

A comparative study of stochastic nature of anthropogenic aerosol using SARIMA and LSTM methods

J. Arul Asir^{1*}, H. Johnson Jeyakumar¹ and C. P. Anil Kumar²

¹P.G and Research Department of Physics, Pope's College, Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli – 627012, Tamil Nadu, India.

²Equatorial Geophysical Research Laboratory, Indian Institute of Geomagnetism, Tirunelveli- 627011, Tamil Nadu, India.

*Corresponding author: arulasirphy@popescollege.edu.in

ABSTRACT

The troposphere is a dynamic region characterized by intricate interactions among natural and anthropogenic processes. Among these, aerosol loading, particularly fine particulate matter (PM_{2.5}), poses significant challenges in understanding the greenhouse effect and complicates air quality assessment due to its spatiotemporal variability. This study presents a comparative stochastic modelling framework for forecasting anthropogenic aerosol concentrations using both parametric and non-parametric approaches over Alandur (13.0° N, 80.10° E), Tamil Nadu, India. Specifically, a Seasonal Autoregressive Integrated Moving Average (SARIMA) model and a Long Short-Term Memory (LSTM) neural network, were employed to forecast PM_{2.5} levels using observational data from 2023 and 2024. SARIMA, a classical statistical model, effectively captured regular patterns and seasonal trends. In contrast, the LSTM model, a deep learning approach, demonstrated enhanced performance in modelling non-linear dependencies and temporal dynamics. Model evaluations based on Root Mean Square Error (RMSE), indicated that LSTM consistently outperformed SARIMA, underlining its effectiveness in capturing the stochastic and complex behaviour of aerosol concentrations. These findings underscore the potential of integrating advanced machine learning techniques with traditional time series models to enhance the accuracy of air quality forecasting and environmental risk mitigation strategies.

Keywords: Aerosol, PM_{2.5} levels, Alandur (Tamil Nadu), Health impact, Seasonal ARIMA, Machine learning, Neural network.

INTRODUCTION

Particulate matter (PM) in the atmosphere, originating from both natural and anthropogenic sources, consists of fine solid or liquid particles suspended in air. These aerosols are either directly emitted or formed via secondary chemical reactions of precursor gases. As a result, PM is a heterogeneous mixture with varied chemical compositions and size distributions, often exhibiting a core-shell morphology where secondary constituents condense around a primary core. Based on aerodynamic diameter, PM is commonly classified into fine particles (PM_{2.5}; ≤ 2.5 μm), ultrafine particles (PM_{0.1}; ≤ 0.1 μm), and nanoparticles (PM_{0.05}; ≤ 0.05 μm or ≤ 50 nm). This spatial-temporal distribution of PM concentration and composition is driven by complex interactions between emission sources, atmospheric transport mechanisms, in-situ chemical transformations, and removal processes such as dry and wet deposition. In addition to primary and secondary sources, background particulate matter representing PM levels in the absence of local anthropological influence, also contributes to the overall spatiotemporal variability (Yang et al., 2011; Celoz and Dabek-Zlotorzynska, 2011; Seinfeld and Pandis, 2016; Guevara, 2016; Donahue et al., 2016). Alongside emission sources and transformation mechanisms, atmospheric conditions play a crucial role in modulating PM behaviour. Meteorological parameters such as wind speed, temperature, and relative humidity, significantly influence the dispersion, accumulation, and chemical transformation of aerosols (Zhang et al., 2018). Recent studies (Bhattarai et al., 2024; Pan et al., 2024; Yim et al., 2024), highlight the growing concern over fine particulate matter (PM_{2.5}) due to its combined impact on air quality degradation and climate alteration. With an

atmospheric lifetime ranging from days to weeks, PM_{2.5} is capable of long-range transport, leading to trans-boundary pollution and contributing to both regional and global environmental challenges. Meteorological influences and public health consequences are linked to PM exposure.

Andreae (2009) established a global correlation between CCN (Cloud Condensation Nuclei) concentrations and Aerosol Optical Thickness (AOT), showing that polluted marine and continental regions have nearly ten times higher CCN concentrations and about five times greater AOT than remote areas. This highlights the crucial role of aerosol optical properties in influencing cloud microphysics. Complementing these findings, Zhang et al. (2012) investigated CCN activation dynamics in Tianjin, China, using a thermal-gradient diffusion chamber (Droplet Measurement Technologies - CCNC), alongside aerosol size and composition analysis. Their findings confirmed that larger or more hygroscopic particles activate at lower supersaturation, while smaller or less hygroscopic particles, require higher supersaturation. Additionally, meteorological factors significantly influence PM_{2.5} concentrations. Zhao et al. (2017), using the WRF-Chem model, found that increased anthropogenic aerosol loading in China, significantly enhanced cloud droplet formation and the liquefaction process in cumulus clouds. Following these results, Cheung et al. (2019) studied PM_{2.5} over northern Taiwan from April 2017 to March 2018. In this study, continental air masses exhibited higher hygroscopicity (K-values), whereas local air masses showed higher Condensation Nuclei and CCN concentrations but lower K-values. These differences were reflected in the PM_{2.5} chemical composition. The study also found that new particle formation enhanced