Journal of Applied Mathematics and Physics, 2025, 13(4), 1323-1351
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

A Mathematical and Computational Model for Multiple COVID-19 Waves Applied to Kenya

Wandera Ogana[1,2*], Victor Ogesa Juma[1,3], Wallace D. Bulimo[4], Vincent Nandwa Chiteri[1]

[1]Department of Mathematics, University of Nairobi, Nairobi, Kenya
[2]African Mathematics Millennium Science Initiative, Nairobi, Kenya
[3]Mathematics Department, University of British Columbia, Vancouver, Canada
[4]Centre for Virus Research and the Department of Epidemiology, Statistics and Informatics, Kenya Medical Research Institute, Nairobi, Kenya
Email: *wogana@gmail.com, vjuma23@math.ubc.ca, Wbulimo@kemri.go.ke, vcnandwa@gmail.com

How to cite this paper: Ogana, W., Juma, V.O., Bulimo, W.D. and Chiteri, V.N. (2025) A Mathematical and Computational Model for Multiple COVID-19 Waves Applied to Kenya. *Journal of Applied Mathematics and Physics*, **13**, 1323-1351.
https://doi.org/10.4236/jamp.2025.134072

Received: February 5, 2025
Accepted: April 20, 2025
Published: April 23, 2025

Abstract

COVID-19 is a disease caused by the novel coronavirus SARS-CoV-2 that emerged at the end of December 2019 and has since spread globally. In Kenya, the virus was first detected on 13th March 2020. Soon after, the Kenyan government implemented non-pharmaceutical interventions (NPIs) to slow the spread of the disease. The pandemic continued to spread and it evolved into several waves over the years despite the discovery of vaccines and treatment. Mathematical models have been developed to help analyse, predict and simulate the dynamics of the pandemic. These models have largely been confined to single waves, without ready extension to multiple waves. In this paper, we develop a mathematical and computational model that can be extended to multiple waves using various concepts. Among these is the application of computational techniques that convert infection curves with negative gradients to those with positive gradients, in the neighbourhood of the change point, namely, where transition occurs from one wave to the next. This effectively generates a new wave. We then introduce a jump mechanism for the susceptible fraction, allowing further computation to align itself with the observed infection curve. To commence the process, we solved the system of governing ordinary differential equations for the period the epidemic spread without intervention and obtained values for the transmission, recovery and death rates that yielded the basic reproduction number, $\mathcal{R}_0 = 2.76$, which is consistent with other related research. We then applied our model to COVID-19 in Kenya and the computation successfully replicated all the waves and also identified the change points located within the months when COVID-19 variants became dominant. The findings strengthen the proposition that the dominant COVID-19 variants were the major drivers of the waves. The techniques can

be extended to new strains of COVID-19, influenza and other respiratory viruses.

Keywords

Mathematical Model, COVID-19 Pandemic, Non-Pharmaceutical Interventions, Delay-Functions, Multiple Waves

1. Introduction

COVID-19 is a disease caused by the novel coronavirus SARS-CoV-2 that emerged at the end of December 2019 and has since spread globally. The disease has hurt the socioeconomic and health structures of many countries. In Kenya, the virus was first detected on 13th March 2020. Soon after, the Kenyan government implemented non-pharmaceutical interventions (NPIs) to slow the spread of the disease [1]. Due to low compliance by the public and a rapid rise in infection, the government imposed more stringent measures that remained in effect from 9th April 2020 to 8th June 2020. Enforcement of the measures adversely affected the country's economy and people's livelihoods. Consequently, the government gradually relaxed some of the mitigation measures during the period from 9th June 2020 to 8th August 2020. Up to this stage, the COVID-19 dynamics was largely influenced by the NPIs imposed by the government [1]. The daily infections continued to go down but suddenly they started to increase in September 2020, fuelled primarily by the 2nd wave of COVID-19. As in many countries, the pandemic in Kenya continued to oscillate in uneven waves of varying amplitudes that were driven by various factors, such as new variants of concern [2], NPIs [3] [4], vaccines and therapy [5] [6], human behaviour [7], health system status [8], and host sensitivity to the virus and disease [9] [10], among others.

The pandemic stimulated extensive research by collaborators from many disciplines determined to address the challenges posed by the disease. Mathematical models were developed to analyse, predict and simulate the dynamics of the pandemic, taking into consideration the myriad drivers [11] [12]. Most widespread are compartmental models in which the human population is traditionally divided into five compartments, namely Susceptible (S), Exposed (E), Infected (I), Recovered (R) and Dead (D) [11]-[14]. By considering the rate of change of individuals in a compartment, and the contribution of appropriate compartments to this change, we obtain a system of ordinary differential equations, containing parameters that define the rate of flow between adjacent compartments. Models that use all of the compartments are called SEIRD. Various acronyms arise, depending on which compartments are in use. The effects of the COVID-19 drivers may be incorporated by creating additional compartments, for example, asymptomatic spread [9] [10], mask-wearing [15], quarantine, vaccination and hospitalization [14] [16]-[18], asymptomatic and varying symptoms [9] [10] [19], and social distancing [20]. This leads to more complicated systems with more ordinary differential equations and increased numbers of parameters. These models can replicate

sub-waves and spikes, due to the NPIs, but they cannot be used to determine how one major wave develops from another.

The emergence of several COVID-19 waves emphasized the need to investigate how these waves developed and whether it was possible to predict their growth and the emergence of future waves. A few compartmental models have made some attempts in this direction. Perakis *et al*. [21] apply the discrete version of the SEIRD model and identify a time, called the change point, that marks the end of one wave and the beginning of the next. They postulate that the recovery and infection rates have jump values at the change point. The method describes well the spikes in one wave but it is less successful in generating another wave. The approach by Ghosh and Ghosh [22] is based on the assumption that the susceptible fraction increases now and again by multiple re-infection of people who have recovered. To achieve this, they add a delay term of the infection, multiplied by a parameter regarded as the rate of re-susceptibility, to the equation involving the Susceptible rate. The same term is subtracted from the equation involving the rate of change of those removed from the system. The results provide a good match for the data from India for a few selected waves. Leonov *et al*. [23] use the SEI model, with the parameters assumed to be piecewise constant. They add an arbitrary function of time to the equation involving the rate of change of infection. This additional term can be considered as a source of infections associated with all other miscellaneous sources. The final solution is obtained from the inverse problem. The method replicates spikes well but it leads to large errors when the gradients of the infection are large, as would be the case involving a new wave. Despite these early efforts, there is still a need for research in this area to come up with reliable and more efficient methods.

The overall aim of this study is to generate the fore section of the current computed wave, from a previous wave, and to apply appropriate conditions that ensure that the model of the whole wave provides the best fit to the data. We apply the techniques to the SIRD system by using the method in [1], together with new concepts in a version of this manuscript which has previously been published as a preprint [24]. To start, we noted that the computed solution of the SIRD system, in the absence of subsequent interventions, has an infection curve that dissipates with time. In addition, a new wave is formed when the observed infection curve diverges from the dissipating computed infection curve. The point of divergence gives us an approximation to the change point, namely the point at which one wave ends and another starts. We then computationally generate the fore section of a new wave by application of a large relaxation force in the neighbourhood of the change point. Some observed infection curves for new waves emanate from a change point in the neighbourhood of which the infection fractions are very low and the susceptible fractions are almost constant. The fore sections of such curves can be approximated by an exponential function. Once the fore section of the current computed wave is generated, we apply a procedure that introduces a jump in the susceptible fraction and perpetuates the growth of a wave in the right direction. Mitigation or relaxation is then undertaken, as appropriate, to yield the rest of the

wave. Using these findings, we have simulated all the COVID-19 waves in Kenya.

We present the paper according to the following outline. In Section 2.1, we describe the SIRD model and present its baseline solution. Section 2.2 contains the derivation of the equations that describe the effects of intervention on the transmission rate. Section 2.3 has a review of the results for the 1st wave as obtained in [25]. In section 2.4 we present the detailed models of the 2nd to 5th waves, starting with the generation of a wave through application of a sufficiently large relaxation. In Section 2.5 we present the detailed models of the 6th and 7th waves, starting with the generation of a wave by assuming exponential infection. In Section 2.6 we introduce error metrics. Section 3.1 contains results and a discussion of the modelled 1st wave compared with observation. Section 3.2 contains results and a discussion of the modelled 2nd to 5th waves compared with observation. Section 3.3 contains results and a discussion of the modelled 6th and 7th waves compared with observation. Consolidated results for all the waves are presented in Section 3.4. Finally, we give a few concluding remarks and recommendations in Section 4.

2. Methods

In this section, we first present the equations for the SIRD model and their baseline solutions. We then summarize the equations that describe the effects of interventions on the disease transmission rate and indicate their impact on development of the 1st wave. We present computational results of the effect of applying large relaxation forces at a point where infection decreases, in the neighbourhood of transition from one wave to another. The results form the basis of generating the 2nd to 5th COVID-19 waves in Kenya. Finally, we derive the equations for the generation of the 6th and 7th waves.

2.1. Baseline Dynamics by SIRD model and present its baseline solution

From the SEIRD compartmental model, we omit the Exposed (E) compartment and remain with the SIRD model and present its baseline solution, where the population is divided into four basic classes, namely: Susceptible (S), Infected (I), Recovered (R) and Dead (D). We assume homogeneous mixing through the population and that the total population, N, is constant over time. Let β be the infection transmission rate from the susceptible to the infected compartment, γ be the recovery rate from the infected to the recovered compartment, and δ the death rate from the infected to the dead compartment. The governing equations are [1] [13] [25]

$$\mathrm{d}S/\mathrm{d}t = -\beta SI \tag{1a}$$

$$\mathrm{d}I/\mathrm{d}t = \beta SI - (\gamma + \delta) I \tag{1b}$$

$$\mathrm{d}R/\mathrm{d}t = \gamma I \tag{1c}$$

$$\mathrm{d}D/\mathrm{d}t = \delta I \tag{1d}$$

where $S(t), I(t), R(t)$ and $D(t)$ represent the fractions of individuals in the

susceptible, infected, recovered and dead compartments, respectively, at any time *t*. System (1) is solved subject to the initial conditions

$$S(0)=S_0,\ I(0)=I_0,\ R(0)=R_0,\ D(0)=D_0 \tag{2}$$

where S_0, I_0, R_0 and D_0 are the initial fractions of the susceptible, infected, recovered and dead individuals, respectively. System (1) is also solved subject to the condition of constant total population that can be put in the form

$$S(t)+I(t)+R(t)+D(t)=1 \tag{3}$$

The unknown parameters β, γ and δ are determined by model fitting of System (1) to COVID-19 data through optimization.

COVID-19 was first detected in Kenya on 13th March 2020 and it spread till 8th April 2020 before any measures were undertaken to control the spread of the disease. This period served as the time during which the disease spread without any intervention and baseline data was being collected. With the death rate, $\delta = 0.015$, approximated from the Case Fatality Rate (CFR), model fitting of System (1) yielded the following values for the remaining parameters during the baseline period [25].

$$\gamma = 0.0518939,\quad \beta = 0.184618 \tag{4}$$

From Equation (4) we obtain the reproduction number, $\mathcal{R}_0$, and the Recovery Days, given by

$$\mathcal{R}_0 = \beta/(\gamma+\delta) = 2.76,\quad \text{Recovery Days} = 1/\gamma = 19.3 \tag{5}$$

The reproduction number gives the approximate number of secondary infections that arise from one infected individual in a population where everyone is susceptible. If we substitute the parameters in Equation (4) into System (1) we can determine the trajectory of the infected fraction, among the other variables.

2.2. The Effect of Interventions on the Transmission Rate

We assume that interventions affect only the transmission rate, $\beta(t)$ while the recovery and the death rates remain constant at $\gamma = 0.0518939$ and $\delta = 0.015$, respectively.

There exist two types of forces in epidemic interventions: *mitigation forces* that reduce the rate and extent of infection and *relaxation forces* that increase the rate and extent. Although mitigation forces are generally perceived as NPIs, arising from the implementation of government policy, we will include among them medical treatment and vaccination, as these actions help to reduce disease spread. Relaxation forces are generally perceived as the lifting of mitigation measures to limit the impact of the disease on society. We will include among them non-compliance with mitigation measures and the emergence of new variants of COVID-19, since both of these dramatically increase the spread of the disease. As noted earlier, some models introduce more compartments to incorporate mitigation and relaxation forces. Our approach is to consider the totality of mitigation and relaxation forces without isolating individual drivers.

In the rest of this subsection, we summarize the effects of interventions on the transmission rate, with details available in [1]. Let an intervention take place on day t_k and let β_b be the incoming transmission rate before and up to t_k. The main objective of intervention is to gradually change the transmission rate from β_b, by a fraction c, so that the transmission rate at the appropriate future time becomes

$$\beta_{op} = \beta\left(t_{k+m}\right) = \left(1-c\right)\beta_b \tag{6}$$

where β_{op} denotes the optimum value of the transmission rate that is achieved due to the intervention at the time t_k, c is the intervention ratio which depends on the intensity of the intervention, and $m > 0$ is the duration, in days, it takes for the transmission rate to achieve the optimum value. Ideally, m should be small so that the effects of the intervention are realized quickly.

From Equation (6) we note that when $0 < c < 1$, then $\beta\left(t_{k+m}\right) < \beta_b$; this corresponds to the intervention being a mitigation since it yields a smaller future transmission rate which represents a reduction by a fraction c of the incoming transmission rate. We call the quantity $100c$ the "percent mitigation". On the other hand when $c < 0$ then $\beta\left(t_{k+m}\right) > \beta_b$; this corresponds to the intervention being a relaxation since it yields a larger future transmission rate which represents an increase by a fraction $|c|$ of the incoming transmission rate. We call $100|c|$ the "percent relaxation". If $c = 0$ then $\beta(t) = \beta_b$ for $t > t_k$. This implies that no intervention has occurred at t_k, since the incoming transmission rate remains unchanged after the supposed intervention. If this transmission rate is used in solving System (1), the resulting infection curve, when $t > t_k$, is called the *non-intervention curve*. We shall see later that the non-intervention curve plays a significant role in identifying the type of intervention appropriate for modelling multiple waves.

If we assume that the intervention at t_k yields a transmission rate, $\beta(t)$, which changes exponentially then it can be shown that [1]

$$\beta(t) = \begin{cases} \beta_b, & t \le t_k \\ \beta_b \mathrm{e}^{g(t)}, & t_k \le t \le t_{k+m} \\ \beta_{op} = \left(1-c\right)\beta_b, & t_{k+m} \le t < \tau \end{cases} \tag{7}$$

where

$$g(t) = \left[\left(t-t_k\right)\ln\left(1-c\right)\right] / \left[t_{k+m} - t_k\right] \tag{8}$$

and τ is the time when the next intervention occurs after the optimum value has been achieved. In solving System (1), we need to take into consideration the value of the transmission rate, $\beta(t)$, according to Equations (7) and (8), depending on the assigned time, t.

2.3. First Wave

Ogana *et al*. [25] developed the model of the 1st wave in three phases: baseline, mitigation and relaxation. We present a summary of the results here because they are crucial to developing the models of subsequent waves. The baseline phase is

covered in Section 2.1. The mitigation phase was from 9th April to 8th June 2020. The incoming transmission rate was the baseline transmission in Equation (4), namely, $\beta_b(t) = 0.184618$ The best model was obtained by solving System (1), while using this value, together with m = 15 and c = 0.41 (41% mitigation) in Equations (7) and (8). From Equation (6), we conclude that the optimum transmission rate during mitigation was $\beta_{op} = 0.108925$.

The relaxation phase was from 9th June to early September 2020, when the second wave formed. The incoming transmission rate was the optimum transmission rate for mitigation, namely. $\beta_b(t) = 0.108925$. The best model was obtained by solving System (1), while using this value, together with $m = 15$ and $c = -0.24$ (24% relaxation) in Equations (7) and (8). From Equation (6), we conclude that the optimum transmission rate during relaxation was

$$\beta_{op} = 0.135067 \tag{9}$$

Consolidation of results from the three phases leads to **Figure 1** in which the modelled percent infection during the 1st wave is compared with data [25]. **Figure 1** also shows the beginning of the observed 2nd wave. Hence the value in Equation (9) serves as the incoming transmission rate for the formation of the second wave.

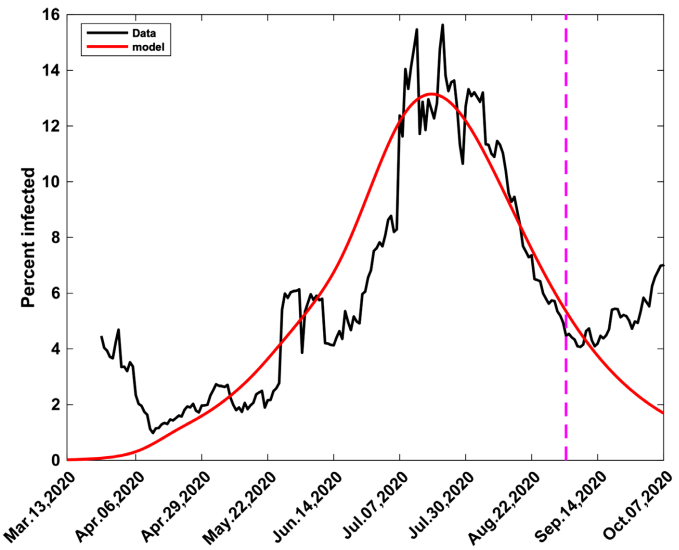


Figure 1. Computed and observed 1st wave of COVID-19 in Kenya. The black curve represents the data while the red represents the model. The purple vertical line intersects the data at the boundary of the 1st and 2nd waves.

2.4. Second to Fifth Waves

The principle behind the modelling of the second to fifth waves is that, under suitable conditions, it is possible to computationally generate a wave by applying a large enough relaxation force, when infection is on the decrease. The generated wave requires some adjustment, however, to make it replicate the observed wave. We illustrate the process by considering what happens between the first and second waves, with the understanding that equivalent results can be obtained between any two adjacent waves.

2.4.1. Application of Relaxation when Infection Decreases

From **Figure 1**, the computed first wave dissipates with time; it represents the path the actual infection would take without further interventions, including the formation of the second wave. Through computation, we were able to generate waves emanating from the 1st wave by applying relaxation forces of large magnitudes. The waves are influenced by the three factors considered below.

1) Change point

We note that the 2nd wave forms when its trajectory diverges from the tail of the computed 1st wave (**Figure 1**). In the neighbourhood of the formation of the 2nd wave, a change occurs in the disease dynamics such that the infection gradient, initially negative, starts to increase and eventually becomes positive once the 2nd wave has formed. The point of divergence of the 2nd wave from the 1st is the same as the time of commencement of the increase in the infection gradient. We use this observation to fix an approximate date for the change point, T_E, between the 1st and 2nd waves, namely, the time when the first wave ends and the second starts. The date can be estimated by drawing a line graph through the 2nd wave data and noting where the line graph meets the computed 1st wave. It may also be obtained by drawing a trendline of the 2nd wave data and noting where this trendline meets the computed 1st wave. We used both approaches and obtained an initial estimate of the change point to be in early September 2020. In modelling multiple waves, other researchers have used different methods to identify the change points [21]-[23].

As the 2nd wave forms, an increase in the infection gradient from negative to positive suggests an increase in the transmission rate. We used this postulation to carry out computational experimentation to determine the effect of relaxation forces in the neighbourhood of the change point. This led to a series of waves. To commence generation of the waves, we chose the change point, T_E, as 8th September 2020 and considered it as an initial estimate of the point at which intervention occurred that yielded the 2nd wave from the 1st. Consequently, the incoming transmission rate, β_b, was chosen as the optimal transmission rate for the 1st wave, as given by Equation (9), so that $\beta_b = 0.135067$, We chose the default value $m = 15$. We substituted these values together with selected relaxation ratios, c, into Equations (7) and (8) and solved System (1) from the change point, T_E, to stability. This yielded a series of waves which emanated from the change point. We adjusted the change point so that the waves could emanate from a point as close as possible to the point of divergence of the 2nd wave from the computed 1st wave. This led to 3rd September 2020 as the best choice for the change point and the solutions obtained by varying c yielded **Figure 2(a)**.

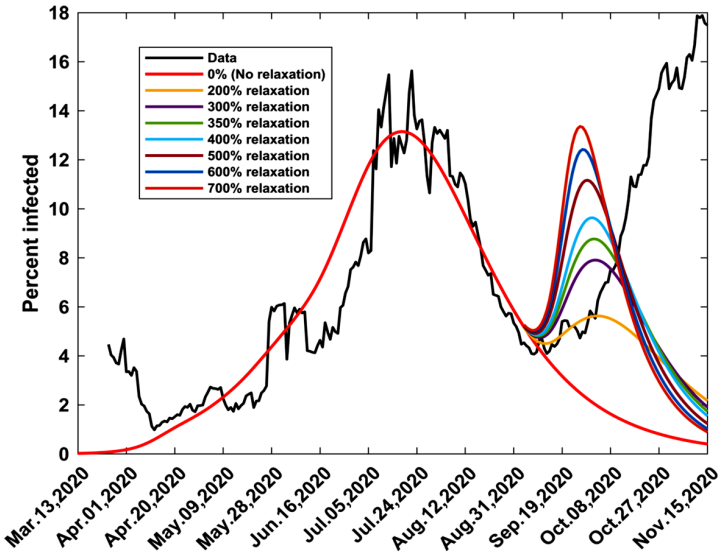
2) Wave base

From **Figure 2(a)**, we note that the computed waves and their apexes are generally above the data and the curvatures of their bases are smaller than that of the observed wave. What we desire are computed waves whose curvatures are as close as possible to that of the observed wave. By adjusting m, we obtain waves whose bases closely match the base of the observed wave. This arose from our observation that varying m, for a fixed c, leads to waves with different curvatures of the

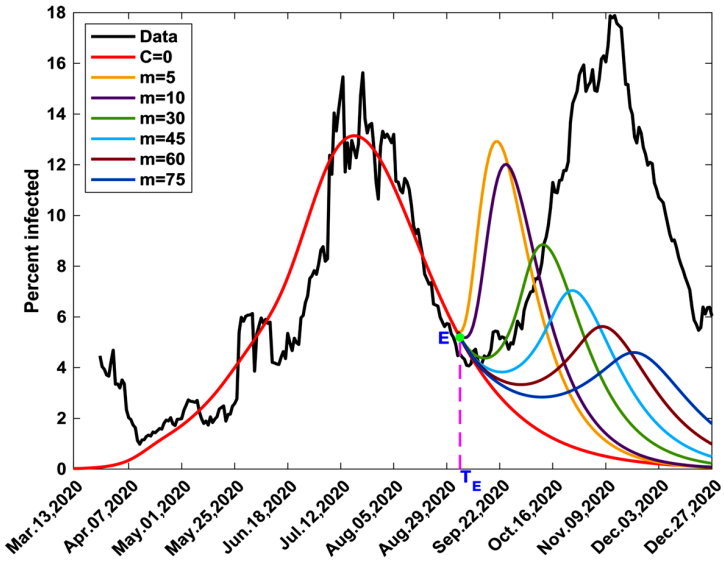
bases (Figure 2(b)). As m increases the base of the wave becomes broader while the crests of the waves become smaller and they move downwards. We found that $m = 30$ yielded waves whose bases best matched the second wave (Figure 2(c)).

3) Wave amplitude

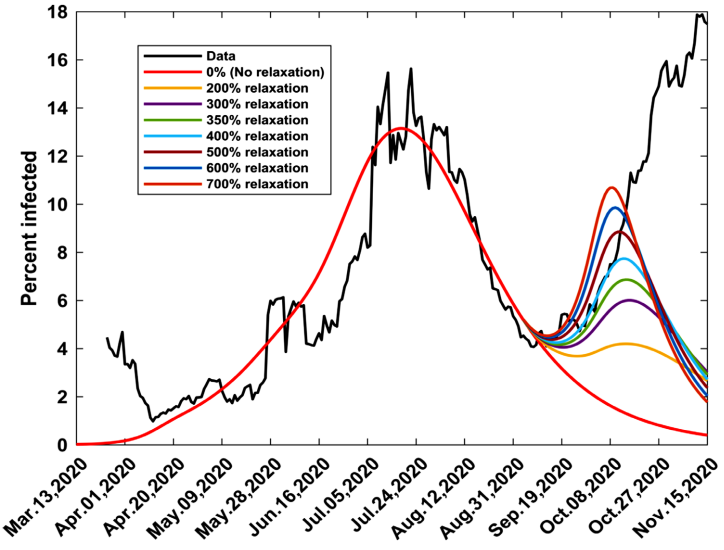
From Figure 2(a) and Figure 2(c) we notice that, for a fixed m, the waves reach a crest and dissipate thereafter. At low relaxation percentages, there does not appear to be any significant effect but, as the percentages increase, ripples begin to form and eventually turn into sharp waves with large crests which lie approximately on the same vertical line. The objective is to find a c for which the computed wave follows the observed wave as much as possible before reaching the crest. Figure 2(c) shows that for $m = 30$ and $T_E = 3^{rd}$ September 2020, the choice $c = -4$ (400% relaxation) leads to such a computed wave.



(a)



(b)



(c)

Figure 2. Effects of the application of relaxation. E is the change point from one wave to another, while T_E is the time at E. (a) Relaxations of varying magnitudes, for $m = 15$, when the infection decreases. (b) Relaxation at the same magnitude, $c = -4$ (400% relaxation), but varying m when the infection decreases. (c) Relaxations of varying magnitudes, for $m = 30$, when the infection decreases.

The phenomena decribed in this sub-section can be replicated at suitable points in the subsequent waves, with equivalent results to **Figure 2**. They form the foundation of our modelling of the 2nd to 5th waves. Every wave will be divided into the fore and back sections, each having different mechanisms of development.

2.4.2. Fore Section of the Wave by Relaxation

The observed 2nd wave starts when the decreasing infection diverges to the right and forms a basin with a concave base, then increases rapidly (**Figure 2**). The portion of the wave from the point of divergence to the point when it starts to increase rapidly is what we identify as the fore section of the wave. Section 2.4.1 shows that the fore section can be modelled by a computed wave generated by the application of a sufficiently large relaxation force. For the model to be successful, the change point and the base geometry should be obtained as accurately as possible. This can be done by the following steps.

i) a) Choose the initial change point, T_E, by examining the time when the trajectory of the observed current wave begins to diverge from the modelled tail of the preceding wave.

b) Choose the default value $m = 15$.

ii) Select a grid of negative values of c, preferably from the interval $[-10, 0)$.

iii) Determine the values at T_E of the disease variables S, I, R and D, and of the incoming transmission rate, β_b.

iv) Solve System (1) for $t > T_E$ while applying Equations (7) and (8).

v) Adjust T_E, m and c, if necessary, and repeat from step iii, till an infection curve is obtained that best fits the data at the fore section of the wave.

2.4.3. Application of Delay-Function

As Figure 2(c) shows, the curve obtained in step (v) of Section 2.4.2 will reach a maximum and dissipate at large times. To avoid this fate and force the model to follow the data upwards, we pick a point on this curve, which lies on, or closest to, the data curve, away from the maximum point. Let this point be G, with associated time T_G and infection I_G (Figure 3). From the preceding wave, or any other suitable previous wave, identify a point B, with associated time T_B and infection I_B. Choose B such that $I_B = I_G$, hence the infection at time T_B equals that at time T_G (Figure 3). Since the infections at the two points are equal and the gradients of the infection curves are in the same direction, we assume the similarity of dynamics at the two points and hence require that the rest of the time-dependent variables should also be equal at the two points. We thus impose the condition

$$\mathcal{V}(T_G) = \mathcal{V}(T_B) \tag{10a}$$

where $\mathcal{V}$ is S, I, R, D and β. The following will thus hold:

$$S_G = S_B,\ I_G = I_B,\ R_G = R_B,\ D_G = D_B \text{ and } \beta_G = \beta_B \tag{10b}$$

where the subscripts denote the values of the quantities at the corresponding points, as shown in Figure 3. Equation (10) is known as the *delay-function condition*, since values at the current time, T_G, are assigned values at a preceding or previous time, T_B. The arc E-G is what we call the fore section of the computed wave. Our objective is to make this arc match the observed infection curve as much as possible, by careful choice of the quantities T_E, m, c, T_G and I_G. If the whole of the preceding wave falls below or above I_G, then it is not possible to obtain the point I_B in the preceding wave, such that $I_G = I_B$, hence we seek I_B from a previous wave.

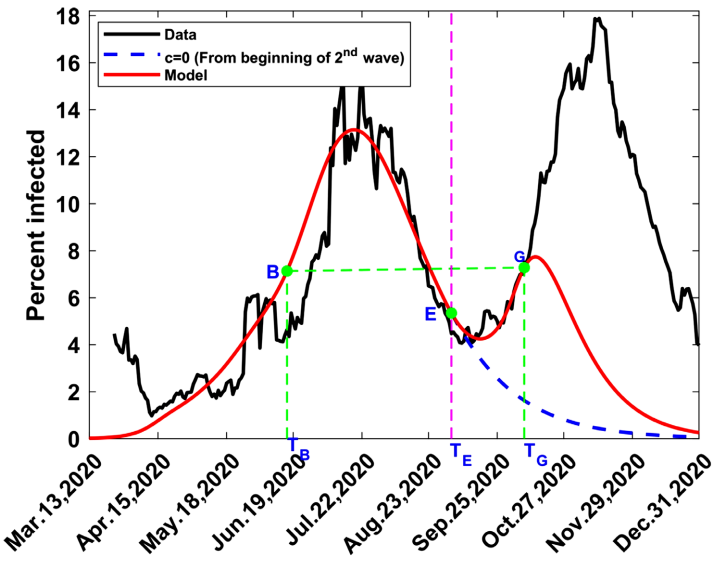


Figure 3. Choice of points for application of delay-function at the current and previous waves. m is the optimal duration of change of the transmission rate. G is the point of application of the delay-function condition, while T_G is the time at G; B is the point in the previous wave where infection at B equals infection at G and T_B is the time at B.

2.4.4. Back Section of the Wave

The back section of the computed wave begins at the end of the fore section, namely at T_G and proceeds for $t > T_G$. The fore section of the wave arises mainly from the effects of the new COVID-19 variants. The back section, on the other hand, is influenced by the new variants and NPIs. We can carry out computation for $t > T_G$ by using the values at T_G as initial values. There is no guarantee that this will yield results that match the data well, since the transmission rate at T_G is obtained through delay-function assignment and it could be different from the transmission rate associated with the data. Given the trajectory of the observed infection, it is important to determine what type of intervention, mitigation or relaxation, should be carried out at T_G or at a suitable subsequent point. The process involves generating a non-intervention curve (see Section 2.2) and comparing it with the data. Since this process will also be required at points other than T_G, we describe it for an intervention at a general point which we lable T_{VN}.

The process involves the following steps.

i) Let β_b be the incoming transmission rate at a general point of intervention, T_{VN}. Choose the default value $m = 15$.

ii) Choose the values of S, I, R and D at T_{VN} as initial values. Generate the non-intervention curve by solving System (1) for $t > T_{VN}$, with $\beta(t) = \beta_b$ and $c = 0$.

iii) Compare the trajectory of the non-intervention curve with data and one of the following two cases will arise:

Case 1: After T_{VN}, the data is predominantly above the non-intervention curve.

This means that the transmission rate for the data is larger than the transmission rate of the model, namely β_b. To obtain a model whose transmission rate is close to that of the data, we apply *relaxation* by inserting negative values of c into Equations (7, 8), and using the resultant transmission rate to solve System (1) for $t > T_{VN}$. By varying c, and adjusting m, if necessary, we can determine the transmission rate that yields an infection curve that best fits the data. This relaxation stays in effect until another intervention is encountered.

Case 2: After T_{VN}, the data is predominantly below the non-intervention curve.

This means that the transmission rate for the data is smaller than the transmission rate of the model, namely β_b. To obtain a model whose transmission rate is close to that of the data, we apply *mitigation* by inserting positive values of c into Equations (7) and (8), and using the resultant transmission rate to solve System (1) for $t > T_{VN}$. By varying c, and adjusting m, if necessary, we can determine the transmission rate that yields an infection curve that best fits the data. This mitigation stays in effect until another intervention is encountered.

In the modelling of the back section, we may use relaxation, mitigation and the delay-function condition, as convenient, to align the model with the data, if the model departs from the expected trend.

2.5. Sixth and Seventh Waves

The methods used to generate the fore sections of the 2nd to 5th waves were applied

to the 6th and 7th waves but they did not succeed, no matter how large a relaxation force was applied. We attributed this to differences in the bases of the two sets of waves. The 2nd to 5th waves start when a decreasing infection diverges to the right and forms a concave shape, with a base, then increases (Figure 2); this geometry allows for sudden and large changes in the gradients of the infection. The 6th and 7th waves, on the other hand, emerge from very low, almost horizontal, infections that gradually increase before rising sharply (Figure 4).

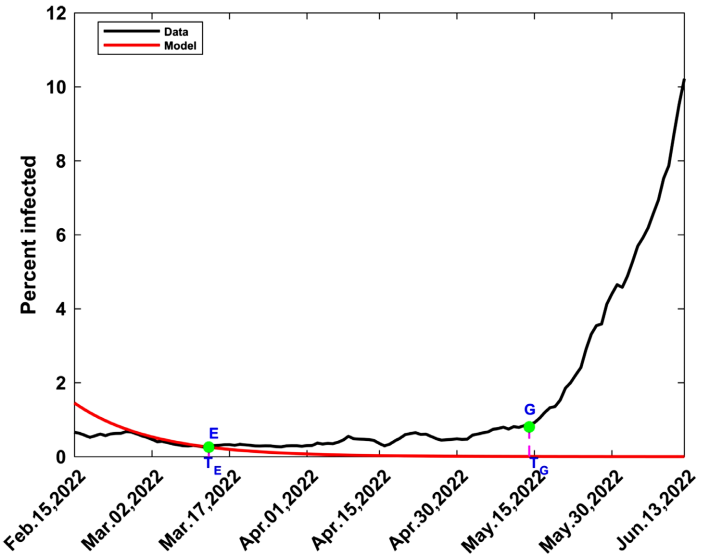


Figure 4. Development of the 6th wave. E is the change point from 5th to 6th wave and T_E the corresponding time. The black curve represents the data corresponding to the 6th wave, while the red curve corresponds to the computed tail of the 5th wave.

2.5.1. Fore Section of the Wave by Exponential Growth

The fore sections of the 6th and 7th waves can be modelled by exponential infection, as shown below. The initial change point, T_E, between the current and previous waves, is determined as before by noting when the trajectory of the current observed wave diverges from the modelled tail of the previous wave. We observed the following properties of the variables I and S at the beginning of the formation of the 6th and 7th waves:

i) Values of $I(t)$ were quite small, < 1%, in the neighbourhood of the change point, and increased slowly away from the change point.

ii) $S(t)$ was approximately constant for many days after the change point to a point close to where a sharp rise in infection commences, labelled G (Figure 4).

The arc E-G forms the fore section of the wave. Using the 2nd property in the preceding paragraph, we assume that S(t) is approximated by the mean value, $\overline{S}$, from T_E to T_G. Equation 1(b) can then be written

$$\mathrm{d}I/\mathrm{d}t = rI \tag{11}$$

where

$$r = \beta\overline{S} - (\gamma + \delta) \tag{12}$$

Equation (11) has the solution

$$I(t) = I(T_E)\mathrm{e}^{r(t-T_E)} \tag{13}$$

For the infection to grow we must have $r > 0$, implying that we must choose a transmission rate β such that

$$\beta > (\gamma + \delta)/\overline{S} \tag{14}$$

Although Equation (14) gives a wide range of choices for β, the value must be selected such that the computed infections from Equation (13) agree with the data as well as possible. This is readily done by trying different values of β and fitting the resulting exponential curve to data till a value of β is reached that yields a curve with the best fit. We denote such a value β_{XP}, to indicate that it is the transmission rate associated with exponential infection. Using Equations (12) and (13), we compute the infected fraction by

$$I(t) = I(T_E)\exp\left\{\left[\beta_{XP}\overline{S} - (\gamma + \delta)\right]\left[t - T_E\right]\right\} \tag{15}$$

Exponential growth cannot be allowed to proceed indefinitely. Just like in the 2nd to 5th waves, we terminate the fore sections of the 6th and 7th waves by enforcing the delay-function condition at the point G, as discussed in Section 2.4.3. Hence Equation (15) is used to determine the infection for t such that $T_E \le t \le T_G$. Within this time interval, $S(t) = \overline{S}$, a constant, while $R(t)$ and $D(t)$ are estimated from

$$R(t_k) = R(t_{k-1}) + \gamma I(t_k),\ D(t_k) = D(t_{k-1}) + \delta I(t_k), \tag{16}$$

through integration of Equations 1(c) and 1(d), respectively.

2.5.2. Back Section of the Wave

The back section of the wave starts at T_G and its modelling proceeds as described in Section 2.4.4, with the following points to be noted:

- The first intervention is a relaxation at T_G. It produces a model, which replicates the left half of the wave.
- The second intervention is a mitigation, near the apex of the wave, and it produces a model which replicates the right half of the wave.

2.6. Error Metrics

Graphical or visual presentation is usually the first step in assessing how well the numerical simulations fit the data. It should, however, be accompanied by a quantitative value that can be used to determine the closeness of the simulation to data. A popular beginning point involves carrying out the regression of the simulated values against the observed data and computing the coefficient of determination, R^2. In our case we use the expression

$$R^2 = 1 - \frac{\mathrm{SSR}}{\mathrm{SST}} \tag{17}$$

where

$$\text{SSR} = \sum_{k=1}^{N_T} \left(I_k - I_k^c \right)^2, \quad \text{SST} = \sum_{k=1}^{N_T} \left(I_k - \bar{I} \right)^2 \tag{18}$$

in which I_k and I_k^c are the observed and simulated infections at node *k*, respectively, $\bar{I}$ is the mean value of the observed infecions during the relevant time period, and N_T is the number of nodes for the relevant time period, and $0 \le R^2 \le 1$. The closer R^2 is to +1 the more accurately the simulation approximates the data.

In addition to R^2, we explore other metrics that can be used to determine how close the simulation is to the observed values. We use an error metric, the root mean square error (RMSE), defined thus [26]-[28]:

$$\text{RMSE} = \sqrt{\frac{1}{N_T} \sum_{k=1}^{N_T} \left(I_k^c - I_k \right)^2} \tag{19}$$

where I_k, I_k^c and N_T are defined immediately after Equation (17). If RMSE = 0 then the simulation perfectly matches the data. This will rarely be the case, hence we use the judgement that the simulation is close to the data if RMSE is close to zero. The actual value of RMSE that is acceptable depends on the dataset and the model; consequently, it is difficult to compare the significance of the RMSE values from different datasets or numerical models. To address these issues, we can obtain the normalized root mean square error, NRMSE, where RMSE is normalized by the mean, range, standard deviation, interquartile range, or square of the data, as convenient [27] [28]. In this paper we apply normalization by the range to obtain [28].

$$\text{NRMSE} = \frac{\text{RMSE}}{I_{\max} - I_{\min}} \tag{20}$$

where $I_{\max}$ and $I_{\min}$ are the maximum and minimum values of the observed infection fractions durng the relevant period. If we multiply the right hand side of Equation (20) by 100 then we express NRMSE as the percentage of the data range taken by RMSE. This enables us to interpret the significance of NRMSE as follows: the simulation is deemed very good if $\text{NRMSE} \le 10\%$, good if 10% < NRMSE < 20%, acceptable if 20% < NRMSE < 30%, and poor if NRSME > 30%. The problem with RMSE and NRMSE is that are biased estimators and the fact that they are close to zero does not guarantee that the simulation fits the data well [27]. It is, therefore, advisable to use RMSE and NRMSE in conjunction with other error metrics. Here we will also use the Hanna-Heinold error metric, HH, given by [29]

$$\text{HH} = \left[\sum_{k=1}^{N_T} \left(I_k^c - I_k \right)^2 \right] \Big/ \left[\sum_{k=1}^{N_T} I_k^c I_k \right] \tag{21}$$

where I_k^c, I_k and N_T are defined immediately after Equation (17). It is shown that HH is an unbiased estimator [27]. The closer HH is to zero, the better the fit of the simulation to the data. The HH error metric has been applied in experimental and numerical investigations involving nonlinear waves, for instance [30]-[32].

3. Results and Discussion

The presentation of results here will be done one wave at a time; thereafter all the waves will be consolidated into one time series. COVID-19 data was obtained from the Ministry of Health, Kenya [33] and Worldometer [34]. We obtained information on SARS-CoV-2 lineages and variants from [2] [35].

3.1. First Wave

The first wave was modelled by Ogana *et al.* [25] and a summary of the results is given in Section 2.3, with the complete wave shown in **Figure 1**. It was driven largely by the global parental SARS-CoV-2 lineage B.1 from March to September 2020 [2] [35]. The fluctuations in the wave were, however, partly influenced by the mitigation measures imposed on 8th April 2020 and the gradual lifting of some of these measures from 8th June 2020.

Section 2.6 yielded R^2 = 0.711 and the error metrics, RMSE = 0.0413, NRMSE = 12% and HH = 0.328. These values show that the simulation provides a fairly good fit to the data, despite the noise in the observations during the first half of the wave.

3.2. Second to Fifth Waves

The procedures are almost identical; the differences are in the dates when major events occur.

3.2.1. Second Wave

Section 2.4.2 yielded the best fit for the fore section (arc E-G in **Figure 5**), obtained using T_E = 03-Sep-20, $\beta_b = 0.13507$, $m = 30$ and $c = -4$ (400% relaxation). From Section 2.4.3, we identified T_G = 08-Oct-20 and I_G = 0.07394. Comparison with previous waves gave T_B = 17-Jun-20 and I_B = 0.07394. Equation (10) assigned values to S_G, I_G, R_G, D_G, and β_G = 0.67533. Section 2.4.4, applied at $T_{VN} = T_G$ = 08-Oct-20 with $c = 0$, $\beta_b = \beta_G = 0.67533$ and $m = 15$ led to Case 1 (blue dashed curve in **Figure 5**). Subsequent relaxation for $t > T_{VN}$ yielded the best fit for the back section (model curve for $t > T_G$ in **Figure 5**), using $c = -0.26$ (26% relaxation) and m = 15. Combination of the fore and back sections resulted in the complete 2nd wave (**Figure 5**).

Section 2.6 yielded R^2 = 0.869 and the error metrics, RMSE = 0.0202, NRMSE = 11%, and HH = 0.0344. These values show that the simulation provides a very good fit to the data.

3.2.2. Third Wave

Section 2.4.2 yielded the best fit for the fore section (arc E-G in **Figure 6**), obtained using T_E = 28-Dec-20, β_b = 0.17018, m = 45 and c = −9 (900% relaxation). From Section 2.4.3, we identified T_G = 10-Feb-21 and I_G = 0.04099. Comparison with previous waves gave T_B = 26-May-20 and I_B = 0.04099. Equation (10) assigned values to S_G, I_G, R_G, D_G, and β_G = 0.10892. An intervention at $T_{VN} = T_G$ = 10-Feb-21 using $c = 0$, $\beta_b = 0.10892$ and m = 15 resulted in a model that followed the

data but began to diverge from the data on 21-Feb-21. Hence Section 2.4.4 was applied at T_{VN} = 21-Feb-21 with $c = 0$, $\beta_b = 0.10892$ and $m = 15$, leading to Case 1 (blue dashed curve in Figure 6). Subsequent relaxation for $t > T_{VN}$ yielded the best fit for the back section (model curve for $t > T_G$ in Figure 6) using $c = -0.5$ (50% relaxation) and $m = 10$. Combination of the fore and back sections resulted in the complete 3rd wave (Figure 6).

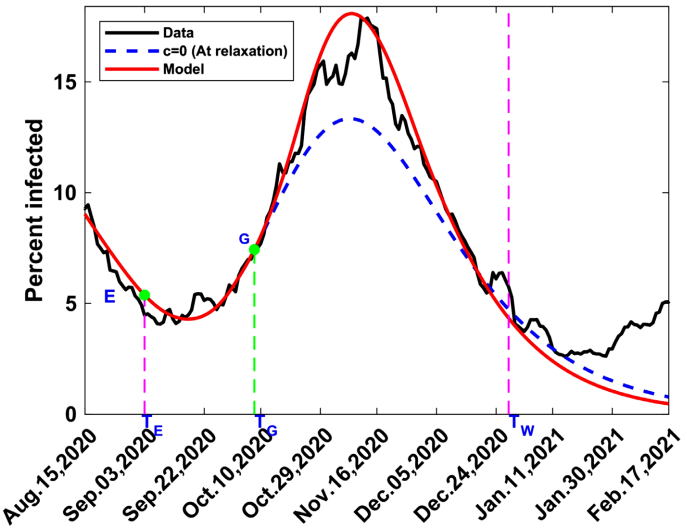


Figure 5. Computed and observed 2nd wave of COVID-19 in Kenya. The black curve represents the data, the red represents the model and the blue dashed curve is the non-intervention curve. E is the change point from the 1st to the 2nd wave, at the corresponding time T_E. G is the point of application of the delay-function condition, at the corresponding time T_G. T_W is the end of the 2nd wave.

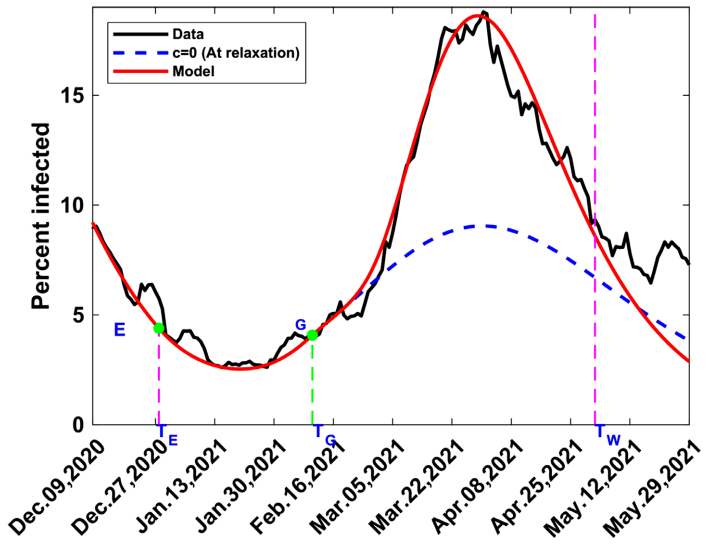


Figure 6. Computed and observed 3rd wave of COVID-19 in Kenya. The black curve represents the data, the red represents the model and the blue dashed curve is the non-intervention curve. E is the change point from the 2nd to 3rd wave, at the corresponding time T_E. G is the point of application of the delay-function condition, at the corresponding time T_G. T_W is the end of the 3rd wave.

Section 2.6 yielded R^2 = 0.903 and the error metrics, RMSE = 0.0194, NRMSE = 8%, and HH = 0.0335. These values show that the simulation provides a very good fit to the data.

3.2.3. Fourth Wave

The 4th COVID-19 wave in Kenya appeared in two spikes of different amplitudes. We modelled each spike separately before combining them to form the complete 4th wave. For convenience of presentation, we will adopt some notations as follows. There is a change point from the 3rd wave to the 1st spike which we label E1; there is another change point from the 1st spike to the 2nd and we label this E2. We let G1 and G2 be the delay-function points in the 1st and 2nd spikes, respectively, and B1 and B2 be the points for application of Equation (10) for values at G1 and G2, respectively.

First spike: Section 2.4.2 yielded the best fit for the fore section (arc E1-G1 in Figure 7) obtained using T_{E1} = 02-May-21, β_b = 0.16339, m = 30 and c = −7 (700% relaxation). From Section 2.4.3, we identified T_{G1} = 01-Jun-21 and I_{G1} = 0.0724. Comparison with previous waves led to T_{B1} = 28-Feb-21 and I_{B1} = 0.0724. Equation (10) assigned values to S_{G1}, I_{G1}, R_{G1}, D_{G1}, and β_{G1} = 1.3071. Section 2.4.4, applied at $T_{VN} = T_{G1}$ = 01-Jun-21 with $c = 0$, $\beta_b = \beta_{G1}$ = 1.3071 and m = 15 led to Case 2 (blue dashed curve in Figure 7). Subsequent mitigation for $t > T_{VN}$ led to the best fit for the back section (model curve from G1 to E2 in Figure 7) using c = 0.5 (50% mitigation) and m = 30. Combination of the fore and back sections resulted in the complete 1st spike (arc E1-E2 in Figure 7).

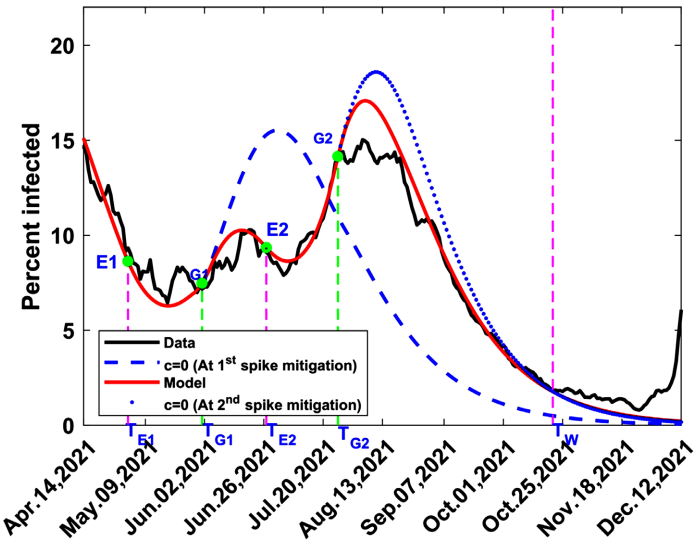


Figure 7. Computed and observed 4th wave of COVID-19 in Kenya. The black curve represents the data and the red represents the model. The blue dashed curve is the non-intervention curve in the 1st spike while the blue dotted curve is the non-intervention curve in the 2nd spike. E1 is the change point from the 3rd wave to the 1st spike of the 4th wave at the corresponding time T_{E1}; G1 is the point of application of delay-function condition in the 1st spike at the corresponding time T_{G1}. E2 is the change point from the 1st to 2nd spike at the corresponding time T_{E2}. G2 is the point of application of the delay-function condition in the 2nd spike at the corresponding time T_{G2}. T_W is the end of the 4th wave.

Second spike: Section 2.4.2 yielded the best fit for the fore section (arc E2-G2 in Figure 7) obtained using T_{E2} = 27-Jun-21, β_b = 0.079341, m = 30 and c = −2.9 (290% relaxation). From Section 2.4.3, we identified T_{G2} = 26-Jul-21 and I_{G2} = 0.143. Comparison with previous waves led to T_{B2} = 14-Mar-21 and I_{B2} = 0.143. Equation (10) assigned values to S_{G2}, I_{G2}, R_{G2}, D_{G2}, and β_{G2} = 0.29571. Section 2.4.4, applied at T_{VN} = T_{G2} = 26-Jul-21 with c = 0, β_b = β_{G2} = 0.29572 and m = 15 led to Case 2 (blue dotted curve in Figure 7). Subsequent mitigation for $t > T_{VN}$ led to the best fit for the back section (model curve for $t > T_{G2}$ in Figure 7), using c = 0.2 (20% mitigation) and m = 15. Combination of the fore and back sections resulted in the complete 2nd spike (model curve for $t > T_{E2}$ in Figure 7). The 1st and 2nd spikes were combined to yield the complete 4th wave (Figure 7).

Section 2.6 yielded R^2 = 0.81 and the error metrics, RMSE = 0.0190, NRMSE = 8%, and HH = 0.0389. These values show that the simulation provides a very good fit to the data.

3.2.4. Fifth Wave

For convenience of presentation, we will adopt some notations as follows. There will be two delay-function points which we label G1 and G2. We let B1 and B2 be the points for application of Equation (10) for values at G1 and G2, respectively.

Section 2.4.2 yielded the best fit for the fore section (arc E-G1 in Figure 8) obtained using T_E = 21-Oct-21, β_b = 0.13071, m = 60 and c = −10 (1000% relaxation). From Section 2.4.3, we identified T_{G1} = 02-Dec-21 and I_{G1} = 0.0147. Comparison with previous waves led to T_{B1} = 27-Apr-20 and I_{B1} = 0.0147. Equation (10) assigned values to S_{G1}, I_{G1}, R_{G1}, D_{G1}, and β_{G1} = 0.67286. Section 2.4.4, applied at T_{VN} = T_{G1} = 02-Dec-21with c = 0, β_b = β_{G1} = 0.67286 and m = 15 led to Case 1 (blue dashed curve in Figure 8). Relaxation was undertaken for t > T_{VN} using β_b = 0.67286, c = −4 (400% relaxation) and m = 15. The resulting model followed data but began to diverge from the data on 14-Dec-21 (model solution in arc G1-G2). Consequently, we chose a 2nd delay-function point at T_{G2} = 14-Dec-21 with I_{G2} = 0.0988. Comparison with previous waves led to T_{B2} = 16-Jul-21 and I_{B2} = 0.0985. Equation (10) assigned values to S_{G2}, I_{G2}, R_{G2}, D_{G2}, and β_{G2} = 0.39473. Section 2.4.4, applied at T_{VN} = T_{G2} = 14-Dec-21 with c = 0, β_b = β_{G2} = 0.39473 and m = 15 led to Case 1 (blue dotted curve in Figure 8). Subsequent relaxation for $t > T_{VN}$led to the best fit for the rest of the back section (model curve for $t > T_{G2}$ in Figure 8), using β_b = β_{G2} = 0.394736, c = −3.5 (350% relaxation) and m = 5. Combination of the fore and back sections resulted in the complete 5th wave (Figure 8).

Section 2.6 yielded R^2 = 0.966 and the error metrics, RMSE = 0.0211, NRMSE = 6%, and HH = 0.0315. These values show that the simulation provides a very good fit to the data.

3.2.5. Observations on the 2nd to 5th Waves

a) Attempts to generate waves from points upstream or downstream of the identified change points led to weak waves which eventually disappeared.

b) When choosing the initial values for the relaxation ratio, c, in Section 2.4.2,

it is advisable to start with $c = 0, -2, -4, -6, -8, -10$, representing 0%, 200%, 400%, 600%, 800%, 1000% relaxation, respectively. Once the suitable change point, T_E, and m have been identified, it will be easy to see the c-curves that immediately enclose the data. Hence a finer grid of c values can be applied to identify the c-value that yields the curve that best fits the data. This procedure helps to minimize computation time.

c) A summary of the computation procedure for the 2nd to 5th waves is provided in the supplement, Table S1.

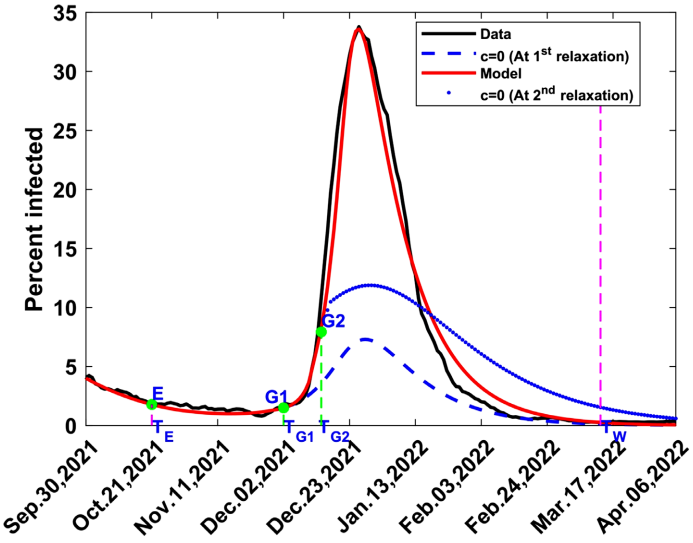


Figure 8. Computed and observed 5th wave of COVID-19 in Kenya. The black curve represents the data while the red represents the model. The blue dashed curve is the non-intervention curve from G1 while the dotted blue curve is the non-intervention curve from G2. E is the change point from the 4th to the 5th wave, at the corresponding time T_E. G1 is the first point of application of the delay-function condition at the time T_{G1}, and G2 is the second point of application of the delay-function condition at the time T_{G2}. T_W is the end of 5th wave.

3.3. Sixth and Seventh Waves

The fore sections of the 6th and 7th waves were generated by exponential approximation, rather than by relaxation, as was the case with the previous waves. The results are, therefore, presented separately in this section.

3.3.1. Sixth Wave

Section 2.4.2 yielded the best fit for the fore section (arc E-G in Figure 9) obtained using T_E = 13-Mar-22, $\bar{S} = 0.00036$, $\beta_{XP} = 238$ and T_G = 14-May-22. Equations (15) and (16) yielded disease variables from T_E to T_G such that $I_G = 0.00817$. Comparison with previous waves gave T_B = 16-Apr-20 and $I_B = 0.00827$. Equation (10) assigned values to S_G, I_G, R_G, D_G, and $\beta_G = 0.84538$. Section 2.4.4, applied at $T_{VN} = T_G$ = 14-May-22 with $c = 0$, $\beta_b = \beta_G = 0.84538$ and $m = 15$ led to Case 1 (blue dashed curve in Figure 9). Subsequent relaxation for $t > T_{VN}$ led to the best fit for the left side of the back section of the wave (model solution from T_G to T_{MT} in

Figure 9), using $c = -0.15$ (15% relaxation) and $m = 15$. We effected another intervention at $T_{VN} = T_{MT}$ = 17-Jun-22, near the apex of the wave. Section 2.4.4, applied at $T_{VN} = T_{MT}$ = 17-Jun-22 with $c = 0$, $\beta_b = 0.169597$ and $m = 15$ led to Case 2 (blue dotted curve in Figure 9). Subsequent mitigation for $t > T_{VN} = T_{MT}$ led to the best fit for the right side of the back section of the wave (model solution for $t > T_{MT}$ in Figure 9), using $c = 0.82$ (82% mitigation) and $m = 15$. Combination of the fore and back sections of the wave yielded the complete 6th wave (Figure 9).

Section 2.6 yielded $R^2 = 0.947$ and the error metrics, RMSE = 0.0093, NRMSE = 6%, and HH = 0.0309. These values show that the simulation provides a very good fit to the data.

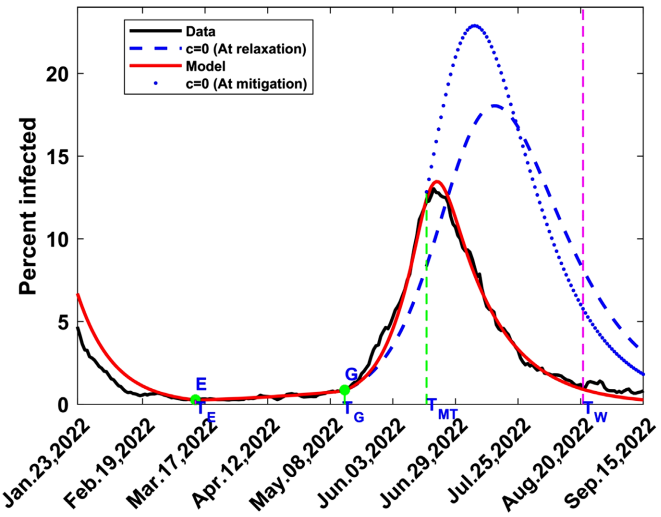


Figure 9. Computed and observed 6th wave of COVID-19 in Kenya. The black curve represents the data while the red represents the model. E is the change point from the 5th to the 6th wave, at the corresponding time T_E. G is the end of the fore section of the wave, where exponential growth terminates and T_G is the corresponding time. The blue dashed curve is the non-intervention curve from T_G. T_{MT} is a point near the apex of the wave where mitigation is applied. The dotted blue curve is the non-intervention curve from T_{MT}. T_W is the end of the 6th wave.

3.3.2. Seventh Wave

Section 2.4.2 yielded the best fit for the fore section (arc E-G in Figure 10), obtained using T_E = 22-Aug-22, $\bar{S} = 0.63$, $\beta_{XP} = 0.11$ and T_G = 10-Oct-22. Equations (15) and (16) yielded disease variables from T_E to T_G such that $I_G = 0.0101$. Comparison with previous waves gave T_B = 17-May-22 and $I_B = 0.0105$. Equation (10) assigned values to S_G, I_G, R_G, D_G, and $\beta_G = 0.029875$. Section 2.4.4, applied at $T_{VN} = T_G$ = 10-Oct-22 with $c = 0$, $\beta_b = \beta_G = 0.029875$ and $m = 15$ led to Case 1 (blue dashed curve in Figure 10). Subsequent relaxation for $t > T_{VN}$ led to the best fit for the left side of the back section of the wave (model solution from T_G to T_{MT} in Figure 10), using $c = -0.2$ (20% relaxation) and $m = 15$. We effected another intervention at $T_{VN} = T_{MT}$ = 05-Nov-22, near the apex of the wave. Section 2.4.4, applied at $T_{VN} = T_{MT}$ = 05-Nov-22 with $c = 0$, $\beta_b = 0.1781$ and m = 15 led to Case 2 (blue dotted curve in Figure 10). Subsequent mitigation for $t > T_{VN} = T_{MT}$ led to

the best fit for the right side of the back section of the wave (model solution for $t > T_{MT}$ in Figure 10), using $c = 0.68$ (68% mitigation) and $m = 10$. Combination of the fore and back sections of the wave yielded the complete 7th wave (Figure 10). We noticed some noise in the data from mid-December 2022 before it stopped being posted in the public portal of the Ministry of Health site on 26 January 2023 [33].

Section 2.6 yielded $R^2 = 0.82$ and the error metrics, RMSE = 0.0186, NRMSE = 11%, and HH = 0.1118. These values show that the simulation provides a good fit to the data, despite the noise in the data towards the end of the wave.

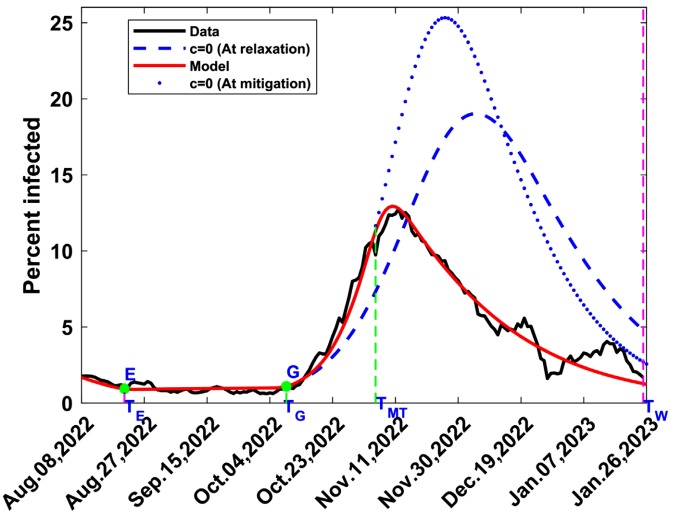


Figure 10. Computed and observed 7th wave of COVID-19 in Kenya. The black curve represents the data while the red represents the model. E is the change point from the 6th to the 7th wave, at the corresponding time T_E. G is the end of the fore section of the wave, where exponential growth terminates, and T_G is the corresponding time. The blue dashed curve is the non-intervention curve from T_G. T_{MT} is a point near the apex of the wave where mitigation is applied. The dotted blue curve is the non-intervention curve from T_{MT}. T_W is the end of the 7th wave.

3.3.3. Observations on the 6th and 7th Waves

a) When estimating the interval $[T_E, T_G]$ within which the susceptible fraction, S, is taken to be constant, in Figure 4, it is sufficient to choose the values of S that agree to one or two significant figures, unless values that agree to more significant figures are available.

b) A summary of the computation procedure for the 6th and 7th waves is provided in the supplement, Table S2.

3.4. Consolidated Results

3.4.1. Complete COVID-19 Waves in Kenya

Consolidation of Figure 1 and Figures 5-10, without the non-intervention curves, yields the complete COVID-19 waves in Kenya, as shown in Figure 11, where the model results are compared with data. The simulation and data are very close graphically.

Section 2.6 yielded $R^2 = 0.882$ and the error metrics, RMSE = 0.02360, NRMSE = 6%, and HH = 0.0711. These values show that the simulation provides a very good fit to the data.

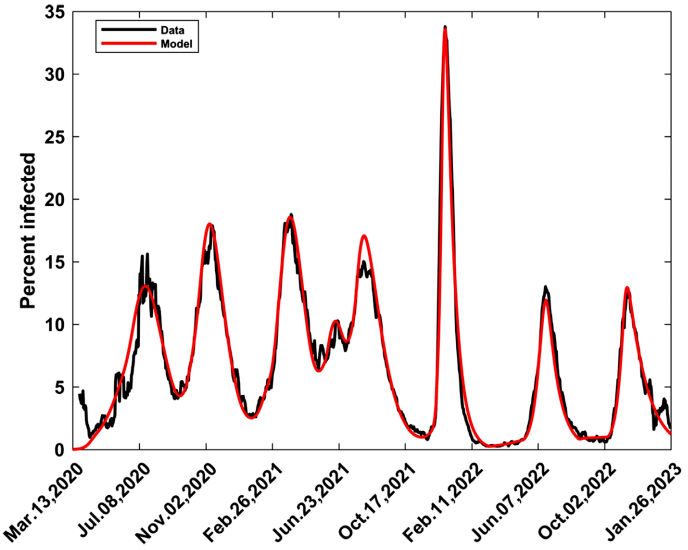


Figure 11. Computed and observed complete COVID-19 Waves in Kenya. The black curve represents the data while the red represents the model.

3.4.2. Wave Amplitudes and Durations

In **Table 1**, we present the amplitudes and durations of the waves. The durations are based on the times between identified change points and they differ slightly from those arrived at from clinical considerations [35]. The strongest wave was the 5th at 33.6% infected and the weakest was the 7th wave at 13.0% infected. The longest lasting was the 1st wave, with a duration of 174 days while the shortest was the 2nd with a duration of 116 days.

Table 1. Kenya COVID-19 wave amplitudes in 7-day averaged percent infected and durations in days.

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
Amplitude (% infected)	13.1	18.0	18.6	17.1	33.6	13.6	13.0
Duration (Days)	174	116	125	172	143	161	157

3.4.3. Change Points and SARS-CoV-2 Variants

In **Table 2**, we present the SARS-CoV-2 variants of concern and the periods when they were dominant, as determined from genomic analysis [2] [35]. We have also indicated the change points at which new waves were computationally generated by the application of large relaxation forces or through exponential approximation. We note that the change points fall in the months when the variants of SARS-CoV-2 started to be dominant. This makes us conclude that the dominant variants

of SARS-CoV-2 were the most likely sources of the large relaxation forces that could turn infections with negative gradients to those with positive gradients, thus leading to new COVID-19 waves. Hence the SARS-CoV-2 variants of concern were, most likely, the main drivers of the COVID-19 waves. There were indeed other relaxation forces arising from the lifting of, or noncompliance to, mitigation measures. In our view such forces were not, on their own, sufficiently large to generate new waves of the epidemic.

Table 2. Dominant SARS-CoV-2 variants and change points.

Wave Number	Dominant SARS-CoV-2 Variant	Period	Change Point
WAVE 1	B.1	MAR-2020 to SEP-2020	13-MAR-20
WAVE 2	Beta	SEP-2020 to NOV-2020	03-SEP-20
WAVE 3	Alpha	DEC-2020 to APR-2021	28-DEC-20
WAVE 4	Delta	MAY-2021 to OCT-2021	02-MAY-21
WAVE 5	Omicron	OCT-2021 to FEB-2022	21-OCT-21
WAVE 6	Omicron, sub-variants BA.4 & BA.5	MAR-2022 to AUG-2022	13-MAR-22
WAVE 7	Omicron BQ1 and BQ1.1	AUG-2022 to JAN-2023	22-AUG-22

4. Conclusions

In this study, we have formulated, analysed and computed a COVID-19 model that is based on the generation of two types of new waves. The first type is a wave generated from a vertically decreasing infection that diverges to the right, and forms a concave shape before increasing upwards; such a wave is modelled by generating a wave using large relaxation ratios ranging from $c = -2.9$ (290% relaxation) to $c = -10$ (1000% relaxation). The second type is a wave that emerges from very low infections with a trajectory that is almost horizontal; such a wave is modelled by an exponential infection. The models were implemented taking into consideration the characteristics of each wave. The simulations provide good to very good fit to the data. No other comparable results, based on compartmental models, are available for modelling the complete COVID-19 waves in Kenya or elsewhere. The closest to this work are results for a few waves in India [22].

The change point that determines the transition from one wave to the next is obtained by noting the tail end of the modelled preceding wave and where the trajectory of the newly observed infection begins to diverge from this tail. After identifying the change point, the transition from one wave to the next has so far been determined empirically, or computationally. Mathematical analysis should be carried out to unravel the theory behind this transition. With increased attention being paid to AI, we should, however, be prepared for the possibility that computational experimentation could yield some exciting results before theoretical justification becomes available.

We have introduced a robust mathematical and computational model that adeptly captures the dynamics of multiple COVID-19 waves and is successfully applied to Kenya. The model was developed after analysis of the wave formation patterns in Kenya. Since such patterns are not unique to Kenya, and the methods section presents descriptions in general parameter terms, it is possible to adapt this model to investigate multiple COVID-19 waves outside Kenya.

Further research is imperative to expand the model's capabilities and enhance its practical application. Future efforts should focus on incorporating real-time data integration to refine predictive accuracy, improving computational algorithms for faster real-time processing, and broadening the model's applicability to include new strains of COVID-19, influenza and other respiratory viruses. Additionally, creating user-friendly interfaces will make this tool accessible to a broader range of policymakers and health professionals, enabling informed decision-making in public health management.

Author Contributions

W.O. conceptualized the study, developed the theory, designed the computation procedures and wrote the initial draft, with input from all the authors. V.O.J. contributed to improvement of the theory, verified analytical methods and carried out the numerical simulations. W.D.B. provided the biological and biochemical interpretation of the model; also provided information on the COVID-19 variants. V.N.C. collected information on mitigation and relaxation forces due to NPIs. All authors discussed the results and contributed to the final manuscript.

Acknowledgements

We acknowledge Lucy Nyanchama for her assistance with data collection.

Data Sources

All the data used is in the public domain [33] [34].

Conflicts of Interest

The authors declare that they have no competing interests.

References

[1] Ogana, W., Juma, V.O. and Bulimo, W.D. (2024) Mathematical Modelling of Non-Pharmaceutical Interventions to Control Infectious Diseases: Application to COVID-19 in Kenya. *Frontiers in Applied Mathematics and Statistics*, **10**, Article 1365184. https://doi.org/10.3389/fams.2024.1365184

[2] Kimita, G., Nyataya, J., Omuseni, E., Sigei, F., Lemtudo, A., Muthanje, E., *et al.* (2022) Temporal Lineage Replacements and Dominance of Imported Variants of Concern during the COVID-19 Pandemic in Kenya. *Communications Medicine*, **2**, Article No. 103. https://doi.org/10.1038/s43856-022-00167-8

[3] World Health Organization and Western Pacific Region (2022) Calibrating Longterm Non-Pharmaceutical Interventions for COVID-19: Principles and Facilitation Tools.

[4] Keeling, M.J., Hill, E.M., Gorsich, E.E., Penman, B., Guyver-Fletcher, G., Holmes, A., *et al.* (2021) Predictions of COVID-19 Dynamics in the UK: Short-Term Forecasting and Analysis of Potential Exit Strategies. *PLOS Computational Biology*, **17**, e1008619. https://doi.org/10.1371/journal.pcbi.1008619

[5] World Health Organization and African Region (2023) COVID-19 Vaccines. https://www.afro.who.int/

[6] CDC Centers for Disease Control and Prevention (2019) COVID-19 Treatments and Preventive Medication. https://www.cdc.gov/covid/index.html

[7] Funk, S., Salathé, M. and Jansen, V.A.A. (2010) Modelling the Influence of Human Behaviour on the Spread of Infectious Diseases: A Review. *Journal of The Royal Society Interface*, **7**, 1247-1256. https://doi.org/10.1098/rsif.2010.0142

[8] Tessema, G.A., Kinfu, Y., Dachew, B.A., Tesema, A.G., Assefa, Y., Alene, K.A., *et al.* (2021) The COVID-19 Pandemic and Healthcare Systems in Africa: A Scoping Review of Preparedness, Impact and Response. *BMJ Global Health*, **6**, e007179. https://doi.org/10.1136/bmjgh-2021-007179

[9] Serhani, M. and Labbardi, H. (2020) Mathematical Modeling of COVID-19 Spreading with Asymptomatic Infected and Interacting Peoples. *Journal of Applied Mathematics and Computing*, **66**, 1-20. https://doi.org/10.1007/s12190-020-01421-9

[10] Anguelov, R., Banasiak, J., Bright, C., Lubuma, J. and Ouifki, R. (2020) The Big Unknown: The Asymptomatic Spread of Covid-19. *BIOMATH*, **9**, Article ID: 2005103. https://doi.org/10.11145/j.biomath.2020.05.103

[11] Cao, L. and Liu, Q. (2022) COVID-19 Modeling: A Review. https://www.medrxiv.org/content/10.1101/2022.08.22.22279022v2

[12] Saldaña, F. and Velasco-Hernández, J.X. (2021) Modeling the COVID-19 Pandemic: A Primer and Overview of Mathematical Epidemiology. *SeMA Journal*, **79**, 225-251. https://doi.org/10.1007/s40324-021-00260-3

[13] Cintra, P.H.P., Citeli M.F. and Fontinele, F.N. (2020) Mathematical Models for Describing and Predicting the COVID-19 Pandemic Crisis. arXiv: 2006.02507. https://www.doi.org/10.48550/arXiv.2006.02507

[14] Melo, L. (2021) Modeling COVID-19 Spread through the SEIRD Epidemic Model and Optimal Control. *Proceedings of GREAT Day*, **2021**, Article 19. https://knightscholar.geneseo.edu/proceedings-of-great-day/vol2021/iss1/19

[15] Eikenberry, S.E., Mancuso, M., Iboi, E., Phan, T., Eikenberry, K., Kuang, Y., *et al.* (2020) To Mask or Not to Mask: Modeling the Potential for Face Mask Use by the General Public to Curtail the COVID-19 Pandemic. *Infectious Disease Modelling*, **5**, 293-308. https://doi.org/10.1016/j.idm.2020.04.001

[16] Leontitsis, A., Senok, A., Alsheikh-Ali, A., Al Nasser, Y., Loney, T. and Alshamsi, A. (2021) SEAHIR: A Specialized Compartmental Model for Covid-19. *International Journal of Environmental Research and Public Health*, **18**, Article 2667. https://doi.org/10.3390/ijerph18052667

[17] Dutta, A. (2022) COVID-19 Waves: Variant Dynamics and Control. *Scientific Reports*, **12**, Article No. 9332. https://doi.org/10.1038/s41598-022-13371-2

[18] Malinzi, J., Juma, V.O., Madubueze, C.E., Mwaonanji, J., Nkem, G.N., Mwakilama, E., *et al.* (2023) COVID-19 Transmission Dynamics and the Impact of Vaccination: Modelling, Analysis and Simulations. *Royal Society Open Science*, **10**, Article ID: 221656. https://doi.org/10.1098/rsos.221656

[19] Morozova, O., Li, Z.R. and Crawford, F.W. (2021) One Year of Modeling and Forecasting COVID-19 Transmission to Support Policymakers in Connecticut. *Scientific Reports*, **11**, Article No. 20271. https://doi.org/10.1038/s41598-021-99590-5

[20] Kimathi, M., Mwalili, S., Ojiambo, V. and Gathungu, D.K. (2021) Age-Structured Model for COVID-19: Effectiveness of Social Distancing and Contact Reduction in Kenya. *Infectious Disease Modelling*, **6**, 15-23. https://doi.org/10.1016/j.idm.2020.10.012

[21] Perakis, G., Singhvi, D., Skali Lami, O. and Thayaparan, L. (2023) COVID-19: A Multiwave Sir-Based Model for Learning Waves. *Production and Operations Management*, **32**, 1471-1489. https://doi.org/10.1111/poms.13681

[22] Ghosh, K. and Ghosh, A.K. (2022) Study of COVID-19 Epidemiological Evolution in India with a Multi-Wave SIR Model. *Nonlinear Dynamics*, **109**, 47-55. https://doi.org/10.1007/s11071-022-07471-x

[23] Leonov, A., Nagornov, O. and Tyuflin, S. (2022) Modeling of Mechanisms of Wave Formation for COVID-19 Epidemic. *Mathematics*, **11**, Article 167. https://doi.org/10.3390/math11010167

[24] Ogana, W., Juma, V.O., Bulimo, W.D. and Chiteri, V.N. (2024) A Mathematical Model for Multiple COVID-19 Waves Applied to Kenya. https://www.medrxiv.org/node/696082.full

[25] Ogana, W., Juma, V.O. and Bulimo, W.D. (2021) A SIRD model and present its baseline solution Applied to COVID-19 Dynamics and Intervention Strategies during the First Wave in Kenya. medRxiv. https://www.medrxiv.org/content/10.1101/2021.03.17.21253626v1

[26] Chai, T. and Draxler, R.R. (2014) Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?—Arguments against Avoiding RMSE in the Literature. *Geoscientific Model Development*, **7**, 1247-1250. https://doi.org/10.5194/gmd-7-1247-2014

[27] Mentaschi, L., Besio, G., Cassola, F. and Mazzino, A. (2013) Problems in RMSE-Based Wave Model Validations. *Ocean Modelling*, **72**, 53-58. https://doi.org/10.1016/j.ocemod.2013.08.003

[28] Otto, S.A. (2019) How to Normalize the RMSE. https://www.marinedatascience.co/blog/2019/01/07/normalizing-the-rmse/

[29] Hanna, S.R. and Heinold, D.W. (1985) Development and Application of Simple Method for Evaluating Air Quality. American Petroleum Institute.

[30] Oosterlo, P., Hofland, B., van der Meer, J.W., Overduin, M. and Steendam, G.J. (2021) Calibration and Preparation of Field Measurements of Oblique Wave Run-Up and Overtopping on Dikes Using Laser Scanners. *Coastal Engineering*, **167**, Article ID: 103915. https://doi.org/10.1016/j.coastaleng.2021.103915

[31] Guillou, N., Chapalain, G. and Sergent, P. (2022) Wave Energy Resource Assessment for Small-Scale WEC near a Harbour. *Journal of Marine Science and Engineering*, **10**, Article 1081. https://doi.org/10.3390/jmse10081081

[32] Bujak, D., Lončar, G., Carević, D. and Kulić, T. (2023) The Feasibility of the ERA5 Forced Numerical Wave Model in Fetch-Limited Basins. *Journal of Marine Science and Engineering*, **11**, Article 59. https://doi.org/10.3390/jmse11010059

[33] (2020) Ministry of Health, Kenya. https://www.health.go.ke/

[34] (2020) Worldometer. https://www.worldometers.info/coronavirus/

[35] Nasimiyu, C., Matoke-Muhia, D., Rono, G.K., Osoro, E., Ouso, D.O., Mwangi, J.M., *et al.* (2022) Imported SARS-CoV-2 Variants of Concern Drove Spread of Infections across Kenya during the Second Year of the Pandemic. *COVID*, **2**, 586-598. https://doi.org/10.3390/covid2050044

Supplement

Table S1. A summary of the computation procedure for the 2nd to 5th COVID-19 Waves in Kenya.

Variables and Parameters		2nd Wave	3rd Wave	4th wave 1st Spike	4th wave 2nd Spike	5th Wave
A. FORE SECTION OF THE WAVE						
A1. Generation of new wave by relaxation						
Change point	T_E	03-09-20	28-12-20	02-05-21	27-06-21	21-10-21
Transmission rate preceding relaxation at T_E; [$=\beta(T_E)$]	β_b	0.13507	0.17018	0.16339	0.079341	0.13071
Relaxation ratio for wave generation	c	−4	−9	−7	−2.9	−10
Duration of optimal change of transmission rate	m	30	45	30	30	60
Use values of S, I, R, D and β at T_E to solve the system till the time for enforcing the delay-function as indicated below.						
A2. Application of delay-function						
Time in current wave at which delay-function is enforced	T_G	08-10-20	10-02-21	01-06-21	26-07-21	02-12-21
Approximate infected ratio at T_G	I_G	0.07394	0.04099	0.0724	0.143	0.0147
Infection in previous wave that approximates I_G	I_B	0.07394	0.04099	0.0724	0.143	0.014699
Time at which I_B occurs.	T_B	17-06-20	26-05-20	28-02-21	14-03-21	27-04-20
Apply delay-function at T_G to, I, R, D and β						
B. BACK SECTION OF THE WAVE						
B1. Intervention (Relaxation/Mitigation)						
From T_G solve the system for L days then apply intervention						
Number of days for solution after T_G	L	0	11	0	0	0
Time at which intervention is enforced	T_{VN}	08-10-20	21-02-21	01-06-21	26-07-21	02-12-21
Transmission rate preceding intervention at T_{VN}; [$=\beta(T_{VN})$]	β_b	0.67533	0.10892	1.3071	0.29571	0.67286
Apply Relaxation/Mitigation from non-intervention curve	c	−0.26	−0.5	0.5	0.2	−4
Duration of optimal change of transmission rate	m	15	10	30	15	15
B2. Application of the 2nd delay-function						
L days from previous relaxation, apply a 2nd delay-function						
Number of days to follow previous relaxation	L	-	-	-	-	12
Time at which the 2nd delay-function is enforced	T_{G2}	-	-	-	-	14-12-21
Approximate infected fraction at T_{G2}	I_{G2}	-	-	-	-	0.0988
Infection in previous wave that approximates I_{G2}	I_{B2}	-	-	-	-	0.0985
Time in previous wave where I_{B2} occurs.	T_{B2}	-	-	-	-	16-07-21
Apply delay-function at T_{B2}; Follow by relaxation at T_{B2}.						
B3. Application of the 2nd relaxation						
Time for 2nd relaxation ($=T_{G2}$)	T_{VN}	-	-	-	-	14-12-21
Transmission rate preceding 2nd relaxation; [$=\beta(T_{VN})$]	β_b	-	-	-	-	0.39473
Relaxation ratio	c	-	-	-	-	−3.5
Duration of optimal change of transmission rate	m	-	-	-	-	5
Use the values of S, I, R, D and β, and the associated Relaxation/Mitigation parameters, at the last intervention point, T_{VN}, and solve the system of ODEs for t > TVN to obtain the rest of the wave						

Table S2. A summary of the computation procedure for the 6th to 7th COVID-19 Waves in Kenya.

Variables and Parameters		6th Wave	7th Wave
A. FRONT SECTION OF COMPUTED WAVE			
A1. Generation of new wave by exponential infection			
Change point	T_E	13-03-22	22-08-22
Infected fraction at Change Point	I_E	0.00255	0.00899
Constant Susceptible fraction for exponential infection	$\overline{S}$	0.00036	0.63
Transmission rate for exponential infection	β_{XP}	238	0.11
Use values of S, I, R, and D at T_E as initial values and solve the system till the time for enforcing delay-function as below.			
A2. Application of delay-function			
Time in current wave at which delay-function is enforced	T_G	14-05-22	10-10-22
Approximate infected ratio at T_G	I_G	0.00817	0.0101
Infection in previous wave that approximates I_G	I_B	0.00827	0.0105
Time at which I_B occurs	T_B	16-04-20	17-05-22
Apply delay-function: Let $V(I_G) = V(I_B)$, where V is S, I, R, D and β			
B. BACK SECTION OF THE WAVES			
B1. Relaxation			
Apply relaxation at the point of delay-function			
Time of at which relaxation is enforced; [$=T_G$]	T_{VN}	14-05-22	10-10-22
Transmission rate preceding relaxation at T_{VN}; [$=\beta(T_{VN})$]	β_b	0.84538	0.029875
Relaxation ratio	c	−0.15	−0.2
Duration of optimal change of transmission rate	m	15	15
B2. Mitigation			
After following the relaxation curve, determine when to apply mitigation			
Time at which mitigation is applied; [$=T_{MT}$]	$T_{VN} = T_{MT}$	17-06-22	05-11-22
Transmission rate preceding mitigation at $T_{VN} = T_{MT}$; [$=\beta(T_{VN})$]	β_b	0.169597	0.1781
Mitigation ratio	c	0.82	0.68
Duration of optimal change of transmission rate	m	15	10
Use the values of S, I, R, D and β, and the associated Mitigation parameters, at the last intervention point, T_{MT}, and solve the system of ODEs for $t > T_{MT}$ to obtain the rest of the wave			