Supplement of Atmos. Chem. Phys., 26, 16125–16149, 2026
https://doi.org/10.5194/acp-26-16125-2026-supplement
© Author(s) 2026. CC BY 4.0 License.



Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

Physics-constrained inverse neural estimation (PINE) of NO_x emission spatial patterns from TROPOMI NO_2 columns over the North China Plain

Yinan Wang et al.

Correspondence to: Yubing Pan (ybpan@ium.cn)

The copyright of individual parts of the supplement might differ from the article licence.

Supplement

This Supplement contains the supplementary figures (Figs. S1–S13), the WRF-Chem configuration (Table S1) and the synthetic test with known emissions (Text S1, Table S2) referenced in the main text.

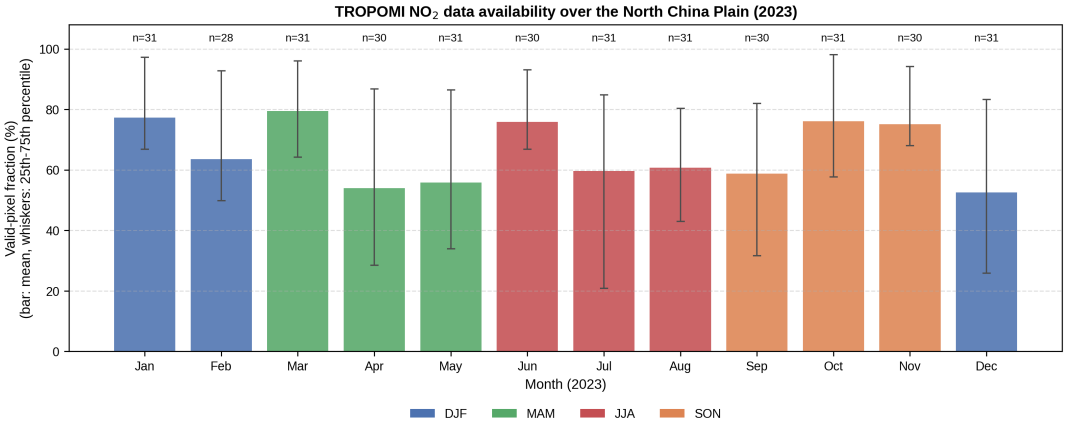


Figure S1. Monthly-mean valid-pixel coverage of TROPOMI (2023, after quality control: qa_value ≥ 0.75, cloud radiance fraction < 0.5). Bars are the monthly mean fraction of the 3621 grid cells with a valid column (52.6 % to 79.5 %); whiskers give the 25th to 75th percentile over days.

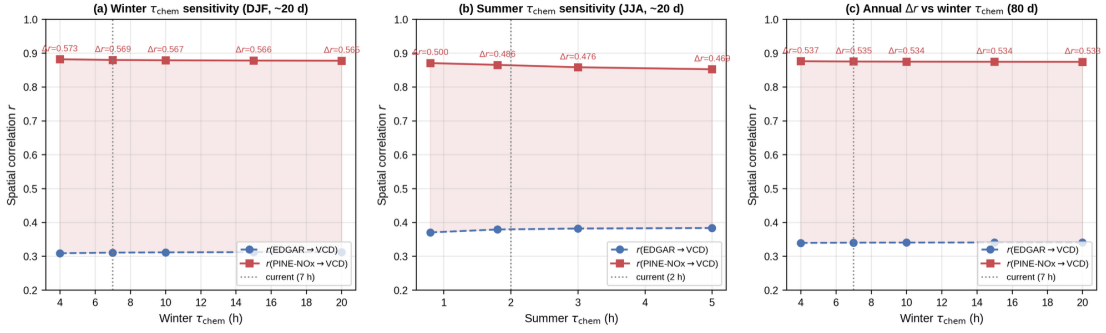


Figure S2. Sensitivity of the reconstruction skill to the chemical lifetime τ_{chem} in the decoder, with the PINE-NOx emissions held fixed. (a) τ_{winter} from 4 to 20 h, DJF days; (b) τ_{summer} from 0.8 to 5 h, JJA days; (c) τ_{winter} against the skill over all 80 sampled days (20 per season). Skill r is the log-space spatial correlation between the simulated and

the observed local column; Δr is the difference between PINE-NOx and EDGAR. Over the winter range the annual Δr varies between 0.533 and 0.537. Dotted lines mark the values used in the decoder.

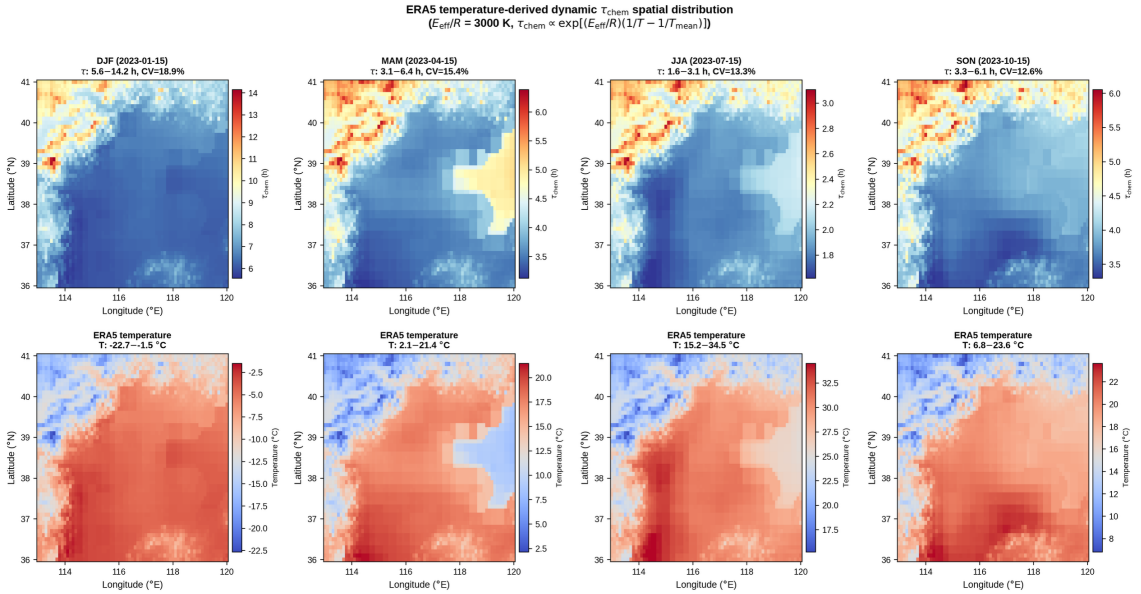


Figure S3. ERA5-temperature-driven dynamic τ_{chem}. Spatial distributions for a representative day in each of the four seasons (top row) and the corresponding ERA5 temperature fields (bottom row). With $E_{eff}/R = 3000$ K, $\tau_{chem} \propto \exp[E_{eff}/R \times (1/T - 1/T_{mean})]$. The seasonal spatial coefficient of variation CV is 18.9% (DJF), 15.4% (MAM), 13.3% (JJA) and 12.6% (SON).

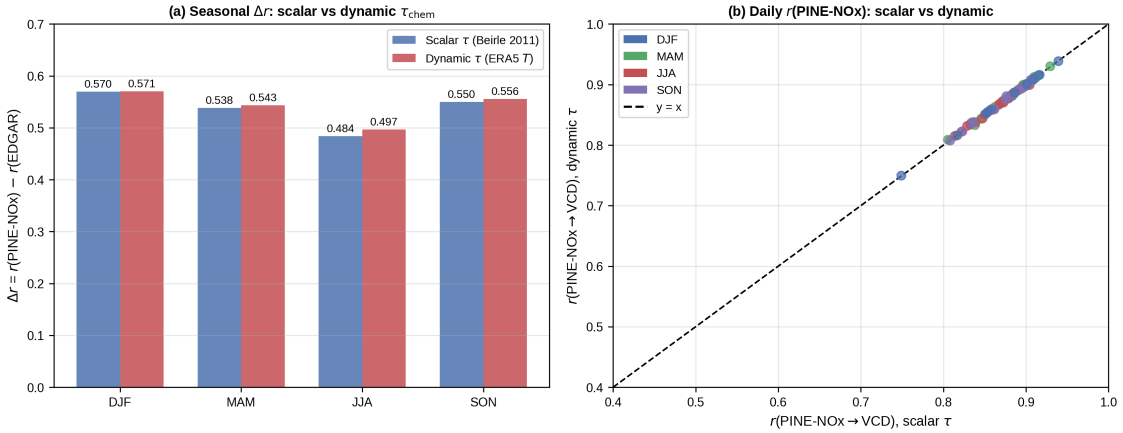


Figure S4. Monthly-mean scalar τ_{chem} against the dynamic τ_{chem} of Fig. S3, 80 sampled days. (a) Seasonal mean Δr for the two schemes; (b) daily reconstruction r of PINE-NOx with the scalar against the dynamic τ_{chem}. The dynamic scheme raises the mean Δr from 0.535 to 0.542, about 1 %.

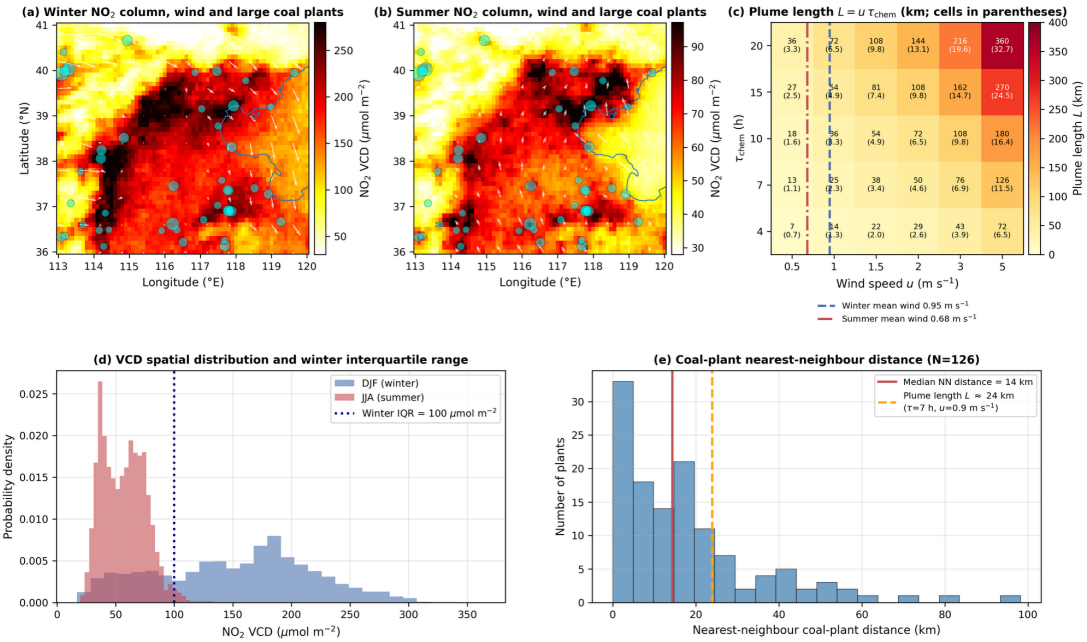


Figure S5. Quantitative assessment of the applicability of the flux-divergence method (FDM) over the North China Plain. (a, b) Winter and summer mean NO_2 column with mean wind and the coal-fired plants of at least 1000 MW (46 of the 126 plants in the domain); (c) plume length $L = u\,\tau_{chem}$ for a range of wind speeds and lifetimes, with the winter and summer mean wind speed (0.95 and 0.68 m s^{-1}); (d) distribution of the winter and summer columns, with the winter interquartile range (100 μmol m^{-2}); (e) nearest-neighbour distance between coal-fired plants (126 plants; median 14 km) against the winter plume length (about 24 km).

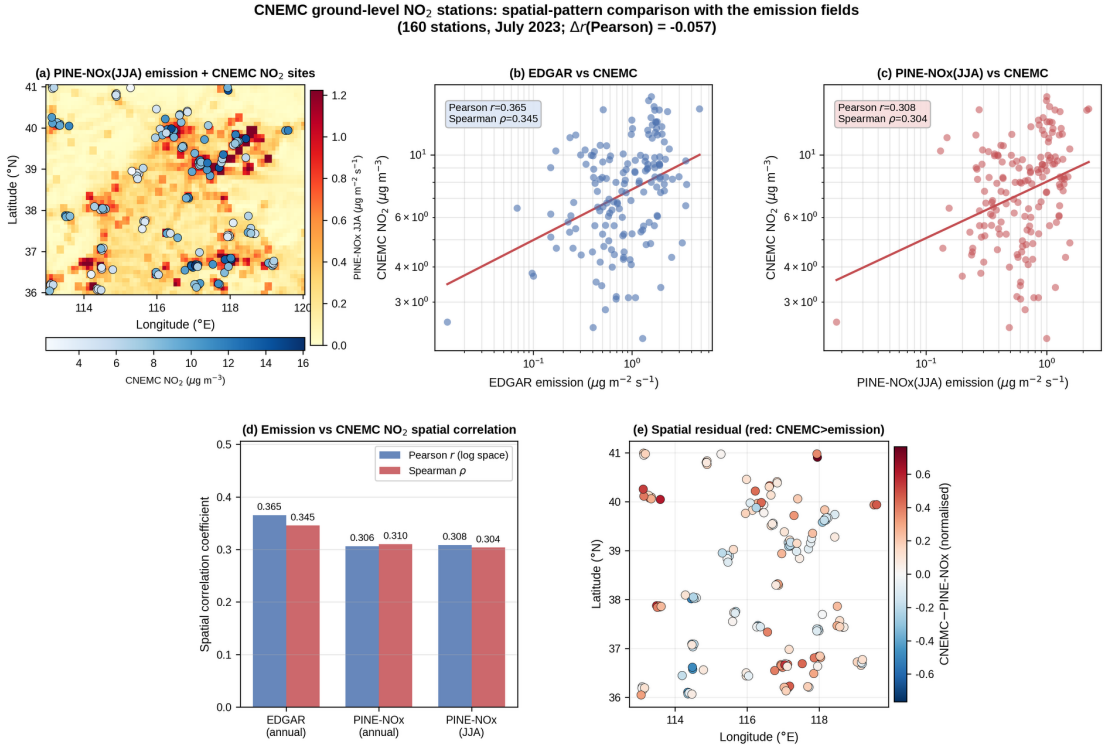


Figure S6. Spatial comparison of CNEMC ground-based stations (n = 160, July 2023) against PINE-NOx and EDGAR emissions in the station cell. (a) PINE-NOx JJA emissions with the station means; (b, c) station NO_2 against EDGAR and against PINE-NOx (JJA) emissions, with log-space Pearson r and Spearman ρ; (d)

correlations for EDGAR (r = 0.365), annual PINE-NOx (0.306) and JJA PINE-NOx (0.308); (e) normalised residual, station NO_2 minus PINE-NOx emission.

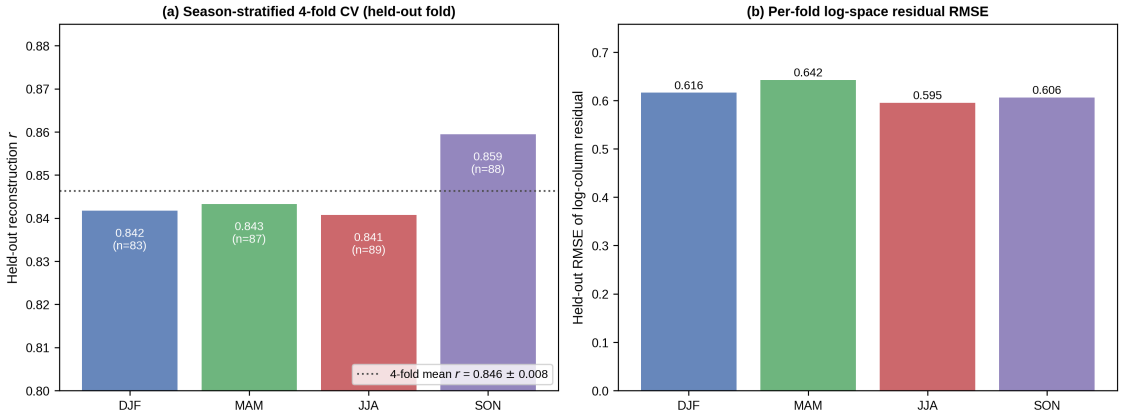


Figure S7. Seasonally stratified four-fold cross-validation (each fold holds out one complete season and retrains the same single-stage PINE-NOx model on the remaining three seasons without any emission label, then evaluates the forward-reconstruction skill on the held-out season that never participated in training). (a) The forward spatial correlation on the held-out season is 0.842 (DJF), 0.843 (MAM), 0.841 (JJA) and 0.859 (SON), with a mean of 0.846 ± 0.008, slightly below the full-set training baseline (r = 0.871). (b) Root-mean-square difference between the simulated and the observed log local column, each demeaned per day, on the held-out season: 0.62 (DJF), 0.64 (MAM), 0.60 (JJA) and 0.61 (SON). The reconstruction skill therefore does not depend on any single season being present in the training set.

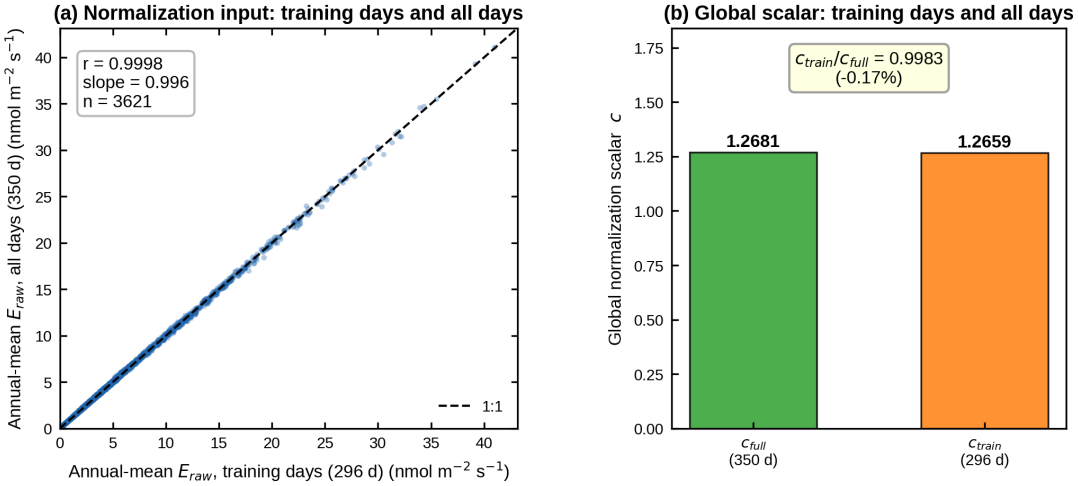


Figure S8. Independence of the domain-wide scalar annual-total normalization. PINE-NOx applies a single domain-wide constant $c = \Sigma E_{EDGAR} / \Sigma \langle E_{raw} \rangle_t = 1.268$ (1.266 when computed from the 296 training days only; panel b) to the network's raw output E_{raw} to normalize the annual total; the annual-mean raw emissions from the training days and from all 350 days agree (r = 0.9998, slope 0.996, n = 3621 cells; panel a). Because c is a single multiplicative constant, the log-space correlation and the demeaned log-RMSE are strictly unaffected by its value ($\log(cE) = \log E + \log c$ is merely a constant shift, which changes neither the Pearson correlation nor the demeaned residual), so the normalization introduces no information leakage under any train/validation split. The agreement of the annual total with EDGAR is therefore a design constraint on the absolute magnitude rather than an independent validation metric.

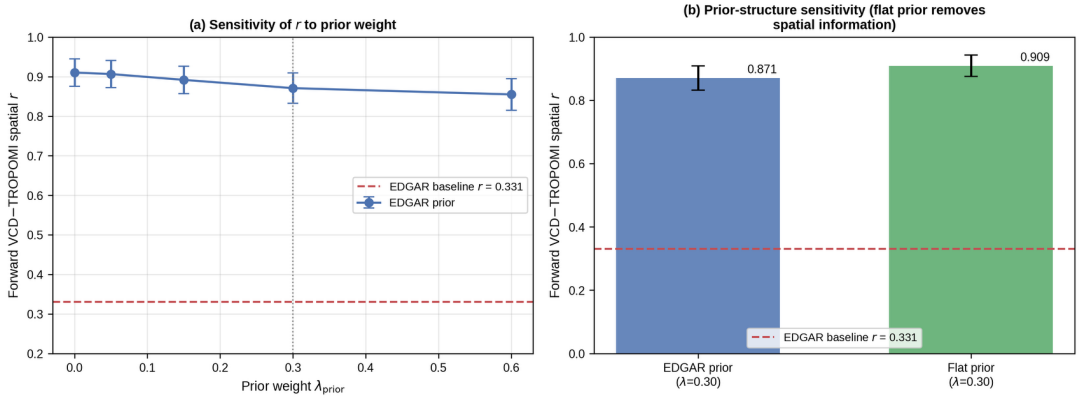


Figure S9. Sensitivity to the prior. (a) Reconstruction r against the prior weight λ_{prior} (0, 0.05, 0.15, 0.30 and 0.60; each a full retraining): 0.910, 0.907, 0.892, 0.871 and 0.855; the value used is 0.30. (b) Replacing the Huber reference prior from the real EDGAR spatial field with a uniform prior stripped of all spatial information (flat prior; domain-wide constant, equal total) and retraining within the same single-stage PINE-NOx framework: under the uniform prior the spatial correlation between the network's inverted forward VCD and TROPOMI is r = 0.909, against r = 0.871 with the EDGAR prior (EDGAR prior through the same decoder: r = 0.331). Points and bars are the mean and the standard deviation over 347 days. The reconstruction skill therefore does not rest on the prior; the emission pattern itself does depend on it (Sect. 3.1).

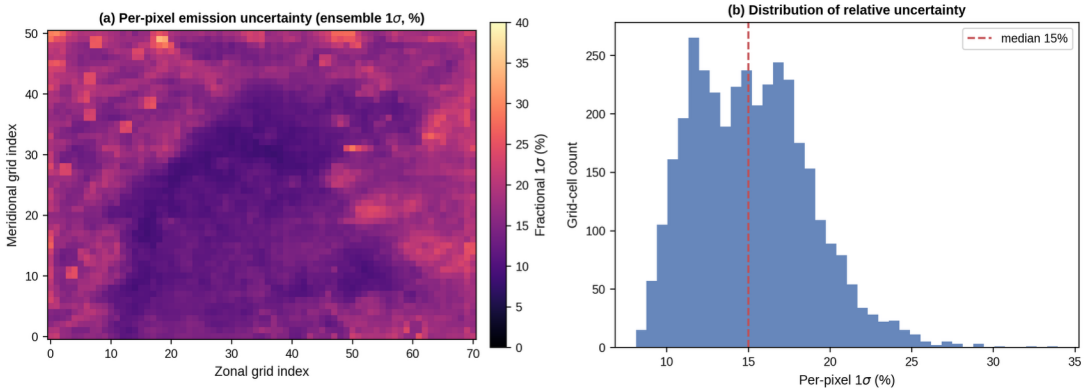


Figure S10. Ensemble epistemic-uncertainty quantification (seasonally stratified four-fold cross-validation). A pixel-wise ensemble across the four held-out-fold models gives a relative emission uncertainty (1σ/mean) with a median of about 15 % (interquartile range 12.3 to 17.5 %; per-cell standard deviation of the natural logarithm of the emission share, averaged over 350 days). The four-fold held-out forward correlation is r = 0.846 ± 0.008. This shows that the local spatial allocation carries appreciable uncertainty; the domain total is fixed by the normalization and is not tested here.

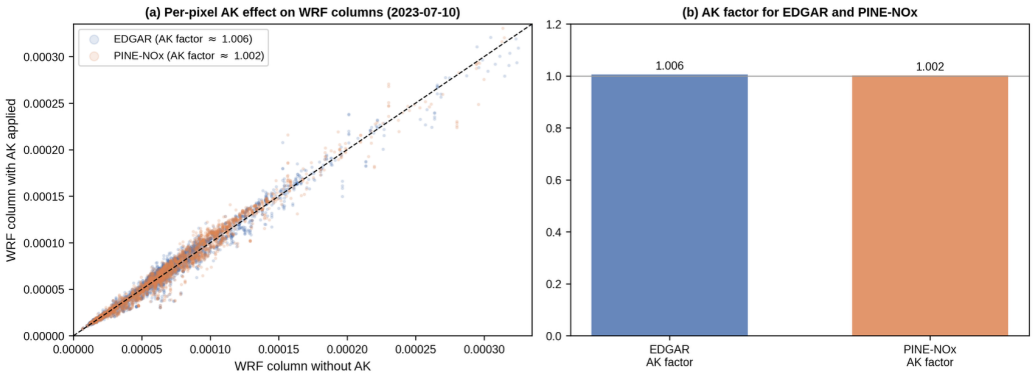


Figure S11. Proof-of-concept averaging-kernel (AK) correction of TROPOMI (overpass on 10 July 2023, 2896 grid cells with a matched pixel). (a) WRF column with the tropospheric AK applied against the column without it, for the EDGAR and the PINE-NOx run; (b) mean ratio of the two, 1.006 and 1.002 (0.894 and 0.890 over 8–31 July); the spatial-pattern correlation between the EDGAR and the PINE-NOx column changes from 0.971 to 0.975. The averaging kernel therefore acts almost identically on the two runs, so it does not change which of them is closer to the observations.

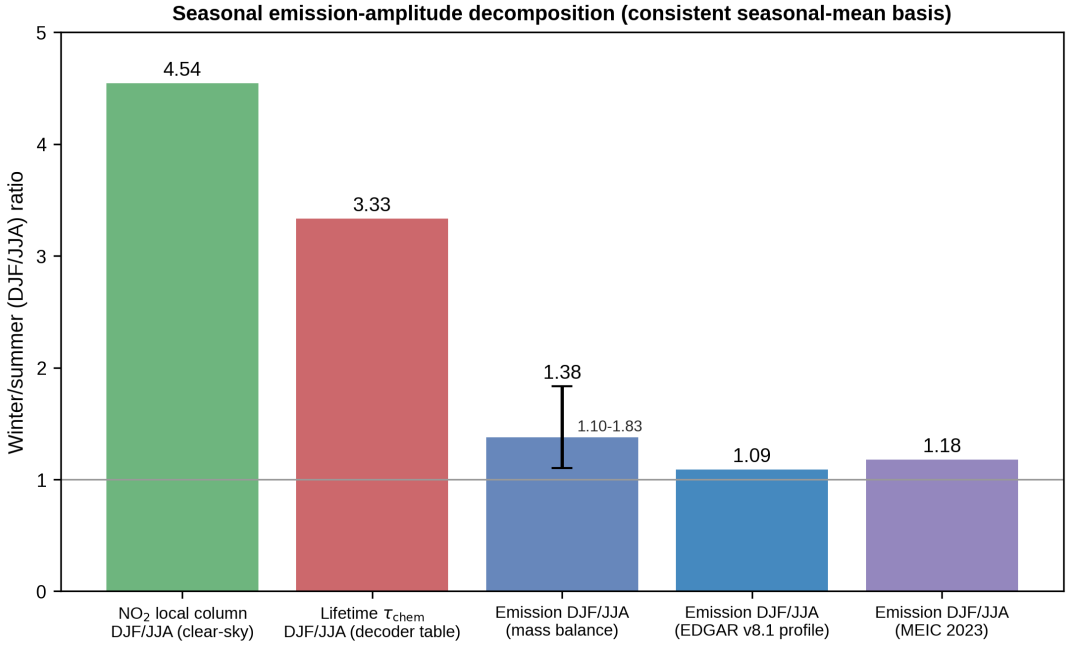


Figure S12. Decomposition of the seasonal amplitude of emissions under a consistent seasonal-mean convention. The clear-sky local NO_2 column winter/summer ratio (DJF/JJA seasonal means) is about 4.5 and the chemical lifetime τ_{chem} of the decoder table winter/summer ratio is about 3.3; dividing the two would imply an emission winter/summer ratio of about 1.4 (interval 1.10 to 1.83 allowing for a ±25% τ uncertainty), too uncertain to be reported as a result, to be compared with 1.09 for the EDGAR v8.1 2022 monthly profile used for the product totals and 1.18 for MEIC 2023. Note that the ratio of the peak month to the trough month is larger than the seasonal-mean ratio (6.5 for the local column, February/August); the two conventions differ and must not be conflated.

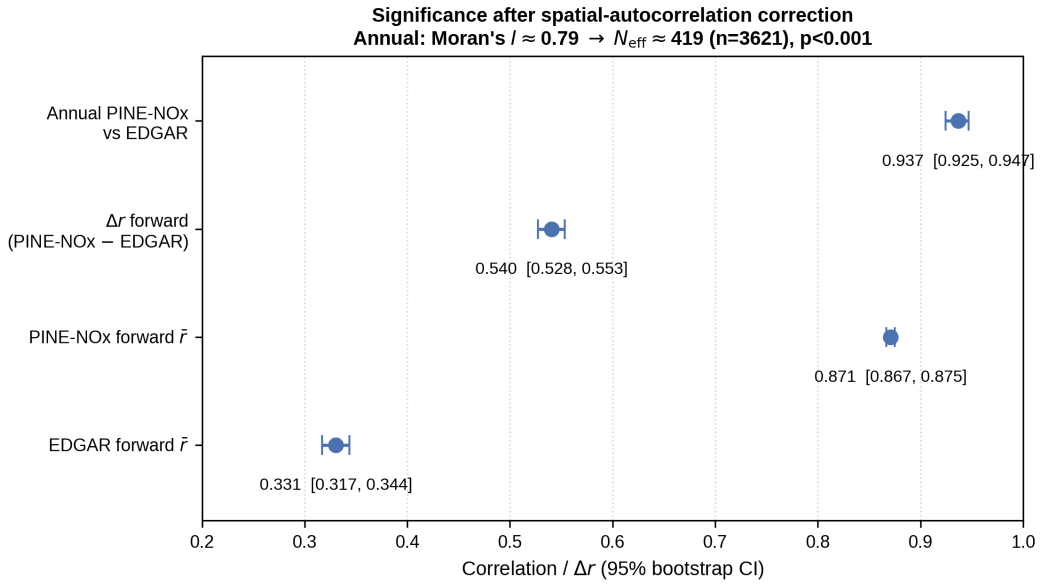


Figure S13. Significance after correction for spatial autocorrelation. For the annual-mean E_{final}–EDGAR field (n = 3621), the log-space correlation is r = 0.937 (95 % interval 0.925–0.947 from a 7 × 7-cell block bootstrap); Moran's I ≈ 0.79 (mean of the two log-emission fields) gives an effective sample size N_{eff} ≈ 419 with p < 0.001. Over 347 days the mean reconstruction r is 0.331 for EDGAR and 0.871 for PINE-NOx; the improvement is Δr = 0.540 (95 % interval 0.528–0.553 from a bootstrap over days); the intervals for the two correlations are 0.317–0.344 (EDGAR) and 0.867–0.875 (PINE-NOx).

Table S1. Configuration of the WRF-Chem runs in Sect. 3.2 and Fig. 7. All runs share this configuration; they differ only in the NOx emissions inside the study domain. The runs of the synthetic test use a simpler configuration, given in Text S1.

Item	Setting
Model	WRF-Chem v4.1.2; the ISORROPIA aerosol routine patched to exit where it would otherwise loop without end (other cases unchanged)
Domain and grid	One domain, Lambert conformal, 101 × 81 points, 27 km; 35 vertical levels, model top 50 hPa
Period	Initialised on 5 July 2023 00 UTC (chemistry from CAMS EAC4) and run to 31 July 2023 18 UTC (about 27 days); first 3 days discarded as spin-up, analysis 8–31 July (24 days)
Meteorology	NCEP GDAS/FNL analyses, 6-hourly; grid nudging of wind, temperature and humidity (3×10^{-4} s^{-1}) above the boundary layer
Physics	WSM3 microphysics; RRTM long-wave and Dudhia short-wave radiation; revised MM5 surface layer; Noah land surface; YSU boundary layer; Kain–Fritsch cumulus
Chemistry	RADM2 gas phase with MADE/SORGAM aerosol; Madronich (TUV) photolysis; Wesely dry deposition; no wet scavenging, cloud chemistry, N_2O_5 heterogeneous chemistry or convective transport of tracers
Chemical initial and boundary conditions	CAMS EAC4 reanalysis for O_3, CO, NO_2, SO_2, HCHO, HNO_3, C_2H_6, C_3H_8 and isoprene; other species from the WRF-Chem default profiles

Anthropogenic emissions	MEIC 2023 monthly emissions of all species (0.25°, five sectors); EDGAR v8.1 NOx outside the MEIC coverage; NOx emitted as 90 % NO and 10 % NO_2 (molar)
NOx inside the study domain	Replaced in the 425 model cells inside the 0.1° study domain by the July-mean pattern of the tested field (EDGAR, EDGAR smoothed with a 2-cell Gaussian kernel, MEIC, PINE-NOx with the EDGAR prior, PINE-NOx with the MEIC prior), scaled to the same NOx mass as MEIC in July; all other species unchanged
Temporal and vertical profiles	EDGAR hourly and day-of-week profiles for China, applied by sector; sector-specific vertical profiles over the lowest 8 model layers
Biogenic emissions	MEGAN v2.04, online (includes soil NO)
Not included	Lightning NOx, open fires, dust and sea salt
Comparison with TROPOMI	Each TROPOMI pixel (qa_value ≥ 0.75, cloud radiance fraction < 0.5) matched to the nearest model column, interpolated in time to the pixel; tropospheric averaging kernel applied; pixels averaged to the 0.1° grid per day
Comparison with stations	Hourly NO_2 in the lowest model layer of the cell containing each of the 160 CNEMC stations

Text S1. Synthetic test with known emissions

This test asks how much of the change from EDGAR the inversion produces when the true emissions are known. WRF-Chem was run for July 2023 (24 analysed days) and January 2023 (23 analysed days) with two NOx emission fields inside the 0.1° study domain: EDGAR itself, so that the true emissions equal the prior, and the MEIC pattern of the same month scaled to the EDGAR total. The simulated tropospheric NO_2 columns were interpolated linearly to the 0.1° grid, sampled with the TROPOMI valid-pixel mask of each day, perturbed with random noise (a lognormal factor with 20 % spread and an additive term of 5×10^{-6} mol m^{-2}) and inverted with the four trained PINE-NOx networks without retraining. The time-mean retrieved fields are compared with the true fields, with MEIC of the same month and with the inversion of the observed TROPOMI columns on the 22 days per month for which both exist (Table S2).

The runs use the domain, vertical grid, physics, chemistry, photolysis and dry deposition of Table S1, the same NCEP GDAS/FNL meteorology and the same CAMS EAC4 chemical initial and boundary conditions (January fields for the January runs), but differ from the runs of Table S1 as follows. They were run without nudging and with WRF-Chem v4.1.2 without the ISORROPIA patch of Table S1. The emissions contain NOx only (90 % NO and 10 % NO_2, molar), all in the lowest model layer: inside the study domain the 0.1° field of the test, mapped onto the model grid with mass conservation, and outside it the EDGAR v8.1 annual NOx emissions. No other anthropogenic species and no biogenic emissions (thus no soil NO) are included. The emissions are constant in time, except in the July run with the MEIC pattern, which uses hourly and day-of-week factors from the EDGAR temporal profiles for China, weighted by the EDGAR sector mix of each grid cell. The July runs cover 5 July 00 UTC to 31 July 18 UTC and the January runs 6 January 00 UTC to 31 January 18 UTC; the first 3 days are discarded (analysis 8–31 July and 9–31 January). The tropospheric NO_2 column (surface to the lapse-rate tropopause, or to 150 hPa where none is found) is the mean of the hourly model output at 05 and 06 UTC, which brackets the TROPOMI overpass (about 13:30 local time); unlike the comparison in Table S1, no averaging kernel is applied and the model is not sampled at the individual pixels.

Three quantities are compared. The first is the share of the domain total held by the 10 % of 0.1° cells with the highest EDGAR emissions (363 cells) and by the half of the cells with the lowest EDGAR emissions. The second is the fraction of the total emission placed in different 0.5° grid cells than in MEIC of the same month, i.e. half the sum of the absolute differences between the two maps of emission shares over the 0.5° blocks whose cells all have MEIC emissions (the measure of Fig. 10). The third is the gain from the observations: this fraction for the inversion of the synthetic columns simulated from EDGAR minus that for the inversion of the observed columns. All values are means over the four network initialisations. Intervals are 95 % bootstrap intervals over 0.5° blocks (2000 draws); for the gains, bias-corrected intervals over days (moving blocks of 3 days) are also given.

In both months the inversion of the columns simulated from EDGAR lowers the share of the highest-emitting cells by about as much as the inversion of the observed columns, although the true emissions equal EDGAR. The size of the change from EDGAR reported in the main text is therefore not interpreted as an observed redistribution. The comparison with MEIC differs between the months. In July the observed columns bring the pattern 5.0 points closer to MEIC than the synthetic EDGAR columns do, about as much as the gain when the MEIC pattern is the true field (5.1 points). In January the gain (0.6 points) cannot be distinguished from zero.

The test has several limitations. It covers two months with 22 matched days each, so the annual fields are not tested directly. The 27 km model columns are smoother than TROPOMI pixels (about 3.5 × 5.5 km), which probably makes the change produced by the method larger in this test than in the inversion of TROPOMI, whereas the control with columns simulated by the PINE decoder (Fig. 10) probably makes it smaller; the test cannot say where between the two the true value lies. No air-mass-factor or other retrieval bias is simulated. The networks were trained on TROPOMI columns and applied to the synthetic columns without retraining. The model configuration is simpler than that of Sect. 3.2 (NOx-only emissions, no nudging, no diurnal emission cycle except in the July run with the MEIC pattern), so the chemical environment and the transport of the simulated columns differ from those of the evaluation runs. MEIC is an inventory and not the truth, so a smaller distance to MEIC shows consistency with an independent estimate, not accuracy.

Table S2. Synthetic test with known emissions (Text S1), July and January 2023, 22 days per month on which both the synthetic and the observed columns are inverted; means over four network initialisations. Shares and fractions in percent of the domain total, changes and gains in percentage points; 95 % intervals over 0.5° blocks in parentheses and, for the gains, bias-corrected intervals over days in brackets.

Quantity	July	January
Share of the 10 % highest-emitting EDGAR cells (top-10 % share), EDGAR = true field of the first test	58.6	58.6
Top-10 % share, MEIC pattern = true field of the second test	27.8	28.3
Top-10 % share, inversion of synthetic columns, true field EDGAR	37.3 (30.1 to 43.6)	33.0 (26.4 to 38.9)
Top-10 % share, inversion of synthetic columns, true field MEIC pattern	35.3 (27.9 to 41.9)	30.9 (24.4 to 37.1)
Top-10 % share, inversion of observed TROPOMI columns	38.7 (31.2 to 45.1)	33.3 (26.6 to 39.2)
Change of the top-10 % share from EDGAR, synthetic columns with true field EDGAR	−21.3 (−23.3 to −19.3)	−25.6 (−27.5 to −23.5)
Change of the top-10 % share from EDGAR, observed columns	−19.9 (−21.6 to −17.8)	−25.3 (−27.5 to −22.9)
Share of the half of the cells with the lowest EDGAR emissions: EDGAR / synthetic (true field EDGAR) / observed	5.8 / 15.3 / 14.2	5.8 / 17.4 / 17.2
Fraction of the emission placed in different 0.5° cells than in MEIC: EDGAR	20.5 (17.2 to 23.9)	21.9 (18.3 to 25.2)
Same fraction, inversion of synthetic columns, true field EDGAR	18.5 (15.1 to 21.9)	17.8 (15.1 to 20.4)
Same fraction, inversion of observed columns	13.5 (11.0 to 16.1)	17.1 (14.2 to 20.0)
Same fraction, inversion of synthetic columns, true field MEIC pattern	13.4 (10.9 to 15.9)	15.5 (12.8 to 18.0)

Gain from the observations: synthetic (true field EDGAR) minus observed	5.0 (2.4 to 7.7) [2.1 to 7.4]	0.6 (−1.2 to 2.5) [−1.2 to 3.3]
Same gain when the true field is the MEIC pattern: synthetic (EDGAR) minus synthetic (MEIC)	5.1 (3.2 to 7.1) [3.2 to 7.8]	2.2 (0.7 to 3.8) [1.1 to 4.0]
Fraction of the emission placed in different 0.5° cells than in the true field, inversion of synthetic columns with true field EDGAR	15.9 (13.5 to 18.6)	19.0 (16.3 to 21.5)
Correlation across 0.1° cells of the departure maps ln(X/EDGAR) of the observed and the synthetic (true field EDGAR) inversions	0.63 (0.57 to 0.69)	0.88 (0.85 to 0.90)