Scientific African 13 (2021) e00914

Contents lists available at ScienceDirect

Scientific African

journal homepage: www.elsevier.com/locate/sciaf

# The dynamics of COVID-19 outbreak in Nigeria: A sub-national analysis

Kayode P. Ayodele[a,*], Hafeez Jimoh[b], Adeniyi F. Fagbamigbe[c], Oluwatoyin H. Onakpoya[d]

[a] *Department of Electronic, and Electrical Engineering Obafemi Awolowo University, Nigeria*
[b] *Intelligent Robotic Systems Laboratory Innopolis University Innopolis City, Russian Federation*
[c] *Department of Epidemiology and Medical Statistics University of Ibadan, Nigeria*
[d] *Department of Medicine Obafemi Awolowo University, Nigeria*

ARTICLE INFO

*Article history:*
Received 10 September 2020
Revised 28 February 2021
Accepted 28 July 2021

Editor: DR B Gyampoh

*Keywords:*
COVID-19
Nigeria
SIRD model
Temperature
Absolute humidity

ABSTRACT

The African health crisis feared at the beginning of the COVID-19 pandemic has not materialized, and there is interest globally in understanding possible peculiarities in COVID-19 outbreak dynamics in the tropics and sub-tropics that have led to a much milder African outbreak than initial projections. Towards this, Susceptible-Infected-Recovered-Dead compartmental models were fitted to COVID-19 data from all Nigerian states in this study, from which four parameters were estimated per state. A density-based clustering method was used to identify states with similar outbreak dynamics, and the stage of the outbreak determined per state. Subsequently, outbreak dynamics were correlated with absolute humidity, temperature, population density and distance to the international passenger travel gateways in the country. The models revealed that while the outbreak is still increasing nationally, outbreaks in at least 12 states have peaked. A total of at least 519,672 confirmed cases were predicted by January 2021, with a worst case scenario of at least 14,785,457. Weak positive correlations were found between COVID-19 spread and absolute humidity (Pearson's Coefficient = 0.136, $p < 0.05$) and temperature (Pearson's Coefficient = 0.021, $p < 0.05$). While many studies have established links between temperature and humidity and COVID-19 spread, the correlation has most usually been negative where it exists. The findings in this study of possible positive correlation is in line with a number of previous studies showing such unexpected correlations in the tropics or subtropics. This highlights even more the importance of additional studies on COVID-19 dynamics in Africa.



## Introduction

The coronavirus disease (COVID-19) first broke out in the city of Wuhan, Hubei Province in China in late December 2019 [20]. With rapid spread across countries especially across Europe, COVID-19 was declared a pandemic by the WHO

* Corresponding author.
*E-mail address:* kayodele@oauife.edu.ng (K.P. Ayodele).

https://doi.org/10.1016/j.sciaf.2021.e00914

on March 12 2020 [20]. It has since spread to nearly all countries and continents. The virus has continued to wreak havoc in different countries at an alarming pace [3]. As of August 12 2020, there are 20,836,339 cases of the virus with nearly 750,000 causalities [32]. The first case of the virus emerged in Nigeria on February 28, 2020 and has increased rapidly within 6 months to 47,743 cases with 979 deaths as of August 12, 2020 [32].

Compared with the number of cases and causalities in the Central America and Europe, Africa currently has a lower burden of COVID-19. These may be ascribed to differences in environmental conditions and the fact that the breakout started later in Africa than most places thereby providing a window of opportunity for preparedness and mitigation efforts such as lock-down. However, Africa has the largest proportion of less developed countries than other continents. The continent nonetheless suffer dearth of medical supplies, very low baseline of and access to hospitalisation capacity, particularly intensive and sub-intensive care. Other parameters such as larger household sizes, higher intergenerational mixing within households, poorer environmental conditions including overcrowded urban settlements, inadequate water and sanitation, pre-existing disease burden with higher prevalence of both undiagnosed, poorly-managed and unmanaged noncommunicable diseases. This health outcomes may be risk factors for COVID-19 severity. Bearing in mind that Africa had the highest burden of infectious diseases, such as HIV, TB, malaria, ebola etc, which might have a negative impact on the longtime severity of COVID-19. There is need for multi-sectoral efforts to stimulate understanding of the spread and severity of the virus in Africa. One of such efforts is modelling of the different characteristics of the virus.

Beside sharing the peculiarities of other Africa countries, the fragile healthcare systems in Nigeria is beginning to be overwhelmed. There are concerns that the current situation may worsen. Nigeria, as the most populous African country, occupies a delicate and strategic position in the continent. An inefficient management of the pandemic may affect other African countries negatively. Very central to the efforts targeted at developing, planning and implementing containment and mitigation measures in Nigeria is understanding and modelling the spread of the virus. Nigeria is however diverse in terms of access to health care, household structures, geographical features, weather, etc. We hypothesized that these differences may affect the spread, recovery from and severity of COVID-19 across the different States in Nigeria. The current study is therefore aimed at modelling and understanding the within-country dynamics of COVID-19 outbreak in Nigeria in terms of number of cases, number of recoveries and number of deaths from COVID-19.

Modelling the spread of the virus worldwide has remained a big task because most parameters about the virus is not known. Since the virus was declared a pandemic, modellers consisting of engineers, mathematicians, Statisticians and data scientists have been presented with daunted task of understanding and modelling the nature, the spread as well as other characteristics of the virus [11].

Several approaches have been engaged in modelling COVID-19 since its outbreak [6,22,33,35,36]. Modelling of an infectious disease, irrespective of its purpose to understand, track and predict its behaviour and behaviour,is very paramount to strategies to control and mitigate the spread of the disease. According to [21], the earliest infectious modelling efforts was in 1662 by John Graunt [21], mathematical modelling of the spread of diseases by Bernoulli in 1766 [23], and the popular foundation of compartmental modelling of epidemics between 1927 and 1933 [14–16]. More recently, different modelling strategies have been developed. The strategies are dynamic and are diverse [23]. Siettos et al. categorised the recent modelling strategies into: (i) statistical-based methods for epidemic surveillance, (ii) mathematical and mechanistic state-space models, and (iii) empirical and machine learning-based methods [23]. The most popular infectious disease models (including those used by the WHO) employed the SIR (Susceptible - Infectious - Recovered), SEIR (Susceptible - Exposed - Infectious - Recovered) and SIRD (Susceptible-Infected-Recovered-Dead) models. They followed establishment of the basic reproduction number, assessment of herd immunity as well as significant clusters. Fong et al. and Wang et al. had used this approach to predict infection rate and spread [11], [29]. The current study utilized the SIRD model.

## Methods

### *0.1. Data sources*

Data on the COVID-19 outbreak in Nigeria were obtained from the COVID-19 microsite of the Nigeria Center for Diseases Control (NCDC) [8]. State-wise data were extracted from individual daily reports for the period between 1st March, 2020 and August 10, 2020. COVID-19 country data for Nigeria were also obtained from the Johns Hopkins University Coronavirus Resource Center repository [28] to validate the extracted NCDC data.

Historical temperature and relative humidity data were obtained from the National Centers for Environmental Information (NCEI) library through the Visual Crossing Weather Data Services web application [5]. Monthly average data for 36 states and the Federal Capital Territory (FCT) were extracted. Absolute humidity was estimated from relative humidity using the Clausius Clapeyron conversion relation [12,13]:

$$AH = \frac{13.2471 RH e^{\left(\frac{17.67T}{T+243.5}\right)}}{273.15 + T} \tag{1}$$

State population and geographical information data were were obtained from the Nigeria National Bureau of Statistics (NBS) [25]. Population values for the period 2011 - 2015 were obtained and the value for 2020 estimated by linear extrapolation.

### 0.2. Compartmental model and notation

The Susceptible-Infected-Recovered-Dead (SIRD) model is a standard infectious disease model for analysing infectious disease outbreak dynamics by tracking the variations with time of the four eponymous variables [10]. In this model, outbreak dynamics are modelled with the following four ordinary differential equations:

$$\frac{dS}{dT} = -\frac{\beta SI}{N} \tag{2}$$

$$\frac{dI}{dT} = \frac{\beta SI}{N} - \gamma I - \alpha I \tag{3}$$

$$\frac{dR}{dT} = \gamma I \tag{4}$$

$$\frac{dD}{dT} = \alpha I \tag{5}$$

where T is time elapsed since the outbreak started, $\alpha$ is the mortality rate, $\beta$ is the effective contact rate, and $\gamma$ is the recovery rate. S, I, R, F, N are the susceptible, infected, recovered, dead, and total populations respectively. Also,

$$S + I + R + D = N \tag{6}$$

In order to more easily compare outbreak dynamics across areas with different populations, the following non-dimensional variation of the SIRD model was adopted in this study:

$$\frac{ds}{dt} = -\rho si \tag{7}$$

$$\frac{di}{dt} = \rho si - \sigma i - \kappa I \tag{8}$$

$$\frac{dr}{dt} = \sigma i \tag{9}$$

$$\frac{d\mathscr{D}}{dt} = \kappa i \tag{10}$$

in which $t = T/\tau$ where $\tau$ is a time scaling constant of convenient value. Also, $s = \frac{S}{N}, i = \frac{I}{N}, r = \frac{R}{N}, \mathscr{D} = \frac{D}{N}$, while $\kappa = \alpha\tau, \rho = \beta\tau$, and $\sigma = \gamma\tau$ are population-normalized versions of the mortality, effective contact, and recovery rates.

The reproduction rate, $R_0$ was estimated as follows:

$$R_0 = \frac{\rho}{\sigma + \kappa} \tag{11}$$

Case fatality rate was computed as the ratio of the number of deaths and the number of confirmed cases. An distinction is made between the number "infected" cases (individuals currently infected) and the "confirmed" cases (cumulative sum of all individuals ever confirmed to be infected, whether or not they have recovered, died, or remain infected).

### 0.3. Data analysis

Visual inspection of daily case reports revealed that the COVID-19 data were generally noisy. One incidence deserves special mention. On August 3, 2020, the number of discharged patients for Lagos state increased by 10946. The mean daily increment prior to that was 13 patients. The following annotation accompanied the 3rd August discharge number: *"Number includes recoveries from treatment centre and community recoveries managed at home"* [7]. Evidently, Lagos state must have taken some time to arrive at a reasonably accurate estimation of the number of cases treated outside official healthcare centers in the state, triggering the adjustment by the NCDC on August 3. This is justifiable, but nevertheless creates certain challenges for COVID-19 outbreak modelling for Lagos and Nigeria.

A one-day spike of that magnitude in discharge numbers is inconsistent with normal infectious disease outbreak dynamics. In this study, two version of the Lagos data were used. The first version employed the raw data as reported by the NCDC, including the spike on 3rd August. In addition, a new dataset was introduced in which the Lagos data was adjusted. Using a ramp function, the 10946-case spike was spread across multiple days. Starting from 1st April and ending on August 3, the 10,946 discharged cases were added to previously recorded discharges for the days, so that the number of cases added for each day increased by a fixed value. Consequently, this study used data for 39 "regions": the FCT, 36 states unmodified, a $38^{th}$ fictitious state termed "Lagos-Adj" for the modified Lagos data, and the cumulative nationwide data for comparison where appropriate.

Daily information for all states were plotted showing confirmed cases, discharged and deaths for a total of 38 states and the national cumulative. Subsequently, each of the 39 datasets were fitted to SIRD models represented by Eqs. 7–10. The value of $\tau$ was set to 180. Parameter estimation for $\kappa, \rho, \sigma$ was carried out on dataset data using the hyperparameter optimization method implemented by the Optuna Python programming language package [1] with a root mean squared log error (RMSLE) cost function, which is robust to the effects of outliers [24].

Compartmental modelling was carried out in the Covsirphy infectious disease modelling library [26]. The SIRD model parameters were assumed to be time-varying throughout the outbreak. However, in order to make the models tractable, a small number of inflection point were assumed, at which model parameters changed. Detection of these inflection points was achieved using phase plane analysis in the SR plane [4] as follows. From Eqs. 2 and 4,

$$\frac{dS}{dR} = -\frac{\beta S}{N\gamma} \tag{12}$$

From which

$$S(R) = Ne^{-\frac{R\beta}{N\gamma}} \tag{13}$$

By plotting this function for each dataset, the inflection points were determined, and the intervening periods between points defined as phases during which the parameters $\kappa, \rho, \sigma$ and $R_0$ were assumed constant. Furthermore, a growth factor time series was computed for each dataset. Where $\Delta C_n$ was increase in the number of confirmed COVID-19 cases between day $n-1$ and day $n$, the growth factor $F_g$ was defined as:

$$F_g = \frac{\Delta C_n}{\Delta C_{n-1}} \tag{14}$$

The current outbreak stage per state was categorized as expansion, contraction or indeterminate on the basis of growth factor values over rolling 7-day epochs, and using the reproduction number, $R_0$. For investigating similarities between the outbreaks in different states, 4-dimensional vectors created from the average values of $\kappa, \rho, \sigma$ and $R_0$ for each state were clustered by means of hierarchical density based clustering [19]. Furthermore, possible relationships between the outbreak and various factors were investigated using linear correlation analysis, including temperature, absolute humidity, and distance from Nigeria's commercial and administrative capitals.

Using the values of the model parameters as at August 10, simulations were run with the SIRD models to project the size of the different populations till the end of December 2020.

## Results

Between March 1, and August 10, a total of 46,866 confirmed cases COVID-19 of were reported in Nigeria, from a total of 321,950 tests conducted. There were confirmed COVID-19 cases in every state and the FCT. As at August 10, 2020, there were 33,346 discharged cases, and 950 confirmed fatalities, leading to a case fatality rate of 2.0% [9]. Figs. 1 and 2 present the cumulative numbers of confirmed cases, recovered (discharged) cases, and deaths over time for 38 states and nationwide.

The values of parameters $\kappa, \rho, \sigma$, and $R_0$ for the duration of outbreak evolution in each state are presented in Fig. 3. To improve the figure, each parameter was scaled by an appropriate factor (indicated at the bottom). There is no plot for Kogi state because the low number of cases prevented the model from converging. Case fatality rates as of August 10, are presented in Fig. 4. There is only one entry for Lagos because the August 3, 2020 adjustment affected neither the total number of confirmed cases nor the deaths from COVID-19.

Figs 5 and 6 present the S-R curves generated for trend analysis of the Lagos and Lagos-Adj datasets respectively. The effect of the spike in discharged cases on August 3, on the trend analysis is evident, as the Lagos outbreak was segmented into 4 phases, rather than the 9 phases identified in the adjusted Lagos data. Similarly, the effect of the adjustment on the reproduction number of Lagos can be seen in Fig. 3 differing by an order of magnitude (1.24 versus 14.77).

Outbreak dynamics in certain states were inordinately affected by the low number of COVID-19 cases. Nowhere is this more evident than in the case fatality rate for Kogi state (40% in Fig. 4), a figure that was obviously impacted by the unreliability of most statistics at small sample sizes [31]. In fact, of the states with the highest CFR (Anambra, Cross River, Kebbi, Sokoto, Taraba, Yobe, Zamfara) only Anambra has more than 100 confirmed cases as of August 10, 2020. Consequently, all states with less than 100 cases were excluded from further analysis. The relationship between CFR and number of confirmed cases was investigated by means of a correlation analysis, and revealed a weak but statistically significant relationship between number of cases and CFR (Pearson coefficient of -0.22; T-statistic = 84.44; p-value = 4.0 x $10^{-}$42).

Table 1 presents the result of the analysis of current outbreak stage per state. Growth factor analysis categorized 13, 5, 17 states as contraction, expansion, and indeterminate stages respectively. Using the more traditional reproduction number approach in which outbreak contraction is determined by $R_0 < 1$, 22 state outbreaks were classified as contracting, while 13 were still expanding. A cluster analysis using the parameters of the SIRD model as feature vectors resulted in five clusters as shown in Fig. 7.

The population densities, average temperature and absolute humidity for the periods between March and August 2020 for each of the 36 states and the FCT are presented in Table 1. Also in Table 1 are the average beeline distances between the respective state capitals and Lagos and Abuja, where the two primary international passenger traffic gateways of the

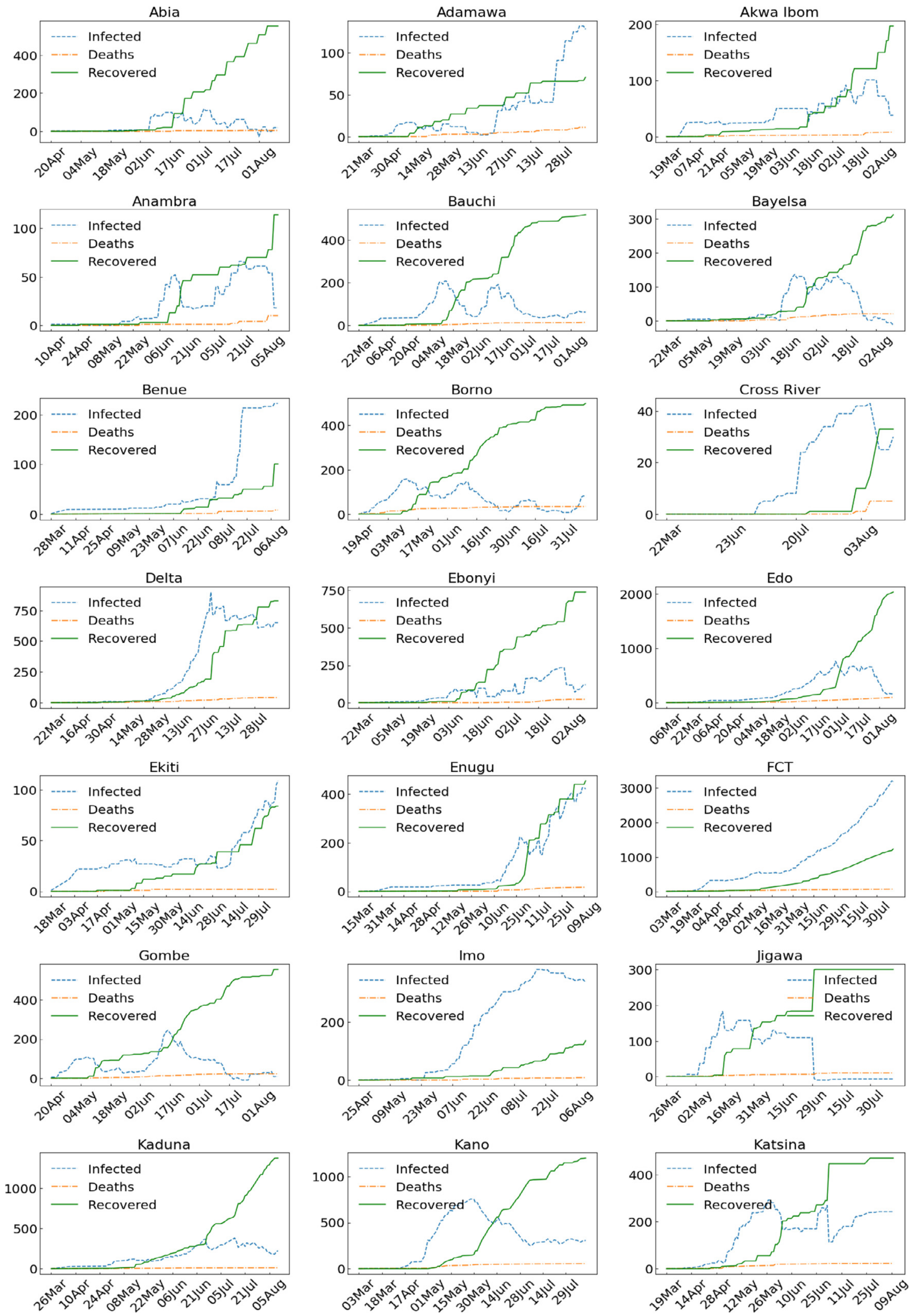


**Fig. 1.** Daily variations in the Infected, Discharge and Death populations for 20 datasets.

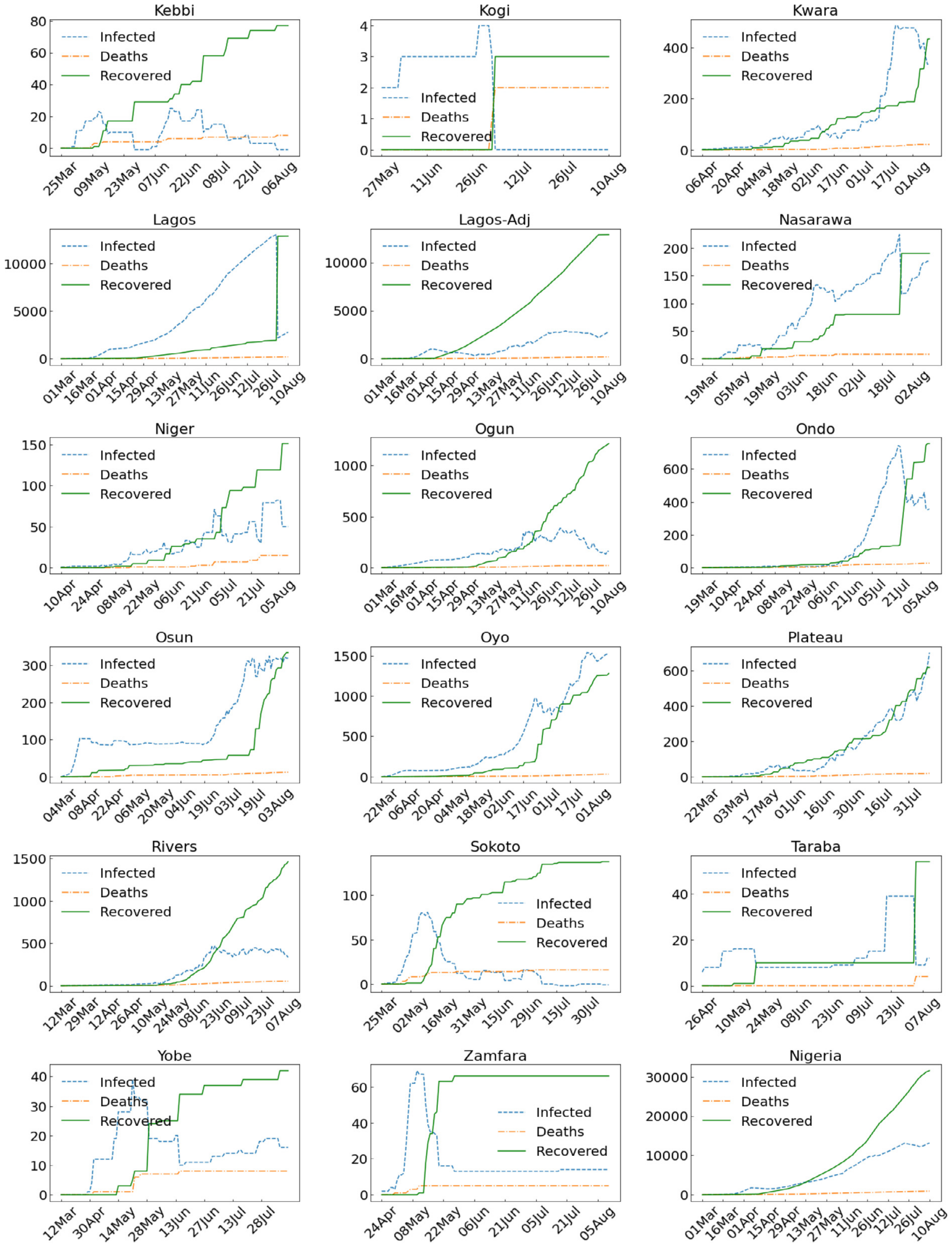


**Fig. 2.** Daily variations in the Infected, Discharge and Death populations for remaining 19 datasets.

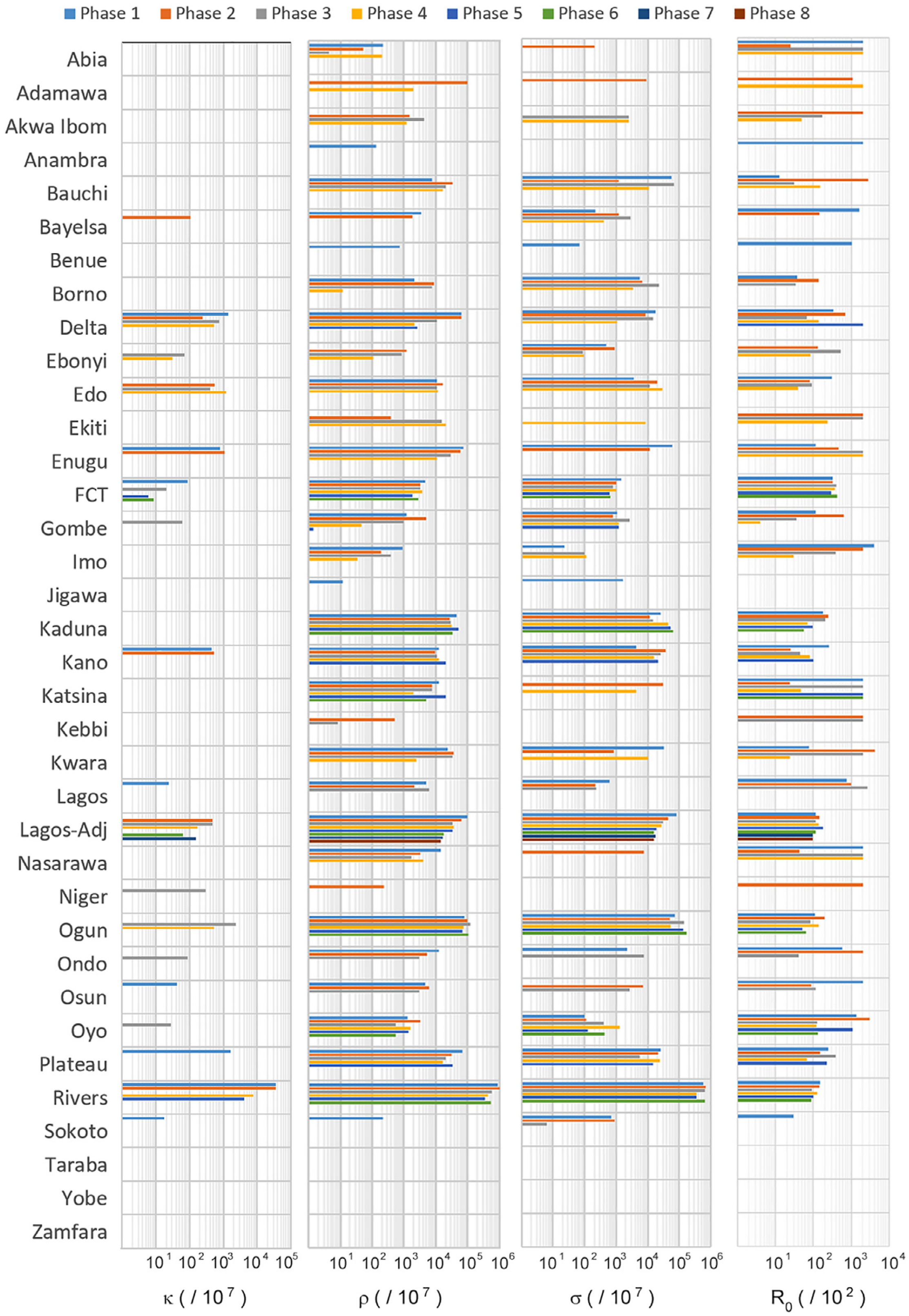


**Fig. 3.** Variations in the parameters $\kappa$, $\rho$, $\sigma$ and $R_0$ for different phases across the different states. Cross River and Kogi States are excluded.

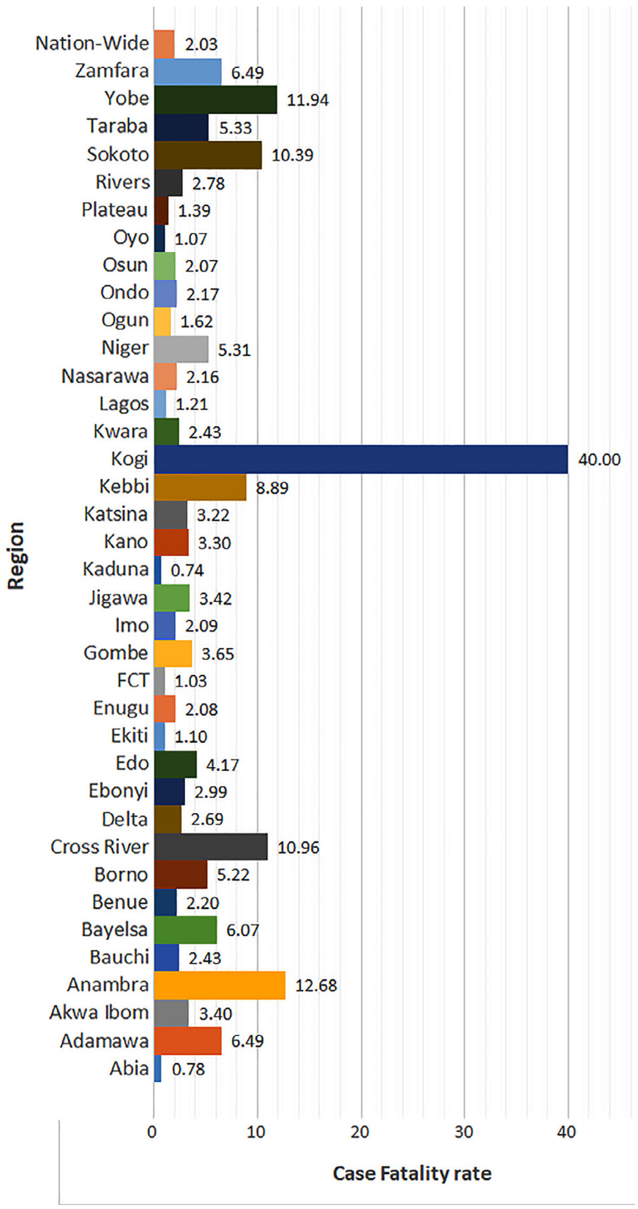


**Fig. 4.** Case fatality rates for 36 states, FCT, and national average.

country are located. The Pearson's coefficient for the correlation analysis between $\rho$ and four variables are presented in Table [2]. From Eqs. 7 and 8, $\rho$ is the parameter in the SIRD model most associated with the rate of outbreak expansion. No relationship was found between $\rho$ and the population densities of states ($p > 0.05$). The table however suggests a weak negative correlation between both temperature and absolute humidity ($p < 0.05$).

Table 4 presents the results of simulations of the SIRD models for states till the end of 2020 assuming that model parameter values on August 10, 2020 stay constant. It should be noted that Kogi and Cross River state simulations failed to converge due to insufficient data, so the Total row is sans those two states.

## Discussion

Generally speaking, variations can be seen in all important COVID-19 outbreak parameters and metrics across Nigeria. Whether those variations are sufficient for accurate inference however depends on important extenuating factors including the current extent of the outbreak per location, and the quality of data collation and management. For the first factor, the outbreaks in certain state are too small to accord statistical significance to certain parameters, most notably in the case of

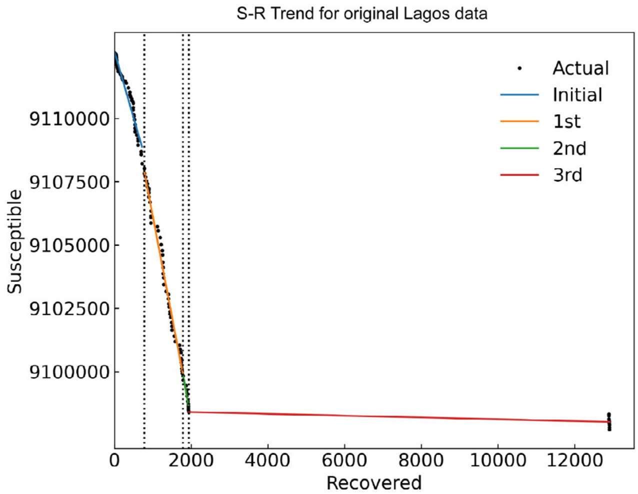


**Fig. 5.** Case fatality rates for 36 states, FCT, and national average.

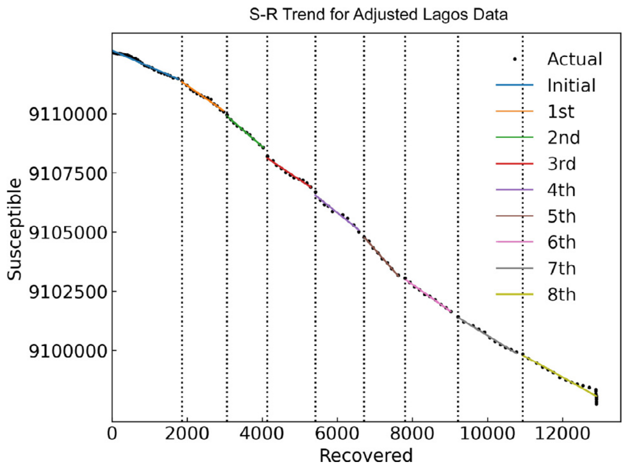


**Fig. 6.** Case fatality rates for 36 states, FCT, and national average.

Kogi state, but also in most states with less than 100 confirmed cases. The question of data quality comes up primarily in the case of Lagos State, since the one-time adjustment of August 3, 2020 significantly affected multiple outbreak parameters not just in Lagos, but across the whole country. The problem can however be glimpsed in sudden spikes and drops in different categories of numerous state data such as Nasarawa, Ondo, Taraba, Yobe and Zamfara states (Figs. 1 and 2). More important than the magnitudes of these occasional spikes and drops is what their existence suggests about the accuracy with which certain records are generated and propagated to the NCDC. The immediate consequence of these extenuating factors is that trends and effects that might otherwise have emerge from analysis may be occluded.

Notwithstanding the noisy nature of the data, some useful trends and patterns emerged. Hierarchical density based clustering identifies underlying patterns and dynamics in data that are often not discernible from visual inspection or intuitive. Nevertheless, hints of geo-cultural influences emerged from the clustering of outbreak dynamics across Nigeria, particularly if Clusters B and C are treated as offshoots of Cluster A. The states making up the clusters generally form a contiguous block linking to either Lagos or Abuja,the commercial and administrative capitals of Nigeria. Cluster D are mainly in the geographical middle of Nigeria, with the exception of Delta State. In addition, most Cluster E states are in a contiguous block in the South Eastern part of the country, with three exceptions. Also notable is the facts that most states in the cluster had late-breaking outbreaks, again alluding to hidden geo-cultural variables.

Elucidation of the full underlying factors responsible for the possible dynamic similarities requires further investigation, but the above-mentioned cluster patterns form a justification for such an endeavour. Very likely, the patterns may relate more to policy, mitigating strategies and cultural aspects of the response of the populations of those states. For this study however, the clusters might provide a qualitative backdrop to discuss the outbreak in each state. For example, Lagos and Ogun have both have more cases per capita than other states in Cluster A. In addition, the high population density of Lagos

**Table 1**
Determination of the stages of outbreaks per state using growth factor and using $R_0$.

| State | Outbreak Stage (Growth Factor) | Outbreak Stage ($R_0$) |
|---|---|---|
| Abia | Indeterminate | Expansion |
| Adamawa | Contraction | Expansion |
| Akwa Ibom | Contraction | Contraction |
| Anambra | Contraction | Contraction |
| Bauchi | Expansion | Expansion |
| Bayelsa | Contraction | Contraction |
| Benue | Contraction | Contraction |
| Borno | Indeterminate | Contraction |
| Delta | Indeterminate | Expansion |
| Ebonyi | Contraction | Contraction |
| Edo | Indeterminate | Contraction |
| Ekiti | Indeterminate | Expansion |
| Enugu | Contraction | Expansion |
| FCT | Expansion | Expansion |
| Gombe | Contraction | Expansion |
| Imo | Contraction | Contraction |
| Jigawa | Indeterminate | Contraction |
| Kaduna | Indeterminate | Contraction |
| Kano | Indeterminate | Contraction |
| Katsina | Contraction | Expansion |
| Kebbi | Indeterminate | Contraction |
| Kwara | Indeterminate | Contraction |
| Lagos | Indeterminate | Contraction |
| Nasarawa | Indeterminate | Expansion |
| Niger | Contraction | Contraction |
| Ogun | Expansion | Contraction |
| Ondo | Contraction | Contraction |
| Osun | Indeterminate | Expansion |
| Oyo | Expansion | Expansion |
| Plateau | Expansion | Expansion |
| Rivers | Indeterminate | Contraction |
| Sokoto | Indeterminate | Contraction |
| Taraba | Contraction | Contraction |
| Yobe | Indeterminate | Contraction |
| Zamfara | Indeterminate | Contraction |
| Nationwide | Indeterminate | Expanding |

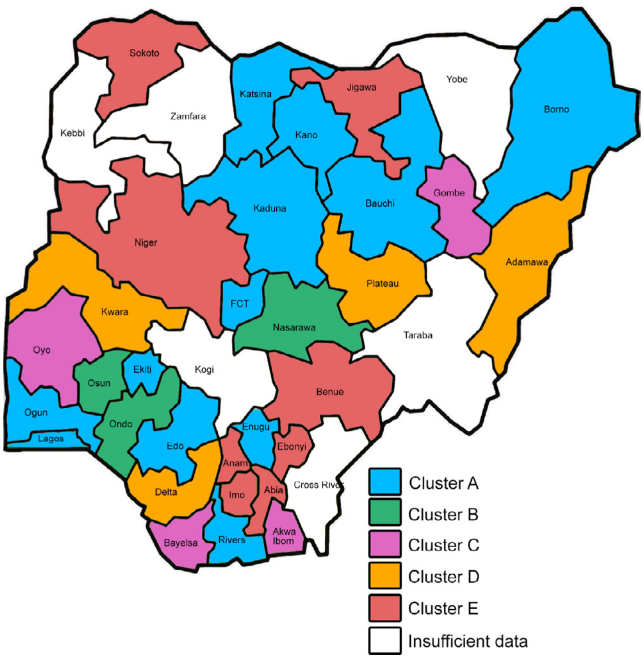


**Fig. 7.** Clustering of COVID-19 outbreak dynamics in different states using density based clustering.

**Table 2**
Average absolute humidity and temperature of different states and distances to Lagos and the FCT.

| State | AH ($g/m^3$) | Temp. ($^\circ$ C) | Distance to Abuja (km) | Distance to Lagos (km) |
|---|---|---|---|---|
| Abia | 21.48 | 27.76 | 392.91 | 465.53 |
| Adamawa | 20.52 | 29.92 | 544.25 | 1046.1 |
| Akwa Ibom | 21.34 | 27.44 | 450.89 | 526.52 |
| Bauchi | 5.51 | 32.83 | 291.89 | 829.5 |
| Bayelsa | | 28.12 | 360.83 | 478.36 |
| Benue | 21.26 | 30.04 | 184.71 | 584.36 |
| Borno | 11.97 | 35.91 | 689.84 | 1226.59 |
| Cross River | 21.34 | 27.44 | 465.2 | 569.63 |
| Ebonyi | 21.30 | 27.93 | 313 | 521.53 |
| Edo | 21.28 | 27.27 | 321.63 | 274.54 |
| Ekiti | 20.63 | 26.90 | 297.72 | 240.32 |
| Enugu | 21.30 | 27.93 | 291 | 453.21 |
| FCT | 19.53 | 27.49 | 0 | 538.04 |
| Gombe | 3.99 | 32.44 | 423.98 | 955.65 |
| Imo | 21.82 | 27.14 | 399.99 | 416.91 |
| Jigawa | 21.23 | 37.8 | 879.13 | 358.87 |
| Kaduna | 15.49 | 27.07 | 163.69 | 635.07 |
| Kano | 15.51 | 29.36 | 345.75 | 836.53 |
| Katsina | 3.72 | 31.94 | 436.69 | 861.06 |
| Kebbi | 23.3 | 32.2 | 673.54 | 521.83 |
| Kwara | 20.05 | 27.72 | 323.17 | 258.52 |
| Lagos | 22.13 | 27.33 | 534.33 | 0 |
| Nassarawa | 18.95 | 36.06 | 128.62 | 609.86 |
| Niger | 19.44 | 27.86 | 121.73 | 494.88 |
| Ogun | 22.04 | 27.54 | 504 | 77.89 |
| Ondo | 20.62 | 26.87 | 323.2 | 219.52 |
| Osun | 20.86 | 27.12 | 352.93 | 194.84 |
| Oyo | 20.85 | 27.11 | 441.7 | 113.57 |
| Plateau | 14.42 | 23.73 | 179.42 | 717.46 |
| Rivers | 21.48 | 26.44 | 480.03 | 442.58 |
| Sokoto | 14.39 | 33.64 | 505.94 | 760.03 |
| Taraba | 21.40 | 33 | 919.97 | 424.03 |
| Yobe | | 3712 | 1109.8 | 571.78 |
| Zamfara | 19.63 | 34.67 | 364.93 | 688.59 |

**Table 3**
Pearson's coefficients and p-values for correlations between $\rho$ and three parameters.

| | Pearson Coefficient | $p$-value |
|---|---|---|
| Distance | -0.184884 | 0.391 |
| Humidity | 0.135733 | 0.035 |
| Temperature | 0.01514 | 0.0162 |
| Population density | 0.0212168 | 0.911391 |

caused low adherence to mitigation strategies in the state. Other Cluster A states may therefore consider Lagos as a limiting case of their own outbreaks.

No correlations were found between the rate of outbreak spread and population density or distance from the primary international ports of entry. However, while the correlations with temperature and absolute humidity were both weak, they were statistically significant ($p < 0.05$). Numerous studies have found associations between COVID-19 spread and both temperature and absolute humidity. Contrary to the findings of the current study, previous studies have mostly found both high temperature and humidity having an inhibitory effect on the spread of COVID-19, for example, [18,30,34].

Notably however, studies in tropical and sub-tropical countries are revealing slightly divergent results. In a study carried out in Jakarta, [27], no correlation was found with humidity. A positive correlation was found with mean temperature (Spearman correlation coefficient of 0.392, $p < 0.01$). Another study in sub-tropical areas in Brazil found that correlations with temperature flattened above $25^0$. Furthermore, [2] provided evidence of positive correlation between COVID-19 spread and both high temperatures and intermediate relative humidity in tropical regions. This backdrop renders the findings in the current study more interesting, as they must be interpreted as additional evidence in an emerging thesis of differing effect of temperature and humidity in the tropics versus in cooler regions.

An important factor in this thesis may be the use of air-conditioning. Higher temperatures lead to more use of air conditioning and there is already some tentative link between air conditioner use and COVID-19 spread [17]. This in fact correlates with a common but anecdotal observation in Nigeria that a disproportionately high percentage of individuals who develop severe complications from COVID-19 are affluent or financially comfortable; a link between high temperatures and COVID-19

**Table 4**
Projected outbreak population sizes by December 31, 2020 using SIRD model parameters as of August 10, 2020.

| State | Infected | Recovered | Fatal | Total Confirmed |
|---|---|---|---|---|
| Abia | 7516 | 70 | 0 | 7586 |
| Adamawa | 25 | 21 | 3 | 49 |
| Akwa Ibom | 40 | 35 | 2 | 77 |
| Anambra | 39 | 13 | 1 | 53 |
| Bauchi | 217 | 769 | 3 | 989 |
| Bayelsa | 3 | 131 | 5 | 139 |
| Benue | 80 | 19 | 1 | 100 |
| Cross River | Did not converge | | | |
| Borno | 0 | 358 | 24 | 382 |
| Delta | 1387 | 358 | 31 | 1776 |
| Ebonyi | 120 | 590 | 126 | 836 |
| Edo | 0 | 2063 | 84 | 2147 |
| Ekiti | 3866 | 2751 | 2 | 6619 |
| Enugu | 1690 | 200 | 9 | 1899 |
| FCT | 49,162 | 15,932 | 293 | 65,387 |
| Gombe | 0 | 311 | 4 | 315 |
| Imo | 167 | 440 | 1 | 608 |
| Jigawa | 103 | 117 | 4 | 224 |
| Kaduna | 0 | 1805 | 6 | 1811 |
| Kano | 266 | 2417 | 50 | 2733 |
| Katsina | 887 | 232 | 13 | 1132 |
| Kogi | Did not converge | | | |
| Kebbi | 28 | 17 | 4 | 49 |
| Kwara | 42 | 595 | 1 | 638 |
| Lagos-Adj | 1363 | 22,871 | 82 | 24,316 |
| Nasarawa | 1410 | 59 | 3 | 1472 |
| Niger | 33 | 29 | 15 | 77 |
| Ogun | 1 | 1233 | 11 | 1245 |
| Ondo | 1 | 1493 | 32 | 1526 |
| Osun | 363 | 1183 | 7 | 1553 |
| Oyo | 2164 | 3144 | 11 | 5319 |
| Plateau | 199,810 | 185,803 | 9 | 385,622 |
| Rivers | 66 | 2621 | 40 | 2727 |
| Sokoto | 12 | 88 | 10 | 110 |
| Taraba | 15 | 10 | 0 | 25 |
| Yobe | 32 | 8 | 7 | 47 |
| Zamfara | 13 | 66 | 5 | 84 |
| Total* | 270,921 | 247,852 | 899 | 519,672 |

*excluding Cross River and Kogi States

spread mediated by air conditioner use would be consistent with that observation. Further investigation into the effect of weather on COVID-19 spread in the tropics is recommended.

A number of revealing contrasts can be made between two pairs of neighbouring states. As the lockdown on movement was gradually lifted on both states, $\rho$ increased before reducing for Ogun State (0.008 -> 0.009 -> 0.012 -> 0.0074 -> 0.011) whereas it reduced monotonically for Lagos, a neighbouring state (0.0095 -> 0.0064 -> 0.0034 -> 0.0036 -> 0.003 -> 0.001763561. This suggests that the lockdown was more effective in Ogun State, but not in Lagos. Consequently, the relaxed lockdown led to slight increase in the Ogun outbreak, which eventually slowed down.

Two other neighbouring states, Kano and Kaduna provide potent examples of effect of mitigation strategies on COVID-19 outbreak. The Kaduna State government was more decisive in enforcing lockdown, contact tracing, and in fact introduced travel restrictions that were stricter than any other state in Northern Nigeria. In contrast, the Kano State government implemented all of those measures in a more relaxed manner. As of August 10, 2020, the case fatality rates for Kaduna and Kano states are 0.74 and 3.3 respectively. This, despite the fact that both states are so close as to be indistinguishable on other important metrics. The two contrasts above and the weak or non-existent correlations with other factors suggest that mitigation strategies are the most important driving factors for the COVID-19 outbreak in various parts of Nigeria.

As shown in Table 1, the outbreaks in at least 12 states using growth factor) or up to 22 (using $R_0$) are contracting, but the national outbreak is still expanding. The projection in Table 4 suggests a total of 519,672 confirmed cases by the end of the year, excluding Cross River and Kogi States. This however includes a somewhat surprising projection of 385,622 confirmed cases for Plateau State, largely due to the fact that Plateau happened to have one of the highest $\rho$ (0.03) of any state on August 10, 2020, along with a low $\sigma$. There is nothing to suggest that such a high effective contact rate will persist for any significant length of time, but in keeping with other states, it was assumed constant for the whole of the simulation period. In addition, the case fatality rate for the simulation was low because many states had no deaths in their most recent phase by August 10, leading to values of $\kappa$ at or close to 0.

**Table 5**
Projected outbreak population sizes by December 31, 2020 assuming a worst case scenario.

| State | Infected | Recovered | Fatal | Total Confirmed |
|---|---|---|---|---|
| Abia | 633,129 | 1,395,163 | 139,509 | 2,167,801 |
| Adamawa | 13,277 | 12,170 | 1218 | 26,665 |
| Akwa Ibom | 21,268 | 19,486 | 1947 | 42,701 |
| Anambra | 20,650 | 18,938 | 1894 | 41,482 |
| Bauchi | 107,181 | 102,867 | 10,212 | 220,260 |
| Bayelsa | 1611 | 1596 | 152 | 3359 |
| Benue | 41,525 | 38,454 | 3844 | 83,823 |
| Borno | 0 | 358 | 24 | 382 |
| Cross River | Did not converge | | | |
| Delta | 482,292 | 545,133 | 54,509 | 1,081,934 |
| Ebonyi | 57,645 | 55,964 | 5664 | 119,273 |
| Edo | 0 | 2063 | 84 | 2147 |
| Ekiti | 339,675 | 734,273 | 73,146 | 1,147,094 |
| Enugu | 408,070 | 550,486 | 55,038 | 1,013,594 |
| FCT | 75,108 | 910,306 | 89,706 | 1,075,120 |
| Gombe | 0 | 311 | 4 | 315 |
| Imo | 83,488 | 78,914 | 7848 | 170,250 |
| Jigawa | 52,921 | 49,348 | 4927 | 107,196 |
| Kaduna | 0 | 1805 | 6 | 1811 |
| Kano | 134,942 | 128,616 | 12,669 | 276,227 |
| Katsina | 359,814 | 378,103 | 37,800 | 775,717 |
| Kogi | Did not converge | | | |
| Kebbi | 14,847 | 13,614 | 1364 | 29,825 |
| Kwara | 21,781 | 20,743 | 2016 | 44,540 |
| Lagos-Adj | 537,276 | 586,037 | 56,392 | 1,179,705 |
| Nasarawa | 263,724 | 402,173 | 40,214 | 706,111 |
| Niger | 17,505 | 16,056 | 1618 | 35,179 |
| Ogun | 3208 | 9124 | 117 | 12,449 |
| Ondo | 536 | 1980 | 81 | 2597 |
| Osun | 158,679 | 161,139 | 16,002 | 335,820 |
| Oyo | 610,508 | 763,833 | 76,078 | 1,450,419 |
| Plateau | 116,648 | 2,169,568 | 197,643 | 2,483,859 |
| Rivers | 33,954 | 33,877 | 3165 | 70,996 |
| Sokoto | 6431 | 5945 | 596 | 12,972 |
| Taraba | 7989 | 7310 | 730 | 16,029 |
| Yobe | 16,785 | 15,463 | 1553 | 33,801 |
| Zamfara | 6958 | 6407 | 639 | 14,004 |
| Total* | 4,649,425 | 9,237,623 | 898,409 | 14,785,457 |

*excluding Cross River and Kogi States

There is a possibility of outbreak increase nationally due to the imminent lifting of all lockdown restrictions. While there is generally high compliance with face mask use and other guidelines in some states such as Osun State, compliance appears to be low in others such as Oyo State. Certain scenarios in a post-lockdown phase suggest that states will experience another expansion phase in the next few months. Of particular concern are social events such as wedding receptions in which jubilant and excitable large crowds converge within indoor environments. Such occasions make super-spreading more likely.

In order to predict a worst-case outcome by January 1, 2020, the SIRD models were simulated to that date again, with a higher value for *rho*. Lagos has by far more COVID-19 cases per capita than any other state. In addition, its high population density makes lockdown restrictions more difficult to adhere to or enforce. The state has also had COVID-19 cases for longer, with more phases that any other state. Consequently, for a worst-case national scenario, the simulation was carried out with the highest value or *rho* (0.011) recorded in Lagos at any point after the initial phase. This resulted in the predictions in Table X, with at least 14,785,457 confirmed cases. This is a high number, but the assumption that led to it is not incredible, since that same $\kappa$ has been observed in Lagos during this same outbreak.

## Conclusion

This study developed analysed COVID-19 outbreak data from Nigerian states and fitted SIRD models to them. Despite the noisy nature of the data, it is evident that the outbreak in Nigeria has been milder than initially anticipated, as is has been in some other tropical and subtropical areas. The finding of positive correlations between COVID-19 spread and both increased temperature and humidity is a surprising one echoed in a small number of previous studies, and underscores the need for more studies on the peculiarities of the COVID-19 outbreak in such areas.

## Declaration of Competing Interest

The authors declare that they have no conflict of interest. No funding or financial assistance was received for this study.

## Acknowledgements

The authors thank Mr. Tayo Badrudeen of Oduduwa University Ile-Ife, for suggestions during the formative stage of this study. The Titan Xp GPU in the workstation used for the study was kindly donated by the NVidia corporation.

## References


[1] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: a next-generation hyperparameter optimization framework, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 2623–2631.
[2] A. Auler, F. Cássaro, V. da Silva, L. Pires, Evidence that high temperatures and intermediate relative humidity might favor the spread of covid-19 in tropical climate: acase study for the most affected brazilian cities, Sci. Total Environ. (2020) 139090.
[3] K. Badu, J.P. Thorn, N. Goonoo, N. Dukhi, A.F. Fagbamigbe, et al., Africa's response to the covid-19 pandemic: areview of the nature of the virus, impacts and implications for preparedness [version 1; peer review: awaiting peer review] (2020).
[4] T.M. Balkew, The sir model when s (t) is a multi-exponential function. (2010).
[5] V.C. Corporation, Weather data services | visual crossing, (https://www.visualcrossing.com/weather/weather-data-services? pln=plan_GqkbrxoxU5fY8e/login).
[6] S. Deb, M. Majumdar, A time series method to analyze incidence pattern and estimate reproduction number of Covid-19, arXiv:2003.10655(2020).
[7] N.C. for Diseases Control, Covid-19 situation report situation report 157 monday 3rd August 2020, 2020a, (https://ncdc.gov.ng/themes/common/files/sitreps/260757f86028314c559c5e463092f42a.pdfa). (Accessed on 08/13/2020).
[8] N.C. for Diseases Control, Ncdc coronavirus Covid-19 microsite, 2020b, (https://covid19.ncdc.gov.ng/gis/b). (Accessed on 08/11/2020).
[9] N.C. for Diseases Control, Ncdc coronavirus Covid-19 microsite, 2020c, (https://covid19.ncdc.gov.ng/gis/c). (Accessed on 08/13/2020).
[10] J. Fernández-Villaverde, C.I. Jones, Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities, Technical Report, National Bureau of Economic Research, 2020.
[11] S.J. Fong, G. Li, N. Dey, R.G. Crespo, E. Herrera-Viedma, Finding an accurate early forecasting model from small dataset: a case of 2019-ncov novel coronavirus outbreak, arXiv preprint arXiv:2003.10776 (2020).
[12] S. Gupta, G.S. Raghuwanshi, A. Chanda, Effect of weather on Covid-19 spread in the US: a prediction model for india in 2020, Sci. Total Environ. (2020) 138860.
[13] H. Herrmann, H. Bucksch, Dictionary Geotechnical Engineering/Wörterbuch GeoTechnik: English-German/Englisch-Deutsch, Springer, Berlin Heidelberg, 2014.
[14] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, Proc. R. Soc. Lond. Ser.A 115 (772) (1927) 700–721.
[15] W.O. Kermack, A.G. McKendrick, Contributions to the mathematical theory of epidemics. ii.the problem of endemicity, Proc. R. Soc. Lond. Ser.A 138 (834) (1932) 55–83.
[16] W.O. Kermack, A.G. McKendrick, Contributions to the mathematical theory of epidemics. III. further studies of the problem of endemicity, Proc. R. Soc. Lond. Ser.A 141 (843) (1933) 94–122.
[17] J. Lu, J. Gu, K. Li, C. Xu, W. Su, Z. Lai, D. Zhou, C. Yu, B. Xu, Z. Yang, Covid-19 outbreak associated with air conditioning in restaurant, guangzhou, china, 2020, Emerg. Infect. Dis. 26 (7) (2020) 1628.
[18] Y. Ma, Y. Zhao, J. Liu, X. He, B. Wang, S. Fu, J. Yan, J. Niu, J. Zhou, B. Luo, Effects of temperature variation and humidity on the death of Covid-19 in wuhan, china, Sci. Total Environ. (2020) 138226.
[19] L. McInnes, J. Healy, S. Astels, Hdbscan: hierarchical density based clustering, J. Open Source Softw. 2 (11) (2017) 205.
[20] W.H. Organization, et al., Who Announces Covid-19 Outbreak a Pandemic, World Health Organization [Internet], 2020.
[21] K.J. Rothman, Lessons from John Graunt, Lancet 347 (8993) (1996) 37–39.
[22] A. Rovetta, L. Bhagavathula, A. S.and Castaldo, Modelling the epidemiological trend and behavior of Covid-19 in italy, medRxiv (2020). 10.1101/2020.03.19.20038968.
[23] C. Siettos, L. Russo, Mathematical modeling of infectious disease dynamics, Virulence 4 (4) (2013) 295–306.
[24] A. Singh, Evaluation metrics for regression models-mae vs mse vs rmse vs rmsle, 2019,
[25] N.B. of Statistics, Annual abstract of statistics, 2016,
[26] C.D. Team, Covsirphy, python package for Covid-19 analysis with sir-derived ode models (version 2.6.0), 2020
[27] R. Tosepu, J. Gunawan, D.S. Effendy, H. Lestari, H. Bahar, P. Asfian, et al., Correlation between weather and Covid-19 pandemic in jakarta, indonesia, Sci. Total Environ. (2020) 138436.
[28] J.H. University, Home - Johns Hopkins coronavirus resource center, 2020, (https://coronavirus.jhu.edu/). (Accessed on 08/11/2020).
[29] H. Wang, Y. Zhang, S. Lu, S. Wang, Tracking and forecasting milepost moments of the epidemic in the early-outbreak: framework and applications to the covid-19, medRxiv (2020a).
[30] J. Wang, K. Tang, K. Feng, W. Lv, High temperature and high humidity reduce the transmission of Covid-19, Available at SSRN 3551767(2020b).
[31] E. Whitley, J. Ball, Statistics review 4: sample size calculations, Crit. Care 6 (4) (2002) 335.
[32] C.C. Worldometer, Worldometer.(2020) 1–22, Doi 10(2020.01)) 23–20018549.
[33] K.L.G.M. Wu J. T., Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-NCOV outbreak originating in Wuhan, China: a modelling study, medRxiv 395(0225) (2020) 689–697.
[34] J. Xie, Y. Zhu, Association between ambient temperature and Covid-19 infection in 122 cities from china, Sci. Total Environ. 724 (2020) 138201.
[35] S.W..R.X. Yang Y., Facing the covid19 outbreak: what should we know and what could we do? J. Med. Virol. 92 (6) (2020) 536–537.
[36] S. Zhao, H. Chen, Modeling the epidemic dynamics and control of Covid-19 outbreak in china, Quant. Biol. 8 (2020) 11–19.