Proceedings of the 9th World Congress on Civil, Structural, and Environmental Engineering (CSEE 2024)
London, United Kingdom – April 14-16, 2024
Paper No. ICEPTP 151
DOI: 10.11159/iceptp24.151

Sustainable AI-Based Prediction of Air Pollution Levels in London

Megha Hegde, Jean-Christophe Nebel, Farzana Rahman
Kingston University, London
Faculty of Engineering, Computer Science and the Environment, Kingston Upon Thames, KT1 2EE, UK
k2162274@kingston.ac.uk; j.nebel@kingston.ac.uk; farzana@kingston.ac.uk

Abstract - Air pollution exposure not only leads to respiratory and cardiovascular diseases, but is also detrimental to cognitive abilities, mental health, and prenatal development. Thus, cities worldwide have invested in sophisticated air pollution monitoring systems to assess and reduce air pollution and its consequences. When excessive build-up of air contaminants occurs, emergency measures must be enacted to reduce human exposure and decrease pollution levels. Predicting such situations a few hours in advance is critical to prevent human health from being compromised. While usage of deep neural networks has become very popular, standard machine learning approaches remain very attractive: they deliver competitive performance, they do not rely on specialised equipment, and their energy consumption is sustainable. Experiments conducted on London air quality data demonstrate that Linear Regression achieves state-of-the-art performance, with 1-hour and 24-hour predictions displaying, respectively, 0.2 and 3.2 mean absolute errors. Moreover, its power usage is a fraction of what is required by its deep learning competitor for both training and predicting, i.e., 1/2840th and 1/126th, respectively. This is significant as they demonstrate air pollution prediction can be sustainable and accurate without prohibitive hardware investments.

Keywords**:** air pollution prediction, sustainable AI, machine learning, energy consumption, air quality, sustainable

1. Introduction

In 1952, the Great Smog of London killed over 10,000 people in 5 days [1]. It led to the first legislation aimed at controlling dangerous emissions in 1956 [2]. Although air quality is now meticulously monitored in the city, 3,600 to 4,100 deaths were attributed to air pollution in 2019 [3]. This situation led Public Health England to categorise human-made air pollution as the most significant environmental risk to public health in the UK. The two most dangerous pollutants are nitrogen dioxide (NO2) and fine particulate matter with a diameter of 2.5 µm or less ($PM_{2.5}$). London authorities put in place an Ultra Low Emission Zone (ULEZ) to reduce air pollution. Since internal combustion engines are the main sources of NO2, its emissions between 2019 and 2022 were reduced by 23%. Unfortunately, $PM_{2.5}$ emissions only decreased by 7% [4]. Indeed, they are much more difficult to both control and predict as only 30% are due to road traffic: over 50% come from regional/international sources and 17% from households burning wood and coal for heating.

Forecasting air pollution proves invaluable in providing information about pollution levels, enabling policymakers to implement measures to mitigate its impact. Thus, many studies have developed air quality forecasting models [5], based on statistical, deterministic, physical, and machine learning (ML) approaches [6]. Methods relying on probability and statistics tend to be intricate and less effective than ML-based models which have demonstrated more reliability and consistency. Among them, Support Vector Regression (SVR) has performed well in predicting pollutants and particulate levels. [6] and [7] respectively. Random Forest (RF) and XGBoost have also proved efficient and able to handle multimodal data such as street map and weather data [8]. Recently, deep learning (DL) approaches have been particularly popular. They include hybrid models such as Convolutional Neural Network- Recurrent Neural Network (RNN), Attention-RNN, and RNN-LSTM (Long Short-Term Memory) [9] and [10]. Since efforts have been focused on predicting air pollution in highly polluted megapoles, such research has been limited in the UK, even in London. Still, in a study comparing 12 ML methods forecasting $PM_{2.5}$ in London, Linear Regression (LR) proved the best standard ML method, but was outperformed by a DL method, i.e., LSTM [11]. On the other hand, similar predictions for Nottingham, reached a different conclusion with LR and SVR achieving better performance than LSTM and Bi-LSTM [12].

While model performance is crucial when choosing a ML approach, its energy and carbon footprints should also be considered. Indeed, the ML community has started reflecting on the balance between performance gains and environmental impact as per [13] and [14] respectively. This issue has particularly been exacerbated with the arrival of large language models, each being responsible for hundreds of tonnes of CO2 equivalent [15]. Moreover, as the trend during the DL era has been to see computational requirements double every 2 months [16] and [17], DL may soon emerge as a counterforce in the battle against climate change [17]. In addition, the cost of specialised hardware, which is required for DL, can be a barrier to entry for individuals and organisations creating inequality within and between communities. This is particularly distressing as air pollution already reinforces socio-economic inequalities [18]. As DL models require more computing power than traditional ML models, and recent studies suggest that, for air pollutant prediction, the performance gain of using DL may be minimal [11] or nonexistent [12], this study will focus on standard ML models using LSTM as a DL reference to identify which is both effective for various prediction horizons and sustainable in term of energy consumption.

2. Methods

The data used in this study consists of hourly $PM_{2.5}$ measurements (in μgm^{-3}) from the Eltham monitoring station in the Royal Borough of Greenwich, London. They were extracted from a wider dataset consisting of values from seven different stations over 120 days, from 1st January to 1st May 2019 [19]. After following the data preprocessing approach reported in [11], the data were restructured into a sliding window format (Fig. 1(b)) before being used for predictions (Fig. 1(a)).

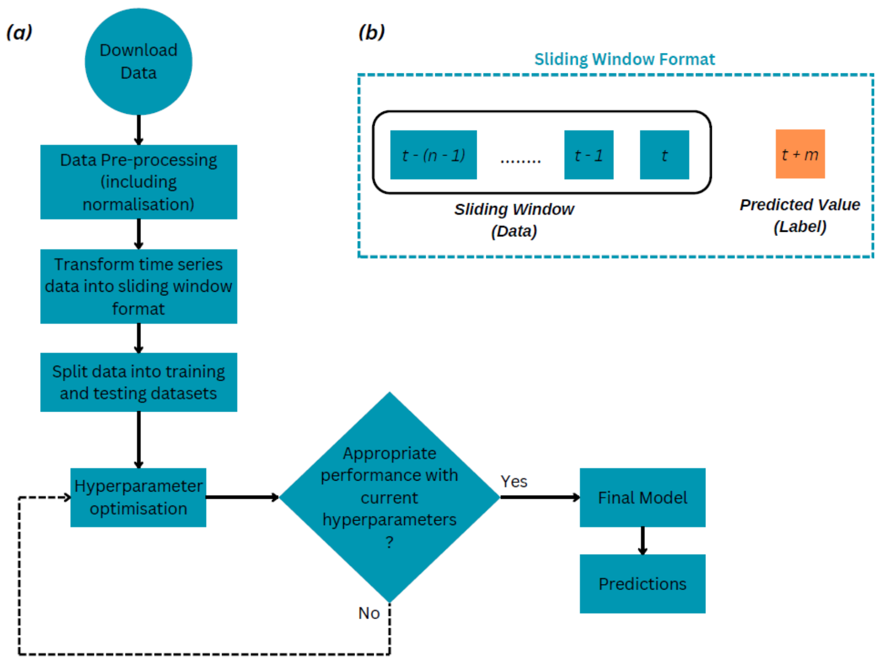

Fig. 1: Flowchart showing methodology (a) and sliding window method (b), where *t* is the current time, *n* is the size of the sliding window, and *m* is the prediction horizon length. For example, when using a sliding window of 3 hours to predict the value in 1 hour's time, *n* and *m* would be equal to 3 and 1 respectively. Thus, prediction of the value at $t + 1$ would rely on values at t, $t - 1$, and $t - 2$.

This study evaluates traditional ML approaches, (i.e., linear regression, RF regression, and XGBoost), and LSTM for forecasting hourly $PM_{2.5}$ concentration in Eltham. It goes further than [11] and their next hour predictions as it investigates predictions for longer horizons from 3 to 24 hours. As the size in hours of the sliding window is an important factor in determining the results from a given model, values from 3 to 24 hours were investigated. Moreover, the hyperparameters for RF regression, XGBoost and LSTM are optimised using a grid search. As per the standards in the field, the root mean squared (RMSE) and mean absolute errors (MAE) are used to evaluate the models.

Since this work aimed to identify an air quality forecasting method that provides the best performance while using as little energy as possible, the energy usage of both training and prediction processes for each model was calculated. As the CodeCarbon package is well-documented [20], it was chosen to estimate energy consumption in kilowatt-hours (kWh).

3. Results

Table 1: Comparative results for 1 hour prediction of traditional machine learning methods and those reported in [11]. Power ratio is defined by the energy consumption of a model during training or predicting divided by that of the best performing LSTM solution.

Model	Hyperparameters	Sliding window size	MAE	RMSE	Power ratio (training)	Power ratio (predicting)
Linear regression	N/A	3 hours	0.239	0.579	1941	88
	N/A	12 hours	0.239	0.581	2951	27
	N/A	19 hours	**0.235**	**0.574**	**2840**	**126**
	N/A	24 hours	0.237	0.577	1268	31
Random forest	estimators: 40, max_depth: 7	3 hours	0.316	0.583	137	49
	estimators: 45, max_depth: 6	12 hours	0.332	0.603	67	47
	estimators: 20, max_depth:6	24 hours	0.35	0.612	2	2
XGBoost	estimators: 100, max_depth: 4, learning rate: 0.1	3 hours	0.327	0.602	278	13
	estimators: 100, max_depth: 2, learning rate: 0.1	12 hours	0.361	0.623	358	15
	estimators: 95, max_depth: 2, learning rate: 0.1	24 hours	0.372	0.630	79	28
LSTM (optimizer: adam, loss: ‘mae’)	units: 39, learning rate: 0.001, batch size: 24	3 hours	0.487	0.821	1.9	0.2
	units: 42, learning rate: 0.001, batch size: 24	12 hours	0.423	0.684	0.9	0.8
	units: 42, learning rate: 0.001, batch size: 24	19 hours	0.398	0.649	1	1
	units: 45, learning rate: 0.001, batch size: 24	24 hours	0.439	0.710	1.4	0.9
Linear regression [11]	N/A	3 hours	0.333	0.579	/	/
Random forest [11]	Not specified	3 hours	0.331	0.591	/	/
XGBoost [11]	Not specified	3 hours	0.345	0.617	/	/
LSTM [11]	Not specified	3 hours	0.292	**0.574**	/	/

In terms of performance, i.e., MAE and RMSE, experiments show that, among the accessed methods for 1 hour predictions, Linear Regression (LR) outperforms the others and LSTM, the best method reported by [11]. One should note they used a much shorter sliding window, i.e., 3 hours. Although a 24-hour window was expected to capture potential daily patterns, 19-hour proved optimal. This shorter window may have helped to prevent overfitting. Concerning energy needs, the DL approach is the worst, consuming 2840 and 126 times more energy than LR for training and predicting, resp. (see Table 1 and Figure 2). Whereas the other models are not as greedy, they are still less sustainable than LR.

Table 2: Predictions using linear regression, with a 19-hour sliding window, for a variety of prediction horizon lengths.

Prediction Horizon	MAE	RMSE
T + 1	0.235	0.574
T + 3	0.788	1.322
T + 6	1.292	1.997
T + 9	1.584	2.432
T + 12	2.120	3.228
T + 15	2.376	3.613
T + 18	2.673	4.111
T + 21	2.923	4.529
T + 24	3.219	4.981

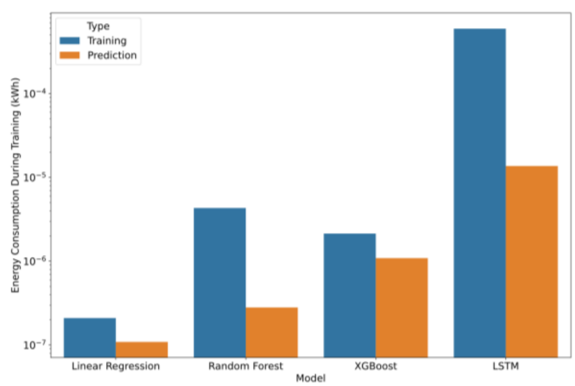

Fig. 2: Comparison of energy consumption - both training and predicting - for each approach using best performing parameters.

Among all methods, linear regression also performs best for longer prediction horizons (data not shown). Moreover, Table 2 reveals that for this model, MAE and RMSE increase linearly with the length of the prediction horizon. These MAE values can be compared to the widths of $PM_{2.5}$ bands used by the UK government to inform public health advice. The narrowest of these bands is 4 μgm^{-3}, suggesting that predictions up to 24 hours may support decision making.

4. Conclusion

Despite the adoption of deep learning solutions in many application areas, this study suggests that linear regression is particularly appropriate to predict air pollution levels. Not only does this approach outperform its competitors in terms of MAE and RMSE, but also it consumes the least energy by a significant margin for both training and predicting. In addition, its predictions for a horizon of up to 24 hours are expected to support decision making to reduce particularly harmful human exposure. Although further investigations should be undertaken, this study supports the aspirations that AI-based solutions are sustainable, affordable, and effective, and that their energy needs must be considered during development.

References

[1] Bell ML, Davis DL, Fletcher T. A retrospective assessment of mortality from the London smog episode of 1952: the role of influenza and pollution. Environ Health Perspect. 2004 Jan;112(1):6-8.

[2] Brimblecombe, P. (2006). The clean air act after 50 years. Weather, 61(11), 311-314.

[3] Dajnak, D., Evangelopoulos, D., Kitwiroon, N., Beevers, S. D., & Walton, H. (2021). London health burden of current air pollution and future health benefits of mayoral air quality policies. City Hall.

[4] Mayor of London, Inner London Ultra Low Emission Zone – One Year Report (February 2023). Peer reviewed by Dr G. Fuller, Imperial College.

[5] Méndez, M., Merayo, M.G. & Núñez, M. Machine learning algorithms to forecast air quality: a survey. Artif Intell Rev (2023).

[6] Kumar, K., Pande, B.P. Air pollution prediction with machine learning: a case study of Indian cities. Int. J. Environ. Sci. Technol. 20, 5333–5348 (2023).

[7] Castelli M, Clemente FM, Popoviˇc A, Silva S, Vanneschi L (2020) A machine learning approach to predict air quality in California. Complexity 2020(8049504):1–23.

[8] Gopalakrishnan V (2021) Hyperlocal air quality prediction using machine learning. Towards data science. https://towardsdatascience.com/hyperlocal-air-quality-prediction-using-machine-learning-ed3a661b9a71

[9] Zaini, N.A., Ean, L.W., Ahmed, A.N. and Malek, M.A., 2022. A systematic literature review of deep learning neural network for time series air quality forecasting. Environmental Science and Pollution Research, pp.1-33.

[10] Zhang, B., Rong, Y., Yong, R., Qin, D., Li, M., Zou, G. and Pan, J., 2022. Deep learning for air pollutant concentration prediction: A review. Atmospheric Environment, p.119347.

[11] A. Utku and U. Can., "Deep Learning Based Air Quality Prediction: A Case Study for London", *Turkish Journal of Nature and Science*, vol. 11, no. 4, pp. 126-134, Dec. 2022.

[12] David A. Wood, Trend decomposition aids forecasts of air particulate matter (PM2.5) assisted by machine and deep learning without recourse to exogenous data, Atmospheric Pollution Research, Volume 13, Issue 3, 2022.

[13] Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., & Pineau, J. (2020). Towards the systematic reporting of the energy and carbon footprints of machine learning. The Journal of Machine Learning Research, 21(1), 10039-10081.

[14] Lucía Bouza, Aurelie Bugeau, and Loic Lannelongue (2023) How to estimate carbon footprint when training deep learning models? A guide and review, Environmental Research Communications, Volume 5, Number 11.

[15] Luccioni S, Viguier S and Ligozat A-L 2022 Estimating the carbon footprint of bloom, a 176b parameter language model arXiv:2211.02001

[16] Mehonic, A., Kenyon, A.J. Brain-inspired computing needs a master plan. Nature 604, 255–260 (2022). https://doi.org/10.1038/s41586-021-04362-W

[17] R.Toews,"Deep Learning'scarbon emissions problem," Forbes, https://www.forbes.com/sites/robtoews/2020/06/17/deep-learnings-climate-change-problem/?sh=4479178c6b43 (accessed Dec. 19, 2023).

[18] Rentschler, Jun; Leonova, Nadia. 2022. Air Pollution and Poverty: PM2.5 Exposure in 211 Countries and Territories. Policy Research Working Paper;10005, World Bank, Washington, DC.

[19] S. Nobell. (2020) Pollution PM2.5 data London 2019 Jan to Apr [Online]. Available: https://www.kaggle.com/datasets/siddharthnobell/pollution-pm25-data-london-2019-jan-to-apr

[20] *Codecarbon* (2020) *CodeCarbon - CodeCarbon 2.3.2 documentation*. Available at: https://mlco2.github.io/codecarbon/index.html (Accessed: 15 December 2023).