NOT PEER-REVIEWED

COVID-19: Finding the End Day

Dr. Sandip Chatterjee

Economist, Allahabad Bank, Rtd.

Visiting Faculty, Bengal Institute of Technology and Management, India

Email: *************, Mobile No.: 9748034344

ABSTRACTS

The study has pivoted on finding a methodology to forecast the end day of the menace of Coronavirus Disease of 2019 (COVID-19) or such pandemic that the planet faces on and often, challenging the core of the civilization. This model has resort to an indirect method to find the end day. As the pandemic grows exponentially, the rate of growth of total cases over previous day reduces asymptotically with herd immunity gaining strength to strength. Instead of finding flat head of the exponential expansion path, the model has looked into close to zero value of daily growth rate to find the end day. ARIMA (p,q,r) model for data smoothing and exponential trend line methodology adopted to find the end day. COVID-19 data for 63 days from March 20, 2020 to May 21, 2020 for seven countries and the globe explored with the proposed methodology. The study has projected toll of COVID-19 using a continuous constant exponential growth/decay model. The end day of the pandemic is projected for the globe when the expansion of the disease would be 0.01% per day. The methodology can be improved further by inclusion of other parameters of social and virology implications.

Keywords: ARIMA(p,q,r), Exponential, Forecast, Trend

1. Introduction

A well-known riddle is gossiped as "A blind-folded blind man is asked to search a black cat in a dark room which is not there". Probably, any effort to work on the COVID-19 reminds us the riddle albeit the cat is very much there. More so, when COVID-19 is rocking the globe, leaving none, not scared of life and livelihood. Lock down has locked up billions of people of the globe without the facility even a convict of capital punishment gets in a jail. This prompted us to take an effort to count the end day based on the relative data [1].

The model is rudimentary as no knowledge of virology, protein science, genetic or chemical biology whatsoever is used in formulation of the statistical forecasting. The issues of lock down and social distancing suggested by scientists and global organizations are also kept away from this model [2]. India, the second largest populous country, known in the world for its inadequate healthcare infrastructure indicated by 0.7 hospital beds, 0.6 doctors, per 1000 people. She resorted to social distancing measures to combat COVID-19 by slamming stringent lockdown and strict quarantine measures in the entire country at an early stage, in addition to wide social awareness initiatives for hand washing with soap and/or to sterilize with alcohol-based sterilizers [3]. Albeit these measures have widely been appreciated by international organizations and nations, any index of measuring

these parameters is expected to suffer from subjectivity due to lack of information and proper bench marking on social distancing and lock down. We have assumed no silver lining of discovery of medicine and/or vaccine to fight Civid-19 till the end day. Notably, in the case with SARS, vaccines were never used as its spread stopped without the need to vaccinate people [4].

The virologists and related scientists are pulling through bewildering times as the economists do amidst economic or financial crises. A group of scientists is tormenting to find out medicine while the other is dogged in finding a vaccine. The third ones are pushing hard to prove that the virus is not made in laboratory. This reminds us the riddle again. Nevertheless, there are some social parasites sneezing conspiracy theories to explore the guilty, interrupting their honey moon days all on a sudden.

The two strong assumptions mentioned above are made, based on the conjecture that by the end day, the globe would calibrate herd immunity to combat the virus. Moreover, herd immunity has also been used for the eradication of many diseases as it provides the basis for vaccines and their applications. [5] The novelty of this model is to resort to an indirect method to find the end day. As the pandemic grows exponentially, singing the triumphing song of *veni vidi vici*, the rate of growth of total cases over previous day reduces asymptotically with herd immunity gaining strength to strength. Instead of finding flat head of the exponential curve, the model has looked into close to zero value of daily growth rate of the disease which is asymptotic to horizontal axis to find the end day.

The data brought out in [1] is used for analysis. In addition to the globe as a whole, the worst affected USA and Canada of American (North) continent, two major victims in the continent of Europe, Spain and Italy; Peru from Latin America and a couple of major Asian prey to COVID-19, Iran and India are considered for the empirical analysis. The historical data-for the period of 20th March, 2020 to 21st May, 2020 is considered for the study.

2   The Methodology

2.1   *The Infection*

The methodology is elaborated with the help of appended Figure 1. Considering an exponential path Aa of COVID-19 spread, with a daily basis growth rate of Bb, asymptotic to X axis, in order to estimate number of days a value on Bb close to zero at point t is considered to determine number of days as Ot. The number of victims is arrived at, considering a constant exponential rate till the end day. When Bb is close to zero at t, Aa becomes flat close to point a. Daily growth rate of COVID-19 being close to zero, we consider the expansion path to become flat, signifying no further major addition of virus infection and cease of the pandemic at the end day.
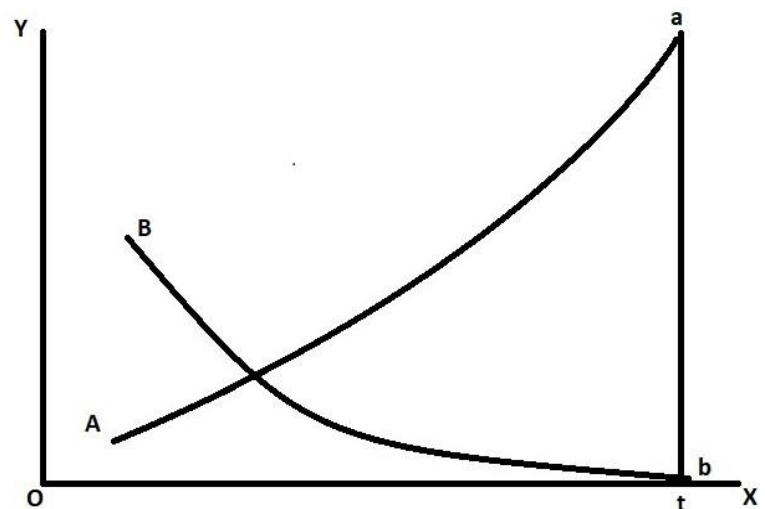

Figure 1: Exponential COVID curve and Asymptotic Daily Growth curve

The daily growth rate of COVID-19 plotted to find best fit trend lines which emerged to be exponential one for all countries (USA, Canada, Spain, Italy, Peru, Iran & India) and the globe. Among others, social issues play a crucial role in pace of spread of the pandemic. To elaborate, there are conflicting issues in India, namely, much talked religious congregation at Nijamuddin, Delhi, congregation of migrated labourers at Mumbai on the one hand and early lock down on the other. The exponential trend lines of daily growth rate of total cases for the seven countries and the globe can be represented as;

$$Y(T) = Ae^{BT} \qquad (1) \quad (A \neq 0)$$

Where Y (T) = value at time "T", A= Initial value, B = rate of growth (when >0) or decay (when <0). Thus, the theme of the model evolves in finding a close to zero daily rate of growth of COVID-19 (this is synonymous to insignificant number of daily increases of infected persons), signifying the end day and calibration of herd immunity to fight the disease.

2.2   *The Toll*

The total toll could have been determined following a similar methodology as in the total number of COVID-19 infected people. However, this is considered redundant in view of strong correlation between total number of infected people and total number of deaths. Thus, we considered a strong relation between number of total cases and toll.

The value of the share (%) of death in total cases at the end day is determined using a continuous exponential growth/decay function as under;

$$S_T = S_0 \, e^{\,kT} \qquad (k > 0 \text{ if growth, else } k < 0)$$

Where,

S_T = ending value of share (share at the end day after further decline)

S_0 = initial value of share (last observed value, the base measuring the decline)

k = continuous growth/decay rate (constant of proportionality)

T = time required to reach end day,

The value of k can be determined from the observations for known value of S_T, S_0 and T.

3   The Crux

The ravage of viruses has challenged human civilization time and again. The toll of Spanish Flu pandemic in 1918-19 accounted 40-50 million lives in a year which the present COVID-19 is hustling at [6]. Often China has emerged as starting place of new viruses, attributed mainly to her food habit and life style. Seemingly, like all other guests, India hosts viruses generously. In last couple of centuries, India has pulled through the devastation of numerous deadly viruses' notable being, Plague, Cholera, Spanish Flu, Influenza etc. which has also figured in Indian literature. These literatures suggest that social distancing used to be followed in those days as well. Presumably, the practice of social distancing became implicit in Indian culture, paving way to even various social taboos. Historically, it is observed that the maiden attack of most of the viruses has caused havoc to the human civilization, reason being everything unknown, reminding the riddle again. COVID-19 studies admit significant mutation in the virus without concluding impact of such mutations on virulence, transmissibility and mortality. Mostly, it is the rising herd immunity that made differences. Thus, this model has stressed more on herd immunity, assuming neither vaccine nor medicine till the end day of COVID-19 disease.

4 The Analysis

The data on key parameters for the countries during the study period is appended in Table 1.

Table 1: The Basic Data

Particulars	As on 20-03-2020			As on 21-05-2020			World Share (%)		CADG (Ratio)	
Country	Victim (V)	Toll (D)	D/V (%)	Victim (V)	Toll (D)	D/V (%)	Victim	Death	Victim	Death
USA	19551	309	1.58	1620902	96354	5.94	31.23	28.83	1.14	1.69
CANADA	1037	12	1.16	81324	6152	7.56	1.57	1.84	1.12	1.94
SPAIN	21571	1093	5.07	280117	27940	9.97	5.40	8.36	0.56	0.75
ITALY	47021	4032	8.57	228006	32486	14.25	4.39	9.72	0.31	0.43
PERU	263	4	1.52	108769	3148	2.89	2.10	0.94	1.83	2.16
IRAN	19644	1433	7.29	129341	7249	5.60	2.49	2.17	0.38	0.32
INDIA	249	5	2.01	118226	3584	3.03	2.28	1.07	1.90	2.11
GLOBE	**275734**	**11457**	**4.16**	**5190496**	**334173**	**6.44**			**0.66**	**0.79**

The studied seven countries share more than half of the COVID-19 cases and its toll as on 21-05-2020. The total data points numbered 63 for the period 20-03-2020 to 21-05-2020 under the study.

### 4.1 *Validity of the Data*

The day to day exponential rise in COVID-19 victims across the globe presumably suffers from serial correlation, requiring data smoothing before formulation of forecasting model [7]. On this endeavour, we have explored customary statistical models to find the best suited one for data adequacy of each country and the globe. Auto Regressive Integrated Moving Average, ARIMA (p,q,r) model were selected being adequate for the data of the respective countries [8]. The key parameters of the resulted model from the comparative analysis are presented in Table 2.

Nation	Parameter	Estimate	Stnd. Error	t	DW TEST	RMSE	RUNS	RUNM	AUTO	MEAN	VAR
					TABLE - 2 : ARIMA MODEL, PARAMETERS & STATISTICS						
USA	AR(1)	1.30	0.04	29.48	1.95	2941.88	OK	OK	OK	OK	OK
	AR(2)	-0.96	0.04	-21.91		MODEL ARIMA (2,2,2)					
	MA(1)	1.49	0.01	187.52							
	MA(2)	-1.02	0.02	-57.42		Significant at 95% confidence level.					
CANADA					2.09	268.69	OK	OK	OK	OK	OK
	MA(1)	0.55	0.11	5.15		MODEL ARIMA (0,2,1)					
						Significant at 95% confidence level.					
SPAIN					1.95	1161.08	OK	OK	**	OK	OK
	AR(1)	-0.36	0.12	-3.03		MODEL ARIMA (1,2,0)					
						Significant at 95% confidence level.					
ITALY	AR(1)	1.23	0.08	15.27	2.27	427.607	OK	OK	*	OK	OK
	AR(2)	-0.85	0.07	-11.79		MODEL ARIMA (2,2,2)					
	MA(1)	1.46	0.11	13.81							
	MA(2)	-0.84	0.09	-8.92		Significant at 95% confidence level.					
PERU	AR(1)	1.03	0.01	128.85	1.99	694.515	OK	OK	OK	OK	***
	MA(1)	0.83	0.07	12.23		MODEL ARIMA (1,1,1)					
						Significant at 95% confidence level.					
IRAN	MA(1)	0.27	0.12	2.28	1.92	225.949	OK	OK	OK	OK	OK
	MA(2)	-0.47	0.12	-3.98		MODEL ARIMA (0,2,2)					
						Significant at 95% confidence level.					
INDIA	AR(1)	2.10	0.03	68.93	2.10	363.199	OK	OK	OK	OK	***
	AR(2)	-1.11	0.03	-34.46		MODEL ARIMA (1,1,1)					
	MA(1)	0.31	0.14	2.25		Significant at 95% confidence level.					
GLOBE	AR(1)	1.24	0.05	26.26	2.06	6674.36	OK	OK	OK	OK	OK
	AR(2)	-0.95	0.05	-20.83		MODEL ARIMA (2,2,2)					
	MA(1)	1.39	0.08	17.53							
	MA(2)	-0.90	0.08	-11.38		Significant at 95% confidence level.					
	Note :	OK - p >= 0.05		* - 0.01 < p <= 0.05,		** - 0.001 < p <= 0.01		*** - p <= 0.001			

Although we started with 63 data points, in the process of moving average in ARIMA (2,2,2) two periods are lost. Further in process of finding daily growth rate over previous day on estimated COVID-19 cases, we lost another period, leaving us with 60 data points. Thus, the daily growth rates calculated on the basis of estimated values of COVID-19 cases made adequate data for further analysis.

With the smoothen data on COVID-19 victims, we worked out daily growth rate of the disease to construct an exponential trend line asymptotic to X axis to find a daily growth rate close to zero to indicate the end day. The details of the exponential trend line model are presented below. The model depicts adequate with significant parameters and statistics. The trend lines are presented in two groups, Graph – 1 captured USA, Canada, Spain and Italy, while, Graph – 2 included trend lines of Peru, Iran, India and the Globe.
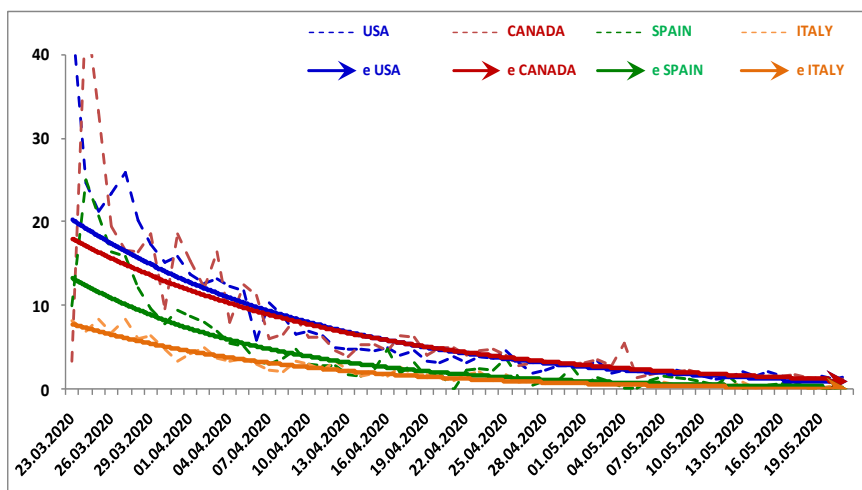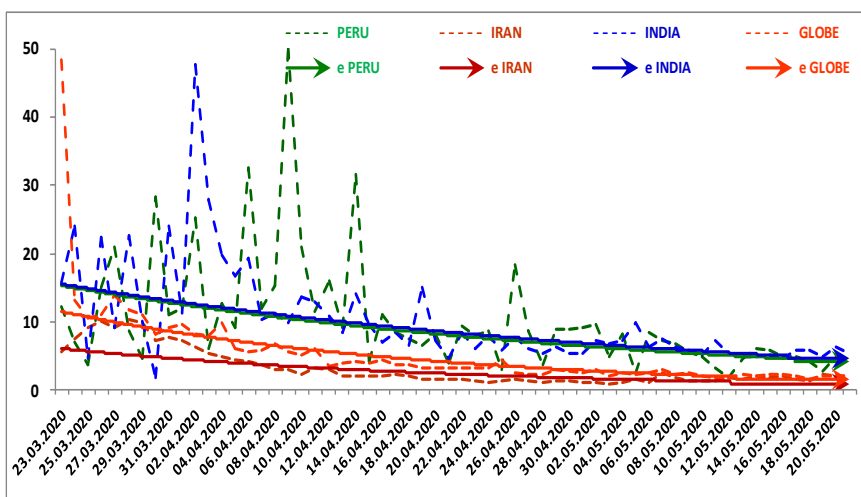

Graph 1: Trend Lines COVID-19 Daily Growth Rates USA, CANADA, SPAIN and ITALY

4.2   *The Spread of the Pandemic and Trend Lines*

The scatter and exponential trend line of daily growth rate of COVID-19 total victims for the seven countries and the Globe and USA are shown in Graph - 1 and Graph - 2, for the period 23rd March, 2020 to 21st May, 2020 i.e. for 60 days. The variation in the trend lines may be attributed to country specific reasons including social factors which are out of the purview of this study.

Graph 2: Trend Lines COVID-19 Daily Growth Rates PERU, IRAN, INDIA and THE GLOBE

In Graph – 1, the trend lines of Italy and Spain have neared the horizontal axis indicating closer to zero value of the daily growth rates of COVID-19 expansion. In Graph -2 trend line of Iran remained closer to the Horizontal axis.

The best suited ARIMA (p,q,r) model for each country facilitated estimated daily COVID-19 victims which are used in formulation of the trend line models for the countries and the Globe to find the end day. The trend line models along with the estimates and statistics are appended Table - 3. The exponential trend line equations are asymptotic to the horizontal axis that facilitates to find a time when the growth rate of the spread is close to zero.

Table – 3 : Trend Lines and Estimates

Parameter	Exp. Equation	R Squared	Constant	Slope
USA	$y_t = 21.47e^{-0.05t}$	**0.92**	21.47	-0.05
			(43.29)	*(-25.57)*
CANADA	$y_t = 18.90e^{-0.05t}$	**0.84**	18.90	-0.05
			(31.52)	*(-17.61)*
SPAIN	$y_t = 14.30e^{-0.07t}$	**0.69**	14.30	-0.07
			(12.70)	*(-11.45)*
ITALY	$y_t = 8.43e^{-0.06t}$	**0.92**	8.43	-0.06
			(26.97)	*(-26.34)*
PERU	$y_t = 15.67e^{-0.02t}$	**0.32**	15.67	-0.02
			(18.64)	*(-5.25)*
IRAN	$y_t = 6.23e^{-0.03t}$	**0.62**	6.23	-0.03
			(15.59)	*(-9.78)*
INDIA	$y_t = 15.87e^{-0.02t}$	**0.39**	15.87	-0.02
			(23.19)	*(-6.13)*
GLOBE	$y_t = 11.86e^{-0.04t}$	**0.83**	10.73	-0.03
			(43.88)	*(-21.85)*

Note : esimations significant at 95% confidence level

Italic figures in parentheses are t values of the statistic

For all the seven countries and the Globe, we have worked out the t when y_t is close to zero. This t indicates the end day of the pandemic, meaning thereby an insignificant increase in the number of COVID-19 victims. All the models and the statistics remained significant at 95% level. The R^2 value both for USA and Italy remained high at 0.92 followed by 0.84 and 0.83 for Canada and the globe. It is 0.69 and 0.62 for Spain and Iran while for India and Peru the figures are 0.39 and 0.32. Albeit it being low for India and Peru the models found to be significant.

5. The Forecast

5.1 *The End of the Pandemic*

Considering the trend line asymptotic to the horizontal axis, the end day is projected at the daily growth rate of 0.01% which the Globe reaches on 30th September 2020. Keeping in view, 'thus far no further', at the end day the figure for India and Peru is kept at 0.34%. The end day is projected to be nearest to Spain on 6th July followed by Italy on 14th July. The details are appended in Table – 4.

Table 4: Forecast – The End Day

Country/ Particulars	Trendline Equation	End Day	Victims on End Day *	Daily Growth	Addition (22-05-2020)	Days to Count (22-05-2020)
USA	$y_t = 21.47e^{-0.05t}$	22-08-2020	20,00,000	0.01	3,79,098	92
CANADA	$y_t = 18.90e^{-0.05t}$	02-09-2020	1,05,000	0.01	23,676	103
SPAIN	$y_t = 14.30e^{-0.07t}$	06-07-2020	2,90,000	0.01	9,883	45
ITALY	$y_t = 8.43e^{-0.06t}$	14-07-2020	2,40,000	0.01	11,994	53
PERU	$y_t = 15.67e^{-0.02t}$	30-09-2020	9,00,000	0.34	7,91,231	131
IRAN	$y_t = 6.23e^{-0.03t}$	30-09-2020	1,80,000	0.02	50,659	131
INDIA	$y_t = 15.87e^{-0.02t}$	30-09-2020	10,50,000	0.34	9,31,774	131
GLOBE	$y_t = 11.86e^{-0.04t}$	30-09-2020	75,00,000	0.01	23,09,504	131
			* rounded off			

Total global victim of COVID-19 is projected at 75,00,000.

5.2   *The Death Toll*

Toll of the pandemic is estimated based on its share in total number of victims. Correlation coefficient between victim and death is presented in Table – 5.

Table 5: Correlation: Victim & Death

Globe and Country-wise correlation coefficient r_{VD} suggests strong association between total COVID-19 cases and number of total deaths for the period of the study. However, the degree of association plummets when the number of daily rise in cases and death are considered for the entire period across the countries. Considering the last 15 days of the study period it is observed that the correlation coefficients r_{VDG-15} has further declined across the globe and the countries.

Country/r	r_{VD}	r_{VDG}	r_{VDG-15}
USA	1.00	0.76	0.67
CANADA	0.98	0.60	0.48
SPAIN	1.00	0.90	0.65
ITALY	1.00	0.93	0.67
PERU	1.00	0.92	0.56
IRAN	0.99	0.52	0.42
INDIA	1.00	0.94	0.73
GLOBE	1.00	0.61	0.62

This phenomenon suggests that the death toll for the end day based on current period share would not be justifiable. However, considering high value of r_{VD} for the entire period of the study we can presume the share of death toll in victims also follows a continuous exponential growth/decay relation.

Considering a continuous exponential growth/decay function for the share (%) of death toll in total cases, we have estimated the death toll at the end day (detailed in Table 6). The continuous growth/decay factor k_i is calculated on the basis of average share of total victims and deaths during the penultimate 15 days (S_0) and last 15 days (S_T). Thus, growth/decay factor for a country worked out as;

$$k_i = (Log_e \, S_T – Log_e \, S_0)/t \text{ Where, } (S_0) \text{ \& } (S_T) \text{ derived as above; } T = 30.$$

While forecasting the share (%) of death toll in total cases at the end day (S_T^E), we have taken into account T as the estimated days to count the end day. Thus,

$$S_T^E = S_T \, e^{\,k_iT} \quad (k_i > 0)$$

The death toll as on the end day is figured out based on the forecasted total victims and share (%).

Table - 6: Forecast of Death Toll

Particulars/Country	USA	CANADA	SPAIN	ITALY	PERU	IRAN	INDIA	GLOBE
Avg. Victim Penaltimate 15 Days	1067500	51957	234505	202678	35184	93524	34668	3229974
Avg. Toll Penaltimate 15 Days	61356	3064	24049	27568	982	5931	1137	226915
Share of Toll to Victms (S0)	5.75	5.90	10.26	13.60	2.79	6.34	3.28	7.03
Avg. Victim Ultimate 15 Days	1457733	73390	271403	222576	81533	115171	84021	4534624
Avg. Toll Ultimate 15 Days	86673	5362	27170	31301	2329	6850	2681	302420
Share of Toll to Victms (ST)	5.95	7.31	10.01	14.06	2.86	5.95	3.19	6.67
ki (Growth/Decay Factor)	0.0005	0.0031	-0.0003	0.0005	0.0003	-0.0009	-0.0004	-0.0008
Esimated Days to reach End Day (T)	92	103	45	53	131	131	131	131
Projected Toll Share on End Day (STE)	6.22	10.06	9.85	14.43	2.99	5.27	3.03	6.04
Estmated Deaths as on End Day	**124405**	**10560**	**28579**	**34625**	**26883**	**9480**	**31839**	**453194**
Addition in Toll till End Day	28051	4408	639	2139	23735	2231	28255	119021

The country-wise projected toll presented in Table – 6 with global toll at about 4.55 lacs.

6   Conclusion

This study has pivoted on finding a methodology to forecast the end day of the menace of COVID-19 disease in particular and such pandemic in general which the planet faces on and often challenging the core of this civilization. Expansion of COVID-19 across all the 110 countries and territories in a quarter of a year has qualified the era to globalization of viruses as well. With a projection of 7.5 million victims and close to half a million tolls, all the 7.8 billion population of the globe got affected both economically and psychologically. The end day is forecasted to come after counting four months and a half more, being scared of life and livelihood. The methodology can be fine-tuned further as we go ahead. Factors like, the social customs and culture, responsive and proactive authority, environment and geographical features, vital statistics and population density, virology & biological parameters etc. can enhance robustness of the methodology. Future research may also include a second surge to be more precise in finding the end day. COVID-19, be it manmade in origin and/or deliberate in spread going with conspiracy theories or naturally originated and/or inadvertently spread, has presaged the globe that the evils of globalization like this pandemic potentially can lead us to a cruel destruction of the civilization if we discount the future.

7   Declarations

7.1   Acknowledgements

7.2   Competing Interests

The author declares that no competing interests exist.

How to Cite:

Sandip Chatterjee. "COVID-19: Finding the End Day". *AIJR Preprints,* 89, version 1, 2020. https://preprints.aijr.org/index.php/ap/preprint/view/89

References

[1]   COVID-19 CORONAVIRUS PANDEMIC, https://www.worldometers.info/coronavirus/,

[2]   UNICEF, Effects of Social Distancing Measures, https://www.unicef.org/innovation/magicbox/covid

[3]   Paital B, Das K, Parida SK, Inter nation social lockdown versus medical care against COVID-19, a mild environmental insight with special reference to India, Science of The Total Environment Volume 728, 1 August 2020, 138914, https://www.sciencedirect.com/science/article/pii/ S0048969720324311

[4]   Petropoulos F, Makridakis S (2020), Forecasting the novel coronavirus COVID-19. PLoS ONE 15(3): e0231236. https://doi.org/10.1371/journal .pone. 0231236

[5]   Suárez, VJC, Hormeño-Holgado AJ, Jiménez M, Tornero-Aguilera, JF, Dynamics of Population Immunity Due to the Herd Effect in the COVID-19 Pandemic, Vaccines 8(2):236 May 2020. https://www.researchgate.net/publication/341481389_Dynamics_of_Population_Immunity_Due_to_the_Herd_Effect_in_the_COVID-19_Pandemic.

[6]   Wikipedia 1918 flu pandemic in India, https://en.wikipedia.org/wiki/1918_flu_pandemic_in_India,

[7]   Taylor JW, Exponential smoothing with a damped multiplicative trend. International Journal of Forecasting, 2003;19(4):715–725, https://www.sciencedirect.com/science/article/abs/pii/ S0169207003000037

[8]   Fattah J, Ezzine L, Aman Z, Moussami H E, Forecasting of demand using ARIMA model, International Journal of Engineering Business Management 10(2):184797901880867· October 2018.