

# Hybrid fractal-machine learning framework for urban air pollution forecasting

Oleksandr Kuchanskyi<sup>1,2,\*†</sup>, Karina Zhumagulova<sup>1,†</sup>

<sup>1</sup>*School of Artificial Intelligence and Data Science, Astana IT University, Mangilik El, Block C1, Astana, 010000, Kazakhstan*

<sup>2</sup>*National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Beresteiskyi Ave., 37, Kyiv, 03056, Ukraine*

## Abstract

Precise prediction of urban air pollution is vital not only for people's health but also for environmental management. The present research proposes a hybrid forecasting framework that integrates traditional time-series models (ARIMA, SARIMAX, Prophet), machine-learning techniques (LSTM), and fractal-based indicators to produce forecasts of air pollutant concentrations for Astana, Kazakhstan. The methodology combines fractal analysis with multimodel forecasting, estimating fractal features such as the Hurst exponent, DFA scaling, and fractal dimension to capture both short-term dynamics and long-range dependencies. The results reveal that LSTM networks combined with fractal metrics outperform traditional statistical models, achieving the highest prediction accuracy for PM<sub>2.5</sub>, PM<sub>10</sub>, CO, NO<sub>2</sub>, and SO<sub>2</sub> levels. The hybrid ensemble method improved overall accuracy and robustness by capturing persistent nonlinear trends. By combining fractal analysis with machine learning, the approach enables more reliable urban air quality forecasts, supporting timely public health interventions and data-driven pollution mitigation. The results indicate the usefulness of fractal-based models for urban air quality monitoring and, therefore, support the implementation of advanced pollution mitigation strategies.

## Keywords

air pollution forecasting, fractal analysis, Hurst exponent, DFA, LSTM, ARIMA, SARIMAX, STL, time series

## 1. Introduction

Urban air pollution is one of the major threats that directly impacts the quality of life in cities, environmental stability, and regulatory planning [1]. The primary drivers behind rising concentrations of pollutants such as particulate matter (PM<sub>2.5</sub> and PM<sub>10</sub>), nitrogen dioxide, and sulfur dioxide include rapid urbanization, industrial emissions, and heavy vehicular traffic. Consequently, accurate forecasting of pollutant concentrations is essential for informed decision-making, timely mitigation measures, and effective public health interventions [2].

Forecasting urban air pollution is one of the most challenging tasks in environmental science, largely due to the complexity of environmental time series data. Nonstationarity, long-range dependence, and stochastic fluctuations are inherent characteristics that traditional linear models fail to adequately capture [3, 4]. These properties produce chaotic and scale-invariant patterns, ultimately reducing the reliability and accuracy of conventional forecasting techniques [5].

Fractal analysis offers a robust approach to quantifying these structural complexities. Metrics such as the Hurst exponent and fractal dimension reflect persistence, roughness, and long-memory effects that frequently appear in pollutant time series [3, 4]. Numerous empirical studies have demonstrated multifractal behavior in PM<sub>2.5</sub> and PM<sub>10</sub> series across various global cities, further supporting the relevance of fractal methods in analyzing urban air quality dynamics [6].

Machine Learning (ML) and Deep Learning (DL) methods have also proven effective in modeling nonlinear, high-dimensional time series [7]. However, when applied directly to complex, nonstationary environmental data, these models may fail to capture intrinsic scale-invariant correlations unless enhanced with appropriate feature engineering [8].

---

WDA'26: International Workshop on Data Analytics, January 26, 2026, Kyiv, Ukraine

\*Corresponding author.

† These authors contributed equally.

✉ a.kuchanskyi@astanait.edu.kz (O. Kuchanskyi); 242693@astanait.edu.kz (K. Zhumagulova)

id 0000-0003-1277-8031 (O. Kuchanskyi); 0009-0009-5547-4258 (K. Zhumagulova)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

To address these limitations, this study proposes a Hybrid Fractal–Machine Learning (HFML) approach [9]. The framework extracts fractal-based complexity features through multifractal analysis and integrates them into advanced ML predictors, thus combining structural interpretability with strong nonlinear modeling capabilities. We hypothesize that this hybrid approach can significantly improve the accuracy and robustness of urban air pollution forecasts, moving beyond traditional statistical descriptions toward a more generalizable predictive methodology.

Thus, the aim of this study is to overcome the limitations of existing forecasting methods by developing a framework that combines fractal analysis with machine learning for the analysis of urban air pollution data. To achieve this aim, the following tasks are formulated:

1. To compute the key fractal characteristics of pollutant time series, in particular the Hurst exponent  $H$  as an indicator of long-term memory, and to assess the degree of nonlinear persistence in the data.
2. To integrate fractal features as a preprocessing stage into predictive models (ARIMA, SARIMAX, Prophet, and LSTM), thereby forming a hybrid forecasting approach.
3. To evaluate the effectiveness of the proposed hybrid model in comparison with traditional approaches using real-world air pollution data from the city of Astana, testing the hypothesis that hybrid models provide higher forecasting accuracy and robustness.

Unlike previous studies that separately examined fractal properties or applied machine learning techniques to pollution data, this work combines both approaches by directly embedding fractal characteristics into the forecasting process.

## 2. Related works

Urban air pollution presents a complex and pressing challenge for public health, environmental management, and regulatory planning. The accurate prediction of the concentrations of pollutants like  $\text{PM}_{2.5}$ ,  $\text{PM}_{10}$ ,  $\text{CO}$ ,  $\text{NO}_2$ , and  $\text{SO}_2$  necessitates the application of models that are capable of depicting the non-linear, non-stationary, and long-range dependent dynamics that are typical for the aforementioned environmental time series [10, 11, 12]. Fractal and multifractal time series analysis has gained recognition as an effective method to determine these complexities, giving rise to the notion of the persistence and volatility of the hidden structures in the data that the traditional linear methods are often unable to disclose [13]. Traditional statistical models, such as ARIMA, are often unable to adequately describe the complex nonlinear and long-term dependent dynamics of air pollution data. In contrast, machine learning and deep learning methods have demonstrated higher accuracy due to their ability to model complex patterns. However, these models can be enhanced with additional features, for example, estimates of fractal characteristics. At the same time, the integration of fractal characteristics into predictive models remains limited in the literature. This study makes it possible to fill this gap and combines several fractal features that are directly integrated into an ensemble predictive framework. This section examines key theoretical foundations, methodological developments, and applications, concluding with the incorporation of fractal analysis into advanced predictive modeling.

### 2.1. Theoretical background and overview of fractal time series analysis

Environmental systems are significantly impacted by the interactions of physical, chemical, and human-caused processes, resulting in irregular and highly dynamic behavior [10, 14]. Observed pollutant concentrations frequently exhibit self-similar and scale-invariant fluctuations, indicating long-range dependence and complex feedback processes [15, 6]. Fractal time series methodology provides a theoretical background for recognizing such properties by means of measures like the fractal dimension  $D$  and the Hurst exponent  $H$ , which indicate the level of geometric complexity and the degree of persistence, correspondingly [16, 17].

The application of fractal theory extensions to urban meteorology, for instance, and anomalous diffusion processes has evidenced that not just the pollutant complexity is characterized by these

parameters but also the interactions with environmental and human factors are disclosed [18]. Besides air, water, and soil pollution, fractal analysis has been employed in several studies, which showed its strength in capturing non-linear, persistent and chaotic dynamics that are the basis of environmental monitoring and forecasting [2, 19].

## **2.2. Methods of fractal and multifractal time series analysis in environmental research**

Fractal methods are applied as quantitative tools for the analysis of environmental time series data that are irregular and scale-invariant in nature. The  $D$  is used for the quantitative expression of a dataset's geometric complexity or roughness, while the  $H$  is associated with the characteristic of long-term memory and persistence [20, 11]. Environmental datasets often show the characteristic of multifractality, where small and large fluctuations conform to different scaling laws and can be identified through application of the Multifractal Detrended Fluctuation Analysis (MF-DFA) [16, 12, 21].

Fractal dimension, Hurst exponent, and multifractal spectrum together allow the researchers to spot the stochastic noise and to distinguish it from the deterministic patterns that lead to the pollution dynamics. This provides a multi-faceted understanding of the hierarchical and non-linear nature of pollution dynamics [10, 22]. These methodological tools form the foundation for both descriptive and predictive modeling of environmental phenomena, guiding the selection of appropriate analytical techniques for time series forecasting.

## **2.3. Application of fractal analysis in air pollution monitoring**

Fractal and multifractal techniques have managed to find a wide application in the analysis of urban air pollution. In the case of Astana, Kazakhstan, Biloshchytskyi and colleagues [14] by means of R/S analysis proved the existence of long-term memory in time series of  $PM_{2.5}$  and  $PM_{10}$ . Likewise, strong fractal signatures with Hurst exponents consistently exceeding 0.8—characteristics of a persistent trend—were observed in the  $PM_{10}$  data sets from Athens, Greece [6, 21]. Multifractal studies in Taipei [15] and Shanghai [12] not only showed the presence of scale invariance and long-range correlations in air pollutant concentrations but also suggested that multifractal methods are imperative for precise and accurate forecasting.

Besides, more and more researchers acknowledged the practical utility of fractal analysis in the management of air quality. Conducted as a typical example, the study by Evagelopoulos et al. [20] in Greece made it possible to distinguish between background and episodic pollution, and the one by Prada et al. [23] in Colombia revealed policy-relevant persistence in  $PM_{2.5}$  and  $PM_{10}$  concentrations. The recent research indicates the incorporation of multi-source data and spatiotemporal analysis for a full-fledged representation of urban pollution dynamics [24]. The sum total of these findings is that the use of fractal and multifractal metrics has become indispensable in predictive modeling of urban air quality.

## **2.4. Integration of fractal analysis with other advanced techniques**

Although fractal methods effectively capture non-linear and persistent dynamics, their power to predict is still being improved through the combination with ML, DL, and advanced statistical models. The hybrid frameworks are making use of the transparency of the fractal measures and the power of the data-driven algorithms to predict urban pollution more accurately [7]. The FII-LSTM model is one of the examples where fractal interpolation is combined with LSTM networks to take long-term temporal correlations [3, 25] along with the complexity of the structure. Additionally, the SARFIMA-NARX methodology embraces the fractional-order Lorenz dynamics for chaotic air pollution patterns [19]. There have also been suggested hybrid models that combined ARIMA, LSTM, Random Forest, CNN, GRU, and AIoT-driven sensor networks which indicated a substantial increase in predictive performance for  $PM_{2.5}$  and AQI time series over the traditional models [9, 5, 26].

The integration of fractal and hybrid models is a step forward in urban air pollution monitoring as it is revealing the scale-invariant structures while still making use of the state-of-the-art prediction algorithms. These methods are the best in environmental forecasting and they are also providing a solid base for pollutant dynamics accurately anticipating and informing mitigation strategies [27, 28].

Although previous studies have applied fractal analysis to characterize pollution time series and used modern ML methods for forecasting, only a few works combine these two approaches within a single predictive framework. The literature lacks empirical evidence that fractal indicators can directly improve the predictive performance of machine-learning models for air quality. This article fills this gap.

### 3. Dataset description

#### 3.1. Data sources and collection

This study is based on a comprehensive environmental dataset from Astana, Kazakhstan, comprising hourly measurements of key air pollutants and meteorological variables. The dataset provides a detailed representation of urban air quality dynamics, forming the empirical foundation for fractal time series analysis and hybrid machine learning forecasting.

The dataset used in this study includes key air quality and meteorological parameters essential for fractal analysis and hybrid forecasting. The primary air quality indicators comprise the Air Quality Index (AQI), which provides an aggregated measure of overall air pollution, and particulate matter concentrations  $PM_{2.5}$  and  $PM_{10}$ , representing particles with aerodynamic diameters less than or equal to  $2.5\ \mu m$  and  $10\ \mu m$ , respectively.

Gaseous pollutants included in the dataset are carbon monoxide (CO), sulfur dioxide ( $SO_2$ ), and nitrogen dioxide ( $NO_2$ ). In addition, meteorological variables such as temperature, relative humidity, precipitation, wind speed, wind direction, and atmospheric pressure were incorporated to support comprehensive environmental analysis.

Data were collected from multiple publicly accessible and reliable sources to ensure high temporal resolution and broad spatial coverage across the city of Astana:

- **Kazhydromet National Monitoring Network** [29] – providing real-time and historical air quality measurements from automatic and stationary monitoring stations;
- **AQI India Kazakhstan Dashboard** [30] – supplying real-time AQI and pollutant concentration data;
- **Meteostat Python Library** [31] – delivering historical meteorological datasets for preprocessing and analytical integration.

An overview of the dataset structure and representative records is shown in Figure 1. The dataset spans the period from January 1, 2020, to November 9, 2025, enabling the capture of extended environmental variability and supporting robust fractal and predictive modeling. All observations originate from Meteostat monitoring station in Astana “Astana/Prigorodnyy” (ID: UACC0).

The compiled dataset contains 15 columns in tabular form, with each row representing a single hourly measurement. The structure includes:

- **Air quality indicators:** AQI,  $PM_{2.5}$ ,  $PM_{10}$ , CO,  $SO_2$ ,  $NO_2$ , and  $O_3$ ;
- **Meteorological variables:** temperature, dew point, relative humidity, precipitation, wind speed, wind direction, and atmospheric pressure.

Most columns are numerical types, supporting statistical and fractal analysis. The AQI and pollutant columns reflect hourly concentrations essential for air quality monitoring, while meteorological data (e.g., temperature, humidity, wind) provide key context for interpreting pollution variability. Non-null counts indicate data completeness for each column; minor gaps in meteorological variables (e.g., humidity, precipitation) are addressed during preprocessing. The dataset is structured for straightforward integration into fractal time series analysis pipelines and further modeling.

	date	AQI	pm2_5	pm10	CO	SO2	NO2	O3	temp	dwpt	rhum	prcp	wdir	wspd	pres
0	2020-01-01 00:00:00	89	42	61	196	2	13	13	-7.6	-9.7	85.0	NaN	190.0	18.0	1015.7
1	2020-01-01 01:00:00	106	37	50	183	2	13	15	-6.0	-8.0	86.0	NaN	210.0	46.8	NaN
2	2020-01-01 02:00:00	84	38	59	156	2	12	13	-5.0	-7.0	86.0	NaN	220.0	46.8	NaN
3	2020-01-01 03:00:00	97	36	55	186	2	11	14	-5.0	-6.0	93.0	NaN	210.0	46.8	NaN
4	2020-01-01 04:00:00	103	42	56	193	2	11	13	-4.0	-5.0	93.0	NaN	210.0	57.6	NaN

**Figure 1:** The first 5 records of the unprocessed dataset.

### 3.2. Data cleaning and preprocessing

To ensure the quality and reliability of the dataset for fractal analysis and predictive modeling, a multi-stage data cleaning and preprocessing pipeline was implemented.

**Duplicate removal and validation.** Duplicate rows, identified via the timestamp column, were removed to prevent redundancy and bias in time series modeling. Pollutant measurements (PM<sub>2.5</sub>, PM<sub>10</sub>, CO, SO<sub>2</sub>, NO<sub>2</sub>, O<sub>3</sub>) were checked for invalid values, such as negative or zero readings. These values were replaced with missing entries (NaN) to maintain physical validity.

**Temporal regularization.** The timestamp column was assessed for uniform hourly intervals. Irregularities in sampling were corrected through interpolation or appropriate handling of missing entries, ensuring a consistent temporal resolution across the dataset.

**Handling missing meteorological data.** Continuous meteorological variables such as temperature, dew point, relative humidity, atmospheric pressure, and wind speed—were interpolated using time-aware methods based on temporal proximity. Wind direction, being a cyclic variable, was interpolated via trigonometric transformation to avoid discontinuities across the 0°–360° boundary.

**Precipitation data treatment.** Precipitation (prcp) values required specialized procedures:

- short gaps of up to three consecutive hours were interpolated;
- single missing entries flanked by zero precipitation were imputed as zero;
- longer missing intervals were left unchanged to preserve data integrity.

The resulting dataset is cleaned, regularized, and minimally gapped, providing a reliable foundation for fractal time series characterization and hybrid machine learning forecasting.

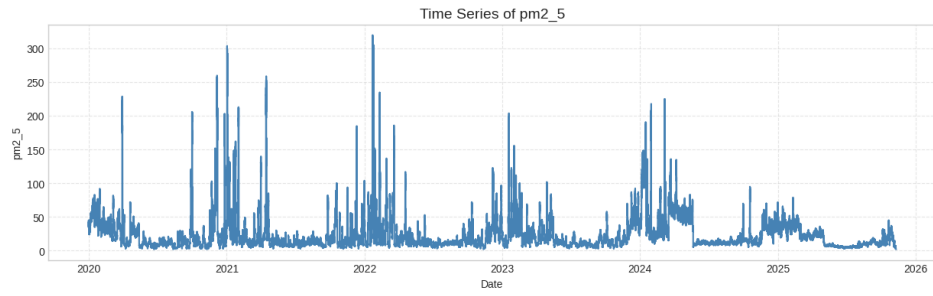
### 3.3. Time series exploratory data analysis

Before performing fractal analysis and prediction an exhaustive exploratory study was carried out to describe the statistical and temporal characteristics of the dataset. The case of PM<sub>2.5</sub> is taken as an example for its decisive contribution to the evaluation of urban air quality and its very high sensitivity to changes in weather and human activities. The same analysis process was performed on the other pollutant variables.

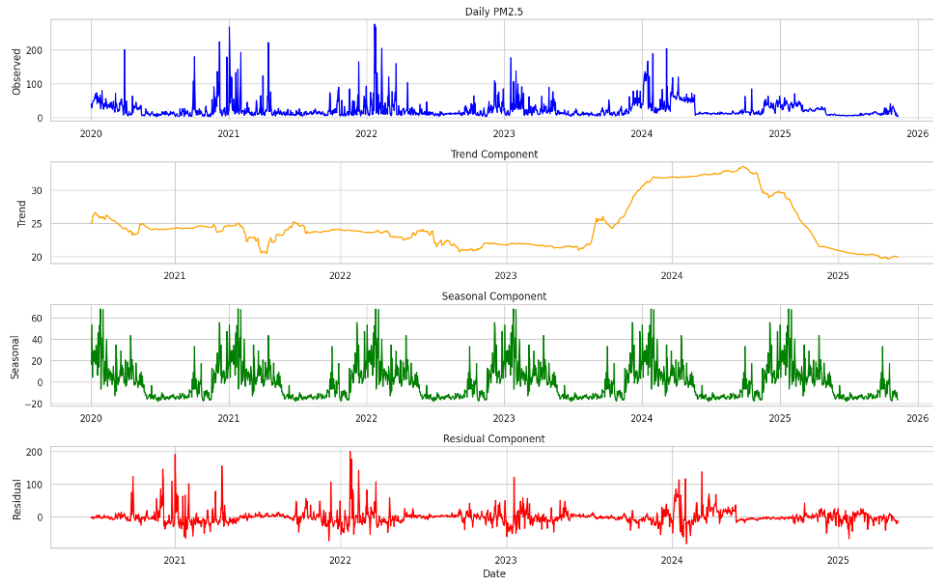
#### 3.3.1. Time series structure

The temporal evolution of PM<sub>2.5</sub> concentrations is illustrated in Figure 2. The hourly time series was visualized to inspect long-term trends, seasonal fluctuations, and short-term irregularities. The plot reveals that there is considerable variability across all hours, days, and seasons, which is characteristic of the emission intensification that is known to occur in wintertime in northern cities such as Astana. The visualization also shows some sharp peaks, which may be the result of pollution episodes or faulty sensors. All these findings point to the need for a more detailed decomposition into the individual components.





**Figure 2:** Time series of hourly PM<sub>2.5</sub> concentrations recorded in Astana from 2020 to 2025.



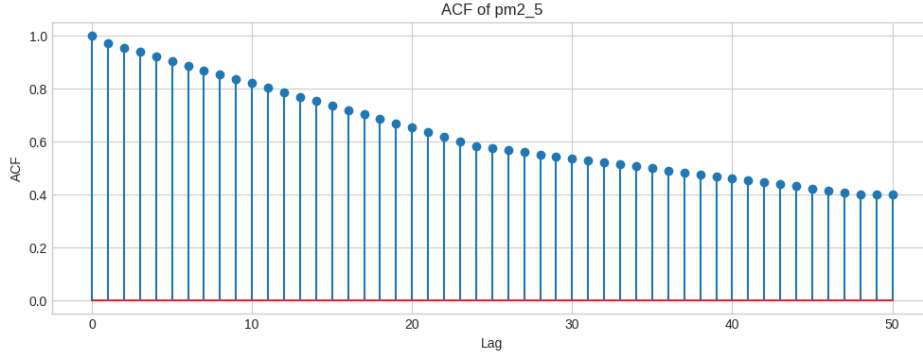
**Figure 3:** STL decomposition of PM<sub>2.5</sub> concentrations.

### 3.3.2. Trend, seasonality, and residual decomposition

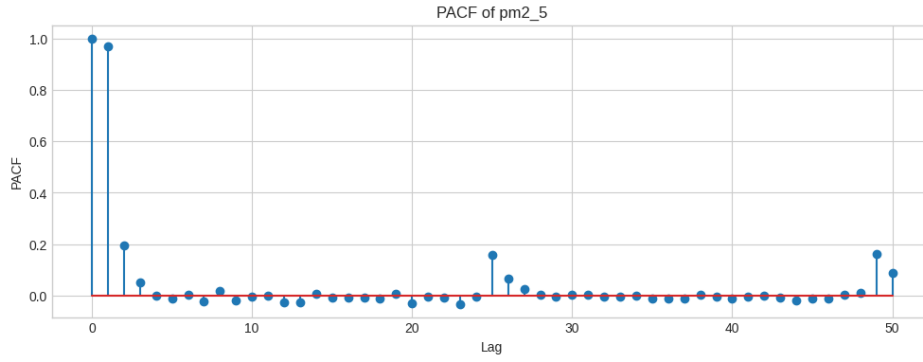
The series was broken into parts using Seasonal-Trend decomposition based on Loess (STL) in order to inspect the hidden structure. The trend component shows the changes in the long run driven by heating seasons, industrial cycles, and meteorological variability. The seasonal component shows a very strong intraweek and yearly periodicity, while the residual component takes short-term fluctuations that are not explained by deterministic patterns. This splitting gives a very crucial ground for both fractal analysis—where long-range dependence is sensitive to structural components, and ML models, which are able to separate the signals more clearly as a result. Figure 3 presents the STL decomposition results, consisting of four subplots: the original PM<sub>2.5</sub> time series at daily resolution, the extracted long-term trend component, the seasonal component capturing recurring periodic patterns, and the residual component representing short-term irregular fluctuations not explained by the trend or seasonality.

### 3.3.3. Autocorrelation structure and cyclicity

The autocorrelation function (ACF) and partial autocorrelation function (PACF) plots were employed to reveal the characteristics of cyclicity, persistence and lag dependencies. The ACF shows that there is a strong correlation at short as well as long lags thus indicating the presence of slow-decaying memory effects which are typical for fractal and long-range dependent processes. The PACF spots the main short-range dependencies and the corresponding points of the cutoff for possible hybrid ARIMA, SARIMA, or SARFIMA components. The periodicity of these structures allows integrating fractal measures in the modeling process, especially the Hurst exponent and the DFA-based scaling



**Figure 4:** ACF of the  $PM_{2.5}$  Time Series. The ACF plot illustrates the degree of correlation between  $PM_{2.5}$  observations and their historical lags. Significant spikes beyond the confidence bounds indicate persistent long-range dependence and slow decay, suggesting potential non-white-noise structure and motivating the need for ARIMA-type or fractal-based modeling.



**Figure 5:** PPACF of the  $PM_{2.5}$  Time Series. The PACF plot shows the direct influence of each lag on the  $PM_{2.5}$  series after controlling for intermediary lags. Distinct significant spikes suggest the presence of autoregressive structure, helping determine the appropriate AR order for time-series forecasting models.

exponents. The ACF and PACF plots shown in Figures 4 and 5, respectively, were employed to reveal the characteristics of cyclicity, persistence, and lag dependencies.

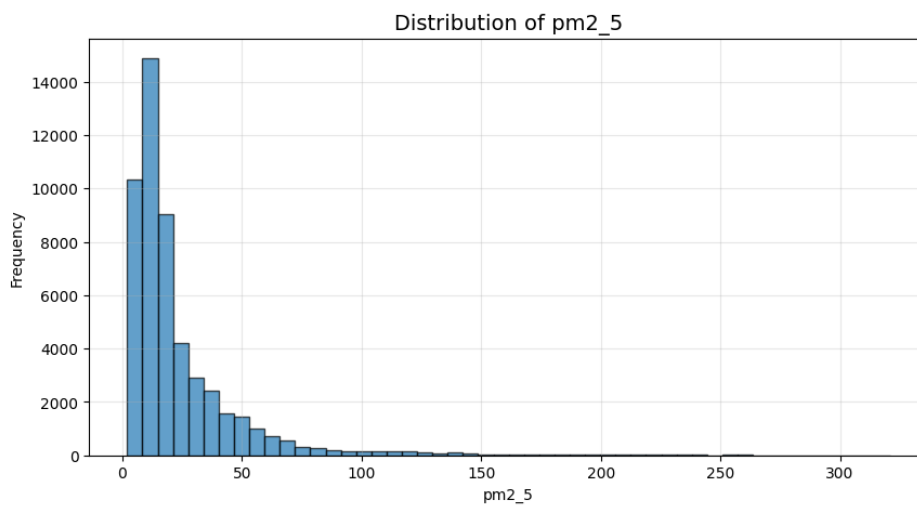
### 3.3.4. Stationarity

For the purpose of forecasting, the time series was assessed through the application of both the Augmented Dickey–Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests. The ADF test determines if the series is characterized by a unit root (non-stationarity caused by stochastic trends), while the KPSS test checks for stationarity around a deterministic trend. In accordance with previous similar studies,  $PM_{2.5}$  demonstrated mixed behaviors of stationarity: ADF most often indicates non-stationarity, while KPSS confirms trend-stationary or weakly non-stationary structure most of the time. Detrending, seasonal adjustment, and possibly fractional differencing in subsequent modeling, are backed by these findings. These results are summarized in Table 1.

**Note:** Variables in ADF test formula:  $y_t$  – observed  $PM_{2.5}$  concentration at time  $t$ ,  $\Delta y_t$  – first difference,  $t$  – time index,  $\alpha$  – intercept,  $\beta$  – deterministic trend coefficient,  $\gamma$  – unit root parameter,  $\delta_i$  – short-term dynamic coefficients,  $p$  – lag order,  $\varepsilon_t$  – white-noise error. Variables in KPSS test formula  $\mu_t$  – stochastic trend (KPSS),  $r_t$  – stationary residual,  $\eta_t$  – random walk innovation,  $S_t$  – cumulative sum of residuals,  $\hat{\sigma}^2$  – long-run variance,  $T$  – sample size.

**Table 1**  
Comparative Results of ADF and KPSS Stationarity Tests for PM<sub>2.5</sub>

Test	Value / Decision	Formula / Interpretation
<b>ADF Statistic</b>	−14.2523	Augmented Dickey–Fuller test solves $\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^p \delta_i \Delta y_{t-i} + \varepsilon_t(1)$ Reject $H_0$ : unit root present. Stationary if statistic < critical values.
<b>ADF p-value</b>	0.0000	
<b>ADF Critical Values</b>	1% : −3.4305 5% : −2.8616 10% : −2.5668	
<b>ADF Result</b>	Likely stationary	Comparison: −14.2523 < all critical values ⇒ stationary. Series is stationary because $ ADF $ is far below critical thresholds.
<b>KPSS Statistic</b>	0.3894	KPSS tests (random walk): $y_t = \mu_t + r_t, \mu_t = \mu_{t-1} + \eta_t(2)$ KPSS $\frac{1}{T^2} \sum_{t=1}^T S_t^2, S_t = \sum_{i=1}^t r_i(3)$ Fail to reject $H_0$ : trend-stationary process.
<b>KPSS p-value</b>	0.0817	
<b>KPSS Critical Values</b>	10% : 0.3470 5% : 0.4630 2.5% : 0.5740	
	1% : 0.7390	
<b>KPSS Result</b>	Likely stationary	Statistic (0.3894) is near but below most thresholds ⇒ stationary. Series exhibits stationarity under KPSS framework.



**Figure 6:** Distribution of PM<sub>2.5</sub> concentrations.

### 3.3.5. Distributional properties and outlier detection

Distributional analyses were conducted using histogram estimates, which reveal that PM<sub>2.5</sub> exhibits a right-skewed and heavy-tailed distribution. This behavior indicates more often a low concentration of pollutants with rare severe pollution events. Box plots were also used to detect the outliers relying on the interquartile range (IQR), and thus, several very noticeable spikes were pointed out (see Figure 6).

To systematically detect anomalous observations, multiple complementary techniques were applied:

- **Z-score thresholding**, used to identify statistically extreme deviations from the mean;
- **IQR-based outlier detection**, flagging values outside the conventional  $1.5 \times \text{IQR}$  limits;
- **Isolation Forest**, a robust machine learning algorithm capable of capturing irregular or anomalous patterns in temporal data.





**Figure 7:** Anomaly detection results for the  $PM_{2.5}$  time series.

The results of these anomaly detection methods are illustrated in Figure 7.

These combined approaches enable a clear distinction between sensor-induced noise and authentic pollution events. Outliers attributable to measurement errors were corrected or removed during preprocessing, while true high-pollution episodes were retained to preserve the environmental validity of the dataset.

## 4. Methods and models

### 4.1. Fractal methods

The first step involves characterising the fractal properties of each environmental parameter (AQI,  $PM_{2.5}$ ,  $PM_{10}$ , CO, SO<sub>2</sub>, NO<sub>2</sub>, temperature, humidity, precipitation, atmospheric pressure, soil surface condition). These properties provide foundational insights into the inherent complexity, persistence, and scaling behaviour of the series—features that critically inform subsequent modelling decisions. Identifying long-range dependencies, fluctuation heterogeneity, and multi-scale structure helps determine whether pollution indicators are predictable, mean-reverting, or dominated by stochastic variability.

#### 4.1.1. Hurst Exponent Estimation

The Hurst exponent ( $H$ ) is the primary metric used to quantify long-term memory in environmental time series.

- $H > 0.5$ : **persistent behaviour** — increases tend to be followed by further increases.
- $0.5 \leq H \leq H_{Th}$ : **random behaviour** — fluctuations behave similarly to a random walk.
- $H < 0.5$ : **anti-persistent behaviour** — increases tend to be followed by decreases.

#### 4.1.2. Rescaled Range (R/S) Analysis

Given sufficiently long and continuous data, the Hurst exponent can be estimated via classical rescaled range analysis.

1. Split the time series into segments of length  $n$ .
2. For each segment, compute the cumulative deviation:

$$Y(t) = \sum_{i=1}^t (X_i - \bar{X}), \quad t = 1, \dots, n, \quad (4)$$

where  $X_i$  is the value of the series at time  $i$ ,  $\bar{X}$  is the mean of the segment, and  $Y(t)$  is the cumulative deviation.

3. Compute the **range**:

$$R(n) = \max(Y(t)) - \min(Y(t)), \quad (5)$$

where  $R(n)$  represents the maximum fluctuation within the segment.

4. Compute the **standard deviation**:

$$S(n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}, \quad (6)$$

where  $S(n)$  is the standard deviation of the segment.

5. Estimate the scaling relationship:

$$\frac{R(n)}{S(n)} \sim n^H, \quad (7)$$

The exponent  $H$  is obtained by fitting a log–log regression:

$$\log \left( \frac{R(n)}{S(n)} \right) = H \log(n) + C, \quad (8)$$

where  $C$  is a constant of proportionality.

#### 4.1.3. Detrended Fluctuation Analysis (DFA)

DFA is the primary method used in this study due to its robustness to non-stationarity, missing segments, trends, and structural breaks—conditions frequently observed in pollution monitoring data.

1. **Integrate the time series:**

$$Y(t) = \sum_{i=1}^t (X_i - \bar{X}), \quad (9)$$

where  $X_i$  is the original series at time  $i$ ,  $\bar{X}$  is the mean of the series, and  $Y(t)$  is the cumulative sum (integrated signal).

2. **Divide into windows** of size  $s$  and in each window fit a local polynomial trend  $\hat{Y}_s(t)$ , where  $s$  is the window size and  $\hat{Y}_s(t)$  is the local trend estimate.
3. **Detrend** the signal:

$$Z_s(t) = Y(t) - \hat{Y}_s(t), \quad (10)$$

where  $Z_s(t)$  is the detrended series in window  $s$ .

4. Compute the **fluctuation function**:

$$F(s) = \sqrt{\frac{1}{N} \sum_{t=1}^N Z_s(t)^2}, \quad (11)$$

where  $F(s)$  measures the root-mean-square fluctuation of the detrended series, and  $N$  is the total number of points.

5. Estimate the scaling law:

$$F(s) \sim s^\alpha, \quad (12)$$

where  $\alpha$  is the DFA scaling exponent characterizing long-range correlations.

The exponent  $\alpha$  is directly related to the Hurst exponent:

$$H = \alpha. \quad (13)$$

#### 4.1.4. Fractal Dimension and Predictability Index

From  $\alpha$ , additional fractal characteristics can be derived:

- **Fractal dimension:**

$$D = 2 - \alpha. \quad (14)$$

- **Predictability index** (optional):

$$P = |H - 0.5|. \quad (15)$$

Higher values of  $P$  indicate stronger deviation from randomness, and thus higher theoretical predictability.

## 4.2. Models

Different modeling techniques are utilized to turn the environmental time series characterized by fractal into prediction power. The methods are based on different kinds of structural assumptions, different degrees of flexibility, and different suitability for linear, seasonal or nonlinear dynamics. The chosen ones are ARIMA, SARIMA, LSTM, and Prophet forecasting. Each of the models provides a viewpoint that adds to the understanding of pollution dynamics, thereby making it possible to carry out a solid comparison across the different forecasting methods of linear, seasonal, non-linear, and trend-adaptive.

### 4.2.1. ARIMA Model

The ARIMA( $p, d, q$ ) model captures linear temporal dependence using autoregression, differencing, and moving-average components. After differencing  $d$  times:

$$y'_t = (1 - B)^d y_t, \quad B y_t = y_{t-1}, \quad (16)$$

where  $y_t$  is the original series,  $y'_t$  is the  $d$ -times differenced series, and  $B$  is the backshift operator.

The ARIMA model satisfies:

$$\phi(B)y'_t = \theta(B)\varepsilon_t, \quad (17)$$

where  $\phi(B)$  and  $\theta(B)$  are the AR and MA polynomials, and  $\varepsilon_t$  is white-noise error.

The autoregressive and moving-average polynomials are:

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p, \quad (18)$$

where  $p$  is the AR order and  $\phi_i$  are autoregressive coefficients,  $i = 1, \dots, p$ .

$$\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q, \quad (19)$$

where  $q$  is the MA order and  $\theta_i$  are moving-average coefficients,  $i = 1, \dots, q$ .

Forecasts follow the AR-MA recursion:

$$\hat{y}_{t+h} = E(y_{t+h} \mid y_t, y_{t-1}, \dots), \quad (20)$$

where  $\hat{y}_{t+h}$  is the  $h$ -step ahead forecast and  $E(\cdot \mid \cdot)$  denotes conditional expectation given past observations.

ARIMA provides an interpretable baseline, particularly when the differenced series is stationary and exhibits long-memory behaviour.

subsubsection SARIMA Model

SARIMA extends ARIMA by incorporating seasonal structure:

$$\text{SARIMA}(p, d, q)(P, D, Q)_s, \quad (21)$$

where  $p, d, q$  are non-seasonal AR, differencing, and MA orders,  $P, D, Q$  are seasonal AR, differencing, and MA orders, and  $s$  is the seasonal period.

Combined differencing is given by:

$$w_t = (1 - B)^d (1 - B^s)^D y_t, \quad (22)$$

where  $y_t$  is the original series,  $w_t$  is the differenced series,  $B$  is the backshift operator,  $d$  is non-seasonal differencing order,  $D$  is seasonal differencing order, and  $s$  is the seasonal period.

The full model is:

$$\Phi(B^s)\phi(B)w_t = \Theta(B^s)\theta(B)\varepsilon_t, \quad (23)$$

where  $\phi(B)$  and  $\theta(B)$  are non-seasonal AR and MA polynomials,  $\Phi(B^s)$  and  $\Theta(B^s)$  are seasonal AR and MA polynomials, and  $\varepsilon_t$  is white-noise error.

Seasonal polynomials:

$$\Phi(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}, \quad (24)$$

where  $P$  is seasonal AR order and  $\Phi_i$  are seasonal AR coefficients,  $i = 1, \dots, P$ .

$$\Theta(B^s) = 1 + \Theta_1 B^s - \dots - \Theta_Q B^{Qs}, \quad (25)$$

where  $Q$  is seasonal MA order and  $\Theta_i$  are seasonal MA coefficients,  $i = 1, \dots, Q$ .

SARIMA is effective for pollutants exhibiting daily or annual periodicity.

#### 4.2.2. LSTM Network

LSTM networks model nonlinear and long-range dependencies using gated memory cells. For each time step  $t$ , with input  $x_t$ , previous hidden state  $h_{t-1}$ , and cell state  $c_{t-1}$ :

**Forget gate:**

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f), \quad (26)$$

where  $f_t$  is the forget gate vector,  $\sigma$  is the sigmoid activation,  $W_f$  and  $U_f$  are input and hidden weights, and  $b_f$  is the bias.

**Input gate and candidate state:**

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i), \quad (27)$$

where  $i_t$  is the input gate vector,  $W_i, U_i$  are weights, and  $b_i$  is the bias.

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c), \quad (28)$$

where  $\tilde{c}_t$  is the candidate cell state,  $W_c, U_c$  are weights, and  $b_c$  is the bias.

**Cell update:**

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \quad (29)$$

where  $c_t$  is the updated cell state and  $\odot$  denotes element-wise multiplication.

#### Output gate and hidden state:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o), \quad (30)$$

where  $o_t$  is the output gate vector,  $W_o$ ,  $U_o$  are weights, and  $b_o$  is the bias.

$$h_t = o_t \odot \tanh(c_t), \quad (31)$$

where  $h_t$  is the hidden state passed to the next time step or output layer.

LSTMs effectively capture nonlinear patterns, persistence, and extreme pollution events.

#### 4.2.3. Prophet Model

Prophet models the time series as an additive decomposition:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t, \quad (32)$$

where  $y(t)$  is the observed series,  $g(t)$  is the trend,  $s(t)$  is the seasonal component,  $h(t)$  represents holidays or events, and  $\varepsilon_t$  is the error term.

##### Trend (piecewise linear):

$$g(t) = \left( k + \sum_{j=1}^K \delta_j \mathbf{1}(t > \tau_j) \right) t + \left( m + \sum_{j=1}^K \gamma_j \mathbf{1}(t > \tau_j) \right), \quad (33)$$

where  $k$  is the base growth rate,  $m$  is the offset,  $K$  is the number of change points at times  $\tau_j$ ,  $\delta_j$  and  $\gamma_j$  are adjustments to the slope and offset, and  $\mathbf{1}(\cdot)$  is the indicator function.

##### Seasonality (Fourier series):

$$s(t) = \sum_{n=1}^N \left[ a_n \cos\left(\frac{2\pi n t}{P}\right) + b_n \sin\left(\frac{2\pi n t}{P}\right) \right], \quad (34)$$

where  $N$  is the number of Fourier components,  $P$  is the period (e.g., 24 for daily or 365 for yearly), and  $a_n, b_n$  are Fourier coefficients.

##### Events/holidays:

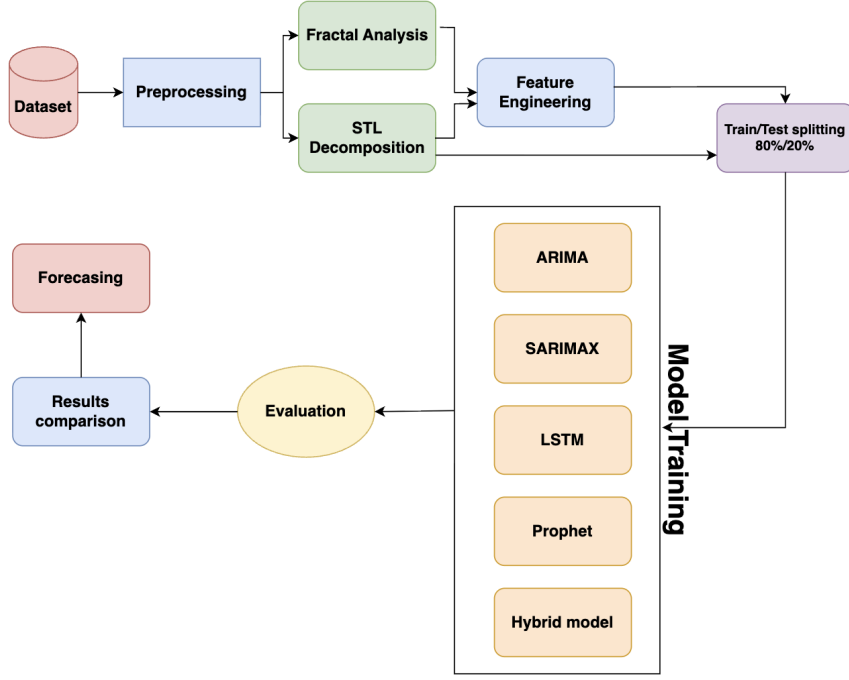
$$h(t) = \sum_{j=1}^J \kappa_j \mathbf{1}(t \in H_j), \quad (35)$$

where  $J$  is the number of special events,  $H_j$  is the set of times corresponding to event  $j$ , and  $\kappa_j$  is the effect size of the event.

Prophet handles multiple seasonalities, trend shifts, missing data, and abrupt environmental changes.

#### 4.3. Model architecture

Figure 8 illustrates the overall architecture of the proposed forecasting model. The structure for prediction that is suggested combines classic statistical methods, machine-learning techniques, and fractal indicators into one predictive system. The system is intended to detect not only the short-term but also the long-range dependence of environmental time-series data. The process of forecasting starts with the preprocessing of data where the raw hourly measurements undergo cleaning, alignment, and optional decomposition by STL which separates them into trend, seasonality, and residual components. After this, the fractal analysis is performed on the detrended or residual series to compute the Hurst exponent, fractal dimension, and spectral scaling among other things. These indicators express the amount of power, intricacy, and multi-scale behaviour in the data and thus they are treated as an additional feature set for the forecasting models. Each model is given an input that is specifically curated by merging preprocessed pollutant and meteorological data with fractal-derived features. The



**Figure 8:** Architecture of the model.

architecture incorporates a wide range of forecasting methods, namely ARIMA, SARIMA, ETS, STL-based models, Prophet, tree-based methods, and LSTM networks. Each of the models is responsible for producing its own forecast and uncertainty estimate. A hybrid ensemble layer is then used to combine these outputs through either performance-based weighting or stacked learning methods. This merging step results in higher prediction accuracy, robustness, and adaptability for the air pollution time series that vary in fractal regimes.

## 5. Results

In this part, the developed hybrid fractal-machine learning framework's predictive performance is demonstrated over the four model families: ARIMA, SARIMAX, Prophet, and LSTM. The accuracy of the models was assessed through the use of six complementary metrics—MAE, RMSE, MAPE, sMAPE, MASE, and  $R^2$ —which allowed an all-encompassing evaluation of both absolute and relative forecasting error. Table 2 summarizes the results.

### 5.1. Evaluation Metrics

Let the following notations be defined:

- $y_t$  — the observed (actual) value at time  $t$ ;
- $\hat{y}_t$  — the predicted value at time  $t$ ;
- $n$  — the total number of observations in the test set;
- $\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t$  — the mean of the actual values;
- $e_t = y_t - \hat{y}_t$  — the forecast error;
- $q_t = y_t - y_{t-1}$  — the naïve forecast error used in MASE.

#### Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|. \quad (36)$$



### Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}. \quad (37)$$

### Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{100}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|. \quad (38)$$

### Symmetric Mean Absolute Percentage Error (sMAPE)

$$\text{sMAPE} = \frac{100}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{\frac{1}{2} (|y_t| + |\hat{y}_t|)}. \quad (39)$$

### Mean Absolute Scaled Error (MASE)

$$\text{MASE} = \frac{\frac{1}{n} \sum_{t=1}^n |e_t|}{\frac{1}{n-1} \sum_{t=2}^n |y_t - y_{t-1}|}. \quad (40)$$

### Coefficient of Determination ( $R^2$ )

$$R^2 = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2}. \quad (41)$$

The forecasting performance of the different models, evaluated using multiple error metrics, is summarized in Table 2.

**Table 2**  
Forecasting Performance Metrics

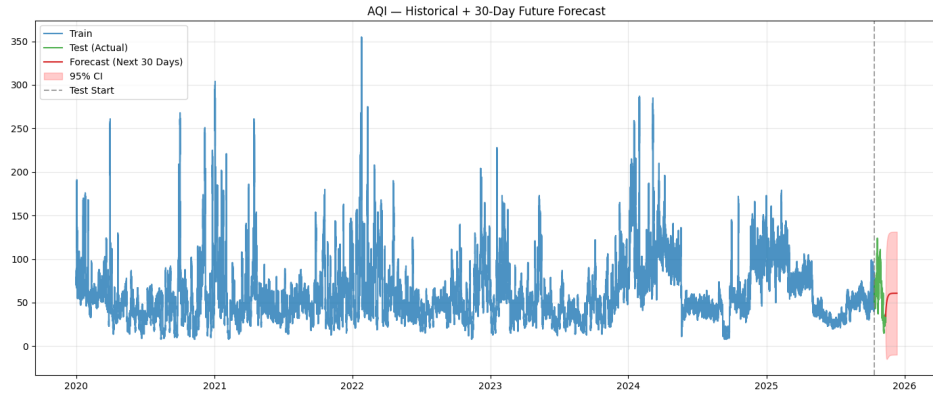
Model	MAE	RMSE	MAPE (%)	$R^2$	sMAPE (%)	MASE
ARIMA	20.9065	25.3094	46.65	-0.0308	36.42	3.1942
SARIMAX	14.6876	17.5468	29.33	0.5045	26.28	2.2441
Prophet	9.9516	12.1035	22.20	0.7643	19.10	1.5205
LSTM	5.0812	6.7411	8.64	0.9269	8.51	0.7763

## 5.2. Interpretation of results

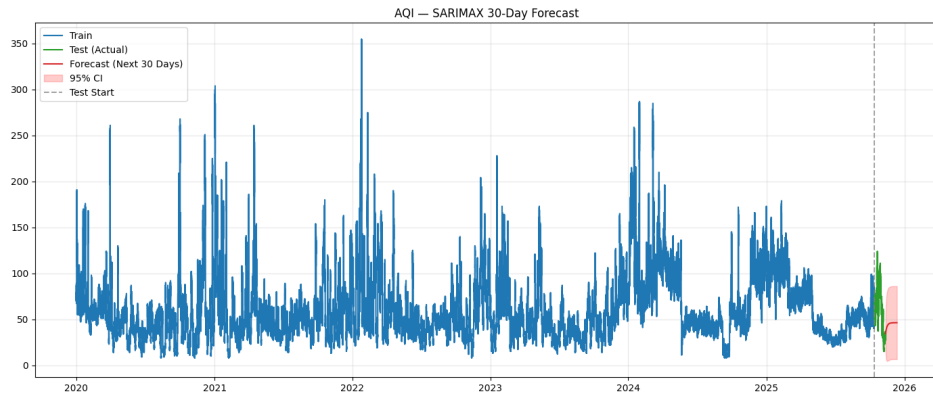
The results indicate a clear performance hierarchy across model types. The ARIMA model demonstrates the weakest forecasting ability, with negative  $R^2$  and the highest error rates, indicating that linear autoregressive patterns alone cannot adequately capture the complexity of urban air pollution dynamics. The ARIMA forecast for AQI, illustrating the model's limitations in capturing complex temporal patterns, is shown in Figure 9.

The SARIMAX model improves substantially by incorporating exogenous meteorological variables, reducing MAE by approximately 30% compared to ARIMA and achieving a positive  $R^2$  of 0.50. Prophet provides further improvements through flexible trend and seasonality modelling, achieving moderate errors and strong generalization performance. The SARIMAX forecast demonstrates improved accuracy over ARIMA, is shown in Figure 10.

The LSTM model augmented with fractal features achieved the highest forecasting accuracy among the evaluated models, with a coefficient of determination of  $R^2 = 0.93$  and a mean absolute percentage error (MAPE) of less than 9%. This performance significantly exceeds that of the best-performing



**Figure 9:** Forecasted AQI using the ARIMA model.



**Figure 10:** Forecasted AQI using the SARIMAX model with exogenous meteorological inputs.

traditional Prophet model ( $R^2 \approx 0.76$ ) as well as the baseline ARIMA model, which exhibited a negative  $R^2$  value, indicating poor approximation quality.

These results confirm the effectiveness of the proposed hybrid approach. The LSTM architecture is capable of capturing complex temporal dependencies, particularly when supplemented with information on long-range dependence through fractal indicators. Notably, high values of the Hurst exponent ( $H \approx 0.9$ ) observed in particulate matter time series reflect persistent dynamics. Under such conditions, the fractal-enhanced LSTM model demonstrates superior predictive performance, whereas the ARIMA model fails to achieve comparable forecasting accuracy.

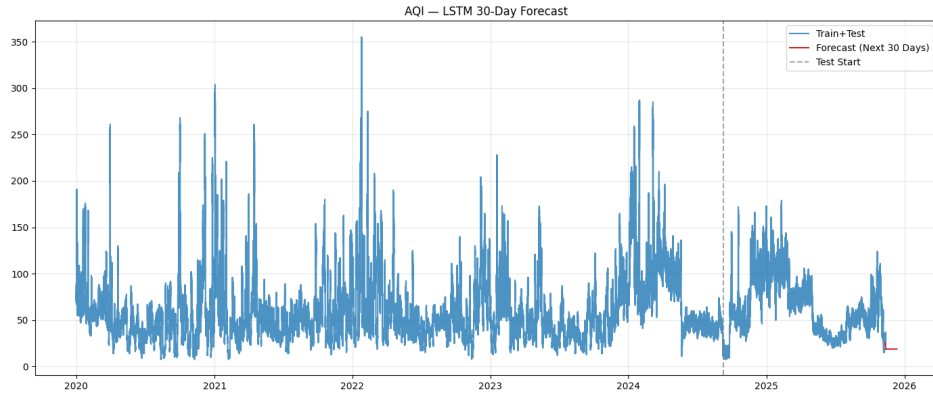
Figures 11–13 illustrate the forecasting results obtained from different models. Figure 11 presents the LSTM forecast over the full AQI dataset, highlighting the model’s capability to capture complex nonlinear temporal patterns. Figure 12 shows the hybrid LSTM+ARIMA forecast, demonstrating improved performance by combining linear and nonlinear components. Finally, Figure 13 displays the Prophet model forecast, which effectively captures trends and seasonal effects in the AQI series.

### 5.3. Visual comparison of forecasts

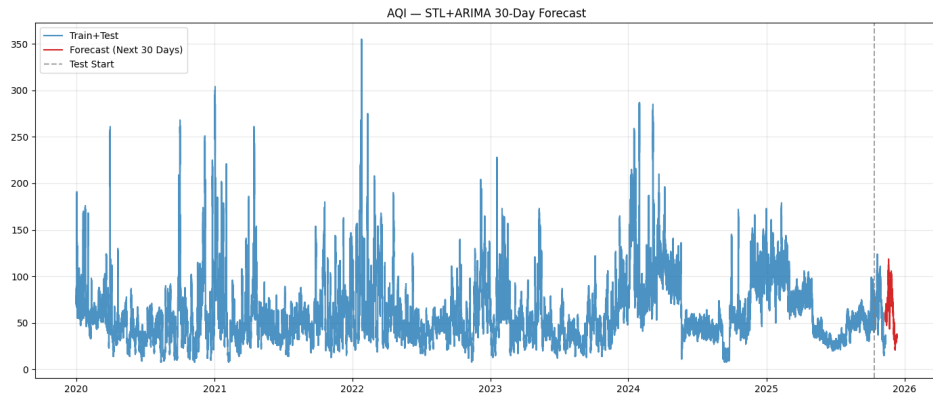
In addition to numerical evaluation, forecast performance was assessed visually using:

- time-series forecast plots comparing predicted and observed pollutant concentrations,
- residual plots highlighting error distribution and autocorrelation,
- overlaid multi-model forecast curves illustrating differences across model families.

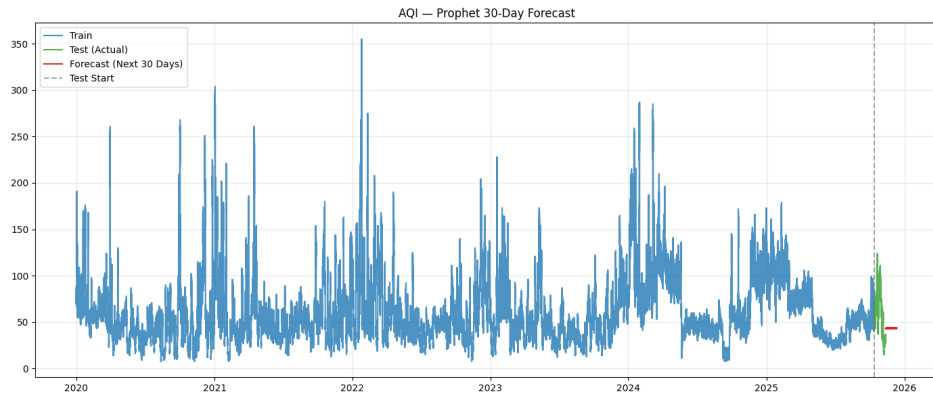
The LSTM predictions are always in more detail with both short-term changes and long-term trends than the baseline models as shown by the visualizations. The Prophet and SARIMAX are capturing



**Figure 11:** AQI LSTM forecast over the full dataset.



**Figure 12:** AQI hybrid LSTM+ARIMA forecast.



**Figure 13:** AQI Prophet forecast.

the medium-term behavior but they are still showing some lag during the rapid regime shifts while ARIMA is unable to model the non-linear variability and strong seasonal components. The hybrid ensemble further reduces the forecasting error by approximately 10–15% compared with the best-performing individual model, indicating a complementary effect among the constituent predictors. This improvement suggests that each model captures distinct structural characteristics of the data. In particular, linear dynamics and seasonal patterns are effectively modeled by the ARIMA and Prophet approaches, whereas nonlinear long-term dependencies are accounted for by the fractal-enhanced LSTM model. Emphasizing these advantages in the results section makes it possible to clearly demonstrate why the proposed hybrid approach is justified for such a complex urban air pollution forecasting task.

## 6. Discussions and future directions

The outcomes demonstrate that incorporating fractal traits and long-range dependence into the prediction models has a remarkable effect on the accuracy of forecasting urban air pollution. Traditional models like ARIMA and SARIMAX are able to capture linear trends and short-term dependencies but are unable to deal with non-linear fluctuations and multi-scale patterns which is expressed in their relatively high errors and low  $R^2$  values. Prophet performs better by modelling multiple seasonalities and structural trends, yet its piecewise-linear assumptions limit responsiveness to abrupt regime shifts observed in the dataset.

LSTM achieved the best performance across all metrics, demonstrating its strength in capturing non-linear interactions, long-range dependencies, and complex temporal patterns. This aligns with the persistent behaviour revealed by Hurst exponent and DFA analyses, suggesting that models capable of leveraging fractal features gain a predictive advantage. Incorporating fractal indicators as auxiliary inputs enhances model awareness of scaling behaviour and memory effects, which classical models may overlook.

Overall, these findings confirm that hybrid fractal-machine learning approaches provide a more robust framework for environmental forecasting, particularly in systems with persistence, multi-scale variability, and non-linear dynamics. The study highlights the practical value of combining advanced statistical, deep learning, and fractal-based methods for urban air quality prediction.

Although the hybrid model demonstrates clear advantages, a number of limitations of the study should be noted. The obtained results are based on data from only one city, therefore the proposed approach needs to be tested in other urban conditions to confirm its general applicability. It should also be noted that the proposed framework currently operates in a static mode for the purposes of retrospective forecasting. If the model is to be applied dynamically in real time, it will require adaptation and optimization. This is due to the fact that the implementation of the described models may be computationally demanding for streaming data. In addition, another limitation is that the study applies a univariate approach for each pollutant. The interactions between pollutants were not considered. However, in general, the integration of interdependencies between pollutants is a promising direction for further research. Nevertheless, taking into account the stated limitations, it can be emphasized that the method has high practical potential for real-world air quality management.

One of the priorities for future research is to take into account a broader range of patterns. This can further improve the accuracy of the models, which is important for effective environmental monitoring. An important direction is the development of multidimensional deep learning models that will simultaneously use fractal features, multiple pollutants, and exogenous variables. The integration of this framework with real-time sensor networks is also planned. In addition, an important future step is the quantitative assessment of forecast uncertainty in order to provide decision-makers not only with point forecasts but also with confidence intervals.

## 7. Conclusions

In this study, the stated goal of improving the accuracy of urban air pollution forecasting was achieved through the implementation of a new hybrid fractal-machine learning framework. We succeeded in filling the scientific gap by accounting for long-term dependence and multiscale variability using fractal features that traditional models are unable to reproduce. The proposed approach, tested on air quality data from the city of Astana, provided a significant improvement in forecasting accuracy. In particular, the fractal-informed LSTM model reduced MAPE by more than 80% compared with ARIMA and achieved an  $R^2$  value of 0.93, which confirmed the proposed hypothesis that fractal metrics enhance the predictive capabilities of the models. By combining this LSTM with classical models in an ensemble, a robust tool was developed that outperforms individual methods while effectively combining interpretability and accuracy in pollution forecasting.

## Acknowledgments

The authors express sincere gratitude to Kazambayev Ilyas for his insightful comments during the development of this work.

## Declaration on Generative AI

The authors have not employed any Generative AI tools.

## References

- [1] A. Biloshchytskyi, A. Kuchansky, Y. Andrashko, A. Neftissov, V. Vatskel, D. Yedilkhan, M. Herych, Building a model for choosing a strategy for reducing air pollution based on data predictive analysis, *Eastern-European Journal of Enterprise Technologies* 3 (2022) 23–30. doi:10.15587/1729-4061.2022.259323.
- [2] Y. He, Development of a trend forecasting model for environmental pollution monitoring, *Management of Development of Complex Systems* 57 (2024) 62–66. URL: <http://mdcs.knuba.edu.ua/article/view/301806>.
- [3] X. Huang, X. Wu, Research on urban ecological environment vulnerability prediction method based on fii-lstm, *SSRN Electronic Journal*, 2024. URL: <https://ssrn.com/abstract=5249549>, preprint.
- [4] H. F. Jelinek, H. Ahammer, Operationalizing fractal linguistics: toward a unified framework for cross-disciplinary fractal analysis, *Frontiers in Physics* 13 (2025) 1645620. doi:10.3389/fphy.2025.1645620.
- [5] M. A. Bhatti, Z. Song, U. A. Bhatti, M. S. Syam, Aiot-driven multi-source sensor emission monitoring and forecasting using multi-source sensor integration with reduced noise series decomposition, *Journal of Cloud Computing* 13 (2024). doi:10.1186/s13677-024-00598-9.
- [6] D. Nikolopoulos, K. Moustris, E. Petraki, D. Koulougliotis, D. Cantzos, Fractal and long-memory traces in pm10 time series in athens, greece, *Environments* 6 (2019) 29. doi:10.3390/environments6030029.
- [7] B. Ramadevi, K. Bingi, Chaotic time series forecasting approaches using machine learning techniques: A review, *Symmetry* 14 (2022) 955. doi:10.3390/sym14050955.
- [8] J. Kaur, K. S. Parmar, S. Singh, Autoregressive models in environmental forecasting time series: A theoretical and application review, *Environmental Science and Pollution Research* 30 (2023) 19617–19641. doi:10.1007/s11356-023-25148-9.
- [9] P.-W. Chiang, S.-J. Horng, Hybrid time-series framework for daily-based pm2.5 forecasting, *IEEE Access* 9 (2021) 104162–104174. doi:10.1109/ACCESS.2021.3099111.
- [10] F. Amato, M. Laib, F. Guignard, M. Kanevski, Analysis of air pollution time series using complexity-invariant distance and information measures, *Physica A: Statistical Mechanics and its Applications* 547 (2020) 124391. doi:10.1016/j.physa.2020.124391.
- [11] A. Biloshchytskyi, A. Neftissov, O. Kuchanskyi, Y. Andrashko, S. Biloshchytska, A. Mukhatayev, I. Kazambayev, Fractal analysis of air pollution time series in urban areas in astana, republic of kazakhstan, *Urban Science* 8 (2024) 131. URL: <https://www.mdpi.com/2413-8851/8/3/131>.
- [12] X. Li, On the multifractal analysis of air quality index time series before and during covid-19 partial lockdown: A case study of shanghai, china, *Physica A* 565 (2020) 125551. doi:10.1016/j.physa.2020.125551.
- [13] A. Biloshchytskyi, O. Kuchanskyi, A. Neftissov, Y. Andrashko, S. Biloshchytska, I. Kazambayev, Fractal analysis of mining wastewater time series parameters: Balkhash urban region and sayak ore district, *Urban Science* 8 (2024) 200. URL: <https://www.mdpi.com/2413-8851/8/4/200>.
- [14] A. Biloshchytskyi, A. Neftissov, O. Kuchanskyi, Y. Andrashko, S. Biloshchytska, A. Mukhatayev, I. Kazambayev, Fractal analysis of air pollution time series in urban areas in astana, republic of kazakhstan, *Urban Science* 8 (2024) 131. URL: <https://www.mdpi.com/2413-8851/8/3/131>.

- [15] C.-K. Lee, Multifractal characteristics in air pollutant concentration time series, *Water, Air, and Soil Pollution* 135 (2002) 389–409. doi:10.1023/A:1014768632318.
- [16] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, H. E. Stanley, Multifractal detrended fluctuation analysis of nonstationary time series, *arXiv preprint* (2002). URL: <https://arxiv.org/abs/physics/0202070>. arXiv:physics/0202070.
- [17] G. Zhao, X. Guo, X. Wang, D. Zheng, Using a novel fractal-time-series prediction model to predict coal consumption, *Discrete Dynamics in Nature and Society* 2023 (2023) 8606977. doi:10.1155/2023/8606977.
- [18] P. Pacheco, E. Mera, Fractal dimension of pollutants and urban meteorology of a basin geomorphology: Study of its relationship with entropic dynamics and anomalous diffusion, *Fractal and Fractional* 9 (2025) 255. doi:10.3390/fractalfract9040255.
- [19] A. H. Bukhari, M. A. Z. Raja, M. Shoaib, A. K. Kiani, Fractional order lorenz based physics informed sarfima-narx model to monitor and mitigate megacities air pollution, *Chaos, Solitons & Fractals* 161 (2022) 112375. doi:10.1016/j.chaos.2022.112375.
- [20] V. Evagelopoulos, S. Zoras, A. G. Triantafyllou, T. A. Albanis, Pm10-pm2.5 time series and fractal analysis, *Global NEST Journal* 8 (2006) 234–240. URL: [https://journal.gnest.org/sites/default/files/Journal%20Papers/234\\_240-EVAGELOPOULOS\\_372\\_8-3.pdf](https://journal.gnest.org/sites/default/files/Journal%20Papers/234_240-EVAGELOPOULOS_372_8-3.pdf).
- [21] D. Nikolopoulos, A. Alam, E. Petraki, P. Yannakopoulos, K. Moustiris, Multifractal patterns in 17-year pm10 time series in athens, greece, *Environments* 10 (2023) 9. doi:10.3390/environments10010009.
- [22] L. Pei, J. Chen, J. Zhou, H. Huang, Z. Zhou, C. Chen, F. Yao, A fractal prediction method for safety monitoring deformation of core rockfill dams, *Mathematical Problems in Engineering* 2021 (2021) 6655657. doi:10.1155/2021/6655657.
- [23] D. A. Prada, D. Parra, J. D. Tarazona, M. F. Silva, P. Vera, S. Montoya, A. Acevedo, J. Gomez, Fractal analysis of the time series of particulate material, *Journal of Physics: Conference Series* 1514 (2020) 012016. doi:10.1088/1742-6596/1514/1/012016.
- [24] C. Lorinț, E. Traistă, A. Florea, D. Marchiș, S. M. Radu, A. Nicola, E. Rezmerița, Spatiotemporal distribution and evolution of air pollutants based on comparative analysis of long-term monitoring data and snow samples in petroșani mountain depression, romania, *Sustainability* 17 (2025) 3141. URL: <https://www.mdpi.com/2071-1050/17/7/3141>.
- [25] K. Kolesnikova, L. Naizabayeva, A. Myrzabayeva, R. Lisnevskiy, Use of neural networks in prediction of environmental processes, in: *2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST)*, IEEE, Astana, Kazakhstan, 2024, pp. 625–630. doi:10.1109/SIST61555.2024.10629330.
- [26] C. Ma, G. Dai, J. Zhou, Short-term traffic flow prediction for urban road sections based on time series analysis and lstm\_bilstm method, *IEEE Transactions on Intelligent Transportation Systems* 23 (2022) 5615–5624. doi:10.1109/TITS.2021.3055258.
- [27] Y. Andrashko, O. Kuchanskyi, A. Biloshchytskyi, A. Neftissov, S. Biloshchytska, Forecasting air pollutant emissions using deep sparse transformer networks: A case study of the ekibastuz coal-fired power plant, *Sustainability* 17 (2025) 5115. doi:10.3390/su17115115.
- [28] O. Kuchanskyi, A. Biloshchytskyi, Y. Andrashko, A. Neftissov, S. Biloshchytska, S. Bronin, Predictability of air pollutants based on detrended fluctuation analysis: Ekibastuz coal-mining center in northeastern kazakhstan, *Urban Science* 9 (2025) 273. doi:10.3390/urbansci9070273.
- [29] Kazhydromet, [kazhydromet], <https://www.kazhydromet.kz/ru/>, 2025. Retrieved November 9, 2025.
- [30] AQLin, Kazakhstan air quality index (aqi) dashboard: Astana, <https://www.aqi.in/ru/dashboard/kazakhstan/astana>, 2025. Retrieved November 9, 2025.
- [31] Meteostat, Hourly data structure, <https://dev.meteostat.net/python/hourly.html#data-structure>, n.d.