Air Quality Forecasting Model Validation

Executive Summary & Detailed Performance Report

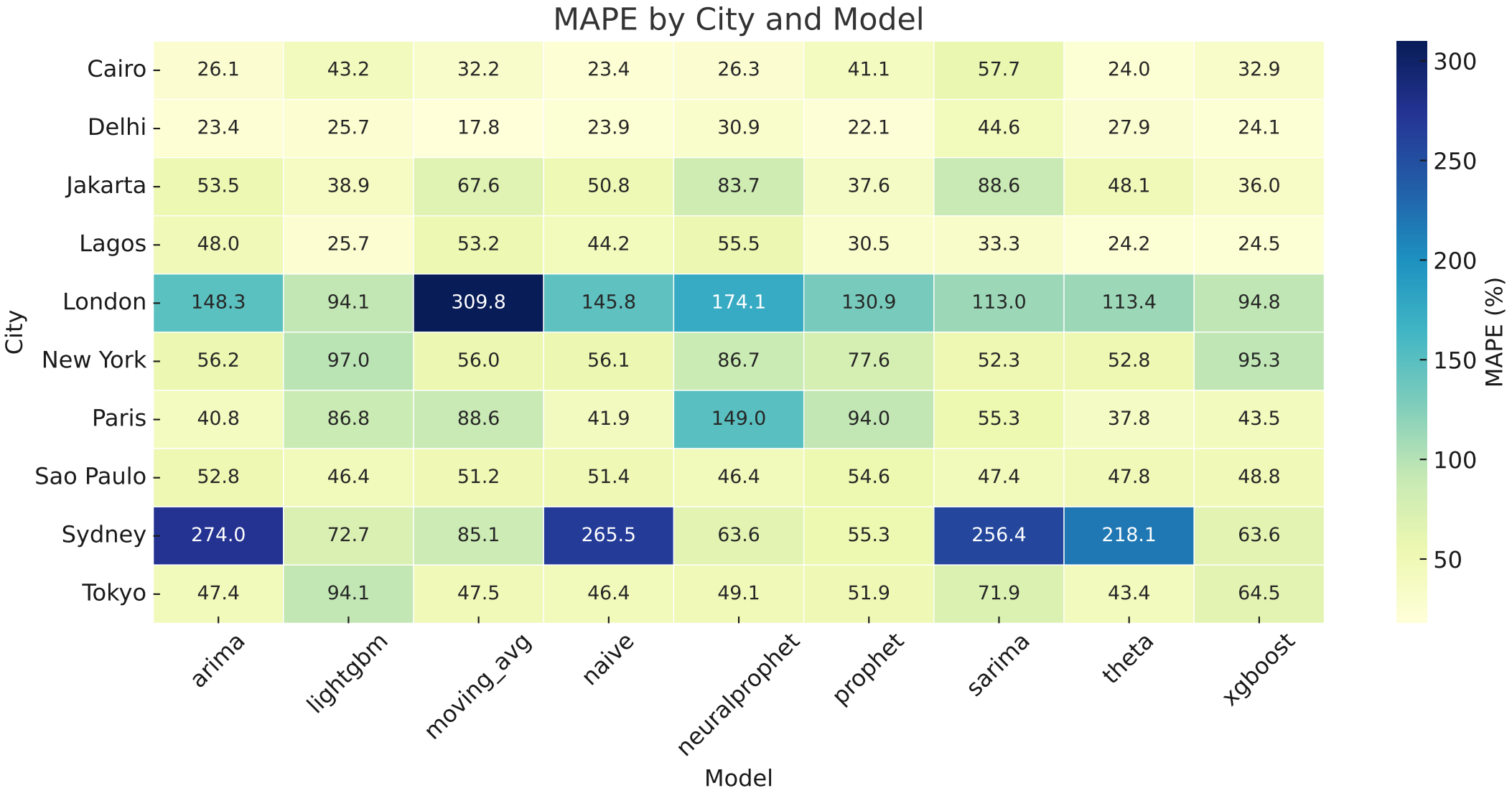
Powered by AQI-Cast | Validation using Open-Meteo PM2.5 Data

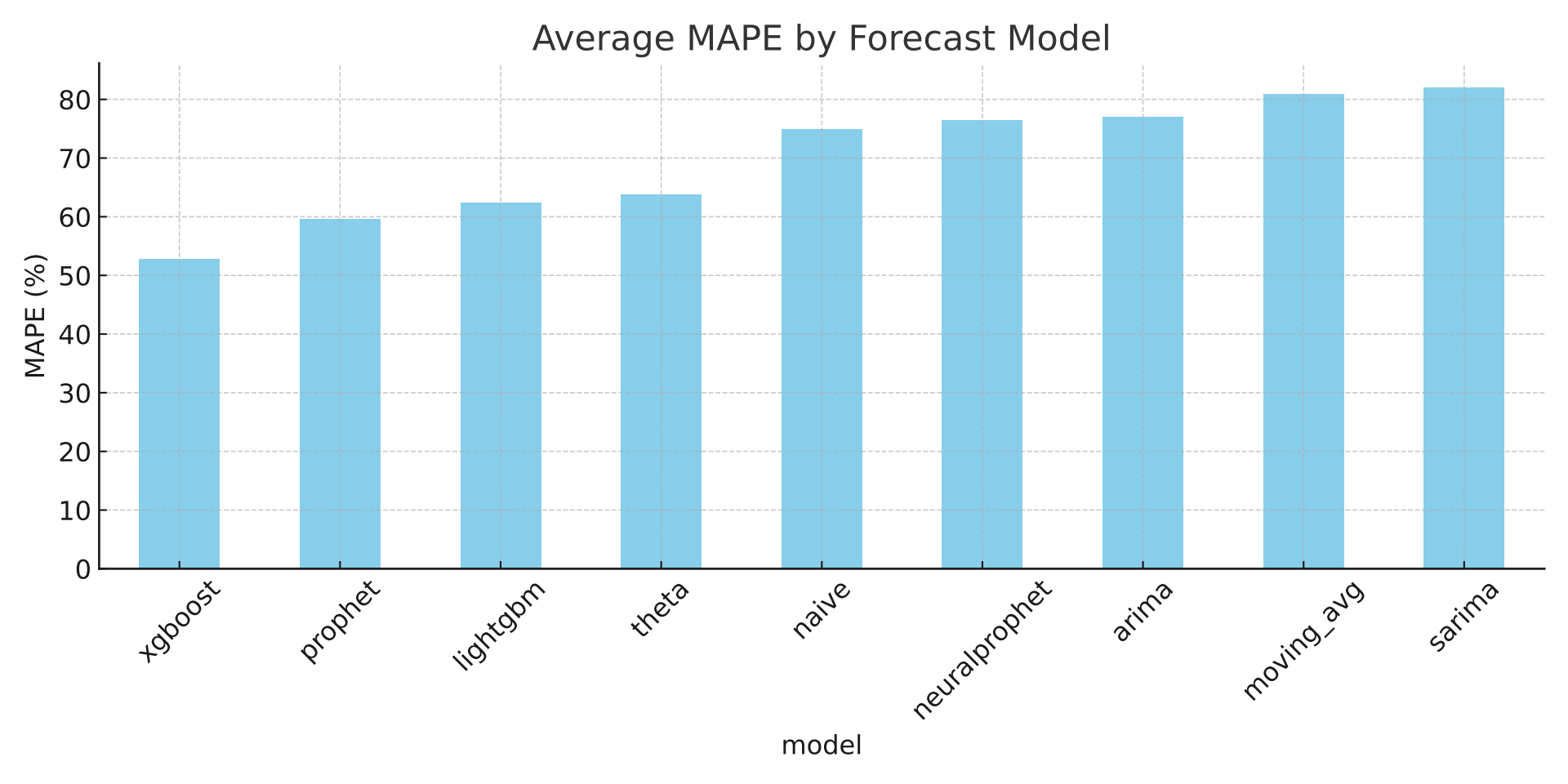
Executive Summary

This report presents a validation study of nine air quality forecasting models across ten globally distributed cities using Open-Meteo PM2.5 data. Each model was evaluated on its ability to forecast the last 7 days of AQI using the preceding 358 days of training data.

Key Findings:
- XGBoost delivered the best overall performance with the lowest MAPE (52.8%), MAE (8.33 µg/m³), and RMSE (10.48 µg/m³).
- Prophet followed closely, while LightGBM, Theta, and ARIMA were competitive but less consistent.
- SARIMA and simple moving average models underperformed, especially in variable cities like Sydney and London.
- Forecasts in cities like Delhi, Cairo, and Lagos were the most accurate across models, likely due to more stable pollution patterns.
- Cities with high variability and lower average AQI (e.g., Sydney, London) showed elevated forecast errors across all models.

Recommendation:
For robust, production-ready AQI forecasting, we recommend using XGBoost as the primary model, optionally blended with Prophet for added interpretability. Fallback to Theta or Naive models can be used for rapid inference or when data is sparse.


MAPE by City and Model
City
Cairo
Delhi
Jakarta
Lagos
London
New York
Paris
Sao Paulo
Sydney
Tokyo
26.1 43.2 32.2 23.4 26.3 41.1 57.7 24.0 32.9
23.4 25.7 17.8 23.9 30.9 22.1 44.6 27.9 24.1
53.5 38.9 67.6 50.8 83.7 37.6 88.6 48.1 36.0
48.0 25.7 53.2 44.2 55.5 30.5 33.3 24.2 24.5
148.3 94.1 309.8 145.8 174.1 130.9 113.0 113.4 94.8
56.2 97.0 56.0 56.1 86.7 77.6 52.3 52.8 95.3
40.8 86.8 88.6 41.9 149.0 94.0 55.3 37.8 43.5
52.8 46.4 51.2 51.4 46.4 54.6 47.4 47.8 48.8
274.0 72.7 85.1 265.5 63.6 55.3 256.4 218.1 63.6
47.4 94.1 47.5 46.4 49.1 51.9 71.9 43.4 64.5
arima
lightgbm
moving_avg
naive
neuralprophet
prophet
sarima
theta
xgboost
Model
MAPE (%)
300
250
200
150
100
50


Average MAPE by Forecast Model
MAPE (%)
0
10
20
30
40
50
60
70
80
xgboost
prophet
lightgbm
theta
naive
neuralprophet
arima
moving_avg
sarima
model

Model-Wise Summary (Avg Across All Cities)

model	MAPE	MAE	RMSE
arima	77.06	12.72	15.68
lightgbm	62.47	9.26	11.53
moving_avg	80.92	10.66	13.02
naive	74.93	12.37	15.36
neuralprophet	76.53	11.48	13.6
prophet	59.56	8.41	10.55
sarima	82.06	16.42	18.91
theta	63.73	11.79	14.67
xgboost	52.81	8.33	10.48

City-Wise Summary (Avg Across All Models)

city	MAPE	MAE	RMSE
Cairo	34.1	5.5	6.82
Delhi	26.73	20.0	25.06
Jakarta	56.09	29.89	35.32
Lagos	37.66	4.62	5.44
London	147.14	8.13	9.09
New York	69.98	5.13	6.77
Paris	70.87	4.29	5.18
Sao Paulo	49.65	10.01	13.78
Sydney	150.47	10.36	11.15
Tokyo	57.36	14.78	18.94