











Research Article

Direct Retrieval of NO₂ Vertical Columns from UV-Vis (390-495 nm) Spectral Radiances Using a Neural Network

Chi Li ¹, Xiaoguang Xu ^{2,3}, Xiong Liu ⁴, Jun Wang,³ Kang Sun ^{5,6}, Jos van Geffen ⁷, Qindan Zhu ⁸, Jianzhong Ma ⁹, Junli Jin ¹⁰, Kai Qin,¹¹ Qin He ¹¹, Pinhua Xie,¹² Bo Ren,¹² and Ronald C. Cohen ^{1,8}

¹Department of Chemistry, University of California, Berkeley, Berkeley, CA, USA

²Joint Center for Earth Systems Technology, University of Maryland, Baltimore County, Baltimore, MA, USA

³Department of Chemical and Biochemical Engineering, Center for Global and Regional Environmental Research, And Informatics Initiative, The University of Iowa, Iowa City, IA, USA

⁴Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

⁵Department of Civil, Structural and Environmental Engineering, University at Buffalo, Buffalo, NY, USA

⁶Research and Education in Energy, Environment and Water Institute, University at Buffalo, Buffalo, NY, USA

⁷Satellite Observations Department, Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands

⁸Department of Earth and Planetary Science, University of California, Berkeley, Berkeley, CA, USA

⁹Chinese Academy of Meteorology Science, China Meteorological Administration, Beijing, China

¹⁰Meteorological Observation Center, China Meteorological Administration, Beijing, China

¹¹School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou, China

¹²Key Laboratory of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, China

Correspondence should be addressed to Chi Li; lynchlee90@gmail.com and Ronald C. Cohen; rccohen@berkeley.edu

Received 24 October 2021; Accepted 10 April 2022; Published 2 May 2022

Copyright © 2022 Chi Li et al. Exclusive Licensee Aerospace Information Research Institute, Chinese Academy of Sciences. Distributed under a Creative Commons Attribution License (CC BY 4.0).

Satellite retrievals of columnar nitrogen dioxide (NO₂) are essential for the characterization of nitrogen oxides (NO_x) processes and impacts. The requirements of modeled *a priori* profiles present an outstanding bottleneck in operational satellite NO₂ retrievals. In this work, we instead use neural network (NN) models trained from over 360,000 radiative transfer (RT) simulations to translate satellite radiances across 390-495 nm to total NO₂ vertical column (NO₂C). Despite the wide variability of the many input parameters in the RT simulations, only a small number of key variables were found essential to the accurate prediction of NO₂C, including observing angles, surface reflectivity and altitude, and several key principal component scores of the radiances. In addition to the NO₂C, the NN training and cross-validation experiments show that the wider retrieval window allows some information about the vertical distribution to be retrieved (e.g., extending the rightmost wavelength from 465 to 495 nm decreases the root-mean-square-error by 0.75%) under high-NO₂C conditions. Applying to four months of TROPOMI data, the trained NN model shows strong ability to reproduce the NO₂C observed by the ground-based Pandora Global Network. The coefficient of determination (R^2 , 0.75) and normalized mean bias (NMB, -33%) are competitive with the level 2 operational TROPOMI product ($R^2 = 0.77$, NMB = -29%) over clear (geometric cloud fraction < 0.2) and polluted (NO₂C ≥ 7.5 × 10¹⁵ molecules/cm²) regions. The NN retrieval approach is ~12 times faster than predictions using high spatial resolution (~3 km) *a priori* profiles from chemical transport modeling, which is especially attractive to the handling of large volume satellite data.

1. Introduction

Nitrogen oxides (NO_x , comprising NO and NO_2) fuel the formation of secondary aerosols [1–3] and ozone [4–6], with broad implications for tropospheric composition, air quality, nitrogen deposition, and climate change [7–9]. Nitrogen dioxide (NO_2) itself is a major pollutant, with elevated concentration causing adverse respiratory diseases [10, 11]. Retrievals of NO_2 from satellite ultraviolet-visible (UV-Vis) spectral measurements are powerful tools for the characterization of the spatiotemporal variation of NO_x . Benchmarked by the Global Ozone Monitoring Experiment (GOME) instrument onboard the European Remote Sensing Satellite (ERS-2), global monitoring of tropospheric NO_2 has been possible since 1995 [12–14], which boosted the investigations of spatial variability and long-term changes of NO_2 abundance [15, 16], various NO_x emission from anthropogenic [17–21] and natural sources [22, 23], NO_x lifetime [24, 25], etc. Moreover, a global geostationary constellation of NO_2 monitoring is emerging [26–28], with hourly and km-scale monitoring capabilities to further facilitate the investigation of diurnal variability of NO_x -relevant processes.

With this promise, retrieval algorithms have been developed and continuously improved to translate the rapidly growing volume (terabytes per day) of earth observation data to physically relevant information (e.g., NO_2 abundance). The majority of existing tropospheric NO_2 retrievals include three steps: (1) NO_2 slant column density (SCD) determination from radiances through spectral fitting [29, 30]; (2) stratospheric-tropospheric component separation [31–33], and (3) air mass factor (AMF) and vertical column density (VCD) calculation [12, 34]. Large uncertainties exist in such retrievals as exemplified by the significant discrepancies among different products from the same sensor, e.g., the ozone monitoring instrument (OMI). Lamsal et al. [35] found that the standard OMI NO_2 product is 22% lower in winter but 42% higher in summer than the DOMINO product [36, 37] over North America, although both were derived from the same SCDs. Regional OMI research products such as the BEHR [38] over North America and POMINO [39] over East Asia also exhibit systematic differences against global operational products due mainly to their locally refined AMF calculations. The AMF [34, 38] relates the physically meaningful VCD with the optically represented SCD (i.e., the total NO_2 amount along all light beam paths received by the satellite instrument). The AMF is the largest source of VCD retrieval errors, containing a structural uncertainty estimated to be in the range of 30–50%, dependent on the representation and treatments of surface reflectivity, clouds, aerosols, and NO_2 vertical profile shape during the retrieval [13, 40, 41]. Higher-resolution inputs of these parameters in regional research products were found to alter the AMF and the retrieved VCD by up to 40–80%, with better resolved NO_2 spatial gradients and local enhancements [38, 39, 42–44]. Valin et al. [45] showed that a model resolution of <12 km is needed to resolve accurate NO_2 gradients (as well as the *a priori* profile) around area sources such as cities (~1 km for point sources such as power plants). But a typical trade-off to provide such high-resolution inputs is

the computational resources required, especially in terms of the *a priori* NO_2 vertical profiles which require the use of a chemical transport model.

Global operational products currently adopt profiles simulated at mesoscales (~1° or ~100 km), which were close to the field of view (FOV) of early sensors like GOME ($40 \times 320 \text{ km}^2$), but are now over 30 times larger than the nadir FOVs of contemporary instruments such as OMI ($13 \times 24 \text{ km}^2$) and the TROPospheric Monitoring Instrument (TROPOMI, $5.5 \times 3.5 \text{ km}^2$). With typical computation environments (e.g., 1 CPU with ~30 cores), the wall time needed to perform such global simulations for 1 month is 2–3 days, which scales as the square of the horizontal resolution [46], making long-term global simulation at local scale (~3 km) almost infeasible. Moreover, the simulated NO_2 profiles also suffer from errors in the model inputs (e.g., emission inventory) and assumptions (e.g., chemical mechanism). In addition, using climatological monthly model profiles was commonly adopted in operational retrievals [33, 44]. Laughner et al. [47] showed that the NO_2 VCDs derived from profiles simulated for the exact day could differ by up to 40% due mainly to day-to-day changes in meteorology. For another instance, lightning NO_x is critical for ozone production in the upper troposphere, while current model parameterization is not adequate to correctly simulate lightning strength [48]. Local VCD errors of up to 100% [49, 50] were reported due to simulated profiles with modeled lightning, limiting the consequent application of these retrievals to constrain lightning. Finally, the retrieved NO_2 VCDs also further impact inverse modeling and data assimilation. A recent study suggested that inconsistent modeled profiles used in the retrieval and in the assimilation alone could increase the *a posteriori* NO_x emission errors by up to 30% over polluted regions [51].

Satellite-observed radiances from nadir observations inherently contain information about the vertical distribution of species, provided that a broader spectral range is used, and the gaseous absorption is strong enough. The theoretical foundation of such vertical sensitivity based on radiative transfer (RT) was introduced in previous studies [52–55]. Briefly, the atmospheric scattering of molecules and aerosols decreases as a (up to -4th) power function of the wavelength in the UV-Vis range, making the sensitivity (e.g., weighting function or AMF) of spectral radiances to gas absorption also an inverse function of wavelength. Relative to higher altitudes in the atmosphere, gases at lower altitudes absorb the photons that are transmitted more strongly compared to scattering, yielding stronger spectral contrast of the AMF. This spectral contrast is more pronounced between more distant wavelengths, and is further enhanced by the temperature (i.e., altitude) dependency of NO_2 absorption strength across the UV-Vis [55]. Physics-based retrievals that exploit these underlying mechanisms have been applied to the retrievals of ozone [54] and SO_2 [52, 53] vertical distributions. For NO_2 , such physics-based approaches have also been developed [56] for its column retrieval; however, they still relied on preassumed profile shapes. A recent sensitivity study based on the optimal estimation (OE) framework [57] suggested that physical

retrieval of tropospheric NO₂ vertical distribution from satellite measurements is only possible under highly polluted conditions (e.g., boundary layer NO₂ column > 2 × 10¹⁶ molecules/cm²) where the information content (i.e., degree of freedom for signal) could reach 2. This low information content plus the high computational expense over a broader spectral range (i.e., 320–500 nm) discouraged attempts to perform purely OE-based determination of NO₂ VCD from satellite radiances. It is still unclear whether the dependence on *a priori* profiles of NO₂ retrievals could be loosened and whether a direct inference of NO₂ VCD from satellite observed radiance is achievable, amid the uncertainty of *a priori* profiles from models.

As an emerging and increasingly attractive tool to predict relevant parameters directly and efficiently from earth observation data [58], machine learning could be an alternative to further investigate this question. Although the interpretability of the prediction model is less than physics-based models, its computational efficiency is especially welcomed by the handling of dense satellite observations, and the model accuracy could be continuously improved by regular updates (iteration) including newly available training data. Capable of resembling underlying nonlinear relationship between relevant variables, the neural network (NN) approach has been applied widely to expedite forward RT modeling [59, 60] and inverse retrieval processes [61–64] by linking satellite signals directly with relevant geophysical parameters using the large volume of a training dataset.

In this paper, we describe the development of an NN-based retrieval approach to directly determine the total vertical column of NO₂ from UV-Vis radiances, using TROPOMI observations as a testbed. We show for the first time that a direct inference of NO₂ column without presimulated *a priori* NO₂ vertical profiles has similar data quality compared to operational global retrievals (i.e., the standard TROPOMI products), under clear-sky and polluted conditions (i.e., total NO₂ ≥ 7.5 × 10¹⁵ molecules/cm² and cloud fraction < 0.2). We built the NN model capabilities based on a synthetic radiance dataset that spans realistic clear-sky scenarios of observing conditions. Section 2 provides a detailed discussion of retrieval sensitivities. We then applied the NN prediction model to four months of TROPOMI data (Section 3) and performed a comparative and evaluation analysis with the standard products and ground-based measurements. Section 4 summarizes the discussion and points to various pathways for future research.

2. Neural Network-Based NO₂ Column Retrieval Model

The NN model (box 2 of Figure 1) to predict NO₂ column density from satellite radiances was generated from a large volume (>360,000 samples) of simulated spectra that span a wide range of realistic observing scenarios (Section 2.1 and box 1 of Figure 1). The model performance was evaluated via a cross-validation approach (Section 2.2 and box 3 of Figure 1).

2.1. RT Simulation. We used the UNified Linearized Vector Radiative Transfer Model (UNL-VRTM) [65, 66] to generate a synthetic dataset of radiances observed by a satellite instrument (i.e., TROPOMI) and the corresponding input variables (box 1 of Figure 1). UNL-VRTM facilitates a user-friendly interface for modifying surface and atmospheric optical parameters, which were fed to the Vector Linearized Discrete Ordinate Radiative Transfer (VLIDORT) [67] RT code to calculate the top of atmosphere (TOA) radiances. UNL-VRTM has been widely used for aerosol retrieval and relevant sensitivity studies based on both band-averaged and hyperspectral radiances [68–71]. To exploit the full information possible for NO₂ retrieval in the UV-Vis measurements, we simulated radiances between 320 and 500 nm, with an interval of 0.01 nm. The simulations were run in scalar-only mode to accelerate the calculation by an order of magnitude across >18,000 wavelengths. Polarization correction factors for each wavelength were interpolated from an additional vector-mode run at 21 wavelengths following [54]. The simulated radiances (I) were then convoluted with the TROPOMI spectral response function (SRF), normalized to solar irradiance (E), and then, recorded as log base values, to mimic the hyperspectral observations ($L_\lambda = \ln(I_\lambda/E_\lambda)$) at ~0.2 nm spectral resolution in conventional retrievals [29, 62, 72].

The input variables of UNL-VRTM in each simulation varied randomly (sampled from uniform distributions) within their realistic ranges according to Table 1. First, solar and satellite position angles were varied in the simulations: solar zenith angle (SZA), satellite/view zenith angle (VZA), and relative azimuth angle (RAA). Surface reflectance spectra (R_λ) were randomly obtained from a land climatological dataset as a combination of MERIS and OMI (Figure S1). The whole atmosphere from sea level to 80 km (0.01 hpa) was divided into 47 layers (30 layers below the assumed tropopause at 12 km) following the hybrid sigma-pressure vertical grid setting in the GEOS-Chem model [51], with the temperature and pressure in each layer consistent with the US standard (1976) atmosphere. The actual layer number in each simulation depended on the surface altitude (H_s) inputs (varied between 0 and 8 km). For aerosols, the size/microphysical properties (water-soluble aerosols from the OPAC dataset [73]) and vertical profile (exponentially decreasing with a scale height at 1 km) were fixed while the optical depth varies in each simulation. Apart from NO₂, vertical columns (G_{ot}) of four additional absorbing gases (SO₂, O₃, H₂O, and CH₂O) also varied with their fixed profile shapes [74] from the US standard (1976) atmosphere encoded in the UNL-VRTM. The O₂-O₂ absorption was also included and implicitly considered based on the pressure in each layer. Raman rotational scattering (Ring effect) [75] is not considered during the RT simulation for this wide spectrum, and we applied the Ring correction (Section 3.1) to ensure consistently Ring-free L_λ in both the simulation and observation.

The NO₂ vertical profile is modeled as composed of two parts: a stratospheric component with fixed vertical distribution following the US standard profile shape between 12 and 80 km, and a tropospheric component with a quasi-Gaussian

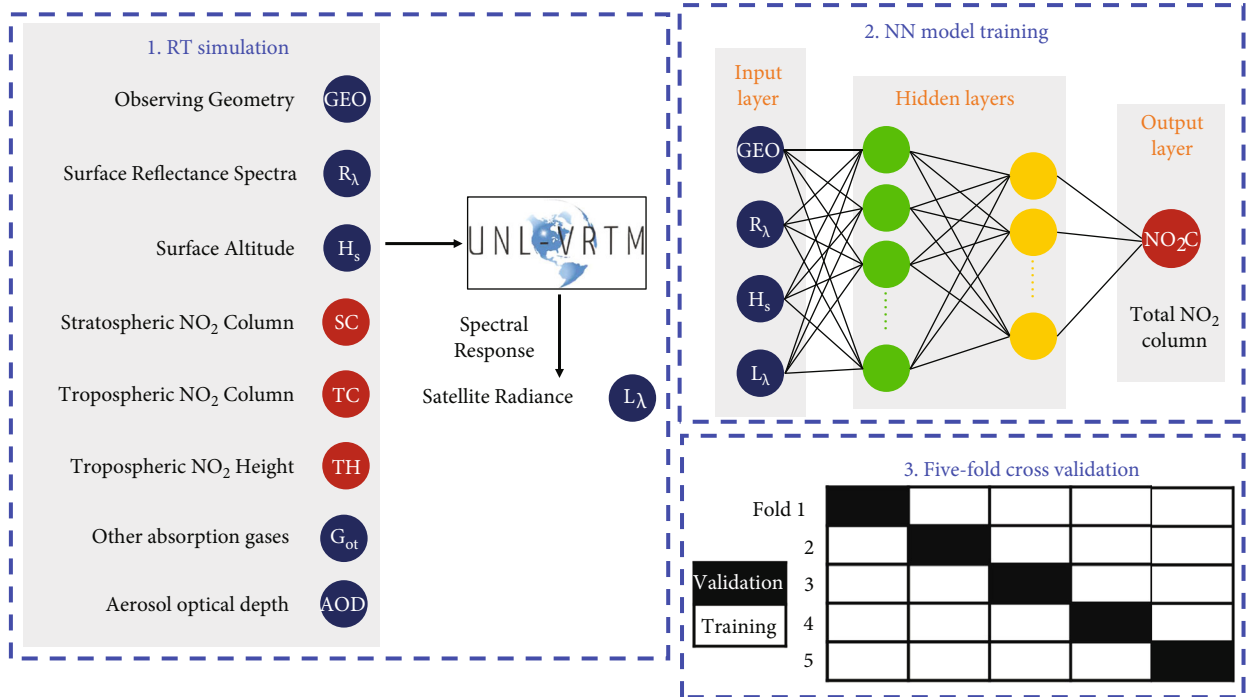


FIGURE 1: Illustration of the neural network model generation to predict total NO_2 column density from satellite radiances. Box 1 includes the details about relevant variables driving the RT simulations (Section 2.1), with parameters in dark blue (red) being potential predictors (predictands) in the NN model; box 2 shows the structure of the NN model; and box 3 presents the conceptual illustration of cross-validation. The total NO_2 column (NO_2C) is the sum of stratospheric (SC) and tropospheric (TC) column. For L_λ and R_λ , the actual inputs in the NN are their associated PC scores (Sections 2.2 and 3.2). Other absorption gases (G_{ot}) and aerosol optical depth (AOD) varied according to Table 1 in the simulation while are not included as predictors in the final NN model according to sensitivity tests in Section 2.4.

TABLE 1: Input variables and their corresponding data ranges or sources in the RT simulation (DU is Dobson unit, $1 \text{ DU} = 2.69 \times 10^{16}$ molecules/ cm^2). Each simulation adopts a random number for each variable. See Section 2.1 and Figure 1 for the definition of each abbreviated variable.

Variable group	NN variable	Range (source)	Notes
GEO	SZA	$0-75^\circ$	
	VZA	$0-75^\circ$	
	RAA	$0-180^\circ$	
R_λ	3 PC scores	From climatology (Figure S1)	
H_s	H_s	$0-8 \text{ km}$	
	TC	$0-2 \text{ DU}$	See Figure S2 about four sub-groups
NO_2	TH	$0-\max(10, 12-H_s) \text{ km}$	
	SC	$0.01-0.2 \text{ DU}$	
	O_3	$200-500 \text{ DU}$	
G_{ot}	SO_2	$0-50 \text{ DU}$	Profile fixed as in US standard (1976)
	HCHO	$0-2 \text{ DU}$	
	H_2O	$(5-50) \times 105 \text{ DU}$	
Aerosols	AOD (550 nm)	$0-1$	

shape [52, 53] between the surface and tropopause (fixed at 12 km). The half width of the tropospheric plume profile is a random number between 0.4 and 1 km for each case. The distribution of NO_2 abundance over different regions is

strongly uneven, and heavy pollution (i.e., tropospheric $\text{NO}_2 > 1016$ molecules/ cm^2) is more likely to occur over urban areas and near the ground [76]. To ensure representativeness, the whole training dataset was also separately

generated for four subgroups (Figure S2), as defined based on specific ranges of tropospheric NO₂ column (TC) and centroid height (TH), as well as H_s . Specifically, the simulations were dominated by urban polluted (low H_s , low TH, and high TC) and remote clean (random H_s and TH, low TC) scenarios, plus some elevated NO₂ enhancements due to lightning and biomass burning (random H_s , high TH, and high TC), and a small number of polluted cases over high altitude surface (high H_s , low TH, and high TC). TH, TC, and H_s varied randomly within the specific range (Table 1) in each subgroup.

Our training set contains 51703 samples with $H_s > 2$ km and bright surface ($R_\lambda > 0.6$ with $<3\%$ spectral variability), which could represent high-altitude snow or cloud surfaces. We directly use the NN from the whole training set (Section 2.2) for retrieving above-cloud NO₂C under fully cloudy assumptions (see details in Section 3.2). As this study was not designed to perform retrievals for cloudy scenes where NO₂ retrievals are highly uncertain, and we focus on discussing the performance of retrievals under clear-sky conditions (Sections 3.3 and 3.4), the lack of cloudy-scene specific retrievals does not alter the conclusions of this paper. Future efforts to improve the treatment of cloudy scenes in the RT simulations and NN retrievals are needed to assess the ability of the NN to describe these scenes (see also discussion in Section 4).

2.2. NN Training and Evaluation. The overall structure of the feed-forward NN used in this study [77] is defined in the box 2 of Figure 1, which composes of an input layer of predictors (dark blue), an output layer of predictands (red), and several hidden layers to mimic their relationship. Each hidden layer includes several computational nodes (neurons, N), and each neuron is modeled as a nonlinear activation function (f) of the weighted (the weights denoted as w) sum of all neurons in the previous layer plus an offset (b):

$$N_{i+1}^j = f \left(\sum_k w_{i,i+1}^{jk} N_i^k + b_{i,i+1}^j \right), \quad (1)$$

where i is the layer index and k (j) is the node index in layer i ($i+1$). The training and application of the NN was done using the Python scikit-learn (sklearn. neural_network.ML-PR regressor) package [78], which attempts to optimize these weights (w) and offsets (b) to minimize the least square difference between predicted and true predictands in the training data. Our numeric tests favored an NN configuration of two hidden layers, with 16 nodes in the first layer and 8 nodes in the second, that had reached the optimal model prediction power. Further increasing node or layer numbers did not improve the performance.

We define the predictand in our study as the total NO₂ vertical column (NO₂C = TC + SC) in the training dataset, reflecting the unknown NO₂ stratospheric/tropospheric separation in real-world retrievals. Trials on retrieving SC or TC were also tested. The performance was significantly weaker, reflecting the low information content about vertical location of NO₂ in most of the satellite observed radiances.

Further processing of the retrieved NO₂C using various available algorithms [31, 32] could separate the stratospheric and tropospheric component, which is beyond the scope of this study. Inherently, all the remaining variables in the RT simulation (i.e., dark blue terms in box 1 of Figure 1) plus the satellite radiances (L_λ) are potentially predictors of the NN model, from which we will determine their actual contribution to model prediction power and select the employed predictors in the final NN model (Section 2.4).

Within the predictors, the spectral surface reflectance (R_λ) from the climatology in Figure S1 and the simulated satellite radiances (L_λ) data contains a vector of 20 and up to 900 wavelengths, respectively, and most of these spectral observations are correlated [79], implying information redundancy. We followed previous studies [62, 72] to use the principal component analysis (PCA) technique to reduce the dimension of these variables and simplify the NN model training while maintaining the information content. The PCA was conducted using the sklearn.decomposition.PCA routine in scikit-learn [78]. We found that the top three leading PC scores of R_λ could reproduce the full (i.e., $>99.99\%$) variability of the employed monthly climatology over land, which were used in the actual NN model experiments. For L_λ , the 15 leading PC scores could explain the full variability in all the training data, and we also only selected several key PC scores that are the most relevant for NO₂ column prediction in the final NN model (Section 2.4).

We used the 5-fold cross-validation technique [64, 80] to evaluate the theoretical model performance for predicting NO₂C in the training dataset (box 3 of Figure 1). Specifically, the whole training data samples were divided into 5 equal-sized groups. In building the NN model (fold 1), the first 20% of the data was used as the evaluation data, and the remaining 80% was used for training. This was repeated five times until the data were fully covered, with every NO₂C record containing a truth and prediction pair, from which the overall evaluation statistics (e.g., R^2) were calculated. The cross-validation helps to minimize bias in training data selection and ensures the representativeness of the evaluation metrics of the whole training records.

2.3. Determination of Optimal Retrieval Spectral Range. The spectral windows for previous NO₂ retrieval were mostly near the peak NO₂ absorption (i.e., the absorption cross section σ) around ~ 430 nm, with the width largely increasing from GOME (425-450 nm) [14] to TROPOMI (405-465 nm) [29]. However, the benefit of wider spectral windows might also be compensated by the uncertainty of other absorptions (especially further from the σ peak) and reduced validity of constant slant column in the spectral fitting [30, 55]. Here, we apply the NN training and cross-validation as a pure data-driven tool to investigate which spectral range provides the maximum useful information.

Using all the potential predictors (dark blue terms in the box 1 of Figure 1 plus the leading 15 PC scores of L_λ), we conducted the NN training and cross-validation adopting radiances covering different spectral ranges. We repeated the experiment 5 times (i.e., vertically aligned circles in

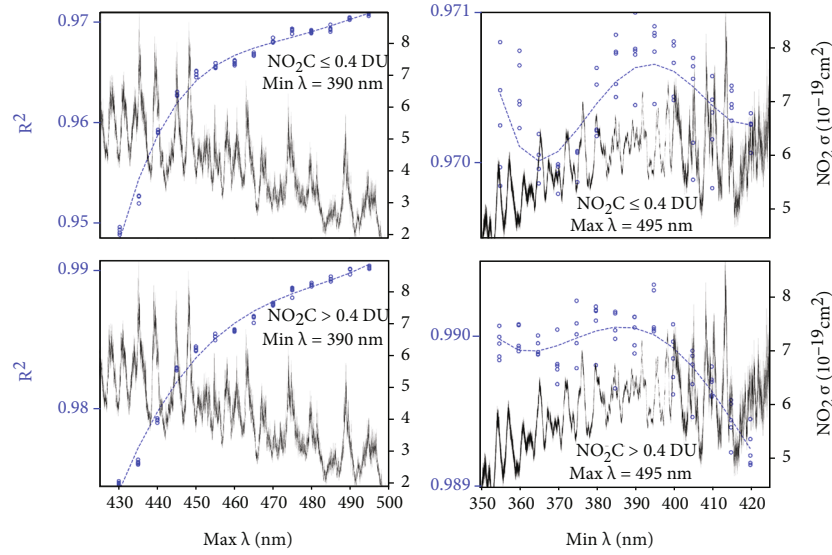


FIGURE 2: Cross-validation R^2 (blue circles, repeated 5 times) for different retrieval windows (sliding at 5 nm interval) using low- NO_2 (a, b) and high- NO_2 (c, d) samples for NN training. (a, c) contain results with fixed start wavelength (390 nm), and (b, d) are R^2 with fixed right wavelength (495 nm). Dashed lines are 4th-order polynomial fits to guide the interpretation. NO_2 absorption cross-section at the temperature of 294 K from Harder et al. [90] is superimposed in black.

Figure 2) for each retrieval window to reduce the impact of NN's internal uncertainty of numeric solution. The results below are based on L_λ from the SRF of the 200th TROPOMI detector row, which represent consistent findings from other rows as expected.

Figure 2 shows the evolution of the coefficient of determination (R^2) following the gradually sliding spectral window of L_λ , and Figure S3 shows the corresponding root-mean-square-error (RMSE). The synthetic data was divided into low- NO_2 ($\text{NO}_2\text{C} \leq 0.4$ DU, (a, b)) and high- NO_2 ($\text{NO}_2\text{C} > 0.4$ DU, (c, d)) subgroups for training. Fixing the left wavelength at 390 nm and varying the max wavelength of the included radiances ((a, c)), we can see a trend of continuously increasing R^2 following the wider spectral range, demonstrating the stronger NN model capability to predict NO_2C from extraspectral information for both low- and high- NO_2 cases. For low- NO_2 , the model improvement vs. wavelength range slows after ~ 465 nm, with additional RMSE reduction of $\sim 0.50\%$ from 465 to 495 nm, compared to the reduction of $\sim 1.85\%$ from 430 to 465 nm. This corresponds to a sharp reduction of NO_2 absorption ($\sigma < 5 \times 10^{-19} \text{ cm}^2$), suggesting strong dependence of the model performance on absorption strength rather than window width. Under high- NO_2 scenarios, the model continues to improve (RMSE reduction of $\sim 0.75\%$ from 465 to 495 nm vs. $\sim 1.95\%$ from 430 to 465 nm), suggesting additional information from extended spectral width is useful despite weaker NO_2 absorption and interferences of other absorptions, e.g., the strong $\text{O}_2\text{-O}_2$ absorption at 477 nm. As previously mentioned, satellite radiances contain sensitivity to the NO_2 vertical location under polluted conditions due to spectrally dependent scattering (i.e., AMF) [57], which is consistent with this more robust increase of R^2 vs. broader spectral range under high- NO_2 scenarios.

However, retrievals under low- NO_2 have a systematic $\sim 2\%$ smaller R^2 than under high- NO_2 conditions, due to the lack of vertical sensitivity that affects the model performance at less-absorbing spectral regions (i.e., weaker ability to distinguish between relatively lower NO_2C aloft and higher NO_2C near the ground). Based on this continuous improvement, we take the ending wavelength in the final NN model as 495 nm, the end of the range in the TROPOMI Band 4 detector.

Similarly, fixing the ending wavelength at 495 nm and sliding the left wavelength (Figures 2(b) and 2(d)) consistently suggests more stable increase of R^2 following wider spectral range used for the high- NO_2 cases, until the min wavelength reaches around 390 nm. Below 390 nm, the interferences from other gases, aerosols, and surface albedo result in plateauing (under high NO_2) or even decrease (under low NO_2) of R^2 . Although the variabilities of these additional gases were included in the NN predictors, they appear not to compensate for the reduction of NO_2 signal over these wavelengths. Including wavelengths below ~ 360 nm under low NO_2 is unstable, as indicated by the increasing scattering of R^2 among 5 repeated experiments. We therefore adopted the 390-495 nm range as the optimal retrieval window for the final NN model configuration. This retrieval window is determined assuming uniform performance of the measured spectral radiance. Inconsistencies between the Band3 and Band4 detectors, as well as degraded precision towards the end of Band4 (e.g., > 492 nm), are expected in realistic measurements. We emphasize the general usefulness of the extended information from wider spectral range in realistic retrievals (see Section 3.3 and Figure S8), while a more comprehensive optimization of retrieval window warrants further investigation and should vary from sensor to sensor.

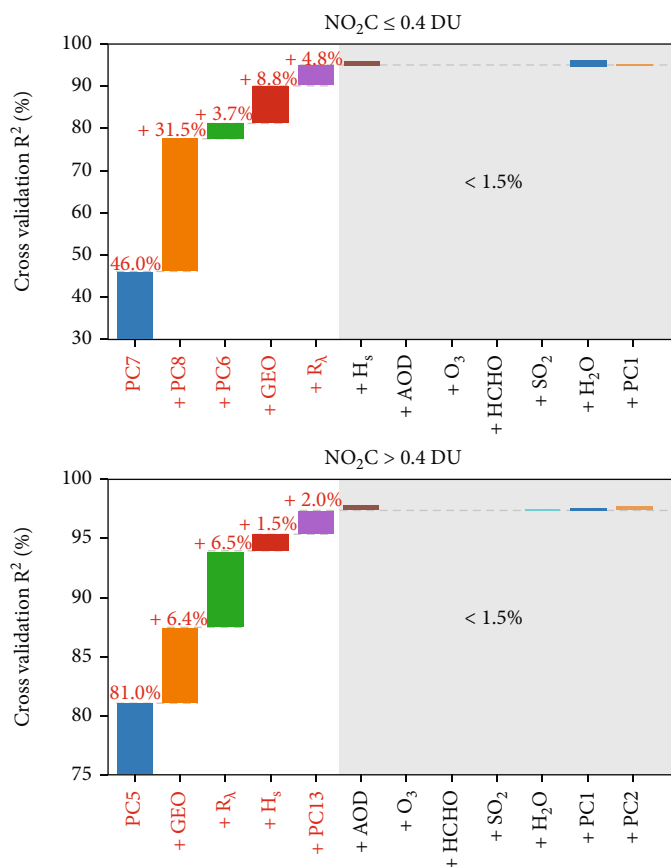


FIGURE 3: Changes in cross-validation R^2 as employing increasing number of predictors in the NN training for low- NO_2 (a) and high- NO_2 (b) cases. The additional variability explained after adding each variable (group) was indicated in red above each waterfall bar, and PC# denotes the #th PC scores or L_λ . All the variables in the shaded area introduce $<1.5\%$ additional variability according to the experiment.

Figure 2 suggests qualitatively contrasting sensitivities of satellite radiances to the spectral range (i.e., NO_2 vertical information) and a necessity of separate consideration, between high- and low- NO_2 cases in the NN model generation. More detailed tests (e.g., Figure S4) consistently indicated that including training records with both high- and low- NO_2 environments resulted in unexpectedly decreasing R^2 vs. wider spectral range over 390–430 nm, reflecting the two different sensitivity regimes that cannot be jointly resolved by one NN model. We therefore separately constructed the NN models for low- and high- NO_2 cases.

2.4. Determination of Sensitive Predictors. Numerical methods like NN tend to be subject to overfitting, i.e., the variability of unrelated predictors falsely contributing to the model prediction. We selected predictors that contribute significantly to NO_2C prediction, using an additive approach. We first ranked the importance of each predictor by the cross-validation R^2 when only that predictor is used in the NN model. We then perform NN training and cross-validation iteratively, by adding one predictor (from the most to least important) at each iteration. The three GEO angles and R_λ PC scores were considered as two variable groups in the experiment. The improvement to R^2 by

adding each predictor is used to evaluate its importance. Again, we exemplify the results below based on L_λ over 390–495 nm and from the SRF of the 200th detector row.

Figure 3 shows the changes in cross-validation R^2 in the additive NN experiment. The key variables (labeled in red) that contribute to the model prediction power were selected to build the final NN models. For low- NO_2 cases, the 7th and 8th PC scores of L_λ from the radiances together could explain nearly 80% of the variability of NO_2C via NN, which is further improved by 4–9%, respectively, by adding the 6th PC score, GEO angles, and R_λ . For the high- NO_2 model, the 5th PC score of L_λ alone can explain more than 80% of the total NO_2C variability, and GEO and R_λ are also significant variables. We found that the 5th PC loading ($R=0.88$) extracted from L_λ of high- NO_2 cases, and the 7th ($R=-0.68$) and 8th ($R=0.55$) PC loadings from the low- NO_2 training set are all highly correlated with the $\text{NO}_2\sigma$. The “earlier appearance” of important PC (i.e., 5th compared to 7th and 8th) in the high- NO_2 model can be explained by the stronger contribution to the variability of L_λ in the training set. Finally, the 13th PC score appears to be the most relevant variable to predict TH in a one-predictor NN test, explaining $\sim 15\%$ TH variability in cross-validation. Hedelt et al. [62] also suggested that volcanic SO_2 vertical information is contained in deep PCs of

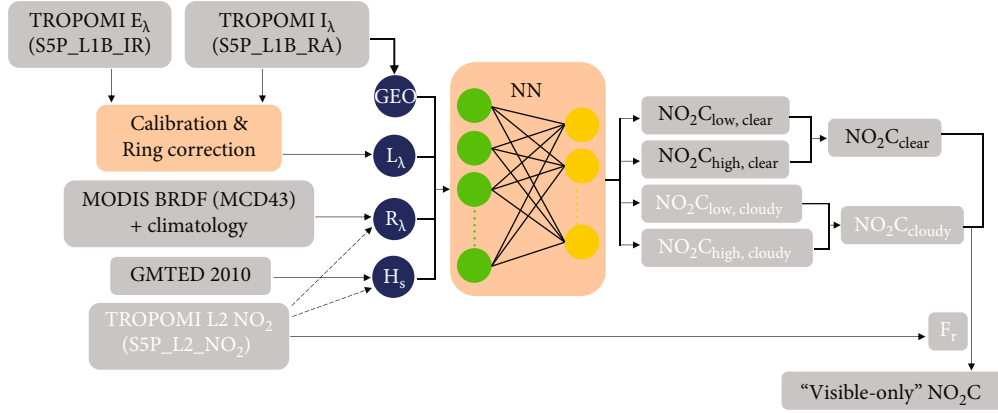


FIGURE 4: Processing flowchart to retrieve NO_2C from TROPOMI radiances (390-495 nm), including datasets (grey shaded) and processing steps (pink shaded). Datasets related with retrievals under clear (cloudy) conditions are labelled with black (white) fonts. See Sections 3.1 and 3.2 for detailed introduction of each processing step.

satellite hyperspectral radiances. Reasonably, we can see that this predictor also helps improve the prediction of NO_2C in high- NO_2 cases.

The importance of surface altitude (H_s) was found different in the low- NO_2 and high- NO_2 NN model due to their different vertical sensitivities. Because the sensitivity to tropospheric NO_2 height is weak, the retrieval under low- NO_2 also becomes almost insensitive to H_s . On the contrary, sensitivity to NO_2 altitude is more pronounced in satellite radiances under high- NO_2 environments, where accurate H_s information also becomes significant. Variabilities of AOD, other absorption gases (G_{ot}), and other PC scores of L_λ were found to contribute <1.5% of additional predicted NO_2C variability and are therefore not included in the final NN model.

3. Application to TROPOMI

The NN model built from the previous section can be applied to TROPOMI radiances and complimentary data to test the validity of retrieved NO_2C . TROPOMI is a hyperspectral backscattering sensor onboard the sun-synchronous Sentinel-5 Precursor (S5P) satellite, achieving daily global coverage with an afternoon (13:30) overpass. TROPOMI has a wide swath of 2600 km and horizontal resolution of $5.5 \text{ km} \times 3.5 \text{ km}$ (since August 6, 2019). Across the two UV-Vis Bands (Band 3: 320-405 nm and Band 4: 405-495 nm), the measured radiances have a spectral resolution of 0.2–0.4 nm and a signal-to-noise ratio of ~ 1000 . These characteristics enable TROPOMI to provide unprecedented urban-scale monitoring of atmospheric composition. Retrievals of the level 2 (L2) operational TROPOMI NO_2 product follow the typical three-step process [13, 29, 81], in which the TM5-MP model is used to feed the *a priori* NO_2 profiles for an AMF calculation. Validation of the L2 products suggest a systematic underestimation of L2 NO_2C by 30% against global Pandora measurements at high- NO_2 (median $\text{NO}_2\text{C} > 9 \times 10^{15}$ molecules/ cm^2) [82].

Details about datasets and variables used in the retrieval are summarized in Table S1. The data processing flowchart

is summarized in Figure 4 and Sections 3.1 and 3.2. We then present a comprehensive evaluation of the retrievals over four months (September/December 2019 and March/June 2020) and three source regions (East Asia, Europe, and North America, Figure S5), using independent satellite and ground-based dataset in Sections 3.3 and 3.4.

3.1. Wavelength Calibration and Ring Correction. We first calculated L_λ from TROPOMI L1B solar irradiance (E) and earthshine radiance (I) data (Bands 3 and 4). Since the whole Band 4 spectra and part of Band 3 (390-495 nm) are used, the wavelength shift and stretch terms for wavelength calibration in TROPOMI standard level-2 (L2) NO_2 data (405-465 nm) are not suitable, and we repeated the calibration process before L2 NO_2 retrievals [29] in this study. Briefly, the shift and stretch terms were first determined for each day and detector row using the solar irradiance observation and reference spectrum (E_0). Then, a further shift (w_λ) and Ring coefficient (C_{ring}) for each TROPOMI pixel were fitted to

$$I(\lambda_n + w_\lambda) = E_0(\lambda_n + w_\lambda) \times P(\lambda_n + w_\lambda) \times \left(1 + C_{\text{ring}} \frac{I_{\text{ring}}(\lambda_n + w_\lambda)}{E_0(\lambda_n + w_\lambda)} \right), \quad (2)$$

where λ_n is the wavelength nodes after the first-step irradiance-based calibration, P is a 3rd-order polynomial function of wavelength that implicitly accounts for the effects of slit function and other errors during the radiometric calibration, I is the radiance intensity measured by TROPOMI, and I_{ring}/E_0 is the simulated sun-normalized Ring spectrum. The employed solar and Ring reference spectra (Table S1) are consistent with the ones used in TROPOMI L2 NO_2 retrievals [29] that were already convoluted with the TROPOMI SRFs of Bands 3 and 4.

The Ring-corrected TROPOMI observation is then calculated based on the calibrated radiance intensity (I_λ) and irradiance (E_λ):

$$L_\lambda = \ln \left[\frac{I_\lambda}{E_\lambda (1 + C_{\text{ring}}(I_{\text{ring},\lambda}/E_{0,\lambda}))} \right]. \quad (3)$$

As we did not include the Raman rotational scattering in the RT simulation, the correction of Ring effect based on the fitted C_{ring} is necessary.

3.2. Visible-Only NO_2C Calculation. The TROPOMI L_λ could then be applied to the NN model to determine NO_2C . For each TROPOMI pixel, we retrieved four NO_2C values, differing by the adopted NN model for high- NO_2 or low- NO_2 conditions and the assumption of pure clear-sky or fully cloudy scene (Figure 4). Under the clear-sky assumption, the R_λ

(calculated using GEO and MODIS BRDF product [83] at 470 nm and spectral shape from the climatology introduced in Figure S1) and H_s (from the GMTED 2010 [84] global elevation dataset) described the surface; while for fully cloudy retrievals, R_λ (cloud albedo with flattened spectral shape) and H_s (converted from cloud pressure) were from the FRESCO+ [81, 85] cloud retrievals using the NO_2 band of TROPOMI (available in the L2 NO_2 product). The PC scores of L_λ (R_λ) were then derived using the precalculated eigenvectors from the training sets (L_λ climatology), before being applied to the NN models.

The NO_2C values determined from applying the high- NO_2 and low- NO_2 NN models, respectively, were first combined under both the clear and cloudy assumptions:

$$\text{NO}_2\text{C} = w_{\text{low}} \cdot \text{NO}_2\text{C}_{\text{low}} + (1 - w_{\text{low}}) \text{NO}_2\text{C}_{\text{high}}, \quad (4)$$

$$w_{\text{low}} = \begin{cases} 1, & \text{if } [\text{NO}_2\text{C}_{\text{low}} < (0.4 - \text{buffer})] \text{ or } [\text{NO}_2\text{C}_{\text{low}} < 0.4 \text{ and } \text{NO}_2\text{C}_{\text{high}} < 0.4] \text{ (Zone I),} \\ 0, & \text{if } \text{NO}_2\text{C}_{\text{low}} \geq 0.4 \text{ and } \text{NO}_2\text{C}_{\text{high}} \geq 0.4 \text{ (Zone II),} \\ \frac{0.4 - \text{NO}_2\text{C}_{\text{low}}}{\text{buffer}}, & \text{if } (0.4 - \text{buffer}) \leq \text{NO}_2\text{C}_{\text{low}} < 0.4 \text{ and } \text{NO}_2\text{C}_{\text{high}} \geq 0.4 \text{ (Zone III),} \\ \text{invalid,} & \text{if } \text{NO}_2\text{C}_{\text{low}} \geq 0.4 \text{ and } \text{NO}_2\text{C}_{\text{high}} < 0.4 \text{ (Zone IV).} \end{cases} \quad (5)$$

This merging strategy was inferred from the 2D density plot of NO_2C from both NN models, which presents a two-mode distribution (Figure S6). One of the two clusters (i.e., Zone I of Figure S6) has retrievals < 0.3 DU from the low- NO_2 NN while the NO_2C predicted by the high- NO_2 NN is not covarying but scattered. This NO_2C from the high- NO_2 NN in this regime is likely irrational considering most pixels are not over emission sources and should have low NO_2C , and NO_2C from the low- NO_2 NN is taken. A smaller number of NO_2C near the other mode (Zone II) are jointly supported by the low- and high- NO_2 NN retrievals (i.e., > 0.4 DU), and the NO_2C from the high- NO_2 NN is selected due to its representativeness. We set a buffer NO_2C (in DU) in Equation (5) to determine a narrow merge zone (i.e., Zone III in Figure S6). Over this regime, the two retrievals were considered indistinguishable and weighted-averaged, with the weights linearly varying between Zones I and II. Zone IV contains retrievals that both NN models predict NO_2C out of its represented ranges, which rarely occurred in the results. Tests varying this buffer NO_2 suggested that a buffer of 0.08 DU resulted in the best agreement vs. ground-based (i.e., Pandonia and MAX-DOAS) measurements in Section 3.3, and this is adopted in the rest of the processing.

Our RT simulation did not explicitly consider cloud scattering and absorption between the bottom surface and TOA due to the strong opaqueness of clouds over the UV-Vis [86]; therefore, the NO_2C from the NN model under both the clear-sky and fully cloudy assumption is “visible-only” columns (i.e., total column above either the surface

or the cloud layer). We merge these two retrievals based on the cloud radiance fraction (F_r) to generate the final NO_2C amount:

$$\frac{1}{\text{NO}_2\text{C}} = \frac{(1 - F_r)}{\text{NO}_2\text{C}_{\text{clear}}} + \frac{F_r}{\text{NO}_2\text{C}_{\text{cloudy}}}. \quad (6)$$

Equation (6) is similar to the processing in the L2 product, where the clear-sky and cloudy AMFs were assumed linearly weighted based on F_r [81]. Although operational retrievals (e.g., the L2 NO_2) report total column above the surface under cloudy scenes, the information under the cloud layer is still contributed by the modeled *a priori* profile. As our purpose is to relax that reliance, we do not attempt to estimate the “ghost column” under cloud, rendering the NO_2C from our approach a pure radiance-based retrieval.

The TROPOMI spectral responses vary subtly between detector rows (each with own SRFs). We generated NN models based on trainings using SRFs of every 50th detector row, and linearly interpolated the retrieved NO_2C to the untrained rows following Hedelt et al. [62]. This interpolation approach was confirmed applicable as no artificial row-based striping was detected in the output daily NO_2C maps.

3.3. Evaluation with the Pandonia Global Network. The Pandonia Global Network (PGN) is the main source of ground truth for our evaluation. PGN provides direct sun-view measurements of total vertical column of various trace

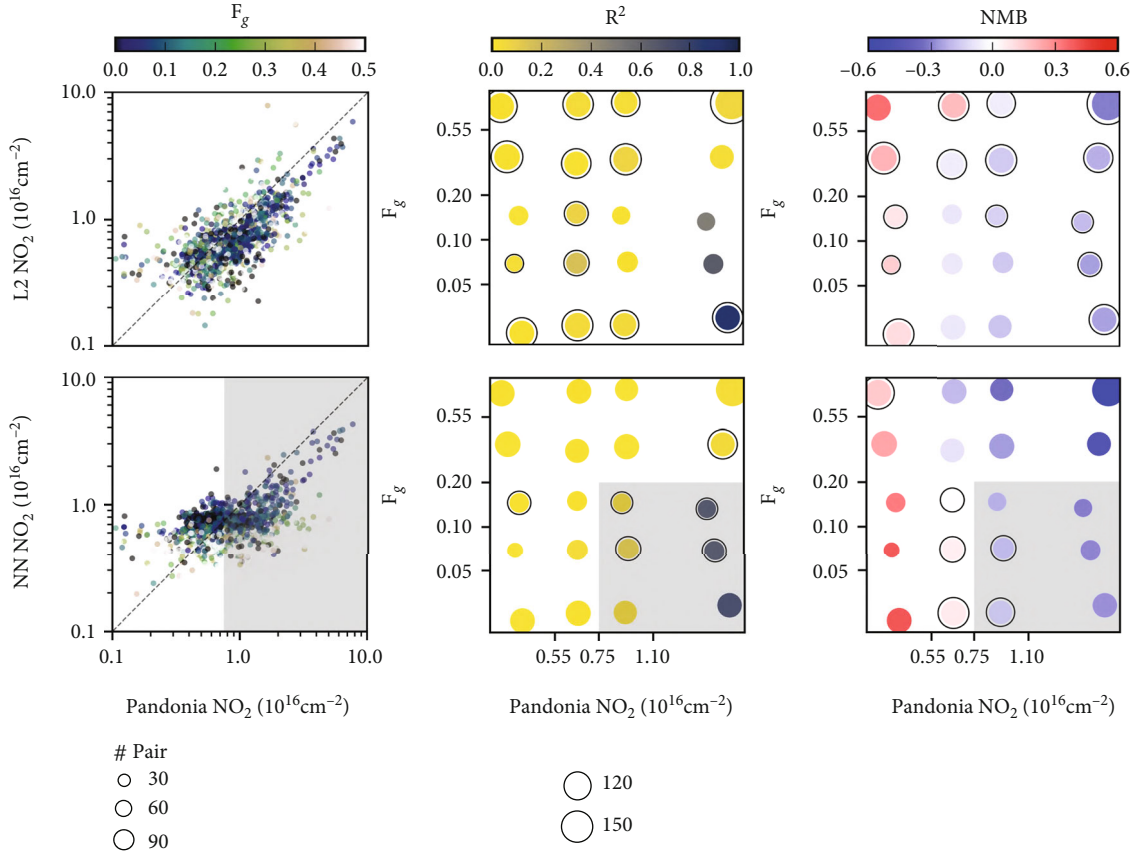


FIGURE 5: Evaluation of the TROPOMI L2 (a–c) operational product and the NN retrievals (d–f) with PGN measurements. (a, d) show scatter plots of all collocated pairs color-coded by the geometric cloud fraction (F_g), (b, e) are R^2 of both retrievals against PGN for 2-D bins (bin boundaries are the axes ticks) of PGN NO_2C and F_g , and (c, f) are the normalized mean biases (NMBs) in each bin. Marker sizes reflect the number of pairs in each bin, and markers with a black outline indicate better performance (e.g., larger R^2 or smaller NMB) than the other retrieval. The shaded area is the “optimal NN regime.”

gases from ground-based standardized Pandora Sun photometers. The Pandora NO_2C has a random error of ~ 0.05 DU (1.35×10^{15} molecules/ cm^2) and has been widely applied to the validation of satellite NO_2 retrievals [82, 87].

The PGN NO_2C were collocated with TROPOMI NO_2C , using conventional spatial and temporal proxy criteria (i.e., the closest pair within 30 min in time and 30 km in distance). All the available TROPOMI retrievals without quality filtering were applied to the collocation due to currently unavailable quality assignments in the NN retrievals (which requires longer-term data collection and more comprehensive evaluations). The evaluation conclusions were also consistent if only selecting high quality PGN data, except for the significantly reduced number of available retrievals under cloudy conditions (i.e., comparing Figure 5 and Figure S7). We present in Figure 5 those without the selection, consistent with the evaluations over China (Section 3.4) that are without available quality flags.

Figure 5 shows the evaluation of the NN retrievals and the TROPOMI L2 NO_2C vs. collocated PGN measurements. Both retrievals partially reproduce the range of PGN NO_2C , with overestimations at low NO_2C and underestimations at high NO_2C . This was similarly observed in a more comprehensive evaluation of the L2 retrievals [82], and could be

partially explained by the TROPOMI FOV that smears the Pandora measured local air mass. Interestingly, the two retrievals show similar distribution of performances over different regimes, as jointly determined by NO_2C and geometric cloud fraction (F_g). Only over high NO_2C ($\geq 1.1 \times 10^{16}$ molecules/ cm^2) and low cloudiness ($F_g < 0.2$) conditions do both retrievals have relatively strong correlations vs. PGN ($R^2 > 0.48$), whereas under the other conditions, both retrievals exhibit low correlations ($R^2 < 0.16$). The distributions of R^2 and normalized mean bias (NMB) from the two independent retrievals highlight the uncertainty of the model simulated *a priori* profile and indicate the profiles only marginally compensate for the weak NO_2 signal.

Without the aid of modeled NO_2 under clouds, the NN retrievals exhibit stronger degradation than the L2 NO_2C under cloudy conditions ($F_g \geq 0.2$). The NN predicted NO_2C also has weaker performances vs. L2 under low NO_2C (i.e., $< 7.5 \times 10^{15}$ molecules/ cm^2) because of the previously mentioned reduced vertical sensitivity (Section 2.3). We identified an “optimal NN regime” (ONR, $\text{NO}_2\text{C} \geq 7.5 \times 10^{15}$ molecules/ cm^2 and $F_g < 0.2$, i.e., shaded in Figure 5), where the NN retrievals reveal robust ability

(i.e., close or slightly better R^2) similar to the L2 products, to reproduce the PGN NO_2C . Within this regime, the NN retrievals have an overall R^2 of 75% and NMB of -33%, both comparable to the L2's R^2 of 77% and NMB of -29%, whereas out of the ONR, the overall R^2 of NN is <1% compared to the L2's R^2 of 17%. Therefore, for the ONR scenario, the extended spectral range provided the trained NN model additional observation-based information to resolve NO_2C variability, that is competitive to a model simulation of *a priori* NO_2 profiles in L2.

This extended spectral information confirmed the contribution to the NN retrieval of the radiances of the TROPOMI Band 4 (401-495 nm, Figure S8). The reduced vertical sensitivity resulted in systematic underestimation of NO_2C vs. PGN across almost all the bins, mainly due to the smaller number (~50%) of high NO_2C ($>1 \times 10^{16}$ molecules/cm²) pixels. As we discussed in Section 2.3, the vertical sensitivity of the high- NO_2 NN model increases with the spectral width of included retrieval window (Section 2.3). Even the ~11 nm spectral information in Band 3 is useful for the NN-based retrieval. We expect the retrievals using the spectral information adopted in L2 (405-465 nm) to perform worse than in Figure S8 if relying on a pure NN-based approach.

3.4. Evaluation with Complementary NO_2 Measurements over China. The PGN network is mostly located over developed regions with smaller NO_2C (Figure S5). Since the NN retrievals reveal a strong dependence on NO_2 abundance, it is also meaningful to evaluate the applicability of the NN retrievals over developing areas with stronger NO_2 pollution (e.g., China).

Over wide areas of China without PGN sites, we used surface NO_2 concentration and MAX-DOAS tropospheric NO_2 column measurements for the evaluation of TROPOMI NO_2C (Figure S5). The hourly surface NO_2 concentrations are from the routine monitoring by the Ministry of Ecology and Environment (MEE) of China [88]. The MAX-DOAS measurements [41, 89] are from three suburban stations (Beijing, Xuzhou and Nanjing, respectively, from North to South in Figure S5).

Figure S9 shows the comparisons of the retrievals against MEE surface NO_2 over China. With a wider dynamic range of NO_2C (i.e., $0 - 6 \times 10^{16}$ molecules/cm²) more frequent occurrence of scenarios favorable for NN retrievals (i.e., within the ONR regime), the NN shows competitive performance vs. the L2 for resolving surface NO_2 variability. We observe similar R^2 for the two retrievals (0.43 and 0.42) in near clear-sky conditions ($F_g < 0.2$). The NN underestimates surface NO_2 in many cases where the L2 does not, forming a discernable population (Figure S9, right), which likely corresponds to cases where NO_2 is strongly concentrated near ground. The underestimation by the NN is consistent with the weaker overall signal of near-ground NO_2 ; thus, higher retrieved NO_2C from the NN still provides a robust indicator of its data quality. For cases with high TROPOMI NO_2C , the R^2 of NN retrievals surpasses that of the L2 product (e.g., above the threshold

$\text{NO}_2\text{C} \geq 6 \times 10^{15}$ molecules/cm², Figure S10). Differences between the two retrievals increase with the NO_2C threshold. This reveals the meaningful vertical sensitivity in the TROPOMI radiances penetrating into the boundary layer that gradually becomes advantageous against the modeled profiles used in L2 for more polluted cases. The NN retrievals are thus especially applicable for estimation of surface NO_2 over China, and other conditions with very high NO_2 amounts.

Figure S11 shows the comparison against MAX-DOAS tropospheric NO_2 columns. Under $F_g < 0.2$, both retrievals are highly correlated ($R^2 > 0.75$) with the MAX-DOAS measurements and are systematically biased low (although the MAX-DOAS data only represents the tropospheric NO_2C). Nonetheless, the NN exhibits less underestimation than the L2 product. We found that switching the L2 dataset from total NO_2C to tropospheric NO_2 column slightly deteriorated the L2 performance.

Figure 6 presents maps of retrieved NO_2C over the East Asia domain, from the two independent retrievals. There are consistent spatial distributions of NO_2C from both products under near clear-sky conditions (i.e., $F_g < 0.2$). NO_2C enhancements are noticed over anthropogenic polluted and heavily industrialized regions, including the North/Northeast China Plain, Inner-Mongolia and Shanxi coal industry center, Fenwei Plain, Sichuan Basin, Yangtze River Delta, Pearl River Delta, Wuhan, Seoul, and Tokyo Metropolitan regions; the rest of the domain is dominated by background NO_2C . The enhanced regions in the NN retrievals (“visible-only” column) are more localized than the L2 (“full” column) product (both with $F_g < 0.2$). The background NO_2C as determined by the NN retrieval are $1-2 \times 10^{15}$ molecules/cm² higher than the L2 product, consistent with its weaker performance and overestimation of NO_2C at $\text{NO}_2\text{C} < 5.5 \times 10^{15}$ molecules/cm² (Figure 5).

The rightmost panels of Figure 6 show the comparison of all collocated daily NO_2C from both products. Two clusters are notable in the scatterplots, one with smaller F_g and tight correlations around the 1-1 line, another with enhanced cloud presence and weakened covariation between the two retrievals (where the visible-only NO_2C from NN is systematically lower). For cases with $F_g < 0.2$, the two products show higher correlations over more severely polluted months (i.e., March and December, $R^2 > 0.65$) than relatively cleaner months (i.e., September, $R^2 < 0.35$). Such enhanced correlation at higher NO_2 concentration again confirms the inherently stronger reliability of both retrievals under polluted and near clear-sky conditions (Figure 5).

In summary, the evaluation of both TROPOMI NO_2C retrievals over China confirms that NN retrievals are accurate over polluted regions, with promising applicability to developing areas around the world.

4. Discussion and Conclusion

It has been proposed that extending the spectral range of NO_2 retrievals introduces additional vertical sensitivities [30, 55, 57], which might facilitate the retrievals without

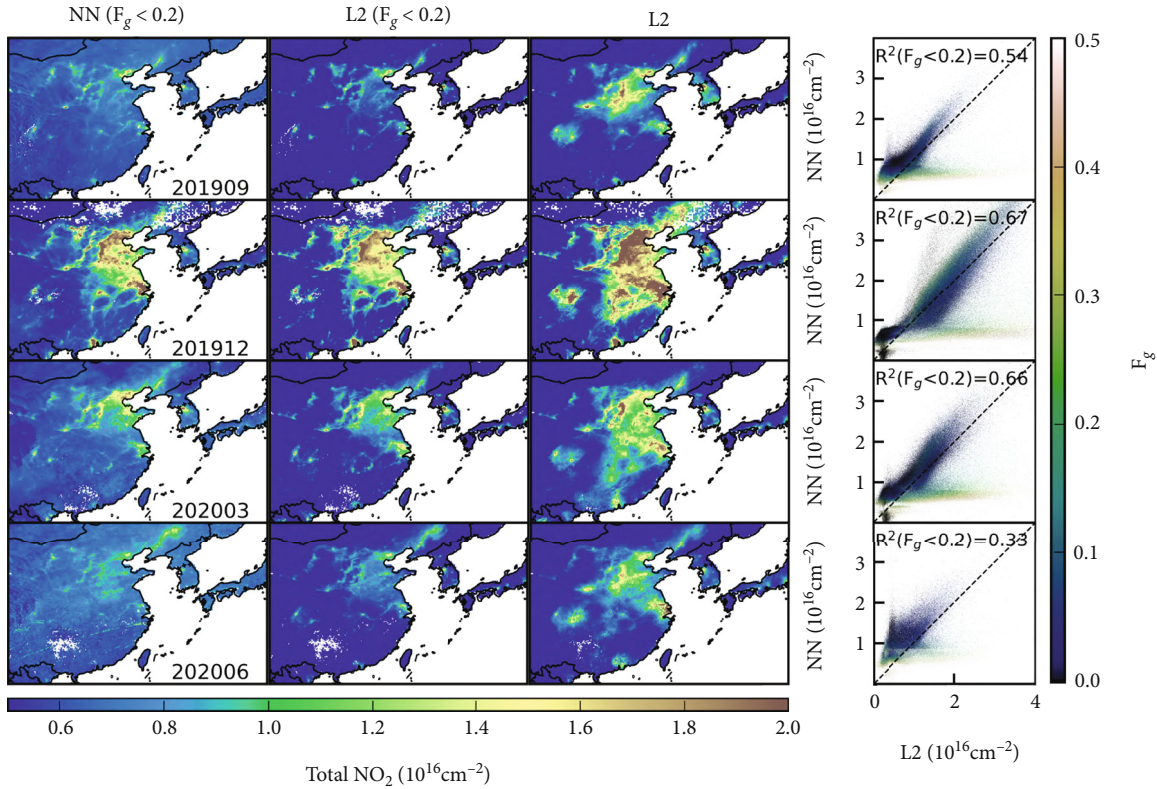


FIGURE 6: Maps of monthly NO_2C from the NN retrievals and from the TROPOMI L2 operational product over East Asia, with each row representing one month of the retrieval experiment. Only cases with $F_g < 0.2$ are included in the NN retrievals (first column), while the L2 products were derived from both near clear-sky (second column) and all-sky cases (third column). The rightmost column shows the scatter plots of all the collocated daily NO_2C from both products, color-coded by the corresponding F_g , with the dashed 1-1 line and the R^2 value for cases with $F_g < 0.2$.

modeling the *a priori* profile—the most time-consuming step. In this study, we used NN as a data-driven tool to investigate this issue from a new perspective and quantitatively tested this idea. To the best of our knowledge, this is the first attempt at a satellite NO_2 retrieval of this form. Our main findings are as follows:

- (1) The NN training experiments confirmed the existence of vertical sensitivity in the satellite radiances utilizing a broader spectral range. The retrieval sensitivity exhibits qualitative contrasts between high- and low- NO_2 cases, indicating a necessity of separate considerations
- (2) The retrieval sensitivities have a reasonable spectral distribution consistent with the relative strength of NO_2 absorption vs. other interfering trace gases, resulting in an optimal retrieval window of 390–495 nm for TROPOMI
- (3) Despite many sources of variabilities in the inputs of the RT simulation, only the key factors (observing geometry, surface reflectivity, surface altitude, and several key PC scores of satellite radiances) significantly contribute to the NO_2C prediction in the optimized NN
- (4) An application and evaluation of the NN model to TROPOMI reveals that NO_2C from the NN retrievals have competitive accuracy relative to the L2 product. The NN retrievals resolve NO_2C variation under less cloudy and more polluted scenarios ($\text{NO}_2\text{C} \geq 7.5 \times 10^{15}$ molecules/ cm^2 and $F_g < 0.2$). In other environments, both the NN and L2 retrievals show distinctive degradation, and the NN retrievals without *a priori* profiles become less reliable than L2. These findings are consistent with the theoretical variation of retrieval sensitivity, as is also revealed in the NN training experiments

Over populated areas, stratospheric NO_2 of $2\text{--}5 \times 10^{15}$ molecules/ cm^2 are persistent [32, 33]. Therefore, this study indicates that the NN retrieval can track tropospheric NO_2 pollution of as low as $\sim 3 \times 10^{15}$ molecules/ cm^2 at similar precision as the L2 data under clear-sky conditions, without the need to simulate *a priori* NO_2 profiles. This conclusion is especially promising for future retrievals using geostationary satellite observations over polluted areas, where the data volumes per day are over 10 times that of polar-orbiting satellites. The wall time for a retrieval using the NN and a 30-core node for 1-month hourly NO_2C at ~ 3 km resolution (typical resolution of future

Geostationary instruments) and over the domain shown in Figure 6 is ~20 hours, ~12 times faster than the time required for model simulation of the *a priori* profile (~10 days). The NN time is dominated by the wavelength calibration (Section 3.1).

For an efficient and effective retrieval for geostationary instruments such as TEMPO [27] or GEMS [26], additional research is needed to improve the machine learning-based NO₂ retrievals. Key ideas to be explored include:

- (1) Extend the training set to represent additional real-world scenarios. The fixed stratospheric NO₂ profile, aerosol optics, and tropopause height in the RT simulations could be better customized to potentially improve performance in regions and seasons where the assumptions were less sound. More sophisticated sampling of RT input variables (e.g., from model simulation over the retrieval domain) other than random sampling should be explored to increase the representativeness. In addition, the separate retrievals assuming clear and fully cloudy conditions could be improved by directly simulating the observed radiances under partial cloudy conditions. Alternative forms of NN structure, such as replacing the radiance PC scores with SCDs after spectral fitting should also be investigated to evaluate which contains more useful information. Finally, the emergence of additional ground-based NO₂C sites and data will facilitate an observation-based training set, as an alternative to RT simulations
- (2) Explore complementary datasets and methods where pure NN-based retrievals are less favorable (i.e., out of the ONR). This will enhance the potential utility of the retrieval products. Adding a climatological background NO₂ or a simulated ghost column, for example, could strengthen the robustness of the retrievals. To separate the stratospheric/tropospheric NO₂ columns, training additional NNs for remote/oceanic scenes or exploring the application of similar L2 stratospheric/tropospheric separation approach should be investigated
- (3) The vertical sensitivity over a wide spectral range indicates the possibility of retrieving NO₂ vertical location under heavily polluted conditions. This possibility should be further exploited, and will be valuable for studies of lightning, biomass burning, and the vertical variation of tropospheric chemistry and for evaluating the *a priori* profiles in operational retrievals

Data Availability

The neural network retrieval code with necessary instructions is available at GitHub (<https://github.com/ChiLi90/ANNNO2>). One-day of input data for testing the code is available at https://mega.nz/file/c10wGKJT#GkY6_HCDLIM88T5vlO7P26bL9oc53Fj7N0z3-oGKp58.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Authors' Contributions

C. Li, X. Liu, and R. C. Cohen designed the study. X. Liu and R. C. Cohen acquired funding for the project. X. Xu and J. Wang provided the UNL-VRTM tool and instructions. For the wavelength calibration, K. Sun shared the algorithm, and J. van Geffen provided the necessary datasets. Q. Zhu collaborated on design of the machine learning approach. J. Ma, J. Jin, K. Qin, Q. He, P. Xie, and B. Ren provide the MAX-DOAS data over China. C. Li performed the radiative transfer simulation, neural network training, and TROPOMI data processing. C. Li interpreted the results and wrote the paper, with editorial contributions from all the other co-authors.

Acknowledgments

We thank all the developers and maintainers of various satellite and other datasets used in this manuscript (Table S1). This work was supported by the Postdoctoral Program in Environmental Chemistry of the Camille and Henry Dreyfus Foundation, the National Aeronautics and Space Administration (grant no. 80NSSC19K0945), and the Smithsonian Institution (grant no. SV383019). J. Wang's participation is made possible by the in-kind (James E. Ashton Professorship) support from The University of Iowa. J. Jin was partially supported by the National Nature Science Foundation of China under the project no. 41805027, and the Ministry of Science and Technology of China under the project no. 2017YFC1501802.

Supplementary Materials

Table S1. Information of used dataset and variables. Figure S1. Surface spectral reflectance ($R\lambda$) climatology used in this study Figure S2. Number of cases for subgroups in the training dataset Figure S3. Similar to Figure 2 but for the root-mean-square-error Figure S4. Cross-validation R2 for different retrieval windows Figure S5. The boundaries of 3 retrieval domains and used surface sites Figure S6. Illustration of the merging method of NO₂C Figure S7. Similar to Figure 5 but limiting to cases with high quality PGN measurements. Figure S8. Similar to Figure 5 but from a separate NN retrieval only using TROPOMI Band 4 Figure S9. Density plot of collocated TROPOMI NO₂C and surface NO₂ over China Figure S10. Evolution of R2 of TROPOMI NO₂C vs. surface NO₂ over China Figure S11. Scatter plot of TROPOMI NO₂C and MAX-DOAS tropospheric NO₂ column. (*Supplementary Materials*)

References

- [1] A. W. Rollins, E. C. Browne, K. E. Min et al., "Evidence for NO_x control over nighttime SOA formation," *Science*, vol. 337, no. 6099, pp. 1210–1212, 2012.

- [2] H. O. T. Pye, H. Liao, S. Wu et al., "Effect of changes in climate and emissions on future sulfate-nitrate-ammonium aerosol levels in the United States," *Journal of Geophysical Research: Atmospheres*, vol. 114, no. D1, 2009.
- [3] A. G. Carlton, H. O. T. Pye, K. R. Baker, and C. J. Hennigan, "Additional benefits of federal air-quality rules: model estimates of controllable biogenic secondary organic aerosol," *Environmental Science & Technology*, vol. 52, no. 16, pp. 9254–9265, 2018.
- [4] Z. Tan, K. Lu, H. Dong et al., "Explicit diagnosis of the local ozone production rate and the ozone-NO_x-VOC sensitivities," *Science Bulletin*, vol. 63, no. 16, pp. 1067–1076, 2018.
- [5] N. Wang, X. Lyu, X. Deng, X. Huang, F. Jiang, and A. Ding, "Aggravating O₃ pollution due to NO_x emission control in eastern China," *Science of the Total Environment*, vol. 677, pp. 732–744, 2019.
- [6] S. E. Pusede and R. C. Cohen, "On the observed response of ozone to NO₂ and VOC reactivity reductions in San Joaquin Valley California 1995–present," *Atmospheric Chemistry and Physics*, vol. 12, no. 18, pp. 8323–8339, 2012.
- [7] D. W. Kicklighter, J. M. Melillo, E. Monier, A. P. Sokolov, and Q. Zhuang, "Future nitrogen availability and its effect on carbon sequestration in Northern Eurasia," *Nature Communications*, vol. 10, no. 1, p. 3024, 2019.
- [8] Y. Li, B. A. Schichtel, J. T. Walker et al., "Increasing importance of deposition of reduced nitrogen in the United States," *Proceedings of the National Academy of Sciences*, vol. 113, no. 21, pp. 5874–5879, 2016.
- [9] W. Battye, V. P. Aneja, and W. H. Schlesinger, "Is nitrogen the next carbon?," *Earth's Future*, vol. 5, no. 9, pp. 894–904, 2017.
- [10] P. Achakulwisut, M. Brauer, P. Hystad, and S. C. Anenberg, "Global, national, and urban burdens of paediatric asthma incidence attributable to ambient NO₂ pollution: estimates from global datasets," *The Lancet Planetary Health*, vol. 3, no. 4, pp. e166–e178, 2019.
- [11] R. Chen, P. Yin, X. Meng et al., "Associations between ambient nitrogen dioxide and daily cause-specific mortality: evidence from 272 Chinese cities," *Epidemiology*, vol. 29, no. 4, 2018.
- [12] R. V. Martin, K. Chance, D. J. Jacob et al., "An improved retrieval of tropospheric nitrogen dioxide from GOME," *Journal of Geophysical Research: Atmospheres*, vol. 107, no. D20, pp. ACH 9-1–ACH 9-21, 2002.
- [13] K. F. Boersma, H. J. Eskes, A. Richter et al., "Improving algorithms and uncertainty estimates for satellite NO₂ retrievals: results from the quality assurance for the essential climate variables (QA4ECV) project," *Atmospheric Measurement Techniques*, vol. 11, no. 12, pp. 6651–6678, 2018.
- [14] J. P. Burrows, M. Weber, M. Buchwitz et al., "The global ozone monitoring experiment (GOME): mission concept and first scientific results," *Journal of the Atmospheric Sciences*, vol. 56, no. 2, pp. 151–175, 1999.
- [15] A. Richter, J. P. Burrows, H. Nüß, C. Granier, and U. Niemeier, "Increase in tropospheric nitrogen dioxide over China observed from space," *Nature*, vol. 437, no. 7055, pp. 129–132, 2005.
- [16] A. Hilboll, A. Richter, and J. P. Burrows, "Long-term changes of tropospheric NO₂ over megacities derived from multiple satellite instruments," *Atmospheric Chemistry and Physics*, vol. 13, no. 8, pp. 4145–4169, 2013.
- [17] A. K. Georgoulias, K. F. Boersma, J. van Vliet et al., "Detection of NO₂ pollution plumes from individual ships with the TROPOMI/S5P satellite sensor," *Environmental Research Letters*, vol. 15, no. 12, article 124037, 2020.
- [18] Z. Jiang, B. C. McDonald, H. Worden et al., "Unexpected slowdown of US pollutant emission reduction in the past decade," *Proceedings of the National Academy of Sciences*, vol. 115, no. 20, pp. 5099–5104, 2018.
- [19] B. N. Duncan, R. V. Martin, A. C. Staudt, R. Yevich, and J. A. Logan, "Interannual and seasonal variability of biomass burning emissions constrained by satellite observations," *Journal of Geophysical Research: Atmospheres*, vol. 108, no. D2, 2003.
- [20] X. Jin, Q. Zhu, and R. Cohen, "Direct estimates of biomass burning NO₂ emissions and lifetimes using daily observations from TROPOMI," *Atmospheric Chemistry and Physics*, vol. 21, no. 20, pp. 15569–15587, 2021.
- [21] Y. Wang, J. Wang, M. Zhou, D. K. Henze, C. Ge, and W. Wang, "Inverse modeling of SO₂ and NO_x emissions over China using multisensor satellite data – Part 2: downscaling techniques for air quality analysis and forecasts," *Atmospheric Chemistry and Physics*, vol. 20, no. 11, pp. 6651–6670, 2020.
- [22] T. Sha, X. Ma, H. Zhang et al., "Impacts of soil NO_x emission on O₃ air quality in rural California," *Environmental Science & Technology*, vol. 55, no. 10, pp. 7113–7122, 2021.
- [23] Y. Wang, C. Ge, L. Castro Garcia, G. D. Jenerette, P. Y. Oikawa, and J. Wang, "Improved modelling of soil NO_x emissions in a high temperature agricultural region: role of background emissions on NO₂ trend over the US," *Environmental Research Letters*, vol. 16, no. 8, article 084061, 2021.
- [24] J. L. Laughner and R. C. Cohen, "Direct observation of changing NO_x lifetime in north American cities," *Science*, vol. 366, no. 6466, pp. 723–727, 2019.
- [25] S. Beirle, K. F. Boersma, U. Platt, M. G. Lawrence, and T. Wagner, "Megacity emissions and lifetimes of nitrogen oxides probed from space," *Science*, vol. 333, no. 6050, pp. 1737–1739, 2011.
- [26] J. Kim, U. Jeong, M. H. Ahn et al., "New era of air quality monitoring from space: geostationary environment monitoring spectrometer (GEMS)," *Bulletin of the American Meteorological Society*, vol. 101, no. 1, pp. E1–E22, 2020.
- [27] P. Zoogman, X. Liu, R. M. Suleiman et al., "Tropospheric emissions: monitoring of pollution (TEMPO)," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 186, pp. 17–39, 2017.
- [28] R. Timmermans, A. Segers, L. Curier et al., "Impact of synthetic space-borne NO₂ observations from the sentinel-4 and sentinel-5P missions on tropospheric NO₂ analyses," *Atmospheric Chemistry and Physics*, vol. 19, no. 19, pp. 12811–12833, 2019.
- [29] J. van Geffen, K. F. Boersma, H. Eskes et al., "S5P TROPOMI NO₂ slant column retrieval: method, stability, uncertainties and comparisons with OMI," *Atmospheric Measurement Techniques*, vol. 13, no. 3, pp. 1315–1335, 2020.
- [30] A. Richter, M. Begoin, A. Hilboll, and J. P. Burrows, "An improved NO₂ retrieval for the GOME-2 satellite instrument," *Atmospheric Measurement Techniques*, vol. 4, no. 6, pp. 1147–1159, 2011.

- [31] J. A. Geddes, R. V. Martin, E. J. Bucsela, C. A. McLinden, and D. J. M. Cunningham, "Stratosphere-troposphere separation of nitrogen dioxide columns from the TEMPO geostationary satellite instrument," *Atmospheric Measurement Techniques*, vol. 11, no. 11, pp. 6271–6287, 2018.
- [32] S. Beirle, C. Hörmann, P. Jöckel et al., "The STRatospheric Estimation Algorithm from Mainz (STREAM): estimating stratospheric NO₂ from nadir-viewing satellites by weighted convolution," *Atmospheric Measurement Techniques*, vol. 9, no. 7, pp. 2753–2779, 2016.
- [33] E. J. Bucsela, N. A. Krotkov, E. A. Celarier et al., "A new stratospheric and tropospheric NO₂ retrieval algorithm for nadir-viewing satellite instruments: applications to OMI," *Atmospheric Measurement Techniques*, vol. 6, no. 10, pp. 2607–2626, 2013.
- [34] P. I. Palmer, D. J. Jacob, K. Chance et al., "Air mass factor formulation for spectroscopic measurements from satellites: application to formaldehyde retrievals from the global ozone monitoring experiment," *Journal of Geophysical Research: Atmospheres*, vol. 106, no. D13, pp. 14539–14550, 2001.
- [35] L. N. Lamsal, R. V. Martin, A. van Donkelaar et al., "Indirect validation of tropospheric nitrogen dioxide retrieved from the OMI satellite instrument: insight into the seasonal variation of nitrogen oxides at northern midlatitudes," *Journal of Geophysical Research: Atmospheres*, vol. 115, no. D5, 2010.
- [36] K. F. Boersma, H. J. Eskes, R. J. Dirksen et al., "An improved tropospheric NO₂ column retrieval algorithm for the ozone monitoring instrument," *Atmospheric Measurement Techniques*, vol. 4, no. 9, pp. 1905–1928, 2011.
- [37] K. F. Boersma, H. J. Eskes, J. P. Veefkind et al., "Near-real time retrieval of tropospheric NO₂ from OMI," *Atmospheric Chemistry and Physics*, vol. 7, no. 8, pp. 2103–2118, 2007.
- [38] J. L. Laughner, Q. Zhu, and R. C. Cohen, "The Berkeley high resolution tropospheric NO₂ product," *Earth System Science Data*, vol. 10, no. 4, pp. 2069–2095, 2018.
- [39] M. Liu, J. Lin, K. F. Boersma et al., "Improved aerosol correction for OMI tropospheric NO₂ retrieval over East Asia: constraint from CALIOP aerosol vertical profile," *Atmospheric Measurement Techniques*, vol. 12, no. 1, pp. 1–21, 2019.
- [40] A. Lorente, K. Folkert Boersma, H. Yu et al., "Structural uncertainty in air mass factor calculation for NO₂ and HCHO satellite retrievals," *Atmospheric Measurement Techniques*, vol. 10, no. 3, pp. 759–782, 2017.
- [41] J. Jin, J. Ma, W. Lin et al., "MAX-DOAS measurements and satellite validation of tropospheric NO₂ and SO₂ vertical column densities at a rural site of North China," *Atmospheric Environment*, vol. 133, pp. 12–25, 2016.
- [42] H. W. L. Mak, J. Laughner, J. Fung, Q. Zhu, and R. Cohen, "Improved satellite retrieval of tropospheric NO₂ column density via updating of air mass factor (AMF): case study of southern China," *Remote Sensing*, vol. 10, no. 11, p. 1789, 2018.
- [43] A. R. Russell, A. E. Perrig, L. C. Valin et al., "A high spatial resolution retrieval of NO₂ column densities from OMI: method and evaluation," *Atmospheric Chemistry and Physics*, vol. 11, no. 16, pp. 8543–8554, 2011.
- [44] L. N. Lamsal, N. A. Krotkov, A. Vasilkov et al., "Ozone monitoring instrument (OMI) Aura nitrogen dioxide standard product version 4.0 with improved surface and cloud treatments," *Atmospheric Measurement Techniques*, vol. 14, no. 1, pp. 455–479, 2021.
- [45] L. C. Valin, A. R. Russell, R. C. Hudman, and R. C. Cohen, "Effects of model resolution on the interpretation of satellite NO₂ observations," *Atmospheric Chemistry and Physics*, vol. 11, no. 22, pp. 11647–11655, 2011.
- [46] S. D. Eastham, M. S. Long, C. A. Keller et al., "GEOS-Chem High Performance (GCHP v11-02c): a next-generation implementation of the GEOS-Chem chemical transport model for massively parallel applications," *Geoscientific Model Development*, vol. 11, no. 7, pp. 2941–2953, 2018.
- [47] J. L. Laughner, A. Zare, and R. C. Cohen, "Effects of daily meteorology on the interpretation of space-based remote sensing of NO₂," *Atmospheric Chemistry and Physics*, vol. 16, no. 23, pp. 15247–15264, 2016.
- [48] L. T. Murray, "Lightning NO_x and impacts on air quality," *Current Pollution Reports*, vol. 2, no. 2, pp. 115–133, 2016.
- [49] J. L. Laughner and R. C. Cohen, "Quantification of the effect of modeled lightning NO₂ on UV-visible air mass factors," *Atmospheric Measurement Techniques*, vol. 10, no. 11, pp. 4403–4419, 2017.
- [50] Q. Zhu, J. L. Laughner, and R. C. Cohen, "Lightning NO₂ simulation over the contiguous US and its effects on satellite NO₂ retrievals," *Atmospheric Chemistry and Physics*, vol. 19, no. 20, pp. 13067–13078, 2019.
- [51] M. J. Cooper, R. V. Martin, D. K. Henze, and D. B. A. Jones, "Effects of a priori profile shape assumptions on comparisons between satellite NO₂ columns and model simulations," *Atmospheric Chemistry and Physics*, vol. 20, no. 12, pp. 7231–7241, 2020.
- [52] K. Yang, X. Liu, P. K. Bhartia et al., "Direct retrieval of sulfur dioxide amount and altitude from spaceborne hyperspectral UV measurements: theory and application," *Journal of Geophysical Research: Atmospheres*, vol. 115, no. D2, 2010.
- [53] C. R. Nowlan, X. Liu, K. Chance et al., "Retrievals of sulfur dioxide from the global ozone monitoring experiment 2 (GOME-2) using an optimal estimation approach: algorithm and initial validation," *Journal of Geophysical Research: Atmospheres*, vol. 116, no. D18, 2011.
- [54] X. Liu, P. K. Bhartia, K. Chance, R. J. D. Spurr, and T. P. Kurosu, "Ozone profile retrievals from the ozone monitoring instrument," *Atmospheric Chemistry and Physics*, vol. 10, no. 5, pp. 2521–2537, 2010.
- [55] L. K. Behrens, A. Hilboll, A. Richter, E. Peters, H. Eskes, and J. P. Burrows, "GOME-2A retrievals of tropospheric NO₂ in different spectral ranges – influence of penetration depth," *Atmospheric Measurement Techniques*, vol. 11, no. 5, pp. 2769–2795, 2018.
- [56] K. Yang, S. A. Carn, C. Ge, J. Wang, and R. R. Dickerson, "Advancing measurements of tropospheric NO₂ from space: new algorithm and first global results from OMPs," *Geophysical Research Letters*, vol. 41, no. 13, pp. 4777–4786, 2014.
- [57] A. Hilboll, A. Richter, and J. P. Burrows, "Vertical information content of nadir measurements of tropospheric NO₂ from satellite," in *EGU General Assembly Conference*, p. 8746, Vienna, May, 2014.
- [58] E. Rolf, J. Proctor, T. Carleton et al., "A generalizable and accessible approach to machine learning with global satellite imagery," *Nature Communications*, vol. 12, no. 1, p. 4392, 2021.

- [59] S. Nanda, M. de Graaf, J. P. Veeffkind et al., "A neural network radiative transfer model approach applied to the tropospheric monitoring instrument aerosol height algorithm," *Atmospheric Measurement Techniques*, vol. 12, no. 12, pp. 6619–6634, 2019.
- [60] T. Le, C. Liu, B. Yao, V. Natraj, and Y. L. Yung, "Application of machine learning to hyperspectral radiative transfer simulations," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 246, article 106928, 2020.
- [61] K. C. Wells, D. B. Millet, V. H. Payne et al., "Satellite isoprene retrievals constrain emissions and atmospheric oxidation," *Nature*, vol. 585, no. 7824, pp. 225–233, 2020.
- [62] P. Hedelt, D. S. Efremenko, D. G. Loyola, R. Spurr, and L. Clarisse, "Sulfur dioxide layer height retrieval from sentinel-5 precursor/TROPOMI using FP_ILM," *Atmospheric Measurement Techniques*, vol. 12, no. 10, pp. 5503–5517, 2019.
- [63] J. Xu, O. Schussler, D. G. L. Rodriguez, F. Romahn, and A. Doicu, "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*, vol. 10, no. 12, pp. 5442–5457, 2017.
- [64] L. She, H. K. Zhang, Z. Li, G. de Leeuw, and B. Huang, "Himawari-8 aerosol optical depth (AOD) retrieval using a deep neural network trained using AERONET observations," *Remote Sensing*, vol. 12, no. 24, p. 4125, 2020.
- [65] J. Wang, X. Xu, S. Ding et al., "A numerical testbed for remote sensing of aerosols, and its demonstration for evaluating retrieval synergy from a geostationary satellite constellation of GEO-CAPE and GOES-R," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 146, pp. 510–528, 2014.
- [66] X. Xu and J. Wang, "UNL-VRTM, a testbed for aerosol remote sensing: model developments and applications," in *Springer Series in Light Scattering: Volume 4: Light Scattering and Radiative Transfer*, A. Kokhanovsky, Ed., pp. 1–69, Springer International Publishing, Cham, 2019.
- [67] R. J. D. Spurr, "VLIDORT: a linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 102, no. 2, pp. 316–342, 2006.
- [68] W. Hou, J. Wang, X. Xu, J. S. Reid, and D. Han, "An algorithm for hyperspectral remote sensing of aerosols: 1. Development of theoretical framework," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 178, pp. 400–415, 2016.
- [69] X. Xu, J. Wang, Y. Wang et al., "Detecting layer height of smoke aerosols over vegetated land and water surfaces via oxygen absorption bands: hourly results from EPIC/DSCOVER in deep space," *Atmospheric Measurement Techniques*, vol. 12, no. 6, pp. 3269–3288, 2019.
- [70] J. Wang, M. Zhou, X. Xu et al., "Development of a nighttime shortwave radiative transfer model for remote sensing of nocturnal aerosols and fires from VIIRS," *Remote Sensing of Environment*, vol. 241, article 111727, 2020.
- [71] F. Zheng, Hou, Sun et al., "Optimal estimation retrieval of aerosol fine-mode fraction from ground-based sky light measurements," *Atmosphere*, vol. 10, no. 4, p. 196, 2019.
- [72] C. Li, J. Joiner, N. A. Krotkov, and P. K. Bhartia, "A fast and sensitive new satellite SO₂ retrieval algorithm based on principal component analysis: application to the ozone monitoring instrument," *Geophysical Research Letters*, vol. 40, no. 23, pp. 6314–6318, 2013.
- [73] M. Hess, P. Koepke, and I. Schult, "Optical properties of aerosols and clouds: the software package OPAC," *Bulletin of the American Meteorological Society*, vol. 79, no. 5, pp. 831–844, 1998.
- [74] R. McClatchey, R. Fenn, and J. Selby, *Optical Properties of the Atmosphere*, Air Force Cambridge Research Laboratories, Office of Aerospace Research, 3rd edition, 1972.
- [75] K. V. Chance and R. J. D. Spurr, "Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum," *Applied Optics*, vol. 36, no. 21, pp. 5224–5230, 1997.
- [76] M. J. Cooper, R. V. Martin, C. A. McLinden, and J. R. Brook, "Inferring ground-level nitrogen dioxide concentrations at fine spatial resolution applied to the TROPOMI satellite instrument," *Environmental Research Letters*, vol. 15, no. 10, p. doi:10.1088/1748-9326/aba3a5, 2020.
- [77] G. E. Hinton, "Connectionist Learning Procedures," in *Machine learning*, pp. 555–610, Morgan Kaufmann, 1990.
- [78] L. Buitinck, G. Louppe, M. Blondel et al., "API design for machine learning software: experiences from the scikit-learn project," 2013, <https://arxiv.org/abs/1309.0238>.
- [79] X. Liu, Q. Yang, H. Li et al., "Development of a fast and accurate PCRTM radiative transfer model in the solar spectral region," *Applied Optics*, vol. 55, no. 29, pp. 8236–8247, 2016.
- [80] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 569–575, 2010.
- [81] J. van Geffen, H. J. Eskes, K. F. Boersma, and J. P. Veeffkind, *TROPOMI ATBD of the Total and Tropospheric NO₂ Data Products, Report S5P-KNMI-L2-0005-RP, version 2.2.0*, KNMI, De Bilt, The Netherlands, 2021, <http://www.tropomi.eu/data-products/nitrogen-dioxide/>.
- [82] T. Verhoelst, S. Compornolle, G. Pinardi et al., "Ground-based validation of the Copernicus sentinel-5P TROPOMI NO₂ measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks," *Atmospheric Measurement Techniques*, vol. 14, no. 1, pp. 481–510, 2021.
- [83] C. B. Schaaf, F. Gao, A. H. Strahler et al., "First operational BRDF, albedo nadir reflectance products from MODIS," *Remote Sensing of Environment*, vol. 83, no. 1-2, pp. 135–148, 2002.
- [84] J. J. Danielson and D. B. Gesch, *Global Multi-Resolution Terrain Elevation Data 2010 (GMTED 2010)*, US Department of the Interior, US Geological Survey, 2011.
- [85] P. Wang, P. Stammes, R. van der A, G. Pinardi, and M. van Roozendael, "FRESCO+: an improved O₃ A-band cloud retrieval algorithm for tropospheric trace gas retrievals," *Atmospheric Chemistry and Physics*, vol. 8, no. 21, pp. 6565–6576, 2008.
- [86] Y. L. Roberts, P. Pilewskie, B. C. Kindel, D. R. Feldman, and W. D. Collins, "Quantitative comparison of the variability in observed and simulated shortwave reflectance," *Atmospheric Chemistry and Physics*, vol. 13, no. 6, pp. 3133–3147, 2013.
- [87] J. Herman, N. Abuhassan, J. Kim et al., "Underestimation of column NO₂ amounts from the OMI satellite compared to diurnally varying ground-based retrievals from multiple PANDORA spectrometer instruments,"

Atmospheric Measurement Techniques, vol. 12, no. 10, pp. 5593–5612, 2019.

- [88] B. Silver, C. L. Reddington, S. R. Arnold, and D. V. Spracklen, “Substantial changes in air pollution across China during 2015–2017,” *Environmental Research Letters*, vol. 13, no. 11, article 114012, 2018.
- [89] Y. Wang, J. Lampel, P. Xie et al., “Ground-based MAX-DOAS observations of tropospheric aerosols, NO₂, SO₂ and HCHO in Wuxi, China, from 2011 to 2014,” *Atmospheric Chemistry and Physics*, vol. 17, no. 3, pp. 2189–2215, 2017.
- [90] J. W. Harder, J. W. Brault, P. V. Johnston, and G. H. Mount, “Temperature dependent NO₂ cross sections at high spectral resolution,” *Journal of Geophysical Research: Atmospheres*, vol. 102, no. D3, pp. 3861–3879, 1997.