INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010
Received: 21.05.2021 | **Accepted:** 12.11.2021 | **Published Online:** 01.12.2021

-RESEARCH ARTICLE-

FORECASTING UNEMPLOYMENT RATE IN SOUTH AFRICA WITH UNEXPECTED EVENTS USING ROBUST ESTIMATORS

Simon Setsweke Nkoane

Sefako Makgatho Health Sciences University, Ga-Rankuwa
E-mail: setsweke.nkwane@ul.ac.za

Solly Matshonasi Seeletse

Sefako Makgatho Health Sciences University, Ga-Rankuwa
E-mail: solly.seleetse@smu.ac.za

─Abstract─

The purpose of the study is to build the time series model and forecast the unemployment rate in South Africa in the presence of the unexpected events or contamination of data using robust estimators. Robust estimators deal with outliers (unexpected events) when identifying the orders, with a view to estimating the parameters of the time series models. Often, time series data are contaminated with anomalies or outliers. The standard methods of parameter estimation such as maximum likelihood (ML), least squares (LS) and method moments (MM) are sensitive to outliers. The quarterly unemployment time series data over time of January 2010 through December 2020 is used. Outliers are identified, and not removed and an ARIMA (1, 1, 1) model is found to be the best suitable model for the unemployment series. An accuracy of the forecast is measured by the standard methods, such as the RMSE, MAPE, and MAE. The forecasting results for the unemployment rate show that the robust estimators perform better in the presence of outliers.

Keywords: Robust estimation, unemployment rate, unexpected events, economy

1. INTRODUCTION

Unemployment rate (UR) in any country deters economic growth, and South Africa (SA) is not an exception. Despite an increase in the number of the graduates, and business opportunities in the country, South Africa's unemployment rate keeps on increasing.

Citation (APA): Feng, S., Salleh, N. Z. B. M., Abdullah, M. B. (2021). Forecasting Unemplyment rate in South Africa with Unexpected Events Using Robust Estimators. *International Journal of Economics and Finance Studies, 13* (2), 199-222. doi:10.34111/ijefs.20212010

INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

South Africa has experienced the establishment of new employer companies and shopping malls . Pasara and Garidzirai (2020) argue that since there was a substantial improvement observed in economic growth, the unemployment rate was supposed to decrease. However, millions of people still find it difficult to get a job. South Africa's unemployment rate stood at 32.5% in the October – December quarter of 2020 (StatsSA, 2020a). This high unemployment rate includes, people with matric, people without matric and graduate unemployment rate. Several studies (Broekhuizen, 2016; Graham, Williams, & Chisoro, 2019; Richa, Rakesh, & Rimzim, 2015) have been conducted to investigate the problems caused as a result of unemployment. Unexpected (unforeseen) events may make things worse by affecting the country's economy negatively. For example, during the Coronavirus pandemic 2019 (Covid-19), in 2020, millions of jobs were lost and no one expected or anticipated this to happen. A Covid-19 pandemic was an unforeseen event. According to the latest Quarterly Labour Force survey, Quarter 2: 2020 South Africa dropped 2.2 million jobs (StatsSA, 2020b) because UR depends on the total number in the labour force. Figure 1 shows that South Africa's unemployment rate has decreased by 6.8% from 30.1% in Quarter 1 to Quarter 2: 2020. Figure 2 shows that the unemployment rate rose by 7.5% to 30.8%. Unemployment rate occurred in Quarter 2 was recorded as the highest rate in Quarterly Labour Force Survey (QLFS), a result of the national lockdown during the Covid-19 pandemic.

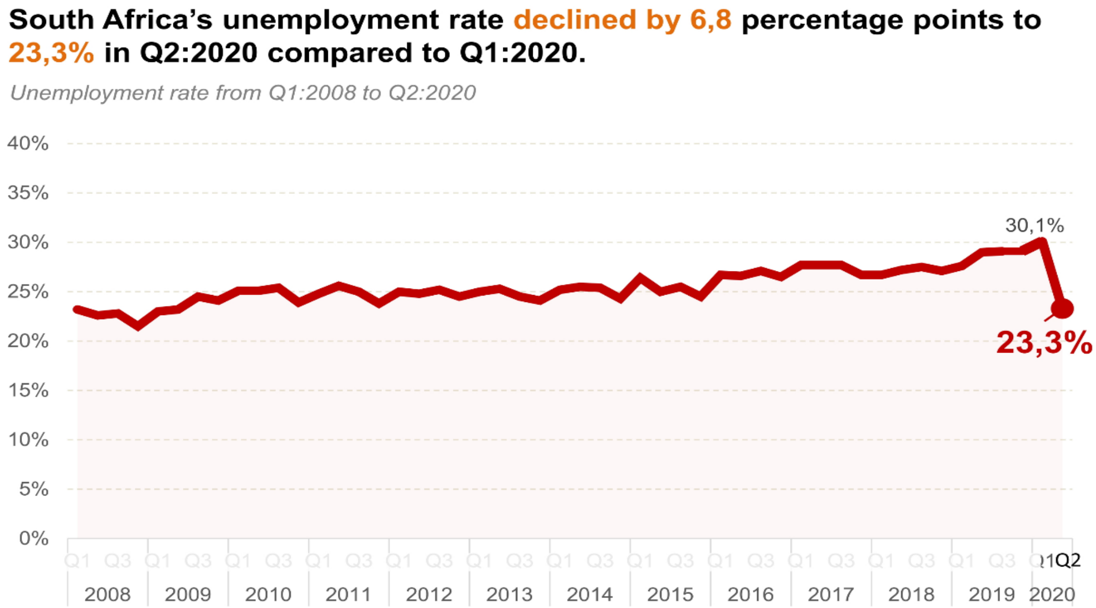

Figure 1: SA's Unemployment Rate

Source: StatsSA (2020a)

Vol: 13 No: 2 Year: 2021 ISSN: 1309-8055 (Online) (pp. 199-222) Doi: 10.34109/ijefs.20212010

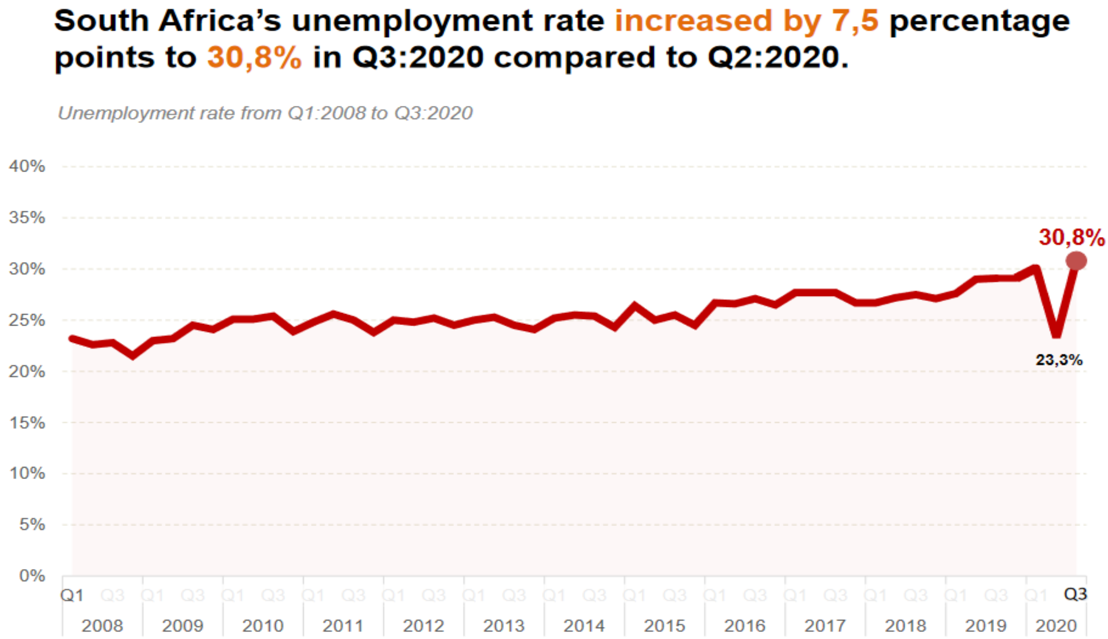

Figure 2: SA's Unemployment Rate

Source: StatsSA (2020a)

Unemployment occurs when a person who is actively searching for employment is unable to find work (Richa et al., 2015). In the presence of unexpected events, forecasting becomes unreliable, and therefore, accurate forecasts in this situation is essential. Forecasting of UR plays an important part in planning. UR forecasting is a basic tool for planning and risk management in tax, finance, education, agriculture and industrial policies (Kumphon, 2013). Several studies on forecasting unemployment rate in different countries (include Indonesia, Thailand, Nigeria and etc.) have been conducted in the past. Kumphon (2013) applied two forecasting approaches, the SARIMA and the Artificial Neuron Network (ANN) method to determine the suitable models for predicting the unemployment rate in Thailand. In their study, SARIMA showed the good performance and the satisfactory model to forecast the UR in Thailand. Mahmudah (2017) applied ARIMA to predict graduate unemployment in Indonesia. The forecasting results showed a tendency increasing unemployment. Adenomon (2017) applied ARIMA model for modelling and forecasting unemployment rate in Nigeria. The forecast of unemployment rates in Nigeria revealed an increase rate from 2015 to 2017. From these studies, time series forecasting models proved to be most effective in forecasting unemployment rate. These studies also revealed the increase of unemployment rate in the future years from the period under consideration.

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

The purpose of the study is to build the time series model and forecast the unemployment rate in South Africa in the presence of the contamination of data using the robust estimators. The overall objectives are to detect the outliers from the time series data, build a suitable mode, and forecast for future unemployment rate using the robust estimators without removing the outliers.

The remaining sections of this study are organized as follows. Section 2 discusses the literature review on unemployment rate and on the usage of robust estimators. Section 3 presents the methodology on building the time series model that will be used to forecast the unemployment rate. Section 4 presents the statistical analysis and discussion of quarterly unemployment rate. Section 5 presents the conclusion while putting forth some recommendations based on the findings of the present study

2. LITERATURE REVIEW

Unemployment is regarded as one of the key socioeconomic challenges in any country. Hence, there have been a fairly large number of studies focusing on unemployment in many countries including South Africa in the past. Unemployment is a global phenomenon which is considered as a pressing contemporary issue around the world (Ramli et al., 2018). Past studies have considered the different aspects of unemployment such as, examining of the determinants or factors of unemployment, examination of the relationship between unemployment and economic growth, and modelling and forecasting unemployment rates. There are several factors that may contribute to the unemployment rate. According to Mabiza, Mahlasela, and Mbohwa (2017) one of the key factors contributing to the issue of unemployment is the lack of experience required by the organizations.

Some studies were conducted to investigate the relationship between unemployment rate and other factors. High unemployment and stagnant economic growth that have been persistent economic challenges in South Africa are studied and explored in these works (Pasara & Garidzirai, 2020). Makaringe and Khobai (2018) investigated the trends and the impact of unemployment on economic growth in South Africa over time period of quarter from 1994 to 2016. The findings confirmed that there is a negative relationship between unemployment and economic growth both in the long and short run.

Modelling and forecasting of unemployment rate has received a considerable amount of attention in existing literature. Ramli et al. (2018) used the two-time series forecasting models ARIMA and Double Exponential Smoothing (Holt's) to predict the unemployment rate in Malaysia. The best model to predict the unemployment in Malaysia is ARIMA (2, 1, 2). The result showed that there is a slight increase of unemployment rate in the next 10 years. Dritsakis and Klazoglou (2018) applied Box-Jenkins methodology on monthly data to the US unemployment rate from January 1955 to July 2017. The results showed that the SARIMA (1, 1, 2) $(1, 1, 1)_{12}$ - GARCH (1, 1) is the best model for US unemployment rate. The forecasting outcome showed that the projected value of unemployment is close to the real value. Didiharyono and Syukri (2020) used the ARIMA for forecasting in an anticipated open unemployment rate in South Sulawesi. Based on the results of the study obtained the best time series model used for forecasting is ARIMA (1, 2, 1) with the small Mean Square value of 2,0474. The ARIMA forecasting results showed that the unemployment rate has increased from year to year. In all these reviewed studies on unemployment, time series analysis was employed to uncontaminated data. In a case where the outliers were detected, the outliers were removed from the data. When the robust estimators are employed, once the outliers are detected they are not removed for the data. This study uses the robust estimators to model and forecast unemployment rate in the presence of outliers. In the next section, the study includes a brief literature review on outliers, and robust estimation.

Hawkins (1980) defines an outlier an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism. There are several terms used to refer to outlier observation such as, anomalous, aberrant, and discordant. Outliers can be caused due to a number of different reasons such as data errors, intentional or motivated misreporting, sampling error, and unexpected events such as weather conditions, strikes, and economical calamities. These causes can be divided into those arising from errors in data and those arising from the inherent variability of the data. Whichever the cause, failing to manage outliers may lead to unreliable results; in time series, outliers can lead to poor forecasts which will then lead to poor planning. Liu et al. (2001) argued that the outlying data in time series potentially could have a rather significant impact on the estimates of the model parameters. In time series, there are four common types of outliers, namely additive outliers (AO), innovative outliers (IO), temporary change (TC) and level shift (LS). An additive outlier indicates an event that affects a series for one time period (Liu et al., 2001). Outliers in time series data were first studied by Fox (1972).

It is important to detect outliers before the analysis of time series data. Outlier detection in time series has been widely studied and several detection methods have been proposed (Abraham & Chuang, 1989; Ljung, 1993). A proper treatment of outliers is important to produce reliable predictions. In a common practice, once the outliers are detected, they are then removed from data. Robust estimators such as M-estimator, GM-estimator, MM-estimator, and least median of squares (LMS) have been applied successfully in the presence of outliers (Bhatia et al., 2016; Hampel et al., 2011; Wang & Lee, 2011). Ahmar et al. (2018) studied the detection and correction of data containing the additive outliers on the model ARIMA (p, d, q). After the detection and correction of data obtained by the iteration of the ARIMA model, the results show that there is an improvement of forecasting error rate data. Abegaz and Mohammed (2018) examined the elasticity of healthcare using OLS, LAD, LMS and M-estimators. For robustness reasons, LAD, LMS and M-estimators were employed. All of these models, OLS, LAD, LMS and M-estimator provided almost consistent results. Grossi and Nan (2019) employed a robust approach to modelling electricity spot prices. Robust estimators are not strongly affected

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

by the presence of spikes and are effective in the prediction of normal prices which are the majority of the data observed on electricity markets. The application of the procedure of the Italian electricity market reveals the forecasting superiority of the robust GM-estimator based on the polynomial weighting function with respect to the non-robust least squares estimator. Bakar and Midi (2019) estimated the usefulness and efficiency of robust methods in contaminated panel data. The findings demonstrate the advantage of using robust estimators (GM- and MM-estimators) in terms of reducing the influence of high leverage points on panel data over the OLS.

Chakraborty et al. (2021) conclude that forecasting unemployment rates is one of the important steps allowing economic entities in the labor market to reduce the uncertainty resulting from the broader socio-economic situation of a country. Therefore, attempts to predict future trends and phenomena that will take place in the labor market are more and more difficult, both for the long-term as well as short and medium-term horizons (Kucharski, 2015). Adenomon (2017) forecasts unemployment rates in Nigeria using ARIMA model for the period of 1972 to 2014. The results of ARIMA model revealed an increasing rate from 2015 to 2017 while a slight decrease in 2018 and recommended that administration should focus on capital project that has the capacity to create employment. Claveria (2019) analyse the panel data of 8 European Countries form 2007-2016, by using ARIMA and predicting the unemployment rate using the degree of agreement from consumer unemployment expectations. Findings confirm that the degree of agreement in consumers' expectations contains useful information to predict unemployment rates, especially for the detection of turning points. Some past studies have employed the ARIMA time series model to forecast the unemployment for various countries Ahmad et al. (2021) in case of Europe, AKTER (2020) for Bangladesh, Walter (2020) in case of India, Didiharyono and Syukri (2020) in South Sulawesi, in Romania (Davidescu, Apostu, & Stoica, 2021), Nguyen et al. (2021) utilized ARIMA and Grey based theory to estimate and forecast the unemployment rate in Vietnam.

Gostkowski and Rokicki (2021) Forecasting Unemployment Rate can be done using different forecasting method such as naive method, regression model, ARIMA, Holt model and Winters model and the authors used these methods and applied them to the monthly data of 2008-2018. Findings show that he suitable methods of forecasting unemployment rate are the quadratic regression model and the Winters multiplicative model. Another studies in case of Pakistan, performs estimation and forecasting of the inflation, interest, literacy and unemployment rate by using nonlinear regression models for the period 2000 to 2019. Nonlinear regression models were selected on basis of goodness of fit, the correlation with the present data and the logical trend of future forecasts. Forecasting rates for the next ten years from 2020 to 2029 have been smoothly determined in the study in light of the afore-mentioned factors (Ghoto, Talpur, & Kamboh, 2021).

Vol: 13 No: 2 Year: 2021 ISSN: 1309-8055 (Online) (pp. 199-222) Doi: 10.34109/ijefs.20212010

Davidescu, Apostu, and Paul (2021) estimating and forcasting unemployment rate in Romanian for the period of 2021-2022 by comparative analysis the different univariate methods of forcasting such as SARIMA, SETAR, ETS, NNAR and Holt–Winters, utilized the data from 2000-2020 in case of Romania. By employing all he above forecasting techniques, final results shows that NNAR model is considered the best model of modeling and forecasting of unemployment rate. Davidescu, Apostu, and Marin (2021) conduct another study to forecast the unemployment rate in Romania from the point of health crisis by applying univariate and multivariate econometrics approach. ARFIMA and SETAR models are used for forcasting by employing data of 2000-2021. Findings confirm that SETAR and VECM provide very similar results in terms of accuracy in case of pre-pandemic period, and further shows that in 2020 reach at 4.1% but during 2021 and 2022 it's decreasing. Unemployment remains a major issue for both developed and developing countries and a driving force to lose their monetary and financial power. Han Lai 2021 predicts the unemployment in light of the effect sof COVID-19 in case of a selected number of developed and developing countries of Asia. Data analysis is done through the ARIMA-ARNN, ARIMA-ANN and ARIMA-SVM approach. Prediction confirmed that unemployment in developing countries is three times higher as compared to advanced countries in the coming years and it will take double the time to address the impacts of Coronavirus in developing countries than in the developed countries of Asia.

From the reviewed literature, most of the studies compared the performances of the robust estimators (M, GM and MM-estimators) against the OLS estimator in the presence of outliers in data. The M, GM, and MM-estimators performed similarly to OLS in a small uncontaminated sample data. Generally, the robust estimators is more effective and useful than the OLS estimator when the dataset has outliers.

3. RESEARCH METHEDOLOGY

Unemployment rate prediction has become critically significant as it can be used by governments to make policy decisions and design accurate policy mechanism and frameworks. The aim is to estimate the most suitable ARIMA model of unemployment rate in South Africa for the period of 1972 to 2014 under the assumption that present unemployment rates depend on the unemployment rate of the preceding or previous year. In this section, we briefly review the nonstationary time series ARIMA models and the estimators in the context of regression, and subsequently apply these estimators to nonstationary time series models.

1. Autoregressive Integrated Moving Average (ARIMA) Models

An ARIMA model studied by George Box and Gwilym Jenkins in 1972 (Box et al., 2015), provides a generalization of the class of ARMA that incorporates a wide range of the non-seasonal and nonstationary series (Box et al., 2015). The ARIMA model combines the past value and the errors. ARIMA models are often expressed in backshift

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

notation. Many previous studies used this model (Adenomon, 2017; Ahmar et al., 2018; AKTER, 2020; Claveria, 2019; Lai et al., 2021; Ramli et al., 2018; Walter, 2020).

The univariate time series process $\{X_t\}$ is said to be ARIMA process of orders p, d, and q if it can be expressed as

$$\phi_p(L)\alpha(L)X_t = \psi_q(L)\varepsilon_t \tag{1}$$

or

$$X_t = \frac{\psi_q(L)}{\phi_p(L)\alpha(L)}\varepsilon_t = \varphi(L)\varepsilon_t \tag{2}$$

where

$$\phi_p(L) = \left(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p\right),$$
$$\psi_q(L) = \left(1 - \psi_1 L - \psi_2 L^2 - \cdots - \psi_q L^q\right)$$

are polynomials in the backshift operator L and all their roots are outside the unit circle; and $\alpha(L) = \Delta(L)$ is a polynomial, a difference operator and all its roots are on the unit circle, and $\{\phi_1, \phi_2, \dots, \phi_p, \psi_1, \psi_2, \dots, \psi_p\}$ are unknown parameters.

The time series model is built in three stages, namely model identification, parameter estimation and model diagnostic (Box- and-Jenkins method). The fitted model is used to predict the future events which is part of forecasting. There are three standard methods for the parameter estimation, namely maximum likelihood (ML), least squares (LS) and method of moments (MM). The stages of the Box-and-Jenkins method are presented in Figure 3 and summarised as follows:

i) *Model Identification:* The appropriate orders p, d, and q of ARIMA model are determined by examining time series plots, using autocorrelation and partial autocorrelation.

ii) *Parameter Estimation*: The parameters of the selected model are estimated.

iii) *Diagnostic Checking*: The fitted model is checked by assessing the residuals.

iv) *Forecasting and Forecast measuring*: The final or selected model is used to forecast future values and the forecast accuracy is measured.

INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

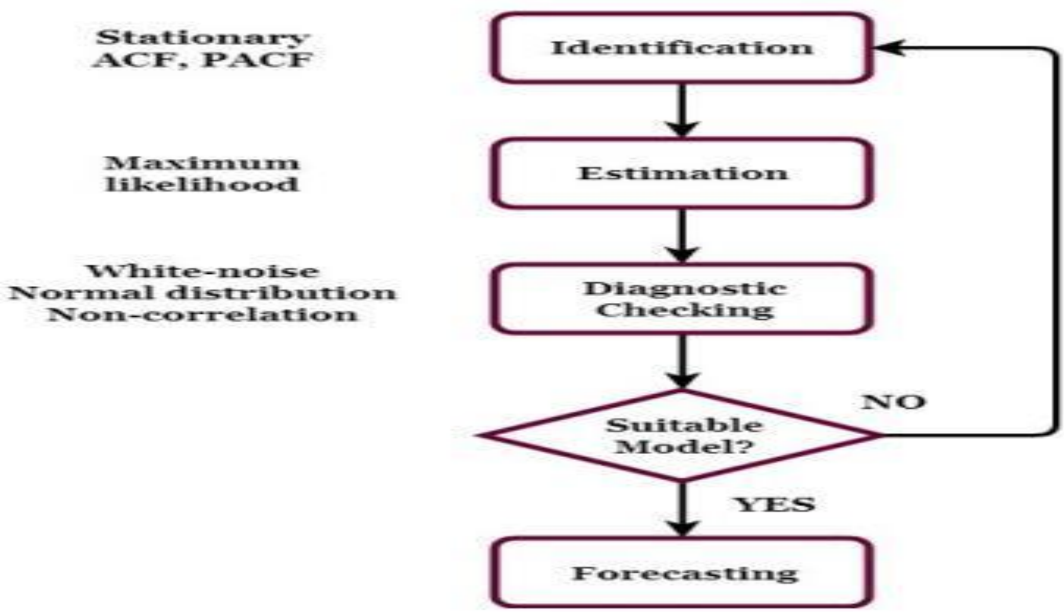

Figure 3: Stages of the Box-and-Jenkins methodology

Source: San-Juan, San-Martín, and Pérez (2012)

2. Robust Regression Estimators

The three standard methods for parameter estimation are sensitive in the presence of outliers. The robust estimators are used at parameter estimate stage in the presence of outliers. Robust estimation of model parameters has been widely studied (Martin & Yohai, 1986; Muler, Pena, & Yohai, 2009; Stockinger & Dutter, 1987). These estimators are briefly described in the context of regression before they can be applied to time series models. The estimators are defined as follows:

1) *M-estimator*, is a class of maximum likelihood (ML) type estimator that was introduced

by Huber (1965) cited in (Hampel, 1992). The M-estimators are regarded as the generalization of MLE. It means that they contain all models that are derived to be ML models. The solutions

$$\hat{\theta}_{ML} = \arg\max_{\theta}\left(\prod_{i=1}^{n} f(x_i, \theta)\right) \tag{3}$$

are called the maximum likelihood estimators (MLEs). The solutions

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

$$\hat{\theta}_{ML} = \arg\max_{\theta} \left(\sum_{i=1}^{n} \rho(x_i, \theta) \right) \tag{4}$$

are called the M-estimators (i.e. they minimize $\sum_{i=1}^{n} \rho(x_i, \theta)$. If ρ is convex, then ψ is the derivative of ρ. The functions ρ and ψ are chosen conveniently. This class of M-estimators was first introduced for location problems by Huber (1965) and the estimators are the solutions $\hat{\theta}$ of the equations,

$$\sum_{i=1}^{n} \psi(X_i - \hat{\theta}) = 0. \tag{5}$$

To define the M-estimator for the general linear regression, the linear regression equation is considered. The *M-estimator* of linear regression model determines the estimator $\hat{\theta}$ by minimizing an objective function of

$$\min \sum_{i=1}^{n} \rho(r_i) = \min \sum_{i=1}^{n} \rho(y_i - X^T\theta) \tag{6}$$

where $\rho(\cdot)$ is a real-valued function that gives the contribution of each residual (it is also called a less rapidly increasing function of the residuals) to the objective function (to bound the influence to large residual).

*2) MM-estimator (*Modified M-estimator or Multi-stage M-estimator) is a special type of M-estimator first proposed by Yohai (1987) for regression to deal with the problem that S-estimator encounters. MM-estimator estimates the regression parameter using the S-estimator which minimizes the scale of the residuals from the M-estimator and then proceeds with M-estimation. The least median of squares (LMS) and M-estimates can also be used as the initial estimates in MM-estimation. The aim of the MM-estimator is to combine estimates that have a high breakdown value and are more efficient for normal errors. The MM-estimator is actually obtained by using a three-step algorithm or procedure proposed by Yohai, Stahel, and Zamar (1991):

Step 1: An initial consistent estimate $\hat{\theta}_0$ by S-estimator with high breakdown point possibly 0.5 is computed, but low normal (Gaussian) efficiency is obtained.

Step 2: A robust M-estimator of scale parameter $\hat{\sigma}$ of the residuals based on the initial value is computed.

Step 3: The M-estimator $\hat{\theta}$ is computed by using a redescending score function ρ_1' and the scale estimate $\hat{\sigma}$ obtained in Step 2.

The MM-estimator is defined as the solution of

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

$$\hat{\theta}_{MM} = \sum_{i=1}^{n} \rho_1' \,(r_i) X_{ij} = 0 \tag{7}$$

or

$$\sum_{i=1}^{n} \rho_1' \left(\frac{y_i - X^T \theta}{\hat{\sigma}} \right) X_{ij} = 0 \tag{8}$$

where $\hat{\sigma}$ is the standard deviation obtained from the residuals of S-estimator and ρ is a Tukey's biweight function defined as follows:

$$\rho_k(r) = \begin{cases} \dfrac{3r^2}{k} - \dfrac{3r^4}{k} + \dfrac{r^6}{k} & if\ |r| \le k, \\ 1 & if\ \ |r| > k \end{cases} \tag{9}$$

where ρ_0 determines the breakdown point (S-estimator) and ρ_1 determines the efficiency (MM-estimator). Both M- and S-estimators use the Tukey's biweight ρ function. The MM-estimator in the final step which must satisfy the following condition:

$$\sum_{i=1}^{n} \rho_1' \left(\frac{y_i - X^T \hat{\theta}}{\hat{\sigma}} \right) \le \sum_{i=1}^{n} \rho_1' \left(\frac{y_i - X^T \hat{\theta}_0}{\hat{\sigma}} \right) \tag{10}$$

and can also be written as $S(\hat{\theta}) \le S(\hat{\theta}_0)$.

3. Robust Regression Estimators

We consider the contaminated ARIMA(p, d, q) process with a fraction of outliers or structural changes defined by Tsay (1988).

$$X_t^* = \left(1 - I_t^{(T)}\right) X_t + I_t^{(T)} \omega \tag{11}$$

where ω is the magnitude of outlier at time $t = T$, X_t is outlier-free series of observation in equation (1), and $I_t^{(T)}$ is an indicator function representing whether or not there is an an outlier at time T such that

$$I_t^{(T)} = \begin{cases} 1, & t = T \\ 0, & t \ne T \end{cases} \tag{12}$$

Two common time series outliers, Additive outlier (AO) also called Type I outlier and Innovative outlier (IO) also called Type II outlier are considered in this study. An AO only affects a single observation, which is either a smaller or larger value compared to

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

the expected value in the data. An IO affects only one residual at the date of the outlier (Deneshkumar & Kannan, 2011).

The AO formula at time $t = T$ is defined as

$$X_t^* = \begin{cases} X_t + \omega, & t = T \\ X_t, & t \neq T \end{cases} = X_t + \omega I_t^{(T)} \tag{13}$$

and the IO formula at time $t = T$ is defined as

$$X_t^* = X_t + \omega \frac{\theta(L)}{\beta(L)} I_t^{(T)} \tag{14}$$

where the ratio $\frac{\theta(L)}{\beta(L)}$ denotes the magnitude and the dynamic pattern of the outlier effect at time point $t = T$ (Kaiser & Herrero, 1999). The AO-model of ARIMA is given by

$$X_t^* = X_t + \omega_t I_t^{(T)} = \varphi(L)\varepsilon_t + \omega I_t^{(T)} \tag{15}$$

The IO-model of ARIMA is given by

$$X_t^* = X_t + \omega \frac{\psi(L)}{\phi(L)\alpha(L)} I_t^{(T)} \tag{16}$$

Then, the M-estimates for ARIMA models are defined as the solutions of the system

$$\sum_{i=1}^{n} \rho\,(r_t) = \sum_{t=1}^{n} \rho\,(\varepsilon_t(\theta)) \tag{17}$$

and the M-estimates of scale (scale equivariant) for ARIMA are now defined as the solutions of the system

$$\sum_{i=1}^{n} \rho\left(\frac{r_t}{\hat{\sigma}}\right) = \sum_{t=1}^{n} \rho\left(\frac{\varepsilon_t(\theta)}{\hat{\sigma}_\varepsilon}\right) = b \tag{18}$$

where ρ is a bounded function and chosen conveniently, b is a positive constant, and $\hat{\sigma}_\varepsilon$ is an estimate of σ. To make the M-estimators of scale consistent of the standard deviation when the data are normal, we require that $E_\Phi\big(\rho(x)\big) = b$ where Φ is the standard normal distribution (Muler et al., 2009). Martin and Yohai (1986) state that the scale estimates have an important role in time series model fitting as the purpose of a scale estimate is to measure the deviations from 0 in the components of a robustly

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

centred sample. The estimator is obtained by solving the partial derivative of equation (18) assuming that $\hat{\sigma}_{\varepsilon}$ is fixed

$$\sum_{t=1}^{n} \psi\left(\frac{\varepsilon_t(\theta)}{\hat{\sigma}_{\varepsilon}}\right)\frac{\partial \varepsilon_t}{\partial \theta} = 0 \tag{19}$$

The robustness properties (includes breakdown point and influence function) and asymptotic properties (including consistency and normality asymptotic) of these estimators are discussed in (Nkoane, 2019).

4. ARIMA Forecasting

The ARIMA models are the most commonly used models for nonstationary time series forecasting. While the exponential smoothing models were based on a description of trend and seasonality of the time series data, ARIMA models aim to describe the autocorrelations in the time series data. An ARIMA forecasting is relatively straightforward, and involves the iterative Box-Jenkins procedure.

The h-step-head forecast at time t based on the ARIMA is given by

$$\hat{\phi}_p(L)\alpha(L)\hat{X}_{t+h|t} = \hat{\psi}_q(L)\varepsilon_{t+h|t} \tag{20}$$

5. Measures of Forecast Accuracy

After the forecasting methods have been established, it is important to examine their levels of accuracy. The accuracy of forecasts is a useful way of choosing the best forecasting technique from different forecasting methods. There are different measures of forecast accuracy, such as the Mean Absolute Deviation (MAD), the Mean Square Error (MSE), and the Sum of Squares Errors (SSE). In fact, the forecasting accuracy measures are divided into scale dependent errors, percentage errors and scale errors. To define these measures, let X_t be the actual value of the time series at time period t, $\hat{X}_t$ be the forecast value of the time series at time period t, and n be the number of time periods

The mean absolute error (MAE) and the root mean squared error (RMSE) are the most commonly used scale-dependent measures and the mean absolute percentage error (MAPE) is the most commonly used percentage measure. They are defined as follows, starting with the forecast error (FE) or residual for the individual forecast values at different time periods are:

$$e_t = X_t - \hat{X}_t \qquad \text{for } i = 1,2,\dots,T \tag{21}$$

This is basically the difference between the actual time series value at time t and the predicted value at time t,$\hat{X}_t$. These error terms were useful to analyse and summarise the

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

accuracy of the forecasts. The study could have used t instead of n to be consistent with sample size (or sampling), but instead, to be consistent it used t for current time (to emphasize time) so that the next time becomes $t+1$. Apart from the direct next period $t+1$, other future periods are $t+2$, $t+3$, and so on. The mean absolute error (MAE) is calculated by the following formula

$$MAE = mean(|e_t|) = mean\left(\left|X_t - \hat{X}_t\right|\right) = \frac{1}{n}\sum_{t=1}^{n}\left(\left|X_t - \hat{X}_t\right|\right), \tag{22}$$

The root mean squared error (RMSE) is derived from the mean square error (MSE). MSE is popular and corrects the 'cancelling out' effects of the previous two error measures and is calculate using the following formula:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(X_t - \hat{X}_t\right)^2, \tag{23}$$

The root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(X_t - \hat{X}_t\right)^2} \tag{24}$$

The MAPE is calculated by the following formula

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|e_t|}{X_t}\cdot 100\% = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|X_t - \hat{X}_t\right|}{X_t}\cdot 100\%\,, \tag{25}$$

4. STATISTICAL RESULTS AND DISCUSSION

The time series data were collected for the quarterly unemployment time series data over time period of January 2010 through December 2020. The data were collected from the official website of Statistics South Africa (StatsSA, 2020b).

6. Descriptive Data Analysis

Table 1 presents the summary statistics of he quarterly unemployment rate. The mean is greater than the median, which means that the distribution of the data is right skewed. This implies that most values with lower rates are clustered on the left of mean and few values with large or extreme rates to the right of distribution. The time series plot in Figure 4 identifies the type of outliers as the Innovational outliers. They occur in quarter 2 of 2020, which is when the government started to implement the lockdown for Covid -19 pandemic.

7. Time Series Analysis

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

Figure 4 shows the quarterly time series plots for quarterly unemployment rate (UR), and the series is not stationary. Figure 5 shows the first differencing quarterly time series plots for quarterly UR and the series is now stationary.

Table 1. Summary Quarterly Unemployment

Statistic	Unemployment percent
Minimum	23.30
First quartile	25.00
Median	25.50
Mean	26.12
Third quartile	27.15
Maximum	30.80

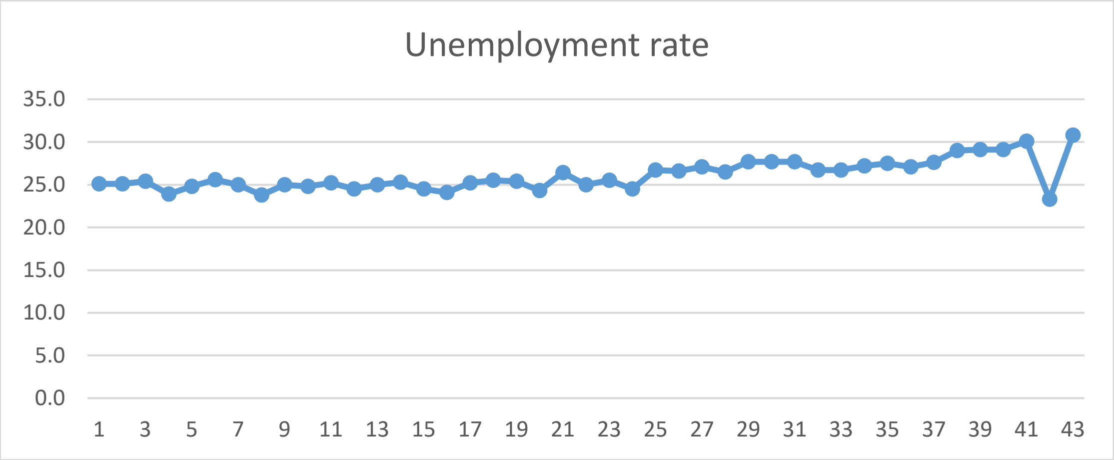

Figure 4. Time Series Plot for UR

Figure 5: First Difference Time Series Plot for UR

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

8. Post ARIMA analysis

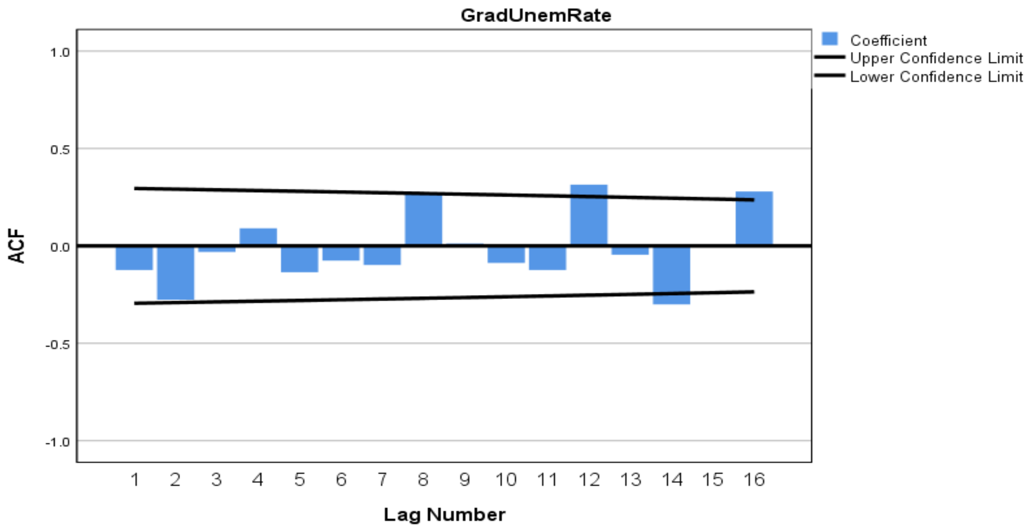

Figure 6. ACF Plot

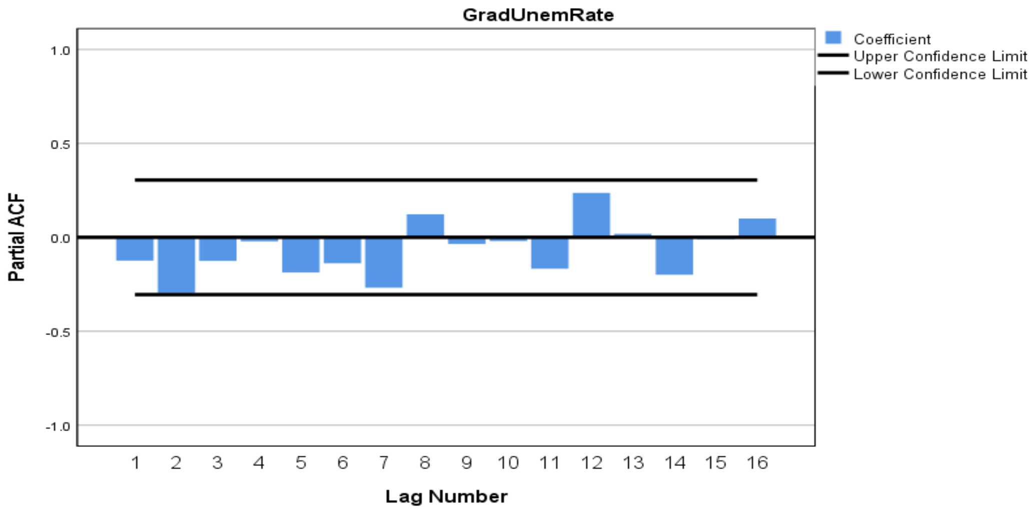

Figure 7. PACF Plot

The post ARIMA analysis is suitable for examining the stability of the model. Figures 6 and 7 present the computed values of Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) of the differenced series with lags up to 16 lies within the 5% significance limits. The correlograms do not show signs of seasonality since the ACF

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

lags are mostly significant at multiples of 2. This implies that the ARIMA model is stable.

Table 2: Estimated Models for Employment

Models	AIC	BIC
$ARIMA$ (1,1,1)	104.19	107.53
$ARIMA$ (2,1,1)	105.44	109.72
$ARIMA$ (0,1,1)	105.77	107.71
$ARIMA$ (2,1,0)	105.27	108.79
$ARIMA$ (2,1,2)	106.49	113.53
$ARIMA$ (1,1,1)	107.71	109.47

Table 2 presents the the estimated models for quarterly UR with the corresponding Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values. The model, ARIMA (1, 1, 1) is found to be the best suitable model for the UR series as its AIC = 104.19 and BIC = 107.53 shown in Table 2 are the lowest. An identified ARIMA (1, 1, 1) model is given by

$$(1-\phi_1 L)(1-L)X_t = (1-\psi_1 L)\varepsilon_t \tag{26}$$

The first differenced time series plot in Figure 5 shows that time series data are contaminated with additive outliers at $t = 42$. Thus

$$I_{AO}^{(42)} = \begin{cases} 1, & t = 42 \\ 0, & t \neq 42 \end{cases} \tag{27}$$

and $\omega = -6.8$. Table 6 presents the type of outlier, the corrupted value and the magnitude of an outlier. In this case, the outliers are not moved, but the parameters are estimated in their presence. The contaminated ARIMA (1, 1, 1) model with AO is given by

$$X_t^* = X_t + \omega I_t^{(T)}, \tag{28}$$

and also written as follows

$$X_t^* = \phi_1 X_{t-1} - \psi_1 \varepsilon_t + \varepsilon_t + \omega I_{AO}^{(42)} \tag{29}$$

The commonly used measures of scale, RMSE, MAPE, and MAE are used for these estimators. Tables 3-5 present the values of the parameter estimates MLE, M, and MM.

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

Table 7 presents the accuracy of the models measured by the RMSE, MAPE and MAE for the quarterly UR

Table 3. Parameter ML Estimates for ARIMA (1, 1, 1) for UR

Model Orders	Estimates (ML)	Standard Error	t-Value	P-Value
AR(1)	0.324	0.243	1.710	0.188
MA(1)	0.785	0.162	3.644	0.000

$$X_t^* = 0.324X_{t-1} - 0.785\varepsilon_t + \varepsilon_t - 6.8I_{AO}^{(42)} \tag{30}$$

Table 4. Parameter M-Estimates for ARIMA (1, 1, 1) for UR

Model Orders	Estimates (M)	Standard Error	t-Value	P-Value
AR(1)	0.423	0.254	1.671	0.102
MA(1)	0.812	0.171	4.614	0.000

$$X_t^* = 0.423X_{t-1} - 0.812\varepsilon_t + \varepsilon_t - 6.8I_{AO}^{(42)} \tag{31}$$

Table 5. Parameter MM-Estimates for ARIMA (1, 1, 1) for UR

Model Orders	Estimates (ML)	Standard Error	t-Value	P-Value
AR(1)	0.452	0.264	1.708	0.096
MA(1)	0.843	0.182	4.720	0.000

$$X_t^* = 0.452X_{t-1} - 0.432\varepsilon_t + \varepsilon_t - 6.8I_{AO}^{(42)} \tag{32}$$

Table 6 Outlier Points

Outlier		
Time	Type	ω
42	AO	-6.8

Table 7 Forecast Accuracy

Estimate	RMSE	MAPE	MAE	Normalized BIC
MLE	0.757	11.598	0.563	-0.294
M	0.703	11.003	0.523	-0.120
MM	0.645	9.345	0.434	-0.113

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

Table 8: ARIMA $(1, 1, 1)$ Forecast for 2021

Quarters 2021	LCL	Forecast	UCL
1	30.4	31.8	33.2
2	30.6	32.0	33.4
3	30.8	32.2	33.6
4	31.0	32.4	33.8

The common practice is that after detecting outliers, the outliers are removed or discarded during data cleaning and analysis. In this case, the forecast, lower and upper confidence limits are computed in the presence of outlier. The forecast of unemployment rates in South Africa for the period 2021 is presented in Table 8. The forecast of unemployment rates in South Africa revealed an increasing rate from Q1 2021 to Q4 2021 with the 95% confidence intervals from model 32.

5. CONCLUSION

The purpose of the study is to build the best time series model to predict the unemployment rate in South Africa in the presence of unforeseen events. During the national lockdown in the wake of the Covid -19 pandemic, in the second quarter, South Africa absolved a large number of jobs, and the unemployment rate also dropped. Forecasting with outliers is challenging as the presence of the outliers may produce poor forecasts. Two robust estimators; M-estimator and MM-estimators are used to estimate the parameters of ARIMA (1, 1, 1) in the presence of outliers. One additive outlier was detected and not removed from the data. The results showed that the M- and MM-estimates are significantly different from the MLE. It has proven that these estimators perform well in the presence of the innovational outliers. The model, ARIMA (1, 1, 1) was also found to be the best suitable model for forecasting the unemployment rate with unexpected events.

According to StatsSA (2020b), during Covid-19, a large number of jobs that were lost are in the trade industry, business services, community services, manufacturing, construction and transportation industry. Most of these sectors are not state-owned public entities, but are privately owned. Therefore, this implies that the stockholders in these entities are affected adversely. Based on these findings of the study, it is recommend that the South African government should strengthen employment by owning as many sectors as possible to ensure all workers are covered in times like these. Government should also review the stockholder policy by considering part-ownership of these entities for subsidy purposes, and for ensuring that they are governed as required by the law or government policy. Government representatives may be deployed on-site so that deviations from required methods of operation do not occur unnoticed. Moreover, it can empower small business services and transportation industrials through funds or loans as most of the workers in these sectors were hard done. These measures should be

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

taken despite the fact that South Africa, like other countries, has already provided temporary social grants for people who were unemployed or lost their jobs during this period. In the long run, this could still pose a problem as the country loaned the money from the World Bank and it is expected to pay it back. Further studies on how to account and cover for the loss of jobs/livelihoods in the presence of unexpected events should be considered.

REFERENCE

Abegaz, K. H., & Mohammed, A. A. (2018). Healthcare expenditure and GDP in Ethiopia from 1995 to 2014: a time-series analysis. *Agriculture & Food Security, 7*(1), 1-7. doi: 10.1186/s40066-018-0199-8

Abraham, B., & Chuang, A. (1989). Outlier Detection and Time Series Modeling. *Technometrics, 31*(2), 241-248. doi: 10.1080/00401706.1989.10488517

Adenomon, M. (2017). Modelling and forecasting unemployment rates in Nigeria using ARIMA model. *FUW Trends in Science and Technology Journal, 2*(1B), 525-531. doi: http://www.ftstjournal.com/uploads/docs/21B%20Article%2035.pdf

Ahmad, M., Khan, Y. A., Jiang, C., Kazmi, S. J. H., & Abbas, S. Z. (2021). The impact of COVID-19 on unemployment rate: An intelligent based unemployment rate prediction in selected countries of Europe. *International Journal of Finance & Economics*, 1-16. doi: https://doi.org/10.1002/ijfe.2434

Ahmar, A. S., Guritno, S., Abdurakhman, Rahman, A., Awi, Alimuddin, . . . Arifin, A. N. M. (2018). Modeling Data Containing Outliers using ARIMA Additive Outlier (ARIMA-AO). *Journal of Physics: Conference Series, 954*, 012010. doi: 10.1088/1742-6596/954/1/012010

AKTER, S. (2020, June 29-30). *Covid-19 and Bangladesh: Threat of Unemployment in the Economy.* Paper presented at the 5. International EMI Entrepreneurship and Social Sciences Congress PROCEEDINGS E-BOOK.

Bakar, N. M. A., & Midi, H. (2019, 2019/01). *The Applications of Robust Estimation in Fixed Effect Panel Data Model.* Paper presented at the 1st Aceh Global Conference (AGC 2018).*Atlantis Press* doi: https://doi.org/10.2991/agc-18.2019.54.

Bhatia, K., Jain, P., Kamalaruban, P., & Kar, P. (2016). Efficient and Consistent Robust Time Series Analysis. *arXiv e-prints, 1,* 1-26. doi: https://arxiv.org/abs/1607.00146

Box, G., Jenkins, G., Reinsel, G., & Ljung, G. (2015). Time Series Analysis: Forecasting and Control. New Jersey: John Wiley & Sons.

Broekhuizen, H. V. (2016). *Graduate unemployment and higher education institutions in South Africa*. Working Paper Series. Bureau for Economic Research and Stellenbosch Economic Retrieved from:https://resep.sun.ac.za/wp-content/uploads/2018/07/wp-08-2016.pdf.

Chakraborty, T., Chakraborty, A. K., Biswas, M., Banerjee, S., & Bhattacharya, S. (2021). Unemployment Rate Forecasting: A Hybrid Approach. *Computational Economics, 57*(1), 183-201. doi: 10.1007/s10614-020-10040-2

Claveria, O. (2019). Forecasting the unemployment rate using the degree of agreement in consumer unemployment expectations. *Journal for Labour Market Research, 53*(3), 1-10. doi: 10.1186/s12651-019-0253-4

Davidescu, A. A., Apostu, S.-A., & Marin, A. (2021). Forecasting the Romanian Unemployment Rate in Time of Health Crisis—A Univariate vs. Multivariate Time Series Approach. *International Journal of Environmental Research and Public Health, 18*(21), 1-31. doi: https://doi.org/10.3390/ijerph182111165

Davidescu, A. A., Apostu, S.-A., & Paul, A. (2021). Comparative Analysis of Different Univariate Forecasting Methods in Modelling and Predicting the Romanian Unemployment Rate for the Period 2021–2022. *Entropy, 23*(3), 1-31. doi: https://doi.org/10.3390/e23030325

Davidescu, A. A., Apostu, S.-A., & Stoica, L. A. (2021). Socioeconomic Effects of COVID-19 Pandemic: Exploring Uncertainty in the Forecast of the Romanian Unemployment Rate for the Period 2020–2023. *Sustainability, 13*(13), 1-22. doi: https://doi.org/10.3390/su13137078

Deneshkumar, V., & Kannan, K. S. (2011). OUTLIERS IN TIME SERIES DATA. *International Journal of Agricultural and Statistics Sciences, 7*(2), 685-691. doi: https://www.researchgate.net/profile/Deneshkumar-V/publication/274721817

Didiharyono, D., & Syukri, M. (2020). Forecasting with arima model in anticipating open unemployment rates in south sulawesi. *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH, 9*(3), 3838-3841. doi: https://osf.io/f5zh6/download

Dritsakis, N., & Klazoglou, P. (2018). Forecasting unemployment rates in USA using Box-Jenkins methodology. *International Journal of Economics and Financial Issues, 8*(1), 9-20. doi: https://econpapers.repec.org/RePEc:eco:journ1:2018-01-2

Fox, A. J. (1972). Outliers in Time Series. *Journal of the Royal Statistical Society: Series B (Methodological), 34*(3), 350-363. doi: https://doi.org/10.1111/j.2517-6161.1972.tb00912.x

Ghoto, A., Talpur, G. M., & Kamboh, S. (2021). Estimation and forecasting of the inflation, interest, literacy and unemployment rate of Pakistan using nonlinear regression models. *Indian Journal of Science and Technology, 14*(3), 251-258. doi: https://doi.org/10.17485/IJST/v14i3.1812

Gostkowski, M., & Rokicki, T. (2021). Forecasting the Unemployment Rate: Application of Selected Prediction Methods. *European Research Studies, 24*(3), 985-1000. doi: http://dx.doi.org/10.35808/ersj/2396

Graham, L., Williams, L., & Chisoro, C. (2019). Barriers to the labour market for unemployed graduates in South Africa. *Journal of Education and Work, 32*(4), 360-376. doi: 10.1080/13639080.2019.1620924

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

Grossi, L., & Nan, F. (2019). Robust forecasting of electricity prices: Simulations, models and the impact of renewable sources. *Technological Forecasting and Social Change, 141*, 305-318. doi: https://doi.org/10.1016/j.techfore.2019.01.006

Hampel, F. R. (1992). Introduction to Huber (1964) robust estimation of a location parameter *Breakthroughs in statistics* (pp. 479-491): Springer. doi:https://doi.org/10.1007/978-1-4612-4380-9_34

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (2011). *Robust statistics: the approach based on influence functions* (Vol. 196): John Wiley & Sons. Retrieved from: https://archive-ouverte.unige.ch/unige:23238.

Hawkins, D. M. (1980). *Identification of outliers* (Vol. 11): London: Chapman and Hall. doi:https://doi.org/10.1007/978-94-015-3994-4

Huber, P. J. (1965). A Robust Version of the Probability Ratio Test. *The Annals of Mathematical Statistics, 36*(6), 1753-1758. doi: http://www.jstor.org/stable/2239116

Kaiser, R., & Herrero, A. M. (1999). *Seasonal outliers in time series*. DES Working Papers Statistics and Econometrics: . University Carlos III of Madrid. Retrieved from:https://ideas.repec.org/p/cte/wsrepe/6333.html.

Kucharski, L. (2015). Long-term unemployment in Poland in 2004-2013. *Economic Studies, 210*, 177-187. doi: http://bazekon.icm.edu.pl/bazekon/element/bwmeta1.element.ekon-element-000171371185

Kumphon, B. (2013). A Forecasting Model for Thailand's Unemployment Rate. *Canadian Center of Science and Education, 7*(7), 10-16. doi: http://dx.doi.org/10.5539/mas.v7n7p10

Lai, H., Khan, Y. A., Thaljaoui, A., Chammam, W., & Abbas, S. Z. (2021). COVID-19 pandemic and unemployment rate: A hybrid unemployment rate prediction approach for developed and developing countries of Asia. *Soft Computing*, 1-16. doi: https://doi.org/10.1007/s00500-021-05871-6

Liu, L.-M., Bhattacharyya, S., Sclove, S. L., Chen, R., & Lattyak, W. J. (2001). Data mining on time series: an illustration using fast-food restaurant franchise data. *Computational Statistics & Data Analysis, 37*(4), 455-476. doi: https://doi.org/10.1016/S0167-9473(01)00014-7

Ljung, G. M. (1993). On Outlier Detection in Time Series. *Journal of the Royal Statistical Society: Series B (Methodological), 55*(2), 559-567. doi: https://doi.org/10.1111/j.2517-6161.1993.tb01924.x

Mabiza, J., Mahlasela, P., & Mbohwa, C. (2017, March 15-17). *Reducing Unemployment Rate in South Africa through Establishment of Graduate Internship Programmes (GIP).* Paper presented at the International MultiConference of Engineers and Computer Scientists IMECS, Hong Kong. Retrieved from: https://www.researchgate.net/publication/316874866

Mahmudah, U. (2017). Autoregressive integrated moving average model to predict graduate unemployment in Indonesia. *Practice and Theory in System of Education, 12*(1), 43-50. doi: 10.1515/ptse-2017-0005

Makaringe, S. C., & Khobai, H. (2018). The effect of unemployment on economic growth in South Africa (1994-2016). *Munich Personal RePEc Archive*, 1-15. doi: https://mpra.ub.uni-muenchen.de/85305/

Martin, R. D., & Yohai, V. J. (1986). Influence Functionals for Time Series. *The Annals of Statistics, 14*(3), 781-818. doi: http://www.jstor.org/stable/3035535

Muler, N., Pena, D., & Yohai, V. J. (2009). Robust estimation for ARMA models. *The Annals of Statistics, 37*(2), 816-840. doi: https://doi.org/10.1214/07-AOS570

Nguyen, P.-H., Tsai, J.-F., Kayral, I. E., & 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), 1-27. doi: https://doi.org/10.3390/su13147879

Nkoane, S. S. (2019). *Robust CM-estimation for univariate time series.* (PhD thesis), Sefako Makgatho Health Sciences University, Ga-Rankuwa, South Africa.

Pasara, M. T., & Garidzirai, R. (2020). Causality Effects among Gross Capital Formation, Unemployment and Economic Growth in South Africa. *Economies, 8*(2), 1-12. doi: 10.3390/economies8020026

Ramli, S. F., Firdaus, M., Uzair, H., Khairi, M., & Zharif, A. (2018). Prediction of the unemployment rate in Malaysia. *International Journal of Modern Trends in Social Sciences, 1*(4), 38-44. doi: http://www.ijmtss.com/PDF/IJMTSS-2018-04-12-04.pdf

Richa, M., Rakesh, G., & Rimzim, M. (2015). Graduate unemployment - A case of Jolgaon city. *IOSR journal of business and management, 17*, 1-9.

San-Juan, J. F., San-Martín, M., & Pérez, I. (2012). An economic hybrid J2 analytical orbit propagator program based on SARIMA models. *Mathematical Problems in Engineering, 2012*. doi: 10.1155/2012/207381

StatsSA. (2020a). Quarterly Employment Statistics (QES), 2nd Quarter. *Statistics South Africa*. Retrieved from: http://www.statssa.gov.za/?p=13690

StatsSA. (2020b). Quarterly Labour Force Survey (QLFS) Report. *Statistics South Africa*. Retrieved from: http://www.statssa.gov.za/publications/P0211/P02112ndQuarter2020.pdf

Stockinger, N., & Dutter, R. (1987). Robust time series analysis: A survey. *Kybernetika, 23*(1), 3-88. doi: http://dml.cz/dmlcz/124955

Walter, D. (2020). Implications of Covid-19 for Labour and Employment in India. *The Indian Journal of Labour Economics, 63*(1), 47-51. doi: 10.1007/s41027-020-00255-0

Wang, F.-K., & Lee, C.-W. (2011). M-estimator with asymmetric influence function for estimating the Burr type III parameters with outliers. *Computers & Mathematics with Applications, 62*(4), 1896-1907. doi: https://doi.org/10.1016/j.camwa.2011.06.033

Vol: 13 **No:** 2 **Year:** 2021 **ISSN**: 1309-8055 (Online) (**pp.** 199-222) **Doi:** 10.34109/ijefs.20212010

Yohai, V. J. (1987). High Breakdown-Point and High Efficiency Robust Estimates for Regression. *The Annals of Statistics*, *15*(2), 642-656. doi: http://www.jstor.org/stable/2241331

Yohai, V. J., Stahel, W. A., & Zamar, R. H. (1991). *A Procedure for Robust Estimation and Inference in Linear Regression*. Paper presented at the Directions in Robust Statistics and Diagnostics, New York.*Springer* doi: https://doi.org/10.1007/978-1-4612-4444-8_20.