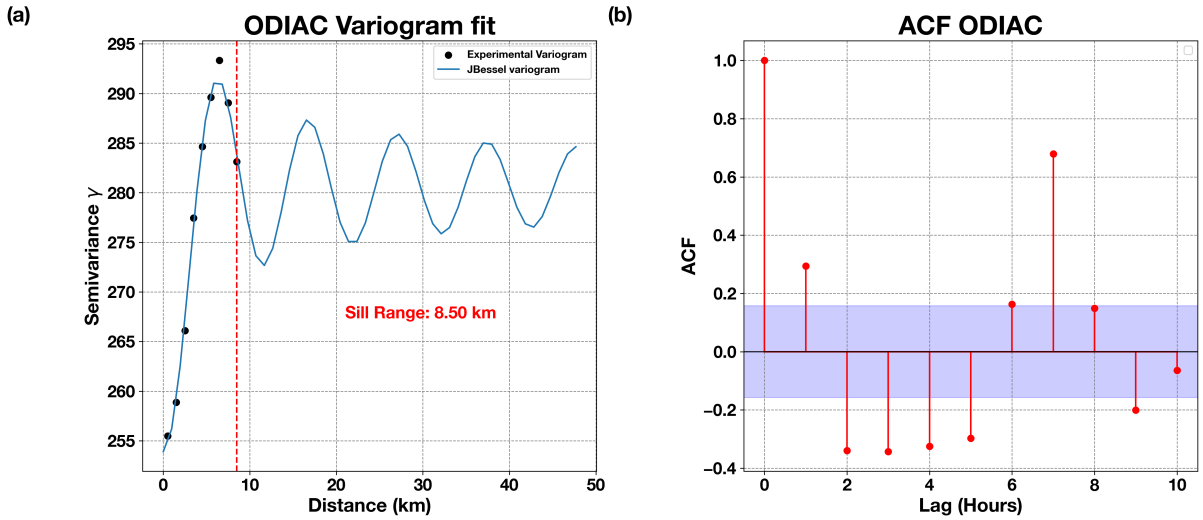


Figure S11. Covariance correlation length scales. (a) A variogram of daytime (3-7 UTC) ODIAC uncertainty over Seoul domain for December 2021. (b) A lagged autocorrelation function of ODIAC uncertainty for daytime (3-7 UTC) in December 2021 over the Seoul domain.

Text S5. Observation Error Covariance

Text S5.1 Relative Residual Error Methods (Rres)

Observed and modeled CO_2 enhancements are expressed as follows (Heald et al., 2004):

$$\boldsymbol{z}_{obs} = \boldsymbol{z}_t + \boldsymbol{\epsilon}_I \tag{S4}$$

$$\boldsymbol{z}_{mod} = \boldsymbol{z}_t + \boldsymbol{b} + \boldsymbol{\epsilon}_M + \boldsymbol{\epsilon}_R \tag{S5}$$

Where $\boldsymbol{z}_{obs}$ include true enhancement ($\boldsymbol{z}_t$) with the instrument error ($\boldsymbol{\epsilon}_I$). $\boldsymbol{z}_{mod}$ (denoted as $\mathbf{H} \cdot \boldsymbol{s}_p$ in this study) is the sum of true value, the model bias ($\boldsymbol{b}$), transport error ($\boldsymbol{\epsilon}_M$), and representative error ($\boldsymbol{\epsilon}_R$). The model bias is the mean difference between the subset of observation points and true enhancement ($\boldsymbol{b} = \overline{\boldsymbol{z}_{mod} - \boldsymbol{z}_t}$). For the sufficient number of observations, $\boldsymbol{\epsilon}_M$, $\boldsymbol{\epsilon}_I$, $\boldsymbol{\epsilon}_R$ have mean values of zero, the model bias can be described as $\boldsymbol{b} = \overline{\boldsymbol{z}_{mod} - \boldsymbol{z}_{obs}}$. Thus, the total of error can be obtained as Eq. (S6):

$$\boldsymbol{\epsilon}_{\sum} = \boldsymbol{\epsilon}_I + \boldsymbol{\epsilon}_M + \boldsymbol{\epsilon}_R = \boldsymbol{z}_{mod} - \boldsymbol{z}_{obs} - \boldsymbol{b} \tag{S6}$$

The RRSD can be obtained dividing standard deviation of $\boldsymbol{\epsilon}_{\sum}$ at each site by the mean enhancement value ($\overline{\boldsymbol{z}_{obs}}$). The matrix $\mathbf{R}$ is a diagonal of the value from Eq. (S7):

$$\mathbf{R} = (RRSD \times \boldsymbol{z}_{obs})^2 \tag{S7}$$

Text S5.2 Summation of the Sources of Error (Rsum)

The Rsum approach, based on Mallia et al. (2020), constructs $\mathbf{R}$ by combining distinct error sources given as Eq. (S8):

$$\mathbf{R} = \mathbf{R}_{particle} + \mathbf{R}_{instrument} + \mathbf{R}_{transPBL} + \mathbf{R}_{transWIND} + \mathbf{R}_{\text{background}} + \mathbf{R}_{biospheric} \tag{S8}$$

$\mathbf{R}_{particle}$ represents the error derived from the release of a finite number of particles within trajectory ensemble. Following the Mallia et al. (2015) we determined this error as the standard deviation of footprint values by simulating STILT 50 times with the same configuration described in Section 2.3. As shown in Table 1, the six observation stations employ different instruments: PICARRO G2401 analyzers at OLP, GWA and NST_L, and CO_2 sensors Li-850 at NST_H, YSB, and SNU. The instrument error ($\mathbf{R}_{instrument}$) was defined using the 1σ values provided by the manufactures. The PICARRO G2401 has a specified accuracy of 17 ppb (Picarro, 2023), while the LI-COR 850 sensor used in SNUCO2M has an accuracy of 1.5 % of the recorded values (LI-COR, 2018). Errors for $\mathbf{R}_{particle}$ and $\mathbf{R}_{instrument}$ were considered to be uncorrelated between observations.

$\mathbf{R}_{transPBL}$ and $\mathbf{R}_{transWIND}$ are defined as the error related to the simulated vertical PBLH and horizontal wind field of the transport model, respectively. Following Kunik et al. (2019), $\mathbf{R}_{transPBL}$ was calculated as 7 % of the mean ΔCO_2 and $\mathbf{R}_{transWIND}$ was calculated as 35 % with a temporal correlation of 2.8 hours. $\mathbf{R}_{background}$ represents the uncertainty arising from the background concentration estimation methods used in the inversion. This value was computed following Mallia et al. (2020) using RMSE between each of three methods adopted in this study, which were then averaged. Finally, $\mathbf{R}_{bio}$ is the error stem from the biogenic fluxes and was estimated as 25 % of the mean of biospheric enhancement ($\mathbf{H} \cdot \boldsymbol{s}_{bio}$) (Mallia et al., 2020). The diagonal elements of the $\mathbf{R}$ matrix are obtained by summing of all these components.

Text S5.3 Diagnostics of Observation Error (RDx)

Desroziers et al. (2005) suggested a diagnostic method of the matrix $\mathbf{R}$ based on the assumptions that the data assimilation process follows linear estimation theory, the errors comprising $\mathbf{R}$ are independent, and the initial $\mathbf{R}$ value accurately represents the true errors. This method defines the $\mathbf{R}$ through the difference between observed ΔCO_2 and simulated ΔCO_2 using posterior emissions and prior emissions, described as Eq. (S9):

$$\mathbf{R} = E\left[(\boldsymbol{z} - \mathbf{H}\cdot\hat{\boldsymbol{s}})(\boldsymbol{z} - \mathbf{H}\cdot\boldsymbol{s}_p)^T\right] = \frac{1}{N-1}\sum_{n=1}^{N}[\{(\boldsymbol{z} - \mathbf{H}\cdot\boldsymbol{s}_p)_n - \overline{(\boldsymbol{z} - \mathbf{H}\cdot\boldsymbol{s}_p)}\}\{(\boldsymbol{z} - \mathbf{H}\cdot\hat{\boldsymbol{s}})_n - \overline{(\boldsymbol{z} - \mathbf{H}\cdot\hat{\boldsymbol{s}})}\}] \quad \text{(S9)}$$

Where N is the number of observation vector for each observation site subset. In this study, we conducted the inversion with $\mathbf{R}$ estimated using the Rsum method and calculated observational error values for each site. Following Bréon et al. (2015) and Sim and Jeong (2025), the matrix $\mathbf{R}$ is diagonal, assuming spatially uncorrelated observational errors.

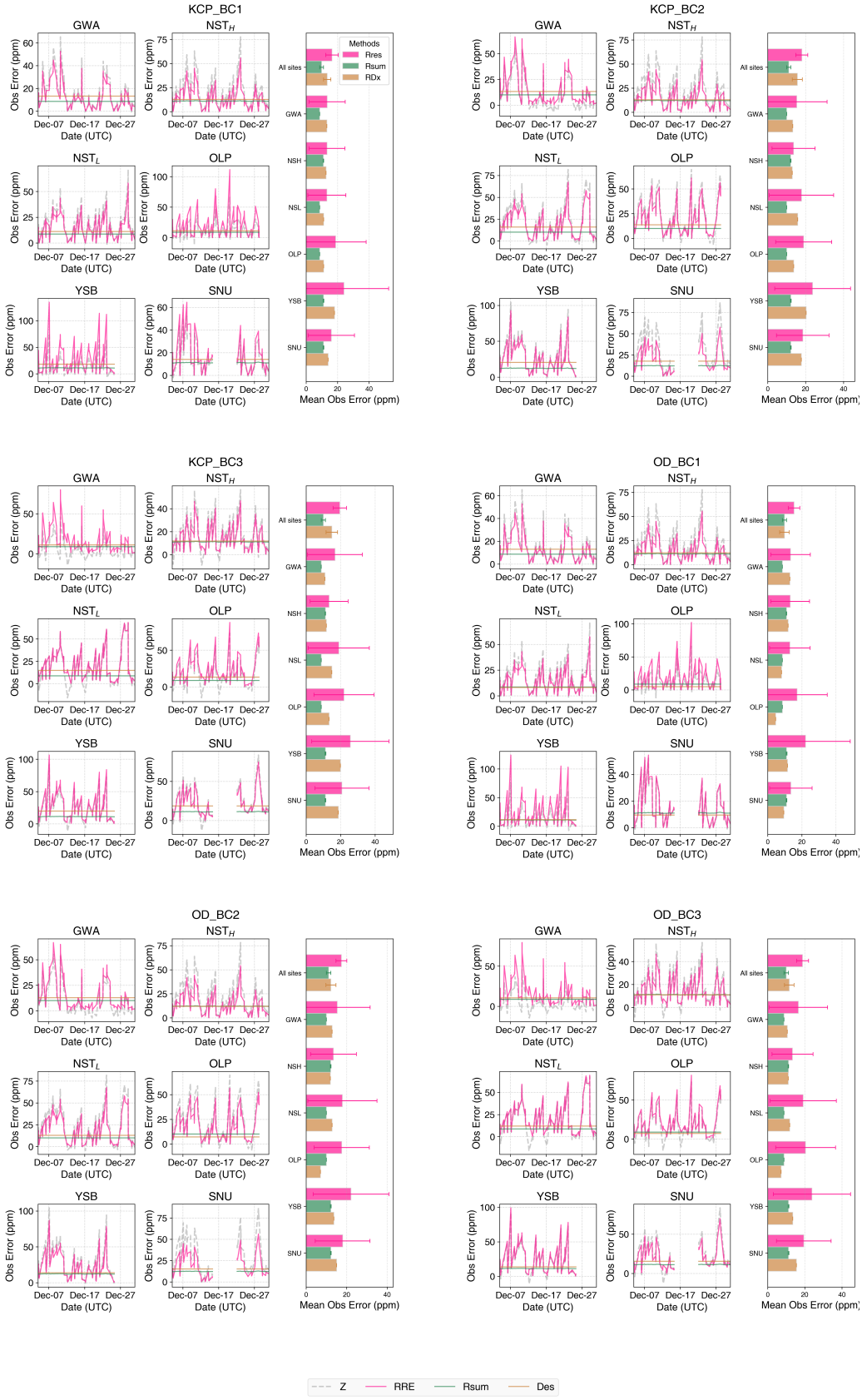


Figure S12. Diagonal elements of observation error covariance for six combinations of prior emissions and background concentration methods. Time series show hourly variations at each site, and bar plots show site-averaged values with error bars representing ± 1 σ.

Text S6. Statistical Analysis of Parameter Assumption Effects on Posterior Estimates

Text S6.1 Influence on Total Emissions

For the daytime total emissions, we evaluated the impact of each parameter assumption on the domain-integrated fluxes using a linear regression model with categorical predictor variables, equivalent to analysis of variance (ANOVA). The main effect regression model can be expressed as Eq. (S10):

$$Y = \beta_0 + \beta_1 P + \beta_2 B + \beta_3 O + \varepsilon \quad \text{(S10)}$$

Where Y represents the daytime total posterior emissions, P, B, and O denote categorical predictors corresponding to the prior inventory, background concentration method, and observation error assumption, respectively. The predictors were implemented through dummy variables, with β_0 representing the reference configuration and β_1-β_3 representing the main effects of each factor and the residual error term ε. The statistical significance of both main effects and interaction effects was evaluated using F-statistic and associated p-values from the regression model.

Effect sizes were quantified using partial eta squared (η_p^2), defined as Eq. (S11), which represents the proportion of variance explained by each parameter assumption (Keppel, 1991; Lakens, 2013):

$$\eta_p{}^2 = \frac{SS_{effect}}{SS_{effect}+SS_{error}} \quad \text{(S11)}$$

Where SS_{effect} represents the sum of squares associated with the subset which has the same parameter assumptions and SS_{error} denotes the residual sum of squares. Larger values indicate that a given assumption explains a larger proportion of the variance in the inversion results, independent of sample size.

To examine whether the influence of one parameter assumption depended on another, an additional model including two-way interaction term was fitted as Eq. (S12):

$$Y = \beta_0 + \beta_1 P + \beta_2 B + \beta_3 O + \beta_4 (P \times B) + \beta_5 (P \times O) + \beta_6 (B \times O) + \varepsilon \quad \text{(S12)}$$

Where β_4-β_6 represent the two-way interaction effects. Since the number of inversion configurations was limited (N=18), this interaction analysis was used primarily as a supplementary diagnostic to identify potential dependencies among assumptions rather than as the primary basis of interpretation.

Text S6.2 Influence on Spatiotemporal Variability

For the spatiotemporal variability, we quantified the similarity between inversion results using correlation-based metrics. Similarity of temporal variability was characterized using the Pearson correlation coefficient (r) between daytime-averaged daily emission time series, while the similarity of spatial distribution was evaluated using the correlation coefficient between spatial emission vectors aggregated over daytime time steps. To isolate the influence of each parameter, we compared correlations between inversion results sharing the same assumption and those using different assumptions. For each inversion case i, the mean correlation for the same assumption ($\bar{r}_i^{same}$) and the mean correlation for different assumptions ($\bar{r}_i^{diff}$) were calculated, and their difference was defined as Eq. (S13):

$$d_i = \overline{r}_i^{same} - \overline{r}_i^{diff} \quad \text{(S13)}$$

The statistical significance of these differences (d_i) was evaluated using t-tests across inversion cases. For the total emission analysis, effect sizes (η_p^2) were used to quantify the influence of parameter assumptions. For the similarity of spatiotemporal variability, the magnitude of the effect was quantified using relative change, calculated as Eq. (S14):

$$Relative\ change\ (\%) = \frac{\bar{r}^{same} - \bar{r}^{diff}}{\bar{r}^{diff}} \times 100 \quad \text{(S14)}$$

In addition, to evaluate how consistently these effects were showed across the 18 inversion cases, we calculated Cohen's d_z, defined as Eq. (S15) (Cohen, 1988; Lakens, 2013):

$$d_z = \frac{\bar{d}}{s_d} \quad \text{(S15)}$$

Where $\bar{d}$ and s_d represent the mean and standard deviation of the within-inversion differences d_i, respectively. Larger values of Cohen's d_z indicate that the effect of differences in assumptions within each factor is more consistently reproduced across inversion cases.

To evaluate interaction effects on temporal variability and spatial distribution between parameter assumptions, interaction contrasts were additionally constructed. For a pair of factor A and B, inversion pairs were sperate in four groups according to whether the two factors shared the same assumption (1) or different assumptions (0). For each inversion case i, the effect of factor A was first calculated separately under two conditions. When factor B shared the same assumption and when factor B used different assumptions. The interaction contrast was then defined as the difference between these two effects as shown in Eq. (S16):

$$I_i = \left(\overline{r_{11,i}} - \overline{r_{01,i}}\right) - \left(\overline{r_{10,i}} - \overline{r_{00,i}}\right) \quad \text{(S16)}$$

Here $\left(\overline{r_{11,i}} - \overline{r_{01,i}}\right)$ represents the effect of factor A under the same assumption condition for factor B, while the second term $\left(\overline{r_{10,i}} - \overline{r_{00,i}}\right)$ indicates the effect of factor A under different assumption condition for factor B. The statistical significance, effect size, and consistency of I_i were evaluated using the same procedure as the main-effect. As with the ANOVA, this interaction analysis was also used as a supplementary diagnostic to examine potential dependencies among parameter assumptions.

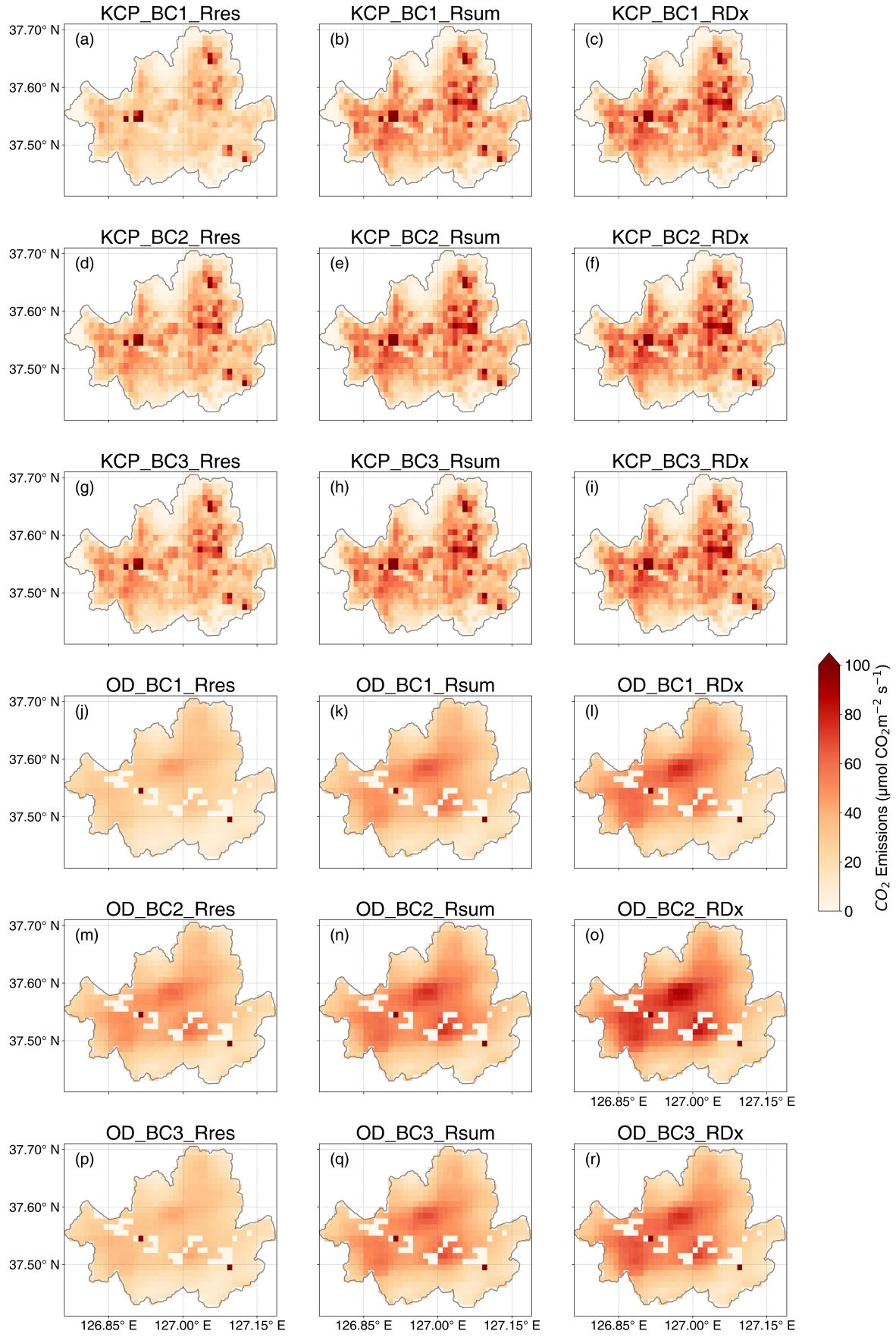


Figure S13. Daytime (3-7 UTC) posterior emission maps from 18 inversion configurations.

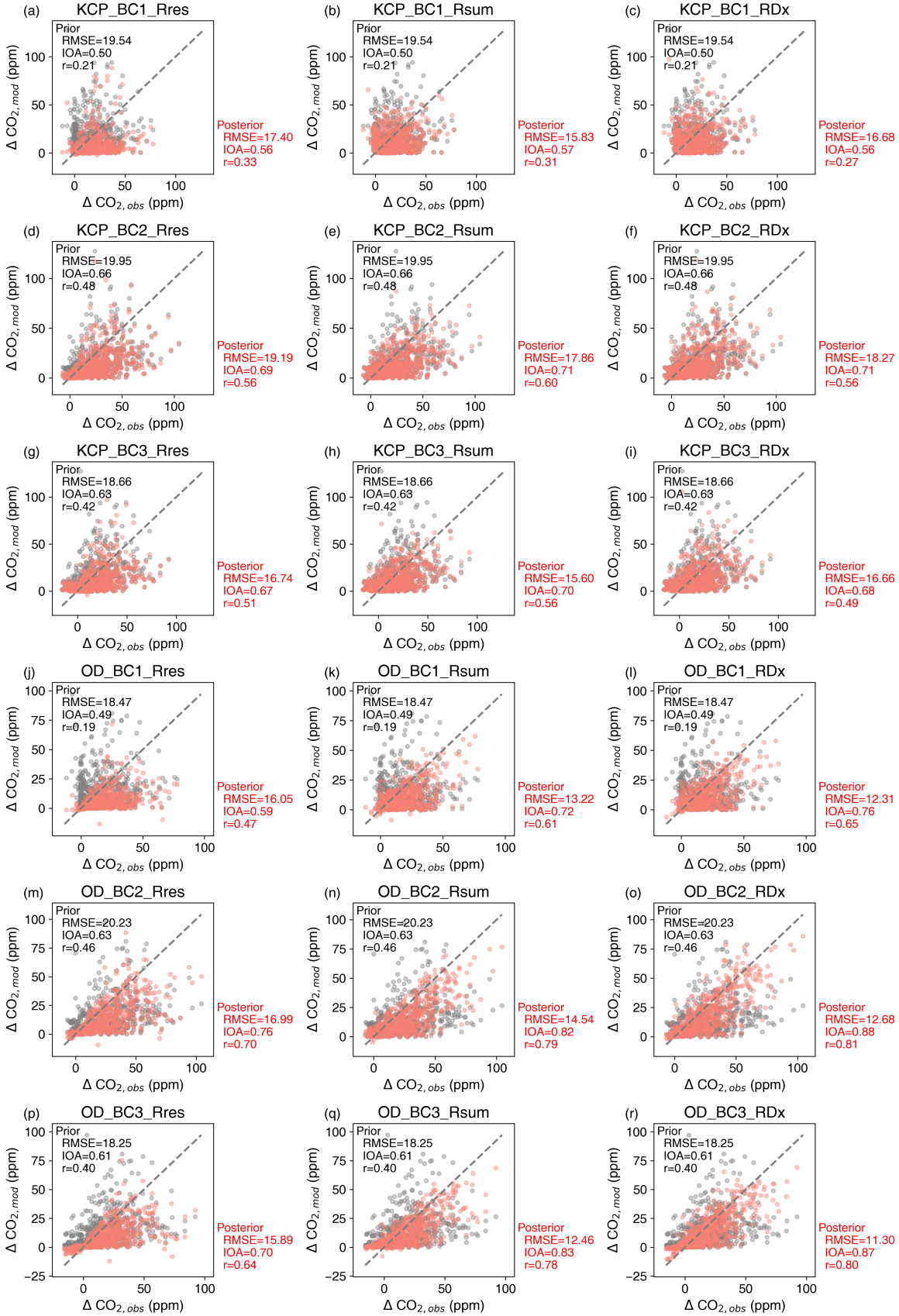


Figure S14. Scatter plots of observed ΔCO_2 ($\Delta CO_{2,obs}$) and simulated ΔCO_2 ($\Delta CO_{2,mod}$) with RMSE, MBE, IOA, and correlation coefficient (r) for priors (gray) and posteriors (red) based on 18 inversion configurations.

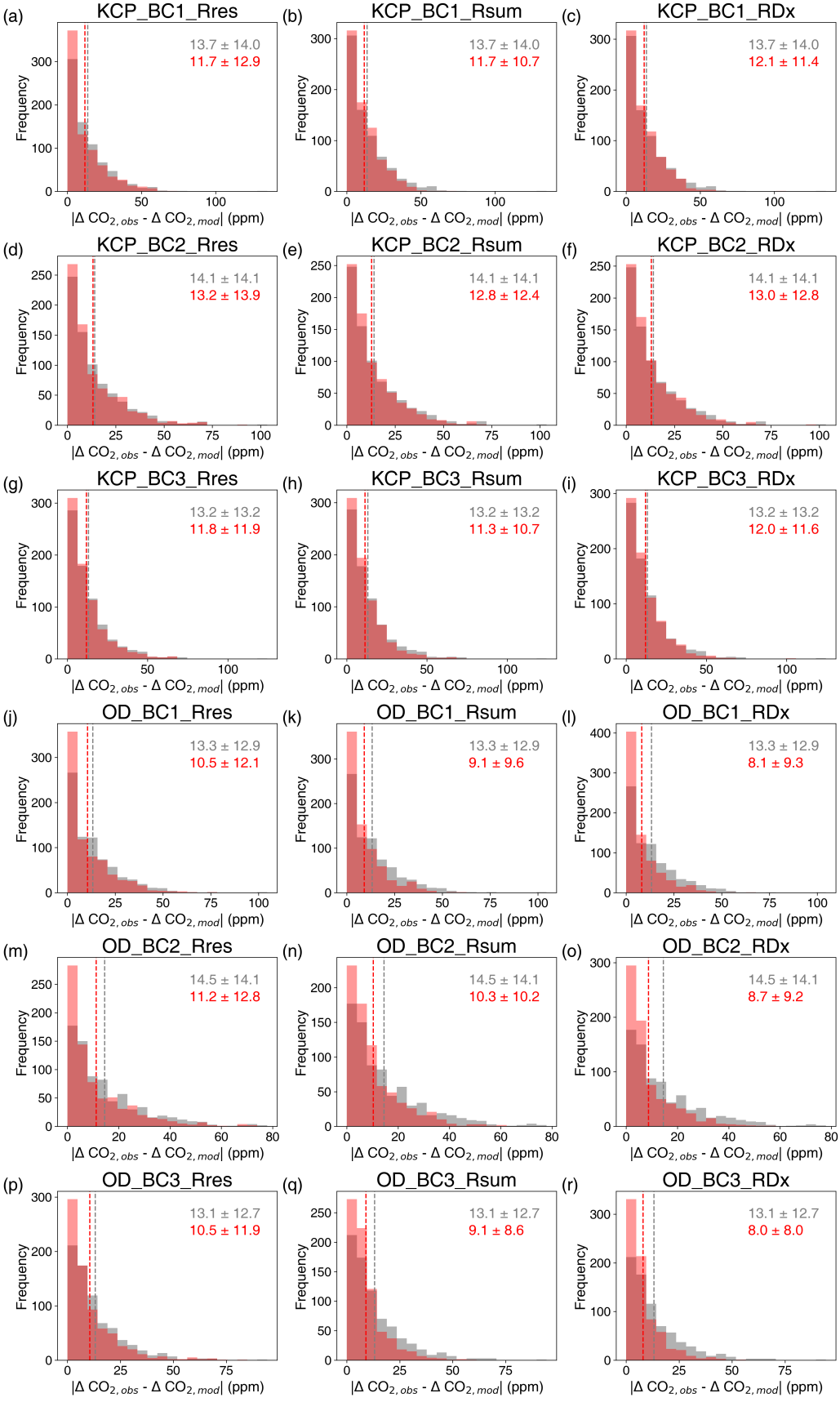


Figure S15. Histograms of absolute differences between observed and modeled ΔCO_2 ($|\Delta CO_{2,\,obs} - \Delta CO_{2,\,mod}|$) for priors (gray) and posteriors (red) based on 18 inversion configurations. MAE and ±1 σ are shown above each plot.

Table S8. Statistical interaction effects of parameter assumptions on posterior emissions. ('Prior' means prior emissions, 'Background' represents background concentrations, and 'ObsErr' indicates observation error).

Information aspect	Parameter	Effect size (η_p^2)	F	p
Total emissions	Prior × Background	0.04	0.14	0.138
	Prior × ObsErr	0.02	0.22	0.217
	Background × ObsErr	0.02	0.44	0.443
		Relative change (%)	Cohen's d_z	p
Temporal variability	Prior × Background	-14.62	-0.57	0.027
	Prior × ObsErr	-14.04	-0.06	0.813
	Background × ObsErr	-1.41	-0.13	0.577
Spatial distribution	Prior × Background	-0.09	-0.07	<0.001
	Prior × ObsErr	1.11	0.36	0.141
	Background × ObsErr	-14.38	-5.86	0.785

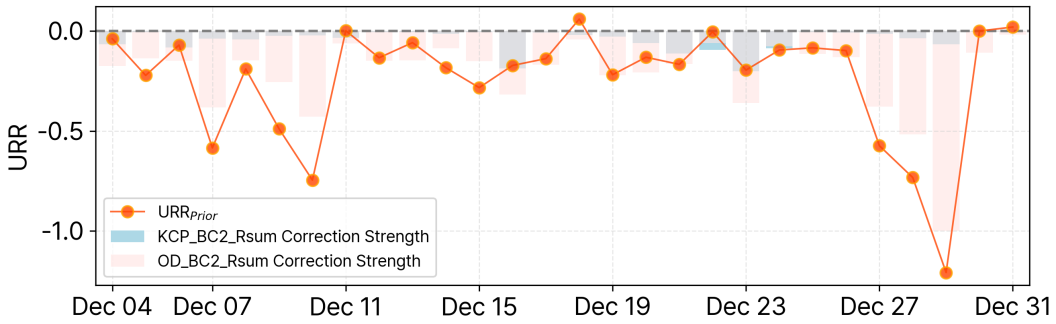


Figure S16. Daily mean Uncertainty Reduction Ratio (URR) time series and correction intensity for different prior emission configurations: KCP_BC3_Rsum and OD_BC3_Rsum. The correction strength (bars) is defined as the normalized daily mean absolute difference between posterior and prior emissions(|posterior-prior|×prior^{-1}).

References

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., Ciais, P., and Meunier, F.: An attempt at estimating Paris area CO2 emissions from atmospheric concentration measurements, Atmos. Chem. Phys., 15, 1707–1724, https://doi.org/10.5194/acp-15-1707-2015 , 2015.

Chandra, N., Lal, S., Venkataramani, S., Patra, P. K., and Sheel, V.: Temporal variations of atmospheric CO_2 and CO at Ahmedabad in western India, Atmos. Chem. Phys., 16, 6153–6173, https://doi.org/10.5194/acp-16-6153-2016 , 2016.

Cho, C. H., Kim, J. S., and Yoo, H. J.: Atmospheric carbon dioxide variations at Korea GAW Center from 1999 to 2006, Asia-Pac. J. Atmos. Sci., 43, 359–365, 2007.

Cohen, J.: Statistical power analysis for the behavioral sciences, Routledge Academic, New York, NY, USA, 1988.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Q. J. Roy. Meteor. Soc., 131, 3385–3396, https://doi.org/10.1256/qj.05.108 , 2005.

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.

Hartmann, D. L.: Matrix methods for analysis of structure in data sets: empirical orthogonal functions, principal component analysis, singular value decomposition, maximum covariance analysis, canonical correlation analysis, etc., ATM 552 Notes, University of Washington, Department of Atmospheric Sciences, available at: https://atmos.washington.edu/~dennis/552_Notes_4.pdf (last access: 6 July 2026), 2016.

He, C., Lu, X., Zhang, Y., Liu, Z., Jiang, F., Sun, Y., Fan, S., and co-authors: Revisiting the quantification of power plant CO_2 emissions in the United States and China from satellite: A comparative study using three top-down approaches, Remote Sens. Environ., 308, 114192, https://doi.org/10.1016/j.rse.2024.114192 , 2024.

Heald, C. L., Jacob, D. J., Jones, D. B., Palmer, P. I., Logan, J. A., Streets, D. G., Sachse, G. W., Gille, J. C., Hoffman, R. N., and Nehrkorn, T.: Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, J. Geophys. Res. Atmos., 109, D23306, https://doi.org/10.1029/2004JD005185 , 2004.

Hong, S. Y. and Lim, J.-O. J.: The WRF single-moment 6-class microphysics scheme (WSM6), Asia-Pac. J. Atmos. Sci., 42, 129–151, 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, D13103, https://doi.org/10.1029/2008JD009944 , 2008.

Jiménez, P. A., Dudhia, J., González-Rouco, J. F., Navarro, J., Montávez, J. P., and García-Bustamante, E.: A revised scheme for the WRF surface layer formulation, Mon. Weather Rev., 140, 898–918, https://doi.org/10.1175/MWR-D-11-00056.1 , 2012.

Kain, J. S. and Fritsch, J. M.: A one-dimensional entraining/detraining plume model and its application in convective parameterization, J. Atmos. Sci., 47, 2784–2802, https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2 , 1990.

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.

Keppel, G.: Design and analysis: A researcher's handbook, Prentice Hall, Englewood Cliffs, NJ, USA, 1991.

Lakens, D.: Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs, Front. Psychol., 4, 863, https://doi.org/10.3389/fpsyg.2013.00863 , 2013.

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., Sweeney, C., Shepson, P. B., Tans, P., and Whetstone, J.: High-resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), Elem. Sci. Anth., 4, 000070, https://doi.org/10.12952/journal.elementa.000070 , 2016.

Lian, J., Lauvaux, T., Utard, H., Bréon, F. M., Broquet, G., Ramonet, M., Xueref-Remy, I., Delmotte, M., and Ciais, P.: Can we use atmospheric CO_2 measurements to verify emission trends reported by cities? Lessons from a 6-year atmospheric inversion over Paris, Atmos. Chem. Phys., 23, 8823–8835, https://doi.org/10.5194/acp-23-8823-2023 , 2023.

Lim, Y. K., Moon, Y. S., Kim, J. S., Song, S. K., and Hong, J. H.: An estimation of mean background concentrations of greenhouse gases observed on Ulleungdo, J. Korean Earth Sci. Soc., 33, 32–38, 2012.

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, 4493, https://doi.org/10.1029/2002JD003161 , 2003.

Liu, S., Feng, Z., Lin, H., Liu, P., Liang, M., Qing, X., Xiong, H., Qiu, S., Li, J., Jiang, K., Hong, H., and Fang, S.: Changes of atmospheric CO_2 in the Tibetan Plateau from 1994 to 2019, J. Geophys. Res. Atmos., 126, e2021JD035299, https://doi.org/10.1029/2021JD035299 , 2021.

LI-COR, Inc.: Using the LI-830 and LI-850 gas analyzers, Publication No. 984-17190, LI-COR, Lincoln, USA, 2018.

Mallia, D. V., Mitchell, L. E., Kunik, L., Fasoli, B., Bares, R., Gurney, K. R., and Lin, J. C.: Constraining urban CO2 emissions using mobile observations from a light rail public transit platform, Environ. Sci. Technol., 54, 15613–15621, https://doi.org/10.1021/acs.est.0c05698 , 2020.

Mallia, D. V., Mitchell, L. E., Gonzalez Vidal, A. E., Wu, D., Kunik, L., and Lin, J. C.: Can we detect urban-scale CO_2 emission changes within medium-sized cities?, J. Geophys. Res. Atmos., 128, e2023JD038686, https://doi.org/10.1029/2023JD038686 , 2023.

McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, Proc. Natl. Acad. Sci. USA, 109, 8423–8428, https://doi.org/10.1073/pnas.1116645109 , 2012.

Morcrette, J.-J., Barker, H. W., Cole, J. N. S., Iacono, M. J., and Pincus, R.: Impact of a new radiation package, McRad, in the ECMWF integrated forecasting system, Mon. Weather Rev., 136, 4773–4798, https://doi.org/10.1175/2008MWR2363.1 , 2008.

National Institute of Meteorological Sciences (NIMS): Construction and validation of WRF-LES-FDDA system, NIMS Tech. Note NIMR-TN-2014-014, Publication No. 11-1360395-000466-01, Seoul, Republic of Korea, 2014.

NOAA: Curve fitting methods applied to time series in NOAA/ESRL/GMD, available at: https://gml.noaa.gov/ccgg/mbl/crvfit/crvfit.html (last access: 6 May 2025), 2021.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Fisher, R. A., Lawrence, P. J., and Muszala, S. P.: Technical description of version 4.5 of the Community Land Model (CLM4.5), NCAR Tech. Note NCAR/TN-503+STR, https://doi.org/10.5065/D6RR1W7M , 2013.

Sim, S. and Jeong, S.: Constraining urban fossil fuel CO_2 emissions in Seoul using combined ground and satellite observations with Bayesian inverse modelling, Atmos. Chem. Phys., 25, 18509–18526, https://doi.org/10.5194/acp-25-18509-2025 , 2025.

Wada, A., Sawa, Y., Matsueda, H., Taguchi, S., Murayama, S., Okubo, S., and Tsutsumi, Y.: Influence of continental air mass transport on atmospheric CO_2 in the western North Pacific, J. Geophys. Res. Atmos., 112, D07311, https://doi.org/10.1029/2006JD007552 , 2007.

Yoo, H.-Y., Lee, K., and Kwon, B.-D.: The optimized analysis zone districting using variogram in urban remote sensing, Korean J. Remote Sens., 24, 107–115, https://doi.org/10.7780/kjrs.2008.24.2.107 , 2008.

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, D16302, https://doi.org/10.1029/2008JD011671 , 2009.