https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

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)

Abstract. Daily-resolution NOx emissions are pivotal for air-quality forecasting, yet static inventories cannot capture day-to-day variability, and conventional satellite inversions are either computationally prohibitive (4D-Var) or circularly dependent on pre-existing emission products. We introduce physics-constrained inverse neural estimation (PINE), instantiated for NOx as PINE-NOx, which retrieves daily NOx emissions over the North China Plain (0.1°, 364 days of 2023) from Sentinel-5P/TROPOMI NO_2 columns. PINE is a physics-constrained autoencoder: a neural encoder maps observed columns to emissions, which a fixed, differentiable transport–chemistry operator decodes back into columns. Physics thus enters structurally through this decoder, not as a soft residual penalty; trained end-to-end to reconstruct observations without emission labels. On 80 season-balanced validation days, PINE-NOx-inferred emissions raise the log-space spatial correlation between simulated and observed columns from 0.395 to 0.837 ($\Delta r = +0.442$, $p < 0.001$), robustly across four encoder backbones and all four seasons. Two independent checks corroborate the inversion: an independent WRF-Chem simulation cuts the summer column bias from +113% to +5%, and the recovered seasonal cycle agrees with the fully independent, bottom-up MEIC inventory ($r = 0.68$), while a prior-removal experiment confirms that the spatial pattern originates from the observations rather than the prior. The inventory provides the first daily-resolved dynamics of North China Plain NOx emissions, a physically disaggregated emission winter/summer ratio of ~1.4, and ~11% spatial redistribution relative to EDGAR. PINE-NOx offers a physically interpretable, label-free and computationally inexpensive paradigm for atmospheric emission inversion.

1 Introduction

The North China Plain (NCP) is one of the most industrially intensive emission regions in the world; its annual-mean NOx emission intensity is roughly 8–10 times the global average, and its NO_2 vertical column densities have long ranked among the highest worldwide (van der A et al., 2017; Zheng et al., 2018; Zhang et al., 2019). Accurately quantifying the day-to-day dynamics of NOx emissions in this region is important for air-quality forecasting, the evaluation of emission-reduction policies and climate simulation. Existing static, annual-mean emission inventories (e.g. EDGAR and the MIX Asian anthropogenic inventory; Li et

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

al., 2017) cannot capture the high-temporal-resolution emission variability revealed by satellite-constrained inversion (Mijling et al., 2013; Kong et al., 2019; Goldberg et al., 2019; Souri et al., 2020), limiting the fine constraint on emission initial fields in high-resolution air-quality forecasting; the dramatic short-term changes in satellite NO_2 during the COVID-19 lockdowns further highlight the importance of daily-resolution emission dynamics for air-quality simulation (Goldberg et al., 2020).

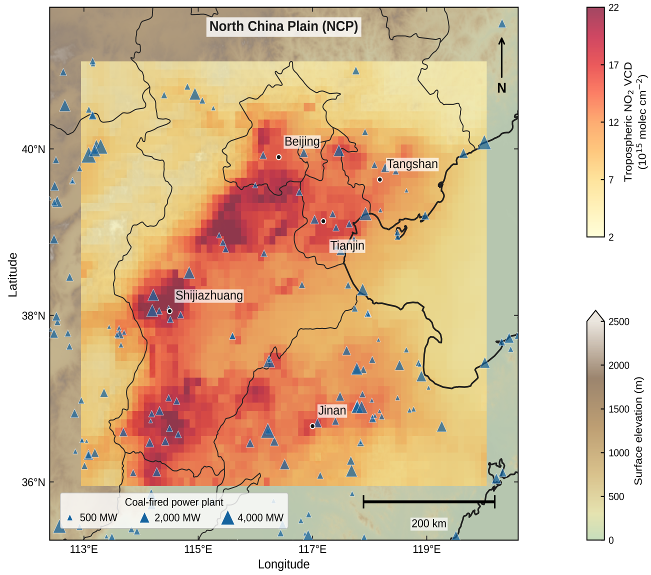

Figure 1. Study region. A shaded-relief map of the North China Plain (lower colour bar: surface elevation) overlaid with the 2023 TROPOMI annual-mean tropospheric NO_2 column (upper colour bar). Triangles mark all 265 coal-fired units, with area directly proportional to installed capacity (33–4540 MW). Major cities and a 200 km scale bar are annotated.

Existing emission-inversion methods all face some degree of limited applicability over the North China Plain (Martin et al., 2003; Streets et al., 2013). The flux-divergence method (Beirle et al., 2011, 2019; de Foy et al., 2015; Liu et al.,2016) relies on assumptions of uniform wind and a clean background, and its signal-to-noise ratio is severely inadequate under the high background concentrations and extremely dense sources of the NCP. A quantitative analysis shows that on calm winter stagnation days (days with ERA5 10 m daily-mean wind speed < 2 m s^{-1}), the domain-mean wind speed is about 1.01 m s^{-1}, the characteristic plume scale is about 25 km, the interquartile range of the background VCD is as large as 104 μmol m^{-2}, and the median nearest-neighbour distance among the 126 coal-fired power plants within the domain is only 14 km, so none of the three necessary conditions is met (Fig. S5). Global chemical-transport-model joint inversion (Miyazaki et al., 2012, 2021) is highly accurate but extremely expensive computationally, with typical run times of hundreds of CPU hours per year; the DECSO local fast inversion (Mijling et al., 2013) is suitable at moderate resolution but lacks the fast, end-to-end inference capability afforded by deep learning.

In recent years, methods that combine physical constraints with data-driven inference, such as physics-informed neural networks (PINNs), differentiable simulation and simulation-based inference, have developed rapidly in Earth system science (Raissi et al., 2019; Cranmer et al., 2020; Karniadakis et al., 2021). The core

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

idea of PINNs is to embed physical laws into the training of neural networks, so that network outputs remain physically consistent even under sparse-data or label-free conditions. Here we extend this idea to atmospheric NOx emission inversion: a differentiable transport-chemistry operator is embedded as a fixed decoder within an autoencoder, so that the neural encoder can invert a physically interpretable emission field under the physical constraint of reconstructing the satellite observations, without any emission labels. The end-to-end fusion of differentiable physical forward modelling with neural-network inversion, together with a systematic comparison of different neural backbones (Ronneberger et al., 2015), the Fourier neural operator (Li et al., 2020), and their spectral fusion for high-resolution emission inversion, has not previously been attempted. We emphasize that our framework differs fundamentally in mechanism from the classical PINN: a PINN approximates the solution of a PDE with a neural network and uses the PDE residual as a soft-constraint loss term, whereas our framework retains an exact numerical solver as a fixed, differentiable decoder and treats physics as a hard constraint, with the neural network performing only the amortized inverse mapping from observations to emissions (after training, a single forward pass suffices for inference, with no per-day re-optimization). It is therefore closer to the differentiable-simulation and inverse-rendering paradigm.

We present the PINE-NOx framework (Fig. 2), which we position as a physics-constrained neural inversion for atmospheric inverse problems: the neural encoder infers emissions from TROPOMI columns, the fixed, differentiable advection-chemistry operator serves as a decoder that renders emissions into columns, and the network is trained label-free to reconstruct the observed columns. The framework makes three core contributions. (1) End-to-end and label-free: it does not rely on any existing emission-inversion product as supervision; emissions are inverted directly as a physical latent variable, eliminating circular reasoning. (2) It systematically compares convolutional, spectral-operator and spectral-fusion encoder backbones in terms of their true forward skill under label-free physical reconstruction, providing a basis for architecture selection in this class of inverse problems. (3) It raises the temporal resolution of NOx emissions from static annual means to the daily scale (day-to-day domain-mean coefficient of variation CV $\approx$ 29%), achieves about 11% spatial redistribution under the TROPOMI constraint, and cross-validates the magnitude and the seasonal dynamics, respectively, against an independent WRF-Chem three-dimensional simulation and the fully independent bottom-up MEIC inventory; once trained, inference takes milliseconds.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

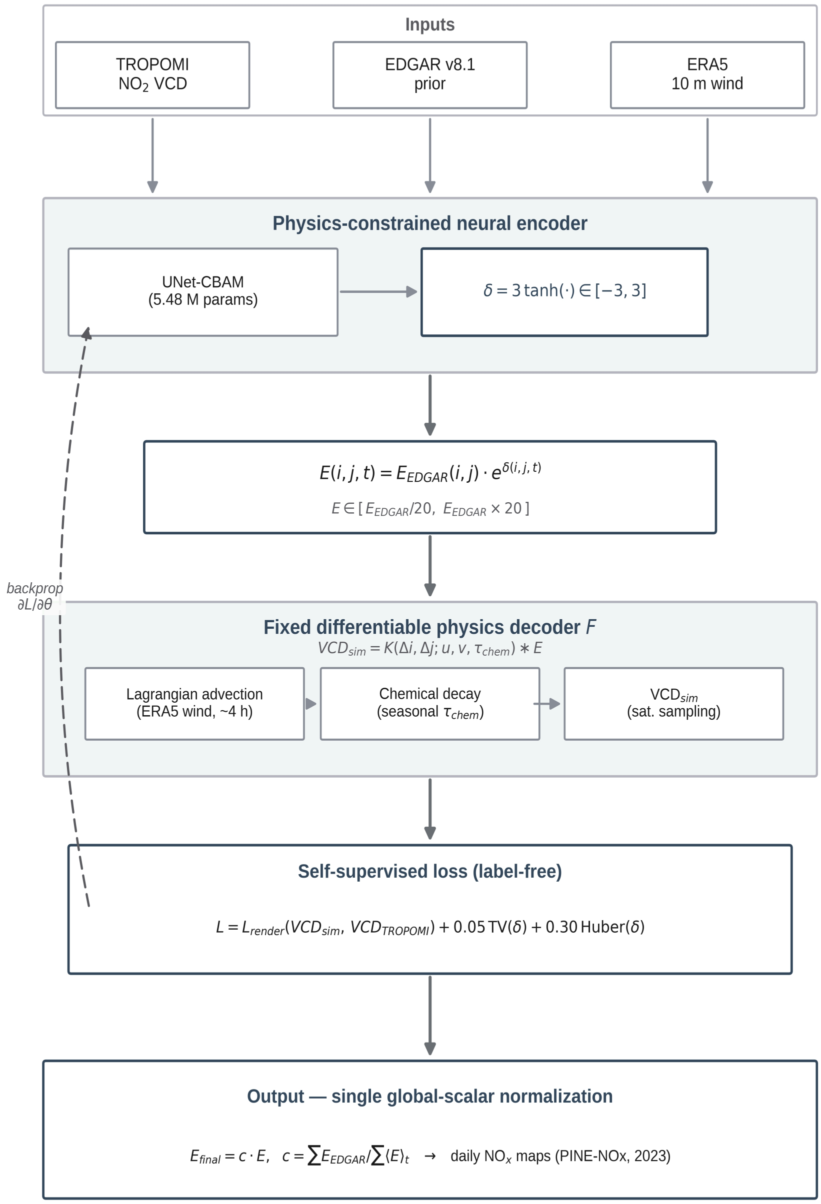

Figure 2. Schematic of the physics-constrained inverse neural estimation (PINE-NOx) framework.

2 Data and methods

The observational and reanalysis data, the prior emission inventory, and the key configuration parameters of
90 PINE-NOx used in this paper are summarized in Table 1.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

Table 1. Summary of PINE-NOx data sources and key configuration parameters.

Item	**Value / description**
Study domain and grid	
Study region	North China Plain (36–41 ° N, 113–120 ° E)
Inversion grid	0.1° × 0.1° regular lat-lon, 51 × 71 (3621 grid cells)
Period	Full year 2023
Observations and forcing data	
TROPOMI NO_2	Sentinel-5P L2 OFFL v2.4.0; QC: qa_value ≥ 0.75, cloud radiance fraction < 0.5; area-weighted resampling
ERA5 reanalysis	u10, v10, T2m, PBLH; daytime daily mean (08:00–16:00 local time)
Prior inventory	EDGAR v8.1 FT2022 NOx annual-mean field, resampled to 0.1°
Independent evaluation data (not ingested)	
MEIC inventory	Multi-resolution Emission Inventory for China (MEIC); monthly NOx, 0.25°→0.1°, domain annual total 2.67 Tg; independent seasonal cross-validation only
CNEMC surface NO_2	China National Environmental Monitoring Centre (CNEMC) surface NO_2; 160 ground stations over the NCP, hourly; qualitative independent spatial-consistency check
Differentiable physical decoder (fixed)	
Advection integration	Time step 450 s (7.5 min), 32 steps, integration time approximately 4 h
Effective chemical lifetime τ_chem	Winter approximately 7 h, summer approximately 1.8 h, linearly interpolated for other months
Magnitude treatment	Rendering loss demeaned per sample over the domain; absolute magnitude provided by annual-total normalization
Neural encoder	
Network architecture	UNet encoder-decoder (5 levels each) + residual connections + CBAM attention, about 5.48M parameters; FNO, UNet-SA and spectral-fusion UNet also compared
Inputs (6 channels)	log E_EDGAR, u10, v10, sin(DoY), cos(DoY), demeaned log VCD
Output	δ = log(E/E_EDGAR), tanh-clipped to [-3, +3], E = E_EDGAR exp(δ)
Label-free training	
Loss function	L = L_render + λ_tv TV(δ) + λ_prior Huber(δ); L_render demeaned per sample over the domain; λ_tv = 0.05, λ_prior = 0.30
Optimization	Adam, lr = 2×10^{-3}, cosine annealing, gradient clipping, about 120 epochs
Data split	Seasonally balanced random hold-out of 20 days per season (80 days in total); the remaining 284 days used for training
Supervisory signal	Reconstruction of the TROPOMI observed VCD only, without any emission label (label-free)
Post-processing	
Annual-total normalization	c = sum(E_EDGAR) / sum(mean_t[E]), E_final = E × c, locking the domain annual total to EDGAR

2.1 Study region and observational data

The study region is the core of the North China Plain (36–41° N, 113–120° E; Fig. 1), discretized on a 0.1° × 0.1° regular latitude-longitude grid (51 × 71 grid cells). Satellite observations are taken from the Sentinel-5P/TROPOMI tropospheric NO_2 Vertical Column Density (VCD) product (Veefkind et al., 2012; Boersma

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

et al., 2018; van Geffen et al., 2022), quality-controlled with qa_value ⩾ 0.75 and cloud radiance fraction < 0.5, and resampled to the study grid by area-weighted averaging (the seasonal-mean spatial distribution is shown in Fig. 3 and the monthly-mean valid-pixel coverage in Fig. S1). Retrieval uncertainties of the tropospheric NO_2 column, which are dominated by the tropospheric air mass factor, have been characterized in detail elsewhere (Boersma et al., 2004; Lorente et al., 2017). Meteorological forcing is from the ERA5 hourly reanalysis (Hersbach et al., 2020), from which we extract daytime daily means (08:00–16:00 local time) of the 10 m wind field (u10, v10), the 2 m surface air temperature and the planetary boundary-layer height (PBLH). The prior emission inventory is the EDGAR v8.1 (Crippa et al., 2026) annual-mean NOx field, resampled to 0.1° resolution and used as the prior base for the physical reconstruction, as the Huber reference constraint, and as an input feature to the neural encoder.

Two further datasets are used solely for independent evaluation and at no stage enter the inversion as a prior, a label or a constraint. The Multi-resolution Emission Inventory for China (MEIC; Zheng et al., 2018; Li et al., 2017) provides a bottom-up, sector-resolved monthly NOx inventory, regridded from 0.25° to the 0.1° study grid (domain annual total 2.67 Tg); because it is constructed fully independently of TROPOMI and, unlike the seasonless EDGAR prior, carries its own bottom-up seasonal cycle, it is used only to cross-validate the realism of the inverted seasonal dynamics (Sect. 3.3, Fig. 12). Hourly surface NO_2 observations from 160 China National Environmental Monitoring Centre (CNEMC) stations within the study domain are used only as a qualitative, independent spatial-consistency check (Figs. S6–S7); because ground-level concentration and the satellite column differ in physical dimension, this comparison is treated as a qualitative consistency test rather than a quantitative validation.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

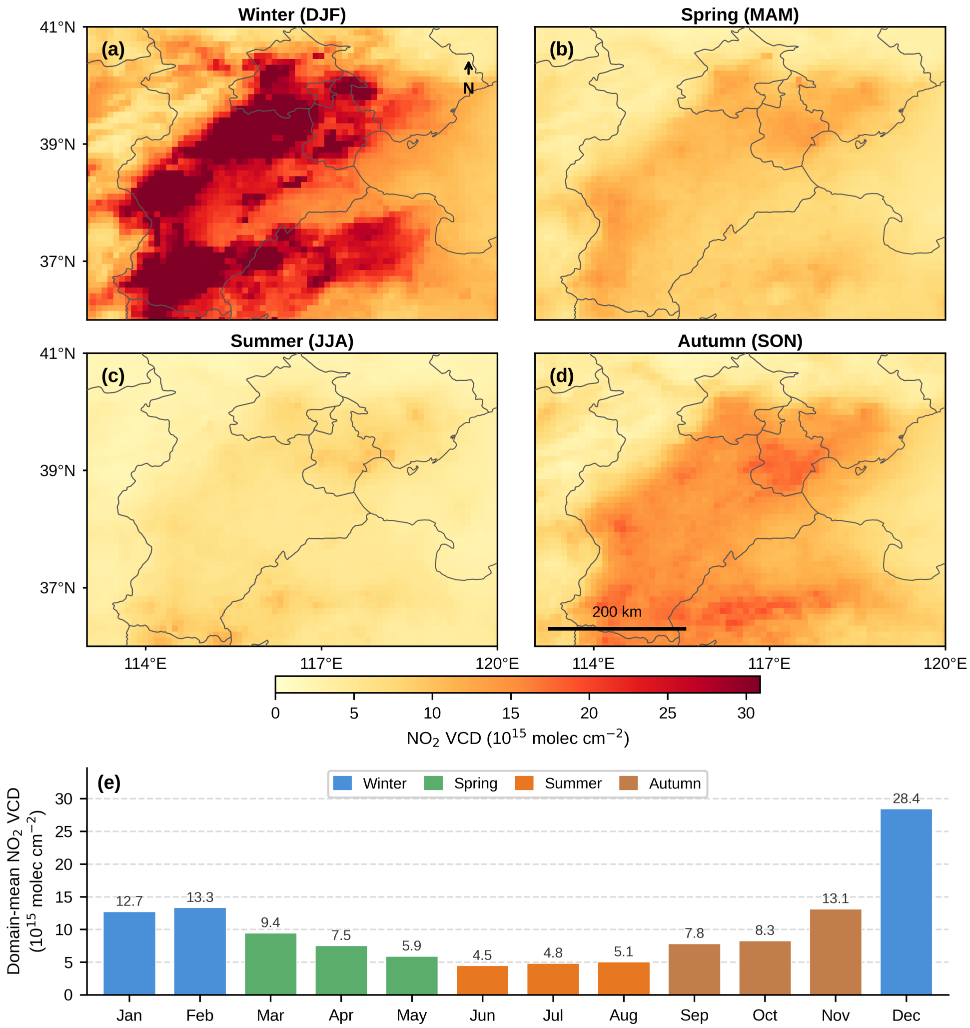

Figure 3. Seasonal climatology of the TROPOMI-observed tropospheric NO_2 column over the North China Plain (2023). (a–d) Winter, spring, summer and autumn means; (e) monthly variation of the regional mean.

2.2 Differentiable atmospheric transport-chemistry decoder

The differentiable physical decoder describes the spatial distribution of the NO_2 VCD formed by a given NOx emission field E(i, j) after atmospheric advection and chemical decay, and serves as the fixed decoder F in the physics-constrained autoencoder. The decoder solves the advection-decay equation in a two-dimensional Eulerian framework: starting from a zero-column initial condition, at each time step it successively applies the local emission source term, the upwind advection term and the first-order chemical-decay term to the column field and integrates iteratively, finally yielding the locally driven simulated column field VCD_sim = F(E). The upwind scheme is differentiable throughout while ensuring numerical stability, and a non-negativity constraint is imposed on the column after each step. The driving wind is the ERA5

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

daytime daily-mean 10 m wind field (u10, v10); the integration time step is 450 s (about 7.5 min) over 32 steps (an integration time of about 4 h, comparable to the typical NOx chemical lifetime). The effective NO_2 chemical lifetime τ_chem follows a monthly-mean parameterization based on the seasonal variation of OH. On the observation side, the TROPOMI column is decomposed into a uniform inflow background and a local emission signal: the background field VCD_bg is robustly estimated as the median of valid pixels in several boundary layers of the study domain (approximating the uniform background of long-range upwind transport), and the local emission signal VCD_local = VCD_TROP - VCD_bg is the target to be reconstructed by the decoder. Because the decoder is fully differentiable, the reconstruction error can be differentiated with respect to the emission field E, providing the analytical gradients required for end-to-end training of the neural encoder (see Sect. 2.3). We note that the decoder drives the transport of a single-overpass snapshot with the ERA5 daytime-mean 10 m wind during the satellite overpass window and does not explicitly represent intraday wind-direction changes; the background field VCD_bg is taken as the robust median of valid pixels in four boundary layers of the study domain, and the sensitivity to the number of layers is reported in the Supplement.

2.3 Neural encoder and self-supervised training based on differentiable physics

Emission inversion in PINE-NOx is performed by a neural encoder. The encoder takes a six-channel spatial field as input: log(E_EDGAR), the ERA5 wind components u10 and v10, the sine and cosine encodings of the day-of-year (sin(DoY), cos(DoY)), and the domain-mean-removed log of the TROPOMI VCD. It outputs a log-increment field δ(i, j), clipped by tanh to [-3, +3] (corresponding to emissions in [0.05, 20] × EDGAR), from which the emissions are obtained as in Eq. (1). The specific encoder backbone is systematically compared in Sect. 3.1; the main results of this paper use UNet-CBAM, whose CBAM attention module follows Woo et al. (2018). Crucially, the emissions output by the encoder are immediately fed into the fixed, differentiable decoder F to generate the simulated column VCD_sim, and the network is trained with reconstruction of the observations as its sole objective, with the loss function given in Eq. (2). Here the rendering term L_render = $\|\log \text{VCD\_sim} - \log \text{VCD\_local}\|^2$, TV(δ) is a total-variation spatial-smoothness regularizer, and H(δ) is a Huber regularizer referenced to the EDGAR prior; after sensitivity testing we set λ_tv = 0.05 and λ_prior = 0.30. Because the decoder is differentiable, the gradient of the reconstruction error is back-propagated end-to-end to the encoder weights, so that the network learns to invert emissions without ever seeing any emission label; this is the core mechanism of the physics-constrained autoencoder. Training uses Adam (Kingma and Adam, 2015) with gradient clipping, converging in about 120 epochs; inference takes about 5 ms per day on our experimental platform. The data are split with a seasonally balanced hold-out validation set: 20 days are randomly held out in each season (80 days in total; fixed random seed = 42) and the remaining days are used for training, to test seasonal generalization. The validation days are interleaved in time with the training days, and the evaluation metric is the demeaned log-space correlation (characterizing the spatial pattern rather than pointwise magnitude), so the temporal autocorrelation between training and validation days has limited influence on the conclusions. This training is a self-supervised

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

paradigm: reconstruction of the observed VCD itself serves as the supervisory signal, and no emission inventory is used as a label at any stage (i.e. label-free).

$$E = E_{EDGAR} \cdot exp(\delta) \quad (1)$$

$$L = L_{render} + \lambda_{tv} \cdot TV(\delta) + \lambda_{prior} \cdot H(\delta) \quad (2)$$

2.4 Identifiability and regularization analysis of the physical constraint

This subsection clarifies, from the perspective of inverse-problem theory, the essential nature of the constraint that the fixed, differentiable decoder F imposes on emission inversion, providing the theoretical basis for the significant improvement in physical self-consistency and for the prior-removal experiment in Sect. 3.1. Its core conclusion is that differentiable physics does not provide a higher-accuracy forward approximation; rather, it transforms an intrinsically ill-posed, over-parameterized inverse mapping into a constrained deconvolution problem that is physically anchored in the observable subspace and closed by the prior in the unobservable subspace.

(1) Linear structure of the decoder. The transport-chemistry decoder used here is a linear operator on the emission field E. Its discrete evolution follows $C_{n+1} = C_n + dt\,(E/M\_NO_2 + A_m C_n)$, where A_m is the combined monthly upwind-advection and first-order chemical-decay operator and the initial column is zero. Expanding step by step and summing over time, the simulated column can be written in the compact form of Eq. (3), where the transport-decay kernel is given in Eq. (4), a linear operator depending only on the wind field and the chemical lifetime. Thus, inverting the emissions E from an observed VCD_TROP reduces mathematically to a deconvolution problem for the linear operator K_m, rather than nonlinear optimization in the general sense.

$$VCD_{sim} = K_m \cdot E \quad (3)$$

$$K_m = \frac{\Delta t}{M_{NO_2}} \sum_{k=0}^{N-1} (I + \Delta t \cdot A_m)^k \quad (4)$$

(2) Ill-posedness of the inverse problem. As an advection-decay smoothing operator, K_m imposes a spatial low-pass action on the emission field: high-frequency spatial structure is strongly attenuated in the forward mapping to columns, so the singular-value spectrum of K_m decays rapidly and forms an approximate null space. Therefore, directly inverting from a finite, noisy VCD_TROP is a typical ill-posed problem, and neither fine-scale spatial structure nor the absolute emission magnitude can be uniquely determined by the observations. PINE-NOx does not invert directly; instead, it regularizes in a variational physical-likelihood-plus-prior form, injecting controllable constraints along the ill-posed directions.

(3) Amortized maximum-a-posteriori estimation. For the observation of any given day, the ideal regularized inversion corresponds to the maximum-a-posteriori (MAP) objective of Eq. (5), where Ω is the valid-observation mask, TV is the total-variation prior, and H is the Huber prior on the log deviation δ from EDGAR. Rather than solving this optimization separately for each day, PINE-NOx trains a single neural encoder g_θ that outputs $E_*(t) \approx E\_EDGAR \exp(g_\theta(x_t))$, amortizing the MAP solution over all days with one shared network. In this framework the encoder approximates the posterior mode, the fixed differentiable physics provides the analytical gradient of the likelihood term, and the prior term provides regularization, the three being co-optimized through end-to-end back-propagation.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

$$E_*(t) = argmin_E \, \|log(K_m \cdot E) - log\, VCD_{local}(t)\|_\Omega^2 + \lambda_{tv} \cdot TV(\delta) + \lambda_{prior} \cdot H(\delta) \quad (5)$$

(4) Identifiability decomposition. The above framework naturally partitions the emission degrees of freedom into two complementary subspaces. The first is the observable subspace, namely the spatial modes spanned by the row space of K_m that contribute significantly to the column; this subspace is directly anchored by the physical-likelihood term, and its inversion is determined entirely by the TROPOMI observations together with the differentiable physics, without relying on any existing emission-inversion product as a label. The second is the unobservable subspace, namely the fine-scale structure that falls into the approximate null space of K_m together with the overall magnitude; this subspace is nearly indistinguishable in the observations and is instead closed by the prior: the total-variation prior suppresses non-physical high-frequency oscillations, the Huber prior softly anchors it to the spatial structure of EDGAR, and the overall absolute magnitude is finally calibrated by the single domain-wide scalar normalization of Sect. 2.5.

(5) Immunity to rank-one perturbations through mean-centring. The rendering loss removes the domain-wide mean from both the simulated and the observed VCD, which is equivalent to orthogonally projecting out the rank-one perturbation subspace spanned by overall gain and background from the data term. The physical meaning is that an error in the absolute value of the chemical lifetime τ_chem manifests mainly as an approximately multiplicative overall scaling of the column, whereas an error in the regional background estimate manifests mainly as an additive overall shift; both are eliminated to first order by demeaning. The loss function is therefore first-order invariant to τ_chem calibration error and background-estimation error, making the inverted spatial pattern robust to these two systematic uncertainties, a property confirmed empirically by the τ_chem sensitivity experiment in Sect. 4.1 (the spatial correlation is essentially unchanged even when the lifetime is perturbed by several-fold).

In summary, the role of the fixed differentiable physics is to reconstruct an over-parameterized inverse mapping, one prone to degenerating into an image memorization of the observations, into a constrained deconvolution that rewards likelihood only for emissions consistent with physical transport; the prior and the domain-wide normalization then close the null-space directions that the observations cannot identify. This identifiability structure, namely spatial pattern from observations and physics and absolute magnitude from the prior constraint, is precisely the theoretical basis on which PINE-NOx can significantly improve the spatial self-consistency of emissions (Sect. 3.1) and pass the WRF-Chem chemical-transport-model consistency check (Sect. 3.2) without using any emission label.

2.5 Domain-wide annual-total emission normalization

The rendering loss of the neural encoder removes the domain-wide mean from both the simulated and the observed VCD, so it constrains only the relative spatial pattern and day-to-day variability of emissions and contains no absolute-magnitude information; the network's inverted emissions thus have an overall-magnitude degree of freedom. To provide an absolute-magnitude reference, a single domain-wide scalar annual-total normalization is introduced as a design constraint, as in Eq. (6) ($c \approx 0.18$ in this study). This normalization only locks the domain's annual total emission to EDGAR and does not change the relative

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

spatial pattern or seasonal cycle (the independence of this normalization from the spatial pattern is shown in Fig. S8). Because the normalization is a single domain-wide constant, all log-space correlations and log-RMSE values are mathematically strictly unaffected by its value (log(E c) = log(E) + log c is merely a constant shift); hence the agreement of the annual total with EDGAR is a design constraint on the absolute magnitude rather than an independent validation. The evaluation of the method's validity rests on the improvement of the spatial pattern, the encoder-architecture comparison and seasonal generalization (Sect. 3.1), and the WRF-Chem chemical-transport-model evaluation (Sect. 3.2).

$$c = \frac{\Sigma_{i,j} E_{EDGAR}}{\Sigma_{i,j}\, mean_t[E]}, \quad E_{final} = c \cdot E \tag{6}$$

3 Results

3.1 Encoder architecture comparison and forward physical self-consistency

Within a unified label-free physical-reconstruction framework, the spatial consistency between the forward-decoder-simulated VCD and the TROPOMI observations was evaluated with the log-space Pearson correlation coefficient; the correlation was computed on 80 validation days held out evenly across the four seasons (20 days per season), without using any emission label at any stage. Driving the decoder with the EDGAR annual-mean emissions, the spatial correlation between the simulated VCD and the observations averages only r = 0.395, reflecting the inherent limitation of a static annual-mean inventory in representing day-to-day spatial patterns. Within this framework we systematically compare the true forward skill of four neural-encoder backbones (Fig. 4a, Table 2): UNet-CBAM (5.48M parameters, r = 0.837), spectral-fusion UNet (6.53M, r = 0.835), UNet-SA (6.27M, r = 0.834) and the Fourier neural operator FNO (2.10M, r = 0.823). All four backbones far exceed the EDGAR prior and differ very little from one another (0.823–0.837), indicating that, under a unified random seed, training under the differentiable-physics constraint is highly robust to the encoder architecture (seasonally stratified cross-validation further supports the generalization robustness of this skill; Fig. S7). UNet- CBAM, with the highest spatial correlation r = 0.837, was selected as the production backbone (Δr = +0.442, consistently positive on all validation days, $p < 0.001$). Unlike the multi-step paradigm of first solving the inversion day by day and then training a network to fit it, here all backbones are trained end-to-end with the same physical-reconstruction objective and ranked by their true forward skill against the observations (rather than by their fidelity to an existing inversion product), making this a fair, circularity-free architecture evaluation. We emphasize that this correlation is a spatial-pattern skill after demeaning in log space; the network targets reconstruction of the observed pattern and does not constrain absolute magnitude, so a magnitude-sensitive absolute RMSE is not used as the primary metric here. We also note that the forward skill above is computed with the same differentiable decoder used in training and is therefore an internal self-consistency metric; external checks of the inverted emissions are provided by the independent WRF-Chem simulation in Sect. 3.2 and the fully independent MEIC inventory in Sect. 3.3.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

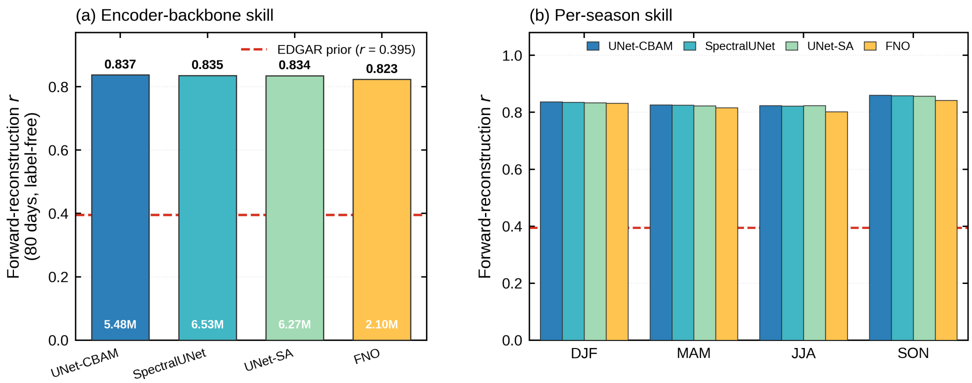

Figure 4. Encoder-backbone ablation. The forward reconstruction skill r is defined as the spatial correlation coefficient between the simulated column rendered by the physical decoder and the TROPOMI observed column in log-column (log-VCD) space. (a) Overall skill of the four encoder backbones, with the parameter count annotated within each bar, against the EDGAR prior baseline (r = 0.395); (b) the same skill decomposed by season.

Table 2. Comparison of PINE-NOx encoder architectures.

Encoder backbone	Parameters (M)	Overall forward r	Δr (vs EDGAR)
UNet-CBAM	5.48	0.837	+0.442
Spectral-fusion UNet (UNet-FNO)	6.53	0.835	+0.440
UNet-SA	6.27	0.834	+0.439
FNO	2.10	0.823	+0.428

Figure 5 illustrates the multi-season forward self-consistency of the inverted emissions for three representative days: a clean summer day (10 August 2023), a transitional autumn day (7 October 2023) and a polluted winter day (18 December 2023). On each day, the simulated columns obtained by driving the decoder with the daily PINE-NOx emissions reproduce the TROPOMI observed morphology (r = 0.872 / 0.888 / 0.792) markedly better than the unadjusted EDGAR prior (r = 0.600 / 0.533 / 0.390), with the improvement holding consistently under both clean and polluted conditions. Extending the skill to the full year (Fig. 6), the monthly forward correlation of the PINE-NOx-inverted emissions is systematically higher than that of the prior, and the annual-mean skill rises from 0.395 for EDGAR to 0.837, about twice the prior. With the validation set held out evenly by season (20 days per season; 80 days in total), the forward skill of UNet-CBAM on the held-out samples is 0.837 in winter (DJF), 0.826 in spring (MAM), 0.824 in summer (JJA) and 0.860 in autumn (SON), with small inter-seasonal fluctuation, indicating that the network's inversion skill does not depend on a particular season and is robust for year-round extrapolation. This robustness stems from the hard constraint of the physical decoder on the advection-chemistry processes and the seasonal chemical-lifetime parameterization, which allow the encoder to generalize without season-specific tuning.

To directly address the central question of whether the spatial-pattern improvement arises solely from the EDGAR prior structure, we designed a prior-removal experiment: when the Huber reference prior is replaced from the real EDGAR spatial field by a uniform prior stripped of all spatial information (a flat, domain-wide constant prior), the spatial correlation between the network's inverted forward VCD and TROPOMI is

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

essentially unchanged (Fig. S9) and remains stable across a full sweep of λ_{prior}. Because a 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 pattern leaks from the prior. An independent cross-comparison with ground-based CNEMC observations (Fig. S6) further supports the spatial plausibility of the inverted emissions.

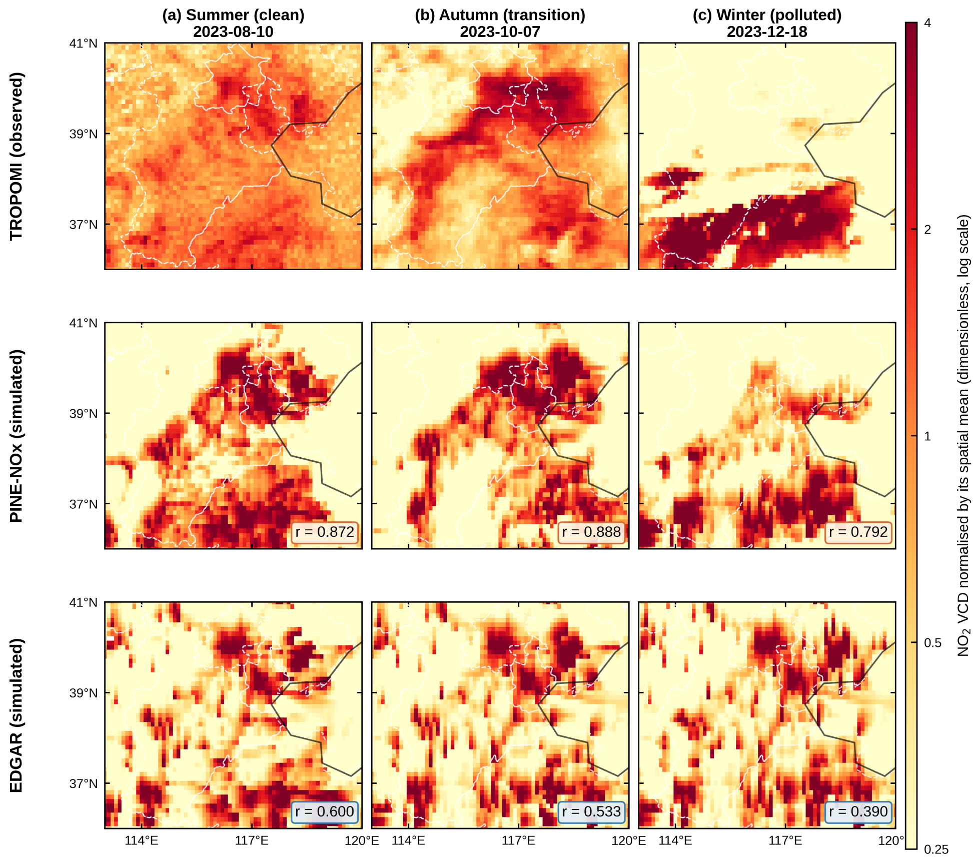

Figure 5. Multi-season forward self-consistency of the inverted emissions. Each column is one of three representative days spanning the seasonal range of the North China Plain: (a) a clean summer day (10 August 2023), (b) a transitional autumn day (7 October 2023) and (c) a polluted winter day (18 December 2023). From top to bottom, the rows are: the TROPOMI observed tropospheric NO_2 column, the column simulated by the differentiable transport-chemistry decoder driven by the daily PINE-NOx emissions, and the column simulated under the unadjusted EDGAR prior. The value in each cell of the lower two rows is the log-space spatial Pearson correlation between the simulated and observed columns for that day. White dashed lines are provincial boundaries.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

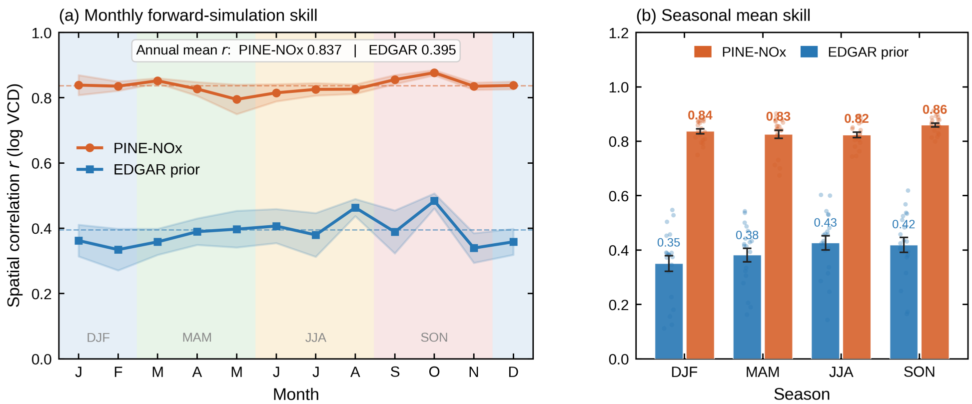

Figure 6. Forward-simulation skill of the inverted emissions over the full year. The skill r is defined as the log-space spatial Pearson correlation between the forward simulation and the TROPOMI observed NO_2 column. (a) Monthly-mean r for PINE-NOx (orange) and the EDGAR prior (blue); the shaded band is the ±1 standard error of the daily values within each month, the background colour blocks indicate the meteorological seasons, and the dashed line and inset box give the annual-mean skill. (b) Season-aggregated mean skill; bars are seasonal means, whiskers are ±1 standard error, and dots are single-day values.

3.2 Independent WRF-Chem chemical transport model evaluation

The forward self-consistency evaluations above are all based on the same differentiable decoder used in the inversion; to provide an external reference independent of that decoder, we carried out a WRF-Chem three-dimensional forward-simulation experiment (July 2023, about 24 days of integration, 27 km horizontal resolution, the RADM2 gas-phase chemical mechanism with the SORGAM aerosol module, CAMS EAC4 chemical boundary conditions and a cold start of the chemical initial field), driven separately by the EDGAR prior and by the PINE-NOx-inverted emissions, to evaluate NO_2 columns over the North China Plain independently (a comparable WRF-Chem–TROPOMI NO_2 cross-evaluation strategy is used by Poraicu et al., 2023). Taking the TROPOMI overpass time (about 13:30 local solar time) as the temporal reference, we computed the monthly-mean tropospheric NO_2 VCD over the NCP for 8–31 July 2023 and compared it directly with TROPOMI observations at the pixel scale (Fig. 7). The results show that EDGAR-driven WRF-Chem systematically overestimates: the domain-mean NO_2 VCD reaches 10.51×10^{15} molec cm^{-2}, about +113% higher than TROPOMI (4.94×10^{15} molec cm^{-2}); when driven instead by PINE-NOx emissions, the domain-mean VCD drops to 5.19×10^{15} molec cm^{-2}, and the bias falls sharply to +5%, in close agreement with the observations. For the spatial pattern, the pixel-level log-space correlation of WRF+PINE-NOx rises from 0.53 for WRF+EDGAR to 0.65, indicating that PINE-NOx corrects the magnitude bias of the WRF-Chem-simulated columns and improves their spatial distribution. We note that the roughly 51% reduction in the simulated column is larger than the reduction of the inverted emissions relative to EDGAR (under the annual-total consistency constraint of Sect. 3.3, summer emissions are lower than EDGAR): in the full gas-phase chemistry of WRF-Chem the column responds nonlinearly to NOx emissions, because as emissions decrease the OH suppression in the high-concentration core weakens and the effective chemical lifetime shortens, so the column decreases super-linearly; hence a column ratio should not be equated directly with an emission

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

ratio. This experiment provides, from the perspective of an external three-dimensional CTM, a consistency reference independent of the inversion decoder. Its independence boundary must be defined objectively: the constraint reference used by WRF-Chem (TROPOMI NO_2 VCD) is the same source as the observations used in the PINE-NOx inversion, so strictly speaking this experiment validates the regional column self-consistency of PINE-NOx within an independent three-dimensional CTM physical framework, rather than a validation fully independent of the observations; the EDGAR overestimate also partly arises because its static 2022 inventory does not reflect the continued decline of NCP NOx in 2023 (van der A et al., 2017; Geng et al., 2021). In addition, this experiment is a roughly 24-day integration with a chemical cold start, and the pixel comparison does not apply the TROPOMI averaging kernel; a proof-of-concept averaging-kernel correction shows that the averaging kernel scales the EDGAR- and PINE-NOx-driven WRF columns by almost the same factor, changing the compared quantities by only about 5% and leaving the spatial-pattern correlation essentially unchanged (Fig. S11), so applying the averaging kernel does not alter the core conclusion that PINE-NOx is closer to the satellite than EDGAR. A truly fully independent emission validation still depends on non-satellite observations such as aircraft and ground-based flux measurements, which we leave for future work (Cooper et al., 2017; Goldberg et al., 2022).

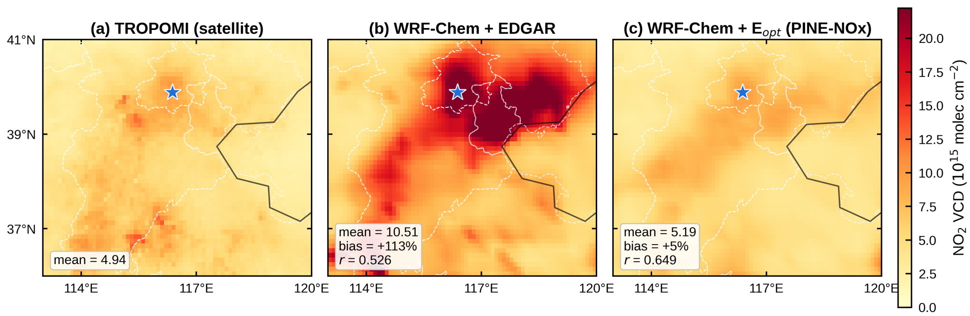

Figure 7. Independent summer chemical-transport-model evaluation (8–31 July 2023). Tropospheric NO_2 column: (a) TROPOMI observations, (b) WRF-Chem simulation driven by the EDGAR prior, and (c) WRF-Chem simulation driven by PINE-NOx emissions, sharing the same absolute colour scale. The blue star marks Beijing; white dashed lines are provincial boundaries.

3.3 Spatiotemporal characteristics of NOx emissions over the North China Plain

The most prominent capability of PINE-NOx is to raise the temporal resolution of emissions from static annual means to the daily scale. Figure 8 shows the domain-mean emissions of all 364 daily inversions in 2023 (normalized to the annual mean). We note that single-day inversions exhibit some day-to-day scatter (day-to-day coefficient of variation CV ≈ 29%) owing to the satellite's single-overpass sampling and observational noise; however, their lag-1 autocorrelation reaches 0.249 (significantly greater than zero), and the 7-day rolling mean depicts a robust weekly-to-monthly evolution, indicating that the inverted signal has genuine temporal structure rather than being purely random inversion error. We therefore use the 7-day rolling mean as the interpretable emission-dynamics product and represent the single-day uncertainty by a ±1σ within-window scatter band; the static EDGAR prior contains no temporal information. The monthly

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

distribution of the daily values (boxes = interquartile range) further reveals a pronounced high-winter, low-summer seasonal cycle, as well as within-month day-to-day scatter that a static inventory cannot represent. This emission field, interpretable at the weekly-to-monthly scale and continuous at the daily scale, is the core increment of this framework relative to existing static inventories and annual-mean inversion products.

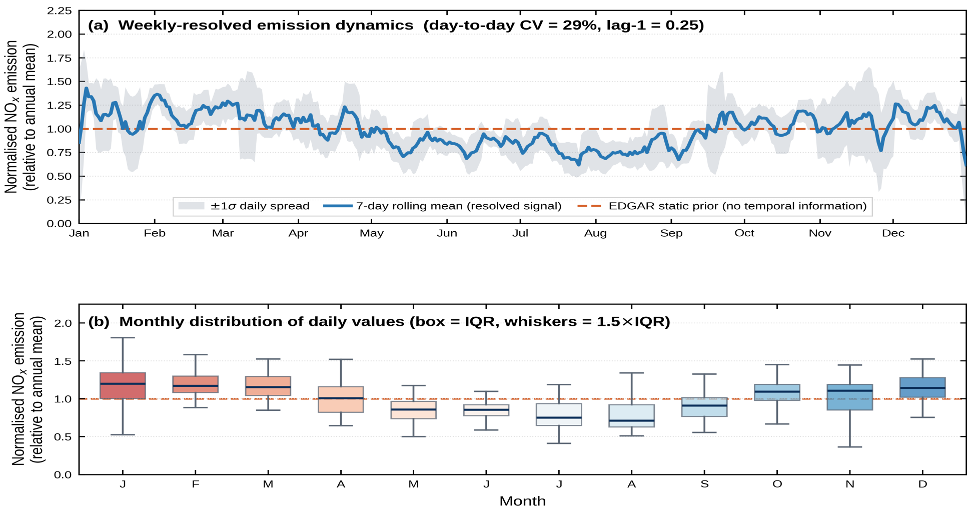

Figure 8. Daily dynamics of the inverted NOx emissions in 2023 (364 daily inversions). (a) Domain-mean emissions normalized to the annual mean: the 7-day rolling mean (blue line, the interpretable weekly-to-monthly signal) with a ±1σ within-day scatter band (grey band), and the EDGAR static prior with no temporal information (dashed line); the day-to-day coefficient of variation CV is approximately 29% and the lag-1 autocorrelation is about 0.25. (b) Monthly distribution of the daily values (boxes = interquartile range, whiskers = 1.5 × IQR), showing the high-winter, low-summer seasonal cycle.

At the seasonal scale, the domain-mean inverted emissions show a pronounced high-winter, low-summer cycle (Fig. 9): although the EDGAR prior has no seasonality (6.14 nmol m^{-2} s^{-1}), the inversion still recovers a clear seasonal cycle from TROPOMI, from a January peak (7.27 nmol m^{-2} s^{-1}) to an August trough (4.84 nmol m^{-2} s^{-1}). To strip out the static annual-mean spatial pattern and examine the seasonal pattern alone, the seasonal mean of E_final in each season was divided by the annual mean to obtain a pure seasonal-modulation field (Fig. 10): winter shows a domain-wide consistent enhancement (domain-mean ratio 1.12), strongest over the Shandong-Bohai industrial and heating belt, whereas spring, summer and autumn are close to or slightly below the annual mean (0.97 / 0.95 / 0.97). Spatially, the inverted emission hotspots are concentrated in three core regions: the central-southern Hebei industrial corridor (Shijiazhuang-Handan-Xingtai, dominated by iron, steel and cement), the Beijing-Tianjin urban agglomeration (with a significant contribution from traffic sources), and the energy-industry base in north-western Shandong (Jinan-Zibo). Owing to advective transport by the prevailing winds, the spatial peak of the VCD is systematically displaced downwind of the emission source, especially in winter, underscoring the necessity of explicitly modelling transport with a differentiable advection decoder.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

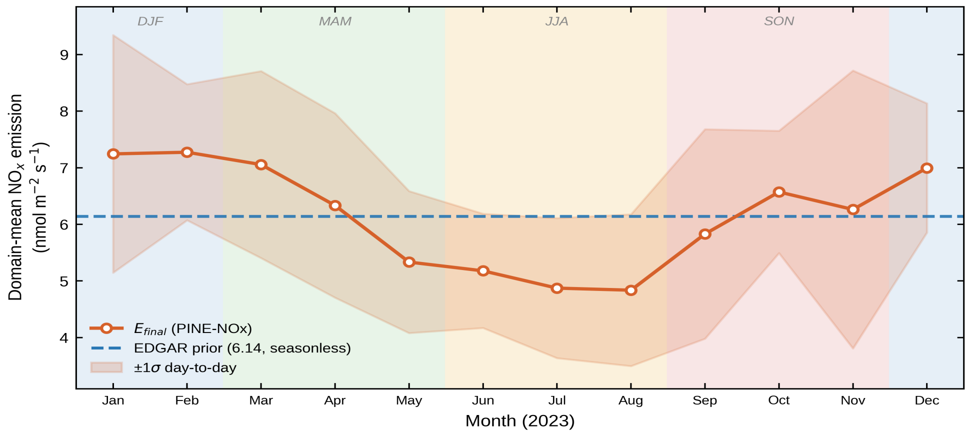

Figure 9. Seasonal cycle of the domain-mean inverted NOx emissions over the North China Plain (2023). Orange line: monthly domain-mean PINE-NOx emissions, with shading for their ±1σ day-to-day scatter; blue dashed line: the EDGAR prior with no seasonality (6.14 nmol m⁻² s⁻¹).

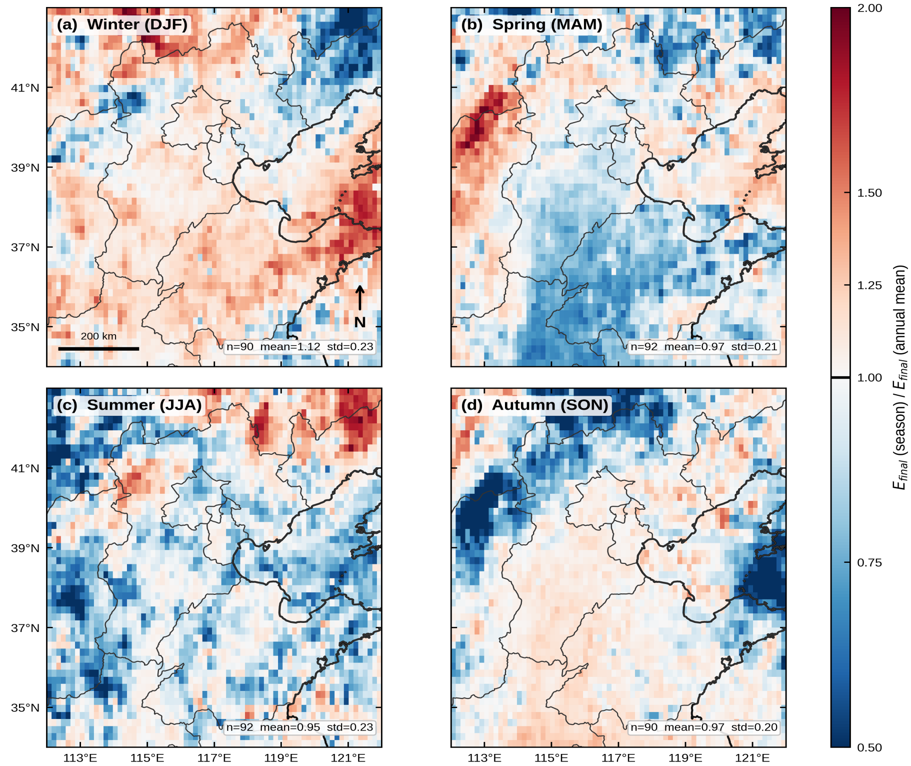

Figure 10. Pure seasonal modulation of the inverted NOx emissions over the North China Plain (2023). Each panel is the seasonal mean of E_final divided by the annual mean of E_final, removing the static annual spatial structure to isolate the seasonal pattern. Values greater than 1.0 (red) indicate enhancement in that season and less than 1.0 (blue) indicate reduction, sharing a diverging colour scale centred on 1.0.

At the annual-mean scale (Fig. 11), the grid-cell log-space correlation between the PINE-NOx annual-mean emission field and the EDGAR prior is as high as r = 0.941, consistent with the expectation that the prior

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

anchors the annual-mean spatial structure. Under the design constraint that the annual total matches the prior, the inversion achieves about 11% spatial redistribution (still significant after correction for spatial autocorrelation, Fig. S13), manifested mainly as a moderate downward adjustment of the pollution cores (urban Beijing, central-southern Hebei point sources) and an upward adjustment of the clean periphery and rural background emissions, consistent with the more diffuse emission distribution than EDGAR implied by the TROPOMI columns. The distinction between the seasonality of the column and that of the emissions must be clarified: under quasi-steady state the tropospheric NO_2 column is approximately proportional to the product of emission flux and effective chemical lifetime (Seinfeld and Pandis, 2016), so most of the several-fold winter-summer difference in the observed column is caused by the seasonal amplification of the chemical lifetime. Decomposing under a consistent seasonal-mean convention: the ratio of the domain's TROPOMI VCD winter/summer (DJF/JJA) seasonal means is about 3.8 (the commonly cited 6 is in fact the December/June peak-to-trough ratio, a different convention that should not be conflated), the winter/summer ratio of the effective chemical lifetime is about 2.8, and dividing the two gives an emission winter/summer ratio of about 1.35 (the full decomposition under a consistent seasonal-mean convention is shown in Fig. S12); this value is self-consistent in magnitude and direction with the domain-mean winter/summer ratio inverted directly from E_final (about 1.42). NCP NOx emissions therefore do have a stronger-in-winter, weaker-in-summer seasonal cycle (domain-mean winter/summer ratio of about 1.4, driven mainly by increased heating-season emissions), but its amplitude is far smaller than the several-fold difference shown directly by the column, most of which is caused by the seasonal amplification of the chemical lifetime. Combining systematic sources such as the monthly-mean scalar chemical lifetime, the overall uncertainty of the inverted emissions is about 12–15% (the seasonally stratified four-fold cross-validation generalization assessment is shown in Fig. S7; the uncertainty quantification is given in Sect. 4.1 and Fig. S10).

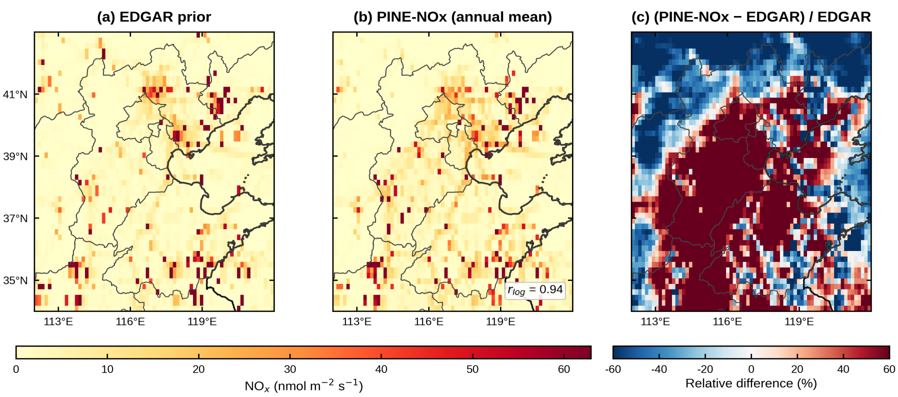

Figure 11. Annual-mean NOx emissions over the North China Plain (2023): (a) the EDGAR prior, (b) the PINE-NOx inversion, and (c) their relative difference (PINE-NOx - EDGAR)/EDGAR. The annual-mean spatial pattern is highly correlated with the prior (log-space correlation r = 0.94), with enhanced emissions in the pollution cores (red) and reduced emissions in the clean periphery (blue), truncated at ±60%.

Finally, we cross-validate the inversion results against the fully independent bottom-up MEIC inventory (Fig. 12). MEIC never enters the inversion, and the EDGAR prior has no seasonality, so this comparison is an

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

independent test of the realism of the inversion's temporal dynamics. In terms of annual total, PINE-NOx and the EDGAR prior are both 3.14 Tg and MEIC is 2.67 Tg, all of the same order. Spatially, the PINE-NOx annual-mean pattern is dominated by the prior (log-space correlation r = 0.94 with EDGAR), and its correlation with the independent and smoother MEIC is r = 0.574. Most convincing is the seasonal shape: after normalizing the monthly emission cycles of PINE-NOx and MEIC each to their own annual mean, the two seasonal shapes agree significantly, even though the winter/summer contrast of PINE-NOx (1.42) is stronger than that of MEIC (1.15). Because neither MEIC nor the EDGAR prior provides any seasonal information to the inversion, the independent agreement between PINE-NOx and MEIC in seasonal shape constitutes strong external corroboration of the realism of the inferred temporal dynamics. We note honestly that the MEIC monthly profile is taken from the representative seasonal distribution of its most recent available year and is used here only for an independent comparison of seasonal shape, not in the inversion, so its agreement with PINE-NOx in seasonal phase is unaffected by small differences in the specific reference year.

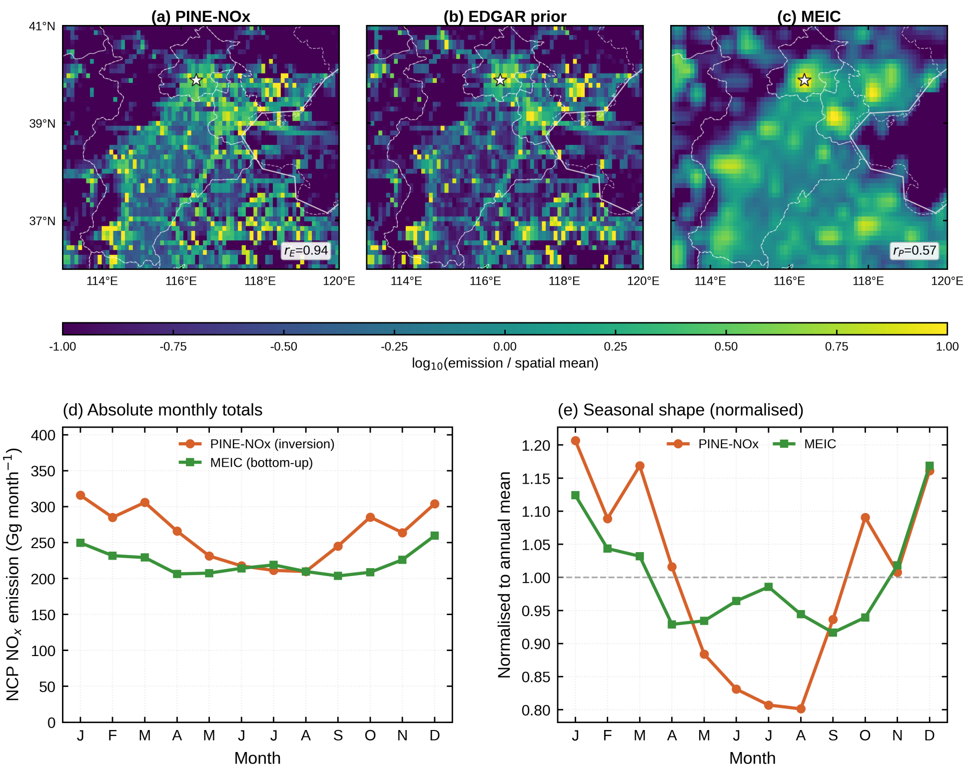

Figure 12. Independent cross-validation against the bottom-up MEIC inventory (2023). (a–c) Annual-mean spatial NOx patterns of PINE-NOx, the EDGAR prior and MEIC, each divided by its own spatial mean (log scale, shared colour scale); insets give the log-space correlations of PINE-NOx with EDGAR (r_E = 0.94) and with MEIC (r_P = 0.57). (d) Domain-total monthly emissions of PINE-NOx (daily inversions aggregated to months) and MEIC. (e) The same monthly cycles normalized to their respective annual means: the seasonal shapes agree (r = 0.68, p =

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

0.015), but the winter-summer contrast of PINE-NOx (1.42) is stronger than that of MEIC (1.15). The white star marks Beijing.

4 Discussion

4.1 Robustness of the chemical-lifetime parameterization

The monthly-mean scalar τ_chem parameterization is a key simplification of the forward model. The decoder renders the column by explicit integration over about 4 h in 32 steps; this window already reaches the order of the summer NOx lifetime and covers a substantial fraction of the winter lifetime; the inversion and the forward evaluation share the same fixed operator, so the conclusions are internally consistent and do not rely on a strict steady-state assumption. To systematically assess the impact of this simplification on the inversion conclusions, we conduct a progressive test at three levels: parameter sensitivity, physical-mechanism refinement and regional applicability. Over the full combinatorial parameter space of τ_winter in {4, 7, 10, 15, 20} h and τ_summer in {0.8, 1.8, 3, 5} h, the variation of Δr does not exceed 0.3% (minimum 0.4512, maximum 0.4532; Fig. S2), demonstrating that the method is highly robust to the choice of the absolute value of τ_chem, a robustness rooted in the framework's optimization target of the demeaned-VCD spatial pattern, which is insensitive to the absolute value of τ_chem and responds only to first order to its spatial gradient. Drawing further on observational evidence that the NOx lifetime varies with temperature and photochemical activity (Laughner et al., 2019; Stavrakou et al., 2013), we introduce a pixel-level dynamic τ_chem driven by the ERA5 daily surface air temperature via the Arrhenius equation (effective activation energy E_eff/R = 3000 K), whose four-season spatial coefficient of variation is about 13–19% (Fig. S3); yet the improvement in Δr is only +1.2% (scalar τ: 0.4521; dynamic τ: 0.4573; Figs. S3–S4), an improvement within the range of current measurement uncertainty. Considering both the very low parameter sensitivity and the limited improvement of the dynamic scheme, the monthly-mean scalar parameterization is fully justified within the accuracy framework of this paper; the emission uncertainty introduced by this τ_chem simplification is about 12–15% (a quantification of the τ-simplification source alone, excluding other sources such as observational inversion, the wind field and normalization), and the dynamic-τ_chem scheme is retained as an explicit direction for future improvement. In addition, a wind-zeroing advection ablation applied to the differentiable decoder shows that advective transport makes a genuine but secondary contribution to the forward skill: with advection switched off, the forward spatial correlation on the 80 validation days drops from 0.837 to 0.746 (still far above EDGAR's 0.395), with the reduction being largest in spring (Δr ≈ 0.13) and smallest in summer (Δr ≈ 0.07), a seasonality consistent with the seasonal variability of wind speed and the seasonal difference in NOx lifetime, indicating that the decoder genuinely exploits transport physics rather than a local rescaling of the prior.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

4.2 Comparison with existing inversion methods

The core advantage of PINE-NOx is the end-to-end fusion of a differentiable physical forward decoder with a neural encoder into a label-free physics-constrained autoencoder, greatly improving computational efficiency while preserving physical interpretability; a multidimensional comparison with representative satellite inversion methods is given in Table 3. Compared with 4D-Var, it requires no complex adjoint-model derivation, substantially lowering the engineering barrier, with single-day inference at the millisecond scale; compared with the FDM, it explicitly handles the diffuse-source signal in high-background regions and, through a regional-applicability assessment (Fig. S5), demonstrates its suitability in the high-density source region of the NCP; compared with methods supervised by existing inversion products, PINE-NOx relies on no emission label, and emissions are inverted directly as a physical latent variable, fundamentally eliminating circular dependence.

Table 3. Comparison of PINE-NOx with representative satellite NOx emission-inversion methods.

Comparison dimension	FDM (flux divergence)	Global CTM assimilation (EnKF / 4D-Var)	DECSO (local fast)	PINE-NOx (this work)
Representative literature	Beirle et al., 2011, 2019	Miyazaki et al., 2012, 2021	Mijling et al., 2013	This work
Physical interpretability	High (analytical)	High (ensemble/adjoint)	Medium	High (differentiable physics + physical regularization)
Requires CTM / adjoint	No	Yes	No	No
Computational cost	Low	Very high (hundreds of CPU h per year)	Medium	Millisecond-scale after training (approximately 5 ms per day)
Temporal resolution	Multi-year / monthly mean	Daily (costly)	Daily (medium resolution)	Daily (continuous, full year)
Fast inference after training (ms)	No	No	No	Yes
Applicability to NCP high-density sources	Limited (insufficient SNR, see Fig. S5)	Applicable	Applicable	Applicable

4.3 Limitations and future directions

The current framework has several limitations worth stating explicitly. First, the forward evaluation uses the demeaned log-space spatial correlation as its metric, characterizing spatial-pattern skill; the network targets reconstruction of the observed pattern and does not constrain absolute magnitude, so the RMSE of the daily absolute magnitude is not its optimization objective, and conclusions should be interpreted primarily in terms of spatial pattern and seasonal cycle. To provide an external independent reference, we carried out a WRF-Chem three-dimensional forward-simulation consistency experiment (see Sect. 3.2 and Fig. 7). We emphasize that the absolute magnitude of the inversion is not determined independently by the observations: the rendering loss is demeaned log, and the domain annual total is locked to EDGAR by design, so the agreement of the annual total with EDGAR is a design constraint rather than an independent validation; what is genuinely observationally constrained is the spatial pattern and the temporal dynamics. Besides WRF-

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

Chem, the fully independent bottom-up MEIC inventory, which never enters the inversion, also corroborates the realism of the inversion's temporal dynamics through its seasonal shape (see Sect. 3.3 and Fig. 12). Second, the monthly-mean scalar chemical lifetime is a simplification, whose sensitivity is examined in Sect. 4.1 and Figs. S2–S4. In addition, advection uses the ERA5 10 m daytime-mean wind to approximate column-mean transport, which may underestimate transport aloft; however, the advection ablation shows this to be a secondary contribution ($\Delta r \approx 0.09$; see Sect. 4.1), and the inversion and the evaluation share the same wind operator, so this simplification has limited influence on the conclusions. Finally, the framework currently targets a single species (NOx) and region (the NCP); extending it to other pollutants (e.g. SO_2, NH_3; for SO_2, for instance, space-based point sources have been catalogued by Fioletov et al., 2016) and regions, and fusing multi-source satellite and ground-based flux observations to constrain the absolute magnitude independently, is a worthwhile future direction.

A further intrinsic limitation concerns temporal sampling. Sentinel-5P/TROPOMI provides only a single overpass per day (about 13:30 local solar time), so each retrieval is an instantaneous snapshot of the NO_2 column at a fixed hour, whereas the quantity we infer is a daily-scale NOx emission rate. Reconstructing a day-scale emission flux from one instantaneous column therefore entails an inherent temporal-sampling information gap: the diurnal cycle of emissions and of the chemical lifetime is not directly observed, and is instead absorbed into the effective daily quantities through the demeaned rendering loss and the monthly-mean lifetime parameterization. A natural way forward is to fuse multiple, temporally complementary observational constraints: continuous ground-based networks provide near-surface, high-frequency anchoring, while the new generation of geostationary spectrometers (GEMS over East Asia, together with TEMPO and Sentinel-4) delivers hourly daytime columns that, combined with polar-orbiting sensors such as OMI and TROPOMI, would substantially densify the temporal and spatial sampling. Embedding such multi-source constraints within the same differentiable, physics-constrained framework is a promising route to a more accurate inventory that explicitly resolves the intraday dynamics of NOx emissions, which we leave for future work.

4.4 Implications for policy applications

The winter/summer ratio of NOx emissions revealed by this study (about 1.4, domain-mean DJF/JJA of E_final; the true seasonal amplitude of emissions after removing atmospheric physical effects) indicates that emission-reduction policies based on annual-mean inventories may systematically underestimate the health-exposure risk of peak winter emissions. The daily-resolution continuous emission field provided by E_final can directly drive air-quality forecasting models, offering a more accurate initial emission constraint for heavy-pollution weather forecasting. Given that NOx is a key precursor of ozone and PM2.5 (Jin and Holloway, 2015; Lu et al., 2019; Geng et al., 2021), a high-temporal-resolution emission constraint has direct value for the refined simulation of precursor-oxidant coupling.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

5 Conclusions

We have proposed and validated the PINE-NOx emission-inversion framework, which fuses a differentiable physical forward decoder and a neural encoder end-to-end into a label-free physics-constrained autoencoder, achieving daily-resolution dynamic inversion of NOx emissions over the North China Plain from TROPOMI and systematically comparing convolutional, spectral-operator and spectral-fusion encoder backbones. The physical-self-consistency evaluation shows that the forward VCD driven by the PINE-NOx-inverted emissions improves its spatial correlation with TROPOMI from $r = 0.395$ to $r = 0.837$ ($\Delta r = +0.442$, $p < 0.001$), with significant improvement in all four seasons (winter (DJF) 0.837, spring (MAM) 0.826, summer (JJA) 0.824, autumn (SON) 0.860), the UNet-CBAM encoder performing best. The prior-removal experiment confirms that the spatial pattern originates from the observational constraint rather than from prior leakage. The inverted inventory reveals a domain-mean winter/summer ratio of about 1.4 for NCP NOx and about 11% spatial redistribution relative to EDGAR. To our knowledge, this framework is the first to characterize the daily-resolution dynamics of NCP NOx in a label-free, physics-constrained manner (day-to-day domain-mean CV ≈ 29%, its weekly-to-monthly evolution being an interpretable signal); the recovered seasonal cycle agrees significantly with the seasonal shape of the fully independent MEIC inventory, and an independent WRF-Chem simulation confirms that driving the model with the inverted emissions reduces the summer domain-mean column bias from +113% to +5%; the three together support the reliability of the inversion. With its label-free, physically interpretable, millisecond-inference characteristics, PINE-NOx provides a new paradigm for atmospheric emission inversion.

Code and data availability

All datasets used in this study are publicly available. Sentinel-5P/TROPOMI tropospheric NO_2 columns are available from the Copernicus Sentinel-5P Data Hub (https://dataspace.copernicus.eu). ERA5 reanalysis fields are available from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu). The EDGAR v8.1 NOx emission inventory is distributed by the European Commission Joint Research Centre (https://edgar.jrc.ec.europa.eu). The MEIC anthropogenic emission inventory is available from the MEIC team, Tsinghua University (http://meicmodel.org.cn). Hourly surface NO2 observations are from the China National Environmental Monitoring Centre (CNEMC, https://www.cnemc.cn). For the independent WRF-Chem evaluation, CAMS EAC4 chemical fields are available from the Copernicus Atmosphere Monitoring Service (https://ads.atmosphere.copernicus.eu) and NCEP GDAS/FNL meteorological data from the NCAR Research Data Archive (https://rda.ucar.edu). Coal-fired power-plant locations and capacities shown in Fig. 1 are from the WRI Global Power Plant Database, which builds on the Global Coal Plant Tracker. WRF-Chem is an open-source community model (https://www2.mmm.ucar.edu/wrf/users/). The PINE-NOx daily NOx emission inventory generated in this study and the figure source data is openly available on Zenodo at https://doi.org/10.5281/zenodo.20783847 (Wang et al., 2026) under the Creative Commons Attribution 4.0 International licence.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

Author contributions

Y. Wang and Y. Pan conceived and designed the study. Y. Wang developed the PINE-NOx framework and the differentiable physical decoder, implemented the software, performed the inversion experiments and the WRF-Chem simulations, carried out the formal analysis and visualization, and wrote the original draft. Y. Pan provided supervision and computational resources, guided the methodology and interpretation of results, and revised the manuscript. D. Lyu provided scientific guidance and advice throughout the study, contributed to the interpretation of the results, and reviewed and edited the manuscript. All authors have read and approved the final version.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Yinan Wang gratefully acknowledges the support of Phase II of the Chinese Meridian Project.

Financial support

This research has been supported by the National Natural Science Foundation of China (Grant No. 42293321).

References

Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., Wang, Y., and Wagner, T.: Pinpointing nitrogen oxide emissions from space, Sci. Adv., 5, eaax9800, https://doi.org/10.1126/sciadv.aax9800, 2019.

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T.: Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space, Science, 333(6050), 1737-1739, https://doi.org/10.1126/science.1207824, 2011.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, J. Geophys. Res., 109, https://doi.org/10.1029/2003JD003962, 2004.

Boersma, K. F., Eskes, H. J., Richter, A., Smedt, I. de, Lorente, A., Beirle, S., van Geffen, J. H. G. M., Zara, M., Peters, E., van Roozendael, M., Wagner, T., Maasakkers, J. D., van der A, R. J., Nightingale, J., Rudder, A. de, Irie, H., Pinardi, G., Lambert, J.-C., and Compernolle, S. C.: Improving algorithms and uncertainty estimates for satellite NO 2 retrievals: results from the quality assurance for the essential climate variables (QA4ECV) project, Atmos. Meas. Tech., 11, https://doi.org/10.5194/amt-11-6651-2018, 6651–6678, 2018.

Cooper, M., Martin, R. V., Padmanabhan, A., and Henze, D. K.: Comparing mass balance and adjoint methods for inverse modeling of nitrogen dioxide columns for global nitrogen oxide emissions, J. Geophys. Res., 122, 4718–4734, https://doi.org/10.1002/2016JD025985, 2017.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

Cranmer, K., Brehmer, J., and Louppe, G.: The frontier of simulation-based inference, Proc. Natl. Acad. Sci. U.S.A., 117, 30055–30062, https://doi.org/10.1073/pnas.1912789117, 2020.

Crippa, M., Guizzardi, D., Pagani, F., Banja, M., Muntean, M., Schaaf, E., Becker, W., Monforti-Ferrario, F., Quadrelli, R., Risquez Martin, A., Grassi, G., Rossi, S., Brandao De Melo, J., Jacome Felix Oom, D., Branco, A., San-Miguel, J., and Vignati, E.: EDGAR v8.1 Global Air Pollutant Emissions, European Commission, Joint Research Centre [data set], https://doi.org/10.2905/JRC.E1ZGQER, 2026.

Fioletov, V. E., McLinden, C. A., Krotkov, N., Li, C., Joiner, J., Theys, N., Carn, S., and Moran, M. D.: A global catalogue of large SO 2 sources and emissions derived from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 16, 11497–11519, https://doi.org/10.5194/acp-16-11497-2016, 2016.

de Foy, B., Lu, Z., Streets, D. G., Lamsal, L. N., and Duncan, B. N.: Estimates of power plant NOx emissions and lifetimes from OMI NO2 satellite retrievals, Atmos. Environ., 116, 1–11, https://doi.org/10.1016/j.atmosenv.2015.05.056, 2015.

Geng, G., Zheng, Y., Zhang, Q., Xue, T., Zhao, H., Tong, D., Zheng, B., Li, M., Liu, F., Hong, C., He, K., and Davis, S. J.: Drivers of PM2.5 air pollution deaths in China 2002–2017, Nat. Geosci., 14, 645–650, https://doi.org/10.1038/s41561-021-00792-32021.

Goldberg, D. L., Anenberg, S. C., Griffin, D., McLinden, C. A., Lu, Z., and Streets, D. G.: Disentangling the Impact of the COVID-19 Lockdowns on Urban NO 2 From Natural Variability, Geophys. Res. Lett., 47, e2020GL089269, https://doi.org/10.1029/2020GL089269, 2020.

Goldberg, D. L., Harkey, M., Foy, B. de, Judd, L., Johnson, J., Yarwood, G., and Holloway, T.: Evaluating NOx emissions and their effect on O3 production in Texas using TROPOMI NO2 and HCHO, Atmos. Chem. Phys., 22, 10875–10900, https://doi.org/10.5194/acp-22-10875-2022, 2022.

Goldberg, D. L., Lu, Z., Streets, D. G., de Foy, B., Griffin, D., McLinden, C. A., Lamsal, L. N., Krotkov, N. A., and Eskes, H.: Enhanced Capabilities of TROPOMI NO_2: Estimating NO_x from North American Cities and Power Plants, Environ. Sci. Technol., https://dx.doi.org/10.1021/acs.est.9b04488, 53, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. de, Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P. de, Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quart. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Jin, X. and Holloway, T.: Spatial and temporal variability of ozone sensitivity over China observed from the Ozone Monitoring Instrument, J. Geophys. Res., 120, 7229–7246, https://doi.org/10.1002/2015JD023250, 2015.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L.: Physics-informed machine learning, Nat. Rev. Phys., 3, 422–440, https://doi.org/10.1038/s42254-021-00314-5, 2021.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

Kinga, D., and Adam, J. B.: A Method for Stochastic Optimization, in: Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015, https://doi.org/10.48550/arXiv.1412.6980, 2015.

Kong, H., Lin, J., Zhang, R., Liu, M., Weng, H., Ni, R., Chen, L., Wang, J., Yan, Y., and Zhang, Q.: High-resolution (0.05° × 0.05°) NO x emissions in the Yangtze River Delta inferred from OMI, Atmos. Chem. Phys., 19, 12835–12856, https://doi.org/10.5194/acp-19-12835-2019, 2019.

Laughner, J. L. and Cohen, R. C.: Direct observation of changing NO x lifetime in North American cities, Science, 366, 723–727, https://doi.org/10.1126/science.aax6832, 2019.

Li, M., Zhang, Q., Kurokawa, J., Woo, J.-H., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D. G., Carmichael, G. R., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., Liu, F., Su, H., and Zheng, B.: MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP, Atmos. Chem. Phys., 17, 935–963, https://doi.org/10.5194/acp-17-935-2017, 2017.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A.: Fourier Neural Operator for Parametric Partial Differential Equations, arXiv [preprint], https://doi.org/10.48550/arXiv.2010.08895, 2020.

Liu, F., Beirle, S., Zhang, Q., Dörner, S., He, K., and Wagner, T.: NO x lifetimes and emissions of cities and power plants in polluted background estimated by satellite observations, Atmos. Chem. Phys., 16, 5283–5298, https://doi.org/10.5194/acp-16-5283-2016, 2016.

Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., Smedt, I. de, van Roozendael, M., Wang, Y., Wagner, T., Beirle, S., Lin, J.-T., Krotkov, N., Stammes, P., Wang, P., Eskes, H. J., and Krol, M.: Structural uncertainty in air mass factor calculation for NO 2 and HCHO satellite retrievals, Atmos. Meas. Tech., 10, 759–782, https://doi.org/10.5194/amt-10-759-2017, 2017.

Lu, X., Zhang, L., Chen, Y., Zhou, M., Zheng, B., Li, K., Liu, Y., Lin, J., Fu, T.-M., and Zhang, Q.: Exploring 2016–2017 surface ozone pollution over China: source contributions and meteorological influences, Atmos. Chem. Phys., 19, 8339–8361, https://doi.org/10.5194/acp-19-8339-2019, 2019.

Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., and Evans, M. J.: Global inventory of nitrogen oxide emissions constrained by space-based observations of NO 2 columns, J. Geophys. Res., 108, https://doi.org/10.1029/2003JD003453, 2003.

Mijling, B., van der A, R. J., and Zhang, Q.: Regional nitrogen oxides emission trends in East Asia observed from space, Atmos. Chem. Phys., 13, 12003–12012, https://doi.org/10.5194/acp-13-12003-2013, 2013.

Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NO x emission estimates derived from an assimilation of OMI tropospheric NO 2 columns, Atmos. Chem. Phys., 12, 2263–2288, https://doi.org/10.5194/acp-12-2263-2012, 2012.

Miyazaki, K., Bowman, K., Sekiya, T., Takigawa, M., Neu, J. L., Sudo, K., Osterman, G., and Eskes, H.: Global tropospheric ozone responses to reduced NO x emissions linked to the COVID-19 worldwide lockdowns, Sci. Adv., 7(24), https://doi.org/10.1126/sciadv.abf7460, 2021.

https://doi.org/10.5194/egusphere-2026-3233
Preprint. Discussion started: 3 July 2026
© Author(s) 2026. CC BY 4.0 License.

Poraicu, C., Müller, J.-F., Stavrakou, T., Fonteyn, D., Tack, F., Deutsch, F., Laffineur, Q., van Malderen, R., and Veldeman, N.: Cross-evaluating WRF-Chem v4.1.2, TROPOMI, APEX, and in situ NO_2 measurements over Antwerp, Belgium, Geosci. Model Dev., 16, 479–508, https://doi.org/10.5194/gmd-16-479-2023, 2023.

Raissi, M., Perdikaris, P., and Karniadakis, G. E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational Physics, 378, 686–707, https://doi.org/10.1016/j.jcp.2018.10.045, 2019.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015, 234–241, https://doi.org/10.1007/978-3-319-24574-4_28, 2015.

Seinfeld, J. H.: Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, New York Academy of Sciences Series, John Wiley & Sons, https://doi.org/10.1063/1.882420, 2016.

Souri, A. H., Nowlan, C. R., González Abad, G., Zhu, L., Blake, D. R., Fried, A., Weinheimer, A. J., Wisthaler, A., Woo, J.-H., Zhang, Q., Chan Miller, C. E., Liu, X., and Chance, K.: An inversion of NO_x and non-methane volatile organic compound (NMVOC) emissions using satellite observations during the KORUS-AQ campaign and implications for surface ozone over East Asia, Atmos. Chem. Phys., 20, 9837–9854, https://doi.org/10.5194/acp-20-9837-2020, 2020.

Stavrakou, T., Müller, J.-F., Boersma, K. F., van der A, R. J., Kurokawa, J., Ohara, T., and Zhang, Q.: Key chemical NO_x sink uncertainties and how they influence top-down emissions of nitrogen oxides, Atmos. Chem. Phys., 13, 9057–9082, https://doi.org/10.5194/acp-13-9057-2013, 2013.

Streets, D. G., Canty, T., Carmichael, G. R., Foy, B. de, Dickerson, R. R., Duncan, B. N., Edwards, D. P., Haynes, J. A., Henze, D. K., Houyoux, M. R., Jacob, D. J., Krotkov, N. A., Lamsal, L. N., Liu, Y., Lu, Z., Martin, R. V., Pfister, G. G., Pinder, R. W., Salawitch, R. J., and Wecht, K. J.: Emissions estimation from satellite retrievals: A review of current capability, Atmos. Environ., 77, 1011–1042, https://doi.org/10.1016/j.atmosenv.2013.05.051, 2013.

van der A, R. J., Mijling, B., Ding, J., Koukouli, M. E., Liu, F., Li, Q., Mao, H., and Theys, N.: Cleaning up the air: effectiveness of air quality policy for SO_2 and NO_x emissions in China, Atmos. Chem. Phys., 17, 1775–1789, https://doi.org/10.5194/acp-17-1775-2017, 2017.

van Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., Linden, M. ter, Ludewig, A., Boersma, K. F., and Veefkind, J. P.: Sentinel-5P TROPOMI NO_2 retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data, Atmos. Meas. Tech., 15, 2037–2060, https://doi.org/10.5194/amt-15-2037-2022, 2022.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., Vries, J. de, Otter, G., Claas, J., Eskes, H. J., Haan, J. F. de, Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sens. Environ., 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

© Author(s) 2026. CC BY 4.0 License.

Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S.: CBAM: Convolutional Block Attention Module, in: Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018, 3–19, https://doi.org/10.1007/978-3-030-01234-2_1, 2018.

Zhang, Q., Zheng, Y., Tong, D., Shao, M., Wang, S., Zhang, Y., Xu, X., Wang, J., He, H., Liu, W., Ding, Y., Lei, Y., Li, J., Wang, Z., Zhang, X., Wang, Y., Cheng, J., Liu, Y., Shi, Q., Yan, L., Geng, G., Hong, C., Li, M., Liu, F., Zheng, B., Cao, J., Ding, A., Gao, J., Fu, Q., Huo, J., Liu, B., Liu, Z., Yang, F., He, K., and Hao, J.: Drivers of improved PM 2.5 air quality in China from 2013 to 2017, Proc. Natl. Acad. Sci. U.S.A., 116, 24463–24469, https://doi.org/10.1073/pnas.1907956116, 2019.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., and Zhang, Q.: Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions, Atmos. Chem. Phys., 18, 14095–14111, https://doi.org/10.5194/acp-18-14095-2018, 2018.