Satellite Hotspot-Based Air Quality Forecasting

UTTHAWIT
RATCHATAPOOM

DEC 20, 2024

2024

Agenda

- Introduction
- Problem Statement
- Procedures
- Finding
- Conclusion
- Recommendation
- Future Works

Introduction: The Silent Killer

Introduction: Sources of PM 2.5 Pollution

- Agricultural Burning: Open burning of agricultural waste, such as crop residue and rice straw
- Forest Fires: Wildfires and forest fires, particularly during dry seasons
- Domestic Biomass Burning: Burning of wood, charcoal, and other biomass fuels for cooking and heating in rural areas
- Industrial Emissions: Industrial activities, especially those involving combustion processes
- Vehicle Emissions: Exhaust fumes from vehicles, particularly diesel-powered vehicles

Problem Statement

- Leveraging Satellite-Based and Weather Data for Daily Air Quality Forecasting in Thailand

Procedure

2024

Data Sources

- Weather Data
  - Temperature (2m, 80m, 120m, 180m)
  - Relative Humidity (2m)
  - Dew Point (2m)
  - Pressure (MSL, Surface)
  - Wind Speed (10 m, 80 m, 120 m, 180 m)
  - Wind Direction (10m, 80 m, 120 m, 180 m)

Data Sources

- Weather Data
  - Data Point: 6,076,656
  - Data Range: Sep 9, 2024 – Dec 11, 2024
  - Coordinate: 13.35 - 97.05, 20.4 - 105.45
  - Source: DWD ICON via Open-Meteo
  - License Type: Commercial Subscription 29€/Month

Data Sources [cont.]

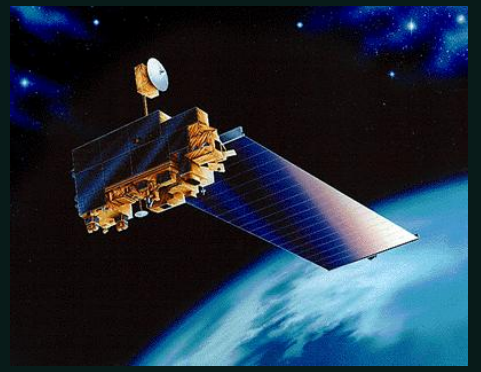
- Air Quality Data
  - PM 2.5 ($\mu g/m^3$) (Prediction Target)
  - UV Index

Data Sources [cont.]

- Air Quality Data
  - Data Point: 6,106,752
  - Data Range: Data Range: Sep 9, 2024 – Dec 11, 2024
  - Coordinate: 13.35 - 97.05, 20.4 - 105.45
  - Source: CAMS (Copernicus Atmosphere Monitoring Service) via Open-Meteo
  - License Type: Commercial Subscription 29€/Month

Data Sources [cont.]

- Satellite Hotspot Data
  - Satellite
    - *Aqua/Terra*
    - *SUOMI*
    - *NOAA 20, 21*
  - Data Point:17,933
  - Data Range: Data Range: Sep 9, 2024 – Dec 11, 2024
  - Source: Fire Information for Resource Management System (FIRMS)



TERRA / AQUA

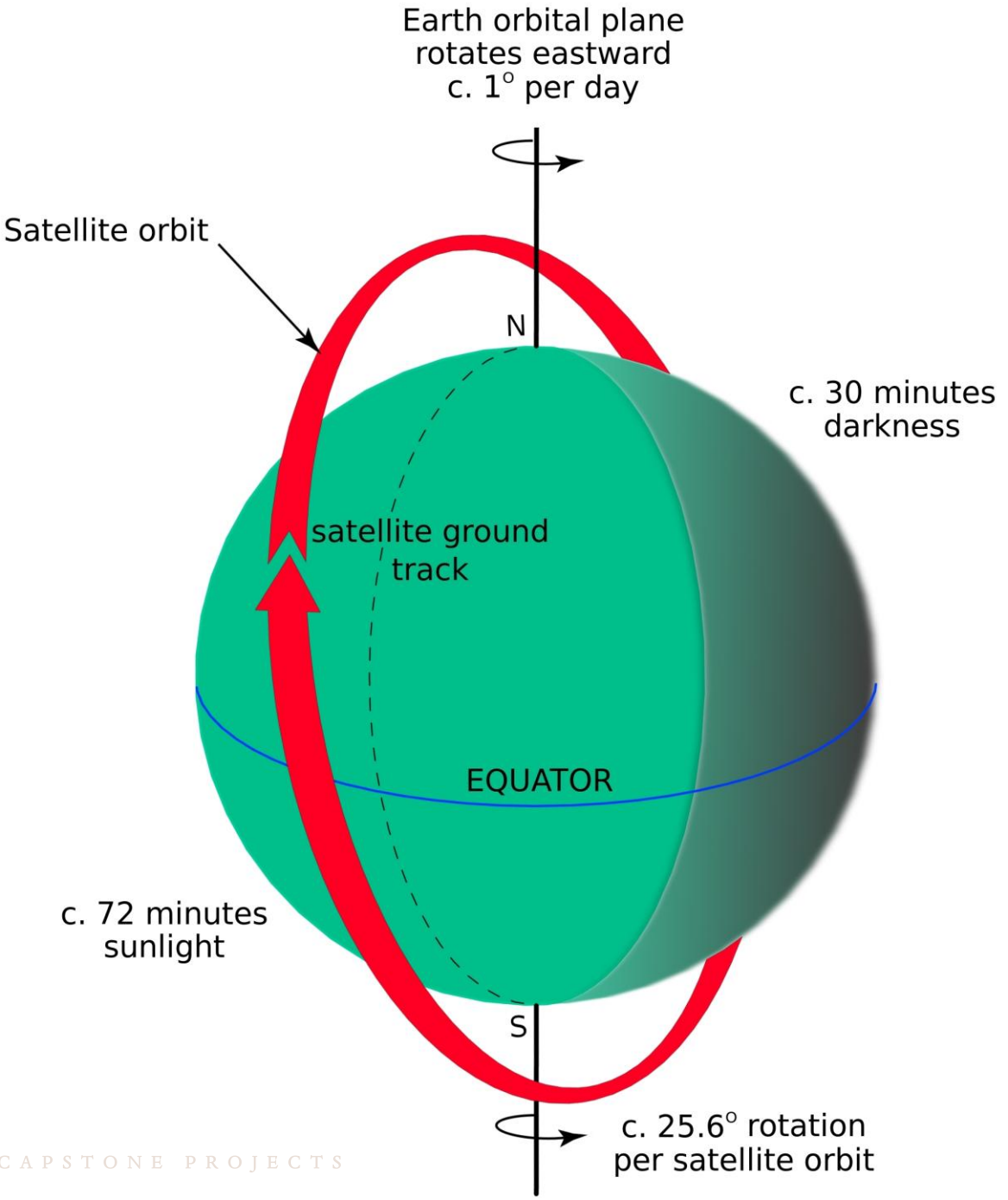
SUOMI

NOAA-20

NOAA-21

Data Sources [cont.]

- Satellite Hotspot Data



Candidate Algorithms

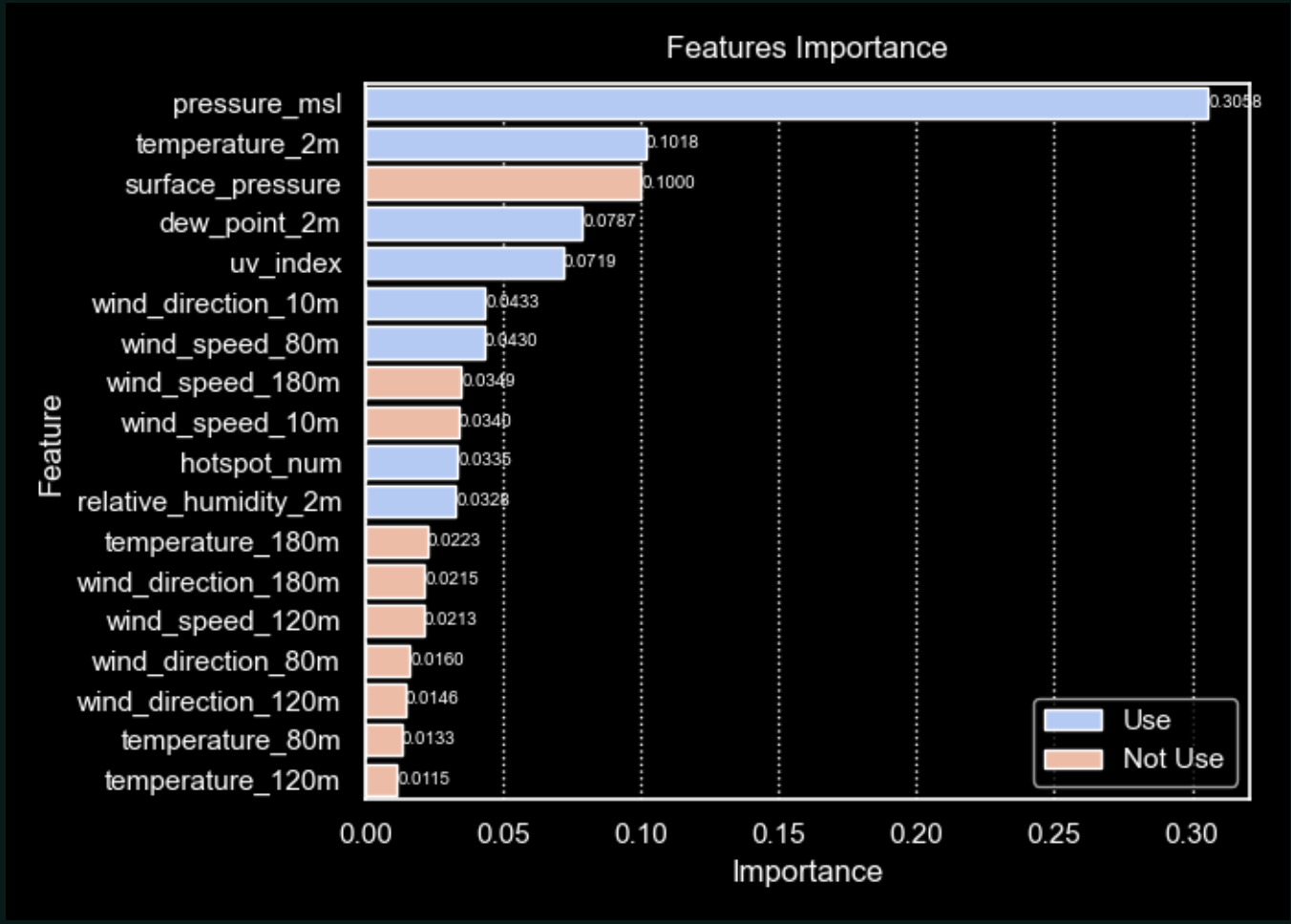
- Classical Regression
  - Linear Regression
  - ElasticNet
  - XGBoost
  - Random Forest Regressor
- Deep Learning
  - Long Short-Term Memory (LSTM)

Findings

2024

Features Importance

1. Pressure hPa (MSL)
2. Temperature C° (2m)
3. Dew Point C° (2m)
4. UV Index
5. Wind Direction Degree (10m)
6. Wind Speed km/h (80m)
7. Hotspot Number
8. Relative Humidity % (2m)

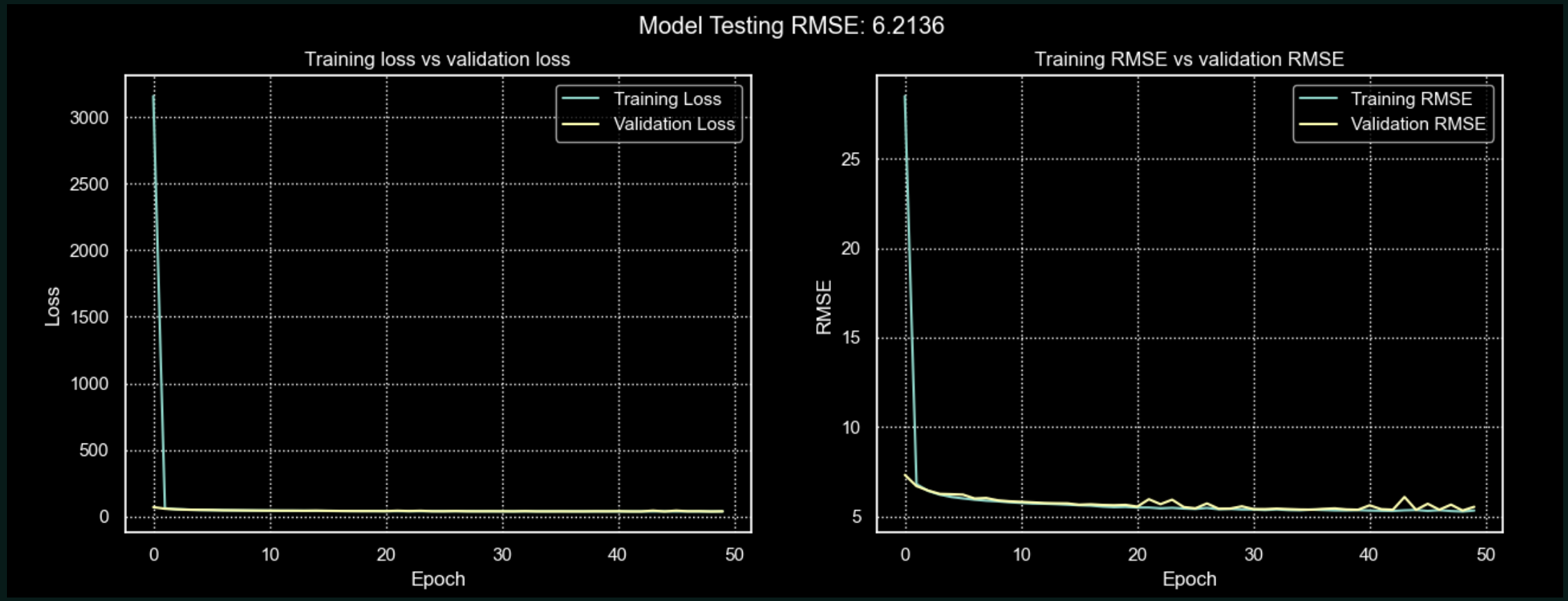


Models Comparison

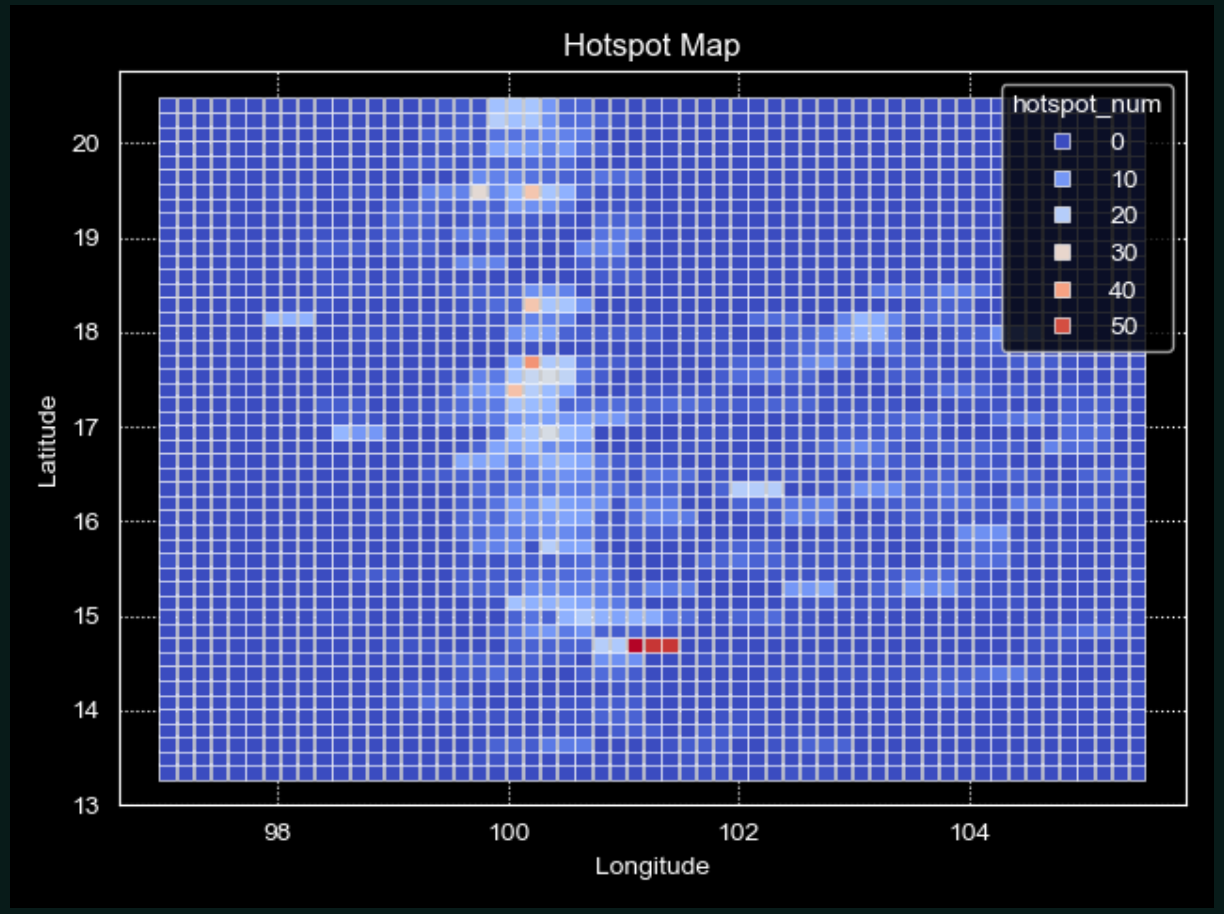
Algorithm Name	Cross Validation Score	Train Score	Test Score	Test RMSE
Random Forest	0.83	0.98	0.82	2.71
XGBoost	0.62	0.64	0.59	4.15
Linear Regression	0.48	0.48	0.46	4.73
ElasticNet	0.46	0.46	0.44	4.82
Long Short-Term Memory				6.10

- Select RandomForestRegressor since it has highest cross validation R^2 score

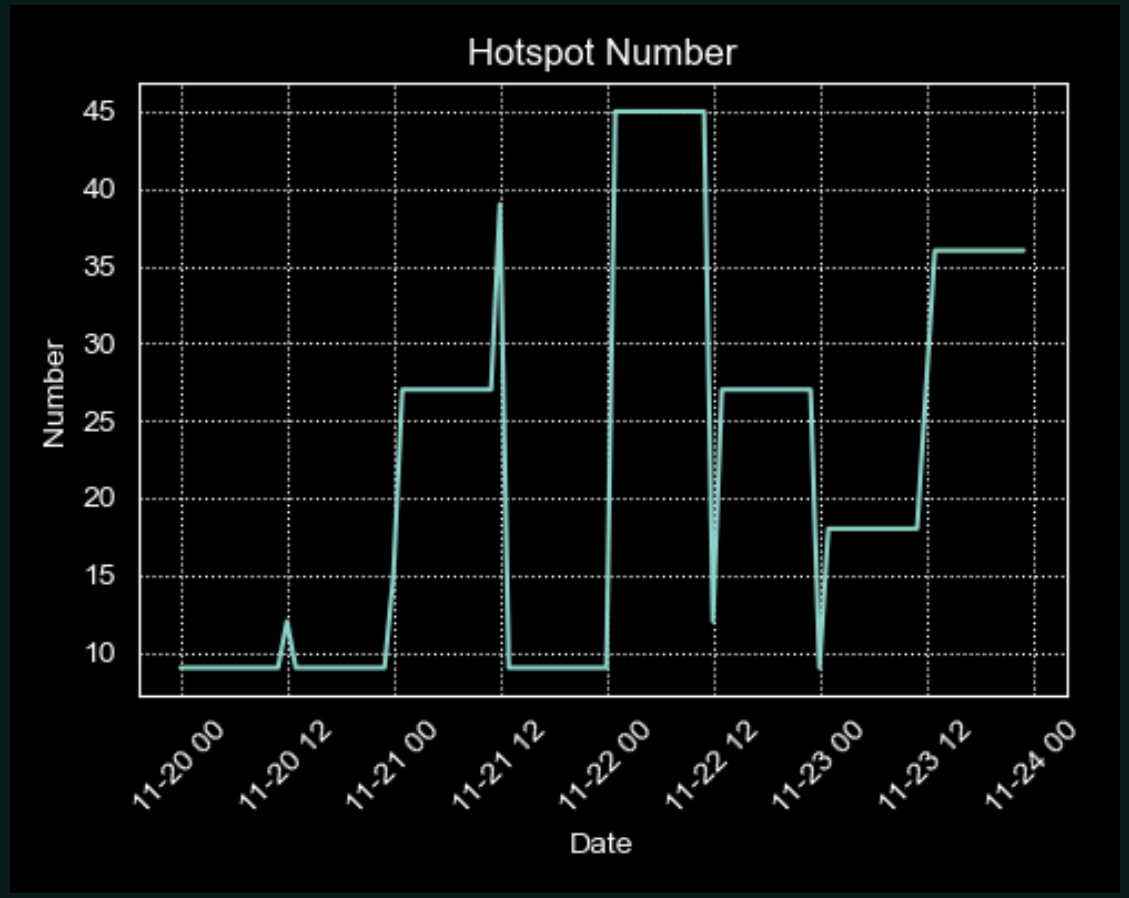
Long Short-Term Memory Model Graph

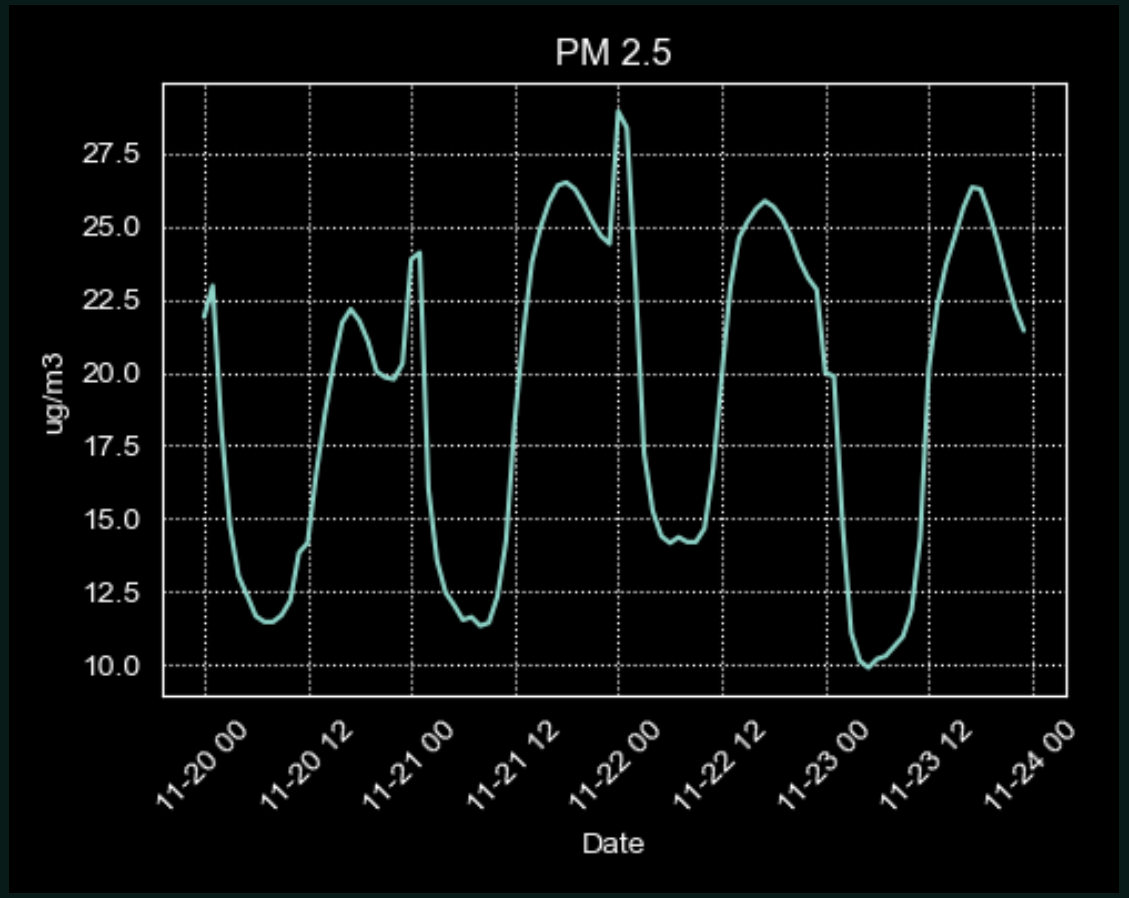


Hotspot of Upper Part of Thailand

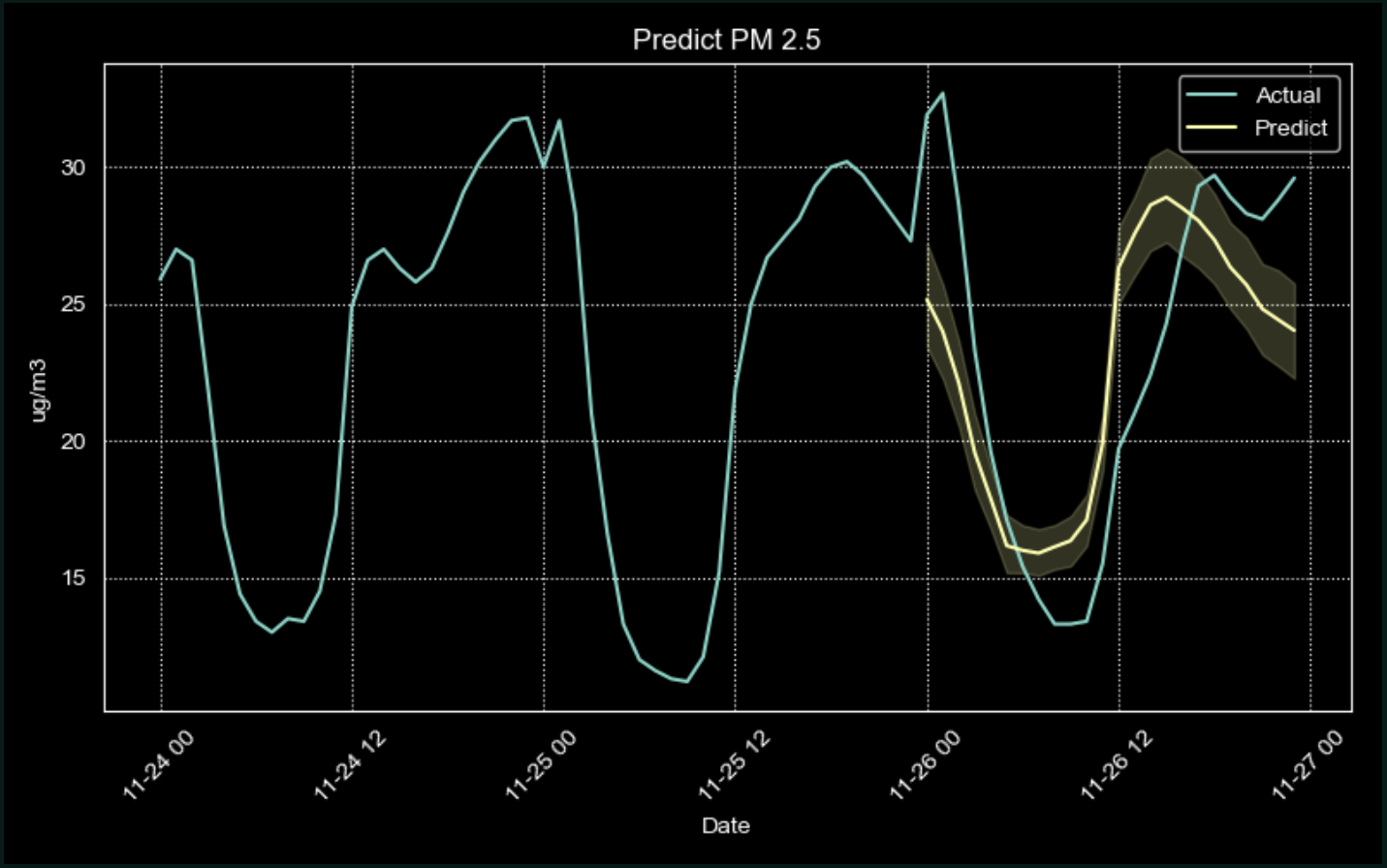


Case: Chiang Mai Hotspot Number

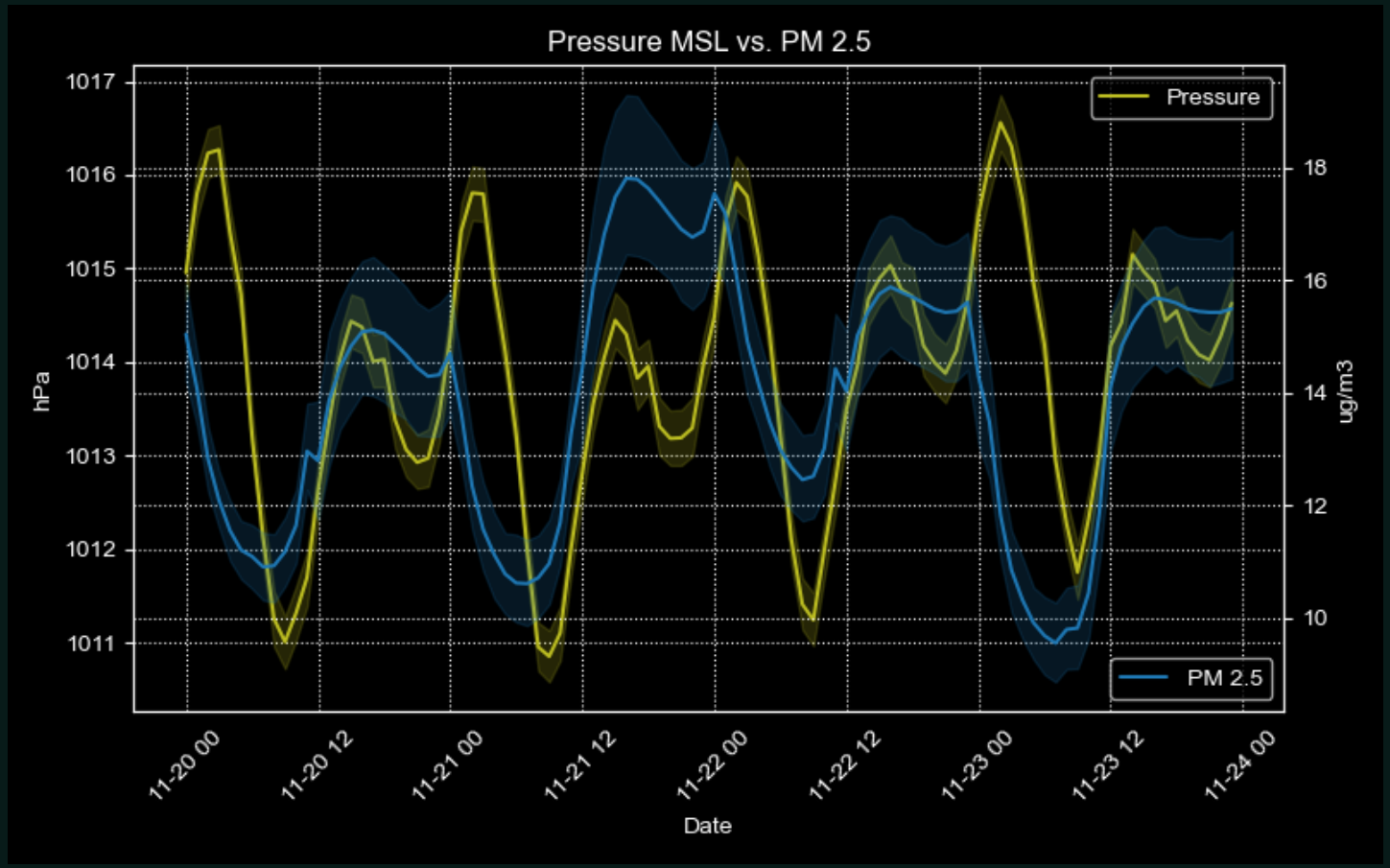




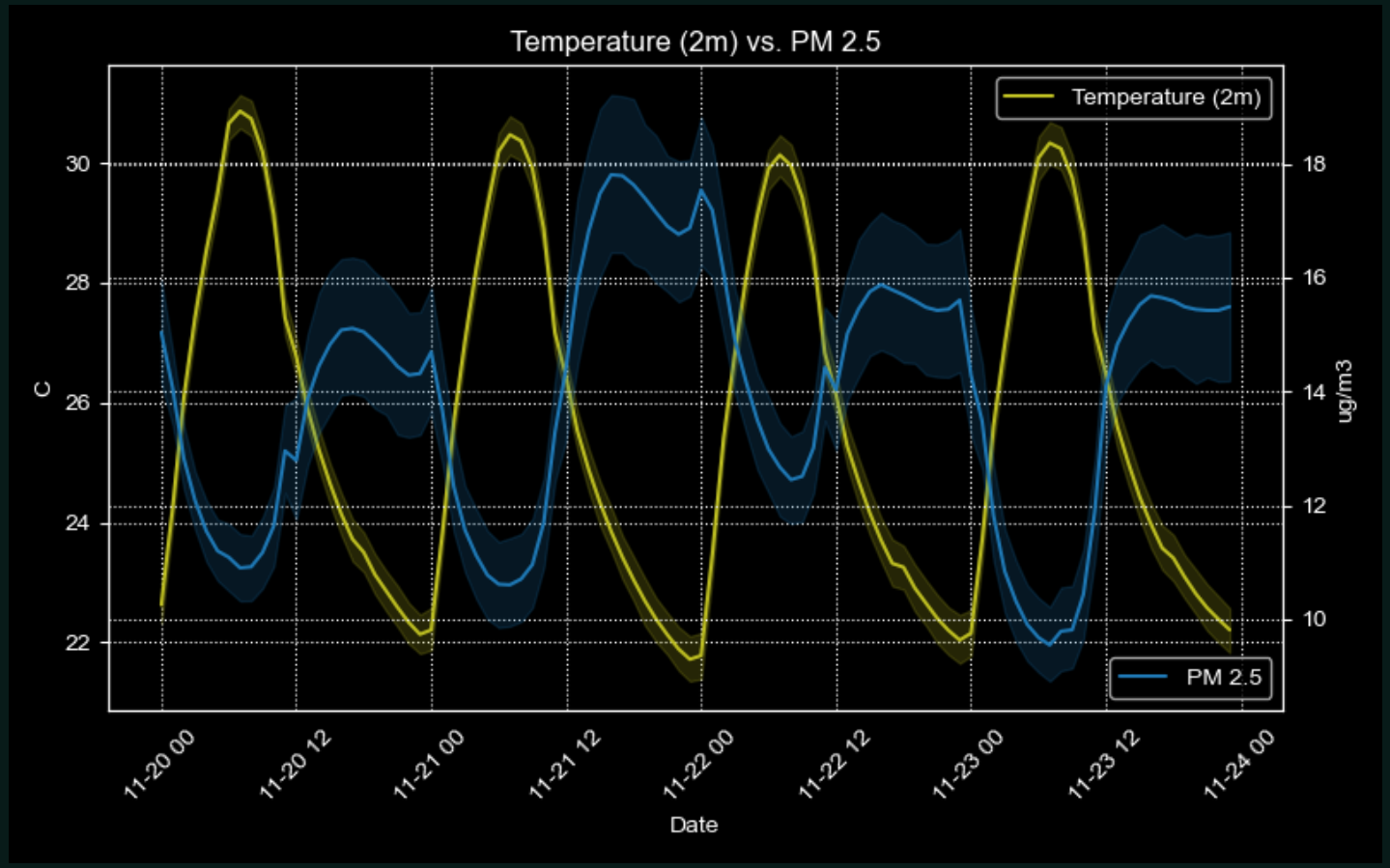
Case: Chiang Mai Hotspot Number



Pressure (MSL) vs PM 2.5



Temperature (2m) vs PM 2.5



Conclusion

- Increased PM2.5 concentrations at night
- PM 2.5 levels are strongly correlated with high atmospheric pressure
- PM 2.5 levels are strongly negative correlated with temperature

Recommendation

- Reduce open burning to mitigate PM2.5
- Stricter Regulations
- Public Awareness Campaigns
- International Cooperation

- Target Audience
  - Government Agencies
  - Media Outlets

Future Works

- Leverage historical data for temporal feature engineering
- Utilize spatial features from adjacent grid areas
- Incorporate historical fire hotspots to predict seasonal fire patterns

Thank you