

## RESEARCH ARTICLE



# Forecasting of Total Column Ozone using Regression Analysis and LSTM-RNN Machine Learning Approach

**OPEN ACCESS****Received:** 31-03-2022**Accepted:** 17-06-2022**Published:** 27-07-2022

**Citation:** Anu TM, Elampari K (2022) Forecasting of Total Column Ozone using Regression Analysis and LSTM-RNN Machine Learning Approach. Indian Journal of Science and Technology 15(29): 1420-1428. <https://doi.org/10.17485/IJST/v15i29.730>

\* **Corresponding author.**

[tmanu196@gmail.com](mailto:tmanu196@gmail.com)

**Funding:** None

**Competing Interests:** None

**Copyright:** © 2022 Anu & Elampari. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](#))

**ISSN**

Print: 0974-6846

Electronic: 0974-5645

**T M Anu<sup>1\*</sup>, K Elampari<sup>2</sup>**

**1** Research Scholar, Reg. No. 18113152132030, Physics Research Centre, S.T. Hindu College, Nagercoil, Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli, 627 012, Tamil Nadu, India

**2** Associate Professor, Department of Physics, S.T. Hindu College, Nagercoil, Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli, 627 012, Tamil Nadu, India

## Abstract

**Objectives:** Atmospheric Ozone plays an important role in global climate change, human health and environmental conditions. Accurate and timely prediction of variation in Total Column Ozone (TCO) concentration is very important for both climatology and environment. In this context, the present study aims to utilize the advancement of Artificial Intelligence in Machine learning and a conventional method to develop models for the prediction of TCO concentration using historical time series data over a tropical region, India.

**Methods:** Long Short-Term Memory (LSTM) deep learning networks are very useful in classifying, processing and making predictions based on time series data. In this work, Multiple Linear Regression (MLR) model, Long Short-Term Memory (LSTM) models are developed for forecasting TCO over a tropical region. The Statistical Parameters such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) were used to analyze and evaluate the performance of the proposed models.

**Findings:** In the present study the TCO concentration varied between 217.4 DU and 288.6 DU. The performances of MLR and LSTM model were compared. From the results, it is evident that the values of RMSE, MAE and MAPE of the MLR model are 3.807924, 2.93766, 1.15 respectively whereas the values for LSTM model are 3.074492, 2.574, 1.01 respectively. LSTM was found to be the accurate model for the study data set and the predicted TCO concentration has a very good correlation with the actual observations and it does not rely on the meteorological parameters. **Novelty:** The LSTM models are usually employed in the field of machine translation, speech recognition and for the prediction of traffic levels, stock levels and air pollutants. This work utilizes the fast numerical machine learning computing library Tensorflow and a high level neural network library Keras that runs on top of Tensorflow to develop a LSTM network model for the prediction of the concentration levels of TCO.

**Keywords:** Total Column Ozone; Regression; Forecasting; Climate Change; Long Short-Term Memory

## 1 Introduction

Global warming and climate change are the current challenges for the living community on the earth. Global warming is primarily caused because of the excessive emission of carbon dioxide into the atmosphere. Also the extensive research on Global warming has highlighted the role of ozone as a core component of the global climate system. Ozone is a gas that is naturally present in our atmosphere and it is found primarily in two of the regions of the atmosphere; about 90% in the stratosphere and the remaining 10% in the troposphere. Total ozone at any location on the globe is defined as the sum of all the ozone in the atmosphere directly above that location and is measured in Dobson Unit (DU)<sup>(1)</sup>. The thickness of Total Column Ozone (TCO) is affected by various aspects like weather systems, changes in the atmospheric chemicals like CH<sub>4</sub>, N<sub>2</sub>O and CFCs due to anthropogenic activities, Quasi-Biennial Oscillation (QBO) and changes in the UV flux of solar insolation. Since TCO is one of the important component of the earth's climate system, it is very essential to predict the variations in the TCO to understand the future changes to our environment.

Regression Analysis, Time Series Decomposition, Moving Average Methods are some of the conventional forecasting models. Machine Learning is an important branch of Artificial Intelligence and Data Science, and it allows machines to learn, think and solve complex problems including predictive analytics. An Artificial Neural Network (ANN) was a mathematical model that was driven by the functional feature of biological neural networks<sup>(2)</sup>. Deep learning is a broad subset of Machine Learning methods based on Artificial Intelligence Neural networks, developed to improve the accuracy of prediction models and was very effective even if the data were diverse, less structured or independent. Deep learning methods were successfully implemented for forecasting stock levels, traffic flow, sales levels, air pollutants, etc. ANN model was implemented in<sup>(3)</sup> to predict the mechanical properties such as plastic behavior, yield strength and tensile strength of aluminum alloys. The phase transitions in frustrated magnetic models were successfully determined by<sup>(4)</sup> using ANN. ANN model was utilized by<sup>(2)</sup> for the prediction of animal category. The authors<sup>(5)</sup> proposed a novel approach based on LSTM-PSO for sale forecasting in E-Commerce. LSTM approach performed high level accuracy for sales forecasting of goods with short term demands<sup>(6)</sup>. The authors<sup>(7)</sup> present an LSTM model for traffic flow prediction. ANN model and Multiple Linear Regression model adopted by<sup>(8,9)</sup> for forecasting surface ozone density, CO, NO<sub>2</sub>, and SO<sub>2</sub> concentration using the historical concentration of data. The authors<sup>(10)</sup> discussed the implementation of LSTM and ARIMA models to forecast the Air pollutants such as CO and NH<sub>3</sub>. The authors<sup>(11)</sup> adopted four different models such as ARIMA, FB Prophet, LSTM and 1D-CNN to predict the concentration of PM<sub>2.5</sub>. In the article,<sup>(12)</sup> the authors proposed the models LSTM and DAE (Deep Auto Encoders) for the prediction of historical PM<sub>2.5</sub> in South Korea. All the authors found that the LSTM model provided better results than conventional models. This work is focused on the prediction of TCO concentration over a tropical region, by using both the conventional and machine learning methods.

## 2 Methodology

The TCO data for a period of 20 years (1<sup>st</sup> January 2000 – 31<sup>st</sup> December 2019) of a tropical region in India, covered by 13.00°N - 80.18°E on the coast of Bay of Bengal is used in this study. The data were obtained from NASA's SBUV - Merged Ozone Data base. (<https://acd-ext.gsfc.nasa.gov/anonftp/toms/sbu/MERGED>). The meteorological data such as Relative Humidity (RH), Pressure (P), Dew Point (DEW), Earth Skin Temperature (EST), Atmospheric Temperature (T), Wind Speed at 50M (WS50M), Wind Speed at 10M (WS10M) and Solar Radiation Insolation Index(SRI) were sourced from NASA's Power Data Access site (<https://power.larc.nasa.gov>).

The analysis of TCO concentration along with meteorological parameters has been carried out by using various statistical and mathematical tools. Correlation Analysis and Time Series Analysis were used to analyse the data for dependency, trend and the influences of meteorological parameters on ozone. The deep learning frameworks Tensorflow and Keras were used to build the proposed LSTM model. Both MLR model and LSTM model are described in the sections (2.1-2.4).

### 2.1 Time Series

A time series is a sequential set of data points, measured typically over successive times. The four main components of a time series are Trend, Seasonal, Cyclic and Irregular. Trend affects the overall movement of average of the data across time series. The medium-term cyclic changes in data describes as cyclic. Seasonal occurs when there is a pattern across time series at a repeated interval. Irregular pattern contributes to randomness in the time series data<sup>(13)</sup>.

### 2.2 Multiple Linear Regression Model

Regression models are pure statistical models based on the empirical relations among different variables and are very useful in Prediction and Data Description. Multiple Linear Regression (MLR) assumes a linear relationship between the independent

and dependent variables<sup>(8)</sup>, and it is used to examine the relationships between the variables. Once each of the independent factors has been determined to predict the dependent variable, the information on the multiple variables can be used to create an accurate prediction on the level of effect they have on the outcome variable. In general, the regression equation may be written as Equation 1.

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_sX_s + e \text{ ---- (1)}$$

Here, Y is the dependent variable, X<sub>1</sub>, X<sub>2</sub>, ..... X<sub>s</sub> are independent variable, b<sub>0</sub>, b<sub>1</sub>, ..., b<sub>s</sub> are called regression coefficients and e is the error in prediction (residual) for each case<sup>(14)</sup>. The coefficient of determination R<sup>2</sup> is calculated to indicate the percent of how much of the total variance is explained by the independent variables.

### 2.3 Long Short -Term Memory Model

Long short-term memory (LSTM) neural network is a widely used RNN architecture in the field of deep learning<sup>(15)</sup>, that mainly solves the processing of sequence data in which each data segment has a correlation with the previous data segment. This idea was first proposed by Hochreite and Schmidhuber and the network consists of an input layer, recurrent hidden layers known as memory blocks and an output layer<sup>(16)</sup>. The memory unit structure of the LSTM network is shown in Figure 1.

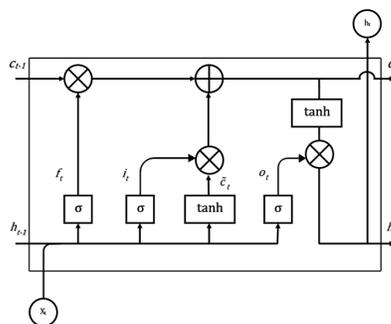


Fig 1. Structure of an LSTM memory unit

The heart of the LSTM is its memory cell or memory module that consists of a loop unit and three gates such as an input gate, a forget gate and an output gate. A nonlinear function controls each gate switch to protect and control the state of the memory unit. In general, the output value of each gates is 0~1, the sigmoid function is used to determine how much information can be given as input to the memory location<sup>(17)</sup>. The input gate is designed to control the writing of input information to the memory and it decides which new messages are to be remembered in the neuron state while the forget gate tells about the information which is to be thrown away from the cell state. The forget gate and output gate determine whether to save or release information from the memory at each decision point<sup>(18)</sup>. The output gate provides the activation signal to the final output of the LSTM block at timestamp ‘t’. The state and status of LSTM cell can be described by a set of mathematical equations. The gates equations for the LSTM are,

$$i_t = \sigma(w_i \bullet (h_{t-1}, x_t] + b_i) \text{ ---- (2)}$$

$$f_t = \sigma(w_f \bullet (h_{t-1}, x_t] + b_f) \text{ ---- (3)}$$

$$o_t = \sigma(w_o \bullet (h_{t-1}, x_t] + b_o) \text{ ---- (4)}$$

Where, i<sub>t</sub> represents input gate, f<sub>t</sub> represents forget gate and o<sub>t</sub> represents the output gate. σ Represents the sigmoid function  $\frac{1}{1+e^{-t}}$ , w<sub>x</sub> represents weight of the respective gate(x) neuron, h<sub>t-1</sub> represents output of the previous LSTM block (at time stamp t-1), b<sub>x</sub> represents biases for the respective gates(x). x<sub>t</sub> represents input at current time stamp. The equations for the state of the cells in the LSTM block, the candidate cell state and the final output is given by,

$$\bar{c}_t = \tanh(w_c \bullet (h_{t-1}, x_t] + b_c) \text{ ---- (5)}$$

$$c_t = f_t * c_{t-1} + i_t * \bar{c}_t \text{ ---- (6)}$$

$$h_t = o_t * \tanh(c_t) \text{ ---- (7)}$$

‘\*’ indicates element wise multiplication.

Where c<sub>t</sub> is the cell state (memory ) at time stamp t,  $\bar{c}_t$  is the candidate for cell state at time stamp t. h<sub>t</sub> hidden cell state. In detail, the model input is given as x = (x<sub>1</sub>, x<sub>2</sub> ..., x<sub>n</sub>), where each x<sub>t</sub> ∈ R<sup>T</sup>, i = 1, 2, .....n; n denotes the number of input

dimensions; T denotes the time lag; and the output sequence is given as  $y = (y_1, y_2, \dots, y_n)$ . A more detailed description of the working of LSTM can be found elsewhere<sup>(19)</sup>. In this paper, the high level neural network library Keras<sup>(20)</sup> and the open source machine learning library Tensor flow were used to build the proposed LSTM model using Anaconda 3 as a back end tool and Python 3.8.8 as a front end tool.

### 2.4 Evaluation of the Performance of developed model

The prediction efficiencies of the models developed were tested using following parameters: Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). These parameters provide a general indication of the strength of the relation between observed and predicted data. These three parameters can be calculated using equations (8-10) respectively<sup>(21)</sup>.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - P_i)^2} \text{ --- (8)}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N abs(O_i - P_i) \text{ --- (9)}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{abs(O_i - P_i)}{O_i} \text{ --- (10)}$$

where  $O_i$  and  $P_i$  denote the observed values and the predicted values; N denotes the number of evaluation samples. MAE and RMSE from the above equations were used to evaluate the absolute error. The performance of a model is said to be high when the values of RMSE and MAE are low. MAPE is used to measure the relative error; the smaller the error value, the closer the predicted result to the actual value.

## 3 Results and Discussion

### 3.1 Descriptive Statistics

The descriptive statistics of TCO data and meteorological parameters for the period from 2000 to 2019 are given in Table 1 . The data contains 7305 measurements gathered between January 2000 and December 2019. For the study region the TCO concentration varies from 217.4 DU to 288.6 DU.

**Table 1.** Descriptive Statistics of TCO and Meteorological Parameters

Parameters	MIN	MAX	MEAN	STD.DEV
TCO	217.4	288.6	260.91	13.21
RH	49.21	92.18	72.53	7.47
P	99.25	101.38	100.45	0.37
DEW	15.79	26.48	22.49	1.93
EST	23.31	34.96	28.86	2.43
T	22.37	34.12	28.05	2.29
WS50M	1.24	16.51	5.39	1.60
WS10M	1.00	13.58	4.38	1.26
SRI	0.16	7.86	5.37	1.49

### 3.2 Correlation Analysis

The Meteorological conditions affect the processes of ozone formation, transportation and dispersion. The changes in TCO vary from place to place depending upon the Meteorological parameters and Geographical factors. The Scatter Diagram and Karl Pearson’s coefficient of correlation methods were used as the measure of correlation. Table 2 presents the correlation between TCO and some meteorological parameters used in this study and scatter plots are shown in Figure 2.

The results of the correlation analysis between the ozone concentration and meteorological parameters are shown for the level of significance  $p < 0.001$ . Correlation between TCO and relative humidity is found to be negative ( $r = -0.3375$ ) which implies that the TCO level is inversely proportional to relative humidity.

The negative correlation between TCO and pressure ( $r = -0.7546$ ), implies that increase in pressure is followed by decrease in ozone concentration. The Pearson’s correlation analysis shows a positive correlation with  $r = 0.6046$  between TCO and Dew point over the study region. The value of correlation coefficients ( $r=0.7449, r=0.8078$ ) indicates that there is a very strong and effective positive relationship between TCO and Earth Skin Temperature as well as Atmospheric Temperature. It is evident

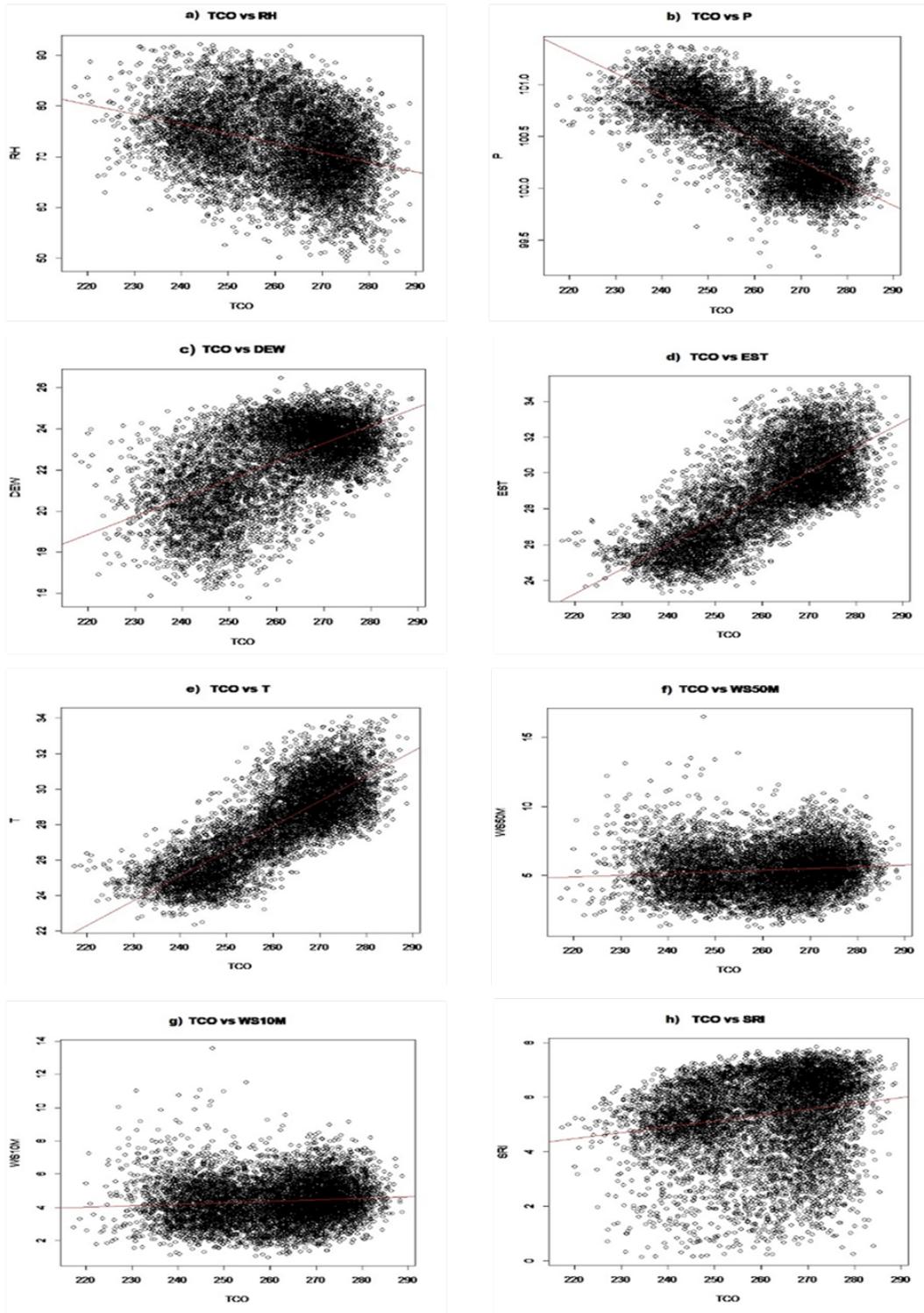


Fig 2. Correlation plot for TCO and Meteorological Parameters

**Table 2.** Correlation between TCO and Meteorological Parameters

Meteorological Parameter	r	t	df	95% Confidence Interval		p
RH	-0.3375	-30.64	7303	-0.3576689	-0.3170259	<2.2e-16
P	-0.7546	-98.27	7303	-0.7642985	-0.7445435	<2.2e-16
DEW	0.6046	64.86	7303	0.5898003	0.6189078	<2.2e-16
EST	0.7449	95.42	7303	0.734522	0.7549431	<2.2e-16
T	0.8078	117.13	7303	0.7997151	0.8156544	<2.2e-16
WS50M	0.0999	8.59	7303	0.0772381	0.1226447	<2.2e-16
WS10M	0.0932	7.99	7303	0.0703959	0.1158629	<2.2e-16
SRI	0.1923	16.74	7303	0.1700815	0.2142519	<2.2e-16

that the ozone variations are directly proportional to the temperature. The results obtained from the correlation of ozone concentrations with wind speed at 50M and 10M range is not significant in this study. The Correlation between TCO and Solar Radiation Insolation index is found to be positive( $r=0.1923$ ) and is shown in Figure 2.

### 3.3 Time Series Analysis of TCO data

The Time series analysis comprises of methods make an attempt to understand the underlying context of the data points and to make forecasting. For the study period, the minimum TCO value is observed in January 2013 and the maximum TCO value is observed in June 2004. In the monthly averages, the lowest concentration was obtained in December and the highest in May. Augmented Dickey-Fuller Test is adopted to check the stationarity of the ozone time series data. The test statistic value of -6.3676 which is less than the value of -3.431 at 1% indicates the data do not have a unit root and are stationary and this series can be used to build the models.

### 3.4 Multiple Linear Regression (MLR) Model

The TCO is estimated by using eight independent variables such as Relative Humidity(RH), Pressure(P), Dew point (DEW), Earth Skin Temperature(EST), Atmospheric Temperature (T), Wind Speed at 50M(WS50M), Wind Speed at 10M(WS10M) and Solar Radiation Insolation Index(SRI). Regression results obtained from the modeling of data set are shown in Table 3 . The MLR model is represented by the mathematical equation(11) and it is used to predict the TCO.

$$y(\text{TCO})=b_0+b_1*(\text{RH})+b_2*(\text{P})+b_3*(\text{DEW})+b_4*(\text{EST})+b_5*(\text{T})+b_6*(\text{WS50M})+b_7*(\text{WS10M})+b_8*(\text{SRI}) \text{ --- (11)}$$

In equation (11), the constants  $b_1, b_2, \dots$  are the regression coefficients of various parameters as described in Table 3 .

**Table 3.** Multiple Regression Analysis

Variable	Symbol	Regression coefficient	Standard error	't' value
Regression constant	$b_0$	-533.107	410.137	-1.300
RH	$b_1$	1.733	1.057	1.639
P	$b_2$	5.880	3.913	1.503
DEW	$b_3$	6.351	1.509	4.209
EST	$b_4$	-9.014	4.536	-1.987
T	$b_5$	3.157	4.803	0.657
WS50M	$b_6$	-22.170	5.293	-4.188
WS10M	$b_7$	27.772	6.598	4.209
SRI	$b_8$	-1.156	1.198	-0.964

Figure 3 compares the actual values (ATCO) and predicted values (PTCO) of TCO by multiple linear regression model. The coefficient of multiple determination  $R^2$  of the regression model is 0.80 and it indicates that 80% of variation in TCO is being explained by the eight variables included in the model.

### 3.5 Long Short – Term Memory Model

The proposed LSTM model is illustrated in the Figure 4.

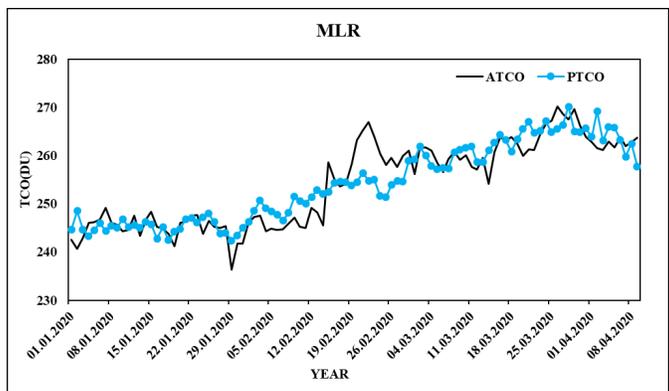


Fig 3. Predicted and actual values of TCO - Multiple Linear Regression Model

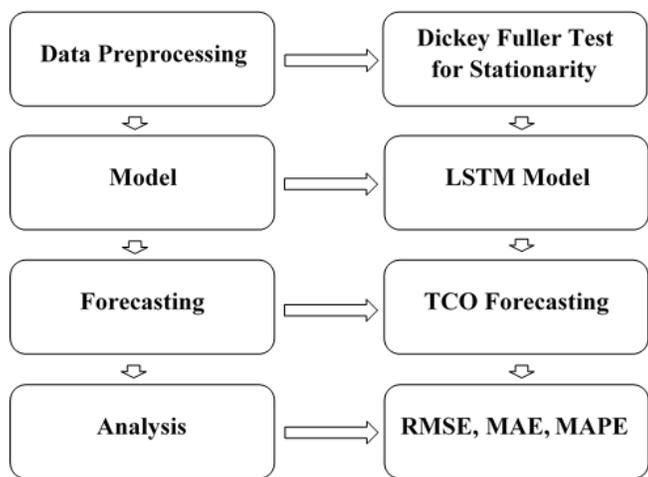


Fig 4. Proposed LSTM model

The historical TCO data is partitioned into three sets; training, testing and validation data and they are used to build and train the model. By means of a random process, the data points are divided into two distinct sets: a training set consisting of 75% of data and the remaining 25% of the data points used for testing. Data for a period of 100 days (100 data points) is used as validation set for verifying the predicted TCO. A time series consists of past 7305 days of observation is used to predict the values of TCO. A multilayered LSTM recurrent neural network is modelled to predict the values of TCO for the next 100 days. The model is defined with 200 neurons in 1<sup>st</sup> hidden layer and 1 neuron in the output layer. To avoid overfitting on the sample data and to increase the generalization capacity of the model, the model is trained for optimal number of epochs<sup>(22)</sup>. The early stopping callback function in keras is utilized to monitor the loss accuracy values and it is found that the optimal number of epochs to train the sample data is 15 and ADAM optimizer is utilized in this model. The model is fit for 15 training epochs with a batch size of 32 and for this model, each time step is one day. LSTM Predicted (PTCO) and actual (ATCO) values of ozone are shown in Figure 5 .

### 3.6 Evaluation Performance of Models

The predicted values by the MLR model and LSTM model were compared with the actual values. The prediction efficiencies of these two models were tested using statistical error analysis methods and the results are tabulated in Table 4 for comparison.

The statistical comparison of two models were used to provide a general indication of the relationship between predicted and actual data. Figure 6 shows the predicted values are generally consistent with the actual values for both MLR and LSTM models but the performance of the LSTM model is found to be better than the MLR model. The R<sup>2</sup> value of these two models

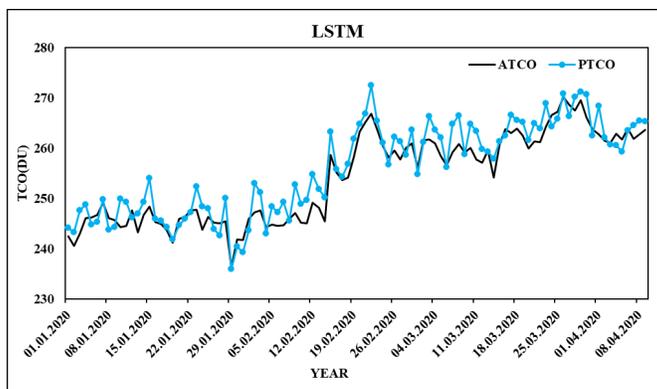


Fig 5. Predicted and actual values of TCO - Long Short – Term Memory Model

Table 4. Statistical Comparison of MLR and LSTM Model

Model	RMSE	MAE	MAPE	R <sup>2</sup>
MLR	3.807924	2.93766	1.152153	0.809301
LSTM	3.074492	2.574	1.01	0.921995

reveals that 92% of explained variance is captured by the LSTM network and 80% of that is captured by MLR model.

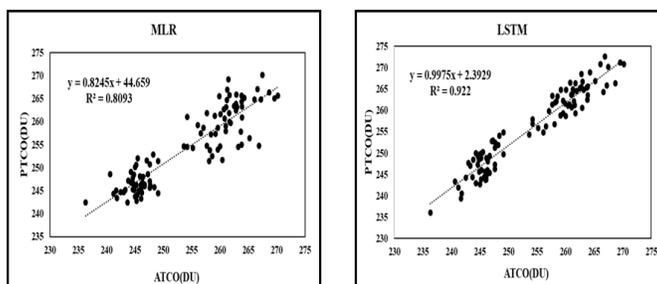


Fig 6. Prediction Performance of MLR and LSTM model

### 4 Conclusion

This study helps to predict the level of TCO concentration over a tropical region by using conventional MLR and LSTM models. The prediction accuracy of the models was assessed by RMSE, MAE and MAPE values. The lower values of the error parameters obtained for LSTM indicates for its improved performance over the conventional MLR. Also the MLR method uses several meteorological parameters as input in the estimation of TCO whereas LSTM uses only the time series for the prediction. Thus, the LSTM model does not rely on meteorological parameters and this model suggests that it is possible to predict the TCO concentration over any region without the meteorological parameters.

### 5 Acknowledgement

The authors are thankful to National Aeronautics and Space Administration (NASA) for providing TCO data and also thankful to Physics Research Centre, S.T. Hindu College, Nagercoil.

### References

- 1) Salawitch RJ, Fahey DW, Hegglin MI, Mcbride LA, Tribett WR, Doherty SJ. Twenty Questions and Answers About the Ozone Layer: 2018 Update, Scientific Assessment of Ozone Depletion: 2018, World Meteorological Organization, Geneva, Switzerland. Geneva, Switzerland. 2019. Available from: <https://csl.noaa.gov/assessments/ozone/2018/twentyquestions>.

- 2) Nasser IM, Abu-Naser SS. Artificial Neural Network for Predicting Animals Category. *International Journal of Academic and Applied Research*. 2019;3(2):18–24. Available from: <http://dstore.alazhar.edu.ps/xmlui/handle/123456789/182>.
- 3) Merayo D, Rodriguez-Prieto A, Camacho AM. Prediction of Mechanical Properties by Artificial Neural Networks to Characterize the Plastic Behavior of Aluminum Alloys. *Materials*. 2020;13(22):5227–5227. Available from: <https://doi.org/10.3390/ma13225227>.
- 4) Corte I, Acevedo S, Arlego M, Lamas CA. Exploring neural network training strategies to determine phase transitions in frustrated magnetic models. *Computational Materials Science*. 2021;198:110702–110702. Available from: <https://doi.org/10.1016/j.commatsci.2021.110702>.
- 5) He QQ, Wu C, Si YW. LSTM with Particle Swarm Optimization for Sales Forecasting. *Electronic Commerce Research and Applications*. 2022;51:101118–101118. Available from: <https://doi.org/10.1016/j.elerap.2022.101118>.
- 6) Shih YS, Lin MH. A LSTM Approach for Sales Forecasting of Goods with Short-Term Demands in E-Commerce. *Intelligent Information and Database Systems*. 2019;p. 244–256. Available from: [https://doi.org/10.1007/978-3-030-14799-0\\_21](https://doi.org/10.1007/978-3-030-14799-0_21).
- 7) Narmadha S, Vijayakumar V. Spatio-Temporal vehicle traffic flow prediction using multivariate CNN and LSTM model. *Materials Today: Proceedings*. 2021. Available from: <https://doi.org/10.1016/j.matpr.2021.04.249>.
- 8) Shams SR, Jahani A, Kalantary S, Moeinaddini M, Khorasani N. The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models for predicting SO<sub>2</sub> concentration. *Urban Climate*. 2021;37:100837–100837. Available from: <https://doi.org/10.1016/j.uclim.2021.100837>.
- 9) Shams SR, Jahani A, Moeinaddini M, Khorasani N, Kalantary S. Forecasting Ozone Density in Tehran Air Using a Smart Data-Driven Approach. *Journal of Health and Safety at Work*. 2021;10(4):406–408. Available from: <http://jhs.w.tums.ac.ir/article-1-6415-en.html>.
- 10) Mani G, Voley R. A comparative analysis of LSTM and ARIMA for enhanced real-time air pollutant levels forecasting using sensor fusion with ground station data. *Cogent Engineering*. 2021;8(1):1936886–1936886. Available from: <https://doi.org/10.1080/23311916.2021.1936886>.
- 11) Garg S, Jindal H. Evaluation of Time Series Forecasting Models for Estimation of PM<sub>2.5</sub> Levels in Air. *2021 6th International Conference for Convergence in Technology (I2CT)*. 2021;p. 1–8. Available from: <https://doi.org/10.1109/I2CT51068.2021.9418215>.
- 12) Xayasouk T, Lee H, Lee G. Air Pollution Prediction Using Long Short-Term Memory (LSTM) and Deep Autoencoder (DAE) Models. *Sustainability*. 2020;12(6):2570–2570. Available from: <https://doi.org/10.3390/su12062570>.
- 13) Yatish HR, Swamy SR. Recent Trends in Time Series Forecasting- A Survey. *International Research Journal of Engineering and Technology*. 2020;7(4). Available from: <https://www.irjet.net/archives/V7/I4/IRJET-V7I41059.pdf>.
- 14) Turabieh H, Sheta A, Braik M, Kovac-Andric E. A Layered Recurrent Neural Network for Imputing Air Pollutants Missing Data and Prediction of NO<sub>2</sub>, O<sub>3</sub>, PM<sub>10</sub>, and PM<sub>2.5</sub>. In: *Forecasting in Mathematics - Recent Advances, New Perspectives and Applications*, IntechOpen. 2021. Available from: <https://doi.org/10.5772/intechopen.93678>.
- 15) Chen Y, Cui S, Chen P, Yuan Q, Kang P, Zhu L. An LSTM-based neural network method of particulate pollution forecast in China. *Environmental Research Letters*. 2021;16(4):044006–044006. Available from: <https://doi.org/10.1088/1748-9326/abe1f5>.
- 16) Moscoso-López JA, González-Enrique J, Urda D, Ruiz-Aguilar JJ, Turias JJ. Hourly pollutants forecasting using a deep learning approach to obtain the AQI. *Logic Journal of the IGPL*. 2022. Available from: <https://doi.org/10.1093/jigpal/jzac035>.
- 17) Liu B, Zhang L, Wang Q, Chen J. A Novel Method for Regional NO<sub>2</sub> Concentration Prediction Using Discrete Wavelet Transform and an LSTM Network. *Computational Intelligence and Neuroscience*. 2021;2021:1–14. Available from: <https://doi.org/10.1155/2021/6631614>.
- 18) Lei F, Dong X, Ma X. Prediction of PM<sub>2.5</sub> concentration considering temporal and spatial features: A case study of Fushun, Liaoning Province. *Journal of Intelligent & Fuzzy Systems*. 2020;39(5):8015–8025. Available from: <https://doi.org/10.3233/JIFS-201515>.
- 19) Chang YS, Chiao HT, Abimannan S, Huang YP, Tsai YT, Lin KM. An LSTM-based aggregated model for air pollution forecasting. *Atmospheric Pollution Research*. 2020;11(8):1451–1463. Available from: <https://doi.org/10.1016/j.apr.2020.05.015>.
- 20) Moolayil J. An introduction to deep learning and keras. *Learn Keras for Deep Neural Networks*;vol. 2019. Apress, Berkeley, CA. . Available from: <https://doi.org/10.1007/978-1-4842-4240-7>.
- 21) Heydari A, Nezhad MM, Garcia DA, Keynia F, De Santoli L. Air pollution forecasting application based on deep learning model and optimization algorithm. *Clean Technologies and Environmental Policy*. 2022;24(2):607–621. Available from: <https://doi.org/10.1007/s10098-021-02080-5>.
- 22) Shin S, Lee Y, Kim M, Park J, Lee S, Min K. Deep neural network model with Bayesian hyperparameter optimization for prediction of NO<sub>x</sub> at transient conditions in a diesel engine. *Engineering Applications of Artificial Intelligence*. 2020;94. Available from: <https://doi.org/10.1016/j.engappai.2020.103761>.