https://doi.org/10.56093/JISAS.V78I1.6
Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 78(1) 2024 47–52

Prediction of Annual Rural Unemployment Rate in West Bengal using Grey Model

Pavithra V., Deb Sankar Gupta, Pradip Basak, Manoj Kanti Debnath and Gobinda Mula

Uttar Banga Krishi Viswavidyalaya, Cooch Behar

Received 10 November 2023, Revised 22 March 2024; Accepted 16 April 2024

SUMMARY

The rural unemployment rate is a critical economic indicator used to assess strength of rural economy in India. Annual estimates of rural UR are released in both usual status (ps+ss) as well as current weekly status (CWS) at the state and national level in India by National Statistics Office (NSO) through Periodic Labour Force Survey (PLFS). At present, the annual rural UR estimates are available for the state of West Bengal from the year 2017-18 to 2021-22. However, there is a notable delay in the publication of UR estimates as compared to the reference period. Therefore, accurate forecasting of the UR is crucial for timely and targeted interventions, and effective policy planning. Conventional forecasting models fail to provide accurate predictions of UR in these type of small time series due to the violation of requirement of number of data points. In contrast, the Grey model requires limited data to establish a differential forecasting model. In this study, application of grey model has been considered to forecast annual rural UR in West Bengal for different age groups as well as gender, and it was found that grey model provides satisfactory forecast.

Keywords: Unemployment rate (UR); Periodic Labour Force Survey (PLFS); Grey model; Forecast.

1. INTRODUCTION

Unemployment rate (UR) is one of the key economic indicators used to gauge the strength of an economy. It is the percentage of unemployed individuals in an economy that are currently in the labour force. UR is one of the most widely monitored indicators by policymakers, investors, and public. In most countries, UR is determined at various levels, including the national, state, and regional levels, through labour-force surveys conducted by national statistical institutes. International organizations such as the Organisation for Economic Co-operation and Development (OECD), International Monetary Fund (IMF), and World Bank also continuously calculate and record the national URs of numerous countries around the world. However, in contrast to this approach, and recognizing the importance of having more frequent labour force data, National Statistical Office (NSO) in India introduced the Periodic Labour Force Survey (PLFS) in April 2017 with two primary objectives: (i) measuring the dynamics in labour force participation and employment status within a short time interval of three months, specifically for urban areas in the current weekly status (CWS), (ii) measuring annual labour force estimates for both rural and urban areas, covering key parameters in both the usual status (ps+ss) and the CWS. In India, annual estimates of UR for both rural and urban areas are available from PLFS at the state and national level since 2017-18. However, there is a considerable time lag in the publication of these estimates. Thus, accurate prediction of UR is essential for economic decision-making and effective policy planning to facilitate the early identification of socio-economic problems and development of strategies to combat the unemployment.

Different forecasting methods such as moving averages, simple exponential smoothing, double exponential smoothing (also known as Holt's method), Holt-Winters Smoothing, Autoregressive Integrated Moving Average (ARIMA) models, Support Vector Machine (SVM), and hybrid models are used to forecast the UR. Voineagu *et al.* (2012) applied econometric

Corresponding Author: Pradip Basak
E-mail address: pradip@ubkv.ac.in

smoothing methods to predict monthly UR using Labour Force Surveys data conducted by the Romanian National Institute of Statistics from 2004 to 2011. Dumicic *et al.* (2015) employed double exponential smoothing and Holt-Winters method, to predict the quarterly UR in selected European nations using the UR data from January-March, 2001 to October-December, 2013. Vicente *et al.* (2015) forecasted the UR in Spain with the ARIMAX model using monthly data of Registered Unemployment Statistics from January, 2004 to December, 2012. Jaffur *et al.* (2016) investigated the 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) developed a hybrid ARIMA-ARNN model to forecast the UR using the monthly data of length more than 400 for Canada, Germany, Japan, Netherlands, Sweden, and Switzerland whereas for New Zealand it is quarterly data of length 132. However, for India in general and West Bengal in particular, annual estimates of rural UR are available for five years from 2017-18 to 2021-22. Therefore, the forecasting methods discussed earlier will fail to provide accurate predictions of UR in these type of small time series data. Deng (1989) proposed Grey model which requires only limited data to build a differential forecasting model. The minimum data requirement for grey models is four observations. Li and Chen (2009) applied Grey model and Grey Markov model to predict UR of graduates in China by using historical data from 1998 to 2006.

Lin (2013) applied 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. Sinha *et al.* (2020) forecasted annual production of rice, wheat, maize, total oilseeds and total pulses in the state of Uttarakhand using the data from 2000 to 2016. Nguyen *et al.* (2021) applied Grey, Grey Verhulst and ARIMA model to forecast the UR in the post COVID-19 period at six different rural and urban regions of Vietnam using the annual UR data from 2014-19. Basak *et al.* (2023) predicted quarterly urban UR in India at the state and national level using Grey model based on quarterly estimates of PLFS data. Therefore, the annual rural URs in West Bengal can be forecasted with reliable accuracy using the Grey model from the limited PLFS data available.

2. DATA DESCRIPTION

The National Statistical Office (NSO), Ministry of Statistics and Programme Implementation (MoSPI), Government of India is the primary source of official statistics in India at the national and state level. In the present study, UR estimates available in the PLFS annual reports from 2017-18 to 2021-22 published by NSO, MoSPI, Government of India are used. This survey data provides estimates of the annual rural UR for different age groups and gender in two approaches:

a. Usual status
b. Current weekly Status (CWS)

In the usual status (ps+ss) approach, UR is described based on principal and subsidiary statuses of the person in the labour force whereas in the CWS approach, the UR gives an average picture of unemployment in a short period of 7 days during the survey period. According to the CWS approach, a person is considered as 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. UR based on both usual status and CWS have significant importance. UR based on usual status approach is appropriate when we are interested in understanding long-term employment patterns and trends. On the other hand, if we need to gauge the immediate state of labour market or to assess the impact of recent events then the UR based on CWS is more suitable. The UR based on usual status is more stable and consistent because it is based on a person's usual labour status over a longer period whereas UR based on CWS approach is more sensitive to short term changes such as seasonal fluctuations or economic shocks.

In PLFS, annual estimates of rural UR in usual status are available for age groups: 15-29 years, 15-59 years, 15 years and above, all age, whereas in CWS approach UR estimates are available for age groups: 15 years and above, all age. The "15 years and above" age group includes persons of age above 15 years whereas "all age" group includes persons of all the ages as per the concepts and definitions of PLFS. The analysis has been carried out using R software.

3. METHODOLOGY:

3.1 Grey model

Grey model is commonly used in time series forecasting when usage of large dataset and statistical assumptions are violated with a limited number of data points. GM (1, 1) type of grey model is the most widely used in the literature, pronounced as "Grey Model First Order One Variable". GM (1, 1) model has a simple form of representation and can only be used in positive data series. Let, the original data series is represented as $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$ which consists of n data points. A new sequence $\mathbf{x}^{(1)} = (x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)})$ can then be generated from $\mathbf{x}^{(0)}$ by accumulated generating operation (AGO) as follows,

$$\mathbf{x}_k^{(1)} = \sum_{j=1}^{k} x_j^{(0)}, \quad k = 1, 2, \ldots, n$$

The new sequence $\mathbf{x}^{(1)} = (x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)})$ can then be approximated as a first order differential equation as,

$$\frac{dx_t^{(1)}}{dt} + \alpha x_t^{(1)} = \beta \tag{1}$$

where, β and α are the developing coefficient and control variable respectively. The AGO is used to identify potential regularities hidden in the data sequences even if the original data are finite, insufficient, and chaotic. The parameters of the grey model $\boldsymbol{\theta} = (\alpha, \beta)'$ are estimated using ordinary least-squares as,

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

where, $\mathbf{B} = \left[-\mathbf{z} \ \mathbf{1}_{\mathrm{n}}\right]$,

$$\mathbf{z} = \left(z_1, z_2, \ldots, z_n\right)',$$

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

$$\mathbf{Y} = (x_2^{(0)}, x_3^{(0)}, \ldots, x_n^{(0)}).$$

Here, δ is adjustment coefficient which is usually specified as 0.5 for convenience. From the Grey differential equation in (1), the solution of $x_k^{(1)}$ at time ($t+1$) is obtained as

$$x_{t+1}^{(1)} = \left[x_1^{(0)} - \frac{\beta}{\alpha}\right] e^{-\alpha t} + \frac{\beta}{\alpha}. \tag{2}$$

Predicted value of the original data series at time ($t+1$) are obtained using Inverse AGO as,

$$x_{t+1}^{(0)} = \left[x_1^{(0)} - \frac{\beta}{\alpha}\right] e^{-\alpha t} (1 - e^{\alpha}) \tag{3}$$

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

$$x_{t+H}^{(0)} = \left[x_1^{(0)} - \frac{\beta}{\alpha}\right] e^{-\alpha(t+H-1)} (1 - e^{\alpha}). \tag{4}$$

3.2 Model accuracy evaluation

Forecasting performance of the Grey model is evaluated in terms of Relative Mean Absolute Percentage Error (RMAPE) and it is computed as

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

where, $x_k^{(0)}$ and $x_k^{(p)}$ represent the actual and predicted value of the annual rural UR at k^{th} year respectively. If RMAPE values are less than 15% then it indicates that Grey model gives a good forecasting accuracy.

4. RESULTS AND DISCUSSIONS

The descriptive statistics of annual rural UR in both usual status and CWS for male, female and all persons of different age groups are presented in Table 1 and 2 respectively. A perusal of Table 1 reveals that the average annual rural UR is highest for age group 15-29 years. Female person has lower UR as compared to the male person for all the age categories. However, variation in female UR is quite high as compared to male person over the different years. It is interesting to note that the difference between male and female UR is quite high for the 15-29 years age group. Estimates of UR are higher in CWS approach as compared to the usual status as well as variability is more in terms of CV. From 2017–18 to 2021–22 average annual rural UR in usual status for male, female and person of all age groups in West Bengal was 4.10, 1.78, and 3.60 respectively, whereas in CWS approach it was 7.98, 5.12, and 7.36 respectively. Even though estimates of UR in 15-59 and all age groups are same for female, however, there is minute difference in estimates of UR in both the age groups for male and person.

Time plots of annual rural UR estimates are generated using the PLFS data from the years 2017-18 to 2021–22 for the male, female, and persons of different age groups in both usual status and CWS for the state of West Bengal. These plots are shown in Fig 1-6. The time plots reveal that annual rural UR was high during the year 2019-20 for the male, female, and persons belonging to all the age groups in both CWS

Table 1. Descriptive statistics of annual rural UR in usual status for male, female and persons of age group 15-29 and 15-59 years

Statistics	15-29 Years			15-59 Years		
	Male	Female	Person	Male	Female	Person
Minimum	10.50	4.60	9.60	4.10	1.30	3.40
Maximum	15.10	7.80	13.70	5.40	2.70	4.80
Mean	12.30	6.14	11.10	4.54	1.82	3.90
Median	12.20	5.60	10.80	4.20	1.80	3.60
Standard deviation	1.76	1.35	1.55	0.55	0.54	0.59
CV	14.28	22.02	13.96	12.22	29.44	15.17

Table 2. Descriptive statistics of annual rural UR in usual status for male, female and persons of age group 15 years and above, and all age

Statistics	15 Years and above			All age group		
	Male	Female	Person	Male	Female	Person
Minimum	3.70	1.30	3.10	3.70	1.30	3.10
Maximum	4.80	2.80	4.40	4.80	2.80	4.40
Mean	4.06	1.78	3.56	4.10	1.78	3.60
Median	3.70	1.70	3.30	4.00	1.70	3.50
Standard deviation	0.51	0.6	0.54	0.46	0.6	0.52
CV	12.63	33.57	15.2	11.31	33.57	14.57

Table 3. Descriptive statistics of annual rural UR in CWS for male, female and persons of different age groups

Statistics	15 Years and above			All age group		
	Male	Female	Person	Male	Female	Person
Minimum	7.10	3.40	6.30	7.10	3.40	6.30
Maximum	10.40	6.70	9.50	10.40	6.70	9.50
Mean	7.94	5.12	7.34	7.98	5.12	7.36
Median	7.40	5.70	7.00	7.60	5.70	7.00
Standard deviation	1.39	1.48	1.27	1.37	1.48	1.27
CV	17.49	28.92	17.25	17.19	28.92	17.20

and usual status which coincided with the COVID–19 pandemics. In CWS approach, the gap between URs of male and female person broadened significantly following COVID–19 pandemics.

The out of sample forecast of annual rural UR in usual status during the year 2022-23 are generated by fitting GM (1, 1) model for male, female and persons of age 15-29 years, 15-59 years, above 15 years and all age group for the state of West Bengal. The forecasted URs are presented in Table 4 along with RMAPE values. Similarly, the out of sample forecast for annual rural UR in CWS approach for male, female and persons of age above 15 years and all age groups are presented in Table 5 along with RMAPE values.

By considering the RMAPE values of the forecasted annual rural UR for male, female, and persons of age 15-29 years, 15-59 years, above 15 years, and all age group, it is evident that the grey model provides a reliable forecast except for female persons of 15-59 years, 15 years and above, and all age groups in the usual status as the forecast RMAPE values are more than 20 % which may be attributed to the high CV of annual rural UR for female persons of all the age groups in the usual status. However, grey model generates a good forecast for annual rural UR in CWS approach for female persons of 15 years and above, and all age group. During the year 2022-23, the annual rural UR in West Bengal in the usual activity status for males, females, and persons in the age group 15-29 years is predicted to be 12.50%, 6.55%, and 11.20%, respectively. In the CWS approach, the annual rural UR in West Bengal for males, females, and persons belonging to the all

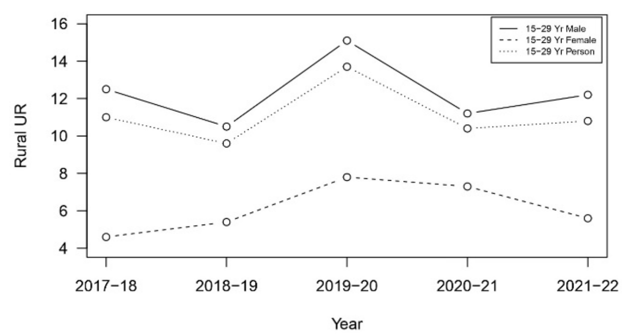

Fig. 1. Time plot of annual rural UR estimates for persons of age 15-29 years during 2017-18 to 2021-22

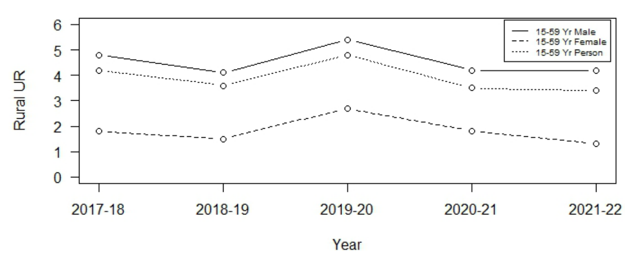

Fig. 2. Time plot of annual rural UR estimates for persons of age 15-59 years during 2017-18 to 2021-22

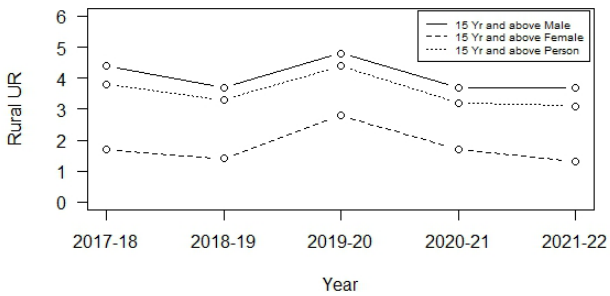

Fig. 3. Time plot of annual rural UR estimates for persons of age 15 years and above during 2017-18 to 2021-22

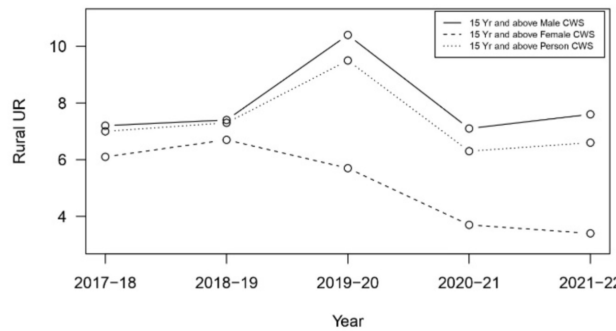

Fig. 5. Time plot of annual rural UR estimates for persons of age 15 years and above in CWS during 2017-18 to 2021-22

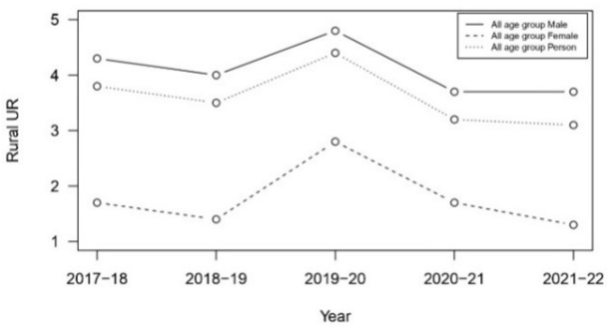

Fig. 4. Time plot of annual rural UR estimates for persons of all age group during 2017-18 to 2021-22

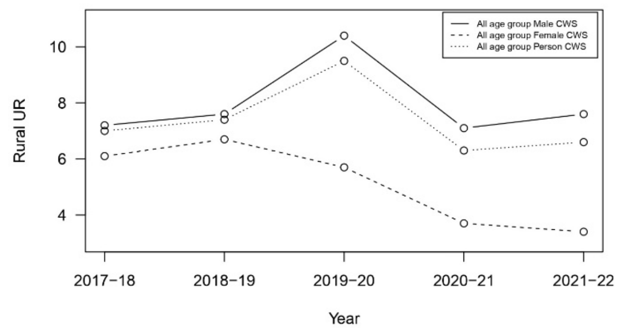

Fig. 6. Time plot of annual rural UR estimates for persons of all age group in CWS during 2017-18 to 2021-22

Table 4. Forecasted annual rural UR in usual status for male, female and all persons of different age group for the year 2022-23 along with RMAPE in West Bengal

Age Group	Male		Female		Person	
	Forecast	RMAPE	Forecast	RMAPE	Forecast	RMAPE
15-29 years	12.54	11.50	6.55	16.08	11.20	11.21
15-59 years	4.27	9.42	1.53	23.34	3.40	10.94
15 years and above	3.72	9.26	1.53	25.73	3.10	11.00
All age group	3.59	7.70	1.53	25.73	3.02	9.81

Table 5. Forecasted annual rural UR in CWS for male, female and all persons of different age group for the year 2022-23 along with RMAPE in West Bengal

Age Group	Male		Female		Person	
	Forecast	RMAPE	Forecast	RMAPE	Forecast	RMAPE
15 years and above	7.51	12.56	2.52	6.28	6.25	11.65
All age group	7.43	12.03	2.52	6.28	6.21	11.38

age group is predicted to be 7.43%, 2.52%, and 6.21% respectively whereas it is predicted to be 3.59% for male, 1.53% for female and 3.02% for person in the usual activity status.

5. CONCLUSION

The study employed Grey model, specifically the GM(1, 1) forecasting model, to predict annual rural unemployment rate of West Bengal using PLFS estimates. This choice was made because the Grey model has shown its capability to generate accurate forecasts even with the limited data available. Policymakers and government organizations can benefit from these advanced estimates of rural UR, which can assist them in developing proactive and focused interventions. The results obtained for various age groups indicate that, in most cases, the Grey model performed well. However, it struggled to provides a reliable forecast for female persons of 15-59 years, 15 years and above, and all age groups in the usual status. Despite this limitation, the Grey model remains a viable option for generating precise out-of-sample forecasts when working with such constrained time series data.

REFERENCES

Basak, P., Ray, M., Sinha, K. and Anuja, A.R. (2023). Prediction of Urban Unemployment Rate in India using Grey Model. *Journal of the Indian Society of Agricultural Statistics*, **77(3)**, 243-248.

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

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

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.

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.

Nguyen, P.H., Tsai, J.F., Kayral, I.E. and Lin, M.H. (2021). Unemployment rates forecasting with grey-based models in the post-COVID-19 period: A case study from Vietnam. *Sustainability*, **13(14)**, 7879.

Li, X. and Chen, W. (2009). A Grey-Markov predication for unemployment rate of graduates in China. In *2009 IEEE International Conference on Grey Systems and Intelligent Services (GSIS 2009)* (pp. 619-624). IEEE.

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.

MoSPI, Government of India (2019). Annual Report: PLFS, 2017-18, Ministry of Statistics and Programme Implementation, New Delhi.

MoSPI, Government of India (2020). Annual Report: PLFS, 2018-19, Ministry of Statistics and Programme Implementation, New Delhi.

MoSPI, Government of India (2021). Annual Report: PLFS, 2019-20, Ministry of Statistics and Programme Implementation, New Delhi.

MoSPI, Government of India (2022). Annual Report: PLFS, 2020-21, Ministry of Statistics and Programme Implementation, New Delhi.

MoSPI, Government of India (2023). Annual Report: PLFS, 2021-22, Ministry of Statistics and Programme Implementation, New Delhi.

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.

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.