Modelling spatio-temporal trends of air pollution in Africa

Paterne Gahungu\textsuperscript{1,2}, Jean Remy Kubwimana\textsuperscript{1}, Lionel Jean Marie Benjamin Muhimpundu\textsuperscript{2}, and Egide Ndamuzi\textsuperscript{3}

\textsuperscript{1}African Institute for Mathematical Sciences, Rwanda
\textsuperscript{2}Institute of Applied Statistics, University of Burundi
\textsuperscript{3}Ecole Normale Superieur de Bujumbura, Burundi

Abstract

Atmospheric pollution remains one of the major public health threat worldwide with an estimated 7 millions deaths annually. In Africa, rapid urbanization and poor transport infrastructure are worsening the problem. In this paper, we have analysed spatio-temporal variations of PM2.5 across different geographical regions in Africa. The West African region remains the most affected by the high levels of pollution with a daily average of 40.856 $\mu g/m^3$ in some cities like Lagos, Abuja and Bamako. In East Africa, Uganda is reporting the highest pollution level with a daily average concentration of 56.14 $\mu g/m^3$ and 38.65 $\mu g/m^3$ for Kigali. In countries located in the central region of Africa, the highest daily average concentration of PM2.5 of 90.075 $\mu g/m^3$ was recorded in N’Djamena. We compare three data driven models in predicting future trends of pollution levels. Neural network is outperforming Gaussian processes and ARIMA models.

1 Introduction

Air pollution in the form of particulate matter is associated with negative health effects and is considered as the largest contributor to premature deaths worldwide \cite{3,21,15,19}. According to the World Health Organization, more than 90 percent of the world’s population are exposed to harmful pollutants with levels exceeding up to five times the new guidelines updated in September 2021 by W.H.O \cite{14}. The major sources of air pollution are attributed to anthropogenic activities like industries, use of highly polluting car for transport, agriculture, and pollution from cooking with fossil fuel. Data on pollution levels and chemical composition are scarce in many low and middle income countries particularly in Africa. This is due to high cost of equipment needed to collect and analyze samples. In most low and middle-income countries, the only data available comes from estimates from satellites. A limited number of research outputs exist in some parts of the world, like the Sub-Saharan African region. The existing literature on air pollution in Africa focus more on the use of low-cost sensors to measure some specific pollutants most likely particulate matter with diameter less than 2.5 micro-meter \cite{13,15,12,4}. Some studies have considered short-term and localized campaigns for source apportionment \cite{19,6}.

This paper focuses on the spatio-temporal variations of air pollution across Africa and explores predictive models for future trends. The remainder of this work is organized as follows. The second session gives a brief description of the data and models used in this study. The third session presents results for spatio-temporal trends and predictions based on three data-driven models: Auto-Regressive Integrated Moving Average, Neural Network, and Gaussian processes. We conclude the work in session four.

2 Data and Methods

2.1 Data

In this work, we focus on particulate matter with a diameter of less than 2.5 micrometers (PM2.5). It is considered the most dangerous pollutant that affects human health with its ability to easily penetrate the lungs. Reliable measurements come from reference monitors. Reference monitors like the Beta Attenuation Monitors (BAMs) are very expensive. This is the main reason why these monitors are not widely used in many African countries, where the data available comes
from low-cost sensors. The table [I] and map [I] below describe the African cities studied based on air pollution data collected by the US Embassies across Africa.

Table 1: The summary of PM2.5 (µg/m³) levels from African cities includes the latitude and longitude of data collection, the time period, and the countries and cities.

| African Cities | Country   | Period          | Latitude | Longitude |
|----------------|-----------|-----------------|----------|-----------|
| Kigali         | Rwanda    | Feb - June 2022 | 1.9441° S | 30.0619° E |
| Conakry        | Guinea    | Jan 2020 - June 2022 | 9.6412° N | 13.5784° W |
| Bamako         | Mali      | Oct 2019 - Jan 2022 | 12.6392° N | 8.8029° W |
| Lagos          | Nigeria   | Feb 2021 - June 2022 | 6.3244° N | 3.3792° E |
| Abuja          | Nigeria   | Feb 2021 - June 2022 | 9.0765° N | 7.3986° E |
| Libreville     | Gabon     | Apr 2021 - Mar 2022 | 0.4162° N | 9.4673° E |
| Algiers        | Algeria   | Feb 2019 - June 2022 | 36.7538° N | 3.0588° E |
| Accra          | Ghana     | Jan 2020 - June 2022 | 5.6037° N | 0.1870° W |
| Kinshasa       | DRC       | Mar - June 2022 | 4.4419° S | 15.2663° E |
| Kampala        | Uganda    | Feb 2017 - June 2022 | 0.3476° N | 32.5825° E |
| Nairobi        | Kenya     | Mar 2021 - June 2022 | 1.2921° S | 36.8219° E |
| N'Djamena      | Chad      | May 2020 - June 2022 | 12.1348° N | 15.0557° E |
| Khartoum       | Sudan     | Jan 2020 - June 2022 | 15.5007° N | 32.5599° E |
| Addis Ababa    | Ethiopia  | Aug 2016 - June 2022 | 8.9026° N | 38.7578° E |
| Antananarivo   | Madagascar| Jan 2020 - June 2022 | 18.8792° S | 47.5079° E |
| Abidjan        | Côte d'Ivoire | Feb 2020 - June 2022 | 5.3600° N | 4.0083° W |

Figure 1: Air pollution monitoring stations at US Embassies in Africa

2.2 Methods

We model the data generation process with the Auto-Regressive Integrated Moving Average (ARIMA) model and two universal function approximators: Neural networks and Gaussian processes.

Auto-Regressive Integrated Moving Average

We model \( y_t \), the observation at time \( t \) as a function of historical values \( y_0, \ldots, y_{t-1} \).

\[
y_t = f(y_0, \ldots, y_{t-1}).
\]

where

\[
f = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + \epsilon_t
\]

\[+ \mu + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \cdots + \theta_q W_{t-q},
\]

where the model coefficients are \( \beta_1, \ldots, \beta_p \), and \( \theta_1, \ldots, \theta_q \), and \( \mu \) is the long run average, \( W_t \) is the shock for the process such that \( W_t = \sigma \times \epsilon_t \) where \( \sigma \) is the conditional standard deviation.

Neural networks

We denote the observations at time \( t \) by \( x_j \) with \( j = 1, \ldots, n \). Here \( n \) is size of the dataset. An approximation of the data generation process using a neural networks can be expressed as follows:

\[
y = f \left( \sum_{j=1}^{n-1} W_j x_j + b \right),
\]

where \( W \) are weights, \( b \) is the bias and \( f \) is the activation function. The following activation functions are commonly used.

- Rectified Linear Unit (ReLU).
  \[
f(.) = \max(0, .).
\]

- Sigmoid function.
  \[
f(.) = \frac{1}{1 + \exp(-.)).
\]

- Hyperbolic tangent.
  \[
f(.) = \tanh(.) = \frac{\exp(.) - \exp(-.))}{\exp(.) + \exp(-.)}.
\]

Gaussian processes

Gaussian processes are considered as infinite-dimensional normal distributions with a mean and a covariance functions denoted by \( m(.) \) and \( k(., .) \) respectively [17].

\[
f(.) \sim GP(m(.), k(., .))
\]

The following covariance functions are commonly used:
3 Results and Discussion

3.1 Trend analysis

Air pollution and its emission in African cities differ based on the country’s population, industrialization, urbanization, economic status, emission sources, physical regions, and meteorology. The trend analysis of PM2.5 levels included sixteen cities from across Africa. The analysis shows that the mass concentrations of PM2.5 in those cities are currently increasing, negatively impacting people’s lives by causing respiratory and cardiovascular diseases. The hourly, daily and seasonal analyses were made based on the cities’ locations.

The figure [2] shows hourly mean concentrations of PM2.5 in the countries from East Africa. From the figure, the highest hourly mean concentration of PM2.5 in the morning hours was observed in Kampala at 8:00 a.m. with the values of 77.46 µg/m³, and the trend started to decline in the period of 8:00 a.m. – 3:00 p.m. After 3:00 p.m., the hourly trend begins to rise as many workers with vehicles return home and emissions from biomass burning as a source of energy at home suggested to contribute.

Kigali ranks second in the morning rush hours when workers are on their way to work, with high levels of hourly mean concentration at 9:00 a.m. PM2.5 levels in Kigali begin to rise again between 4:00 p.m. and late at night, with the highest hourly mean concentration of 54.82 µg/m³ recorded at 0:00 a.m. These differences in PM2.5 levels in Kigali are thought to be caused by traffic, industries, and combustion activities. Other developing cities, such as Nairobi, Antananarivo and Addis Ababa have PM2.5 concentrations that are less than the ones for Kigali and Kampala. Though the mass concentrations of PM2.5 in Nairobi are not as high as in other cities such as Kampala, Kigali, and Addis Ababa, according to [3], mineral dust and traffic-related factors account for 74% of annual PM2.5 concentrations of the entire country.

In their findings have shown that PM2.5 concentrations in Nairobi, particularly in the street and areas adjacent to a city road, have exceeded WHO guidelines of 25 µg/m³ as a 24-hour mean. The East African traffic network is expanding on a daily basis, resulting in high levels of traffic activity; however, the majorities of traffic in these cities are old and in poor condition. These issues contribute to the respiratory problems that pedestrians may experience as a result of the improved road networks, particularly during morning and evening rush hours.

Environmental degradation has put pressure on West African countries as a result of rapid population growth, urbanization, and a growing economy. Small particle concentrations in West Africa are frequently much higher than the WHO recommended limits [14]. These fine particles are caused by human activities such as charcoal fires, waste combustion in cities, and Savannah fires. Other particles originate in North Africa as a result of the wind blowing dust from the Sahara desert.

The figure [3] depicts the hourly mean mass concentrations of PM2.5 in West African cities. According to the figure Bamako experiences the highest hourly mean PM2.5 concentrations due to its location. Emissions in Bamako have been linked to traffic, biomass burning, dust, and desert dust events. All of these factors are thought to have contributed to this city having the highest observable PM2.5 levels of any West African city studied.

Higher morning peak time averages in Bamako was observed at 9:00 a.m., with values of 85.89 µg/m³, and during the evening hours, the trend began to rise from 3:00 p.m. to 9:00 p.m. where the highest values of 115.95 µg/m³ was recorded at 9:00 p.m. This has resulted in this city having the highest PM2.5 emissions among the remaining West African cities considered in the study.

Nigeria, as one of Africa’s fastest developing countries, is being affected by the rise in mass concentrations of PM2.5 caused by human activities and industrialization in cities like Lagos and Abuja. According to the World Bank [7], Nigeria had the highest number of premature deaths due to ambient PM2.5, particularly in the cities mentioned above. Road transport, industrial emissions, and power generation are the primary sources of air pollution in Nigeria, though more research is needed to determine the contribution of each sector separately.

Because some cars in Nigeria are
older than five years, the quality of imported fuels (diesel and gasoline), and the unlimited means of transportation in cities, road transport contributes much more than other sources. All of these factors contributed to Abuja and Lagos having the highest mass concentrations of PM2.5 in the morning rush hours when compared to other remaining cities such as Accra, Conakry, and Bamako, particularly between the hours of 7:00 a.m. and 9:00 a.m. Between 6:00 p.m. and 10:00 p.m. in Abuja, PM2.5 emissions begin to rise due to mineral dust, vehicle exhaust, industrial emissions, and the use of firewood as a cooking fuel.

In Guinea, particularly in Conakry, some of the proposed air pollution emissions are caused by unregulated combustion and processing emissions from industrial sites, unregulated emissions from gasoline vehicles and residential wood burning. According to [20], cement manufacturing plants, electric power plants, brick manufacturing operations, steel smelters, and medical waste incinerators are among the industries that are thought to contribute to the mass concentration of PM2.5 in Conakry. Conakry has low hourly emissions when compared to other cities such as Bamako, Lagos, and Abuja. At 9:00 a.m., the highest hourly mean mass concentration of PM2.5 was recorded, with a value of 40.25 µg/m³ whereas the highest hourly mean concentrations of PM2.5 in the morning hours in Bamako, Lagos, and Abuja are 85.89 µg/m³ at 9:00 a.m., 57.06 µg/m³ at 8:00 a.m., and 51.94 µg/m³ at 7:00 a.m respectively. Other cities, such as Abidjan and Accra, which are close to the Atlantic Ocean, have emission rates similar to Conakry, as shown by the hourly trends in the figure [3].

![Figure 3: Hourly PM2.5 concentrations in West Africa](image)

The hourly mean variation of PM2.5 in figure [4] compares African cities from central region. According to the graph [4], the highest hourly mean mass concentrations of PM2.5 were observed in N'Djamaena and Khartoum. These two cities are close to the Sahara desert, which influences PM2.5 emissions in neighboring countries. Like in other cities, in the morning hours from 6:00 am to 9:00 am in N'Djamaena, the emission has increased up to 104.53 µg/m³ at 8:00 am, while in the evening, the highest hourly mean concentration of 139.83 µg/m³ was observed at 8:00 pm. In Khartoum, the highest peak of 64.36 µg/m³ was observed at 8:00 a.m., and levels of PM2.5 decreased after 8:00 a.m. until noon. The levels gradually increased at a slow rate during the afternoon hours until the evening. The highest peak in the evening and late night was at 9:00 a.m., with a mass concentration of 76.94 µg/m³.

Though PM2.5 concentrations in Kinshasa and Libreville in figure [4] are not as high as in other cities, the sources of PM2.5 emissions include road transportation, poorly maintained vehicles, smoke from open-air barbeques, burning trash, and non-standard fuels such as gasoline and diesel [11]. Gabon as a developing country, it has policies in place to control the quality of imported gasoline and diesel, but no clear policy in place to control vehicle emission standards and air quality regulations.

![Figure 4: Hourly PM2.5 concentrations in Central Africa](image)

Fuel combustion, power generation, and industrial facilities are some of the primary sources of PM2.5 in Algiers. Residential fireplaces and wood burning as a source of energy at home are two other sources that have increased emissions. This rise in PM2.5 mass concentrations in Algiers is linked to various cases of mortality and morbidity throughout the country. From figure [5], the highest and lowest hourly mean mass concentration of PM2.5 of 22.45 µg/m³ and 18.94 µg/m³ was observed at 10:00 a.m. and 7:00 a.m. respectively. The increase in PM2.5 mass concentrations in Algiers from 7:00 a.m. to 10:00 a.m. is associated with an increase in traffic activities, particularly traffic movements as people commute to work. PM2.5 levels dropped from 10:00 a.m. throughout the afternoon hours to 19.44/m3 at 6:00 p.m. Then, from 7:00 a.m. until late, the levels of PM2.5 in Algiers increased, reaching 21.908 µg/m³ at 2:00 a.m.
The daily trend analysis in figure [6] shows that weekdays (Monday through Friday) have high PM2.5 concentrations during the weekends (Saturday and Sunday) in some cities. These differences in emission during the weekdays and weekends are linked to the air pollution resulted from human and traffic activities in east Africa. Human activities have influenced the rate of emissions during the week, and the decrease on weekends is due to minimal activities. In Kigali, Nairobi, and Antananarivo, the highest peak was observed on weekdays rather than weekends, whereas weekends had higher levels of PM2.5 than weekdays in Kampala and Addis Ababa. The lower levels of PM2.5 in Antananarivo on weekends compared to weekdays were caused by a decrease in traffic activity. The daily variation of PM2.5 concentration in Kampala, Kigali, and Nairobi shows that the highest levels of PM2.5 were observed on Saturday, Wednesday, and Tuesday, with values of 57.35 µg/m³, 42.97 µg/m³, and 17.40 µg/m³, respectively. This implies that because these cities are among the most populated in East Africa, the daily mass concentration of PM2.5 varies due to daily activities. In Addis Ababa, weekday emissions were lower than on weekends, with Saturday having the highest daily mass concentration of 25.51 µg/m³.

According to figure [8], the concentration of PM2.5 was higher on weekdays than on weekends in Khartoum and N’Djamen. The average mass concentration of PM2.5 in Khartoum was 78.78 µg/m³ during the week and 72.979 µg/m³ at the weekend. The average PM2.5 level in N’Djamen was 92.26 µg/m³ during the week and 84.64 µg/m³ on weekends. When these results are compared, weekday emissions are more significant than weekends because they are the days with the most traffic and the influence of agricultural and construction activities is visible on weekdays. Although daily emissions are significant in Khartoum and N’Djamen, they are not the same in Libreville and Kinshasa. The highest mass concentration in PM2.5 was 13.68 µg/m³ in Libreville on Thursday, and 4.62 µg/m³ in Kinshasa on Monday. The daily emissions in these cities were not as high as in Khartoum and N’Djamen.

The figure [7] shows the weekly average PM2.5 for cities in West African countries. The mass concentration of PM2.5 is higher in cities such as Bamako, Lagos, and Abuja than in Accra, Abidjan, and Conakry. In these coastal cities, PM2.5 levels are lower than in other cities, with the exception of Lagos, Africa’s most populous city, where daily emissions that pollute the air are high.

Figure 5: Hourly PM2.5 concentrations in North Africa

Figure 6: Daily PM2.5 concentrations in East Africa

Figure 7: Daily PM2.5 concentrations in West Africa

Figure 8: Daily PM2.5 concentrations in Central Africa

Figure 9: shows the daily variation of PM2.5 levels in Algiers. The overall weekdays mean concentration is 20.83 µg/m³ and 20.76 µg/m³ on weekends. This indicates
that emissions were higher during the week than on weekends. The highest value of 21.33 μg/m³ PM2.5 was observed on Monday, while the lowest value of 20.409 μg/m³ was observed on Friday.

Figure 9: Daily PM2.5 concentrations in North Africa

The figure [10] depicts the seasonal variation of PM2.5 concentrations in East African cities. During the dry season, mineral dust is one of the suggested contributors to the increase in particulate matter due to unpaved road surfaces, especially when the wind blows dust. During the dry season, Savannah fires are associated with carbonaceous aerosols detected in this region of East Africa.

In figure [10] seasonal variations in PM2.5 levels were higher in Kàmpala and Addis Abàba than in other East African cities such as Nairobi and Antananarivo. The highest measured PM2.5 levels in Kàmpala during the dry season were observed in July, and as the country entered the short rainy season, emissions began to decline until the start of the short dry season. The monthly mean concentration of PM2.5 in the dry season is 67.29 μg/m³ in July and 73.51 μg/m³ in January whereas the lowest values observed in April and May were 38.37 μg/m³ and 42.06 μg/m³, respectively. Similarly, seasonal variation was observed in Addis Abàba, with PM2.5 averaging 36.71 μg/m³ during the dry season and 21.17 μg/m³ during the rainy season.

There are only two seasons in Madagascar: a hot, rainy season from November to April and a cooler, dry season from May to October. The highest level of PM2.5 in November was 36.11 μg/m³ during the hot, rainy season, while it was 31.84 μg/m³ during the other season. Nairobi appears to have lower PM2.5 levels than other cities. This implies that the factors that would be expected to cause seasonal variations (mineral dust suspension, open-air waste burning, and agricultural burning) have emitted fewer pollutants than in other cities. PM2.5 concentrations in Nairobi are higher in July and August than in other months.

Western countries experience the Harmattan season, which occurs between November and March during the dry season. The dry and dusty northeasterly trade wind blows from the Saharan desert across west Africa into the Gulf of Guinea during this season. The Harmattan brings desert-like weather conditions that reduce humidity, raise daily temperatures, prevent rainfall formation, and produce large clouds of dust that cause dust storms or sandstorms. Tornadoes can form when the Harmattan interacts with monsoon winds. Moist equatorial air masses from the Atlantic and Pacific oceans bring annual monsoon rains from May to September. This wind blows from the south-west and brings rain between May and September (wet monsoon).

The seasonal variation of PM2.5 levels in West African cities is described in the figure [11]. The decline of PM2.5 concentrations in the period of April to July is due to the rainy season across the region. In the rainy season, Conakry had the lowest monthly average PM2.5 concentrations of 8.40 μg/m³ and 9.37 μg/m³ in September and August, while in the dry season, Abuja, Bamako, and Lagos had the highest monthly average PM2.5 concentrations of 120.84 μg/m³, 106.92 μg/m³ and 93.02 μg/m³ in March, February, and January, respectively.

The monthly variations of PM2.5 levels in cities in Africa’s central region are shown in figure [12]. According to the figure, the highest levels of PM2.5 in N’Djamena
were observed in March and February. After March, emissions decrease as the country enters the rainy season, which reduces the amount of dust in N’Djamena. The rate of emission in Libreville is lower than in Khartoum and N’Djamena, indicating that these two countries which are close to the Sahara desert have high monthly mean mass concentrations of PM2.5. Khartoum has a subtropical desert climate, with relative rainfall from the African monsoon from the south in July and September. High emissions occur in December and February, when the temperature is warm during the day and cool at night. This temperature variation, especially during the day, raises the rate of PM2.5 emissions in Khartoum.

The seasonal spatio-temporal variation of PM2.5 in Algiers is represented in figure [13]. The monthly average value of PM2.5 ratios falls month after month. Many sandstorms hit Algiers in January, February, and March, raising PM2.5 levels. The monthly mean mass concentrations of PM2.5 in February, January, and March were 22.86 µg/m³, 23.0055 µg/m³, and 24.58 µg/m³, respectively. PM2.5 levels have risen in the months of May, June, and July with monthly average of µg/m³, µg/m³, and µg/m³, respectively because the weather is warm. The lowest monthly mean PM2.5 concentration of 15.79 µg/m³ was observed in November, as the country entered the winter season.

### 3.2 Forecasting

The figures from [14 - 29] are for forecasting trends in real-time PM2.5 data in all of the cities studied. To train models, data was divided into two sets: training and testing. Both models were used to forecast PM2.5 levels in various African cities.
Figure 17: Forecasting with ARIMA, Neural Networks, and Gaussian Processes on Lagos PM2.5 data

Figure 18: Forecasting with ARIMA, Neural Networks, and Gaussian Processes on Abuja PM2.5 data

Figure 19: Forecasting with ARIMA, Neural Networks, and Gaussian Processes on Libreville PM2.5 data

Figure 20: Forecasting with ARIMA, Neural Networks, and Gaussian Processes on Algiers PM2.5 data

Figure 21: Forecasting with ARIMA, Neural Networks, and Gaussian Processes on Accra PM2.5 data

Figure 22: Forecasting with ARIMA, Neural Networks, and Gaussian Processes on Kinshasa PM2.5 data

Figure 23: Forecasting with ARIMA, Neural Networks, and Gaussian Processes on Kampala PM2.5 data

Figure 24: Forecasting with ARIMA, Neural Networks, and Gaussian Processes on Nairobi PM2.5 data
4 Model Comparison

The prediction of air pollution, particularly particulate matter less than 2.5 microns in diameter, is location dependent because the sources of emissions differ from city to city and country to country. The forecasting results show that neural network and ARIMA models outperformed Gaussian Process Regression algorithms in adapting to real-time PM2.5 trends. The models were compared using three statistical metrics: root mean square, mean absolute error, and mean absolute percentage error to determine the accuracy of the model in adapting to actual PM2.5 levels. The table compares the performance metrics of ARIMA, Neural Networks, and Gaussian Processes.

Table 2: The Root Mean Square Errors (RMSEs) of ARIMA, Neural Networks, and Gaussian Processes.

| US Embassy in cities | ARIMA | ANN | GPR |
|----------------------|-------|-----|-----|
| Kigali               | 18.441| 17.770| 26.295 |
| Conakry              | 10.544| 11.386| 27.336 |
| Bamako               | 7.989 | 8.138| 29.383 |
| Lagos                | 11.494| 11.932| 42.882 |
| Abuja                | 28.168| 29.253| 67.153 |
| Libreville           | 3.756 | 4.244| 3.888 |
| Algiers              | 5.785 | 6.260| 10.954 |
| Accra                | 15.328| 16.672| 39.869 |
| Kinshasa             | 0.789 | 0.773| 1.631 |
| Kampala              | 12.179| 13.056| 27.059 |
| Nairobi              | 4.265 | 4.480| 2.692 |
| N’Djamena            | 108.71| 123.67| 165.84 |
| Khartoum             | 60.488| 60.170| 231.69 |
| Addis Ababa          | 5.801 | 6.237| 15.368 |
| Antananarivo         | 4.282 | 4.534| 5.664 |
| Abidjan              | 7.660 | 7.594| 20.509 |
According to the RMSE results, both the ARIMA and the Neural Network outperform the Gaussian process. This means that neural networks, as a machine learning approach, can be used to analyze and predict time series data, as well as to solve real-world problems like air pollution. Although the results of the Gaussian process regression are not as precise as those of the other models, the study shows that they can be used to forecast air quality data.

5 Conclusion

This study has focused on the analysis of PM2.5 variations across Africa. As it has been demonstrated in this paper, data-driven models such as Auto-Regressive Integrated Moving Average, Neural Network, and Gaussian Process Regression can be used to forecast PM2.5 trends in Africa. Future work should consider other air pollutants that affect human health, such as PM10, SO2, NOx, CO, and O3.

6 Acknowledgments

The authors, P.G. and J-R.K., would like to thank the African Institute for Mathematical Sciences for their assistance, as well as the financial support provided by the Canadian government through Global Affairs Canada and the International Development Research Centre.

References

[1] N. Belarbi, M. Belamri, B. Dahmani, and M. Benamar. Road traffic and pm10, pm2.5 emission at an urban area in Algeria: Identification and statistical analysis. Pollution, 6(3):651–660, 2020.

[2] L. Croitoru, J. C. Chang, and A. Kelly. The cost of air pollution in Lagos. 2020.

[3] G. Ezeh, O. Abiye, and I. Obioh. Elemental analyses and source apportionment of pm2.5 and pm2.5–10 aerosols from Nigerian urban cities. Cogent Environmental Science, 3(1):1323376, 2017.

[4] P. Gahungu and J. R. Kubwimana. Trend analysis and forecasting air pollution in Rwanda. arXiv preprint arXiv:2205.10024, 2022.

[5] S. M. Gaita, J. Boman, M. J. Gatar, J. B. Pettersson, and S. Janhäll. Source apportionment and seasonal variation of pm2.5 in a sub-saharan African city: Nairobi, Kenya. Atmospheric Chemistry and Physics, 14(18):9977–9991, 2014.

[6] E. Kalisa, E. G. Nagato, E. Bizuru, K. C. Lee, N. Tang, S. B. Pointing, K. Hayakawa, S. D. Archer, and D. C. Lacap-Bugler. Characterization and risk assessment of atmospheric pm2.5 and pm10 particulate-bound pahs and npahs in Rwanda, central-east Africa. 2018.

[7] K. KARIN and C. SHUBHAM. Air pollution: A silent killer in Lagos. https://blogs.worldbank.org/africacan/air-pollution-silent-killer-lagos, 2020.

[8] K.-H. Kim, E. Kabir, and S. Kabir. A review on the human health impact of airborne particulate matter. Environment international, 74:136–143, 2015.

[9] P. L. Kinney, M. G. Gichuru, N. Volavka-Close, N. Ngo, P. K. Ndiba, A. Law, A. Gachanja, S. M. Gaita, S. N. Chillrud, and E. Sclar. Traffic impacts on pm2.5 air quality in Nairobi, Kenya. Environmental science & policy, 14(4):369–378, 2011.

[10] J.-F. Léon, A. B. Akpo, M. Bedou, D. Djossou, M. Bodjrenou, V. Yoboué, and C. Lioussé. Pm 2.5 surface concentrations in southern-west African urban areas based on sun photometer and satellite observations. Atmospheric Chemistry and Physics, 21(3):1815–1834, 2021.

[11] C. McFarlane, P. K. Isevulambire, R. S. Lumbuenamo, A. M. E. Ndinga, R. Dhammapala, X. Jin, V. F. McNeill, C. Malings, R. Subramanian, and D. M. Westervelt. First measurements of ambient pm2.5 in Kinshasa, Democratic Republic of Congo and Brazzaville, Republic of Congo using field-calibrated low-cost sensors. Aerosol and Air Quality Research, 21(7), 2021.

[12] C. McFarlane, G. Raheja, C. Malings, E. K. Appoh, A. F. Hughes, and D. M. Westervelt. Application of gaussian mixture regression for the correction of low cost pm2.5 monitoring data in Accra, Ghana. ACS Earth and Space Chemistry, 5(9):2268–2279, 2021.

[13] D. Okure, J. Ssematimba, R. Sserunjogi, N. L. Gracia, M. E. Soppelsa, and E. Bainomugisha. Characterization of ambient air quality in selected urban areas in Uganda using low-cost sensing and measurement technologies. Environmental Science & Technology, 56(6):3324–3339, 2022.

[14] W. H. Organization et al. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization, 2021.

[15] G. Raheja, K. Sabi, H. Sonla, E. K. Gbedjjangni, C. M. McFarlane, C. G. Hodoli, and D. M. Westervelt. A network of field-calibrated low-cost sensor measurements
of pm2.5 in Lomé, Togo, over one to two years. ACS Earth and Space Chemistry, 6(4):1011–1021, 2022.

[16] R. Rajak and A. Chattopadhyay. Short and long term exposure to ambient air pollution and impact on health in India: a systematic review. International journal of environmental health research, 30(6):593–617, 2020.

[17] C. E. Rasmussen. Gaussian processes in machine learning. In Summer school on machine learning, pages 63–71. Springer, 2003.

[18] S. Schwander, C. D. Okello, J. Freers, J. C. Chow, J. G. Watson, M. Corry, and Q. Meng. Ambient particulate matter air pollution in Mpererwe district, Kampala, Uganda: A pilot study. Journal of environmental and public health, 20(14), 2014.

[19] M. Shehab and F. Pope. Effects of short-term exposure to particulate matter air pollution on cognitive performance. Scientific reports, 9(1):1–10, 2019.

[20] J. P. Weinstein, S. R. Hedges, and S. Kimbrohugh. Characterization and aerosol mass balance of pm2.5 and pm10 collected in Conakry, Guinea during the 2004 harmattan period. Chemosphere, 78(8):980–988, 2010.

[21] C. M. Wong, H. K. Lai, H. Tsang, T. Q. Thach, G. N. Thomas, K. B. H. Lam, K. P. Chan, L. Yang, A. K. Lau, J. G. Ayres, et al. Satellite-based estimates of long-term exposure to fine particles and association with mortality in elderly Hong Kong residents. Environmental health perspectives, 123(11):1167–1172, 2015.