

2023

## NeSNet: A Deep Network for Estimating Near-Surface Pollutant Concentrations

Prasanjit Dey

*Technological University Dublin, Ireland, d22124678@mytudublin.ie*

Bibhash Pran Das

*National Institute of Technology Rourkela, India*

Yee Hui Lee

*Nanyang Technological University, Singapore*

*See next page for additional authors*

Follow this and additional works at: <https://arrow.tudublin.ie/scschcomart>



Part of the [Computer Engineering Commons](#)

---

### Recommended Citation

Dey, Prasanjit; Pran Das, Bibhash; Lee, Yee Hui; and Dev, Soumyabrata, "NeSNet: A Deep Network for Estimating Near-Surface Pollutant Concentrations" (2023). *Articles*. 199.

<https://arrow.tudublin.ie/scschcomart/199>

This Article is brought to you for free and open access by the School of Computer Science at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact [arrow.admin@tudublin.ie](mailto:arrow.admin@tudublin.ie), [aisling.coyne@tudublin.ie](mailto:aisling.coyne@tudublin.ie), [vera.kilshaw@tudublin.ie](mailto:vera.kilshaw@tudublin.ie).



This work is licensed under a [Creative Commons Attribution-Share Alike 4.0 International License](#).

Funder: Science Foundation Ireland; ADAPT SFI Research Centre, UCD

---

**Authors**

Prasanjit Dey, Bibhash Pran Das, Yee Hui Lee, and Soumyabrata Dev

# NeSNet: A Deep Network for Estimating Near-Surface Pollutant Concentrations

Prasanjit Dey , Bibhash Pran Das, Yee Hui Lee , *Senior Member, IEEE*, and Soumyabrata Dev , *Member, IEEE*

**Abstract**—With the threat of atmospheric pollution on the rise in recent years, round-the-clock monitoring of the concentration of atmospheric gases has become utterly necessary. As opposed to traditional in situ measurement strategies, satellite monitoring offers a convenient alternative for truly global coverage. However, satellite measurements do not provide information about the vertical profile of concentration, and estimation methods must be used to deduce near-surface concentration. Existing works that address this problem often adopt approaches that use auxiliary variables such as meteorological parameters and population density information along with vertical column density (VCD) measurements. In remote areas where such information is not available, these methods are likely to fail. In our work, we propose a near-surface network, a convolutional neural network that has been designed to perform the estimation of near-surface concentrations of atmospheric trace gases using only VCD values. We demonstrate the working of our method for nitrogen dioxide (NO<sub>2</sub>), sulfur dioxide (SO<sub>2</sub>), and ozone (O<sub>3</sub>). The proposed method shows RMSE scores of 6.272, 7.20, and 16.03 μg/m<sup>3</sup> for SO<sub>2</sub>, NO<sub>2</sub>, and O<sub>3</sub>, respectively. We also perform a detailed analysis of the impact of various factors on model performance. In the future, this method also use to determine the concentration of additional air pollutants including PM<sub>2.5</sub> and PM<sub>10</sub>. To possibly improve the effectiveness of the model, other meteorological variables, such as temperature, relative humidity, wind speed, and wind direction can be incorporated.

**Index Terms**—Atmospheric pollutants, ground observations, nitrogen dioxide, ozone, satellite measurements, sulfur dioxide.

## I. INTRODUCTION

**A**TMOSPHERIC pollution has become a serious threat in recent years. The advancements and expansion of industrial activity and civilization have been the major catalysts. With serious consequences like climate change and global warming,

Manuscript received 15 November 2022; accepted 27 January 2023. Date of publication 13 February 2023; date of current version 20 April 2023. This work was supported by Science Foundation Ireland under Grant 13/RC/2106\_P2 at the ADAPT SFI Research Centre at University College Dublin. ADAPT, the SFI Research Centre for AI-Driven Digital Content Technology, is funded by Science Foundation Ireland through the SFI Research Centres Programme. (Prasanjit Dey and Bibhash Pran Das contributed equally to this work.) (Corresponding author: Soumyabrata Dev.)

Prasanjit Dey is with the School of Computer Science, Technological University Dublin, D07 EWV4 Dublin, Ireland (e-mail: d22124678@mytudublin.ie).

Bibhash Pran Das is with the National Institute of Technology Rourkela, Rourkela 769005, India (e-mail: 118ei0355@nitrkl.ac.in).

Yee Hui Lee is with the Nanyang Technological University, Singapore 639798 (e-mail: eyhlee@ntu.edu.sg).

Soumyabrata Dev is with the ADAPT SFI Research Centre, School of Computer Science, University College Dublin, D04V1W8 Dublin, Ireland (e-mail: soumyabrata.dev@ucd.ie).

Digital Object Identifier 10.1109/JSTARS.2023.3244719

the onset of which is already being observed, keeping a check on atmospheric pollutant levels is now more important than ever. Trace gases play a major role in atmospheric chemistry. Many of these are also regarded as major atmospheric pollutants. The concentration of gases, such as (SO<sub>2</sub>), (O<sub>3</sub>), (NO<sub>2</sub>), etc., are indicators of air quality. Therefore, in this study, we primarily concern ourselves with concentrations of NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub>.

Oxidation of atmospheric NO<sub>2</sub> results in the formation of nitrogen aerosols, which impacts the particulate matter (PM) concentrations. Moreover, NO<sub>2</sub> is a precursor to O<sub>3</sub>. NO<sub>2</sub> in itself is an abundant atmospheric pollutant and has been proven to be harmful to humans. It may result in the contraction of various cardiovascular and respiratory ailments [1]. A high concentration of NO<sub>2</sub> contributes to acid rain, which causes corrosion. Stratospheric ozone is responsible for blocking UV rays. However, in the lower atmosphere, excess ozone concentration results in a decrease in agricultural yield [2] and also has an adverse effect on human health. Like NO<sub>2</sub>, sulfur dioxide also causes the formation of secondary pollutants. SO<sub>2</sub> causes respiratory illness, especially in children and the elderly. It also contributes to the formation of smog and acid rain.

The near-surface concentration of gases including NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub> have traditionally been measured from ground monitoring stations for accuracy in measurements [3]. This approach, however, is not devoid of demerits. The greatest among them is the fact that for monitoring concentrations in an area, all necessary monitoring equipment must be set up in the locality. This proves to be a challenge in remote corners of the world. Moreover, even for areas that have monitoring stations, the measurements are accurate only till certain distances from the monitoring stations. Therefore, high-resolution concentration measurements are usually only available in urban areas [4], [5]. In such a scenario, satellite measurements offer a lucrative alternative [6], [7]. Atmosphere observation instruments such as Ozone Monitoring Instrument (OMI) aboard NASA's Aura [8] satellite and Tropospheric Ozone Monitoring Instrument (TROPOMI) aboard ESAs Sentinel 5-p satellite provide round-the-clock measurements of atmospheric trace gas concentrations [9], [10]. The measurements are in the form of vertical column density (VCD) [11]. Although VCD values have a strong correlation with near-surface concentration, these measurements do not provide information about the vertical distribution of concentration. This means that the exact surface concentration of atmospheric gases cannot be measured from satellite instruments. It is because of this reason that various methods have been

proposed for the estimation of the near-surface concentration of atmospheric gases from tropospheric or total VCD.

Early works proposed in this area were mostly based on atmospheric chemistry models and simple regression models [12], [13]. In the last decade, the use of machine learning methods for the estimation of various atmospheric gas concentrations has been on the rise. In recent years, deep learning is also slowly making its way into the area [14], [15], [16]. In the work by Lamsal et al. [17] GEOS-Chem atmospheric chemistry model is used for the simulation of surface NO<sub>2</sub> concentration over the United States and Canada. Gu et al. [18] use RAMS-CMAQ air quality modeling system for a similar study over China. In the last few years, the use of machine learning models for the estimation of near-surface NO<sub>2</sub> concentration has garnered attention. Most of the methods that have been proposed use a number of features along with tropospheric VCD of NO<sub>2</sub> to perform the estimation. For example, in their work, Kang et al. [19] used a large number of satellite data variables along with auxiliary variables such as population density, wind speed, direction, etc., to construct estimation models. Support vector regressor (SVR), random forest (RF), XGBoost, and Light GBM were used. The introduction of deep learning to this field is relatively new. In their work, Li and Wu [20] proposed the use of residual deep neural nets for imputing missing data in tropospheric VCD of NO<sub>2</sub> and performing an estimation of near-surface concentration. Chan et al. [21] proposed a neural-network-based method for the estimation of NO<sub>2</sub> concentration over Germany. A few other deep learning methods have been proposed for similar tasks [22], [23], [24].

In this work, we propose a Near-Surface Network (NeSNet), a deep learning method using convolutional neural networks for the estimation of the near-surface concentration of nitrogen dioxide, sulfur dioxide, and ozone using tropospheric/total VCD measurements from TROPOMI aboard ESAs Sentinel 5-p satellite. The study is carried out over the landmass of Ireland. Although CNNs have been used in other domains in remote sensing [25], [26], to the best of our knowledge, they have not been applied to this particular task. The novel contributions of this work are as follows:

- 1) application of convolutional neural networks in the domain of satellite-based remote sensing;
- 2) extensive study of the performance of deep learning models for estimation of near-surface concentrations for multiple trace gases;
- 3) independence of the proposed method from ground-based measurements, thus making it suitable for prediction over remote locations and, therefore, giving true global coverage;
- 4) univariate approach to the estimation problem, i.e., we only use VCD values as inputs to our model.

The network has been designed for univariate estimation of near-surface concentrations. We only consider VCD values as input to the model. However, we also experiment with using satellite altitude as an auxiliary input to the model and record the change in performance. Here, we incorporated altitude to investigate how the model might behave at higher altitudes.

## II. DATASET

In this work, we consider the area over Ireland for the estimation of the near-surface concentration of NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub> using satellite measurements. Therefore, the dataset created for this purpose has two facets, ground monitoring station measurements and satellite measurements of tropospheric or total VCD. The estimation is carried out using data over a period of 16 months ranging from January 1, 2020, to May 1, 2021, for three different regions for three pollutants. The level of pollution is dependent on the location and the type of pollutant. Therefore, we have studied three pollutants in different regions according to their higher impact on the air. The data preparation strategy for satellite and ground data are individually described in the following sections.

### A. Satellite Data

Satellite measurements of VCD for NO<sub>2</sub> and total VCD for SO<sub>2</sub> and O<sub>3</sub> are obtained from the TROPOMI aboard ESAs Sentinel 5-p satellite [27]. TROPOMI is a space-borne, nadir-viewing imaging spectrometer capable of measuring wavelength bands between ultraviolet and shortwave infrared. It operates with a push broom configuration and has a swath width of 2600 km on the earth's surface. The L2-processed products for each of these gases are available in the form of netCDF files with a hierarchical data structure. The tropospheric VCD values are available for various ground pixels for each day in the period of study. However, for SO<sub>2</sub> and O<sub>3</sub>, only total VCD data are available for the period. Therefore, for these two gases, estimation is made with total VCD instead of tropospheric VCD. We also note here that for a single day, multiple sets of measurements may be available at different times of the day for a given area. In our work, we will ignore concentration variations within the day and only concern ourselves with the estimation of daily average concentrations. To meet our needs, we extract the VCD values and perform a regridding procedure in order to obtain the data in the form of a matrix that represents a geospatial grid over the area of study. The exact geospatial coordinates that mark the area of study vary for each of these gases and are chosen according to the available ground monitoring stations.

For NO<sub>2</sub>, the coordinates of the lower-left and upper-right corners of the bounding box are (51.8285°N, -9.4003°E) and (54.323°N, -6.032°E), respectively. Inside this bounding box, a grid is defined such that each cell in the grid represents a certain geographic area. The resolution of the grid is fixed at 0.05° × 0.05°. Therefore, the grid has 49 rows and 67 columns. The following Algorithm I is a step-by-step description of the subsequent procedure.

With this procedure, a dataset of 485 matrices is obtained where each matrix represents gridded, satellite-measured VCD values.

For SO<sub>2</sub>, the lower-left and upper-right corners of the bounding box for the area of study are (51.795°N, -9.0893°E) and (55.004°N, -6.105°E), respectively. Defining a grid with resolution 0.05° × 0.05° gives 64 rows and 59 columns in the grid. The regridding procedure as described for NO<sub>2</sub> is followed to obtain daily average VCD matrices of size 64 ×

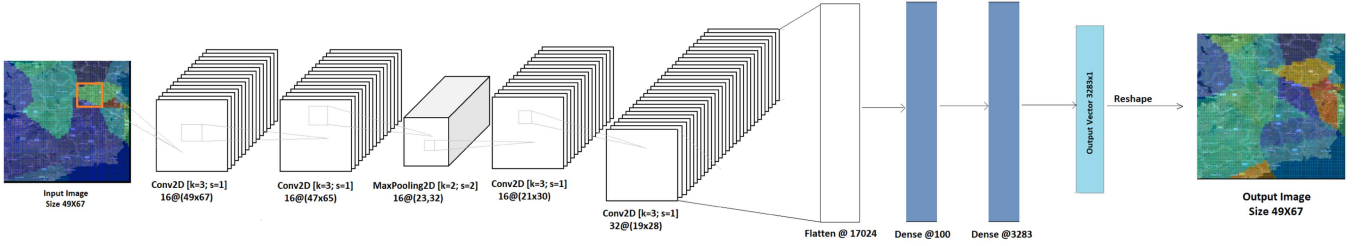


Fig. 1. Proposed network architecture. The shapes for different layers shown here correspond to that of the model for  $\text{NO}_2$ . The parameters denote the stride of the convolution operation and  $k$  denotes the kernel size.

---

### Algorithm 1: Dataset Preparation Procedure.

---

```

for day  $\in$  Period of study do
  matrices  $\leftarrow$  empty set
  for file  $\in$  files for day do
    VCDMatrix  $\leftarrow$  empty array of size  $49 \times 67$ 
    for cell  $\in$  VCDMatrix do
      cell  $\leftarrow$  VCD value of the nearest ground pixel
        in satellite data according to
        geographical distance.
    end for
    Append VCDMatrix to matrices
  end for
  MatrixForDay  $\leftarrow$  Mean{matrices}
  Add MatrixForDay to Dataset
end for

```

---

59. Similarly for  $\text{O}_3$ , choosing a bounding box of lower-left corner ( $51.8285^\circ\text{N}$ ,  $-10.3401^\circ\text{E}$ ) and upper-right corner ( $55.422^\circ\text{N}$ ,  $-6.022^\circ\text{E}$ ) gives matrices of size  $71 \times 86$ .

### B. Ground Data

Since the problem is one of estimation of near-surface concentrations of atmospheric trace gases, we require ground monitoring station data in order to train our model. These data are obtained from the Environmental Protection Agency of Ireland's monitoring stations. For  $\text{NO}_2$ , data from 29 active monitoring stations across Ireland are available during the period of January 1, 2020, to May 1, 2021. For  $\text{SO}_2$ , 14 stations are available, whereas for  $\text{O}_3$ , 20 stations are available. After obtaining the data for the entire period of study for all three gases, a regridding procedure similar to the one used for satellite data is used. The different bounding boxes as mentioned in the previous section are used to discern the geographical area of interest. A grid with resolution  $0.05^\circ \times 0.05^\circ$  is defined for this area. The concentration values are then assigned to the grid cells by using geographical distance to determine the nearest monitoring station. In the end, we obtain 485 matrices each for  $\text{NO}_2$ ,  $\text{SO}_2$ , and  $\text{O}_3$  having sizes  $49 \times 67$ ,  $64 \times 59$ , and  $71 \times 86$ , respectively.

## III. METHOD

Upon inspection of near-surface concentrations, we observe that certain local patterns are exhibited, i.e., close geographical

locations seem to have similar concentration values. Exploiting these local patterns in concentration may prove to be beneficial for an estimation model. Based on this, we propose a deep convolutional neural network to perform a regression of the near-surface concentration values from VCD measurements.

### A. Convolutional Neural Networks

Convolutional neural networks, brought into popularity by LeCun et al. [28], have emerged as the go-to algorithm for image tasks. They are based on the principle of filters that share their weights for individual layers. A single convolutional layer may contain several learnable filters. These filters are applied to an image by convolution using sliding windows that convolve with all regions of the input image with an overlapping distance called the stride and produce outputs known as feature maps. The weights of the filters are learned during model training. Multiple convolutional layers can be stacked to build a hierarchical feature extractor.

The convolutional operation between the input feature maps and a convolutional layer within the CNN architecture is given in

$$h_j^{(n)} = \sum h_k^{(n-1)} * w_{kj}^{(n)} + b_j^{(n)} \quad (1)$$

where  $*$  denotes a 2-D convolution,  $h_j^{(n)}$  is the  $j$ th feature map output in the  $n$ th hidden layer,  $h_k^{(n-1)}$  is the  $k$ th channel in the  $(n-1)$ th hidden layer, is the  $k$ th channel in the  $j$ th filter in the  $n$ th layer and  $b_j^{(n)}$  is its corresponding bias term.

CNNs owe their performance superiority mainly to the following characteristics:

- 1) extraction of local patterns in a hierarchical form by the means of stacked convolution layers;
- 2) translation invariance by virtue of the convolution operation;
- 3) reduced number of parameters as compared to feed-forward networks owing to weight sharing.

In light of these characteristics, we use a convolutional neural network model for further experiments on our task. The architecture of this model is described in the subsequent section.

### B. Proposed Network Architecture

The proposed network architecture is depicted in Fig. 1, and is made of four convolutional layers, two dense layers, and a max-pooling layer. All convolution layers used in the models

TABLE I  
COMPARISON OF PERFORMANCE OF THE PROPOSED MODEL WITH OTHER METHODS

Metric	Proposed method			Linear regression			XGBoost			LGBM		
	NO <sub>2</sub>	SO <sub>2</sub>	O <sub>3</sub>	NO <sub>2</sub>	SO <sub>2</sub>	O <sub>3</sub>	NO <sub>2</sub>	SO <sub>2</sub>	O <sub>3</sub>	NO <sub>2</sub>	SO <sub>2</sub>	O <sub>3</sub>
RMSE	7.20	6.272	16.03	8.595	7.99	19.36	8.534	7.980	19.488	8.538	7.981	19.530
MAE	4.94	2.965	11.804	6.271	4.791	14.04	6.217	4.796	14.157	6.221	4.796	14.194
R <sup>2</sup> <sub>DAvg</sub>	0.649	0.779	0.734	0.0806	0.001	0.106	0.046	-0.015	0.079	0.044	-0.014	0.071

The entries marked in blue represent the best score under the concerned metric among all methods for that gas.

have filters of size  $3 \times 3$ . The input layer feeds the satellite measurements matrix to the first convolution layer which has 16 filters of size  $3 \times 3$ . The second convolution layer also has 16 filters. The output of this layer is passed to a max-pooling layer with pool size 2 and strides 2. Therefore, this layer effectively reduces the dimension of the feature map by half. The output of this layer is then fed to another convolution layer with 16 filters followed by a convolution layer with 32 filters. This makes up the convolutional block of the model. The output of this convolution block is then flattened to make it suitable for feed-forward processing. This output is then passed to a dense layer having 100 units and a linear activation function. Finally, for regressing the ground concentration value, a dense layer with the same number of units as the number of grid cells in the ground concentration matrix is used. Hence, the length of the output vectors of the models for NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub> are 3283, 3776, and 6106, respectively.

#### IV. RESULTS

In this section, we describe the procedure for training the model. Subsequently, we discuss model validation and compare the performance of the proposed method against previously used methods. We also perform a detailed analysis of the effect of various parameters on model performance. The Python code for this work has been documented in a reproducible manner and is publicly available at <https://github.com/Bibhash123/NESNet>.

##### A. Model Training and Validation

In this study, we have generated satellite VCD data to calculate the concentration of surface pollution namely NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub>. We have used this satellite VCD data in the CNN model to evaluate and predict the concentration of near-surface air pollutants in the various areas near the station. The proposed architecture are trained individually for NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub>. Models with the proposed architecture are defined such that for each of these models, the output shape is the same as the number of grid cells for which concentration is to be estimated. The loss function for training is chosen as the mean-squared error. This is because we want our estimated values to be close to the true concentration values and, therefore, are interested in minimizing the deviation. Also during training, we observe the root-mean-squared error value at each epoch. To achieve proper convergence of the model an Adam optimizer with a learning

rate of 0.001 is used. Adam optimizer has gained widespread popularity in recent years due to its reliable performance.

As mentioned previously, our dataset for each of these gases has 485 samples. In order to train and validate our models with this data, a fivefold cross-validation strategy is used. This implies that in each fold, 80% of the data is used for training and 20% is used for validation. The model is trained for 50 epochs. However, most of the time, the model seems to start overfitting to the training data after 30 epochs. To tackle this problem, model checkpoints are used to save the model with the best validation root-mean-squared error. We also use early stopping with a patience of 5 to stop the training if the validation RMSE does not improve for 5 consecutive epochs. For NO<sub>2</sub>, the trained model gives a cross-validation RMSE score of  $7.20 \mu\text{g}/\text{m}^3$ . For SO<sub>2</sub>, this score is  $6.272 \mu\text{g}/\text{m}^3$ , whereas for O<sub>3</sub>, it is  $16.03 \mu\text{g}/\text{m}^3$ . We also validate our models using a few other metrics. The proposed method shows a mean absolute error (MAE) of  $4.94 \mu\text{g}/\text{m}^3$ ,  $2.965 \mu\text{g}/\text{m}^3$ , and  $11.804 \mu\text{g}/\text{m}^3$  for NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub>, respectively. Apart from observing the deviation of the predicted near-surface concentrations from the true values, we are also interested in analyzing how well the predicted values correlate with the true values. Therefore, Pearson's correlation coefficient seems to be an appropriate metric. However, in our case, we are trying to predict the daily average near-surface concentration because of which the satellite and ground concentration data are related for individual days. We, therefore, feel that it is more reasonable to calculate Pearson's correlation coefficient between predicted and true values individually for each day. To report the performance, all of these coefficients are averaged out. Throughout this article, we refer to this modified correlation metric as R<sup>2</sup><sub>DAvg</sub>. With the proposed method, R<sup>2</sup><sub>DAvg</sub> of 0.648, 0.779, and 0.734 are obtained for NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub>, respectively. Traditionally, various regression algorithms, boosting algorithms, and simple neural nets have been used for the task of concentration estimation. For example, Kang et al. [19] have used XGBoost and light GBM to perform an estimation using satellite VCD data as well as other auxiliary variables. Therefore, the performance of our model has been compared to that of other algorithms for different metrics in Table I. It is clear that in the case of univariate estimation, the proposed method performs all other previously used methods. Near-surface air pollutants such as NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub> gases have a significantly negative effect on human health and sustainable development. To identify the level of air pollutants, government agencies have defined the air quality index (AQI). It is calculated based on the level of several air pollutants. In this

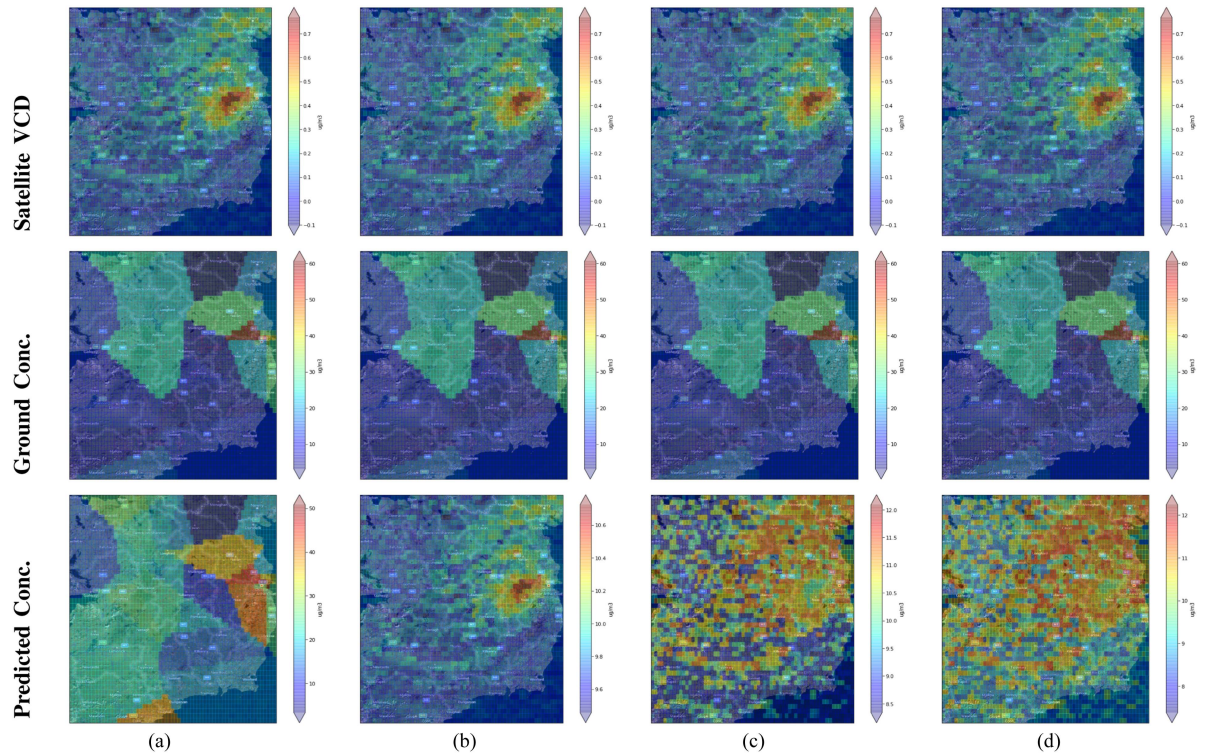
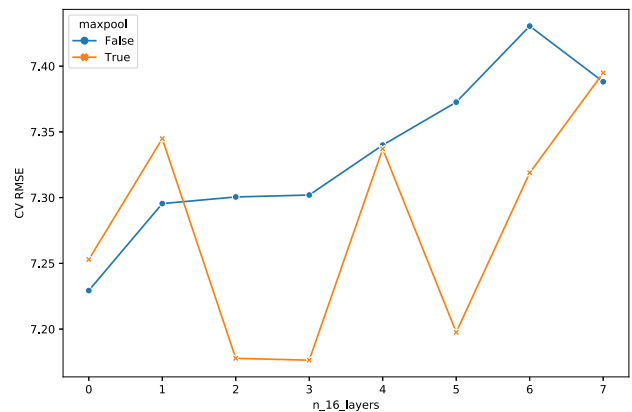


Fig. 2. Visualization of satellite concentration values, ground concentrations, and model prediction of ground concentrations for different estimation models. (a) Proposed method. (b) Linear regression. (c) XGBoost regression. (d) LGBM regression.

work, we employed the CNN model to analyze and forecast the concentration of near-surface air pollutants. The AQI is calculated using this concentration of air pollutants. Fig. 2 shows the satellite VCD, Ground concentration, and prediction concentration of three pollutants near the surface for the proposed model as well as Linear Regression XGBoost Regression, and LGBM regression. We clearly observe that local patterns in near-surface concentrations are preserved in the predictions for the proposed method. There are no such patterns that can be observed between ground concentration and prediction concentration for linear regression, XGBoost, or LGBM.

### B. Impact of Filter Layers on Model Performance

In the network architecture presented in this article, four convolutional layers with a max-pooling layer between the second and third convolutional layers are used. This architecture was arrived at via an experiment. In this analysis, we perform two subexperiments. In both of these, the number of filter layers in the model is gradually increased and the model is retrained. The cross-validation RMSE score is recorded in each case. To vary the filter layers initially, a model with one convolutional layer with 32 filters and 2 dense layers is defined. On top of this block, we gradually keep adding 16 filtered convolutional layers. In the first subexperiment, only convolutional layers are added, whereas in the second subexperiment after every three network layers, a max-pooling layer is added. The variations of the RMSE score obtained in both these cases are shown in Fig. 3(a). In the no max-pooling layer case, we can see that in general, the RMSE



(a)

Layers	Model Architecture
0	C(32)→F
1	C(16)→C(32)→F
2	C(16)→C(16)→M→C(32)→F
3	C(16)→C(16)→M→C(16)→C(32)→F
4	C(16)→C(16)→M→C(16)→C(16)→C(32)→F
5	C(16)→C(16)→M→C(16)→C(16)→C(16)→M→C(32)→F
6	C(16)→C(16)→M→C(16)→C(16)→C(16)→M→C(16)→C(32)→F
7	C(16)→C(16)→M→C(16)→C(16)→C(16)→M→C(16)→C(16)→C(32)→F

(b)

Fig. 3. (a) Model cross-validation RMSE score with changing architecture. The model used for this experiment is trained on data for NO<sub>2</sub>. (b) Architectures of the convolutional block in the model for a particular number of 16 filtered layers (Only the case with max-pooling is shown). C(x) denotes the convolution layer with x filters, F is a flattened layer, and M denotes a max-pooling layer.

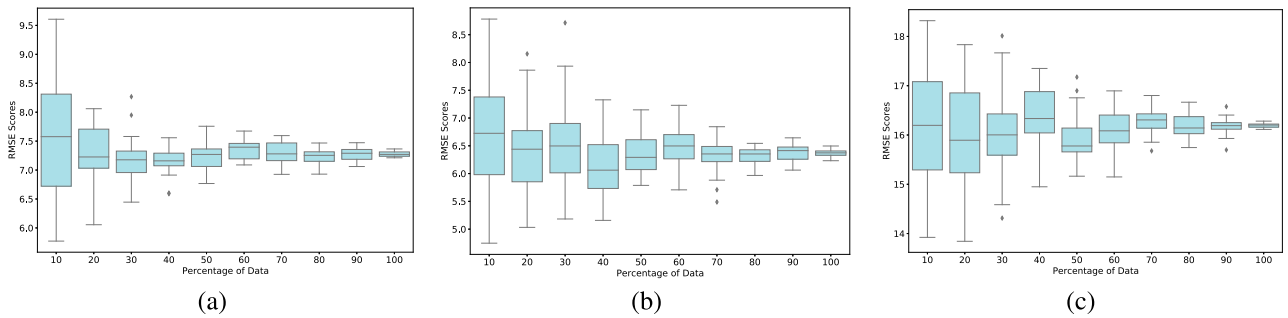


Fig. 4. (a) Dataset size versus RMSE score for  $\text{NO}_2$ . (b) Dataset size versus RMSE Score for  $\text{SO}_2$ . (c) Dataset size versus RMSE score for  $\text{O}_3$ .

TABLE II  
MEAN RMSE AND IQR OF RMSE SHOWN BY THE PROPOSED METHOD FOR DIFFERENT DATASET SIZES

% of Data	$\text{NO}_2$		$\text{SO}_2$		$\text{O}_3$	
	Mean RMSE	IQR	Mean RMSE	IQR	Mean RMSE	IQR
10	7.632	1.683	6.702	1.417	16.204	1.803
20	7.319	0.694	6.457	1.055	15.969	1.710
30	7.202	0.3806	6.575	0.898	16.044	0.947
40	7.164	0.2439	6.145	0.824	16.432	0.852
50	7.250	0.3332	6.372	0.565	15.974	0.609
60	7.354	0.2886	6.429	0.439	16.119	0.602
70	7.303	0.3077	6.320	0.300	16.277	0.317
80	7.241	0.165	6.315	0.222	16.203	0.365
90	7.266	0.190	6.377	0.230	16.185	0.141
100	7.279	0.077	6.370	0.085	16.193	0.065

score deteriorates (increases) with an increase in model depth. On the other hand, fluctuations in the RMSE score are observed in the case of max-pooling. However, we note that optimum performance is obtained in the case where three 16-filtered layers are used. As shown in Fig. 3(b), this architecture corresponds to the one that has been presented as the proposed architecture in this work.

### C. Impact of Dataset Size on Model Performance

In the majority of past applications of deep learning models, it has very well been established that the size of the training dataset is pivotal to the model performance. To validate this idea for our proposed model, we carry out an experiment to observe variations in the performance of the model with changing dataset size. Our dataset originally has 485 samples. In this experiment, we perform a random sampling of the data to select training sets with varying sizes and then retrain the model for each sampled dataset using a threefold cross-validation strategy. The steps of the experiment are detailed as follows.

- 1) 10% of the total dataset is selected by means of random sampling.
- 2) A model with the proposed architecture is trained and validated on the sampled data using a threefold cross-validation strategy. The average RMSE score is recorded.

- 3) Steps 1) and 2) are repeated 20 times. In each repetition, the randomly sampled dataset varies. Therefore, a spread of RMSE score values is obtained.
- 4) The values obtained in step (4) are plotted in a box plot.
- 5) Steps 1)–4) are repeated by gradually varying the sampling size from 10% to 100% in steps of 10%.

The box plots obtained in this manner for  $\text{NO}_2$ ,  $\text{SO}_2$ , and  $\text{O}_3$  are shown in Fig. 4. It is evident from the plots for all three cases that as the dataset size increases, the spread of the RMSE score decreases. This means that for larger training dataset sizes, the model is more robust to variations in the data the most robust model being obtained when using the entire dataset. Numerical evidence of this observation is shown in Table II. The interquartile range (IQR) is the difference between the RMSE values Q2 and Q1, where Q1 represents the point for which the 25 percentile of the cases have a lower RMSE and Q2 represents the point for which the 75 percentile of the cases have a lower RMSE. The IQR, therefore, gives a measure of the spread of the RMSE score. It is clear that as the dataset size increases, the IQR decreases.

### D. Inclusion of Altitude Parameter on Model Performance

Throughout this work, we have explored a univariate approach to the estimation of near-surface concentrations of  $\text{NO}_2$ ,  $\text{SO}_2$ , and  $\text{O}_3$ . In our proposed method, we have only considered VCD



TABLE III  
PERFORMANCE EVALUATION WHEN SATELLITE ALTITUDE DATA IS USED AS AN  
AUXILIARY VARIABLE FOR MODEL TRAINING COMPARED TO WHEN ONLY  
VCD DATA WITHOUT ALTITUDE IS USED

Methods		RMSE	MAE	R2 <sub>DAvg</sub>
With altitude	Linear regression	8.603	6.279	0.037
	XGBoost	8.549	6.234	0.049
	LGBM	8.554	6.238	0.047
	NeSNet (Proposed)	7.599	5.306	0.647
Without altitude	Linear regression	8.595	6.271	0.0806
	XGBoost	8.534	6.217	0.046
	LGBM	8.538	6.221	0.044
	NeSNet (Proposed)	7.20	4.94	0.649

values as inputs to our model. In this section, we discuss the possibility of including other variables such as temperature, relative humidity, wind speed, and wind direction as predictors as well. To that end, we conduct an experiment by including the altitude of the satellite at respective scan lines as another input in our model. Here, we incorporated altitude to investigate how the model might behave at higher altitudes. However, in the data available to us, the altitude values were only available for NO<sub>2</sub>. Therefore, the results of this experiment are shown for NO<sub>2</sub>.

For the experiment, a slightly modified architecture of the proposed neural net is considered. We take two independent convolutional blocks of the same architecture as that of the one in the proposed method. The inputs to the first convolutional block are the VCD matrices and the input to the second convolutional block are satellite altitude matrices. These satellite altitude matrices are obtained by adopting the same gridding procedure that is used for obtaining VCD matrices (described in the dataset section). The outputs of both these convolutional blocks are then concatenated and passed to the block of fully connected layers of the same structure as for the proposed model. This approach gives us a model with two input heads and a single-output head. One more aspect that is considered here is the fact the scale of values for VCD and altitude do not match and altitude values are particularly high in magnitude. To scale both types of inputs, standardization is used. The necessity of standardization was ascertained by the observation that model performance deteriorated significantly when standardization was not done. The results obtained in this experiment are shown in Table III.

## V. CONCLUSION

In this work, we have proposed a convolutional-neural-net-based method for the estimation of the near-surface concentration of NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub> using measurements of VCD from Sentinel 5-p TROPOMI. Our method performs this estimation using only satellite VCD measurements and does not rely on any ground-based measurements. In our experiments, we have shown that our method performs significantly better

than methods that have been traditionally used in this domain. Our model gives RMSE scores of 6.272, 7.20, and 16.03  $\mu\text{g}/\text{m}^3$  for SO<sub>2</sub>, NO<sub>2</sub>, and O<sub>3</sub>, respectively. Also, we clearly establish that exploiting local patterns in concentrations can give good performance even when a univariate estimation is performed. While the proposed method gives a good correlation between predicted and true near-surface concentrations for individual days, other methods show drastically poor correlations for the univariate estimation case.

In the future, this work may be extended to perform similar estimations for other atmospheric pollutants, such as PM<sub>2.5</sub> and PM<sub>10</sub>. To possibly improve the effectiveness of the model, other meteorological variables such as temperature, relative humidity, wind speed, and wind direction can be incorporated. There is no meteorological information in the VCD data. To do this, we must gather meteorological data from a ground-based sensor. In the future, we develop an algorithm to correlate meteorological data to the VCD data.

## REFERENCES

- [1] C.-F. Wu et al., "Association of short-term exposure to fine particulate matter and nitrogen dioxide with acute cardiovascular effects," *Sci. Total Environ.*, vol. 569–570, pp. 300–305, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0048969716312578>
- [2] R. Van Dingenen, F. J. Dentener, F. Raes, M. C. Krol, L. Emberson, and J. Cofala, "The global impact of ozone on agricultural crop yields under current and future air quality legislation," *Atmospheric Environ.*, vol. 43, no. 3, pp. 604–618, 2009. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1352231008009424>
- [3] S. Dev, Y. H. Lee, and S. Winkler, "Color-based segmentation of sky/cloud images from ground-based cameras," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 1, pp. 231–242, Jan. 2017.
- [4] J. Wu, F. Orlandi, D. O'Sullivan, E. Pisoni, and S. Dev, "Boosting climate analysis with semantically uplifted knowledge graphs," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 4708–4718, May 2022.
- [5] J. Wu, F. Orlandi, D. O'Sullivan, and S. Dev, "Linkclimate: An interoperable knowledge graph platform for climate data," *Comput. Geosci.*, vol. 169, 2022, Art. no. 105215.
- [6] X. Y. Zhang, L. M. Zhao, M. M. Cheng, and D. M. Chen, "Estimating ground-level ozone concentrations in eastern China using satellite-based precursors," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4754–4763, Jul. 2020.
- [7] S. Bhattacharjee and J. Chen, "Prediction of satellite-based column CO<sub>2</sub> concentration by combining emission inventory and LULC information," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8285–8300, Dec. 2020.
- [8] M. Schoeberl et al., "Overview of the EOS aura mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1066–1074, May 2006.
- [9] L. Shikwambana, P. Mhangara, and N. Mbatha, "Trend analysis and first time observations of sulphur dioxide and nitrogen dioxide in South Africa using TROPOMI/Sentinel-5 P data," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 91, 2020, Art. no. 102130.
- [10] J. Veeffkind et al., "TROPOMI on the ESA Sentinel-5 precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications," *Remote Sens. Environ.*, vol. 120, pp. 70–83, 2012.
- [11] C. Li et al., "Version 2 Ozone monitoring instrument SO<sub>2</sub> product (OMSO2 V2): New anthropogenic SO<sub>2</sub> vertical column density dataset," *Atmospheric Meas. Techn.*, vol. 13, no. 11, pp. 6175–6191, 2020.
- [12] B. Mijling and R. J. van der A, "Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale," *J. Geophys. Res., Atmospheres*, vol. 117, no. D17, 2012, Art. no. D17302. [Online]. Available: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JD017817>
- [13] E. V. Novotny, M. J. Bechle, D. B. Millet, and J. D. Marshall, "National satellite-based land-use regression: NO<sub>2</sub> in the United States," *Environ. Sci. Technol.*, vol. 45, no. 10, pp. 4407–4414, 2011, doi: [10.1021/es103578x](https://doi.org/10.1021/es103578x).

- [14] P. Dey, S. Chaulya, and S. Kumar, "Hybrid CNN-LSTM and IoT-based coal mine hazards monitoring and prediction system," *Process Saf. Environ. Protection*, vol. 152, pp. 249–263, 2021.
- [15] K. Kumari et al., "UMAP and LSTM based fire status and explosibility prediction for sealed-off area in underground coal mine," *Process Saf. Environ. Protection*, vol. 146, pp. 837–852, 2021.
- [16] P. Dey et al., "T-SNE and variational auto-encoder with a Bi-LSTM neural network-based model for prediction of gas concentration in a sealed-off area of underground coal mines," *Soft Comput.*, vol. 25, pp. 14183–14207, 2021.
- [17] L. Lamsal et al., "Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument," *J. Geophys. Res.*, vol. 113, 2008.
- [18] J. Gu et al., "Ground-level NO<sub>2</sub> concentrations over China inferred from the satellite OMI and CMAQ model simulations," *Remote Sens.*, vol. 9, no. 6, 2017, Art. no. 519.
- [19] Y. Kang et al., "Estimation of surface-level NO<sub>2</sub> and O<sub>3</sub> concentrations using TROPOMI data and machine learning over east Asia," *Environ. Pollut.*, vol. 288, 2021, Art. no. 117711. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0269749121012938>
- [20] L. Li and J. Wu, "Spatiotemporal estimation of satellite-borne and ground-level NO<sub>2</sub> using full residual deep networks," *Remote Sens. Environ.*, vol. 254, 2021, Art. no. 112257. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0034425720306301>
- [21] K. L. Chan, E. Khorsandi, S. Liu, F. Baier, and P. Valks, "Estimation of surface NO<sub>2</sub> concentrations over Germany from TROPOMI satellite observations using a machine learning method," *Remote Sens.*, vol. 13, p. 969, 2021.
- [22] S. M. Cabaneros, J. K. Calautit, and B. Hughes, "Spatial estimation of outdoor NO<sub>2</sub> levels in central London using deep neural networks and a wavelet decomposition technique," *Ecol. Model.*, vol. 424, 2020, Art. no. 109017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0304380020300892>
- [23] T. Li, Y. Wang, and Q. Yuan, "Remote sensing estimation of regional NO<sub>2</sub> via space-time neural networks," *Remote Sens.*, vol. 12, no. 16, 2020, Art. no. 2514. [Online]. Available: <https://www.mdpi.com/2072-4292/12/16/2514>
- [24] T. Li and X. Cheng, "Estimating daily full-coverage surface ozone concentration using satellite observations and a spatiotemporally embedded deep learning approach," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 101, 2021, Art. no. 102356. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0303243421000635>
- [25] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
- [26] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, Feb. 2017.
- [27] K. N. M. I., "(KNMI), Copernicus sentinel data processed by ESA," 2019, sentinel-5P TROPOMI Tropospheric NO<sub>2</sub> 1-Orbit L2 5.5 km × 3.5 km, *Goddard Earth Sci. Data Inf. Serv. Center (GES DISC), Greenbelt, MD, USA*, 2019xs. Accessed: Jun., 1 2021, doi: [10.5270/S5P-s4ljg54](https://doi.org/10.5270/S5P-s4ljg54).
- [28] Y. LeCun et al., "Convolutional networks for images, speech, and time-series," *Handbook Brain Theory Neural Netw.*, vol. 3361, no. 10, 1995.



**Prasanjit Dey** received the M.Tech. degree in information technology from the Maulana Abul Kalam Azad University of Technology, Kolkata, India, in 2016. He is currently working toward the Ph.D. degree in time-series data prediction using deep learning models with the School of Computer Science, Technological University Dublin, Dublin, Ireland.

His research project is on monitoring and short-term forecasting of atmospheric air pollutants using deep neural networks. After his M.Tech. degree, he was a Project Associate with the CSIR-Central Institute of Mining and Fuel Research, India. His research interest includes predicting time-series data using deep learning and machine learning algorithms.



**Bibhash Pran Das** received the B.Tech. degree in electronics and instrumentation engineering from the National Institute of Technology Rourkela, Rourkela, India, in 2022.

He is currently a Software Engineer with Microsoft, Hyderabad, India. His research interests include deep learning and computer vision.



**Yee Hui Lee** (Senior Member, IEEE) received the B.Eng. (Hons.) and M.Eng. degrees from the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore, in 1996 and 1998, respectively, and the Ph.D. degree from the University of York, York, U.K., in 2002, all in electronic engineering.

Since 2002, she has been a Faculty Member with Nanyang Technological University, where she is currently an Associate Professor with the School of Electrical and Electronic Engineering. Her research interests include the channel characterization, rain propagation, antenna design, electromagnetic bandgap structures, and evolutionary techniques.



**Soumyabrata Dev** (Member, IEEE) received the B.Tech. degree (*summa cum laude*) in electronics and communication engineering from the National Institute of Technology Silchar, Silchar, India, in 2010, and the Ph.D. degree in electrical and electronic engineering from Nanyang Technological University (NTU), Singapore, in 2017.

He is currently an Assistant Professor with the School of Computer Science, University College Dublin, Dublin, Ireland, and an SFI Funded Investigator with ADAPT SFI Research Centre, Dublin.

In 2015, he was a Visiting Student with Audiovisual Communication Laboratory (LCAV), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He has authored more than 120 publications in leading journals and conferences. His research interests include remote sensing, statistical image processing, machine learning, and deep learning.