Hourly near-ground NO₂ concentration retrieval from geostationary satellite observations

Jie Song[1], Lin Zang[2,3*], Feiyue Mao[1], Yi Zhang[1], Jiangping Chen[1]

[1] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China – songjie2016@whu.edu.cn, maofeiyue@whu.edu.cn, ivory2008@whu.edu.cn, chen_jp@whu.edu.cn
[2] Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430079, China
[3] Key Laboratory of Polar Environment Monitoring and Public Governance (Wuhan University)，Ministry of Education, Wuhan 430079, China - zanglin2018@whu.edu.cn

Key words: NO₂, stationary satellite, infrared spectrum, hourly.

Abstract

Nitrogen dioxide (NO₂) is an important contributor to the formation of acid rain, photochemical smog and aerosol particles, which seriously endangers public health. At present, remote sensing of polar-orbiting satellites is a conventional means to obtain large-scale NO₂ distribution, but it cannot capture the rapid change because of long revisit periods. The Advanced Himawari Imager (AHI) on the Himawari-8 geostationary satellite has the advantage of high time resolution, which makes it possible to realize near-real-time atmospheric monitoring. Here, based on the absorption characteristics of NO₂ in infrared radiation, hourly near-surface NO₂ concentrations are retrieved based on the brightness temperature from AHI and auxiliary information such as meteorology and aerosol. The results of 10-fold cross-validation show that NO₂ estimations from satellite are in good agreement with in-situ measurements, and their determination coefficient (R^2) can reach 0.79. Due to different emission and atmospheric diffusion conditions at different time, the model performance presents a diurnal variation of high accuracy in the noon and afternoon but low accuracy in the morning. Based on the retrieval dataset, it is found that high NO₂ concentrations are mainly concentrated in the densely populated and industrial areas such as the North China area. In addition, NO₂ pollution mainly occurs in autumn and winter, and the average NO₂ concentration in winter is about 1.63 times that in summer in 2021. This study provides a new insight for satellite retrieval of NO₂, which is of great significance for real-time pollution monitoring and public health protection.

1. Introduction

Environmental pollution is of worldwide concern. As a trace gas, Nitrogen dioxide (NO₂) plays a key role in the chemistry of the atmosphere. It participates in the control of the strong oxidant (e.g., O₃ and OH), which determines the oxidizing capacity of the atmosphere (Chapleski et al., 2016; Gligorovski et al., 2015). In addition, the mixture of NOₓ (NOₓ = NO + NO₂) with VOC produces photochemical smog, posing a serious threat to human health and crop production (Tie et al., 2002). Therefore, studying the spatiotemporal distribution and evolution of NO₂ concentration near the ground is necessary for the stable development of the economy and society.

The main sources of NO₂ are emissions from fossil fuel combustion and biomass burning (van der A et al., 2008) . But for a long time, the large-scale distribution of NO₂ could only be analyzed by global chemistry transport models, because ground-based or air-borne measurement campaigns were temporally and spatially limited (Toenges-Schuller et al., 2006). An important step in filling the gap has been made by the satellite remote sensing. A number of space-borne sensors have been launched to detect nitrogen dioxide concentrations, including Global Ozone Monitoring Experiment instrument (GOME) (Loyola et al., 2007), Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) (Rozanov et al., 1992), Ozone Monitoring Instrument (OMI) (Ma et al., 2015), GOME-2, TROPOspheric Monitoring Instrument (TROPOMI) (Borsdorff et al., 2018), Environmental trace gases monitoring instrument (EMI) (Xiao et al., 2023), etc. These loads significantly improve the detection ability of trace gases such as NO₂, which is of great significance for understanding the spatial and temporal distribution of polluted gases. However, these loads are generally carried on polar-orbiting satellites with long revisit periods, which cannot meet the demands of capturing the rapid changes of NO₂ distribution.

With the development of satellite technology, a new generation of geostationary meteorological satellites, such as Himawari-8 and Fengyun 4A, have been successfully launched, and they provide the possibility for near-real-time atmospheric monitoring. Taking Himawari-8 as an example, the temporal resolution of full-disk observation is 10 min (Shang et al., 2017). However, at present, these satellites do not carry sensors specifically for NO₂ retrieval. Traditional NO₂ retrieval is mainly based on the differential absorption spectroscopy (DOAS) (Fu et al., 2009; Vidot et al., 2010). The basic principle of the algorithm is that using the narrow molecular absorption bands to identify trace gases and their absorption strength to retrieve tropospheric and stratospheric trace-gas concentrations. The traditional spectral window used for NO₂ retrieval is 400 ~ 500 nm. In addition to ultraviolet radiation, NO₂ also has a certain ability to absorb infrared radiation. Specifically, NO₂ has weak absorption at 3.3 ~ 3.5 μm and relatively strong absorption at 6.1 ~ 6.5 μm (Sur et al., 2017). Therefore, we here investigate the feasibility of retrieving near-ground NO₂ concentration based on the infrared spectrum information from Himawari-8, which may provide a new insight for capturing the rapid changes of NO₂ distribution.

2. Data and methods

2.1 Dataset

As mentioned above, the infrared absorption of NO₂ occurs in the ranges of 3.3 ~ 3.5 μm and 6.1 ~ 6.5 μm, and the latter can be also weakly absorbed by sulfur dioxide (SO₂). Besides, SO₂ has a strong absorption at 7.14 ~ 7.69 μm (main absorption band) (J. D. Li et al., 2022), Here, we employed AHI infrared bright temperature observations including tbb7 (λ ≈ 3.9 μm), tbb8 (λ ≈ 6.2 μm), tbb9 (λ ≈ 6.9 μm), tbb10 band (λ ≈ 7.0 μm), and the ratio of tbb8/tbb10 to retrieve NO₂ concentration. Among them, the last two parameters are used to filter out the absorption

interference of SO_2. The time resolution of AHI is 10 min, and this study sampled once every hour between 9 and 16 o 'clock local time. Notably, only NO_2 concentration in the clear sky was retrieved, and the cloud mask was provided by AHI. And, in order to remove the scattering interference of aerosol particles on light, we also utilized AHI's aerosol optical depth (AOD) products (Hong et al., 2020).

The distribution of pollutants is not only determined by emissions, but also regulated by weather and terrain (He et al., 2017; R. Li et al., 2019; Wei et al., 2023). Here, parameters including boundary layer height (BLH, unit: m), 10-m wind speed (uwind and vwind, unit: m/s), 2-m temperature (Temp, unit: K), surface relative humidity (RH, unit: %), surface pressure (SP, unit: Pa), Column-integrated water vapor (PW, unit: mm), digital elevation model (DEM, unit: m) and normalized difference vegetation index (NDVI) were used to characterize meteorological and topographic effects. Among them, the meteorological parameters were derived from the ERA5 reanalysis data published by the European Centre for Medium-Range Weather Forecasts (ECMWF). The atmospheric photochemical oxidation process was also considered by incorporating ozone changes provided by ERA-5, i.e., surface ozone mass mixing ratio (O3mr, unit: kg/kg) and column-integrated ozone concentration (O3all, unit: ppm). The DEM data was provided by the United States Geological Survey (USGS), and NDVI was derived from MODIS products published by the National Aeronautics and Space Administration (NASA) data center. The in-situ NO_2 observations was provided by the National Environmental Air Quality Monitoring Network (NEAQMN). More details about data used are shown in **Table 1**. The data is covered from January to December in 2021. In order to ensure the effectiveness of model training, only the regions with dense distribution of sites are focused, covering 100°E ~ 130°E and 20°N ~ 44°N.

Data type	Specific Parameter	Spatial Resolution	Temporal Resolution
Bright Temperature	tbb7, tbb8, tbb9, and tbb10 (BT_{TOA})	0.05°	10 min
Cloud Mask	Clear sky	0.05°	10 min
Aerosol	AOD	5 km	1 h
Other auxiliary data	Temp	0.25°	1 h
	RH	0.25°	1 h
	SP	0.25°	1 h
	uwind/vwind	0.25°	1 h
	BLH	0.25°	1 h
	O3mr	0.25°	1 h
	PW	0.25°	1 h
	O3all	0.25°	1 h
	NDVI	0.05°	16 days
	DEM	90 m	--
Station observation	Measured NO_2	site	5min

Table 1. The data set used in this study

2.2 Data matching

In view of the heterogeneity of the data used in this study (points or planes) and the differences in spatial and temporal resolution, in order to form a complete dataset as the basic sample for inversion, data preprocessing and matching need to be carried out: Firstly, the grid meteorological and geographical data were resampled to 0.05°, as same as that of the bright temperature from Himawari-8. Secondly, each station of NO_2 observation was taken as the center point, and all gridded parameters were averaged within the spatiotemporal window of 0.05° and 1 h. In this study, 371576 matching samples were obtained.

2.3 Methodology

2.3.1 NO_2 retrieval by LightGBM model

Gradient Boosting Decision Tree (GBDT) is an important class of machine learning, which mainly uses weak classifiers (decision trees) for iterative training to obtain the best model. LightGBM (Light Gradient Boosting Machine) is an enhanced GBDT model, and compared to traditional GBDT, it has the advantages of faster training speed, lower memory consumption, distributed support and rapid processing of massive data (Liang et al., 2022; Xinwei et al., 2021; Yukun & eng, 2019). Here, we use LightGBM as the basic model to realize the relationship learning between NO_2 concentration and explanatory variables (Formula 1). The specific retrieval process is shown in Figure 1.

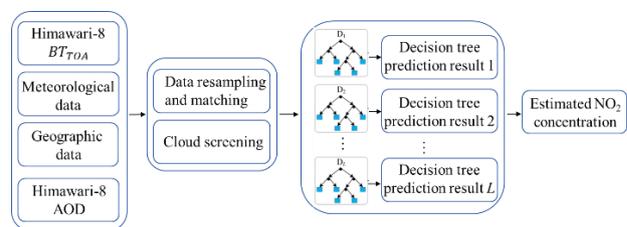$$NO_2 = f(BT_{TOA}, RH, TEMP, SP, uwind, vwind, DEM, \\ BLH, O3mr, PW, O3all, NDVI, AOD) \quad (1)$$

Figure 1. Schematic diagram of NO_2 concentration retrieval based on LightGBM model

2.3.2 Verification of model performance

The 10-fold cross validation (10-fold CV) method was utilized for model performance evaluation: (1) randomly dividing the dataset into ten equally sized subsets, (2) selecting one subset as the test set while the other nine subsets for training the model, and (3) repeating the process ten times to ensure that each subset is used as the test set once. Finally, the determination coefficients (R^2), relative root mean square error (rRMSE), and relative mean absolute error (rMAE) between satellite estimation and in-situ measurements were calculated to quantify the model accuracy (Chen et al., 2022).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{obs} - y_i^{est})^2}{\sum_{i=1}^{n}(y_i^{obs} - \overline{y^{obs}})^2} \quad (2)$$

$$rRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^{obs} - y_i^{est})^2}}{\overline{y^{obs}}} \quad (3)$$

$$rMAE = \frac{\frac{1}{n}\sum_{i=1}^{n}|y_i^{obs} - y_i^{est}|}{\overline{y^{obs}}} \quad (4)$$

Where y_i^{obs} is the in-situ observation; y_i^{est} is the model estimation; $\overline{y^{obs}}$ is the average of in-situ observed NO_2 concentration; n is the sample size.

3. Results and analysis

3.1 Model performance evaluation

The results of 10-fold CV show (Figure 2) that satellite retrieval and ground-based observation have a high consistency, with the overall R^2, rRMSE and rMAE of 0.79, 0.33 and 0.23. And all of R^2 is more than 0.70 during the daytime. Compared to existing NO_2 retrievals based on polar-orbiting satellite-observed NO_2 columns, our model may not be the most accurate. For example, Zhang et al. (2022) estimated near-surface NO_2 distributions in North China using tropospheric NO_2 columns from TROPOMI, achieving an R^2 of 0.85. However, our study is able to provide hourly datasets while ensuring acceptable model performance. Furthermore, the introduce of thermal infrared information reduces the reliance on satellite sensors specialized for trace gases in traditional retrievals and enables multi-sensor joint retrievals like Himawari-8 and Fengyun-4A.

Notably, the model has good performance in the noon and afternoon, and relatively poor performance in the morning. which are related to the accuracy of satellite observations and atmospheric conditions. Generally speaking, the higher altitude angle of the sun at noon is conducive to ensuring the stable observation of the satellite. At the same time, the higher noon temperature and boundary layer are conducive to the diffusion and mixing of pollutants, which is more consistent with the physical premise of uniform mixing of pollutants adopted in the satellite inversion algorithm. In addition, the atmospheric humidity is relatively low at noon, and the liquid phase reaction related to NO_2 is weak. These factors ensure a better model performance in the noon period than in the morning. more detailed explanations can be found in (Zang et al., 2018).

The seasonal verification shows (Figure 3) that the model exhibits excellent performance in autumn and winter ($R^2 = 0.79$), followed by spring ($R^2 = 0.74$), and the worst in summer ($R^2 = 0.66$). The seasonal difference may be caused by several potential factors: firstly, there is more rainfall in summer and less in winter. Rainfall has a wet removal effect on pollutants in the atmosphere. Secondly, in summer, photochemical reactions are more active in the atmosphere than that in winter, and convective activity is also strong during summer times, resulting in a relatively short lifespan of NO_2. These seasonal changes may make it difficult for the model to accurately capture changes in gas concentration in summer, while it is easier to achieve accurate predictions in winter.
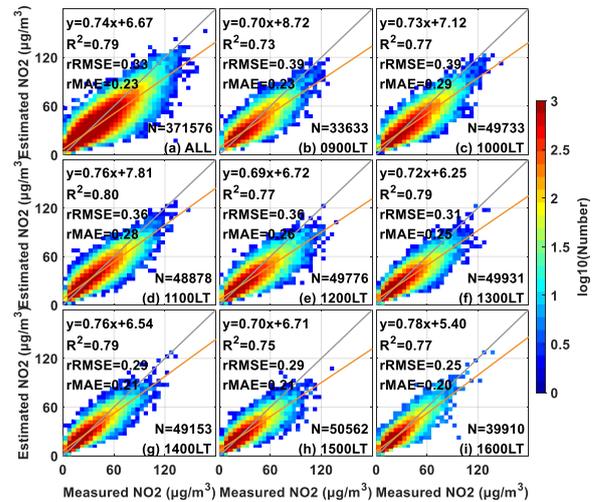

Figure 2. Model performance quantified by 10-fold CV method at different times in 2021: (a) statistics for all matched samples and (b-i) for different times during the daytime (09: 00-16: 00 LT). The gray solid line represents the 1:1 line, the orange solid line represents the linear regression fitting line, and the color-bar represents the logarithm of the sample size.
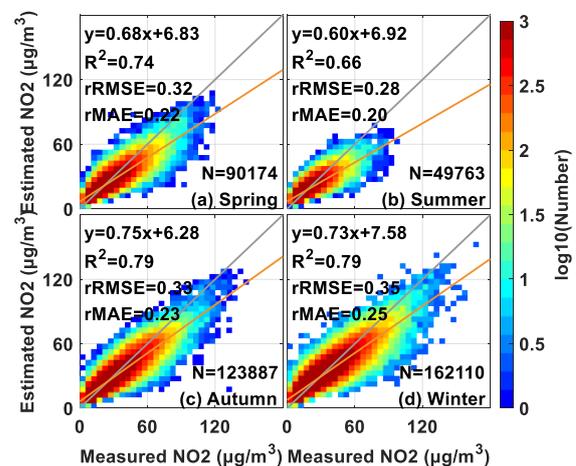

Figure 3. Model performance quantified by 10-fold CV method at different seasons in 2021, and (a-d) represents four seasons: spring (Mar., Apr. and May.), summer (Jun., Jul. and Aug.), autumn (Sept., Oct. and Nov.), and winter (Jan., Feb., and Dec.). The gray solid line represents the 1:1 line, the orange solid line represents the linear regression fitting line, and the color-bar represents the logarithm of the sample size.

3.2 Spatiotemporal distribution of near-surface NO_2

Figure 4 shows the distribution of near-surface NO_2 concentrations. According to statistics, the average concentration of NO_2 in the whole research area is 17.64 μg/m^3, and its high values are mainly concentrated in densely populated and industrial areas such as Beijing-Tianjin-Hebei, the North China Plain, and Jiangsu-Zhejiang-Shanghai. As mentioned earlier, the main sources of NO_2 are emissions from fossil fuel combustion and biomass burning. At the same time, the meteorological

conditions in these areas are more complex, which is not conducive to the diffusion and removal of pollutants.

From a temporal perspective, the concentration of NO_2 is higher at 9 am in the morning (22.97 μg/m^3). This may be due to the low traffic flow and weak solar radiation in the early morning, which is not conducive to the decomposition of NO_2. As time passed, the concentration began to gradually decrease around 10 a.m (21.80 μg/m^3). This may be related to the enhancement of solar radiation and the rise in temperature and boundary layer height. However, in the afternoon, the concentration of NO_2 showed an increasing trend again, and the concentration is 14.32 μg/m^3 at 16:00. This may be related to the accumulation of industrial waste gas and vehicle exhaust emissions on the road. In order to improve the air quality, strict pollution control measures need to be taken to reduce the emissions of NO_2. At the same time, it is crucial to strengthen the construction of monitoring and warning systems in order to timely detect and respond to pollution incidents.
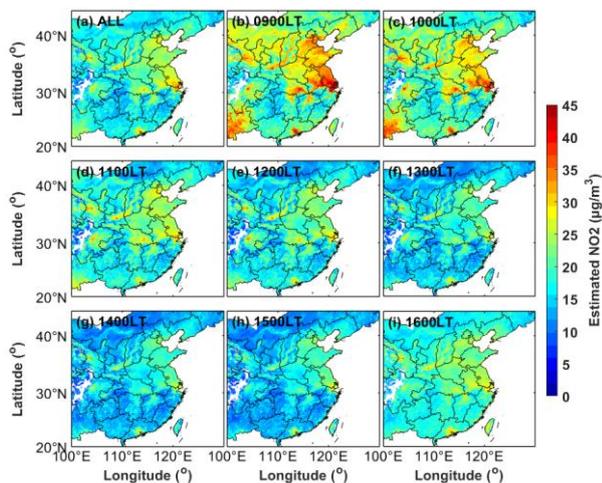

Figure 4. Hourly spatial distributions of near-surface NO_2 concentration in 2021: (a) annual average of all estimates and (b-i) annual average at different times during the daytime (09: 00-16: 00 LT), with the color-bar representing NO_2 concentration.

As shown in Figure 5, the concentration distribution of NO_2 near the surface has significant seasonal variation characteristics, from low to high in summer (13.11 μg/m^3), spring (17.03 μg/m^3), autumn (17.11 μg/m^3), and winter (21.39 μg/m^3). The increase in temperature and radiation in summer promotes photochemical reactions, leading to the decomposition of NO_2 into ozone. And the higher humidity in summer is also in favor of the transformation from gaseous to solid nitrate through heterogeneous liquefaction processes. In addition, sufficient rainfall in summer contributes to the removal of pollutants. The strong atmospheric convection generated by hot weather and strong monsoon weather enhances air mixing and the upward elevation of the atmospheric boundary layer, and also promotes the diffusion of pollutants. These factors result in generally low near-surface NO_2 concentrations in summer. However, the concentration of NO_2 near the surface is highest in winter, with a wider range of high values. This is related to winter heating and industrial emissions. In addition, the low temperature in winter weakens the chemical conversion of NO_2 to other nitrogen compounds, and lower surface temperatures lead to reduced vertical air mixing and promote local accumulation of pollutants.
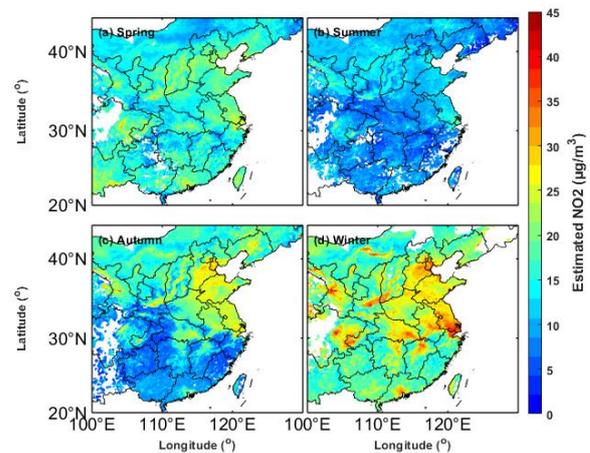

Figure 5. Seasonal distribution map of near-surface NO_2 concentration in 2021, where (a-d) represents four seasons: spring, summer, autumn, and winter, and the color of the colorbar represents NO_2 concentration

4. Conclusions

Remote sensing of polar-orbiting satellites is a conventional means to obtain large-scale NO_2 distribution. However, because NO_2 exists in the atmosphere for a short time, the traditional methos is difficult to meet the demands of capturing its rapid changes. Here, based on the absorption characteristics of NO_2 in the infrared spectrum, we use LightGBM machine learning model to carry out hourly near-surface NO_2 concentration retrieval from geostationary satellite in central and eastern China. The main conclusions are as follows:

(1) Satellite-observed infrared bright temperature has the potential to retrieve NO_2 concentration with high accuracy. The R^2 between satellite estimation and in-situ observation generally exceeds 0.75, and the rRMSE remains below 0.40. From a daily perspective, the model performs best around 11:00 ($R^2 = 0.80$); from a seasonal perspective, the performance is poor in relatively clean summer, while it is excellent in heavy polluted winter ($R^2 = 0.79$).

(2) The distribution of NO_2 show a significant spatiotemporal heterogeneity. High concentrations of NO_2 mainly congregated in densely populated and industrial areas such as Beijing-Tianjin-Hebei, the North China Plain and Jiangsu-Zhejiang-Shanghai. From a daily perspective, NO_2 pollution is more serious in the morning and evening; from a seasonal perspective, pollution is worse in autumn and winter. This is determined by both anthropogenic emissions and atmospheric conditions.

The grid NO_2 concentration product obtained in this study has high temporal resolution, which is conducive to capturing the fine evolution of pollutants and providing support for regional pollution prevention and control. Although our model has achieved satisfactory performance, there are still lots of uncertainties related to satellite observations, reanalysis data, as well as uneven sample collection. In the future, more multi-source data will be combined to constrain and correct the above potential uncertainties, so as to further improve the accuracy of NO_2 estimation.

References

Borsdorff, T., Aan de Brugh, J., Hu, H., Aben, I., Hasekamp, O., & Landgraf, J. 2018. Measuring Carbon Monoxide With TROPOMI: First Results and a Comparison With ECMWF-IFS Analysis Data. *Geophysical Research Letters, 45*(6), 2826-2832.

Chapleski, R. C., Zhang, Y. F., Troya, D., & Morris, J. R. 2016. Heterogeneous chemistry and reaction dynamics of the atmospheric oxidants, O_3, NO_3, and OH, on organic surfaces. *Chemical Society Reviews, 45*(13), 3731-3746.

Chen, B., Song, Z. H., Shi, B. L., & Li, M. J. 2022. An interpretable deep forest model for estimating hourly PM_{10} concentration in China using Himawari-8 data. *Atmospheric Environment, 268*.

Fu, C. L., Dong, H., Hua, T. J., & Lin, S. 2009. Overview of tropospheric NO_2 vertical column density retrieval from space measurement. *Journal of Remote Sensing, 13*(3), 343-354.

Gligorovski, S., Strekowski, R., Barbati, S., & Vione, D. 2015. Environmental Implications of Hydroxyl Radicals (•OH). *Chemical Reviews, 115*(24), 13051-13092.

He, J. J., Gong, S. L., Yu, Y., Yu, L. J., Wu, L., Mao, H. J., . . . Li, R. P. 2017. Air pollution characteristics and their relation to meteorological conditions during 2014-2015 in major Chinese cities. *Environmental Pollution, 223*, 484-496.

Hong, J., Mao, F. Y., Min, Q. L., Pan, Z. X., Wang, W., Zhang, T. H., & Gong, W. 2020. Improved $PM_{2.5}$ predictions of WRF-Chem via the integration of Himawari-8 satellite data and ground observations. *Environmental Pollution, 263*.

Li, J. D., Ding, Y. J., Li, Z., & Peng, Z. M. 2022. Simultaneous measurements of SO_2 and SO_3 in the heterogeneous conversions of SO_2 using QCL absorption spectroscopy. *Applied Physics B-Lasers and Optics, 128*(3).

Li, R., Wang, Z. Z., Cui, L. L., Fu, H. B., Zhang, L. W., Kong, L. D., . . . Chen, J. M. 2019. Air pollution characteristics in China during 2015-2016: Spatiotemporal variations and key meteorological factors. *Science of the Total Environment, 648*, 902-915.

Liang, Z. P., Zou, T. X., Gong, J. L., Zhou, M., Shen, W. J., Zhang, J. T., . . . Lu, Y. H. 2022. Evaluation of Soil Nutrient Status Based on LightGBM Model: An Example of Tobacco Planting Soil in Debao County, Guangxi. *Applied Sciences-Basel, 12*(23).

Loyola, D., Erbertseder, T., Balis, D., Lambert, J. C., Spurr, R., Van Roozendael, M., . . . Lerot, C. (2007, Sep). *Operational Monitoring of the Antarctic Ozone Hole: Transition from GOME and SCIAMACHY to GOME-2.* Paper presented at the Symposium for the 20th Anniversary of the Montreal Protocol 2007, Athens, GREECE.

Ma, M. L., Shi, R. H., & Gao, W. (2015, Aug 11-12). *Validation of the OMI-TOMS and OMI-DOAS total ozone column data using ground-based observations over China.* Paper presented at the Conference on Remote Sensing and Modeling of Ecosystems for Sustainability XII, San Diego, CA.

Rozanov, V. V., Timofeev, Y. M., Polyakov, A. V., Burrows, J., & Chance, K. 1992. ON THE POSSIBILITIES OF PRECISION-MEASUREMENTS OF O3 AND NO2 FROM SPACE BY SOLAR-RADIATION ABSORPTION-SPECTRA. *Izvestiya Akademii Nauk Fizika Atmosfery I Okeana, 28*(5), 500-505.

Shang, H. Z., Chen, L. F., Letu, H. S., Zhao, M., Li, S. S., & Bao, S. H. 2017. Development of a daytime cloud and haze detection algorithm for Himawari-8 satellite measurements over central and eastern China. *Journal of Geophysical Research-Atmospheres, 122*(6), 3528-3543.

Sur, R., Peng, W. Y., Strand, C., Spearrin, R. M., Jeffries, J. B., Hanson, R. K., . . . Sridharan, A. K. 2017. Mid-infrared laser absorption spectroscopy of NO_2 at elevated temperatures. *Journal of Quantitative Spectroscopy & Radiative Transfer, 187*, 364-374.

Tie, X. X., Zhang, R. Y., Brasseur, G., & Lei, W. F. 2002. Global NO_x production by lightning. *Journal of Atmospheric Chemistry, 43*(1), 61-74.

Toenges-Schuller, N., Stein, O., Rohrer, F., Wahner, A., Richter, A., Burrows, J. P., . . . Elvidge, C. D. 2006. Global distribution pattern of anthropogenic nitrogen oxide emissions: Correlation analysis of satellite measurements and model calculations. *Journal of Geophysical Research-Atmospheres, 111*(D5).

van der A, R. J., Eskes, H. J., Boersma, K. F., van Noije, T. P. C., Van Roozendael, M., De Smedt, I., . . . Meijer, E. W. 2008. Trends, seasonal variability and dominant NO_x source derived from a ten year record of NO_2 measured from space. *Journal of Geophysical Research-Atmospheres, 113*(D4).

Vidot, J., Jourdan, O., Kokhanosvky, A. A., Szczap, F., Giraud, V., & Rozanov, V. V. 2010. Retrieval of tropospheric NO_2 columns from satellite measurements in presence of cirrus: A theoretical sensitivity study using SCIATRAN and prospect application for the A-Train. *Journal of Quantitative Spectroscopy & Radiative Transfer, 111*(4), 586-601.

Wei, J., Li, Z. Q., Wang, J., Li, C., Gupta, P., & Cribb, M. 2023. Ground-level gaseous pollutants (NO_2, SO_2, and CO) in China: daily seamless mapping and spatiotemporal variations. *Atmospheric Chemistry and Physics, 23*(2), 1511-1532.

Xiao, C. L., Hua, T. J., Peng, W. Y., Chao, Y., Jun, L., & Fu, C. L. 2023. Tropospheric NO2 changes in global typical regions during the COVID-19 pandemic from GF-5 EMI observations. . *National Remote Sensing Bulletin, 27*(08), 1807-1820.

Xinwei, L., Wubin, H., & Yingsha, J. 2021. Study of the Classified Identification of the Strong Convective Weathers

Based on the LightGBM Algorithm. *Plateau Meteorology, 40*(04), 909-918.

Yukun, C., & eng, Z. M. 2019. Ultra-short-term Prediction of Wind Farm Power Generation Based on Principal Component Analysis and LightGBM. *35*(06), 562-566.

Zang, L., Mao, F. Y., Guo, J. P., Gong, W., Wang, W., & Pan, Z. X. 2018. Estimating hourly PM_1 concentrations from Himawari-8 aerosol optical depth in China. *Environmental Pollution, 241*, 654-663.

Zhang, C. X., Liu, C., Li, B., Zhao, F., & Zhao, C. H. 2022. Spatiotemporal neural network for estimating surface NO2 concentrations over north China and their human health impact. *Environmental Pollution, 307*.