

Research Article

Open Access



- ¹ Department of Spatial Epidemiology, ICAR- National Institute of Veterinary Epidemiology and Disease Informatics, Bengaluru, Karnataka, India.
- ² Virology Lab, ICAR- National Institute of Veterinary Epidemiology and Disease Informatics, Bengaluru, Karnataka, India.
- ³ Chanre Rheumatology and Immunology Centre and Research, Bengaluru, Karnataka, India.

Contacts of Authors



* To whom correspondence should be addressed: Bharath P. C. Thyagaraju

Citation: Bharath Prasad Cholanayakanahalli Thyagaraju, Srikantha Gowda, Sharanagouda Patil, Chandrashekar Srikantiah, and Kuralayanapalya Puttahonnappa Suresh (2020). Future trends of COVID-19 disease outbreak in different states in India: a periodic regression analysis. Highlights in BioScience Volume 3. Article ID 20224. doi:10.36462/H.BioSci.20224

Received: August 4, 2020

Accepted: September 7, 2020

Published: October 1, 2020

Copyright: © 2020 Thyagaraju *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and supplementary materials.

Funding: The authors have no support or funding to report.

Competing interests: The authors declare that they have no competing interests.

Future trends of COVID-19 disease outbreak in different states in India: a periodic regression analysis

Bharath Prasad Cholanayakanahalli Thyagaraju^{1,*}, Srikantha Gowda¹, Sharanagouda Patil², Chandrashekar Srikantiah³, and Kuralayanapalya Puttahonnappa Suresh¹

Abstract

COVID-19 (Coronavirus disease 19) is the deadliest pandemic, and by August 2, >18.2 million population worldwide were infected with SARS-CoV-2 virus causing burden on human life and economic loss. Disease outbreak analysis has become a priority for the Indian government to initiate necessary healthcare measures in lowering the impact of this deadly pandemic viral disease. In this study, time series data for COVID-19 disease was extracted from the website www.covid19india.org, analysed by using periodic regression model, the expected number of cases till 02 October 2020 was predicted and to develop a stochastic models using periodic regression in the top 15 highly infected states in India. The analysis reported increasing pattern at initial days of prediction and showed a decreasing trend for the number of reporting cases, which may reduce in future days for states like West Bengal, Karnataka, Uttar Pradesh, Bihar, Telangana, Assam and Odisha. However, for the states of Maharashtra, Tamil Nadu, Gujarat, Rajasthan, Haryana and Madhya Pradesh, showed a rapid phase of increase in disease outbreak that is likely to infect more population and indicates the pandemic nature of this disease over a period. Presently, Delhi shows a drastic reduction in the number of cases, that may increase in the future, which can be controlled if appropriate preventive measures are followed strictly and effectively. Our model highlights that continuous and constant efforts are needed for the prevention of new infections of the disease in all states that helps to effectively mitigate the disease and to allocate scarce resources effectively in the future that could improve the economic wealth in India.

Keywords: COVID-19, Disease, Outbreaks, Periodic regression, SARS-CoV-2.

Introduction

The novel corona virus (severe acute respiratory syndrome coronavirus 2 [SARS-CoV-2]), COVID-19 pandemic in India is part of the worldwide viral disease pandemic, has continued to spread across the world like a forest fire [1,2]. According to WHO reports, COVID-19 disease has affected more than 210 countries globally [1]. This virus has a high transmission rate due to absence of suitable drugs, the infectious COVID-19 disease has ruined the normal life of many people around the world. The first case of COVID-19 was reported in India on 30 January 2020, which basically originated from Wuhan city, China. At present, more than half a million populations are infected by corona virus, and have deceased > 6.92 lakh, and forced about 8 billion people to remain at their houses [3].

As on 02 August 2020, the Ministry of Health and Family Welfare (MoHFW) has confirmed a total number of 1,805,838 active cases, 1,188,389 recoveries that include 1 migration and 38,176 deaths in the country [4]. India, presently has the largest number of active cases for COVID-19 infections in Asia, [5] and is at the 3rd place after USA (United States of America) and Brazil, with the highest number of total confirmed cases [6]. The case mortality rate (CMR) in India is relatively much lower at 2.90%, when compared to the global CMR at 4.7%, as on 03 August 2020 [7]. Eleven states in India accounts for more than half of the total reported cases – Maharashtra, Tamil Nadu, Andhra Pradesh, Delhi, Karnataka, Uttar Pradesh, West Bengal, Telangana, Gujarat, Bihar and Rajasthan [8]. As on 10 June 2020, India's recovery rate for the first time exceeded the number of COVID-19 cases and showed 50% decrease in total infections [9]. It is noteworthy that by early July 2020, the recovery rate has crossed 60%. Although, active cases have continued to upsurge persistently.

Disease data are categorized by seasonal variations. Periodic or cyclic phenomena are mainly close to not only biological datasets, but also to non-biological datasets [10]. Several kinds of disease data tend to fluctuate at regular time intervals and also displays characteristic of periodic/cyclic nature [11]. The cyclical character of outbreaks in nature is exhibited in time and space. An earlier report indicated that the biological cycles may be subdivided into two groups, i.e., physical and non-physical environment [10]. In the first group, biological outbreaks are determined physiologically. Meanwhile, in the second group, cycles mainly depend on several other potential risk factors, which can modify or change the disease outbreaks.

Disease modeling and analysis of incidences will help in projecting the disease probability and guide in early preparedness of the disease by taking effective control measures and to develop appropriate preventive healthcare procedures [12]. Periodic regression curve corresponds to few variables in time and is frequently repeated at fixed time intervals for predicting the spreading of COVID-19 disease [13]. Well-tested and effective viral vaccines against this deadly virus have not been developed and henceforth, this study plays a key role in mitigating this pandemic and flattening this epidemic curve/peak. It will lead to a bigger picture in taking aggressive and timely control measures in improving infrastructure, service facilities, viral vaccines and effectively controlling related epidemics in future.

In our study, we estimated the present and future trend patterns of COVID-19 disease cases and develop a stochastic model using periodic regression approach for the next sixty days in 15 highly infected states in India until 02 October 2020.

Materials and Methods

Data source

The time series data on the total number of COVID-19 cases reported was extracted from cloud source database from the website www.covid19india.org and tabulated for further analysis [3]. The data for model development were updated until 02 August 2020.

Model Development and statistical analysis using Periodic Regression

The analysis and interpretation of periodicity, in the time-series data the important determinants were the dimension of the cycle or fundamental time period, its amplitude or the range from the minimum to the maximum response and angular point in time during the periodic cycle when the response is maximum [10]. These parameters were assessed by utilizing many statistical packages. Time series data or an outbreak Y_t ($t=1, \dots, N$) observed at equal intervals of time was shown as $Y_t = Y_t + \varepsilon_t$, where Y_t is unobserved fixed value at time $\{t\}$ and $\{\varepsilon_t\}$ is a sequence of random errors identical and independently distributed with expectation 0 and variance σ^2 .

To discover variability of the disease outbreak data whether it has periodic/cyclic constituents, the series was approximated by finite Fourier series of the form, if the number of data was even or odd [10].

We computed the future trend for COVID-19 disease outbreaks for the next ninety days using periodic regression model. Statistical analysis was carried out by using R-software version 3.6.3 CRAN (Comprehensive R Archive Network) and regression curves were obtained.

Results

The detailed periodic regression analysis was performed for the top 15 Indian states, which accounts for 93% of total COVID-19 cases reported in India and the probable number of cases for the period of the next sixty days was predicted and represented (**Table 1**). Most of the states in India showed the significance for intercept, X value, sinus and cosinus values and are given in **Table 1**. COVID-19 disease data were analyzed to showcase future trends of corona disease infections for the period of two months (until 02 October 2020). The periodic regression analysis curve showed the baseline of outbreaks, upper bound line which was 95% confidence interval from the baseline and the observed line indicated actual coronavirus infections that are likely to occur during the next sixty days in 15 states in India.

The present and future COVID-19 disease curves for highly infected top 15 states in India are represented in **Figure 1**. The trend analysis of Maharashtra, Tamil Nadu, Gujarat, Rajasthan, Haryana and Madhya Pradesh (**Figure 1**) states revealed that the COVID-19 infections are increasing at a rapid phase. Currently, it is observed that the outbreaks curve shows that the disease infection has crossed the upper bound line and indicates the pandemic nature of this disease over a period of time. This trend pattern shows that the coronavirus disease may reach critical conditions in the upcoming days if strict measures are not imposed to control this pandemic in states listed above.

The trend analysis of West Bengal, Telangana Assam, Karnataka and Odisha states revealed that COVID-19 disease infections will increase till mid of September and decreases slowly by late September and early October. The observed trend line shows increasing and decreasing pattern over a period of time, it may tend to follow below the upper bound line and the number of cases reporting may reduce in future days if the control measures are followed effectively and strictly (**Figure 1**).

Table 1. Periodic regression analysis values of COVID-19 disease outbreaks in the top 15 highly infected states in India.

States	Parameter	Estimate	Standard deviation	t-value	p-value	R ²	Adj. R ²
Maharashtra	Intercept	-1524.25	140.41	-10.85	<0.001**	0.8998	0.8973
	x	1.90	1.90	31.50	<0.001**		
	c1	95.81	95.81	5.49	<0.001**		
	s1	99.99	99.99	2.55	0.012*		
Tamil Nadu	Intercept	-1033.35	117.42	-8.80	<0.001**	0.8346	0.8305
	x	1.61	1.61	22.51	<0.001**		
	c1	78.67	78.67	3.87	<0.001**		
	s1	84.45	84.45	-0.69	0.491 ^{ns}		
Andhra Pradesh	Intercept	436.968	81.22	-5.38	<0.001**	0.5553	0.5445
	x	1.11	1.11	10.84	<0.001**		
	c1	54.41	54.41	4.16	<0.001**		
	s1	58.42	58.42	1.83	0.069 ^{ns}		
Delhi	Intercept	-194.303	67.54	-2.87	0.005*	0.6877	0.6800
	x	0.85	0.85	16.42	<0.001**		
	c1	50.02	50.02	-1.25	0.213 ^{ns}		
	s1	53.50	53.50	-2.51	0.013*		
Karnataka	Intercept	-636.579	101.40	-6.27	<0.001**	0.6095	0.5999
	x	1.38	1.38	12.08	<0.001**		
	c1	68.51	68.51	4.99	<0.001**		
	s1	72.53	72.53	2.58	0.011*		
Uttar Pradesh	Intercept	-307.077	42.57	-7.21	<0.001**	0.7510	0.7449
	x	0.58	0.58	17.91	<0.001**		
	c1	28.66	28.66	3.72	<0.001**		
	s1	30.52	30.52	2.18	0.031 ^{ns}		
West Bengal	Intercept	-326.815	42.76	-7.64	<0.001**	0.7222	0.7154
	x	0.58	0.58	16.97	<0.001**		
	c1	28.82	28.82	3.28	0.001**		
	s1	30.65	30.65	3.59	<0.001**		
Telangana	Intercept	-324.613	52.32	-6.20	<0.001**	0.6516	0.6431
	x	0.69	0.69	14.13	<0.001**		
	c1	37.46	37.46	3.72	<0.001**		
	s1	37.01	37.01	-0.45	0.651 ^{ns}		
Gujarat	Intercept	-73.863	9.68	-7.62	<0.001**	0.9644	0.9635
	x	0.13	0.13	52.31	<0.001**		
	c1	6.55	6.55	8.75	<0.001**		
	s1	6.94	6.94	-1.83	0.069 ^{ns}		
Bihar	Intercept	-216.621	33.54	-6.45	<0.001**	0.6372	0.6284
	x	0.45	0.45	13.49	<0.001**		
	c1	22.77	22.77	3.86	<0.001**		
	s1	23.92	23.92	4.62	<0.001**		
Rajasthan	Intercept	-88.154	13.26	-6.64	<0.001**	0.8657	0.8625
	x	0.18	0.18	26.57	<0.001**		
	c1	8.88	8.88	4.03	<0.001**		
	s1	9.54	9.54	3.47	<0.001**		
Assam	Intercept	-222.315	29.43	-7.55	<0.001**	0.6844	0.6767
	x	0.40	0.40	15.30	<0.001**		
	c1	20.09	20.09	3.53	<0.001**		
	s1	21.01	21.01	2.98	0.003*		
Haryana	Intercept	-169.338	19.47	-8.69	<0.001**	0.8021	0.7973
	x	0.26	0.26	21.66	<0.001**		
	c1	13.31	13.31	-1.81	0.073 ^{ns}		
	s1	13.85	13.85	0.64	0.519 ^{ns}		
Odisha	Intercept	-135.761	17.29	-7.84	<0.001**	0.7444	0.7381
	x	0.23	0.23	17.35	<0.001**		
	c1	11.76	11.76	4.81	<0.001**		
	s1	12.34	12.34	2.61	0.010		
Madhya Pradesh	Intercept	-53.396	13.99	-3.81	<0.001**	0.7434	0.7371
	x	0.19	0.19	18.00	<0.001**		
	c1	9.54	9.54	3.66	<0.001**		
	s1	9.96	9.96	4.95	<0.001**		

Note: ns=Non-significant; *Significant level at 95% (P<0.05); **Significant level at 99% (P<0.01).

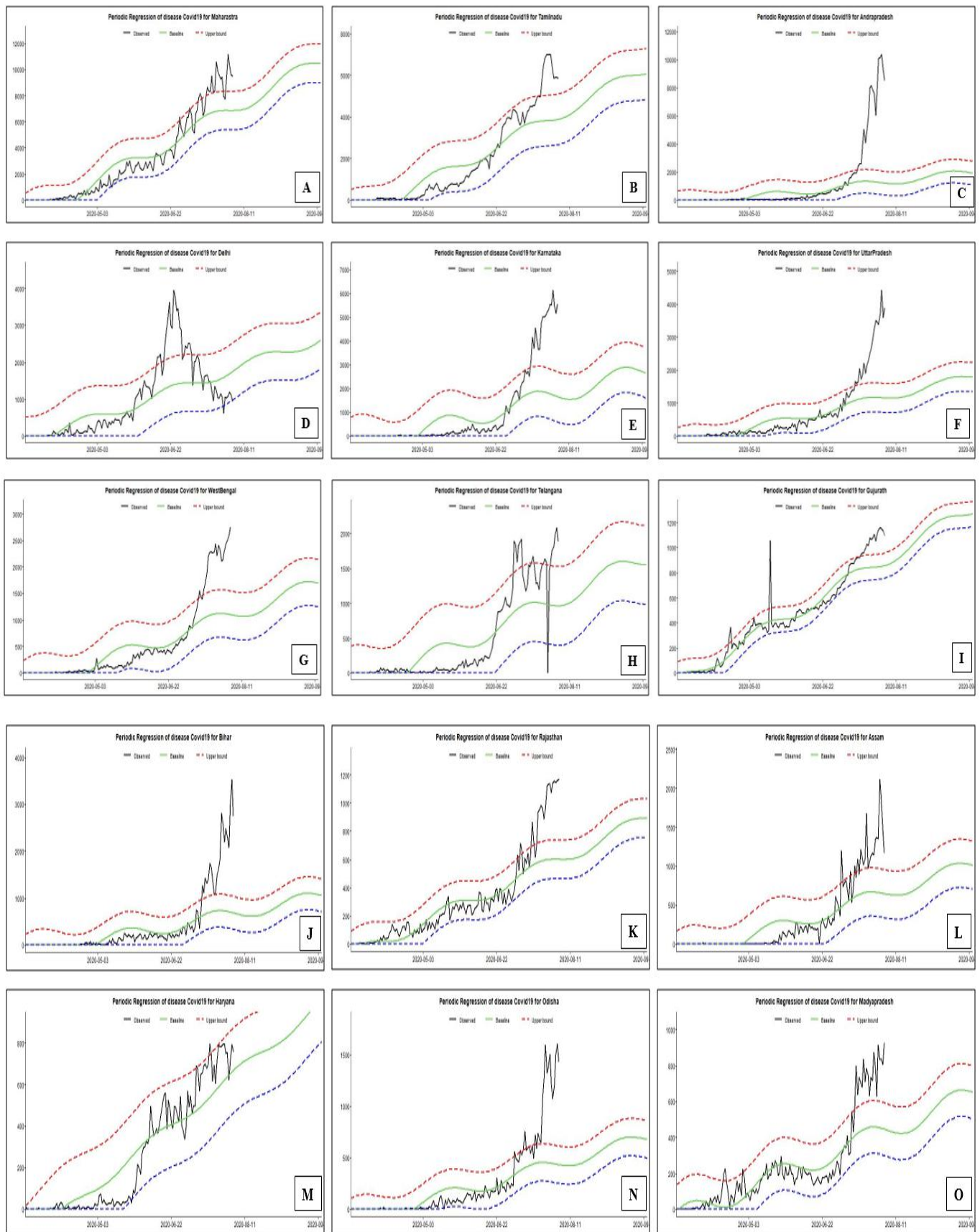


Figure 1: Periodic regression analysis of COVID-19 disease outbreaks in the top 15 highly infected states in India- Maharashtra (A), Tamil Nadu (B), Andhra Pradesh (D), Delhi (C), Karnataka (E), Uttar Pradesh (F), West Bengal (G), Telangana (I), Gujarat (H), Bihar (J), Rajasthan (K), Assam (L), Haryana (M), Odisha (N) and Madhya Pradesh (O).

The trend analysis of states like, Uttar Pradesh and Bihar revealed that COVID-19 disease shows a steady state in a number of cases reported till the end of prediction cycle (till 02 October 2020), the observed line may tend to follow the baseline and the overall infections may not be that severe in future.

Presently, the trend analysis of Delhi revealed that COVID-19 disease infection has drastically reduced. The observed curve shows that the infection rate has decreased over a time period, the trend line may tend to follow below the baseline and the number of cases reported is reduced that may be due to strictly following the rules imposed by GOI to effectively mitigate the disease. According to periodic regression analysis, the number of cases may increase anytime in future days between August and October if government relaxes the rules imposed on the movement of population and travelers from different states.

Discussion

Our current situation demands prediction of disease trend patterns to take appropriate healthcare interventions for controlling the disease in the future. The analysis predicts very frightening outcomes, which define worsen conditions in India, particularly in states of Maharashtra, Tamil Nadu, Karnataka, Andhra Pradesh, Uttar Pradesh, West Bengal, Telangana, Bihar, Haryana and Madhya Pradesh. Based on the predictions resulted of our study, public health officials shall tailor their aggressive planning and interventions to introduce rapid viral infection control measures at hospital levels as well as at community level to restrict the COVID-19 pandemic and were similar to study analysed at global level and extracted data upon Machine Learning approach using Artificial Intelligence techniques for the top 10 countries [14]. However, with limited resources available, globally COVID-19 infections present a significant challenge for modellers, government and civil societies to mitigate the effect of health crisis. In our study, we adopted the periodic regression model to determine the future trends of the disease for top 15 highly infected states in India. According to results of our study, it is expected that a greater number of inhabitants would be infected and become ill in future. Fortunately, India took strict control measures like enforced quarantine, lockdown, curfews and travel restrictions etc., to restrict the spread of infection at the early stage of a pandemic, but due to the relaxation of lockdown population starts to mix without proper care and may trigger a rapid increase in viral transmission and spread of disease.

The major limitations for the control of COVID-19 disease infection in developing country like India are poor vaccination cover, lack in financial support and insufficient medical infrastructure, which delays in building of herd immunity [15]. Some evidence advises that people who recover from COVID-19 will develop immunity to SARS-CoV-2 virus [16]. However, at this juncture the duration of susceptibility is unclear. If immunity remains longer than the disease outburst, then declining immunity will not disturb the dynamics of the epidemics. The current pattern shows that there will be a geometric progression in the next few days' even if government has taken stringent control measures and there are likely more chances of slipping into the exponential cases in the long run as per results of periodic regression. Hospitals, extension of

medical facility and testing capabilities, contact tracing of all suspected cases in all affected states is critical and should be continued at a very rapid pace to mitigate viral transmission. However, with the current planning and interventions, the GOI is looks forward to reduce/flatten the epidemic curve. For effective pattern of trend to be seen, the updated data at interval of fifteen days' must be used to carry out the analysis, as the peak trends to fluctuate at regular time intervals and also shows the characteristic of periodic nature. The curve may flatten in the future as number of COVID-19 disease cases may increase or decrease for each state according to the effective control measures followed by respective states.

Conclusion

Our study shows the future trends for COVID-19 disease infections using periodic regression model and indicates that there would be an exponential increase in the number of cases for the next few months (until 02 October 2020). In addition, the death rate (2.12%) is comparatively decreasing with respect to the number of positive corona cases, which is promising. There is a likely chance of exponential increase in the number of cases which might be due to the relaxation of lockdown rules, failed to track the COVID-19 cases as population starts mixing without proper social distancing and not wearing a face mask, increased number of shops, malls, markets opened and migration of populations across different states without proper safety measures. However, during 50 plus days of lockdown period, India has given a buffer time for the hospitals and healthcare organizations in each state to prepare themselves with appropriate healthcare measures if the pandemic explodes in future. Periodic regression is the best fitted model for any observed data, the slight difference in the future trend curve may arise due to intrinsic characteristics of any model and parametrization. While the Periodic regression can also be used for long-term prediction with non-linear cyclic trend pattern. Our findings are helpful for policy makers to take informed decisions, planners to improve the risk management, assist in developing preventive measures and allocation of scarce resources effectively in the future.

Acknowledgement

The authors are thankful to Deputy Director General, Indian Council of Agricultural Research (ICAR), New Delhi, and Director, ICAR-NIVEDI, Bengaluru and staff of Spatial Epidemiology lab for providing the necessary support, encouragement and valuable suggestions in writing this article.

References

1. Report of the WHO – world joint mission on Coronavirus Disease 2019 (COVID-19) [Internet]. World Health Organization; 2020. [cited, 2020 August 02]. Available from: <https://covid19.who.int/>.
2. Singh RK, Rani M, Bhagavathula AS, Sah R, Rodriguez-Morales AJ, Kalita H, et al. Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model. *JMIR Public Health and Surveillance*. 2020;6(2): e19115.

3. Data on number of Covid-19 cases reported in India [Internet]. A cloud source data; 2020. [cited 2020 August 02]. Available from: <https://www.covid19india.org/>.
4. Data on confirmed cases, recoveries and death of Covid-19 in India [Internet]. Ministry of Health and Family Welfare; New Delhi; 2020. [cited 2020 August 03]. Available from: <https://www.mohfw.gov.in>
5. India most infected by Covid-19 among Asian countries, leaves Turkey behind [Internet]. Hindustan Times, New Delhi; 2020 [cited 2020 May 30]. Available from: <https://www.hindustantimes.com/india-news/>.
6. Sagar K. India becomes third worst affected country by coronavirus, overtakes Russia [Internet]. Deccan Herald. New Delhi; 2020 [cited 2020 July 5]. Available from: <https://m.dailyhunt.in/news/india/english/deccan+herald-epaper-deccan/>.
7. Another grim milestone: India No.3 in total coronavirus cases [Internet]. Hindustan Times, New Delhi; 2020 [cited 2020 July 6]. Available from: <https://www.hindustantimes.com/india-news/>.
8. Infections over 1 lakh, five cities with half the cases: India's coronavirus story so far [Internet]. The Week; 2020 [cited 2020 May 20]. Available from: <https://www.theweek.in/news/india/>.
9. Kumar S. Covid-19: Number of recoveries exceed active cases for first time. [Internet]. Hindustan Times, New Delhi; 2020 [cited 2020 June 11]. Available from: <https://www.hindustantimes.com/india-news/>
10. Bliss CI. *Statistics in Biology*. McGraw-Hill Book Company, New York, USA, 1970.
11. Little TM, Hills FJ. *Agricultural experimentation: Design and Analysis*. John Wiley and Sons, Inc, New York, USA; 1978.
12. Yan P. Distribution theory, stochastic processes and infectious disease modelling. In *Mathematical epidemiology 2008* (pp. 229-293). Springer, Berlin, Heidelberg.
13. Krishnamoorthy P, Kurli R, Patil SS, Roy P, Suresh KP. Trends and future prediction of livestock diseases outbreaks by periodic regression analysis. *Indian Journal of Animal Sciences*. 2019; 89(4): 369–76.
14. Ye QH, Qin LX, Forgues M, He P, Kim JW, Peng AC, et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nature Medicine* 2003; 9(4): 416-23.
15. Mai MV, Krauthammer M. Controlling testing volume for respiratory viruses using machine learning and text mining. In: *AMIA annual symposium proceedings*, vol 2016. American Medical Informatics Association. 2016;p 1910
16. Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. *Clinical Infectious Diseases*. 2020;ciaa344.