Evolution of the COVID Pandemic: A Technique for Mathematical Analysis of Data

K. Tennakone*[a,b], L. Ajith DeSilva[c], K. J. Wimalasena[d] and Z. J. Welchel[c]

[a] *National Institute of Fundamental Studies, Sri Lanka*

[b] *Department of Physics, Georgia State University, United States*

[c] *University of West Georgia, United States*

[d] *University of Peradeniya, Sri Lanka*

Abstract

The analysis of COVID-19 statistics in different regions of the world with the intention of understanding the trend of progression of the pandemic is a task of paramount importance. Publicly available data includes cumulative number of cases, new cases each day, and the mortality. Extracting information from this data necessitates mathematical modeling. In this note a simple technique is adopted to determine the trend towards stabilized elimination of the infection, as implicated by saturation of the cumulative number of cases. Results pertaining to several representative regions of the world are presented. In several regions, evidence there to the effect that the pandemic will come to an end. The estimated saturation values of the cumulative numbers are indicated.

Keywords: Coronavirus, Mathematical modelling, COVID modelling, COVID cumulative numbers

*Email: ktenna@yahoo.co.uk

1. Introduction

A question of utmost importance to individual nations and the whole world is how the COVID-19 pandemic would progress and whether available statistics can be utilized to determine the future trend in terms of mathematical models. In this note, we present a simple technique of analyzing publicly available COVID-19 data based on the cumulative number and the new cases each day for several geopolitical regions. Analysis indicates a current general trend of cumulative COVID-19 cases approaching an equilibrium – possibly a sign that the pandemic is slowly moving towards a halt, perhaps because of the effectiveness of control measures and development of heard immunity. In few regions, slow exponential growth seems to continue.

2. Method and Discussion

The publicly available data on progression of COVID in different geographical regions of the world are:

(1) Cumulative number of cases N at the beginning of a given date (defined as at an instant of time t).

(2) New cases dN reported at the end of each day (after a time interval dt = 1, therefore dN is effectively $\frac{dN}{dt}$). The rate of growth can also be computed from plots of *N* vs *t.*

OriginLab data analysis software is used plot curves and perform correlation analysis. The very initial growth of the epidemic in many regions has been exponential[1-2], therefore we assume that it evolves according to the dynamical system equation,

$$\frac{dN}{dt} = N\,F(N) \tag{1}$$

where the growth constant $F(N)$ is a function of N, approximating to a constant around $N{\sim}0$. A basic assumption of the model is that the rate of increase of the number of infective (number of reported in a day) is a function of the cumulative number of cases at that instant. Implying progression of the epidemic at a given instant is governed by its history.

From (1)

$$\frac{d}{dN}\left(\frac{dN}{dt}\right) = N\frac{dF}{dN} + F(N) = N\frac{dF}{dN} + N^{-1}\frac{dN}{dt} \tag{2}$$

The pandemic stops brewing and reach an equilibrium when $\frac{dN}{dt} = 0$ after attaining a certain cumulative number $N = N_E$. The point $N = N_E$ is asymptotically stable[3] at this point, provided $\frac{d}{dN}\left[\frac{dN}{dt}\right] < 0$. Clearly this condition satisfied if,

$$\frac{dF}{dN} < 0 \text{ at } N = N_E \tag{3}$$

Eqn. (1) also embodies the requirement that point that point N = 0 corresponds to an unstable equilibrium. To illustrate these ideas, consider the logistic model, where $F(N) = \alpha - \beta N$, (α and β are positive constants), with the equilibrium point $N_E = \frac{\alpha}{\beta}$ as $\frac{dF}{dN} = -\beta$, the equilibrium is stable. The cumulative number of cases in some epidemics approximately agree with the logistic

model[4-7]. Our analysis shows that COVID-19 cumulative numbers deviate from the logistic variation.

The plot of $F(N)$ vs N is same as the plot of $N^{-1}\frac{dN}{dt}$ vs N, if $F(N)$ turns out be monotonically decreasing and reaching zero, either cutting the N −axis or approaching N −axis asymptotically, the pandemic will practically reach an equilibrium. In the simple logistic model $F(N)$ is a negative slope straight line, cutting N −axis at the point of equilibrium. Possible scenarios are depicted in the Fig. 1 below, which portray different modes of variation of $F(N)$ with N.

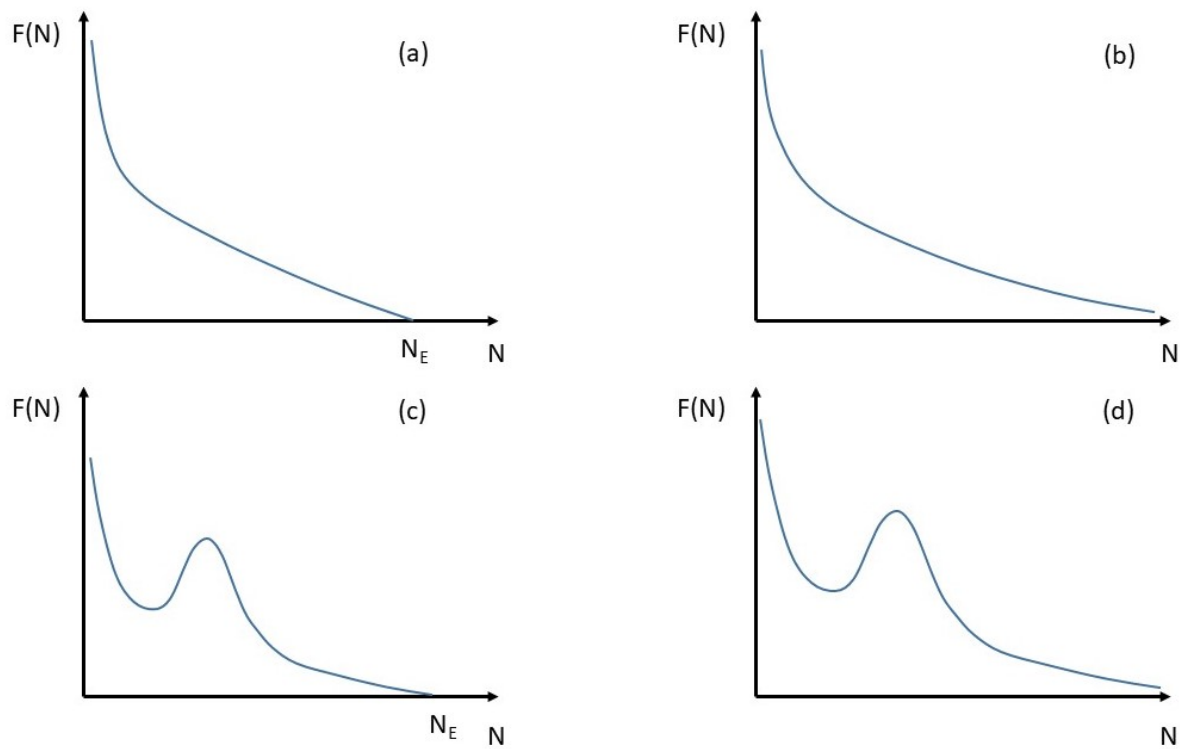
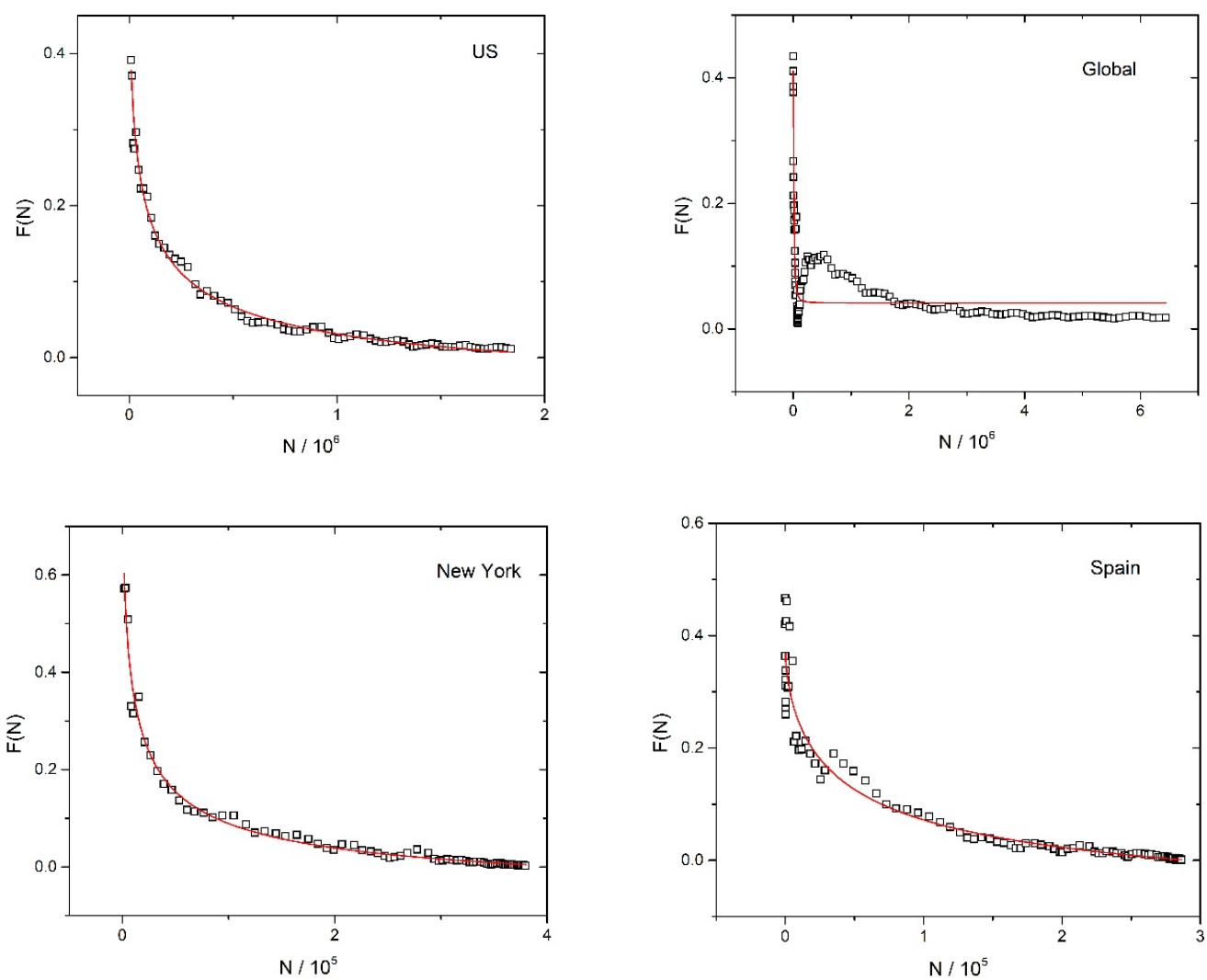
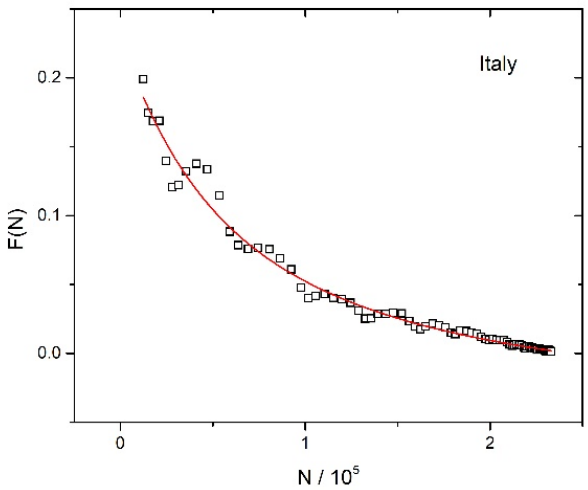


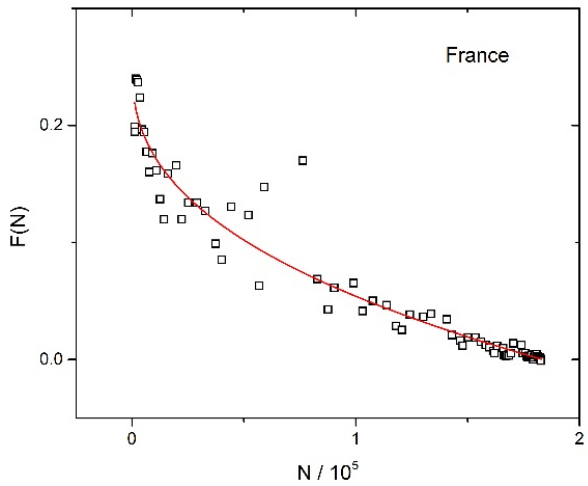
Fig. 1: Sketch of different modes of variation $F\ (N)$ with N.

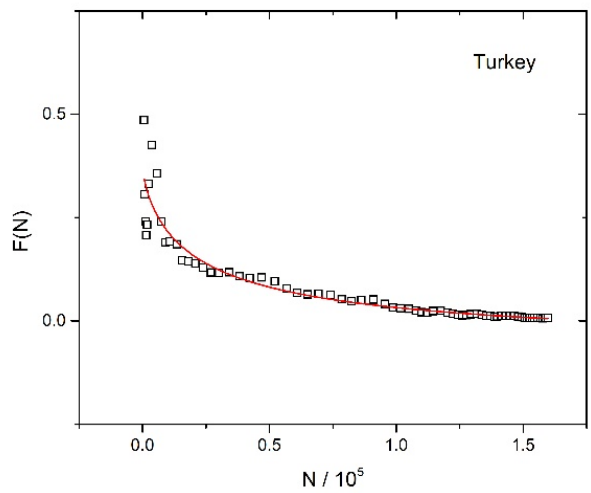
In the figures (a) and (b) $F(N)$ decreases continuously, and in (a) a stable equilibrium is reached at $N = N_E$, whereas in (b) equilibrium is approached asymptomatically. In figures (c) and (d), the monotonic decrease of N occurs after reaching a peak. In dealing with real data, the behaviour of $F(N)$ at points close to $N = 0$ is obviously erratic, but does interfere with the subsequent trend.

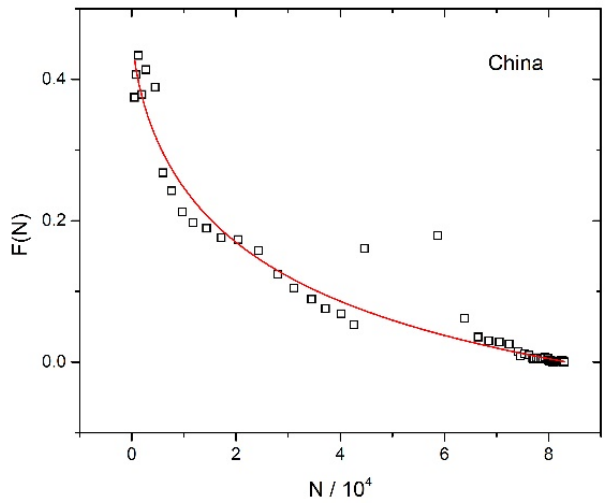
The plots of $F(N)$ vs N for several representative geopolitical regions are presented below (Fig.2). All computations are based on publicly available data[8].

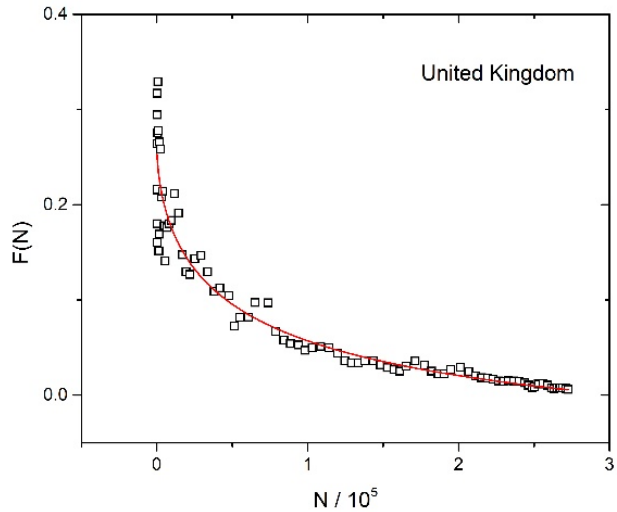


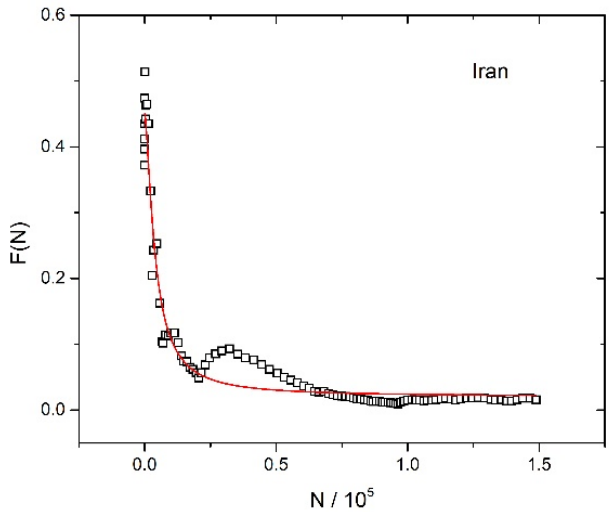

Italy
F(N)
N / 10^5

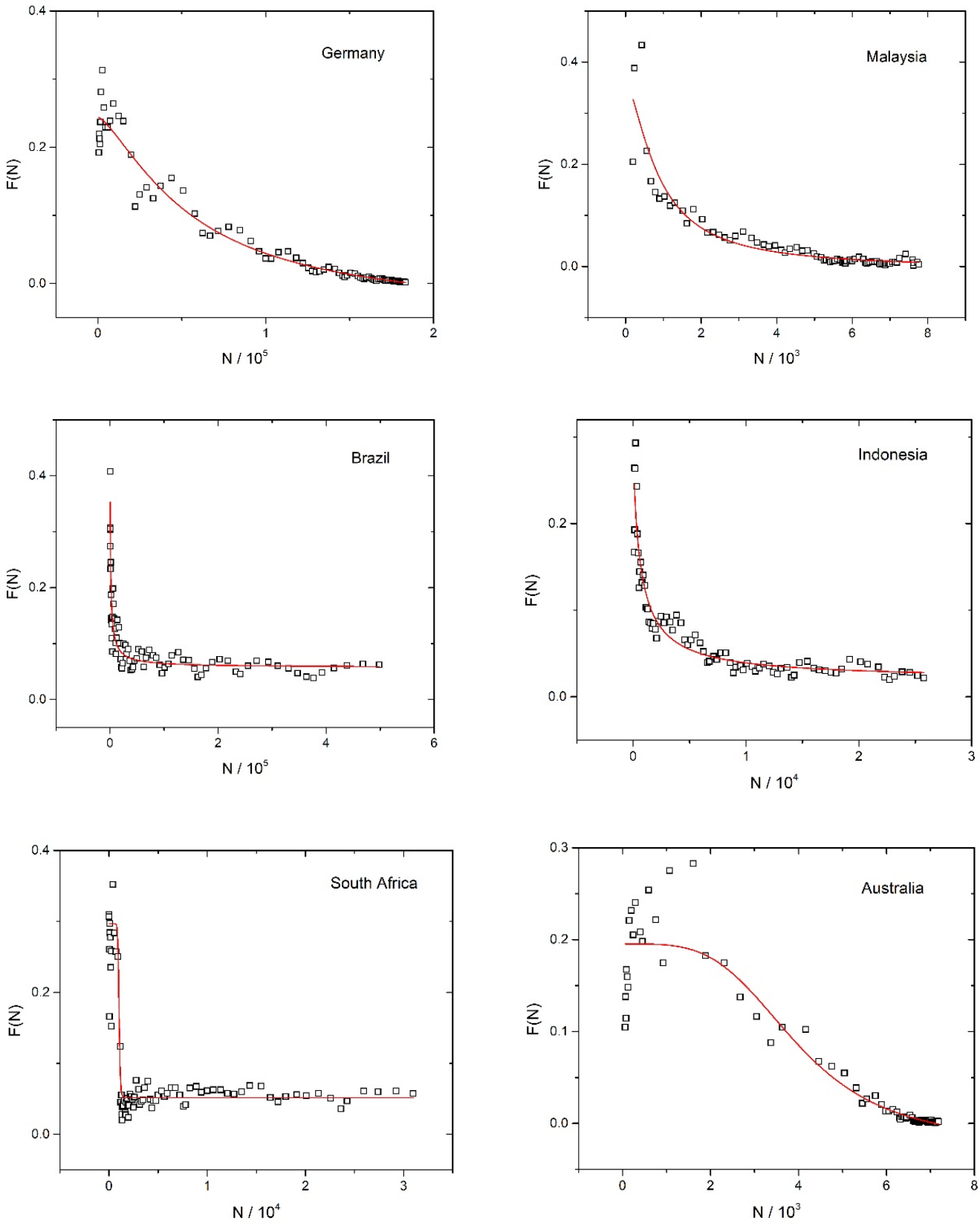

France
F(N)
N / 10^5


Turkey
F(N)
N / 10^5


China
F(N)
N / 10^4


United Kingdom
F(N)
N / 10^5


Iran
F(N)
N / 10^5


Germany
F(N)
N / 10^5
Malaysia
F(N)
N / 10^3
Brazil
F(N)
N / 10^5
Indonesia
F(N)
N / 10^4
South Africa
F(N)
N / 10^4
Australia
F(N)
N / 10^3

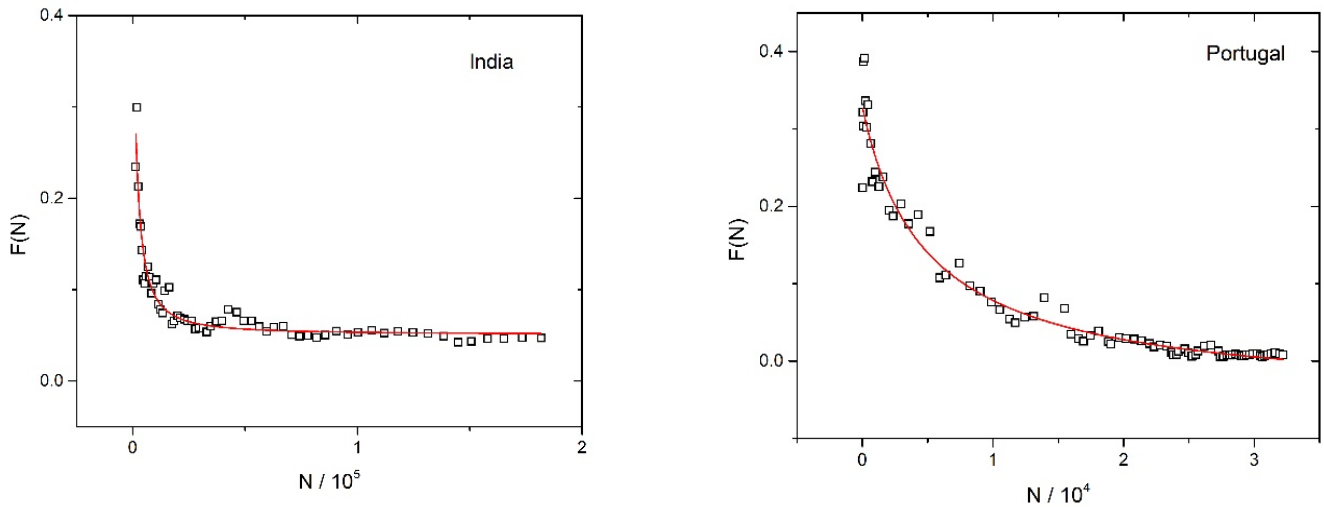


Fig.2: Variation of the function $F(N)$ for different regions with number of cumulative cases N.

In most the above plots the later trend in variation of $F(N)$ with N is monotonic decrease, indicating that $\frac{dF(N)}{dN}$ remains negative and approach $N-$ axis towards a stable equilibrium. In some regions, identification of functional form $F(N)$ using OriginLab software, enabled determination of the point at which $F(N)$ will cut the $N-$ axis allowing determination of the saturation of the cumulative number of COVID cases. In few plots the current trend approximates to a straight line nearly parallel to $N-$ axis, indicating that slow exponential growth continue.

Fig. 3 shows the plot of $F(N)$ vs N for Sri Lanka. Although there is no matured COVID epidemic in Sri Lanka, the current trend is a monotonically decreasing $F(N)$.

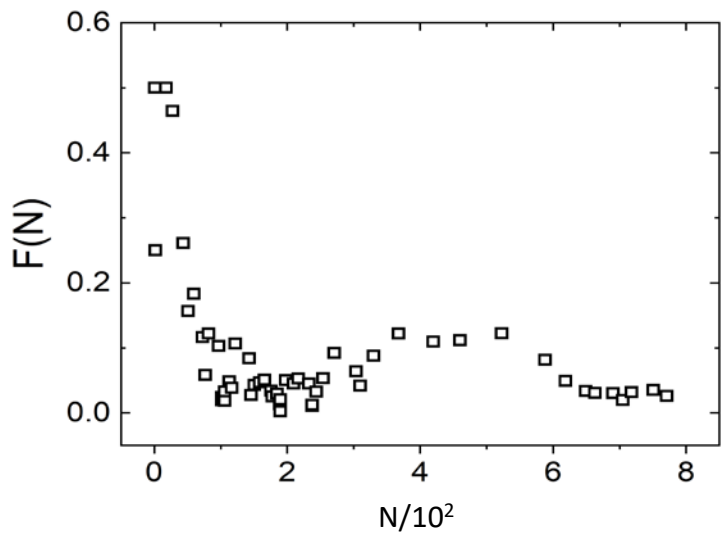


Fig. 3: Plots of $F(N) = N^{-1}\frac{dF(N)}{dN}$ vs *N* for cumulative cases in Sri Lanka.

When curve fitting is attempted to determine the functional form of $F(N)$, a good fit was obtained for the dose-response four parameter logistic function[9-10],

$$F(N) = A_2 + \frac{(A_1 - A_2)}{1+\left(\frac{N}{N_0}\right)^p} \tag{4}$$

The median values of the parameters A_1, A_2, N_0, P for a number of regions where the R^2 confidence level in curve fitting is greater than 90% are given in the Table.1.

Table (1): The values of the parameters A_1, A_2, N_0, P defined in equation (4) and values of N_E for different regions derived from the plots $F(N)$ vs N.

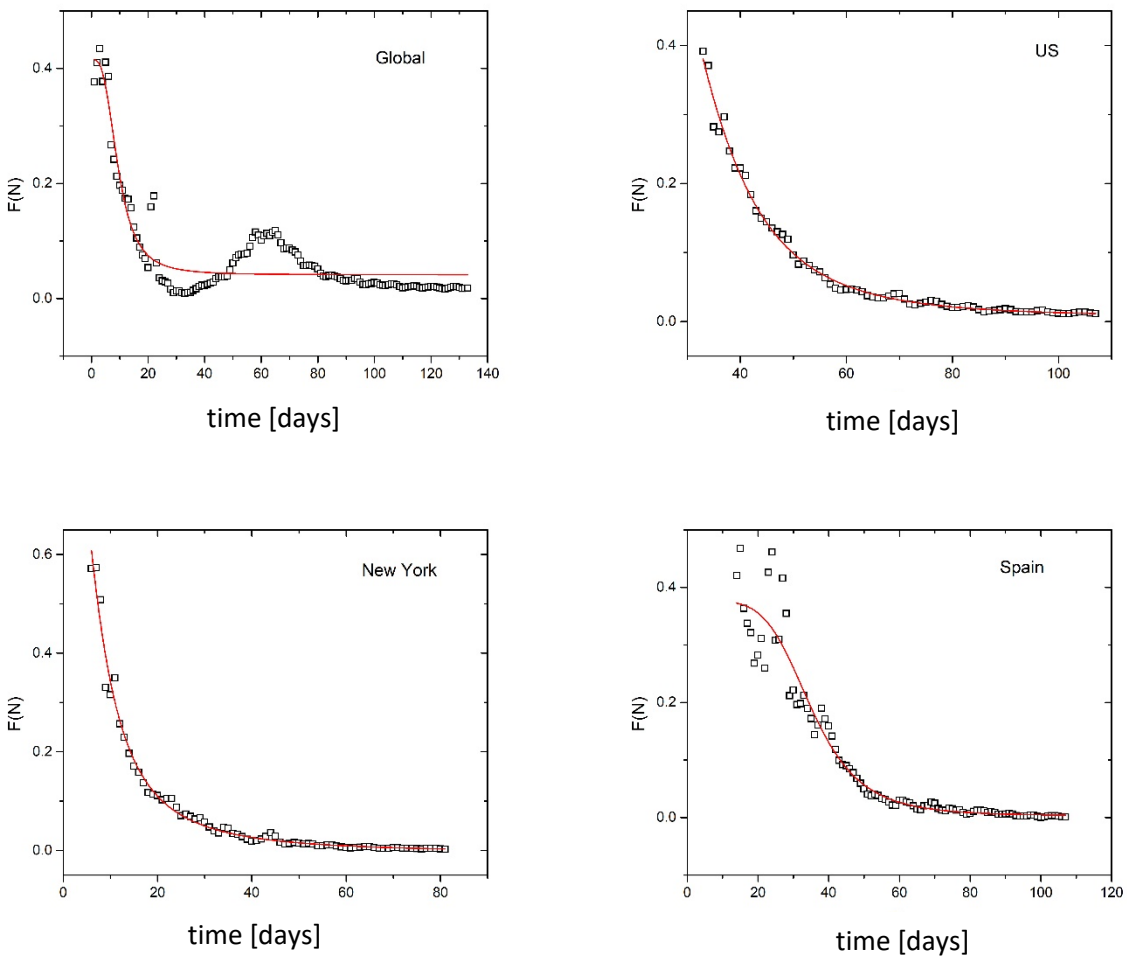
Region/county	A_1	A_2	N_0	P	N_E
Global	0.41406	0.04119	10617.30848	1.68395	--
US	0.55506	-0.06345	47521.15992	0.56162	2259510.992
New York	0.86337	-0.07446	8117.73888	0.61879	425953.854
Spain	0.37295	-0.11479	48228.20465	0.65413	292168.0176
Italy	0.22506	-0.049	61952.99835	1.11492	243174.8192
France	0.25309	-8.01084	$8.89446*10^8$	0.40713	183570.4111
Turkey	0.37143	-0.07666	22761.7411	0.76677	178223.6071
China	0.47111	-0.25691	33634.27007	0.66663	83524.85003
UK	0.25993	-0.10101	66763.62443	0.61732	308672.6637
Iran	0.45116	0.02046	4012.63294	1.48895	--
Germany	0.24438	-0.04552	56128.33271	1.38784	188400.0239
Malaysia	0.36245	-0.0048	865.73732	1.51162	15127.47387
Brazil	0.42691	0.05718	1149.5979	0.87332	--
Indonesia	0.28212	0.01943	669.88712	0.92724	--
South Africa	0.29688	0.05154	1025.47591	15.71136	--
Australia	0.19538	-0.02119	3964.11687	3.79703	7115.843504
India	0.3979	0.05092	2129.04562	1.28128	--
Portugal	0.3302	-0.05394	5112.11772	0.95988	33756.35386

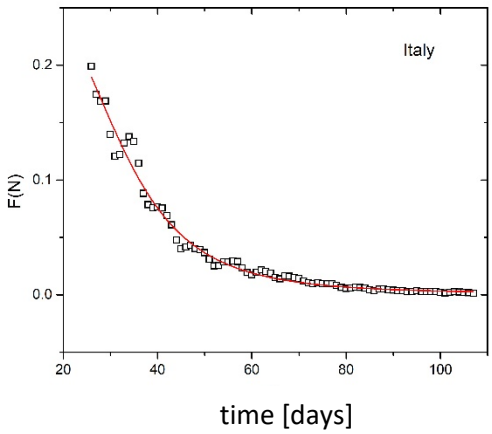
In cases where $A_2 = (-A_0)$ is negative, it follows from (4) that $F(N)$ vanishes (equivalently $\frac{dN}{dt}$) when,

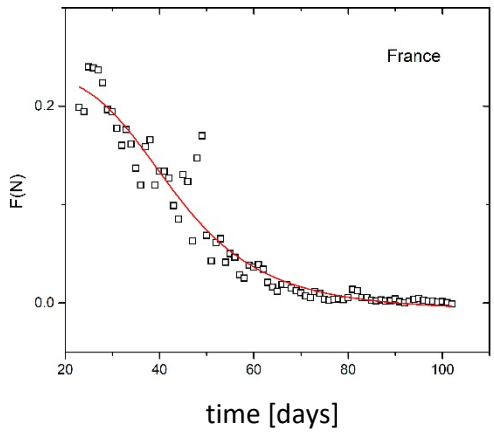
time [days]

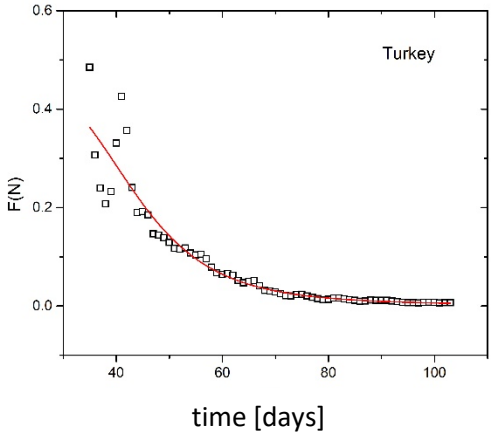
$$N = N_E = N_0 \left(\frac{A_1}{A_0}\right)^{\frac{1}{p}} \tag{5}$$

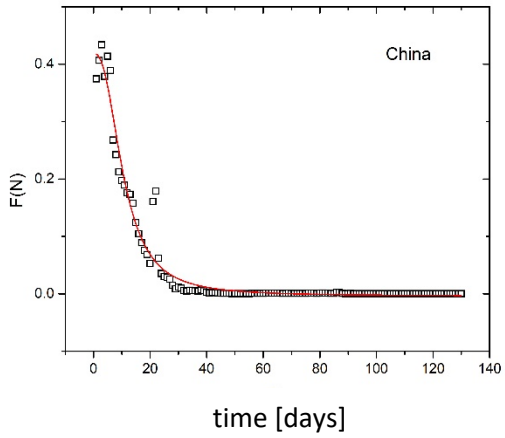
The values of N_E for different regions are also indicated in Table.1. Fig.4 shows how the function $F(N)$ for different regions vary with time. According to data analysis the value of the parameter A_2 is positive for Iran, Brazil, India, Indonesia and South Africa and the fitting does not suggest an approach to an equilibrium at present.

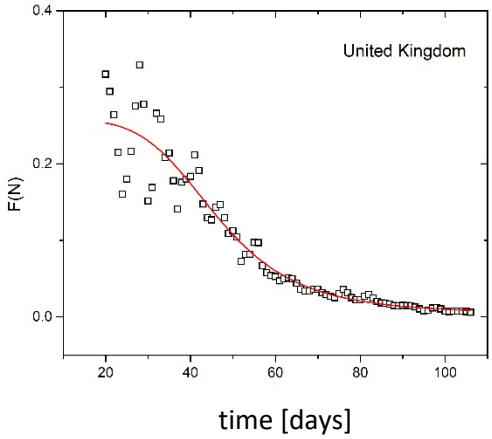


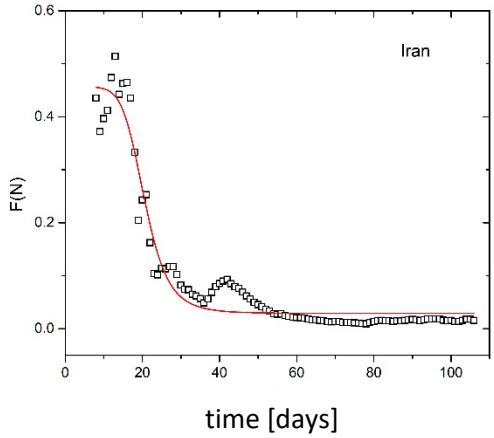

Italy
F(N)
time [days]

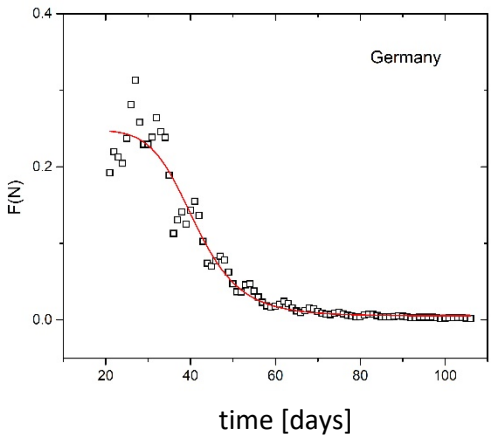

France
F(N)
time [days]

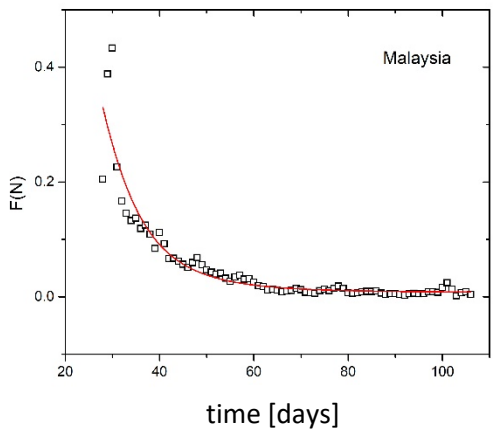

Turkey
F(N)
time [days]

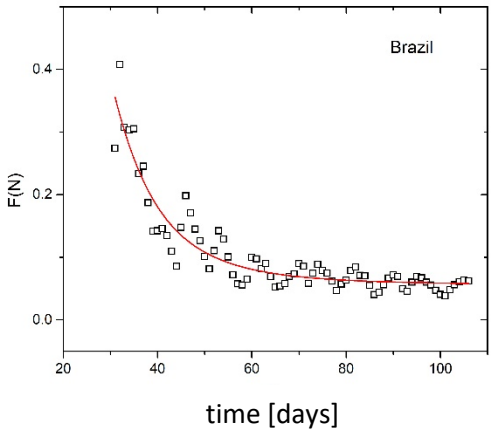

China
F(N)
time [days]

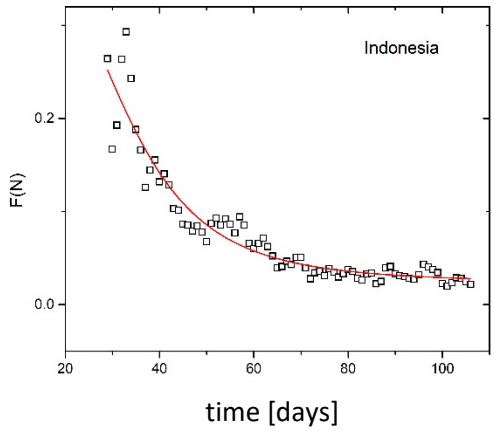

United Kingdom
F(N)
time [days]

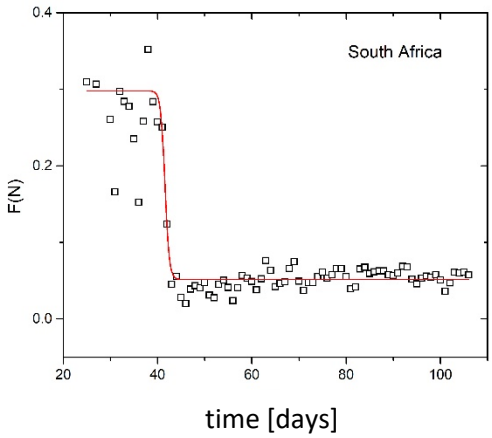

Iran
F(N)
time [days]

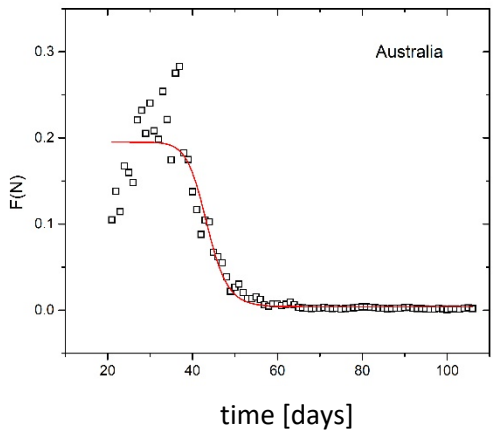

Germany
F(N)
time [days]


Malaysia
F(N)
time [days]


Brazil
F(N)
time [days]


Indonesia
F(N)
time [days]


South Africa
F(N)
time [days]


Australia
F(N)
time [days]

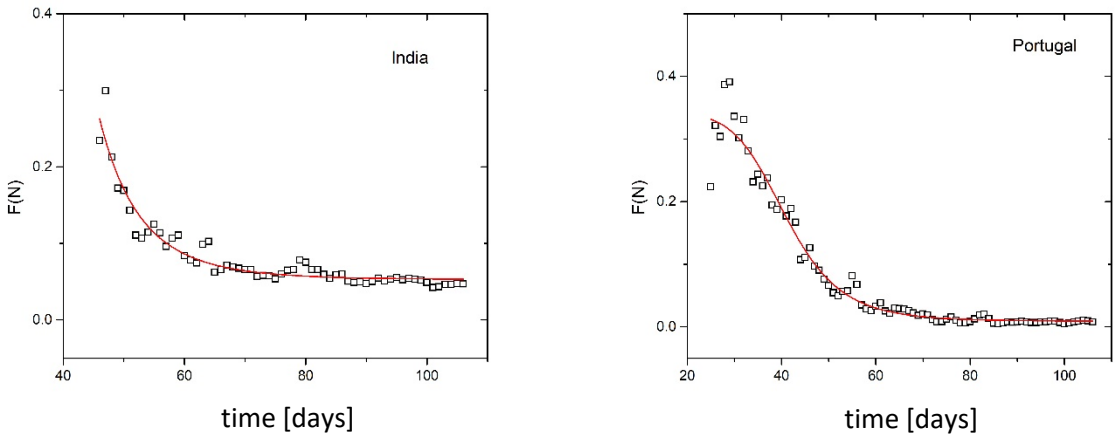


Fig.4: Variation of the function $F(N)$ for different regions with time.

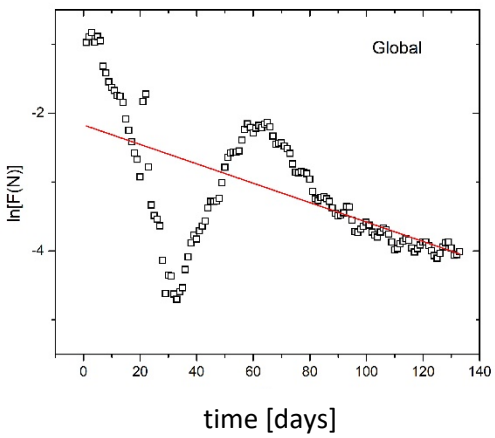
As seen from Fig.5, plot of $\ln[F(N)]$ vs time (t) fits to a straight-line after initial stages of the epidemic, so that

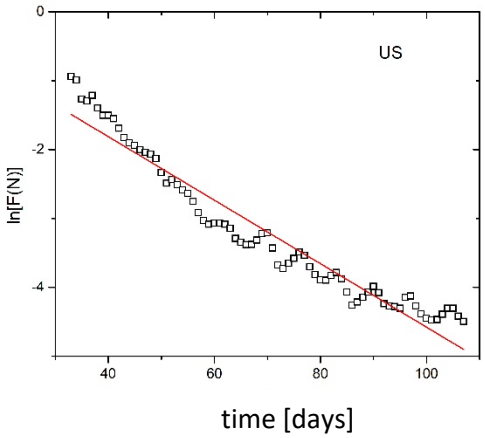
$$\ln\left(\frac{1}{N}\frac{dN}{dt}\right) = \ln[\,F(N)] = a + bt \tag{6}$$

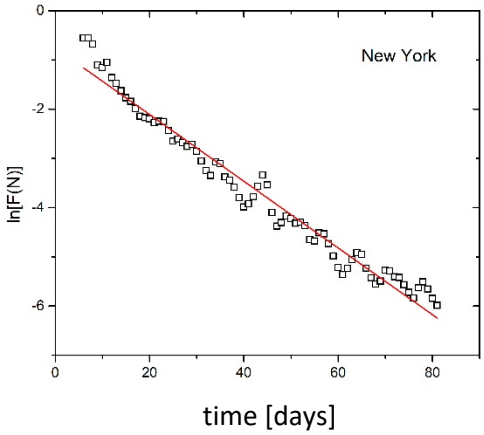
When (6) is integrated we obtain,

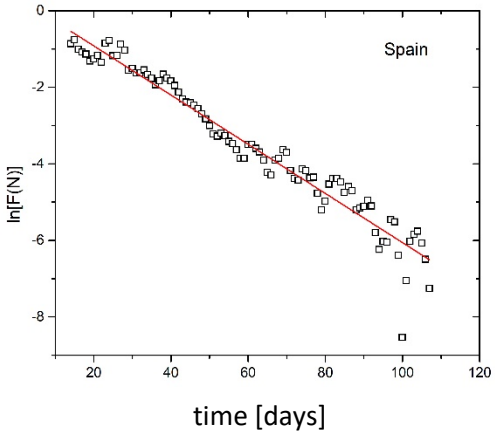
$$\ln N = \ln N_E + (\frac{1}{b})e^{(a+bt)} \tag{7}$$

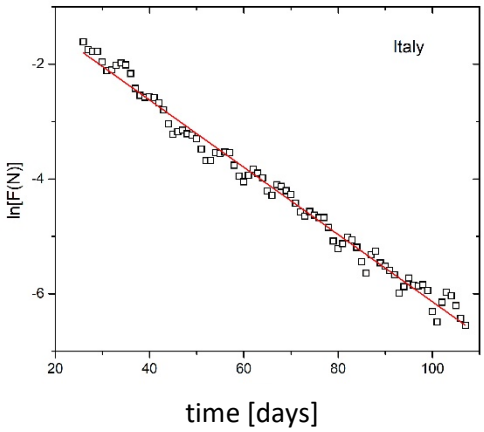
where *a* and *b* are constants.

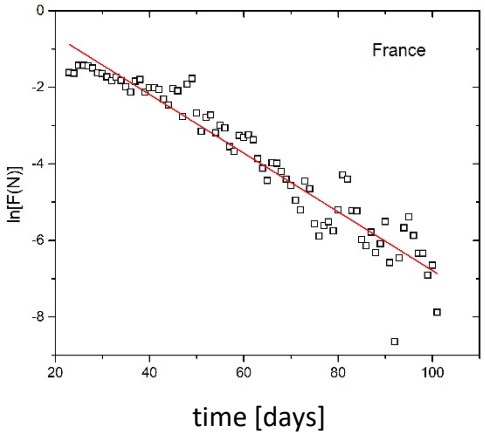

Global
ln[F(N)]
time [days]

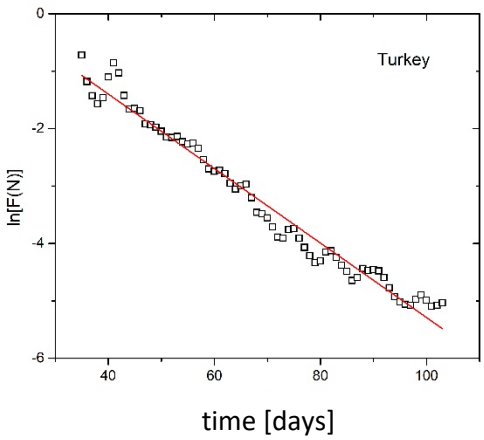

US
ln[F(N)]
time [days]

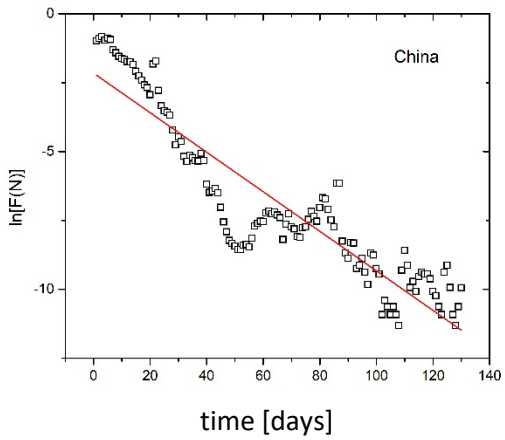

New York
ln[F(N)]
time [days]

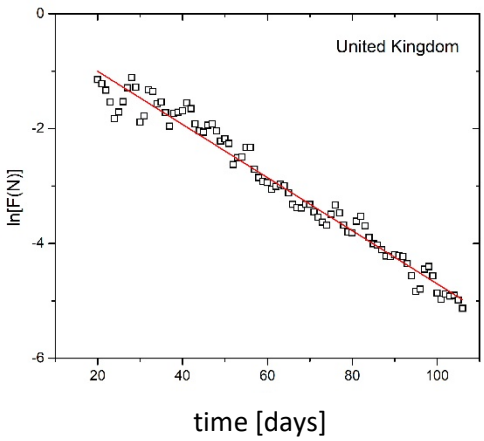

Spain
ln[F(N)]
time [days]

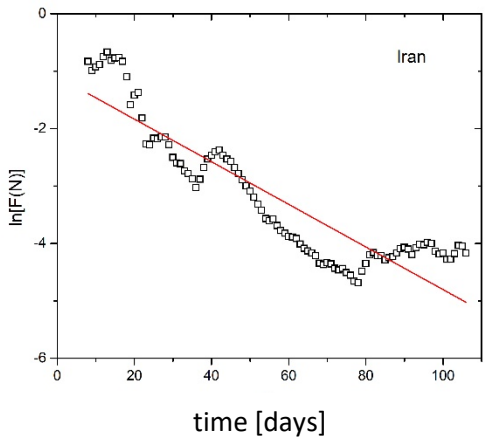

Italy
ln[F(N)]
time [days]

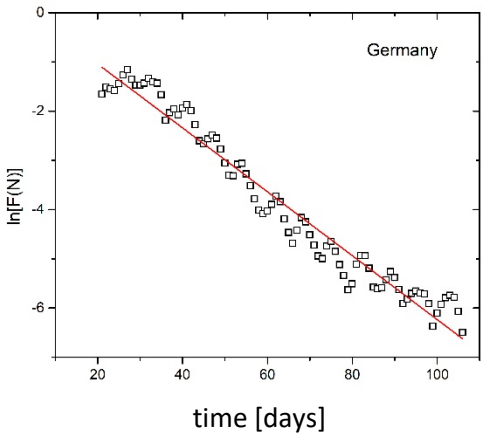

France
ln[F(N)]
time [days]

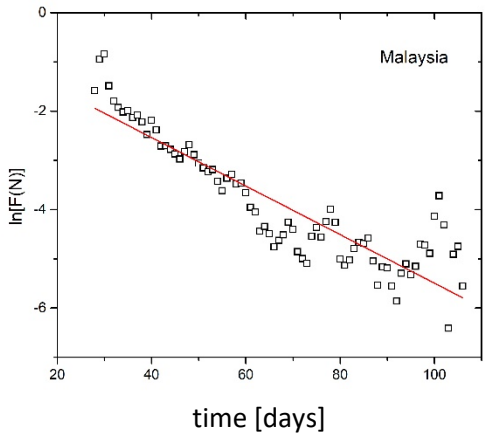

Turkey
ln[F(N)]
time [days]


China
ln[F(N)]
time [days]


United Kingdom
ln[F(N)]
time [days]


Iran
ln[F(N)]
time [days]


Germany
ln[F(N)]
time [days]


Malaysia
ln[F(N)]
time [days]

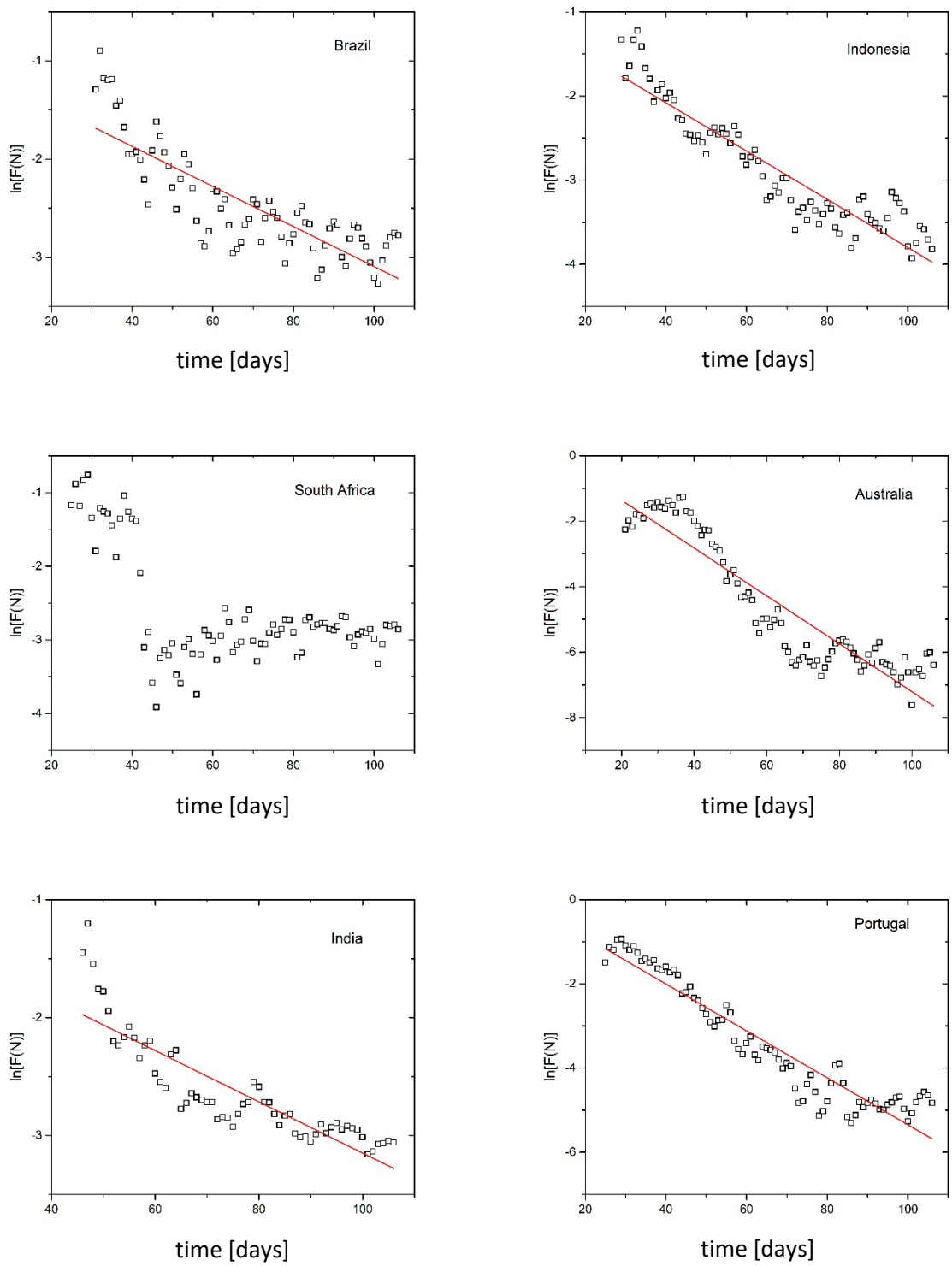


Fig.5: Plot of $\ln[F(N)]$ vs time.

Table (2): Values of the parameters *a* and *b* defined in eqn. (7) and value of the relaxation time b^{-1}.			
	a	*b*	b^{-1}(days)
Global	-2.17298	-0.01408	71.02273
US	0.03133	-0.0461	21.69197
New York	-0.75288	-0.06779	14.75144
Spain	0.36705	-0.06424	15.56663
Italy	-0.28148	-0.05857	17.07359
France	0.88323	-0.07671	13.03611
Turkey	1.19755	-0.06489	15.4107
China	-2.15384	-0.07175	13.93728
UK	-0.07713	-0.04627	21.61228
Iran	-1.09328	-0.03711	26.94691
Germany	0.25788	-0.06492	15.40357
Malaysia	-0.56168	-0.04934	20.26753
Brazil	-1.05468	-0.0204	49.01961
Indonesia	-0.93308	-0.02869	34.85535
South Africa	-	-	-
Australia	0.10604	-0.07313	13.67428
India	-0.97645	-0.02174	45.99816
Portugal	0.22618	-0.05572	17.94688

The values of the parameters *a* and *b* are given in Table .2, parameter *b* being always negative the equilibrium $N = N_E$ is approached asymptotically and the logarithm of the cumulative number of cases decaying with a relaxation time $\frac{1}{b}$ varying in the range $\sim 13 - 49$ days (Table 2). The value of $-\frac{1}{b}$ highest in Brazil and India, where the progress in achieving the equilibrium is slow and lowest in France, Australia and China. Since this is $\ln[N]$, the cumulative number of cases, approach the equilibrium very slowly. The first sketch of Figs. 2, 4 and 5 presents global data analysis. Here the value of $-\frac{1}{b}$ is ~ 71 days, indicating a general tendancy towards stabilization of the pandemic – the log scale cumulative numbers approaching a saturation value with relaxation time of 71 days.

Conclusion

The analysis suggests that fitting COVID-19 cumulative cases *N* to an equation of the form $\frac{dN}{dt} = N\,F(N)$ gives insight into the progression pandemic in different regions. As $\frac{dN}{dt}$ is essentially the number of cases reported each day, the function $F(N)$ can be easily plotted. Since $F(N) = \frac{d}{dt}[\ln(N)]$, this quantity can be readily ascertained from logarithmic plots of the cumulative number. The variation of $F(N)$ with N depicts the progression towards stabilization of the cumulative number. If $F(N)$ approaches N −axis with negative slope, indication is that the equilibrium point N_E within reach. This tendency is visible in majority of the curves presented - most probably a signature of success of the containment measures.

Forecasting the progress of the pandemic is possible by fitting data to determine the functional form of $F(N)$ and best fit was obtained for the four parameter dose-response logistic function

given by eqn.(4). Analysis enabled determination of saturation value of the cumulative number for most regions.

More information about the evolution of the pandemic could follow on updating of the plots. A limitation of the model is assumption of the equation (1) where $F(N)$ not an explicit function of t. Dynamical systems interpretations, apply only under the above assumption. Relaxation of control measures, mutations of the virus or influences of the climate may alter the observed trends $F(N)$ forecasted here.

Conflicts of Interest

Authors declare that there are no conflicts of interest.

Funding

This study is not funded by organization or an individual.

Disclaimer

Authors are not responsible for authenticity of data used, any errors that had in occurred in calculations or interpretations and making decisions based on this study.

Acknowledgement

Authors wish to thank Prof. D.Tantrigoda, University of Sri Jayewardhanapura**,** Sri Lanka for valuable discussions, constructive comments and reading the manuscript.

References

1. S. Ansumali, M. K. Prakash, A very Flat Peak: Exponential growth phase of COVID-19 is mostly followed by a prolonged linear growth phase, medRxiv preprint doi: https://doi.org/10.1101/2020.04.07.20055772

2. B. F. Maier, D. Brockmann, Effective containment explains subexpoential growth in recent confirmed COVID-19 cases in China, Science 27 (2020) 444. doi:10.1126/science.abb4557

3. C. P. Simon, J. A. Jacques, Reproduction number and stability of equllibria of SI models for Heterogeneous Populations, SIAM J. Appl. Maths. 52 (1992) 541. https://doi.org/10.1137/0152030

4. B. Pell, Y. Kuang, C. Vibound, G. Chowell, Using phenomenological models for forecasting the 2015 Ebola challenge. Epidemics 22(2018) 62. DOI: 10.1016/j.epidem.2016.11.002

5. K. Burghardt, C. Verzijl, J. Huang, M. Ingram, B. Song, M. Hasne, Testing Modeling Assumptions in the West Africa Ebola Outbreak, Sci.Rep. 6 (2016) 34598. DOI: 10.1038/srep34598

6. G. Chowell, C. Viboud, L. Simonsen, S. M. A. Vespignam, Perspectives on model forecasts of 2014-2015 Ebola epidemic in West Africa:l essons and the way forward. BMC Medicine 15 (2017) 42. https://doi.org/10.1186/s12916-017-0811-y

7. E. Buyuktahtakin, E. des Bordes, E.Y. Kibis, A new epidemics-logistic model: Insights into controlling Ebola virus disease in West Africa, Eur.J. Oper. Res. 265 (2018) 1046. DOI: 10.1016/j.ejor.2017.08.037

8. All Data Retrieved from JHU Data Repository: *https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv* *(Last access June/03/2020)*

9. A. F. Brouwer, M.H. Weler,M.C. Eisenberg,R.Meza, J.N.S. Eisenberg, PLOS Computational Biology, (2017) 1-28. https://doi.org/10.1371/journal.pcbi.1005481

10. W. W. Focke, I vader Westhuzen,N. Musee, M.Theodor, Kinetic interpretation of log-logistic dose-time response curves, Scientific Reports 7 (2017) 2234. https://doi.org/10.1038/s41598-017-02474-w