

<
Previous PDF

Next PDF
>

Alexandria Engineering Journal (2021) 60, 2021–2032

HOSTED BY



Alexandria University

Alexandria Engineering Journal

[www.elsevier.com/locate/aej](http://www.elsevier.com/locate/aej)  
[www.sciencedirect.com](http://www.sciencedirect.com)



# Spatiotemporal prediction of air quality based on LSTM neural network



Dewen Seng<sup>a,b,\*</sup>, Qiyan Zhang<sup>a</sup>, Xuefeng Zhang<sup>a,b</sup>, Guangsen Chen<sup>a</sup>,  
Xiyuan Chen<sup>a,b</sup>

<sup>a</sup> School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China  
<sup>b</sup> Key Laboratory of Complex Systems Modeling and Simulation, Ministry of Education, Hangzhou 310018, China

Received 27 October 2020; revised 1 December 2020; accepted 9 December 2020  
Available online 19 December 2020

## KEYWORDS

Air quality prediction;  
Deep learning;  
LSTM;  
Supervised learning;  
Time series

**Abstract** Accurate monitoring of air quality is of great importance to our daily life. By predicting the air quality in advance, we can make timely warnings and defenses to minimize the threat to life. With a large number of environmental data, the air quality prediction based on deep learning technology is studied in depth. Based on long short-term memory (LSTM), a comprehensive prediction model with multi-output and multi-index of supervised learning (MMSL) was proposed. The particle concentration data (mainly PM<sub>2.5</sub>, means particles with aerodynamic diameter ≤ 2.5 mm) of the present monitoring station, as well as that of the nearest neighbor stations, the meteorological data, and the gaseous pollutant data in the air (mainly CO, NO<sub>2</sub>, O<sub>3</sub>, SO<sub>2</sub>) of the same period were integrated. All data were converted into the supervised learning format and normalized. The LSTM was used for training to obtain the predicted values of air quality pollution indicators (PM<sub>2.5</sub>, CO, NO<sub>2</sub>, O<sub>3</sub>, SO<sub>2</sub>). In the present study, the representative stations of the 35 monitoring stations in Beijing were selected, and input the air quality sequences of the representative stations with different data characteristics into the model to obtain the predicted concentration values of the air quality indicators of the representative stations, then calculated the average value as the overall air quality prediction result of Beijing. The air quality time series datasets collected from 35 air quality mon-