# Abstract

# Introduction

The Earth's atmosphere is the best of all the atmospheres in our solar systems. It is made up of many layers. Since technology has been getting better, people have been moving to cities, which has made pollution worse. There are many signs of pollution, such as ozone, nitrogen dioxide, PM2.5, and sodium dioxide. Ozone is a big part of the pollution that is always in the air in cities. [1] Ozone is a greenhouse gas and a pollutant of the air in cities. It has very bad effects on both climate change and people's health. In the past few years, a lot has been done to lower surface ozone levels by putting in place strict measures to control ozone precursors' emissions. Carbon emissions around the world are also linked to ozone. In terms of controlling emissions, methane, carbon monoxide, and volatile organic compounds, which are precursors to ozone, have a lot to do with ozone and how to control them. [2] Ozone has been found to be a major oxidant, and it is a part of photochemical smog, which is one of the main pollutants that lowers the quality of the air.

Ozone plays a unique role in absorbing certain wavelengths of incoming solar ultraviolet light. One reason ozone is a serious environmental problem is because it is not directly emitted into the air, which makes it hard to predict and control. [3] All life is shielded from the sun's damaging radiation by the ozone layer, but human activities have worn down this barrier. Less UV protection from the ozone layer results from this decrease in ozone concentration. This constant decrease causes higher risks for skin cancer and cataract rates. The combustion of fossil fuels resulted in higher concentrations of trace gases like nitrous oxide and carbon monoxide. As a result of the buildup, spread, and transformation of these air pollutants, the quality of the atmosphere has decreased. [4] Climate change and air pollution are both on the rise, causing environmental conditions to deteriorate. When temperatures increase, climate change leads to a weakening of the ozone layer.

Recent patterns and distribution of field studies have shown that there is an increase in mortality rate during the summer smog due to high ground level ozone concentration. By the use of strict emission control measures for Ozone precursors, a substantial effort has been made to lower tropospheric ozone concentrations. A monitoring station has been built in accordance with ozone concentration forecasts to predict higher geographical distribution change that aids in ozone reduction. Also, it is crucial to execute accurate regional estimates for ozone concentration in order to reduce greenhouse gas emissions and ensure public health.

Several technologies can be used to figure out how much ozone is in the air. There are two main ways to figure out how much pollution is in the air: numerical methods and data-driven methods. The Goddard Earth Observing System with Chemistry, the Weather Research and Forecasting Model with Chemistry, and the Community Multistage Air Quality Model are all used as numerical models. This model allows for clear and strict logic and a strong ability to explain. This also made it hard to make long-term predictions because the cost of computing was high and it was hard to get enough data. By taking the valuable information from a vast quantity of data about the target variable, data-driven models are a potential technique to create precise predictions about the target variable. Data-driven models may be divided into shallow machine learning models and deep learning models, according to several earlier research. The quantity of ozone in the air in an hour may be predicted using machine learning in just a few simple steps. The static Multi Layer Perceptron model.

However as technology improves, new ways are being made to improve ozone monitoring and forecasting. For example, AI and machine learning algorithms are used to make predictions about ozone concentrations more accurate. These technologies can look at a lot of data from many different sources, like ground-based monitoring stations, satellites, and weather models, to make ozone concentration maps that are more accurate and complete. This can help figure out where there is a lot of ozone pollution and guide policy decisions that aim to cut ozone emissions. [7]

Machine learning is extensively used as an empirical method to forecast the ozone concentration. Throughout the time, various ensemble methods have been used. Support vector machines and random forests are some models that have been used in research. Modelling the ozone's fluctuations and making accurate forecasts are two of the most crucial duties for the researchers. There are both deterministic and statistical (black box) models. A lot of physical and chemical interactions between predictor variables must be taken into account during the relatively complex process of using partial differential equations to create a deterministic model to forecast ozone concentrations in a specific area. This process also necessitates a lot of accurate input data (such as emissions, meteorology, and land cover). These are the key reasons why creating and maintaining deterministic models is expensive. [8] Several studies have improved the ozone concentration forecast system during the last ten years.

Artificial neural networks are also used to predict pollutants like particulate matter, sulphur dioxide, etc. This paper compares three predictive models: the autoregressive-moving average with exogenous inputs (ARMAX), multilayer perceptron, and the finite impulse response (FIR) neural network. The goal is to figure out the hourly ozone concentration 24 hours in advance. They looked at the highest levels of ozone in the summer between 1996 and 1999 at three Spanish monitoring sites in cities and towns. The MLP neural networks performed better than the linear ARMAX models, which performed better than the dynamic FIR neural networks, based on the five performance criteria that were used. [10]

**Objects and the author’s contribution**

In this paper, we suggest analysing Indian air quality data using ozone concentration data obtained from a monitoring site. This includes forecasting ozone concentration for data on the quality of the air in India using several machine learning and deep learning techniques. All forecasting models may be compared, and the results can be used to make decisions in the future. For the purpose of predicting ozone, the suggested models will learn from regional patterns and long-term spatiotemporal distributions of air quality data. Our goal is to assess the state of ozone in the vicinity of a reputable monitoring station and the relationships between other pollutants and ozone concentration.

# Literature Review

  Brian S. et. al. has trained a deep learning model which is a hybrid combination of Recurrent Neural Network and Long short Term Memory models. The author had done forecasting the prediction for 8hr of Ozone concentration. During the Training of model, Data had been collected from Kuwait, small country located in the Persian Gulf. The Differential optical absorption spectroscopy analyzers were used to collect data for tropospheric Ozone. They have reduced the dimension of data with the help of PCA and correlation filters. The Mean Absolute Error was used as a regression metric. This allowed air station to accurately predict air pollution concentration while only monitoring key factors for real time analysis.[11]

Bilge Ozbay et. al. have applied multivariate statistical methods in predicting ozone ( O3) measurements at the ground level of the troposphere as the function of pollution and atmospheric considerations. Various correlation between ozone and other pollutants were investigated using statistical correlation methods like bivariate and Pearson correlation methods. The paper contain two main techniques. One of the way to use (MLR) Multiple Linear Regression for prediction Ozone concentration while this also focused on PCA for reducing number of variables for prediction. The Model made for both annual and seasonal trend for forecasting the pollutant to test the model efficiency in different condition. Annual and Seasonal Trends were evaluated individually in MLR models and calculated R2 values were found as 0.90, 0.92, 0.85 respectively. The Warming condition got the highest R2 Score 0.92.[12]

Ning Jin et. al. proposed a novel model in the log short term memory model. The model is an enhanced using nested LSTM layer and using multiple task multiple channels for forecasting AQI data. The data was pre-processed using Discrete Stationary Wavelet Transform that decompose the real data into multiple sub-signals including eliminate lower frequency and diminish higher frequency components. Many models like SVR, LSTM, NLSTM and MLP are used to compare the efficiency of proposed model. They have done prediction for six air pollutants from which the results for ozone was 0.99 of R2 score. They have used Beijing dataset from 12 observing station for UCI ML repository [13].

Ebrahim et al. have a one of its first kind of Convolutional Neural Network based regression model that is used to predict the real time hourly concentration in Seoul, South Korea for 2017. This paper emphasis on how we can also use convolutional layers for regression as CNN captures temporary variations of the input data by convolution through time series using kernel of fixed size. The small changes have resulted in better correlation that in term gives beneficial results for the prediction system. The model is made of five convolutional layers with continuation by a fully connected layer and then output layer. The ReLU activation function is implemented to normalize input data.The Index of agreement and Correlation coefficient, MAE and RMSE are used for evaluation. The IOA achieved was 0.87 and correlation coefficient of 0.7. the Root mean square error was 12.01 for entire 2017 year.[14]

Kabesok Ko.et al. has worked on a dataset about ozone concentration. The ddata set used here contain planetary

# Reference

1. Tree-based ensemble deep learning model for spatiotemporal surface ozone (O3) prediction and interpretation Zhou Zang
2. Regional prediction of ground-level ozone using a hybrid sequence-tosequence deep learning approach
3. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone
4. Seasonal ground level ozone prediction using multiple linear regression (MLR) mode
5. Hybrid deep learning model for ozone concentration prediction: comprehensive evaluation and comparison with various machine and deep learning algorithms
6. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone
7. Long time series ozone prediction in China: A novel dynamic spatiotemporal deep learning approach
8. Hourly ozone prediction for a 24-h horizon using neural networks
9. Balaguer Ballester, E., Camps i Valls, G., Carrasco-Rodriguez, J.L., Soria Olivas, E., del Valle-Tascon, S., 2002. Effective 1-day ahead prediction of hourly surface ozone concentrations in eastern Spain using linear models and neural networks. Ecological Modelling 156, 27–41.
10. Deep Air Quality Forecasting Using Hybrid Deep Learning Framework
11. Brian S. Freeman, Graham Taylor, Bahram Gharabaghi & Jesse Thé (2018) Forecasting air quality time series using deep learning, Journal of the Air & Waste Management Association, 68:8, 866-886, DOI: [10.1080/10962247.2018.1459956](https://doi.org/10.1080/10962247.2018.1459956)
12. Bilge Özbay, Gülşen Aydın Keskin, Şenay Çetin Doğruparmak, Savaş Ayberk,Multivariate methods for ground-level ozone modeling,Atmospheric Research, Volume 102, Issues 1–2, 2011, Pages 57-65, ISSN 0169-8095, <https://doi.org/10.1016/j.atmosres.2011.06.005>. (<https://www.sciencedirect.com/science/article/pii/S0169809511001839>)
13. N. Jin, Y. Zeng, K. Yan and Z. Ji, "Multivariate Air Quality Forecasting With Nested Long Short Term Memory Neural Network," in IEEE Transactions on Industrial Informatics, vol. 17, no. 12, pp. 8514-8522, Dec. 2021, doi: 10.1109/TII.2021.3065425.
14. A real-time hourly ozone prediction system using deep convolutional neural network Ebrahim Eslami, Yunsoo Choi\*, Yannic Lops, Alqamah Sayeed Department of Earth and Atmospheric Sciences, University of Houston, TX 77004 \*corresponding author, [ychoi23@central.uh.edu](mailto:ychoi23@central.uh.edu)