



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Original articles

Maximal reproduction number estimation and identification of transmission rate from the first inflection point of new infectious cases waves: COVID-19 outbreak example

J. Waku^a, K. Oshinubi^{b,*}, J. Demongeot^b^a UMMISCO UMI IRD 209 & LIRIMA, University of Yaoundé I, P.O Box 337 Yaoundé, Cameroon^b Laboratory AGEIS EA 7407, Team Tools for e-Gnosis Medical & Labcom CNRS/UGA/OrangeLabs Telecom4Health, Faculty of Medicine, University Grenoble Alpes (UGA), 38700 La Tronche, France

Received 30 November 2021; received in revised form 15 February 2022; accepted 16 February 2022

Available online 25 February 2022

Abstract

The dynamics of COVID-19 pandemic varies across countries and it is important for researchers to study different kind of phenomena observed at different stages of the waves during the epidemic period. Our interest in this paper is not to model what happened during the endemic state but during the epidemic state. We proposed a continuous formulation of a unique maximum reproduction number estimate with an assumption that the epidemic curve is in form of the Gaussian curve and then compare the model with the discrete form and the observed basic reproduction number during the contagiousness period considered. Furthermore, we estimated the transmission rate from identification of the first inflection point of a wave of the curve of daily new infectious cases using the Bernoulli S–I (Susceptible–Infected) equation. We applied this new method to the real data from Cameroon COVID-19 outbreak both at national and regional levels. High correlation was observed between the socio-economic parameters and epidemiology parameters at regional level in Cameroon. Also, the method was applied to the second wave COVID-19 outbreak for the world data which is a period the phenomena we are considering were observed. Lastly, it was observed that the models presented results correspond with the epidemic dynamics in Cameroon and World data. We recommend that it is important to study what happened during the growth inflection point as some countries data did not climax.

© 2022 International Association for Mathematics and Computers in Simulation (IMACS). Published by Elsevier B.V. All rights reserved.

Keywords: Bernoulli; COVID-19; Epidemic; Inflection point; Reproduction number

1. Introduction

In recent time it was discovered that the COVID-19 virus is still mutating, hence it is possible as we approach different seasons due to the movement of people from one place to another and increase in travels that the transmission rate from one person to the other may increase exponentially, but the aggressive vaccination campaign by policy makers, government introduction of lockdowns and some level of restriction in order to mitigate the virus

* Corresponding author.

E-mail addresses: jules.waku@gmail.com (J. Waku), Kayode.Oshinubi@univ-grenoble-alpes.fr (K. Oshinubi), Jacques.Demongeot@univ-grenoble-alpes.fr (J. Demongeot).

<https://doi.org/10.1016/j.matcom.2022.02.023>

0378-4754/© 2022 International Association for Mathematics and Computers in Simulation (IMACS). Published by Elsevier B.V. All rights reserved.

spread may help to reduce the death rate. The United Kingdom health security agency reported on November 25, 2021 that a new COVID-19 variant that was first identified in Botswana is the most challenging variant yet and has doubled the number of mutations of Delta variant. The study of the spread of COVID-19 pandemic has been interest of researchers, most especially the socio-economic gap the situation has exposed in many countries.

Since the seminal article of D. Bernoulli in 1760 [1], and the improvements made by D. Ross [24], McKermack [17], and then by many other more recent authors [3–5,14–16,21], one of the central problems remains the estimate of the most important parameters of the equation governing the growth of observed infectious cases, namely the reproduction number each day during the contagious period of an infectious patient and the transmission rate.

In the first case, we propose a new method of estimating the maximum number of daily reproduction numbers and, in the second case, we describe a new method of calculating the rate of transmission by identifying the inflection points of the curve of new infectious cases observed.

In Section 2, we describe the method of estimating the maximum daily reproduction number, then in Section 3, the method of identifying inflection points of the curve of new infectious cases in the Bernoulli model. In Section 4, we apply these methods to data from the COVID-19 epidemic in Cameroon. Finally, in Section 5, we propose a generalization to other countries, in Section 6 a discussion, and eventually, we describe in Section 7 some perspectives of this work.

2. Maximal daily reproduction number estimation

Let us suppose that we consider an infectious disease outbreak in a country after a first random endemic period and fix the start of the epidemic wave at day t_0 when the number of the new cases X_{t_0} has reached for the first time a threshold, for example equal to 100 new cases of infectious. We suppose that, if the duration of the infectivity (or contagiousness) period of an individual is equal to r days, we have recorded the new cases at times $t_0 - 1, \dots, t_0 - r$. Then, for the days $j \geq t_0$, we have:

$$X_j = \sum_{k=1,r} R_k X_{j-k}, \tag{1}$$

where the convolution equation (1) can be deconvolved in order to get the R_k 's, i.e., the numbers of people infected by an infectious individual during the k th day of his period of contagiousness [8]. R_k is called the daily reproduction number at day k of the contagiousness period of an individual.

The deconvolution equation is given in matrix form by:

$$R = M^{-1}X, \tag{2}$$

where $X = (X_j, \dots, X_{j-r+1})$ is the vector of new infectious cases at days $j, \dots, j - r + 1$ and $R = (R_1, \dots, R_r)$ the vector of the daily reproduction numbers along the contagiousness period of length r . Both are r -dimensional vectors and M is defined as the following $r-r$ matrix:

$$M = \begin{pmatrix} X_{j-1} & X_{j-2} & \dots & X_{j-r} \\ X_{j-k-1} & X_{j-k-2} & \dots & X_{j-k-r} \\ X_{j-r} & X_{j-r-1} & \dots & X_{j-2r+1} \end{pmatrix} \tag{3}$$

Remark. The number R_j depends on the individual i and on the day j , and we would consider in fact the number R_{ikj} of people infected at day j by a given infectious individual i during the k th day of his period of contagiousness of length r_i . In the following, for the sake of simplicity, we neglect the dependency on i, j and r , which is reasonable during exponential growth [8].

2.1. Continuous observation time

Let us suppose that the observation time is continuous and that the new cases of infectious $X(t)$ depend on the distribution of the quantities $O(s) = R(s)/E(R)$, where the quantity $E(R) = \int_{\mathfrak{R}} R(t)dt$:

$$X(t) = \int_{\mathfrak{R}} R(s)X(t-s)ds = E(R) \int_{\mathfrak{R}} X(t-s)O(s)ds \tag{4}$$

(a) we will suppose in a first case that the curve $O(s)$ is a Gaussian curve $\mathcal{N}(m, \sigma)$ then we have:

$$R(s) = R_{\max} \exp(-(s - m)^2/2\sigma^2), \tag{5}$$

where R_{\max} is the maximum of the daily reproduction numbers, and we will show that the continuous equivalent of Eq. (1) can be solved if $X(t)$ is supposed to be exponential: $X(t) = 100 e^{\beta(t-t_0)} = c e^{\beta t}$, for $t \geq t_0$, where $c = 100 e^{-\beta t_0}$. Eq. (5) can be written as:

$$c e^{\beta t} = \int_{\mathfrak{R}} R_{\max} \exp(-(s - m)^2/2\sigma^2) c e^{\beta(t-s)} ds,$$

from which we deduce:

$$\begin{aligned} 1 &= R_{\max} \int_{\mathfrak{R}} \exp(-(s - m)^2/2\sigma^2 - \beta s) ds \\ 1 &= R_{\max} \int_{\mathfrak{R}} \exp(-(s - m + \beta\sigma^2)^2/2\sigma^2 + \beta^2\sigma^2/2 - \beta m) ds \\ 1 &= R_{\max} \sigma (2\pi)^{1/2} \exp(\beta^2\sigma^2/2 - \beta m) \int_{\mathfrak{R}} \exp(-(s - m + \beta\sigma^2)^2/2\sigma^2) / \sigma (2\pi)^{1/2} ds \end{aligned}$$

Hence, we obtain:

$$1 = R_{\max} \sigma (2\pi)^{1/2} \exp(\beta^2\sigma^2/2 - \beta m) \tag{6}$$

and if we know σ ($4\sigma \approx r$, the duration of the contagiousness period), $m \approx 2\sigma$ and β (the transmission rate), we get R_{\max} , the value at day m (the middle of the contagiousness period) of the maximal daily reproduction number. (b) if the curve of the daily reproduction numbers $R(s)$ is the sum of two Gaussian-shaped curves, the first corresponding to the activation of the innate immunity and the second to the activation of the adaptive immunity, we have:

$$R(s) = R_{\max1} \exp(-(s - m_1)^2/2\sigma_1^2) + R_{\max2} \exp(-(s - m_2)^2/2\sigma_2^2)$$

In this case, Eq. (5) can be written as:

$$c e^{\beta t} = \left[\int_{\mathfrak{R}} R_{\max1} \exp(-(s - m_1)^2/2\sigma_1^2) + R_{\max2} \exp(-(s - m_2)^2/2\sigma_2^2) \right] c e^{\beta(t-s)} ds \tag{7}$$

Then, we obtain:

$$1 = R_{\max1} \exp(\beta^2\sigma_1^2/2 - \beta m_1) + R_{\max2} \exp(\beta^2\sigma_2^2/2 - \beta m_2) \tag{8}$$

formula from which we can calculate the value of a parameter in function of the value of others. If $4\sigma_1 + 4\sigma_2$ can be estimated by the duration of the contagiousness period and if we suppose that σ_1 is close to σ_2 , with $m_1 = 2\sigma_1$ and $m_2 = 4\sigma_1 + 2\sigma_2$, then the only remaining unknown parameters are $R_{\max1}$ and $R_{\max2}$. If these two maximal daily reproduction numbers are supposed to be close, the Eq. (8) allows the final calculation of a unique R_{\max} , which could be compared to the one obtained by the discrete method of deconvolution presented in [8].

2.2. Discrete observation time

(1) If the observation time is discrete, we replace $O(s)=R(s)/E(R)$, with $E(R) = \int_{\mathfrak{R}} R(t)dt$ by $O_k = R_k/E(R)$, with $E(R) = \sum_{k=1,r} R_k = R_0$, where R_0 is the basic reproduction number and using Eq. (1), the Eq. (4) by:

$$X_j = \sum_{k=1,r} R_k X_{j-k} X(t) = E(R) \sum_{k=1,r} X_{j-k} O_k,$$

where the distribution $\{O_k\}_{k=1,r}$ can have several uni- or bi-modal shapes estimated by deconvoluting Eq. (1) [8].

3. Estimation of transmission rate from identification of the first inflection point of a wave of the curve of daily new infectious cases

Following the same approach as in [9], we want now to derive the inflection point during the outbreak of COVID-19 pandemic using Bernoulli equation and to analyze this Susceptible–Infected (SI) compartment model

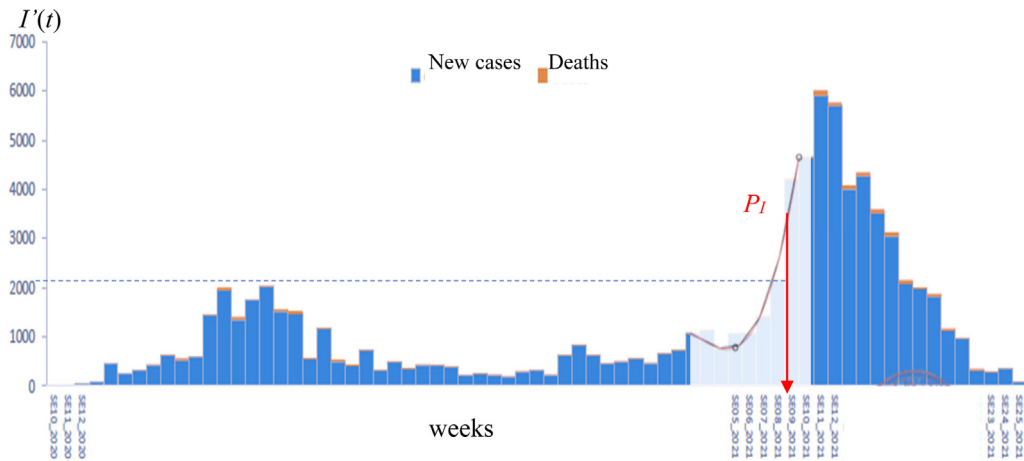


Fig. 1. COVID-19 weekly new cases in Cameroon showing the inflection point P_I (red arrow) during the first half part of the second wave between week 5 of 2021 and week 12 of 2021.

Source: From [2].

for calculating the time of the inflection point of the curve of new infectious cases. First, let us recall that the SI Bernoulli differential equation model of an outbreak epidemic is given as:

$$\frac{dS}{dt} = -\frac{\beta}{S_o} S(t) I(t) + dS(t), \quad \frac{dI}{dt} = \frac{\beta}{S_o} S(t) I(t) - \nu I(t), \tag{9}$$

where d is the susceptible demographic balance, ν the specific mortality rate due to the disease, β/S_o the disease transmission rate, $S(t)$ the number of susceptible individuals, $I(t)$ the number of infectious individuals at time $t \geq 0$, and the initial conditions of the model are: $S(0) = S_o > 0$ and $I(0) = 1$.

Graphical representation of the inflection point P_I of the curve of $I'(t)$ is given Fig. 1, for the first half part of the second wave of the COVID-19 weekly new reported cases between week 5 2021 and week 12 2021 (data from [2]). The red curve represents the cubic spline approximation of the raw data with the following statistics: p -value = 0.03, Multiple R-squared MR = 0.90, Relative Mean Square Error RMSE = 0.247.

During the same period, it is also possible to smooth the death data (from [25]) with a cubic spline with the following statistics: p -value = 0.12, Multiple R-squared MR = 0.99, Relative Mean Square Error RMSE = 4.380.

3.1. Mathematical formulation

We want to formulate the problem in such a way that we can identify all parameters from the set $\emptyset = \{(S_o, I_o, K, \beta)\} \subset (0, \infty)^4$ if we know the reported daily infectious cases $I'(t)$ for all times $t \geq 0$. It should be noted that it is not sufficient alone to know the value of $I'(t)$ at any time $t \geq 0$, because the knowledge about the value of S_o , the number of susceptible people in the population before the epidemic outbreak is also necessary to allow to obtain a precise information about the values of I_o, K, β (see [6,8,9]).

Let us consider the solution of the Bernoulli equation with $d = \nu = 0$ (fertility and mortality rates negligible):

$$I(t) = \frac{S_o e^{\beta(t-a)}}{1 + e^{\beta(t-a)}}, \quad S(t) = \frac{S_o}{1 + e^{\beta(t-a)}},$$

where S_o, β and a represent respectively the susceptible size at endemic start, transmission rate and endemic start time.

For any time t , we have:

$$I(t) + S(t) = S_o \left[\frac{e^{\beta(t-a)}}{1 + e^{\beta(t-a)}} + \frac{1}{1 + e^{\beta(t-a)}} \right] = S_o e^{\beta(t-a)} = \frac{I(t)}{S_o - I(t)}$$

If we consider that the epidemic wave starts at time 0, where $I(0) = 1$, we have:

$$1 = \frac{S_o e^{-\beta a}}{1 + e^{-\beta a}}$$

Then, $S(0) = S_o - 1$ and a is given by the following equation:

$$a = \log(S_o - 1)/\beta \tag{10}$$

3.2. Derivation of the equation of the inflection point P_I

From Bernoulli equation, where $v = 0$; we have:

$$I'(t) = \frac{dI}{dt} = \frac{\beta}{S_o} SI = \frac{\beta}{S_o} \times \frac{S_o}{1 + e^{\beta(t-a)}} \times \frac{S_o e^{\beta(t-a)}}{1 + e^{\beta(t-a)}} = \frac{\beta S_o e^{\beta(t-a)}}{(1 + e^{\beta(t-a)})^2}$$

The sufficient existence condition for a point of inflection of order 2 for $I'(t)$ in the case that $I(t)$ is 3-times continuously differentiable in a certain neighborhood of a point t_i , with $I'''(t_i) = 0$, and $I''(t_i) \neq 0$. Then, $I'(t)$ has an inflection point (P_I) of order 2 at t_i , and by differentiating $I'(t)$ twice, we get:

$$I''(t) = \frac{d^2 I}{dt^2} = \beta^2 S_o \frac{e^{\beta(t-a)} - 2e^{2\beta(t-a)}}{(1 + e^{\beta(t-a)})^3}$$

$$I'''(t) = \frac{d^3 I}{dt^3} = \beta^3 S_o \frac{e^{\beta(t-a)}(1 - 6e^{\beta(t-a)} + 2e^{2\beta(t-a)})}{(1 + e^{\beta(t-a)})^4}$$

Then, the equation giving β and a from its lowest root $x_i = 3 - \sqrt{7} = e^{\beta(t_i-a)} = \frac{I(t_i)}{S_o - I(t_i)}$ (depending on $I(t_i)$) is:

$$1 - 6x_i + 2x_i^2 = 1 - 6e^{\beta(t_i-a)} + 2e^{2\beta(t_i-a)} = 0 \tag{11}$$

If we assume now $I(0) = 100$, then from the same calculations, we can identify the values of β and a from the Eqs. (10) and (11).

3.3. Identification method

If we calculate the value of the time t_i at which the first P_I of a wave is observed and if we know the value of S_0 , then we can estimate the maximal reproduction number as $I''(t_i)$ and identify β and a from the Eqs. (10) and (11). The steps of the identification method are as follows:

1. From an interpolation of the curve of the daily reported new cases (see for example red curve in Fig. 1) of a particular country epidemic wave, we calculate the lowest t_i of the wave at which $I'''(t_i) = 0$,
2. From the value of t_i , we estimate the maximal reproduction number as $I''(t_i)$,
3. From the knowledge of $S(0)$, the number of susceptible people in the population at the start of a particular country epidemic wave, we can get S_o ,
4. From Eqs. (10) and (11), we can identify β and a .

4. Application to the COVID-19 outbreak in Cameroon

We will use in the following the Cameroon’s data concerning the COVID-19 outbreak and from the Cameroon Situation Reports (weekly regional new cases and deaths [2]) and from the public database Worldometers (weekly national new cases and deaths [25]). We will analyze in the following the Cameroon data at national and regional levels (see Fig. 3), by comparing the 10 regions from a set of socio-economic and epidemiologic variables (Table 3) whose data are coming from the public database of Globaldatalab [11].

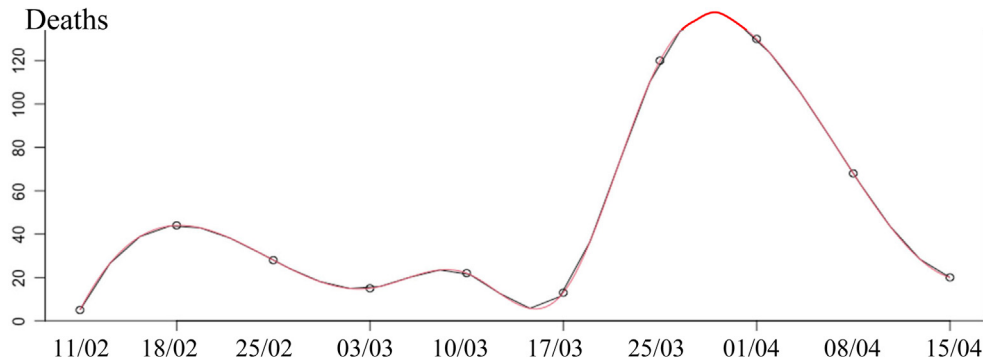


Fig. 2. Spline graph for COVID-19 deaths in Cameroon during the second wave between week 5 of 2021 and week 14 of 2021.



Fig. 3. The 10 regions of Cameroon.
Source: From [25].

4.1. Estimation of the maximal basic reproduction number R_{0max} and of the maximal daily reproduction number R_{max}

4.1.1. Study in Cameroon

(1) We first calculate the time t_i of the first inflection point P_I of the second wave of the new infectious cases $X(t)$ for the Cameroon. For that, we have smoothed the raw data of the weekly reported new infectious cases in Fig. 2 by using a cubic spline and the maximal reproduction number is estimated by the maximal value of its slope divided

by $X(t)$ at $t = t_i = 4$ (weeks), equal to $R_{0,max} = 2.04$, close to the maximal value of the effective reproduction number given for this period by the public database Renkulab [23].

(2) From Eq. (6), we can estimate the maximal daily reproduction number R_{max} :

$$1 = R_{max}\sigma(2\pi)^{1/2} \exp(\beta^2\sigma^2/2 - \beta m),$$

Then, from [19] we estimate $\sigma = 1$, $m = 2$, $\beta = \beta/S_o = 0.4$. Then, we have:

$$R_{max} = 1/(2\pi)^{1/2} \exp(0.8 - 0.08) \approx 0.9,$$

which corresponds in case of a Gaussian shape for the daily reproduction numbers R_j 's along the contagiousness period of length $r \approx 4\sigma$ to a basic reproduction number $R_0 = \sum_{k=1,r} R_k \approx 2.1$.

4.1.2. Calculation of the distribution of daily reproduction numbers at the beginning of the pandemic in Cameroon

The numbers of new cases during the first wave between March 15 and March 28 2020 are the following:

March15 1, 1, 5, 3, 0, 14, 13, 0, 16, 10, 9, 0, 16, 0 **March28**

Then, the matrix M defined in Section 2 is equal to:

$$M = \begin{bmatrix} 16 & 0 & 9 & 10 & 16 & 0 & 13 \\ 0 & 9 & 10 & 16 & 0 & 13 & 14 \\ 9 & 10 & 16 & 0 & 13 & 14 & 0 \\ 10 & 16 & 0 & 13 & 14 & 0 & 3 \\ 16 & 0 & 13 & 14 & 0 & 3 & 5 \\ 0 & 13 & 14 & 0 & 3 & 5 & 1 \\ 13 & 14 & 0 & 3 & 5 & 1 & 1 \end{bmatrix}$$

and we have:

$$M^{-1} = \begin{bmatrix} 0.00582788 & -0.00882952 & 0.009303 & -0.03084686 & 0.0216235 & -0.02815066 & 0.06042457 \\ -0.00882952 & 0.0070809 & -0.019324 & 0.00609393 & -0.01580307 & 0.03737718 & 0.03900754 \\ 0.009303 & -0.019324 & -0.00781664 & -0.00403656 & 0.02691476 & 0.06319195 & -0.03605914 \\ -0.03084686 & 0.00609393 & -0.00403656 & 0.06311531 & 0.04014842 & -0.01719015 & -0.05720378 \\ 0.0216235 & -0.01580307 & 0.02691476 & 0.04014842 & -0.02492553 & -0.01022751 & -0.04545266 \\ -0.02815066 & 0.03737718 & 0.06319195 & -0.01719015 & -0.01022751 & -0.07132353 & 0.01670959 \\ 0.06042457 & 0.03900754 & -0.03605914 & -0.05720378 & 0.04545266 & 0.01670959 & 0.05053995 \end{bmatrix}$$

Because the vector of new infectious cases $X = \begin{bmatrix} 0 \\ 16 \\ 0 \\ 9 \\ 10 \\ 16 \\ 0 \end{bmatrix}$, hence $R = M^{-1}X = \begin{bmatrix} -0.65306968 \\ 0.60814399 \\ 0.93470577 \\ 0.7919826 \\ -0.30440867 \\ -0.80012803 \\ -0.07788649 \end{bmatrix}$ and we can represent

the evolution of the daily reproduction numbers R_j 's in Fig. 4, which shows a unimodal distribution on the first days ($j = 1, \dots, 4$) of the infection period, with a sum of the R_j 's equal to 0.4, close to the effective reproduction number R_e in [23].

4.1.3. Study in other countries

We calculated the R_{max} using same procedure as above for six other countries using the transmission rate of the first wave in 2020: Australia = 4.6214×10^{-8} , Italy = 1.2292×10^{-8} , Spain = 3.3182×10^{-7} , United State = 3.1828×10^{-7} , United Kingdom = 1.1526×10^{-7} and Canada = 2.0731×10^{-7}) given in [22], with $\sigma = 1$ and $\sigma = 2$, corresponding to a duration of the contagiousness period of respectively 4 and 8 days. We also calculated R_{max2} by using Eq. (8) presented in Table 2. We used the values of R_{max} in Table 1 as R_{max1} while still maintaining the values for the transmission rate and σ .

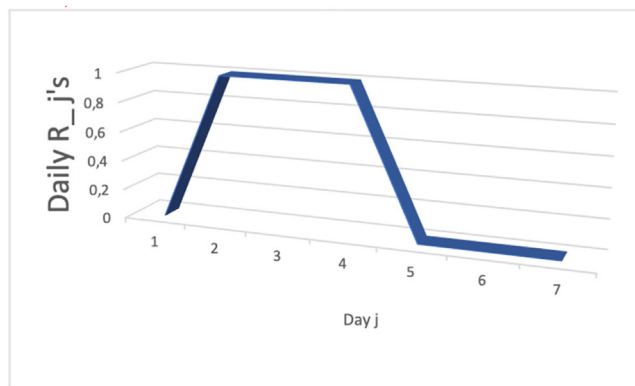


Fig. 4. The distribution of the daily reproduction numbers R_j 's on the infection period.

Table 1
Result for R_{\max} for different countries.

Countries	$\sigma = 1$	$\sigma = 2$
United Kingdom	0.39978	0.20031
Australia	0.40018	0.20071
Italy	0.39927	0.19980
Spain	0.40136	0.20190
Canada	0.40018	0.20098
United States	0.40126	0.20180

Table 2
Result for $R_{\max 2}$ while using R_{\max} in Table 1 as $R_{\max 1}$ for different countries.

Countries	$\sigma = 2$	$\sigma = 1$
United Kingdom	0.79842	0.59769
Australia	0.79742	0.59608
Italy	0.79970	0.59974
Spain	0.79440	0.59129
Canada	0.79673	0.59498
United States	0.79466	0.59169

4.2. Identification of the parameters β and a

From Eqs. (10) and (11), we get:

$$a = \text{Log}(S_o - 1)/\beta \text{ and } 1 - 6e^{\beta(t_i-a)} + 2e^{2\beta(t_i-a)} = 0,$$

with $S_o = 27,000,000$ and $t_i = 4$. Then $a = 7/\beta$ and $1 - 6e^{4\beta-7} + 2e^{8\beta-14} = 0$, from which we get: $\beta = 2.0$ and $a = 3.5$, what is consistent for a with the previous estimation of t_i , but overvalued for β , perhaps due to an overestimation of S_o . Indeed, if S_o corresponds to the population of the 3 regions (Littoral, Center and Northwest) giving the essential of the infectious cases (for example, $325/371 = 88\%$ of all the new cases during the week of 23/06/2021). More, if we consider the cumulative number of infectious during the whole week 8 of 2021, there was 2000 cases. If we consider that the susceptible population in contact with these infectious is $2000 \times 50 = 100,000$, then from Eq. (10) $a = 11.5/\beta$, and from Eq. (11), $\beta = 2.4$ and $a = 4.7$, which shows the sensitivity of calculation of β to initial conditions concerning the susceptible population.

Table 3
Values of epidemiologic parameters for different regions of Cameroon.

Cameroon regions	Discrete R	Transmission rate β	R_{\max} $\sigma = 1$	R_{\max} $\sigma = 2$
Adamaoua	1.30175	1.50	2.60143	0.89396
Center	1.04135	1.75	2.85711	0.47851
Est	2.18015	1.50	2.60143	0.89396
Extreme_Nord	1.74880	1.50	2.60143	0.89396
Littoral	0.91404	1.75	2.85711	0.47851
Nord	1.31392	1.50	2.60143	0.89396
Nord_Ouest	3.64468	1.50	2.60143	0.89396
Ouest	3.38611	1.50	2.60143	0.89396
Sud	1.19674	1.50	2.60143	0.89396
Sud_Ouest	0.88279	1.50	2.60143	0.89396

4.3. Study at the regional level

4.3.1. Estimation of the epidemiologic parameters

By using the methods proposed in Section 3, we obtain the value of the main epidemiologic parameters for each region of Cameroon (Table 3) and we observe a big difference between the urban regions, essentially the regions of Yaoundé and Douala (i.e. respectively the Center and Littoral regions) whose urban density is high and the other mainly rural regions.

4.3.2. Correlations between epidemiologic, socio-economic and demographic parameters

We have observed the epidemiologic variables (numbers of weekly reported new infectious cases, cured and deceased patients) during the 19 first weeks of 2021 (cf. in Fig. 1, data from [25]). The parameters list is given in Table 4 and their values are extracted from Table 3 and from [11,23]. The correlations between the parameters of Table 1 considered as variables on the 10 regions of Cameroon can be analyzed thanks to Figs. 5 and 6, which show that the epidemiologic parameters (from parameter Dose 1 to parameter Recovered) are significantly anti-correlated (pink color) with many socio-economic and demographic parameters, namely the three first ones (Subnational Human Development Index, International Wealth Index and Gross National Income per capita) and to those linked to the house holding, showing the influence of a good socio-economic state on the resistance against the virus.

By exploiting the data corresponding to the parameters of Table 4, we can analyze the correlations existing between three families of parameters observed in the different regions of Cameroon during the second wave of the COVID-19 outbreak. In Fig. 5, we see that the socio-economic factors are highly positively correlated with each other, but negatively correlated with demographic parameters and not correlated with epidemiologic parameters, except those corresponding to vaccination and death which are anticorrelated with the signs of wealth. This can be explained by a higher socio-cultural level, which pushes individuals to monitor their health more and to get vaccinated. In Fig. 6, we can observe a greater number of demographic parameters that are anti-correlated with socio-economic factors. These are in particular infant mortality, mean age, fertility and poverty indicators, which have positive or zero correlation coefficients between them and very negative with wealth indicators, such as GNI or possession of material goods or land (television and house), because in general a rise in the standard of living causes fertility and infant mortality decrease.

We observe in Fig. 7 a high correlation between Sud-Ouest and Littoral with the score 0.918 and the least correlations are between Center and Adamaoua with the score 0.242, between Est and Center with the score 0.282 and between Nord and Center with the score 0.253.

We observed in Figs. 8 to 12 that for cured cases, there is high correlations between Ouest and Center with score 0.9308 and also between Littoral and Nord-Ouest with score 0.936967. The least correlation is between Extreme-Nord and Sud with score 0.04.

We observed in Fig. 13 three anticorrelations: (i) between Nord and Adamaoua with score -0.00967 , (ii) between Nord and Sud-Ouest with score -0.18 , and (iii) between Sud-Ouest and Ouest with score -0.09988 . The highest correlation is Nord-Ouest and Center with score 0.845.

Table 4

List of the parameters considered for the different regions of Cameroon.

Socio-economic parameters	
HDI	Subnational Human Development Index
IWI	International Wealth Index
GNI	Gross National Income per capita
PoorHighest	Percentage poor households (IWI value under 70)
PoorLowest	Percentage poor households (IWI value under 35)
MeanEducation	Mean years education population aged 20+
Household_TV	Percentage households with a TV
Household_C	Percentage households with a computer
Household_W	Percentage households with piped water
Household_E	Percentage households with electricity
Household_P	Percentage households with a phone
Household_I	Percentage households with internet access
HouseholdSize	Average household size
Patrilocality	Patrilocality index (positive values patrilocal)
Dependency	Dependency ratio
Demographic parameters	
FertilityRate	Total fertility rate
InfantMortality	Infant mortality rate
Population	Total area population in millions
PopulationU	Urban population size
Epidemiologic parameters	
Dose1	Number of vaccinated people that have received first dose
Dose2	Number of vaccinated people having received second dose
Pfizer	Number of vaccinated people with Pfizer vaccine
Sinopharm	Number of vaccinated people with Sinopharm vaccine
Test	Number of tested people
PositiveCases	Number of people that tested positive
ASYM	Patients that are asymptomatic and mild symptoms
Hmoderate	Patients with moderate symptoms
Hsevere	Patients with severe symptoms
Deaths	Number of deaths recorded
Recovered	Number of recovery recorded
R_1	Daily reproduction number of the first day of infection
R _{max1}	R maximum first wave
R _{max2}	R maximum second wave
BETA	Transmission rate

5. Application to the COVID-19 outbreak in the world

We present in Fig. 14 the global daily new cases of COVID-19 in the world and then smooth the weekly data from October 26 2020 to January 17 2021. The red curve in Fig. 15 represents the cubic spline approximation of the raw data with the following statistics for the second wave: p -value = 0.05592, Multiple R-squared MR = 0.9949, Relative Mean Square Error RMSE = 41426.05.

We calculate the time t_i of the first inflection point P_I of the second wave (which presents a shoulder and no maximum) of the new infectious cases $X(t)$ for the entire world: this time corresponds to a second wave shoulder, in which the maximum of the new cases is not existing for this wave, hence cannot be used for identifying the epidemiologic parameters as in [10,18,22]. For that, we have smoothed the raw data of the weekly reported new infectious cases in Fig. 13 by using a cubic spline and we have found $t_i = 4$ (weeks).

Then, from Eqs. (10) and (11), we get:

$$a = \log(S_o - 1)/\beta \text{ and } 1 - 6e^{\beta(t_i - a)} + 2e^{2\beta(t_i - a)} = 0$$

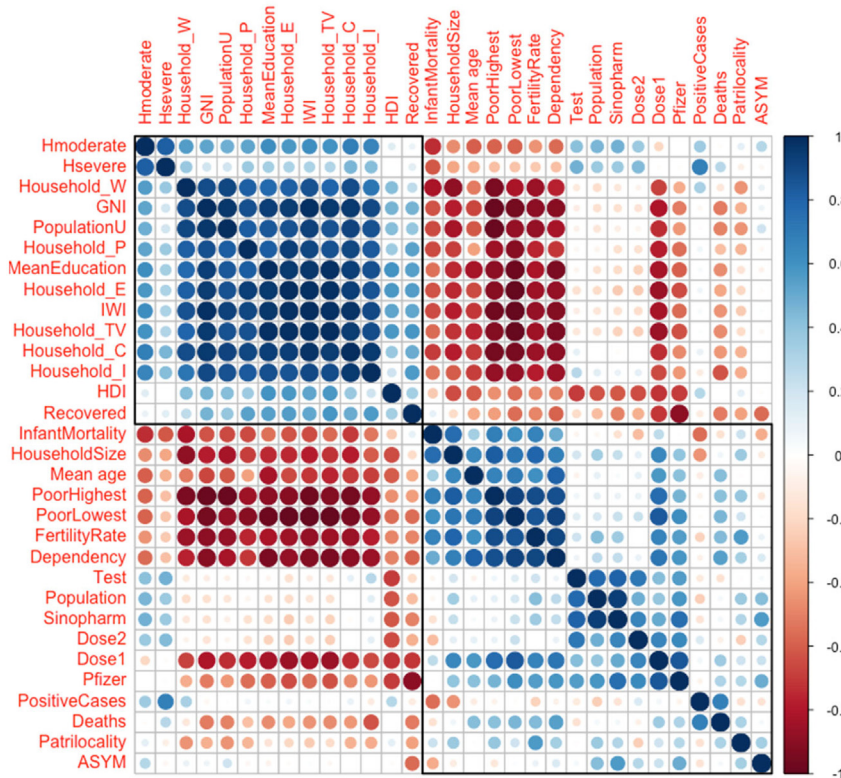


Fig. 5. Correlations between various epidemiologic, socio-economic and demographic parameters observed in the different regions of Cameroon during the second wave. The false color scale indicates the intensity of the correlation between -1 and 1 . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with $S_0 = 7,000,000,000$ and $t_i = 4$. Then $a = 10/\beta$ and $1 - 6e^{4\beta-10} + 2e^{8\beta-20} = 0$, from which we get:

$$\beta = 2.575 \text{ and } a = 3.883,$$

which is consistent with the values of the effective reproduction number calculated in [26].

6. Discussion

We have introduced in the paper two new methods, the first for estimating the maximum number of daily reproduction numbers and the second for calculating the rate of transmission by identifying the inflection points of the curve of new infectious cases observed in the framework of the classical Bernoulli model. Then, we have applied these methods to the COVID-19 outbreak in Cameroon, and eventually to the whole world data.

The analysis of the results shows that the model is very sensitive to the length of the contagiousness period r , taken equal to 4σ , in the case of a Gaussian distribution $\mathcal{N}(m, \sigma)$ for the daily reproduction number. When the value of the standard deviation σ increases, the larger its value, the smaller the value of the maximum R_{\max} of the daily reproduction numbers and vice-versa. In the case of a bimodal distribution, if $R_{\max1} \geq 0.5$, $R_{\max2}$ will be negative and if we assume that the values of $R_{\max1}$ or $R_{\max2}$ and values of transmission rate for the countries in consideration are close, then the values of $R_{\max1}$ or $R_{\max2}$ are close for all the countries.

The results shown in Tables 1 and 2 are similar to the effective values of R_e given in [23] and to the values of the deterministic R_0 calculated in [8] during the contagiousness period in consideration. This last remark shows the consistency of the results given in the present paper with those obtained from the classical methods of estimation of the effective and deterministic reproduction numbers.

The results obtained by neglecting the mortality due to the COVID-19 outbreak and to the demography of the susceptible population are given on Fig. 16.

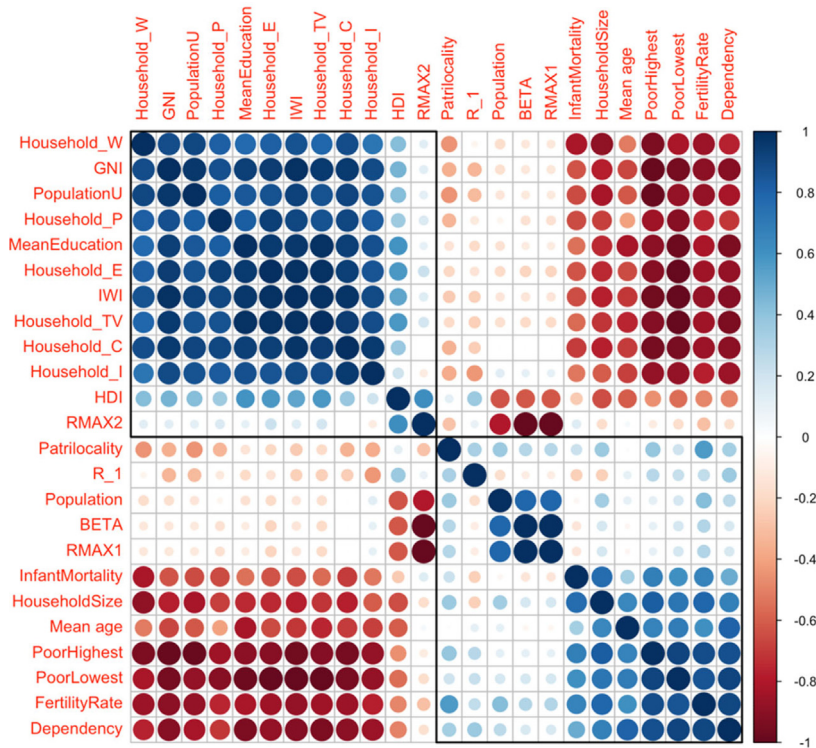


Fig. 6. Correlations between various epidemiologic, socio-economic and demographic parameters observed in the different regions of Cameroon during the second wave. The false color scale indicates the intensity of the correlation between -1 and 1 . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

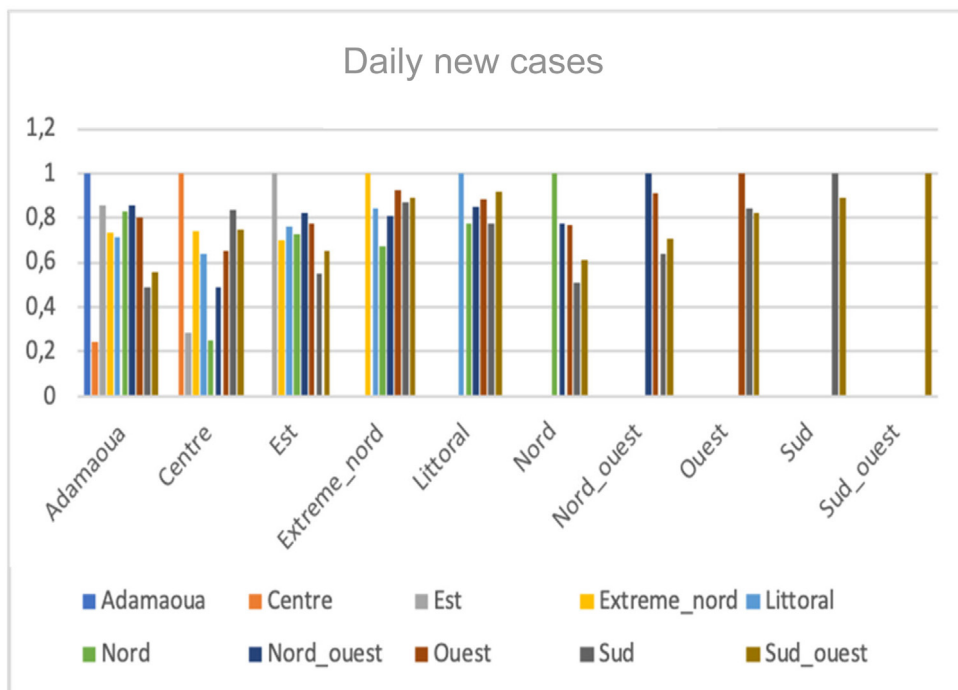


Fig. 7. Correlations of the new cases numbers between the 10 regions of Cameroon.

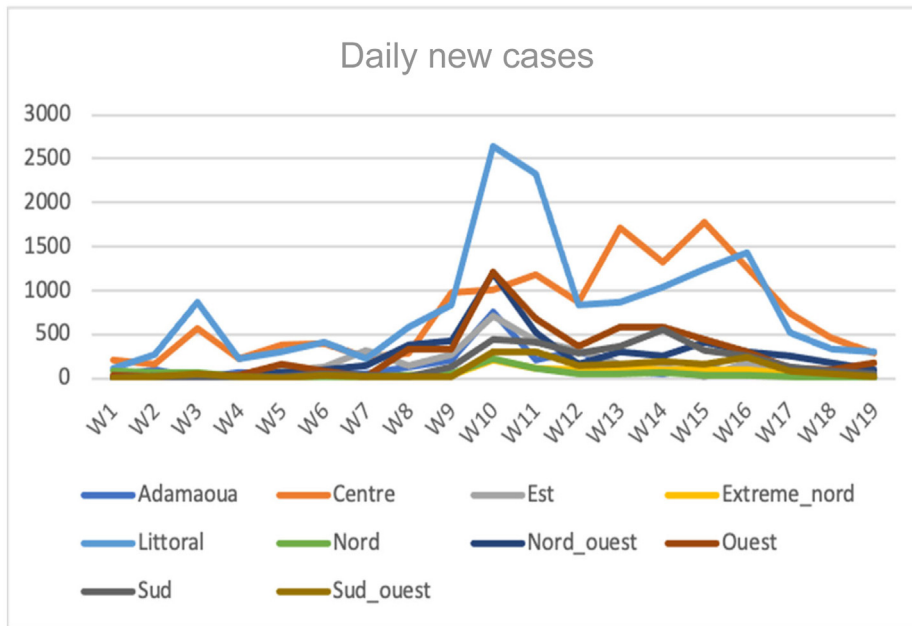


Fig. 8. Evolution of the daily new cases numbers in the 10 regions of Cameroon along the second wave between weeks 1 and 10 of 2021.

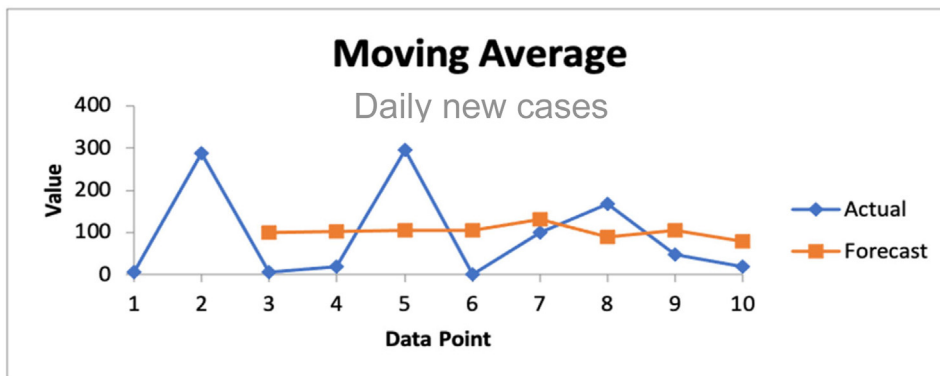


Fig. 9. Evolution of the daily new cases numbers in Cameroon along the second wave between weeks 1 and 10 of 2021.

The Fig. 16 shows slight differences at the beginning of the epidemic wave between Model 1, corresponding to Eq. (9) in which we have neglected the fertility and the mortality ($d = \nu = 0$) and Model 2, corresponding to Eq. (9) in which we have fixed the demographic balance of susceptibles d to 0.01 and the specific mortality due to the COVID-19 outbreak ν to 0.01, which are realistic values for Cameroon.

7. Perspectives and conclusion

In this article, we have given a new method for calculating the inflection point of the curve during the exponential phase of a wave. We have then deduced the value of the reproduction number at this inflection point. After that, we have calculated for each region of Cameroon the value of three sets of parameters, one devoted to socio-economic state, another to demography and the last to epidemiology, among which the value of the reproduction number at inflection. We have systematically correlated these sets of parameters and have discussed the relationships between them: the main conclusion is that socio-economic factors are (i) highly positively correlated with each other, (ii) not correlated with epidemiologic parameters (except those corresponding to vaccination and death negatively correlated

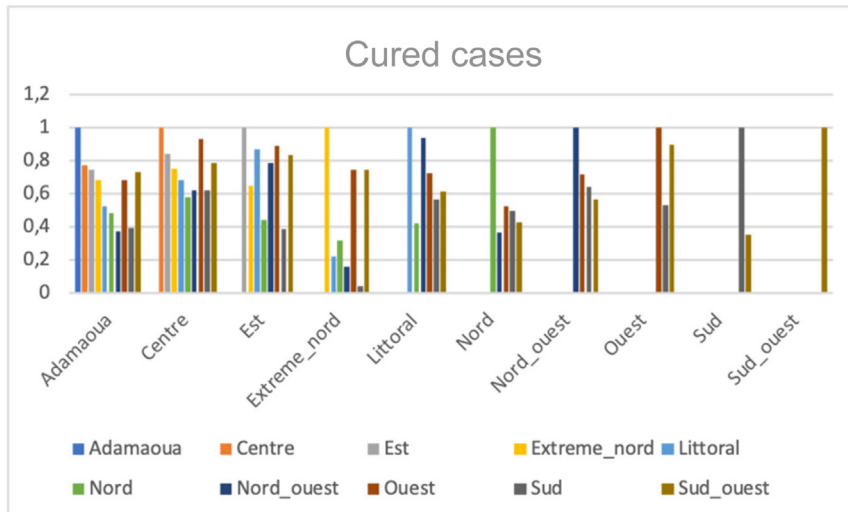


Fig. 10. Correlations of the cured cases numbers between the 10 regions of Cameroon.

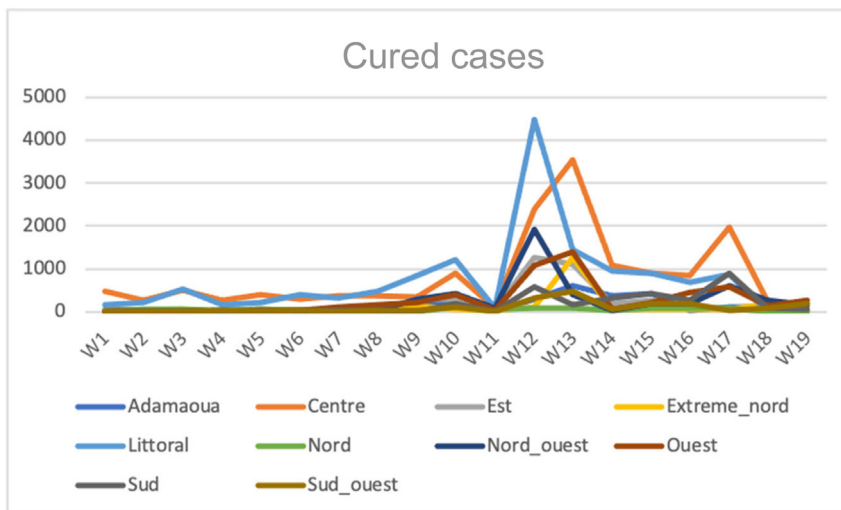


Fig. 11. Evolution of the cured cases numbers in the 10 regions of Cameroon along the second wave between weeks 1 and 10 of 2021.

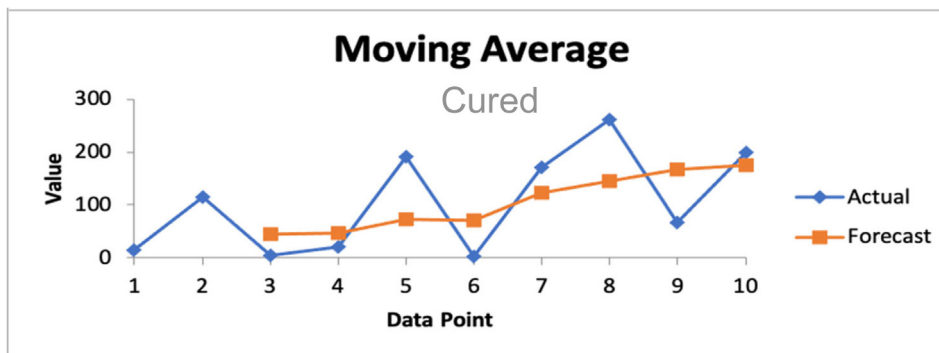


Fig. 12. Evolution of cured cases numbers in Cameroon along second wave between weeks 1 and 10 of 2021.

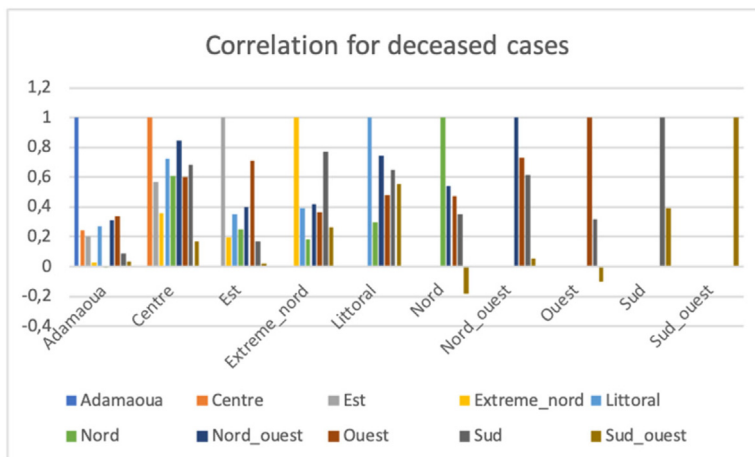


Fig. 13. Correlations of the deceased cases numbers between the 10 regions of Cameroon.

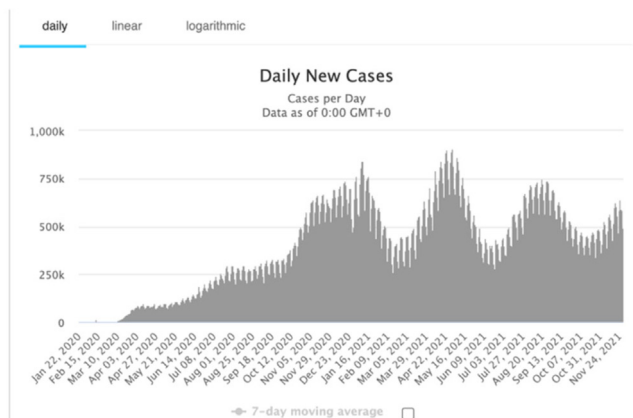


Fig. 14. World COVID-19 new cases.
Source: From [25].

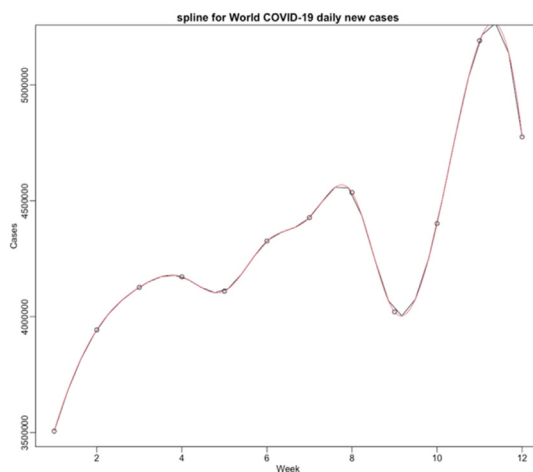


Fig. 15. Spline graph for world COVID-19 daily new cases during the second wave.

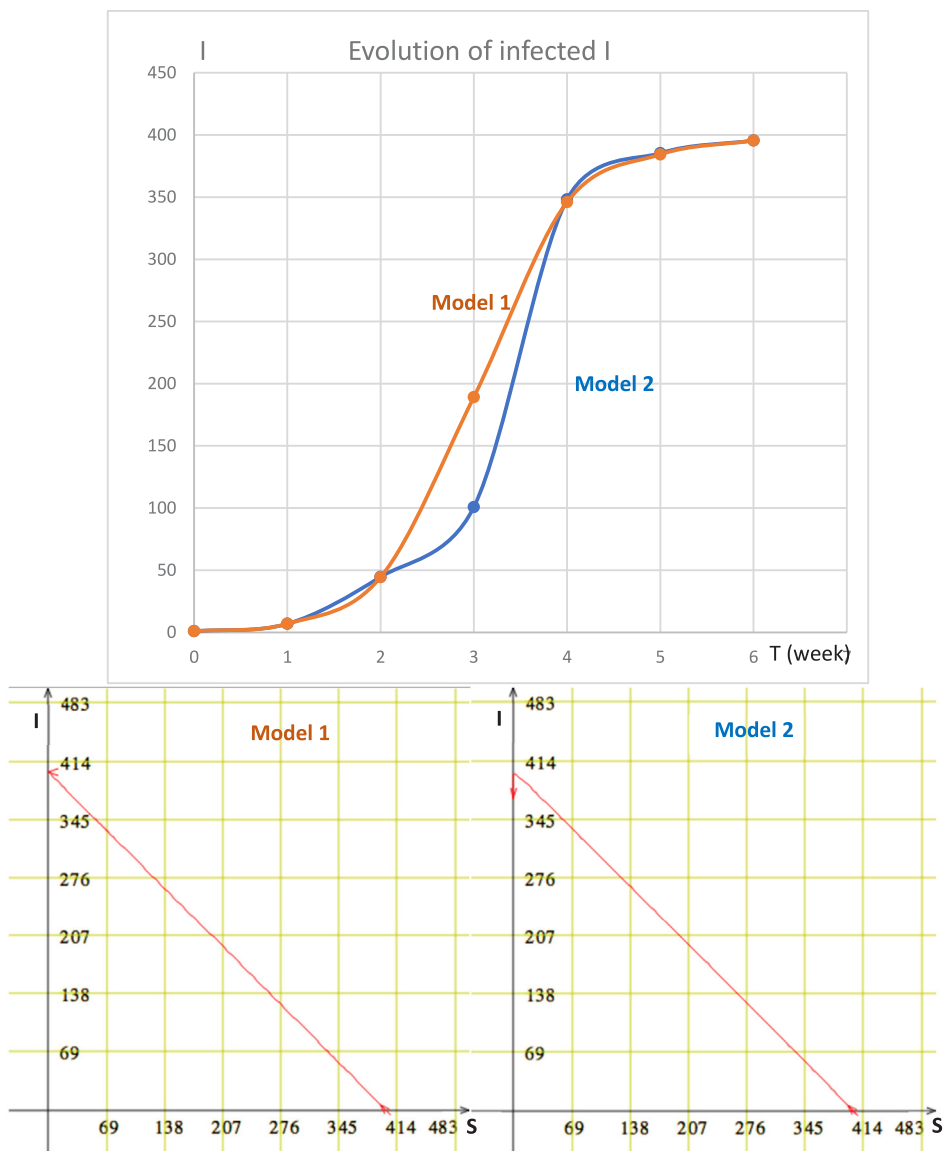


Fig. 16. Top: simulation of the evolution of the number of new infected I in Model 1 (Eq. (9) with $d = \nu = 0$) and Model 2 (Eq. (9) with $d = \nu = 0.01$). Bottom: simulation of Models 1 and 2 with same initial conditions for infected ($I(0) = 1$) and susceptible individuals ($S(0) = 400$).

with the signs of wealth) and (iii) negatively correlated with demographic parameters (namely infant mortality, mean age, fertility and poverty indicators).

In [7,12,13,20], we have already shown different complementary methods, which are available for Cameroon. In a further study, we will use the ARIMA model to analyze COVID-19 incidence pattern for all the waves in Cameroon [7], by taking into account the unidentified infectious cases [13]. We will also identify the main geoclimatic and socio-economic factors explaining the disparities between the different regions of Cameroon.

CRedit authorship contribution statement

J. Waku: Conceptualization, Methodology, Validation, Investigation, Data curation, Visualization, Supervision, Project administration. **K. Oshinubi:** Conceptualization, Methodology, Software, Validation, Formal analysis, Data

curation, Writing – original draft, Writing – review & editing, Visualization. **J. Demongeot:** Conceptualization, Methodology, Validation, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to acknowledge the Petroleum Technology Development Fund (PTDF) Nigeria doctoral fellowship in collaboration with Campus France Africa Unit.

Funding

No specific funding was received for this research.

Appendix A. Supplementary data

The data and code used for the analysis can be found in the link below: <https://github.com/Honkay/Point-of-Inflexion-and-maximal-reproduction-estimation.git>.

References

- [1] D. Bernoulli, Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir, Mem. Acad. Roy. Sci. Paris (1760).
- [2] Cameroon situation reports, 2022, Available online: <https://reliefweb.int/report/cameroon/> (Accessed 10 February 2022).
- [3] J. Demongeot, J. Gaudart, A. Lontos, E. Promayon, J. Mintsas, M. Rachdi, Least diffusion zones in morphogenesis and epidemiology, Int. J. Bifurcation Chaos 22 (2012) 1250028.
- [4] J. Demongeot, J. Gaudart, J. Mintsas, M. Rachdi, Demography in epidemics modelling, Commun. Pure Appl. Anal. 11 (2012) 61–82.
- [5] J. Demongeot, Q. Griette, P. Magal, Computations of the transmission rates in SI epidemic model applied to COVID-19 data in mainland China, R. Soc. Open Sci. 7 (2020) 201878.
- [6] J. Demongeot, Q. Griette, P. Magal, Computations of the transmission rates in SI epidemic model applied to COVID-19 data in mainland China, R. Soc. Open Sci. 7 (2020) 201878.
- [7] J. Demongeot, K. Oshinubi, M. Rachdi, L. Hobbad, M. Alahiane, S. Iggui, J. Gaudart, I. Ouassou, The application of ARIMA model to analyse Covid-19 incidence pattern in several countries, J. Math. Comput. Sci. 12 (2022) 10.
- [8] J. Demongeot, K. Oshinubi, M. Rachdi, H. Seligmann, F. Thuderoz, J. Waku, Estimation of daily reproduction rates in COVID-19 outbreak, Computation 9 (2021) 109.
- [9] A. Ducrot, P. Magal, T. Nguyen, G.F. Webb, Identifying the number of unreported cases in SIR epidemic models, Math. Med. Biol. 37 (2020) 243–261.
- [10] A. Ducrot, P. Magal, T. Nguyen, G.F. Webb, Identifying the number of unreported cases in sir epidemic models, Math. Med. Biol. 37 (2020) 243–261.
- [11] Globaldatalab, 2021, Available online: <https://globaldatalab.org/profiles/region/CMRr110/> (Accessed 10 October 2021).
- [12] Q. Griette, J. Demongeot, P. Magal, A robust phenomenological approach to investigate COVID-19 data for France, Math. Appl. Sci. Eng. 2 (2021) 149–160.
- [13] Q. Griette, J. Demongeot, P. Magal, What can we learn from COVID-19 data by using epidemic models with unidentified infectious cases? Math. Biosci. Eng. 19 (2022) 537–594.
- [14] Z. Liu, P. Magal, O. Seydi, G. Webb, Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data, Math. Biosci. Eng. 17 (2020) 3040–3051.
- [15] Z. Liu, P. Magal, O. Seydi, G. Webb, Understanding unreported cases in the Covid-19 epidemic outbreak in Wuhan, China, and importance of major public health interventions, Biology 9 (2020) 50.
- [16] P. Magal, O. Seydi, G. Webb, Y. Wu, A model of vaccination for dengue in the Philippines 2016–2018, Front. Appl. Math. Stat. 7 (2021) 760259.
- [17] A.G. McKendrick, Applications of mathematics to medical problems, Proc. Edinb. Math. Soc. 44 (1925) 1–34.
- [18] T. Nguyen, Identifiability of a Class of SIR Models and Applications (Ph.D. thesis), Bordeaux, 2021, <http://www.theses.fr/2021BORD0164>.
- [19] L. Nkague Nkamba, T.T. Manga, Modelling and prediction of the spread of COVID-19 in Cameroon and assessing the governmental measures (march–september 2020), COVID 1 (2021) (2020) 622–644.
- [20] K. Oshinubi, M. Rachdi, J. Demongeot, Analysis of reproduction number R_0 of COVID-19 using current health expenditure as gross domestic product percentage (CHE/GDP) across countries, Healthcare 9 (2021) 1247.

- [21] K. Oshinubi, M. Rachdi, J. Demongeot, Modelling of COVID-19 pandemic vis-à-vis some socio-economic factors, *Front. Appl. Math. Stat.* 7 (2021) 786983.
- [22] A. Rahman, M.A. Kuddus, Modelling the transmission dynamics of COVID-19 in six high-burden countries, *BioMed Res. Int.* (2021) 5089184.
- [23] Renkulab, 2021, Available online: https://renkulab.shinyapps.io/COVID-19-Epidemic-Forecasting/_w_f89fdf90/.
- [24] R. Ross, An application of the theory of probabilities to the study of a priori pathometry, *Proc. R. Soc. Series A* 92 (1916) 204–230.
- [25] Worldometers, 2021, Available online: <https://www.worldometers.info/coronavirus/> (Accessed 14 November 2021).
- [26] C.J. Yu, Z.X. Wang, Y. Xu, M.X. Hu, K. Chen, G. Qin, Assessment of basic reproductive number for COVID-19 at global level: A meta-analysis, *Medicine* 100 (2021) e25837.