Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 1):
pp. 6857-6866
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.01.710

Forecasting The Scope Of Third Wave Of COVID - 19 In Uttarakhand, India

Omdeep Gupta[1), a)], Dr. Pradeep Joshi[1,b)], Poonam Verma[2, c)] Dr. Rupa Khanna Malhotra[3]

[1]School of Management, Graphic Era Hill University, Dehradun, Uttarakhand, India.
[2]School of Computing, Graphic Era Hill University, Dehradun, Uttarakhand, India.
[3]School of Commerce, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India.b
[a)]ogupta@gehu.ac.in
[b)]pjoshi@gehu.ac.in
[c)] pverma@gehu.ac.in

Abstract. In the last one and half year the whole world had faced the challenge of Covid – 19 and the situation is still risky according to the predictions of World Health Organization (WHO). Governments need to forecast about the strength and behaviour of pandemic to minimize the effects of pandemic. In the current study, Linear Regression, Polynomial Regression and Holt's Linear Model is applied to predict the magnitude and time period of third wave of COVID – 19 in Uttarakhand.Results depict that the peak of the third wave in Uttarakhand is expected to be in Mid-September using Linear Regression. This study is valuable to government, medical institutions and other agencies involved in controlling pandemic by improving decision making and providing medical facilities and treatment to the people.

Keywords – Forecasting, Covid – 19, Uttarakhand, India.

INTRODUCTION

Last year, in March 2020, Covid – 19 hit India and because of stringent measures like weekly lockdown and containment zones, India was able to control it quite better in comparison to other developed countries (Ghosh, Nandi and Mallick, 2020). By the month of November – December things were quite better and in following months the different institutions including schools started opening (Dubey, 2020). But, suddenly, in the month of April it surged again and within a month it was touching its peak (Times of India, 2021). Moreover, the surge was so acute in these two months in India that average positive per day tripled in comparison to last year. The behaviour of pandemic changed in one more way this year that it affected larger number of middle-aged persons this year in comparison to previous year. Further, its affect was also found to be more severe in comparison to last year as the death toll this year was significantly high. And most

surprisingly, this second wave fell within two months as the situation was quite under control as it approached first week of June (Firstpost, May 2021). But, by that time the pandemic had already done a huge damage within a short span of two months. And the main reason was that none has thought of this kind of behaviour of pandemic and hence there was little preparation for it. Therefore, it is essential to predict the strength and behaviour of pandemic as it will help in improving decision making related to resource allocation and medical facilities.

One very important observation about Covid -19 is the variation in its impact across different states of India due to various socio – economic and environmental reasons (Ghosh and Chakraborty, 2020). Moreover, there is huge diversity across different sates of India in terms of climate, pollution, culture, lifestyle, and income and education of people. Therefore, it is difficult to generalize the results of whole nation to a particular state. This study has focused on Uttarakhand state of India. Uttarakhand is a small state of India with area 53,483 sq.km. and a population of 101.17 lakh with a literacy rate of 79.63%. It is situated at the bottoms of the Himalaya Mountain and is majorly a mountainous state (Uttarakhand official website). As most of the previous research have focused on forecasting with respect to entire country, therefore this study attempts to study the pandemic with respect to the Uttarakhand state of India. Keeping in view the above issues, this study aimed at forecasting the magnitude and time - period of third wave of Covid – 19 with respect to Uttarakhand.

Globally, as of 30th June 2021, there were 18.20 million established cases of COVID-19, containing 3.95 million deaths, and 1.85 billion vaccinated according to the data of WHO (WHO report). If we look at figures in India, total number of cases were 3,03,62,848 as on 30th June 2021. Almost 96.92% (2,94,27,330) have recovered and number of deaths have been 1.31% (3,98,454). Regarding vaccination, till now total 33,28,54,527 doses have been administered. As far as Uttarakhand is concerned total number of cases have been 3.39,933 out of which 3,30,593 (97%) were discharged and 7,095 (2.1%) died. The death ratio in Uttarakhand have been 2.1% which is surprisingly second highest after Punjab (2.7%) (PIB Bulletin, 30th June, 2021).

REVIEW OF LITERATURE

Graichen (2021) compared the consequent waves from a European perspective and found that though some countries were able to manage the first wave well, but second wave proved to be hard test because of number of factors including reduced restrictions, seasonal changes, and festival gatherings. The importance of comparison across the waves can be understood from the fact that the treatment has also been different across the waves. For example, second and third wave in Europe saw more use of corticosteroids compared with first waveTaboada et al. (2021). The reason for changes in treatment can be attributed to large number of continuing consequences of COVID-19 apart from the initial wave of serious contaminations. These consequences may also include acute respiratory distress syndrome (ARDS) than can show severe hypoxaemia which has also been described as ‘oxygen levels unsuited for life’ (Baker, Safayynia and Evered, 2021). Soriano et al. (2021) identified the characteristics of third wave that included higher

percentage of case related to family contacts, use of rapid antigen tests for quick analysis and lower proportion of clinically serious cases and death rates. Further, identification of more transmissible virus strains and the impact of vaccination was found in latest waves. Similarly, Seong et al. (2021), compared second wave and third in South Korea and found that third wave had longer duration, higher fatality rate and different transmission chains. Above studies conducted in diverse parts of the globe showed that the behaviour of pandemic has been different across the waves in terms of intensity, symptoms, and treatment. In India, and particularly Uttarakhand the difference was observed between first and second wave due t the fact that the second wave has been much more damaging with the fast-spreading variants that has infected more of younger people (India TV, 2021).

Since Covid – 19 has spread its imprints across the world, some of the researchers have conducted prediction related studies, limited of them are with reference to India and none of them is with respect to Uttarakhand. Costris, Schwartz and Smith (2020) studied the trustworthiness of different mathematical models put forward during various pandemics and found that models developed in early stages could be useful for long term forecasting. Tomar and Gupta (2020) applied data-driven assessment methods like long short-term memory (LSTM) for estimation of the count of COVID-19 cases in India 30 days ahead and for understanding the outcome of precautionary actions like public separation and lockdown on the spread of COVID-19. The estimation of different parameters by the given method was precise within a certain range. Torrealba, Conde and Hernandez (2020) conducted a study to predict the number of cases in Mexico with the help of mathematical (Gompertz and Logistic) and computational (Artificial Neural Network) models by using only the established cases. The findings showed a good match between the experimental data and those found by the mathematical and computational models. Mohamadou, Halidou and Kapen (2020) reviewed the techniques used in the previous works related to the recognition of COVID-19 using mathematical modeling and Artificial intelligence (AI). It was found that mathematical modeling was mostly done by Susceptible-Exposed-Infected-Removed (SEIR) and Susceptible-infected-recovered (SIR) models while Convolutional Neural Network (CNN) was mostly used for AI implementations. Ghosh and Chakraborty, (2020) forecasted the count of infections for each state in India for the next 30 days using logistic, the exponential, and the susceptible-infectious-susceptible models and classified the states into severe, moderate, and controlled category. Huang et al. (2020)combined the pandemicforecasting model with worldepidemic data and measured the effect of ecological factors to establish global prediction system. This system predicted COVID-19 status for each country and for each day by using epidemiological susceptible–infectious–recovery (SIR).

MATERIALS AND METHODS

In this experiment, we have utilized Jupyter Notebook for Python programming. The data related to COVID-19 was collected for Uttarakhand state. The epidemiological records till 30th June 2021 were taken from John Hopkins University website. The data of COVID-19 is available for confirmed, deceased, recovered from 30th January 2020. Since the data

did consist of few empty cells and negative values, the pre-processing of the data was required. The positivity rate and active cases were calculated from the given three parameters respectively. For experimentation, linear regression, polynomial regression and Holt winter's model have been used to carry out predictions.

The values of the different parameters used in the data for the prediction have been estimated using different pre-defined function of Python's library. Using the data of Confirmed, Recovered and Deceased of the Uttarakhand state, other parameters have been calculated.

The basic steps for the prediction model were as given below:

1) Data Acquisition
2) Data Pre-processing
3) Model Fitting using CSV time series data
4) Parameter Selection
5) Prediction using Selected Parameters

FINDINGS

This section describes the predictions related outcomes obtained by using Linear Regression, Polynomial Regression and Holts' Winter Model.

LINEAR REGRESSION PREDICTION

Linear Regression is a linear approach for modelling the relationship between independent and dependent variables (Abolmaali, 2021). The findings of the linear regression for the confirmed cases are shown in graph below (Figure 1).

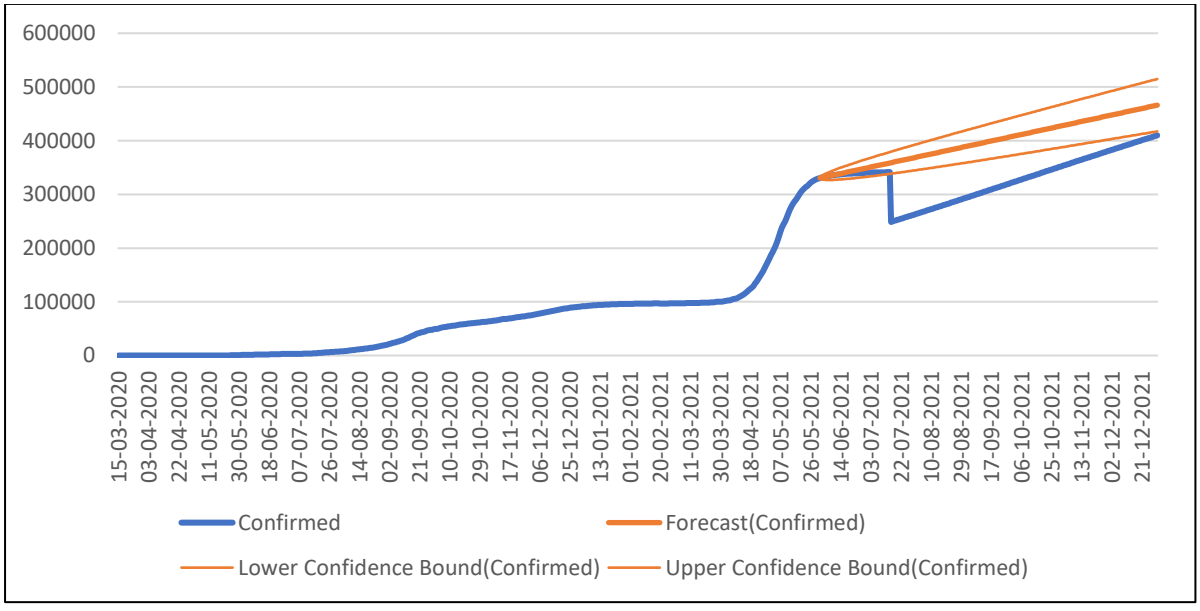

FIGURE 1. Prediction of Linear Regression for Confirmed cases

Linear Regression for the recovered cases is shown below (Figure 2). The statistical value of the SMAPE for the model is 0.01 and the value of RMSE for the model was found to be 131507375.

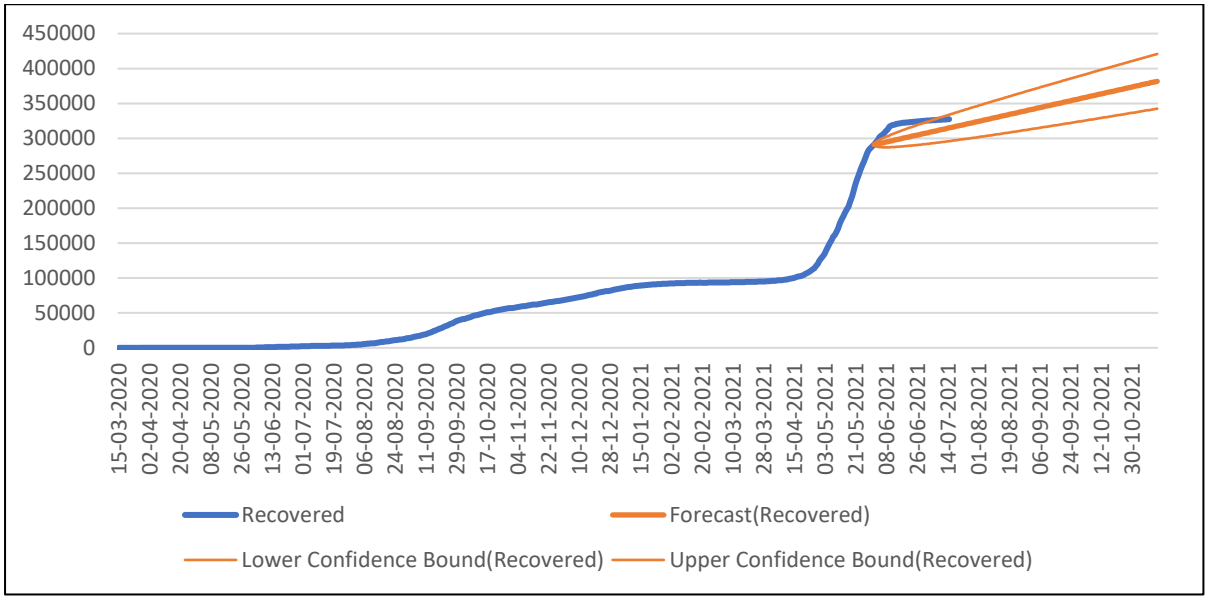

FIGURE 2. Prediction of Linear Regression for Recovered cases

The result of linear regression for prediction of the deceased cases is shown below (Figure 3). The results obtained are from 1st June 2021 and as shown in the figure, it can be clearly observed that the pattern of the Covid-19 is following the upper bound of the predicted values. The prediction ranges from the 1st June 2021 to 14th Nov 2021.

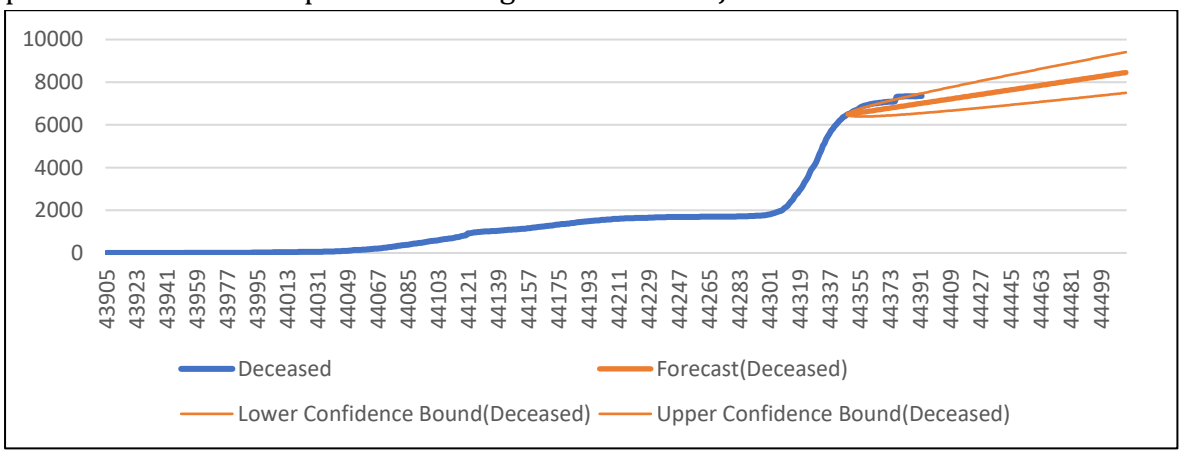

FIGURE 3. Prediction of Linear Regression for Deceased cases

The graph above shows a trend with an improvement in the number of recovered cases. According to the linear regression the number of daily recovered cases will 381606 by 14th Nov 2021. To reduce the effect of the pandemic, it becomes necessary to reduce the number of deceased patients which is seen to be 8455 as shown by the model.

POLYNOMIAL REGRESSION PREDICTION

Polynomial Regression is a form of linear regression that forms a relationship between the independent and the dependent variables(Fritz, 2021). The basic target of this regression is to model the expected value of dependent variable in terms of an independent variable x. In the present model, we have taken the degree as 4 for the values of the dependent variables such as Date, vaccination doses, recovered, confirmed cases for the active cases.The statistical value of RMSE for the model was found to be 148861.63. (Figure 4).

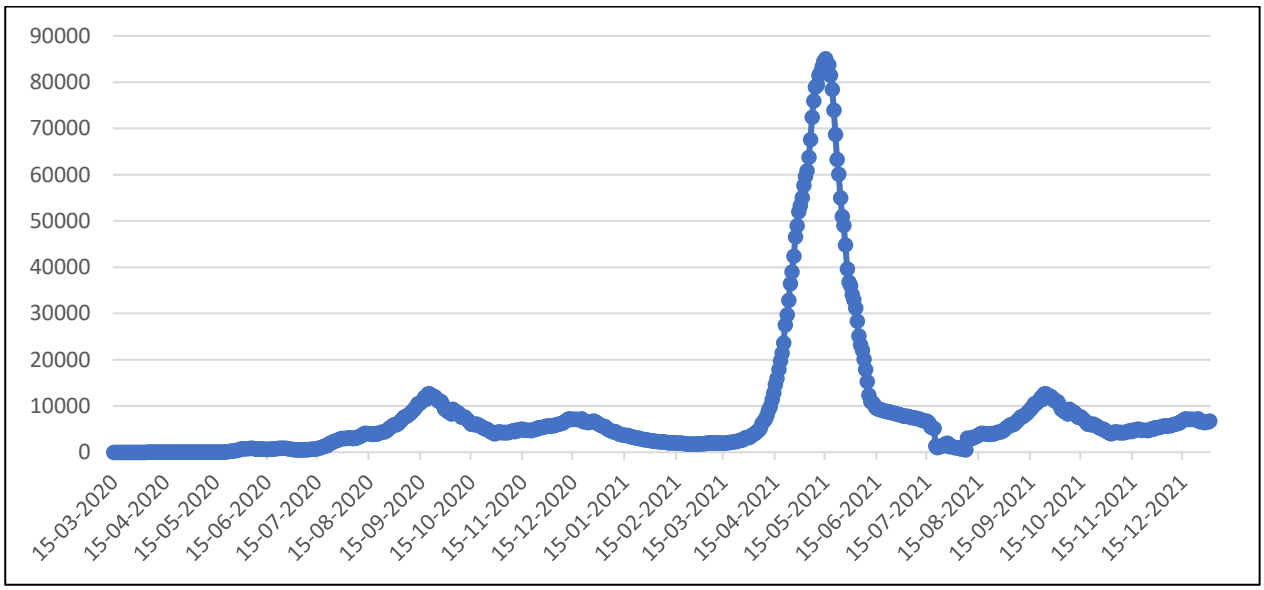

FIGURE 4. Prediction of Active Cases using Polynomial Regression for upcoming months

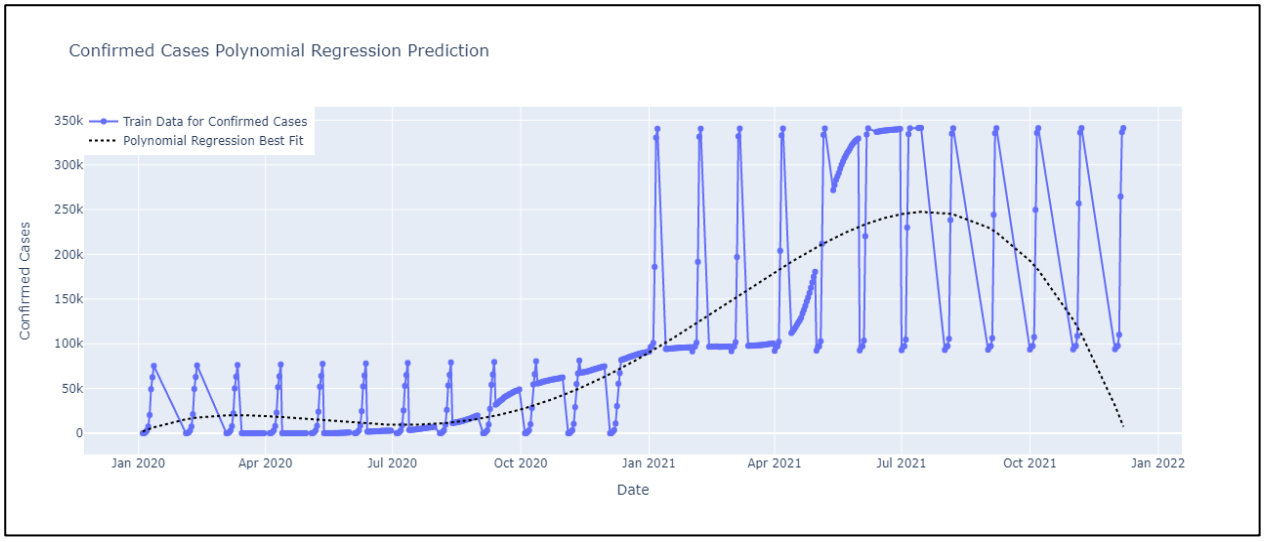

FIGURE 5. Prediction of Confirmed Cases using Polynomial Regression

The prediction for Confirmed Cases in upcoming months in Uttarakhand is shown in figure 5 and date wise predicted number of cases are shown below in Table 1.

TABLE 1. Prediction for Confirmed Cases in upcoming months in Uttarakhand

Date	No. of Cases	Date	No. of Cases	Date	No. of Cases
21-08-2021	3850	07-10-2021	8319	23-11-2021	4752
22-08-2021	4073	08-10-2021	9342	24-11-2021	4674
09-09-2021	7649	26-10-2021	5340	12-12-2021	6293
10-09-2021	7715	27-10-2021	5099	13-12-2021	6370
11-09-2021	8042	28-10-2021	5002	14-12-2021	6675
12-09-2021	8261	29-10-2021	4895	15-12-2021	6883
26-09-2021	12032	12-11-2021	4495	29-12-2021	6643

HOLT'S WINTER MODEL

Holt's two-parameter model is a popular smoothing model for forecasting data with trend.

Additive Model: $\hat{y}_{t+h|t}=\ell_t+hb_t+s_{t+h-m(k+1)}$

Multiplicative Model: $\hat{y}_{t+h|t}=(\ell_t+hb_t) \times s_{t+h-m(k+1)}$

Multiplicative model is preferred over additive model, as the variation changes in proportion to the series. Holt's model has three separate equations used to generate a final forecast (Sulandari, Suhartono and Rodrigues, 2021).

$$\ell_t=\alpha(y_t / s_{t-m})+(1-\alpha)(\ell_{t-1}+b_{t-1}) \text{.....eq. i}$$

$$b_t=\beta^*(\ell_t-\ell_{t-1})+(1-\beta^*)b_{t-1} \text{.....eq.ii}$$

$$s_t=\gamma(y_t / \ell_{t-1}-b_{t-1})+(1-\gamma)s_{t-m} \text{.....eq. iii}$$

First equation (eq. i) deals with a basic smoothing equation that directly adjusts the last smoothed value for last period's trend, which is further updated over time through the second equation (eq. ii). Finally, the third equation is used to generate the final forecast (eq. iii0. This model is popularly recognised as trend-enhanced exponential smoothing (Wongoutong, 2021). It is an extension of the Holt's Linear Model that deals with the seasonal changes occurring over the pattern or trend.

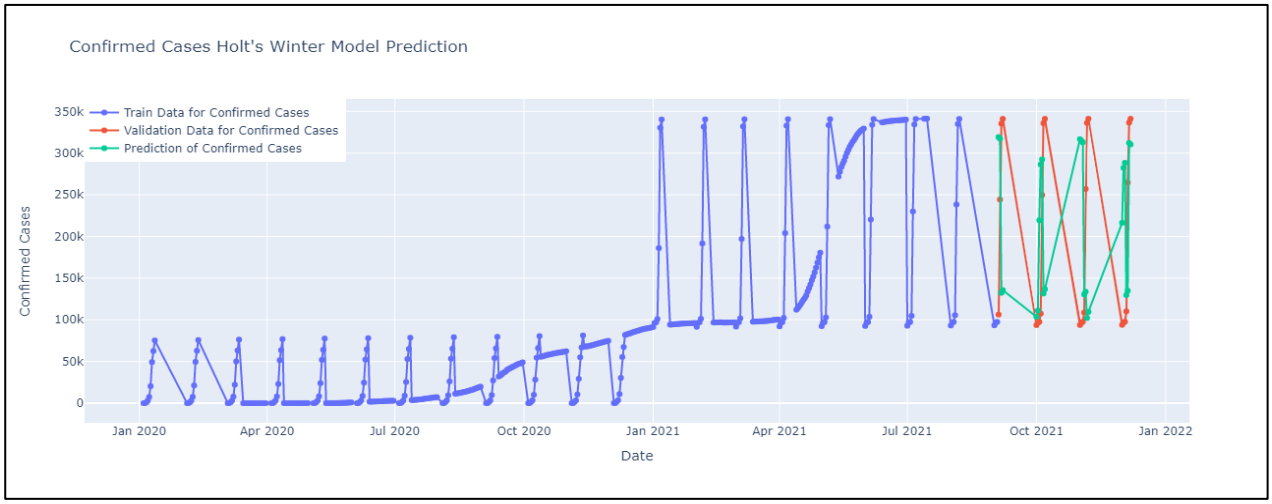

FIGURE 6. Prediction of the Confirmed Cases using Holt's Winter Model for the upcoming months

TABLE 2. Prediction of Maximum Confirmed Cases using Holt Winter Model

Dates	Prediction of Maximum Confirmed
4th Sept 20	31.9239 k
1st Oct 20	10.3600 k
1st Nov 20	31.6697 k

We have considered Seasonal Degree=12. Both Upper Bound and Lower Bound has been depicted in the graph. (figure 6). The Holt's winter model considers seasonal, trend while predicting the cases (Table 2). The statistical value of RMSE for the model was found to be 194213.75.

CONCLUSION

The results obtained for Uttarakhand shows that there will be a peak of COVID-19 cases on daily infections.However, unlock 3.0 has been witnessing a surge in its value, thereby indicating an increase in the number of individuals getting infected as more relaxations permit people to move around more freely and thereby, they come in contact with the infected person leading to an increased susceptibility. However, the recovery rate is also on constant rise, thereby leading to dip in the rate of deceased patients. This paper has utilized basic models for forecasting from the available data of Confirmed, Recovered, Deceased and Vaccination cases. Linear Regression and Polynomial Regression have been able to describe the most possibly higher and lower COVID-19 cases likely to occur in upcoming months in the state of Uttarakhand.

One of the limitations of such forecasting techniques is the change in result due to interplay of different variables. In case ofCovid-19 the variants of the virus due to mutations play an important role. The availability and reliability of the specific variant related data is always a challenge. Other factors that affect the spread of the virus include festivals, social gatherings and other socio-cultural activities and therefore they may affect the result also. Since supervised algorithms require large training data, lack of such large data can make the results prone to overfitting.Future scope may include the application of unsupervised and semi supervised algorithms for better prediction.

REFERENCES

1. Abolmaali, S. (2021). A comparative study of SIR Model, Linear Regression, Logistic Function and ARIMA Model for forecasting COVID-19 cases. MedRxiv
2. Baker HA, Safavynia SA, Evered LA. (2021) The 'third wave': impending cognitive and functional decline in COVID-19 survivors. Br J Anaesth. 2021 Jan;126(1):44-47. doi: 10.1016/j.bja.2020.09.045.
3. Costris, C., Schwartz. E.J. and Smith, R. (2020), Predicting COVID-19 using past pandemics as a guide: how reliable were mathematical models then, and how reliable will they be now? Mathematical Biosciences and Engineering, 17(6): 7502 - 7518. DOI: 10.3934/mbe.2020383

4. Dubey, J. (2020), Covid-19 Factoid: India's active cases drop by 135,000 in November, Business Standard, 30th November, 2020, https://www.business-standard.com/article/current-affairs/covid-19-factoid-india-s-active-cases-drop-by-135-000-in-november-120113000562_1.html, accessed on 15th June, 2021
5. Firstpost (May, 2021), COVID-19 tracker, 4 May, 2021: India records 3,57,229 new cases, 3,449 deaths in 24 hrs, Firstpost, May 04, 2021, https://www.firstpost.com/india/covid-19-tracker-4-may-2021-india-records-357229-new-cases-3449-deaths-in-24-hrs-9592691.html, accessed on 30th May, 2021
6. Fritz, M. (2021). Improved Output Gap Estimates and Forecasts Using a Local Linear Regression. In Engineering Proceedings (Vol. 5, No. 1, p. 32). Multidisciplinary Digital Publishing Institute.
7. Ghosh, A., Nundy, S. and Mallick, T. K. (2020), How India is dealing with COVID-19 pandemic, Sensors International, Vol.1, https://doi.org/10.1016/j.sintl.2020.100021.
8. Ghosh, P., Ghosh, R., & Chakraborty, B. (2020). COVID-19 in India: Statewise Analysis and Prediction. JMIR public health and surveillance, 6(3), e20341. https://doi.org/10.2196/20341
9. Graichen H. (2021). What is the difference between the first and the second/third wave of Covid-19? - German perspective. Journal of orthopaedics, 24, A1–A3. https://doi.org/10.1016/j.jor.2021.01.011
10. https://covid19.who.int/. Accessed on 30th June 2021.
11. https://uk.gov.in/pages/display/115-state-profileAccessed on 25th June 2021.
12. Huang, J., Zhang, L., Liu, X., Wei, Y., Liu, C., Lian, X., Huang, Z., Chou, J., Liu, X., Li, X., Yang, K., Wang, J., Liang, H., Gu, Q., Du, P., & Zhang, T. (2020). Global prediction system for COVID-19 pandemic. Science bulletin, 65(22), 1884–1887. https://doi.org/10.1016/j.scib.2020.08.002
13. India TV, (2021), Second wave affecting young more than old: Uttarakhand State Control Room Data, India TV, May 23, 2021 8:37 IST, https://www.indiatvnews.com/news/india/second-wave-affecting-young-more-than-old-uttarakhand-state-control-data-706392 accessed on 10th of June 2021
14. Mohamadou, Y., Halidou, A. &Kapen, P.T. (2020) A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. ApplIntell 50, 3913–3925. https://doi.org/10.1007/s10489-020-01770-9
15. PIB's Bulletin on Covid-19 (30th June, 2021), https://pib.gov.in/PressReleasePage.aspx?PRID=1731608, accessed on 30th June, 2021
16. Seong, H., Hyun, H.J., Yun, J.G., Noh, J.Y., Cheong, H.J., Kim, W.J., Song, J.Y., (2021), Comparison of the second and third waves of the COVID-19 pandemic in South

Korea: Importance of early public health intervention, International Journal of Infectious Diseases, Vol. 104, pp. 742-745
17. Soriano, V., Mendoza, C., Gallego, Felix G., Corral, O., and Barreiro, P. (2021). Third wave of COVID-19 in Madrid, Spain, International Journal of Infectious Diseases, Vol. 107, pp. 212-214, https://doi.org/10.1016/j.ijid.2021.04.074.
18. Sulandari, W., Suhartono, S., and Rodrigues, P. C. (2021). Exponential Smoothing on Modeling and Forecasting Multiple Seasonal Time Series: An Overview. Fluctuation and Noise Letters, 2130003
19. Taboada, M., González, M., Alvarez, A., Eiras, M., Costa, J., Álvarez, J., &Seoane-Pillado, T. (2021), First, second and third wave of COVID-19. What have we changed in the ICU management of these patients? The Journal of infection, 82(6), e14–e15. https://doi.org/10.1016/j.jinf.2021.03.027
20. Times of India, (2021), April horror: How second wave of Covid hit Delhi and Mumbai, Times of India, May 12, 2021, https://timesofindia.indiatimes.com/india/april-horror-how-covids-second-wave-hit-delhi-and-mumbai/articleshow/82576919.cms, accessed on 20th June 2021
21. Tomar, A., & Gupta, N. (2020). Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. The Science of the total environment, 728, 138762. https://doi.org/10.1016/j.scitotenv.2020.138762
22. Torrealba-Rodriguez, O., Conde-Gutiérrez, R. A., & Hernández-Javier, A. L. (2020). Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. Chaos, solitons, and fractals, 138, 109946. https://doi.org/10.1016/j.chaos.2020.109946
23. Wongoutong, C. (2021). The Effect on Forecasting Accuracy of the Holt-Winters Method When Using the Incorrect Model on a Non-Stationary Time Series. Thailand Statistician, 19(3), 565-582.