

World unemployment problem and the unemployment rate forecasting with ARIMA model

Feilu Sun

Hailiang Experimental Middle School, Zhuji, 311800, China

kongyan@ldy.edu.rs

Abstract. In a rapidly changing global economic landscape, the level of unemployment remains one of the most important indicators of economic health, social stability and political dynamics. Due to the complexity and diversity of the unemployment problem, there are still some shortcomings and controversies in the existing research. Therefore, the aim of this study is to explore and predict unemployment in depth. This study mainly explores the history, current situation and influencing factors of unemployment statistics in depth, so as to make corresponding predictions and analyses. The research methods include literature review and R for relevant statistics and prediction. According to the forecast results, the world unemployment problem is still relatively serious, and the unemployment rate of the world population will remain relatively high in the coming years. However, due to the existence of various force majeure factors, the future unemployment rate may still show a relatively large increase or decline. This study demonstrated the applicability of the ARIMA model in predicting the values of this variable. These models do a good job of capturing patterns and trends, and the predicted values are reliable. At the same time, this study breaks through the traditional thinking, explores the causes of unemployment from multiple angles and analyzes the future social situation, which also has certain implications for further exploring the unemployment problem.

Keywords: World unemployment, ARIMA model, forecast.

1. Introduction

Unemployment has always been a huge problem, especially in the rapid development of society or in the face of emergencies. Therefore, from history to the present society, the emergence, improvement and application of various methods have gradually enriched the research on unemployment models to explore and improve the unemployment problem. William Beveridge's exploration of structural unemployment and its impact on the efficiency of labor markets in the 1930s helped develop quantitative models, such as the use of correlation analysis and regression analysis to predict unemployment levels. Decades later, the Phillips curve depicted an inverse relationship between unemployment and inflation, triggering a debate about the theoretical trade-offs involved [1]. By the 1970s, structuralist economists emphasized the impact of labor market characteristics, demographic characteristics, and institutional factors on unemployment. Swedish economist Gosta Rehn has demonstrated the lasting impact of the size of the young cohort on career trajectories [2]. This sociological perspective enriches forecasting methods.

With the acceleration of globalization in the 1990s, economists such as Richard Freeman promoted cross-country comparative analysis of labor markets, highlighting the large differences in unemployment rates across countries in terms of labor laws, educational systems, and social security systems [3]. More recently, rapid advances in automation and artificial intelligence have forced economists to reexamine their assumptions about technological unemployment. Some, such as Carl Benedikt Frey, predict a mass exodus of workers [4], while others are more optimistic about the prospects for job creation.

Nkwatoh set out to evaluate the validity of forecasts from various unemployment rate models, ranging from traditional time series models such as ARIMA to more sophisticated ARCH models. These models have a better ability to capture and predict patterns in time series data, and Nkwatoh applied them in the context of unemployment [5]. This illustrates the adaptability and relevance of these models in different economic environments. Dumi et al. used smoothing techniques to remove noise from time series data to clearly reveal the underlying trend in order to find the most accurate way to forecast future unemployment rates [6]. Thus improving the accuracy of the forecast.

The recent global outbreak and persistence of the novel coronavirus has led to sudden and dramatic changes in employment levels, accelerating the adoption of remote work [7] and challenging the traditional understanding of employment and unemployment. Trade wars, political tensions, and changing international alliances may affect economic policies and, consequently, job opportunities [8]. The rise of automation, artificial intelligence, and other technological advances may lead to job displacement in some industries while creating opportunities for others [9]. The trend towards gig economy work, part-time work and contract employment is growing and its transitory nature makes it challenging to accurately measure and predict unemployment as traditional models may not account for unemployment in non-traditional jobs [10]. These changes may introduce uncertainty into the market, and a sudden rise or fall in employment may cause the data to become volatile, so that traditional models that rely on past historical data to predict future trends may not work well.

Events such as pandemics, natural disasters, or geopolitical tensions can disrupt global supply chains, which can lead to temporary job losses in sectors that depend on global supply chains, so forecasting models need to be adaptive to take these external shocks into account. Governments around the world have been implementing policies to offset the economic impact of the pandemic and other challenges. These policies, from stimulus packages to job guarantees, affect the unemployment rate. Forecasting models need to account for potential policy changes and their effects. At the same time, employment in traditional industries may decline while new technology-driven industries may emerge, which also has an impact on the unemployment rate. Therefore, forecasting models need to take into account the dynamic nature of the job market in the face of technological change.

Therefore, this paper aims to provide an in-depth forecast of the overall world unemployment rate using historical data and recent events. At the same time, in addition to the numbers, reveal the basic patterns, challenges and opportunities that prevail in the unemployment situation. The analysis aims to capture the broader story of the transition from economic to social.

2. Methods

2.1. Data sources

The paper is based on Data from the World Bank Open Data website, which contains more than 7,000 indicators from the World Bank database, hundreds of which go back 50 years.

2.2. Variable selection

The data used in this paper total world unemployment (as a percentage of the total labour force) from 1991 to 2022 (simulated ILO estimates). Table 1 shows the corresponding percentage of world unemployment for each year.

Table 1. Percentage of unemployed people per year

year	percentage	year	percentage
1991	4.9	2007	5.9
1992	5.1	2008	5.9
1993	5.4	2009	6.5
1994	5.6	2010	6.3
1995	5.8	2011	6.2
1996	5.9	2012	6.1
1997	5.9	2013	6.1
1998	6.1	2014	6.0
1999	6.2	2015	6.0
2000	6.1	2016	6.0
2001	6.2	2017	5.9
2002	6.4	2018	5.7
2003	6.5	2019	5.5
2004	6.4	2020	6.9
2005	6.3	2021	6.2
2006	6.1	2022	5.8

2.3. Method introduction

In this paper, the above methods will be summarized and optimized, and the data will be analyzed, predicted and evaluated by the following methods. Firstly, historical data are collected, outliers and missing values are removed, and future predicted values are obtained by weighting through smoothing techniques to reduce the volatility of the time series and make it adapt to changes.

Secondly, the order of the ARIMA model is determined according to the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the observed data. The ACF provides autocorrelation values for any series with lagged values, takes the time series into account when looking for correlations, and helps identify the order of autoregressive (AR) and moving average (MA) components in the model. PACF Instead of looking for a correlation between lags like ACF and current, it looks for a correlation between the residual and the next lagged value.

Then, the parameter estimation is completed by obtaining the value of the entire parameter by reasoning over the sample. Next, we test whether the residual series satisfies the assumption of white noise (randomness). If the residual series satisfies white noise, then the residual follows a normal distribution, and the regression model has a better degree of fit, which can predict the unknown data more accurately.

The ARIMA model is then used to forecast future unemployment rates. ARIMA model is to use the autoregression, moving average, difference and other transformations of time series data to fit the data with fewer parameters, and establish a model that can describe the characteristics of the data, so as to predict the future changes of the data.

Finally, the model is evaluated. Using metrics such as root mean square error (RMSE) and mean absolute percentage error (MAPE), the predictions of the ARIMA model are compared with actual observations and their accuracy is evaluated.

3. Results and discussion

Figure 1 plots the world total unemployment rate (as a percentage of the total Labour force) from 1991 to 2022 (modelled ILO estimates) as a line graph to facilitate comparison. Because the unemployment rate is affected by many factors, it fluctuates up and down over time, but the overall trend is still upward.

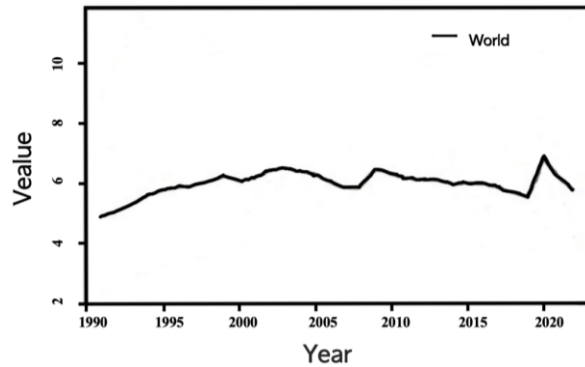


Figure 1. Unemployment line graph.

Then, the Arima function selects the appropriate Arima model for each variable from the prediction package. This function automatically selects the order of ARIMA models based on ACF and PACF plots. The ACF is presented in Figure 2 and the PACF in Figure 3 below.

ACF - World

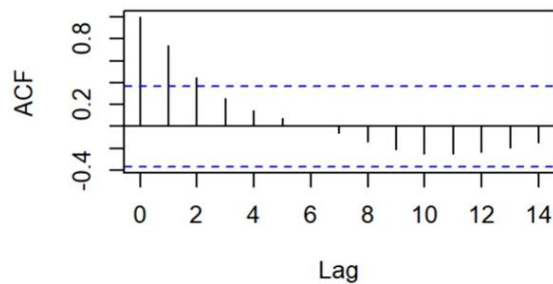


Figure 2. ACF plot.

By looking at the ACF of the data, we can determine the order of the ARIMA model. For ACF, the numbers are relatively flat. From the figure 3, it is observed that it is not white noise because the values are not all in the dotted lines.

PACF - World

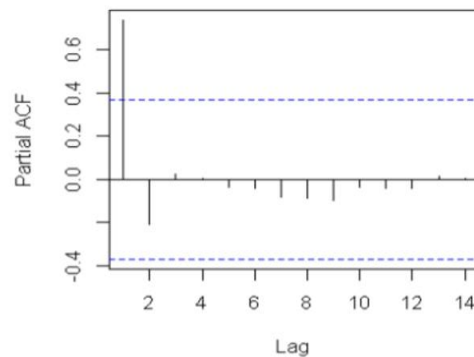


Figure 3. PACF plot.

For PACF, the data are also relatively flat. From the figure 4, it is observed that it is not white noise, because some values exceed the dotted line. At the same time, qqplot can also be used to test the residual hypothesis and judge whether the data obey the normal distribution. As can be seen from the figure below, it basically follows a normal distribution with little data fluctuation and a smooth data set. It can be seen that the modelling effect is good. Table 4 show the Q-Q Plot.

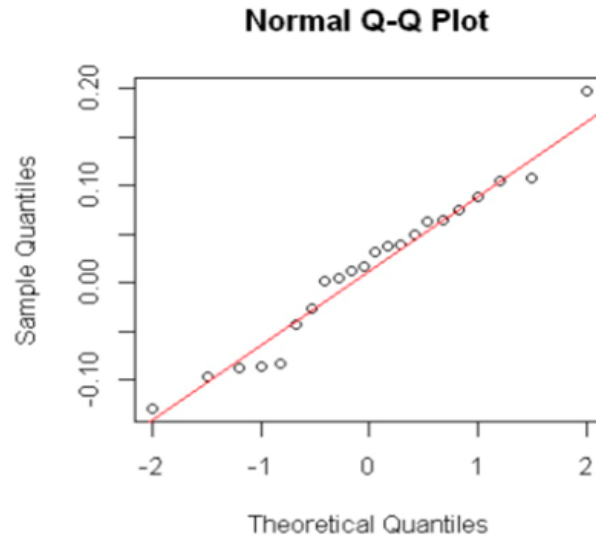


Figure 4. Q-Q Plot.

Finally, a model is constructed to forecast and evaluate the world unemployment rate. Table 5 below presents the forecast results. In model evaluation, root-mean-square error is a common index to measure the accuracy of prediction model. It can help us verify whether the predicted value is consistent with the actual value, and compare the difference. The small RMSE value indicates that the prediction model is close to the actual value, and the accuracy of the model is relatively high. The same is true of the MAPE. Therefore, it can be seen from the following data that the prediction accuracy is generally good.

Table 2. Results and evaluation

Year	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2023	5.997	5.888	6.107	5.830	6.165
2024	5.955	5.656	6.253	5.497	6.412
2025	5.980	5.539	6.422	5.305	6.656
2026	5.984	5.453	6.514	5.173	6.795
2027	5.984	5.380	6.588	5.060	6.908
2028	5.984	5.314	6.654	4.960	7.008

RMSE: 0.172
 MAPE: 2.103 %

4. Conclusion

Therefore, the world unemployment problem is still relatively serious, and according to the forecast results, the unemployment rate of the world population will remain relatively high in the coming years. At the same time, relatively large increases and decreases in unemployment may occur due to the spread of various types of epidemics, the generation of natural disasters, the progressive development of science and technology, changes in policies, changes in the population base, changes in employment patterns and other factors.

The results show that the ARIMA model is suitable for predicting the values of these variables. These models perform well in capturing patterns and trends in the data, and the forecasts generated are reliable. The study also highlights the importance of model evaluation and residual hypothesis testing to ensure forecast accuracy.

Future research could explore the use of other time series forecasting methods, such as machine learning algorithms, to compare their performance with ARIMA models. Future research could also consider using multiple data sets and nonstationary tests.

In summary, this study demonstrates the applicability of the ARIMA model in predicting the value of this variable. These models do a good job of capturing patterns and trends. The predicted values are reliable, but unknown variables may affect the forecasts.

References

- [1] Li X Y and Li J 2007 Beveridge Curve Analysis of unemployment in China. *Population of Northwest China*, 28(2), 5.
- [2] Erixon L 1997 The Father of the Swedish Model in Memory of Gosta Rehn (1913-96). *Economic & Industrial Democracy*, 18(4), 635-650.
- [3] Freeman R B 2004 SSRN-Trade Wars: The Exaggerated Impact of Trade in Economic Debate by Richard Freeman. *Blackwell Publishing Ltd*, 27(1), 1-23.
- [4] Huria S 2021 Book Review: The Technology Trap: Capital, Labor, and Power in the Age of Automation by Carl Benedikt Frey. *Working paper*.
- [5] Nkwatoh L S 2019 Can cameroon become an emerging economy by the year 2035? projections from univariate time series analysis. *Journal of economics and international finance*.
- [6] Dumii K, Asni A E and Muk B 2015 Forecasting Unemployment Rate in Selected European Countries Using Smoothing Methods. *International journal of social education economics & management engineering*.
- [7] Lu Y Y and Yan Z Q 2021 The impact of unemployment on China's economy in the context of COVID-19. *Heilongjiang Science*, 12(8), 3.
- [8] Li C A and Liu N 2021 Study on the impact of "normal" trade war on China's economy and employment. *Working paper*.
- [9] Cynthia K 2020 Unemployment problem of the artificial intelligence research. *Journal of cooperation in economy and technology*, 2.
- [10] Nadou 2021 Is the gig economy really the future of work. *Working paper*.