University of Tehran Press

Pollution

Print ISSN: 2383-451X
Online ISSN: 2383-4501

https://jpoll.ut.ac.ir/

Prediction of Air Pollutants Concentration Emitted from Kirkuk Cement Plant Based on Deep Learning and Gaussian Equation Outputs

Qayssar Mahmood Ajaj[1, 2] | Helmi Zulhaidi Mohd Shafri[1]✉ | Mohammad Firuz Ramli[3] | Aimrun Wayayok[4]

1.Department of Civil Engineering and Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia
2. Department of Surveying Techniques Engineering, Technical Engineering College of Kirkuk, Northern Technical University, Kirkuk 36001, Iraq
3. Department of Environmental Sciences, Universiti Putra Malaysia, Serdang 43400, Malaysia
4. Department of Biological and Agricultural Engineering, Universiti Putra Malaysia, Serdang 43400, Malaysia

Article Info

Article type:
Research Article

Article history:
Received: 4 Dec 2022
Revised: 7 Jan 2023
Accepted: 19 Jan 2023

Keywords:
Deep learning (DL)
Convolution neural network (CNN)
Pollution prediction
Gaussian Plume Model
Gases

ABSTRACT

Researchers are interested in developing techniques to monitor, manage and predict the risks of gases and particles emitted from cement factories, which have a direct and negative impact on human health. Deep learning (DL) is a critical component of data mining, which further involves statistics and prediction. In this study, we developed a deep learning prediction model called the Deep Pollutant Prediction Model (DPPM). The data used for DPPM are separated into two types: observed data from a pollution monitoring station of the Institute of Mental Health in Ahmedabad City, India coded as (GJ001), to validate the model and simulated data generated using the Gaussian Plume Model for the hypothetical receptor (Laylan District, Kirkuk, Iraq) to predict the pollution that emitted from Kirkuk Cement Plant 5 km apart from the study area. The findings indicated that the DPPM has high efficiency in both Allahabad and Laylan stations, with more closed results for the data in the Laylan station, which is based on the Gaussian equation simulated data. Since the highest loss function value in the Laylan is 0.0221 of the CaO parameter, while it is 4.466 of the AQI parameter for the Allahabad Station, and the smallest loss function value in the Laylan is equal to 0.0041of both Fe2O3 and MgO parameters, it corresponds to 0.038 of Xylene for the Allahabad station. The results of the study proved that data continuity and non-volatility produce excellent outcomes for DPPM.

Cite this article: Mahmood Ajaj, Q., Shafri, H.Z.M., Firuz Ramli, M., & Wayayok, A. (2023). Prediction of Air Pollutants Concentration Emitted from Kirkuk Cement Plant Based on Deep Learning and Gaussian Equation Outputs. *Pollution*, 9 (2), 810-820.
http://doi.org/10.22059/POLL.2023.352186.1714

 Publisher: University of Tehran Press.
DOI: http://doi.org/10.22059/POLL.2023.352186.1714

INTRODUCTION

Several gases, including ozone (O_3), carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen oxides (NO_x), as well as suspended particles (PM), volatile organic compounds (VOC_s), certain metals, and other pollutants that emitted from sources of pollution, all have a negative and direct impact on human health in urban environments (Zhu et al., 2018). Many cities around the world suffer from high concentrations of air pollutants due to the notable

*Corresponding Author Email: *helmi@upm.edu.my*

economic expansion and development, therefore, assessment and prediction of these pollutants are necessary to assess the risks resulting from it (J. Zhang & Ding, 2017). Methods for predicting air pollutants, in general, are divided into two main types: statistical and deterministic methods (Li et al., 2017).

Artificial intelligence (AI) contributes to the development of effective prediction models in various fields of science by assisting in the rapid discovery of optimal solutions (Oprea et al., 2017). Researchers have long been interested in air quality prediction techniques in order to improve modeling using machine learning, particularly deep learning, and data processing in order to provide decision-makers with high-accuracy results and information (Schürholz et al., 2020).

Deep learning is a branch of machine learning that employs a network of neurons structured in layers, with three layers: input, output, and hidden (Mohsen et al., 2018). For the purpose of learning complex prediction models, such as multi-layer neural networks with numerous hidden units, DL refers to a collection of learning algorithms rather than a single method (Emmert-Streib et al., 2020). Deep learning can understand context information from various data sources and the relations between data and deep learning-based techniques' promising abilities drive researchers to use deep architectures in preference prediction problems as well (Khan et al., 2021).

Libraries are collections of pre-written code that users can utilize to improve their operations in programming languages. Python deep learning, AI, and machine learning libraries are used by developers to complete complex tasks (Davies, 2018).

Prediction of air pollutants has received much study and involves a number of problems and difficulties. Several research of various countries on the prediction of air pollutants depending on deep learning are discussed in this section. Convolution neural network (CNN) model practically proved its efficiency through the test of input data, since Mao & Lee, (2019) forecasted air quality by predicting the hourly concentration of air contaminants such as ozone, particle matter $PM_{2.5}$ using deep learning techniques. As well, Q. Zhang et al., (2020) suggested a model hybrid deep learning model that forecasts air quality at high resolution by merging Convolutional Neural Networks (CNN) with Long Short Term Memory (LSTM). Ma et al., (2020) presented an approach that uses ambient concentrations of 18 chemical indicators to rapidly assess the coefficients of air quality equations, using data generated by simulating a complex atmospheric chemical transport model. This method makes use of deep learning techniques and chemical indicators of pollutant formation (CTM).

Whenever two or more models are integrated, a more efficient model can be generated by comparing the combined model (hybrid) to each individual model independently. Heydari et al., (2022) anticipated and analyzed Combined Cycle Power Plant air pollution by creating a novel hybrid intelligence model based on LSTM and MVO to. Bekkar et al., (2021) combined historical data on pollutants, meteorological information, and $PM_{2.5}$ concentrations at nearby stations to use CNN-LSTM to predict the hourly forecast of $PM_{2.5}$ concentrations in Beijing, China. Deep learning recently has been utilized for various types of pollutants prediction as Isam Drewil & Jabbar Al-Bahadili, (2021) indicate that the most of the researchers utilize more advanced and intelligent approaches to deal with the problem of air pollution detection in the early stages, while only a few researchers use basic strategies. Muthukumar et al., (2022) employed the advanced deep predictive Convolutional LSTM (ConvLSTM) model along with the cutting-edge Graph Convolutional Network (GCN) architecture to predict temporal and spatial hourly $PM_{2.5}$ in the Los Angeles City over time.

The aim of this paper is to build a prediction model named Deep Pollutant Prediction Model for pollutants emitted from Kirkuk Cement Plant based on deep learning and simulated output Gaussian Plume Model of pollution concentration.

MATERIALS AND METHODS

This section describes the region of study as well as the parts of the proposed method for building a smart model based on deep learning to predict air pollutants, as follows:

Study Area

Laylan is a sub-district of Kirkuk Governorate, Iraq located 19 km southeast of the city of Kirkuk as illustrated in Fig 1. The geographical location of the Laylan district ranges between latitudes (44˚20′ E, 44˚ 45′ E) and longitudes (35˚ 08′ N, 35˚ 25′N). Laylan district is famous for agriculture, livestock, and industry. Where the Kirkuk Cement Plant is located to the west of Laylan district, about 5 km as indicates in figure (1), and it is considered the most environmentally affected area. The cement plant releases large quantities of pollutants that affect human health, especially respiratory diseases, and water pollution, according to the environmental report of the Iraqi Ministry of Environment. The population of Laylan district is about 16,000 and it is multi-national. It is also characterized by its picturesque nature, which made it a local tourist entertainment area. The Laylan District considers a Hypothetical receptor to apply Gaussian Plume Model and calculate the simulated data.

Dataset Description

Because the study area lacked long-term pollution data, the proposed model was applied to a dataset from the Institute of Mental Health in Ahmedabad City, India (GJ001). The GJ001 pollution dataset contains 12 parameters which are PM_{10}, PM, NO, NO_2, CO, SO_2, O_3, Benzene, Toluene, and Xylene, Air Quality Index (AQI), and compounds for 2009 observed days. Initially, this dataset was processed using the statistical missing values approach in order to accommodate the suggested deep learning model. Laylan District is selected to be a point for calculating the concentration of pollutants from the source (cement plant) using the Gaussian Plume Model

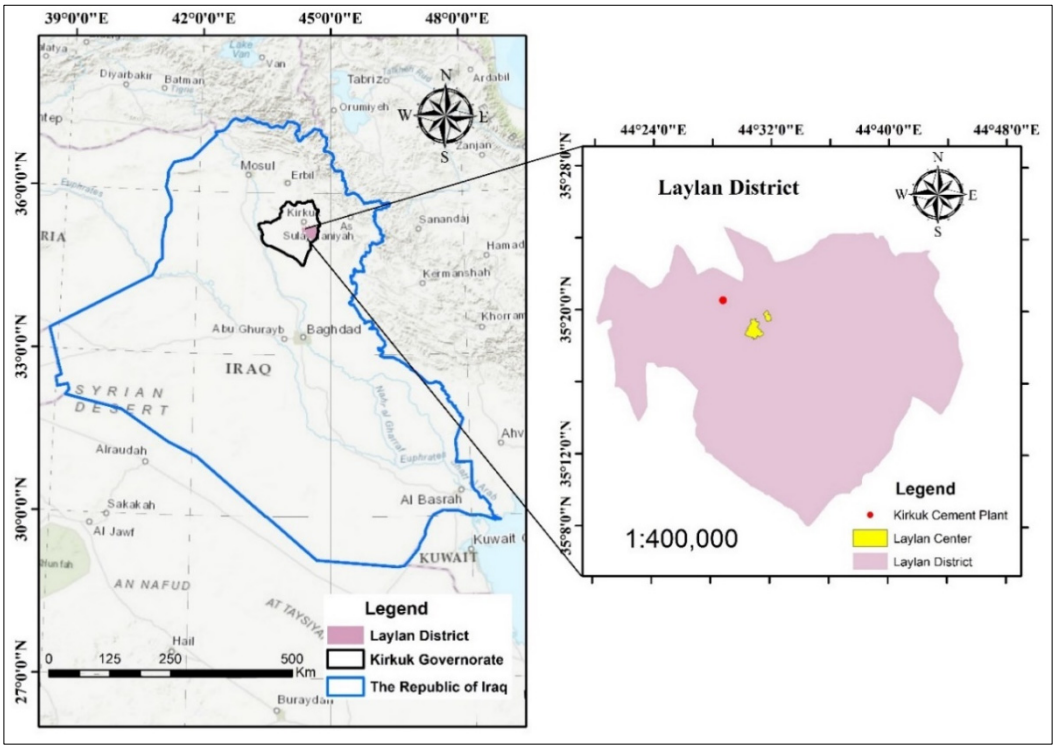

Fig. 1. The study area (Laylan District, Kirkuk, Iraq)

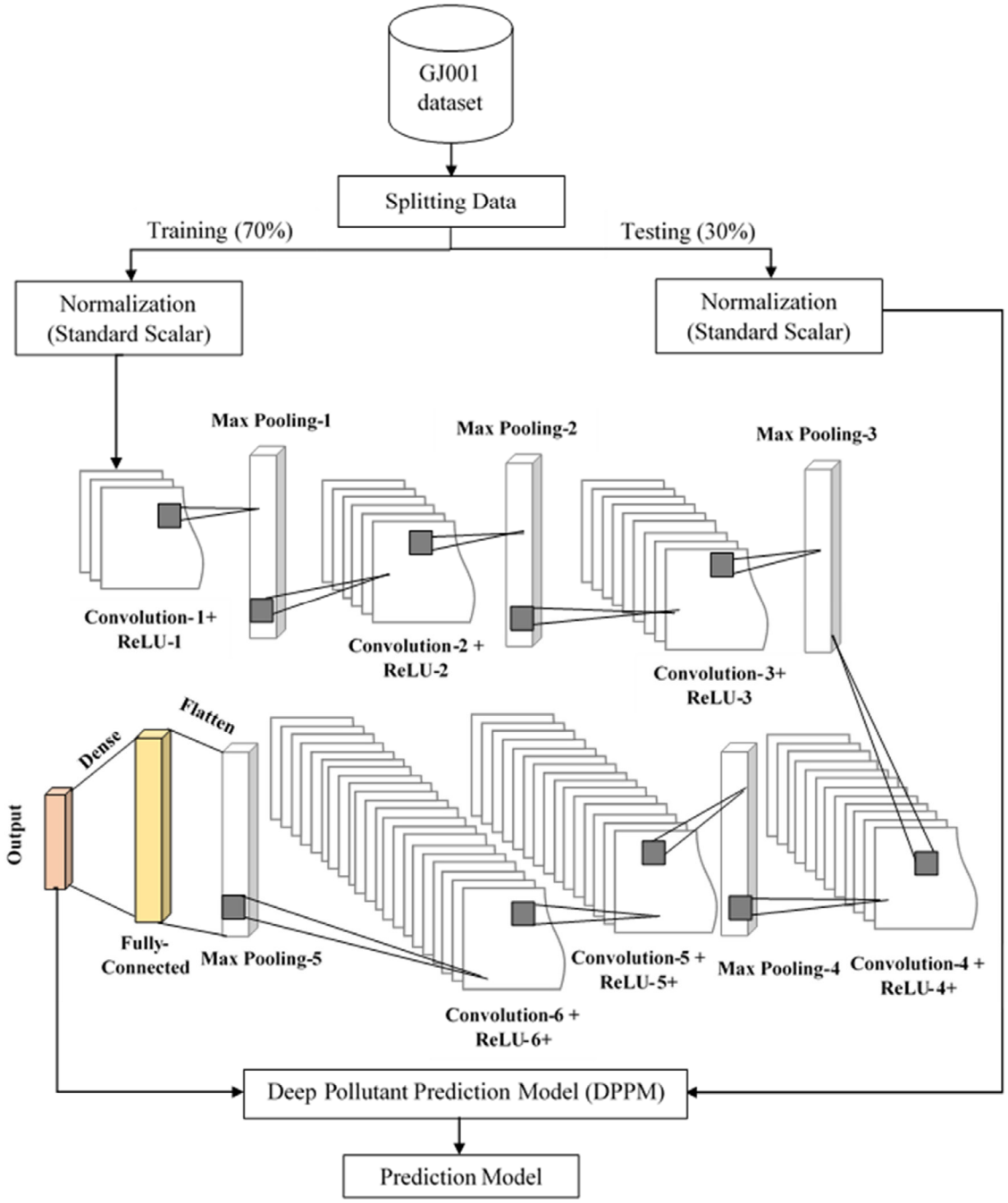

Fig. 2. Flowchart of the proposed model.

and the wind speed data obtained from the RETScreen program, which provides meteorological data from virtual stations of NASA. The simulated data were calculated for three years 2019, 2020, and 2021. The simulated data include 7 parameters which are Calcium Carbonate ($CaCO_3$), Silicon dioxide (SiO_2), Aluminum oxide (Al_2O_3), Iron (III) Oxide (Fe_2O_3), Calcium oxide (CaO), Magnesium oxide (MgO), and Sulfur trioxide (SO_3). These emissions are components of cement material, which are released in various amounts from the stack of the cement plant.

Proposed model

Building the suggested model is a crucial procedure since the model must be successfully represented, which has a significant effect on the success of the planned model when implemented applications. The suggested model is designed to predict pollution components emitted from

Table 1. Proposed model layers

NO.	Layer Type	Filters	Size/Stride	Activation Function
1	Convolutional	16	3/1	ReLU
2	Max Pooling	–	1/1	–
3	Convolutional	32	3/1	ReLU
4	Max Pooling	–	1/1	–
5	Convolutional	64	3/1	ReLU
6	Max Pooling	–	1/1	–
7	Convolutional	128	3/1	ReLU
8	Max Pooling	–	1/1	–
9	Convolutional	265	3/1	ReLU
10	Convolutional	512	3/1	ReLU
11	Max Pooling	–	1/1	–
12	Flatten	–	–	–
13	Dense	–	–	Softmax

cement plants in a short time and with reliable data. Figure (2) shows the overall procedures of DPPM model.

Data Splitting

Data splitting is a key step for model validation since it splits a given dataset into training and testing sets (Wu et al., 2013). The training data is then used to fit and evaluate statistics and machine learning models. If it Hold-out a set of data for validation separate from the training set, it may analyze and compare the accuracy of multiple models' predictions without being worried about potential overfitting of the training set (Castiñeira et al., 2020). The dataset is split into 70% for training and 30% for testing in this model.

Data Normalization

Normalization is the process of modifying data in order to compare statistics from various measures accurately by reducing artifactual biases in the source observations (Weiss et al., 2017). Feature normalization is an essential process of data preprocessing for applying machine learning and deep learning to multiple datasets (Ferreira, P., Le, D. C., & Zincir-Heywood, 2019). Data normalization could be effective in the data processing stage despite requiring a significant increase in processing and memory capability (Pires et al., 2020). In this study, we used StandardScaler to normalize the dataset. Standard Scaler provides standard data by eliminating the mean and scaling to unit variance (Towfek El-Kenawy, 2019).

Architecture of DPPM

Deep learning has increasingly acquired popularity as a consequence of its capacity to scale to big data, perform feature engineering from beginning to end, and provide accuracy in learning supervised and unsupervised data (Dargazany et al., 2018). Deep learning is a subset of machine learning that is based on artificial neural networks. Learning is possible in a supervised, semi-supervised, or unsupervised environment (Bagherzadeh & Asil, 2019). The DPPM consists of 13 layers as follows:

- ***Convolutional neural network (CNN)*** (6) layers.
- ***Max Pooling*** (5) layers.
- ***Flatten*** (1) layer
- ***Dense*** (1) layer.

Table (1) explains these layers in some detail.

To modify the data (reducing the error rate and adjusting the weights) from the associated errors, the Adam optimizer was used. Recently, the Adam optimization algorithm—an adaptation of the Stochastic Gradient Descent algorithm—has seen widespread use in deep learning applications, particularly in computer vision and natural language processing.

The learning rate used in the model was (0.001) and the learning rate affects the convergence of the improvement process as it balances the effect of the curvature of the cost function. When the learning rate is very small then the update will be small and the optimization is slow especially when the cost curve is low and the update is likely to settle at the local minimum. And when the learning rate is very large then the upgrading will be large and the improvement varies especially when the curvature of the cost function is high. If the learning rate is chosen well, the updates will be appropriate and the optimization will converge to a good set of parameters.

The activation function of deep neural networks has a significant influence on the training procedure's performance (Hayou et al., 2019). The kernel allows natural methods to manage learning capacity and reduce overfitting (Mairal, 2016). The kernel size and number of filters have a substantial influence on network accuracy (Agrawal & Mittal, 2020). When a large amount of information and a more complicated kernel are employed, the strides influence how a convolution process works with a kernel (Mureşan & Oltean, 2018).

The validity of the model is evaluated by the value of the loss functions which play an important role in statistical models. Loss functions specify a target by which the performance of the model is evaluated and the parameters used by the model are determined by reducing the chosen loss function. The loss function denotes the degree of disagreement between the model's predicted and true values (Zhong & Zhao, 2020).

Due to the unstable values in the data, which is thought to be noisy data, the proposed model is built with a low learning rate to prevent the system from learning much and a large number of kernel revolutions to increase the probability of evaluation and thereby increase the reliability of the results.

Evaluation Metrics

The system has been evaluated using two metrics. These metrics are explained in this section as follows:

Matching Rate

It is defined as the probability that any value in that class has been properly predicted. Eq. (1) yields the following result:

$$Matching\ Rate = \frac{Number\ of\ tests\ is\ correct}{Total\ number\ of\ tests} \tag{1}$$

Loss Function

This value is computed as shown in Eq. (2) as follows:

$$Loss = \frac{Number\ of\ tests\ is\ wrong}{Total\ number\ of\ tests} \tag{2}$$

RESULTS AND DISCUSSION

To assess the validity of the proposed model (Deep Pollutant Prediction Model), data from the GJ001 station in Allahabad City, as well as simulated data from Laylan City retrieved from the Gaussian Plume Model, were utilized. The loss function index was used in this work to evaluate the performance of the DPPM.

GJ001 dataset results

Table 2. The loss error of DPPM for the GJ001 dataset

Parameters	Loss Function testing 30%	Matching Percentage
$PM_{2.5}$	0.668	99.332
PM_{10}	1.141	98.859
NO	1.140	98.86
NO_2	0.575	99.425
NO_X	0.470	99.53
CO	0.220	99.78
SO_2	0.548	99.452
O_3	0.388	99.612
Benzene	0.051	99.949
Toluene	0.275	99.725
Xylene	0.038	99.962
AQI	4.466	95.534

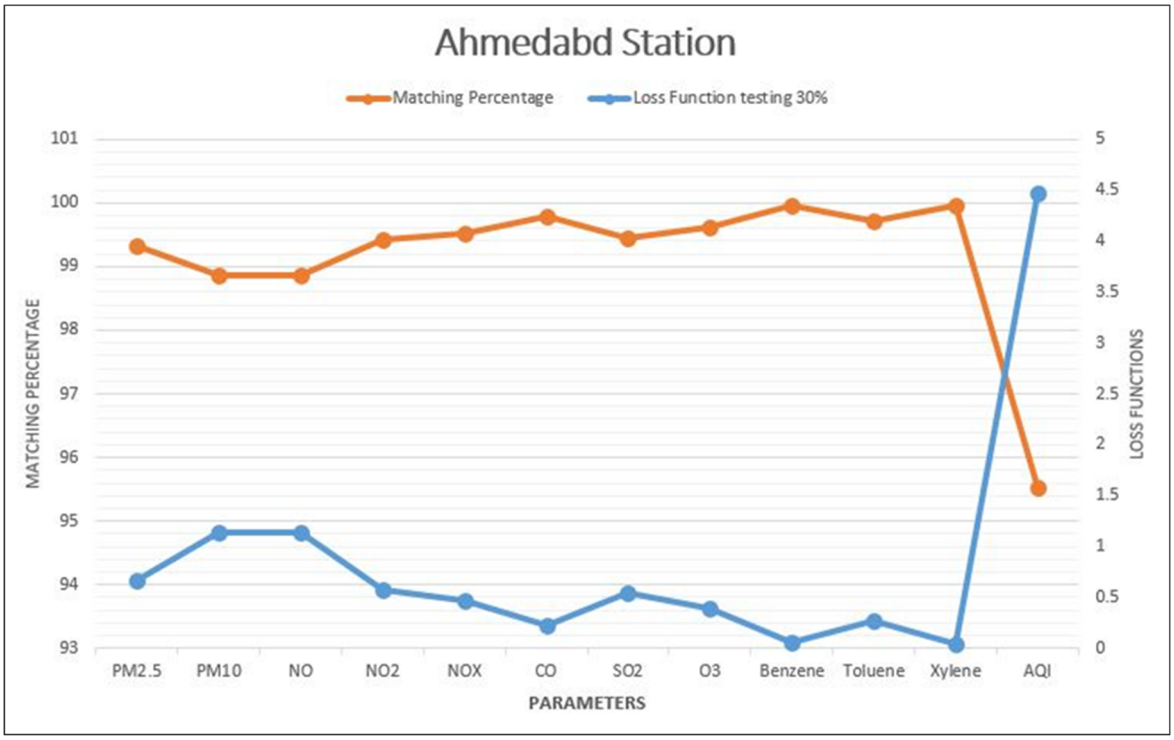

Fig. 3. Loss functions and matching percentage chart of GJ001 parameters

Despite the reality that the data for Allahabad are noisy and unstable DPPM demonstrated high efficiency in the prediction process for xylene and benzene parameters, with loss functions of 0.038 and 0.051, respectively. The $PM_{2.5}$, NO_2, NO_X, CO, SO_2, O_3, and Toluene parameters have excellent DPPM efficiency, with loss functions less than 1%, as shown in table 2. The loss functions for the PM_{10} and NO parameters are quite similar and indicate very good outcomes, with 1.41 and 1.4, respectively. The one parameter with a massive loss function relative to the others is the AQI, which is evident by the fact that it is mathematically dependent on the sum of several other parameters, including O_3, particulate matter ($PM_{2.5}$ and BM_{10}), CO, SO_2, and NO_2.

Despite the inconstancy of the measured values in the station and the missing values procedures done on the data preprocessing, these findings assessed the effectiveness of DPPM.

Table (2) illustrates the loss function error of DPPM for each parameter included in the GJ001 dataset. Figure (3) illustrates the loss functions and matching percentage of parameters of GJ001 station.

Simulated output results

The Gaussian output (simulated data) results demonstrate the high efficiency of DPPM with very low function loss. Table (3) shows the output of the Gaussian equation with the loss functions and their matching percentage for each parameter. The findings are relatively constant with little or no significant vacillation, which contributes to the effectiveness of the model. The smallest loss function was 0.0041 for the elements Fe_2O_3 and MgO, while the loss function close to them, which is 0.0042 for Al_2O_3 and 0.0045 for SO_3, as well as the loss function for SiO_2 was 0.0073. While the largest values of the loss functions were 0.0174 and 0.0221 for $CaCO_3$ and CaO respectively, they express the efficiency of DPPM due to its smallness values. Figure (4) states the loss functions and matching percentage of parameters of Laylan station.

The uncertainty level of the DPPM predictions represents by loss function which in general

Table 3. The loss error of DPPM for Laylan District from (Simulated output)

Parameters	Loss Function testing 30%	Matching Percentage
$CaCO_3$	0.0174	99.9826
SiO_2	0.0073	99.9927
Al_2O_3	0.0042	99.9958
Fe_2O_3	0.0041	99.9959
CaO	0.0221	99.9779
MgO	0.0041	99.9959
SO_3	0.0045	99.9955

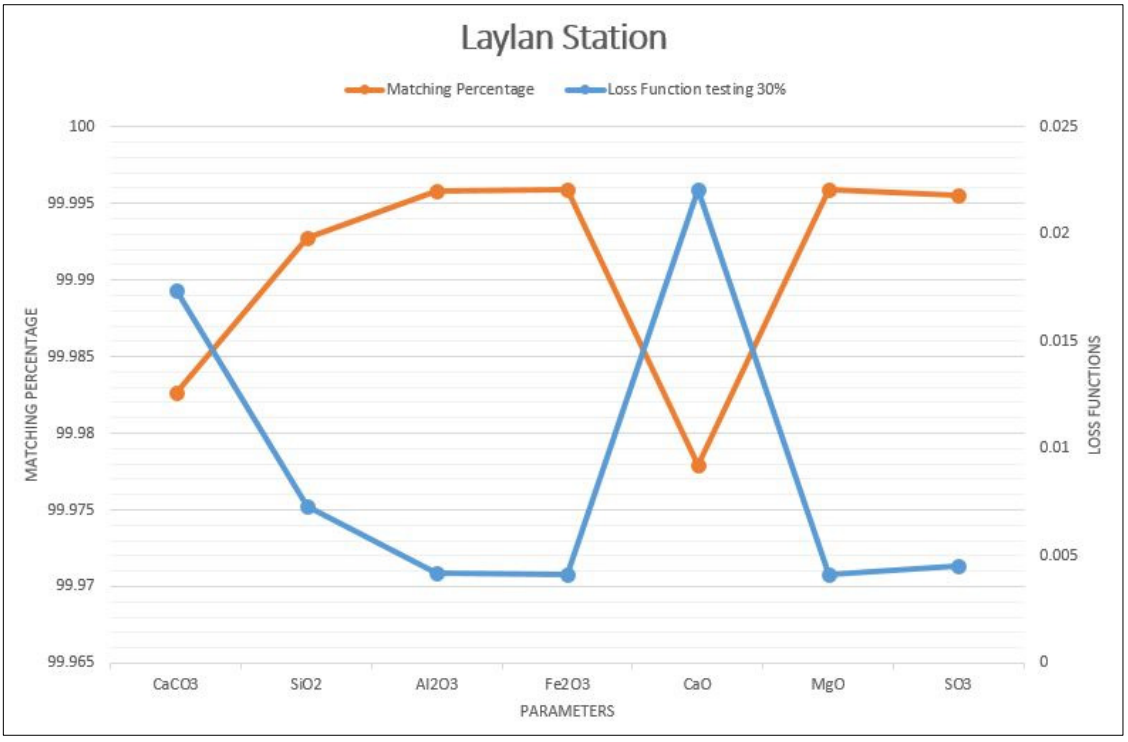

Fig. 4. Loss functions and matching percentage chart of Laylan parameters

was less than 1% for simulated data of Gaussian Plume Model. Whereas at the Ahmedabad station, many of the contaminated elements had loss functions of less than 1%, with the exception of two elements, NO and PM_{10}, which were slightly more than 1%, and the last element, whose loss function value was more than 4%, which remains because it is calculated from the total other contaminated elements. Based on these findings, we infer that the suggested DPPM model is quite efficient.

CONCLUSION

In this study, we developed a proposed model based on deep learning and named it as Deep Pollutant Prediction Model. This model has an efficient and high precision structure and it is containing from six layers of CNN with filter sizes of 16, 32, 64, 128, 256, and 512 with kernel size equal to (3). The strides and padding of CNN layers are the same and equal to (1). The activation function of CNN is the ReLU function and the input shape is equal to the number of items in the data. Furthermore, it has five Maxpooling layers, one Flatten layer, and one Dense layer.

Despite the reality that the data for Allahabad are noisy and unstable which is effect on model, DPPM demonstrated high efficiency in the prediction process for xylene and benzene parameters, with loss functions of 0.038 and 0.051, respectively. The $PM_{2.5}$, NO_2, NO_X, CO, SO_2, O_3, and Toluene parameters have excellent DPPM efficiency, with loss functions less than 1%. The loss functions for the PM_{10} and NO parameters are quite similar and indicate very good outcomes, with 1.41 and 1.4, respectively.

The one parameter with a massive loss function relative to the others is the AQI, which is evident by the fact that it is mathematically dependent on the sum of several other parameters, including O_3, particulate matter ($PM_{2.5}$ and PM_{10}), CO, SO_2, and NO_2. Despite the inconstancy of the measured values in the station and the missing values procedures done on the data preprocessing, these findings assessed the effectiveness of DPPM.

Seven parameters were used in the implementation of DPPM for the Laylan station which are $CaCO_3$, SiO_2, Al_2O_3, Fe_2O_3, CaO, MgO, and SO_3. These parameters are the simulated data of Gaussian Plume Model outputs. The loss functions for all the parameters in the DPPM findings for the Laylan station were better than those for the parameters for the Allahabad station. Where elements Fe_2O_3 and MgO had the lowest loss function value of 0.0041, and element CaO had the greatest loss function value of 0.0221.

GRANT SUPPORT DETAILS

The present research did not receive any financial support.

CONFLICT OF INTEREST

The authors declare that there is not any conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/ or falsification, double publication and/or submission, and redundancy has been completely observed by the authors.

LIFE SCIENCE REPORTING

No life science threat was practiced in this research.

REFERENCES

Agrawal, A., & Mittal, N. (2020). Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy. *Visual Computer*, *36*(2), 405–412. https://doi.org/10.1007/s00371-019-01630-9

Bagherzadeh, J., & Asil, H. (2019). A review of various semi-supervised learning models with a deep learning and memory approach. *Iran Journal of Computer Science*, *2*(2), 65–80. https://doi.org/10.1007/s42044-018-00027-6

Bekkar, A., Hssina, B., Douzi, S., & Douzi, K. (2021). Air-pollution prediction in smart city, deep learning approach. *Journal of Big Data*, *8*(1), 1–21. https://doi.org/10.1186/s40537-021-00548-1

Castiñeira, D., Schlosser, K. R., Geva, A., Rahmani, A. R., Fiore, G., Walsh, B. K., Smallwood, C. D., Arnold, J. H., & Santillana, M. (2020). Adding continuous vital sign information to static clinical data improves the prediction of length of stay after intubation: A data-driven machine learning approach. *Respiratory Care*, *65*(9), 1367–1377. https://doi.org/10.4187/respcare.07561

Dargazany, A. R., Stegagno, P., & Mankodiya, K. (2018). WearableDL: Wearable Internet-of-Things and Deep Learning for Big Data Analytics - Concept, Literature, and Future. *Mobile Information Systems*, *2018*(Dl). https://doi.org/10.1155/2018/8125126

Davies, E. R. (2018). Face detection and recognition. In *Computer Vision* (Issue June). https://doi.org/10.1016/b978-0-12-809284-2.00021-6

Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., & Dehmer, M. (2020). An Introductory Review of Deep Learning for Prediction Models With Big Data. *Frontiers in Artificial Intelligence*, *3*(February), 1–23. https://doi.org/10.3389/frai.2020.00004

Ferreira, P., Le, D. C., & Zincir-Heywood, N. (2019). Exploring Feature Normalization and Temporal Information for Machine Learning Based Insider Threat Detection. *2019 15th International Conference on Network and Service Management (CNSM)*, (pp. 1-7). IEEE.

Hayou, S., Doucet, A., & Rousseau, J. (2019). On the impact of the activation function on deep neural networks training. *36th International Conference on Machine Learning, ICML 2019*, *2019-June*, 4746–4754.

Heydari, A., Majidi Nezhad, M., Astiaso Garcia, D., Keynia, F., & De Santoli, L. (2022). Air pollution forecasting application based on deep learning model and optimization algorithm. *Clean Technologies and Environmental Policy*, *24*(2), 607–621. https://doi.org/10.1007/s10098-021-02080-5

Isam Drewil, G., & Jabbar Al-Bahadili, R. (2021). Multicultural Education Forecast Air Pollution in Smart City Using Deep Learning Techniques: A Review. *Multicultural Education*, *7*(5), 38–47. https://doi.org/10.5281/zenodo.4737746

Khan, Z. Y., Niu, Z., Sandiwarno, S., & Prince, R. (2021). Deep learning techniques for rating prediction: a survey of the state-of-the-art. In *Artificial Intelligence Review* (Vol. 54, Issue 1). Springer Netherlands. https://doi.org/10.1007/s10462-020-09892-9

Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., & Chi, T. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environmental Pollution*, *231*, 997–1004. https://doi.org/10.1016/j.envpol.2017.08.114

Ma, M., Dong, Z., Jang, C., Zhu, Y., & Zheng, H. (2020). Deep learning for prediction of the air quality response to emission changes. *Environmental Science & Technology*, *54*(14), 8589–8600. https://doi.org/10.1021/acs.est.0c02923.Deep

Mairal, J. (2016). End-to-end kernel learning with supervised convolutional kernel networks. *Advances in Neural Information Processing Systems*, *Nips*, 1407–1415.

Mao, Y., & Lee, S. (2019). Deep Convolutional Neural Network for Air Quality Prediction. *Journal of Physics: Conference Series*, *1302*(3), 032046. https://doi.org/10.1088/1742-6596/1302/3/032046

Mohsen, H., El-Dahshan, E.-S. A., El-Horbaty, E.-S. M., & Salem, A.-B. M. (2018). Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, *3*(1), 68–71. https://doi.org/10.1016/j.fcij.2017.12.001

Mureşan, H., & Oltean, M. (2018). Fruit recognition from images using deep learning. *Acta Universitatis Sapientiae, Informatica*, *10*(1), 26–42. https://doi.org/10.2478/ausi-2018-0002

Muthukumar, P., Cocom, E., Nagrecha, K., Comer, D., Burga, I., Taub, J., Calvert, C. F., Holm, J., & Pourhomayoun, M. (2022). Predicting PM2.5 atmospheric air pollution using deep learning with meteorological data and ground-based observations and remote-sensing satellite big data. *Air Quality, Atmosphere and Health*, *15*(7), 1221–1234. https://doi.org/10.1007/s11869-021-01126-3

Oprea, M., Popescu, M., Mihalache, S. F., & Dragomir, E. G. (2017). Data mining and ANFIS application to particulate matter air pollutant prediction. A comparative study. *ICAART 2017 - Proceedings of the 9th International Conference on Agents and Artificial Intelligence*, *2*(Icaart), 551–558. https://doi.org/10.5220/0006196405510558

Pires, I. M., Hussain, F., Garcia, N. M., Lameski, P., & Zdravevski, E. (2020). Homogeneous data normalization and deep learning: A case study in human activity classification. *Future Internet*, *12*(11), 1–14. https://doi.org/10.3390/fi12110194

Schürholz, D., Kubler, S., & Zaslavsky, A. (2020). Artificial intelligence-enabled context-aware air quality prediction for smart cities. *Journal of Cleaner Production*, *271*, 121941. https://doi.org/10.1016/j.jclepro.2020.121941

Towfek El-Kenawy, E.-S. M. (2019). A Machine Learning Model for Hemoglobin Estimation and Anemia Classification. *International Journal of Computer Science and Information Security (IJCSIS)*, *17*(2). https://sites.google.com/site/ijcsis/

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E. R., & Knight, R. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome*, *5*(1), 1–18. https://doi.org/10.1186/s40168-017-0237-y

Wu, W., May, R. J., Maier, H. R., & Dandy, G. C. (2013). A benchmarking approach for comparing data splitting methods for modeling water resources parameters using artificial neural networks. *Water Resources Research*, *49*(11), 7598–7614. https://doi.org/10.1002/2012WR012713

Zhang, J., & Ding, W. (2017). Prediction of air pollutants concentration based on an extreme learning machine: The case of Hong Kong. *International Journal of Environmental Research and Public Health*, *14*(2), 1–19. https://doi.org/10.3390/ijerph14020114

Zhang, Q., Lam, J. C., Li, V. O., & Han, Y. (2020). Deep-AIR: A Hybrid CNN-LSTM Framework forFine-Grained Air Pollution Forecast. *ArXiv Preprint ArXiv:2001.11957*, 1–7. http://arxiv.org/abs/2001.11957

Zhong, Y., & Zhao, M. (2020). Research on deep learning in apple leaf disease recognition. *Computers and Electronics in Agriculture*, *168*(August 2019), 105146. https://doi.org/10.1016/j.compag.2019.105146

Zhu, D., Cai, C., Yang, T., & Zhou, X. (2018). A machine learning approach for air quality prediction: Model regularization and optimization. *Big Data and Cognitive Computing*, *2*(1), 1–15. https://doi.org/10.3390/bdcc2010005