The incubation period of coronavirus disease (COVID-19): A tremendous public health threat—Forecasting from publicly available case data in India

Rinshu Dwivedi¹ | Ramesh Athe² | Kavi Mahesh² | Praveen Kumar Modem²

¹Indian Institute of Information Technology, Trichy, Tamil Nadu, India
²Indian Institute of Information Technology, Dharwad, Karnataka, India

Correspondence
Ramesh Athe, Indian Institute of Information Technology, Dharwad, 580029, Karnataka, India.
Email: dr.athe9@gmail.com

The World Health Organization (WHO) declared the Coronavirus Disease (COVID-19) a pandemic due to the huge upsurge in the number of reported cases worldwide. The COVID-19 pandemic in India has become a public health threat, and if we go by the number of confirmed cases then the situation seems to be a matter of grave concern. According to real-time data, the numbers of confirmed cases are growing exponentially. No doubt, substantial public health interventions both at the national and state levels are implemented immediately by the Government of India; there is a need for improved preparedness plans and mitigation strategies along with accurate forecasting. The present study aims to forecast the COVID-19 outbreak infected cases in India. The data have been obtained from https://www.covid19india.org, https://www.worldometers.info/coronavirus, and ICMR reported publicly available information about COVID-19 confirmation cases. We have used the double exponential smoothing method for forecasting the trends in terms of confirmed, active, recovered and death cases from COVID-19 for emergency preparedness and future predictions. Findings reveal that the estimated value of point forecast is just 8.22% of the total number of confirmed cases reported on a daily basis across the country. It was observed that the deaths were lower for the states and union territories with a higher detection rate. It is suggested that by keeping in view the limited healthcare resources in the country, accurate forecasting, early detection, and avoidance of acute care for the majority of infected cases is indispensable.

KEYWORDS
coronavirus disease (COVID-19), emergency-preparedness, forecasting, India, public-health

INTRODUCTION

Worldwide, over the past two decades, a larger number of individuals and animals have been affected with three epidemics caused by the family of coronavirus (Severe Acute Respiratory Syndrome (SARS)-2003, Middle East Respiratory Syndrome (MERS)-2012, and Coronavirus Disease (COVID-19)) (Banik et al., 2015; Hilgenfeld & Peiris, 2013; Paraskevis et al., 2020). However, significant genetic dissimilarities have been documented between pathogens of these three epidemics, particularly amid MERS and COVID-19 (Dhama et al., 2020; Kolifarhood et al., 2020; WHO, 2020; Yang et al., 2020). Initially, the hotspots for these epidemics were Middle East, Saudi Arabia, and China (Sutton & Subbarao, 2015; Wu et al., 2017), as they are transmitted from animal to human, and later transmissions of pathogens were reported from human to human in other countries as well. Epidemiological evidence of COVID-19 outbreak began in Wuhan, China from December 12, 2019 (Wu & McGoogan, 2020). The WHO declared COVID-19 a pandemic on March 11, 2020 due to a huge upsurge in the number of reported cases for COVID-19 as compared to SARS and MERS.
The index case of the COVID-19 pandemic in India was reported from Thrissur, Kerala on January 30, 2020. Being one of the largest populated countries, and having resource constraints, the case detection for COVID-19 was initially low (Guo et al., 2020; Singhal, 2020). There were also notable variations in the predicted estimates of infection cases. India has observed a 14 h voluntary public curfew on March 22, 2020 followed by a lockdown in 75 districts where confirmed cases of COVID-19 have been reported. The country has introduced the first phase of complete lockdown for 21 days amidst rapidly surging confirmed cases emphasizing the sustained risk of further spread across the nation from March 24, 2020. The outbreak has been declared an epidemic among the majority of the Indian states and Union territories, under the Epidemic Diseases Act, 1897 (Dikid et al., 2013). All the tourist visas have been suspended, along with the closure of all educational institutions, modes of transportation, and a majority of commercial establishments have been shut down to avoid the spread. Various phases of lockdown have been rolled out to curb the spread of virus. As per the available data from the Ministry of Health and Family Welfare (MoHFW), the country has a total of 13,835 confirmed cases, along with 1767 recoveries and 452 deaths on April 17, 2020 (MoHFW COVID-19, n.d.).

According to real-time data, COVID-19 cases are growing exponentially. Healthcare system in India is embracing itself for may be its biggest battle to treat the infected cases. It is, therefore, important to forecast the precise increase in the expected number of cases to comprehend what is required, and to arrest the perturbing trends. Accurate forecasting dissemination can play a critical role in apprising the governments and healthcare professionals what to expect and which measures to impose to motivate the public to adhere with them. In this paper, we have estimated the possible forecasted cases across the country from the publicly available case data by using the double exponential smoothing (DES) method. We assume that the publicly available daily case data are legitimate subject to certain reporting biases if any. The objective is not to attempt for meticulous precision but to merely provide useful insights on COVID-19 in India.

1.1 Data and methods

The data were collected from the publically available secondary sources (https://www.covid19india.org; https://www.worldometers.info/coronavirus/; Ministry of Health and Family Welfare (MoHFW), and Indian Council of Medical Research (ICMR, n.d.) reports) on COVID-19 (2020) confirmed, active, recovered, and death cases. We assume on the basis of available literature that the mean duration between the onsets of symptoms and in case of no recovery, resultant death may occur in 14 days (Cascella et al., 2020). We collected and analysed the data from inception to April 14, 2020 (First phase of lockdown), by using a DES method, a family of forecasting models.

DES technique, also known as averaging, is used for the detection of notable variations in the data by considering the most recent data. This method uses the historical data for forecasted values. This is one of the most efficient short-term forecasting methods that apply to univariate time series data as a sequence of observations. The present study has used Holt's two-parameter DES model, which allows to separately smooth the trends by using the second smoothing constant. In this method, we use weighted averages of past observations to forecast the new values. The main idea here is to give more priority to the recent values in the series. As observation gets older (in terms of time), the importance of these values get exponentially smaller.

The DES method combines error, trend, and seasonal components in a smoothing calculation. Each term can be combined either additively, multiplicatively, or be left out of the model (Hox & Boeije, 2005; Hyndman et al., 2008; Verity et al., 2020). The DES method is a revised smoothing method than simple exponential smoothing method, which is applicable to the time series data with trend and no seasonality. Mainly this method is a linear combination of previous value of series for generating and predicting the future values (Bajric et al., 2013; Brown-Smoothing, 1963; Newbold & Granger, 1974). The application of the DES method requires three equations,

\[
\text{Forecast} = \text{Estimated level} + \text{trend at most recent time point}
\]

\[F_{t+k} = L_t + T_t \times k\]

where \(F_{t+k}\) represents the forecast for m-steps ahead period and \(K\) represents the number of period forecasts.

\[\text{Estimated level} = L_t = \alpha \times Y_t + (1 - \alpha) \times (L_{t-1} + T_{t-1})\]

\[\text{Estimated trend} = T_t = \beta \times (L_t - L_{t-1}) + (1 - \beta) \times T_{t-1}\]

Forecasting in R software is done with both simple and DES methods using a forecast package, which is available in base R (Hyndman & Khandakar, 2007; Montgomery et al., 2015). The advantages associated with this model are its better precision power as compared to other methods. It takes into account the most recent observations and weights them accordingly for reasonable future predictions. As such this method is best suited for short-term forecasting, as it assumes trends and patterns that will be similar to the current patterns and trends. However, major limitations associated with the DES methods are its inflexibility in terms of using few statistical data, especially for long-term accurate predictions. It can be used only when the time series data have minimal variations, with no seasonality.

2 RESULTS

The study estimates suggested that the total number of COVID-19 infections in India could be around 22,780.15 with 95% CI (18,723.44, 26,836.85) as of April 24, 2020. As a result, the confirmed rate of disease for the country as a whole will increase approximately at the rate of 8.22% daily. The base data were taken from March 11, 2020 to April 14, 2020 for confirmed cases on the daily basis. On
the basis of this data, we have predicted these confirmed cases from April 15, 2020 to April 24, 2020 and presented in Table 1. Moreover, we have estimated the number of active, recovered, and deaths due to COVID-19. The rate of increase in active cases were 8.92% (of 19,276.74 estimated active infections), for recovered cases it was 9.75% (of 2,827.29 estimated recovered infections), and the death

| Date         | Confirmed cases |         | Active cases |         |
|--------------|----------------|---------|--------------|---------|
|              | Point estimate | 95% confidence intervals | Point estimate | 95% confidence intervals |
| April 15, 2020 | 12,612.22     | 12,340.91 | 12,883.52    | 10,692.31 | 10,432.07 | 10,952.54 |
| April 16, 2020 | 13,741.99     | 13,247.05 | 14,236.93    | 11,646.13 | 11,183.33 | 12,108.93 |
| April 17, 2020 | 14,871.76     | 14,078.17 | 15,665.35    | 12,599.96 | 11,862.58 | 13,337.33 |
| April 18, 2020 | 16,001.53     | 14,855.55 | 17,147.51    | 13,553.78 | 12,490.98 | 14,616.58 |
| April 19, 2020 | 17,131.30     | 15,588.37 | 18,674.22    | 14,507.61 | 13,077.61 | 15,937.61 |
| April 20, 2020 | 18,261.07     | 16,281.93 | 20,240.21    | 15,461.44 | 13,627.59 | 17,295.28 |
| April 21, 2020 | 193.90.84     | 16,939.79 | 21,841.88    | 16,415.26 | 14,144.34 | 18,686.18 |
| April 22, 2020 | 20,520.61     | 17,564.65 | 23,476.56    | 17,369.09 | 14,630.41 | 20,107.76 |
| April 23, 2020 | 21,650.38     | 18,158.62 | 25,142.13    | 18,322.91 | 15,087.79 | 21,558.04 |
| April 24, 2020 | 22,780.15     | 18,723.44 | 26,836.85    | 19,276.74 | 15,518.10 | 23,035.38 |

**TABLE 2** Forecasting rate of recovered and death cases from double exponential smoothing method

| Date         | Recovered cases |         | Death cases |         |
|--------------|----------------|---------|-------------|---------|
|              | Point estimate | 95% confidence intervals | Point estimate | 95% confidence intervals |
| April 15, 2020 | 1,507.63      | 1,460.04 | 1,555.21    | 426.72 | 412.95 | 440.48 |
| April 16, 2020 | 1,654.25      | 1,563.55 | 1744.95     | 458.92 | 437.51 | 480.33 |
| April 17, 2020 | 1800.88       | 1,660.73 | 1941.04     | 491.13 | 458.52 | 523.74 |
| April 18, 2020 | 1947.51       | 1751.87 | 2,143.15    | 523.34 | 477.09 | 569.59 |
| April 19, 2020 | 2094.14       | 1837.49 | 2,350.79    | 555.54 | 493.74 | 617.35 |
| April 20, 2020 | 2,240.77      | 1917.98 | 2,563.56    | 587.75 | 508.76 | 666.74 |
| April 21, 2020 | 2,387.39      | 1993.71 | 2,781.11    | 619.96 | 522.31 | 717.60 |
| April 22, 2020 | 2,534.03      | 2064.94 | 3,003.12    | 652.16 | 534.53 | 769.79 |
| April 23, 2020 | 2,680.66      | 2,131.95 | 3,229.37    | 684.37 | 545.51 | 823.23 |
| April 24, 2020 | 2,827.29      | 2,194.93 | 3,459.65    | 716.58 | 555.32 | 877.84 |

**FIGURE 1** Forecasting of confirmed and active cases through double exponential method
rate has increased at 7.01% (of 716.58 estimated deaths) (Tables 1 and 2). The point forecast values for confirmed cases from 15 to 24 April, 2020 were 12,612.22, 13,741.99, 14,871.76, 16,001.53, 17,131.30, 18,261.07, 19,390.84, 20,520.61, 21,650.38, and 22,780.15. These values indicate the point forecast for the confirmed cases, which will be increasing 1,129.77 per day.

Figure 1a,b shows the number of confirmed and active cases for COVID-19. The figure panel shows a trend line depicted through red and blue colours. The red colour curve indicates the number of confirmed and active cases from 11 March to April 14, 2020 based on the actual data. This data were the base for the predicted values of confirmed and active cases in India, which was portrayed via blue colour trend line, that is, predicted region (Hyndman & Khandakar, 2007). It is predicting the number of confirmed and active cases in India from 15 to April 24, 2020. As per our predictions from 15 to April 24, 2020, the count of confirmed and active cases in India will fall in this region. The number of confirmed cases is presented at Y axis, and X axis represents day from March 11, 2020 to April 24, 2020. Similarly, recovered and death cases are depicted in Figure 2a,b.

3 | DISCUSSION

The COVID-19 pandemic in India has become a public health threat, and if we go as per the forecasts then the situation seems to be a matter of grave concern. This new virus represents tremendous concerns for the public health scenario in India. We neither have a vaccine, nor do we have a medication to treat from this virus. Our study emphasises upon the early detection and avoidance of critical care for a majority of infected patients to avoid overcrowding of the limited health care resources available in the country. The estimates from the present study have been cross checked with the actual cases reported on April 14 and 15, 2020. Our forecasts for the similar dates were approximately tallying with actual figures having minor variation.

Based on the high detection rate, it may enhance the testing capacity, which is essentially required. Late detection may put the cases in greater need of mechanical ventilation and ICU care, which impose a highly significant economic burden on the healthcare system (Ferguson et al., 2020).

Moreover, we observed that there is a need for the study based on the clinical characteristics of the deceased from COVID-19 to help the epidemiologists, and public health practitioners to address the concerns more effectively. Rapid assessments and clinical trials are critical to control the risk and to ensure efficient utilization of the available resources. There are also concerns that the total number of infections could be much higher as India’s testing rates are among the lowest in the world. No doubt, substantial public health interventions have been immediately implemented by the Government of India, there is a need for improved preparedness plans and mitigation strategies equipped for rapid deployment. Delay in required actions today can mean the difference between a manageable situation and a hopelessly overburdened healthcare system. Ultimately, we have to respect these restrictions implemented by the government in terms of a second phase of lockdown to help in reducing the growth of further infections.

4 | CONCLUSION

Accurate forecasting can play a critical role in understanding disease dynamics, comparing interventions, and providing critical data and information to the healthcare practitioners and government authorities. Health being a state subject in India vested a prominent position to the state and local officials to take key decisions in terms of nonpharmaceuticals public health interventions (NPIs) such as social distancing, home quarantine, and extension of lockdown measures. The NPI measures can be effective in reducing the reproduction of the infection cases only if sustained over time. By having a glimpse over
what we are going to encounter in the coming time, the accurate forecasting methods can significantly contribute in designing a more unified approach to fight with the pandemic. It will also help the policymakers to introduce better regulatory mechanisms such as fiscal measures to reduce the spread of viruses. However, forecasting alone should never be the only solution to this pandemic until there is a desired change in knowledge, attitude, and practices of the population towards this disease. The unified decision-making frameworks may help in reducing the distress, over optimism, and make us mentally prepared to address this biggest public health threat worldwide, especially for the highly populated countries such as India. We hope that forecasting COVID-19 will be more helpful for adapting better intervention policies.

**DATA AVAILABILITY STATEMENT**

The data that support the findings of this study are openly available in [repository name(s) “covid19india & coronavirus”] at https://www.covid19india.org and https://www.worldometers.info/coronavirus, reference number [5 & 6]. These data were derived from the following resources and is available in the public domain: https://www.covid19india.org and https://www.worldometers.info/coronavirus. The data used for the present study are available in the public domain. The supplementary material is not available but references to the repository is already provided in the article.

**ORCID**

Rinshu Dwivedi https://orcid.org/0000-0001-5912-2764  
Ramesh Athre https://orcid.org/0000-0002-2915-5095

**REFERENCES**

Bajracharya, H., Bajracharya, I., & Pasic, M. (2013). A comparison of slope seasonality and exponential smoothing forecasting models. [repository name], [publication date]. https://doi.org/10.2507/28th.daaam.proceedings.013. https://www.daaam.info/Downloads/Pdfs/proceedings/proceedings_2017/013.pdf

Banik, G. R., Khandaker, G., & Rashid, H. (2015). Middle east respiratory syndrome coronavirus “MERS-CoV”: Current knowledge gaps. Paediatric Respiratory Reviews, 16(3), 197–202.

Boshnakov, G. N. (2016). Introduction to Time Series Analysis and Forecasting, 2nd Edition, Wiley Series in Probability and Statistics, by Douglas C. Montgomery, Cheryl L. Jennings and Murat Kuluahci (eds). Published by John Wiley and Sons, Hoboken, NJ, USA, 2015. Total number of pag. Journal of Time Series Analysis, 37(6), 864–864. https://dx.doi.org/10.1111/jtsa.12203.

Cascella, M., Rajnik, M., Cuomo, A., Dulebohn, S. C., & Di Napoli, R. (2020). Features, evaluation and treatment coronavirus (COVID-19). In StatPearls [Internet]: COVID-19 India. StatPearls Publishing.

COVID-19. (2020). Corona virus pandemic. Retrieved from https://www.worldometers.info/coronavirus

Dhama, K., Sharan, K., Tiwari, R., Sircar, S., Bhat, S., Malik, Y. S., Karam Pal, S., Chaicumpa, W., Bonilla-Aldana, D. K., & Rodriguez-Morales, A. J. (2020). Coronavirus COVID-19 disease 2019-19. Clinical microbiology reviews, 33(4). https://doi.org/10.1128/CMR.00020-20

Dikid, T., Jain, S. K., Sharma, A., Kumar, A., & Narain, J. P. (2013). Emerging & re-emerging infections in India: An overview. The Indian Journal of Medical Research, 138(1), 19.

Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L. C., van Elsland, S., & Thompson, H. (2020). Report 9: Impact of Non-pharmaceutical Interventions (NPIS) to Reduce COVID19 Mortality and Healthcare Demand. 1–20. Imperial College London, https://doi.org/10.25561/77482.

Guo, Y. R., Cao, Q. D., Hong, Z. S., Tan, Y. Y., Chen, S. D., Jin, H. J., Tan, K. S., Wang, D. Y., & Yan, Y. (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak—an update on the status. Military Medical Research, 7(1), 1–10.

Healy M. J. R., Brown R. G. (1964). Smoothing, Forecasting and Prediction of Discrete Time Series. Journal of the Royal Statistical Society. Series A (General), 127(2), 292. http://dx.doi.org/10.2307/2344012.

Hilgenfeld, R., & Peiris, M. (2013). From SARS to MERS: 10 years of research on highly pathogenic human coronaviruses. Antiviral Research, 100(1), 286–295.

Hox, J. J., & Boeije, H. R. (2005). Data collection, primary versus secondary. Encyclopedia of Social Measurement, 1, 593–599. https://doi.org/10.1016/B0-12-369398-5/00041-4

Hyndman, R., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). Linear Innovations State Space Models. Forecasting with exponential smoothing: The state space approach, 1, (35–51). Springer Series in Statistics: Springer Science & Business Media. https://www.springer.com/gp/book/9783540471916.

Hyndman, R. J., Khandakar, Y. (2008). Automatic Time Series Forecasting: The forecastPackage for R. Journal of Statistical Software, 27, (3). http://dx.doi.org/10.18637/jss.v027.i03.

Indian Council of Medical Research (ICMR). (2020) Retrieved from https://www.icmr.gov.in/.

Kolifarhood, G., Aghaali, M., Saadati, H. M., Taherpour, N., Rahimi, S., Izadi, N., & Nazari, S. S. H. (2020). Epidemiological and clinical aspects of COVID-19: a narrative review. Archives of Academic Emergency Medicine, 8(1):e41. https://doi.org/10.22037/aeem.v8i1.620.

Ministry of Health and Family Welfare (MoHFW) Covid 19. (n.d.) Retrieved from https://www.mohfw.gov.in/

Newbold, P., & Granger, C. W. (1974). Experience with forecasting univariate time series and the combination of forecasts. Journal of the Royal Statistical Society: Series A (General), 137(2), 131–146.

Paraskevis, D., Kostaki, E. G., Magiorkinis, G., Panayiotakopoulos, G., Sourvinos, G., & Tsiouras, S. (2020). Full-genome evolutionary analysis of the novel coronavirus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. Infection, Genetics and Evolution, 79, 104212.

Singhal, T. (2020). A review of coronavirus disease-2019 (COVID-19). The Indian Journal of Pediatrics, 87(4), 1–6. https://doi.org/10.1007/s12098-020-03262-6.

Sutton, T. C., & Subbarao, K. (2015). Development of animal models against emerging coronaviruses: From SARS to MERS coronaviruses. Virology, 479, 247–258.

Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., & Dighe, A. (2020). Estimates of the severity of coronavirus disease 2019: A model-based analysis. The Lancet Infectious Diseases, 20, 669–677.

World Health Organization. (2020). Coronavirus disease 2019 (COVID-19): situation report, 72. Wu, T., Perrings, C., Kinzig, A., Collins, J. P., Minteer, B. A., & Daszak, P. (2017). Economic growth, urbanization, globalization, and the risks of emerging infectious diseases in China: A review. Ambio, 46(1), 18–29.

Wu, Z. & McGoogan, J. M. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA, 2020;323(19):1196–1202. https://doi.org/10.1007/s12098-020-03262-6

Yang, Y., Peng, F., Wang, R., Guan, K., Jiang, T., Xu, G., Sun, J., & Chang, C. (2020). The deadly coronaviruses: The 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China. Journal of Autoimmunity, 109, 102434.
AUTHOR BIOGRAPHIES

Dr Rinshu Dwivedi, Assistant Professor, Indian Institute of Information Technology, Trichy, Tamil Nadu, India-620009. Email: rinshudwivedi999@gmail.com

Dr Ramesh Athe, Assistant Professor, Indian Institute of Information Technology, Dharwad, Karnataka, India-580009. Email: dr.athe9@gmail.com

Prof. Kavi Mahesh, Director, Indian Institute of Information Technology, Dharwad, Karnataka, India-580009.

Mr Praveen Kumar Modem, B.Tech Student, Indian Institute of Information Technology, Dharwad, Karnataka, India-580029. Email: modemiiit@gmail.com

How to cite this article: Dwivedi R, Athe R, Mahesh K, Modem PK. The incubation period of coronavirus disease (COVID-19): A tremendous public health threat—Forecasting from publicly available case data in India. J Public Affairs. 2021; 21:e2619. https://doi.org/10.1002/pa.2619