

# Modelling COVID-19 cases in Nigeria: Forecasts, uncertainties, projections and the link with weather

Adeyeri O.E. (✉ [cyndyfem@gmail.com](mailto:cyndyfem@gmail.com))

Department of Meteorology and Climate Science, Federal University of Technology, Akure, Nigeria  
<https://orcid.org/0000-0002-9735-0677>

Oyekan K.S.A.

Nigerian Meteorological Agency, Abuja, Nigeria

Ige S.O.

Nigerian Meteorological Agency, Abuja, Nigeria

Akinbobola A.

Department of Meteorology and Climate Science, Federal University of Technology, Akure, Nigeria

Okogbue E.C.

Department of Meteorology and Climate Science, Federal University of Technology, Akure, Nigeria

---

## Research Article

**Keywords:** COVID-19, Projection, Epidemic curve, Reproduction number, Infectious disease, Nigeria

**Posted Date:** July 14th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-41193/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The World Health Organization (WHO) declared COVID-19 a global pandemic on 11 March 2020 due to its global spread. In Nigeria, the first case was documented on 27 February 2020. Since then, it has spread to most parts of the country. This study models, forecasts and projects COVID-19 incidence, cumulative incidence and death cases in Nigeria using six estimation methods i.e. the attack rate, maximum likelihood, exponential growth, Markov chain monte Carlo (MCMC), time-dependent and the sequential Bayesian approaches. A sensitivity analysis with respect to the mean generation time is used to quantify the associated reproduction number uncertainties. The relationship between the COVID-19 incidence and five meteorological variables are further assessed. The result shows that the highest incidences are recorded in days with either religious activities or market days while the weekday trend decreases towards the weekend. It is also established that COVID-19 incidence significantly increases with increasing sea level pressure (0.7 correlation coefficient) and significantly decreases with increasing maximum temperature (-0.3 correlation coefficient). Also, selecting an optimal period for reproduction number estimates reduces the variability between estimates. As an example, in the EG approach, the epidemic curve that optimally fits the exponential growth is between 1- and 53-time units with reproduction number estimate of 1.60 [1.58; 1.62] at 95% confidence interval. However, this optimal reproduction number estimate is different from the default reproduction number estimate. Using the MCMC approach, the correlation coefficients between the observed and forecasted incidence, cumulative death and cumulative confirmed cases are 0.66, 0.92 and 0.90 respectively. The projections till December shows values approaching 1,000,000, 120,000 and 3,000,000 respectively. Therefore, timely intervention and effective preventive measures are immediately needed to mitigate a full-scale epidemic in the country.

## 1. Introduction

Coronavirus disease (COVID-19) is an infectious disease caused by severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2). Wuhan Municipal Health Commission, China first reported COVID-19 as a cluster of cases of pneumonia on 31 December 2019 [1]. The novel Coronavirus was officially confirmed in Thailand on 13 January 2020, subsequently spreading to Japan and South Korea [1]. COVID-19 has now become a global pandemic due to the alarming spread and severity level [2]. There have been 9,473,214 confirmed cases and 484,249 deaths worldwide as of 26 June 2020.

The disease has spread to at least ten (10) South-East Asian countries (686,192 cases and 19,651 deaths), fifty-four (54) European countries (2,619,753 cases and 195,535 deaths), twenty-two (22) Western Pacific countries (210,315 cases and 7,394 deaths), fifty-four (54) countries in the Americas region (4,709,927 cases and 233,628 deaths), twenty-two (22) countries in the Eastern Mediterranean region (987,534 cases and 22,464 deaths) and forty-seven (47) African countries with 258,752 cases and 5,564 deaths [3].

Although the COVID–19 cases in the continent are gradual relative to other parts of the world, the infection has risen exponentially in recent weeks and still spreading. The first case of Africa’s COVID–19 was reported in Egypt on 14 February 2020 [1] while Nigeria registered its first case on 27 February 2020. Since the index case was first reported, the number of new cases from COVID–19 community transmissions has steadily increased. As of 27 June 2020, Nigeria’s cumulative reported cases of COVID–19 had risen to 23,298 with 554 deaths, 8,253 discharged and 14,491 active cases currently receiving treatment [4].

In response to the spread, the Federal and State governments in Nigeria have introduced some measures to curb the spread of the virus and protect the citizens. This includes the suspension of all activities and religious meetings, the permanent closure of public and private schools and universities, the ban on domestic and international flights, building of more isolation centres, closure of borders etc.

Even though data paucity at the onset of this pandemic hindered optimal forecasting and projection for the country, the epidemiological characteristics of COVID–19 dynamism in West Africa and Nigeria, in particular, is different from other parts of the world. Yet, the uncertainties surrounding the modelling of the COVID–19 spread in Nigeria cannot be underemphasized. Nevertheless, few studies have focused on the modelling and forecasting of COVID–19 cases in Nigeria. For example, Ayinde et al [5] compared different statistical models and estimators in forecasting COVID–19 cases without accounting for the uncertainties associated with it. Adegboye et al [6] investigated the early COVID–19 transmissibility using the Bayesian method while incorporating serial distribution uncertainties, however without projections. Although this method assumes a population random mixing, the susceptible depletion is not considered. Consequently, the uncertainty surrounding the COVID–19 transmissibility is still inherent. Therefore, it is important to correct these lapses for a robust forecast and projection which could translate to better intervention and effective preventive and mitigating measures against the outbreak in the country.

This study contributes to the existing knowledge of COVID–19 dynamism in Nigeria by modelling and forecasting the incidence, cumulative incidence and cumulate death using the attack rate, maximum likelihood, exponential growth, Markov chain Monte Carlo, time-dependent and the sequential Bayesian approaches of estimating the COVID–19 reproduction number. Subsequently, these variables are projected for the future. The uncertainty associated with the reproduction numbers as well as the reproduction ratio sensitivity to the estimation time-period is quantified whilst the correlation between COVID–19 incidence and some meteorological variables are examined.

## **2. Data And Methods**

Daily Nigeria confirmed and death cases arising from COVID–19 are extracted from the archives of the Nigerian Centre for Disease Control (NCDC) [4] for 117 days i.e. from February 28 (onset of the outbreak) to June 22 2020. The first death was pronounced on March 23 2020. However, this current analysis focuses on the onset of the locally confirmed case (March 17) till June 22. The spatial representation of the confirmed cases and death across the country is presented in Fig. 1 [7].

To establish the relationship between COVID–19 incidence and meteorological variables in the country’s epidemic centre, daily observed maximum temperature (°C), atmospheric pressure at sea level (hPa), humidity (%), rainfall (mm) and maximum sustained wind speed (Km/h) are retrieved from the archives of the Trans-African HydroMeteorological Observatory.

Subsequently, based on the epidemic curve from the initial and time-dependent reproduction number estimations (as described below), the disease transmissibility and death cases are modelled and forecasted. For all time-dependent reproduction number estimations, a weekly sliding window is applied to the incidence series [8,9]. The model outputs are validated for one month. After a performance check, the best model is used to project the future incidence, cumulative incidence and cumulate death till December 2020.

## ***2.1 Estimating reproduction numbers***

The reproduction number is the average number of healthy people infected per sick person. During an epidemic, reproduction numbers can be estimated at any time. This study assesses the initial and time-dependent reproduction numbers using six methods described below.

### **Attack Rate (AR)**

The AR is the proportion of the population who are eventually infected. This is connected to the basic reproduction number and given as [10],

$$R_0 = - \frac{\log\left(\frac{AR}{S_0}\right)}{AR - (1 - S_0)} \quad (1)$$

where  $S_0$  is the initial proportion of the susceptible population. In this method, there must be a closed population with homogenous mixing and no intervention is provided throughout the epidemic. The method is modelled after the Susceptible - Infectious–Recovered (SIR) model. For this method, the population size is pegged at the country’s population of 200 million people.

From above, the reproduction number throughout an outbreak is estimated based on a single serial interval distribution, thereby not accounting for the serial interval distribution uncertainties.

Therefore, five time-dependent reproduction number estimations are further examined and explained below.

### **Sequential Bayesian (SB)**

This method estimates the initial reproduction number sequentially by approximating the SIR model using the Poisson process such that the mean  $Nte^{y(R-1)}$  approximates the incidence (N) at time t+1,  $N(t+1)$  [11], where  $1/\gamma$  is the average period of the infectious duration. The method initializes daily reproduction number (R) distribution with the previous day posterior distribution [12] and the generation time is explicitly estimated using an exponential distribution. However, this method assumes a population random mixing while the susceptible depletion is not considered.

## Time-dependent reproduction numbers (TD)

Reproduction numbers in TD are calculated by averaging all reproduction numbers over the transmission networks which are compatible with observations [13]. This is given as;

$$Q_{kl} = \frac{N_l w(t_k - t_l)}{\sum_{k \neq p} N_k w(t_k - t_p)} \quad (2)$$

where  $Q_{kl}$  is the probability that case k with onset-time  $t_k$  is infected by case l with onset-time  $t_l$ . N is the incidence and w is the generation time distribution. Therefore, the effective reproduction number ( $R_l$ ) for case l is  $\sum_k Q_{kl}$ . This is averaged over all cases with identical onset date which is represented as;

$$R_t = \frac{1}{N_t} \sum_{\{t_l=t\}} R_l \quad (3)$$

With TD, the importation cases during the outbreak can be taken into consideration.

## Exponential Growth (EG)

The EG rate is the change in the number of new epidemic cases per person per unit of time. This is estimated using Poisson regression (pr) [14]. In this case, the initial reproduction ratio can be linked to the EG rate during the early phases of the outbreak [15]. However, there is no assumption on population mixing. The reproduction number (R) following EG is computed as [12];

$$R = \frac{1}{m(-pr)} \quad (4)$$

where m is the function for generating generation-time distribution moments.

The generation-time distribution is derived from the time lag between all infectee or infector pairs [16].

## Maximum Likelihood (ML)

In ML, the reproduction number (R) is estimated by maximizing its log-likelihood. This is given as [17];

$$\sum_{t=1}^T \log \left[ \frac{e^{-\mu_t} \mu_t^{N_t}}{N_t!} \right] \quad (5)$$

Where  $\mu_t = R \sum_{k=1}^t N_{t-k} w_k$

ML is calculated based on the exponential growth period. The best exponential growth period is selected based on the deviance of the coefficient of determination. ML also assumes that the number of secondary cases as a result of an index case is Poisson-distributed with an expected reproduction number. However, there is no assumption on population mixing.

## Markov Chain Monte Carlo R estimation (MCMC)

In this method, the serial interval distribution is estimated from the exposure data at prescribing intervals using the MCMC while the reproduction number is calculated from the serial interval posterior distribution. The serial interval posterior distribution is based on an uncertainty bootstrap approach with 100000 simulations and 1000 resamples. This accounts for the reproduction number uncertainties. The details of this method are presented in Thompson et al [18].

## 2.2 Projections

The data from daily historical incidence, reproduction number and the serial interval is used to simulate probable epidemic trajectories and to project the future daily incidence. This assumes a branching process where daily incidence follows a Poisson process which is determined by the daily infectiousness. The projected daily incidence  $\varphi$  is calculated as [19];

$$\varphi_t = \sum_{s=1}^{t-1} y_s \delta(t-s) \quad (6)$$

where,

$$\delta(t; s) = \frac{s^t e^{-s}}{t!}$$

$e$  is the Euler's number,  $t$  is the number of times an event occurs in an interval,  $s$  is the average number of events in an interval (rate parameter),  $\delta$  is the probability mass function of the serial interval,  $y_s$  is the incidence time. The number of simulated epidemic curves is set to 10000.

The ratio of RMSE to the standard deviation of the observations (RSR) [20], the relative index of agreement [21], percentage bias and the Pearson's correlation coefficient is used to quantify the error between the observed and forecasted incidence at a 30-day projection window.

Additionally, the overall infectivity due to previously infected people at any time step  $t$  is calculated by weighting the sum of the formerly infected individuals by their infectivity at time  $t$  which is given by a discrete serial interval distribution [8].

On the effect of meteorological variables on the spread of COVID-19, Lagos state, which is the epicentre of COVID-19 cases in Nigeria is taken as an example to assess this. Pearson's correlation is used to understand the relationship between the outbreak of the local cases and the meteorological variables. However, a 7-day lag period is adopted between the meteorological variables and the onset of the local incidence. Subsequently, a calendar plot is used to verify the days of the week with the highest locally confirmed cases.

### **3. Result And Discussion**

In this session, the results of the COVID-19 incidence and death cases, sensitivity analysis as well as the future projection is presented.

#### ***3.1 Temporal trend of COVID-19***

The temporal distribution of the COVID-19 incidence between February 27 and June 24 is presented in Fig. 2. From the onset of the outbreak until mid-April, the number of confirmed cases is between 0 and 100. However, from May, there is a gradual increase in the incidence for most days in May with Sunday, May 30 having the highest case and Tuesday 5<sup>th</sup> and 12<sup>th</sup> having the lowest cases. However, the highest incidences are recorded on June 18<sup>th</sup> with over 730 cases. The lowest confirmed cases in June is on Tuesday, June 2<sup>nd</sup>. The average weekly analysis for the period of the study shows that Friday has the highest range of incidence with values between 100 to 300. Sunday has the lowest range of between 100 to 220 confirmed cases.

However, the trend of incidence decreases from weekdays to weekends. The sharpest increase is from Tuesday to Wednesday. The temporal distribution of the confirmed cases depends on the lifestyle and adherence of people to the quarantine and lockdown measures put in place by the federal government. For example, the highest cases are recorded in days with either religious activities or market days. Moreover, the weekday trend decreasing towards the weekend shows that most people are indoors on weekends due to the total lockdown adopted by many states in the federation. Despite the lockdown,

many institutions offer skeletal services during the week. This could contribute to the high incidences during the weekdays. Additionally, the infectivity increases with time irrespective of the incidence pattern (Fig. 3).

## ***3.2 Reproduction numbers***

Using both the initial and time-dependent approaches for estimating the reproduction number for the entire study period, there are different estimates recorded (Fig. 4) with relatively high variations even though the analysis is carried out over the same dataset. This ranges from 0.9 for MCMC to 11 for SB at a 95% confidence interval.

Particularly for SB, although it is a time-dependent approach, it fails after the first gap due to periods with zero incidence [12]. However, this does not affect the other time-dependent methods inasmuch the periods with zero incidences is shorter than the maximum generation time. While the AR approach needs the least input information, it is not useful during an epidemic because it considers only the initial reproduction number to estimate the epidemic curve for the whole period [12]. This is mostly inaccurate during an epidemic. Additionally, for the entire period of the outbreak, AR assumes that there should not be any form of intervention to curb the outbreak. This is most unlikely. Generally, when  $R > 1$ , the disease spreads in the population.  $R < 1$  means the disease has not started spreading. However, when  $R$  is large, it becomes impossible for the disease to be controlled which eventually leads to a pandemic. As an example, an  $R$  value of 1.2 on the 10<sup>th</sup> day for the MCMC (Fig. 4) means an average of 1.2 persons are infected for each infected person on that day. To contain the spread of the outbreak, the herd immunity threshold which denotes the percentage of the population in need of immunization must be greater than  $1 - R^{-1}$  [22]. Given this, with  $R_{in}$  MCMC equalling 1.06 means 11.8 million people (5.9% of the total population) needs to be immunized to contain the outbreak. Subsequently, the epidemic curve generated by all methods (except SB and AR) does not overlap while maintaining proximity with the incidence (Fig. 5). Furthermore, the generation distribution time may be uncertain during an evolving outbreak. Nevertheless, the period when the outbreak growth becomes exponential has to be specified in the EG and ML approaches. This is usually the period between the first case and when the incidence is maximum. Whilst this outbreak is still emerging in Nigeria, choosing a date could be misleading, thereby increasing the uncertainty. To reduce this uncertainty, a sensitivity analysis is carried out. A deviance R-squared statistic is checked over the entire outbreak period. When the modelled incidence fits the observation optimally, the largest R-squared is recorded. As an example, in the EG approach (Fig. 6; 7), the epidemic curve that optimally fits the exponential growth is between 1- and 53-time units with reproduction number estimate of 1.60 [1.58; 1.62] at 95% confidence interval. However, this optimal reproduction number estimate (Fig. 6) is different from the default reproduction number estimate shown in Table 1.

Table 1 Initial reproduction number as estimated by the different methods

Methods	ML	EG	TD	SB	AR	MCMC
Lower bound	1.182	1.245	1.080	2.265	1.000	0.980
Upper bond	1.229	1.254	1.736	11.195	1.000	1.150
R (95% CI)	1.206	1.249	1.391	6.000	1.000	1.063

It is worthy to note that choosing an optimal period reduces the variability between estimates [12]. The variation in estimates according to changes in the generation time distribution shows an oscillating pattern of estimates with the mean generation time (Fig. 8) [23]. However, the general pattern shows an increase in estimates with the mean generation time [24]. The epidemic curve shows a range between 1.28 and 1.40 for the reported reproduction ratio and the mean generation time increasing from 7.80 to 12.80 days (Fig. 8). Therefore, it is pertinent to carry out a sensitivity analysis with respect to the mean generation time to quantify the associated uncertainties. As in the case of complexities resulting from overlapping generations, a joint-estimation of the reproduction number and generation time distribution could be adopted [25]. Alternatively, periods with exponential growth could be chosen [12].

### ***3.3 Comparing methods and projections***

Based on the goodness of fit, the observed and forecasted incidence are compared in time. Fig. 9 shows variant degrees of accurate forecasting power of each model. However, the AR and ML approach constantly underestimates the observed incidence while SB overestimates it. Satisfactory results are observed in TD and MCMC approaches.

Using the MCMC approach, errors in the modelled incidence, cumulative death and cumulative confirmed cases are quantified based on a 30-day projection window using four error metrics (Table 2). The result is satisfactory with Pearson's correlation coefficients of 0.66, 0.92 and 0.90 values for incidence, cumulative incidence cases and cumulative death respectively. Based on this result, the projections are made until December (Fig. 10). There is a continuous increase in the incidence, cumulative incidence and cumulative death with values approaching 1000000, 120000 and 3000000 respectively.

Table 2 Error metrics between the observed and predicted incidence at 30 days projection window

Cases	Confirmed	Cumulative Confirmed	Cumulative Death
PBIAS %	-6	-3	17.4
RSR	0.77	0.11	1.6
rd	0.76	0.79	0.79
r	0.66	0.92	0.9

RSR is the ratio of RMSE to the standard deviation of the observations, rd is the relative index of agreement, PBIAS is the percentage bias and r is the Pearson's correlation coefficient

### ***3.4 Relationship between meteorological parameters and COVID-19***

The result of the Pearson's correlation between the outbreak and the meteorological variables in Lagos considering a 7-day lag period (Fig. 11) shows that the correlation between the atmospheric pressure at sea level and the COVID-19 incidence is highest with a value of 0.71 at  $P \leq 0.001$ . This means COVID-19 incidence increases significantly with increasing surface pressure. However, it decreases with increasing maximum temperature ( $-0.34$  at  $P \leq 0.01$ ). The lowest correlation is with the maximum sustained wind ( $-0.03$  at no significant level). In general, the maximum temperature and atmospheric pressure at sea level are the main meteorological variables affecting the COVID-19 incidence in the country's epicentre. The interactions between these meteorological variables are not easily dissociated. For example, there is a negative correlation between rainfall and temperature during the wet season [26], while atmospheric pressure at sea level and temperature is negatively correlated. Historically, the highest maximum temperature occurs between March and May [26], however, the highest atmospheric pressure at sea level occurs in July [27]. Given this relationship, more incidence and subsequent transmission of COVID-19 is expected in the coming months. Considering the present realities and future projections, there is a need for crucial preparations and assessments of the consequences and complications that could arise from the COVID-19 pandemic as well as dedicating relevant resources to understanding and managing its probable effects on the country's economy, livelihood and wellbeing of the citizens.

## **4. Conclusion**

There is a continuous increase in the incidence, cumulative confirmed cases and cumulative death. This increase, especially in the incidence, is disturbing. Therefore, the health officials, policymakers, the general public and the disease control agencies are reminded of the various risks (e.g. economic recession, social discomfort and insecurity) associated with further increment and spread of COVID-19 cases. Timely intervention, holistic planning, effective preventive measures are urgently needed to mitigate a full-blown epidemic in the country.

## **Declarations**

*Funding:* This research received no external funding.

*Conflict of Interest:* The authors declare no conflict of interest

## **References**

- [1] WHO. World Health Organization Novel Coronavirus Disease (2019-nCoV) Situation. [June 26, 2020]; Available from: <https://www.who.int/20200121-sitrep-1-2019-ncov.pdf>.
- [2] Bedford J, Enria D, Giesecke J, Heymann DL, Ihekweazu C, Kobinger G et al. COVID-19: towards controlling of a pandemic. *Lancet* 2020;395(10229):1015–8. [https://doi.org/10.1016/S0140-6736\(20\)30673-5](https://doi.org/10.1016/S0140-6736(20)30673-5).
- [3] WHO. Coronavirus disease 2019 (COVID-19) situation report. [June 26, 2020]; Available from: <https://www.who.int/20200626-covid-19-sitrep-158.pdf>.
- [4] NCDC. COVID-19. [June 27, 2020]; Available from: <http://covid19.ncdc.gov.ng>.
- [5] Ayinde K, Lukman AF, Rauf RI, Alabi OO, Okon CE, Ayinde OE. Modeling Nigerian Covid-19 cases: A comparative analysis of models and estimators. *Chaos Solitons Fractals* 2020;138:109911. <https://doi.org/10.1016/j.chaos.2020.109911>.
- [6] Adegboye OA, Adekunle AI, Gayawan E. Early Transmission Dynamics of Novel Coronavirus (COVID-19) in Nigeria. *Int J Environ Res Public Health* 2020;17(9). <https://doi.org/10.3390/ijerph17093054>.
- [7] Oramfe. COVID-19 Infected and Death. Nigeria; 2020.
- [8] Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am J Epidemiol* 2013;178(9):1505–12. <https://doi.org/10.1093/aje/kwt133>.
- [9] Kamvar ZN, Cai J, Pulliam JRC, Schumacher J, Jombart T. Epidemic curves made easy using the R package incidence. *F1000Res* 2019;8:139. <https://doi.org/10.12688/f1000research.18002.1>.
- [10] Dietz K. The estimation of the basic reproduction number for infectious diseases. *Stat Methods Med Res* 1993;2(1):23–41. <https://doi.org/10.1177/096228029300200103>.
- [11] Bettencourt LMA, Ribeiro RM. Real time bayesian estimation of the epidemic potential of emerging infectious diseases. *PLoS ONE* 2008;3(5):e2185. <https://doi.org/10.1371/journal.pone.0002185>.
- [12] Obadia T, Haneef R, Boëlle P-Y. The R0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC Med Inform Decis Mak* 2012;12:147. <https://doi.org/10.1186/1472-6947-12-147>.
- [13] Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol* 2004;160(6):509–16. <https://doi.org/10.1093/aje/kwh255>.
- [14] Boëlle PY, Bernillon P, Desenclos JC. A preliminary estimation of the reproduction ratio for new influenza A(H1N1) from the outbreak in Mexico, March-April 2009. *Euro Surveill* 2009;14(19). <https://doi.org/10.2807/ese.14.19.19205-en>.

- [15] Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc Biol Sci* 2007;274(1609):599–604. <https://doi.org/10.1098/rspb.2006.3754>.
- [16] Svensson A. A note on generation times in epidemic models. *Math Biosci* 2007;208(1):300–11. <https://doi.org/10.1016/j.mbs.2006.10.010>.
- [17] White LF, Pagano M. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. *Stat Med* 2008;27(16):2999–3016. <https://doi.org/10.1002/sim.3136>.
- [18] Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, Demarsh PA et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics* 2019;29:100356. <https://doi.org/10.1016/j.epidem.2019.100356>.
- [19] Selvadurai APS. Poisson's equation. In: Selvadurai APS, editor. *Partial Differential Equations in Mechanics 1: Fundamentals, Laplace's Equation, Diffusion Equation, Wave Equation*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2000, p. 503–647.
- [20] Moriasi D, Arnold J, van Liew M, Bingner R, Harmel R, Veith T. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* 2007;50:885.
- [21] Adeyeri OE, Laux P, Lawin AE, Oyekan KSA. Multiple bias-correction of dynamically downscaled CMIP5 climate models temperature projection: a case study of the transboundary Komadugu-Yobe river basin, Lake Chad region, West Africa. *SN Appl. Sci.* 2020;2(7). <https://doi.org/10.1007/s42452-020-3009-4>.
- [22] Fine P, Eames K, Heymann DL. "Herd immunity": a rough guide. *Clin Infect Dis* 2011;52(7):911–6. <https://doi.org/10.1093/cid/cir007>.
- [23] Cauchemez S, Bhattarai A, Marchbanks TL, Fagan RP, Ostroff S, Ferguson NM et al. Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proceedings of the National Academy of Sciences* 2011;108(7):2825–30. <https://doi.org/10.1073/pnas.1008895108>.
- [24] Nishiura H, Castillo-Chavez C, Safan M, Chowell G. Transmission potential of the new influenza A(H1N1) virus and its age-specificity in Japan. *Euro Surveill* 2009;14(22). <https://doi.org/10.2807/ese.14.22.19227-en>.
- [25] Cauchemez S, Boelle P-Y, Donnelly CA, Ferguson NM, Thomas G, Leung GM et al. Real-time estimates in early detection of SARS. *Emerging Infect Dis* 2006;12(1):110–3. <https://doi.org/10.3201/eid1201.050593>.
- [26] Adeyeri OE, Laux P, Lawin AE, Ige SO, Kunstmann H. Analysis of hydrometeorological variables over the transboundary Komadugu-Yobe basin, West Africa. *Journal of Water and Climate Change* 2019.

## Figures

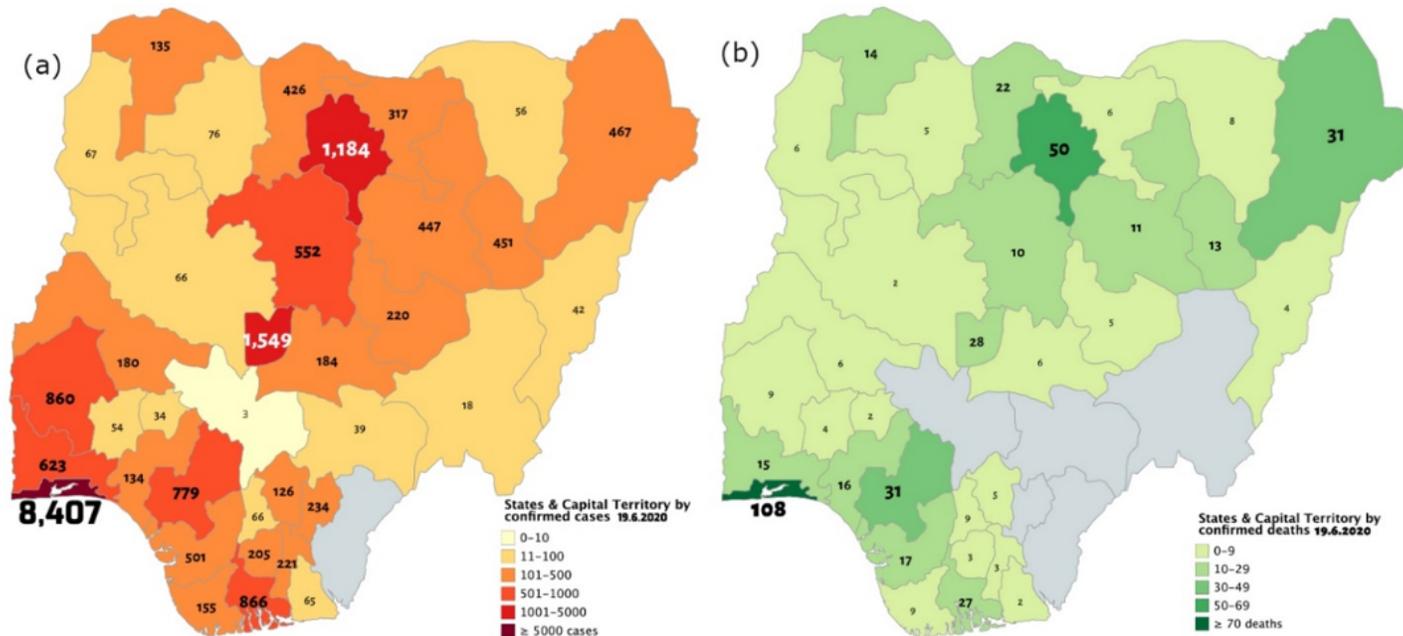


Figure 1

Spatial distribution of Nigerian COVID-19 confirmed (a) and death (b) cases on the 19th of June 2020. Adopted after Oramfe [7].

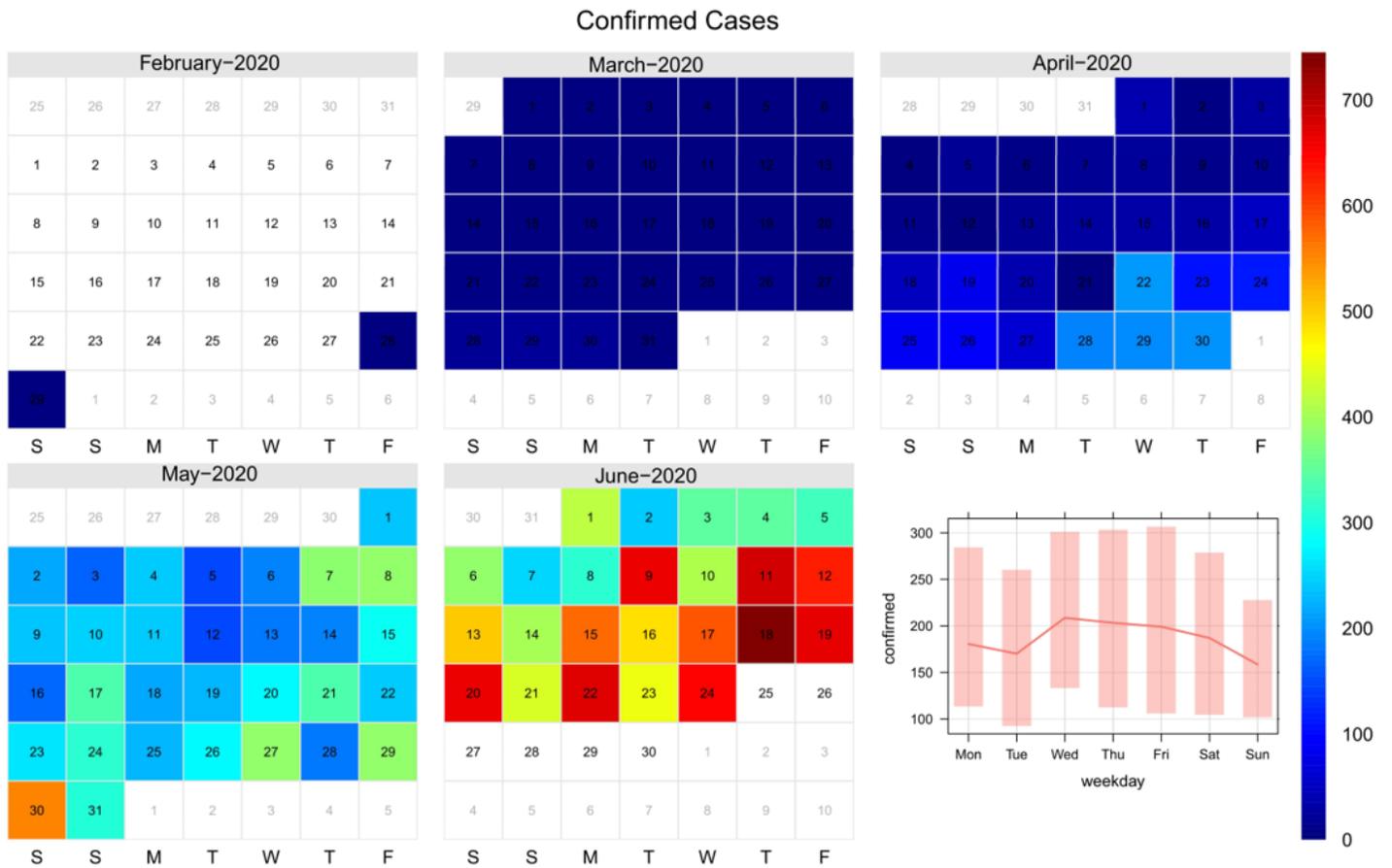


Figure 2

Temporal distribution of incidence

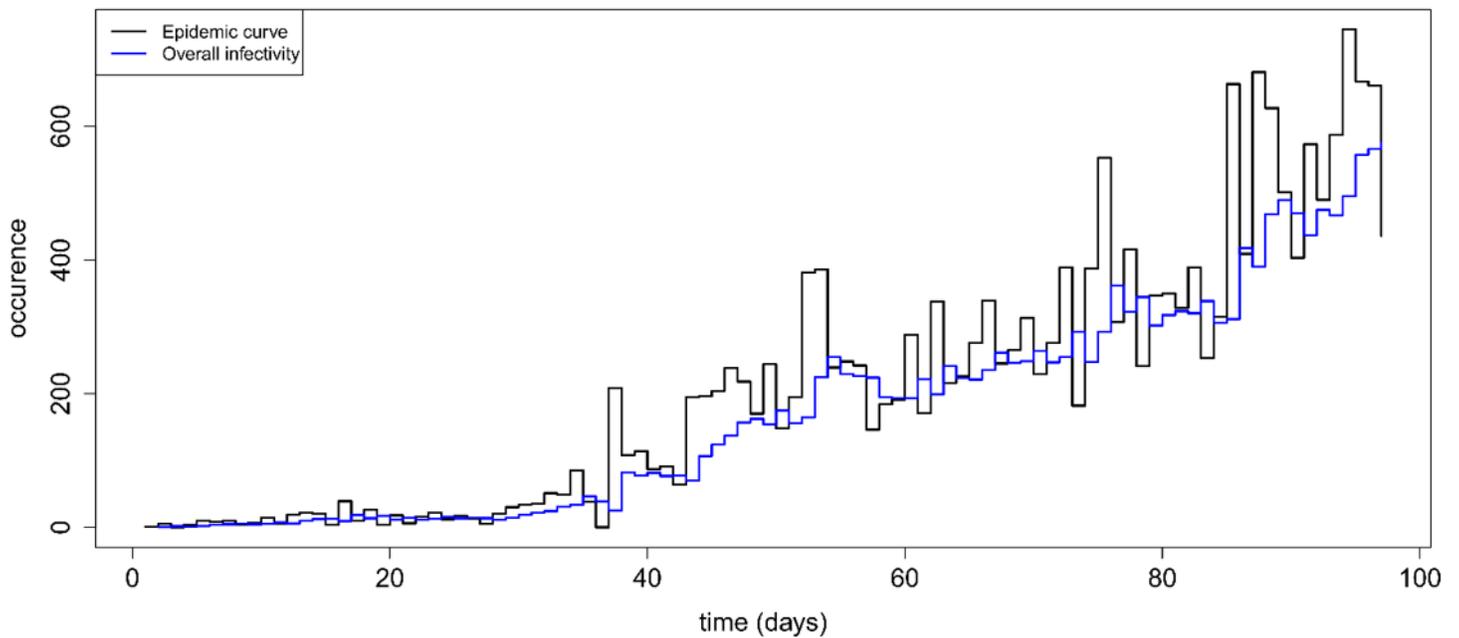


Figure 3

## Infectivity and Epidemic curve

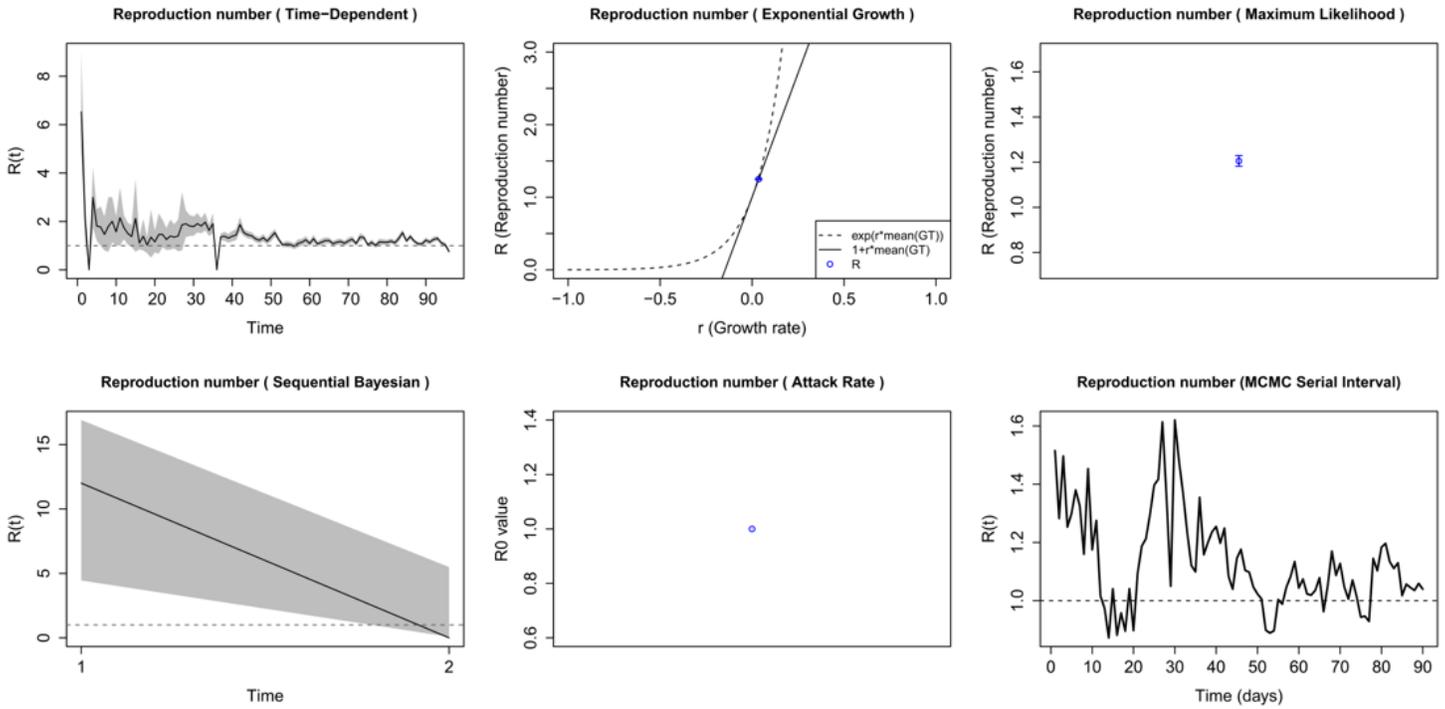


Figure 4

## Reproduction number estimates by different methods

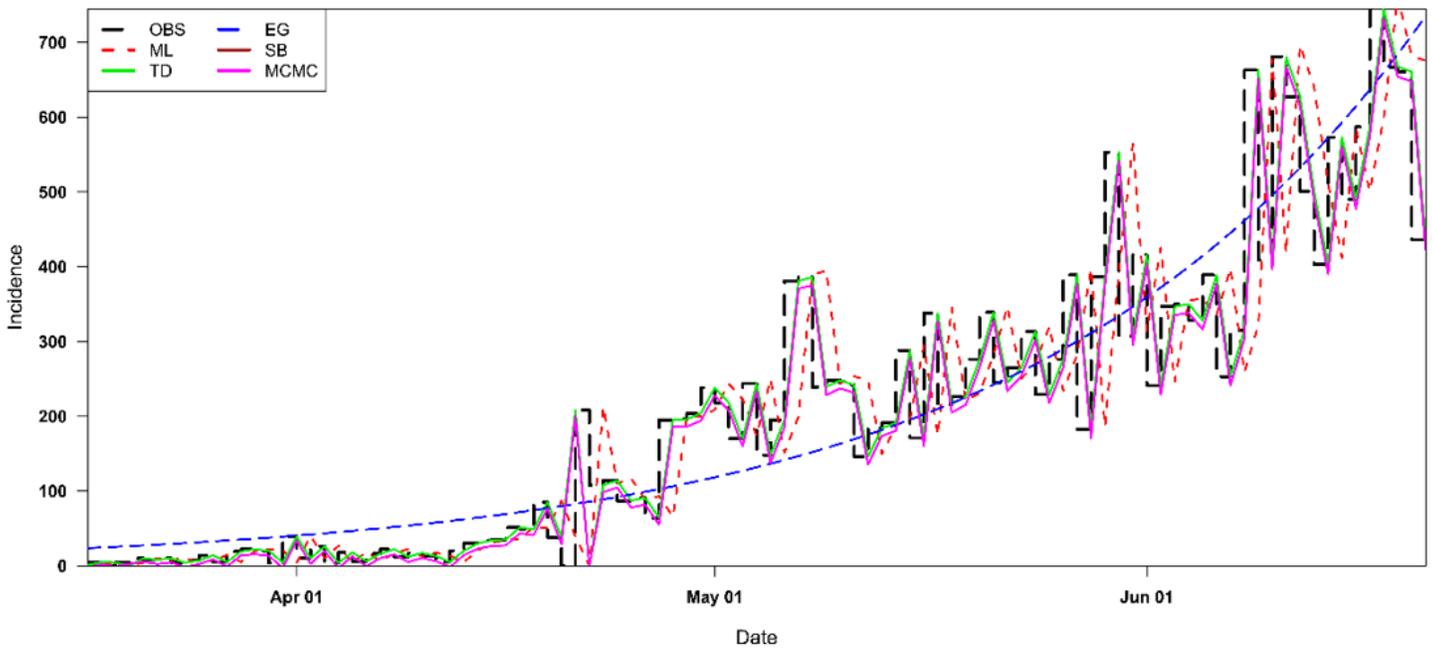
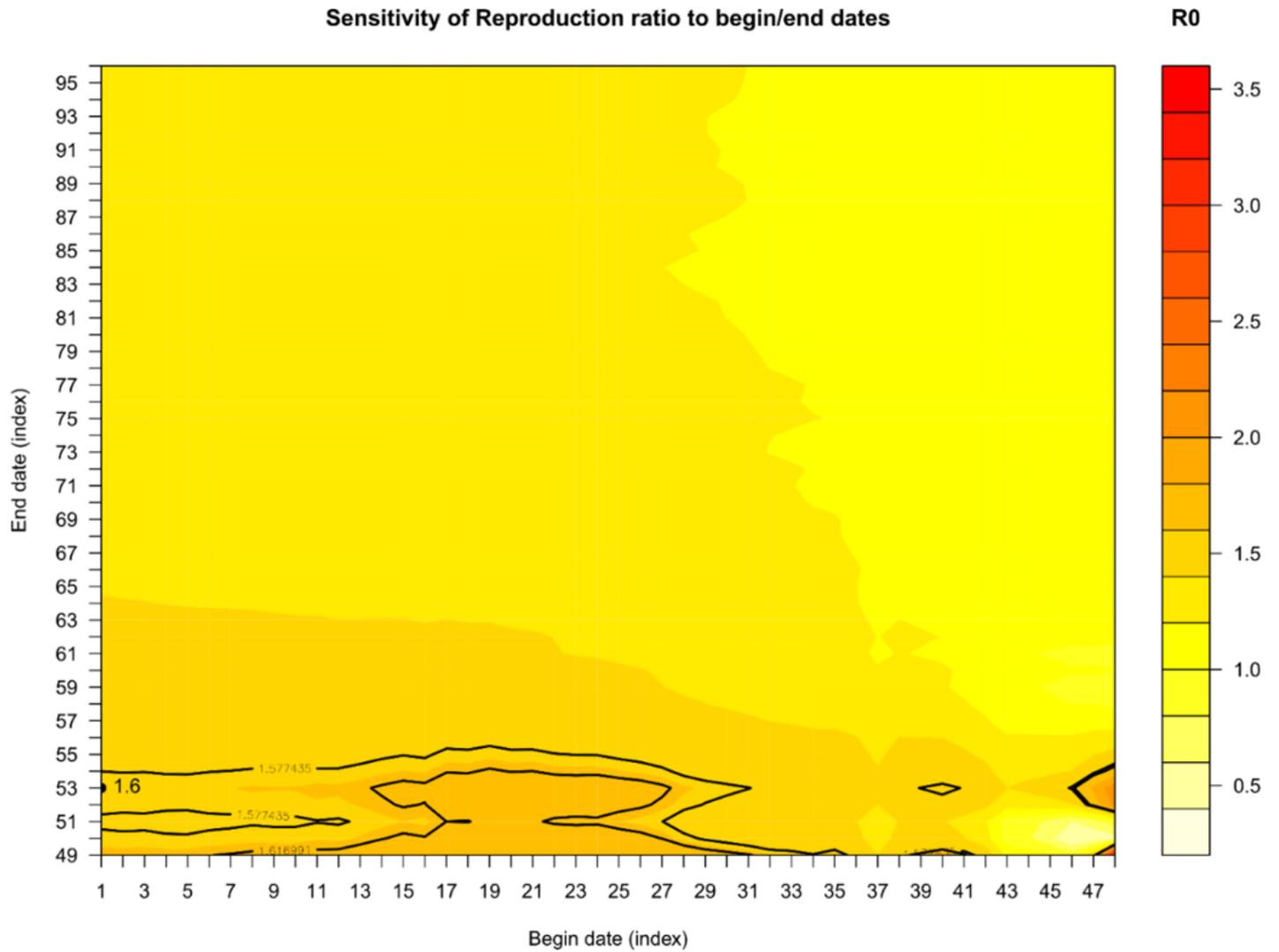


Figure 5

Estimates of the reproduction ratio and goodness of fit for the observed incidence (step function) and model-predicted incidence for each method



**Figure 6**

Reproduction ratio sensitivity to the estimation time-period. The best fit is shown as a dot and the solid black lines show the 95%CI.

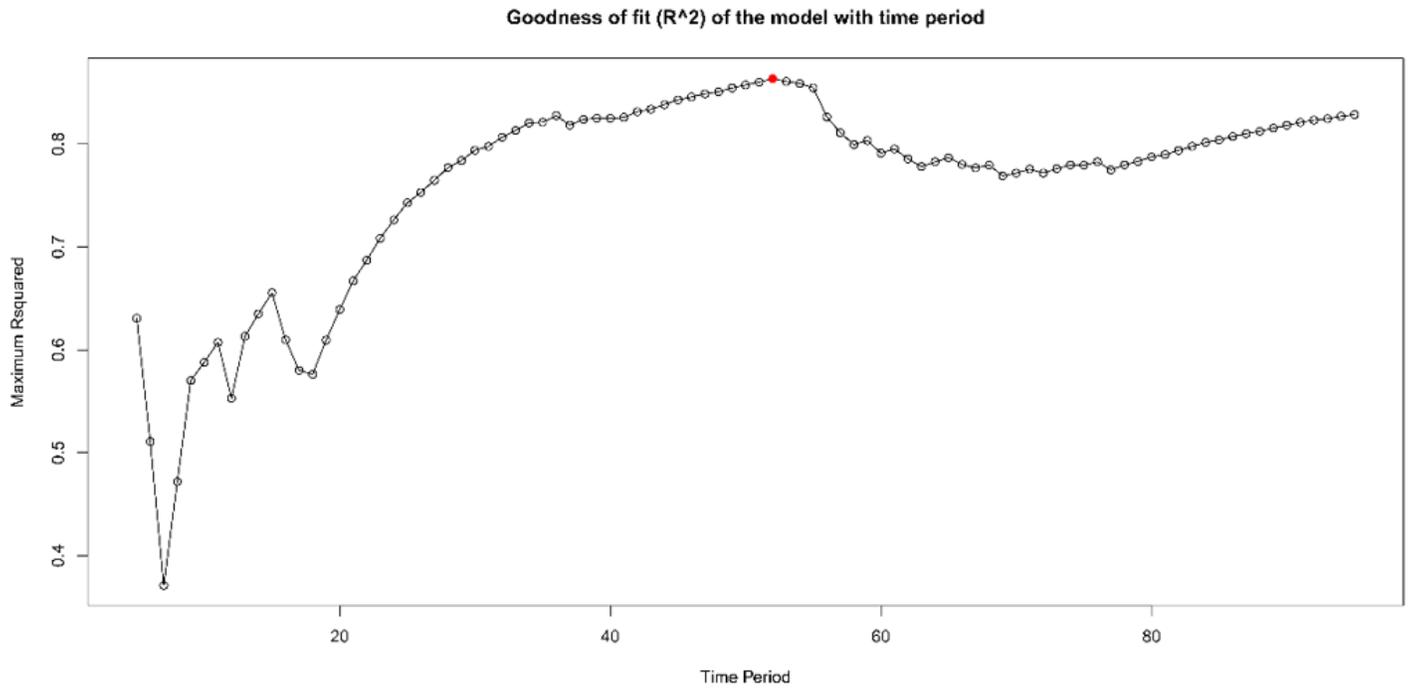


Figure 7

Maximum deviance R-squared statistic for time-periods of increasing duration. The red dot corresponds with the best value.

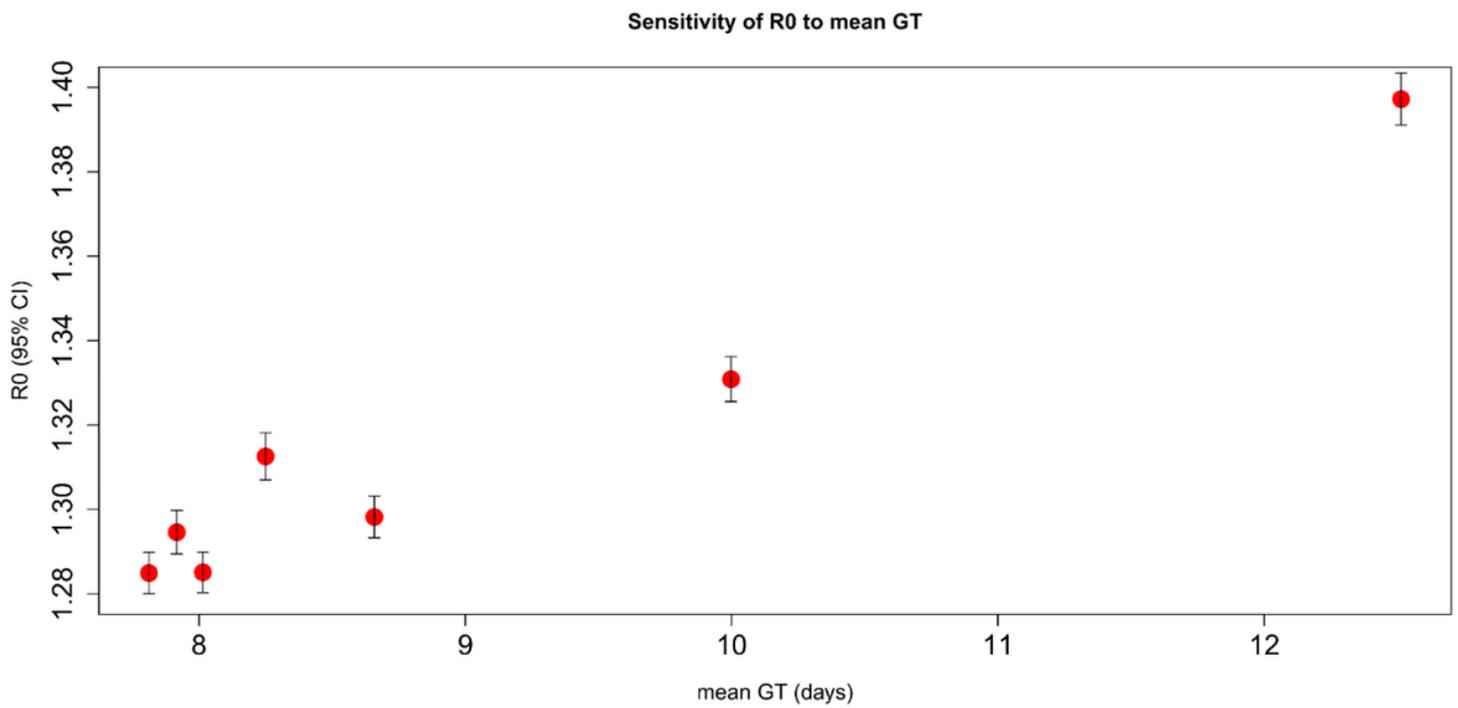
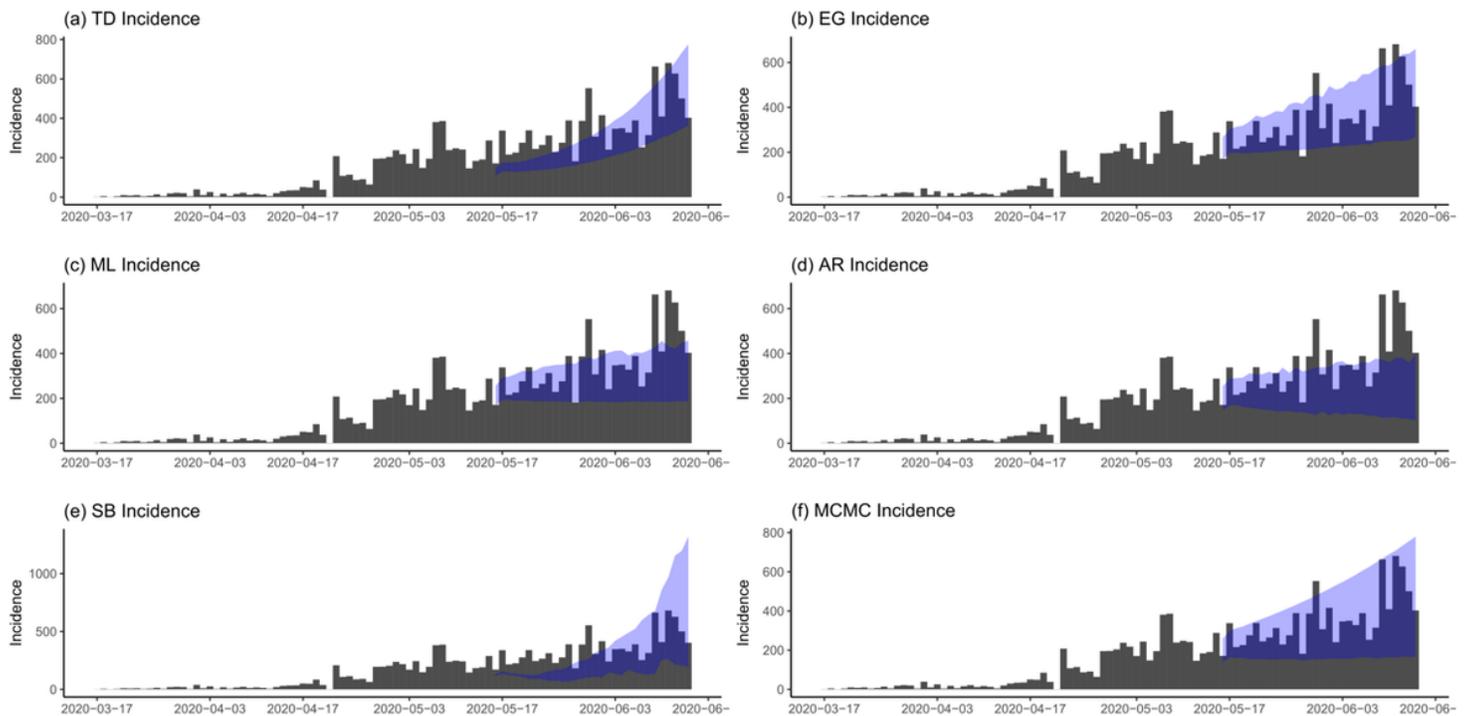


Figure 8

Reproduction number sensitivity to the generation time distribution. The vertical bars are the 95% confidence interval



**Figure 9**

Observed incidence (bar plot) and model-predicted incidence for each method Using the MCMC approach, errors in the modelled incidence, cumulative death and cumulative confirmed cases are quantified based on a 30-day projection window using four error metrics (Table 2). The result is satisfactory with Pearson's correlation coefficients of 0.66, 0.92 and 0.90 values for incidence, cumulative incidence cases and cumulative death respectively. Based on this result, the projections are made until December (Fig. 10). There is a continuous increase in the incidence, cumulative incidence and cumulative death with values approaching 1000000, 120000 and 3000000 respectively.

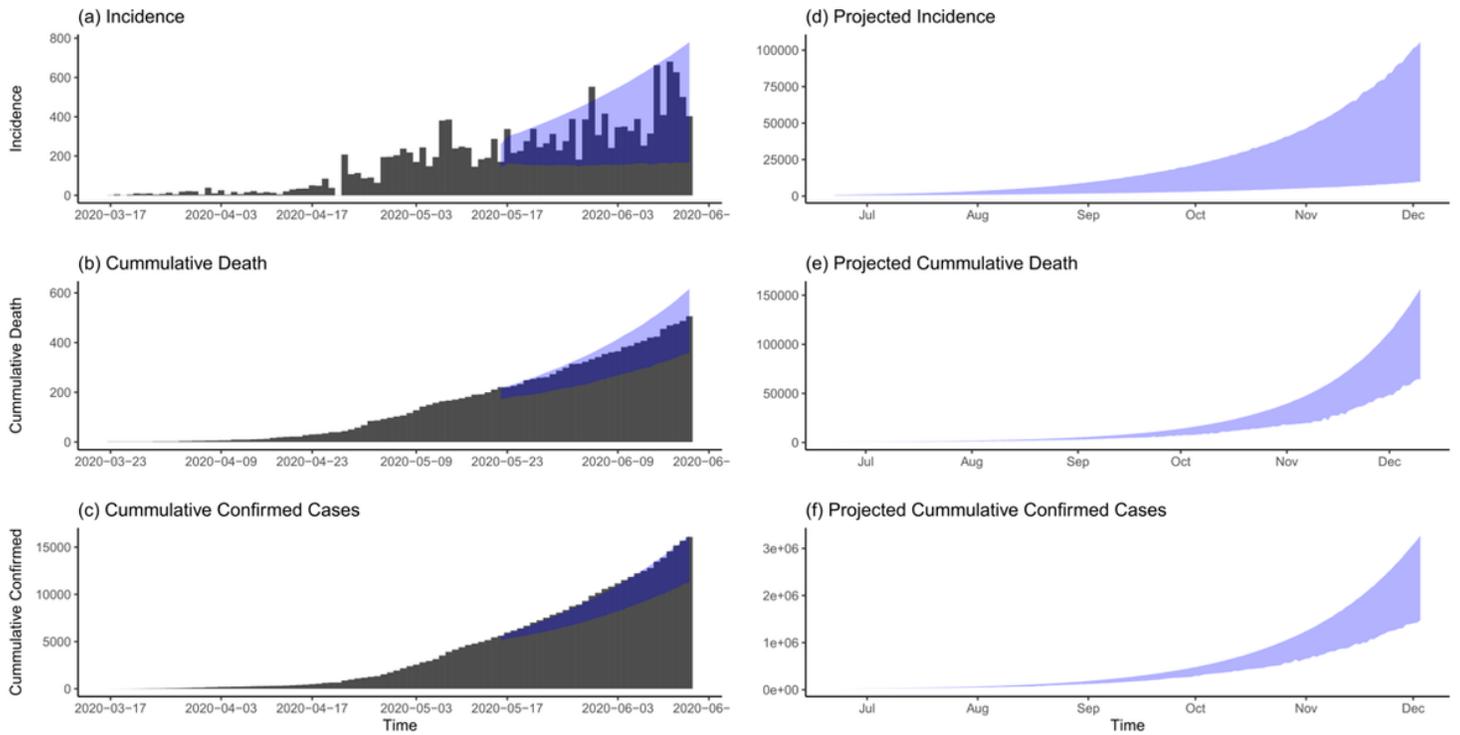


Figure 10

MCMC incidence, cumulative death, cumulative confirmed cases and their projections

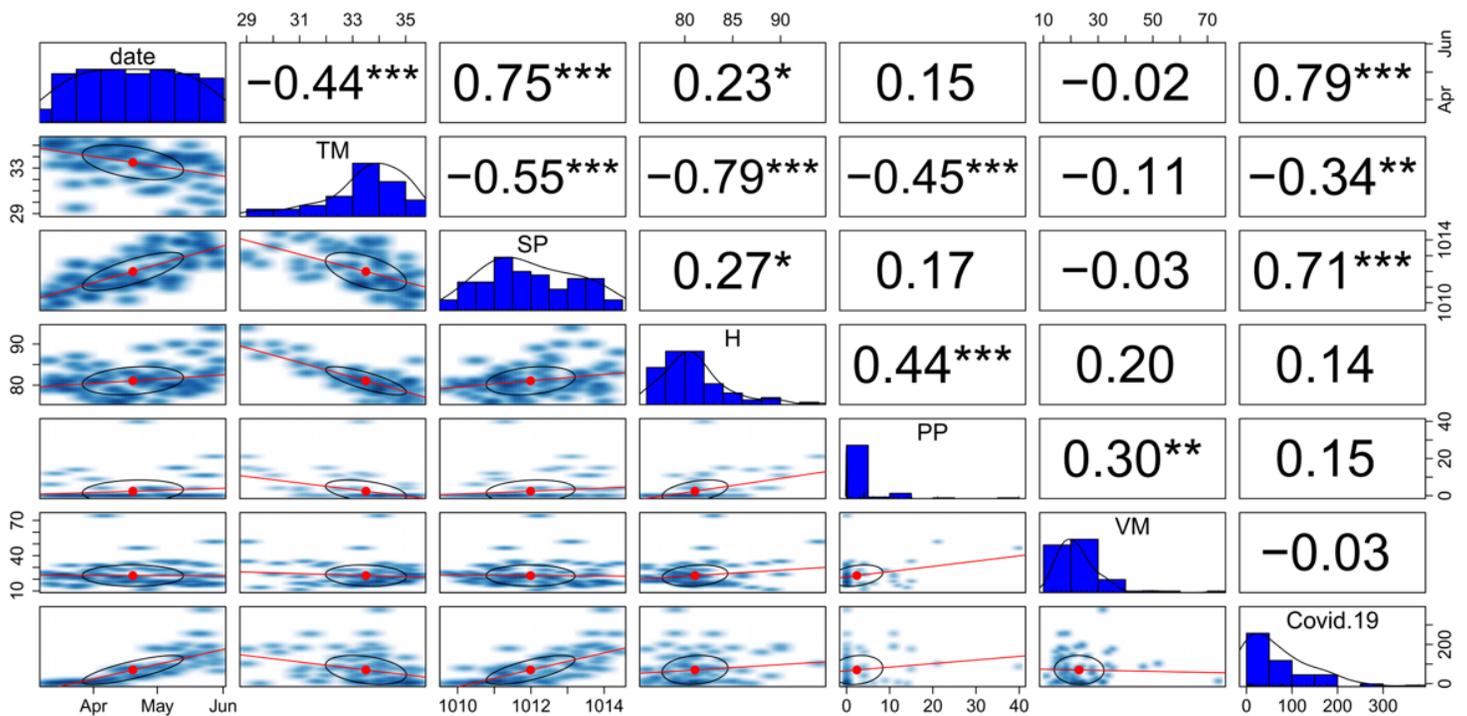


Figure 11

Correlation plot between daily meteorological parameters and COVID-19 incidence between March and May with a 7-day lag in COVID confirmed cases. \* represents  $P \leq 0.05$ , \*\* represents  $P \leq 0.01$ , \*\*\* represents  $P \leq 0.001$  otherwise, the correlation is not significant. TM is daily observed maximum temperature ( $^{\circ}\text{C}$ ), SP is atmospheric pressure at sea level (hPa), H is humidity (%), PP is rainfall (mm) and VM is maximum sustained wind speed (Km/h)