



Predicting COVID-19 Cases From Atmospheric Parameters Using Machine Learning Approach

S. T. Ogunjo¹ , I. A. Fuwape^{1,2} , and A. B. Rabi³

¹Department of Physics, Federal University of Technology Akure, Akure, Nigeria, ²Office of the Vice Chancellor, Michael and Cecilia Ibru University, Ughelli, Nigeria, ³Centre for Atmospheric Research, National Space and Research Development Agency, Anyigba, Nigeria

Special Section:

The COVID-19 pandemic and environmental conditions in Africa

Key Points:

- This study explored the potential of predicting COVID-19 cases using various machine learning algorithms
- Different machine learning algorithms suggests different lags for efficient prediction
- Temperature was found to be the best predictor for COVID-19 cases within the study locations

Correspondence to:

S. T. Ogunjo,
stogunjo@futa.edu.ng

Citation:

Ogunjo, S. T., Fuwape, I. A., & Rabi, A. B. (2022). Predicting COVID-19 cases from atmospheric parameters using machine learning approach. *GeoHealth*, 6, e2021GH000509. <https://doi.org/10.1029/2021GH000509>

Received 3 SEP 2021

Accepted 28 FEB 2022

Abstract The dynamical nature of COVID-19 cases in different parts of the world requires robust mathematical approaches for prediction and forecasting. In this study, we aim to (a) forecast future COVID-19 cases based on past infections, (b) predict current COVID-19 cases using PM2.5, temperature, and humidity data, using four different machine learning classifiers (Decision Tree, K-nearest neighbor, Support Vector Machine, and Random Forest). Based on RMSE values, k-nearest neighbor and support vector machine algorithms were found to be the best for predicting future incidences of COVID-19 based on past histories. From the RMSE values obtained, temperature was found to be the best predictor for number of COVID-19 cases, followed by relative humidity. Decision tree models was found to perform poorly in the prediction of COVID-19 cases considering particulate matter and atmospheric parameters as predictors. Our results suggests the possibility of predicting virus infection using machine learning. This will guide policy makers in proactive monitoring and control.

Plain Language Summary The world is racing to combat the COVID-19 pandemic. Predicting the number of cases within a location is important in curtailing the spread of the virus. In this study, we used machine learning to predict future infections based on past infection cases and atmospheric predictors. Our results showed that the approach is effective in predicting the COVID-19 cases within the locations considered.

1. Introduction

Different approaches have been proposed to forecast COVID-19 cases. Cao and Francis (2021) used vector autoregression models to forecast the number of COVID-19 cases in Indiana, Pennsylvania using the concentration of the virus in waste water. Information theory approach has also been proposed for the investigation and prediction of COVID-19 (Fernandes et al., 2021). One of the most common approaches to modeling COVID-19 is the use of Susceptible-Exposed - Infectious - Recovered (SEIR) compartmental models. The basic SEIR model have been developed for Indonesia (Annas et al., 2020), India (G. Pandey et al., 2020), United Kingdom (Peiliang & Li, 2020), China (Wan et al., 2020). Dansu and Ogunjo (2021) extended the SEIR model to investigate inter-community transmission of COVID-19 with two Nigerian states as a typical case. In order to capture the special dynamics of COVID-19, several modifications have been proposed for the standard SEIR compartmental model. López and Rodo (2021) introduced the confined, quarantined and dead people into the model to account for the peculiarity of COVID-19. By considering travel restrictions and social distancing, and reclassifying exposed and infected people a new SEIR model was proposed (Li et al., 2021). SEIR models and the modified variants have several limitations including fixed rates and does not cover all the factors responsible for the virus. Hence, the need for improved approaches to modeling and forecasting the disease.

Machine learning and artificial intelligence algorithms have been used to study the evolution of COVID-19 cases in different countries. The performance of different deep learning algorithms in forecasting the number of COVID-19 cases was investigated across six different countries (Zeroual et al., 2020). The variational autoencoder was found to outperform other models including recurrent neural network, long-short term memory, and gated recurrent units. Machine learning algorithms have been used to identify age, days in hospital, Lymphocyte and Neutrophils as important predictors in determining COVID-19 fatalities (Smith & Alvarez, 2021). In a similar research across 10 countries, Long-short term memory was found to perform better in six countries while Gated Recurrent Units was effective in four countries (ArunKumar et al., 2021). The forecast performance of stochastic and neural network models in COVID-19 case for 8 European countries has been carried out (Kirbaş

© 2022 The Authors. GeoHealth published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

et al., 2020). The long-short term memory was found to show the best performance. A novel support vector regression model was developed to predict the spread, growth rate, and end of the COVID-19 across different countries (Yadav et al., 2020). The bidirectional long-short term memory was also compared with other state of the art forecasting technique and found to be more effective (Said et al., 2021). Three artificial neural network algorithms, Radial Basis-Function, Fuzzy Cluster-Means, and Non-linear Autoregressive-Network with Exogenous Inputs were used to spatial forecast COVID-19 cases in Iraq (Yahya et al., 2021). A deep neural network was developed to predict COVID-19 cases across the United States and European countries using gini coefficients, percentage of tested population, and urban population (Hashim et al., 2021).

The performance of machine learning algorithms is heavily dependent on the model predictors. There exist a bidirectional relationship between atmospheric parameters and aerosols with COVID-19 infections. On the one hand, the spread of the virus and the associated non-pharmaceutical interventions have led to changes in atmospheric weather conditions and aerosol propagation in many parts of the world (Fuwape et al., 2021). Furthermore, changing weather has been reported to have significant health impact on humans and animals (Orimoloye et al., 2019; Ropo et al., 2017). Temperature and particulate matter up to 15-day lag have been found to be associated with an increase in COVID-19 cases in Italy (Stufano et al., 2021). The initial outbreak of the pandemic in India has been reportedly associated with increase in temperature and humidity (A. Pandey et al., 2022). Nonlinear relationship was observed between atmospheric parameters (temperature and humidity) and COVID-19 in several cities within the United States of America (Runkle et al., 2020). Association between atmospheric parameters and aerosols with incidences of COVID-19 have been confirmed in Algeria (Rahal et al., 2021), Egypt (Anis, 2020), Turkey (Şahin, 2020), and Indonesia (Tosepu et al., 2020).

The highlighted researches considered 1-day lag cases as inputs to the machine learning model. However, this might not yield this best results for forecasting as the virus have been known to have latency period between 7 and 14 days. Also, 1-day lag cases might not give sufficient information for the models. It is essential to consider n-days lag cases for better prediction and forecasting. Furthermore, these studies did not consider any other predictors. It has been reported that atmospheric parameters including aerosols are responsible for the spread of the virus. Hence, it is pertinent to investigate prediction of COVID-19 cases using air pollutants and atmospheric parameters. In this study, we aim to investigate the performance of different machine learning classifiers in the forecasting of COVID-19 cases using the number of cases from n-days. Also, we investigated the performance of the machine learning classifiers with atmospheric parameters and air pollutants as predictors.

2. Methodology

2.1. Data and Study Area

For this study, six locations within Nigeria were considered based on their geographical locations. The locations are classified into two - Northern stations (Kebbi, Kano, Abuja) and Southern stations (Delta, Edo, and Osun). The northern and southern stations have different climatic regimes. The air pollution in the northern region is largely driven by dust from the Bodele region in Chad (Sunnun et al., 2013). The dust system, driven by large scale oscillations reaches the southern parts of the country (Anuforom, 2007) and transported as far as the Amazon basin in South America (Koren et al., 2006). The weather dynamics of the southern part is driven largely by the Atlantic ocean (Ogunjo et al., 2019). This is largely responsible for the low temperature range throughout the year (Eludoyin et al., 2014). The major sources of pollution in the southern part are biomass burning and gas flaring (Ezeh et al., 2017; Ologunorisa, 2001). Daily COVID-19 cases for each of the locations were obtained from the National Centre for Disease Control (www.ncdc.gov) while the atmospheric data (temperature and relative humidity) and particulate matter (PM_{2.5}, PM_{1.0}, PM_{10.0}) were obtained from the ongoing campaign of the Centre for Atmospheric Research, National Space Research and Development Agency using Purple Air sensors. The Purple Air sensors were provided courtesy of the Alliance for Education, Science, Engineering and Design in Africa (AESEDA), Penn State University, USA. The data for particulate matter and atmospheric variables were retrieved from the Purple air network (www.purpleair.com). The research was conducted during the harmattan season in Nigeria from 1 November 2020 till 31 March 2021.

2.2. Machine Learning Algorithms

Machine learning approach has been chosen due to its various applications in various fields, ability to identify patterns, does not require specific distribution of the underlying data, and reliable results. Four machine learning algorithms (Decision Tree, Random Forest, Support Vector Machine, and k Nearest Neighbor) were considered in this study. In all of the algorithms, 80% of the data was used for training while 20% was used for testing. The root mean square error (RMSE) was considered as the test statistics due to its ability to compensate for large errors and has the same unit as the dependent variable.

In Decision Tree (DT), a series of decision based on given conditions are used to arrive at a conclusion. The internal nodes are the available choices at a particular point in the tree. The result that will result in the subdivision of the tree into n -subsets is called the root node. The results from the root and internal node culminate in branches. Branching or splitting is based on a set of conditions. In this study, we used the gini index for splitting. The DT algorithm has been used for COVID-19 diagnosis from chest X-ray imaging (Yoo et al., 2020), quantify the impact of mandatory lockdown on COVID-19 cases (Karnon, 2020), and predict COVID-19 cases and fatality based on age and gender (Bhatnagar et al., 2020). The DT approach has been found to be better than k Nearest Neighbor and other methods in predicting recoveries of infected patients from COVID-19 (Muhammad et al., 2020; Pourhomayoun & Shakibi, 2021).

k Nearest Neighbor (KNN) involves creating a space for training the data set. When a new sample to be trained is introduced into the sample space, the distance to the nearest neighbor in that space is estimated. Then, the status of the sample is determined by the number of neighbors in the vicinity. In this study, the nearest neighbor was estimated using the kd tree approach and a total of 48 neighbors. An enhanced version of KNN has been proposed for the improved detection of COVID-19 infections (Shaban et al., 2020). An improved COVID-19 detection method based on genome sequence was performed using KNN (Arslan & Arslan, 2021). Using age and gender, the KNN classification method was found to be superior in predicting the recovery of infected patients (Romadhon & Kurniawan, 2021).

Support Vector Machine (SVM) are non-parametric approaches to classification of data points. In SVM, a boundary line is drawn for the classification. Points close to this boundary are called support vectors. The classification is then performed by the linear combination of the boundaries (Yadav et al., 2020). In this study, the radial basis function was used as the kernel with a polynomial function of degree 3. The SVM has been coupled with particle swarm optimization for the detection of COVID-19 virus from chest X-ray images (Dixit et al., 2021). The spread of COVID-19 across different regions of the world has been predicted based on SVM (Yadav et al., 2020). The SVM algorithm has been deployed for the real-time prediction of COVID-19 infection, recoveries, and fatalities (V. Singh, Poonia, et al., 2020). For better performance, the SVM method has been coupled with least squares for the prediction of COVID-19 trajectory (S. Singh, Parmar, et al., 2020).

Random Forest (RF) is a DT based algorithm. The decisions are made from a randomly selected subset of the training data. The decision from various decisions are then used in making the final output. In this study, 40 "trees" were used with the gini measure. The RF algorithm was fine tuned with adaboost for the prediction of infected patients' health (Iwendi et al., 2020). Spatio-temporal near future prediction of COVID-19 was implemented worldwide using random forest with good results (Yeşilkanat, 2020). RF has been found to outperform other machine learning algorithms in the prediction of COVID-19 (Prakash et al., 2020). In India, RF was found to outperform other algorithms for the prediction of cases, fatalities, and recoveries (Gupta et al., 2021).

3. Results and Discussion

Using n -days lag as predictors, the step ahead prediction of COVID-19 cases was made at the different locations (Figure 1). The lag with minimum RMSE values represents the amount of previous rates that is needed to make an informed decision by the machine learning algorithms. In Kebbi State (Figure 1a), all the models showed the same lag values of 7 days. This means that values for the last 7 days is required by the models to make the best predictions in Kebbi State. In terms of RMSE, KNN and SVM presented identical values while the largest error of 4 infections were shown by RF. In Kano State (Figure 1b), DT and KNN were observed to have the same number of lags (5 days). However, KNN presents the lowest RMSE value of 6 cases amongst the four models while RF exhibited the worst estimate at 10 cases. All the models agree in terms of lag (4 days) at Abuja (Figure 1c). The

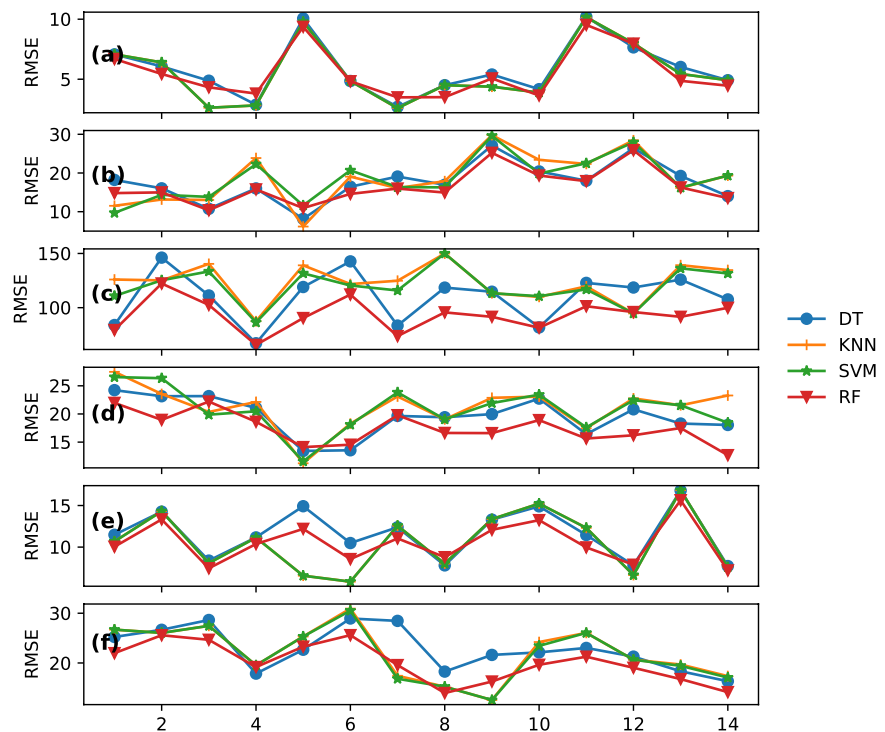


Figure 1. Root mean square error values for the different machine learning algorithms at (a) Kebbi (b) Kano (c) Abuja (d) Delta (e) Edo (f) Osun States.

RMSE values were observed to be 67, 87, 86, and 65 for DT, KNN, SVM, RF models respectively. Thus, RF outperformed all the other models in Abuja. In Delta State, all the models have the minimum RMSE values at 5-days lag except RF which showed 14-days lag. This implies that the RF algorithm will need much more information than the other algorithms in Delta State for best prediction. In Delta State, the best RMSE values was obtained for KNN. KNN and SVM showed identical 6-days lag values while DT and RF also showed identical 14-days lag values in Edo State. KNN and SVM also showed identical 9 days lag in Osun State while DT and RF showed 14-days and 8-days lag respectively. In Edo and Osun State, the RMSE values for KNN and SVM were found to be identical. Generally, KNN and SVM showed identical performances in three of the locations considered, KNN showed superior performance in two of the locations while DT outperformed other models in only one location. In terms of RMSE values, the performance of the models in increasing order was observed to be Kebbi, Edo, Kano, Delta, Osun, and Abuja.

The possibility of estimating COVID-19 cases from particulate matter (PM_{2.5}) was considered at zero lag. In Figure 2, the linear relationship between the predicted and measured values are presented for the various locations under consideration. The RMSE values are shown in each plot. The greatest error was observed for Edo in all machine learning algorithms considered. In Kebbi, KNN outperformed the other models with an error of about 9 cases while DT has the worst performance. Kebbi was found to have the least error amongst all the locations. In Kano, DT exhibited the lowest RMSE values among the models, similar to the results in the number of cases. With an RMSE of about 25 cases, RF was the best performing model in Abuja. However, SVM showed similar performance while the worst performing model was KNN. The best performing models were observed to be KNN, RF, and RF in Delta, Edo, and Osun States respectively. Generally, RF was seen to have the greatest performance at zero lag with superior results in four locations.

In order to determine the effect of lag on the performance of the models, the number of COVID-19 cases were predicted under different lags for the three predictors. The results are shown in Figure 3 and Table 1. The highest and lowest lags were found at 14-day and 1-day lags respectively. In Kebbi State, all the models showed the same 7-days lag for PM_{2.5} as predictor except RF which showed a 2-day lag. The RMSE values were in the range 8.16–8.73 with KNN and DT reporting the best and worst performance respectively. This implies that KNN will give the best prediction of COVID-19 infection in Kebbi state with an error of about 8 cases given values

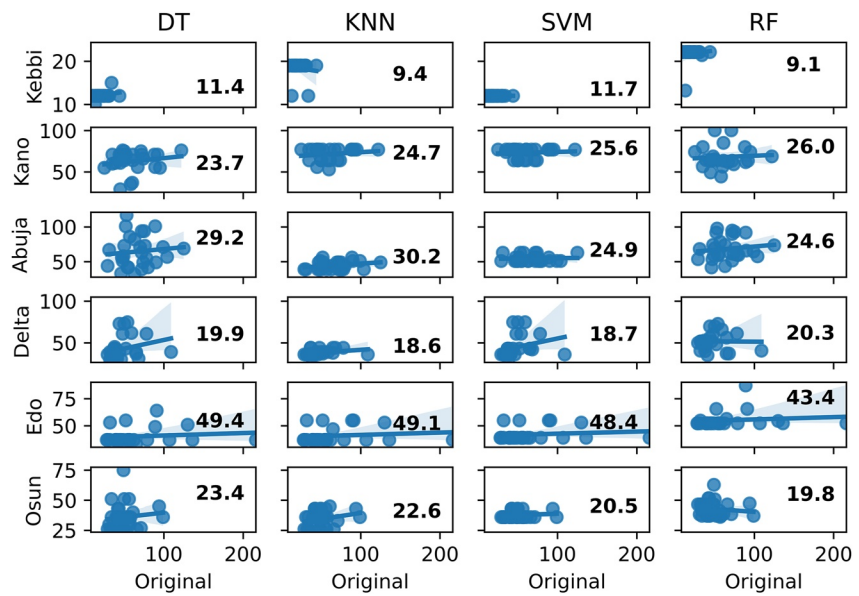


Figure 2. Linear plot of predicted and measured COVID-19 cases using zero lag particulate matter as predictor with root mean square error as a measure.

for 7 previous days. All the models were in agreement for the number of lags with temperature and humidity at 3-days and 13 days respectively. RF and DT were the best performing models with RMSE values of 3.08 and 4.25 in temperature and humidity respectively while SVM has the worst performance in the two parameters. In Kano State, DT and RF are in agreement with the required number of lags in both PM2.5 and temperature. The highest and lowest lags with PM2.5 as predictors were observed in KNN and SVM respectively. Considering temperature as predictors, KNN showed the highest number of lags but lowest RMSE values to emerge as the best performing model. In the case of humidity, KNN and SVM showed lags of 5-days while RF has the highest lag. There are a number of agreements in n-lags for Abuja. DT/KNN agrees for PM2.5, KNN/SVM and DT/RF pairs showed the same number of lags for temperature, and KNN/SVM agrees for humidity. SVM, KNN, RF were the best performing models for PM2.5, temperature, and humidity respectively in Abuja while DT showed the worst

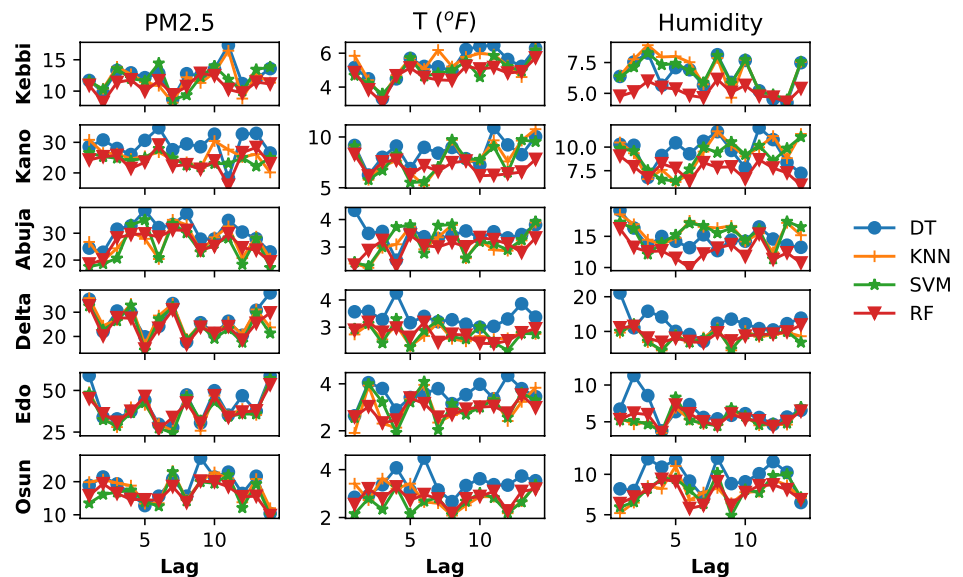


Figure 3. Performance evaluation of different machine learning algorithms at different lags using three predictors.

Table 1
Optimal Lags for Best Prediction Using Different Predictors

Location	Parameter	DT	kNN	SVM	RF
Kebbi	PM2.5	7	7	7	2
	Temperature	3	3	3	3
	Humidity	13	13	13	13
Kano	PM2.5	11	14	9	11
	Temperature	2	6	5	2
	Humidity	3	5	5	14
Abuja	PM2.5	2	2	14	1
	Temperature	4	2	2	4
	Humidity	8	12	12	6
Delta	PM2.5	8	5	5	5
	Temperature	10	5	12	11
	Humidity	7	4	4	7
Edo	PM2.5	7	7	7	6
	Temperature	1	1	4	3
	Humidity	4	4	4	4
Osun	PM2.5	14	14	14	14
	Temperature	8	8	8	8
	Humidity	14	1	9	6

performance among all the predictors. In Abuja, temperature was found to be the best predictors with an error of about 3 cases.

The performance of the models in the southern locations (Delta, Edo, and Osun States) were also considered. The models were in agreement with a lag of 5 when PM2.5 was considered as the only predictor in Delta state, except DT which showed a lag of 8. The RMSE values were found to be in the range of 14.27–17.73 with SVM and DT having the lowest and highest values respectively. In the case of temperature in Delta State, KNN has the lowest lag while SVM has the highest lag. This implies that KNN requires COVID-19 information from 5 days to make the best predictions while SVM needs 12 days of information. The best and worst performing models are SVM and DT respectively. For humidity, KNN/SVM needs for 4 days of information while DT/RF requires 7 days of information. In this case, KNN has the best performance while DT has the worst performance of the four models considered. DT, KNN, and SVM requires PM2.5 information from prior 7 days to effectively predict the COVID-19 cases in Edo State while RF requires 6 days worth. The best result was obtained in SVM with an error of about 24 cases. DT and KNN were found to have 1-day lag while SVM and RF requires 4 and 3 days lag respectively when predictions were made with temperature. The RMSE values were between 2 and 3 cases in all the models with KNN outperforming others. All the models were in agreement with respect to humidity in Edo state, giving a lag of 4 days. In Osun State, the models agree about the number of days for PM2.5 and temperature at 14 and 8 days lag respectively. In the case of humidity, DT reports the highest lag of 14 while KNN reports 1 day lag.

Generally, the best performing model in the southern location was SVM outperforming in five out of the nine cases, closely followed by KNN with 4 out of the nine cases. In all the southern locations, DT has the worst performance of all the models considered. However, in the northern locations, the best performing model was RF closely followed by KNN. Furthermore, based on RMSE values, temperature is a better predictor due to the low values reported. This is closely followed by relative humidity while PM2.5 is the worst predictor with consistently high RMSE values in all locations considered.

The global COVID-19 pandemic requires several approaches to understand, mitigate, and curtail its impact on world population and economy. There has been both pharmaceutical and non-pharmaceutical approaches to limiting the spread of the virus within population. Despite the lax enforcement of non-pharmaceutical interventions such as travel restrictions, the transmission level and fatality rate of the virus within Africa remains low compared to the rest of the world. This has been attributed to the youthful population (Njenga et al., 2020) and experience with pandemics (Musa et al., 2021). Merow and Urban (2020) posited that seasonality will drive the spread of COVID-19 globally. Considering the uncertainty surrounding the driving force of the infections globally, it will be pertinent to explore the role of atmospheric and aerosols. Furthermore, the potential to predict future occurrences of the infections based on past information about infection numbers or atmospheric conditions will help in mitigating the spread within a location. This study has shown that using machine learning algorithms the number of infections can be predicted with minimal error using previous 3–5 days case numbers. Furthermore, it was shown that previous days information about atmospheric and particulate matter are better predictors than 1-day data for COVID-19 cases. This information is important for better management and mitigation of the virus within the locations considered in this study.

4. Conclusion

In this study, we have examined the potential of machine learning approaches in predicting COVID-19 cases using atmospheric parameters within selected locations in Nigeria. Four machine learning techniques were considered: Decision tree, *k* Nearest Neighbor, support vector machine, and random forest. First, we determined the effect

of previous n-days COVID-19 cases on the forecast performance of the machine learning techniques. We found that for some locations the same number of lags were reported, however, in other locations different lags were obtained. Both KNN and SVM were found to have superior performance in this scenario. Furthermore, we evaluated the forecast capabilities of the machine learning techniques in COVID-19 cases prediction using atmospheric parameters as predictors at different lags. Decision tree method was found to have the worst performance of the four methods considered in this study. Our results presents a new approach to the study of COVID-19 virus by showing the amount of information required for effective prediction in machine learning algorithms. This is particularly important for the planning and management of the pandemic in tropical Nigeria. This research can be extended to consider predictors such as human mobility data. Furthermore, the possibility of other machine learning algorithms for prediction of COVID-19 can be explored. The study of this approach in other locations across the world will create global synergy in the fight against the virus.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

- (1) [Dataset] The particulate matter (PM1.0, 2.5, 10.0) and atmospheric parameters (temperature and humidity) were obtained from [purpleair.com](https://www.purpleair.com) (<https://www.purpleair.com/sensorlist?exclude=true&nwlat=9.37944961738998&selat=5.6830568335046365&nwlng=1.5013524736630188&selng=9.784479996099975&sensorsActive2=604800>).
- (2) [Dataset] The COVID-19 statistics were downloaded from <https://covid19.ncdc.gov.ng/state/>.

Acknowledgments

We acknowledge the Centre for Atmospheric Research and their partners for promoting high standards of atmospheric observatory practice as well as the Federal Government of Nigeria for continuous funding of the Nigerian Space programme (www.carnasrda.com). All data used in this study are publicly available.

References

- Anis, A. (2020). The effect of temperature upon transmission of Covid-19: Australia and Egypt case study. Available at SSRN 3567639.
- Annas, S., Pratama, M. I., Rifandi, M., Sanusi, W., & Side, S. (2020). Stability analysis and numerical simulation of seir model for pandemic Covid-19 spread in Indonesia. *Chaos, Solitons & Fractals*, *139*, 110072. <https://doi.org/10.1016/j.chaos.2020.110072>
- Anuforum, A. C. (2007). Spatial distribution and temporal variability of harmattan dust haze in sub-sahel west Africa. *Atmospheric Environment*, *41*(39), 9079–9090. <https://doi.org/10.1016/j.atmosenv.2007.08.003>
- Arslan, H., & Arslan, H. (2021). A new Covid-19 detection method from human genome sequences using cpg island features and knn classifier. *Engineering Science and Technology, an International Journal*, *24*(4), 839–847. <https://doi.org/10.1016/j.jestch.2020.12.026>
- ArunKumar, K., Kalaga, D. V., Kumar, C. M. S., Kawaji, M., & Brenza, T. M. (2021). Forecasting of Covid-19 using deep layer recurrent neural networks (rnns) with gated recurrent units (grus) and long short-term memory (lstm) cells. *Chaos, Solitons & Fractals*, *146*, 110861. <https://doi.org/10.1016/j.chaos.2021.110861>
- Bhatnagar, V., Poonia, R. C., Nagar, P., Kumar, S., Singh, V., Raja, L., & Dass, P. (2020). Descriptive analysis of Covid-19 patients in the context of India. *Journal of Interdisciplinary Mathematics*, *24*(3), 489–504. <https://doi.org/10.1080/09720502.2020.1761635>
- Cao, Y., & Francis, R. (2021). *On forecasting the community-level Covid-19 cases from the concentration of sars-cov-2 in wastewater* (p. 147451). Science of The Total Environment.
- Dansu, E. J., & Ogunjo, S. T. (2021). *Analysis of infectious disease problems (Covid-19) and their global impact*. In (Chap. Dynamics of inter-community spread of COVID-19). Springer, Nature.
- Dixit, A., Mani, A., & Bansal, R. (2021). *Cov2-detect-net: Design of Covid-19 prediction model based on hybrid de-pso with svm using chest x-ray images*. Information Sciences.
- Eludoyin, O. M., Adelekan, I. O., Webster, R., & Eludoyin, A. O. (2014). Air temperature, relative humidity, climate regionalization and thermal comfort of Nigeria. *International Journal of Climatology*, *34*(6), 2000–2018. <https://doi.org/10.1002/joc.3817>
- Ezeh, G., Abiye, O., & Obioh, I. (2017). Elemental analyses and source apportionment of pm2. 5 and pm2. 5–10 aerosols from nigerian urban cities. *Cogent Environmental Science*, *3*(1), 1323376. <https://doi.org/10.1080/23311843.2017.1323376>
- Fernandes, L. H., de Araújo, F. H., Silva, M. A., & Acioli-Santos, B. (2021). Covid-19 lethality in brazilian states using information theory quantifiers. *Physica Scripta*, *96*(3), 035003. <https://doi.org/10.1088/1402-4896/abd5ee>
- Fuwape, I., Okpalaonwuka, C., & Ogunjo, S. (2021). Impact of Covid-19 pandemic lockdown on distribution of inorganic pollutants in selected cities of Nigeria. *Air Quality, Atmosphere & Health*, *14*(2), 149–155. <https://doi.org/10.1007/s11869-020-00921-8>
- Gupta, V. K., Gupta, A., Kumar, D., & Sardana, A. (2021). Prediction of Covid-19 confirmed, death, and cured cases in India using random forest model. *Big Data Mining and Analytics*, *4*(2), 116–123. <https://doi.org/10.26599/bdma.2020.9020016>
- Hashim, S., Farooq, S., Syriopoulos, E., Cremer, K. d. I. L., Vogt, A., de Jong, N., et al. (2021). *Machine learning model for predicting number of Covid19 cases in countries with low number of tests*. medRxiv. Retrieved from <https://www.medrxiv.org/content/early/2021/07/14/2021.07.12.21260298>
- Iwendi, C., Bashir, A. K., Peshkar, A., Sujatha, R., Chatterjee, J. M., Pasupuleti, S., et al. (2020). Covid-19 patient health prediction using boosted random forest algorithm. *Frontiers in Public Health*, *8*, 357. <https://doi.org/10.3389/fpubh.2020.00357>
- Karnon, J. (2020). A simple decision analysis of a mandatory lockdown response to the Covid-19 pandemic. *Applied Health Economics and Health Policy*, *18*(3), 329–331. <https://doi.org/10.1007/s40258-020-00581-w>
- Kırbaş, İ., Sözen, A., Tuncer, A. D., & Kazancıoğlu, F. Ş. (2020). Comparative analysis and forecasting of Covid-19 cases in various European countries with arima, narnn and lstm approaches. *Chaos, Solitons & Fractals*, *138*, 110015. <https://doi.org/10.1016/j.chaos.2020.110015>

- Koren, I., Kaufman, Y. J., Washington, R., Todd, M. C., Rudich, Y., Martins, J. V., & Rosenfeld, D. (2006). The bodélé depression: A single spot in the sahara that provides most of the mineral dust to the amazon forest. *Environmental Research Letters*, *1*(1), 014005. <https://doi.org/10.1088/1748-9326/1/1/014005>
- Li, W., Gong, J., Zhou, J., Zhang, L., Wang, D., Li, J., et al. (2021). An evaluation of Covid-19 transmission control in wenzhou using a modified seir model. *Epidemiology and Infection*, *149*, e2. <https://doi.org/10.1017/s0950268820003064>
- López, L., & Rodo, X. (2021). A modified seir model to predict the Covid-19 outbreak in Spain and Italy: Simulating control scenarios and multi-scale epidemics. *Results in Physics*, *21*, 103746. <https://doi.org/10.1016/j.rinp.2020.103746>
- Merow, C., & Urban, M. C. (2020). Seasonality and uncertainty in global Covid-19 growth rates. *Proceedings of the National Academy of Sciences*, *117*(44), 27456–27464. <https://doi.org/10.1073/pnas.2008590117>
- Muhammad, L., Islam, M. M., Usman, S. S., & Ayon, S. I. (2020). Predictive data mining models for novel coronavirus (Covid-19) infected patients' recovery. *SN Computer Science*, *1*(4), 1–7. <https://doi.org/10.1007/s42979-020-00216-w>
- Musa, H. H., Musa, T. H., Musa, I. H., Musa, I. H., Ranciaro, A., & Campbell, M. C. (2021). Addressing Africa's pandemic puzzle: Perspectives on Covid-19 transmission and mortality in sub-saharan Africa. *International Journal of Infectious Diseases*, *102*, 483–488. <https://doi.org/10.1016/j.ijid.2020.09.1456>
- Njenga, M. K., Dawa, J., Nanyingi, M., Gachohi, J., Ngere, I., Letko, M., et al. (2020). Why is there low morbidity and mortality of Covid-19 in Africa? *The American Journal of Tropical Medicine and Hygiene*, *103*(2), 564–569. <https://doi.org/10.4269/ajtmh.20-0474>
- Ogunjo, S. T., Fuwape, I. A., & Olusegun, C. F. (2019). *Impact of large scale climate oscillation on drought in west africa*. arXiv preprint arXiv:1901.10145.
- Ologunorisa, T. E. (2001). A review of the effects of gas flaring on the Niger delta environment. *The International Journal of Sustainable Development and World Ecology*, *8*(3), 249–255. <https://doi.org/10.1080/13504500109470082>
- Orimoloye, I. R., Mazinyo, S. P., Kalumba, A. M., Ekundayo, O. Y., & Nel, W. (2019). Implications of climate variability and change on urban and human health: A review. *Cities*, *91*, 213–223. <https://doi.org/10.1016/j.cities.2019.01.009>
- Pandey, A., Kadir, T., Kumar, R., & Sinwar, D. (2022). Analyzing effects of temperature, humidity, and urban population in the initial outbreak of Covid19 pandemic in India. In *Data engineering for smart systems* (pp. 469–478). Springer.
- Pandey, G., Chaudhary, P., Gupta, R., & Pal, S. (2020). *Seir and regression model based Covid-19 outbreak predictions in India*. arXiv preprint arXiv:2004.00958.
- Peiliang, S., & Li, K. (2020). *An seir model for assessment of current Covid-19 pandemic situation in the UK*. medRxiv.
- Pourhomayoun, M., & Shakibi, M. (2021). Predicting mortality risk in patients with Covid-19 using machine learning to help medical decision-making. *Smart Health*, *20*, 100178. <https://doi.org/10.1016/j.smhl.2020.100178>
- Prakash, K. B., Imambi, S. S., Ismail, M., Kumar, T. P., & Pawan, Y. (2020). Analysis, prediction and evaluation of Covid-19 datasets using machine learning algorithms. *International Journal*, *8*(5), 2199–2204. <https://doi.org/10.30534/ijeter/2020/117852020>
- Rahal, F., Rezak, S., & Baba Hamed, F. Z. (2021). Impact of meteorological parameters on the Covid-19 incidence: The case of the city of oran, Algeria. *Journal of Clinical & Experimental Investigations*, *12*(1), em00762. <https://doi.org/10.29333/jcei/9562>
- Romadhon, M. R., & Kurniawan, F. (2021). A comparison of naive bayes methods, logistic regression and knn for predicting healing of Covid-19 patients in Indonesia. In *2021 3rd east Indonesia conference on computer and information technology (eiconcit)* (pp. 41–44).
- Ropo, O. I., Perez, M. S., Werner, N., & Enoch, T. (2017). Climate variability and heat stress index have increasing potential ill-health and environmental impacts in the east London, south Africa. *International Journal of Applied Engineering Research*, *12*(17), 6910–6918.
- Runkle, J. D., Sugg, M. M., Leeper, R. D., Rao, Y., Matthews, J. L., & Rennie, J. J. (2020). Short-term effects of specific humidity and temperature on Covid-19 morbidity in select us cities. *The Science of the Total Environment*, *740*, 140093. <https://doi.org/10.1016/j.scitotenv.2020.140093>
- Şahin, M. (2020). Impact of weather on Covid-19 pandemic in Turkey. *The Science of the Total Environment*, *728*, 138810. <https://doi.org/10.1016/j.scitotenv.2020.138810>
- Said, A. B., Erradi, A., Aly, H. A., & Mohamed, A. (2021). Predicting Covid-19 cases using bidirectional lstm on multivariate time series. *Environmental Science and Pollution Research*, *28*(40), 1–10. <https://doi.org/10.1007/s11356-021-14286-7>
- Shaban, W. M., Rabie, A. H., Saleh, A. I., & Abo-Elsoud, M. (2020). A new Covid-19 patients detection strategy (cpds) based on hybrid feature selection and enhanced knn classifier. *Knowledge-Based Systems*, *205*, 106270. <https://doi.org/10.1016/j.knsys.2020.106270>
- Singh, S., Parmar, K. S., Makkhan, S. J. S., Kaur, J., Peshoria, S., & Kumar, J. (2020). Study of arima and least square support vector machine (ls-svm) models for the prediction of sars-cov-2 confirmed cases in the most affected countries. *Chaos, Solitons & Fractals*, *139*, 110086. <https://doi.org/10.1016/j.chaos.2020.110086>
- Singh, V., Poonia, R. C., Kumar, S., Dass, P., Agarwal, P., Bhatnagar, V., & Raja, L. (2020). Prediction of Covid-19 corona virus pandemic based on time series data using support vector machine. *Journal of Discrete Mathematical Sciences and Cryptography*, *23*(8), 1–15. <https://doi.org/10.1080/09720529.2020.1784535>
- Smith, M., & Alvarez, F. (2021). Identifying mortality factors from machine learning using shapley values—a case of Covid19. *Expert Systems with Applications*, *176*, 114832. <https://doi.org/10.1016/j.eswa.2021.114832>
- Stufano, A., Lisco, S., Bartolomeo, N., Marsico, A., Lucchese, G., Jahantigh, H., et al. (2021). Covid19 outbreak in lombardy, Italy: An analysis on the short-term relationship between air pollution, climatic factors and the susceptibility to sars-cov-2 infection. *Environmental Research*, *198*, 111197. <https://doi.org/10.1016/j.envres.2021.111197>
- Sunnu, A., Resch, F., & Afeti, G. (2013). Back-trajectory model of the saharan dust flux and particle mass distribution in west Africa. *Aeolian Research*, *9*, 125–132. <https://doi.org/10.1016/j.aeolia.2013.02.005>
- Tosepu, R., Gunawan, J., Effendy, D. S., Ahmad, L. O. A. I., Lestari, H., Bahar, H., et al. (2020). Correlation between weather and Covid-19 pandemic in jakarta, Indonesia. *The Science of the Total Environment*, *725*, 138436. <https://doi.org/10.1016/j.scitotenv.2020.138436>
- Wan, K., Chen, J., Lu, C., Dong, L., Wu, Z., & Zhang, L. (2020). When will the battle against novel coronavirus end in wuhan: A seir modeling analysis. *Journal of global health*, *10*(1). <https://doi.org/10.7189/jogh.10.011002>
- Yadav, M., Perumal, M., & Srinivas, M. (2020). Analysis on novel coronavirus (Covid-19) using machine learning methods. *Chaos, Solitons & Fractals*, *139*, 110050. <https://doi.org/10.1016/j.chaos.2020.110050>
- Yahya, B. M., Yahya, F. S., & Thannoun, R. G. (2021). Covid-19 prediction analysis using artificial intelligence procedures and gis spatial analyst: A case study for Iraq. *Applied Geomatics*, *1–11*(3), 481–491. <https://doi.org/10.1007/s12518-021-00365-4>
- Yeşilkanat, C. M. (2020). Spatio-temporal estimation of the daily cases of Covid-19 in worldwide using random forest machine learning algorithm. *Chaos, Solitons & Fractals*, *140*, 110210. <https://doi.org/10.1016/j.chaos.2020.110210>
- Yoo, S. H., Geng, H., Chiu, T. L., Yu, S. K., Cho, D. C., Heo, J., et al. (2020). Deep learning-based decision-tree classifier for Covid-19 diagnosis from chest x-ray imaging. *Frontiers of Medicine*, *7*, 427. <https://doi.org/10.3389/fmed.2020.00427>
- Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting Covid-19 time-series data: A comparative study. *Chaos, Solitons & Fractals*, *140*, 110121. <https://doi.org/10.1016/j.chaos.2020.110121>