

Article

Not peer-reviewed version

# The Spatiotemporal Distribution of NO<sub>2</sub> in China Based on Refined 2DCNN-LSTM Model Retrieval and Factor Interpretability Analysis

[Ruming Chen](#) , [Bin Chen](#) <sup>\*</sup> , Zhihao Song , [Jiashun Hu](#) , Xingzhao Zhou , Lin Zhao

Posted Date: 23 August 2023

doi: 10.20944/preprints202308.1565.v1

Keywords: nitrogen oxide retrieval; 2DCNN-LSTM; machine learning; factor interpretability



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# The Spatiotemporal Distribution of NO<sub>2</sub> in China Based on Refined 2DCNN-LSTM Model Retrieval and Factor Interpretability Analysis

Ruming Chen <sup>1,2</sup>, Jiashun Hu <sup>1,2</sup>, Zhihao Song <sup>1,2</sup>, Yixuan Wang <sup>1,2</sup>, XingZhao Zhou <sup>1,2</sup>, Lin Zhao <sup>1,2</sup> and Bin Chen <sup>1,2,\*</sup>

<sup>1</sup> College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China

<sup>2</sup> Collaborative Innovation Center of Western Ecological Security, Lanzhou 730000, China

\* Correspondence: chenbin@lzu.edu.cn

**Abstract:** With the advancement of urbanization in China, effective control of pollutant emissions and air quality have become important goals in current environmental management. Nitrogen dioxide (NO<sub>2</sub>), as a precursor of tropospheric ozone and fine particulate matter, plays a significant role in atmospheric chemistry research and air pollution control. However, the uneven ground monitoring stations and low temporal resolution of polar-orbiting satellites set challenges for accurately assessing near-surface NO<sub>2</sub> concentrations. To address this issue, a spatiotemporal refined NO<sub>2</sub> retrieval model was established for China using the geostationary satellite Himawari-8. The spatiotemporal characteristics of NO<sub>2</sub> were analyzed and its contribution factors were explored. Firstly, seven Himawari-8 channels sensitive to NO<sub>2</sub> were selected by using the forward feature selection based on information entropy. Subsequently, a 2DCNN-LSTM network model was constructed, incorporating the selected channels and meteorological variables as retrieval factors to estimate hourly NO<sub>2</sub> in China from March 2018 to February 2020 (with a resolution of 0.05°, per hour). The performance evaluation demonstrated that the full-channel 2DCNN-LSTM model had good fitting capability and robustness (R<sup>2</sup>=0.74, RMSE=10.93), and further improvements were achieved after channel selection (R<sup>2</sup>=0.87, RMSE=6.84). The 10-fold cross-validation results indicated that the R<sup>2</sup> between retrieval and measured values was above 0.85, the MAE was within 5.60, and the RMSE was within 7.90. R<sup>2</sup> varied between 0.85 and 0.90, showing better validation at mid-day (R<sup>2</sup>=0.89) and in spring and fall transition seasons (R<sup>2</sup> =0.88 and R<sup>2</sup> =0.90). To investigate the cooperative effect of meteorological factors and other air pollutants on NO<sub>2</sub>, statistical methods (Beta coefficients) were used to test the factor interpretability. Meteorological factors as well as other pollutants were analyzed. From a statistical perspective, PM<sub>2.5</sub>, Boundary Layer Height, and O<sub>3</sub> were found to have the largest impacts on near-surface NO<sub>2</sub> concentrations, with each standard deviation change in these factors leading to 0.28, 0.24, and 0.23 in standard deviations of near-surface NO<sub>2</sub>, respectively. Findings of the study contributed to a comprehensive understanding of the spatiotemporal distribution of NO<sub>2</sub> and provide a scientific basis for formulating targeted air pollution policies.

**Keywords:** nitrogen oxide retrieval; 2DCNN-LSTM; machine learning; factor interpretability

## 1. Introduction

Nitrogen oxides (NO<sub>x</sub>) refers to a group of oxidizing compounds produced from the reaction of nitrogen (N<sub>2</sub>) and oxygen (O<sub>2</sub>) under high-temperature and high-pressure conditions. They include nitric oxide, nitrogen dioxide, dinitrogen tetroxide, dinitrogen pentoxide, and nitrogen pentoxide. NO<sub>x</sub> is a byproduct of industrial production, transportation, and energy consumption activities. The strong reactivity and solubility of NO<sub>x</sub>- organic compound systems are important reasons contributing to the lively oxidative properties of tropospheric atmosphere. The emission of NO<sub>x</sub> has an impact on the environment and human health. As a main component of air pollution, NO<sub>x</sub> can

react with other primary pollutants to produce secondary pollutants such as ozone, sulfuric acid, sulfate, and fine particulate matter. Under solar radiation,  $\text{NO}_x$  undergoes a series of photochemical reactions with hydrocarbon compounds, resulting in photochemical smog. The atmosphere polluted by  $\text{NO}_x$  is a heterogeneous, complex, liquid-particulate mixture[1], Epidemiological studies have shown evidence that short and long term exposure to pollutants increases the risk of cardiovascular events, including thrombosis, arrhythmias, acute arterial vasoconstriction, and systemic inflammation[2]. Local inflammation mediated by nitrogen dioxide enhances the permeability of the alveolar-capillary barrier, thereby aggravating the pulmonary congestion symptoms in susceptible individuals[3]. Major urban agglomerations in China, such as the Beijing-Tianjin-Hebei region and the Yangtze River Delta, experience severe  $\text{NO}_x$  pollution[4–7], Therefore, It's crucial to study the elaborate retrieval of  $\text{NO}_2$  concentrations and influencing factors for scientifically mitigating air pollution.

$\text{NO}_2$  concentration trends can be used to assess the effectiveness of air pollution regulations and the impact of sudden events (such as epidemic lockdowns) on emissions[8,9]. Some research used the XGBoost model to explain the nonlinear relationship between near-surface  $\text{NO}_2$  measurements and influencing factors for the retrieval of  $\text{NO}_2$  concentrations[10]. Several studies have developed numerical models based on transport and photochemical processes to assess high-resolution  $\text{NO}_x$  distributions at the urban scale[11,12]. and used the 3D-CTMS model to construct a framework for summer  $\text{NO}_x$  emissions in China[13]. Determining the driving factors of  $\text{NO}_2$  column concentration using Geo-detector[14]. Some studies used a stable isotope and Bayesian-based SIAR model to identify the source of nitrate in rivers[15]. Some study analyzed the source of tropospheric nitrogen dioxide measurements in northeastern China using MAX-DOAS[16].

In urban agglomerations, most  $\text{NO}_2$  is generated from exhaust emissions in the internal combustion engines of motor vehicles, and it is a tracer for anthropogenic fuel consumption related to cities and transportation. Studies analyzed the multi-scale spatiotemporal variations of  $\text{NO}_2$  emissions in the Beijing-Tianjin-Hebei region using trajectory, vehicle specifications, and highways network information[4], and have explored the influence of meteorological and socioeconomic factors on vertical column concentrations of  $\text{NO}_2$  in coastal ports of China[17]. Such research has unique advantages in formulating targeted emission reduction policies for different regions.

Although significant progress has been made in the retrieval and source apportionment of  $\text{NO}_2$  from satellite data, they are mainly limited to urban or regional scales, and their coarse spatial and temporal resolutions restrict their application in meteorology, atmospheric physics, environmental research and assessment of  $\text{NO}_2$  pollution evolution[18], resulting in uncertainty in the overall mechanisms of  $\text{NO}_2$  impacts.

$\text{NO}_x$  has become an important air pollutant with significant impacts on environmental effects and human health. While extensive studies have been conducted to estimate concentration distributions and the mechanisms of influence, the existing models have limitations due to the neglect of temporal and spatial characteristics and biased factors selection, which hampers estimation effect. A study on the construction of a multi-scale time-lag correlation network for detecting air pollution interactions between neighboring cities due to time-lag effects and transmission patterns[19]. In this study, the spatiotemporal distribution of  $\text{NO}_2$  in China was assessed using multi-source data, what's more, independent validation and 10-fold cross-validation were both used to measure the model robustness. Finally, the contributions of meteorological factors and other pollutants were investigated by calculating the factor interpretability. A 2DCNN-LSTM model integrating spatiotemporal data was constructed to better estimate near-surface  $\text{NO}_2$  in China. The research results are expected to provide scientific basis for the effective development of air pollution measurements. The structure of the paper is as follows: Section II describes the research area, data, and methods; Section III presents the research results; Section IV discusses and summarizes the findings, providing conclusive opinions and future research directions.

2. Materials and Methods

2.1. Data Sources and Preparation

The study utilized pollutant station observation data, top-of-atmosphere radiation (TOAR) data of the Himawari-8 satellite, meteorological variables, and geospatial data. The period covered in this study was from March 2018 to February 2020 (winter refers to December of that year to February of the following year, spring refers to March to May, summer refers to June to August, autumn refers to September to November, and all times are Beijing time).

2.1.1. NO<sub>2</sub> Monitoring Stations and National Highways Map in China

Hourly near-surface NO<sub>2</sub> concentrations and other pollutants (O<sub>3</sub>, SO<sub>2</sub>, CO, PM<sub>2.5</sub>, PM<sub>10</sub>) are provided by the China National Environmental Monitoring Center (CNEMC) (<http://www.cnemc.cn/>). The Air Quality Monitoring Network covers four levels: national, provincial, municipal, and county. In terms of monitoring functions, the network includes urban air quality monitoring, regional air quality monitoring, background air quality monitoring, etc. During the study period, a total of 1,436 nationally supervised automatic air monitoring stations, including monitoring in 338 cities above prefecture level, regional (rural) monitoring and background environmental monitoring, were implemented.

National primary highways in China include international highways, national defense highways, etc. The total length of China's national highways is approximately 299,000 kilometers, and the length of expressways is approximately 162,000 kilometers. These two types of highways together form the national highway network. The number of motor vehicles in China has continued to grow in recent years, and vehicular emissions have become one of the important sources of air pollution in China. The large motor vehicle emissions generated by this large-scale highway operation can be used to visually measure the NO<sub>2</sub> results against the analytical model estimates. The distribution of air quality monitoring stations and national highways in China is shown in Figure 1.



**Figure 1.** Distribution of land cover types, air quality monitoring stations (black dots) and national highways (brown lines) in China.



2.1.2. Himawari-8 TOAR Data

Himawari-8 is a geostationary satellite launched by the Japan Meteorological Agency (JMA) in 2014 and put into operation in July 2015, providing data coverage for one-third of the Earth (80°E-160°W, 60°N-60°S) with its nadir point located at 140.7°E on the equator. The satellite carries various remote sensing spectral payloads, including a high-resolution satellite camera, microwave radiometers and a laser altimeter in the visible, near-infrared, and infrared imaging bands. Among them, the Advanced Himawari Imager (AHI) provides data information for 16 channels, including three visible channels, three near-infrared channels, and ten infrared channels. The high-resolution, high-accuracy, and all-day remote sensing observation data provided by Himawari-8 are widely used in surface environmental monitoring, meteorological forecasting, natural resource management, urban planning, and other fields.

2.1.3. Meteorological and Geographic Data

Meteorological and geographical factors can influence the transport and diffusion of pollutants. The variations of meteorological signals (temperature, humidity, wind speed, pressure, diffusion conditions, Land Use or Cover Change (LUCC), etc.) are a part of climate, environmental, and ecological changes. ERA5 meteorological data is a global meteorological reanalysis dataset developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). It uses advanced global atmospheric models and a large amount of observation data. The data products include various meteorological variables such as 2m temperature, wind, precipitation, humidity, and radiation. ERA5 assimilates a large amount of remote sensing information, as well as conventional observations from the ground and high altitude. It covers the time range from 1979 to the present, with high spatiotemporal resolutions (temporal resolution down to hourly and spatial resolution of 0.25°×0.25°).

The meteorological data used in this study were the hourly ERA5 reanalysis data for the study period (March 2018 to February 2020, consistent with estimated period). The Boundary layer height (BLH), relative humidity (RH), surface pressure (SP), 2m temperature (TM), longitude and latitude wind components (U10 and V10) were selected as the meteorological features affecting NO<sub>2</sub> diffusion. Geographic data refers to LUCC, including different types of land use or cover types and their spatiotemporal distribution. Details of the involved variables, such as units and sources, are presented in Table 1.

**Table 1.** Detailed information of the data used in the study.

Variables	Implication	Time series length	Unit	Spatial resolution	Temporal resolution	Data source
NO <sub>2</sub>	NO <sub>2</sub> observation data	March 2018 to February 2020	μg/m <sub>3</sub>	site	Hourly	CNE MC
TOAR1	AHI blue band (0.46μm)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR2	AHI green band (0.51μm)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR3	AHI red band (0.64μm)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR4	AHI Near-infrared band (0.86μm)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR5	AHI Near-infrared band (1.5μm)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR6	AHI Near-infrared band (2.3μm)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA

TOAR7	AHI Infrared band (3.9 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR8	AHI Infrared band (6.2 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR9	AHI Infrared band (6.9 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR10	AHI Infrared band (7.3 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR11	AHI Infrared band (8.6 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR12	AHI Infrared band (9.6 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR13	AHI Infrared band (10.4 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR14	AHI Infrared band (11.2 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR15	AHI Infrared band (12.4 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
TOAR16	AHI Infrared band (13.3 $\mu$ m)	March 2018 to February 2020	/	0.05°×0.05°	Hourly	JMA
BLH	Boundary layer height	March 2018 to February 2020	m	0.25°×0.25°	Hourly	ERA-5
TM	2m temperature	March 2018 to February 2020	K	0.1°×0.1°	Hourly	ERA-5
RH	Relative humidity	March 2018 to February 2020	%	0.25°×0.25°	Hourly	ERA-5
U10	10m u component of wind	March 2018 to February 2020	m/s	0.1°×0.1°	Hourly	ERA-5
V10	10m v component of wind	March 2018 to February 2020	m/s	0.1°×0.1°	Hourly	ERA-5
SP	Surface pressure	March 2018 to February 2020	Pa	0.1°×0.1°	Hourly	ERA-5
LUCC	The type of surface	March 2018 to February 2020	/	0.05°×0.05°	Yearly	NAS A

## 2.2. Methods

### 2.2.1. Data Matching

Bilinear interpolation was employed to adjust the spatial resolution of meteorological and geographic data to match the 0.05°×0.05° spatial resolution of Himawari-8 TOAR data. Based on a 0.05°×0.05° grid, the NO<sub>2</sub> hourly average data (70°E-140°E, 15°N-55°N) recorded by the CNEMC, TOAR data, and ERA5 meteorological data were matched. If multiple sites existed within a given grid, the average of the NO<sub>2</sub> hourly concentrations at stations were used[20,21].

### 2.2.2. Feature Selection Based on Information Entropy

Forward feature selection is a feature selection method based on subset search. It starts with an initial set of features and gradually adds features until the desired number of features is attained. Information entropy serves as an indicator of variable uncertainty. The process of forward feature selection based on information entropy can be divided into the following steps: for each feature, calculate the conditional entropy of the given the target variable (Equation 1), and subtract it from the entropy of the target variable to obtain the information gain (Equation 2). The information gain measures the extent to which integrating a feature reduces model uncertainty.

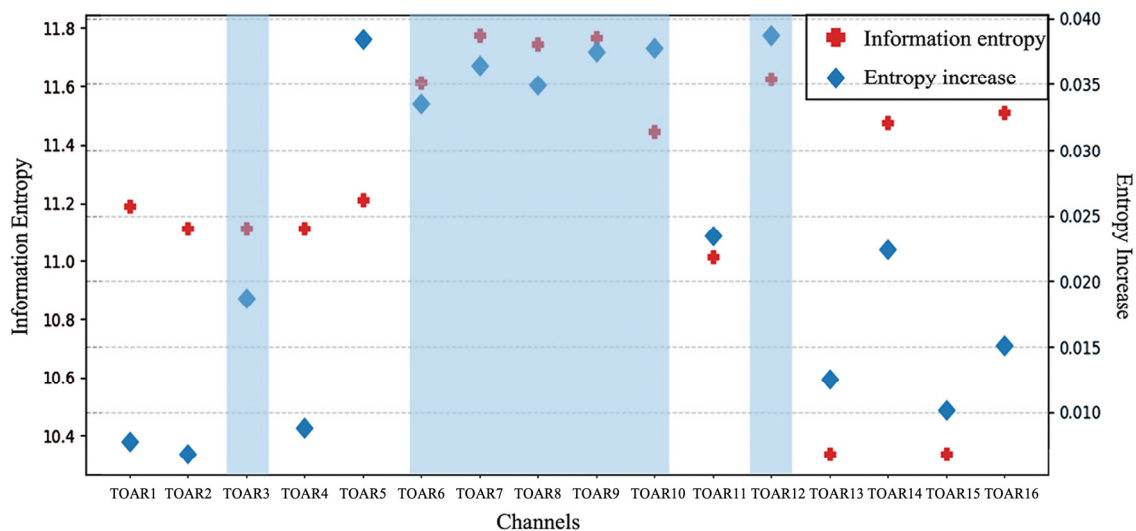
Formulas for the calculations are as follows:

$$H(x_i) = - \sum_{i=1}^n p(x_i) \log_2 P(x_i) \quad (1)$$

$$Gain(x_i) = H(x_i) - H(x_{i-1}) \quad (2)$$

By utilizing forward feature selection based on information entropy, channels highly correlated with the target variable were selected, reducing the number of channels, so as to improving the model generalization ability. Information entropy and information gain were used as indicators for effectively selecting the channels involved in the modeling process, identifying the combination of spectral channels combination, thereby leading to precise NO<sub>2</sub> estimation and an enhancement in model accuracy.

The information gain and entropy of each channel was computed (Figure 2). Based on the selection of a high-resolution channel combination (TOAR6, TOAR7, TOAR8, TOAR9, TOAR10, TOAR12), the feature selection problem seeks an optimal balance between model complexity and the model ability to describe the dataset, choosing a few, appropriate features can both avoid overfitting and increase model interpretation. TOAR3 with higher score in the visible channels, to ensure that the information from the visible, near-infrared, and infrared channels was all used as modeling variables. However, TOAR5 maintains less information entropy compared to other near-infrared channels (TOAR6 and TOAR7), so it was not taken into consideration. Finally, TOAR3, TOAR6, TOAR7, TOAR8, TOAR9, TOAR10, and TOAR12 were selected as the seven channels to participate in the modeling process.



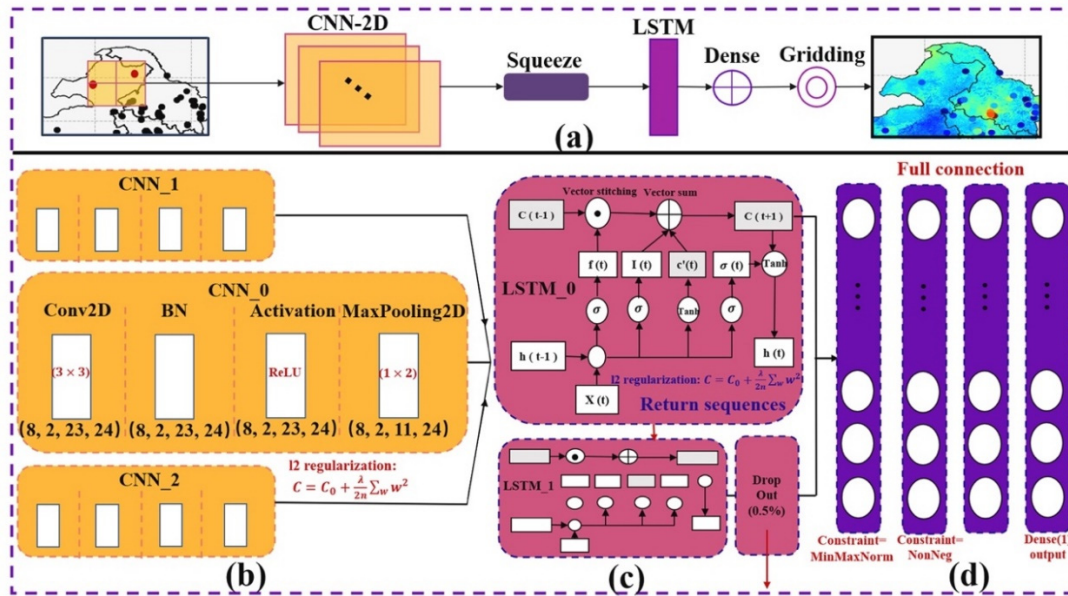
**Figure 2.** Results of forward feature selection screening channels based on information entropy; Entropy and information gain contained in 16 channels of Himawari-8 satellite. Left axis: information entropy; Right axis: information gain. The red cross scatter indicates the information entropy value of each channel, and the dark blue diamond indicates the information gain of each channel. The seven channels covered by blue regions are subsets of the selected optimal feature channels.

### 2.2.3.2. DCNN-LSTM

Convolutional Neural Network (CNN) is a deep learning model widely employed in image processing, computer vision and related domains, known for its powerful feature extraction capabilities. The CNN model utilizes convolutional layers as the core for extracting spatial features from data tensors. Long Short-Term Memory (LSTM) is a special type of neural network, which makes recursive calls based on sequential data, and has extensive applications in time series processing[22]. The LSTM's core is the cell state from time step  $t-1$  to  $t$ , and the gate mechanism is

used to control the flow and loss of features, addressing the issue of gradient vanishing caused by the short-term dependence of data[23].

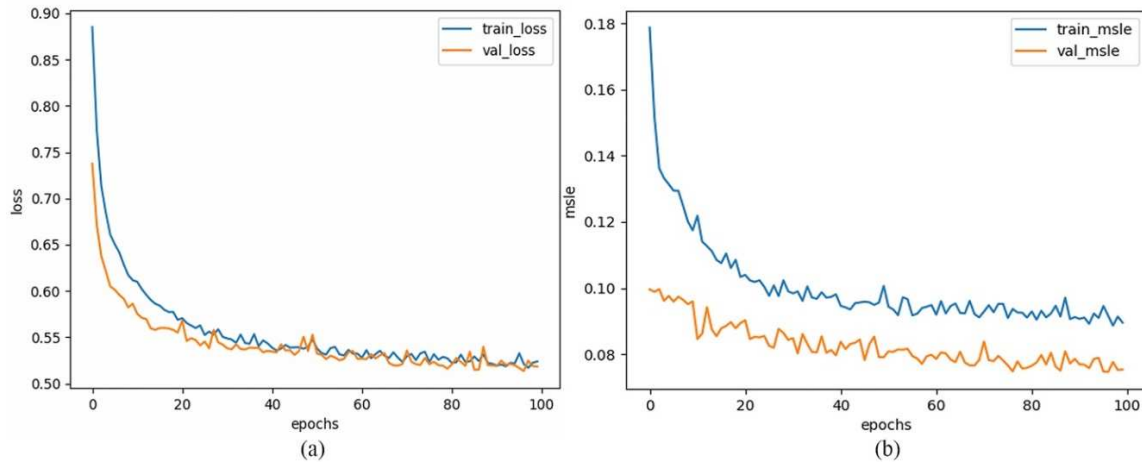
In the field of remote sensing retrieval using multi-variable modeling, the CNN-2D convolutional layer extracts spatial features using its kernels, while the LSTM captures long-term dependency features of time series. The fully connected layer maps feature to target variables. Finally, the flatten layers are coupled to construct the 2DCNN-LSTM model (Figure 3). This model holds considerable promise in processing data encompassing both spatial and temporal attributes.



**Figure 3.** The proposed network architecture. (a) General overview of the end-to-end deep neural network model 2DCNN-LSTM. (b) Consisting of three portion of CNN receiving four-dimensional tensors with spatial and temporal information as inputs, and each portion consists of a two-dimensional convolutional layer used to extract the spatial features, a Batch Normalization layer, an activation layer, and a pooling layer. (c) Two connected LSTM layers of cell states from  $t-1$  to  $t$ , a dropout layer is inserted between the LSTMs to randomly inactivate neurons to suppress overfitting. (d) Four portions of the Dense layer that extracts feature associations so that they are mapped onto the output space, with the first two Dense layers imposing a MinMaxNorm constraint and a NonNeg constraint, respectively.

Traditional remote sensing retrieval methods based on physical framework involve the measurement of physical parameters. On the contrary, the 2DCNN-LSTM model does not rely on any prior knowledge. This type of CNN model can directly learn feature relationships from input[24,25]. The CNN-2D convolutional layer has strong adaptability and can handle different types of data environments[24,26]. Compared with one-dimensional machine learning models, the CNN-2D layer can handle high-dimensional data and has stronger expressive power for complex nonlinear problems[27]. In practical applications, the optimization of hyperparameters in CNN model is crucial. The model solved the optimisation problem by constructing a minimisation function through the Adam optimiser, using ReLU as the activation function to introduce non-linearity into the neural network, allowing it to learn more complex functions. The weights of the 2DCNN-LSTM model were automatically adjusted during training, with L2 regularization incorporated into the loss function to gauge model complexity, thus mitigating noise from training data. After numerous independent validation and parameter adjustments, the hyperparameters of the 2DCNN-LSTM model were determined (Table 2). The loss curves (Figure 4) showed that both the training loss and validation loss have converged, and the difference was small, indicating that the hyperparameter combination was appropriate with the model reaching good-fit.





**Figure 4.** Loss falling curves; (a) Loss curves; blue line: training error; orange line: validation error. (b) Mean square logarithmic error (MSLE) curves; blue line: training mean square logarithmic error; orange line: validation mean square logarithmic error.

**Table 2.** 2DCNN-LSTM model hyperparameter settings.

Key hyperparameters	Value
Loss	Mean Absolute Error
Optimizer	Adam
Learning Rate	0.0009
Epoch	100
Batch size	8
Activation functions	ReLU
Regularizing functions	Regularizers.L2 (0.005)
Hidden layers	30
Dropout	0.05
Trainable params	39708

#### 2.2.4. Integration Gradient Approximation and Beta Coefficients

The algorithm of the 2DCNN-LSTM model consists of many stacked nonlinear functions, making it difficult to visualize its internal mechanisms. Computing the interpretability similar black-box models is important, as it helps to investigate the impact of independent variables on the overall effectiveness of the network structure and enhances confidence in the retrieval results. This study introduces a method based on integration gradient approximation and standardized regression coefficients (Beta coefficients) to interpret the 2DCNN-LSTM model. The integration gradient approximation estimates the contribution of each input feature to the output by calculating the integration of the feature gradients between the input and output (Equation 3).

$$IGA = (x_i - x'_i) \times \int_{a=0}^1 \frac{\partial F(x'_i + a \times (x_i - x'_i))}{\partial x'_i} da \quad (3)$$

When calculating the Beta coefficients, the multiple regression model data was standardized to eliminate the influence of differences in dimensions and magnitudes, making different independent

variables comparable. Then used the absolute values of the standardized regression coefficients to compare the effects of each independent variable on the dependent variable.

Assuming the existence of a linear regression model with  $p$  independent variables (Equation 4), the standardized regression coefficients were calculated (Equation 5).

$$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \quad (4)$$

$$\beta_i^* = \frac{\beta_i \times S_{xi}}{S_y} \quad (5)$$

where  $S_y$  is the standard deviation of dependent variable  $y$ ,  $S_{xi}$  is the standard deviation of independent variable  $x_i$ ,  $\beta_i$  is the original regression coefficient, and  $\beta_i^*$  is the standardized regression coefficient.

### 3. Results

#### 3.1. Model Performance Evaluation

The evaluation of the NO<sub>2</sub> retrieval model is an important part of this study, using two commonly used model performance validation methods: individual validation and 10-fold cross-validation, to assess the accuracy and robustness of the model.

Firstly, individual validation was performed by dividing the dataset into a training set (80%) and a validation set (10%), with the remaining data used as the testing set (10%) to prevent overfitting of the model on validation set. The model configuration was adjusted based on the validation data until reaching optimal performance, and finally the independent validation indicators were obtained from the model's performance on the testing set.

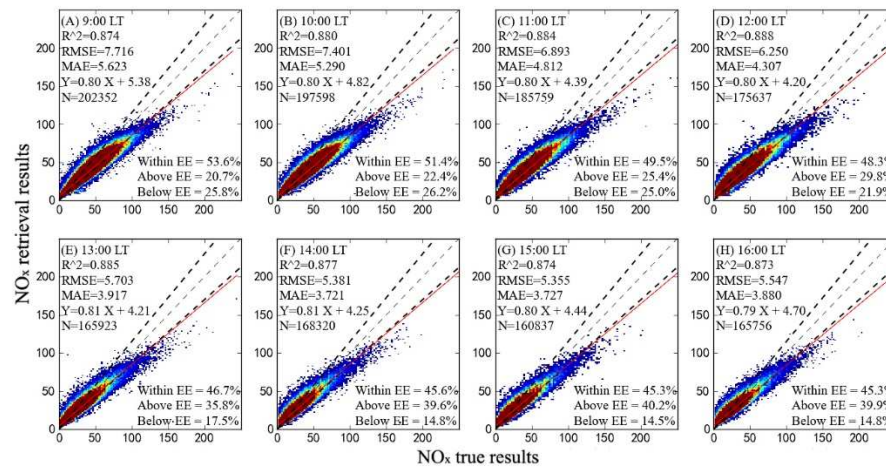
The computed results showed that the feature-selected 2DCNN-LSTM model exhibited high accuracy, with an  $R^2$  value of 0.87, indicating that it fits well with the target data, and the selected feature combination had a strong explanatory capacity. Furthermore, the RMSE was 6.84, indicating a small average error between the model's validation results and the actual observations. Additionally, using all 16 channels of Himawari-8 in the model yielded an  $R^2$  of 0.74 and an RMSE of 10.93, demonstrating that the channel selection process significantly improved the performance of 2DCNN-LSTM while saving computational resources.

Subsequently, 10-fold cross-validation was performed by dividing the dataset into ten subsets, with one subset randomly chosen as the validation set, and the remaining nine subsets served as the training set. The hyperparameters were kept consistent throughout these ten validation processes, and the average performance across the ten folds was finally used as a comprehensive measurement of the model's quality.

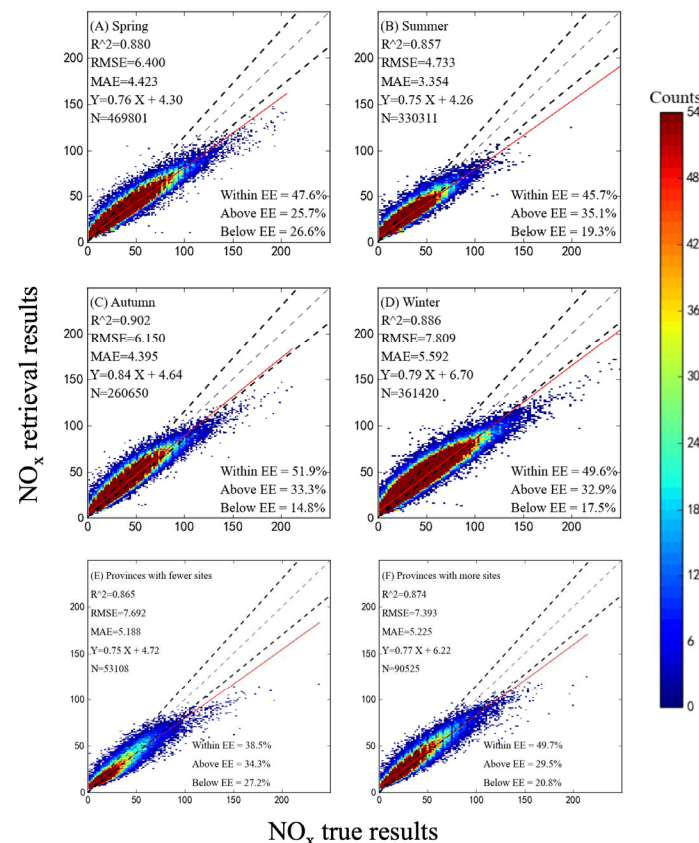
The 10-fold cross-validation method was commonly used over individual validation method, due to its ability to mitigate the randomness caused by a single split of the training and validation sets, and fully utilize available dataset for model performance validation. This multiple splitting process helps to avoid the construction of incorrect models with inadequate generalization capabilities due to specific data splits.

To validate the accuracy of the model, the 8-hour daytime (Figure 5) and the seasonal results (Figure 6(A - D)) were validated for further analysis. The x-axis represented the observed hourly NO<sub>2</sub> concentrations, while the y-axis represented the near-surface concentration results obtained from the 2DCNN-LSTM model. The regression evaluation indicators, including  $R^2$ , RMSE, and MAE, were used to assess the fitting effect and accuracy of the model[28]. We clustered based on the density of observation stations in various provinces of China (Figure 6 (E-F)) to form low-density measurement areas (Tibet, Ningxia, Qinghai, and Xinjiang) and high-density measurement areas (Jiangsu, Zhejiang, Beijing, and Guangdong), and model performance was assessed. Preliminary findings suggested that differences in station density directly impacted sample size, thus affecting model estimation results. A higher-density areas with larger training sets facilitated the algorithm's ability to discern correct patterns and trends, thereby improving accuracy as the training set size increases.

The  $R^2$  of low monitoring station density regions was lower than the high density. This comparison demonstrated the direct impact of station density differences on the model's estimation performance, but this difference was less than 0.1, indicating that it did not affect the main conclusions of the model.



**Figure 5.** The 8-hour daytime 10-fold cross-validation results based on 2DCNN-LSTM (hourly cross-validation results (A - H)). The black dashed line represents the expected error line, the light dashed line represents the 1:1 line, and the red solid line represents the linear regression fit line; N represents the sample size obtained each time.



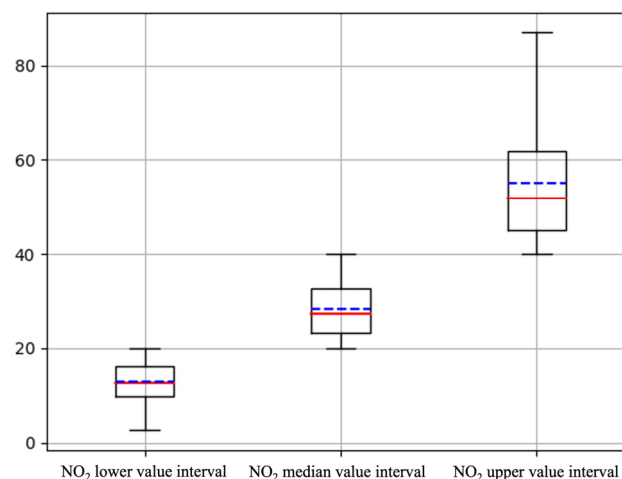
**Figure 6.** Cross-validation results based on 2DCNN-LSTM (cross-validation results for seasons from spring to summer (A - D), cross-validation results with different station density regions (E - F)). (E) The region of high station density consists of the provinces ( Jiangsu, Zhejiang, Beijing and Guangdong). (F) The region of low station density consists of the provinces (Tibet, Ningxia, Qinghai and Xinjiang). The black dashed line represents the expected error line, the light dashed line

represents the 1:1 line, and the red solid line represents the linear regression fit line; N represents the sample size obtained each time.

The cross-validation results showed that the retrieval results were well consistent with the observed values, with  $R^2$  above 0.85 and MAE within 5.70. The RMSE was within 7.90. From morning to evening, the  $R^2$  ranged from 0.87 to 0.89, with the best performance around noon ( $R^2 = 0.89$ ). From spring to summer, the  $R^2$  ranged from 0.86 to 0.90, with good performance observed during the transitional seasons of spring and autumn ( $R^2 = 0.88$  and  $R^2 = 0.90$ ). It is noteworthy that favorable meteorological diffusion conditions (Such as higher Boundary layer height and high temperature) during midday lead to enhanced pollutant dispersion, as shown in the boxplot analysis in Figure 7. the low-value range of  $\text{NO}_2$  exhibited reduced dispersion compared to other concentration ranges, with the median proximity to the mean line. This quasi-normal distribution with low and unbiased dispersion diminished the retrieval error of the model. Hence, during midday, the performance of the model had improved.

Additionally, the BLH is generally higher in summer than that in autumn and winter[29], while in autumn, the correlation between  $\text{NO}_2$  vertical diffusion and BLH is higher. The contribution of meteorological factors was higher in autumn than that in other seasons (the beta coefficient of autumn BLH was 0.27, which was higher than the 0.24 for the whole year). Therefore, the autumn 2DCNN-LSTM model performed better than the annual model, with the highest accuracy ( $R^2 = 0.90$ ), this was because being at the transition seasons when pollutant concentrations are relatively low, reducing model retrieval error and difficulty. In summer, the overall average  $\text{NO}_2$  concentration ( $8.58 \pm 6.75 \mu\text{g}/\text{m}^3$ ) was lower than in other seasons, resulting in higher accuracy in predicting the target variable. However, the  $R^2$  was relatively lower, indicating a lower explanatory capability of the model for the  $\text{NO}_2$  concentrations in summer. Furthermore, the results showed that around 50% of the estimated concentrations lied within the Expected Error. The consistency of accuracy and reliability in the daytime 8-hour cross-validation and seasonal validation further confirmed the stability and applicability of the model. Overall, the 2DCNN-LSTM model performed well on both hourly and seasonal scales and was able to analyze the temporal characteristics of  $\text{NO}_2$  concentrations.

The estimation of high  $\text{NO}_2$  concentrations was influenced by multiple complex factors, including meteorological conditions and emissions sources, which may bring errors when simulating these extreme scenarios. The box plots reflected the interquartile range and depicting the data's dispersion, where the height of the box visually represents the data's dispersion, with  $\text{NO}_2$  concentration showing a right-skewed distribution and data dispersion in the high-value range. The configuration of the box plots for high  $\text{NO}_2$  concentrations indicated that the concentration data became more dispersed with higher values, and outliers tended to be concentrated at the higher end. The model's performance under extreme and high-concentration conditions may deviate from expectations, which can be attributed to the multifaceted influences on the estimation of high  $\text{NO}_2$  concentrations, encompassing factors such as meteorological conditions and emissions sources.





**Figure 7.** The boxplots for different NO<sub>2</sub> concentration intervals. The red x solid line is the median, the blue dashed line is the mean, and the top and bottom of the box are the upper quartile (Q3) and lower quartile (Q1) of the data, respectively.

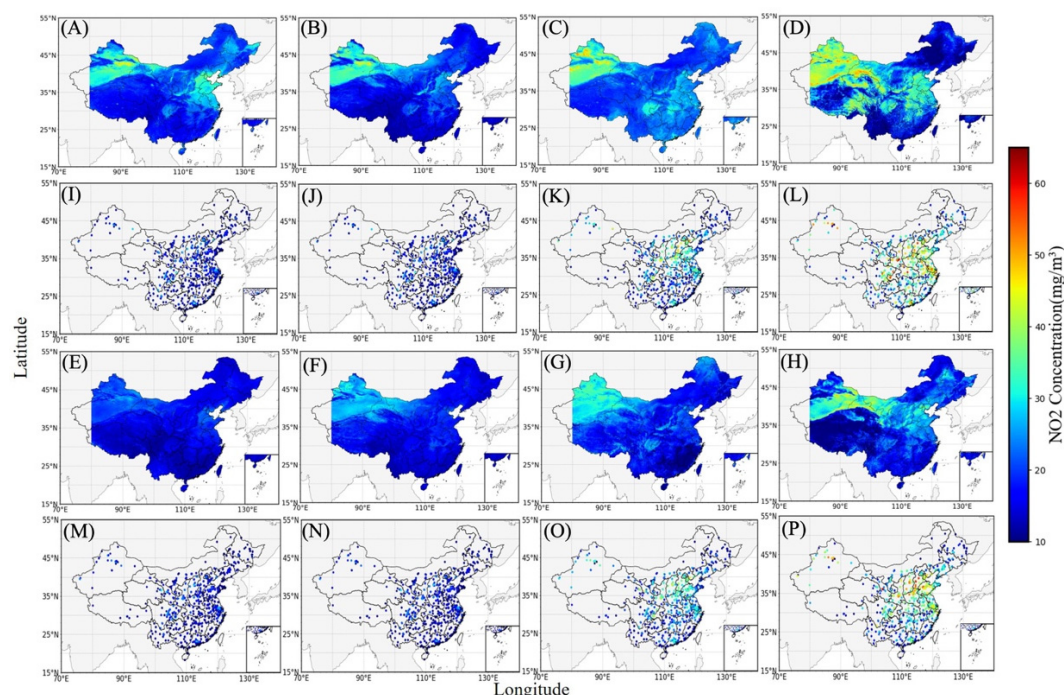
### 3.2. Retrieval Results

Based on the evaluation of the model, further analysis was conducted concerning the estimation results, with a particular focus on the spatial and temporal variations of NO<sub>2</sub> distribution in different seasons from 2018 to 2019 in China.

Atmospheric NO<sub>2</sub> is primarily derived from natural sources, while in megacity clusters and industrial regions in China, NO<sub>2</sub> is mostly emitted from human activities such as fuel combustion, mobile sources, and stationary sources such as industrial ovens. Near-surface NO<sub>2</sub> concentrations are generally higher in these zones due to thick population density, heavy traffic volume and high industrial emissions.

The retrieval NO<sub>2</sub> from 2DCNN-LSTM was in good agreement with the spatial distribution of near-surface observations. There were significant differences in the distributions of NO<sub>2</sub> across seasons (Figure 8). Concentrations were relatively low in spring and summer, while they were rather higher in autumn and winter, especially in highly industrialized and urbanized areas. These seasonal differences were related to meteorological conditions, population density, and industrial activities. Specifically, the near-surface NO<sub>2</sub> retrieval results based on satellite and 2DCNN-LSTM in China exhibited the following spatiotemporal and seasonal characteristics:

The seasonal characteristics of near-surface NO<sub>2</sub> showed a low in summer and a peak in winter. The higher temperatures and increased radiation in summer promoted photochemical reactions, leading to NO<sub>2</sub> transferred into ozone and other nitrogen-containing secondary aerosols, resulting in generally low near-surface NO<sub>2</sub> concentrations in summer. Additionally, abundant rainfall and high humidity in summer facilitated the removal of pollutants, while the hot weather and strong atmospheric convection enhanced air mixing and the upward lifting of the atmospheric boundary layer, leading to decreased joint air stability and easy diffusion of pollutants. These factors resulted in the lowest average NO<sub>2</sub> concentrations in summer. In winter, near-surface NO<sub>2</sub> concentrations were highest, with a larger range of high value, mainly distributed in the Beijing-Tianjin-Hebei region, Yangtze River Delta, Pearl River Delta, Sichuan Basin, and Xinjiang. This was related to coal combustion for heating and industrial emissions in winter. Additionally, the low temperatures in winter weakened the chemical conversion of NO<sub>2</sub> to other nitrogen compounds, and the lower surface temperature led to a decrease in the air vertical mixing and an increase in air stability, making it difficult for pollutants to diffuse. Pollution became worse when a temperature inversion occurred, resulting in the highest average near-surface NO<sub>2</sub> concentration ( $35.11 \pm 23.27 \mu\text{g}/\text{m}^3$ ) in winter. In the transitional seasons of spring and autumn, the average NO<sub>2</sub> concentrations were similar, ranging from  $12.20 \pm 8.43 \mu\text{g}/\text{m}^3$  in spring to  $23.49 \pm 12.35 \mu\text{g}/\text{m}^3$  in autumn. The seasonal variation of NO<sub>2</sub> levels was as follows: summer < spring < autumn < winter, with autumn and winter showing significantly higher levels compared to spring and summer.



**Figure 8.** Spatial distribution of NO<sub>2</sub> retrievals (A- H) versus spatial distribution of stations (I-P) for each season from 2018 to 2019 (unit:  $\mu\text{g}/\text{m}^3$ ).

From a time-series perspective, the pollutant concentrations in each season of 2019 were lower than 2018. Compared to 2018, the average NO<sub>2</sub> concentration in 2019 decreased by 7.0%, while in spring, summer, autumn, and winter, it decreased by 7.5%, 7.1%, 2.9%, and 9.7%, respectively. In 2019, the high value zones showed a significant reduction in range and exhibited a continuous downward trend in concentrations. This trend was closely related to the ongoing environmental governance efforts[30] and the promotion of new energy vehicles in China[31]. China had been actively promoting environmental governance by implementing a series of measures to reduce pollutant emissions, such as strengthening the control of industrial and vehicular emissions, improving emission control in coal-fired electric power plants, and enhancing the monitoring, early warning systems for atmospheric pollutants, and other vigorous initiatives. Furthermore, China had been promoting the popularization and use of electric vehicles, encouraging the use of clean energy transportation through providing subsidies and building charging facilities.

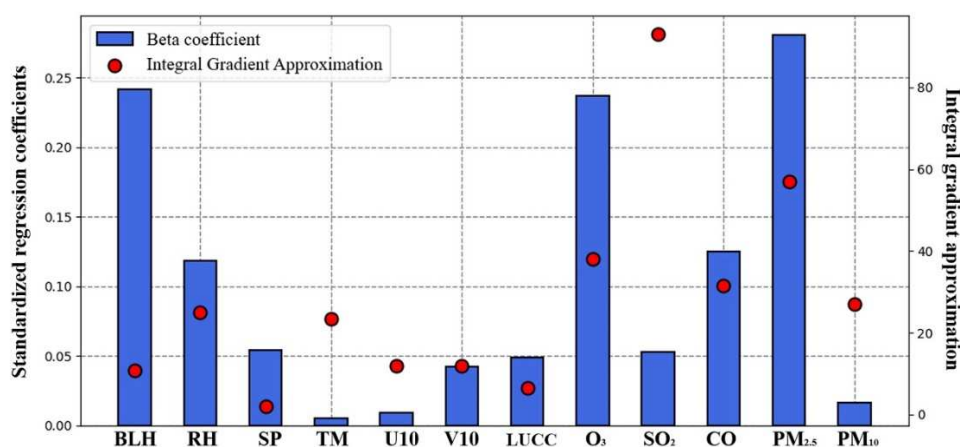
### 3.3. Factor Interpretability

The distribution of near-surface NO<sub>2</sub> in China was influenced by meteorological conditions and other pollutants. Assessing factor interpretability was meaningful for interpreting the mechanisms of black-box model and improving optimization. The following analysis examined the relationships between the considered factors, including meteorological variables (BLH, RH, SP, TM, U10, V10), and other pollutants variables (O<sub>3</sub>, SO<sub>2</sub>, CO, PM<sub>2.5</sub>, PM<sub>10</sub>), as well as their contributions to the 2DCNN-LSTM model capability (Figure 9). Although other pollutant factors were not considered during the model construction, validation, and retrieval processes, exploring these influences and contributions provides guidance for coordinated pollution control among different pollutants.

The nitrate component in PM<sub>2.5</sub> is a product of NO<sub>2</sub> chemical transformation. Nitrate compounds in the atmosphere or adsorbed on particles can be reconverted to gaseous reactive NO<sub>2</sub>, which is an important NO<sub>2</sub> regeneration process. According to the computed beta coefficients, PM<sub>2.5</sub> was statistically the most influential factor on near-surface NO<sub>2</sub>, with a change of one standard deviation in PM<sub>2.5</sub> corresponding to a change of 0.28 standard deviations in near-surface NO<sub>2</sub>. Additionally, the variation in BLH affected the air vertical mixing, and the combined effects of NO<sub>2</sub> reactions generated

secondary pollutants mainly  $O_3$ , both contribute strongly to near-surface  $NO_2$  concentrations, with changes of 0.24 and 0.23 standard deviations, respectively, for one standard deviation change in these factors.

By applying feature perturbation, the factors disrupted in order from the loop were put into the model, and the difference between the MSE of the retrieval result and the original was calculated, so as to explore the contribution of one factor to the retrieval capability of the 2DCNN-LSTM model. The results showed that  $SO_2$  contributed the most to the model's retrieval capability.



**Figure 9.** Model predictability measures for the feature variables, left axis: standardized regression coefficients; right axis: approximation of the integral gradient. The blue bars indicate the corresponding Beta coefficients of each factor and red dots indicate the corresponding integral gradient values of each factor.

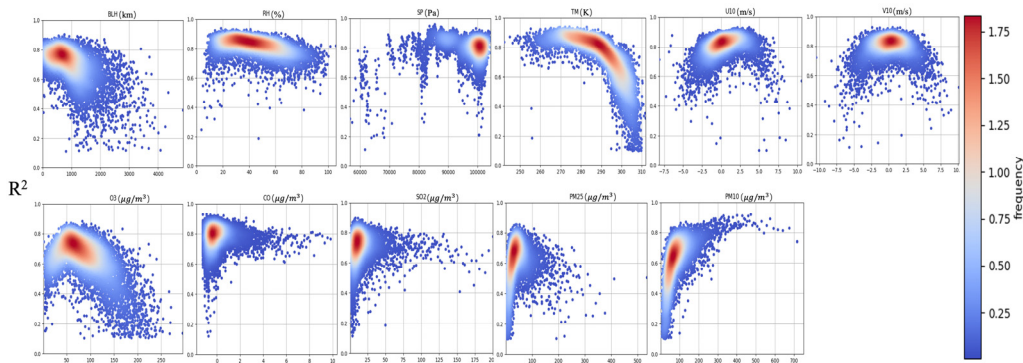
#### 4. Discussion

It is essential to formulate more effective policies to reduce  $NO_2$  emissions, along with strengthening monitoring network and improving the air quality monitoring accuracy, especially in highly industrialized and urbanized zones. Compared to traditional station-based monitoring,  $NO_2$  retrieval models have advantages in comparing retrieval results. With full coverage and high resolution, these models can provide more accurate and comprehensive information on  $NO_2$  distribution. Such comprehensive analysis can reveal broader pollution problems and provide more targeted data support for environmental protection and policymaking.

What's more, the estimated distribution of  $NO_2$  showed good correspondence with highways networks, as illustrated by the distribution of national highways in China (Figure 1). For instance, the  $NO_2$  high value zone across the Gansu province in each season corresponds well to G30 (Lianyungang-Horgos Highway), and the  $NO_2$  high value zone along the Yellow River in Inner Mongolia exists obviously in each season, and its distribution of  $NO_2$  peaks corresponds well to G110 (Beijing-Qingtongxia Highway) in the winter. The  $NO_2$  peak zones on the east side of Tibet and the south side of Qinghai in winter correspond well with G109 (Beijing-Lhasa Highway) and G318 (Shanghai-Neramu Highway). Beijing-Tianjin-Hebei, Yangtze River Delta, Pearl River Delta, and Sichuan Basin where the transportation routes connect closely are the obvious emission peak areas.

PAHs and black carbon in  $PM_{2.5}$  and  $PM_{10}$  are characterised by their composition, distribution, sources and health risks[32], Being potentially affected by regulatory actions[33]. As mentioned earlier, other pollutants and meteorological conditions can have an impact on pollutant concentrations. In order to figure out the influence of studied parameters on model, Figure 10 displayed the variation trend of the  $R^2$  concerning different feature factors, offering insight into how each parameter's changes influence spatiotemporal patterns. The results revealed that the  $R^2$  of the estimated  $NO_2$  concentration varies with fluctuations in meteorological elements and other pollutant concentrations. In general, various factors influenced the  $R^2$  of the model.  $R^2$  increased with SP, U10,

and PM<sub>10</sub>, and decreased with BLH, RH, and TM. The variation in R<sup>2</sup> with V10 was less apparent, and other factors exhibited a trend of initially increasing to a maximum before decreasing.



**Figure 10.** Trends in the goodness of fit (R<sup>2</sup>) of feature factors.

The article conducted NO<sub>2</sub> retrieval research based on satellite remote sensing technology, utilizing a fine and grid-based model to engage in an overview on the spatiotemporal distribution of pollutants in the monitoring area, and explored factor interpretability. However, there were still some research limitations. It is worth considering that the combined effects of other factors such as traffic flow and industrial emissions on NO<sub>2</sub> concentrations have not been fully considered in this study. Meanwhile, discussing the correspondence between NO<sub>2</sub> retrieval results and highways network in China, a relatively simple and intuitive graphical comparison method was used, in order to more accurately measure the correspondence between the two, it is possible to construct relevant indicators to quantify the spatial and temporal correlations, as well as to reveal the causal relationship[34,35] of the interactions between different regions. In future studies, spatio-temporal correlations can be explored and interactions between regions can be revealed through causal analyses. With these improvements, the depth and breadth of this study can be further expanded, leading to a more comprehensive understanding of the driving mechanisms and influencing factors of NO<sub>2</sub> concentration changes.

In addition, future research should focus on further improving and refining the model by expanding more factors, increasing research time series and validating with multi-source monitoring data. Efforts should also be made to enhance retrieval and source apportionment methods, and to strengthen model development and evaluation, so as to gain a more accurate understanding of the NO<sub>2</sub> distribution and factor contributions. Due to the different behavior based on the seasons and regions, the clustering can be performed on different seasons and regions to train dependent 2DCNN-LSTM models with varying parameters, thus using a non-linear system with varying parameters to determine the intrinsic order certainty between the explanatory variables and the target variables in each season or districts, in order to better analyse and compare differences between seasons, geographic area. Furthermore, further research is needed to improve related techniques and methods, enhance resolution in retrieval process, improve data quality, and strengthen collaborative research with meteorology, environment and other fields, to promote pollution management and air quality improvement.

**5. Conclusion**

This study aimed to estimate near-surface hourly NO<sub>2</sub> concentrations in China using geostationary satellite Himawari-8 data with a spatial resolution of 0.05°, and to perform retrieval estimation using a parameter-optimized 2DCNN-LSTM model. The results demonstrated that the model exhibited good fitting capacity and robustness and had high accuracy in estimating near-surface NO<sub>2</sub> concentrations. In terms of model performance evaluation, the channel-selected showed further improvement compared to the original, indicating the significant impact of satellite channel selection on model performance. The 10-fold cross-validation demonstrated that the results had R-



squared above 0.85 and MAE within 5.60, and RMSE within 7.90. the best validation performance was observed around noon and during the transitional seasons of spring and autumn. By analyzing the statistical methods and feature perturbation, the extent of meteorological and other pollutant factors affecting on the retrieval accuracy of the 2DCNN-LSTM model was explored. The results showed that BLH, O<sub>3</sub>, SO<sub>2</sub>, and PM<sub>2.5</sub> had statistically significant contributions to the model retrieval capability and played important roles in the estimation of NO<sub>2</sub> concentrations.

Results of the study have existing linkages with atmospheric chemistry studies and air pollution control, which are being analysed below: Firstly, the results extend the understanding of the spatial and temporal distribution of pollutants, especially in the interpretation of the changes in NO<sub>2</sub> concentrations during urbanisation. In addition, the effects of meteorological factors and other air pollutants on NO<sub>2</sub> concentrations were analysed in depth, which helped to reveal the sources and mechanisms of changes of NO<sub>2</sub> in the atmosphere. The results of these studies can provide a scientific basis for further improving air quality management and formulating corresponding policies. Finally, from the practical application point of view, this study is of great significance for monitoring and predicting air quality conditions. By establishing a spatio-temporal high-fine-degree NO<sub>2</sub> detection model, the air pollution levels in different regions and seasons can be more accurately understood, which can provide support for the relevant departments to formulate precise pollution control measures.

**Author Contributions:** Conceptualization, B.C.; Methodology, R.C. and B.C.; Software, R.C.; Validation, R.C.; Formal Analysis, B.C. and R.C.; Investigation, B.C. and R.C.; Resources, B.C. R.C., J.H., Z.S., Y.W. X.Z, and L.Z.; Data Curation, B.C. R.C., Z.S., J.H., X.Z, and L.Z.; Writing – Original Draft Preparation, B.C. and R.C.; Writing – Review & Editing, B.C. R.C., J.H., Z.S., and Y.W.; Visualization, R.C.; Supervision, B.C.; Project Administration, B.C.; Funding Acquisition, B.C.

**Funding:** The work Supported by the National Natural Science Foundation of China (Grant number 41775021), the Fundamental Research Funds for the Central Universities (Grant number lzujbky-2022-ct06), and the Gansu Provincial Science and Technology Plan (Grant number 23JRRA1038).

**Data Availability Statements:** The data for the article comes from the third-party data, and the download address is given below: global land cover data with 30 meter resolution: <http://data.starcloud.pcl.ac.cn/zh/resource/2>; Hourly near-surface pollutant concentrations were provided by the China National Environmental Monitoring Center (CNEMC) (<http://www.cnemc.cn/>); The Himawari-8 TOAR data provided by the Japan Meteorological Agency, (<https://www.eorc.jaxa.jp/ptree/index.html>); The ERA-5 data are available(<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview>).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sun, Q., X. Hong, and L.E. Wold. Cardiovascular Effects of Ambient Particulate Air Pollution Exposure of the article. *Circulation*. **2010**, 121, 2755-2765.
2. Brook and D. R. Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association of the article. *Circulation*. **2004**, 109, 2655-2671.
3. Ciampi, Q., et al. Nitrogen dioxide component of air pollution increases pulmonary congestion assessed by lung ultrasound in patients with chronic coronary syndromes of the article. *Environmental Science and Pollution Research*. **2022**, 29, 26960-26968. 10.1007/s11356-021-17941-1.
4. Cheng, S., et al. Multiscale spatiotemporal variations of NO(x) emissions from heavy duty diesel trucks in the Beijing-Tianjin-Hebei region of the article. *Sci Total Environ*. **2022**, 854, 158753. 10.1016/j.scitotenv.2022.158753.
5. Liu, J. and W. Chen. First satellite-based regional hourly NO(2) estimations using a space-time ensemble learning model: A case study for Beijing-Tianjin-Hebei Region, China of the article. *Sci Total Environ*. **2022**, 820, 153289. 10.1016/j.scitotenv.2022.153289.
6. Meng, K., et al. Spatio-temporal variations in SO(2) and NO(2) emissions caused by heating over the Beijing-Tianjin-Hebei Region constrained by an adaptive nudging method with OMI data of the article. *Sci Total Environ*. **2018**, 642, 543-552. 10.1016/j.scitotenv.2018.06.021.

7. Chen, L., et al. Vertical profiles of O<sub>3</sub>, NO<sub>2</sub> and PM in a major fine chemical industry park in the Yangtze River Delta of China detected by a sensor package on an unmanned aerial vehicle of the article. *Sci Total Environ.* **2022**, 845, 157113. 10.1016/j.scitotenv.2022.157113.
8. Cooper, M.J., et al. Global fine-scale changes in ambient NO<sub>2</sub> during COVID-19 lockdowns of the article. *Nature.* **2022**, 601, 380-387. 10.1038/s41586-021-04229-0.
9. Dong, L., et al. Analysis on the Characteristics of Air Pollution in China during the COVID-19 Outbreak of the article. *Atmosphere.* **2021**, 12, 205.
10. Chi, Y., et al. Machine learning-based estimation of ground-level NO<sub>2</sub> concentrations over China of the article. *Sci Total Environ.* **2022**, 807, 150721. 10.1016/j.scitotenv.2021.150721.
11. Dai, Y., et al. Chemistry, transport, emission, and shading effects on NO<sub>2</sub> and O<sub>3</sub> distributions within urban canyons of the article. *Environ Pollut.* **2022**, 315, 120347. 10.1016/j.envpol.2022.120347.
12. Liu, J., et al. The influence of solar natural heating and NO<sub>x</sub> photochemistry on flow and reactive pollutant exposure in 2D street canyons of the article. *The Science of the total environment.* **2021**, 759, 143527. 10.1016/j.scitotenv.2020.143527.
13. Chen, Y., et al. Development of an integrated machine-learning and data assimilation framework for NO<sub>x</sub> emission inversion of the article. *Sci Total Environ.* **2023**, 871, 161951. 10.1016/j.scitotenv.2023.161951.
14. Guo, X., et al. Analysis of the Spatial&ndash;Temporal Distribution Characteristics of NO<sub>2</sub> and Their Influencing Factors in the Yangtze River Delta Based on Sentinel-5P Satellite Data of the article. *Atmosphere.* **2022**, 13, 1923.
15. Ji, X., et al. Nitrate pollution source apportionment, uncertainty and sensitivity analysis across a rural-urban river network based on delta(15)N/delta(18)O-NO<sub>3</sub>(-) isotopes and SIAR modeling of the article. *J Hazard Mater.* **2022**, 438, 129480. 10.1016/j.jhazmat.2022.129480.
16. Liu, F., et al. Source analysis of the tropospheric NO<sub>2</sub> based on MAX-DOAS measurements in northeastern China of the article. *Environ Pollut.* **2022**, 306, 119424. 10.1016/j.envpol.2022.119424.
17. Zhang, Y., et al. Spatiotemporal variations of NO<sub>2</sub> and its driving factors in the coastal ports of China of the article. *Sci Total Environ.* **2023**, 162041. 10.1016/j.scitotenv.2023.162041.
18. Liu, J. and W. Chen. First satellite-based regional hourly NO<sub>2</sub> estimations using a space-time ensemble learning model: A case study for Beijing-Tianjin-Hebei Region, China of the article. *Science of The Total Environment.* **2022**, 820, 153289. 10.1016/j.scitotenv.2022.153289.
19. Zhang, Z., et al. Multiscale time-lagged correlation networks for detecting air pollution interaction of the article. *Physica A: Statistical Mechanics and its Applications.* **2022**, 602, 127627. <https://doi.org/10.1016/j.physa.2022.127627>.
20. Chen, B., et al. Estimation of near-surface ozone concentration and analysis of main weather situation in China based on machine learning model and Himawari-8 TOAR data of the article. *Science of the Total Environment.* **2023**, 864, 160928. 10.1016/j.scitotenv.2022.160928.
21. Song, Z.H., et al. High temporal and spatial resolution PM<sub>2.5</sub> dataset acquisition and pollution assessment based on FY-4A TOAR data and deep forest model in China of the article. *Atmospheric Research.* **2022**, 274, 10.1016/j.atmosres.2022.106199.
22. Hochreiter, S. and J. Schmidhuber. Long Short-Term Memory of the article. *Neural Computation.* **1997**, 9, 1735-1780. 10.1162/neco.1997.9.8.1735.
23. Yin, H., et al. Rainfall-runoff modeling using LSTM-based multi-state-vector sequence-to-sequence model of the article. *Journal of Hydrology.* **2021**, 598, 126378. 10.1016/j.jhydrol.2021.126378.
24. Paoletti, M.E., et al. Deep learning classifiers for hyperspectral imaging: A review of the article. *ISPRS Journal of Photogrammetry and Remote Sensing.* **2019**, 158, 279-317. 10.1016/j.isprsjprs.2019.09.006.
25. Ahmed, R., et al. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization of the article. *Renewable and Sustainable Energy Reviews.* **2020**, 124, 109792. 10.1016/j.rser.2020.109792.
26. Zhang, F., et al. Ultrasonic adaptive plane wave high-resolution imaging based on convolutional neural network of the article. *NDT & E International.* **2023**, 138, 102891. 10.1016/j.ndteint.2023.102891.
27. Tahmasebi, P., et al. Machine learning in geo- and environmental sciences: From small to large scale of the article. *Advances in Water Resources.* **2020**, 142, 103619. 10.1016/j.advwatres.2020.103619.
28. Bin, C., et al. Estimation of Atmospheric PM<sub>10</sub> Concentration in China Using an Interpretable Deep Learning Model and Top-of-the-Atmosphere Reflectance Data From China's New Generation Geostationary Meteorological Satellite, FY-4A of the article. *Journal of Geophysical Research-Atmospheres.* **2022**, 127, 10.1029/2021jd036393.
29. Guo, J.P., et al. The climatology of planetary boundary layer height in China derived from radiosonde and reanalysis data of the article. *Atmospheric Chemistry and Physics.* **2016**, 16, 13309-13319. 10.5194/acp-16-13309-2016.
30. Hansen, M.H., H.T. Li, and R. Svarverud. Ecological civilization: Interpreting the Chinese past, projecting the global future of the article. *Global Environmental Change-Human and Policy Dimensions.* **2018**, 53, 195-203. 10.1016/j.gloenvcha.2018.09.014.

31. Bian, X.H., et al. Prospect Analysis for the Complementary Development of Gas-Fueled and Electric Vehicles in China of the article. *3rd Annual International Conference on Sustainable Development [ICSD]*. **2017**, 111, 252-258.
32. Ambade, B., et al. Health Risk Assessment, Composition, and Distribution of Polycyclic Aromatic Hydrocarbons (PAHs) in Drinking Water of Southern Jharkhand, East India of the article. *Archives of Environmental Contamination and Toxicology*. **2021**, 80, 1-14. 10.1007/s00244-020-00779-y.
33. Ambade, B., et al. COVID-19 lockdowns reduce the Black carbon and polycyclic aromatic hydrocarbons of the Asian atmosphere: source apportionment and health hazard evaluation of the article. *Environment, Development and Sustainability*. **2021**, 23, 12252-12271. 10.1007/s10668-020-01167-1.
34. Charakopoulos, A.K., G.A. Katsouli, and T.E. Karakasidis. Dynamics and causalities of atmospheric and oceanic data identified by complex networks and Granger causality analysis of the article. *Physica A: Statistical Mechanics and its Applications*. **2018**, 495, 436-453. <https://doi.org/10.1016/j.physa.2017.12.027>.
35. Yuan, K., et al. Causality guided machine learning model on wetland CH<sub>4</sub> emissions across global wetlands of the article. *Agricultural and Forest Meteorology*. **2022**, 324, 109115. <https://doi.org/10.1016/j.agrformet.2022.109115>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.