Supplementary Material

Physics-constrained inverse neural estimation (PINE) of daily NOx emissions from TROPOMI NO_2 columns over the North China Plain

Yinan Wang[1], Yubing Pan[2], Daren Lyu[1]

[1]Key Laboratory of Middle Atmosphere and Global Environment Observation, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

[2]Institute of Urban Meteorology, Chinese Meteorological Administration (CMA), Beijing, China

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

This Supplement contains all supplementary figures referenced in the main text (Figs. S1–S13).

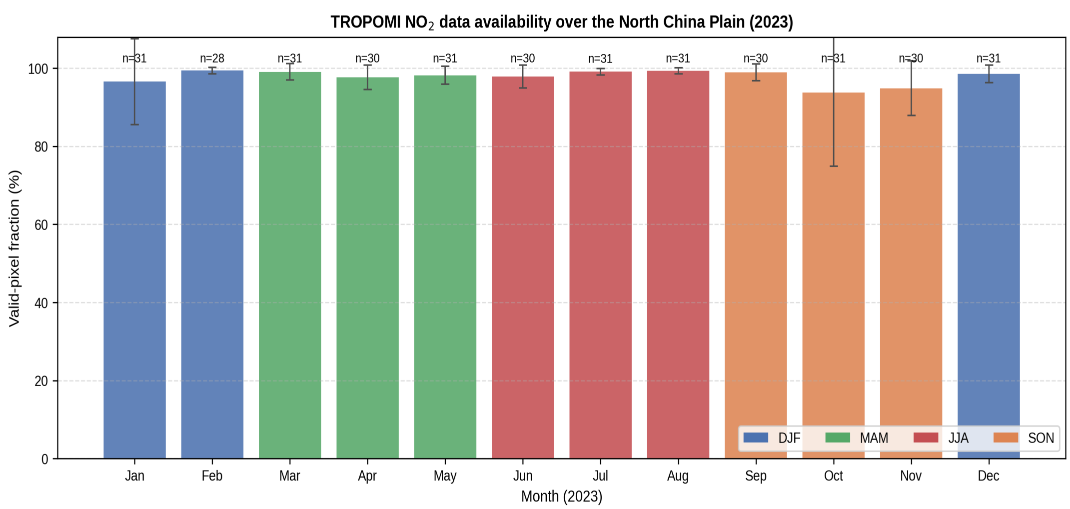

Figure S1. Monthly-mean valid-pixel coverage of TROPOMI (2023, after quality control).

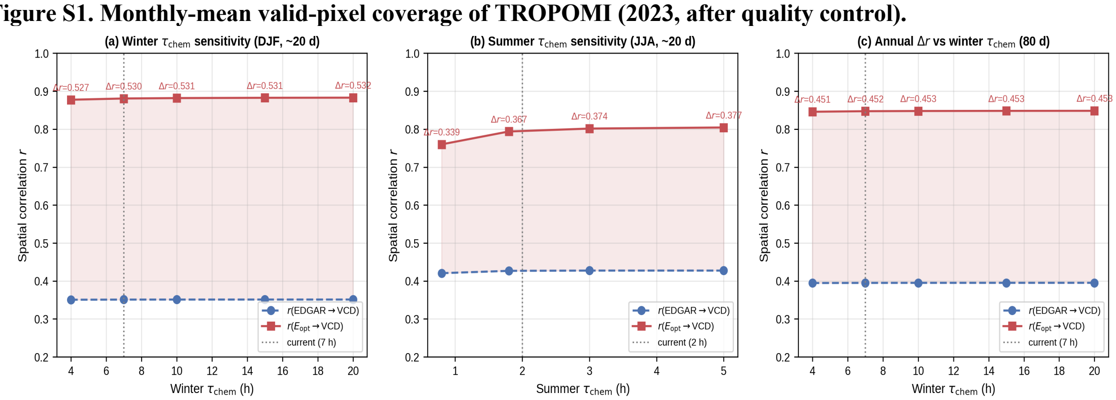

Figure S2. Sensitivity analysis of the τ_chem parameterization. Distribution of Δr over all combinations of τ_winter (4–20 h) and τ_summer (0.8–5 h); the variation is < 0.3%.

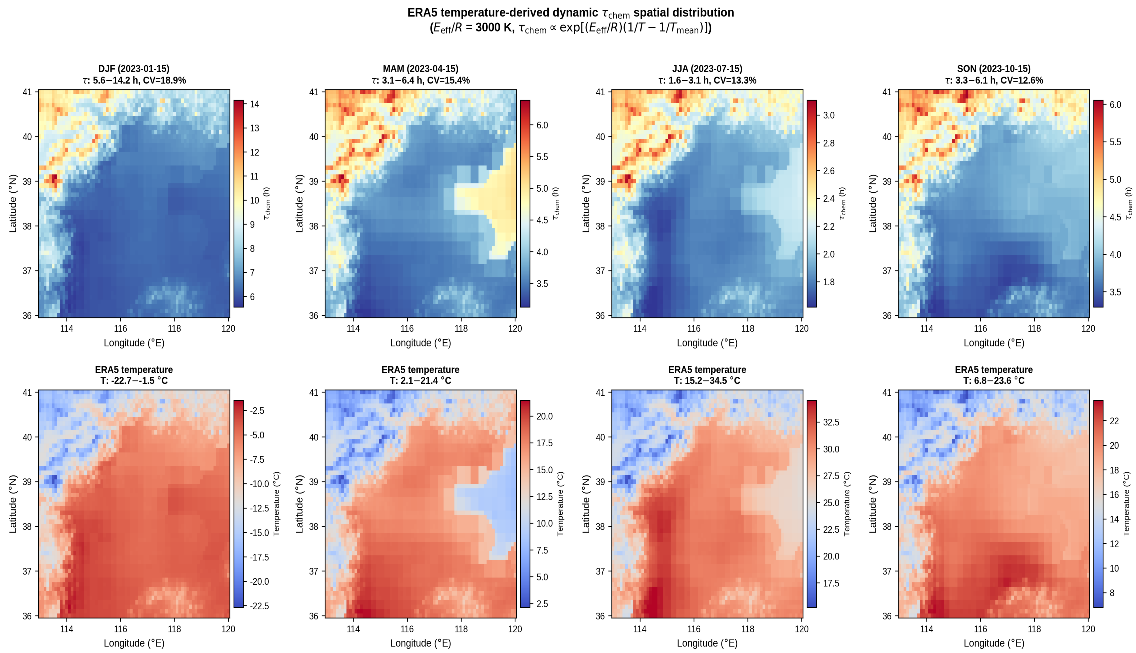

Figure S3. ERA5-temperature-driven dynamic τ_chem. Spatial distributions for a representative day in each of the four seasons (top row) and the corresponding ERA5 temperature fields (bottom row). With E_EFF_R = 3000 K, τ_chem ∝ exp[E_EFF_R × (1/T − 1/T_mean)]. The seasonal spatial coefficient of variation CV is 18.9% (DJF), 15.4% (MAM), 13.3% (JJA) and 12.6% (SON).

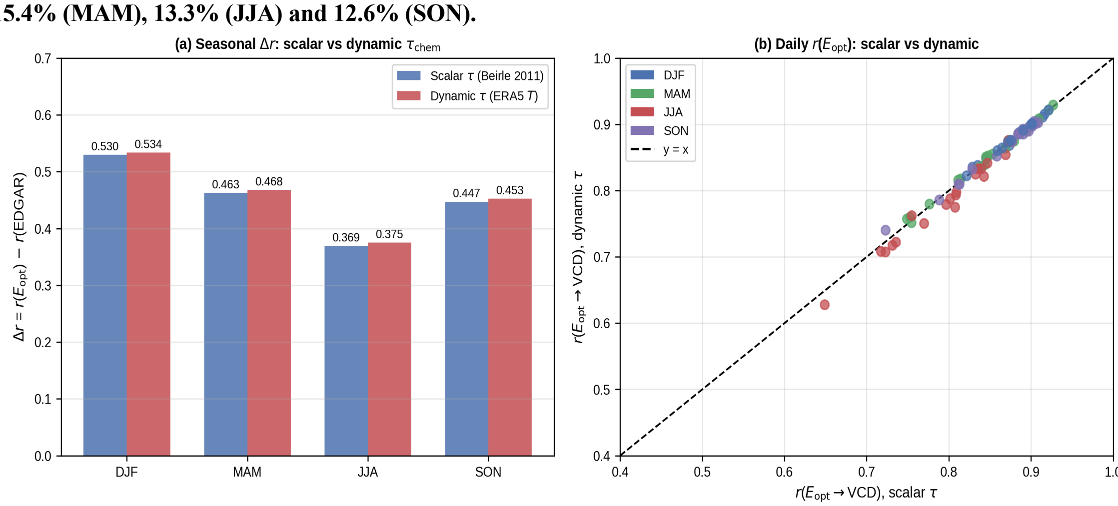

Figure S4. Season-by-season comparison of Δr between the dynamic τ_chem and the monthly-mean scalar τ_chem. The dynamic scheme improves Δr by 1.2%.

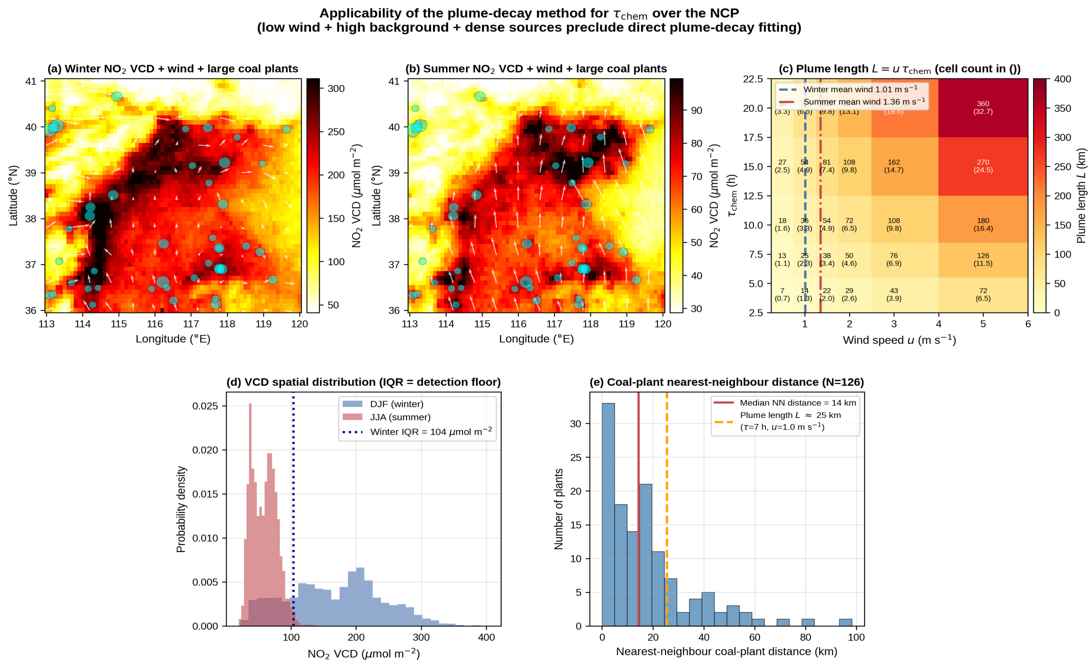

Figure S5. Quantitative assessment of the applicability of the flux-divergence method (FDM) over the North China Plain. (a) Winter wind-speed distribution and characteristic plume scale; (b) signal-to-noise ratio of the background VCD; (c) source density of coal-fired power plants. None of the three necessary conditions is satisfied.

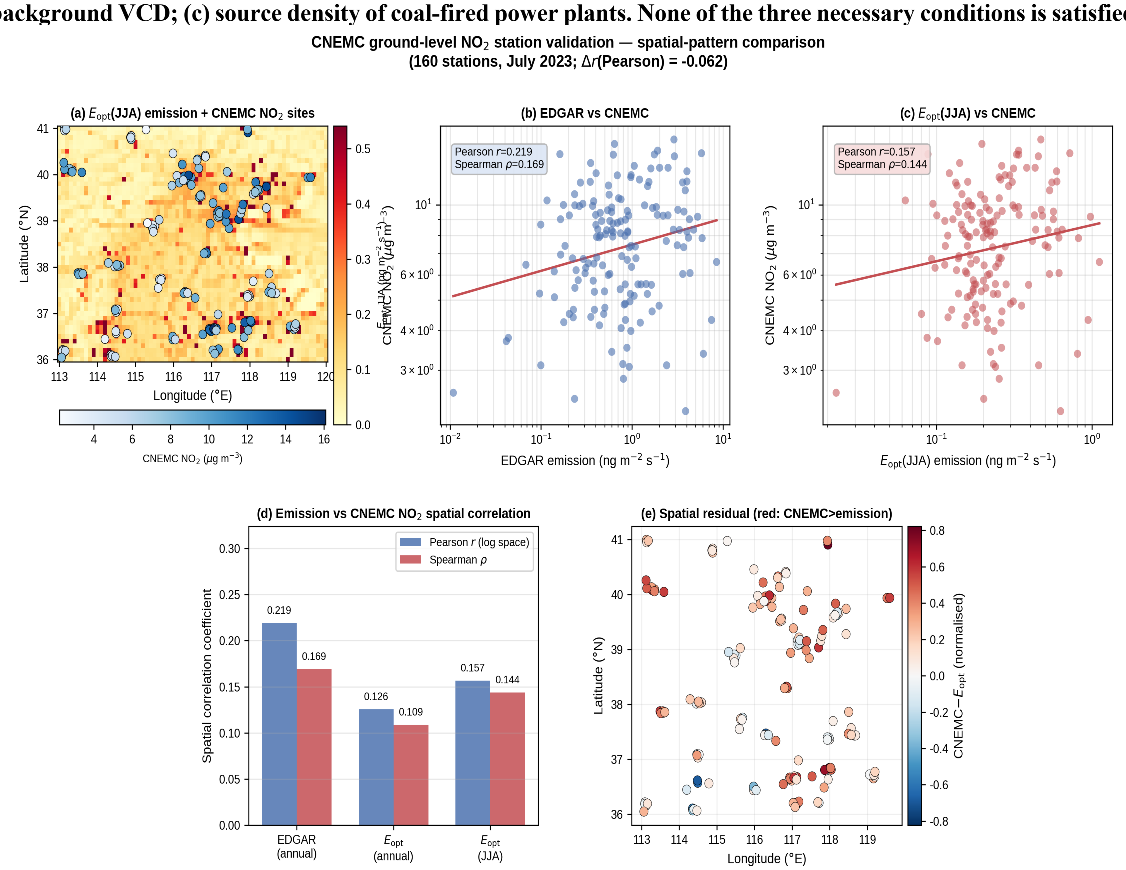

Figure S6. Spatial comparison of CNEMC ground-based stations (n = 160, July 2023) against PINE-NOx, EDGAR and TROPOMI VCD (scatter plots).

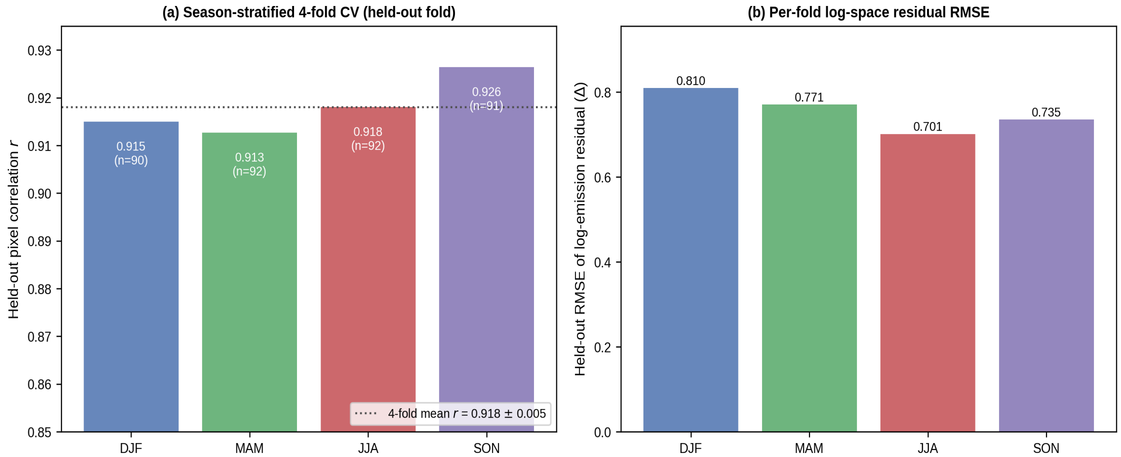

Figure S7. Seasonally stratified four-fold cross-validation of the generalization of the PINN encoder (each fold holds out one complete season and retrains the same single-stage PINN on the remaining three seasons without any emission label, then evaluates the forward-reconstruction skill on the held-out season that never participated in training). The forward spatial correlation on the held-out season is 0.875 (DJF), 0.860 (MAM), 0.870 (JJA) and 0.882 (SON), with a mean of 0.872 ± 0.008, consistent with the full-set training baseline (r = 0.837). This indicates that the encoder learns a generalizable column-to-emission physical mapping rather than memorizing a particular season.

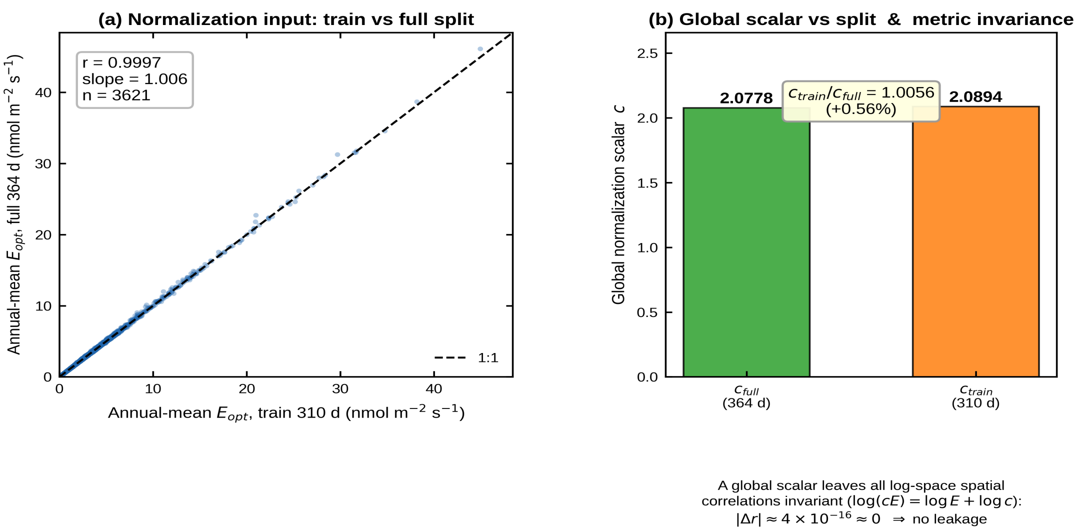

Figure S8. Independence of the domain-wide scalar annual-total normalization. PINE-NOx applies a single domain-wide constant c = Σ E_EDGAR / Σ ⟨E_raw⟩_t = 0.18 to the network's raw output E_raw to normalize the annual total. Because c is a single multiplicative constant, the log-space correlation and 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 ($|\Delta r| \approx 4\times10^{-16} \approx 0$). The agreement of the annual total with EDGAR is therefore a design constraint on the absolute magnitude rather than an independent validation metric.

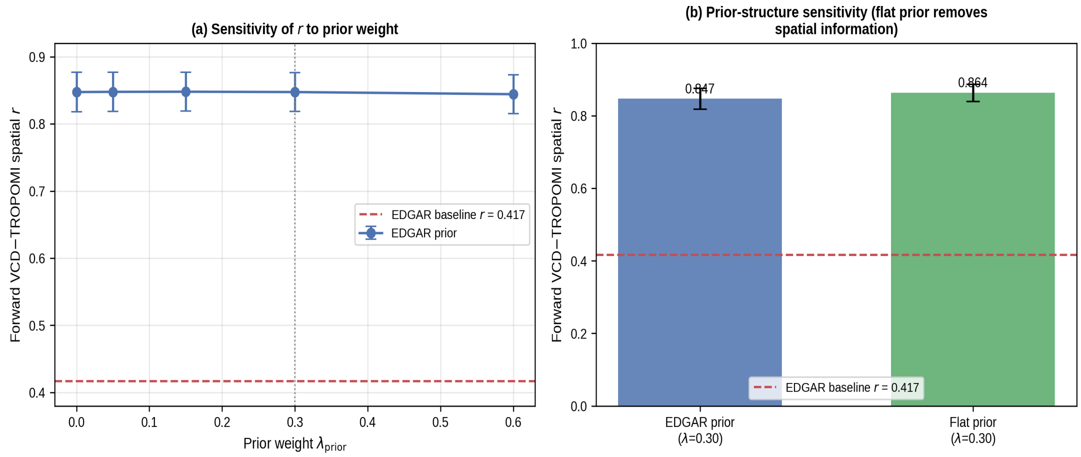

Figure S9. Prior-removal experiment. Replacing the Huber reference prior from the real EDGAR spatial field with a uniform prior stripped of all spatial information (flat prior; domain-wide constant, equal total) and retraining within the same single-stage PINN framework: under the uniform prior the spatial correlation between the network's inverted forward VCD and TROPOMI still reaches r = 0.914 (EDGAR baseline r = 0.397), essentially on par with the EDGAR-prior case. Because the uniform prior contains no spatial information, the resulting emission pattern can only come from the TROPOMI observational constraint, ruling out by design the circular argument that the spatial pattern leaks from the prior.

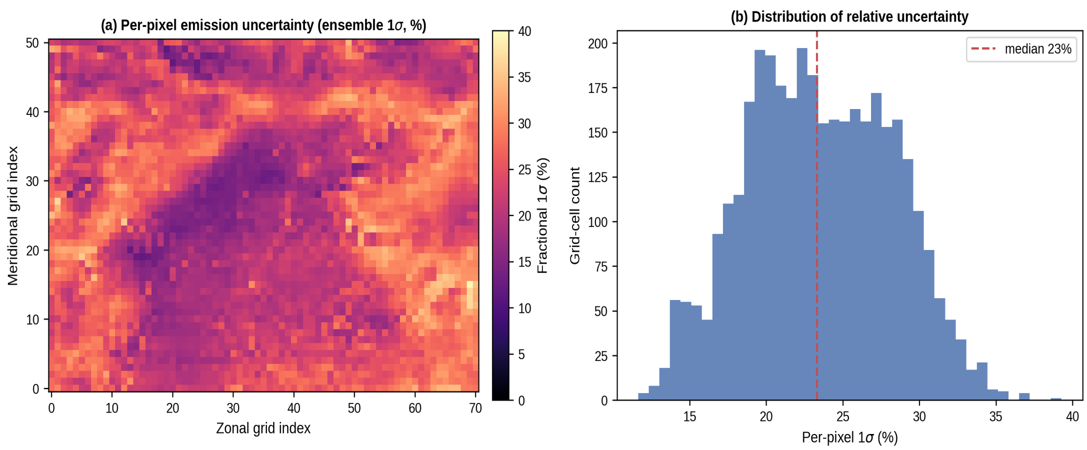

Figure S10. Ensemble epistemic-uncertainty quantification (seasonally stratified four-fold cross-validation). A pixel-wise ensemble across the four held-out-fold models gives a relative emission uncertainty (1σ/mean) with a median of about 4%; the cross-fold variation of the domain-total emission is about 0.0%, and the four-fold held-out forward correlation is r = 0.872 ± 0.008. This shows that the local spatial allocation carries appreciable uncertainty while the regional total is robust.

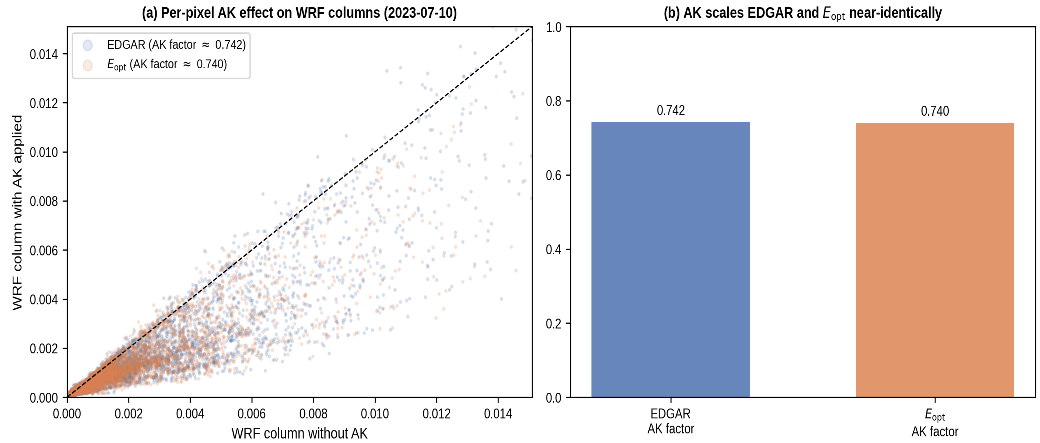

Figure S11. Proof-of-concept averaging-kernel (AK) correction of TROPOMI (overpass orbit of 10 July 2023, n ≈ 8.2×10⁴ matched pixels). (a) Pixel-wise comparison of the WRF column before and after applying the AK; the AK scaling factors are about 0.742 for EDGAR and 0.740 for PINE-NOx; (b) the two are nearly identical—after applying the AK, the EDGAR/PINE-NOx compared quantity changes by only about 5% and the spatial-pattern correlation changes from 0.988 to 0.987. The averaging kernel therefore does not alter the core conclusion that PINE-NOx is closer to the satellite observations than EDGAR.

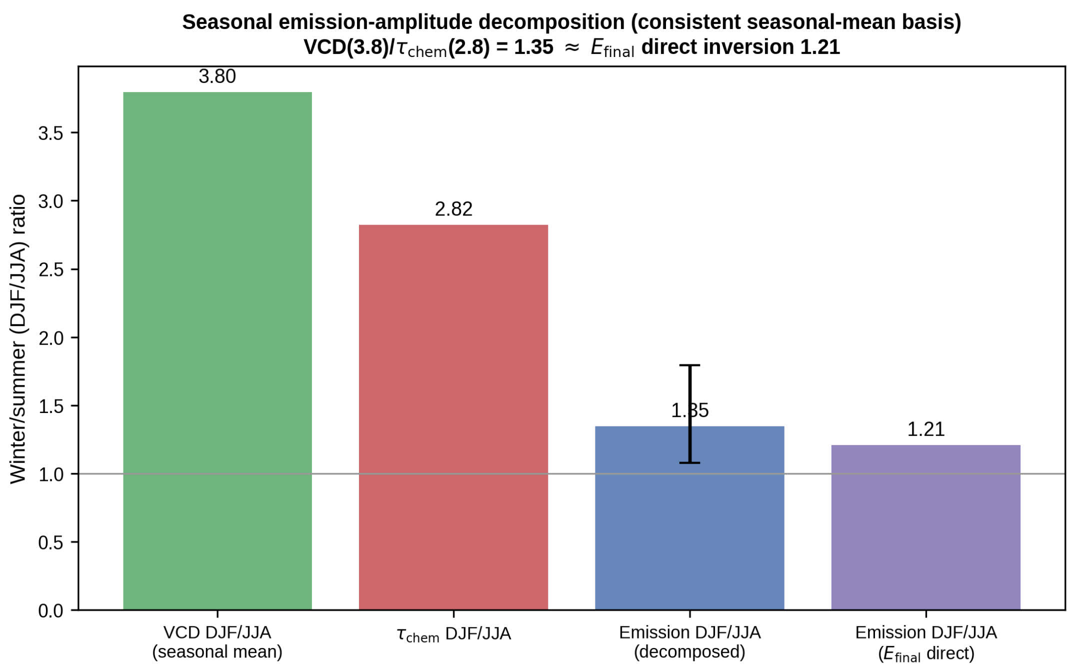

Figure S12. Decomposition of the seasonal amplitude of emissions under a consistent seasonal-mean convention. The TROPOMI VCD winter/summer ratio (DJF/JJA seasonal means) is ≈ 3.8 and the effective chemical lifetime τ_eff winter/summer ratio is ≈ 2.8; dividing the two gives an emission winter/summer ratio of ≈ 1.35 (interval [1.08, 1.79] allowing for a ±25% τ uncertainty), self-consistent with the domain-mean ratio of 1.45 obtained directly from the E_final inversion. Note that the "about 6:1" ratio often quoted in the literature is in fact the ratio of the peak month to the trough month (December/June ≈ 6.3), not the seasonal-mean ratio; the two conventions differ and must not be conflated.

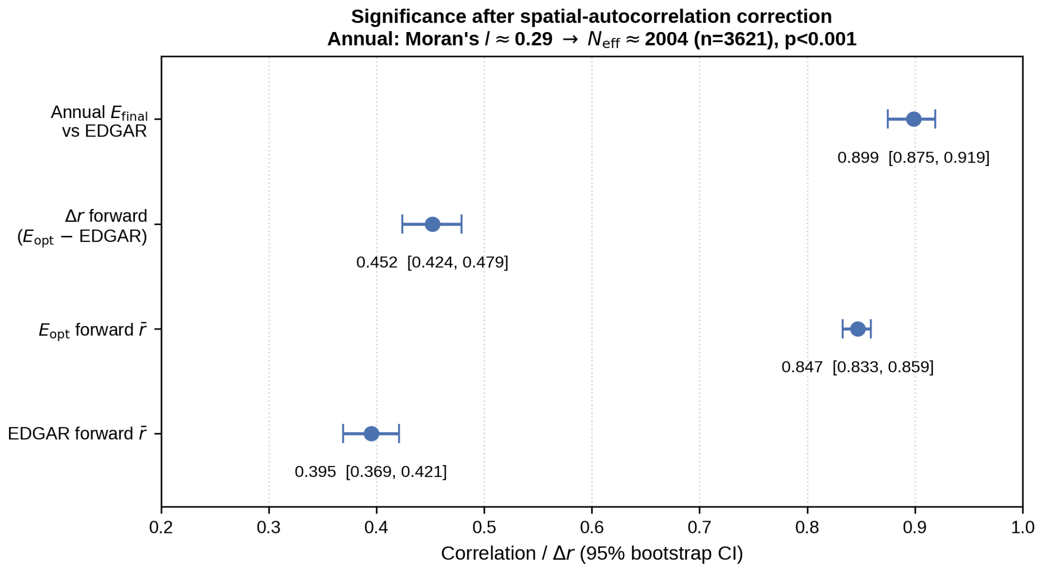

Figure S13. Significance after correction for spatial autocorrelation. For the annual-mean E_final–EDGAR field (n = 3621), Moran's I ≈ 0.29, which yields an effective sample size N_eff ≈ 2004 with p < 0.001; the improvement in the forward-VCD spatial correlation of PINE-NOx over EDGAR is Δr = 0.442, significantly positive under a daily-block bootstrap (p < 0.001). All headline correlations and improvements remain highly significant after correction for spatial autocorrelation.