A gapless 0.05° daytime-hourly tropospheric NO_2 dataset (2019–2024) over key Asian hotspots reconstructed using physics-aware deep learning

Hongrui Gao[1], Qin He[1*], Kai Qin[1*], Jhoon Kim[2], Diego Loyola[3], Pravash Tiwari[1], Lingxiao Lu[1], Jason B. Cohen[1]

[1]School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou, China

[2]Department of Atmospheric Sciences, Yonsei University, Seoul, South Korea

[3]German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Oberpfaffenhofen, Germany

**Correspondence to:* heqin@cumt.edu.cn / qinkai@cumt.edu.cn

Abstract. High-frequency observations of tropospheric nitrogen dioxide (NO_2) support regional air quality assessment and analysis of its photochemical evolution. Although the Geostationary Environment Monitoring Spectrometer (GEMS) and other new-generation geostationary satellites provide hourly observations during daytime, their spatial coverage is often severely compromised by cloud obstruction and inherent retrieval limitations. Furthermore, their short observation history hinders long-term, cross-regional environmental assessments. To address this, this study presents a gapless, 0.05° daytime-hourly tropospheric NO_2 vertical column density (VCD) dataset for representative hotspots in Asia spanning 2019 to 2024, reconstructed using a physics-aware deep learning framework. The framework integrates partial convolutions to handle irregularly missing satellite observations and proposes a Physics-aware Normalization (PhysNorm) module. Specifically, the model inputs comprise the 0.05° TROPOMI tropospheric NO_2 VCD and its validity mask, together with 0.25° predictors including ERA5 meteorological variables, EAC4 total-column NO_2, land cover, elevation, and the cosine of the solar zenith angle. The supervised training target is the hourly GEMS tropospheric NO_2 VCD from 2023. PhysNorm dynamically modulates 0.05° high-resolution feature maps using 0.25° low-resolution physical features, thereby ensuring rigorous physical continuity while filling data gaps. Validation results show that the model performs exceptionally on the test set ($R^2 = 0.889$) and maintains high generalization stability in an independent validation on 2024 data, which was not used for training. Cross-validation against an independent polar-orbiting satellite instrument (GOME-2C) confirms the reliability of the dataset in reconstructing pollution hotspots and its applicability to periods without GEMS observations (2019–2022). Comparison with independent ground-based Pandora observations showed an overall correlation of $R = 0.644$ for 2019–2024. Using this reconstructed dataset, this study finely delineates the periodic diurnal characteristics of NO_2 in typical Asian regions and characterizes inter-annual variations in NO_2 concentration patterns over the past six years. This newly generated dataset provides a spatiotemporally continuous daytime record, overcoming the limitations of cloud cover and short observational histories, thereby facilitating long-term, high-resolution air quality assessments and epidemiological studies. This dataset is available at https://doi.org/10.5281/zenodo.20427767 (Gao et al., 2026).

1. Introduction

Nitrogen dioxide (NO_2), a critical tropospheric trace gas and the primary constituent of nitrogen oxides (NO_x = NO + NO_2), predominantly originates from anthropogenic fossil fuel combustion, biomass burning, and biogenic soil emissions (Beirle et al., 2011; Purchase et al., 2023; Chang et al., 2025). As a crucial precursor to tropospheric ozone (O_3) and secondary aerosols, NO_2 profoundly impacts regional and global air quality (Wang et al., 2021; Baruah et al., 2022). Furthermore, by participating in the hydroxyl radical (OH) cycle, NO_2 directly modulates the atmospheric oxidative capacity and the transformation rates of other pollutants (Monks et al., 2009). Elevated NO_2 concentrations pose significant threats to human health, causing damage to the respiratory system and increasing the risk of asthma and cardiovascular diseases (Burnett et al., 2004; Samoli et al., 2006; Kampa and Castanas, 2008). Because of its short lifetime (typically ranging from a few hours to a day), the spatial distribution of NO_2 exhibits pronounced spatial heterogeneity and temporal variability, making it a widely adopted indicator for assessing atmospheric environmental changes and evaluating the effectiveness of emission control policies (Lamsal et al., 2013a; Goldberg et al., 2019). Given these environmental and health implications, continuous and accurate monitoring of NO_2 through diverse technical approaches remains a core research imperative in the atmospheric and environmental sciences.

Traditional ground-based monitoring networks provide NO_2 observations with high temporal resolution and accuracy. However, they are unevenly distributed, and their spatial coverage remains inherently limited, rendering them inadequate for large-scale monitoring requirements (Wei et al., 2022; Ahmad et al., 2024; Gu et al., 2025). In contrast, satellite remote sensing, characterized by its extensive and continuous observational capabilities, has undergone substantial development over the past few decades, emerging as a mainstream approach for regional and global NO_2 monitoring (Safieddine et al., 2013; Liu et al., 2021c; Li and Wu, 2021). Since the Global Ozone Monitoring Experiment (GOME) first measured global NO_2 column densities in the 1990s (Burrows et al., 1999), a succession of sensors, including SCIAMACHY, OMI, and GOME-2 has been launched (Bovensmann et al., 1999; Levelt et al., 2006a, b; Callies et al., 2000). These instruments enabled the study of global distribution and long-term NO_2 trends (Richter et al., 2005; Boersma et al., 2011a; Russell et al., 2012; Duncan et al., 2016; Krotkov et al., 2016; Liu et al., 2021b; Chan et al., 2023). In October 2017, the TROPOspheric Monitoring Instrument (TROPOMI) was launched onboard the Sentinel-5 Precursor satellite (Veefkind et al., 2012a). TROPOMI improved the spatial resolution from the 13 × 24 km^2 offered by OMI to 3.5 × 5.5 km^2, while maintaining daily global coverage. However, LEO satellites are limited by their orbital characteristics and can only provide daily observations. This sampling frequency is insufficient to resolve the diurnal dynamics of NO_2 driven by morning and evening traffic peaks, industrial activity cycles, and photochemical processes (He et al., 2025). To address this, Geostationary Earth Orbit (GEO) satellites have been deployed to increase temporal resolution. In February 2020, the Geostationary Environment Monitoring Spectrometer (GEMS) onboard the GEO-KOMPSAT-2B (GK-2B) satellite became operational (Kim et al., 2020a), providing 6–10 hourly observations per day of air quality over East and Southeast Asia (Choi et al., 2024; He et al., 2024a). Together with the Tropospheric Emissions: Monitoring of Pollution (TEMPO launched on 7 April 2023) over North America and Copernicus Sentinel-4 over Europe (launched on 1 July 2025 onboard EUMETSAT's Meteosat Third Generation Sounder (MTG-S1)), GEMS forms part of the global geostationary constellation for monitoring air quality, promising near-continuous, high-frequency observations over the major emission regions (Ingmann et al., 2012; Zoogman et al., 2017).

Nevertheless, the ideal scenario of global hourly monitoring remains a distant prospect. First, regardless of whether considering the pioneering GEMS or the subsequent TEMPO and Sentinel-4, the available temporal record of these instruments is relatively short, constraining the application of high-temporal-resolution data to longer-term atmospheric environment studies, policy assessments, and health impact analyses. Second, data completeness is a major challenge. The daily frequency of valid

observations varies due to seasonal changes in solar illumination and daylight duration. In addition, due to cloud contamination and high aerosol loading, the number of valid pixels in retrieved products can be significantly reduced (Boersma et al., 2004; Belle and Liu, 2016; Lee et al., 2025), making spatiotemporal gap-filling an essential step to enhance data completeness and usability. These considerations motivate a compelling research topic: based on the existing LEO satellite data, combined with other auxiliary information, to construct a temporally extensible and spatially near-seamless hourly tropospheric NO_2 dataset in time. To address this question, a method based on TROPOMI data and a series of auxiliary data to simulate the observation capability of GEMS is proposed in this study. It attempts to extend the hourly data to the time periods when GEMS does not have observations, even before launch, so as to provide more complete data support for scientific research in Asia, and contribute experience for future multi-sensor collaboration.

There are already a considerable number of studies related to the construction of long-term, high-coverage remote sensing datasets. Early spatiotemporal gap-filling approaches for individual satellite sensors were rooted primarily in geostatistical principles, such as Kriging interpolation and its derivatives (Hengl et al., 2007; Zammit-Mangion et al., 2018; Chen et al., 2020). Data Interpolating Empirical Orthogonal Functions (DINEOF), which reconstructs missing information by extracting the dominant spatiotemporal modes of the dataset, has been applied not only in oceanography (Alvera-Azcárate et al., 2005), but also to atmospheric products (Jiang et al., 2022; Qin et al., 2024). Supported by auxiliary data, Land Use Regression (LUR) models and Geographically Weighted Regression (GWR) have been extensively utilized in urban-scale air quality assessments by establishing statistical relationships between pollutant concentrations and spatial variables such as land-use types and population density (Hoek et al., 2008; Beelen et al., 2013; Meng et al., 2015; Qin et al., 2017). However, these traditional methods exhibit limitations when addressing complex, multi-source, heterogeneous data fusion tasks. In recent years, machine learning algorithms have been widely adopted in the fusion and prediction of remote sensing data (Peng et al., 2024), including Random Forest (RF), Gradient Boosting Decision Trees (GBDT), and Light Gradient Boosting Machine (LightGBM) (He et al., 2020; Kang et al., 2021; Fu et al., 2023; Ma et al., 2023; Shao et al., 2023; Wang et al., 2023). Also, within the atmospheric remote sensing domain, deep learning has been successfully deployed across various tasks, including the improvement of satellite retrieval algorithms (Loyola, 2006; Xu et al., 2017; She et al., 2020; Rao et al., 2022; Fan et al., 2023; She et al., 2024; Wei et al., 2025), the reconstruction of data gaps caused by cloud cover and other factors (Yu and Liu, 2021; Li et al., 2023; He et al., 2024b; Qu et al., 2025; Zhang et al., 2025a; Wang, 2026), and the establishment of more accurate mapping relationships between satellite column observations and surface-level concentrations (Sun et al., 2019; Scheibenreif et al., 2022; Xing et al., 2024; Li et al., 2025; Handschuh et al., 2026). Despite these advances, existing deep learning models often struggle to effectively manage non-random missing patterns in satellite observations. Simplistic interpolation or masking strategies ignore the spatial continuity between valid observational pixels and their surrounding missing regions. Moreover, conventional deep learning architectures typically integrate meteorological variables or physical constraints through straightforward channel-level concatenation, lacking an in-depth characterization of the dynamic modulation between atmospheric physical processes and pollutant distributions.

To address this, this study proposes PhysNorm-Net, a physics-aware network to reconstruct a gapless 0.05° daytime-hourly tropospheric NO_2 dataset covering the period from 1 January 2019 to 31 December 2024 over four Asian hotspot domains: the North China Plain (NP; 34–44° N, 112–122° E), the Sichuan Basin and its surroundings (CC; 26–36° N, 100–110° E), Korea and Japan (KJ; 32–42° N, 126–136° E), and northern India (NI; 20–30° N, 75–85° E). This architecture is designed to fuse multi-source heterogeneous satellite data with physical and meteorological priors, achieving high-precision, spatiotemporally continuous tropospheric NO_2 vertical column density (VCD) reconstruction. The value and major contributions of this dataset are summarized as follows: 1) High spatiotemporal completeness and resolution: It provides seamless hourly NO_2 records during daytime at 0.05° resolution covering both pre- and post-GEMS periods, effectively mitigating data gaps caused by cloud contamination and

observation limitations. 2) Physical consistency: Instead of simple statistical interpolation, the dataset generation is strictly constrained by meteorological and chemical priors, ensuring that the reconstructed concentration fields adhere to atmospheric physical continuity. Independent validation confirms its robust inter-annual stability. 3) Enhanced utility for long-term environmental analysis: The dataset enables fine-grained delineation of previously obscured periodic diurnal characteristics and supports the characterization of inter-annual variations in NO_2 concentration patterns over the past six years.

2. Study area and data

2.1 Study area

This study selects four typical regions encompassing diverse climatic zones, emission characteristics, and topographical complexities, as detailed in Table 1.

Table 1: Four representative study regions and their geographical extents selected in this study.

Region	Abbreviation	Latitude	Longitude
North China Plain	NP	34–44°N	112–122°E
Sichuan Basin and surroundings	CC	26–36°N	100–110°E
Korea and Japan	KJ	32–42°N	126–136°E
Northern India	NI	20–30°N	75–85°E

The North China Plain (NP) is one of China's largest plains, characterized by intensive heavy industrial and traffic emissions. The Sichuan Basin and its surrounding areas, encompassing major urban centers such as Chengdu and Chongqing (CC), are defined by an enclosed topographic configuration that severely restricts atmospheric ventilation and promotes pollutant accumulation, thereby producing a topographically driven atmospheric environment fundamentally distinct from that of the other study regions. The Korea and Japan (KJ) region spans the Korean Peninsula and the Japanese archipelago, where NO_2 distributions are significantly influenced by transboundary transport and maritime climatic conditions. Northern India (NI), situated in the Indo-Gangetic Plain, experiences frequent winter haze episodes and agricultural biomass burning. Given the high spatial heterogeneity of NO_2 distributions, the selection of these four representative regions maximizes information density and gradient diversity within the training sample.

2.2 GEMS tropospheric NO_2 data

The Geostationary Environment Monitoring Spectrometer (GEMS), an ultraviolet-visible (UV-Vis) imaging spectrometer onboard the GEO-KOMPSAT-2B (GK-2B) satellite, was launched in February 2020 and became operationally active in November 2022 following in-orbit testing. GEMS provides spectral measurements from 300 to 500 nm with a spectral resolution of approximately 0.6 nm. It achieves a spatial resolution of 3.5 km × 7 km (over Seoul at nadir) and a temporal resolution of one hour. The observational domain of GEMS covers 5°S–45°N and 75°E–145°E, delivering 6–10 measurements per day over the Asian region (Kim et al., 2020b).

This study utilizes the GEMS Level 2 NO_2 tropospheric VCD v3.0 product. To ensure data quality, quality control was applied, retaining only pixels where the VCD is greater than zero, the cloud fraction is less than 0.3, and the solar zenith angle is less than 70°. Thus, negative and zero GEMS retrievals were treated as missing values rather than being replaced with zero. This criterion was adopted because the supervised target was defined as a nonnegative tropospheric NO_2 column field and because the subsequent logarithmic transformation requires nonnegative input values. Pixels removed by this criterion were excluded from the training

loss and did not act as zero-valued training labels. The quality-controlled GEMS data were then regridded to a uniform 0.05° × 0.05° latitude-longitude grid using bilinear interpolation to facilitate spatial alignment with other datasets. GEMS data from January 1, 2023, to December 31, 2023, were used as the training labels for the model, with data of 2024 reserved for additional validation. Furthermore, all the GEMS data mentioned subsequently refer to the nearest whole-hour time point of their average transit time in that area.

2.3 TROPOMI tropospheric NO_2 data

The TROPOspheric Monitoring Instrument (TROPOMI) is an imaging spectrometer onboard the European Space Agency's (ESA) Sentinel-5 Precursor (S5P) satellite, which was launched on October 13, 2017 (Veefkind et al., 2012b; Ialongo et al., 2020). S5P operates in a sun-synchronous polar orbit at an altitude of approximately 824 km, achieving daily global coverage with a local equator crossing time of around 13:30. TROPOMI employs a push-broom imaging technique with a wide swath of 2600 km. Its spectral range covers the ultraviolet, visible, near-infrared, and shortwave infrared bands (270–2385 nm), with the NO_2 retrieval primarily relying on the visible band (405–465 nm) (Geffen et al., 2019) and the cloud retrieval relying on UV/visible/NIR bands (Loyola et al., 2018). The instrument's spatial resolution was enhanced from 7 × 3.5 km^2 to 5.5 × 3.5 km^2 (across-track × along-track) on August 6, 2019 (Liu et al., 2021a), making it the highest-resolution operational sensor for NO_2 monitoring on a polar-orbiting platform.

The TROPOMI NO_2 product has been validated against ground-based Pandora spectrometers, Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) observations, and aircraft campaigns, demonstrating high accuracy (Griffin et al., 2019; Ialongo et al., 2020; Verhoelst et al., 2021). This study uses the TROPOMI Level 2 tropospheric NO_2 VCD products from 2019 to 2024, with data from 2023 employed for model training. For TROPOMI data, a quality assurance flag (qa_value) greater than 0.75 was used as the filtering criterion. The filtered data were subsequently regridded to the same 0.05° × 0.05° grid using the HARP toolkit (https://github.com/stcorp/harp). In addition to the qa_value threshold, TROPOMI retrievals with tropospheric NO_2 VCD values less than or equal to zero were treated as missing. This filtering maintained consistency with the nonnegative model-input domain and the logarithmic normalization used in the framework.

2.4 GOME-2 tropospheric NO_2 data

GOME-2 (Global Ozone Monitoring Experiment-2) is an ultraviolet-visible imaging spectrometer onboard the MetOp series of satellites operated by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). The GOME-2 series consists of three satellites: MetOp-A (launched in October 2006), MetOp-B (launched in September 2012), and MetOp-C (launched in November 2018). GOME-2 has a local overpass time of approximately 09:30 and provides daily global coverage (Munro et al., 2016).

The GOME-2 NO_2 products (Valks et al., 2011) have undergone long-term validation and improvement. Comparison studies with ground-based MAX-DOAS, Pandora observations, and other satellite products have demonstrated good consistency at both global and regional scales (Pinardi et al., 2010; Liu et al., 2019). This study utilizes the GOME-2C Level 2 daily tropospheric NO_2 product provided by EUMETSAT AC-SAF (Hassinen et al., 2016). The nominal native ground-pixel size of GOME-2C is 80 km × 40 km for forward-scan observations and 240 km × 40 km for back-scan observations (Chan et al., 2023). The data were regridded to a 0.25° × 0.25° resolution using the HARP toolkit, and pixels with a cloud fraction greater than 0.3 were filtered out. For consistency with the GEMS and TROPOMI processing, GOME-2 retrievals with tropospheric NO_2 VCD values less than or equal to zero were also excluded. This product is primarily used for the cross-comparison and independent validation of the model's prediction results in subsequent chapters. Given the significant difference between GOME-2's overpass time and those of

TROPOMI and GEMS, its data offer an independent reference for assessing the diurnal variation of NO_2 and the model's temporal generalization capability.

2.5 Pandora ground-based NO_2 observations

Pandora is a ground-based ultraviolet–visible spectrometer system that measures atmospheric trace gases, including NO_2, using direct-sun and scattered-sky observations (Herman et al., 2009). The Pandonia Global Network (PGN) coordinates the processing and distribution of Pandora measurements to support air-quality monitoring and satellite validation (Verhoelst et al., 2021; Park et al., 2026). Its high temporal resolution provides a ground-based reference for evaluating the temporal variability of satellite observations and reconstructed products. This study uses the rnvh3p1-8 tropospheric NO_2 column product derived from sky-scanning observations. Observations available during 2019–2024 were obtained as an external reference independent of model training. Usable observations were identified at 16 stations: Beijing-RADI in NP; Busan, Daegu, Fukuoka, Incheon-ESC, Jeonju, Kobe, Seosan, Seoul, Seoul-KU, Seoul-SNU, Suwon-USW, Ulsan, and Yongin in KJ; and Haldwani-ARIES and Nainital-ARIES in NI. No eligible station was identified in CC for this evaluation. Observational coverage varied among stations and years.

The evaluation used finite positive tropospheric NO_2 columns with quality flags 10 or 11, corresponding to preliminary high- and medium-quality estimates, respectively, that had not undergone formal quality assurance. Low-quality observations (flag 12) were excluded, and no additional upper concentration threshold was applied. Column amounts were converted from mol/m^2 to $Pmolec/cm^2$. At Seoul-SNU, Pandora 149 was used preferentially, and Pandora 163 supplied only hours without valid observations from Pandora 149, avoiding duplicate station-hour samples. The rnvh3 tropospheric column represents approximately the surface-to-4 km column; differences in vertical sensitivity and spatial sampling relative to satellite retrievals should therefore be considered when interpreting the comparisons.

2.6 ERA5 meteorological reanalysis data

Meteorological conditions play a decisive role in influencing the spatiotemporal heterogeneity of tropospheric NO_2 through physical transport and chemical modulation processes. This study uses the fifth generation of the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis, ERA5 (Hersbach et al., 2020). This dataset is produced using the Integrated Forecasting System (IFS) Cycle 41r2, which assimilates a vast range of global observations through a 4D-Var data assimilation system. It provides data at a spatial resolution of 0.25° × 0.25° and an hourly temporal resolution. ERA5 has been widely adopted in atmospheric chemistry research, serving as a meteorological driver for chemical transport models (CTMs), as auxiliary input for satellite retrievals, and as feature variables in machine learning models (Zhang et al., 2025b; Mols et al., 2026).

Both single-level and pressure-level variables from ERA5 were utilized in this study. The single-level variables include boundary layer height (blh), surface pressure (sp), and total cloud cover (tcc). Boundary layer height is a key physical parameter governing near-surface NO_2 concentrations (Stull, 2012; von Engeln and Teixeira, 2013) and indirectly influences satellite-retrieved tropospheric NO_2 VCDs through the lower-tropospheric contribution to the column. Surface pressure not only indicates the evolution of weather systems but also directly affects the normalization of a priori profiles and the weighting of lower-layer contributions in the calculation of air mass factor (AMF) (Boersma et al., 2007). Total cloud cover modulates the photolysis rate of NO_2 and photochemical equilibrium by altering shortwave radiation reaching the lower atmosphere (van Weele and Duynkerke, 1993), while also being closely linked to precipitation and local thermal circulations (Sfica et al., 2021). For multi-level variables, specific humidity (q), temperature (T), and horizontal wind components (u, v) were extracted at four levels: 500 hPa, 850 hPa, 925 hPa, and 1000 hPa. Lower-level variables can capture thermodynamic and dynamic conditions near NO_2 emission sources, whereas the 500 hPa level represents the large-scale circulation background and vertical transport potential in the mid-troposphere (Wang

et al., 2026). Temperature and specific humidity jointly determine the thermodynamic conditions and oxidant environment for atmospheric chemical reactions (Mentel et al., 1996; Wang et al., 2020). The wind components provide vector information on the horizontal transport of pollutants. All ERA5 variables used are hourly data at a 0.25° resolution, serving as part of the low-resolution channels input to the model.

2.7 EAC4 reanalysis data

The ECMWF Atmospheric Composition Reanalysis 4 (EAC4) is a core data product of the Copernicus Atmosphere Monitoring Service (CAMS), providing a spatiotemporally consistent reanalysis of global atmospheric composition from 2003 to the present (Inness et al., 2019). EAC4 is based on the atmospheric chemistry module of the ECMWF's IFS and assimilates multi-source satellite observations with model forecasts using a 4D-Var scheme to produce chemically consistent and spatiotemporally complete atmospheric composition fields (Flemming et al., 2017).

This study employs the total column NO_2 (tcno2) variable from EAC4 as an input feature. Because tcno2 includes both tropospheric and stratospheric NO_2, it is not directly equivalent to the GEMS tropospheric VCD used as the supervised target. EAC4 is therefore used as a spatially complete, coarse-resolution chemical background prior rather than a target-equivalent representation of tropospheric NO_2. Although it differs from the satellite-observed tropospheric NO_2 concentration, this product provides a spatiotemporally complete a priori estimate of the NO_2 field, thereby enriching valuable physical prior information available to the model. The original EAC4 product has a temporal resolution of 3 hours and a spatial resolution of 0.75° × 0.75°. To align with the overall framework, EAC4 data were resampled to a 1-hour temporal resolution and a 0.25° × 0.25° spatial resolution using linear interpolation, making it consistent with the ERA5 reanalysis data.

2.8 MCD12C1 land cover data

Land use/land cover (LULC) type is a critical factor influencing the spatial distribution of NO_2. Different land use types correspond to distinct emission source characteristics and intensities: urban and built-up areas concentrate emissions from traffic, industry, and residential activities; agricultural lands are associated with emissions from farm machinery and soil nitrogen cycles; and forests and grasslands are primarily influenced by natural emissions and biomass burning (Lamsal et al., 2013b; Song et al., 2019). Furthermore, land use types indirectly affect the accuracy of satellite retrievals and local meteorological conditions by influencing surface albedo and roughness (Auer, 1978; Noguchi et al., 2014). Consequently, incorporating land use information as a spatial prior helps the model to better understand the heterogeneity of NO_2 spatial patterns.

This study uses the annual MODIS Version 6.1 Climate Modeling Grid (CMG) land cover product, MCD12C1, for the years 2019–2023 (Friedl and Sulla-Menashe, 2022). The Leaf Area Index (LAI)-based classification scheme (LCType3) was selected. Given the mismatch in spatial resolution between MCD12C1 (0.05°) and physical driver fields (0.25°), the land cover data were spatially aggregated to the 0.25° grid by averaging.

2.9 ETOPO 2022 elevation data

Topography is another essential spatial prior variable. Elevation is directly related to atmospheric pressure and air density (Zhou et al., 2009), while topographical variations influence local circulation patterns (e.g., valley and slope winds) and pollutant transport pathways (Gohm et al., 2009). Topography is also closely linked to the spatial distribution of emission sources, as human activities and NO_2 emissions are predominantly concentrated in low-lying plains and basins, whereas emissions are sparse in high-altitude mountains and plateaus (Lamsal et al., 2013b; Yuan et al., 2025). Additionally, terrain elevation affects the calculation of the AMF in satellite retrievals, with higher uncertainty typically observed over high-altitude areas (Boersma et al., 2011b).

Therefore, integrating topographical information into the model enhances its predictive capability for NO_2 distribution, especially in regions with complex terrain.

This study utilizes the Earth TOPOgraphy (ETOPO) 2022 global relief model (NOAA National Centers for Environmental Information, 2022). As the latest version in the ETOPO series, it integrates multi-source topographic and bathymetric data, including satellite altimetry, shipborne sonar, airborne gravity, and terrestrial digital elevation models, to provide a high-resolution global relief dataset for land and ocean (MacFerrin et al., 2025). This study utilizes the 60-arc-second resolution version, which was resampled to a 0.25° × 0.25° resolution via spatial aggregation (averaging) to align with the grid of the other physical variables.

3. Methodology

3.1 Multi-source data integration framework

This study integrates multi-source datasets with heterogeneous origins, spatiotemporal resolutions, and physical units, constructing an input tensor system comprising two 0.05° high-resolution input channels (the TROPOMI tropospheric NO_2 VCD and its validity mask), 33 low-resolution predictor channels at 0.25° resolution, and one supervised target channel consisting of the quality-controlled 0.05° hourly GEMS tropospheric NO_2 VCD. Table 2 systematically summarizes the source, original resolution, physical significance, and normalization strategy of all input variables. All reconstructed hourly fields correspond to daytime GEMS observation hours.

Table 2: Summary of input variables for PhysNorm-Net.

Source	Variable Name	Abbreviation	Spatial Resolution	Temporal Resolution	Channels	Normalization Method
TROPOMI	Tropospheric NO_2 VCD	TROPOMI NO_2	0.05°	daily	1	Log+Z-Score
	mask	mask	0.05°	daily	1	/
GEMS	Tropospheric NO_2 VCD	GEMS NO_2	0.05°	hourly	1 (label)	Log+Z-Score
EAC4	NO_2 Total Column	tcno2	0.25°	hourly	1	Z-Score
ERA5	Surface Pressure	sp	0.25°	hourly	1	Z-Score
	Total Cloud Cover	tcc			1	Z-Score
	Boundary Layer Height	blh			1	Z-Score
	Specific Humidity	q			4	Z-Score
	Temperature	T			4	Z-Score
	U-component of wind	u			4	Z-Score

	V-component of wind	v			4	Z-Score
ETOPO2022	Digital Elevation Model	DEM	0.25°	/	1	Min-Max
MCD12C1	Land Cover Type	/	0.25°	yearly	11	Convert to decimals
Computed	Cosine of Solar Zenith Angle	$\cos(\theta_{SZA})$	0.25°	hourly	1	/

Note: The 4 channels for ERA5 multi-level variables correspond to 1000, 925, 850, and 500 hPa pressure levels. The 11 channels for MCD12C1 represent the 11 land cover types.

To ensure numerical stability during neural network training and eliminate the influence of dimensional discrepancies on gradient propagation, differentiated normalization strategies were tailored to the characteristics of each variable type. Prior to normalization, the 2023 dataset was randomly partitioned by date into training (256 days), validation (54 days), and test (55 days) sets. This split was used for model development and for evaluating performance on held-out dates within 2023. Because randomly selected dates may have weather conditions similar to those of nearby training dates, the 2023 test set is not interpreted as a strictly temporally independent evaluation. The complete 2024 dataset was retained outside the 2023 data split and used for independent-year validation.

The probability distribution of tropospheric NO_2 column densities exhibits a pronounced positive skew, with several orders of magnitude separating high-value tails (corresponding to pollution events) from background concentrations. A direct linear scaling would compress the information content of high-concentration pixels while disproportionately amplifying minor variations in low-concentration regions, thereby hindering the model's ability to learn the spatial features of both pollution hotspots and background areas. To address this, TROPOMI and GEMS tropospheric NO_2 VCDs were first scaled by a factor of 10^{-15} (converting to units of Pmolec/cm^2), followed by logarithmic transformation and Z-score normalization:

$$x_{\log} = \ln\left(x_{\text{trop}} + \epsilon\right) \tag{1}$$

$$x_{\text{norm}} = \frac{x_{\log} - \mu_{\log}}{\sigma_{\log}} \tag{2}$$

Here, ϵ is a small constant to prevent undefined logarithms for zero values, while $\mu_{\log}$ and $\sigma_{\log}$ represent the global mean and standard deviation of the log-transformed training data, respectively. The logarithm transform effectively compresses the dynamic range of the data, making its distribution more Gaussian-like while preserving the relative relationships between concentration gradients. The subsequent Z-score standardization centers the data and scales it to a standard deviation of one, aligning it with the standard input assumptions of deep neural networks (LeCun et al., 2012).

For ERA5 and EAC4 reanalysis variables, Z-score normalization was uniformly applied:

$$x' = \frac{x - \mu}{\sigma} \tag{3}$$

Where μ and σ are the spatiotemporal statistics of each variable computed over the training set. It is emphasized that all normalization statistics are derived exclusively from the training set and subsequently applied to the validation and test sets to strictly prevent data leakage.

Since DEM data possess well-defined physical boundaries (from sea level to the highest peak) and its absolute values carry physical meaning, a min-max normalization was applied to preserve this characteristic while mapping values to the [0, 1] interval:

$$z' = \frac{z - z_{\min}}{z_{\max} - z_{\min}} \tag{4}$$

Where $z_{\max}$ and $z_{\min}$ are the global minimum and maximum DEM values.

Land use data from MCD12C1 are provided as area fractions (original values as percentages from 0–100); these were directly converted to decimals in the [0, 1] range by dividing by 100. Normalization of both DEM and land use data was performed prior to spatial resampling.

For temporal encoding, while spatiotemporal prediction tasks typically employ periodic functions (e.g., sine/cosine transforms) to explicitly embed hour, day, or year cycles, such purely mathematical encodings may lack a direct connection to atmospheric physical processes. Therefore, this study proposes using the cosine of the solar zenith angle ($\cos(\theta_{\mathrm{SZA}})$) as a substitute for conventional time encoding. For a given UTC time, regions at different longitudes experience different local solar times and corresponding SZAs. The SZA implicitly encodes the longitude-time coupling, enabling the model to understand local illumination conditions without needing to learn complex time zone conversion logic. Furthermore, the SZA reflects both diurnal and seasonal variations (depending on solar declination), allowing a single variable to represent periodicity across multiple time scales and reducing input feature redundancy.

The calculation process is as follows: based on the UTC timestamp and the central coordinates of the region, the solar elevation angle (α) is first computed using a solar position algorithm (Reda and Andreas, 2004), from which the solar zenith angle (θ_{SZA}) is derived:

$$\theta_{\mathrm{SZA}} = 90° - \alpha \tag{5}$$

To constrain input values within the numerical range preferred by neural networks, the cosine is taken:

$$f_{\mathrm{SZA}} = \cos(\theta_{\mathrm{SZA}}) \tag{6}$$

This value approaches 0 near sunrise/sunset and approaches 1 at solar noon in tropical regions (or less than 1 at mid-to-high latitudes), offering intuitive physical interpretation. Given that the spatial resolution of data such as EAC4 is 0.25°, corresponding to a grid scale of approximately 25–30 km, the spatial variation of the SZA within this scale is negligible (< 0.1°). Therefore, f_{SZA} is treated as a spatially uniform field and appended to the low-resolution tensor.

3.2 Overall architecture of PhysNorm-Net

Let the high-resolution TROPOMI NO_2 observations be denoted as $X_{\mathrm{obs}} \in \mathbb{R}^{H\times W}$. Due to cloud cover, aerosol contamination, and sensor swath gaps, the observation field is spatially sparse. Therefore, a binary mask $M \in \{0,1\}^{H\times W}$ is introduced to mark the locations of valid pixels (1 for valid, 0 for missing). Concurrently, to incorporate the dynamic constraints of atmospheric transport and dispersion, a set of spatiotemporally aligned, low-resolution physical and topography data is utilized, denoted as $P \in \mathbb{R}^{h\times w\times C}$, where $h < H$, $w < W$, and C is the number of physical variable channels. The objective of this study is to learn a nonlinear mapping function $\mathcal{F}$ that infers the complete, high-spatiotemporal-resolution NO_2 field $\hat{Y}$ by fusing the sparse observations with the dense physical prior:

$$\hat{Y} = \mathcal{F}(\mathrm{X}_{\mathrm{obs}}, P, M) \tag{7}$$

The challenge inherent in this problem lies in the significant modality gap between X_{obs} and P: the former contains high-frequency spatial textures but is incomplete, while the latter is physically continuous but spatially coarse. To address this multi-modal fusion challenge, the PhysNorm-Net is designed. As illustrated in Figure 1, the network's main body inherits the classic U-Net architecture (Ronneberger et al., 2015), comprising a contracting path (encoder) and an expansive path (decoder), with skip connections to preserve multi-scale features. However, two critical asymmetric modifications are introduced to the standard U-Net to accommodate irregularly missing data and to integrate physical constraints:

The Main Observation Stream, which forms the U-Net backbone, is responsible for extracting multi-scale features from the spatially incomplete TROPOMI field. Unlike traditional Convolutional Neural Networks (CNNs) that use standard convolutional layers, this stream is constructed from a stack of Partial Convolution layers. This enables the network to adaptively handle data missing in arbitrary shapes, progressively in-painting the feature maps during the encoding process.

The Physics-Conditional Stream operates as a parallel feature extractor, deriving latent physical information from the low-resolution data. This stream employs a deep residual network (ResNet) structure to prevent gradient vanishing during deep feature extraction and to ensure the effective propagation of physical information.

During the decoding (upsampling) phase, features from the physics stream are not simply concatenated with the observation features. Instead, they are used to dynamically modulate the statistical distribution of the observation features through a novel Physics-Aware Normalization (PhysNorm) module. This conditional affine transformation fuses the two streams of information to reconstruct the high-resolution tropospheric NO_2 field. This design ensures that the generated pollutant concentration field not only retains the fine details from the satellite observations but also adheres macroscopically to the prevailing meteorological dispersion patterns.

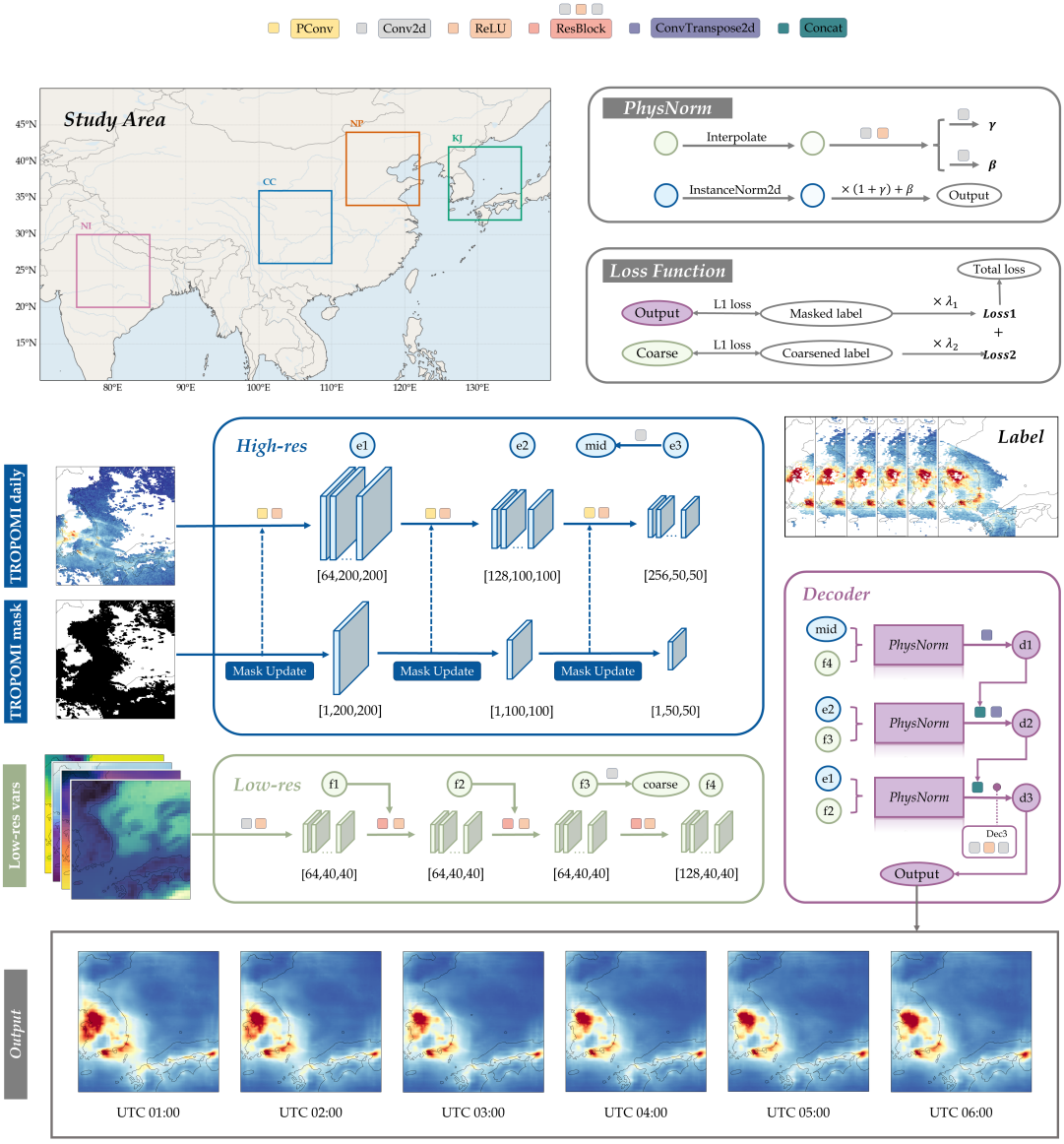


Figure 1: Architecture diagram of PhysNorm-Net.

3.3 Observation encoding via Partial Convolution

Spatial gaps in satellite remote sensing imagery exhibit high irregularity, with patterns, areas, and locations varying across samples. Standard convolution operations implicitly assume all input pixels are valid and incorporate padded values (e.g., zeros or mean-filled values) in missing regions into feature calculations. This leads to the contamination of features from valid areas by missing areas, an error that propagates and diffuses throughout the feature map as network depth increases. To resolve this issue, this study adopts the Partial Convolution (PConv) layer, proposed by Liu et al. (2018), as the fundamental operator in the encoder. The core mechanism of PConv is to perform convolution exclusively on the valid pixels defined by a mask and to re-weight the result based on the proportion of valid pixels.

Assume the input feature at layer l is $X^{(l)}$, the corresponding binary mask is $M^{(l)}$, the convolution kernel is W, and the bias is b. The partial convolution output at position (i,j) is computed as:

$$x'_{(i,j)} = \begin{cases} W^T\left(X^{(l)} \odot M^{(l)}\right)_{(i,j)} \cdot \dfrac{\mathrm{sum}(1)}{\mathrm{sum}(M^{(l)})_{(i,j)}} + b, & \text{if } \mathrm{sum}\left(M^{(l)}\right)_{(i,j)} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

Here, $\odot$ denotes element-wise multiplication, $sum(1)$ is the total number of pixels in the convolutional window, and $sum\left(M^{(l)}\right)_{(i,j)}$ is the number of valid pixels within that window. The scaling factor $\frac{\mathrm{sum}(1)}{\mathrm{sum}\left(M^{(l)}\right)_{(i,j)}}$ automatically corrects for the reduction in signal strength caused by the diminished amount of valid information. Simultaneously, the mask is updated after each layer: a position in the next layer's mask is marked as valid if at least one valid pixel existed within its receptive field in the current layer. This hierarchical update mechanism enables valid information to propagate layer-by-layer toward missing regions as network depth increases, providing the decoder with structurally complete semantic features.

Specifically, the high-resolution encoder employs a three-stage downsampling architecture. The first stage uses a 7×7 large-kernel convolution (64 channels) to rapidly expand the receptive field while preserving the original resolution, which is suitable for capturing medium-scale NO_2 spatial structures. The subsequent two stages use 3×3 convolutions (128 and 256 channels, respectively) to perform 2x downsampling. The feature maps and updated masks at each stage are preserved for use in the decoder via skip connections.

3.4 Residual encoding of physical priors

The physical prior encoding module is tasked with extracting a hierarchical feature representation of the atmospheric state from the low-resolution tensor $P \in \mathbb{R}^{h\times w\times C}$. The design objectives for this module are: (a) to capture the nonlinear mapping relationships between meteorological fields, surface characteristics, and NO_2 distribution; and (b) to generate multi-level physical features for fusion at different resolution scales in the decoder. Traditional deep convolutional networks, after alternating between multiple layers of linear convolution and nonlinear activation, are prone to gradient degradation, which can lead to the over-abstraction or even loss of spatially continuous information present in the original meteorological data. Therefore, this module adopts a residual network (ResNet) architecture, consisting of an initial convolutional layer followed by three residual blocks. The introduction of residual connections alleviates gradient vanishing in deep networks (He et al., 2015) while facilitating the learning of incremental functions near identity mappings from input to output. This network produces three levels of physical feature maps $\left\{F_{\mathrm{phy}}^{(1)}, F_{\mathrm{phy}}^{(2)}, F_{\mathrm{phy}}^{(3)}\right\}$ (with 64, 64, and 128 channels, corresponding to f2, f3, f4 in Figure 1), which encode hierarchical information from raw meteorological variables to abstract physical patterns.

Specifically, the low-resolution input is first mapped to a latent feature space via an initial convolutional layer, followed by deep feature extraction through a stack of residual blocks. If the input feature to the k-th residual block is $F_{\text{phy}}^{(k-1)}$, the feature propagation process can be described as:

$$F_{\text{phy}}^{(k)} = \text{ReLU}\left(F_{\text{phy}}^{(k-1)} + \mathcal{R}_k\left(F_{\text{phy}}^{(k-1)}\right)\right) \tag{9}$$

where $\mathcal{R}_k$ represents the residual mapping composed of two consecutive convolutional layers with embedded nonlinear activation functions. The identity mapping term $F_{\text{phy}}^{(k-1)}$ allows fundamental physical information provided by the low-resolution branch to be directly transmitted to deeper network layers. The residual term $\mathcal{R}_k(\cdot)$ is consequently forced to learn only refinements (or perturbations). The physical prior encoder does not merely output a single global feature vector but extracts multi-scale hierarchical features under different receptive fields. These features are then routed to and precisely aligned with the PhysNorm modules at different depths of the high-resolution decoder.

Additionally, this module includes an auxiliary prediction head (a 1×1 convolution layer) that directly predicts a coarse-resolution NO_2 field $\hat{Y}_{\text{coarse}}$ from the intermediate feature layer $F_{\text{phys}}^{(2)}$. This auxiliary output provides an additional supervisory signal during training (detailed in Section 3.6), ensuring that the physical features possess direct predictive capability for NO_2 concentrations.

3.5 Physics-Aware Normalization (PhysNorm) module and multi-scale decoder

A core challenge of this research is how to effectively fuse low-resolution physical information during the decoder's process of recovering high-resolution details. Traditional fusion methods, such as concatenation or addition, treat physical information as an additional feature channel, relying on the network to implicitly learn its relationship with satellite observations. This often leads to the deep learning network prioritizing the observational data and ignoring the physics stream, causing the model to degrade into image interpolation. Inspired by Spatially-Adaptive (De)normalization (SPADE) from the image generation domain (Park et al., 2019), this study proposes the Physics-Aware Normalization module. This module treats the meteorological field as an environmental context that dynamically modulates the normalized statistics of the satellite feature map via an affine transformation. The workflow of PhysNorm involves three main steps:

Destylization: First, the input satellite feature map (F_{sat}, corresponding to mid, e2, e1 in Figure 1) undergoes parameter-free Instance Normalization (IN). This step removes biases in absolute concentration that are specific to a particular time and region by subtracting the mean and dividing by the standard deviation, thereby retaining the pure relative structural information:

$$\hat{F}_{\text{sat}} = \frac{F_{\text{sat}} - \mu(F_{\text{sat}})}{\sigma(F_{\text{sat}})} \tag{10}$$

Parameter Mapping: Using the multi-scale features (F_{phy}) extracted by the physics encoder, two separate shallow convolutional networks, ϕ_γ and ϕ_β, learn the scaling (γ) and shifting (β) coefficients corresponding to the specific meteorological conditions:

$$\gamma = \phi_\gamma\left(F_{\text{phy}}\right), \beta = \phi_\beta\left(F_{\text{phy}}\right) \tag{11}$$

Here, γ and β are tensors of the same spatial dimensions as $\hat{F}_{\text{sat}}$.

Physics Injection: The learned parameters are then used to modulate the normalized features:

$$F_{\text{out}} = \hat{F}_{\text{sat}} \cdot (1 + \gamma) + \beta \tag{12}$$

The scaling factor γ and shifting factor β control the local concentration contrast and background level, respectively. For instance, under conditions of strong winds and high dispersion, where the pollutant distribution tends to be more uniform, the

model would predict smaller γ values to suppress local peaks. In areas with a low boundary layer height or high emissions, the model would generate positive β values to holistically elevate the predicted concentrations in that region. Through PhysNorm, the model can be compelled to rely on the physical state (low-resolution input) to determine the numerical range and distribution intensity of the final output, fundamentally resolving the information imbalance issue in modality fusion.

The decoder adopts a symmetric three-stage upsampling architecture, progressively recovering spatial resolution via transposed convolutions. PhysNorm is deployed at each decoder stage, modulated by the deep, intermediate, and shallow features from the physical prior module, respectively. This hierarchical design allows physical information of varying abstraction levels to act on corresponding spatial scales: deep features modulate weather-scale patterns, while shallow features modulate variations at the local scale.

In this architecture, the PhysNorm module modulates encoder features prior to concatenation. The final output layer is a lightweight convolutional head (two 3×3 convolutional layers, 128→32→1 channels), which maps the fused features to a single-channel NO_2 prediction field.

3.6 Dual-task loss function

To maintain physical consistency while ensuring high-resolution reconstruction accuracy, this study proposes a dual-task constrained loss function. The total loss is a linear combination of a fine-grained reconstruction loss and a coarse-grained physical consistency loss:

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{fine}} + \lambda_2 \mathcal{L}_{\text{coarse}} \tag{13}$$

The fine-grained reconstruction loss ($\mathcal{L}_{\text{fine}}$) supervises the final high-resolution output of the U-Net. Since the ground truth labels (GEMS data) themselves contain missing values, the loss is computed only over valid pixels. An L1 loss is adopted to enhance robustness against outliers in satellite data.

The physical consistency loss ($\mathcal{L}_{\text{coarse}}$) supervises the coarse-resolution prediction from the physical prior module. The GEMS labels are downsampled to the low-resolution grid after handling invalid values, and the loss is computed in the same masked L1 form. This auxiliary task ensures that the physical features possess direct predictive capability for NO_2 and provides a more direct gradient signal to the shallow layers of the network, alleviating the gradient dilution problem in deep networks. In our setup, both λ_1 and λ_2 are set to 1.0.

3.7 Evaluation metrics

This study employs multiple evaluation metrics to assess model performance, all of which are calculated based on valid observation pixels. Detailed descriptions follow:

The coefficient of determination (R^2) measures the extent of the variance between predictions and original observations. It is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{14}$$

where y_i and $\hat{y}_i$ are the observed and predicted values for the i-th sample, $\bar{y}$ is the mean of the observed values, and n is the total number of samples. R^2 ranges from $-\infty$ to 1, with values closer to 1 indicating a better fit. To evaluate the statistical significance of R^2, the F-test is employed to compute its associated probability (p-value). The F-statistic is defined as:

$$F = \frac{R^2/k}{(1 - R^2)/(n - k - 1)} \tag{15}$$

where k is the number of independent variables in the regression model. Under the null hypothesis ($R^2_{true} = 0$), F follows an F-distribution with degrees of freedom $(k, n-k-1)$. The p-value is computed via the survival function of the F-distribution, and p<0.05 indicates that R^2 is significantly different from zero at the 95% confidence level.

When validating against another independent source (e.g., the LEO satellite GOME-2), which may have inherent differences in retrieval algorithms and observation geometries, this study uses the Pearson correlation coefficient (R) as the core metric for cross-validation:

$$R = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}} \tag{16}$$

where $\bar{\hat{y}}$ is the mean of predicted values. Pearson R ranges from −1 to 1, and values closer to ±1 indicate stronger linear correlation.

Root Mean Square Error (RMSE) quantifies the overall magnitude of the prediction error. It is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{17}$$

RMSE has the same units as the observations (Pmolec/cm^2), facilitating a direct assessment of the error's actual magnitude.

Mean Absolute Error (MAE) measures the average magnitude of the errors, assigning equal weight to each error. It is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{18}$$

Compared to RMSE, MAE is more robust to outliers. Using both metrics can reveal characteristics of the error distribution: a RMSE value significantly larger than the MAE suggests the presence of a few samples with large errors.

Normalized Mean Bias (NMB) normalizes the mean bias relative to the mean of the observations, enabling comparable assessments across different regions and time periods. It is defined as:

$$NMB = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{\sum_{i=1}^{n} y_i} \tag{19}$$

NMB expresses the relative degree of systematic bias as a percentage, eliminating the influence of varying absolute NO_2 concentrations across different regions or seasons.

4. Results

4.1 Overall dataset accuracy and spatiotemporal consistency

To characterize model performance on held-out dates within the 2023 training year, Figure 2 presents the overall scatter density plots for the training, validation, and test sets. The predicted dataset achieved high prediction accuracy across all three sets, with R^2 values of 0.939, 0.882, and 0.889, respectively (all p-values < 0.01). Furthermore, during the training and testing process involving tens of millions of samples, the performance metrics on the validation and test sets were highly consistent, with no significant signs of overfitting. This indicates that the designed PhysNorm-Net model can effectively learn the non-linear mapping relationship between multi-source physical drivers and pollutant concentrations. Although this evaluation directly examines model performance under naturally occurring TROPOMI gaps wherever valid GEMS references are available, reconstructed pixels for which GEMS itself is unavailable cannot be directly verified against GEMS. Therefore, the reported metrics should not be interpreted as a pixelwise validation of the entire gapless product.

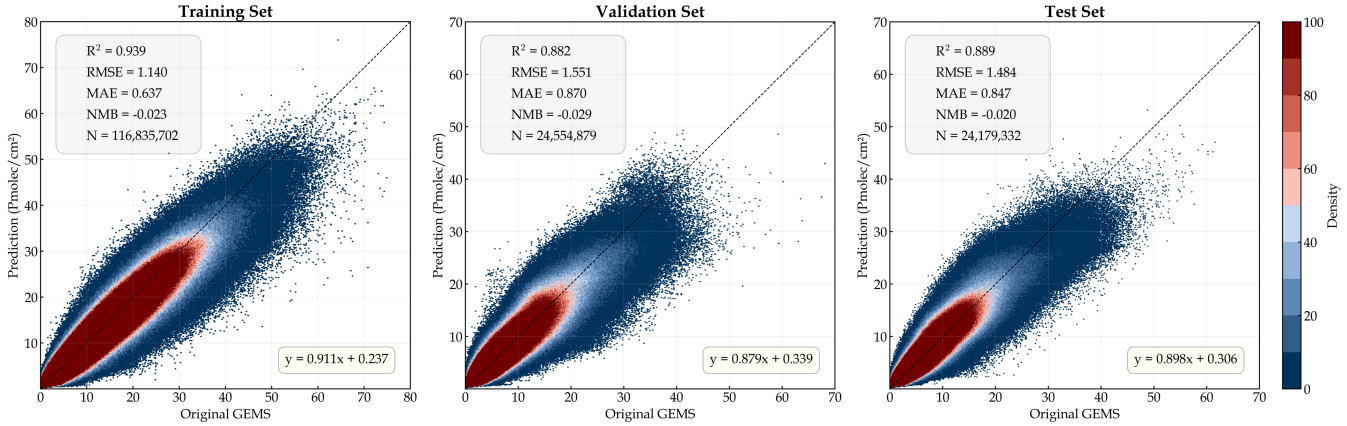


Figure 2: Scatter density plots of model predictions versus original GEMS observations.

Notably, the predicted dataset does not show a systematic underestimation or overestimation trend, except for a slight underestimation in regions with high tropospheric NO_2 column densities (>40 Pmolec/cm^2). A value-stratified evaluation of valid GEMS pixels in the fixed 2023 test set indicates that the reconstruction remains well constrained across the main moderate-to-high concentration range. For observed VCDs between 5 and 40 Pmolec/cm^2, which account for 21.73% of valid test pixels, the aggregated NMB is −0.044, with an MAE of 1.994 Pmolec/cm^2 and an RMSE of 2.753 Pmolec/cm^2. Within the 20–40 Pmolec/cm^2 range, the mean bias is −3.239 Pmolec/cm^2 (NMB = −0.126; N = 317,663), showing that the model retains useful discrimination of elevated pollution conditions. A more pronounced compression occurs only in the extreme tail above 40 Pmolec/cm^2, which contains 15,037 valid pixels (0.062% of the test set): the mean bias is −12.203 Pmolec/cm^2 (NMB = −0.264; MAE = 12.442 Pmolec/cm^2). This phenomenon is relatively common in current deep learning-based atmospheric remote sensing retrieval studies (Wei et al., 2021). The main reasons are twofold: first, the L1 loss functions used for model optimization tend to fit the global mean, with a relatively insufficient penalty weight for extreme high values; second, high-concentration samples constitute a very small fraction of the overall data distribution and often correspond to local, sudden emission events or adverse stagnant meteorological conditions, which low-resolution data cannot fully capture. Nevertheless, considering that the core concentration range of tropospheric NO_2 accounts for the vast majority of samples, the dataset demonstrates high reliability in this range, thereby supporting regional air quality assessment and spatiotemporal evolution analysis, including the characterization of the occurrence, spatial extent, and temporal evolution of elevated pollution conditions. However, analyses requiring the absolute magnitude of individual extreme peaks should account for the quantified negative bias.

Regarding spatial distribution details, Figure 3 uses April 26, 2023, as an example to show a single-day spatial comparison across the four typical regions. It is visually evident that the predicted dataset not only successfully fills large data gaps caused by factors like cloud cover but also accurately reproduces the spatial patterns of the original observations and the locations of heavy pollution plumes. In some local areas with complex terrain or dense emission sources (e.g., urban clusters in the NP region), the concentration gradients in the predictions are slightly smoother compared to the original high-resolution satellite observations. This is due to an unavoidable spatial smoothing effect introduced by the physical prior module when fusing the 0.25° physical driving fields with the 0.05° remote sensing data. However, this trade-off effectively eliminates non-physical abrupt changes in the original satellite products caused by retrieval noise and surface albedo errors, while also significantly mitigating common sensor striping artifacts. As a result, the final generated spatial distribution retains sufficiently accurate local hotspot information while adhering to the physical laws of atmospheric diffusion and transport.

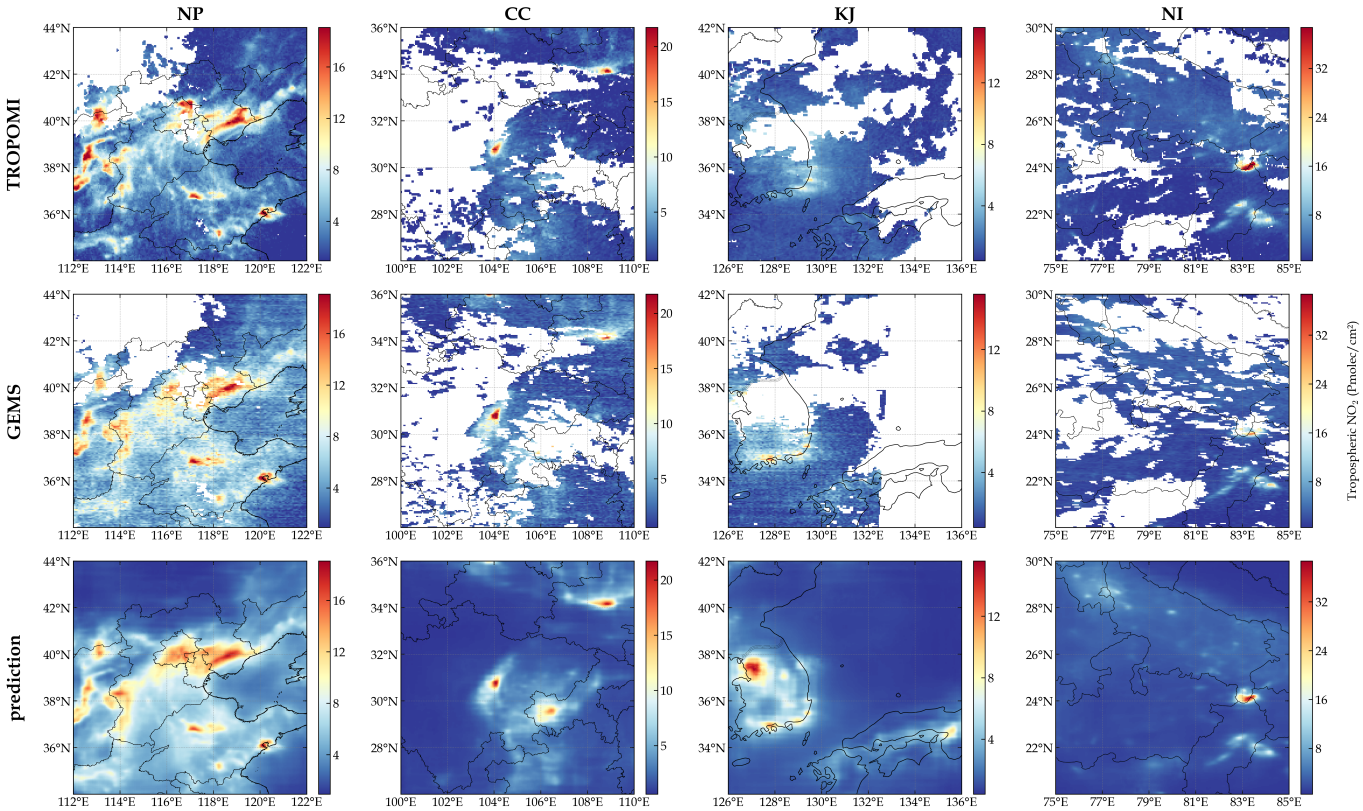


Figure 3: Comparison of single-day spatial distributions of NO_2 concentration from TROPOMI, GEMS, and model predictions on April 26, 2023.

Figure 4 shows the complete hourly GEMS observations and model predictions from 23:00 (UTC) on January 9, 2023 to 08:00 (UTC) on January 10, 2023, clearly demonstrating that the model not only has the ability to fill spatial gaps but also can predict the complete observation time series, thereby obtaining a more coherent dynamic change process.

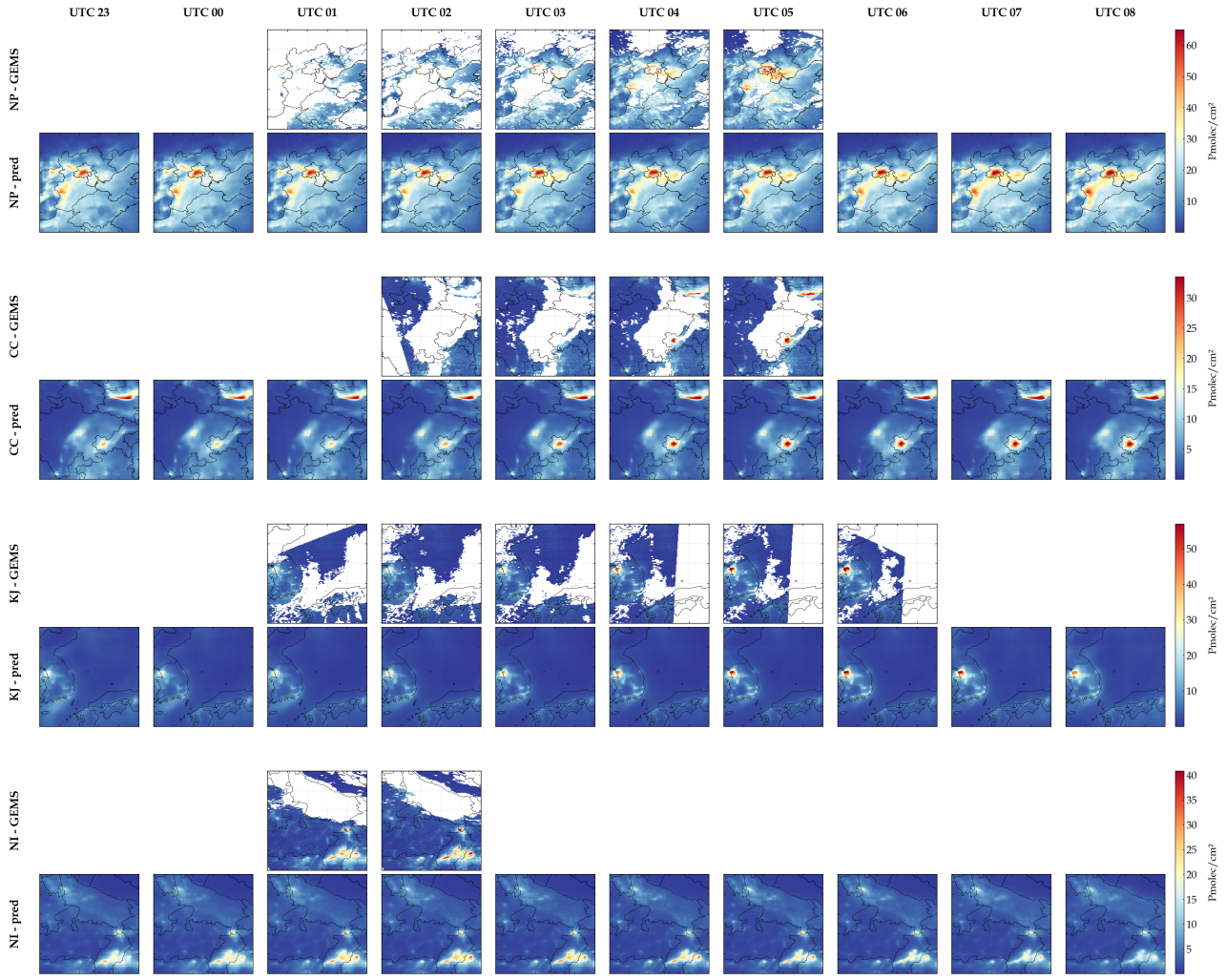


Figure 4: Hourly spatial distributions of NO_2 concentration from GEMS, and model predictions.

To assess reconstruction performance under TROPOMI input gaps, two complementary evaluations were conducted using the held-out 2023 test set. First, pixels with valid GEMS observations but missing TROPOMI inputs were evaluated, with the original TROPOMI availability mask retained during inference. The reconstruction achieved an R^2 of 0.801 and an RMSE of 1.235 Pmolec/cm^2 over this subset. Second, a controlled masking experiment assessed sensitivity to spatially contiguous input gaps using 60 temporally distributed test scenes. Because GEMS serves as the supervised target and evaluation reference rather than an inference input, artificial gaps were introduced by masking originally valid TROPOMI observations within contiguous blocks and setting the corresponding validity-mask values to indicate missing inputs, while retaining all other inputs. Four block scales were tested: 0.25°, 0.5°, 1°, and 2°. For the scenes eligible at all four scales, predictions within the artificially masked areas were evaluated against valid co-located GEMS observations. Relative to the corresponding unmasked predictions at the same locations, the RMSE increased by 0.015 Pmolec/cm^2 (1.2%), 0.853 Pmolec/cm^2 (35.6%), 1.327 Pmolec/cm^2 (64.9%), and 0.723 Pmolec/cm^2 (40.7%), respectively. These results indicate a small performance loss for the 0.25° gaps and greater degradation for the larger tested gaps, although the response was not monotonic with gap size. Together, these evaluations quantify reconstruction performance under naturally occurring and artificially imposed TROPOMI input gaps where valid GEMS references are available.

Reconstructed pixels where GEMS itself is unavailable cannot be directly verified against GEMS. Therefore, these results should not be interpreted as pixelwise validation of the entire gapless product.

To further validate the model's temporal extrapolation capability and long-term application potential, this study applied the model to the data of 2024, which was not involved in any training phase, for independent validation. The performance was compared by region with that of 2023, and the results are shown in Table 3.

Table 3: Comparison of model performance on 2023 and 2024 datasets by region.

Metric	NP		CC		KJ		NI	
	2023	2024	2023	2024	2023	2024	2023	2024
R^2	0.921	0.855	0.879	0.771	0.897	0.817	0.819	0.650
RMSE	1.722	2.339	0.829	1.006	0.884	1.017	1.096	1.403
MAE	1.055	1.461	0.507	0.610	0.492	0.571	0.561	0.815
NMB	-0.016	0.076	-0.023	0.119	-0.035	0.098	-0.041	0.226

The 2024 evaluation produced lower R^2 values and higher RMSE and MAE values than the internal 2023 test set but retained predictive skill in all four regions, with R^2 values above 0.81 in NP and KJ. These results indicate that the learned relationships remain partly transferable to an unseen year, although performance varied by region and decreased most in NI. Notably, the NMB changed from negative values in all four regions in 2023 (-0.041 to -0.016) to positive values in all four regions in 2024 (0.076 to 0.226), indicating a consistent positive mean bias in the independent-year evaluation. Because the model was trained on 2023 data and the normalization statistics for the dynamic variables were derived from that training set, this sign reversal may reflect an interannual shift in the predictor-target relationship relative to the training distribution; however, evaluation of only one unseen year is insufficient to establish persistent temporal drift. The comparatively larger decrease in NI may be partly associated with the smaller number of valid GEMS labels available for this region, which provides less information for model training. Overall, the 2024 results support partial temporal transferability of the proposed framework while revealing a year-dependent bias that should be considered when applying it to longer-term historical reconstruction.

4.2 Assessment of spatial hotspot capture and historical reconstruction capability using GOME-2

In order to evaluate the hotspot reconstruction capability and historical prediction accuracy of the dataset predicted by PhysNorm-Net, this section introduces the GOME-2 instrument as a completely independent reference object. Although GOME-2C is independent of model training, it is not a retrieval-harmonized reference. Differences in AMF formulation, a priori NO_2 profiles, cloud treatment, viewing geometry, and spatial sampling among GOME-2C, GEMS, and TROPOMI may introduce systematic differences among the retrieved tropospheric NO_2 VCDs. The following cross-satellite comparisons are therefore interpreted primarily as an assessment of consistency in spatial patterns and variability, rather than as an estimate of an absolute sensor-independent NO_2 scale.

Figure 5 presents selected examples of morning NO_2 spatial distributions across the four study regions. In the selected NP case, both the reconstruction and GOME-2C observations show an extensive area of elevated NO_2 over the North China Plain, with finer spatial structures visible in the reconstruction and available GEMS observations. In the CC case, enhanced columns around Chengdu, Chongqing, and the northeastern part of the domain are evident in both the reconstruction and satellite observations, although their magnitudes and spatial extents differ. In the KJ case, the reconstruction captures enhancements over western and southern South Korea that are also apparent in the available GEMS observations, whereas GOME-2C shows a smoother distribution and a weaker enhancement near Seoul. In the NI case, both the reconstruction and GOME-2C indicate elevated columns in the southeastern part of the domain, while differences remain in the intensity and extent of individual features.

These cases illustrate the reconstruction's ability to represent broad regional pollution patterns, while also revealing differences among the reconstructed fields and satellite observations.

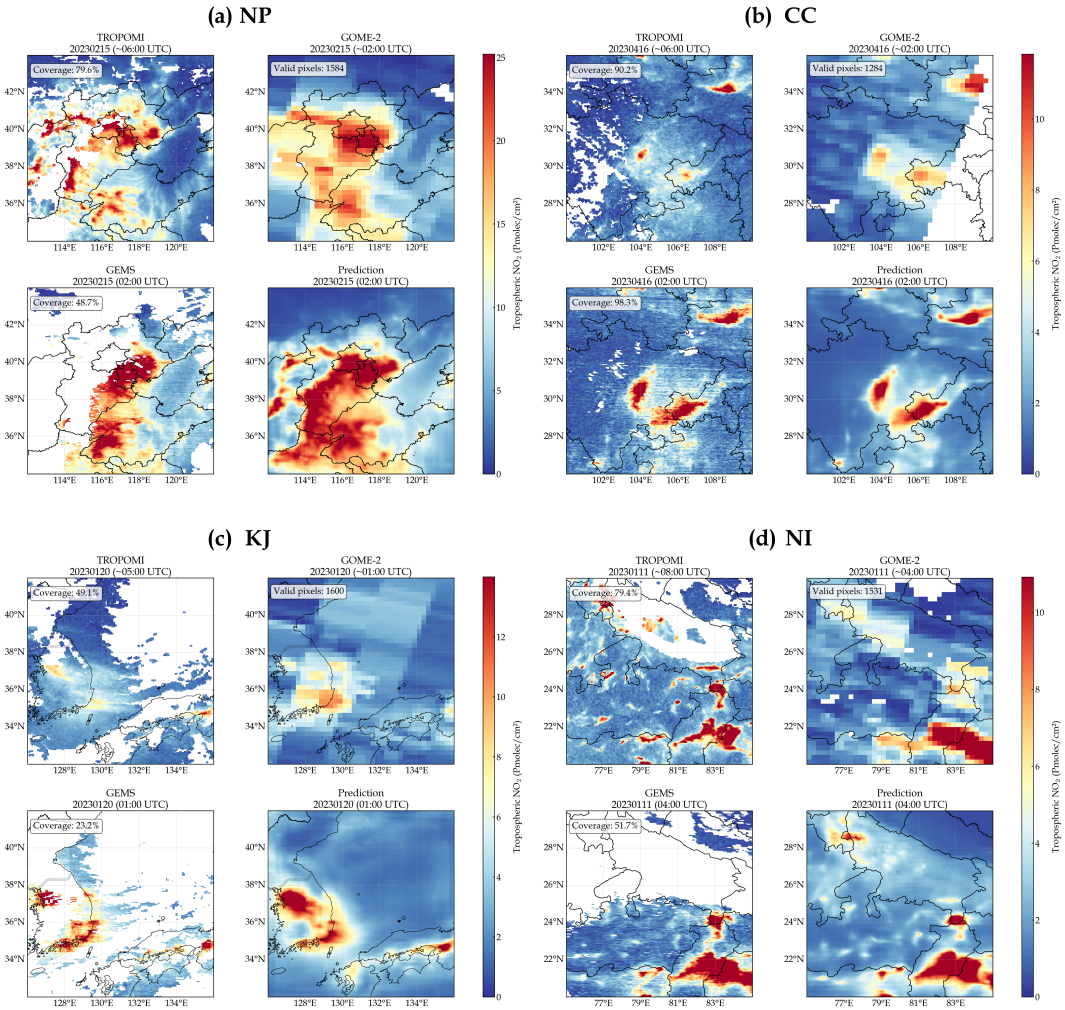


Figure 5: Comparison of tropospheric NO_2 spatial distributions from TROPOMI, GOME-2C, GEMS, and model predictions. GEMS and predictions are shown at the hourly time corresponding approximately to the GOME-2C overpass; TROPOMI observations are shown at their own overpass times.

Figure 6 displays the validation results of the predictions and original GEMS against GOME-2C from 2019 to 2024, across the four study regions and at the combined all-region scale. Since GEMS data are only available for 2023–2024, the comparison for 2019–2022 is only shown for predictions versus GOME-2C.

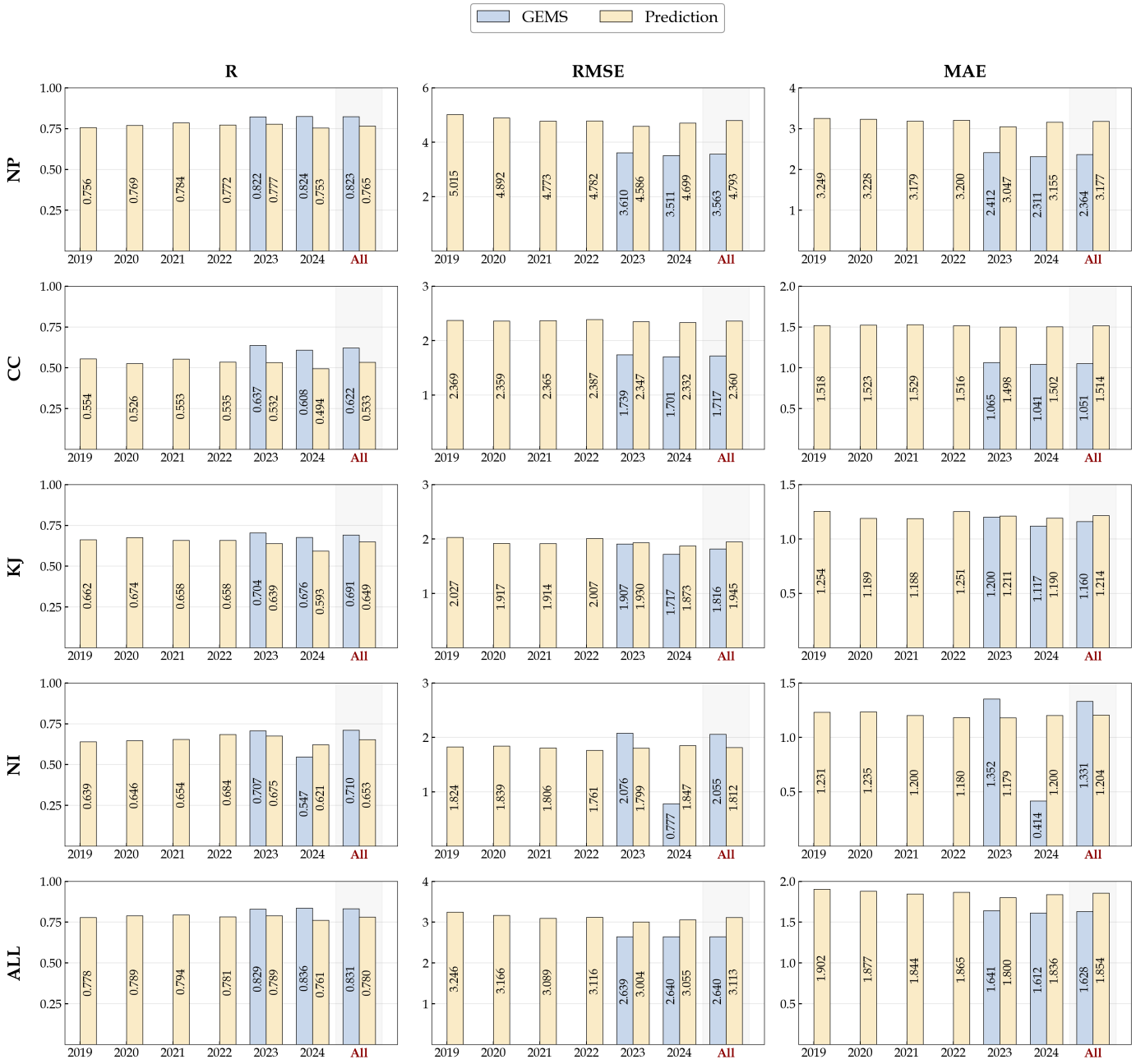


Figure 6: Annual and overall validation results of GEMS and predictions against GOME-2C for each region.

At the all-region scale, the predictions exhibit good inter-annual stability from 2019 to 2024. The annual R for the four combined regions ranges from 0.76 to 0.79, with the RMSE stable within 3.00–3.25 Pmolec/cm^2, and the MAE maintained between 1.80–1.90 Pmolec/cm^2. For the entire six-year period, the overall statistics for predictions against GOME-2C are R = 0.780, RMSE = 3.113 Pmolec/cm^2, and MAE = 1.854 Pmolec/cm^2. Notably, during the 2019–2022 retrospective period, the overall correlation and error levels are comparable to those of 2023–2024, indicating that the dataset predicted by the model has good stability in temporal extrapolation and does not show significant performance degradation when backtracking to non-training years.

At the regional scale, different regions exhibit clear differences. The NP region shows the highest correlation, with a six-year combined R of 0.765 for the predictions, but it also has the largest error magnitude, with RMSE and MAE of 4.794 and 3.177 Pmolec/cm^2, respectively. This suggests that the dataset can effectively capture the dominant structure of NO_2 spatial variations in the North China Plain but is accompanied by larger absolute errors in the context of high values and strong gradients. In contrast,

595 the CC region exhibits the lowest correlation, with a multi-year combined R of only 0.533 for the predictions, while its RMSE and
596 MAE are 2.360 and 1.514 Pmolec/cm^2, respectively. This indicates that although the absolute error is not high, the spatial
597 consistency is relatively weak in this region. The KJ and NI regions have relatively low overall error levels. The predicted RMSE
598 and MAE for the KJ region are 1.945 and 1.214 Pmolec/cm^2, and for the NI region, they are 1.812 and 1.204 Pmolec/cm^2. The R
599 for both regions is approximately 0.65, indicating that the dataset maintains a relatively stable fitting ability in low-to-medium
600 concentration backgrounds.

601 In the 2023–2024 period when GEMS observations are available, Figure 6 further validates the high consistency and
602 comparability between the predictions and the GEMS products. At the all-region scale, the correlation of GEMS with GOME-2C
603 (R = 0.829–0.836) indeed shows a slight advantage, while the predictions' R value for the same period also reaches a high level of
604 0.761–0.789. This demonstrates that the output dataset has successfully internalized and inherited the capability of GEMS data to
605 capture large-scale NO_2 spatial variability. A detailed regional comparison further highlights the robust performance of the
606 predictions across different pollution backgrounds. In KJ, the performance gap between predictions and GEMS is minimized, with
607 very close RMSE (approx. 1.87–1.93 vs. 1.72–1.91 Pmolec/cm^2) and MAE values. In NI, the predictions also show excellent
608 reliability: in 2023, its RMSE (1.799 Pmolec/cm^2) and MAE (1.179 Pmolec/cm^2) were both lower than those of GEMS for the
609 same period (RMSE = 2.076 Pmolec/cm^2, MAE = 1.352 Pmolec/cm^2).

610 Overall, the independent validation based on GOME-2C demonstrates that the predictions from PhysNorm-Net can stably
611 maintain a high overall correlation and a relatively controlled level of error during the 2019–2024 period. Furthermore, it did not
612 show significant performance degradation in the retrospective years without GEMS label support. This indicates that the
613 reconstructed dataset not only effectively inherits the spatiotemporal distribution information from GEMS but also has the potential
614 for extension to historical periods, providing a reliable foundation for constructing long-term, hourly tropospheric NO_2 datasets.

615 **4.3 Evaluation against ground-based Pandora observations**

616 Valid Pandora observations collected within ±30 minutes of each hourly product timestamp were averaged and matched to
617 the nearest 0.05° grid cell. Each hourly average required at least one valid Pandora observation, and each station contributed at
618 most one sample per hour. Comparisons used finite positive product values. Annual and combined R, RMSE, and MAE were
619 calculated from the collocated station-hour pairs. Reconstructed values were also evaluated separately where the corresponding
620 GEMS scene was present but the collocated pixel lacked a valid observation.

621 The evaluation yielded 83,262 collocated station-hour pairs across 16 stations during 2019–2024 (Table 4). The combined R,
622 RMSE, and MAE were 0.644, 7.200 Pmolec/cm^2, and 4.896 Pmolec/cm^2, respectively. Annual R ranged from 0.539 to 0.696,
623 while RMSE ranged from 6.528 to 9.052 Pmolec/cm^2. These comparisons extend the ground-based assessment to the historical
624 reconstruction period and to hours without valid GEMS observations. However, the number of contributing stations increased from
625 two in 2019 to 16 in 2024. Differences among annual statistics therefore reflect changes in station composition and observational
626 coverage as well as reconstruction performance, and should not be interpreted as a temporal trend in product accuracy.

627 **Table 4. Annual and combined evaluation of the reconstructed NO_2 product against Pandora observations during 2019–**
628 **2024.**

Year	Stations	N	R	RMSE	MAE
2019	2	588	0.539	9.052	5.696
2020	3	2,884	0.665	8.635	5.616
2021	8	11,942	0.638	8.569	5.878

2022	9	22,139	0.617	7.208	4.950
2023	11	17,133	0.696	6.868	4.837
2024	16	28,576	0.598	6.528	4.389
2019–2024	16	83,262	0.644	7.200	4.896

For 2023–2024, GEMS and the reconstruction achieved R values of 0.755 and 0.712, respectively, against Pandora on 16,407 common-valid samples, with corresponding RMSE values of 6.186 and 6.464 Pmolec/cm^2 (Table 5). On the 21,234 samples where the collocated GEMS pixel lacked a valid observation, the reconstruction achieved R = 0.612 and RMSE = 7.121 Pmolec/cm^2. The correlations for these gap-filled samples were 0.659 in 2023 and 0.570 in 2024, compared with Pandora–GEMS correlations of 0.804 and 0.699 on the respective valid-GEMS subsets. The lower combined correlation and higher RMSE for the gap-filled samples indicate reduced agreement under these sampling conditions. However, the subsets differ in observational conditions, so these differences cannot be attributed solely to gap filling. The station-based results also do not establish uniform accuracy across all reconstructed pixels or study regions.

Table 5. Evaluation against Pandora observations during 2023–2024: (a) comparison of GEMS and the reconstruction on common-valid samples; (b) evaluation of the reconstruction where collocated GEMS pixels lack valid observations.

(a) Common-valid samples

Period	N	GEMS R	GEMS RMSE	Reconstruction R	Reconstruction RMSE
2023	6844	0.804	6.493	0.765	6.584
2024	9563	0.699	5.957	0.651	6.377
2023–2024	16407	0.755	6.186	0.712	6.464

(b) GEMS-gap samples

Period	N	Reconstruction R	Reconstruction RMSE
2023	7,693	0.659	7.204
2024	13,541	0.570	7.073
2023–2024	21,234	0.612	7.121

4.4 Spatiotemporal analysis of the dataset predicted by PhysNorm-Net

Leveraging the predicted hourly tropospheric NO_2 dataset from 2019 to 2024, this section systematically analyzes the spatiotemporal features of tropospheric NO_2 across the four study regions, focusing on diurnal variations and inter-annual evolution at different spatial scales.

Figure 7 presents the multi-year average Month-Hour 2D heatmaps of NO_2 for the four regions. In the East Asian regions (NP, CC, KJ), an NO_2 valley typically appears between 12:00–14:00 local time (LT), with high-value bands in the morning (around 08:00 LT) and evening (after 17:00 LT). Seasonally, they always exhibit high values in winter and low values in summer. This macroscopic feature aligns well with physical expectations: the bimodal morning and evening peaks are primarily attributed to concentrated emissions from rush-hour traffic coupled with a shallower planetary boundary layer, while the midday trough is due to strong photochemical dissociation of NO_2 induced by intense solar radiation. The NI region exhibits a seasonal phase distinct from that of East Asia, with its highest concentrations occurring not in winter but in the spring (March–May, the pre-monsoon period). This unique characteristic may reflect a combination of intense surface heating, frequent agricultural biomass burning, and potentially weakened photochemical loss due to aerosol-radiation feedback from heavy dust aerosols, collectively leading to

extremely high NO_2 loads. This evolution of diurnal patterns with the seasons highlights the necessity of high-frequency and high-coverage observations for capturing complex regional pollution mechanisms.

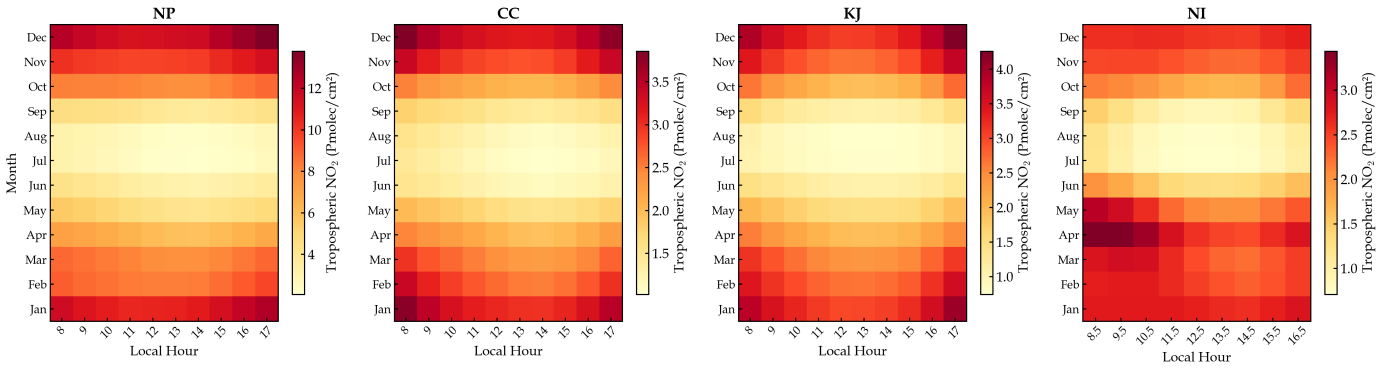


Figure 7: Month-Hour 2D distribution of tropospheric NO_2 in the four study regions from 2019 to 2024.

Beyond the regional macroscopic patterns, Figure 8 focuses on four representative cities, including Beijing, Chengdu, Seoul, and Delhi, to reveal fine-grained diurnal structures and their inter-annual evolution under the combined influence of anthropogenic activity, atmospheric chemistry, and meteorological conditions. The diurnal amplitude in Beijing is higher in winter, whereas strong photochemical loss and deeper boundary-layer mixing suppress the overall NO_2 levels and weaken the amplitude in summer. From 2019 to 2024, the seasonal diurnal curves in Beijing show a generally parallel downward shift, indicating a sustained decrease in the reconstructed NO_2 column burden across daytime hours. Chengdu exhibits a relatively flat diurnal variation, which may be associated with the enclosed topography of the Sichuan Basin, restricted horizontal ventilation, and limited boundary-layer development. The comparatively smaller decrease in Chengdu and the temporary rebound in 2021 may reflect the combined effects of local anthropogenic activity, regional transport, and inter-annual variations in basin meteorology. Seoul exhibits an afternoon peak in winter and spring, indicating that its NO_2 burden may be influenced by regional transport from surrounding industrial areas in addition to local traffic emissions. Seoul shows the largest six-year decrease among the four cities, consistent with the declining trend recorded by the official surface NO_2 monitoring network in the Seoul metropolitan area. Delhi is strongly modulated by the South Asian monsoon system. Its diurnal amplitude is largest during the pre-monsoon spring, with high morning values followed by a sharp afternoon decrease. During the summer monsoon, daytime concentrations decline to their annual minimum under the combined effects of convective mixing, wet removal, atmospheric transport, and seasonal changes in anthropogenic and natural sources.

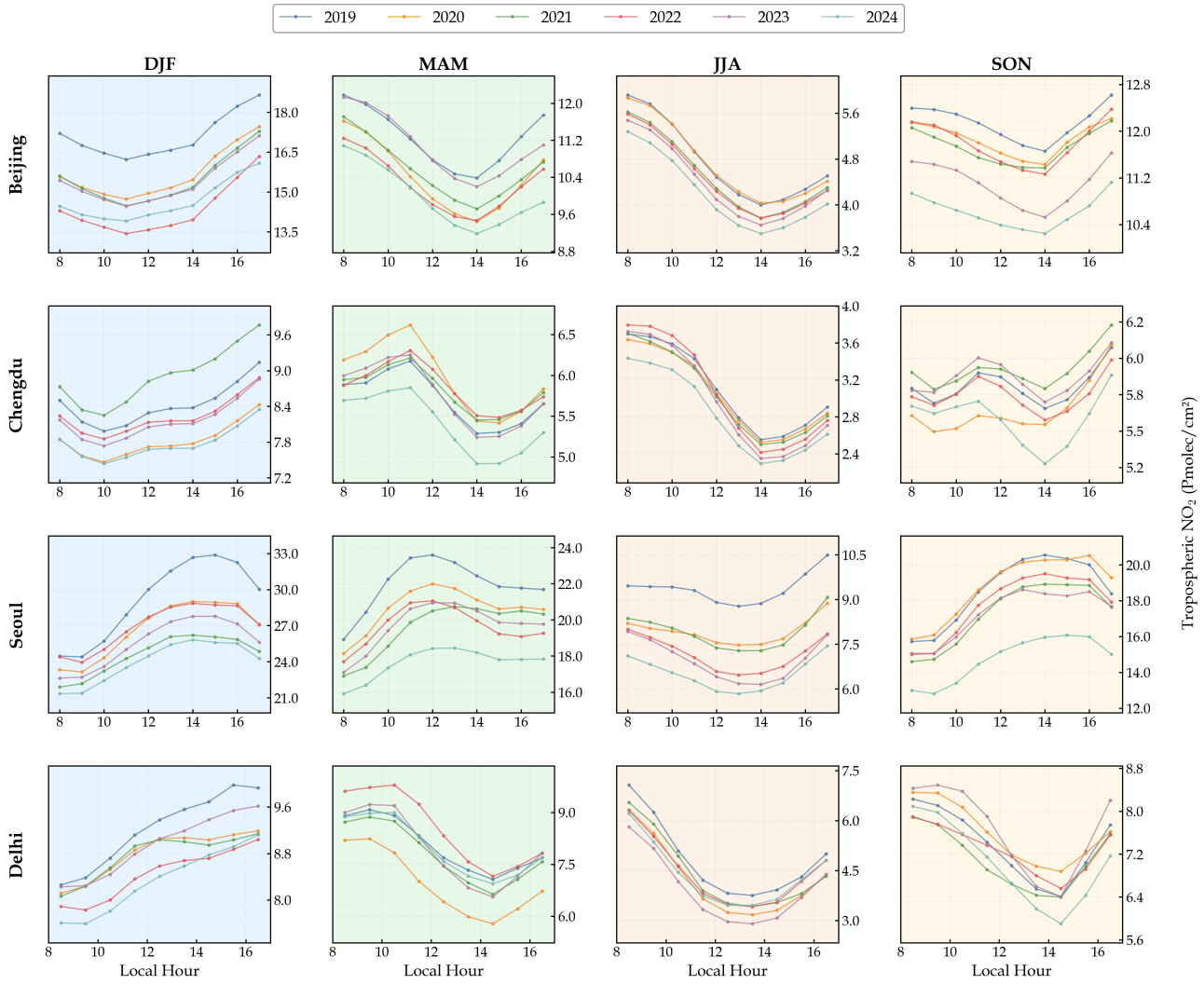


Figure 8: Inter-annual evolution of seasonal diurnal variations of tropospheric NO_2 in Beijing, Chengdu, Seoul, and Delhi from 2019 to 2024.

It should be emphasized that the inter-annual differences in the reconstructed tropospheric NO_2 VCDs should not be interpreted as direct estimates of changes in NO_x emissions. Tropospheric NO_2 columns are jointly affected by emissions, chemical lifetime, boundary-layer mixing, horizontal and vertical transport, and meteorological factors such as wind, temperature, humidity, and solar radiation (Seo et al., 2024; Yang et al., 2024; Wang et al., 2025). Accordingly, all the analyses describe inter-annual increases and decreases in the reconstructed NO_2 column burden rather than inferring direct changes in NO_x emissions. Socioeconomic activities and emission-control policies are therefore considered only as contextual information for interpreting the observed patterns. The reconstructed dataset characterizes when and where the NO_2 column burden changes, but it does not by itself quantitatively attribute those changes to specific drivers or separate policy-related effects from regional transport and meteorological variability.

Figures 9 and 10 use the hourly structure of the reconstructed dataset to compare inter-annual changes during full daytime (daily, 08:00–17:00 LT), morning (08:00–10:00 LT), midday (11:00–13:00 LT), and afternoon (14:00–17:00 LT) periods. This time-of-day-resolved analysis provides information that cannot be obtained from polar-orbiting satellite observations made at a single local overpass time. The broad 2019–2024 patterns are consistent with a recent TROPOMI-based analysis reporting

substantial NO_2 VCD decreases in North and East China and South Korea, together with more heterogeneous changes across India (Huber et al., 2026).

The NP region exhibited an overall decrease in reconstructed NO_2 VCD from 2019 to 2024 across the morning, midday, and afternoon periods. The 2019–2020 decrease was slightly larger in the morning, whereas the pronounced decrease from 2023 to 2024 occurred during all three periods. Figure 10 further shows that the strongest six-year decreases extended from the Beijing–Tianjin–Hebei urban corridor towards Shandong, with a broadly consistent spatial footprint throughout the daytime.

In the CC region, the decrease in 2020 and the subsequent rebound in 2021 occurred during all three daytime periods. Spatially, the principal urban centers, including Chengdu, Chongqing, and Xi'an, showed localized decreases, whereas several surrounding industrial and urban areas exhibited increases. In contrast to the relatively consistent daytime decrease in NP, the six-year decrease in CC was most pronounced during the afternoon.

The KJ region showed an overall decrease from 2019 to 2024 during the morning, midday, and afternoon periods, together with a modest rebound in 2022 and a pronounced decrease in 2024. Figure 10 indicates that the six-year decreases were concentrated over South Korea during all three periods, although localized inter-annual variations remained near Incheon, coastal areas, and the surrounding waters.

The NI region exhibited the most spatially heterogeneous pattern among the four study regions. Regional mean reconstructed NO_2 VCD increased markedly in 2022 during all three daytime periods and subsequently returned to approximately its 2019 level by 2024. Figure 10 reveals a complex mixture of increases and decreases along the Indo-Gangetic Plain, with clear differences in their spatial distribution among the morning, midday, and afternoon periods.

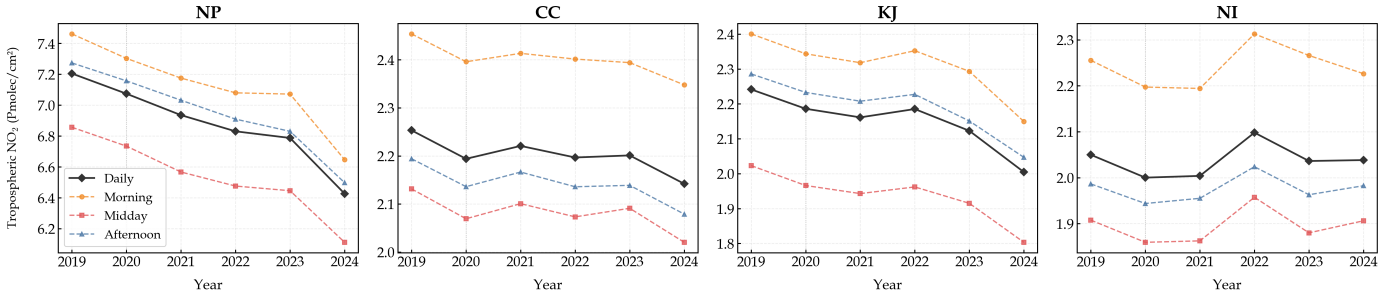
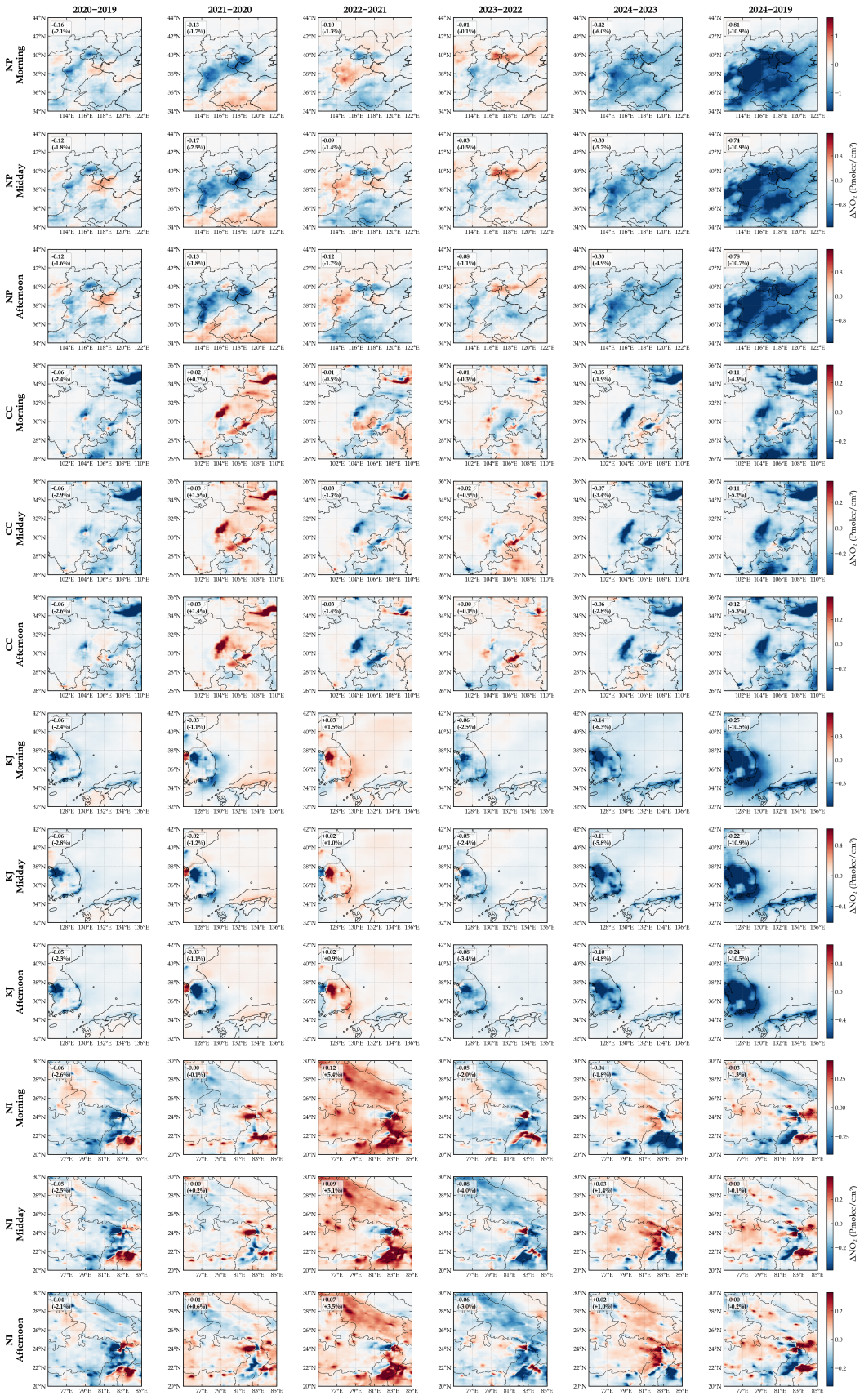


Figure 9: Inter-annual variation of tropospheric NO_2 in the four study regions during full daytime, morning, midday, and afternoon periods from 2019 to 2024.



711

712 **Figure 10: Spatial distribution of inter-annual differences in tropospheric NO_2 for different time slots in the four study**
713 **regions from 2019 to 2024.**

5. Discussion

5.1 Physical interpretability and feature importance analysis of PhysNorm-Net

The following diagnostics are reported to build trust in the dataset by demonstrating that the reconstruction responds to physical drivers, rather than as a stand-alone mechanistic study. This study first verified the functional integrity and physical responsiveness of the dual-branch architecture through a model diagnostic experiment (Fig. 11). The experiment compared the outputs of the normal prediction mode (using both satellite observations and physical prior inputs) with a "blind test" mode (where satellite branch inputs were zeroed out with a full missing-data mask, forcing the model to rely solely on the physical branch). This was done to examine the model's sensitivity to changes in physical inputs. The diagnostic results show that the model can still output the general distribution of NO_2 even when relying exclusively on the physical branch, which ensures predictive capability during periods without TROPOMI observations. The inclusion of the satellite branch then effectively corrects the concentration details and gradients. From a functional validation perspective, this finding confirms that the physical branch plays a primary role in providing foundational information, while the satellite branch is mainly responsible for the fine-grained refinement of high-resolution details.

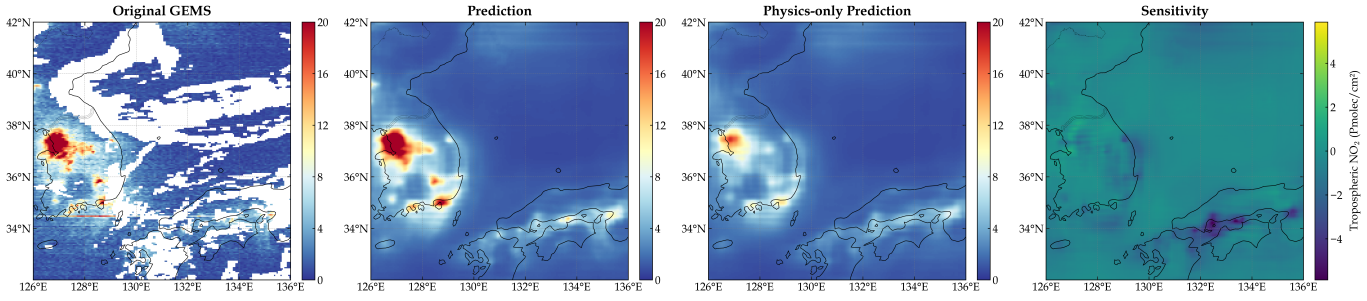


Figure 11: Functional diagnostic experiment of the PhysNorm-Net dual-branch architecture. The example shown corresponds to the KJ region at 04:00 UTC on 11 November 2024.

Building on this diagnostic foundation, to further unveil the model's decision-making mechanism, we quantitatively analyzed the relative contributions of each input variable during daytime hours in 2023 using a zero-out perturbation method (Zhang and Gao, 2020). This method calculates a variable's importance score by zeroing out its specific channel and measuring the resulting deviation from the baseline output, quantified by MAE. Figure 12 reports the three highest-ranked individual input channels for each bin. The dominant variable combinations differ significantly across the four regions, indicating that the model does not simply rely on a fixed set of factors but instead forms region-specific control structures.

Because several input variables are physically and statistically correlated, such as specific humidity and temperature across multiple pressure levels, the individual-channel scores are not additive and should not be interpreted as unique causal contributions. When correlated variables contain overlapping information, the model may compensate for the perturbation through the remaining channels, or the perturbation may disrupt information shared across several channels. Therefore, the resulting rankings represent conditional model sensitivity under the applied perturbation, rather than independent explanatory importance; the rankings in Fig. 12 are interpreted as diagnostic evidence of relative model sensitivity under the sampled 2023 conditions, and are considered together with the physical context of each region and season.

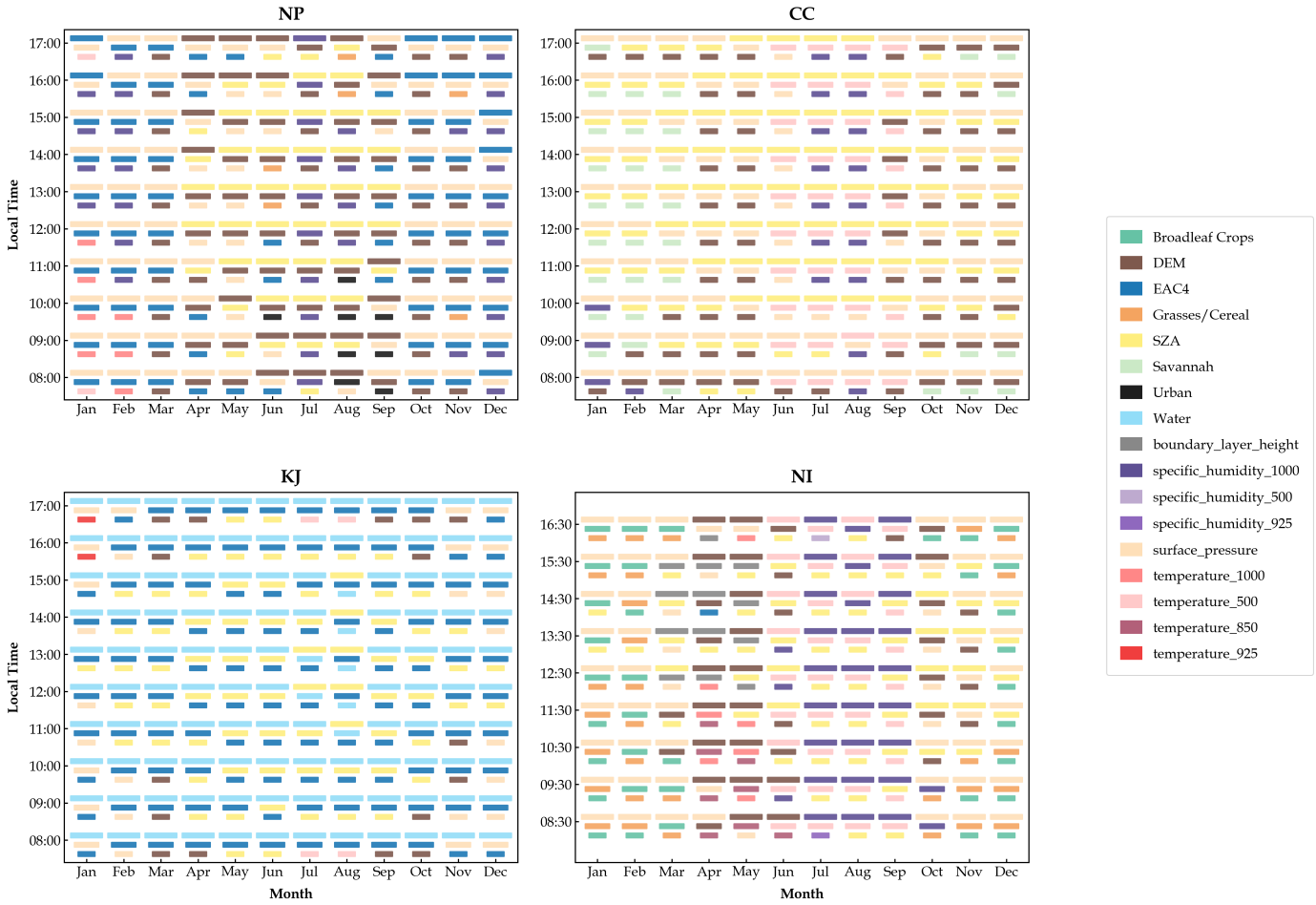


Figure 12: Monthly and hourly feature importance rankings for the four regions based on the zero-out perturbation method.

As shown in Fig. 12, in the NP region, the most stable and prominent variables during winter and late autumn are EAC4 and surface pressure, with the third position often supplemented by near-surface or mid-tropospheric thermodynamic variables. This channel-level pattern has a clear physical meaning: in East Asian winters, frequent temperature inversions, stable weather conditions, and high relative humidity collectively suppress vertical mixing and promote the accumulation of NO_2 in the near-surface layer (Chai et al., 2014; Xu et al., 2011). Concurrently, the increased importance of EAC4 during certain winter periods indicates greater model sensitivity to the large-scale chemical background represented by EAC4. In summer (JJA), SZA becomes more prominent and shows a strong daytime peak around noon. This aligns with the physical processes of enhanced convective mixing and accelerated photochemical decomposition in the warm season (Liu et al., 2022).

In the CC region, the rankings are dominated by surface pressure and SZA. Surface pressure is the top-ranked input channel throughout the daytime in winter and remains prominent during the morning and late-afternoon periods in transitional months, whereas SZA is generally top-ranked from spring to early autumn, particularly around midday. DEM frequently appears among the three highest-ranked channels outside midsummer, while the savannah land-cover channel is more prominent during colder months. During summer, the 500 hPa temperature and 1000 hPa specific humidity channels also become prominent. These seasonal and diurnal shifts indicate changes in model sensitivity between circulation and ventilation constraints and radiation and thermodynamic conditions, consistent with the meteorological characteristics of the Sichuan Basin (Sun et al., 2020; Feng et al., 2022).

In the KJ region, the water-body land-cover channel is the top-ranked variable during almost all months and daytime hours. This implies that in a peninsula-archipelago system with complex coastlines and strong land-sea segmentation, the water body variable effectively encodes the regional land-sea distribution, which in turn determines the spatial extent of emission sources and sea breeze influence. EAC4 also remains among the three highest-ranked channels in nearly all month–hour bins, generally occupying the second or third position. SZA and surface pressure account for most of the other highly ranked channels, with their prominence varying across seasons and daytime hours. Overall, the channel-level results indicate persistent model dependence on both the land-sea spatial pattern and the large-scale chemical background, together with seasonal and diurnal modulation by radiation and meteorological conditions.

The NI region exhibits a strong seasonal switch. During winter and late autumn, surface pressure is almost consistently ranked first, indicating that NO_2 distribution in this region is primarily controlled by boundary layer stability and ventilation conditions during the cold and transitional seasons. Concurrently, the land use types grasslands/cereal crops and broadleaf crops become prominent in winter. These agricultural land types are widespread in the Indo-Gangetic Plain, and their high importance reflects the impact of agricultural burning—the widespread burning of crop residues in this region is a major seasonal source of NO_x (Cusworth et al., 2018; Jethva et al., 2018). Around the monsoon season, the dominant variables shift to specific humidity, temperature, and SZA, with these thermodynamic and radiation-related channels appearing more frequently among the top-ranked variables, whereas the prominence of individual land-use channels decreases. This reflects the profound influence of monsoon dynamics on vertical mixing, wet scavenging, and the large-scale redistribution of pollutants (Ravindra et al., 2003).

The rankings in Fig. 12 quantify the model's dependence on individual input channels rather than the aggregate contributions of broader physical processes. Furthermore, the zero-out perturbation method measures model dependency rather than strict physical causality. Therefore, the importance of static variables such as land use and DEM should be interpreted as the model's reliance on spatial encoding and regional templates.

During the above feature-importance analysis, EAC4 was identified as an important chemical prior. We therefore further quantified the dependence of the reconstructed values on EAC4 using all 1,450 scenes in the held-out 2023 test set. With the model weights and all other inputs fixed, the de-standardized EAC4 field was perturbed by ±20%, and the results were pooled over GEMS-valid pixels for which the TROPOMI input was missing. The baseline reconstruction for this subset achieved R^2=0.801 and an RMSE of 1.235 Pmolec/cm^2. Increasing and decreasing EAC4 by 20% changed the pooled mean reconstruction by +4.02% and −4.60%, respectively, corresponding to normalized sensitivity to EAC4 perturbations of 0.216 (estimated using a central difference as $E = \frac{\bar{\hat{y}}_{+20\%} - \bar{\hat{y}}_{-20\%}}{0.4\bar{\hat{y}}_{\text{baseline}}}$, where $\bar{\hat{y}}_{+20\%}$, $\bar{\hat{y}}_{-20\%}$, and $\bar{\hat{y}}_{\text{baseline}}$ denote the mean reconstructed NO_2 VCD over the same pixel subset with EAC4 increased by 20%, decreased by 20%, and unchanged, respectively). This dimensionless metric quantifies the relative response of the pooled mean reconstruction to the imposed EAC4 perturbations, rather than prediction bias relative to GEMS. Its value of 0.216 indicates that the relative output response is approximately 21.6% of the relative input perturbation, as estimated across the tested ±20% range. The region-specific normalized sensitivities were 0.134, 0.269, 0.220, and 0.227 for CC, KJ, NI, and NP, respectively. This attenuated and regionally variable response indicates that EAC4 provides an important large-scale chemical constraint but is not transferred proportionally to the model output.

5.2 Limitations and future work

Although the validation results support several uses of the reconstructed dataset, users should consider the following limitations.

The training domains were selected to emphasize representative high-emission and complex-terrain regions, which increases the representation of polluted and high-gradient conditions but limits model generalizability to very clean backgrounds, such as remote oceanic regions. Uncertainty under low-column conditions also arises from treating satellite NO_2 VCD retrievals less than or equal to zero as missing values. Although this treatment is consistent with the nonnegative model domain and the logarithmic transformation, negative retrievals form part of the random retrieval-error distribution, and their exclusion may introduce an upward conditional-sampling bias near the satellite retrieval detection limit. The reported evaluations therefore characterize the positive-retrieval subset, and the product should not be interpreted as an unbiased estimate under ultra-low background conditions. Future work will examine adaptive sampling, loss weighting, sign-preserving preprocessing, and alternative transformation strategies.

The framework was trained primarily on 2023 data from four hotspot regions. Longer training records may improve robustness to extreme meteorological conditions and inter-annual shifts in the emission baseline. Because the physicochemical evolution of tropospheric NO_2 depends on local meteorology, atmospheric chemistry, transport, and emission characteristics, direct application of the model beyond the training domains may be affected by domain shift. Transfer to a new region should therefore be accompanied by region-specific validation, with fine-tuning using local reference data when performance degradation is observed. Regions characterized by substantially different emission structures, meteorological regimes, topography, or seasonal behavior may require more extensive retraining (Tiwari et al., 2026). Independent evaluation should be conducted before the reconstructed product is used for quantitative analyses in a new domain.

At the data-product level, integrating the 0.25° physical prior fields with the 0.05° observations may introduce spatial smoothing. Although this integration can suppress retrieval noise, it may also weaken localized gradients around individual emission sources. Related uncertainty arises from the EAC4 chemical prior because EAC4 tcno2 is a coarse-resolution, 3-hourly total-column quantity, whereas the target is an hourly tropospheric VCD. The column-type mismatch may transmit EAC4 background or stratospheric biases into the reconstruction, while linear temporal interpolation may smooth rapidly evolving chemical features between EAC4 analysis times. These effects are expected to be most relevant where TROPOMI information is absent or sparse. Future work will examine higher-temporal-resolution and vertically resolved tropospheric chemical priors and methods that explicitly represent uncertainties associated with temporal interpolation and column-type mismatch.

Independent evaluation of the 2019–2022 historical reconstruction remains limited by the temporal and spatial coverage of the reference observations. GOME-2C provides a satellite comparison near its approximately 09:30 LT overpass, while Pandora supplies additional hourly comparisons at the available ground stations. However, Pandora coverage is concentrated in KJ, with limited coverage in NP and NI and no eligible station in CC in this evaluation. These observations therefore do not provide comprehensive validation of daytime variability throughout all study domains. Cross-sensor retrieval differences add further uncertainty because the TROPOMI input, GEMS training target, and GOME-2C external reference differ in their retrieval configurations, observation times, and spatial sampling. Differences in AMF formulation, a priori NO_2 profiles, cloud treatment, viewing geometry, and spatial footprints may affect both the relationship learned between TROPOMI and GEMS and the discrepancies observed in the GOME-2C evaluation. Future work will assess cross-sensor consistency using retrieval-harmonized products or comparison approaches that explicitly account for these differences. The continued operation of GEMS, TEMPO, and Sentinel-4 will also provide opportunities to examine cross-sensor transferability and extend the framework to broader daytime-hourly NO_2 applications.

6. Conclusions

This study presents the first seamless, 0.05° daytime-hourly tropospheric NO_2 vertical column density dataset covering typical Asian hotspots from 2019 to 2024. To overcome the core challenges of spatial discontinuity and short historical records inherent in high-frequency geostationary satellite data, this dataset was reconstructed using an advanced physics-aware deep learning framework (PhysNorm-Net).

The high quality and physical reliability of this dataset are ensured through a rigorous multi-modal fusion mechanism. First, this study utilizes a Partial Convolutional network with dynamic mask updates to effectively isolate mean-value contamination from invalid pixels, enabling the precise extraction of high-resolution spatial textures. Second, deviating from traditional channel-wise concatenation, PhysNorm module uses multi-level meteorological, topographical, and chemical prior fields as environmental context to adaptively control the affine transformation of high-resolution features. This ensures that the reconstructed concentration fields strictly adhere to the physical laws of advection-diffusion and photochemical reactions. Third, the loss function adopts a combination of fine-grained reconstruction and coarse-grained physical consistency constraints to further mitigate the gradient dilution problem in deep networks.

Quantitative assessments confirm that the reconstructed dataset achieves extremely high fitting accuracy on the test set (R^2 = 0.889) and can effectively recover morning and evening pollution hotspots missed by polar-orbiting instruments like TROPOMI due to their spatiotemporal coverage limitations. More importantly, the long-term retrospective validation against GOME-2C data (R = 0.780), not used in the PhysNorm-Net training, demonstrates that the dataset possesses robust temporal stability and high fidelity in historical periods (2019–2022) prior to the operational deployment of GEMS. Pandora observations provide additional ground-based evaluation during the historical reconstruction period, although their uneven station coverage, together with the restricted overpass sampling of GOME-2C, limits comprehensive validation of daytime variability across all study domains.

Through in-depth analysis, this newly generated long-term hourly dataset has proven its unique scientific utility in capturing complex atmospheric dynamics. It not only clearly reveals the monsoon-driven seasonal phase reversal in North India and the typical bimodal and single-trough diurnal pattern in East Asia, but also characterizes the sustained decreases in reconstructed NO_2 in the North China Plain and the Korea-Japan region since 2019, together with localized inter-annual rebounds in specific years (e.g., 2021). By combining sparse satellite retrievals with continuous physical priors, the dataset provides a continuous record that can support studies of long-term air pollution variability, health exposure, and epidemiological applications across the four study regions.

Data availability

TROPOMI (https://doi.org/10.5270/S5P-9bnp8q8) and MCD12C1 (https://doi.org/10.5067/MODIS/MCD12C1.061) data used in this study are available on the EARTHDATA platform. GEMS data can be accessed at https://nesc.nier.go.kr/en/html/datasvc/data.do?pageIndex=1&outputInnb=64&atrb=NO2_Trop. The GOME-2 data are provided by the EUMETSAT AC-SAF as described in https://acsaf.org/products/nto_no2.php. ERA5 on single and pressure levels can be accessed at https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview and https://cds.climate.copernicus.eu/datasets/reanalysis-era5-pressure-levels?tab=overview, respectively. Pandora rnvh3p1-8 observations are available from the Pandonia Global Network through https://api.pandonia-global-network.org/docs. EAC4 data can be accessed at https://ads.atmosphere.copernicus.eu/datasets/cams-global-reanalysis-eac4?tab=overview. ETOPO data can be found at https://www.ncei.noaa.gov/products/etopo-global-relief-model. The gapless 0.05° daytime-hourly tropospheric NO_2 dataset (2019–2024) produced in this study is available at https://doi.org/10.5281/zenodo.20427767 (Gao et al., 2026).

Author contributions

Hongrui Gao: Methodology, Investigation, Formal Analysis, Visualization, Data Curation, Writing – Original Draft; Qin He: Conceptualization, Investigation, Writing – Review & Editing. Kai Qin: Funding Acquisition, Project Administration, Supervision. Jhoon Kim: Writing – Review & Editing. Diego Loyola: Writing – Review & Editing. Pravash Tiwari: Writing – Review & Editing. Lingxiao Lu: Writing – Review & Editing. Jason B. Cohen: Supervision, Writing – Review & Editing.

Competing interests

The authors declare that they have no conflict of interest.

Financial support

This work was supported by the National Natural Science Foundation of China [Grant Numbers 42375125].

References

Ahmad, N., Lin, C., Lau, A. K. H., Kim, J., Zhang, T., Yu, F., Li, C., Li, Y., Fung, J. C. H., and Lao, X. Q.: Estimation of ground-level NO_2 and its spatiotemporal variations in China using GEMS measurements and a nested machine learning model, Atmospheric Chemistry and Physics, 24, 9645–9665, https://doi.org/10.5194/acp-24-9645-2024, 2024.

Alvera-Azcárate, A., Barth, A., Rixen, M., and Beckers, J. M.: Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic Sea surface temperature, Ocean Modell., 9, 325–346, https://doi.org/10.1016/j.ocemod.2004.08.001, 2005.

Auer, A. H.: Correlation of Land Use and Cover with Meteorological Anomalies, Journal of Applied Meteorology (1962-1982), 17, 636–643, https://doi.org/10.1175/1520-0450(1978)017%3C0636:COLUAC%3E2.0.CO;2, 1978.

Baruah, U. D., Robeson, S. M., Saikia, A., Mili, N., Sung, K., and Chand, P.: Spatio-temporal characterization of tropospheric ozone and its precursor pollutants NO2 and HCHO over South Asia, Science of The Total Environment, 809, 151135, https://doi.org/10.1016/j.scitotenv.2021.151135, 2022.

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.-Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K. T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., Birk, M., Cyrys, J., Von Klot, S., Nádor, G., Varró, M. J., Dėdelė, A., Gražulevičienė, R., Mölter, A., Lindley, S., Madsen, C., Cesaroni, G., Ranzi, A., Badaloni, C., Hoffmann, B., Nonnemacher, M., Krämer, U., Kuhlbusch, T., Cirach, M., De Nazelle, A., Nieuwenhuijsen, M., Bellander, T., Korek, M., Olsson, D., Strömgren, M., Dons, E., Jerrett, M., Fischer, P., Wang, M., Brunekreef, B., and De Hoogh, K.: Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project, Atmos. Environ., 72, 10–23, https://doi.org/10.1016/j.atmosenv.2013.02.037, 2013.

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, 1737–1739, https://doi.org/10.1126/science.1207824, 2011.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO_2 from OMI, Atmospheric Chemistry and Physics, 7, 2103–2118, https://doi.org/10.5194/acp-7-2103-2007, 2007.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the ozone monitoring instrument, Atmospheric Measurement Techniques, 4, 1905–1928, https://doi.org/10.5194/amt-4-1905-2011, 2011a.

912 Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep,
913 M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO2 column retrieval algorithm for the
914 ozone monitoring instrument, Atmospheric Measurement Techniques, 4, 1905–1928, https://doi.org/10.5194/amt-4-1905-2011,
915 2011b.

916 Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.:
917 SCIAMACHY: Mission objectives and measurement modes, 1999.

918 Burnett, R. T., Stieb, D., Brook, J. R., Cakmak, S., Dales, R., Raizenne, M., Vincent, R., and Dann, T.: Associations between
919 Short-Term Changes in Nitrogen Dioxide and Mortality in Canadian Cities, Archives of Environmental Health: An International
920 Journal, 59, 228–236, https://doi.org/10.3200/AEOH.59.5.228-236, 2004.

921 Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., DeBeek, R., Hoogen, R.,
922 Bramstedt, K., Eichmann, K.-U., Eisinger, M., and Perner, D.: The global ozone monitoring experiment (GOME): Mission concept
923 and first scientific results, 1999.

924 Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2-metop's second-generation sensor for operational
925 ozone monitoring, ESA Bull., 102, 28–36, 2000.

926 Chai, F., Gao, J., Chen, Z., Wang, S., Zhang, Y., Zhang, J., Zhang, H., Yun, Y., and Ren, C.: Spatial and temporal variation of
927 particulate matter and gaseous pollutants in 26 cities in China, Journal of Environmental Sciences, 26, 75–82,
928 https://doi.org/10.1016/S1001-0742(13)60383-6, 2014.

929 Chan, K. L., Valks, P., Heue, K.-P., Lutz, R., Hedelt, P., Loyola, D., Pinardi, G., Van Roozendael, M., Hendrick, F., Wagner, T.,
930 Kumar, V., Bais, A., Piters, A., Irie, H., Takashima, H., Kanaya, Y., Choi, Y., Park, K., Chong, J., Cede, A., Frieß, U., Richter, A.,
931 Ma, J., Benavent, N., Holla, R., Postylyakov, O., Rivera Cárdenas, C., and Wenig, M.: Global ozone monitoring experiment-2
932 (GOME-2) daily and monthly level-3 products of atmospheric trace gas columns, Earth System Science Data, 15, 1831–1870,
933 https://doi.org/10.5194/essd-15-1831-2023, 2023.

934 Chang, B., Liu, H., Zhang, C., Xing, C., Tan, W., and Liu, C.: Relating satellite NO2 tropospheric columns to near-surface
935 concentrations: implications from ground-based MAX-DOAS NO2 vertical profile observations, Npj Clim. Atmos. Sci., 8, 1,
936 https://doi.org/10.1038/s41612-024-00891-z, 2025.

937 Chen, T.-H., Hsu, Y.-C., Zeng, Y.-T., Candice Lung, S.-C., Su, H.-J., Chao, H. J., and Wu, C.-D.: A hybrid kriging/land-use
938 regression model with asian culture-specific sources to assess NO2 spatial-temporal variations, Environmental Pollution, 259,
939 113875, https://doi.org/10.1016/j.envpol.2019.113875, 2020.

940 Choi, H., Liu, X., Jeong, U., Chong, H., Kim, J., Ahn, M. H., Ko, D. H., Lee, D.-W., Moon, K.-J., and Lee, K.-M.: Geostationary
941 environment monitoring spectrometer (GEMS) polarization characteristics and correction algorithm, Atmospheric Measurement
942 Techniques, 17, 145–164, https://doi.org/10.5194/amt-17-145-2024, 2024.

943 Cusworth, D. H., Mickley, L. J., Sulprizio, M. P., Liu, T., Marlier, M. E., DeFries, R. S., Guttikunda, S. K., and Gupta, P.:
944 Quantifying the influence of agricultural fires in northwest India on urban air pollution in Delhi, India, Environ. Res. Lett., 13,
945 044018, https://doi.org/10.1088/1748-9326/aab303, 2018.

946 Duncan, B. N., Lamsal, L. N., Thompson, A. M., Yoshida, Y., Lu, Z., Streets, D. G., Hurwitz, M. M., and Pickering, K. E.: A
947 space-based, high-resolution view of notable changes in urban NOx pollution around the world (2005–2014), Journal of
948 Geophysical Research: Atmospheres, 121, 976–996, https://doi.org/10.1002/2015JD024121, 2016.

949 von Engeln, A. and Teixeira, J.: A planetary boundary layer height climatology derived from ECMWF reanalysis data, Journal of
950 Climate, 26, 6575–6590, https://doi.org/10.1175/JCLI-D-12-00385.1, 2013.

951 Fan, Y., Sun, L., and Liu, X.: GOCI-II geostationary satellite hourly aerosol optical depth obtained by data-driven methods:
952 Validation and comparison, Atmos. Environ., 310, 119965, https://doi.org/10.1016/j.atmosenv.2023.119965, 2023.

953 Feng, X., Wang, S., and Guo, J.: Temperature inversions in the lower troposphere over the sichuan basin, China: Seasonal feature
954 and relation with regional atmospheric circulations, Atmospheric Research, 271, 106097,
955 https://doi.org/10.1016/j.atmosres.2022.106097, 2022.

Flemming, J., Peuch, V. H., and Jones, L.: Ten years of forecasting atmospheric composition at ECMWF, in: ECMWF Newsletter, vol. 152, 2017.

Friedl, M. and Sulla-Menashe, D.: MODIS/Terra+ Aqua land cover type yearly L3 global 0.05 Deg CMG V061, NASA EOSDIS Land Processes Distributed Active Archive Center (DAAC) data set, MCD12C1. 061, 2022.

Fu, J., Tang, D., Grieneisen, M. L., Yang, F., Yang, J., Wu, G., Wang, C., and Zhan, Y.: A machine learning-based approach for fusing measurements from standard sites, low-cost sensors, and satellite retrievals: Application to NO2 pollution hotspot identification, Atmos. Environ., 302, 119756, https://doi.org/10.1016/j.atmosenv.2023.119756, 2023.

Gao, H., He, Q., Qin, K., Kim, J., Loyola, D., Tiwari, P., Lu, L., and Cohen, J. B.: A gapless 0.05° hourly tropospheric NO2 dataset (2019–2024) over key Asian hotspots reconstructed using physics-aware deep learning, https://doi.org/10.5281/zenodo.20427767, 2026.

Geffen, V., Eskes, H., Boersma, K., Maasakkers, J., and Veefkind, J.: TROPOMI ATBD of the total and tropospheric NO2 data products, DLR document, 2019.

Gohm, A., Harnisch, F., Vergeiner, J., Obleitner, F., Schnitzhofer, R., Hansel, A., Fix, A., Neininger, B., Emeis, S., and Schäfer, K.: Air pollution transport in an alpine valley: Results from airborne and ground-based observations, Boundary-Layer Meteorol, 131, 441–463, https://doi.org/10.1007/s10546-009-9371-9, 2009.

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., 53, 12594–12601, https://doi.org/10.1021/acs.est.9b04488, 2019.

Griffin, D., Zhao, X., McLinden, C. A., Boersma, F., Bourassa, A., Dammers, E., Degenstein, D., Eskes, H., Fehr, L., Fioletov, V., Hayden, K., Kharol, S. K., Li, S.-M., Makar, P., Martin, R. V., Mihele, C., Mittermeier, R. L., Krotkov, N., Sneep, M., Lamsal, L. N., Linden, M. ter, Geffen, J. van, Veefkind, P., and Wolde, M.: High-Resolution Mapping of Nitrogen Dioxide With TROPOMI: First Results and Validation Over the Canadian Oil Sands, Geophysical Research Letters, 46, 1049–1060, https://doi.org/10.1029/2018GL081095, 2019.

Gu, J., Liang, X., Song, S., Li, Y., Chen, L., Tao, J., and Tian, Y.: Harmonizing satellite and ground NO2 observations in China: A multi-sensor framework for scenario-specific calibration, ISPRS Journal of Photogrammetry and Remote Sensing, 230, 486–494, https://doi.org/10.1016/j.isprsjprs.2025.09.028, 2025.

Handschuh, J., Baier, F., Molina García, V., Friedl, P., and Loyola, D.: TROPOMI-Based PM2.5 Estimates and Their Evaluation During a High-Pollution Event in Germany, Remote Sensing, 18, 562, https://doi.org/10.3390/rs18040562, 2026.

Hassinen, S., Balis, D., Bauer, H., Begoin, M., Delcloo, A., Eleftheratos, K., Gimeno Garcia, S., Granville, J., Grossi, M., Hao, N., Hedelt, P., Hendrick, F., Hess, M., Heue, K.-P., Hovila, J., Jønch-Sørensen, H., Kalakoski, N., Kauppi, A., Kiemle, S., Kins, L., Koukouli, M. E., Kujanpää, J., Lambert, J.-C., Lang, R., Lerot, C., Loyola, D., Pedergnana, M., Pinardi, G., Romahn, F., van Roozendael, M., Lutz, R., De Smedt, I., Stammes, P., Steinbrecht, W., Tamminen, J., Theys, N., Tilstra, L. G., Tuinder, O. N. E., Valks, P., Zerefos, C., Zimmer, W., and Zyrichidou, I.: Overview of the O3M SAF GOME-2 operational atmospheric composition and UV radiation data products and data availability, Atmospheric Measurement Techniques, 9, 383–407, https://doi.org/10.5194/amt-9-383-2016, 2016.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition, https://doi.org/10.48550/arXiv.1512.03385, 10 December 2015., 2015.

He, Q., Qin, K., Cohen, J. B., Loyola, D., Li, D., Shi, J., and Xue, Y.: Spatially and temporally coherent reconstruction of tropospheric NO_2 over China combining OMI and GOME-2B measurements, Environ. Res. Lett., 15, 125011, https://doi.org/10.1088/1748-9326/abc7df, 2020.

He, Q., Qin, K., Cohen, J. B., Li, D., and Kim, J.: Quantifying uncertainty in ML-derived atmosphere remote sensing: Hourly surface NO2 estimation with GEMS, Geophysical Research Letters, 51, e2024GL110468, https://doi.org/10.1029/2024GL110468, 2024a.

He, Q., Qin, K., Li, X., Lu, L., Wong, M. S., and Cohen, J. B.: Diurnal NOx emission underestimation constrained using overlapping TROPOMI swaths, Atmospheric Environment, 358, 121354, https://doi.org/10.1016/j.atmosenv.2025.121354, 2025.

He, Z., Fan, G., Li, X., Gong, F.-Y., Liang, M., Gao, L., and Zhou, M.: Spatio-temporal modeling of satellite-observed CO2 columns in China using deep learning, Int. J. Appl. Earth Obs. Geoinf., 129, 103859, https://doi.org/10.1016/j.jag.2024.103859, 2024b.

Hengl, T., Heuvelink, G. B. M., and Rossiter, D. G.: About regression-kriging: From equations to case studies, Comput. Geosci., 33, 1301–1315, https://doi.org/10.1016/j.cageo.2007.05.001, 2007.

Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M., and Abuhassan, N.: NO2 column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/10.1029/2009JD011848, 2009.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., and Schepers, D.: The ERA5 global reanalysis, Quarterly journal of the royal meteorological society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hoek, G., Beelen, R., De Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., and Briggs, D.: A review of land-use regression models to assess spatial variation of outdoor air pollution, Atmos. Environ., 42, 7561–7578, https://doi.org/10.1016/j.atmosenv.2008.05.057, 2008.

Huber, D. E., Kerr, G. H., Nawaz, M. O., Runkel, S., Anenberg, S. C., and Goldberg, D. L.: Global NO_2 changes between 2019 and 2024 as observed by TROPOMI in urban areas and emerging hotspots, Atmospheric Chemistry and Physics, 26, 3783–3803, https://doi.org/10.5194/acp-26-3783-2026, 2026.

Ialongo, I., Virta, H., Eskes, H., Hovila, J., and Douros, J.: Comparison of TROPOMI/Sentinel-5 Precursor NO_2 observations with ground-based measurements in Helsinki, Atmospheric Measurement Techniques, 13, 205–218, https://doi.org/10.5194/amt-13-205-2020, 2020.

Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H., and Courrèges-Lacoste, G. B.: Requirements for the GMES atmosphere service and ESA's implementation concept: Sentinels-4/-5 and -5p, Remote Sensing of Environment, 120, 58–69, https://doi.org/10.1016/j.rse.2012.01.023, 2012.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Jethva, H., Chand, D., Torres, O., Gupta, P., Lyapustin, A., and Patadia, F.: Agricultural Burning and Air Quality over Northern India: A Synergistic Analysis using NASA's A-train Satellite Data and Ground Measurements, Aerosol Air Qual. Res., 18, 1756–1773, https://doi.org/10.4209/aaqr.2017.12.0583, 2018.

Jiang, Y., Gao, Z., He, J., Wu, J., and Christakos, G.: Application and Analysis of XCO2 Data from OCO Satellite Using a Synthetic DINEOF–BME Spatiotemporal Interpolation Framework, Remote Sens., 14, 4422, https://doi.org/10.3390/rs14174422, 2022.

Kampa, M. and Castanas, E.: Human health effects of air pollution, Environ. Pollut., 151, 362–367, https://doi.org/10.1016/j.envpol.2007.06.012, 2008.

Kang, Y., Choi, H., Im, J., Park, S., Shin, M., Song, C.-K., and Kim, S.: Estimation of surface-level NO2 and O3 concentrations using TROPOMI data and machine learning over East Asia, Environ. Pollut., 288, 117711, https://doi.org/10.1016/j.envpol.2021.117711, 2021.

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., Jeong, M.-J., Park, S. K., Lee, K.-M., Song, C.-K., Kim, S.-W., Kim, Y. J., Kim, S.-W., Kim, M., Go, S., Liu, X., Chance, K., Miller, C. C., Al-Saadi, J., Veihelmann, B., Bhartia, P. K., Torres, O., Abad, G. G., Haffner, D. P., Ko, D. H., Lee, S. H., Woo, J.-H., Chong, H., Park, S. S., Nicks, D., Choi, W. J., Moon, K.-J., Cho, A., Yoon, J., Kim, S., Hong, H., Lee, K., Lee, H., Lee, S., Choi, M., Veefkind, P., Levelt, P. F., Edwards, D. P., Kang, M., Eo, M., Bak, J., Baek, K., Kwon, H.-A., Yang, J., Park, J., Han, K. M., Kim, B.-R., Shin, H.-W., Choi, H., Lee, E., Chong, J., Cha, Y., Koo, J.-H., Irie, H., Hayashida, S., Kasai, Y., Kanaya, Y., Liu, C., Lin, J., Crawford, J. H., Carmichael, G. R., Newchurch, M. J., Lefer, B. L., Herman, J. R., Swap, R. J., Lau, A. K. H., Kurosu, T. P., Jaross, G., Ahlers, B., Dobber, M., McElroy, C. T., and Choi, Y.: New era of air quality monitoring from space: Geostationary environment monitoring spectrometer (GEMS), https://doi.org/10.1175/BAMS-D-18-0013.1, 2020a.

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., Jeong, M.-J., Park, S. K., Lee, K.-M., Song, C.-K., Kim, S.-W., Kim, Y. J., Kim, S.-W., Kim, M., Go, S., Liu, X., Chance, K., Miller, C. C., Al-Saadi, J., Veihelmann, B., Bhartia, P. K., Torres, O., Abad, G. G., Haffner, D. P., Ko, D. H., Lee, S. H., Woo, J.-H., Chong, H., Park, S. S., Nicks, D., Choi, W. J., Moon, K.-J., Cho, A., Yoon, J., Kim, S., Hong, H., Lee, K., Lee, H., Lee, S., Choi, M., Veefkind, P., Levelt, P. F., Edwards, D. P., Kang, M., Eo, M., Bak, J., Baek, K., Kwon, H.-A., Yang, J., Park, J., Han, K. M., Kim, B.-R., Shin, H.-W., Choi, H., Lee, E., Chong, J., Cha, Y., Koo, J.-H., Irie, H., Hayashida, S., Kasai, Y., Kanaya, Y., Liu, C., Lin, J., Crawford, J. H., Carmichael, G. R., Newchurch, M. J., Lefer, B. L., Herman, J. R., Swap, R. J., Lau, A. K. H., Kurosu, T. P., Jaross, G., Ahlers, B., Dobber, M., McElroy, C. T., and Choi, Y.: New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS), https://doi.org/10.1175/BAMS-D-18-0013.1, 2020b.

Krotkov, N. A., McLinden, C. A., Li, C., Lamsal, L. N., Celarier, E. A., Marchenko, S. V., Swartz, W. H., Bucsela, E. J., Joiner, J., Duncan, B. N., Boersma, K. F., Veefkind, J. P., Levelt, P. F., Fioletov, V. E., Dickerson, R. R., He, H., Lu, Z., and Streets, D. G.: Aura OMI observations of regional SO_2 and NO_2 pollution changes from 2005 to 2015, Atmospheric Chemistry and Physics, 16, 4605–4629, https://doi.org/10.5194/acp-16-4605-2016, 2016.

Lamsal, L. N., Martin, R. V., Parrish, D. D., and Krotkov, N. A.: Scaling Relationship for NO_2 Pollution and Urban Population Size: A Satellite Perspective, Environ. Sci. Technol., 47, 7855–7861, https://doi.org/10.1021/es400744g, 2013a.

Lamsal, L. N., Martin, R. V., Parrish, D. D., and Krotkov, N. A.: Scaling Relationship for NO2 Pollution and Urban Population Size: A Satellite Perspective, Environ. Sci. Technol., 47, 7855–7861, https://doi.org/10.1021/es400744g, 2013b.

LeCun, Y. A., Bottou, L., Orr, G. B., and Müller, K.-R.: Efficient BackProp, in: Neural Networks: Tricks of the Trade: Second Edition, edited by: Montavon, G., Orr, G. B., and Müller, K.-R., Springer, Berlin, Heidelberg, 9–48, https://doi.org/10.1007/978-3-642-35289-8_3, 2012.

Levelt, P. F., Hilsenrath, E., Leppelmeier, G. W., van den Oord, G. H. J., Bhartia, P. K., Tamminen, J., de Haan, J. F., and Veefkind, J. P.: Science objectives of the ozone monitoring instrument, IEEE Transactions on Geoscience and Remote Sensing, 44, 1199–1208, https://doi.org/10.1109/TGRS.2006.872336, 2006a.

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., Vries, J. de, Stammes, P., Lundell, J. O. V., and Saari, H.: The ozone monitoring instrument, IEEE Transactions on Geoscience and Remote Sensing, 44, 1093–1101, https://doi.org/10.1109/TGRS.2006.872333, 2006b.

Li, L. and Wu, J.: Spatiotemporal estimation of satellite-borne and ground-level NO2 using full residual deep networks, Remote Sensing of Environment, 254, 112257, https://doi.org/10.1016/j.rse.2020.112257, 2021.

Li, T., Wu, J., and Wang, T.: Generating daily high-resolution and full-coverage XCO2 across China from 2015 to 2020 based on OCO-2 and CAMS data, Science of The Total Environment, 893, 164921, https://doi.org/10.1016/j.scitotenv.2023.164921, 2023.

Li, T., Wu, J., Wang, Y., and Su, Y.: Improved seamless mapping of surface O3 concentrations using an integrated deep learning framework, Npj Clim. Atmos. Sci., 8, 124, https://doi.org/10.1038/s41612-025-01007-x, 2025.

Liu, G., Shih, K. J., Wang, T.-C., Reda, F. A., Sapra, K., Yu, Z., Tao, A., and Catanzaro, B.: Partial convolution based padding, https://doi.org/10.48550/arXiv.1811.11718, 28 November 2018., 2018.

Liu, S., Valks, P., Pinardi, G., De Smedt, I., Yu, H., Beirle, S., and Richter, A.: An improved total and tropospheric NO_2 column retrieval for GOME-2, Atmospheric Measurement Techniques, 12, 1029–1057, https://doi.org/10.5194/amt-12-1029-2019, 2019.

Liu, S., Valks, P., Pinardi, G., Xu, J., Chan, K. L., Argyrouli, A., Lutz, R., Beirle, S., Khorsandi, E., Baier, F., Huijnen, V., Bais, A., Donner, S., Dörner, S., Gratsea, M., Hendrick, F., Karagkiozidis, D., Lange, K., Piters, A. J. M., Remmers, J., Richter, A., Van Roozendael, M., Wagner, T., Wenig, M., and Loyola, D. G.: An improved TROPOMI tropospheric NO_2 research product over Europe, Atmospheric Measurement Techniques, 14, 7297–7327, https://doi.org/10.5194/amt-14-7297-2021, 2021a.

Liu, S., Valks, P., Beirle, S., and Loyola, D. G.: Nitrogen dioxide decline and rebound observed by GOME-2 and TROPOMI during COVID-19 pandemic, Air Qual Atmos Health, 14, 1737–1755, https://doi.org/10.1007/s11869-021-01046-2, 2021b.

Liu, X., Ran, L., Lin, W., Xu, X., Ma, Z., Dong, F., He, D., Zhou, L., Shi, Q., and Wang, Y.: Measurement report: Variations in surface SO_2 and NO_x mixing ratios from 2004 to 2016 at a background site in the North China Plain, Atmospheric Chemistry and Physics, 22, 7071–7085, https://doi.org/10.5194/acp-22-7071-2022, 2022.

Liu, Y., Zhao, J., Song, K., Cheng, C., Li, S., and Cai, K.: Spatiotempo ral evolution analysis of NO2 column density before and after COVID-19 pandemic in henan province based on SI-APSTE model, Sci Rep, 11, 18614, https://doi.org/10.1038/s41598-021-97745-y, 2021c.

Loyola, D.: Applications of neural network methods to the processing of earth observation satellite data, Neural Networks, 19, 168–177, https://doi.org/10.1016/j.neunet.2006.01.010, 2006.

Loyola, D., Gimeno García, S., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J. D., Pedergnana, M., Doicu, A., Molina García, V., and Schüssler, O.: The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor, Atmospheric Measurement Techniques, 11, 409–427, https://doi.org/10.5194/amt-11-409-2018, 2018.

Ma, Y., Liu, T., Yu, Z., Jiang, C., Xu, G., and Lu, Z.: All-weather precipitable water vapor map reconstruction using data fusion and machine learning-based spatial downscaling, Atmos. Res., 296, 107068, https://doi.org/10.1016/j.atmosres.2023.107068, 2023.

MacFerrin, M., Amante, C., Carignan, K., Love, M., and Lim, E.: The Earth Topography 2022 (ETOPO 2022) global DEM dataset, Earth System Science Data, 17, 1835–1849, https://doi.org/10.5194/essd-17-1835-2025, 2025.

Meng, X., Chen, L., Cai, J., Zou, B., Wu, C.-F., Fu, Q., Zhang, Y., Liu, Y., and Kan, H.: A land use regression model for estimating the NO2 concentration in shanghai, China, Environ. Res., 137, 308–315, https://doi.org/10.1016/j.envres.2015.01.003, 2015.

Mentel, T. F., Bleilebens, D., and Wahner, A.: A study of nighttime nitrogen oxide oxidation in a large reaction chamber—the fate of NO2, N2O5, HNO3, and O3 at different humidities, Atmospheric Environment, 30, 4007–4020, https://doi.org/10.1016/1352-2310(96)00117-3, 1996.

Mols, A., Boersma, K. F., Denier van der Gon, H., and Krol, M.: An improved bayesian inversion to estimate daily NO_x emissions of paris from TROPOMI NO_2 observations between 2018–2023, Atmospheric Chemistry and Physics, 26, 1497–1513, https://doi.org/10.5194/acp-26-1497-2026, 2026.

Monks, P. S., Granier, C., Fuzzi, S., Stohl, A., Williams, M. L., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., Bey, I., Blake, N., Blake, R. S., Carslaw, K., Cooper, O. R., Dentener, F., Fowler, D., Fragkou, E., Frost, G. J., Generoso, S., Ginoux, P., Grewe, V., Guenther, A., Hansson, H. C., Henne, S., Hjorth, J., Hofzumahaus, A., Huntrieser, H., Isaksen, I. S. A., Jenkin, M. E., Kaiser, J., Kanakidou, M., Klimont, Z., Kulmala, M., Laj, P., Lawrence, M. G., Lee, J. D., Liousse, C., Maione, M., McFiggans, G., Metzger, A., Mieville, A., Moussiopoulos, N., Orlando, J. J., O'Dowd, C. D., Palmer, P. I., Parrish, D. D., Petzold, A., Platt, U., Pöschl, U., Prévôt, A. S. H., Reeves, C. E., Reimann, S., Rudich, Y., Sellegri, K., Steinbrecher, R., Simpson, D., Ten Brink, H., Theloke, J., Van Der Werf, G. R., Vautard, R., Vestreng, V., Vlachokostas, Ch., and Von Glasow, R.: Atmospheric composition change – global and regional air quality, Atmos. Environ., 43, 5268–5350, https://doi.org/10.1016/j.atmosenv.2009.08.021, 2009.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview, Atmospheric Measurement Techniques, 9, 1279–1301, https://doi.org/10.5194/amt-9-1279-2016, 2016.

NOAA National Centers for Environmental Information: ETOPO 2022 15 arc-second global relief model, https://doi.org/10.25921/FD45-GT74, 2022.

Noguchi, K., Richter, A., Rozanov, V., Rozanov, A., Burrows, J. P., Irie, H., and Kita, K.: Effect of surface BRDF of various land cover types on geostationary observations of tropospheric NO_2, Atmospheric Measurement Techniques, 7, 3497–3508, https://doi.org/10.5194/amt-7-3497-2014, 2014.

Park, J.-U., Lim, S., Hanisco, T. F., Abuhassan, N., Place, B. K., Pandey, A., Cede, A., Tiefengraber, M., Gebetsberger, M., Park, J., Choi, J., Crawford, J. H., Song, C.-K., and Kim, S.-W.: Global analysis of nitrogen dioxide and formaldehyde column densities from the Pandora global network: Variability and implications for satellite validation, Remote Sensing of Environment, 335, 115249, https://doi.org/10.1016/j.rse.2026.115249, 2026.

Park, T., Liu, M.-Y., Wang, T.-C., and Zhu, J.-Y.: Semantic image synthesis with spatially-adaptive normalization, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2332–2341, https://doi.org/10.1109/CVPR.2019.00244, 2019.

Peng, Z., Zhang, B., Wang, D., Niu, X., Sun, J., Xu, H., Cao, J., and Shen, Z.: Application of machine learning in atmospheric pollution research: A state-of-art review, Science of The Total Environment, 910, 168588, https://doi.org/10.1016/j.scitotenv.2023.168588, 2024.

Pinardi, G., Lambert, J. C., Granville, J., Van Roozendael, M., Delcloo, A., De Backer, H., Valks, P., and Hao, N.: Overview of the validation of GOME-2 total and tropospheric NO2 columns, in: EUMETSAT meteorological satellite conference, 2010.

Purchase, M. L., Bending, G. D., and Mushinski, R. M.: Spatiotemporal Variations of Soil Reactive Nitrogen Oxide Fluxes across the Anthropogenic Landscape, Environ. Sci. Technol., 57, 16348–16360, https://doi.org/10.1021/acs.est.3c05849, 2023.

Qin, K., Rao, L., Xu, J., Bai, Y., Zou, J., Hao, N., Li, S., and Yu, C.: Estimating Ground Level NO2 Concentrations over Central-Eastern China Using a Satellite-Based Geographically and Temporally Weighted Regression Model, Remote Sens., 9, 950, https://doi.org/10.3390/rs9090950, 2017.

Qin, K., Gao, H., Liu, X., He, Q., Tiwari, P., and Cohen, J. B.: The global daily High Spatial–Temporal Coverage Merged tropospheric NO_2 dataset (HSTCM-NO_2) from 2007 to 2022 based on OMI and GOME-2, Earth Syst. Sci. Data, 16, 5287–5310, https://doi.org/10.5194/essd-16-5287-2024, 2024.

Qu, Y., Wei, J., Xing, H., Shi, X., Ao, Z., and Meng, X.: Global Estimates of Daily Gapless Atmospheric XCH_4 Concentrations From Satellite and Reanalysis Data During 2003–2020, IEEE Trans. Geosci. Remote Sens., 63, 1–12, https://doi.org/10.1109/TGRS.2025.3593486, 2025.

Rao, L., Xu, J., Efremenko, D. S., Loyola, D. G., and Doicu, A.: Aerosol Parameters Retrieval From TROPOMI/S5P Using Physics-Based Neural Networks, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 6473–6484, https://doi.org/10.1109/JSTARS.2022.3196843, 2022.

Ravindra, K., Mor, S., Ameena, Kamyotra, J. S., and Kaushik, C. P.: Variation in Spatial Pattern of Criteria Air Pollutants Before and During Initial Rain of Monsoon, Environ Monit Assess, 87, 145–153, https://doi.org/10.1023/A:1024650215970, 2003.

Reda, I. and Andreas, A.: Solar position algorithm for solar radiation applications, Solar Energy, 76, 577–589, https://doi.org/10.1016/j.solener.2003.12.003, 2004.

Richter, A., Burrows, J. P., Nüß, H., Granier, C., and Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, Nature, 437, 129–132, https://doi.org/10.1038/nature04092, 2005.

Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, https://doi.org/10.48550/arXiv.1505.04597, 18 May 2015., 2015.

Russell, A. R., Valin, L. C., and Cohen, R. C.: Trends in OMI NO_2 observations over the united states: Effects of emission control technology and the economic recession, Atmospheric Chemistry and Physics, 12, 12197–12209, https://doi.org/10.5194/acp-12-12197-2012, 2012.

Safieddine, S., Clerbaux, C., George, M., Hadji-Lazaro, J., Hurtmans, D., Coheur, P.-F., Wespes, C., Loyola, D., Valks, P., and Hao, N.: Tropospheric ozone and nitrogen dioxide measurements in urban and rural regions as seen by IASI and GOME-2, Journal of Geophysical Research: Atmospheres, 118, 10,555-10,566, https://doi.org/10.1002/jgrd.50669, 2013.

Samoli, E., Aga, E., Touloumi, G., Nisiotis, K., Forsberg, B., Lefranc, A., Pekkanen, J., Wojtyniak, B., Schindler, C., Niciu, E., Brunstein, R., Dodič Fikfak, M., Schwartz, J., and Katsouyanni, K.: Short-term effects of nitrogen dioxide on mortality: an analysis within the APHEA project, Eur. Respir. J., 27, 1129–1138, https://doi.org/10.1183/09031936.06.00143905, 2006.

Scheibenreif, L., Mommert, M., and Borth, D.: Toward Global Estimation of Ground-Level NO_2 Pollution With Deep Learning and Remote Sensing, IEEE Trans. Geosci. Remote Sens., 60, 1–14, https://doi.org/10.1109/TGRS.2022.3160827, 2022.

Seo, S., Valks, P., Lutz, R., Heue, K.-P., Hedelt, P., Molina García, V., Loyola, D., Lee, H., and Kim, J.: Tropospheric NO_2 retrieval algorithm for geostationary satellite instruments: applications to GEMS, Atmospheric Measurement Techniques, 17, 6163–6191, https://doi.org/10.5194/amt-17-6163-2024, 2024.

Sfica, L., Beck, C., Nita, A.-I., Voiculescu, M., Birsan, M.-V., and Philipp, A.: Cloud cover changes driven by atmospheric circulation in europe during the last decades, International Journal of Meteorology, 41, 221–2230, 2021.

Shao, Y., Zhao, W., Liu, R., Yang, J., Liu, M., Fang, W., Hu, L., Adams, M., Bi, J., and Ma, Z.: Estimation of daily NO2 with explainable machine learning model in China, 2007–2020, Atmos. Environ., 314, 120111, https://doi.org/10.1016/j.atmosenv.2023.120111, 2023.

She, L., Zhang, H. K., Li, Z., de Leeuw, G., and Huang, B.: Himawari-8 aerosol optical depth (AOD) retrieval using a deep neural network trained using AERONET observations, Remote Sensing, 12, 4125, https://doi.org/10.3390/rs12244125, 2020.

She, L., Li, Z., De Leeuw, G., Wang, W., Wang, Y., Yang, L., Feng, Z., Yang, C., and Shi, Y.: Time series retrieval of multi-wavelength aerosol optical depth by adapting transformer (TMAT) using himawari-8 AHI data, Remote Sensing of Environment, 305, 114115, https://doi.org/10.1016/j.rse.2024.114115, 2024.

Song, W., Jia, H., Li, Z., Tang, D., and Wang, C.: Detecting urban land-use configuration effects on NO2 and NO variations using geographically weighted land use regression, Atmospheric Environment, 197, 166–176, https://doi.org/10.1016/j.atmosenv.2018.10.031, 2019.

Stull, R. B.: An introduction to boundary layer meteorology, Springer Science & Business Media, 2012.

Sun, Y., Zeng, Q., Geng, B., Lin, X., Sude, B., and Chen, L.: Deep Learning Architecture for Estimating Hourly Ground-Level $PM_{2.5}$ Using Satellite Remote Sensing, IEEE Geosci. Remote Sens. Lett., 16, 1343–1347, https://doi.org/10.1109/LGRS.2019.2900270, 2019.

Sun, Y., Niu, T., He, J., Ma, Z., Liu, P., Xiao, D., Hu, J., Yang, J., and Yan, X.: Classification of circulation patterns during the formation and dissipation of continuous pollution weather over the sichuan basin, China, Atmospheric Environment, 223, 117244, https://doi.org/10.1016/j.atmosenv.2019.117244, 2020.

Tiwari, P., Cohen, J. B., Gao, H., Lu, L., Wang, J., Dubovik, O., and Qin, K.: Microphysical evolution and column loading drive nonlinear regional contrast in black carbon top-of-atmosphere forcing, Atmospheric Chemistry and Physics, 26, 9149–9180, https://doi.org/10.5194/acp-26-9149-2026, 2026.

Valks, P., Pinardi, G., Richter, A., Lambert, J.-C., Hao, N., Loyola, D., Van Roozendael, M., and Emmadi, S.: Operational total and tropospheric NO_2 column retrieval for GOME-2, Atmospheric Measurement Techniques, 4, 1491–1514, https://doi.org/10.5194/amt-4-1491-2011, 2011.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., 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 Sensing of Environment, 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012a.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., 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 Sensing of Environment, 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012b.

Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J.-C., Eskes, H. J., Eichmann, K.-U., Fjæraa, A. M., Granville, J., Niemeijer, S., Cede, A., Tiefengraber, M., Hendrick, F., Pazmiño, A., Bais, A., Bazureau, A., Boersma, K. F., Bognar, K., Dehn, A., Donner, S., Elokhov, A., Gebetsberger, M., Goutail, F., Grutter de la Mora, M., Gruzdev, A., Gratsea, M., Hansen, G. H., Irie, H., Jepsen, N., Kanaya, Y., Karagkiozidis, D., Kivi, R., Kreher, K., Levelt, P. F., Liu, C., Müller, M., Navarro Comas, M., Piters, A. J. M., Pommereau, J.-P., Portafaix, T., Prados-Roman, C., Puentedura, O., Querel, R., Remmers, J., Richter, A., Rimmer, J., Rivera Cárdenas, C., Saavedra de Miguel, L., Sinyakov, V. P., Stremme, W., Strong, K., Van Roozendael, M., Veefkind, J. P., Wagner, T., Wittrock, F., Yela González, M., and Zehner, C.: Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO_2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks, Atmospheric Measurement Techniques, 14, 481–510, https://doi.org/10.5194/amt-14-481-2021, 2021.

Wang, J.: Global daily 1 km gapless XCO_2 (2003−2023) derived from multi-satellite observations and a spatiotemporal deep learning framework, Environ. Impact Assess. Rev., 117, 108146, https://doi.org/10.1016/j.eiar.2025.108146, 2026.

Wang, J., Li, J., Ye, J., Zhao, J., Wu, Y., Hu, J., Liu, D., Nie, D., Shen, F., and Huang, X.: Fast sulfate formation from oxidation of SO2 by NO2 and HONO observed in beijing haze, Nature Communications, 11, 2844, 2020.

Wang, M., Duan, Y., Zhang, Z., Yuan, Q., Li, X., Han, S., Huo, J., Chen, J., Lin, Y., Fu, Q., Wang, T., Cao, J., and Lee, S.: Reduction in vehicular emissions attributable to the Covid-19 lockdown in Shanghai: insights from 5 years of monitoring-based machine learning, Atmos. Chem. Phys., 23, 10313–10324, https://doi.org/10.5194/acp-23-10313-2023, 2023.

Wang, S., Cohen, J. B., Guan, L., Lu, L., Tiwari, P., and Qin, K.: Observationally constrained global NOx and CO emissions variability reveals sources which contribute significantly to CO2 emissions, npj Clim Atmos Sci, 8, 87, https://doi.org/10.1038/s41612-025-00977-2, 2025.

Wang, S., Cohen, J. B., Qin, K., Guan, L., Tiwari, P., Feng, Y., Li, X., He, Q., Li, Z., Fan, C., and Wang, D.: Synergistic analysis and vertical transport attribution of carbon monoxide surface concentrations and satellite column loadings over China, Remote Sensing of Environment, 344, 115562, https://doi.org/10.1016/j.rse.2026.115562, 2026.

Wang, Y., Yuan, Q., Li, T., Zhu, L., and Zhang, L.: Estimating daily full-coverage near surface O3, CO, and NO2 concentrations at a high spatial resolution over China based on S5P-TROPOMI and GEOS-FP, Isprs J. Photogramm. Remote Sens., 175, 311–325, https://doi.org/10.1016/j.isprsjprs.2021.03.018, 2021.

van Weele, M. and Duynkerke, P. G.: Effect of clouds on the photodissociation of NO2: Observations and modelling, Journal of atmospheric chemistry, 16, 231–255, https://doi.org/10.1007/BF00696898, 1993.

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., and Cribb, M.: Reconstructing 1-km-resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications, Remote Sensing of Environment, 252, 112136, https://doi.org/10.1016/j.rse.2020.112136, 2021.

Wei, J., Liu, S., Li, Z., Liu, C., Qin, K., Liu, X., Pinker, R. T., Dickerson, R. R., Lin, J., Boersma, K. F., Sun, L., Li, R., Xue, W., Cui, Y., Zhang, C., and Wang, J.: Ground-level NO2 surveillance from space across China for high resolution using interpretable spatiotemporally weighted artificial intelligence, Environ. Sci. Technol., 56, 9988–9998, https://doi.org/10.1021/acs.est.2c03834, 2022.

Wei, W., Bao, Z., Su, Y., Wang, P., and Bai, S.: Research of rapid retrieving models for GOSAT XCO2 based on machine learning, Comput. Geosci., 29, 35, https://doi.org/10.1007/s10596-025-10376-4, 2025.

Xing, J., Baek, B. H., Li, S., Wang, C.-T., Song, G., Ma, S., Zheng, S., Liu, C., Tong, D., Woo, J.-H., Liu, T.-Y., and Fu, J. S.: A Physically Constrained Deep-Learning Fusion Method for Estimating Surface NO_2 Concentration from Satellite and Ground Monitors, Environ. Sci. Technol., 58, 21218–21228, https://doi.org/10.1021/acs.est.4c07341, 2024.

Xu, J., Schüssler, O., Rodriguez, D. G. L., Romahn, F., and Doicu, A.: A Novel Ozone Profile Shape Retrieval Using Full-Physics Inverse Learning Machine (FP-ILM), IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10, 5442–5457, https://doi.org/10.1109/JSTARS.2017.2740168, 2017.

Xu, W. Y., Zhao, C. S., Ran, L., Deng, Z. Z., Liu, P. F., Ma, N., Lin, W. L., Xu, X. B., Yan, P., He, X., Yu, J., Liang, W. D., and Chen, L. L.: Characteristics of pollutants and their correlation to meteorological conditions at a suburban site in the North China Plain, Atmospheric Chemistry and Physics, 11, 4353–4369, https://doi.org/10.5194/acp-11-4353-2011, 2011.

Yang, L. H., Jacob, D. J., Dang, R., Oak, Y. J., Lin, H., Kim, J., Zhai, S., Colombi, N. K., Pendergrass, D. C., Beaudry, E., Shah, V., Feng, X., Yantosca, R. M., Chong, H., Park, J., Lee, H., Lee, W.-J., Kim, S., Kim, E., Travis, K. R., Crawford, J. H., and Liao, H.: Interpreting Geostationary Environment Monitoring Spectrometer (GEMS) geostationary satellite observations of the diurnal variation in nitrogen dioxide (NO_2) over East Asia, Atmospheric Chemistry and Physics, 24, 7027–7039, https://doi.org/10.5194/acp-24-7027-2024, 2024.

Yu, M. and Liu, Q.: Deep learning-based downscaling of tropospheric nitrogen dioxide using ground-level and satellite observations, Science of The Total Environment, 773, 145145, https://doi.org/10.1016/j.scitotenv.2021.145145, 2021.

Yuan, L., Han, W., Meng, J., Wang, Y., Yu, H., and Li, W.: Uncovering the impact of urban functional zones on air quality in China, Atmospheric Chemistry and Physics, 25, 10421–10442, https://doi.org/10.5194/acp-25-10421-2025, 2025.

Zammit-Mangion, A., Cressie, N., and Shumack, C.: On Statistical Approaches to Generate Level 3 Products from Satellite Remote Sensing Retrievals, Remote Sens., 10, 155, https://doi.org/10.3390/rs10010155, 2018.

Zhang, M., Li, S., Xing, J., Song, G., Li, S., Dong, J., Zheng, S., Han, G., and Yang, J.: Enhancing spatiotemporal coverage of satellite-derived high-resolution NO2 data with a super-resolution model, Remote Sens. Environ., 328, 114897, https://doi.org/10.1016/j.rse.2025.114897, 2025a.

Zhang, Q., Boersma, K. F., van der Laan, C., Mols, A., Zhao, B., Li, S., and Pan, Y.: Estimating the variability in NO_x emissions from wuhan with TROPOMI NO_2 data during 2018 to 2023, Atmospheric Chemistry and Physics, 25, 3313–3326, https://doi.org/10.5194/acp-25-3313-2025, 2025b.

Zhang, X. and Gao, J.: Measuring Feature Importance of Convolutional Neural Networks, IEEE Access, 8, 196062–196074, https://doi.org/10.1109/ACCESS.2020.3034625, 2020.

Zhou, Y., Brunner, D., Boersma, K. F., Dirksen, R., and Wang, P.: An improved tropospheric NO_2 retrieval for OMI observations in the vicinity of mountainous terrain, Atmospheric Measurement Techniques, 2, 401–416, https://doi.org/10.5194/amt-2-401-2009, 2009.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M. R., Arola, A., Baker, B. D., Canova, B. P., Chan Miller, C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, A., González Abad, G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., Kim, J., Krotkov, N. A., Lamsal, L., Li, C., Lindfors, A., Martin, R. V., McElroy, C. T., McLinden, C., Natraj, V., Neil, D. O., Nowlan, C. R., O'Sullivan, E. J., Palmer, P. I., Pierce, R. B., Pippin, M. R., Saiz-Lopez, A., Spurr, R. J. D., Szykman, J. J., Torres, O., Veefkind, J. P., Veihelmann, B., Wang, H., Wang, J., and Chance, K.: Tropospheric emissions: Monitoring of pollution (TEMPO), Journal of Quantitative Spectroscopy and Radiative Transfer, 186, 17–39, https://doi.org/10.1016/j.jqsrt.2016.05.008, 2017.