COPY RIGHT

Paper Authors

Karthik Boppana

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

Mathematical Modelling of Epidemic Data:Case study of COVID-19

Karthik Boppana

Student, Department of Civil Engineering, PSG College of Technology, Coimbatore, India

Abstract

The 2019-Novel-Coronavirus (COVID-19) has affected more than 180 countries and 93,77,126 cases confirmed (As of June 24 2020). Understanding the transmission dynamics of the infection in each country which are affected on a daily basis and evaluating the effectiveness of control policies is critical for our further actions. To date, the statistics of COVID-19 reported cases show more than 80 percent of infected had a mild case of disease, while around 14 percent of infected experienced a severe one and about 5 percent are categorized as critical disease victims. To help estimate uncertainties in observational data and uncertainties in calculation based on observational data; To characterize numerical output from mathematical models and help understand the model's behaviour and to assess the model to stimulate important features of the natural system (model validation) is an important task.A mathematical model for creating a proper emphasis on when actually the COVID cases will tend to decline based on the number of cases everyday is very crucial; however as with all mathematical models, the predictive ability of the model is limited by the accuracy of the available data and to the so-called level of abstraction used for modelling the problem.

Article Info: **Pandemic, Coronavirus,Bailey'sModel, Regression Model**

Introduction

Since the outbreak of the Coronavirus COVID-19 in January 2020, the virus has affected many countries and taken the lives of several thousands of people worldwide. By March 2020, the World Health Organization (WHO) declared the situation a pandemic, the first of its kind in our generation. To date, many countries and regions have been locked-down and applied strict social distancing measures to stop the virus propagation. From a strategic and healthcare management perspective, the propagation pattern of the disease and the prediction of its spread over time is of great importance, to save lives and to minimize the social and economic consequences of the disease. Within the scientific community, the problem of interest has been studied in various communities including mathematical epidemiology, biological systems modelling,signal processing and control engineering.In this study, epidemic outbreaks are studied froman interdisciplinary perspective, by using an extension of the susceptible-exposed-infected-recovered (SEIR) model, which is a mathematical compartmental model based on the average behaviour of a population under study. The objectiveis to

provide a better understanding of the significance of mathematical modelling for epidemic diseases. It is shown by simulation, how social measures such as distancing, regional lockdowns and public health vigilance, can influence the model parameters, which in turns change the mortality

rates and active contaminated cases over time.

It should be highlighted that mathematical models applied to real-world systems (social, biological, economical, etc.) are only valid under their assumptions and hypothesis. Therefore, the research and similar ones that address epidemic patterns, do not convey direct clinical information and dangers for the public, but should rather be used by healthcare strategistsfor better planning and decision making.

In Section II, a brief introduction to mathematical modelling of biological systems is presented, to highlight the scope of the present study and to open perspectives for the interested researchers, who may be less familiar with the context. The

proposed model for the outspread of the Coronavirus is presented in Section III. The article is concluded with some general remarks and future perspectives.

II. AN INTRODUCTION TO MATHEMATICAL EPIDEMIOLOGY AND COMPARTMENTAL MODELLING

A. Mathematical modelling

A model is an entity that resembles a system or object incertain aspects, but is easier to work with as compared to theoriginal system. Models are used for the following

1) Identification and better understanding of systems,

2) Simulation of a system's behaviour,

3) Prediction of its future behaviour, and ultimately

4)   System   control.

While various types of models are used for biological systems, we are commonly interested in mathematical models as they permit the prediction and possible control of biologicalsystems. In choosing among different available models, the widely accepted principle is the model parsimony, which simply means that "a model should be as simple as possible andas complex as necessary!". The model parsimony, is also an important factor for estimating the unknown model parameters using real data. A more accurate model with fewer number of parameters is evidently preferred over a less accurate andmore complex model. But how should one select between a more accurate complex model and a less accurate simpler one? Measures such as the Akaike information criterion (AIC),the Bayesian information criterion (BIC) and the minimum description length (MDL), address the balance between the number of observations and the model unknown parameters to select between competing models with variable number ofparameters and different levels of accuracy. Finally, the physical interpretability of the model parameters and the ability to estimate the parameters such that the model matchesreal-world data, is what makes the whole modelling framework meaningful.

International Journal for Innovative
Engineering and Management Research
A Peer Reviewed Open Access International Journal

www.ijiemr.org

B. From stochastic infection propagation models to ordinary differential (difference) equation modelling

The outbreak of a contagious disease in a large population is a stochastic event. Starting from a single infected individual, the infection is transmitted to others in a stochastic manner, either by direct contact, proximity, or environmental traces (infected objects left over in the environment). The new infected generation in turns transmits the infection (again probabilistically) to the healthy individuals that they meet or encounter. During the primary stages of an epidemic outbreak, healthy infected individual encounters are statistically independent. As a result, the chance of multiple infected people meeting a single healthy individual is probabilistically low.

Therefore, assuming that each infected individual contaminates R0 new people on average (known as the reproduction number), if R0 > 1 the disease spreads exponentially from one time step to another (for example on a daily basis). However, in a finite population, the exponential growth can not continue for ever. Depending on the population size and contact patterns, the probability of infected people encountering independent healthy individuals decreases. Therefore, after the initial outbreak that exponentially spreads among the population, the infected population tend to encounter each other and repeated healthy ones (the healthy individuals already contacted by another infected person).

Hence, the stochastic model of infection propagation, somehow saturates. The probabilistic models used for modelling such epidemic spread are commonly based on the branching process and a Poisson distribution for the probability of contact between infectious and healthy subjects which help us understand more about the disease dynamics and help us understand more about the spread of diseases. This in turn helps to analyse the situation more wisely.

C. Mathematical epidemiology

In order to model the propagation of epidemic diseases in a population, certain disease- and population-specific assumptions are required. The most common assumptions in this context include:

• The diseases are contagious and transfer via contact.

• A disease may or may not be fatal.

• There may be births during the period of study, and the birth may (or may not) be congenitally transferred from the mother to the baby.

• The disease can have an exposure period, during which the contaminants carry and spread the disease, but do not have visible symptoms.

• Catching the disease may or may not result in short-term or long-term immunity. Depending on the case, the recovered patients can again become susceptible to the disease.

• Interventions such as medication, vaccination, lockdown, quarantine and

social distancing can change the pattern of propagation.

III. PROPOSED EPIDEMIC MODEL

All of the countries affected by the Covid-19 epidemic have the same primary goal: curb the spread of Sars-CoV-2 and thus prevent new diseases. While certain countries have achieved significant successes through consistent social distancing and restricted public life, others are still in a phase of unchecked expansion.An unusual perspective on the number of cases is offered by comparing the daily new infections with the number of all confirmed cases: Here, the slowdown in growth becomes more apparent than when viewed on the timeline - where there is a downward trend, it is visible as a steeply decreasing curve. The epidemics tend to grow exponentially at first, and also that exponential growth is really hard to understand. But regarding the news - we rather know where we're heading, and if we're making detectable progress. Are we winning or losing?

Because of course, we can't have exponential growth forever - at some point the disease will run out of new people to infect, either because most people have already been infected, or because we as a society managed to get it under control. But the scary part is to drawn predictions when you're in the middle of an exponential, it's essentially not possible to tell when it's going to end. Are we in for 10 times as many cases as we currently have? Or 100 times as many? Or 1000? Exactly when exponential growth ends is important, because it hugely determines how many people fall ill, yet so little reporting

actually focuses at all on how to tell if exponential growth is ending. Graphically visualizing the COVID-19 epidemic on a global scale is very important.

Trajectory of World COVID-19 Confirmed Cases (2020-06-25)
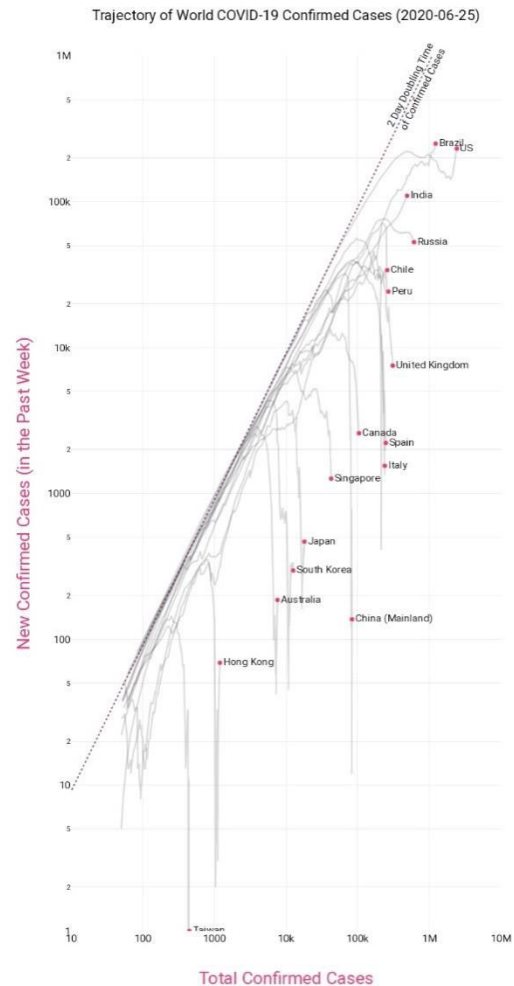
Credits – aatishb.com

The graph shows all countries travelling along the trajectory of exponential growth,and it makes it obvious which ones have managed to stop the exponential spread ofdisease - they plummet downwards off the main sequence in a way that find super compelling.And thefigure also makes it abundantly clear that, even if a country doesn't have lots of cases right now, covid-19 is probably going to follow this same trajectory there and end up

spreading - until that country hits emergency.

So how did we make this graph? Well, there are three key ideas:

1.)The first is to plot on a logarithmic scale, since that's the natural scale for exponential growth - note that the scale grows by multiples of 10, so 10, 100, 1000, rather than 10, 20, 30.This scales up small numbers and scales down large numbers, making the growth equally apparent
on all scales, and lets us compare the growth in countries with very different numbers of cases.

Which brings us to the second idea: 2.) Understanding changes early, by looking at change itself.

For example, if we look at the growth of cases in South Korea, you can see that at first they're exponential, and later, the growth slows down. But when we're halfway up this curve, it's hard to tell by eye that it's slowing down - it still looks exponential. If instead you chart the number of new cases in the last week, in other words, the rate of growth, it's much easier to see that the growth is starting to slow down. When the number of new cases each week flattens out or goes down, we've escaped the exponential growth zone.

The third idea behind the graph is from physics: 3.) Not to plot against time. Usually, when we see exponential growth, the number of cases is plotted versus time. But the spread of the virus doesn't have to

limit itself to time(if it's March or April); it only requires about two things:

- how many cases there are right now

- how many new cases there will be - that is, the growth rate.

The defining feature of exponential growth is that the number of new cases is proportional to the number of existing cases, which means that if we plot new cases vs total cases, exponential growth appears as a straight line. So the number of new cases (the growth rate) is on the y axis, and the cumulative number of cases is on the x axis, both on logarithmic scales.

This gives us a graph that shows where all countries are in their COVID-19 scenarios, it makes it obvious that the disease is spreading in the same manner everywhere we're all headed on the same trajectory, just shifted in time; and it makes it obvious where public health measures like testing, isolation, social distancing, and contact tracing have started to control the disease, and where they either aren't working or haven't hadtimeto show up in the data.

The graph gives us some sense of what's actually happening in these uncertain times.

That said, this graph also has a number of caveats & limitations -One important aspect of this logarithmic graph is that [Logarithmic scales distort] 10,000 looks really close to 1,000 on a log scale; this kind of distortion might allow people to take COVID-19 less seriously. But however, this can be curtailed by graph to

the proper scale. Also, the log scale on the x axis makes it harder to see a resurgence of new infections after a significant downturn. In such a case a normal plot compared with time is better for that.

Another important caveat is that this graph (& basically every other COVID-19 graph) is not actually showing the true number of cases, just the number of detected cases. The true number of cases is unknown but certainly much higher than the number detected
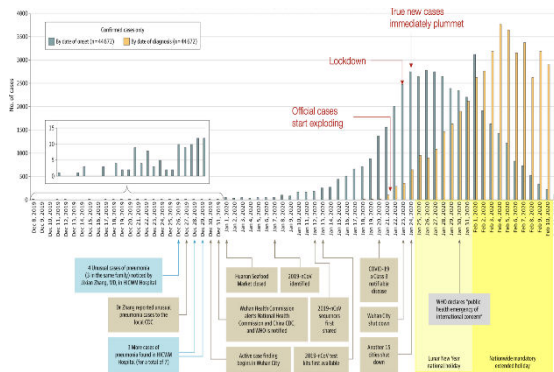
Credits – Thomas Pueyo at medium.com

True Growth Rate vs Tested Growth Rate

In reality, COVID-19 cases spread at a slower
rate than what the data implies. It's a subtle idea, but the data reflect not just an increase in cases, but also an increase in the number of tests performed. The data we're using is incomplete, as it relies on daily reports from healthcare systems around the world. Also, different countries have dramatic differences in the resources that are available or dedicated for testing.

Finally, the trends in the plot are delayed a few days, since plotting the average growth

rate over last week (there's too much variability in the data to plot daily growth rates). This is actually kind of a good thing because it means that it's a pessimistic graph, it doesn't get too excited too soon, and so a downward trend on the graph is much more likely to be a real downward trend.And a real downward trend is what isneeded.

A lot of the daily news just reports recent data points. Yet to understand where we're headed, it's not enough to know just where we are today - we need to be talking about the trends: how many new cases there are today relative to the number of new cases yesterday, or last week. Charting the rate of change empowers us all to more clearly see what the future holds.

The next caveat is that unlike most other COVID graphs, time is not plotted on the x axis, which might be confusing. Instead, time is shown through an animation or a computer-generated stimulation model.

The computer-generated stimulation model who look something like the following images which show the updated cases based on every day count.
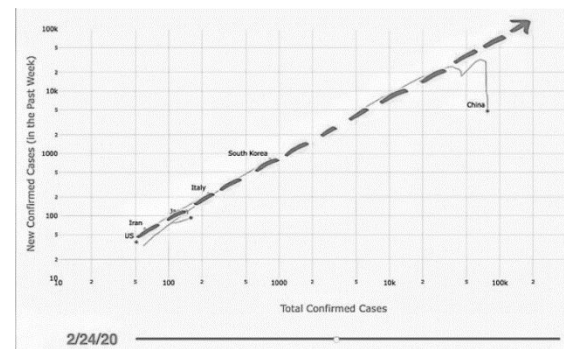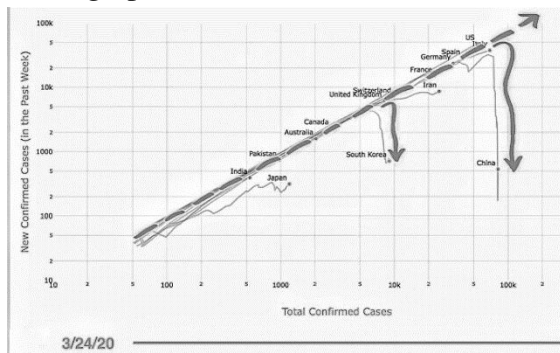

Credits – aatishb.com

The above graph is the snap of the stimulation graph which is generated and the graph corresponds to the day 24th February 2020. The next graphis the snap taken from the time generated model corresponding to the day 19th march 2020.
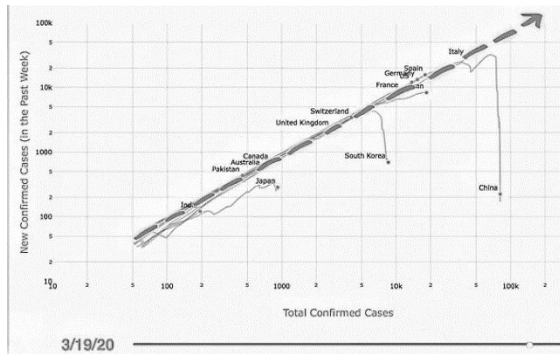
Credits – aatishb.com

By 24th march 2020 countries like China and South Korea have escaped the exponential growth rate which is indicated in the graph.

Credits – aatishb.com

The above graphs are taken from the stimulation model which has been generated by taking the log scale instead of taking the conventional linear scale which makes it more reliable.

Analysing the data using the model

All of the countries affected by the Covid-19 epidemic have the same primary goal: curb the spread of Sars-CoV-2 and thus prevent new diseases. While certain countries have achieved significant successes through consistent social distancing and restricted public life, others are still in a phase of unchecked expansion. An unusual perspective on the number of cases is offered by comparing the daily new infections with the number of all confirmed cases: Here, the slowdown in growth becomes more apparent than when viewed on the timeline - where there is a downward trend, it is visible as a steeply decreasing curve. Of course, the graphic always shows only a snapshot; Only the future will tell whether success will last.

Group 1: On the road to success

Number of new Covid-19 infections each day compared to the number of all cases; Countries with over 5000 cases, logarithmic scales.
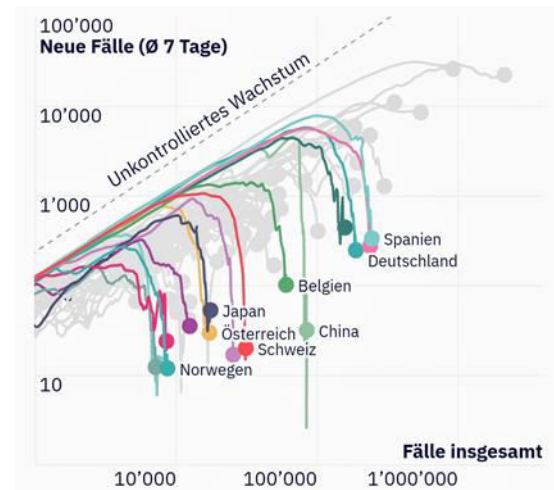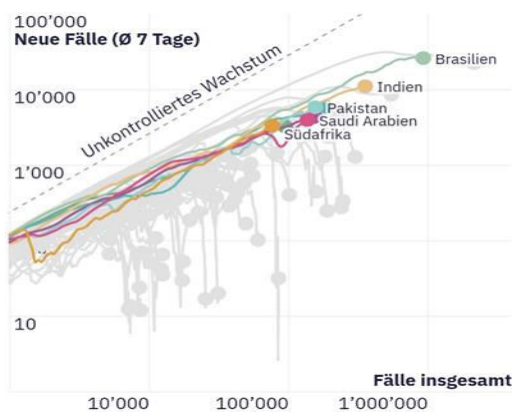

Credits – TagesAnzeiger.ch

The graph shows when countries have left the path of exponential growth in the confirmed Covid-19 case numbers. The trick: New cases are compared to the accumulated on a double logarithmic scale. Nations in which the development runs along the dashed, diagonal slope are traveling at an exponential pace. The

comparison of the development of the number of cases can be shown most clearly if one shows them on a logarithmic axis. It is not the numerical value of a quantity that is plotted on such a scale, but the logarithm of the same. The Y axis - and in this special case also the X axis - is compressed and enables larger value ranges to be displayed clearly.All countries with over 5000 confirmed cases are displayed, the members of the group are highlighted in colour.The "Success Course" group includes those countries that have managed to grow from exponential growth to a case doubling time of more than 200 days - they arehighlighted in colour in the graphic. The epidemic has left its mark in these countries, but today there are few new infections.Covid-19 was initially able to spread rapidly from the epicentres of China and Italy , but was contained by some drastic measures. In this group, Switzerland is

particularly noticeable with a steep drop in the curve.

Group 2: Unbraked

Number of new Covid-19 infections each day compared to the number of all cases; Countries with over 5000 cases, logarithmic scales.
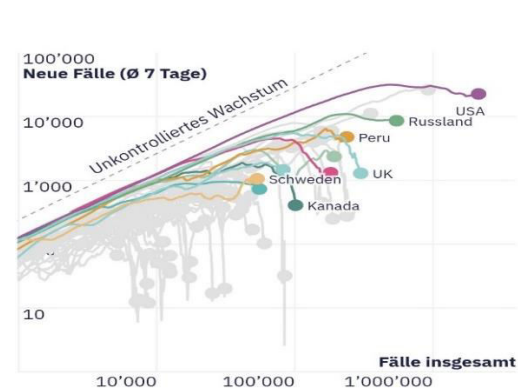

Credits – TagesAnzeiger.ch

The group of countries with continued unchecked growth of infections includes those with a doubling time of less than 30 days - of which, due to space constraints, only the nine countries with more than 50,000 cases are highlighted. Their curves grow exponentially, and most of them are not yet weakening.The "brakes" include Brazil and India, two of the countries most currently affected. The virus is also spreading rapidly in other large areas with large numbers of cases, such as Pakistan, Saudi Arabia and South Africa. The curves of these countries with a total of over 87,000 Covid-19 deaths are very close togetherthese also include Chile, Mexico, Columbia and Bangladesh.

Group 3: On the brink

Number of new Covid-19 infections each day compared to the number of all cases; Countries with over 5000 cases, logarithmic scales. The situation for countries in this group is uncertain: They have a doubling time of the confirmed number of cases of between 30 and 200 days - in the graph, countries are highlighted with more than 50,000 cases.
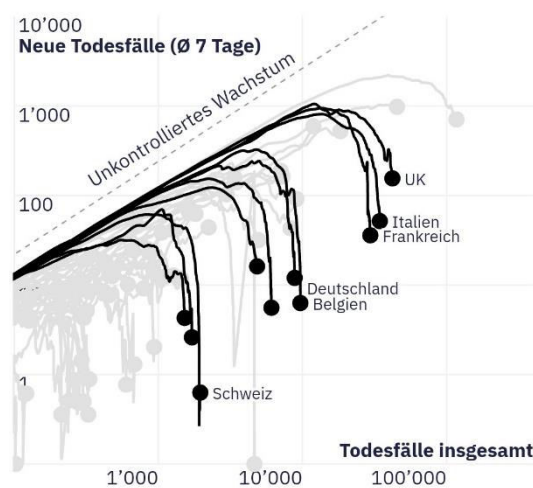

Credits –TagesAnzeiger.ch

While some members of this group show a slow but clearly recognizable downward trend (Canada, USA and Great Britain), the curve of other countries is still irregular and without any recognizable direction (Sweden and Russia).

Deaths

Number of daily deaths caused by Covid-19 compared to the number of all deaths caused by Covid-19; by country, logarithmic scales.

Credits – TagesAnzeiger.ch

The deaths can be viewed in the same way: the number of deaths per day compared to the number of all deaths. The graphic highlights countries that have more than 1,500 deaths caused by Covid-19 and have managed to keep daily growth below 0.5 percent of the total.

Findings and Discussion

The first three cases for COVID 19 were reported in India from Kerala on 30 January 2020, however they were of epidemiologically linked with other cases. Kerala later became COVID 19 free, before actual epidemic in I India started. Actual epidemic in India started on 2nd

March and since then the number of confirmed cases have been rising the same can be seen in Figure 1 The graph is showing the day wise new COVID cases with reference to seven days moving average of doubling time.
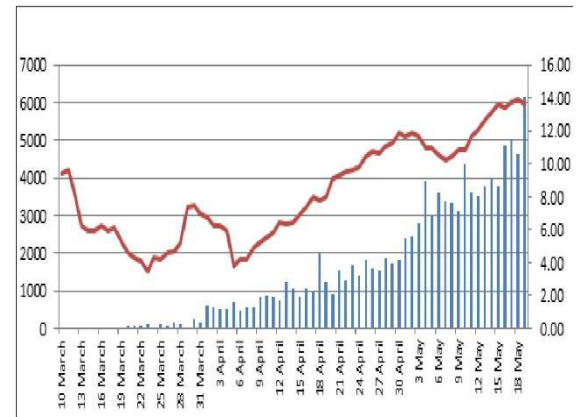

Figure I.Seven Days Moving Average of Doubling Time and New COVID 19 Cases per Day, India In the Figure 1. It is showing that on 19 May 2020, in India, total 6147 cases were positive for COVID 19, whereas the cumulative number is 106475 and the 7 days moving average of doubling time is 13.5.

In the graph below the number of cumulative cases since 1 March 2020 to 19 May 2020 are shown in blue colour bar and the numbers of removed tases are shown in red colour bar.

The line diagram in the same graph is showing the BAILEY'S Model Relative Removal Rate (BMRRR). In the graph the BMR is reached to 42. It will further depend on how different control and preventive measures are being taken for COVID 19 in the country.

Decrease in the BAILEY's rate will indicate spread of virus in the community and reasons for the same need to be ascertained specially to find out whether there are issues related to the implementation of the public health

measures. Important one to find whether infection is spreading in newer areas previously unaffected. Increase in the rate may indicate successful implementation of public health measure and also introduction of better therapeutic options or introduction of vaccine.
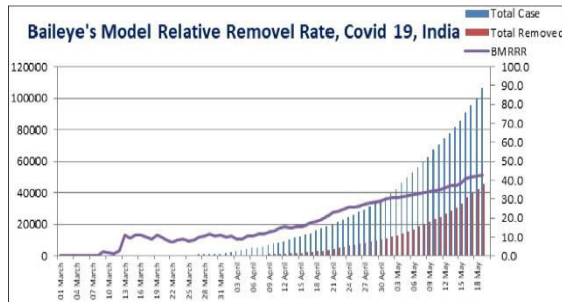

Figure 2.Bars shows the cumulative number of COVID:19 cases from 1-March 2020 to 19 May 2020 along with number of removed cumulative cases. The line diagram shows the change in BAILEY Model Relative Removal Rate.

According to this mathematical calculation, when the number of infected is equal to the number of removed patients, the coefficient will reach 100% threshold and the epidemic will distinguish. That means the transmission of the virus has more or less stopped. Taking the rate to higher and higher level is reflection of moving forward right direction and success of control measures being taken.
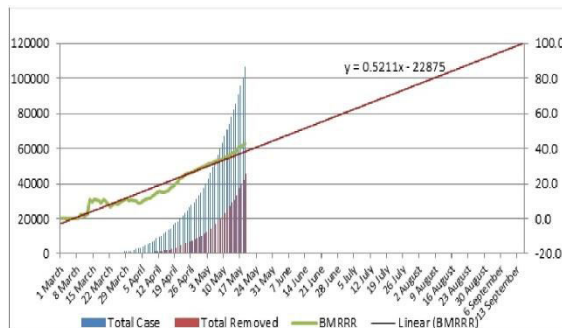

Figure 3. Regression Analysis (Linear) of Bailey's Relative Removal Rate, COVID 19, India

In the above graph, Regression Analysis (Linear) of Bailey's Relative Removal Rate (BMRRR), COVID 19, India has been done and it is showing that the linear line reaching to 100 in the mid of September, 2020. So it may be interpreted that at that point of time the number of infected will be equal to the number of removed patients and that's why the coefficient will reach 100% threshold.

However, we must understand that this is a Stochastic model and outcome will depend on variance structure around it. There are two main stochasticity, demographic and environmental.So the decision makers must try to Control the variables related to these two factors so that Bailey's Relative Removal Rate (BMRRR) continues to go up.

This can be an important tool in the hand of Central, State and District Authorities in decision-making and taking appropriate action at this time in the virus containment, epidemic situation management and control the pandemic spreading in the country. These predictions have far-reaching consequences regarding how quickly and how quickly and how strongly Government move to curb an Epidemic.

Benefit of using this mathematical model is that it takes into consideration of various factors including spread of disease(field activities and other public health measures), clinical care/recovery rate, effectiveness of any treatment/vaccine might be introduced later over a long period of time and not just short-term fluctuations.

Limitations

Based on the collected secondary data for a specific period of time to fit and estimate the basic case number, infection rate, and

recovery rate of COVID-19.When we apply any mathematical model, making some assumptions for a certain period of time, impact of few factors are being ignored such as population birth rate and natural mortality. Natural calamity, unpredictable population movement and important national or international events may have significant influence on this model. With the continuous release of epidemic data these important indicators may undergo significant changes in the spread of COVID-19 among the population.

Conclusion

The Prediction of the epidemic duration and size should be interpreted as applicable only to the confirmed cases based on the previous record. The issues regarding the accuracy reliability and standard of reported data is complex and are beyond the scope of this work, which is more oriented on the mathematical modelling side. This is a simple model which can be utilized at Centre, State and District level for better epidemic situation management and control.

As per the guidelines of Govt. of India and various states, faculty of Community Medicine Department from various Medical College is supporting State/District Public Health Authority in data analysis and interpretation. This faculty can support analysis and interpretation of State/District data (whenever numbers of cases are high) by using this model and also help in interpretation and relevant decision making. This will further help government to take long-term disease prevention and intervention measures.

References

1.) World Health Organisation, accessed on 24th June 2020.
<https://www.who.int/>

2.) Davit Gondauri, Etaterine Mikautadze, Mikheil Batiashvili,*Research on COVID-19 Virus Spreading Statistics*, Modestum, 2020

3.) Organisation WH. Coronavirus disease 2019 Situation Report – 70. 30 March 2020

4.) Andrew Tindall,*Coronavirus (COVID-19): a case study in emerging disease*, News & Insights, 2020

5.)Aatish Bhatia, 2020, *Minute Physics and World Data provided by John Hopkins University*, accessed 24th June 2020.

6.) Bailey N,*The total size of a general stochastic epidemic*, Biometrika, 1953