

FORECASTING OF DAILY PM₁₀ USING LONG SHORT TERM MEMORY NEURAL NETWORK IN GANGA NAGAR, MEERUT INDIA FOR HEALTH AND AGRICULTURE APPLICATIONS

VIBHA YADAV¹ AND BISHAL KUMAR MISHRA²

School of Agricultural Sciences, IIMT University, Meerut 250 001, U.P., India

(Received 1 February, 2023; Accepted 28 April, 2023)

ABSTRACT

Air quality is known to significantly affect health, forecasting it is a very important task. A highly industrialized region in India, Meerut has one of the most extensive agricultural applications. Delhi, India's central pollution control board keeps time series data. In order to predict PM_{2.5} one day in advance, a Long Short Term Memory Network is used. The findings demonstrate that PM_{2.5} is predicted more accurately. This study is interesting because it can be used by government agencies, businesses, and citizens alike to make informed decisions.

KEY WORDS : PM_{2.5}, Forecasting, LSTM.

INTRODUCTION

The hazard of air pollution to human health and the environment is spreading around the globe. Rapid urbanization and exploding industrialization are factors in the emission of several air pollutants into the environment. Particulate matter, Carbon Oxides, Sulfur Oxides, Nitrogen Oxides, Ozone, and Hydrocarbons are a few of the several pollutants that affect the environment. Particulate matter (PM) is one of the main air pollutants produced by dust particles that come from a variety of sources, including the burning of coal and oil as well as mining operations. The size of the particle is significant when analyzing the behavior of a contaminant in the atmosphere. PM may be divided into PM₁₀ and PM_{2.5}. Less than 10 µm particles often float in the air whereas bigger than 10 µm particles tend to settle quickly. Because of the strong correlation between these PM and cardiovascular and pulmonary conditions, it is crucial to anticipate the pollutant concentration using time and location identification. Several tools used by various writers are provided.

In order to forecast daily values of CO, PM₁₀, NO, NO₂, NO_x, SO₂, H₂S, and O₃ various combinations of

input variables, used by Baawain *et al.* (2014) who constructed an ANN model (AP and meteorological variables). The regression (R) value for the training and test sets is greater than 70%. For Egypt, Elminir and Galil (2006) created an ANN model to forecast daily PM₁₀, NO₂, and CO. WS, RH, T, and WD are the ANN's inputs, while the output data are PM₁₀, NO₂, and CO. The accuracy of the model's predictions for PM₁₀, NO₂, and CO is 96%, 97%, and 99.6%, respectively. Using the suggested particle swarm optimization (PSO) and ANN model, Filho and Fernandes (2013) performed daily prediction of CO, SO₂, NO₂, and PM₁₀ levels. The statistics on pollutants are sourced from So Paulo, Brazil. The findings demonstrate that the PSO-ANN technique produces fair and reliable forecasts. MLP and RBF were employed by Skrzypski and Szakiel (2008) to estimate the daily PM₁₀ concentration. Daily rainfall totals, WS, SR, vertical wind speeds, T, and atmospheric pressure are the input variables. The prediction accuracy of MLP and RBF models are 3.7% and 1.9% respectively. Cellular neural network (CNN) was utilized by Sahin *et al.* (2011) to forecast the daily values of SO₂ and PM₁₀ for Istanbul, Turkey. The CNN algorithm is used to figure out missing air pollution data and can forecast SO₂

(^{1,2}Assistant Professor)

more accurately than PM10 can. The CNN model's R value ranges from 54% to 87%.

From the explanation above, it is clear that the effect of prediction of air pollutant is important so in this study PM_{2.5} is forecasted using LSTM network.

This paper is organized as follows: Methodology are shown in Section 2. Research and discussions are presented in Section 3 followed by Conclusions in Section 4.

METHODOLOGY

Data Collection

A real-world dataset from more than one year is used in this study. The daily value of PM_{2.5} for Ganga Nagar Meerut India is collected from Central

Pollution Control Board (CPCB) Delhi India as shown in Fig.1.

Long Short Term Memory Network

Long short-term memory networks, or LSTMs, are employed in deep learning. Many recurrent neural networks (RNNs), particularly those used for sequence prediction, are capable of learning long-term dependencies. On time series forecasting tasks, the effectiveness of LSTMs has been shown. Additionally, they have been used in a variety of applications. Long-term trends and a large number of dependencies are intended for LSTM responses. The core idea of an LSTM is to use gates to regulate information flow across the network, enabling it to

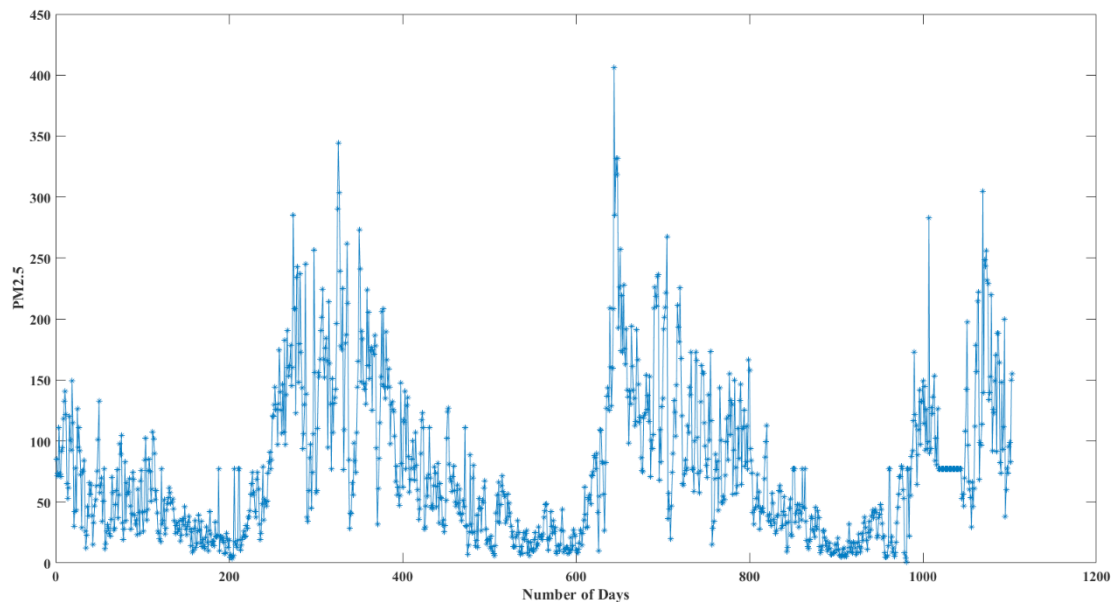


Fig. 1. Daily value of PM_{2.5}

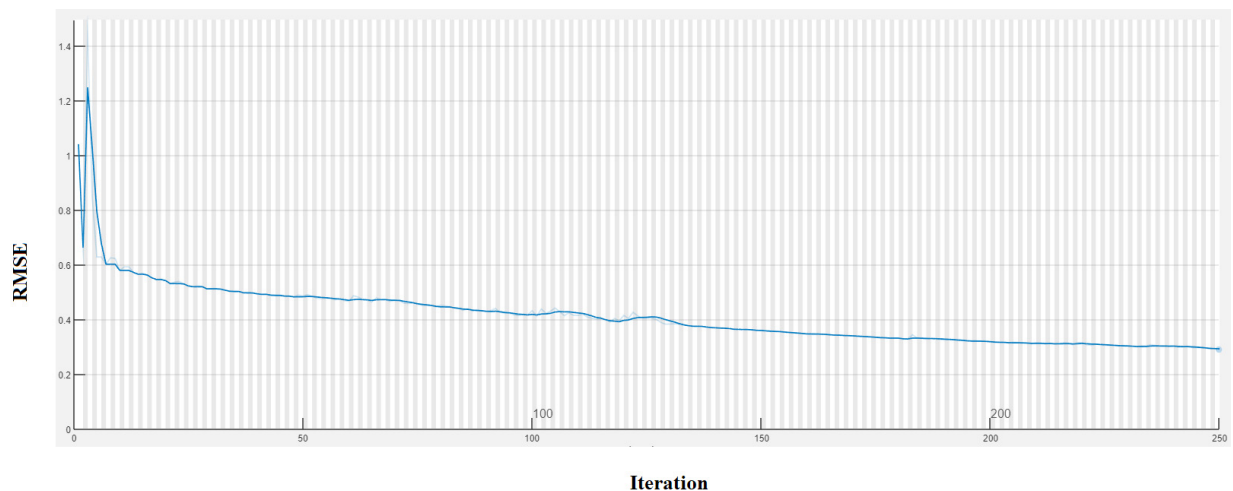


Fig. 2. Root Mean Square Error

recall specific information for longer periods of time. An LSTM model is trained using particular knowledge and data, which in this study includes historical information about air pollutants. The output at time step t is mapped to the input, hidden state, and memory cell state by the LSTM function. The hidden state $h(t-1)$ and memory cell state $c(t-1)$ from the previous time step are also used by the LSTM model along with the input $x(t)$ at each time step t to calculate the output $y(t)$ at the current time step. The input and previous hidden state are used to update the hidden state and memory cell state at each time step, enabling the LSTM to maintain a certain level of information for a long time.

RESULTS AND DISCUSSION

The proposed LSTM-based PM_{2.5} forecasting model was trained and evaluated on a compiled dataset of PM_{2.5} of Ganga Nagar, Meerut India. The root mean square error (RMSE) decreases as number of iteration increases. The Error and comparison between forecast and observed PM_{2.5} value is shown in Fig. 3. The forecast value of PM_{2.5} is shown Fig. 4.

CONCLUSION

With this work, we have demonstrated that short-term pollution forecasts can be successfully

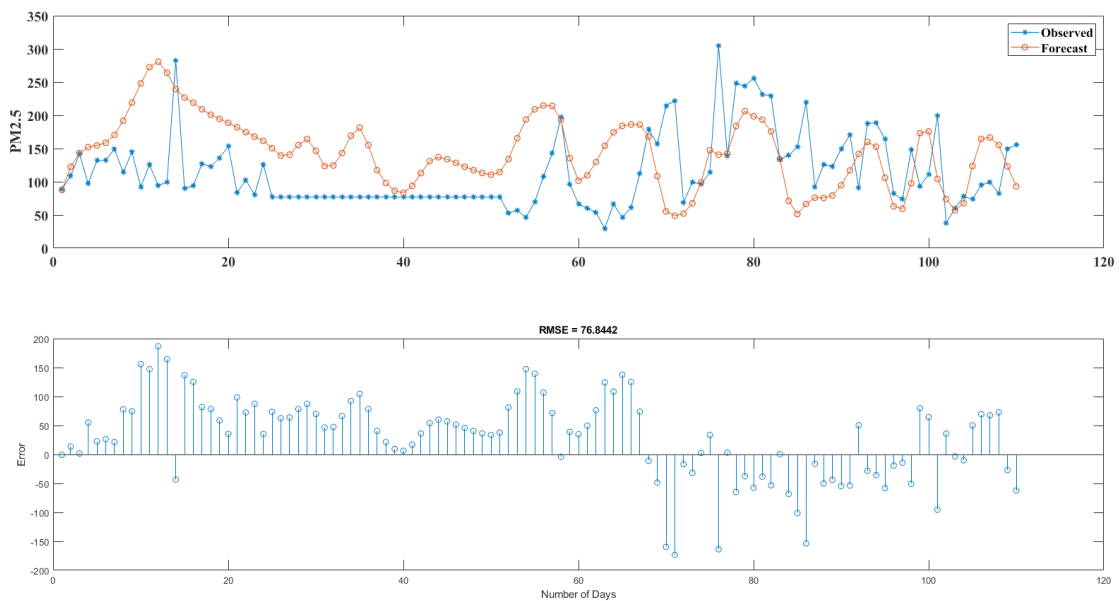


Fig. 3. Error and comparison between forecast and observed PM_{2.5} value

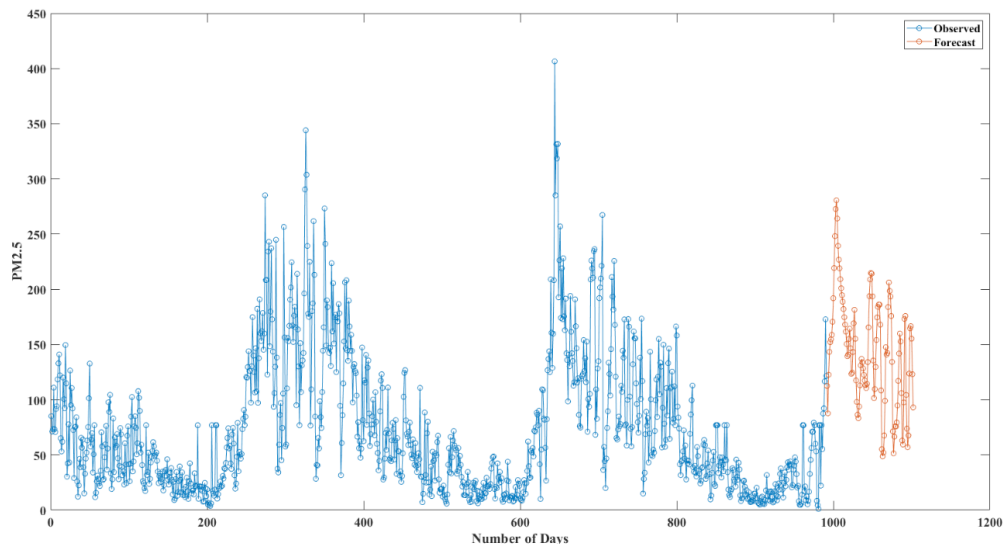


Fig. 4. Forecast value of PM_{2.5}

generated using LSTMs, allowing citizens, governments, and businesses to use this tool as a decision-aid. Implementing high-resolution local prediction systems will be one of the research's future uses, allowing us, for instance, to be aware of the level of pollution in schools and its significance for children's health. The idea is to install sensors in key locations to monitor and forecast pollutants and provide advice on air quality, enabling decision-making and problem-solving. For the benefit of future generations, we must create a healthier environment.

ACKNOWLEDGEMENT

The authors would like to thank Central Pollution Control Board, New Delhi for providing the time series data for this study.

REFERENCES

- Baawain, S.M. and Serihi, A.S.A. 2014. Systematic Approach for the Prediction of Ground-Level Air Pollution (around an Industrial Port) Using an Artificial Neural Network. *Aerosol and Air Quality Research*, 14: 124-134.
- Chow, S.K.H., Lee, E.W.M. and Li, D.H.W. 2012. Short-term prediction of photovoltaic energy generation by intelligent approach. *Energy and Buildings*. 55: 660-7.
- Elminir, K.H. and Galil, A.H. 2006. Estimation of air pollutant concentration from meteorological parameters using Artificial Neural Network. *Journal of Electrical Eng.* 57(2): 105-110.
- Filho, A.S.F. and Fernandes, M.F. 2013. Time-series Forecasting of pollutant concentration levels using particle swarm optimisation and artificial neural network. *Journal of Quim Nova*. 36(6): 783-789. <http://www.cpcb.gov.in/CAAQM/fmUserAvgReportCriteria.aspx>.
- Mishra, D. and Goyal, P. 2015. Development of artificial intelligence based NO₂ forecasting models at Taj Mahal, Agra. *Atmospheric Pollution Research*. 6: 99-106.
- Skryzypski, J. and Szakiel, J.E. 2008. Neural Network Prediction Models as a Tool for Air Quality Management in Cities. *Journal of Environment Protection Engineering*. 34: 130-137.
- Sahin, A.U., Bayat, C. and Ucan, N.O. 2011. Application of cellular neural network (CNN) to the prediction of missing air Pollutant data. *Journal of Atmospheric Research*. 101: 314-326.
- Yadav, A.K., Malik, H. and Chandel, S.S. 2014. Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models. *Renewable and Sustainable Energy Reviews*. 31: 509-519.