Review of: "A Mathematical Characterisation of COVID-19 in Mauritius"

Matthew Merski[1]

1 University of Warsaw

Potential competing interests: No potential competing interests to declare.

In "A Mathematical Characterization of COVID-19 in Mauritius" by S. Z. S. Hassen, the author performs retrospective analysis of several waves of Sars-CoV-2 infections in Mauritius, estimating a number of descriptive statistical terms for the course of the pandemic in Mauritius. The author attempts to use the analysis to inform future epidemics. Generally, this is a good attempt at analysis, but many of the estimates and assumptions utilized in this analysis lack support or are phrased inelegantly. Many numerical values are cited without reference to their origin and statistical justification for choice of model and goodness of fit for reported results is missing.

Intro:

1) One of the stated goals of the manuscript is to "*accurately estimate … the reproduction number*" in order to better deal with future epidemics. While this is a laudable goal, for many viruses and specifically Sars-CoV-2, the number of new infections arising from any single infected individual is highly variable and is better described by non-Gaussian distributions, such as a skewed or ??? distribution which suggests that reducing the number of, in the common parlance, "superspreader events", would be generally more effective and more focused than targeting all cases generally (with the caveat that it isn't simple to identify which individuals are most likely to be the most infectious). Furthermore, there is some natural variation in estimates of the reproduction number [Journal of Travel Medicine, 2020, 1–4]. The author also is a bit lax with their definition of "reproduction number" stating that "*This figure is known as the basic reproduction number and can be accurately estimated by using data available in the first days….. The reproduction number is a function of time and will decrease…*" This is confusing. The author clarifies the differences later by introducing R and R_0 but this should be done in the intro as well. The author suggests that the basic reproductive number can be calculated from the early portion of the infection wave. However, this is generally untrue as 1) it is difficult to estimate this number because of imperfections in identifying the number of cases (*i.e.* missed cases) especially for Sars-CoV-2 because many, perhaps even half of all cases are asymptomatic [International Journal of Infectious Diseases 94, 154–155] and 2) more cases will be missed when the health care system is strained, especially during periods of peak infection (*i.e.* disease waves) [Bulletin of Mathematical Biology 76, 245–260]. However, there are clear infectivity differences between different diseases, such as highly contagious measles and the less contagious HIV. The author should spend some time to clarify these basic premises in the introduction.

2) Estimates of case fatality rate and case mortality rate are also useful to inform public health decisions. The author

compares these between Mauritius, India, the Fiji Island, the US, & UK. However, CFR & CMR are strongly influenced by many factors including climate and population density as well as those that are political, social, or economic and comparisons of different countries MUST take these factors into account as well, which makes it exceedingly difficult to compare two countries in a statistically unbiased manner. As such Table 2 doesn't add much to the manuscript and due to its excessive size for the information provided, this table should be removed. There are other points in the manuscript where such comparisons are made and need to also be addressed before the conclusions paragraph.

Modeling of the first COVID-19 wave:

1) The statement "*the nonlinear least square solver computed*" is vague. While its good to include the parameters used, the solver and other programs and methods used must be explicitly identified.

2) There are an entire family of logistic curves based on the work of Richards [J Theor Biol (2010) 267:417-25]. The author however seems to simply choose a single model without any thought to selection of this specific model. And while use of a logistic model for this data is not unreasonable, the choice of the specific model used must be addressed and justified with some kind of suitable statistical support (error minimization, predictive power, Akaike criterion, as available).

3) The author reports the average growth rate in figure 1 to be obtained from the derivative of the model curve at a given time point. Essentially the author is taking a tangent to the model curve, but this is an instantaneous rate, not correctly an "average" rate and the nomenclature should be changed, most easily done by removing the word "average" in the figure and associated text.

Relationship between ceiling value and growth rate:

1) The author makes the following statement: "*The measures taken by authorities around the world did not however separately affect the three parameters* [M,α t_0] *described above and instead their action(s) distorted the bell-shaped curve into an asymmetric distribution. The level of distortion observed was a function of the extent of actions taken by authorities. For example, a strict lock down reduced α considerably more than social distancing measures such social gathering restrictions would. The use of masks and travel restrictions also affected α and t0. It is difficult to quantify the individual effect of each of these measures on the curve, as they were usually imposed in parallel.*" This statement appears to be self-contradictory, both that governmental decisions did and did not affect the parameters. This needs to be clarified and examples with references need to be included to support the edited statement.

2) The author states: "*Furthermore, the slope α is in fact an average of different α's at different moment in time as a consequence of continuously changing enforcement by the authorities.*" This statement is also confusing. All measurements are in some extent or another an average of multiple similar terms, due to stochastic factors and inhomogeneities in the sample so its not clear what the author intends to mean here.

3) "*We gathered data for the first wave*" should be more correctly "We considered data for the first wave"

4) "*It was noted that quick effective intervention by the authorities generally resulted in symmetric curve with a clear peak whereas delayed intervention tended to distort the curve and result in a prolonged trailing end of the bell-shaped curve.*"

This statement requires supporting references

5)"*In some cases, countries developed a plateau rather than a peak which meant that*

they had a really hard time managing the pandemic with large number of cases over a long duration of time". This statement requires supporting references.

6) " *The United States in particular suffered from this, primarily due to the fact that large segments*

of the population did not abide by instructions for social distancing until it was too late". This statement requires supporting references.

7) The author normalizes the value M for each country, however what this was normalized to (total population? total cases? population density?) is not mentioned.

8) Figure 2 is a theoretical model, however this fact is not terribly obvious. It behooves the author to produce any experimental support for the stated hypothesis that: "*A surprising correlation is found between the final number of infections and the slope of the curve.*" Analysis of the waves from different countries that the author mentions in this manuscript could provide this support but it does not, in fact, seem surprising that a different slope would be found for an infection wave that will total 100,000 cases as compared to one that will total 500 cases.

Sections 4:

1) The author claims that an R_0 greater than one implies exponential growth. This is incorrect as many diseases have less than exponential growth rates [Proc Natl Acad Sci U S A. (1989) 86: 4793–4797)]. Nor should a value less than one imply exponential decline as the author also mentions many instances of long tailed infection waves in this manuscript.

2) The author's implication that there is a proportion P of the population which represent the amount of immune individuals which must be present to stop exponential growth is also flawed as SARS-CoV-2 and many other coronaviridae fail to produce long lasting immunity after infection, which makes a traditional SIR analysis less applicable [Lancet, 401, P833-842 (2023)] and the text should be modified to address this.

3) The author's assertion that the log-linear number of cases early in the first wave of the the pandemic supports exponential growth despite the caveat that there are statistical errors due to the small total number of daily cases is excessively confident in flawed data and should be removed.

4) The author uses European data to estimate R_0 for Mauritius, despite previously having stated the importance of the social habits of the population. This number also varies significantly across the EU (which included the UK at the time) so the author needs to justify why this is an acceptable estimate despite these caveats.

Modeling of the second COVID-19 wave:

1) The author argues that a logistic curve can only account for a natural spread of the virus. This is incorrect as a standard

SIR model will give a logistic curve, it only requires that an isolated/unchanging population with three sub populations: susceptible, infected, and removed. The basic SIR model does not adequately handle population exchange between the isolated and another population, which is perhaps what the author is aiming for, but a logistic growth curve is not necessarily unable to handle a situation like this, especially since many can have different rates for the upswing and downswing portions of the wave [PLoS ONE 12: e0178691].

2) It is unclear what more "contagious variant" the author is referencing, possibly the Delta variant, which was first detected in India but its origin has not been as of yet proven. This needs to be clarified and proper citation support given.

3) The author assumed that 80% of the second wave cases had occurred by July 2022 and based their modeling on that fact. More than 6 months have elapsed since then. Was that assumption true? Did the actual number of cases fall within the expected uncertainty of the model? As the author states that the intention of this manuscript is to inform public health decisions using these models, the author should extrapolate the model forward in time (after July 2022) and compare that with the actual number of cases and comment on what the model gets right and where improvements could be made.

Sections 6:

1) Remove the inelegant phrase "creating a psychosis around the virus".

2) The author mentions a number of anecdotes that suggest the number of COVID-19 deaths was not properly recorded in several countries. These need citations from independent, permanently identifiable sources or should be removed. Likewise, the CMR numbers reported require citations.

Conclusion:

1) The author offers minimal statistical analysis comparing their model to the actual data so the conclusion that they "*successfully fitted mathematical models to the two COVID-19 waves*" is unjustified.