On evolution model for SARS-Cov-2 - infected population. The case of New Zealand

Estate V. Khmaladze [1] **and Giorgi Kvizhinadze** [2]

[1] *Victoria University of Wellington, PO Box 600, Wellington, New Zealand, e-mail:* Estate.Khmaladze@vuw.ac.nz

[2] *Capital and Coast District Health Board, Private Bag 7902, Wellington, New Zealand, e-mail:* Giorgi.Kvizhinadze@ccdhb.org.nz

Abstract: The work proposes a mathematical model of the process of Covid-19 epidemic as it evolved in New Zealand. The model uses a system of differential equations which emanate from natural assumptions on some probability measure and evolution of this measure on evolving family of simplexes.

The authors did not aim at mathematical complications – the model is simple and easy to follow. The aim rather was practical – to come to justifiable estimations of important parameters like the rate of infection as function of time, thus quantifying effectiveness of the Government measures. Another parameters estimated were the probability distribution of detection times and recovery times.

Keywords and phrases: Populations, Queues, Covid-19, Convolutions, Distribution of incubation times, Detection times, Distribution of recovery times, Mixtures of distribution.

1. General outlook

SARS-CoV-2 (Covid-19) infection arrived in New Zealand at the end of February 2020. At that time there was already much discussion and much anxiety around epidemics in China, and also in Italy. The public opinion was prepared for the idea that strong measures will be necessary.

It turned out that the Government understood well what was about to unfold if no measures were taken, and had a plan. According to this plan, there were four levels of social restrictions, from level 1, least restrictive, to level 4, the complete lockdown with strict limits on contacts between people everywhere: in public transport, in supermarkets, between neighbours and within extended families. Even in small shops one in, one out policy was implemented, cafes and restaurants were closed. Older people were promised essential supplies to be delivered to their homes. In principle, police could enforce the restrictions. This is not a social research, but as citizens of the country, we did not feel much interference from the police.

In parallel, testing and contact tracing capacities were ramped-up. At the beginning of the epidemic, on average, there were around 2000 tests conducted

daily. This number has increased to more than 5000 daily tests in the lockdown period, cf. NZ MOH-1 [2020]. Similarly, over the same time period, contact tracing capacity has increased from the ability to trace contacts of couple of hundred daily cases to the ability to trace contacts of up to 1000 daily cases, cf. NZ MOH-2 [2020]. These measures: closure of the boarder, social distancing, timely testing of large numbers of people and quick and effective contact tracing all contributed to the final result which made New Zealand one of the most effective countries in the fight against Covid-19.

At the announcement of the Government's policy, Prime Minister called the epidemic the "textbook case". The events and the handling of these events have shown that this was correct. To come up with an adequate quantitative model became a task for probabilistic and statistical community of New Zealand. The cleaner and closer to "textbook" it gets, the more useful it becomes. In this work we did not try to fit the data as such, but to explain the data in its several aspects.

There is a temptation and, we think, a danger, in work like this: it is tempting to write a model of how we think the epidemic should evolve in clean, almost a laboratory conditions. We, however, devote only one section, section 2, to this. Mostly, we analyse the process of eradication of the infection as it evolved in reality of the country, along with the data which described this process.

2. The model: simplexes, differential equations

Consider a population of people, who may be vulnerable to disease, like, in our case, the population of New Zealand. Let N denote the size of this population. Suppose we start observations at some moment called $t = 0$; say, the day zero is chosen as 1 January 2020. We describe each individual in this population by three moments in time: S is the moment when the unaffected individual becomes infected; if the person was never infected, we say the S is infinite; T is the incubation period, the time which elapses from infection until the symptoms become apparent and the person is ill; Z is the time until recovery or, in rare unhappy cases death. In other words, S is calendar time after zero, when the person becomes infected, $S + T$ is the calendar time when the person becomes ill, and $S + T + Z$ is the calendar time of recovery.

These S, T and Z are for us non-negative random variables, and below we say what we assume about their distribution(s). Therefore, our model starts with N triplets $(S_i, T_i, Z_i)_{i=1}^N$. Now introduce several point processes, which are of central importance in description of the epidemic. Three of them will be counting processes, i.e. of cumulative character, with non-decreasing trajectories, and two

of them will be queues, that is, differences of counting processes. Namely,

$$\begin{aligned} X_0(t) &= \sum_i \mathbb{I}_{(S_i \leqslant t)}, \\ X_{1a}(t) &= \sum_i \mathbb{I}_{(S_i+T_i \leqslant t)}, \\ X_{2a}(t) &= \sum_i \mathbb{I}_{(S_i+T_i+Z_i \leqslant t)}. \end{aligned} \tag{1}$$

Think about X_0 as the main, "driving" counting process, with S_i labeled in increasing order, so that S_1 is the moment when the first infection occurs, S_2 is the second such moment, and so on. Think about the pair (T_i, Z_i) as a 2-dimensional mark, associated with S_i. Summation runs over all individuals in our initial population. Here $\mathbb{I}_A$ is indicator function of an event A. For example,

$$\mathbb{I}_{(S_i \leqslant t)} = \begin{cases} 1, \text{ if } S_i \leqslant t, \\ 0, \text{ if } S_i > t, \end{cases} \qquad \text{and} \quad \mathbb{I}_{(S_i+T_i > t)} = \begin{cases} 1, \text{ if } S_i + T_i > t \\ 0, \text{ if } S_i + T_i \leqslant t \end{cases}$$

are indicator functions of events $S_i \leqslant t$ and $S_i + T_i > t$.

According to their definitions,

$X_0(t)$ is the number of individuals, infected up to time t,
$X_{1a}(t)$ is the number of individuals, which became symptomatic up to time t,
$X_{2a}(t)$ is the number of individuals who were ill and recovered up to time t.

The process X_0, although of major importance, is not observable. One can make guesses about the moments S_i-s, not of many of them, and this would not be what one calls observation. However, the processes X_0, X_{1a} and X_{2a} are interconnected and, as the model below shows, evolution of X_0 can be inferred from observations of X_{1a} and X_{2a}.

Now introduce two more processes,

$$\begin{aligned} X_1(t) &= \sum_i \mathbb{I}_{(S_i \leqslant t)} \mathbb{I}_{(S_i+T_i > t)} = X_0(t) - X_{1a}(t), \\ X_2(t) &= \sum_i \mathbb{I}_{(S_i+T_i \leqslant t)} \mathbb{I}_{(S_i+T_i+Z_i > t)} = X_{1a} - X_{2a}(t), \end{aligned}$$

which are differences of our counting processes and, thus, form queues, see, e.g. Breamud [1981]. Their heuristic meaning is also clear:

$X_1(t)$ is the number of infected but still asymptomatic individuals at time t,
$X_2(t)$ is the number of individuals who are ill (symptomatic) at the time t.

In contrast to the process X_2, which was observable and the data accurately reported, the process X_1 was, although important, unobservable.[1]

[1] In New Zealand, community testing has been done to a certain degree, especially towards the end of lockdown. However, there was no case detected as the result of it, probably, due to small prevalence of infection.

Later we will use the same notations for populations themselves, not just for size of these populations; we may say, for example, that individual belongs to population $X_1(t)$ meaning that at time t the individual was infected but not symptomatic.

Let us think about t as time measured in days; intensities (or rates) below will, therefore, be daily intensities or rates.

Distribution of the triplets. Assume now that within each triplet (S_i, T_i, Z_i), the random variables T_i and Z_i are independent from S_i and independent from each other. Also assume that distribution of all $T_i, i = 1, 2, \dots,$ is the same and G is the distribution function of each of them. Similarly, assume that $Z_i, i = 1, 2, \dots,$ are identically distributed with distribution function F. The last assumptions are some simplifications, as distributions G and F may depend, for example, on the age of the individual, and also on the viral load. However, for adequate description of the populations this may not be very important.

The spread of our triplets in the positive quadrant $\mathbb{R}^3_+$ can be fully described by the counting measure ν_N: for any set $A \in \mathbb{R}^3_+$,

$$\nu_N(A) = \sum_i \mathbb{I}_{((S_i,T_i,Z_i)\in A)}.$$

For example, on rectangles we have

$$\nu_N((0,t] \times (0,x] \times (0,z]) = \sum_i \mathbb{I}_{(S_i \leqslant t)} \mathbb{I}_{(T_i \leqslant x)} \mathbb{I}_{(Z_i \leqslant z)}$$

and

$$\nu_N((0,t] \times \mathbb{R} \times \mathbb{R}) = \sum_i \mathbb{I}_{(S_i \leqslant t)} = X_0(t).$$

The assumption of independence of S_i, T_i and Z_i implies that the expected value of ν_N is the product measure,

$$E\nu_N((0,t] \times (0,x] \times (0,z]) = H(t)G(x)F(z) = Q(t,x,z),$$

where $H(t)$ is the expected value of $X_0(t)$:

$$E\nu_N((0,t] \times \mathbb{R} \times \mathbb{R}) = EX_0(t) = H(t).$$

Introduce the following simplexes, which depend on t, and their complements:

$$\begin{aligned} A_{1a}(t) &= \{(s,x,z) : s + x \leqslant t\}, \\ A_1(t) &= \{(s,x,z) : s \leqslant t, s + x > t\}, \\ A_{2a}(t) &= \{(s,x,z) : s + x + z \leqslant t\}, \\ A_2(t) &= \{(s,x,z) : s + x < t, s + x + z > t\}, \end{aligned}$$

Note that $\nu_N(A_{1a}(t))$ and $\nu_N(A_1(t))$ are nothing else but the processes $X_{1a}(t)$, and $X_1(t)$, respectively. Indeed, if the triplet $(S_i, T_i, Z_i) \in A_{1a}(t)$ then the corresponding person belongs to the sub-population $X_{1a}(t)$ and if $(S_i, T_i, Z_i) \in A_1(t)$ then the corresponding person belongs to the sub-population $X_1(t)$. Similarly,

$$\nu_N(A_{2a}(t)) = X_{2a}(t), \quad \text{and} \quad \nu_N(A_2(t)) = X_2(t).$$

Therefore, to describe evolution of our point processes is the same as to describe the evolution of ν_N on the simplexes above.

Expected values of X_{1a} and X_1. To draw parallels, note that the assumption of independence of S_i and T_i and i.i.d-ness of T_i-s, will render the pairs (S_i, T_i)-s a classical model of a stationary population with S_i-s as the moments of birth and T_i-s as the lifetimes. If the population is not stationary and its growth rate can change in time, the expected value $H(t) = EX_0(t)$ should be anticipated to follow differential equation

$$dH(t) = m(t)dt + r(t)H(t)dt \tag{2}$$

(out of many sources, we refer, e.g., to Bjornstad [2018], Pollard [1973] and Khmaladze [2013], Lecture 14). In the first term, $m(t)$ is the (daily) immigration rate, or for us, the daily immigration rate of infected people. Even in absence of the exact data, one could model $m(t)$ as a given fraction of the total arrivals to New Zealand. However, Ministry of Health knows, and has published, the dates of arrival of such people. We consider the case of the time varying $m(t)$, because due to the Government measures, the daily arrivals indeed changed in time down to zero; it may change again and gradually increase. Some finer points of this data we consider in Section 3. In the second term, $r(t)$ denotes the growth rate, or for us, the infection rate of the population. Heuristically, $r(t)$ is the expected number of people a randomly selected infected person will infect on the day t. In fixed social conditions and habits of life this intensity depends on how aggressive the virus is and can be assumed constant. However, social conditions in New Zealand did not stay the same – we moved through levels. Therefore it also should be considered variable in time. We discuss the concept of infection rate again in Section 3.

No assumptions about the "lifetimes" T_i-s participate in equation (2). If we turn from the "birth" or arrival process X_0 to "death" or exit process X_{1a} formed by those with confirmed infection, and then to the population X_1, the usual model for X_0 and X_{1a}, or the model for their expected values will be

$$dH(t) = m(t)dt + r(t)Q_1(t)dt, \quad dQ_{1a}(t) = \mathsf{d}(t)Q_1(t)dt, \tag{3}$$

where Q_{1a} and Q_1 are expected values of X_{1a} and X_1, respectively. Here $\mathsf{d(t)}$ would be what is called the exit rate, and the heuristic reasoning behind the equations (3) would be that a randomly selected individual from the population $X_1(t)$ in the time interval $[t, t + dt)$ will create a new infected individual with

probability $r(t)dt$ and will exit with probability $\mathsf{d}(\mathsf{t})dt$, or will do neither of these with probability $1-(r(t)+\mathsf{d}(t))dt$. Taking the difference of the two equations, we obtain very familiar equation for the expected population size:

$$dQ_1(t) = m(t)dt + \rho(t)Q_1(t)dt,$$

where $\rho(t) = r(t) - \mathsf{d}(t)$ is the Malthusian population growth rate. This indeed would be familiar equation, but not a very useful one, in the present situation.

We see two reasons for saying this. The first is that the rate $\rho(t)$ is much less interesting object than the infection rate $r(t)$. It is the infection rate which describes how aggressive the virus is and how favourable the social conditions are for its spread. The second is that we can not observe the size of population $X_1(t)$. What we have are the observations of the exit process $X_{1a}(t)$: the accurate count of daily new cases and therefore cumulative number of cases have been an object of acute interests and was counted with all possible accuracy.

Therefore, what we found ourselves in was the following situation: given observations on the exit process $X_{1a}(t)$ can we say something about the rate of increase $r(t)$ of the unobservable arrival process $X_0(t)$? The answer is "yes" if we have correct mathematical connection between the two processes.

For this we retain the equation for H in (3), or

$$dH(t) = m(t)dt + r(t)[H(t) - Q_{1a}(t)]dt, \tag{4}$$

and then note that if T_i-s are independent and identically distributed with distribution G, then Q_{1a} is convolution of H and G:

$$Q_{1a}(t) = \int_0^t H(t-s)dG(s) = \int_0^t G(t-s)dH(s) \tag{5}$$

so that

$$dQ_{1a}(t) = \int_0^t g(t-s)dH(s).$$

The assumption that the incubation times are identically distributed is not the assumption that the population X_1 is stationary. The rate $r(t)$ still can be variable, within this model, and was variable in reality. Substitution of (5) into (4) gives us the eventual form of equation for H:

$$dH(t) = m(t)dt + r(t)\int_0^t [1 - G(t-s)]dH(s)dt \tag{6}$$

Equation (6) is the Volterra integral equation of the second kind. Very clear introduction to its theory can be found, e.g., in the classical monograph Tricomi [1985], Ch I. It is well known that the solution for H, or rather for its derivative $h(t) = dH(t)/dt$, can be explicitly represented through the resolvent of this

equation. However, if we think about $dH(t)$ as a "small increment forward", from t to $t+dt$, then the recurrent nature of (6) becomes very clear: to calculate $dH(t)$ one needs only past $dH(s)$ with $s < t$. Therefore, solution for h becomes, numerically, simple and convenient procedure, which we followed.

Expected value of $X_{2a}(t)$. Again, from the definition of $\nu(A_{2a}(t))$ it follows that

$$Q_{2a}(t) = E\nu(A_{2a}(t)) = \int_0^t F(t-s)dQ_{1a}(s).$$

To express it in terms of H, introduce the convolution of distributions F and G:

$$F \star G(t) = \int_0^t F(t-s)g(s)ds.$$

Then for both Q_{2a} and Q_2 we have

$$Q_{2a}(t) = \int_0^t F \star G(t-s)dH(s)$$
$$Q_2(t) = \int_0^t [F(t-s) - F \star G(t-s)]dH(s)$$

In differential form the expression for Q_{2a} was convenient for numerical calculations:

$$dQ_{2a}(t) = \int_0^t f \star g(t-y)dH(y)\,dt. \tag{7}$$

The choice of distributions G and F. The presence of convolution here has some bearings on the practical question of how to choose densities f and g in our model. The data on duration of incubation period, at least at present, is scant. It would be highly speculative to try and estimate g from this data. The situation with data on duration of illness may be better, but reliable estimation of density in purely non-parametric fashion requires many data, cf., e.g., Silverman [1986]. We can, however, significantly reduce the difficulty by choosing both g and f from a family of distributions, which is closed under convolution. These are families of distributions $\{L_\alpha, \alpha \geqslant 0\}$, indexed by a scalar parameters α, and such that convolution of L_α and L_β is the distribution $L_{\alpha+\beta}$.

Now imagine that we choose as distribution of incubation periods, G, some L_α, and choose as distribution of the period of illness, F, some L_β. If the family $\{L_\alpha, \alpha \geqslant 0\}$, is sufficiently rich family, then to make such choice will be possible and will require only estimation of two numbers, α and β, from the data, which statistically is much more simple and stable process. And at the same time, we will know the distribution $L_{\alpha+\beta}$ straightaway.

One such family, we believe, is available for us. It is the family of Gamma distributions with the same scale parameter λ:

$$l_\alpha(x) = \frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}$$

is the density of L_α.

3. Statistical discussions. What was found.

In this section we consider the differences between relatively idealised model of Section 2 and reality of data. Let us do this more or less along the time-line of the process and start with immigration.

Immigration. The data on when the people, who later proved to be infected, arrived in the country is fully available. There have been 570 such persons. Until when? later arrivals? We did not try to smooth this data in any way and used it as it is. Therefore the derivative of H, denoted h, as a solution of equation (6) looks "smoother" than m, but can not be very smooth, see Figure 6. There was, however, a question whether some infections have been secondary, that is, acquired after the arrival. This could affect the estimation of $r(t)$. To safeguard from this, one can use the density g of detection times and cut off the manifestation times, which happen to be longer than 90%-quantile point of our g. This reduced the number 570 to 529 of those who can be prudently assumed to acquire infection overseas. Figure 1 shows the date of arrival (on x-axis) and the date infection was confirmed (on y-axis). The green diagonal line shows the 90% cut-off line.

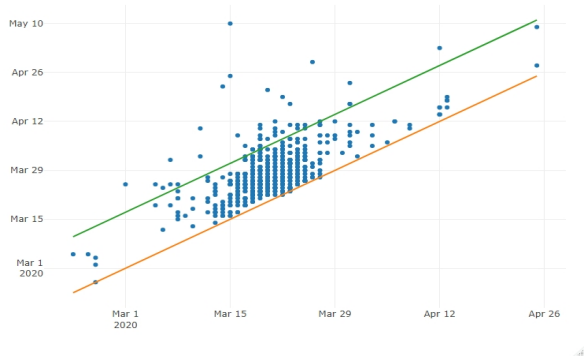

FIG 1. *Arrivals of persons who were later found infected. The date of arrival is on x-axis.*

Distribution of incubation time T_i. We introduced these as times from the moment of infection to the moment of becoming symptomatic, i.e. ill. These

are important moments in the management of epidemic because from this moment on, up to recovery, the person had to be isolated. In practice, however, infection was detected often through contact tracing and testing process, without being necessarily symptomatic. All such persons also had to be isolated from population. Therefore, the estimated distribution of T_i-s we denoted G *is not* the distribution of incubation time, but the distribution of the detection time. It would be very interesting from medical point of view to evaluate the distribution of the incubation times, but the distribution of detection times is also very interesting object: it gives numerical description of effectiveness of the policy of tracing of the infection. It shows that the policy was effective. The authorities would not spare resources in testing, large number of people have been tested and, relative to this number, not so many detected. Thus one can expect that this G will be close to exponential distribution, and Figure 2 shows its graph. The distribution of incubation times could not have been exponential – the process of incubation should have been happening and symptoms could not have been appearing spontaneously.

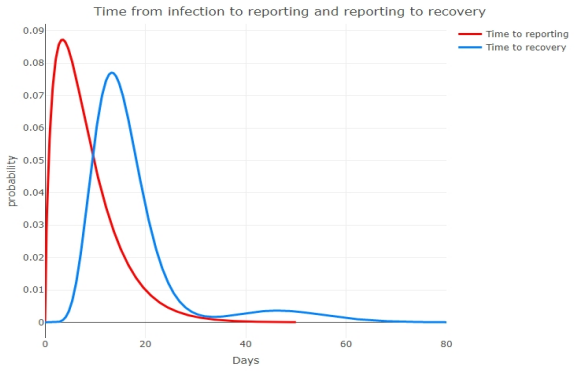

FIG 2. *More to the left is the graph of density g of detection times (red line). It is the Γ-density with shape parameter $\alpha = 1.72$ and scale parameter $\lambda = 0.21$. More to the right (blue line) is the graph of density f of recovery times. As we say in "Distribution of recovery time Z_i", it is a mixture of two distributions with weights $p = 0.92$, for those who followed "normal" recovery process, and $1-p = 0.08$ for those who met with complications. Both admixtures are Γ-densities with scale parameter $\lambda = 0.6$ and shape parameters $\alpha = 9$ and 28.8, respectively.*

The rate of infection. This is important characteristic of the process. Our $r(t)$ can change from day to day. It describes not an individual, but the social conditions of life. Intuitively speaking, it is the expected number of people an infected person will infect in the day t. So, it measures how aggressive the illness is and how much do people mix with each other. While what the virus is, it is – we do not presume the strain has changed *en masse* during the several months of epidemic, the social condition have been changing. As we said, complete lockdown was declared in New Zealand on 24 March, 2020. In the mind of

everybody, this would mean that $r(t)$ abruptly changed from some noticeable level to nearly zero.

The change indeed took place, but it was surprising to see that the process was not very rapid even in law-obedient and morally prepared population. It is not strange to see that the rate started to decline few days earlier than the declaration of lockdown: people started keeping some distance from each other. However, it was surprising to see that, numerically, sharper change in $r(t)$ can not be achieved even with some flexibility in the choice of g.

It was not possible to achieve the level of agreement with the data, which is demonstrated in Figures 7 and 8, as soon as we tried to change $r(t)$ more sharply.

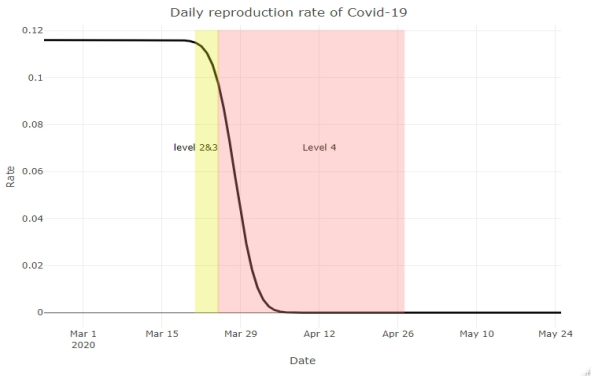

FIG 3. *Explicit mathematical form of $r(t)$, in our view, is not so important, given that it stays on some relatively high level up to declaration of the lockdown, and then decreases relatively quickly to essentially smaller values – as quickly as the data allows us to assume. In our case $r(t) = r_1[1 - \Phi_{\mu,\sigma^2}(t)] + r_0$, where $\Phi_{\mu,\sigma^2}(t)$ denotes the normal distribution function with expected value and variance μ and σ^2, respectively. We have $r_1 = 0.116, r_0 = 0$, while $\mu = 35$, which is close to the number of days from the ban on international flights and the day of lockdown, and $\sigma^2 = 4$.*

We have to conclude, therefore, that some social inertia exists, in spite of best intentions of everybody, and too rapid changes is very difficult to achieve, at least without resorting to the methods on martial law. In bigger populations this phenomenon will be, presumably, pronounced more sharply. In New Zealand, however, without any martial law, and in less than two weeks time, the rate $r(t) = 0$ was achieved. The fitted graph is shown in Figure 3

In conclusion, it may be good to notice, for not necessarily statistically minded reader, that there is difference between interpretation of $r(t)$ and what can be infectiousness of an individual. How infectious an individual can be, certainly, depends on the individual, on the viral load, on the stage of infection, see, e.g., To *et al.* [2020], on the social activity of the person, the nature of the job, and so on. However, in population, the notion "randomly selected individual"

smoothes over individual variations and makes analysis simple, useful, and yet statistically correct.

Distribution of recovery time Z_i. The smooth black line in Figure 4 is asymmetric. It is also shown as the most right line in Figure 6. What should be responsible for this asymmetry is the density f of recovery times.

Formally, there are many ways to model asymmetric distributions. However, it would be much more informative to find the one which agrees with medical intuition. Neither of the authors are medical doctors, but we had an advantage of discussions with medical practitioners. The model we used was the mixture $pf_1(z) + (1-p)f_2(z)$. Here the density $f_1(z)$ describes recovery times for people who in this process did not experience complication. This was the majority of cases, $p = 0.92$. The density $f_2(z)$, which also belongs to Γ-family, describes recovery times of the cases with complications. Existence of the cases with complications does not constitute a discovery, but it was not very clear whether they can "blend" with the general population or not, whether there is some sort of dichotomy – either cases with complications or cases without. As soon as one assumes the existence of the mixture, the data then dictates that the mode of $f_2(z)$ is about 45 days, more than twice more than the mode of distribution of recoveries without complications.

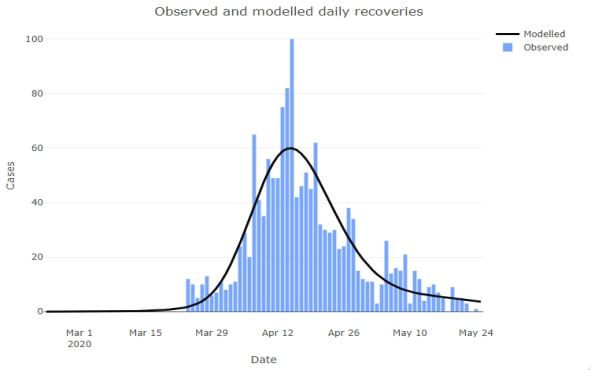

FIG 4. *The graph shows the daily recovery rate and its model. In our notations it is $dX_{2a}(t)$ and $dQ_{2a}(t)/dt$, see equation (7). Very noticeable asymmetry, which we comment on in "Distribution of recovery time". We show this curve also in comparison with $h(t)$ and $dQ_{1a}(t)/dt$ on Figure 6*

Somewhat better judgement of fit can be made from the graphs of cumulative quantities $X_{da}(t)$ and $Q_{2a}(t)$ in Figure 5

Fitting the model to the data.

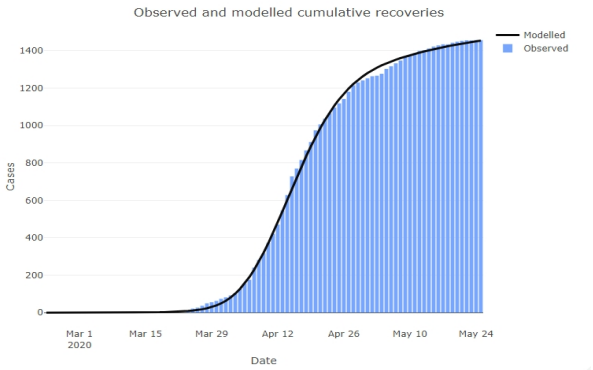

FIG 5. *The graph shows cumulative number of recovered people and its model. In our notations these are* $X_{2a}(t)$ *and* $Q_{2a}(t)$.

The main source of information we had are given in Figures 7 and Figure 6 and their cumulative versions in Figures 8. and 5. We also should add the immigration data, shown in Figure 1. From the shapes of these empirical curves it may not be very visible what are the forms of the main curves of the model, which are $H(t), Q_{1a}(t)$ and $Q_{2a}(t)$. That is because the model did not attempt to "directly" approximate empirical data, but rather try to detect what may be revealed from them.

On Figure 6 we show the graphs of the "speed" $h(t) = dH(t)/dt$ of arrivals of new infections, unobservable but calculable as the solution of (4), and then the speed of exit $dQ_{1a}(t)/dt$ and $dQ_{2a}(t)/dt$ as solutions of (5) and (7), respectively. the graph of $h(t)$ is irregular in shape because of irregularity of the shape in daily immigration $m(t)$. The next convolution smoothes the graph of $dQ_{1a}(t)/dt$ very much.

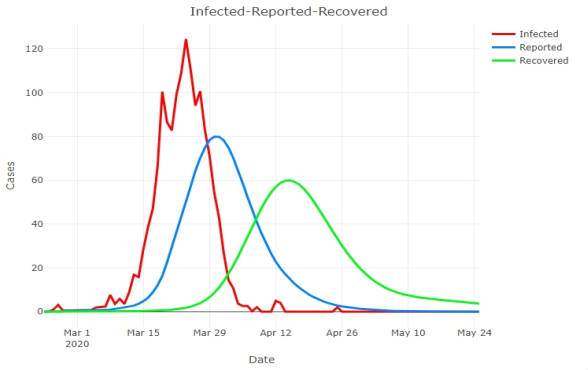

FIG 6. *The graph on the left, in red, shows the graph of* $h(t) = dH(t)/dt$ *which is solution of equation* (4)*, or equivalently, equation* (6)*. It is irregular in shape because of irregularity of the shape in daily immigration* $m(t)$*. The unimodal graph in the middle, in blue, shows the derivative of* $Q_{1a}(t)$*, see equation* (5) *and the display formula immediately after it. The graph on the right, in green, is the derivative of* Q_{2a}*, see* (7)*. It is smooth but asymmetric – as discussed in sub-section "Distribution of recovery time* Z_i*".*

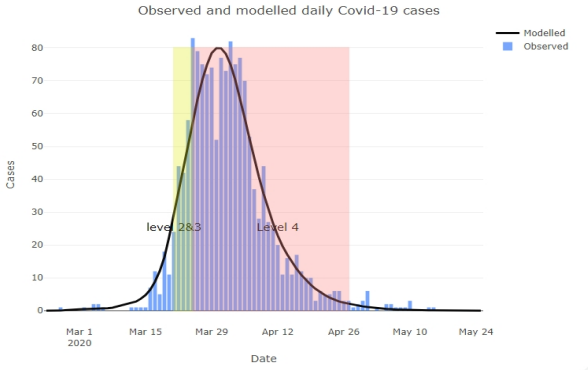

FIG 7. *The graph shows daily reported cases and its expected value, from the model. In our notations these are daily increments of* X_{1a} *and* Q_{1a}*, respectively. Overlaid on the graph are periods of New Zealand levels 2 and 3, and then the strongest level 4 (the "lockdown") of social restrictions.*

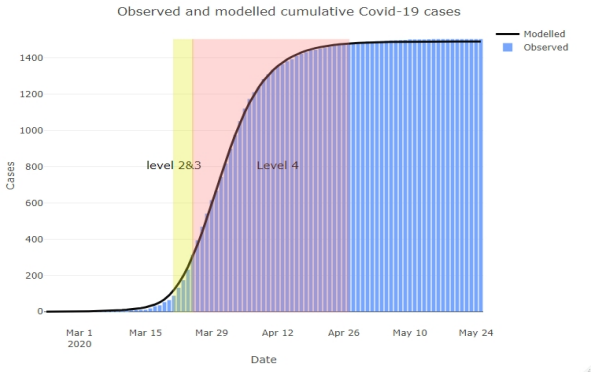

FIG 8. *The graph shows cumulative number of reported cases and its expected value, from the model. In our notations these are X_{1a} and Q_{1a}, respectively. Again, overlaid on the graph we show periods of New Zealand levels 2 and 3, and then the strongest level 4, of social restrictions. From 13 May, 2020, the restrictions have been relaxed, relatively quickly, back to level 2*

References

Ottar N. Bjornstad (2018), *Epidemics: Models and data using R*, Springer

Pierre Brémaud (1981), *Point processes and queues. Martingale Dynamics*, Springer-Verlag, 1981

P.E. Greenwood, L.F. Gordillo, Stochastic epidemic modelling, *Mathematical and Statistical Estimation Approaches in Epidemiology*, 2009, 31-52.

Estate Khmaladze (2013), *Statistical Methods with Applications to Demography and Life Insurance*, CRC Press, Chapman and Hall Book

NZ Ministry of Health, Confirmed and probable cases of Covid-19 in New Zealand;
https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus/covid-19-current-situation/covid-19-current-cases

NZ Ministry of Health, Media release, 5 April, 2020, Contact tracing at forefront of fight against COVID-19;
https://www.health.govt.nz/news-media/media-releases/contact-tracing-forefront-fight-against-covid-19

J. H. Pollard (1973), *Mathematical Models for the Growth of Human Population,* Cambridge University Press

B. W. Silverman (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall/CRC Press

Kelvin Kai-Wang To, *et al.*, Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study, *The Lancet, Infectious Diseases*, 2020, 20, pp. 566 - 574

F. G. Tricomi (1985), *Integral equations,* Dover Publications, Inc., unaltered and unabridged republication of the 1957 edition from Interscience Publishers, Inc.