https://doi.org/10.7250/CONECT.2026.040

MODELING AND PREDICTION OF PM2.5 USING DIFFERENT DEEP LEARNING TECHNIQUES: A COMPARATIVE ANALYSIS

Murad A. YAGHI[1*], Huthaifa AL-OMARI[2], Salem ALEMAISHAT[3]

[1] *Data Science and Artificial Intelligence Department, Al-Hussein Technical University, King Hussein Business Park, Amman, Jordan*

[2,3] *Computer Science Department, Al-Hussein Technical University, King Hussein Business Park, Amman, Jordan*

[*] ***Corresponding author.** Email address: murad.yaghi@htu.edu.jo*

Abstract – Accurate forecasting of fine particulate matter ($PM_{2.5}$) concentrations is essential for public health protection and environmental management. While deep learning approaches have shown promise for $PM_{2.5}$ prediction, consistent comparative evaluations under standardized conditions remain limited. This paper presents a comprehensive performance analysis of three widely adopted sequence forecasting architectures: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Neural Basis Expansion Analysis for Time Series (N-BEATS), for short-term $PM_{2.5}$ concentration prediction. All models were trained and evaluated under identical experimental conditions using a 70/15/15 train-validation-test split across four input window lengths (5, 10, 15, and 20 hours). Our results demonstrate that extending the temporal context systematically improves predictive accuracy, though marginal gains diminish beyond 15–20 time steps. N-BEATS consistently outperformed both recurrent architectures across all metrics and window sizes, achieving a root mean square error (RMSE) of 55.28 $\mu g/m^3$ and R^2 of 0.5116 at the 20-step horizon representing a 14 % reduction in RMSE compared to GRU and superior reproduction of hazardous pollution episodes. Additionally, N-BEATS exhibited substantially faster training times and near-constant computational costs across window sizes, whereas LSTM and GRU training times increased three-fold. The feed-forward, block-based architecture of N-BEATS enables highly parallelized computation while its interpretable basis-function decomposition better captures abrupt nonlinear patterns inherent in pollution dynamics. These findings establish N-BEATS as a computationally efficient and accurate choice for real-time air quality forecasting systems, while highlighting the importance of standardized evaluation protocols for advancing atmospheric time-series prediction.

Keywords – GRU; LSTM; N-BEATS; particulate matter; time series forecasting