MODELLING AND FORECASTING COVID-19 CASES IN INDIA USING ARTIFICIAL NEURAL NETWORKS

*Dr. Smartson. P. NYONI[1], Thabani NYONI[2], Tatenda. A. CHIHOHO[3]

[1]*ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe*
[2]*Department of Economics, University of Zimbabwe, Harare, Zimbabwe*
[3]*Department of Economics, University of Zimbabwe, Harare, Zimbabwe*
**Corresponding Author*

ABSTRACT

India has been lately witnessing a daily surge in the COVID-19 infected cases. The country is currently amongst the worst-hit nations worldwide. In order to prevent the further worsening of the situation, forecasting the future course of the pandemic is inevitable. In this research article, the ANN approach was applied to analyze daily new cases of COVID-19 in India. The employed data covers the period 30 January 2020 to 30 October 2020 and the out-of-sample period ranges over the period November 2020 to April 2021. The residuals and forecast evaluation criteria (Error, MSE and MAE) of the applied model indicate that the model is stable in forecasting daily new cases of COVID-19 in India. The results of the study indicate that COVID-19 daily new infections are likely to, generally, remain quite very high; over the next 6 months in India. In order to reverse the forecasted scenario, the administration and healthcare personnel need to formulate and implement robust control measures. On the other side of the same coin, the general public needs to seriously, responsibly and religiously adhere to the set guidelines to slow down the spread of the pandemic and ultimately prevent it from escalating into a catastrophe.

Keywords: - *ANN, COVID-19, Forecasting*

INTRODUCTION

Since the first emergence of the Coronavirus Disease 2019 (COVID-19) in Wuhan, Hubei Province, China, in December 2019 (Wu *et al*., 2020; Zhou *et al*., 2020; Zhu *et al*., 2020), the scourge has proliferated globally and has affected literally all nations till date (Worldometers, 2020). The global outbreak and severity of this contagious disease (Li *et al*., 2020) prompted the World Health Organization (WHO) to declare it as a Public Health Emergency of International Concern (PHEIC) on January 30, 2020, and then subsequently, to classify it as a pandemic on March 11, 2020 (WHO, 2020). The COVID-19 pandemic was first confirmed in India on January 30, 2020, in the state of Kerala. A total of 82087 confirmed cases and 2648 deaths in the country were already reported by May 15, 2020 (India COVID-19 Tracker, 2020). As of August 2020, COVID-19 has infected approximately 20 million people across the globe with 90 countries in community transmission stage (WHO, 2020), leading to significant efforts towards control (Rawaf *et al*., 2020), modelling (Barkur & Vibha, 2020; Chatterjee *et al*., 2020a; Giordano *et al*., 2020), search for a cure (Le *et al*., 2020) for COVID-19 across the world and India (Chatterjee *et al*., 2020; Singhal, 2020). Today India has emerged as the latest global

hotspot of the pandemic owing to the high rise in cases and the fact that most of the population of the country lives in densely packed cities (Worldometers, 2020). No doubt the spread of this virus outbreak has seriously disrupted life, economy and health of citizens and this has become a great concern for everyone, as to how long will this scenario last and when the disease could be controlled (Sardar *et al*., 2020). Modeling and forecasting of the COVID-19 pandemic in India, just like in any other affected country; remains important, until a dedicated vaccine is developed for the disease.

REVIEW OF RELATED STUDIES

Table 1 below is a summary of relevant previous studied done in India:

Table 1: Studies Reviewed

Author/s (Year)	Study Period	Method	Major Findings
Patrikar *et al*. (2020)	January – March	Modified SEIR model	Social distancing seems to be working for India
Salgotra *et al*. (2020)	January – May 2020	Genetic Programming	The proposed GEP-based models are highly reliable for time series prediction of COVID-19 cases in India
Rafiq *et al*. (2020)	January – May	State-space Model	The proposed state-space model can very well capture the disease variations with high accuracy
Goswami *et al*. (2020)	January – May	General Additive Model	Statistically significant linear trend found for the daily-confirmed cases of COVID-19
Rath *et al*. (2020)	January – June	Multiple Linear Regression	The proposed model is reliable for forecasting COVID-19 daily cases in India
Malavika *et al*. (2020)	January – June	SIR and Logistic Growth models	The Logistic Growth model is suitable for short-term forecasts of COVID-19 cases while the SIR model performs better for long-term predictions
Khan & Gupta (2020)	January – April	ARIMA and NAR models	The best ARIMA model was the ARIMA (1, 1, 0), while the preferred NAR model architecture constituted 10 neurons
Khajanchi & Sarkar (2020)	January – June	Compartmental Mathematical Model	Short-time prediction shows the increasing trend of daily and cumulative cases of COVID-19 for the 4 (studied) states of India
Suvarna *et al*.	January	Gaussian	The proposed model is reliable for both long and short term forecasts of COVID-19

(2020)	– June	Additive Model	cases. COVID-19 cases will continue on an upward trend even beyond December 2020 into around March 2021.
Dhamodharavadhani *et al*. (2020)	January – May	Neural Network Models	The PNN and RBFNN-based MRP model performed better than the other models for COVID-19 forecasting
Tamang *et al*. (2020)	January – March	Neural Networks	COVID-19 Cases are predicted to rise in India
Marimuthu *et al*. (2020)	January – April	SEIR model	Rising trend of COVID-19-TB co-infected patients was predicted: hence the importance of primary prevention measures, especially in TB patients
Arora *et al*. (2020)	January – May	Deep Learning Models	The proposed models are reliable
Sarkar *et al*. (2020)	January – April	SEIR model	Quarantining susceptible individuals can effectively control COVID-19
Mahajan *et al*. (2020)	January – May	SIPHERD	Increasing tests per day, social distancing and strict lockdown are effective in controlling spread of COVID-19
Ray *et al*. (2020)	January – May	SEIR model	The proposed model is effective
Roy *et al*. (2020)	January – May	ARIMA	The proposed model was reliable in forecasting future infections of COVID-19

Source: Literature Review (2020)

METHOD

The modeling and forecasting of COVID-19 cases has been attempted by a wide range of researchers from the very beginning of cases in India (Kotwal *et al*., 2020). Given the fact that ANN techniques are one of the most efficient methods in processing huge data sets (Khan & Gupta, 2020; Dhamodharavadhani *et al*., 2020; Tamang *et al*., 2020; Arora *et al*., 2020), this paper applies the multi-layer perceptron neural network type of the ANN approach in order to predict COVID-19 cases in India. The study particulary applies the ANN (12, 12, 1) model and chooses the more efficient hyperbolic tangent function as the activation function.

Data Issues

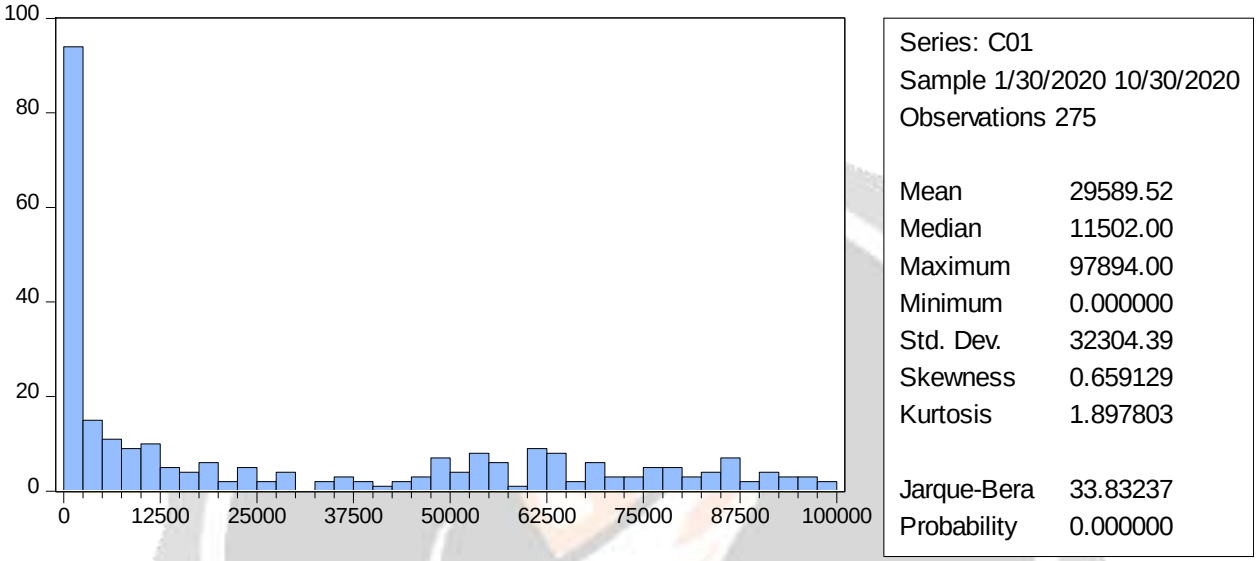
This study is based on daily new cases of COVID-19 infections (reffered to as series, C, in this study) for all age groups in India. The data covers the period 30 January 2020 to 30 October 2020 while the out-of-sample forecast covers the period November 2020 April 2021. All the data

employed in this paper was gathered from the COVID-19 data repository prepared by the CSSE at JH University.

FINDINGS OF THE STUDY

Descriptive Statistics

Figure 1: Descriptive statistics



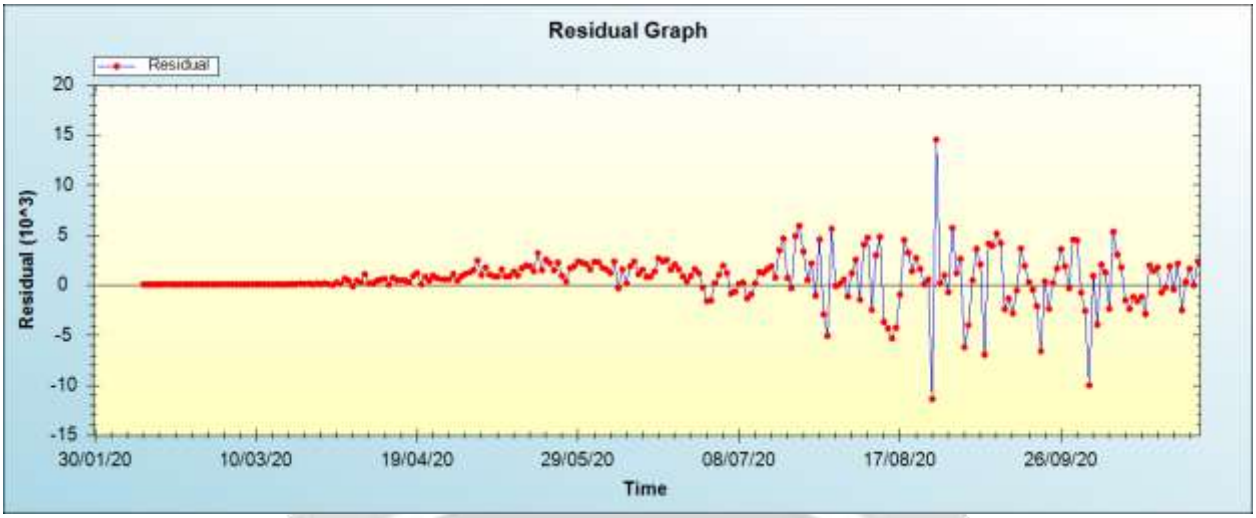
ANN MODEL SUMMARY FOR COVID-19 DAILY CASES IN INDIA

Table 2: ANN model summary

Variable	C
Observations	263 (After Adjusting Endpoints)
Neural Network Architecture:	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Function	Hyperbolic Tangent Function
Back Propagation Learning:	
Learning Rate	0.005
Momentum	0.05
Criteria:	
Error	0.045025
MSE	5996261.442071
MAE	1594.765639

Residual Analysis for the ANN model

Figure 2: Residual analysis



In-sample Forecast for C

Figure 3: In-sample forecast for the C series

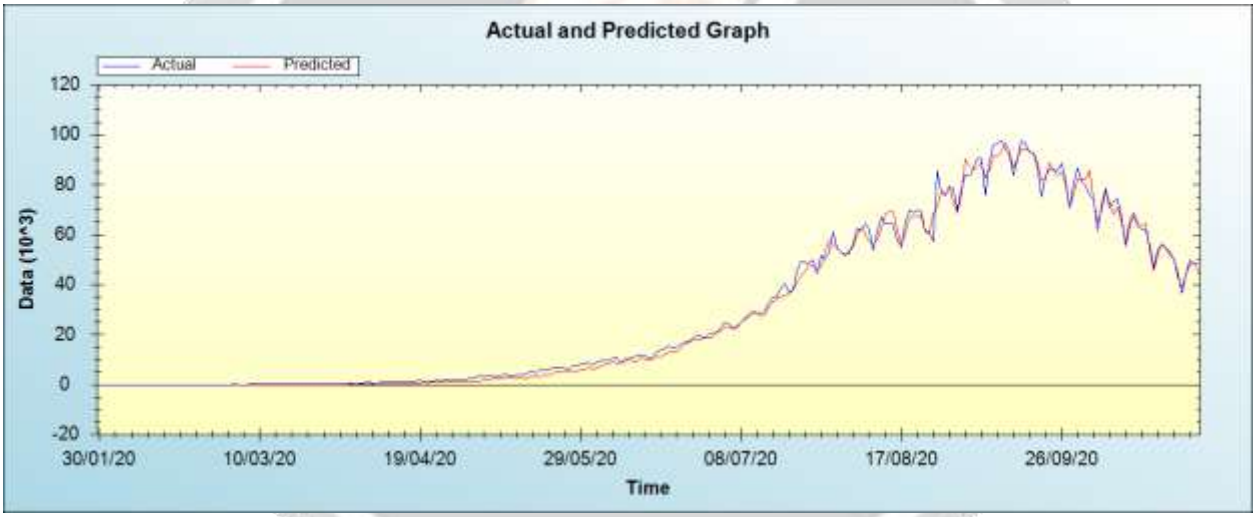
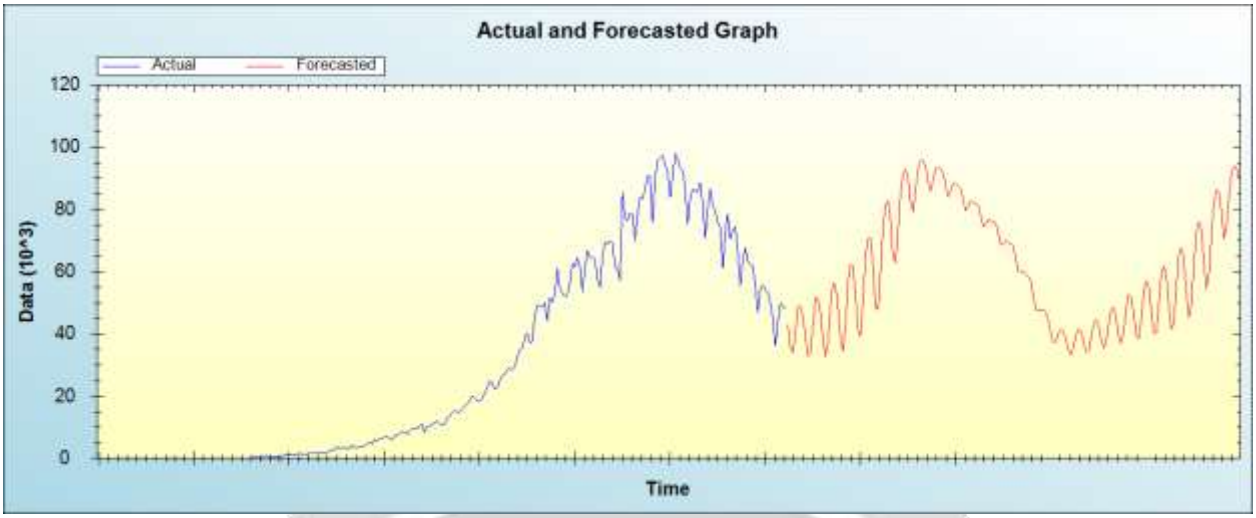


Figure 3 shows the in-sample forecast for C series.

Out-of-Sample Forecast for C: Actual and Forecasted Graph

Figure 4: Out-of-sample forecast for C: actual and forecasted graph



Out-of-Sample Forecast for C: Forecasts only

Table 3: Tabulated out-of-sample forecasts

Day/Month/Year	Forecasts
31/10/20	43109.7924
01/11/20	36010.8898
02/11/20	33707.4065
03/11/20	40338.4084
04/11/20	48527.4481
05/11/20	49062.9651
06/11/20	45742.8945
07/11/20	39447.2645
08/11/20	33006.9959
09/11/20	33571.9096
10/11/20	43272.1393
11/11/20	51461.3038
12/11/20	51435.4261
13/11/20	46769.5776
14/11/20	37935.1363

15/11/20	32811.2946
16/11/20	36938.0280
17/11/20	49341.1289
18/11/20	56255.3411
19/11/20	55701.8530
20/11/20	48914.9528
21/11/20	38187.3765
22/11/20	34692.4020
23/11/20	42705.3613
24/11/20	56796.3079
25/11/20	62438.0063
26/11/20	61779.9097
27/11/20	52726.9646
28/11/20	40851.9863
29/11/20	39136.3191
30/11/20	50422.6514
01/12/20	65134.8236
02/12/20	70686.1188
03/12/20	70735.5462
04/12/20	61209.8537
05/12/20	48291.3459
06/12/20	47874.2036
07/12/20	60319.6019
08/12/20	74752.6678
09/12/20	81582.0216
10/12/20	82848.1750

11/12/20	76671.2339
12/12/20	64338.5732
13/12/20	63011.6549
14/12/20	74137.1358
15/12/20	85528.0200
16/12/20	91744.3922
17/12/20	92769.4246
18/12/20	90107.4926
19/12/20	82058.6750
20/12/20	79341.5690
21/12/20	86849.3243
22/12/20	93348.6344
23/12/20	95677.9437
24/12/20	95657.0241
25/12/20	93971.5358
26/12/20	88739.2605
27/12/20	85696.9391
28/12/20	89647.9721
29/12/20	93094.5121
30/12/20	93370.2146
31/12/20	92860.6850
01/01/21	91559.5523
02/01/21	87267.2031
03/01/21	84073.7655
04/01/21	86229.2529
05/01/21	88430.3008

06/01/21	88050.4894
07/01/21	87570.8133
08/01/21	86749.3842
09/01/21	82899.7847
10/01/21	79603.1613
11/01/21	80988.3044
12/01/21	82697.7272
13/01/21	82126.1771
14/01/21	81861.8248
15/01/21	81484.1899
16/01/21	77871.8146
17/01/21	74472.1203
18/01/21	75458.9237
19/01/21	76858.9106
20/01/21	76154.1487
21/01/21	76031.6458
22/01/21	75837.3312
23/01/21	72153.8719
24/01/21	68547.1175
25/01/21	69140.0416
26/01/21	70102.8519
27/01/21	69162.1708
28/01/21	68889.5566
29/01/21	68282.1635
30/01/21	64030.9691
31/01/21	60060.7291

01/02/21	60019.2051
02/02/21	60347.9043
03/02/21	59233.1379
04/02/21	58419.2838
05/02/21	56624.8899
06/02/21	51744.6460
07/02/21	47810.6052
08/02/21	47334.4632
09/02/21	47942.9083
10/02/21	47707.0895
11/02/21	46691.4356
12/02/21	44215.4953
13/02/21	39947.8307
14/02/21	37101.1358
15/02/21	37725.5287
16/02/21	40218.9237
17/02/21	41477.9917
18/02/21	40759.5817
19/02/21	38145.5307
20/02/21	34687.0487
21/02/21	33348.8919
22/02/21	35848.6530
23/02/21	39867.1664
24/02/21	41651.5993
25/02/21	40580.4631
26/02/21	37398.1204

27/02/21	34270.1037
28/02/21	34431.2601
01/03/21	38695.4655
02/03/21	43358.5476
03/03/21	44715.2318
04/03/21	42706.8074
05/03/21	38451.9371
06/03/21	35552.7250
07/03/21	37362.2575
08/03/21	43272.2664
09/03/21	47950.0652
10/03/21	48438.1952
11/03/21	44964.1921
12/03/21	39420.1808
13/03/21	37047.6712
14/03/21	40830.6353
15/03/21	48287.3156
16/03/21	52560.4688
17/03/21	51930.5984
18/03/21	46533.7200
19/03/21	39871.3334
20/03/21	38480.6446
21/03/21	44655.1696
22/03/21	53433.0115
23/03/21	57033.9392
24/03/21	55167.4467

25/03/21	47605.9685
26/03/21	40233.7557
27/03/21	40301.0491
28/03/21	49179.4914
29/03/21	58858.9805
30/03/21	61810.6004
31/03/21	58825.2572
01/04/21	49235.3827
02/04/21	41493.1506
03/04/21	43420.7449
04/04/21	54826.4039
05/04/21	65013.1245
06/04/21	67827.4462
07/04/21	64343.4732
08/04/21	53477.8339
09/04/21	45282.7349
10/04/21	49138.8828
11/04/21	61938.2069
12/04/21	72620.7912
13/04/21	76209.1835
14/04/21	73675.3906
15/04/21	63557.1635
16/04/21	54439.4615
17/04/21	59293.3504
18/04/21	71368.3471
19/04/21	82034.6632

20/04/21	86417.8056
21/04/21	85830.5077
22/04/21	79472.0245
23/04/21	70845.5816
24/04/21	74262.2474
25/04/21	83801.8097
26/04/21	91209.4485
27/04/21	93869.8879
28/04/21	93766.7310
29/04/21	90393.6804

Figure 5: Graphical presentation of out-of-sample forecasts

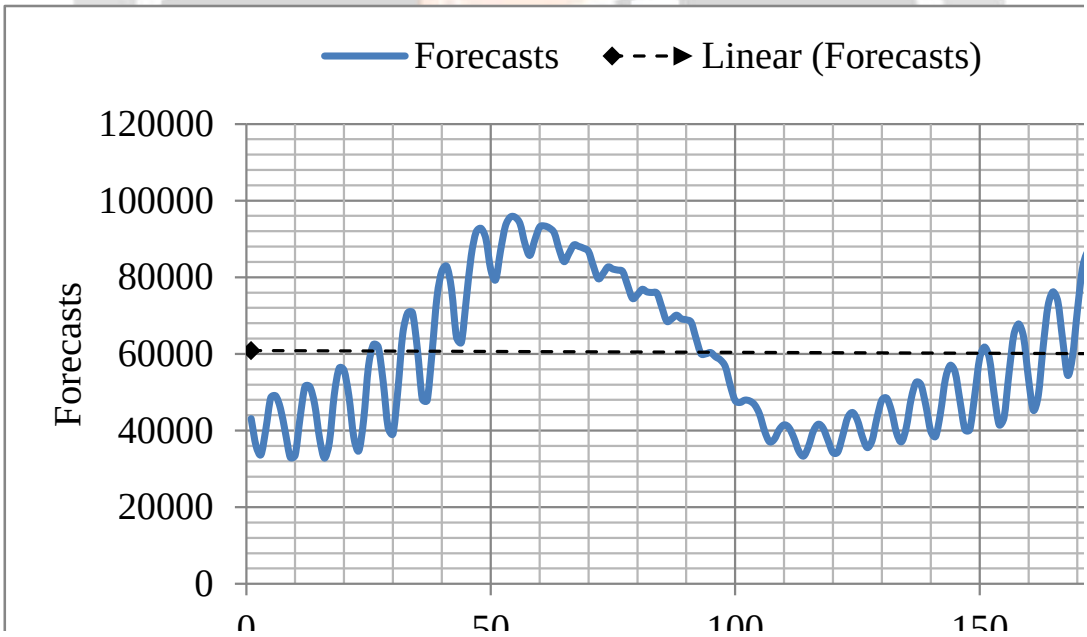


Figure 1 shows the descriptive statistics of the series under consideration. Worthy to note is the fact that the average daily new COVID-19 cases for the country; over the study period was around 29589 cases per day. The maximum is 97894 cases. Table 2 is the ANN model summary. Figure 2 shows the residual analysis of the model: it is clear that our model is stable. Figure 3 shows in-sample forecasts while figure 4 and figure 5 as well as table 3 show out-of-sample predictions. Our results indicate that the COVID-19 pandemic is likely to remain serious in India. The fact that daily new cases will remain around 60000 implies that there is still need for the implementation of serious control and preventive measures, especially on the part of the government. Any deviation from serious control and preventive measures is likely to result in a

more catastrophic wave of infections. The predictions of this are consistent with previous studies such as Khajanchi & Sarkar (2020), Suvarna *et al*. (2020) and Tamang *et al*. (2020).

CONCLUSION & RECOMMENDATIONS

COVID-19, declared as a pandemic by WHO, is currently the most aggressive disease in the world, affecting everyone in this world, in one way or the other. India, the second most populous country in the world, has not been spared by this scourge. Indeed, COVID-19 has become a major threat in India. It has become inevitable to forecast the daily new COVID-19 cases in the country: this information is important for policy makers to be in a position to control the pandemic, particulary within the next 6 months. The study made use of a simple ANN model, based on 275 observations of daily new COVID-19 cases. We strongly recommend that the government of India should ensure strict adherence to lock-down measures while creating awareness about the COVID-19 pandemic. We also urge Indians to be more responsible, especially with regards to social distancing, washing of hands and wearing of masks.

REFERENCES

[1] Arora, P., Kumar, H., & Panigrahi, B. K. (2020). Prediction and Analysis of COVID-19 Positive Cases Using Deep Learning Models: A Descriptive Case Study of India, *Chaos, Solitons and Fractal*, 139: 1 – 9.

[2] Barkur, G., & Vibha, G. B. K. (2020). Sentiment Analysis of Nationwide Lockdown due to COVID-19 Outbreak: Evidence From India, *Asian Journal of Psychiatry*, pp: 1 – 9.

[3] Chatterjee, K., *et al*. (2020). Healthcare Impact of COVID-19 Epidemic in India: A Stochastic Mathematical Model, *Medical Journal of Armed Forces India*, pp: 1 – 14.

[4] Chatterjee, P., *et al*. (2020b). The 2019 Novel Coronavirus Disease (COVID-19) Pandemic: A Review of the Current Evidence, *Indian Journal of Medical Research*, 151 (2): 147 – 162.

[5] COVID-19 Repository By the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.

[6] Dhamodharavadhani, S., Rathipriya, R., & Chatterjee, J. M. (2020). COVID-19 Mortality Rate Prediction for India Using Statistical Neural Network Models, *Frontiers in Public Health*, 8: 1 – 12.

[7] Dong, E., *et al*. (2020). An Interactive Web-based Dashboard to Track COVID-19 in Real Time, *Lancet Infectious Diseases*, 20 (5): 533 – 534.

[8] Giordano, G., *et al*. (2020). Modelling the COVID-19 Epidemic and Implementation of Population-wide Interventions in Italy, *Nature Medicine*, 26: 855 – 860.

[9] Goswami, K., Bharali, S., & Hazarika, J. (2020). Projections for COVID-19 Pandemic in India and the Effect of Temperature and Humidity, *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14: 801 – 805.

[10] India COVID-19 Tracker (2020). https://www.covid19india.org/ Retrieved 2020-06-03.

[11] Khajanchi, S., & Sarkar, K. (2020). Forecasting the Daily and Cumulative Number of Cases for the COVID-19 Pandemic in India, *Chaos*, 30: 1 – 17.

[12] Khan, F. M., & Gupta, R. (2020). ARIMA and NAR Based Prediction Model for Time Series Analysis of COVID-19 Cases in India, *Journal of Safety Science and Resilience*, 1: 12 – 18.

[13] Kotwal, A., *et al*. (2020). Predictive Models of COVID-19 in India: A Rapid Review, *Medical Journal of Armed Forces India*, pp: 1 – 10.

[14] Le, T. T. *et al*. (2020). The COVID-19 Vaccine Development Landscape, *Nature - Review of Drug Discovery*, 19 (5): 305 – 306.

[15] Li, Q., *et al*. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-infected Pneumonia, *New England Journal of Medicine*, 89 (15): 1 – 12.

[16] Mahajan, A., Sivadas, N. A., & Solanki, R. (2020). An Epidemic Model SIPHERD and its Application for Prediction of the Spread of COVID-19 Infection in India, *Chaos, Solitons and Fractals*, 140: 1 – 6.

[17] Malavika, B., *et al*. (2020). Forecasting COVID-19 Epidemic in India and High Incidence States Using SIR and Logistic Growth Models, *Clinical Epidemiology and Global Health*, pp: 1 – 8.

[18] Marimuthu, Y., *et al*. (2020). COVID-19 and Tuberculosis: A Mathematical Model Based Forecasting in Delhi, India, *Indian Journal of Tuberculosis*, 67: 177 – 181.

[19] Patrikar, S., *et al*. (2020). Projections for Novel Coronavirus (COVID-19) and Evaluation of Epidemic Response Strategies for India, *Medical Journal of Armed Forces India*, 76: 268 – 275.

[20] Rafiq, D., Suhail, S. A., & Bazaz, M. A. (2020). Evaluation and Prediction of COVID-19 in India: A Case of Worst Hit States, *Chaos, Solitons and Fractals*, 139: 1 – 9.

[21] Rath, S., Tripathy, A., & Tripathy, A. R. (2020). Prediction of New Active Cases of Coronavirus Disease (COVID-19) Pandemic Using Multiple Linear Regression Model, *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14: 1467 – 1474.

[22] Rawaf, S., Yamamoto, H. O., & Rawaf, D. (2020). Unlocking Towns and Cities: COVID-19 Exit Strategy, *East Mediterranean Health Journal*, 26 (5): 499 – 502.

[23] Ray, D., *et al*. (2020). Predictions, Role of Interventions and Effects of a Historic National Lockdown in India's Response to the COVID-19 Pandemic: Data Science Call to Arms, *Harvard Data Science Review*, 1: 1 – 15.

[24] Roy, S., Bhunia, G. S., & Shit, P. K. (2020). Spatial Prediction of COVID-19 Epidemic Using ARIMA Techniques in India, *Modeling Earth Systems and Environment*, pp: 1 – 7.

[25] Salgotra, R., Gandomi, M., & Gandomi, A. H. (2020). Time Series Analysis and Forecast of the COVID-19 Pandemic in India Using Genetic Programming, *Chaos, Solitons and Fractals*, 138: 1 – 15.

[26] Sardar, T., *et al*. (2020). Assessment of Lockdown Effect in Some States and Overall India: A Predictive Mathematical Study on COVID-19 Outbreak, *Chaos, Solitons and Fractals*, 139: 1 – 10.

[27] Sarkar, K., Khajanchi, S., & Nieto, J. J. (2020). Modelling and Forecasting the COVID-19 Pandemic in India, *Chaos, Solitons and Fractals*, 139: 1 – 16.

[28] Singhal, T. (2020). A Review of Coronavirus Disease-19 (COVID-19), *Indian Journal of Pediatrics*, 87: 281 – 286.

[29] Suvarna, B., *et al*. (2020). Machine Learning Algorithm for Predicting Number of COVID-19 Cases, *Journal of Xi'an University of Architecture & Technology*, XII (VI): 1175 – 1186.

[30] Tamang, S. K., Singh, P. D., & Datta, B. (2020). Forecasting of COVID-19 Cases Based on Prediction Using Artificial Neural Networks Curve Fitting Technique, *Global Journal of Environmental Science and Management*, 6 (4): 1 – 12.

[31] WHO (2020). https://www.who.int/emergencies/ Retrieved 2020-04-01.

[32] Worldometers (2020). https://www.worldometers.info/coronavirus/ Retrieved 2020-09-06.

[33] Wu, F., *et al*. (2020). A New Corona Virus Associated With Human Respiratory Disease in China, *Nature*, 579 (7798): 265 – 269.

[34] Zhou, P., *et al*. (2020). A Pneumonia Outbreak Associated With a New Coronavirus of Probable Bat Origin, *Nature*, 579 (7798): 270 – 273.

[35] Zhu, N., *et al*. (2020). A Novel Coronavirus From Patients With Pneumonia in China, *New England Journal of Medicine*, 89 (15): 1 – 6.