Journal of Applied and Natural Science
16(2), 704 - 712 (2024)
ISSN : 0974-9411 (Print), 2231-5209 (Online)
journals.ansfoundation.org

Research Article

$PM_{2.5}$ concentration prediction using Generative adversarial network: A novel approach

Shrabani Medhi*
Department of Computer Science and Engineering, Girijananda Chowdhury Institute of Management and Technology, Assam Science and Technology University, (Assam), India
Minakshi Gogoi
Department of Computer Science and Engineering, Girijananda Chowdhury University, (Assam), India

*Corresponding author. E-mail: shrabani26@gmail.com

Article Info
https://doi.org/10.31018/jans.v16i2.5510
Received: February 25, 2024
Revised: May 07, 2024
Accepted: May 12, 2024

How to Cite
Medhi, S. & Gogoi, M. (2024). $PM_{2.5}$ concentration prediction using Generative adversarial network: A novel approach. *Journal of Applied and Natural Science*, 16(2), 704 - 712. https://doi.org/10.31018/jans.v16i2.5510

Abstract
Over the last few years, air pollution has become a matter of great concern. Numerous machine learning and deep learning techniques have been applied to predict $PM_{2.5}$ (Particulate Matter$_{2.5}$). However, deterministic models perform forecasting based on the mean of probable outputs and cannot handle the uncertainties in real-life situations. With the aim of solving the low accuracy of $PM_{2.5}$ concentration prediction during uncertainties, the present study proposed an innovative probabilistic model-Prob $PM_{2.5}$ which predicts one day ahead $PM_{2.5}$ concentration for time series data, which is multivariate in nature. First, a comprehensive correlation analysis between the meteorological features and $PM_{2.5}$ concentration is done. Finally, the Conditional GAN framework is used to train the ProbPM$_{2.5}$ with the help of adversarial training. The proposed framework that transformed a deterministic model into a probabilistic model provided improved performance. Comparative analysis with conventional models, such as LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) reveals that ProbPM$_{2.5}$ outperforms during testing, showcasing resilience in the face of unforeseen events like COVID-19. Hence, the proposed method could perform improved characterization of time series characteristics of the air pollutant changes in order to obtain better accuracy of $PM_{2.5}$ concentration prediction

Keywords: Air pollution monitoring, Gated Recurrent Unit, Generative adversarial network, Long-short term memory, Multivariate time series data, $PM_{2.5}$ prediction

INTRODUCTION

Air pollution has increased tremendously in the last few decades because of increased growth, urbanization, and improved lifestyles in cities. In most Indian cities, air pollution is rising day by day due to rapid urbanization (Selokar *et al.*, 2020). Thus, addressing this issue is vital (Bhadauria *et al.*, 2023; Khanna *et al.*, 2013). $PM_{2.5}$ (Particulate Matter $_{2.5}$) is the most dangerous (Evans *et al.*, 2013) pollutant among all. They are suspended particles with a diameter less than or equal to 2.5 microns. It has a trivial diameter, large surface area and strong activity. It can easily absorb various kinds of toxic and harmful substances. It stays in the atmosphere long and has a large diffusion rate. That is why it greatly affects human health and the air quality. Literature advocates that for every 10 mg/m^3 increase in $PM_{2.5}$, it can upsurge the cardiovascular disease rate by 12 ~ 14%. $PM_{2.5}$ contains numerous organic compounds, such as hydrocarbons and formaldehyde and inorganic compounds, such as SO4^2- (Sulfates), NO3^- (Nitrates), etc. (Feng *et al.*, 2016; O'Donnell *et al.*, 2011). That is why it is essential to understand the temporal and spatial evolution and forecasting of $PM_{2.5}$ is very important so that effective measures can be taken by the concerned authorities to deal with the problem(Medhi and Gogoi, 2021).

With the improvement in lifestyle and increased environmental protection awareness, real-time monitoring of $PM_{2.5}$ can no longer satisfy people. The prediction of $PM_{2.5}$ for the future is of great concern, so that effective measures can be taken beforehand. So, it is important to monitor $PM_{2.5}$ concentration and perform predictions based on historical data. Past literature uses classic algorithms such as ARIMA (Autoregressive Integrated Moving Average) (Abhilash *et al.*, 2018), machine learning algorithms, LSTM (Long Short-Term Memory), etc., for predicting time series data. An overview of forecasting in time series data is revealed in paper (Mahalakshmi *et al.*, 2016). Numerous research studies have utilized point prediction methods, which offer simplicity and ease of understanding. Real life is full of uncertainties like COVID-19 pandemic, which cannot be reflected by these deterministic models (Dutta *et al.*, 2023). Probabilistic models have been designed to handle these issues. They quantify the uncertainties of the predictions by forming probability distributions over possible outcomes(Gneiting and Katzfuss, 2014). In the current scenario, GAN (Generative Adversarial Network) is among the most powerful models for performing predictions. The model uses a generator and discriminator, which is adversarial in nature and helps increase the model's accuracy. GAN has been extensively used in image generation, but not much is done concerning time series data.

Timely and accurate $PM_{2.5}$ concentration prediction will help the government manage major air pollution in an emergency and provide a scientific basis to take measures and decisions for production, emission and traffic restrictions. The government can also formulate prevention and control measures by analyzing the changing trend of $PM_{2.5}$ concentration based on the prediction information.

The existing prediction models based on machine learning predict $PM_{2.5}$ concentration based on historical data of the target prediction site. It cannot fully consider the spatial relationship between the prediction site and the surrounding monitoring sites. Further, the model fails to produce accurate predictions in case of uncertainties. To grasp the intricate probability distribution within air pollution data, which inherently involves uncertainties, GANs emerge as an immensely potent tool. However, mastering them presents a formidable challenge. The selection of model architecture and hyperparameters requires meticulous attention, given the volatility of the training process. (Bai *et al.*, 2021; Goodfellow *et al.*, 2016). The present study aimed to introduce an innovative probabilistic predictive model, $ProbPM_{2.5}$ to predict $PM_{2.5}$ concentration using Conditional Generative Adversarial Network.

MATERIALS AND METHODS

This experiment used a framework set up using Keras and Tensorflow(Hany and Walters, 2019). Experiments were performed using LSTM, GRU (Saif-ul-Allah *et al.*, 2022) and the proposed model and analysis were based on the results (Hochreiter and Schmidhuber, 1997; Staudemeyer and Morris, 2019). In the LSTM model(Sun and Li, 2020), Bidirectional LSTM (Kim *et al.*, 2023) was utilized in the 1st layer. Adam algorithm with a learning rate (0.001) was used as an optimizer. Batch size of 64 and 50 epochs was used during training. In the GRU model, two layers of GRU were used. Adam algorithm with a learning rate 0.0001 was used as the optimizer, and batch size of 128 and 50 epochs was used for training.

Dataset

This study used the continuous Ambient air quality monitoring station (CAAQMS) data for Guwahati. Guwahati has been chosen as the focus area of the study because of its recognition as one of the cities having the highest recorded levels of Black Carbon globally (Barman and Gokhale, 2019). The concentration of particulate matter in Guwahati exceeds permissible thresholds significantly, posing a severe threat to the health of both adults and children(Amnuaylojaroen & Parasin, 2023; Oliveira *et al.*, 2016). The Pollution Control Board of Assam (PCBA), headquartered in Bamunimaidan, Guwahati, is accountable for monitoring the city's ambient air quality. Since 2008, the PCBA has consistently reported $PM_{2.5}$ concentrations well above the recommended limits (Kioumourtzoglou *et al.*, 2016).

The first data set contained CAAQMS data of Guwahati city from January 2019 to December 2022 (3 years). The data count was 33067, which contained hourly data. Some data was lost due to missing values. The second data set contained CAAQMS data of Delhi city from January 2016 to December 2022 (4 years). The data count was 66090. Both the datasets were collected from the Central Pollution Control Board, India (*CCR*, n.d.) (https://airquality.cpcb.gov.in/ccr/#/caaqm-dashboard-all/caaqm-landing/data).

The parameters used in the study are given in Table. 1. A descriptive statistic of the available meteorological conditions, criteria gases and particulates measures: count, mean, standard deviation, minimum, 25%, 50%, 75%, maximum, skewness, kurtosis and variance for Guwahati and Delhi city were calculated. There was no high skewness value in the datasets. It showed that no

Table 1. Summary of measurement site and observed variables.

Measurement Site	Type	Variables
Guwahati City and Delhi City	Meteorological conditions	Relative humidity, Wind speed, Wind direction, Temperature, Rainfall, Pressure
	Criteria gases	NO_2, SO_2, NO, NOx, NH_3, CO, Ozone, Benzene, Eth-Benzene, MP-Xylene
	Particulates	PM2.5, PM10

sharpness was observed with the increase in the data. The high value of kurtosis in $PM_{2.5}$ indicates the presence of data discontinuities. The aim is to predict 1 day ahead $PM_{2.5}$ concentration for classification and regression. 75% of the dataset was used for training, 20% for testing and 5% for validation.

Principle

GAN architecture was used to predict $PM_{2.5}$ concentration. Conditional GAN was used to train a probabilistic $PM_{2.5}$ forecasting model with the help of adversarial training (Bai *et al.*, 2021). ProbPM2.5 was utilized as the generator. The gradient required for optimizing $ProbPM_{2.5}$ during training was provided by discriminator. The Conditional GAN utilized historical data {Xt, …, Xo} as its condition to forecast P(Xt+1 | Xt,…, Xo).

The value function for training $ProbPM_{2.5}$ is given by

Framework for conversion of deterministic model to probabilistic model

Addressing the numerous challenges associated with multivariate time series data is imperative. A more intricate architecture is necessary to predict future values and handle feature dependencies in multivariate time series data. As discussed earlier, ensuring a stable training process for GANs necessitates meticulous

$$\min\max V(Di, PB2.5) = Ext+1 \sim Pdata(xt+1)[\log[(Di(xt+1\,|xt,\dots,x0))] + Ez{\sim}Pz(z)[\log\left(1 - Di\left(PB2.5((z|xt,\dots,x0)\right)\right)] \quad \text{Eq. 1}$$

model architecture and selection of hyperparameters. Yet, finding an optimal architecture for the generator and discriminator, especially for multivariate time series data, can be exceptionally laborious or unattainable. To mitigate this challenge, a novel framework is introduced. This framework constructs a probabilistic $PM_{2.5}$ predictor by leveraging a deterministic $PM_{2.5}$ predictor, employing the architecture of GANs. The search for suitable generator and discriminator architectures was conducted independently to streamline the GAN architecture search process. The proposed framework and the adversarial training setup are illustrated in Fig. 1. Initially, an optimal architecture for the deterministic $PM_{2.5}$ forecasting model is sought. If a suitable deterministic model is already available, the first step can be bypassed, and the existing model can be utilized. For hyperparameter tuning, Bayesian Optimization used the Bayes Theorem to search for suitable parameters. Learning rate between 0.0001 and 0.0009 was used, epochs between 100 to 200 and batch size of 64 to 512 was used. The model was finally trained using Adam algorithm as an optimizer, 0.0001 as the learning rate, 100 epochs and a batch size of 128. Then, a noise vector Nz was integrated into the deterministic model. After conducting various experiments, it was observed that optimal outputs were obtained when a noise vector was inserted into the advanced layers of the network. In the initial stages, the network's earlier layers were employed to learn the depiction of the input window. Subsequently, the model underwent training using GAN to acquire the probabilistic model, $ProbPM_{2.5}$. To function as the discriminator, an appropriate $PM_{2.5}$ classifier must be sought during the GAN training process. The

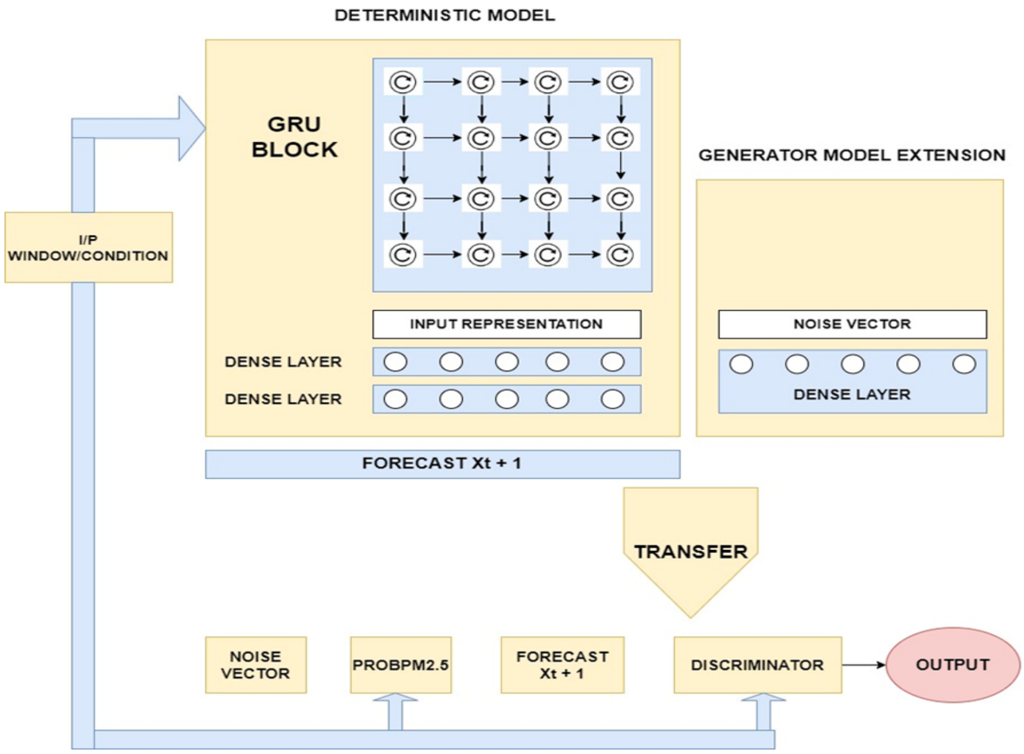


Fig. 1. *Proposed framework and adversarial training setup*

search space of the GAN architecture is reduced to the discriminator only. This helps to find a discriminator structure efficiently that can train $ProbPM_{2.5}$, delivering superior performance compared to the deterministic models alone. The framework works like the below:

Choose a suitable deterministic model for $PM_{2.5}$ prediction. This can involve either utilizing an existing model or searching for an ideal deterministic $PM_{2.5}$ predictor model.

Construct the generator using the architecture and hyperparameters derived from the selected deterministic $PM_{2.5}$ predictor. Embed a noise vector into the advanced layers of the network to facilitate the generation of synthetic samples.

Explore various discriminator architectures to find an optimal one, then train the $ProbPM_{2.5}$ model using the selected discriminator.

Model setup

An architecture search was run for each experiment to search for an optimal deterministic model. Mean Absolute Error (MAE) was utilized as a loss function to train the deterministic model. The input window representation was learnt using Gated Recurrent Unit (GRU). The depiction of the input window was then fed through two dense layers to map it to prediction of $PM_{2.5}$. $ProbPM_{2.5}$ was built by concatenating the noise vector to the output of the window representation of GRU and then the Multi-Layer Perceptron block was extended as shown in Fig. 2. In the final step, an optimal architecture for the discriminator was searched by using a genetic algorithm and the $ProbPM_{2.5}$ was trained. Concatenation of X_{t+1} to the end of the input window was done by the discriminator to construct $\{X_{t+1}, X_t, \dots X_0\}$. After that, a GRU block was employed, and its output was fed through two layers of MLP using PyTorch (Hany & Walters, 2019).

Data analysis

Following evaluation metrics (Steurer *et al.*, 2021) were utilized to evaluate the efficiency of the models.

Negative form of Continuous ranked probability score (CRPS*)

The negative form of CPRS* (Zamo & Naveau, 2018) (Berrisch & Ziel, 2023; Hersbach, 2000) was used to reflect the calibration and sharpness of a probabilistic model. It is given below

Here, Y and Y' represent independent copies of a random variable generated by the probabilistic predictor G, while y denotes the ground truth. It is vital to find a direct way to compare deterministic and probabilistic models. And it is done by CRPS*.

Mean absolute error (MAE)

In place of CRPS*, Mean Absolute Error (MAE) (Medhi *et al.*, n.d.)is used for the deterministic model. MAE is denoted as

Where x is the actual value and x' is the predicted value

$$CRPS*(Ge, y) = Eg\mid Y - y\mid - \left(\frac{1}{2}\right) Eg\mid Y - Y'\mid \quad \text{Eq. 2}$$

Root Mean Square Error (RMSE)

RMSE is used to analyze the performance of the models (Chai & Draxler, 2014, 2014). It is denoted by

Where yi is the actual $PM_{2.5}$ concentration value, yp is the predicted $PM_{2.5}$ concentration value, and n is the number of data points.

RESULTS AND DISCUSSION

Correlation analysis between $PM_{2.5}$ concentration and meteorological features

$$MAE(x, x^{\wedge\prime}) = E\mid x^{\wedge\prime} - x\mid \quad \text{Eq. 3}$$

The influence of meteorological features on $PM_{2.5}$ concentration is very important and it is also very complex (Y. Liu *et al.*, 2021; Wang *et al.*, 2019). If each feature is considered separately, it becomes difficult to reflect

$$RMSE = \sqrt{}((1)/n \textstyle\sum\_(i=1)^\wedge n(yi - yp)) \quad \text{Eq. 4}$$

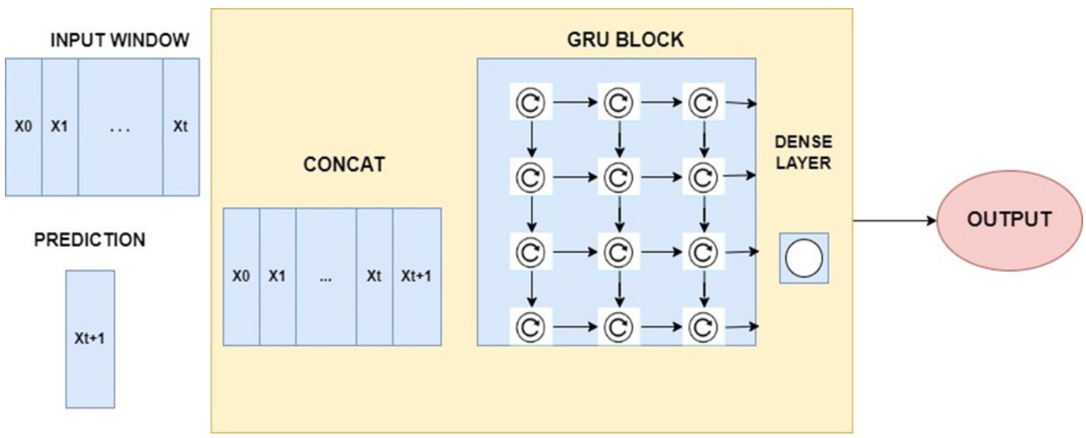


Fig. 2. *Discriminator architecture of conditional GAN*

the coupling effect of multiple features on $PM_{2.5}$ concentration well. The correlation between $PM_{2.5}$ concentration and meteorological features is depicted in Fig. 3. The analysis revealed positive correlations between $PM_{2.5}$ concentration and relative humidity, temperature, and pressure. Conversely, negative correlations were observed between $PM_{2.5}$ concentration, wind speed, wind direction, and rainfall.

$PM_{2.5}$ prediction using LSTM

The result of LSTM for the Delhi and Guwahati datasets are depicted in Fig. 4 (a) and (b). In both scenarios, all the predicted values were slightly lower than the actual $PM_{2.5}$ concentration. But, the prediction was slightly better for the Delhi dataset than the Guwahati dataset. During the training phase, the RMSE value was slightly higher for the Guwahati dataset than the Delhi dataset. During the testing phase also, the RMSE value was slightly higher for the Guwahati dataset than the Delhi dataset. For both datasets, the model performed better in the training phase than in the testing phase. LSTM model performed better with the Delhi dataset with RMSE value of 5.86, contrary to the Guwahati dataset, which had an RMSE value of 6.86. LSTM performed worst among all the models.

$PM_{2.5}$ prediction using GRU

The result of GRU model for the Delhi and Guwahati datasets is shown in Fig. 5 (a) and (b). Like LSTM, in GRU model also, all the predicted values were slightly lower than the actual $PM_{2.5}$ concentration in both cases (Tran *et al.*, 2023). Again, for the Delhi dataset the prediction was slightly better compared to Guwahati dataset. During both training and testing, the RMSE value was slightly higher for the Guwahati dataset than the Delhi dataset. For both datasets, the model performed better in the training phase than in the testing phase. GRU model performed better with Delhi dataset with RMSE value of 4.38, contrary to Guwahati dataset with an RMSE value of 5.38. Compared to LSTM model, GRU model predicted more accurately the $PM_{2.5}$ concentration. Among all the models, GRU performed best during the training phase. However, during the testing phase, it performed better than LSTM, but prediction

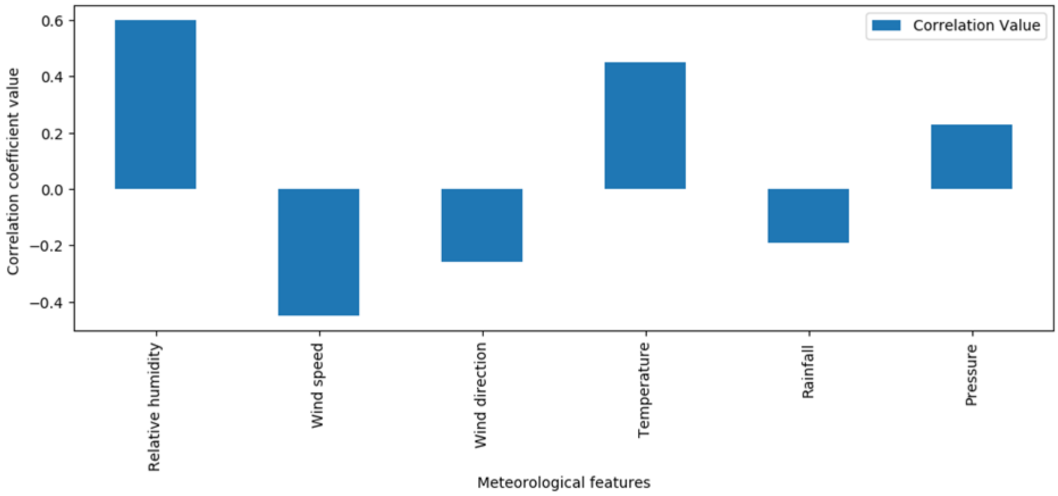
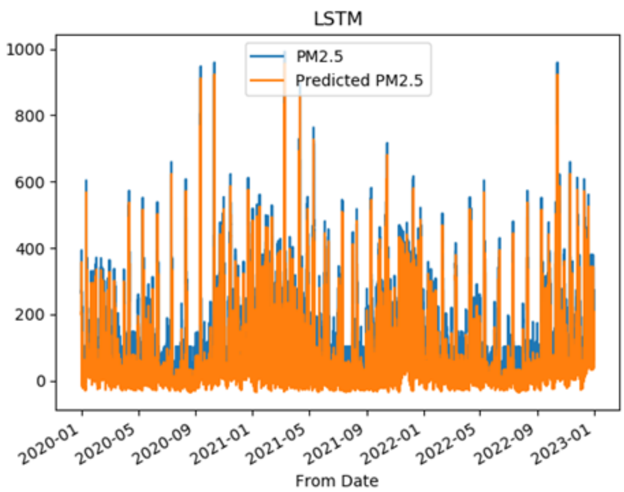


Fig. 3. *Correlation coefficients between $PM_{2.5}$ concentration and meteorological features*



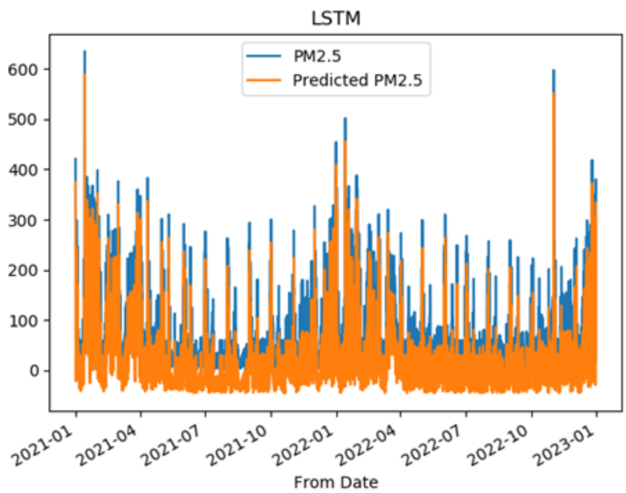


Fig. 4. *Test data plot using LSTM for (a) Delhi Dataset (b) Guwahati Dataset*

accuracy was less compared to ProbPM$_{2.5}$ model.

The deterministic model served as a framework for capturing a robust representation from the input time window. Trained to estimate the mean of potential predictions, these models inherently encoded a clear indication of the target distribution's location. Subsequently, leveraging these indicators, the Multi-Layer Perceptron block adeptly translated the noise vector z into a probabilistic distribution of forthcoming PM$_{2.5}$ concentration values. The performance of GRU and LSTM decreased due to the occurrence of the unexpected event COVID-19 when the concentration of PM$_{2.5}$ decreased suddenly due to fewer pollution-causing activities due to the lockdown. However, ProbPM$_{2.5}$ performed well despite the unexpected event.

PM$_{2.5}$ prediction using ProbPM$_{2.5}$

The loss plot of ProbPM$_{2.5}$ model for Guwahati dataset and Delhi dataset is given in Fig 6(a) and (b), respectively. For Guwahati dataset, D_loss is the loss path of the discriminator and G_loss is the loss path of the generator. Over time, it was noticed that the discriminator's loss initially surpassed that of the generator; however, gradually, the discriminator's loss approached zero. This trend was consistent even when examining the Delhi dataset, where the discriminator's loss remained higher than that of the generator. But it converged to zero faster as compared to the Guwahati dataset. The proposed framework underwent experimentation on the datasets to identify optimal hyperparameters, as outlined in Table 2. In Table 3, experiment's results are summarized, presenting CRPS* of the optimal deterministic model and ProbPM$_{2.5}$ model for the two datasets. To calculate CRPS* for ProbPM$_{2.5}$, 200 times sampling was done. In Table 4, a comparison was made between the RMSE values during training and testing period for all three models. When the testing dataset was considered, it was found that ProbPM$_{2.5}$ performed the best. LSTM and GRU's performance decreased due to the unexpected event COVID-19 (Biswas and Pathak, 2022; Mangayarkarasi *et al.*, 2021). The proposed model exhibited superior performance compared to the traditional baseline models.

Based on the conducted experiments, it was noted that ProbPM$_{2.5}$ surpassed the deterministic model in performance on the Delhi dataset despite having a nearly alike structure. Despite the dataset containing a sub-

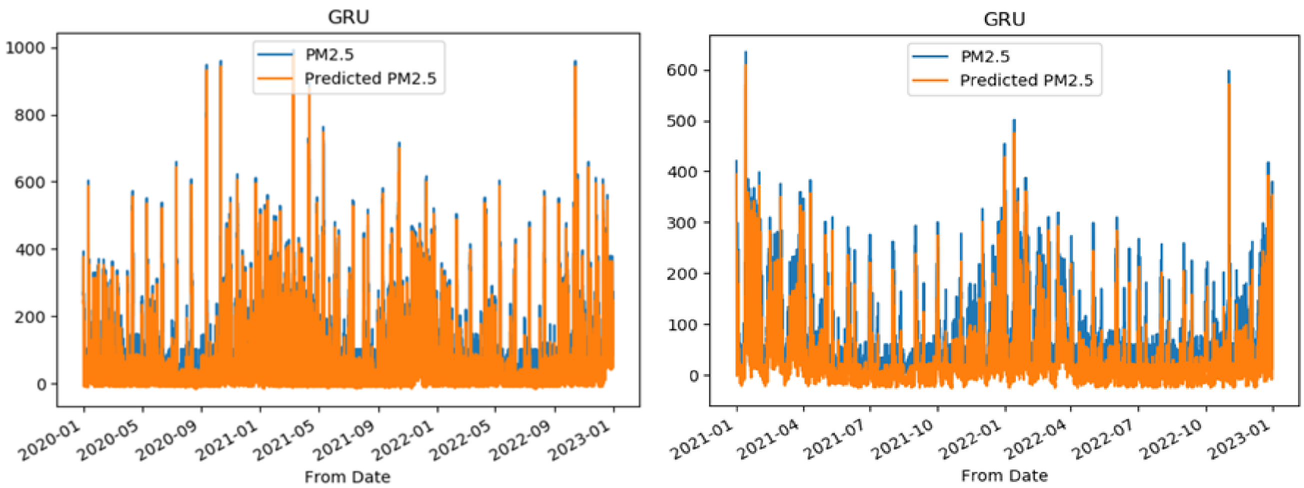


Fig. 5. *Test data plot using GRU for (a) Delhi Dataset (b) Guwahati Dataset*

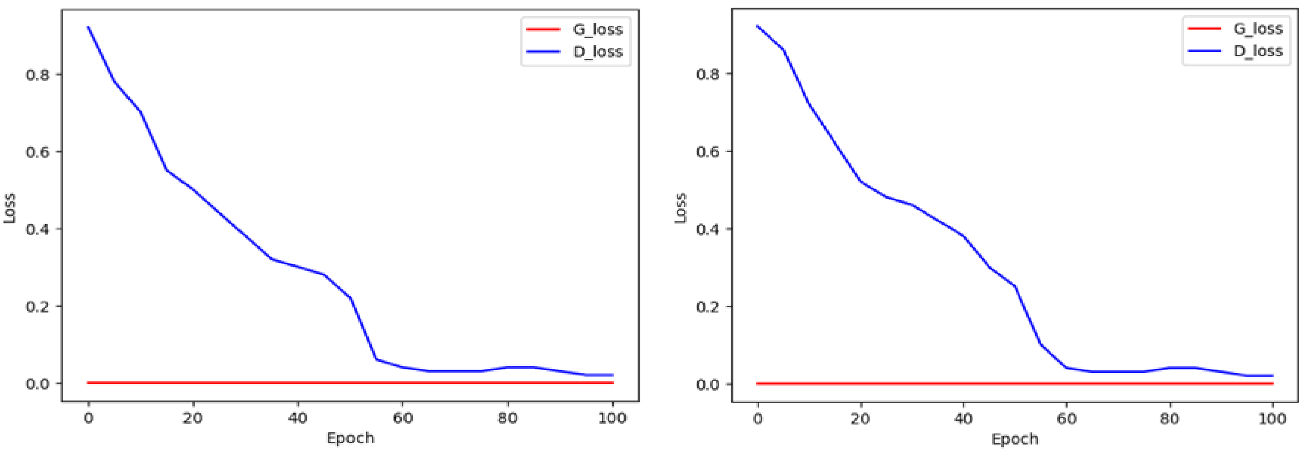


Fig. 6. *Proposed model loss plot for (a) Delhi Dataset and (b) Guwahati Dataset*

Table 2. Hyperparameter optimal values for the generator and discriminator for each experiment

	Delhi Dataset	
	Generator Hyperparameters	Discriminator Hyperparameters
Input Window Size	180	--
Noise Size	320	--
Number of GRU layers	1	3
Number of GRU cells in each layer	125	158
	Guwahati Dataset	
	Generator Hyperparameters	Discriminator Hyperparameters
Input Window Size	174	--
Noise Size	195	--
Number of GRU layers	1	1
Number of GRU cells in each layer	125	128

Table 3. CRPS* of deterministic model and ProbPM2.5

	Best Deterministic Model	**ProbPM2.5**
CRPS* for Delhi dataset	247.5	240.6
CRPS* for Guwahati dataset	1.07×10^{-2}	7.63×10^{-3}

Table 4. RMSE (train data and test data) of deterministic model and ProbPM2.5

	LSTM	**GRU**	**ProbPM2.5**
RMSE (Train data)	1.52	1.20	1.79
RMSE (Test data)	6.86	5.38	4.65

stantial number of features, ProbPM$_{2.5}$ demonstrated the capability to offer precise predictions for multivariate time series datasets (Muruganandam and Arumugam, 2023). In the Guwahati dataset, ProbPM$_{2.5}$ outperformed its parallel deterministic model in spite of having similar structural similarities. It was also observed that the proposed model worked well even though the Guwahati dataset size was smaller than the Delhi dataset. It was also noted that the proposed framework successfully transformed the deterministic model into a more accurate probabilistic model. GAN is sensitive to the architecture of its components (A. Liu *et al.*, 2019), but ProbPM$_{2.5}$ still worked well when it was defined by engaging the deterministic architecture model.

Conclusion

The paper presents three main contributions: i) an innovative probabilistic model, ProbPM$_{2.5}$ was introduced to predict the PM$_{2.5}$ concentration using multivariate time series data. Conditional GAN was used to set up training, ii) A framework was proposed that transformed deterministic PM$_{2.5}$ prediction model into probabilistic model, and iii) Experiments were conducted on two datasets. Results show that the probability model outperformed the deterministic models. The results indicated that ProbPM$_{2.5}$ effectively learnt intricate patterns from datasets, discerning dependencies among numerous attributes and accurately predicting PM$_{2.5}$ concentration. The experiments conducted on the proposed framework demonstrated a systematic approach for transforming deterministic models into probabilistic models, resulting in enhanced accuracy. The promising outputs of the experimentations suggest a great potential for probabilistic prediction utilizing GANs and further research can be done with this approach. Further complex architectures for discriminators and generators can be experimented. The proposed framework simplified the process of selecting model architecture and hyperparameters. Leveraging an efficient deterministic model as a foundation facilitated the creation of more effective probabilistic models. Experimental results consistently demonstrate the superiority of the proposed model over others, highlighting its efficacy and potential for practical applications.

ACKNOWLEDGEMENTS

The authors express their profound gratitude to the Department of Computer Science and Engineering, Girijananda Chowdhury University and Assam Science and Technology University, Assam, India, especially the Central Pollution Control Board, India, for support during data collection.

Conflict of Interest

The authors declare that they have no conflict of interest.

REFERENCES

1. Amnuaylojaroen, T. & Parasin, N. (2023). Future health risk assessment of exposure to $PM_{2.5}$ in different age groups of children in Northern Thailand. *Toxics*, *11*(3), 291.
2. Bai, T., Luo, J., Zhao, J., Wen, B. & Wang, Q. (2021). Recent advances in adversarial training for adversarial robustness. *arXiv Preprint arXiv:2102.01356*.
3. Barman, N. & Gokhale, S. (2019). Urban black carbon-source apportionment, emissions and long-range transport over the Brahmaputra River Valley. *Science of the Total Environment*, *693*, 133577.
4. Berrisch, J. & Ziel, F. (2023). CRPS Learning. *Journal of Econometrics*, *237*(2), 105221. https://doi.org/10.1016/j.jeconom.2021.11.008
5. Bhadauria, N., Chauhan, A., Ranjan, R. & Jindal, T. (2023). An assessment of seasonal, monthly and diurnal variations of ambient air quality in the Gurugram city (Haryana). *Journal of Applied and Natural Science*, *15* (1), Article 1. https://doi.org/10.31018/jans.v15i1.4142
6. Biswas, J. & Pathak, B. (2022). Effects of COVID-19 pandemic lockdown: A satellite data-based appraisal of air quality in Guwahati, Assam. *Materials Today: Proceedings*, *65*, 2794–2800. https://doi.org/10.1016/j.matpr.2022.06.218
7. Chai, T. & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific Model Development Discussions*, *7*(1), 1525–1534.
8. CPCB (2024). Central Pollution Control Board (CPCB). (n.d.). Retrieved January 2, 2024, from https://cpcb.nic.in/
9. Dutta, A., Jinsart, W., Das, U. C. & Dutta, G. (2023). An empirical analysis of Delhi's air quality throughout different COVID-19 pandemic waves. *Journal of Applied and Natural Science*, *15*(1), Article 1. https://doi.org/10.31018/jans.v15i1.4271
10. Evans, J., van Donkelaar, A., Martin, R. V., Burnett, R., Rainham, D. G., Birkett, N. J. & Krewski, D. (2013). Estimates of global mortality attributable to particulate air pollution using satellite imagery. *Environmental Research*, *120*, 33–42.
11. Feng, S., Gao, D., Liao, F., Zhou, F. & Wang, X. (2016). The health effects of ambient PM2.5 and potential mechanisms. *Ecotoxicology and Environmental Safety*, *128*, 67–74. https://doi.org/10.1016/j.ecoenv.2016.01.030
12. Gneiting, T. & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application*, *1*, 125–151.
13. Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep learning*. MIT press.
14. Hany, J. & Walters, G. (2019). *Hands-On Generative Adversarial Networks with PyTorch 1. x: Implement next-generation neural networks to build powerful GAN models using Python*. Packt Publishing Ltd. https://books.google.com/books?hl=en&lr=&id=URbEDwAAQBAJ&oi=fnd&pg=PP1&dq=select+optimal+discriminator+in+gan+using+pytorch&ots=WgywmebEz3&sig=Ci7g85UESAxhJ0lpwO_kRTPyUxk
15. Hersbach, H. (2000). Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather and Forecasting - WEATHER FORECAST*, *15*, 559–570. https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2
16. Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.
17. Khanna, D. R., Nigam, N. S. & Bhutiani, R. (2013). Monitoring of ambient air quality in relation to traffic density in Bareilly City (U.P.), India. *Journal of Applied and Natural Science*, *5*(2), Article 2. https://doi.org/10.31018/jans.v5i2.359
18. Kim, B., Kim, E., Jung, S., Kim, M., Kim, J. & Kim, S. (2023). PM2.5 Concentration Forecasting Using Weighted Bi-LSTM and Random Forest Feature Importance-Based Feature Selection. *Atmosphere*, *14*, 968. https://doi.org/10.3390/atmos14060968
19. Kioumourtzoglou, M.-A., Schwartz, J. D., Weisskopf, M. G., Melly, S. J., Wang, Y., Dominici, F. & Zanobetti, A. (2016). Long-term PM2. 5 exposure and neurological hospital admissions in the northeastern United States. *Environmental Health Perspectives*, *124*(1), 23–29.
20. Liu, A., Liu, X., Fan, J., Ma, Y., Zhang, A., Xie, H. & Tao, D. (2019). Perceptual-sensitive gan for generating adversarial patches. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 1028–1035. https://ojs.aaai.org/index.php/AAAI/article/view/3893
21. Liu, Y., Shi, G., Zhan, Y., Zhou, L. & Yang, F. (2021). Characteristics of PM2. 5 spatial distribution and influencing meteorological conditions in Sichuan Basin, southwestern China. *Atmospheric Environment*, *253*, 118364.
22. Mahalakshmi, G., Sridevi, S. & Rajaram, S. (2016). A survey on forecasting of time series data. *2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16)*, 1–8.
23. Mangayarkarasi, R., Vanmathi, C., Khan, M. Z., Noorwali, A., Jain, R. & Agarwal, P. (2021). COVID19: Forecasting Air Quality Index and Particulate Matter (PM2. 5). *Computers, Materials & Continua*, *67*(3). https://research.vit.ac.in/publication/covid19-forecasting-air-quality-index-and-particulate-matter-pm25/pdf/publisher-pdf-fulltext-covid19-forecasting-air-quality-index-and-particulate-matter-pm25.pdf
24. Medhi, S. & Gogoi, M. (2021). Visualization and Analysis of COVID-19 Impact on PM2. 5 Concentration in Guwahati city. *2021 International Conference on Computational Performance Evaluation (ComPE)*, 012–016.
25. Muruganandam, N. S. & Arumugam, U. (2023). Dynamic Ensemble Multivariate Time Series Forecasting Model for PM2. 5. *Computer Systems Science & Engineering*, *44*(2). https://cdn.techscience.cn/ueditor/files/csse/TSP_CSSE-44-2/TSP_CSSE_24943/TSP_CSSE_24943.pdf
26. O'Donnell, M. J., Fang, J., Mittleman, M. A., Kapral, M. K. & Wellenius, G. A. (2011). Fine Particulate Air Pollution (PM2.5) and the Risk of Acute Ischemic Stroke. *Epidemiology (Cambridge, Mass.)*, *22*(3), 422–431. https://doi.org/10.1097/EDE.0b013e3182126580
27. Oliveira, M., Slezakova, K., Delerue-Matos, C., Pereira, M. C. & Morais, S. (2016). Assessment of air quality in preschool environments (3–5 years old children) with emphasis on elemental composition of PM10 and PM2. 5. *Environmental Pollution*, *214*, 430–439.
28. Saif-ul-Allah, M. W., Qyyum, M. A., Ul-Haq, N., Salman, C. A. & Ahmed, F. (2022). Gated recurrent unit coupled

with projection to model plane imputation for the PM2. 5 prediction for Guangzhou City, China. *Frontiers in Environmental Science*, *9*, 816616.

29. Selokar, A., Ramachandran, B., Elangovan, K. N. & Dattatreya Varma, B. (2020). PM 2.5 particulate matter and its effects in Delhi/NCR. *Materials Today: Proceedings*, *33*, 4566–4572. https://doi.org/10.1016/j.matpr.2020.08.187
30. Staudemeyer, R. C. & Morris, E. R. (2019). *Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks* (arXiv:1909.09586). arXiv. http://arxiv.org/abs/1909.09586
31. Steurer, M., Hill, R. J. & Pfeifer, N. (2021). Metrics for evaluating the performance of machine learning based automated valuation models. *Journal of Property Research*, *38*(2), 99–129. https://doi.org/ 10.1080/09599916.2020.1858937
32. Sun, W. & Li, Z. (2020). Hourly PM2.5 concentration forecasting based on mode decomposition-recombination technique and ensemble learning approach in severe haze episodes of China. *Journal of Cleaner Production*, *263*, 121442. https://doi.org/10.1016/j.jclepro.2020.121442
33. Tran, H. D., Huang, H.-Y., Yu, J.-Y. & Wang, S.-H. (2023). Forecasting hourly PM2.5 concentration with an optimized LSTM model. *Atmospheric Environment*, *315*, 120161. https://doi.org/10.1016/j.atmosenv.2023.120161
34. Wang, H., Li, J., Peng, Y., Zhang, M., Che, H. & Zhang, X. (2019). The impacts of the meteorology features on PM2. 5 levels during a severe haze episode in central-east China. *Atmospheric Environment*, *197*, 177–189.
35. Zamo, M. & Naveau, P. (2018). Estimation of the continuous ranked probability score with limited information and applications to ensemble weather forecasts. *Mathematical Geosciences*, *50*(2), 209–234.