Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 77(3) 2023 243–248

Prediction of Urban Unemployment Rate in India using Grey Model

Pradip Basak, Mrinmoy Ray, Kanchan Sinha and Anuja A.R.
[1] *Uttar Banga Krishi Viswavidyalaya, Cooch Behar*
[2] *ICAR-Indian Agricultural Statistics Research Institute, New Delhi*
[3] *ICAR-Central Marine Fisheries Research Institute, Kochi*

Received 19 October 2022; Revised 27 June 2023; Accepted 03 August 2023

SUMMARY

Urban Unemployment Rate (UR) is a crucial indicator representing the livelihood of people in India. In India, the quarterly estimates of urban UR in the Current Weekly Status (CWS) are released by National Statistics Office (NSO) through Periodic Labour Force Survey (PLFS). At present, the urban UR estimates are available in India from the quarter April-June 2018 to January-March 2023 at the state and national level. Accurate forecasting of the UR is essential for early identification of the socio-economic problems so that timely and targeted intervention, and proper policy planning can be done to reduce the same. Time series methodology utilised so far for the forecasting of UR require monthly or quarterly data of sufficient length. Therefore, the usual methods of forecasting of UR may not yield reliable forecast in this type of small time series as the assumption on the data requirement will be violated. As a superiority to conventional statistical models, grey models require very limited data to build a forecast model (Deng, 1989). In this article, application of grey model has been considered on the quarterly estimates of urban UR for forecasting the unemployment in urban India. The Grey model shows excellent performance in forecasting the urban UR at the national level and at the state level, it shows good performances for most of the states.

Keywords: Grey model; Periodic Labour Force Survey; Urban unemployment rate; Forecast model.

1. INTRODUCTION

The unemployment in any nation has several adverse effects on the economy. According to International Labour Organisation (ILO, 2022), it is projected that the status of unemployment at global level in 2022 stands at 207 million. India has been facing the menace of unemployment for years due to increasing working population, technological advancements causing substitution of labour by capital and uneven development pattern (Chand *et al.*, 2017). Furthermore, the problem of unemployment leads to the problem of poverty (Dev and Venkatanarayana, 2011). People unemployed for a long time may indulge in illegal and wrong activities for earning money which increases crime in the country (Mittal *et al.*, 2019). Due to unemployment the workforce that could have been gainfully employed to generate resources actually gets dependent on the remaining working population, thus escalating socio-economic costs for the state. Thus, unemployment indicates the health of the economy and UR is the most frequent measure of unemployment. Considering the importance of availability of labour force data at more frequent time intervals, National Statistical Office (NSO) launched Periodic Labour Force Survey (PLFS) on April, 2017 with the twin objectives: (i) to estimate the key employment and unemployment indicators (viz. Worker Population Ratio (WPR), Labour Force Participation Rate (LFPR), UR) in the short time interval of three months for the urban areas only in the CWS, and (ii) to estimate employment and unemployment indicators in both usual status (ps+ss) and CWS in both rural and urban areas annually. According to NSO, UR is defined as the percentage of unemployed persons in the labour force. The estimates of UR in CWS gives an average picture of unemployment in a short period of 7 days during the survey period. According to the CWS approach, a person was considered as

Corresponding author: Kanchan Sinha
E-mail address: kanchan.sinha@icar.gov.in

unemployed in a week if he/she did not work even for 1 hour during the week but sought or was available for work at least for 1 hour during the week. The Quarterly Bulletin of PLFS provides estimates of labour force indicators including UR in the short time interval of three months for the urban areas only in the CWS starting from the quarter April-June, 2018. Therefore, the quarterly estimates of urban UR in the CWS are available from the quarter April-June, 2018 to January-March, 2023 at the state and national level in India.

Though PLFS is a comprehensive data set employed to estimate trends in labour market, the available data has lag between survey and reporting (9 months for the quarterly PLFS for urban sector) (Economic survey, 2021-22). Many advanced economies such as USA, United Kingdom and Japan provides their published data in less than two months of survey as far as labour indicators are concerned (Economic Survey, 2021-22). Therefore, many a times, proxies such as work demand under MGNREGA or EPFO subscriptions are frequently employed to assess recent trends in employment (Economic survey, 2021-22). This also signifies the need to have a reliable and working forecasting model to predict the future trends of labour markets, especially the UR, with more accuracy. There are different methods to forecast UR, say for example, moving averages, simple exponential smoothing, double exponential smoothing (also known as Holt's method), Holt-Winters Smoothing and Autoregressive Integrated Moving Average (ARIMA) model. Voineagu *et al.* (2012) employed econometric smoothing methods based on a modified version of Holt's approach to predict monthly UR using data from a series of Labour Force Surveys conducted by the Romanian National Institute of Statistics between 2004 and 2011. Dumicic *et al.* (2015) used many approaches, including double exponential smoothing and the Holt-Winters method, to predict the UR in selected European nations using quarterly UR data from January-March 2001 to October-November 2013. Vicente *et al.* (2015) applied the ARIMAX model to the monthly dataset of Registered Unemployment Statistics from January 2004 to December 2012, published by the Spanish Ministry of Employment and Social Security, as well as explanatory variables from both the demand and supply sides of the labour market. Jaffur *et al.* (2016) studied the out-of-sample forecasting ability of several linear and nonlinear univariate time series models, such as ARIMA and GARCH, using the monthly seasonally adjusted Canadian URs from 1980 to 2013. Chakraborty *et al.* (2020) proposed an integrated strategy by combining the ARIMA and autoregressive neural network models, that can more reliably forecast URs. They have used the monthly data on URs of length more than 400 for Canada, Germany, Japan, Netherlands, Sweden, and Switzerland whereas for New Zealand it is quarterly data of length 132. Therefore, it is evident that all these methods of forecasting UR need time series data of sufficient length.

In India, quarterly estimates of urban UR are available for twenty quarters (April-June, 2018 to January-March, 2023). Therefore, the typical approaches for predicting UR may not be applicable to this type of small time series data, as the assumption about required number of observations will be violated. As an advantage over conventional statistical models, the Grey model takes less data to construct a differential forecasting model (Deng, 1989). The minimum number of observations required to build a Grey model is four. Kayacan *et al.* (2010) examined the accuracy of several Grey models for forecasting the United States dollar to Euro exchange rate using data from 01.01.2005 to 30.12.2007. Lin (2013) used the Grey model to predict the diffusion of mobile cellular broadband and fixed broadband in Taiwan based on annual mobile cellular broadband penetration data from 2005 to 2011. Different applications of the Grey model, its modification and comparison with other methods are available in literatures, see for example Muqtadir *et al.* (2016) and Yang *et al.* (2018). Sinha *et al.* (2020) applied the Grey model for forecasting the annual production of rice, wheat, maize, total oilseeds and total pulses in the Uttarakhand state of India using yearly data from 2000 to 2016. Therefore, application of the Grey model may be envisaged to forecast the urban UR at the state and national level in India based on limited quarterly data available.

2. DATA DESCRIPTION

This section presents the basic sources of data i.e., survey data which is being used to forecast the urban UR at the state and national level. The quarterly estimates of urban UR are collected from the Quarterly Bulletin, PLFS published by NSO, India. The estimates are available for the April-June, 2018 to January-March, 2023 for different age groups and gender at the state and national level. Different age groups are

15-29 years, 15 years and above, and all age group. Since the persons belonging to age group 15-29 years is the country's youth population and the UR was highest in this age group, therefore, gender specific analysis was focused for this age group.

3. METHODOLOGY

3.1 Grey model

Deng (1989) proposed the Grey model as a solution to the "short sample size" and "poor information quality" related uncertainty issue. In grey models, it is assumed that all input data values are positive and that the sampling frequency of the time series is fixed. The most prevalent type of grey model in the literature is GM (1,1), which is pronounced as "Grey Model First Order One Variable." This is a model for predicting small time series. The coefficients of the differential equations of the GM (1,1) model fluctuate with time. In other words, the model is updated whenever the prediction model receives new data. Since estimates of urban UR based on PLFS are available for twenty quarters at the state and national levels, and all the data points are positive, therefore, grey models can be used to predict the urban UR. Let the original data sequence be $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, ..., x_n^{(0)})$. The original time series data sequence, $\mathbf{x}^{(0)}$ is then subjected to an operation, called Accumulating Generation Operation (AGO) to smooth the randomness present in the data. The new data sequence $\mathbf{x}^{(1)} = (x_1^{(1)}, x_2^{(1)}, ..., x_n^{(1)})$ is generated from $\mathbf{x}^{(0)}$ by the AGO as follows:

$$\mathbf{x}_k^{(1)} = \sum_{i=1}^{k} x_i^{(0)}, \quad k = 1, 2, ..., n.$$

The differential equations of the GM (1,1) model are given by

$$\frac{dx_k^{(0)}}{dk} + \alpha x_k^{(1)} = \beta,$$

where, β and α are the developing coefficient and control variable respectively. The parameters of the grey model $\boldsymbol{\theta} = (\alpha, \beta)'$ can be estimated as follows:

$$\boldsymbol{\theta} = \left(\mathbf{B}'\mathbf{B}\right)^{-1} \mathbf{B}'\mathbf{Y}$$

where, $\mathbf{B} = \left[-\mathbf{z}\ \mathbf{1}_n\right]$, $\mathbf{z} = \left(z_1, z_2, ..., z_n\right)'$,

$$z_k = \delta x_k^{(1)} + (1-\delta)x_{k-1}^{(1)}, \; k = 2, 3, ..., n,$$

$\mathbf{1}_n$ is the unit vector of length n and $\mathbf{Y} = (x_2^{(0)}, x_3^{(0)}, ..., x_n^{(0)})$.

Here δ is adjustment coefficient and usually specified as 0.5 for convenience, but this is not the optimal setting. From the Grey differential equation, the solution of $x_k^{(1)}$ at time $(i+1)$ is given by

$$x_{i+1}^{(p)} = \left[x_1^{(0)} - \frac{\beta}{\alpha}\right] e^{-\alpha i} + \frac{\beta}{\alpha}.$$

To obtain the predicted value of the original data at time $(i+1)$, the Inverse AGO is used to establish the following grey model,

$$x_{i+1}^{(p)} = \left[x_1^{(0)} - \frac{\beta}{\alpha}\right] e^{-\alpha i}(1 - e^{\alpha}),$$

and the predicted value of the original data at time $(i+H)$ is given by

$$x_{i+H}^{(p)} = \left[x_1^{(0)} - \frac{\beta}{\alpha}\right] e^{-\alpha(i+H-1)}(1 - e^{\alpha}).$$

3.2 Model accuracy evaluation method

In order to evaluate the forecasting performance of the Grey model, Relative Mean Absolute Percentage Error (RMAPE) is used as a criterion. The RMAPE is computed as

$$RMAPE = \frac{1}{n-1} \sum_{i=2}^{n} \frac{\left|x_i^{(0)} - x_i^{(p)}\right|}{x_i^{(0)}} \times 100\%,$$

where, $x_i^{(0)}$ and $x_i^{(p)}$ represent the actual and predicted value of the urban UR, respectively.

4. RESULTS AND DISCUSSIONS

The time plots of quarterly urban UR estimates are generated using the PLFS data from the quarter April-June, 2018 to January-march, 2023 for the male, female and persons of age15-29years.The time plots are generated for the major Indian states like Gujarat, Maharashtra, Tamil Nadu, Uttar Pradesh and West Bengal as well as at the all-India level. These plots are presented in the Fig.1 to Fig.3.

The time plot in Fig. 1 shows that there was a decreasing trend in urban UR for male persons of age 15-29 years from the quarter April-June, 2018 to October-December, 2019 followed by a drastic increase during the quarter April-June, 2020 which coincided with the covid-19 pandemic period. However, subsequent to this, urban URs show a decreasing trend at the state and national level. Similar patterns were also observed for urban URs of female and all persons of age 15-29 years. The urban UR was found to be lowest in Gujarat for male, female and persons of age 15-29 years for

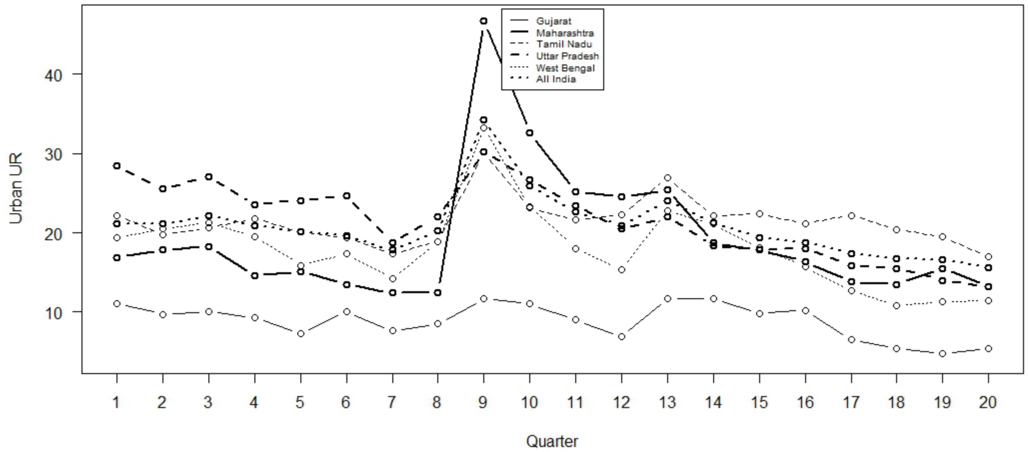

Fig. 1. Time plot of quarterly urban UR estimates for male persons of age 15-29 years from April-June, 2018 to January-march, 2023 for the major Indian states as well as at all India

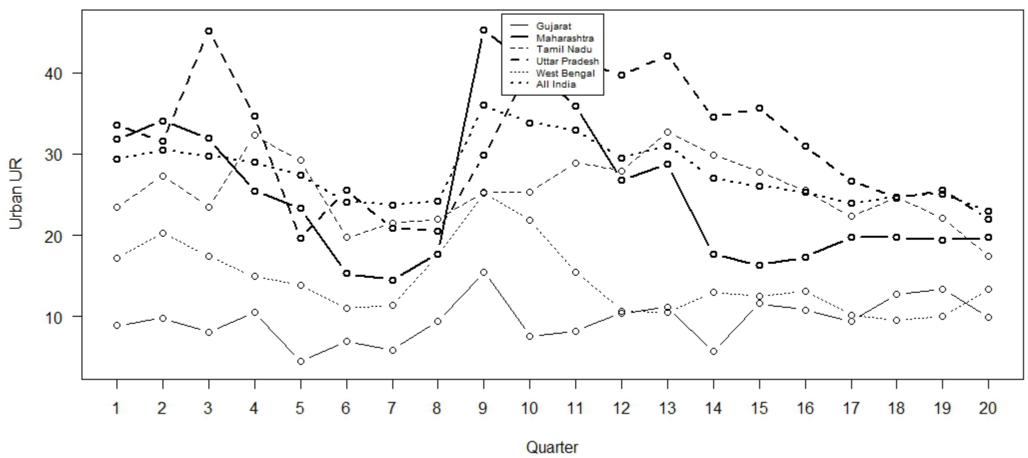

Fig. 2. Time plot of quarterly urban UR estimates for female persons of age 15-29 years from April-June, 2018 to January-march, 2023 for the major Indian states as well as at all India

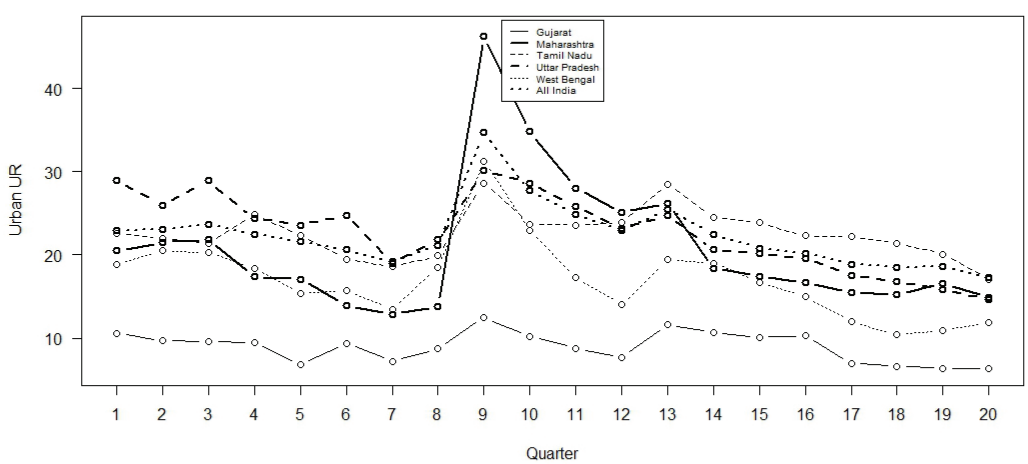

Fig. 3. Time plot of quarterly urban UR estimates for all persons of age 15-29 years from April-June, 2018 to January-march, 2023 for the major Indian states as well as at all India

Table 1. Forecasted urban UR for male, female and all persons of age 15-29 years for the quarter April-June, 2023 along with RMAPE

State	15-29 years male		15-29 years female		15-29 years person	
	Forecast April-June 2023	RMAPE	Forecast April-June 2023	RMAPE	Forecast April-June 2023	RMAPE
Andhra Pradesh	21.35	14.76	25.50	12.82	9.01	20.05
Assam	17.99	16.32	54.66	20.04	8.59	16.03
Bihar	23.55	13.17	36.81	13.86	9.61	12.39
Chhattisgarh	30.61	11.44	33.91	11.75	13.45	16.51
Delhi	11.19	32.04	19.06	35.72	5.89	24.36
Gujarat	7.33	21.37	11.73	23.85	3.76	20.44
Haryana	23.52	12.43	44.01	27.13	11.38	16.77
Himachal Pradesh	24.25	22.77	41.54	29.85	11.49	15.14
Jharkhand	21.01	26.97	24.45	29.07	10.52	35.31
Karnataka	12.28	15.73	20.57	41.72	5.60	23.73
Kerala	29.81	15.90	46.14	6.48	14.16	22.67
Madhya Pradesh	20.78	13.11	23.56	14.20	9.59	19.43
Maharashtra	18.28	30.60	19.82	27.80	9.10	46.08
Odisha	26.85	11.20	34.45	15.36	11.33	16.59
Punjab	19.65	11.56	27.14	14.36	8.29	10.40
Rajasthan	30.70	10.71	54.01	15.56	14.14	13.12
Tamil Nadu	21.42	9.59	23.92	13.29	8.97	19.25
Telangana	20.66	9.70	23.30	10.50	8.82	17.14
Uttarakhand	28.25	22.20	28.22	17.42	11.82	24.33
Uttar Pradesh	15.35	9.75	29.23	23.76	7.69	14.03
West Bengal	14.27	18.87	10.89	19.18	5.56	24.78
All-India	18.36	11.93	25.33	9.47	8.50	18.16

most of the quarters. However, it is interesting to note that the gap between male and female UR is quite high for majority of the states. The out of sample forecast values of urban UR during the quarter April-June, 2023 for male, female and all persons of age 15-29 years are generated from the GM (1,1) model and are presented in Table 1 along with RMAPE value of the prediction.

From Table 1, it is evident that RMAPE values of predicted urban UR for the male persons of age 15-29 years indicate that grey model gives satisfactory forecast for the states of Andhra Pradesh, Bihar, Chhattisgarh, Haryana, Madhya Pradesh, Odisha, Punjab, Rajasthan as well as at all India level. However, it gives good performance for the state of Tamil Nadu, Telangana and Uttar Pradesh. For female persons of age 15-29 years, RMAPE values depict that grey model generates satisfactory forecast of urban UR for the states of Andhra Pradesh, Bihar, Chhattisgarh, Madhya Pradesh, Punjab, Tamil Nadu and Telangana. For the state of Kerala and at all-India level it generates good forecast. However, for persons of age 15-29 years, grey model generates satisfactory forecast of urban UR for the state of Bihar, Himachal Pradesh, Punjab, Rajasthan and Uttar Pradesh. During the quarter April-June, 2023, maximum and minimum urban UR for male person of age 15-29 years is predicted to be 30.70 and 7.33 which corresponds to the state of Rajasthan and Gujarat respectively. For female person of age 15-29, maximum and minimum UR is expected to be 54.66 and 10.89 during the quarter April-June, 2023 corresponding to the state of Assam and West Bengal respectively. From these results, it is quite evident that urban UR for female persons is expected to be much higher than male persons except for the state of West Bengal. In particular, for the state of Assam this difference in urban UR between male and female persons is quite high. Earlier literatures also agree with the result as in general it is reported that female is more unemployed compared to male and the gap is more evident in urban areas. (Sinha, 2013; Nepram *et al.*, 2021)

5. CONCLUSION

In this study, grey model has been utilised to predict the urban UR at the state and national level using PLFS estimates. As the grey model provides reasonable quarterly estimates of urban UR with the limited available data, the model's potential use in policy planning is more promising. These advanced estimates of quarterly urban UR generated from the grey model will aid policymakers and Government agencies in designing proactive and targeted interventions. At the state level, majority of the states showed satisfactory performance in terms of forecast accuracy. However, the performance of the grey model may not be sufficient at the state level when estimates of urban UR have significant variations in the past, as shown by the findings for several states. Still grey model is a viable solution for generating out of sample forecast in this type of small time series data with reasonable precision. However, there is possibility of considerable development in this aspect.

ACKNOWLEDGEMENTS

The authors are profoundly thankful to the editors and the anonymous referee for their valuable comments and suggestions which led to improvements in the article. The authors also gratefully acknowledge the valuable contributions made by Late Dr. Hukum Chandra, National Fellow and Principal Scientist, ICAR-IASRI, New Delhi.

REFERENCES

Chakraborty, T., Chakraborty, A.K., Biswas, M. Banerjee, S. and Bhattacharya, S. (2021). Unemployment Rate Forecasting: A Hybrid Approach, *Computational Economics*, **57**,183-201.

Chand, K., Tiwari, R. and Phuyal, M. (2017). Economic Growth and Unemployment Rate: An Empirical Study of Indian Economy, *PRAGATI: Journal of Indian Economy*, **4**, 130-137.

Deng, J. L. (1989). Introduction to grey system theory. *The Journal of Grey System*, **1**, 1-24.

Dev, S. M. and Venkatanarayana, M. (2011). Youth employment and unemployment in India. *Mumbai: Indira Gandhi Institute of Development Research*. Available at http://www.igidr.ac.in/pdf/publication/WP-2011-009.pdf accessed on October 10, 2022.

Dumicic, K., Ceh Casni, A. and Zmuk, B. (2015). Forecasting unemployment rate in selected European countries using smoothing methods. *World Academy of Science, Engineering and Technology: International Journal of Social, Education, Economics and Management Engineering*, **9**(4), 867-872.

Government of India (2022). Economic Survey 2021-22. Economic Division, Department of Economic Affairs, Ministry of Finance, Government of India, New Delhi.

International Labour Organisation (2022). World Employment and Social Outlook Trends 2022. International Labour Office, Geneva.

Kayacan, E., Ulutas, B. and Kaynak, O. (2010). Grey system theory-based models in time series prediction. *Expert Systems with Applications*, **37**, 1784-1789.

Khan Jaffur, Z.R., Sookia, N.U.H., Nunkoo Gonpot, P., and Seetanah, B. (2017). Out-of- sample forecasting of the Canadian unemployment rates using univariate models. *Applied Economics Letters*, **24**(15), 1097-1101.

Lin, C.S. (2013). Forecasting and analyzing the competitive diffusion of mobile cellular broadband and fixed broad-band in Taiwan with limited historical data. *Economic Model*, **35**, 207-213.

Mittal, M., Goyal, L.M., Sethi, J.K. and Hemanth, D.J. (2019). Monitoring the Impact of Economic Crisison Crime in India using Machine Learning. *Computational Economics*, **53**, 1467-1485.

Muqtadir, A., Suryono and Gunawan, V. (2016). The Implementation of Grey Forecasting Model for Forecast Result's Food Crop Agricultural. *Scientific Journal of Informatics*, **3**(2), 159-166.

National Statistical Office (2019). *Quarterly Bulletin, Periodic Labour Force Survey (PLFS): January-March, 2019*. National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India.

National Statistical Office (2020). *Quarterly Bulletin, Periodic Labour Force Survey (PLFS): July-September 2020*. National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India.

National Statistical Office (2021). *Quarterly Bulletin, Periodic Labour Force Survey (PLFS): July-September 2021*. National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India.

National Statistical Office (2022). *Quarterly Bulletin, Periodic Labour Force Survey (PLFS): July-September 2022*. National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India.

National Statistical Office (2023). *Quarterly Bulletin, Periodic Labour Force Survey (PLFS): January-March 2023*. National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India.

Nepram, D., Singh, S.P., and Jaman, S. (2021). The Effect of Government Expenditure on Unemployment in India: A State Level Analysis. *The Journal of Asian Finance, Economics and Business*, **8**(3), 763-769.

Sinha, K. and Sahu, P.K. (2020). Forecasting Short Time Series using Rolling Grey Bayesian Framework. *International Journal of Statistical Sciences*, **20**(2), 207-224.

Sinha, P. (2013). *Combating youth unemployment in India*. New Delhi: Friedrich-Ebert-Stiftung, Department for Global Policy and Development.

Vicente, M.R., López-Menéndez, A.J., and Pérez, R. (2015). Forecasting unemployment with internet search data: Does it help to improve predictions when job destruction is skyrocketing? *Technological Forecasting and Social Change*, **92**, 132-139.

Voineagu, V., Pisica, S. and Caragea, N. (2012). Forecasting Monthly Unemployment by Econometric Smoothing Techniques. *Journal of Economic Computation and Economic Cybernetics Studies and Research*, **46**(3), 255-267.

Yang, X., Zou, Kong, D. and Jiang, G. (2018). The analysis of GM (1, 1) grey model to predict the incidence trend of typhoid and paratyphoid fevers in Wuhan City, China, *Medicine*, **97**(34), 1-5.