Modelling and Analyzing Covid19 Pandemic (Preliminary Draft)

Miguel H. Schindler, DeltaP mschindler@deltap.com Cipolletti, Río Negro Argentina 2-Apr-20 (updated on 10-Apr-20)

Abstract

Unlike in previous epidemic outbreaks, there are significant amounts of data available, from different countries in the world, to describe the developing situation of CoVid19 pandemic.

There is a necessity to improve the quality of the virus propagation models in order to predict, recommend and evaluate measures to control or moderate its impact.

A preliminary evaluation of the cumulative affected cases in Semi-Log plots shows that there is no straight line as expected in exponential growth.

On the other side, the same data in Log-Log plots shows a consistent straight alignment suggesting that the expansion follows a power or allometric law.

A computer SIR type numerical model was developed using a nodal network of city or country size number of elements, each of them connected to a definite number of contacts.

The results of the model closely match the behavior of the existing field reports. The numerical model allows modifying parameters starting at any point in time in order to consider changes resulting from actions impacting on social behavior.

Introduction

The purpose of this paper is to present a method to describe and analyze the virus propagation behavior and to predict the effects of actions such as social distancing or others on the developing course of an epidemic:

- Reproduce expansion behavior from existing data.

- Recognize changes of the actual situation comparing to the expected behavior predicted in the model.

- Predict and evaluate the effectiveness of control or mitigation measures.

- Validate acquired data, possibly detecting cases where the acquisition is incorrect or incomplete.

- Enhance the value of each taken measurement as a control point of the simulation.

- Provide a "type – curve" normalizing method for better understanding and diagnosing the model.

Data Reports

Data from reported country records was collected from "*Worldometers*" web site is shown in Figure 1.

In order to compare the curves, the time was shifted to the onset of local transmission.
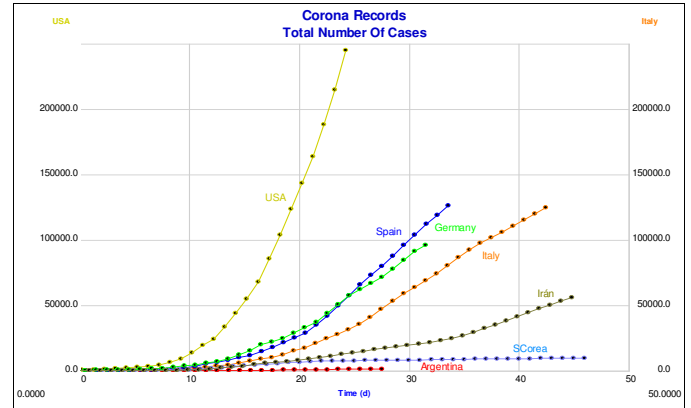

Figure 1. Cartesian Graph Actual Records

Figure 2 shows the cumulative positive cases in a SemiLog Plot. The total cases have consistent curvature and there is no straight alignment as expected in case of exponential growth. A real case report is included in Figure 2a.
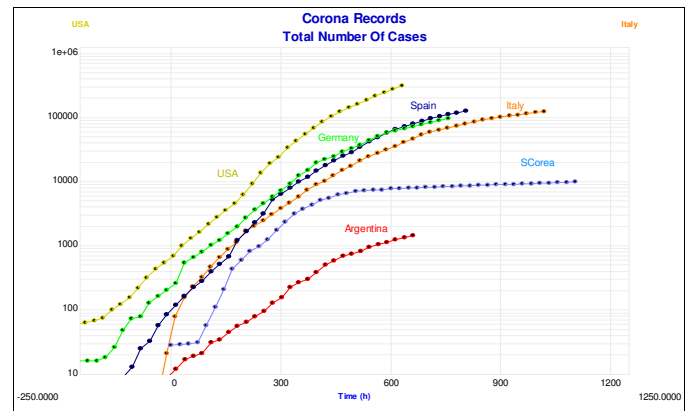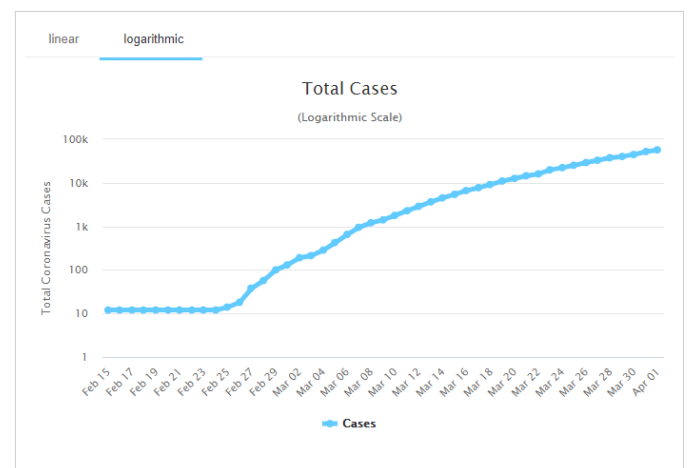

Figure 2. Semi-Log Plot

Figure 2a. France's Semi-Log Plot

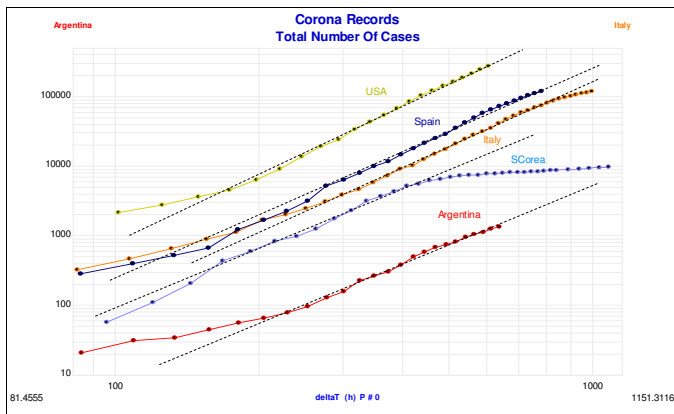In Figure 3 the same data is plotted in Log-Log format.

Figure 3. Log-Log Plot Actual Records

Straight line behavior is detected in the Log-Log plot. The straight line portions show a remarkable tendency to align to the same slope.

The straight line portion can be described by a power or allometric law: Total cases = α . time ^ β

This slope is close to $\beta \approx 3$, consistent with a cubic expansion.

Model Description

An SIR type computer model was implemented in C++, with a network of people connected to a fixed number of neighbors.

Each node represents a person switching its state in Immune-Susceptible-Infected-Recovered-Deceased.

Each person is connected or linked to neighbors in the network.

At every time step, each node is exposed to contagion from neighbors. The probability of switching from Susceptible to Infected is $P = 1 - (1-P_i)^n$. Where n is the number of infected contacts and P_i is the probability of transmitting the virus in that time step. The result is defined evaluating a random function.

After infection, the subject becomes contagious in a period between incubation and healing elapsed times.

The process runs for a definite number of time steps.

The model was experimented using 1, 2 and 3 dimensional arrays. It was found that the straight line portions of the total cases output in Log-Log plots have the same power exponent as the number of dimensions of the array.

Taking into account that the exponent of the power law of the cases of the country reports is 3, a 3-dimensional topology was selected.

The nodal network of people was implemented selecting a cubic matrix topology. **Figure 4**

This configuration enables for multiple options or contacts as possible contagion sources. It was selected 6 links per node. It is possible to select less links keeping the 3 power exponent, as long as there is at least 1 link for each dimension.
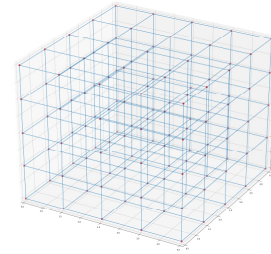

Figure 4. Network.

The matrix is initialized considering the total population TP = n . n . n. Each person is assigned a status, either Immune or Susceptible.

A fixed number of initial nodes is switched to Infected.

The simulation runs for a definite number of time steps.

In order to keep the probability of infections uniform, the probability was adjusted as follows:

$$(1 - P_i) = (1 - P_{TS})^{(\gamma \frac{t}{T} + \delta)}$$

where

$$\gamma = -3.64\, P_i$$

and

$$\delta = 1 - \gamma$$

Simulation Model Outputs

The results of the model will be shown in graphs below.
Figure 5 is a Cartesian plot describing the process from initial infection to total saturation of susceptible nodes.
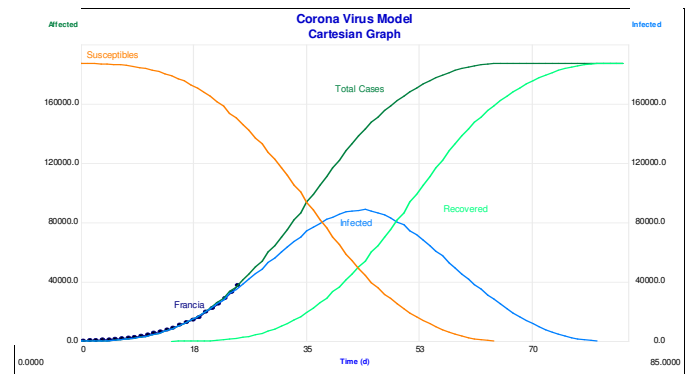

Figure 5. Model Cartesian Graph

Figure 6 shows the cumulative positive cases in a Log-Log Plot. After a short initial transient period, the evolution of total cases follows a straight line. This pattern continues in time until the infection reaches barriers to its propagation or unitl there are no more susceptible nodes connected to infected nodes.
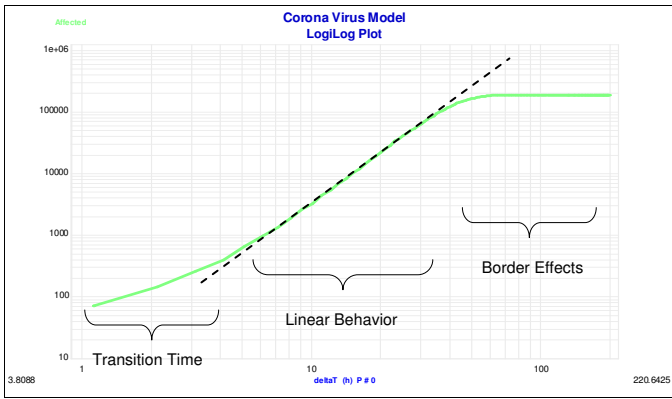
Figure 6. Model Log-Log Plot

The scalar parameter α is affected by the number of outbreaks, the number of initial nodes infected, the incubation time and the contagion probability.

$$\alpha = f(N, T_{incubation}, Pi, N^{\underline{o}}\ initially\ infected)$$

Then the total cases during the straight line expansion for a single epidemic growth can be calculated as:

Eq.$_1$ Tc= αt^3

Normalized Diagnostic Plot

Given the fact that simulations depict the same Log-Log plot slope (or the same allometric power parameter), it is convenient for comparison purposes to normalize the total cases evolution curve, dividing by a factor of (time ^ β).
The resulting curve will be horizontal in the straight line portion of the propagation.
The reason for this artifact is that horizontal lines are readily identified by the human eye.

- Enable to know if the initial transient period has ended.

- If acquired data moves away from the horizontal, it can describe changes, like new sources, propagation barriers or effects of policies applied on the social behavior.
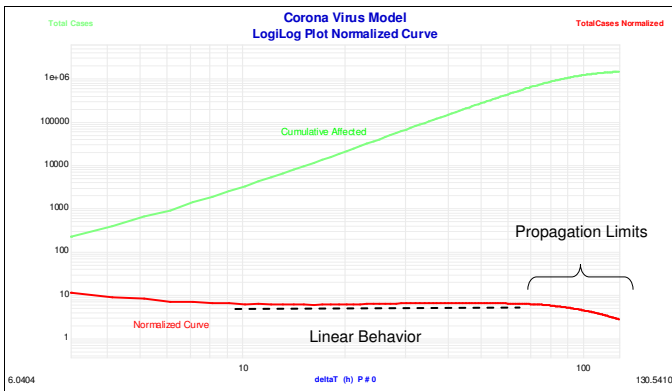

Figure 6. Normalized Total Cases Curve

In Figure 6, the normalized curve is included. After the initial transient period, there is a horizontal line on the normalized

curve that corresponds to the linear or allometric portion of the propagation. At longer times, the descent on the normalized curve indicates that the propagation process is reaching a limit or saturating the available susceptible population.

Parameter Sensitivities in Log-Log Plot.

Figure 6a, shows sensitivity to population size. Cases from the USA were matched with the models.

It has been found that in smaller populations, the saturation cycle takes less time. In order to match the virus growth rate, the model was enabled to define multiple cities at the same time.

N is the number of epidemic growths

Eq.$_2$

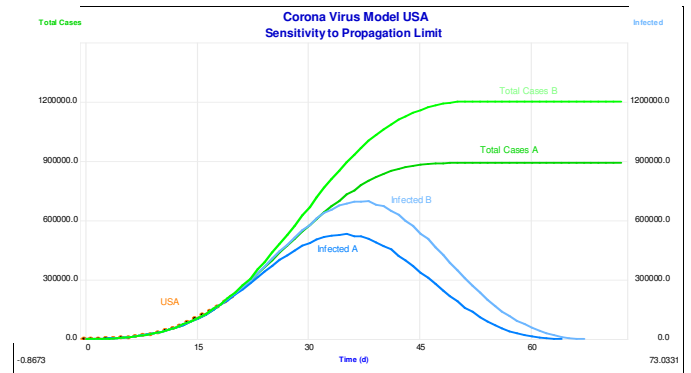$$Tc = \sum_0^N \alpha\ (t - t_i)^3$$

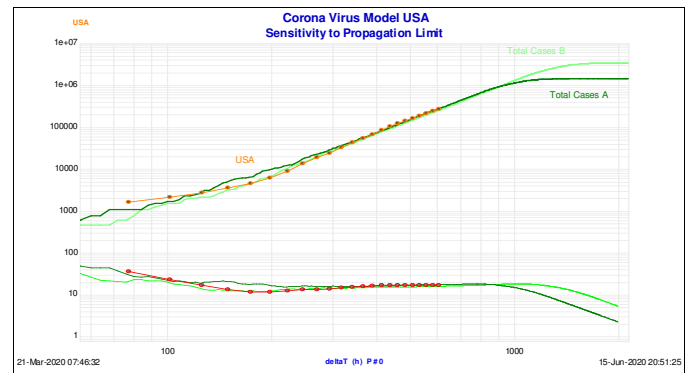**Figure 6a. Cartesian sensitivity to population size**

Figure 6b. Log-Log sensitivity to City Size

In Figure 6b, the same cases are shown in Log-Log plot.
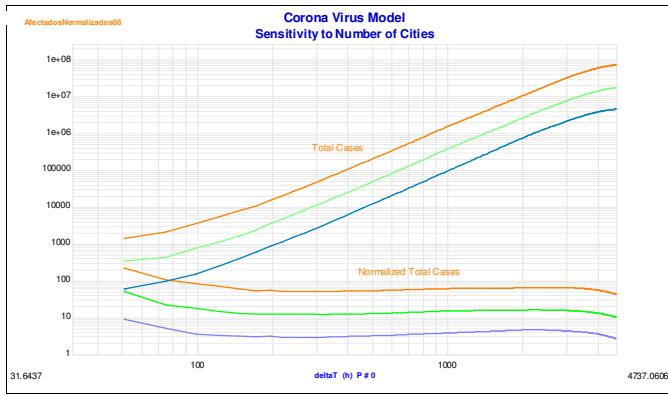Larger population sizes will divert from the horizontal line at longer elapsed times.

Figure 6c. Log-Log sensitivity to Number of Cities

Changing the amount of simultaneous epidemic growths allows matching country reports by means of defining multiple cities or neighborhoods.
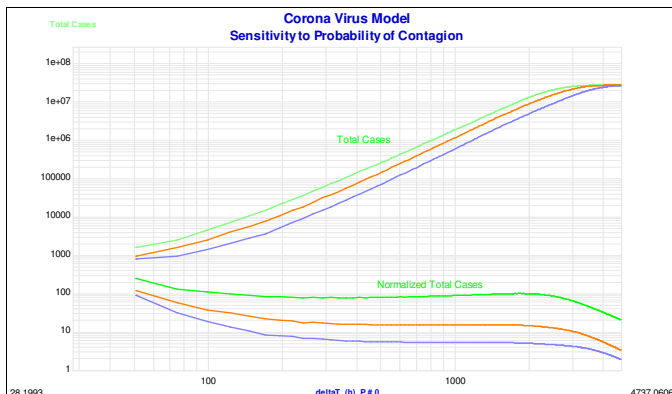

Figure 6d. Log-Log sensitivity to Contagion Probability

In Figure 6d, the probability of transmitting the disease was set at 85, 35 and 20%.

By means of calculating the normalized curve position it is possible to estimate the number of active epidemic growths.

Applications of the Method

In this section data from specific countries is matched with the model presented.

South Korea

In Figure 7a, the curve with dotted markers indicates the total historical confirmed cases.

Both models match the initial behavior adequately.

At the time indicated with the arrow, the change indicates the successful result of the limiting measures taken to control the epidemic. This effect was accounted in the model, changing the probability of transmitting the virus.

It is important to take notice, that, even when applying these control actions, there are ongoing active cases.

Without these actions, the total cases could have evolved according to the curve shown above.

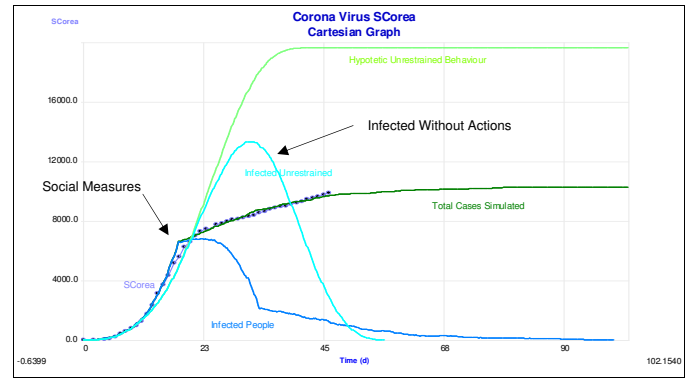The ripple in the simulated curves is caused by the stochastic process of infecting the nodes.

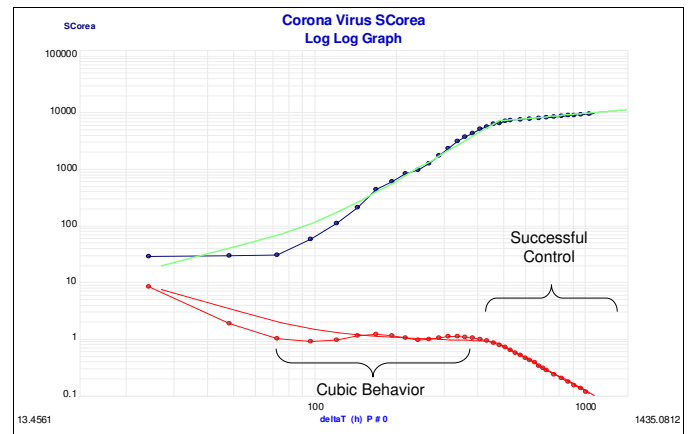**Figure 7a. South Korea Cartesian**

Figure 7b. South Korea Log-Log Plot

In Figure 7b, the cubic behavior is reached after 70 h.

After 400 h, the downward trend of the normalized curve indicates that control measures are effective.

Italy

In Figure 8a, the curve with dotted markers indicates the total historical confirmed cases.

At the time pointed with the arrow, the change indicates the effects of the limiting measures taken to control the epidemic.
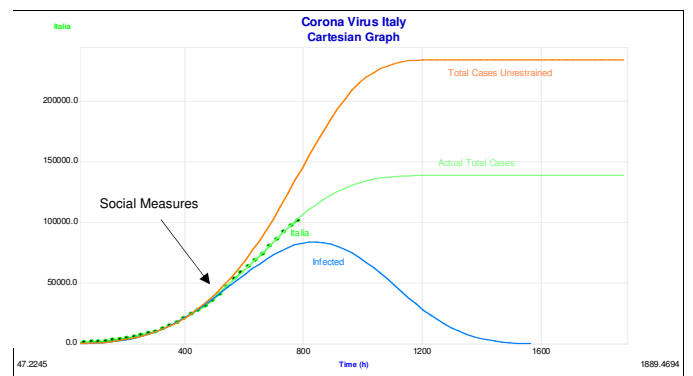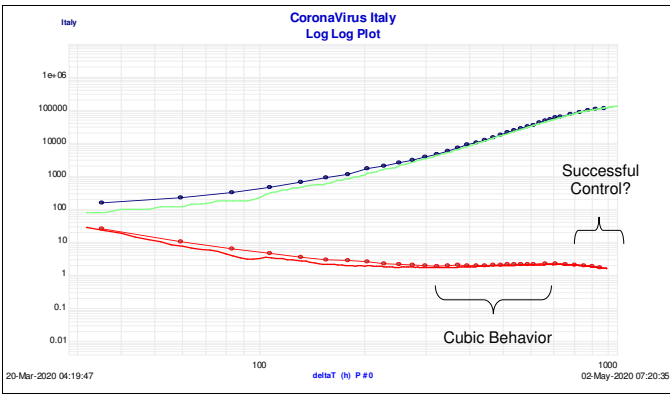

Figure 8a. Italy Cartesian

Figure 8b. Italy Log-Log Plot

After 600 h, the downward trend of the normalized curve seems to indicate that control measures are effective.

Germany

In Figure 9a, the curve with dotted markers indicates the total historical confirmed cases.

At the time pointed with the arrow, the changes indicate that social distancing measures are starting to limit the expansion of the epidemic. In both cases A and B, the contagion rate was reduced to the same value to account for social measures. Even if social measures are kept, if the expansion is not containined, it could evolve to case B or higher.
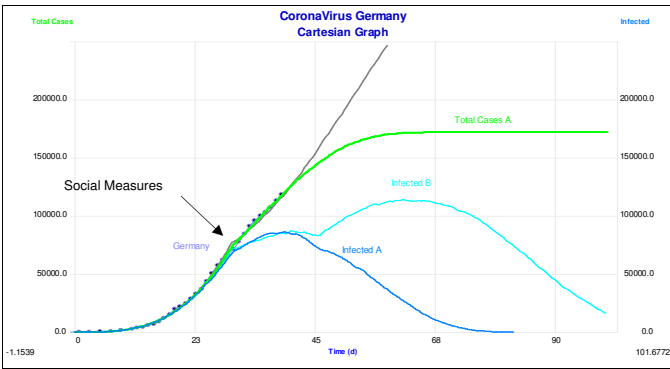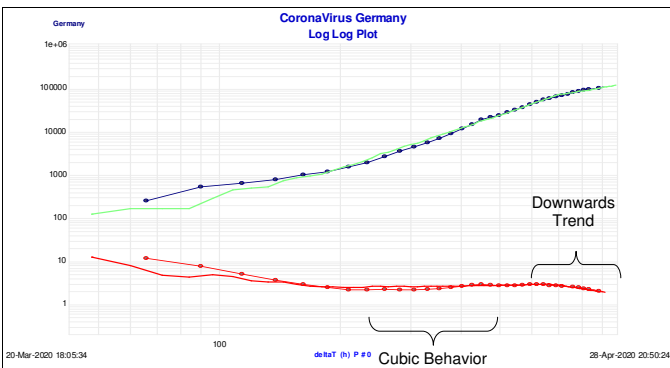

Figure 9a. Germany Cartesian

Figure 9b. Germany Log-Log Plot

After 600 h the downward trend of the normalized curve seems to indicate that control measures are effective.

USA

In Figure 10a, the curve with dotted markers indicates the total historical reported cases.

There is a slight downward movement in the normalized curve that indicates that control measures are starting to affect contagion rates.

In figure 10a there is a sensibility to propagation limits (before social measures were taken).
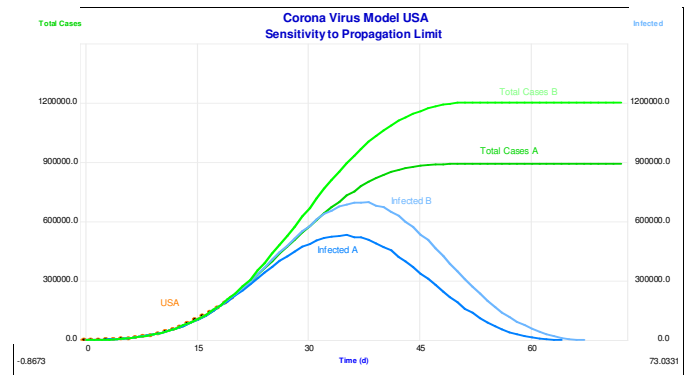

Figure 10a. USA Cartesian

Figure 10b shows the impact of social measures on number of infected. In Case A, the measures stay in place until the end of the epidemic. In Case B they are interrupted. The output of the model reflects the increase in total and infected cases.
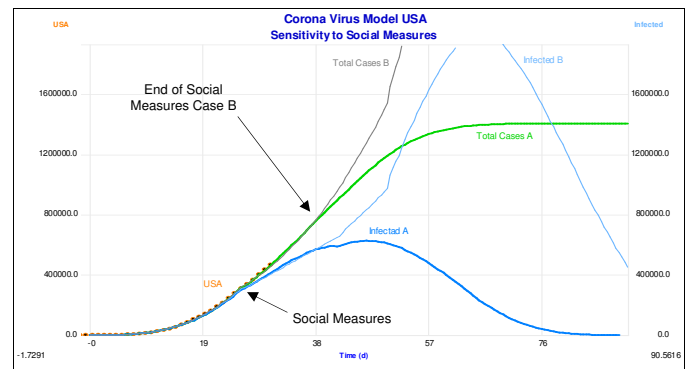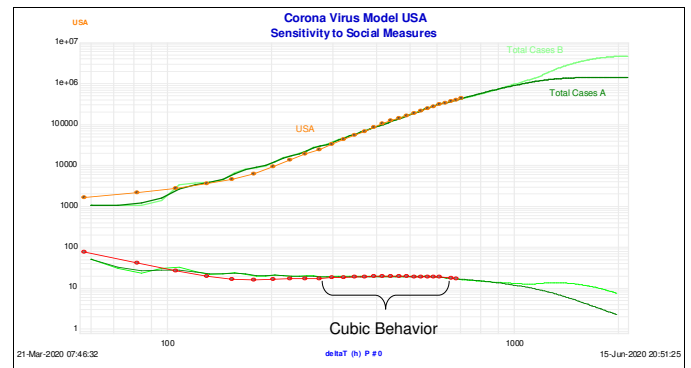

Figure 10b. USA Cartesian

Figure 10c. USA Log-Log Plot

Argentina

In Figure 11a, the curve with dotted markers indicates the total historical reported cases.
Argentina has taken early social distancing measures.
There is a slight downward movement in the normalized curve which might indicate that control measures are starting to contain the propagation of the epidemic. Even if social measures are kept, if the expansion is not contained, it could evolve into Case B or higher.
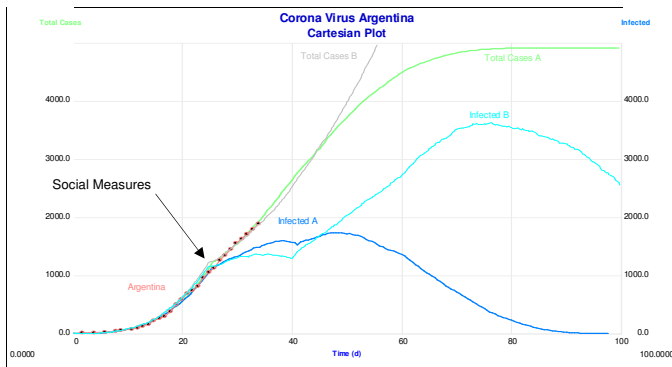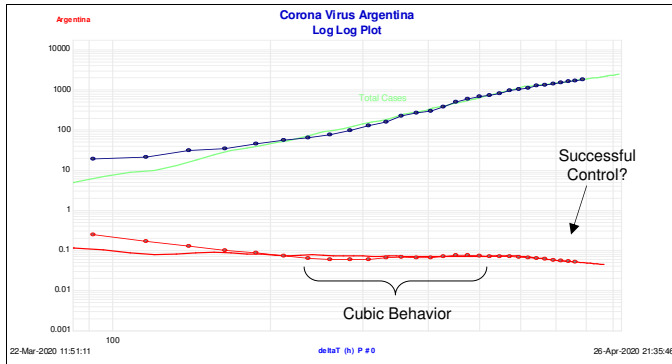

Figure 11a. Argentina Cartesian

Figure 11b. Argentina Log-Log Plot

Conclusions

The propagating behavior of the epidemic follows an allometric law with a cubic power coefficient.
A three-dimensional node array was used to best describe the personal contact network.
The presented numerical model has the capability of accurately reproducing the data analyzed.
The unrestrained model behavior could expand up to the total susceptible population.
It makes it possible to estimate the number of independent epidemic growths.
The normalized curve presented in the Log-Log plot improves monitoring and diagnosing the evolution of the epidemic.

Difficulties and Future Work

Investigate if the scale of human personal contact network correlates with model matrix dimensions.
Detailed case analysis.

Nomenclature

α = allometric scalar factor
β = allometric power factor
t = time from start of the contagion
Pi = probability of transmitting the virus in a reported time T for each contact
P_{TS} = probability of transmitting the virus in a simulation time step for each contact
N = number of independent epidemic growths
z = number of contacts of each node
Tc = Total Affected Cases

References

1. FIRST ATTEMPTS TO MODEL THE DYNAMICS OF THE CORONAVIRUS OUTBREAK 2020 THOMAS GÖTZ Mathematical Institute University of Koblenz–Landau D–56070 Koblenz 10Feb20
2. POWER-LAW DISTRIBUTION IN THE NUMBER OF CONFIRMED COVID-19 CASES. Bernd Blasius. Institute for Chemistry and Biology of the Marine Environment University of Oldenburg, Oldenburg, Germany
3. AN ARIMA MODEL TO FORECAST THE SPREAD AND THE FINAL SIZE OF COVID-2019 EPIDEMIC IN ITALY. Gaetano Perone. University of Bergamo, Italy