Faculty of Graduate Studies

Master Program in Applied Statistics and Data Science

Comparison of ARIMA and Artificial Neural Networks Models for Forecasting Unemployment Rate Behavior in Palestine

Prepared by

Mu'men Hasan

Student number: 1205365

Supervisor

Dr. Hassan Abu Hassan

Co-supervisor

Dr. Mohsen Ayyash

Master Thesis

Birzeit University

Palestine

Faculty of Graduate Studies

Master Program in Applied Statistics and Data Science

Comparison of ARIMA and Artificial Neural Networks Models for Forecasting Unemployment Rate Behavior in Palestine

Prepared by

Mu'men Hasan

Student number: 1205365

Supervisor

Dr. Hassan Abu Hassan

Co-supervisor

Dr. Mohsen Ayyash

This thesis was submitted in partial fulfillment of the requirements for the master's degree in applied statistics and data science from the faculty of graduate studies at Birzeit University, Palestine.

January 2024

BIRZEIT UNIVERSITY

Faculty of Graduate Studies

Master Program in Applied Statistics and Data Science

Comparison of ARIMA and Artificial Neural Networks Models for Forecasting Unemployment Rate Behavior in Palestine

Prepared by

Mu'men Hasan

Student number: 1205365

This Thesis was defended successfully on January 27, 2024, and approved by

Committee Members:		**Signature**
Dr. Hassan Abu Hassan	(Supervisor)	
Dr. Mohsen Ayyash	(CoSupervisor)	Mohsen Ayyash
Dr. Tareq Sadeq	(Internal Examiner)	
Dr. Sameera Awawda	(Internal Examiner)	

Acknowledgments

It's a great pleasure for me to thank Dr. Hassan Abu Hassan and Dr. Mohsen Ayyash who kindly agreed to supervise my thesis. I would like to express my deep appreciation to Dr. Abu Hassan and Dr. Ayyash for their great support, encouragement, and understanding. It was a great honor for me to work with Dr. Abu Hassan and Dr. Ayyash and to learn from their experience and scientific knowledge. I also thank Dr. Tareq Sadeq and Dr. Sameera Awawda for accepting to be committee members and also for their comments and suggestions.

Dedication

I dedicate this thesis to my parents, my brother and sisters, and friends who have supported me throughout my academic journey.

I would like to thank my mother; Without her endless love and encouragement, I would never have been able to complete my graduate degree.

All gratitude to my best friends Mohammad Shalalfeh and Moath Hoshiea.

With all my love and respect.

Abstract

Palestine exhibits a higher rate of unemployment with an increasing trend over the past few decades that may have adverse impacts on the overall economy and the entire population, welfare, and life satisfaction among those out of work. The key purpose of the current study is to forecast the behavior of the Palestinian unemployment rate using seasonal autoregressive integrated moving average (SARIMA) and neural network auto-regression (NNAR) models. Furthermore, the current study compares the forecasting performance obtained from SARIMA and NNAR methods. The current study utilizes quarterly unemployment rates in Palestine over the period from 1996Q1 – 2023Q2 provided by the Palestinian Labor Force Survey (PLFS) published by the Palestinian Central Bureau of Statistics (PCBS). In this context, this study divides the data into two datasets, namely, training (in-sample) and testing (out-of-sample). The training dataset covers the period from 1996Q1 to 2017Q4 while the testing dataset covers the remaining period. This study found that SARIMA(0,1,1)(0,1,1)[4] is the best classical linear time series model to forecast the out-of-sample unemployment rate in Palestine. Alternatively, the findings of this study revealed that the best nonlinear model using neural networks was NNAR(1,1,10)[4]. This study demonstrated that NNAR(1,1,10)[10] outperformed SARIMA(0,1,1)(0,1,1)[4] in terms of various measures of forecasting accuracy including MAE, MAPE, RMSE, and MASE, predictive values, and the DM test for predictive accuracy, for both in-sample and out-of-sample datasets. In this context, the emerged results of the present study about the best models using the NNAR(1,1,10)[4] model, show the predicted Palestinian unemployment rate in the third quarter of 2023 is 24.1% compared to 24.7% in the second quarter of 2023, which exhibited a slight downward trend while for the overall forecasting period, it is expected to range from 23.2 to 27.5%. The key findings of this study highlight that univariate forecasting of the Palestinian unemployment rate using NNAR methods outperforms the results obtained from the SARIMA models. Finally, the study shows several important potential recommendations that can be forwarded for both policy implications and future research, Further work is needed to compare more forecasting methods including Holt-Winters, self-exciting threshold auto-regressive models, machine learning, and hybrid methods. Moreover, to increase the accuracy of the forecasting models, it is important

to include important factors of unemployment in future work such as inflation, interest rate, gross domestic product, and percentage of tertiary education.

الملخص

تُظهر فلسطين معدلات بطالة عالية مع اتجاه متزايد خلال العقود القليلة الماضية، وهذا قد يكون له آثار سلبية على الاقتصاد العام وعلى السكان والرفاهية والرضا عن الحياة بين العاطلين عن العمل. الهدف الرئيسي من هذه الدراسة هو التنبؤ بسلوك معدل البطالة في فلسطين باستخدام نموذج الانحدار الذاتي والمتوسطات المتحركة الموسمي (SARIMA)، ونموذج الشبكات العصبية الاصطناعية (ANNs) وفحص أداء التنبؤ الذي تم الحصول عليه من نماذج SARIMA و ANNs من خلال الانحدار الذاتي للشبكة العصبية (NNAR). الدراسة الحالية تستخدم معدلات البطالة الربعية في فلسطين في الفترة من الربع الأول في العام 1996 الى الربع الثاني من العام 2023 الذي تم من خلال مسح القوى العاملة الفلسطينية (PLFS) ونشره الجهاز المركزي الفلسطيني للإحصاء (PCBS). وفي هذا السياق، قسمت هذه الدراسة البيانات إلى مجموعات بيانات تدريبية وبيانات الاختبار. تغطي مجموعة البيانات التدريبية الفترة من الربع الأول للعام 1996 إلى الربع الرابع للعام 2017، بينما تغطي مجموعة بيانات الاختبار الفترة المتبقية. وجدت هذه الدراسة أن SARIMA (0,1,1) (0,1,1) [4] هو أفضل نموذج زمني خطي كلاسيكي للتنبؤ بمعدل البطالة في عينة الاختبار في فلسطين. من ناحية أخرى، كشفت نتائج هذه الدراسة أن أفضل نموذج غير خطي باستخدام الشبكات العصبية هو NNAR (1,1,10) [4]. اما بما يتعلق بمقارنة أداء التنبؤ، أظهرت هذه الدراسة أن NNAR (1,1,10) [10] تفوق في الأداء على SARIMA (0,1,1) (0,1,1) [4] من حيث مقاييس دقة التنبؤ (مثل RMSE و MAE و MAPE و MASE) والقيم التنبؤية واختبار DM للدقة التنبؤية لكل من العينات التدريبية والاختبارية. وفي هذا السياق، أظهرت نتائج الدراسة الحالية حول أفضل النماذج باستخدام نموذج NNAR(1,1,10) [4] أن معدل البطالة الفلسطيني المتوقع في الربع الثالث من عام 2023 سيكون 24.1% مقارنة في 24.7% في الربع الثاني من عام 2023، وايضا أظهر النموذج اتجاه هابط طفيف لمعدلات البطالة ومن المتوقع أن يتراوح من 23.2% إلى 27.5٪ ما بين الربع الثالث من العام 2023 والربع الثاني من العام 2025. وتسلط النتائج الرئيسية لهذه الدراسة الضوء على أن التنبؤ احادي المتغير بمعدل البطالة الفلسطيني باستخدام نموذج NNAR يفوق النتائج من التي تم الحصول عليها من نماذج SARIMA. أخيراً، تُظهر الدراسة العديد من التوصيات المحتملة المهمة التي يمكن تقديمها لكل من الآثار المترتبة على السياسية والأبحاث المستقبلية، انه هناك حاجة إلى مزيد من العمل لمقارنة طرق التنبؤ بما في ذلك Holt-Winters، SETAR، والتعلم الآلي، والأساليب المدمجة. ايضاً، لزيادة دقة نماذج التنبؤ، من المهم إدراج عوامل مهمة للبطالة في العمل المستقبلي مثل الناتج المحلي الإجمالي، ومعدل التضخم، وسعر الفائدة، والنسبة المئوية للتعليم العالي.

Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

PCBS	Palestinian Central Bureau of Statistics
PLFS	Palestinian Labor Force Survey
CPI	Consumer Price Index
GDP	Gross Domestic Product
AR	Autoregressive
MA	Moving Average
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
SARIMA	Seasonal Autoregressive Integrated Moving Average
ANNs	Artificial Neural Networks
NNAR	Neural Network Autoregression
ACF	Autocorrelation Function
PACF	Partial Autocorrelation Function
EACF	Extended Autocorrelation Function
AIC	Akaike Information Criterion
BIC	Bayes Information Criterion
ME	Mean Error
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
MAE	Mean Absolute Error
MASE	Mean Absolute Scaled Error
MSE	Mean Squared Error
R^2	R Squared
ILO	International labor organization
ADF	Augmented Dickey-Fuller
Q*	Ljung-Box Test Value
SVM	Support Vector Machines
RF	Random Forests
DT	Decision Trees

CHAPTER ONE

INTRODUCTION

1.1 Introduction

This chapter is devoted to showing the introduction of this thesis. The Second Section explains the study background. The Third Section motivates the problem under investigation in this study. The Fourth Section shows the main purpose and objectives of this thesis. The Fifth Section states the hypothesis of this study. The Sixth Section shows the significance of study. The Seventh Section illustrates the study's limitations. The Last Section shows the Organization of this thesis.

1.2 Study Background

Unemployment is a serious economic problem that receives great concern from policymakers worldwide, especially in developing countries because it may have social and economic implications. Accordingly, unemployment can have negative effects on the overall economy and on the entire population, welfare, and life satisfaction of those who are unemployed (Abugamea, 2018).

As one of the developing countries, Palestine is characterized by long-term and high unemployment rates in the past few decades. According to the Palestinian Central Bureau of Statistics (PCBS, 2022), the unemployment rate scored 26.4% in 2021 and it is more pronounced among females (42.9% for females versus. 22.4% for males), youth population aged 15 – 24 years (37.2% for males and 64.5% for females) and in the Gaza Strip (46.9%) as compared to the West Bank (15.5%). The main cause of these high unemployment rates emerged primarily due to the consequences of Israeli occupation policies that disrupted the Palestinian economy, the prolonged siege to the Gaza Strip, and the restricted movement of Palestinian workers to the Israeli labor market (Abugamea, 2018). A study by Daoud (2006) suggested that unemployment in Palestine exhibits various characteristics, such as a high incidence, prolonged duration, significant volatility, a close association with political circumstances, and a distinct contrast between the West Bank and the Gaza Strip. In the same way, the World Bank (2012) established that the Palestinian labor market from 2000 onwards has displayed persistent high unemployment rates, a continuous decrease in youth employment and economic engagement, and notably low levels of female labor participation.

Several research's on unemployment forecasting have been conducted with different approaches that have been proposed in recent years to inform policy decisions and overcome this problem. These approaches can be categorized into two general types: statistical methods and soft computing techniques. Accordingly, different statistical methods have been applied to forecast unemployment rates across different countries worldwide (Dritsakis & Klazoglou, 2018; Nikolaos, Stergios, Tasos, & Ioannis, 2016; Nkwatoh, 2012; Barnichon & Garda, 2016). Traditional statistical techniques included autoregressive integrated moving average (ARIMA), exponential smoothing, and generalized autoregressive conditional heteroskedasticity (GARCH) volatility models, among others (Adebiyi, Adewumi, & Ayo, 2014; Tabachnick & Fidell, 2001). Meanwhile, the ARIMA model (i.e., Box-Jenkins model) is frequently employed in forecasting and is widely considered as the most effective method for forecasting across various fields, it finds extensive application in handling time series data. ARIMA models rely on previous values of the series and previous error terms for forecasting (Mishra & Desai, 2005; Meyler, Kenny, & Quinn, 1998; Tabachnick & Fidell, 2001; Box & Jenkins, 1976). Nevertheless, ARIMA models exhibit greater resilience and efficiency when it comes to short-term forecasting, especially when compared to more complicated structural models. (Meyler, Kenny, & Quinn, 1998).

On the other hand, Machine learning (ML) and artificial neural networks (ANNs) belong to the realm of soft computing techniques, which are considered an alternative method to classical time series models. These methods provide more accurate forecasting results and are widely used in multidisciplinary fields like economics, social sciences, engineering, business, finance, etc. that mostly have a nonlinear behavior (Tansel, et al., 1999; Khashei & Bijari, 2010). Artificial neural networks (ANNs) are data-driven and adaptive techniques that operate with minimal initial assumptions. Furthermore, ANNs are regarded as effective predictors, more accurate, and very efficient models in solving real-world problems especially those that have non-linear patterns that allow making generalized observations from the findings learned from original data. Peláez (2006) used ANNs to forecast unemployment rates in the USA. The neural networks Autoregression (NNAR) model is one of the most commonly used univariate time series forecasting methods, which depends on the past values of the values of the series in the hidden input layers. This method has been widely applied to forecast unemployment rate in different countries worldwide and contrasted with classical time series models in terms of forecasting accuracy and performance (Vicente et al., 2015; Davidescu et al., 2021; Dumičić et al., 2015; Yamacli & Yamacli, 2023). This study focuses on comparing the forecasting performance of the Palestinian unemployment rate between seasonal ARIMA and NNAR models.

1.3 Problem Statement

The unemployment is a crucial labor market metric, signifying a disparity between labor supply and demand. The implications of this indicator have substantial social and economic ramifications, playing a pivotal role in macroeconomic assessments and serving as a key factor when evaluating a country's economic performance from a labor standpoint. (Kavkler, et al., 2009).

Furthermore, the Palestinian labor market faces significant challenges such as Israeli occupation and its associated restrictions on movements, lower rates of labor force participation, especially among females, and higher rates of unemployment as compared to other neighboring and developed countries. Therefore, It is essential to investigate the behavior of unemployment rates in Palestine to inform policy decisions and reinforce measures to alleviate it.

Moreover, given that most economic problems are non-linear, ARIMA models may fail to provide more accurate and efficient prediction results as compared to NNAR models.

Therefore, there is an increasing requirement to tackle strongly non-linear, time-changing challenges, as various real-world scenarios like unemployment rates exhibit nonlinearity and unpredictable fluctuations over time (Khashei, Bijari, & Ardali, 2009; Khashei & Bijari, 2010).

Finally, the current literature showed inconclusive results regarding the best methods to forecast the unemployment rate and some studies attributed it to geographical and time horizon variations (Ahmad et al., 2017; Davidescu et al., 2021; Katris, 2019). Thus, this study compares the performance of ARIMA and NNAR models for a case of the unemployment rate in Palestine.

1.4 Study Aim and Objective

This study mainly aims to forecast the behavior of unemployment rates in Palestine using ARIMA and NNAR models. The more specific objectives are:

1. To forecast the unemployment rates in Palestine using ARIMA models.
2. To forecast the unemployment rates in Palestine using NNAR models.
3. To compare the performance of the prediction models obtained from ARIMA and NNAR models.

This study aims to analyze and understand the short-run variation in the unemployment rate. We aim to investigate the temporary fluctuations and patterns within the time series data, focusing on identifying and characterizing short-term dynamics.

1.5 Study Hypotheses

The current thesis is intended to seek an answer to the following main hypothesis:

H_0: There is no difference in the forecasting performance and accuracy between ARIMA models and NNAR.

This null hypothesis assumes that both ARIMA and NNAR perform equally well in forecasting, with no significant difference in their ability to predict future outcomes (i.e., unemployment rate). To test this hypothesis, statistical methods such as hypothesis testing, cross-validation, or forecast accuracy metrics will be employed to compare the performance of the two forecasting methods. If the null hypothesis is rejected, it would indicate that there is a significant difference in the forecasting performance and accuracy between ARIMA and NNAR.

1.6 Significance of Study

The significance of studying unemployment lies in its far-reaching impact on individuals, communities, and economies. Therefore, studying unemployment is essential for understanding the dynamics of labor markets, promoting economic stability, addressing social challenges, formulating effective policies, developing human capital, and facilitating economic and financial planning. Hence, this research project will contribute to the current knowledge about unemployment in various ways. First, it will provide a comprehensive picture of the behavior and nature of unemployment in Palestine. Second, the use of Artificial Neural Networks (ANNs) for forecasting unemployment may bring improved accuracy, timely predictions, and the capacity to comprehend intricate connections between variables as compared to traditional time series ARIMA models, which will enhance policy formulation, risk assessment, and planning as well as ultimately supporting policymakers in making informed decisions. Third, studying unemployment has significant contributions and impacts on society in various ways, including but not limited to supporting economic stability, increasing labor force participation, boosting productivity, stimulating economic growth, social inclusion, and improving well-being.

In summary, unemployment forecasting plays a vital role in informing policy decisions and helps society by facilitating effective policy planning, resource allocation, early intervention, skill development, and the design of social safety nets. By leveraging accurate forecasts, therefore, policymakers and the government of Palestine can better address unemployment challenges and support individuals and communities in their transition to a more resilient and prosperous future. Moreover, by addressing unemployment and its associated challenges, society can foster equitable, resilient, and prosperous communities that provide opportunities for all individuals to thrive.

1.7 Limitations of Study

This study has some potential limitations. To mention a few, the data have outliers and thus affected the performance of ARIMA and SARIMA models. Furthermore, this study did not take into account the spatial and gender differences in the unemployment rate in Palestine given that there are significant variations in the unemployment rate by gender and geographic areas (e.g., West Bank and Gaza Strip).

1.8 Thesis Organization

This thesis is organized as follows. The Second Chapter is devoted to showing the literature review. The Third Chapter explains the data and methodology of this study. The Fourth Chapter illustrates the statistical results and discussions. The Last Chapter pertains to the conclusions and recommendations based on the study findings.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter is devoted to providing an overview of unemployment and to summarize the findings of the empirical studies that aimed to forecast the unemployment rate across different countries worldwide. The Second Section pertains to the overview of unemployment. The Third Section provides an extensive review of the literature regarding the methods used in forecasting the unemployment rate. The last section concludes the chapter and summarizes the main findings of the reviewed studies.

2.2 Overview of Unemployment

Unemployment emerged due to economic deficiencies in the employment sector, manifesting both during the transition to a market economy and in times of economic expansion (Bădulescu, 2006). Unemployment stands as a noteworthy labor market problem, indicating a disparity between the availability of labor and its demand. This metric carries substantial social and economic repercussions and is a key component in assessing macroeconomic growth, making it a crucial factor when evaluating a country's economic performance in terms of its workforce. (Barrow & Kourentzes, 2018)

In this context, unemployment emerges as a significant worldwide social concern, impacting each nation to different extents, contingent on their economic progress. The expansion of the population results in a larger labor force, but in the short term, available job opportunities may fall short of meeting this growing demand (Cai & Wang, 2010). The efforts to modify the economic structure, reform the education system, and create specialized fields do not adequately address the requirements for economic transformation. The rural labor force's skill set falls short of meeting the job market demands, exacerbating the unemployment problem. A potential solution to this challenge involves implementing an early warning system for unemployment, emphasizing the importance of forecasting. (Bussiere & Fratzscher, 2006).

Following the international guideline, individuals identified as unemployed are those aged between 15 and 74 who meet three specific criteria simultaneously: they do not currently hold a job, are ready to commence employment among the next two weeks, and have actively sought employment at any time in the past four weeks. The unemployment rate quantifies the proportion of the unemployed within the labor force, which encompasses all individuals in a country who are capable of contributing their labor to produce goods and services within the designated time period, encompassing both employed and unemployed individuals. (International Labor Organisation, 2013).

2.3 Literature Review

In recent years there has been a growing interest in modeling and predicting macroeconomic indicators such as gross domestic product, consumer price index, inflation rate, and unemployment rate using time series analyses and machine learning approaches like NNAR models. Accordingly, several studies compared the forecasting accuracy and performance of models for unemployment rates worldwide (Gogas, Papadimitriou , & Sofianos, 2022; Yamacli & Yamacli, 2023).

Previous research has applied ARIMA models to forecast unemployment rates in various countries worldwide showing evidence of the usefulness and effectiveness of such classical prediction methods and out-of-sample forecasts (Vicente, López-Menéndez, & Pérez, 2015; Yamacli & Yamacli, 2023; Dumičić, Čeh Časni, & Žmuk, 2015; Khan Jafur, Sookia, Nunkoo Gonpot, & Seetanah, 2017). For example, a study by (Yamacli & Yamacli, 2023) forecasted the unemployment rate in Turkey over the period 1 August 2008 to 31 August 2022 using ARIMA and ANN and compared the forecasting accuracy obtained from both methods. They demonstrated that the perfect model for predicting the unemployment rate in Turkey was ARMA (2,1) because it has the lowest forecasting accuracy measures including R^2, RMSE, MAE, and MAPE. However, the prediction error during COVID-19 period obtained from ANNs was less than those obtained from ARMA(2,1) indicating that ANNs are more accurate to forecast unemployment rate under economic uncertainty.

Therefore, the conventional ARIMA models were reasonable for forecasting stochastic time series data. However, (Nagao, Takeda, & Tanaka, 2019) indicated that nonlinear short-run forecasting of the seasonally adjusted unemployment rate in the U.S. outperforms linear models.

Within the European Union, predictions for the unemployment rate are made utilizing the Box–Jenkins and TRAMO/SEATS methodologies. (Gagea & Balan, 2008; Mladenovic, Ilic, & Kostic, 2017). Applying a spatial GVAR model in the regions of Germany (Schanne, Wapler, & Weyh, 2010). In Greece, both dynamic and static processes have been implemented through SARIMA models. (Dritsaki, 2016; Dritsakis, Athianos, Stylianou, & Samaras, 2018). Slovakia has utilized GARCH and ARIMA models to predict the unemployment rate. (Rublikova & Lubyova, 2013).

The use of machine learning such as ANNs, deep learning, and support vector machines (SVM) in forecasting unemployment rates were also evident in the literature that account for the non-linear pattern of the data (Gogas, Papadimitriou , & Sofianos, 2022; Chakraborty, Chakraborty, Biswas, Banerjee, & Bhattacharya , 2021; Firmino, de Mattos Neto, & Ferreira, 2014). The findings of some previous studies indicated that ANNs showed more accurate forecasting results during the asymmetries of the business cycle or during economic uncertainty (Feuerriegel & Gordon, 2019; Chakraborty, Chakraborty, Biswas, Banerjee, & Bhattacharya , 2021). These findings were confirmed from the U.S and Turkish unemployment rates as well (Yamacli & Yamacli, 2023; Peláez , 2006). Furthermore, a study by (Gogas, Papadimitriou , & Sofianos, 2022) forecasted the unemployment rate in the euro area utilizing monthly unemployment rates over the period from April 1998 to September 2019. The study applied three different machine learning approaches including decision trees, random forests, and support vector machines. The findings revealed that RF with higher forecasting accuracy as compared to DT and SVM methodologies.

A study by Davidescu et al. (2021) compared the predicting performance of monthly unemployment rate in Romania over the period 2000 – 2020 Derived through diverse methods, which encompass exponential smoothing models, NNAR, SARIMA, and the self-exciting threshold autoregressive model. Based on the evaluation of in-sample forecast performance using accuracy measures of RMSE, MAE, MAPE, It was evident that the multiplicative Holt-Winters model was one of the models that performed exceptionally well. In assessing the testing sample forecasting performance, the RMSE and MAE values indicated that the NNAR model surpassed the performance of other models. Nevertheless, in terms of MAPE evaluation, the SARIMA model demonstrated higher forecast accuracy. Additionally, the Diebold-Mariano test results, conducted for a single forecast horizon with testing sample methods, highlighted variances in the forecasting efficacy between SARIMA and NNAR. Meanwhile, the NNAR model was deemed the superior choice for the modeling and prediction of the unemployment rate.

On the other hand, some studies adopted a hybrid approach to forecast the unemployment rate that combines linear and non-linear models to reduce bias and variances of the forecasting error of component models. (Chakraborty, Chakraborty, Biswas, Banerjee, & Bhattacharya , 2021) Projected the unemployment rate through a hybrid approach in seven countries: Canada, Germany, Japan, New Zealand, Netherlands, Switzerland, and Sweden. The findings indicated that the hybrid models outperform the conventional time series models but it cannot show the explosive behavior of the variances over time. They showed that ARIMA models were shown to be useful in forecasting stochastic time series while ANNs showed promising forecasting results over the past few decades. However, ANNs failed to produce the optimal network architecture (Chakraborty, Chakraborty, Biswas, Banerjee, & Bhattacharya , 2021). Therefore, NNAR has been recently introduced to overcome this limitation of ARIMA models. NNAR consists of a neural network with only one hidden layer, where the inputs to the network are lagged values of the time series that combines elements of autoregressive models and feed-forward neural networks for time series analysis (Teräsvirta, Van Dijk, & Medeiros, 2005). The primary advantage of this approach lies in its minimal complexity and straightforward interpretability compared to the invention of artificial neural networks (ANN). (Hyndman & Athanasopoulos, 2018).

Furthermore, certain studies employed a multivariate approach to predict the unemployment rate. An investigation conducted in Turkey explores the correlation between unemployment and inflation over the period from 1988 to 2002. The study employs a Vector Autoregressive Model (VAR) along with impulse response analysis to elucidate this relationship. The findings indicate an inverse relationship between unemployment and inflation. (ERDAL, DOĞAN, & KARAKAŞ, 2015). Additionally, research conducted in Western Balkan countries, focusing on Albania, Serbia, Macedonia, Montenegro, Bosnia-Herzegovina, and Kosovo, examines the influence of specific macroeconomic indicators on the unemployment rate. This study spans from 2000 to 2017 and employs the vector autoregressive model (VAR) as its methodology. The findings reveal that all analyzed macroeconomic variables—including inflation, interest rates, GDP, and FDI—significantly impact the unemployment rate within this group of nations. (Vladi & Eglantina, 2019)

Nonetheless, Katris (2019) highlighted that there is no universal and best statistical approach to forecast unemployment rate and the choice of the best model depends on factors such as the forecasting horizon and geographic location. In this study, results of traditional linear ARIMA and nonlinear NNAR models are compared to fill gap in the literature comparing the performance of these models with respect to modeling unemployment rate in Palestine.

2.4 Chapter Summary

This chapter discussed the overview of the unemployment problem in the labor market and its definition according to the International Labor Organization. Moreover, an extensive review of the literature was conducted and revealed that most studies indicated that non-linear models including NNAR models outperformed classical ARIMA forecasting. However, The reviewed studies do not provide definitive results for forecasting the unemployment rate. Table 2.1 summarizes the findings from the related literature. The next Chapter shows the methodology of the study.

Table 2.1. Summary of the Findings of the Empirical Studies.

Author/s (year)	Country	Data frequency	Methods	Main Findings
Davidescu et al. (2021)	Romania	Monthly 2000 - 2020	Holt-Winters (additive, multiplicative, and ETS); SARIMA; NNAR; SETAR	- Holt-Winters multiplicative method was better than others for training dataset. - NNAR models outperform SARIMA models in the testing dataset based on RMSE, MAE, and DM test.
Yamacli & Yamacli (2023)	Turkey	Monthly 01.01.2008-31.08.2022	ARIMA and NNAR	- ARMA(2,1) has better performance as compared to ANN model based on RMSE, MAE, MAPE, MASE, and R-squared when there is no economic crisis. - ANN has better performance than ARMA(2,1) when there is economic crisis.
Didiharyono D. & Muhammad Syukri (2020)	South Sulawesi, Indonesia	Yearly 1986 – 2004 ,& Every 6 months 2005 - 2018	ARIMA	- The ARIMA (1,2,1) model yielded the most effective time series forecast, as evidenced by its small Mean Square value.
Nikolaos Dritsakis ,& Paraskevi Klazoglou (2018)	USA	Monthly January 1995 – July 2017	ARIMA, ARCH ,& GARCH	- The SARIMA(1,1,2)(1,1,1)12 − GARCH(1,1) model demonstrated superior performance in projecting US unemployment.
Michał Gostkowski ,& Tomasz Rokicki (2021)	Poland	Monthly January 2008 – December 2018	The naive method, the regression model, ARIMA, Holt model and Winters model.	- The findings suggest that the quadratic regression model and the Winters multiplicative model are the most appropriate methods for forecasting the unemployment rate.
Chakraborty et al. (2020)	Canada, Germany, Japan, Netherlands, New Zealand, Sweden, and Switzerland	Monthly data except New Zealand quarterly data	ARIMA & NNAR. (A Hybrid Approach)	- The findings regarding the asymptotic stationarity of the hybrid approach, incorporating Markov chains and nonlinear time series analysis techniques. These findings ensure that the proposed model does not exhibit 'explosive' behavior or an increasing variance over time.
Gogas et al. (2022)	Euro-Area	Monthly April 1998 – September 2019	decision trees (DT), random forests (RF), and support vector machines (SVM)	- RF model outperforms the other models by reaching a full-dataset forecasting accuracy of 88.5% and 85.4% on the out-of-sample.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter pertains to the data used and methods employed in this study. The data used in this study in addition to the description of the two broad techniques, namely, ARIMA, and ANNs approaches to predict the unemployment rates in Palestine. In the ARIMA model, the following steps are performed: Exploring data, dividing data into two subsets, namely, training, and test data, stationarity test, transforming non-stationary data, model identification, model fit, model diagnostic, and forecasting with accuracy check. In the ANNs model, the following steps are performed: dividing the data, fitting the ANN model, model diagnostic, and forecasting. Furthermore, evaluation of the model performance obtained from ARIMA and ANNs models will be compared to determine the superior model by calculating the Mean Absolute Error, Mean Absolute Percentage Error, and Root Mean Square Error.

3.2 Data Description

This study utilizes secondary quarterly unemployment data during the period from the first quarter of 1996 to the second quarter of 2023 that was extracted from the quarterly reports of the Palestinian Labor Force Survey (PLFS) and press releases published via the Palestinian Central Bureau of Statistics (PCBS).

The main indicator of unemployment used as a primary outcome is the unemployment rate as per the definition of the International Labor Organization (ILO Standards ICLS-19th).

The dataset consists of 110 observations, and includes the following variables:

1. Date/Time: Each observation in the dataset is associated with a specific quarter and year, allowing for temporal analysis.
2. Unemployment Rate: This is the main variable of interest, representing the unemployment rate for each quarter.

3.3 Time Series Models

3.3.1 Overview

The data that can be obtained from observations through sequential collection over time are referred to the time series data. It is a series of data points gathered across successive time intervals such as days, months, years, or even minutes or seconds. It can be used in various disciplines, including finance, economics, meteorology, engineering, and the social sciences, frequently employing time series data. In addition to assisting in the discovery of patterns, trends, and seasonality in the data, it allows for the investigation of how variables change over time.

Time series analyses encompass the use of statistical and mathematical methods to comprehend and analyze the patterns and behavior of the data. It includes tasks like predicting future values, identifying anomalies, figuring out trends and seasonality, and determining how certain events or other circumstances will affect the data. To assess and generate predictions based on time series data, time series models, including autoregressive integrated moving average or exponential smoothing methods, are commonly employed. (Cryer & Chan, 2008)

Referred to as a stochastic process, the set of random variables $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, ...\}$ serves as a depiction of an observed time series. In the context of such a stochastic process, the mean function can be written as:

$$E(Y_t) = \mu_t \text{ for } t = 0, \pm 1, \pm 2, \pm 3, ... \quad (3.3.1.1)$$

Where at each time point, the expected value of the process, denoted as μ_t, may vary.

The autocovariance function; $\gamma_{t,s}$, of the stochastic process; Y_t, is given by:

$$\gamma_{t,s} = Cov\ (Y_t, Y_s) \qquad for\ t, s = 0, \pm 1, \pm 2, \ldots \quad (3.3.1.2)$$

where $Cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E(Y_t Y_s) - \mu_t\, \mu_s$.

The autocorrelation function, $\rho_{t,s}$, of the stochastic process; Y_t, is given by

$$\rho_{t,s} = Corr\ (Y_t, Y_s) \qquad for\ t, s\ =\ 0, \pm 1, \pm 2, \ldots \qquad (3.3.1.3)$$

Where $Corr\ (Y_t, Y_s) = \frac{Cov\ (Y_t, Y_s)}{\sqrt{Var\ (Y_t, Y_t) Var(Y_s, Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}$

The analyses of time series data can be categorized into univariate and multivariate time series. When the analyses of time series data include only one series it is called univariate time series analyses. On the other hand, when time series analyses involve more than time series, it is called multivariate time series analyses. The focus of this study is on the univariate time series analysis. (Cryer & Chan, 2008)

3.3.2 Stationary Test

To derive statistical insights about a stochastic process from an observed dataset of that process, it is usually required to apply specific simplifying, and typically reasonable, assumptions about its properties. Among these, the fundamental assumption that holds the most significance is that of stationarity. The core concept of stationarity posits that the probabilistic rules governing the behavior of a process remain consistent over time. (Cryer & Chan, 2008)

Some time series data may be presented with stationary or weakly stationary conditions while others may be presented with a non-stationary nature which makes the prediction challenging. For stationary time series data, predictions can be simply executed while non-stationary time series data requires some manipulations to the original data such as transforming or differencing to make it stationary. Therefore, stationarity of the time series must be checked before making inferences or forecasting of the time series data. Accordingly, different statistical tests can be performed to examine the stationarity of the data in the time series that are presented below:

First method: Unit root test

Prior to applying a specific model to the time series data, it is essential to examine the stationarity of the series. A time series demonstrates stationarity when both its mean and autocovariance remain consistent over the entire series. This implies that the statistical distribution of any combination of the variables in the time series remains consistent over time. Therefore, the stochastic process, denoted as "y_t" is considered stationary if:

i. $E(y_t) = \mu \text{ , constant for all } t \text{ values}$ (3.3.2.1)

ii. $Cov(y_t, y_{t-i}) = Cov(y_s, y_{s-i}) \text{ for any period } t, s$ (3.3.2.2)

In Equation 3.3.2.1, it implies that y_t maintains a consistent finite mean μ throughout the process, and 3.3.2.2 requiring the autocovariance of the process to be independent of time and dependent solely over time lag (i), the vectors y_t and y_{t-i} are distinct. Consequently, a process is considered stationary when its first and second moments stay consistent throughout time. (Cryer & Chan, 2008)

Examine a simple AR (1) process:

$$y_t = \rho y_{t-1} + e_t \quad (3.3.2.3)$$

In this context, where y_t represents the dependent variable, t denotes the time index, ρ stands for the parameter under estimation, and e_t is considered a white noise. If $|\rho| > 1$, y_t becomes a non-stationary series, leading to an increase in the variance of y_t over time, approaching infinity. Conversely, if $|\rho| < 1$, y_t remains a stationary series. Consequently, Assessing stationarity involves examining in case the absolute value of ρ is consistently less than one through testing. (Cryer & Chan, 2008)

The test hypothesis is:

H_0: Time series is not stationary ($\rho=1$)

H_1: Time series is stationary ($\rho<1$)

Second method: Augmented Dickey-Fuller (ADF) Test

ADF test is one of the most commonly used tests for stationarity of the time series data that has been developed by Dickey and Fuller (1979, 1981).

The conventional Dickey-Fuller test involves estimating equation (3.3.2.3) after subtracting $y_{(t-1)}$ from both sides of the equation. (Cryer & Chan, 2008)

$$\Delta y_t = \beta y_{t-1} + e_t \quad (3.3.2.4)$$

Where $\beta = \rho - 1$, $\Delta y_t = y_t - y_{t-1}$.

$$H_0: \beta = 0$$

$$H_1: \beta < 0 \qquad (3.3.2.5)$$

and assessed utilizing the standard t-ratio for β:

$$t_\beta = \frac{\hat{\beta}}{SE(\hat{\beta})} \qquad (3.3.2.6)$$

In which:

- $\hat{\beta}$ represents the approximation of β

- $SE(\hat{\beta})$ is the standard error of the coefficient.

When considering the null hypothesis of the unit root test, it's important to note that the DF test statistic doesn't adhere to the standard Student's t-distribution; instead, it follows an asymptotic t-distribution.

The earlier explanation of the basic Dickey-Fuller unit root test is suitable exclusively for cases where the series conforms to an AR(1) process. If the series exhibits correlation at lags beyond the first order, it breaches the assumption of disturbances being white noise, e_t. To address this, Introducing a parametric correction for higher-order correlation, the Augmented Dickey-Fuller test assumes that the series follows an AR(p) process. It incorporates lagged difference terms of the dependent variable y on the right-hand side of the test regression.

$$\Delta y_t = \rho y_{t-1} + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \cdots + \beta_p \Delta y_{t-p} + W_t \qquad (3.3.2.7)$$

This extended specification is subsequently employed to test equation (3.3.2.5) using the t-ratio (3.3.2.6). An important finding established by Fuller is that the asymptotic distribution of the t-ratio for ρ remains unaffected by the number of lagged first differences incorporated in the ADF regression. Furthermore, although the assumption that y_t adheres to an autoregressive (AR) process may initially appear restrictive, as per Said and Dickey (1984), they demonstrate that the ADF test retains asymptotic validity even in the presence of a moving average (MA) component, provided that a sufficient number of lagged difference terms are included in the test regression. (Cryer & Chan, 2008)

However, The Augmented Dickey-Fuller test is a statistical test utilized to ascertain the stationarity of a given time series. The ADF test is based on the autoregressive model and helps to identify the presence of a unit root in the time series. A unit root indicates that the series is non-stationary, meaning it has a trend or shows some form of dependence on its past values.

H_0: The time series has a unit root and is non-stationary.

H_1: The time series is stationary.

If the computed t-statistic is lower than the critical values or the p-value exceeds 0.05, then the null hypothesis is not rejected, indicating that the time series possesses a unit root and is non-stationary. On the other hand, if the computed t-statistic exceeds the critical values or the p-value is below 0.05, the null hypothesis is rejected, signifying that the time series is stationary. (Cryer & Chan, 2008)

3.3.3 Seasonal Unit Roots

Similar to the ADF tests, it is crucial to verify that the residuals obtained from estimating the HEGY equation exhibit characteristics of white noise. Therefore, This involves evaluating the suitable lag length for the dependent variable, aiming to confirm the presence of serially uncorrelated residuals. Subsequently, we assess whether deterministic components should be included in the model. (Charemza & Deadman, 1992)

The examination for seasonal integration using the HEGY test entails estimating the subsequent regression (a specialized scenario for quarterly data):

$$\Delta^4 Y_t = \alpha + \beta t + \sum_{j=2}^{4} b_j Q_{jt} + \sum_{j=1}^{4} \pi_i W_{it-1} + \sum_{\ell}^{k} \gamma_\ell \Delta^4 Y_{t-\ell} + a_t \qquad (3.3.3.1)$$

where Q_{jt} is a seasonal dummy, and the W_{it} are given below.

$$W_{1t} = (1 + B)(1 + B^2)Y_t \quad (3.3.3.2)$$

$$W_{2t} = -(1 - B)(1 + B^2)Y_t \quad (3.3.3.3)$$

$$W_{3t} = -(1 - B)(1 + B)Y_t \quad (3.3.3.4)$$

$$W_{4t} = -B(1 + B)(1 + B)Y_t \quad (3.3.3.5)$$

Following ordinary least squares (OLS) estimation, assessments are carried out to evaluate the hypotheses $\pi 1 = 0$, $\pi 2 = 0$, and the combined hypothesis $\pi 3 = \pi 4 = 0$. The HEGY test serves as a combined evaluation for LR (zero frequency) unit roots and seasonal unit roots. In cases where none of the πi values are found to be equal to zero, this indicates that the series exhibits stationarity at both seasonal and nonseasonal frequencies. (Charemza & Deadman, 1992)

3.3.4 Box-Jenkins method

Forecasting is employed as a decision support tool at multiple levels to assist in financial planning, strategic development, and anticipating future circumstances. In its essence, forecasting involves predicting future outcomes by leveraging past and current data, as well as analyzing trends. The most common forecasting technique is the Box-Jenkins method, a method that is based on autoregressive integrated moving average methods, which are mathematical models that capture the temporal dependencies and patterns in time series data.

The Box-Jenkins method consists of several steps:

1. **Identification**: In this step, the underlying properties of the time series data are identified. This entails scrutinizing the autocorrelation and partial autocorrelation plots to ascertain the order of differencing, autoregressive, and moving average components within the ARIMA model.
2. **Estimation**: Once the model order is determined, the model parameters are estimated using maximum likelihood estimation or other suitable methods. This involves fitting

the ARIMA model to the historical data and estimating the coefficients of the autoregressive and moving average terms.

3. **Diagnostic checking**: After parameter estimation, the residuals (the differences between the predicted and actual values) are analyzed to ensure that they satisfy certain assumptions, such as being normally distributed with zero mean and constant variance. Diagnostic tests are performed to detect any remaining patterns or systematic deviations in the residuals.

4. **Forecasting**: Once the model is validated, it can be used to forecast future values of the time series. The model generates point forecasts as well as prediction intervals that provide a range of possible values for the forecasted variable. The method begins by assuming that if the time series is stationary, it can be estimated using an ARMA model, and if it is non-stationary, an ARIMA model can be used to approximate the process that generated it. Figure 3.1 illustrates the process of forecasting using Box-Jenkins model.

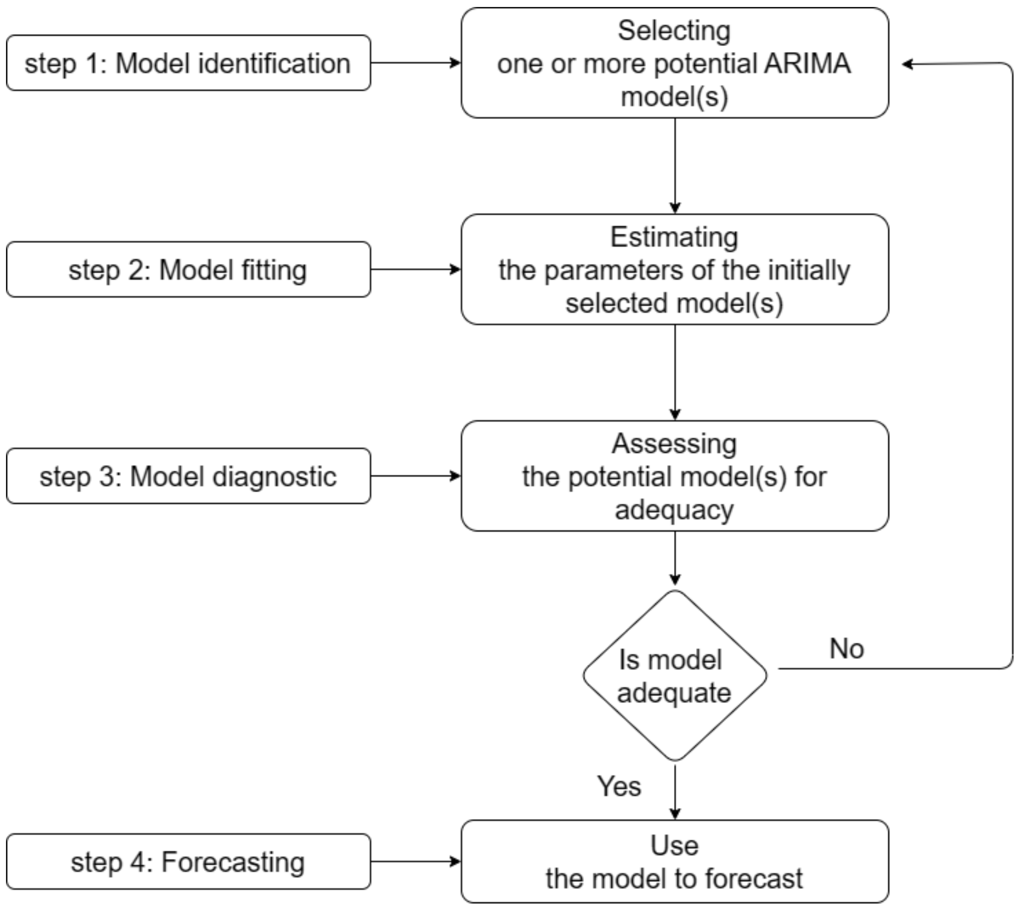

Figure 3.1 Flowchart of Box-Jenkins Method

3.3.5 Autoregressive Moving Average Model (ARMA)

1) Autoregressive Processes

In an AR(p) process, each observation in the time series is expressed as a linear combination of the previous p observations and a random error term. The term "autoregressive" refers to the fact that the current observation depends on its past values.

A p^{th} order regressive process AR(p) y_t is given by:

$$y_t = \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + e_t \qquad (3.3.4.1)$$

where:

- y_t indicates the time series value at time t.
- $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ refer to previous time series values at various lags.
- $\emptyset_1, \emptyset_2, \dots, \emptyset_p$ are the parameters associated with the respective past values.
- e_t is a random error term.

A linear combination of the p most recent past values of itself and an innovation term e_t constitutes the present values of the series y_t, which encapsulates all novel elements in a time series that are not accounted for by the preceding values. (Cryer and Chan, 2008).

A higher order of p implies a longer memory of past values influencing the current observation.

2) Moving Average Processes

We refer to this type of series as an MA(q), which stands for 'moving average of order q'. The term 'moving average' originates from the process of deriving y_t by using the assigned weights $1, -\theta_1, -\theta_2, \dots, -\theta_q$ to the variables $e_t, e_{t-1}, e_{t-2}, \dots, e_{t-q}$ and subsequently adjusting the weights and employing them to $e_{t+1}, e_t, e_{t-1}, \dots, e_{t-q+1}$ to obtain y_{t+1} and so on (Cryer and Chan, 2008).

A q^{th} order Moving average process MA(q) y_t is given by:

$$y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \qquad (3.3.4.2)$$

Where:

- y_t indicates the time series value at time t.
- e_t, e_{t-1}, e_{t-2},…, e_{t-q} are the error terms or shocks at different time lags.
- θ_1, θ_2, ..., θ_q are the parameters associated with the respective error terms.

Estimating the parameters of an MA process typically involves methods like least squares estimation or maximum likelihood estimation. A higher order of q indicates a longer memory of past shocks influencing the current observation.

3) ARMA

The mixed Auto Regressive Moving Average model combines elements of both Auto Regressive and Moving Average components. The model assumes that the series is stationary. (Cryer and Chan, 2008).

In an ARMA(p, q) model, the current observation of the time series is expressed as a linear combination of its past values (AR component) and past error terms (MA component). The AR component models the dependence on previous observations, while the MA component accounts for the influence of past shocks or random errors.

The general notation for an ARMA(p, q) process is:

$$y_t = \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \qquad (3.3.4.3)$$

Where:

- y_t indicates the time series value at time t.
- $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ are the past values of the time series at different lags.
- $\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ are the parameters associated with the respective past values.
- $e_t, e_{t-1}, e_{t-2}, \ldots, e_{t-q}$ are the error terms or shocks at different time lags.
- $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters associated with the respective error terms.

The order of the ARMA model is indicated as (p, q), with p representing the order of the autoregressive component and q representing the order of the moving average component. The choice of p and q depends on the characteristics of the time series data and can be determined through statistical techniques such as information criteria or by analyzing the autocorrelation and partial autocorrelation functions.

4) Autoregressive Integrated Moving Average Model (ARIMA)

The Auto-Regressive Integrated Moving Average Model is used to model non-stationary Time series data that has been transformed to stationary through differencing (Cryer and Chan, 2008).

It is a combination of three components: autoregressive (AR), integrated (I), and moving average (MA).

The general notation for an ARIMA(p, d, q) process is:

$$w_t = \emptyset_1 w_{t-1} + \emptyset_2 w_{t-2} + \cdots + \emptyset_p w_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (3.3.4.4)$$

Where $w_t = y_t - y_{t-d}$

Where *d* is the difference number in the equation.

Definition

A time series y_t is considered to adhere to an ARIMA model if and only if the d^{th} difference is denoted by

$$w_t = \nabla^d y_t \quad (3.3.4.5)$$

is a stationary ARMA model. Thus, if w_t is ARMA(p,q) then y_t is ARIMA (p,d,q).

5) Seasonal Autoregressive Integrated Moving Average Model (SARIMA)

The implication of seasonality in ARIMA models is crucial when analyzing time series data such as the unemployment rate. This integration transforms the ARIMA model into a SARIMA process. SARIMA is a more comprehensive version of ARIMA that accommodates seasonal and non-seasonal patterns observed in the data.

Much like the ARIMA model, SARIMA presupposes that forthcoming forecasted values are outcomes of a combination that is linear, involving past values and preceding errors. It captures the seasonal component, making it particularly useful in situations where seasonality plays a significant role. Sometimes commonly known as the Multiplicative ARIMA model, SARIMA is represented as ARIMA(p, d, q)(P, D, Q)S. The lag-based representation of this model is as follows:

$$\emptyset(L)\varphi(L^S)(1-L)^d(1-L^S)^D y_t = \theta(L)\vartheta(L^S)e_t \tag{3.3.4.6}$$

Using L of order p and q respectively the model includes the following AR and MA characteristic polynomials:

$$\emptyset(L) = 1 - \emptyset_1 L - \emptyset_2 L^2 - \cdots - \emptyset_p L^p \tag{3.3.4.7}$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q \tag{3.3.4.8}$$

Furthermore, polynomial functions with seasonal orders P and Q, as illustrated below:

$$\varphi(L^S) = 1 - \varphi_1 L^S - \varphi_2 L^{2S} - \cdots - \varphi_P L^{(P)S} \tag{3.3.4.9}$$

$$\vartheta(L^S) = 1 - \vartheta_1 L^S - \vartheta_2 L^{2S} - \cdots - \vartheta_Q L^{(Q)S} \tag{3.3.4.10}$$

Where:

- y_t indicates the time series value at time t.
- e_t error terms characterized by white noise.
- p, d, q - the order of non-seasonal AR, differencing, and non-seasonal MA respectively.
- P,D,Q - the order of seasonal AR, differencing, and seasonal MA respectively.
- S-seasonal order, in this case S=4 for quarterly data.
- L – lag operator $L^k y_t = y_{t-k}$

3.3.6 SARIMA Order Selection

Plotting the ACF and PACF is the method used to determine the lags for the seasonal and non-seasonal autoregressive and moving average components (p, P, q, and Q). These plots expose the internal correlations among time series observations at various time intervals, Providing information on both seasonal and non-seasonal lags. The plot of both ACF and PACF exhibit peaks and truncations at certain lags, where “k” represents the non-seasonal lags and “k_s” denotes the seasonal lags. The model's order is determined by the count of significant peaks observed in these plots. (Lo Duca, 2021)

3.3.7 Choosing the best SARIMA model

When examining the autocorrelation and partial autocorrelation function plots, it becomes evident that there may be various SARIMA(p, d, q)(P, D, Q) models featuring distinct significant lags for p, P, q, and Q. Consequently, the selection of the SARIMA model with the perfect lag lengths for both seasonal and non-seasonal components necessitate the application of specific criteria. In this study, the Akaike Information Criterion (AIC) and the Schwartz and Bayes Information Criterion (BIC) were utilized for this purpose. The calculation of these criteria is as follows:

$$AIC = -2log\, R + 2k \qquad (3.3.4.11)$$

$$BIC = -logR + klogR \qquad (3.3.4.12)$$

In this context, the total number of parameters in the model is represented by k = p + P + q + Q, and R signifies the likelihood function. To Identify the best model, the SARIMA(p, d, q)(P, D, Q) is characterized by low AIC and BIC values.

3.3.8 Model checking

Model diagnostics play a crucial role in time series analysis and are typically carried out through residual analysis. The examination of residuals from a fitted model is of particular importance. These residuals should exhibit the characteristics of white noise, which include normal distribution with a mean of zero, constant variance, and the absence of autocorrelation issues.

3.3.9 Residual Analysis

In cases where residuals do not conform to the characteristics of white noise, several issues can arise. The estimated parameter variances can experience bias and inconsistency, leading to tests becoming unreliable during model estimation. Additionally, Predictions produced by these models may exhibit inefficiency, primarily because of the elevated variance in forecast errors. Therefore, it is crucial to prioritize performing residual analysis before utilizing models for specific purposes. (Cryer & Chan, 2008)

3.3.10 Residual autocorrelation test

Autocorrelation analysis of residuals is conducted using both graphical methods and statistical tests. One approach to examining the autocorrelation structure of residuals is to create plots of their autocorrelation and partial autocorrelation. The plots assist in revealing the presence of autocorrelation within the residuals, indicating that there might be unaccounted information in the model. Another approach involves plotting residuals against their respective time-lags. Plot e_t on x-axis and e_{t-1} on y-axis. Namely, graphing the subsequent observations (e_1, e_2), (e_2, e_3), … ,(e_n, e_{n+1}).

When the points in a plot of autocorrelation and partial autocorrelation of residuals predominantly fall in quadrants one and three, it suggests the presence of positive autocorrelation. Conversely, if most of the points are concentrated in quadrants two and four, it indicates negative autocorrelation. On the other hand, when the points are uniformly scattered across all quadrants, it suggests that the residuals are random in nature.

3.3.11 Univariate Residuals autocorrelation test (portmanteau test)

This test is employed to examine the autocorrelation pattern in the residuals. The hypothesis of the test is:

H_0: The residuals are not serially correlated

H_1: at least one successive residual is serially correlated

To examine this null hypothesis, Box and Pierce (1970) introduced the Q-statistics.

$$Q = n \sum_{m=1}^{M} r_m^2 \qquad (3.3.4.13)$$

In which:

-m is the lag length,

-n is the number of observations,

-and r_m is the autocorrelation function of the residuals series at lag m.

Under the null hypothesis, Q is asymptotically distributed χ^2_{m-p-q}.

In a finite sample, accurate approximation of the Q-statistic may not be achievable with... χ^2_{m-p-q}. The adjusted Q-statistic proposed by Ljung and Box is:

$$\tilde{Q} = n(n+2)\sum_{m=1}^{M} r_m^2\,(m-k)^{-1} \qquad (3.3.4.14)$$

where M = time lag and r_m = the accumulated sample autocorrelations (Box et al., 2015).

Therefore, if the model is accurate then, $\tilde{Q} \sim \chi^2_{m-p-q}$. The test selection involves rejecting the null hypothesis at the significance level of α, if $\tilde{Q} > \chi^2_{m-p-q}(1-\alpha)$. Indicating the presence of autocorrelation in residuals, thus violating the assumption.

3.3.12 Residuals normality test

The standard test for checking the normality of residuals is the Shapiro-Wilk test. Shapiro-Wilk normality hypothesis test:

H_O: the data are normally distributed

H_1: the data are not normally distributed

The test statistic is:

$$Z = \frac{\left(\sum_{t=1}^{n} a_t y_t\right)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} \qquad (3.3.4.15)$$

Where:

- y_t represents the value of the time series at time t.
- $\bar{y}$ represents the sample mean.

The coefficients a_t are given by

$$(a_1, a_2, \ldots, a_n) = \frac{M^T V^{-1}}{C} \qquad (3.3.4.16)$$

where *C* is a vector norm

$$C = \|V^{-1}m\| = (m^T V^{-1} V^{-1} m)^{\frac{1}{2}} \qquad (3.3.4.17)$$

and the vector m

$$m = (m_1, m_2, \dots, m_n)^T \qquad (3.3.4.18)$$

The composition includes expected values of order statistics obtained from independently and identically distributed random variables sampled from the standard normal distribution. Ultimately, the covariance matrix of the normal order statistics is represented by V.

Thus, If the p-value is below the selected alpha level, the null hypothesis is rejected, indicating evidence that the tested data is not normally distributed.

3.4 Neural Networks

Time series data can exhibit characteristics that are not consistently stationary, linear, or normal. When attempting to make the data stationary, it is possible to lose important information. However, non-parametric methods like Neural Networks and Spectral Analysis do not rely on these assumptions. Artificial Neural Networks (ANNs) are forecasting techniques that utilize simplified mathematical models inspired by the brain. ANNs enable the exploration of complex non-linear relationships between the response variable and its predictors, offering several advantages over traditional Time series models, including the ability to handle missing values. (Ciaburro & Venkateswaran, 2017)

3.4.1 Neural Network Architecture

A neural network can be conceptualized as a network composed of interconnected "neurons" organized in layers. The lowermost layer consists of the predictors or inputs, while the uppermost layer represents the forecasts or outputs. Additional layers, known as hidden layers, may exist in between and contain hidden neurons. The most basic neural networks lack hidden layers and essentially function as linear regressions. Figure 3.2 illustrates the neural network depiction of a linear regression with five predictors. The predictors are associated with coefficients known as "weights". The predicts are derived by combining the inputs in a linear manner. Within the neural network framework, the weights are determined using a "learning algorithm" that minimizes a "cost function," such as the Mean Squared

Error (MSE). However, in this straightforward example, linear regression can be employed as a more efficient method for training the model. (Ciaburro & Venkateswaran, 2017)

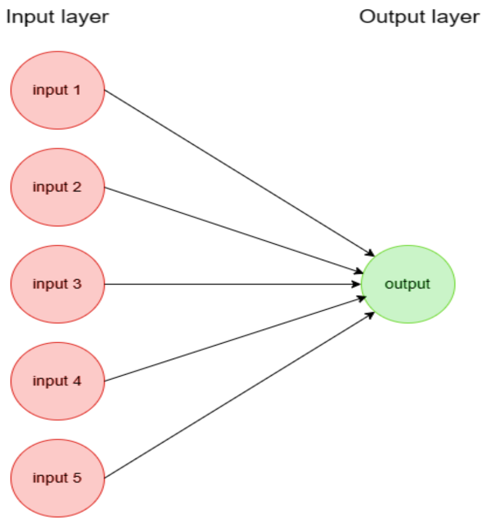

Figure 3.2 Neural network depiction of a linear regression with five predictors

When a hidden layer with intermediate neurons is introduced, the neural network transitions into a non-linear model. Figure 3.3 provides a simple illustration of this concept.

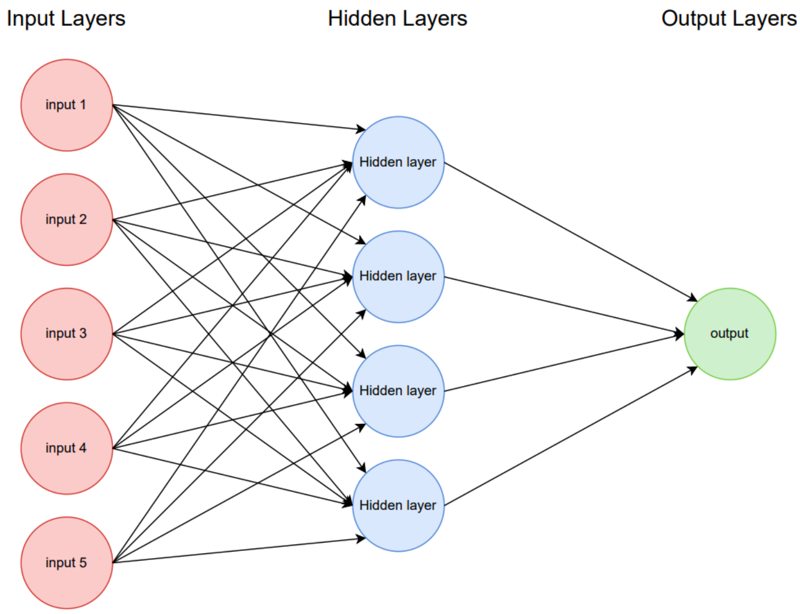

Figure 3.3 Neural network transitions into a non-linear model

In a multilayer feed-forward network, information flows through each layer of nodes, with each layer receiving inputs from the preceding layers. The outputs produced by the nodes in one layer act as inputs for the next layer. To calculate the inputs for each node, a weighted linear combination of the inputs is created. This combined result is subsequently subjected to a nonlinear function, altering it before being outputted.

The relationship between the output y_t with the input $(y_{t-1}, y_{t-2}, \dots, y_{t-p})$ can be represented as:

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{j=1}^{p} \beta_{ij} y_{t-i}\right) + e_t \qquad (3.4.1.1)$$

Where:

$$\alpha_j \text{ as } (\mathrm{j} = 0,1,2, \dots, \mathrm{q}) \ and\ \beta_{ij}\ as\ (i = 0,1,2, \dots p; j = 0,1,2, \dots q)$$

Are the model parameters which are weighted linear combinations called connection weights. p is the number of input nodes and q as the number of hidden nodes.

The hidden layer g is modified using a non-linear function.

$$g(x) = \frac{1}{1+e^{-x}} \qquad (3.4.1.2)$$

Which reduce the influence of extreme input values, this mechanism enhances the network's resilience to outliers. Therefore model 3.4.1.1 maps nonlinear past values $(y_{t-1}, y_{t-2}, \dots, y_{t-p})$ to future value y_t:

$$y_t = f\left(y_{t-1}, y_{t-2}, \dots, y_{t-p}, w\right) + e_t \qquad (3.4.1.3)$$

with “w” as a vector of all parameters and f is the function of the network and connection weights. This equation illustrates that the neural network functions as a nonlinear autoregressive model. The lagged values of the time series are the inputs in a neural network.

3.4.2 Neural network autoregression

In the context of time series data, the past values of the time series can be utilized as inputs to a neural network, similar to how lagged values are employed in a linear autoregression model. This model is known as a neural network autoregression or NNAR model. The notation NNAR(p, k) signifies a Neural Network Autoregression model with p-lagged inputs and k nodes in the hidden layer. (Ciaburro & Venkateswaran, 2017)

Figure 3.4 provides a simple illustration of this concept.

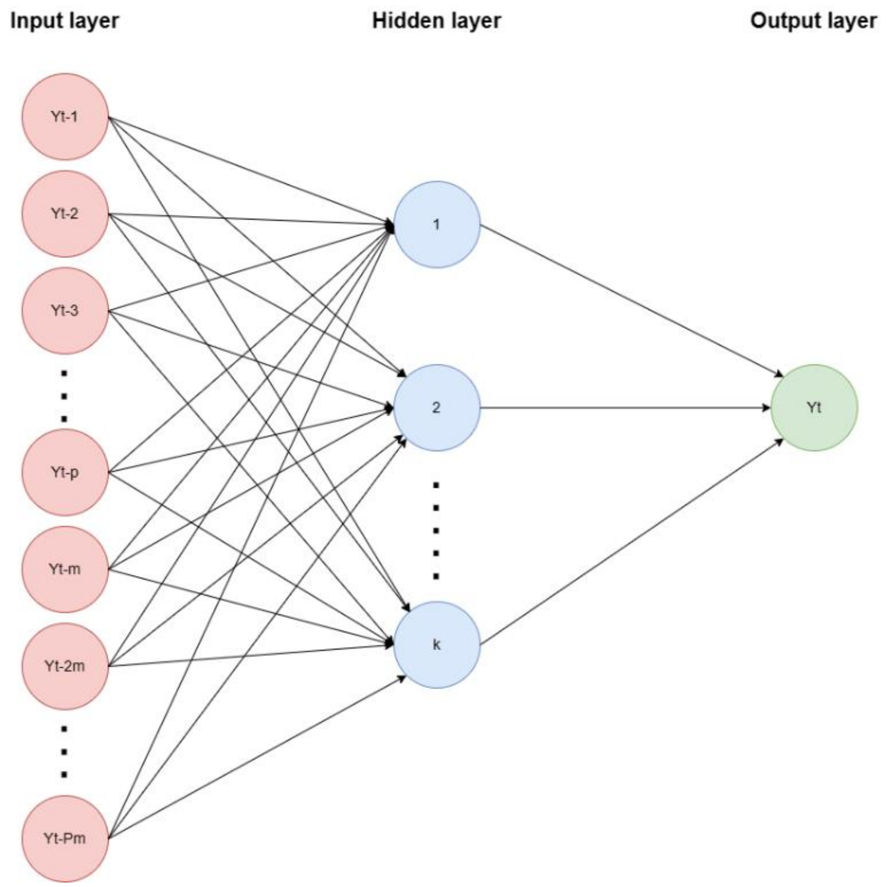

Figure 3.4 A diagram represent the NNAR (p,P,k)m model

In the case of a seasonal NNAR, denoted as NNAR $(p, P, k)_m$, P represents the number of seasonal lags. Selecting the appropriate configuration for the hidden layer, which depends on the data, is crucial to prevent overfitting. The choice of p helps determine the nonlinear autocorrelation structure of the series.

The estimation of parameters involves selecting those that minimize the accuracy metrics, and the model's performance is assessed using out-of-sample predictions. The neural network makes forecasts iteratively, one step at a time. Therefore, when making multiple-step forecasts, the initial forecast is used as the subsequent input in combination with the historical data until all forecasts have been generated. Figure 3.5 illustrates the process of forecasting using NNAR model.

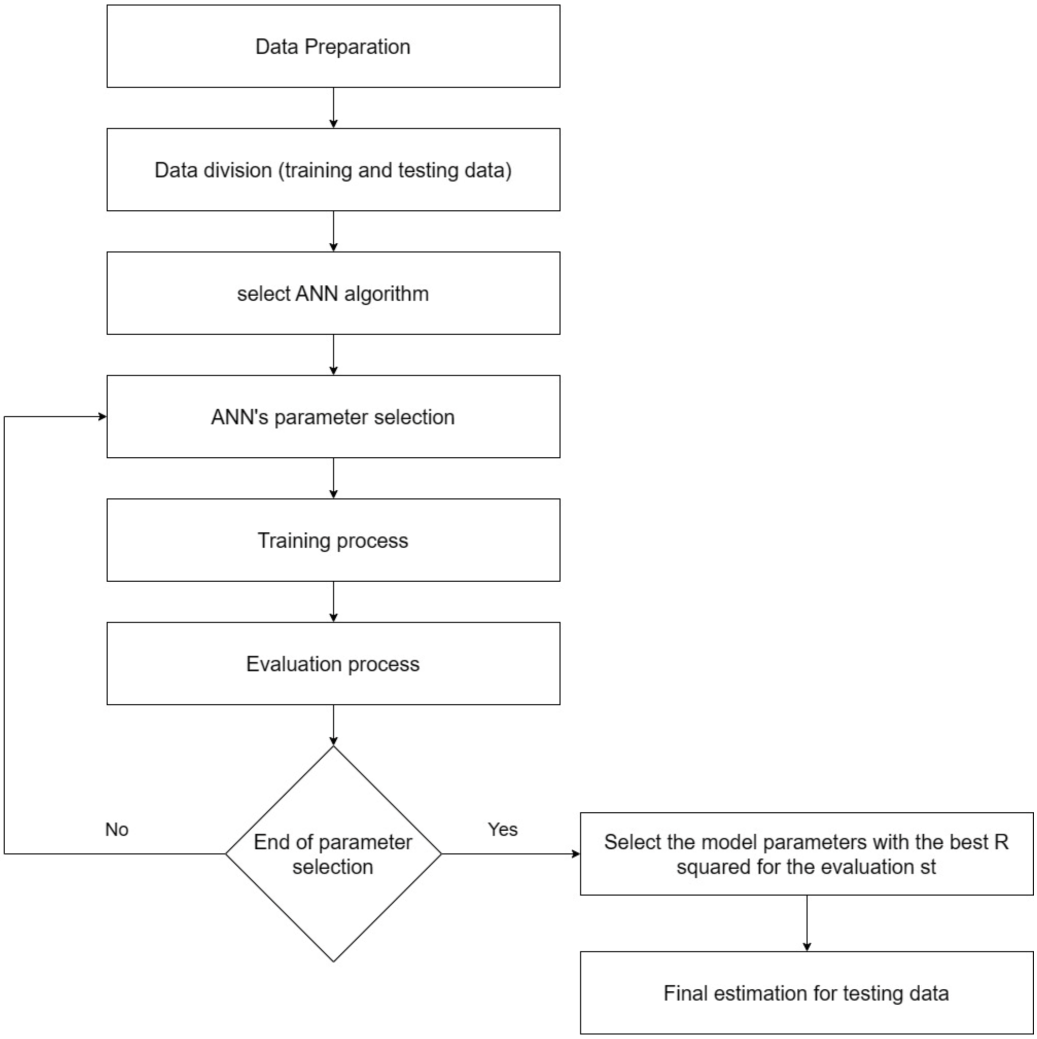

Figure 3.5 illustrates the proposed NNAR model of this study.

3.5 Forecasting accuracy

Forecasting accuracy serves as a benchmark for evaluating the performance of forecasting models. It assesses the effectiveness of a forecasting model in predicting future values. To gauge the forecast ability of each model, the study compares their performance using error statistics. Specifically, three error statistics are employed: Root Mean Squared Error, Mean Absolute Error, and Mean Absolute Percent Error. A smaller value of these error statistics indicates better forecast performance for the model. (Kozuch, Cywicka, & Adamowicz, 2023)

After conducting the analysis on unemployment rate. The validation set technique, which involves splitting the data into the training and test set, is used. Where the training set was used to obtain the model parameters, and the test set was used to confirm the accuracy level of the data set. The next step is to predict future values.

Suppose $y_{t+h/t}$ represents the forecast h steps ahead of y_{t+h}, the associated forecast error can be defined as $e_{t+h/t} = y_{t+h} - y_{t+h/t}$. Subsequently, the evaluation statistics for forecasting, relying on N h-step ahead predictions, can be formulated as follows:

$$Root\ Mean\ Square\ Error\ (RMSE) = \sqrt{\frac{1}{N}\sum_{j=t+1}^{t+N}(e_{j+h/j})^2} \quad (3.5.1)$$

$$Mean\ Absolute\ Error\ (MAE) = \frac{1}{N}\sum_{j=t+1}^{t+N}\left|e_{j+h/j}\right| \quad (3.5.2)$$

$$Mean\ Absolute\ Percent\ Error\ (MAPE) = \frac{1}{N}\sum_{j=t+1}^{t+N}\left|\frac{e_{j+h/j}}{y_{j+h}}\right| \quad (3.5.3)$$

Applied to gauge the disparity between the forecasted and observed values in the series, the Root Mean Square Error serves as an indicator of predictive accuracy. The MAE calculates the average absolute errors to assess the proximity of predicted values to the actual values. It provides a measure of how closely the predicted values align with the true values. Additionally, the MAPE quantifies accuracy by expressing the errors as a percentage of the values being measured. Hence, a superior forecasting capability of the model is indicated by the presence of smaller RMSE, MAE, and MAPE values. (Kozuch, Cywicka, & Adamowicz, 2023)

3.6 Comparison between Time Series and NNAR Models

The following table provides a summary of the strengths and weaknesses of both the ARIMA model and the ANN model.

Table 3.1 Comparison of models

Model	strengths	weaknesses
ARIMA	The model is characterized by its simplicity. Moreover, it possesses the capability to dynamically determine model parameters and demonstrates a high computational speed.	The model lacks the ability to capture the underlying relationships between variables or analyze the interplay between factors. Additionally, its suitability is limited to short-term predictions.
ANN	The model exhibits a fast calculation speed and excellent capability for nonlinear fitting. Importantly, it eliminates the need for establishing a mathematical model.	The model lacks the capability to express and analyze the relationship between the input and output of the predicted system.

3.7 Comparing forecasting predictive performance (accuracy)

To compare forecasting performance accuracy Diebold Mariano test was used, also known as the DM test, which is a statistical test that is used to compare the accuracy between two forecasting models.

DM test is used to choose the most accurate model, suppose that we have two models, model 1 and model 2, if the errors in model 1 are smaller than model 2 we can infer that model 1 is more accurate. The test uses the mean squared error (MSE) to measure the forecast errors, smallest value of MSE is better for forecast.

The Diebold Mariano hypothesis is:

H_0: Both forecasts have the same accuracy

H_1: The forecasts do not have the same accuracy

If the p-value is less than 0.05 we reject the null hypothesis that both forecast models have the same accuracy, we conclude that one forecast is significantly more accurate than the other. (Diebold , 2015)

3.8 Chapter summary

In this chapter, we begin by exploring the data description employed in this study. Then, we discuss an overview of time series models and explain the significance of concepts such as stationarity. Furthermore, this chapter offers a comprehensive review of ARIMA and NNAR models. Lastly, we discuss the comparative analysis of forecasting predictive performance through the Diebold Mariano test.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

This chapter pertains to the statistical modeling of the unemployment rate in Palestine. The Second Section describes the current data. The Third Section shows the results of time series models including ARIMA and SARIMA models. The Fourth Section shows the results of NNAR models. The Fifth Section compares the forecasting accuracy of the unemployment rate in Palestine obtained from ARIMA and NNAR models. The Sixth Section provides the forecasted unemployment rate in Palestine for the next eight quarters (i.e., 2023Q3 to 2025Q2). The Seventh Section discusses the results of this study and compares them with the current literature. The Last Section summarizes the chapter.

4.2 Data Description

Figure 4.1 displays the time series plot depicting the unemployment rate in Palestine, which exhibited a general rising trend throughout the study duration. Table 4.1 shows the descriptive statistics of the unemployment rate in Palestine over the study period from 1996Q1 to 2023Q2. The findings indicated that the average of the unemployment rate was 23.9% (SD = 5.0%). The highest unemployment rate (35.6%) was encountered in the third quarter of 2002, which might be attributed to the second Intifada whereby Israel has imposed movement restrictions. However, the lowest unemployment rate (8.8%) was in the second quarter of 2000.

Furthermore, to detect an outlying case, the current study computed the interval: [Q1-1.5IQR, Q3+1.5IQR] whereby any case that lies outside this interval is considered an outlier. The data on unemployment rate over the study period exhibited the presence of outliers because some cases lie outside the interval: [15.73, 33.13].

Table 4.1 Descriptive Statistics of the Unemployment Rate in Palestine (1996Q1 – 2023Q2)

Variable	Mean	SD	Median	Min	Max	Range	[Q1-1.5IQR, Q3+1.5IQR]
Unemployment rate, %	23.9	5.0	24.7	8.8	35.6	26.8	[15.73, 33.13]

SD: Standard deviation; IQR: Interquartile range

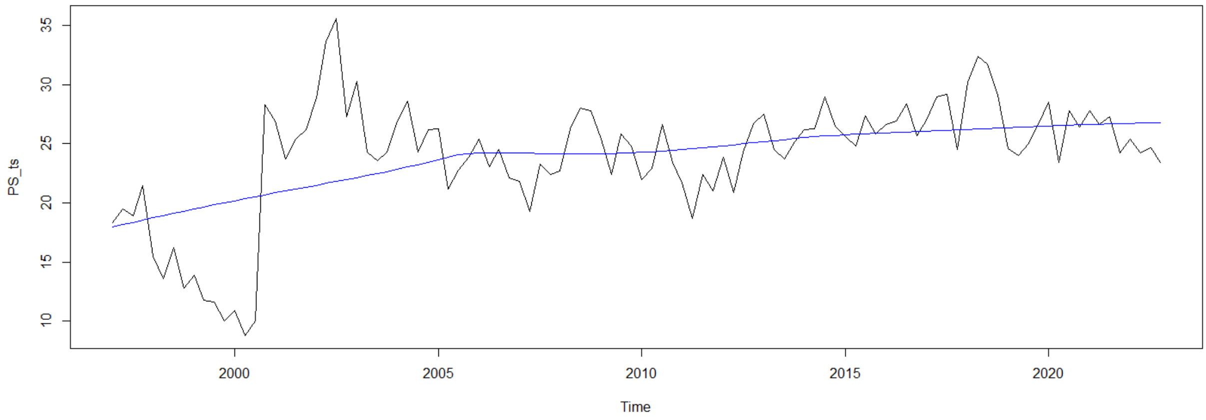

Figure 4.1 Plot of Unemployment Rate in Palestine (1996-2023; quarterly).

Figure 4.2 illustrates that the Palestinian unemployment rate displayed seasonal variations between 1996 and 2023, with peaks occurring during the third quarter of the year. The chart illustrates the quarterly unemployment rate's evolution, emphasizing a noticeable recurring pattern in the data, a pattern that was further corroborated through the autocorrelation plot.(Figure 4.3).

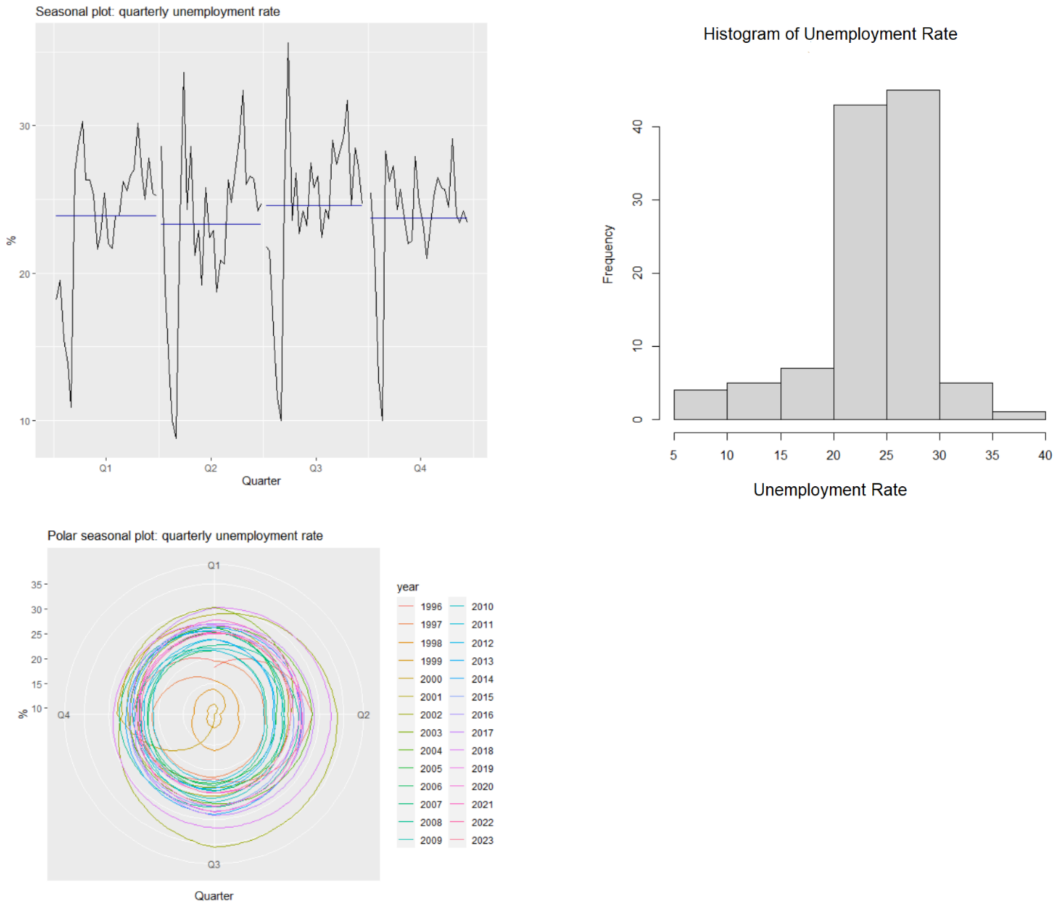

Figure 4.2 Distribution and Seasonal Trends in the Quarterly Unemployment Rate in Palestine (1996Q1 – 2023Q2).

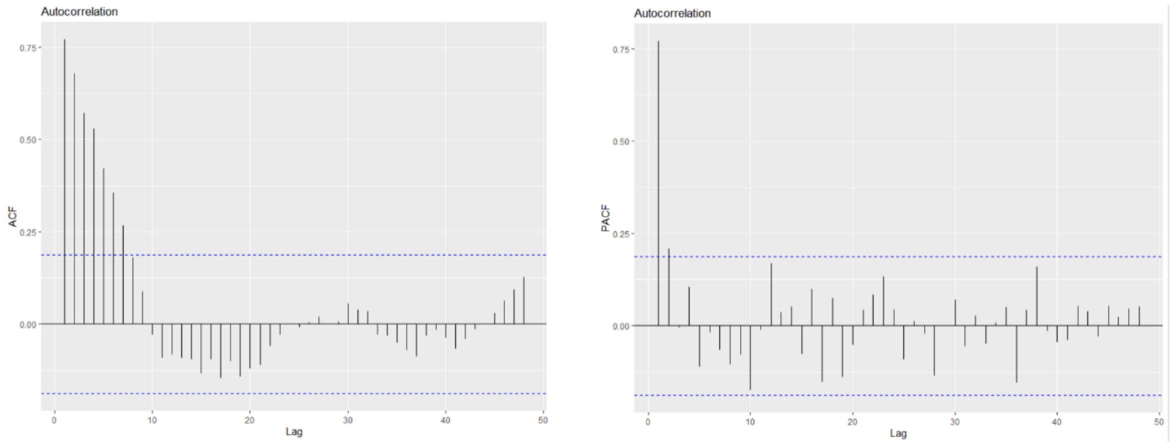

Figure 4.3. ACF and PACF Plots of Quarterly Unemployment Rate in Palestine (1996Q1 – 2023Q2).

4.3 SARIMA Model

For SARIMA model, In the current study, the data was divided into training and testing datasets. The training dataset covers the period from 1996Q1 to 2017Q4 while the testing dataset covers the remaining period. Table 4.2 and Figure 4.4 illustrate the descriptive statistics of the unemployment rate in Palestine for the training dataset. The mean unemployment rate in the training dataset was 23.2% (SD = 5.2%) with a minimum rate of 8.8% and a maximum rate of 35.6%. The training series displayed a significant seasonal pattern throughout the period from 1996Q1 to 2017Q4 as indicated in Figure 4.4.

Table 4.2 The descriptive Statistics of the Training Dataset

Training dataset	**Mean**	**SD**	**Median**	**Min**	**Max**	**Skewness**	**Kurtosis**
Unemployment rate, %	23.2	5.2	24.3	8.8	35.6	- 0.95	1.02
SD: Standard Deviation							

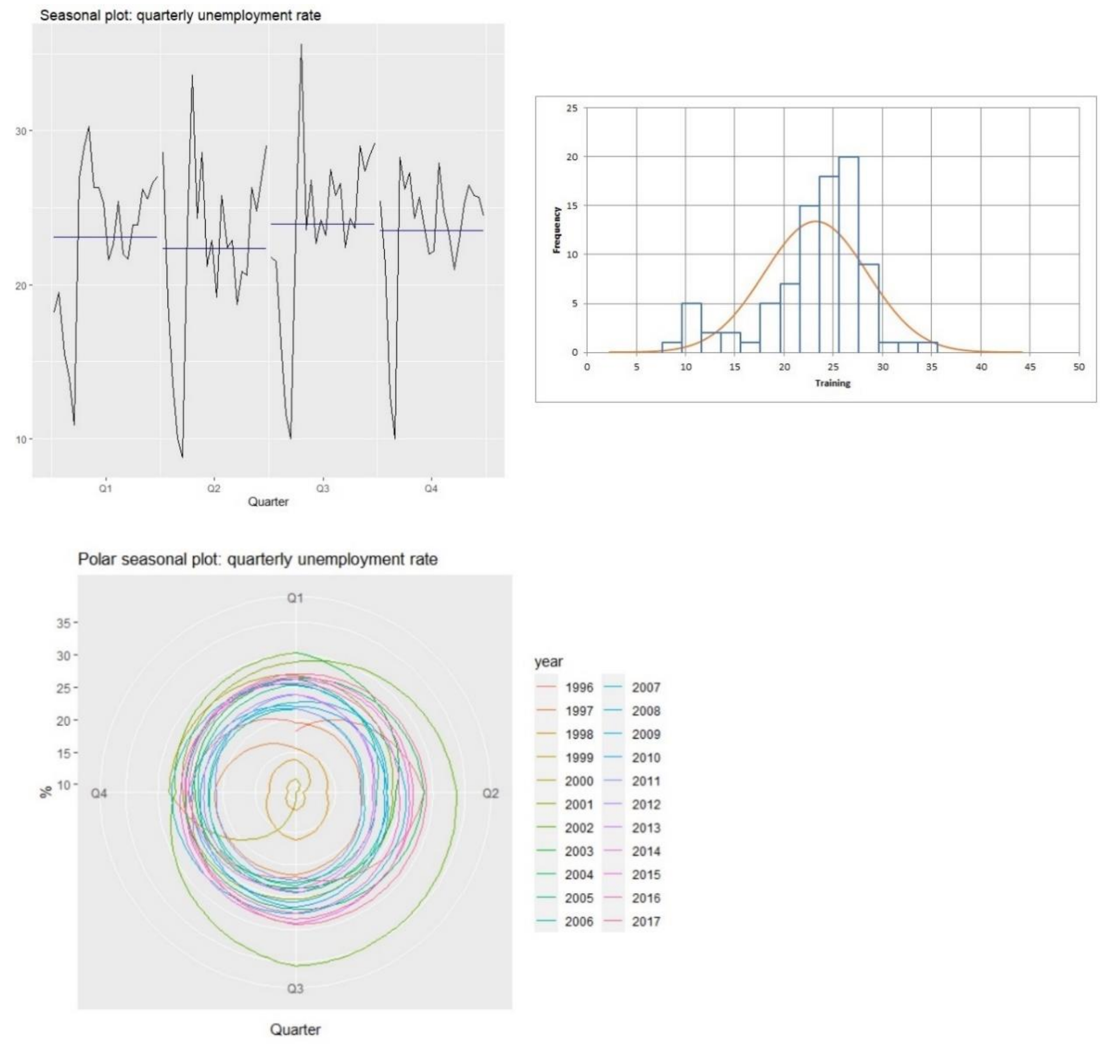

Figure 4.4 Seasonal Plots and Histogram of the Training Dataset (Palestine)

4.3.1 Stationary Test

To identify an appropriate time series model, it is crucial to investigate stationarity through the Augmented Dickey-Fuller (ADF) and Philips-Perron (PP) tests. After visually inspecting the ACF and PACF for training dataset of quarterly Palestinian unemployment rate, it became apparent that the autocorrelation coefficients exhibited a gradual decrease. This suggests the presence of a non-stationary and rather consistent seasonal pattern within the time series as indicated in Figure 4.5.

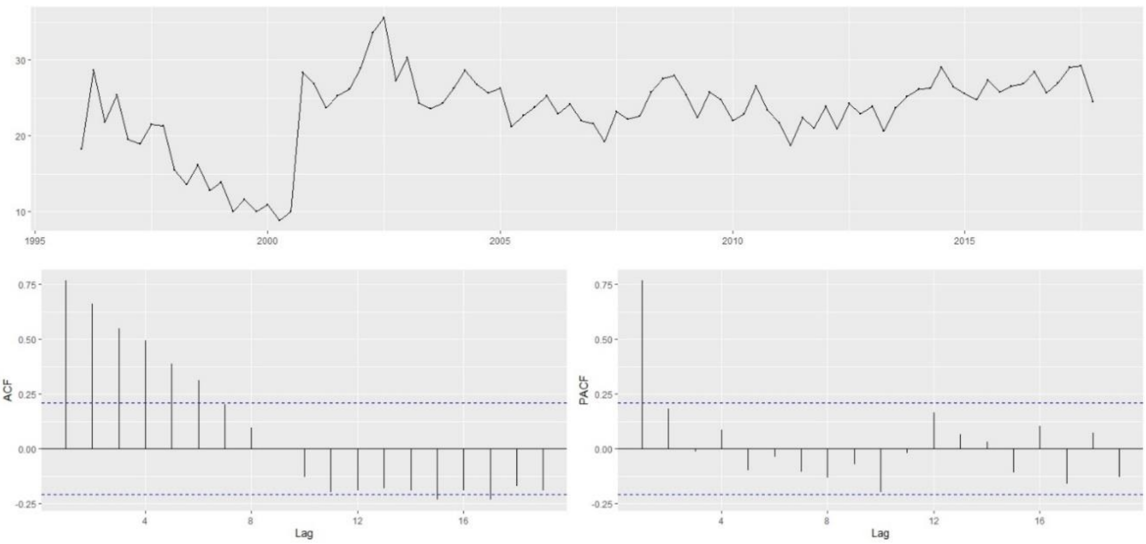

Figure 4.5 Time Series, Autocorrelations, and Partial Autocorrelations Plots of the Training dataset (Palestine)

The first difference plot for the time series distinctly indicates that the unemployment rate of the first difference displays a stationary mean time series. As a result, the original quarterly unemployment rate is deemed non-stationary. As an alternative approach, we examined the existence of unit roots by initially conducting the ADF and PP tests on the series in its original form and then on the series in its first differences. The empirical findings pertaining to the unemployment rate are presented. Table 4.5, demonstrating that the unemployment rate series becomes stationary when considering first differences, signifying it is integrated of order 1.

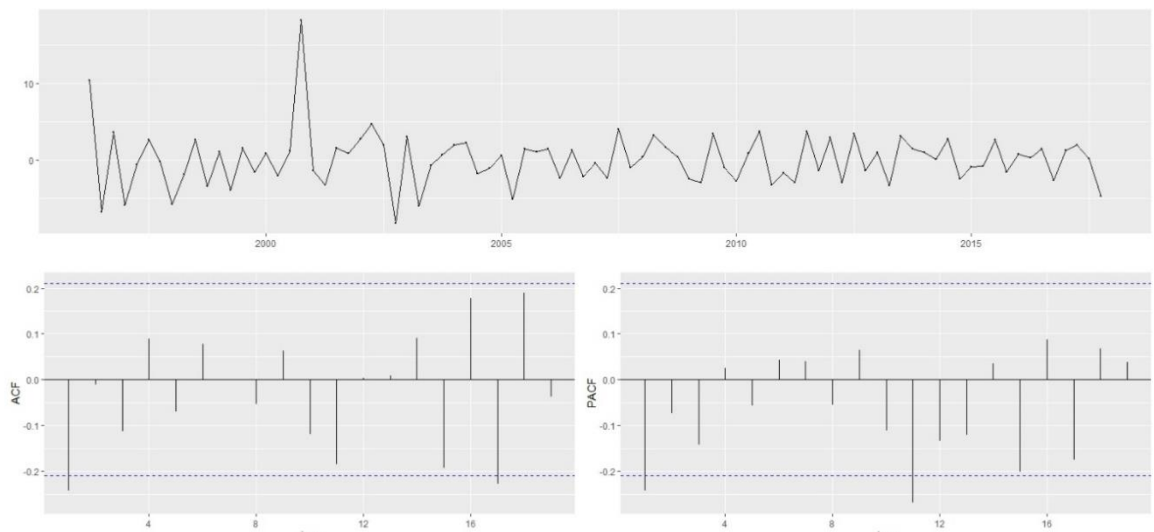

Figure 4.6 Time Series, Autocorrelations, and Partial Autocorrelations Plots of the First Difference of Unemployment Rate for Training Dataset (Palestine)

Table 4.3 Unit Root Test of the Training Dataset for the Unemployment Rate

			Unemployment Rate
ADF	Level	Test statistic	-2.697
		p-value	0.289
	1st difference	Test statistic	-4.531***
		p-value	0.01
PP	Level	Test statistic	-23.103**
		p-value	0.027
	1st difference	Test statistic	-102.47***
		p-value	< 0.001

ADF: Augmented Dickey-Fuller; PP: Phillips-Perron
*** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

To validate the results of unit root analysis, the current analysis involved examining the potential presence of structural breaks in the unemployment rate in Palestine of the training dataset. Consequently, Zivot-Andrews test was applied and the empirical findings indicate the test statistic for both scenarios (i.e., trend and trend and intercept) not exceeded the critical values of the test in absolute terms (-3.96 for trend and -4.73 for both intercept and trend). Therefore, this indicates that there is not enough evidence to reject both null hypotheses that unemployment rate has structural break in trend and in both trend and intercept over the entire time period (i.e., the mean is not constant over the entire time series) as displayed in Table 4.4. This implies that unemployment rate maintains a unit root with a structural break in the trend, as well as in both the intercept and trend components. Therefore, the original unemployment rate time series for training dataset in Palestine exhibits a non-stationary process and it is integrated of order 1; I(1).

Table 4.4 Zivot-Andrews Test for Structural Breaks in the Unemployment Rate.

	Trend(a)	Both(b)
Minimum t-stat (p-value)	- 3.96 (0.409)	- 4.73 (0.083)
Critical values		
1%	- 4.93	-5.57
5%	- 4.42	-5.08
10%	- 4.11	-4.82
Potential brake point	2012Q2	2009Q3

a: accommodating for breaks in trend.
b: accommodating for breaks in both trend & intercept.

On the other hand, considering the observed seasonal pattern in the unemployment series during the training dataset for Palestinian unemployment rate, this study chooses a seasonal ARIMA model (SARIMA). Therefore, it is crucial to determine whether seasonal differencing is required. To explore the existence of stochastic seasonality within the data, the Hylleberg, Engle, Granger and Yoo Test (HEGY) for seasonal unit roots test was utilized. The empirical results of the HEGY test indicate the acceptance of the presence of only non-seasonal unit root while seasonal unit root is not presented in the training dataset as indicated in Table 4.5. Hence, there is no need for seasonal differencing and thus it can be concluded that the original unemployment rate in the training dataset is non-stationary and lacks stochastic seasonality. Therefore, having an integration order of 1; denoted as I (1).

Hence, it can be concluded that the initial unemployment rate is a non-stationary series without seasonal unit root. However, the graphical presentation of the unemployment rate suggests that there is seasonal peaks or anomalies in the data and thus HEGY test might be affected. In this study, the Palestinian unemployment rate exhibited dramatic increase in two periods related to the political situation and Intifadah (1996Q2 and 2002Q3). Therefore, the author decided to take the first seasonal difference and compares the results with not integrated SARIMA model.

Table 4.5 Results from the HEGY Test for a Seasonal Unit Root in the Unemployment Rate within the training dataset.

Null Hypothesis	Unemployment Rate
Non-seasonal unit root (zero frequency)	0.218
Seasonal unit root (1 quarter per cycle)	< 0.001
Seasonal unit root (2 quarters per cycle)	< 0.001
Seasonal unit root (3 quarters per cycle)	< 0.001
Seasonal unit root (4 quarters per cycle)	< 0.001

Note: The HEGY test was conducted with the inclusion of intercept, trend, and seasonal dummies. A maximum of eight lags was considered, following the Schwarz criterion, and 1000 simulations were carried out.
Deterministic terms: constant + trend + seasonal dummies
Lag selection criterion and order: fixed, 0;
P-values: based on response surface regressions

4.3.2 Model identification

Model identification involves establishing the parameter values p, q, d, P, Q, and D for the first and seasonal differences of the unemployment rate in Palestine. The ACF plot signifies the existence of a non-seasonal MA component of order 1, denoted as MA(1). On the other hand, the PACF plot indicates the significance of seasonal lags at 4, 5, and 12, capturing potential components of order 1 for seasonal autoregressive. However, since the autocorrelation at the seasonal lags 4, 8 is negative, it suggests the possibility of modeling combinations of both seasonal and non-seasonal autoregressive parts or it may lack to non-seasonal component only (see Figure 4.6). Therefore, the AR order might be 1 or 2. However, when considering first seasonal difference, the order of auto-regressive process might be 0 or 1 as shown in Figure 4.7.

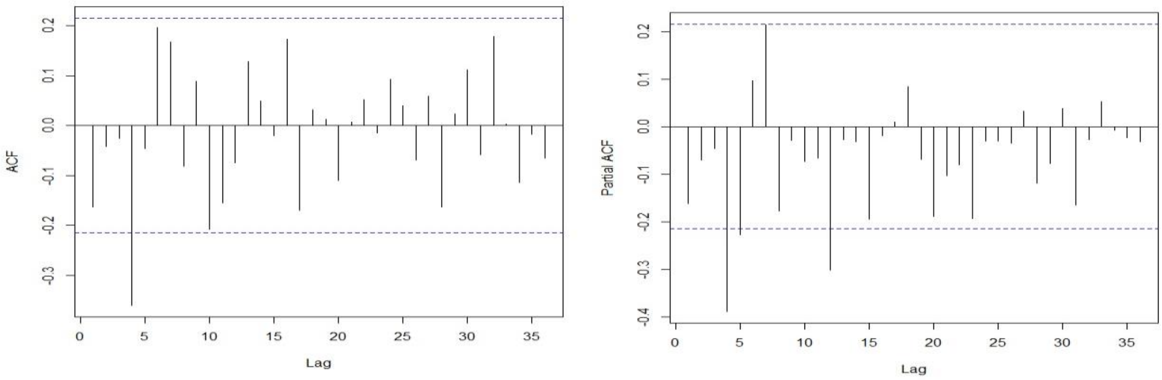

Figure 4.7. Chart displaying the Autocorrelation and Partial Correlation for the first and Seasonal Differenced Unemployment Rate in Palestine.

The EACF table, there is an obvious triangular region of zeros shown in the sample EACF that will indicate a quite clear model with a specific p=1 and q=0 that would be appropriate for the time series data.

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	o	o	o	o	o	o	o
1	o	o	o	o	o	o	o	o	o	o	x	o	o	o
2	x	o	o	o	o	o	o	o	o	o	x	o	o	o
3	x	o	o	o	o	o	o	o	o	o	x	o	o	o
4	x	o	x	o	o	o	o	o	o	o	o	o	o	o
5	x	o	o	o	o	o	o	o	o	o	o	o	o	o
6	x	o	o	x	x	o	o	o	o	o	o	o	o	o
7	x	x	x	x	x	o	o	o	o	o	o	o	o	o

Given the graphical diagnostics of the differenced stationary series of the unemployment rate in Palestine, the current study also compared various SARIMA models to choose the most suitable models through comparing goodness of fit measures including AIC, AICc, and BIC whereby the most optimal model was identified based on the lowest goodness of fit measures. The findings suggest that the best model is ARIMA(1,1,1)(0,0,1)[4] because it has the lowest values of AIC, BIC, and AICc as displayed in Table 4.6. Nonetheless, when considering the first seasonal difference, the findings suggested that taking the first seasonal difference provides lower model fit indexes in terms of AIC, AICc, and BIC and the best model was found to be ARIMA(0,1,1)(0,1,1)[4] as it exhibits the lowest model fit indexes as compared to seasonal ARIMA models and as exhibited in Figure 4.7.

Table 4.6 Model Fit Indices of Suggested SARIMA Models

Model	AIC	AICc	BIC
ARIMA(2,1,2)(1,0,1)[4]	469.66	471.07	486.92
ARIMA(1,1,0)(1,0,0)[4]	466.60	466.89	473.99
ARIMA(1,1,1)(1,0,0)[4]	464.65	465.14	474.52
ARIMA(1,1,1)(0,0,1)[4]	**464.49**	**464.98**	**474.36**
ARIMA(1,1,1)(1,0,1)[4]	466.49	467.23	478.82
ARIMA(0,1,0)(0,0,1)[4]	470.05	470.19	474.98
ARIMA(0,1,0)(1,0,1)[4]	471.98	472.27	479.38
ARIMA(0,1,2)(1,0,1)[4]	469.36	470.10	481.69
ARIMA(0,1,2)(0,0,1)[4]	467.38	467.87	477.24
ARIMA(0,1,1)(1,0,1)[4]	467.67	468.16	477.54
ARIMA(2,1,1)(0,0,1)[4]	469.74	470.48	482.07
ARIMA(2,1,1)(1,0,1)[4]	467.16	468.21	481.95
ARIMA(2,1,1)(2,0,1)[4]	469.12	470.54	486.39
ARIMA(0,1,1)(0,1,1)[4]	**456.47**	**456.78**	**463.73**
ARIMA(1,1,1)(1,1,1)[4]	460.06	460.84	472.15
ARIMA(1,1,1)(0,1,1)[4]	458.31	458.82	467.99

AIC: Akaike Information Criteria;
AICc: Akaike Information Criteria correction for small sample size;
BIC: Bayesian Information Criteria.

4.3.3 Model Fitting and Diagnostics

Given that the best model has been identified, the next step is to estimate the model parameters using the maximum likelihood estimation method. Table 4.7 Presents the results of the calculated SARIMA(0,1,1)(0,1,1)[4] and SARIMA(1,1,1)(0,0,1)[4] and it was evident that all parameters of the model exhibit statistical significance at a 5% level of significance.

Table 4.7 Parameters Estimate for both SARIMA Models.

	Estimate	**S.E**	**Critical value**	**P-value**
Parameters Estimate of the ARIMA(1,1,1)(0,0,1)[4] Model				
AR1	0.719	0.095	7.541	4.681×10^{-14} ***
MA1	-0.969	0.046	-21.080	$< 2.2 \times 10^{-16}$ ***
SMA1	0.197	0.115	1.715	0.086*
Parameters Estimates of the ARIMA(0,1,1)(0,1,1)[4] Model				
MA1	-0.280	0.117	-2.389	0.017**
SMA1	-0.939	0.133	-7.053	1.748×10^{-12} ***

S.E: Standard errors.
*** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

In addition to conventional tests such as t-test to evaluate statistical significance of model parameters and F-test to evaluate model's overall validity, the process of selecting the best model also involves evaluating the performance of residuals. To accomplish this, an analysis of the residual series was carried out to ensure that it exhibits behavior resembling a white noise process, which is a core assumption for many time series statistical models.

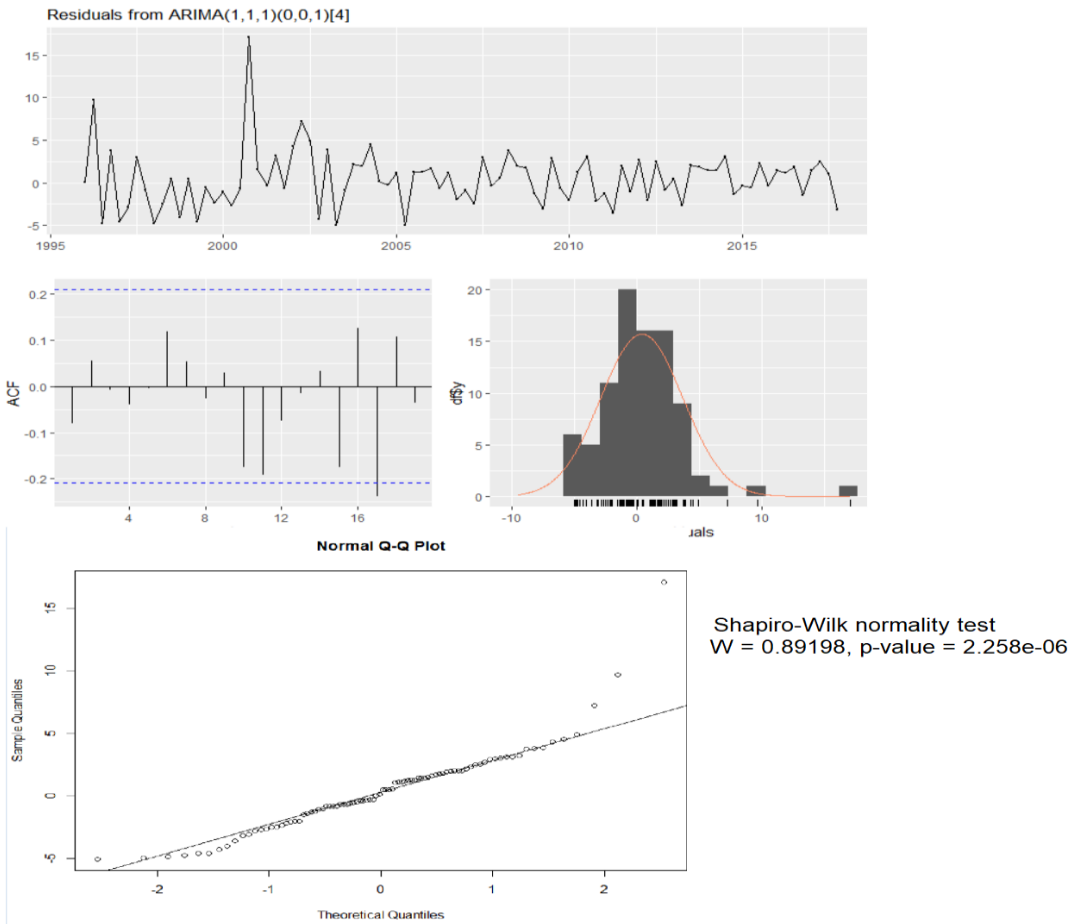

Figure 4.8 Residuals Diagnostic Plots for ARIMA(1,1,1)(0,0,1)[4]

The results of the Ljung-Box test show that the p-values connected to the test statistic surpassed the 5% significance level for all lag orders, as displayed in the table 4.8. This indicates that no statistically significant autocorrelation has been identified in the residuals. (see Figure 4.8 and Figure 4.9).

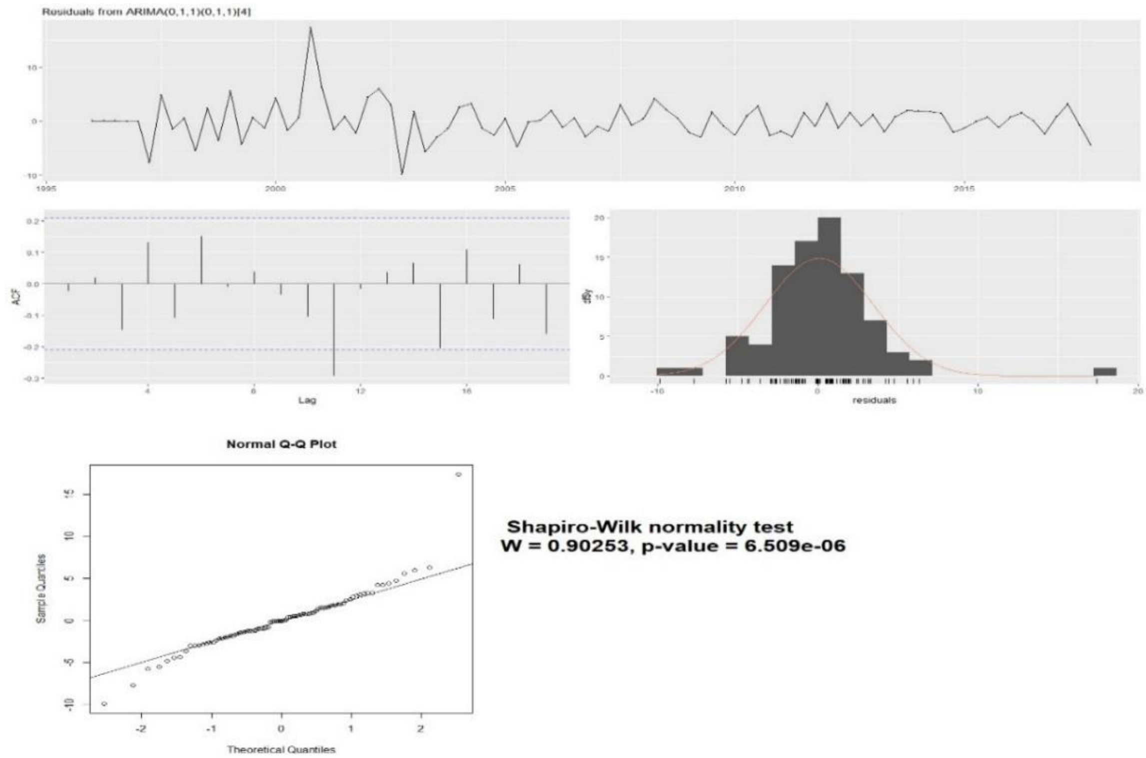

Figure 4.9 Residuals Diagnostic Plots for ARIMA(0,1,1)(0,1,1)[4]

To examine the presence of the Autoregressive Conditional Heteroscedasticity in the residuals, the ARCH-LM test was conducted and displayed in Table 4.8. The results provide evidence on the absence of ARCH in the residuals at 5% level of significance. However, it exhibits significant heteroscedasticity in residuals at 10% level significance. As for normality of the residuals, the findings indicate that the residuals are not normally distributed which might affect the accuracy of the forecasting results. It is important to note that all suggested models examined in Table 4.6 exhibited non-normal residuals making it difficult to rely on the forecasting results obtained from classical seasonal ARIMA models. Nonetheless, ARIMA(0,1,1)(0,1,1)[4] exhibited the most reliable model for forecasting since it does not form an ARCH model. This is probably due to the presence of outliers in the current data that might affect the ARIMA model's performance.

Table 4.8. Residuals Diagnostics using Ljung-Box and ARCH-LM tests for ARIMA(0,1,1)(0,1,1)[4] Model.

Lag	Ljung-Box	p-value	ARCH-LM Test	p-value
4	0.972	0.914	51.424	0.052
8	2.532	0.960	34.482	0.067
12	9.060	0.698	22.350	0.099
16	13.248	0.655	7.160	0.993

Table 4.9 illustrates the accuracy of the forecasting unemployment rate in Palestine in terms of RMSE, MAE, MAPE, and MASE. The findings show that accuracy measures obtained from ARIMA(0,1,1)(0,1,1)[4] are lowest than those obtained from ARIMA(1,1,1)(0,0,1)[4] model. Figure 4.10 shows that unemployment rate exhibited a slight stable pattern without seasonality over the forecasting period (i.e., 2018Q1 – 2023Q2), which did not match the actual unemployment rates. On the other hand, Figure 4.11 shows the forecasted unemployment rate using ARIMA(0,1,1)(0,1,1)[4] model and exhibits a slight decreasing trend with a clear pattern of seasonality over the forecasting period (i.e., 2018Q1 – 2023Q2) and fluctuating around the rate of 24.9 to 27.5%, which is best match the actual unemployment rate as compared to ARIMA(1,1,1)(0,0,1)[4] model.

Table 4.9. Forecasting Performance of the Estimated SARIMA Models

	ARIMA(0,1,1)(0,1,1)[4]		ARIMA(1,1,1)(0,0,1)[4]	
Accuracy Measure	**Training dataset**	**Testing dataset**	**Training dataset**	**Testing dataset**
RMSE	3.371	2.787	3.296	3.015
MAE	2.322	2.236	2.377	2.365
MAPE	10.804	8.237	10.907	8.582
MASE	0.654	0.629	0.669	0.640

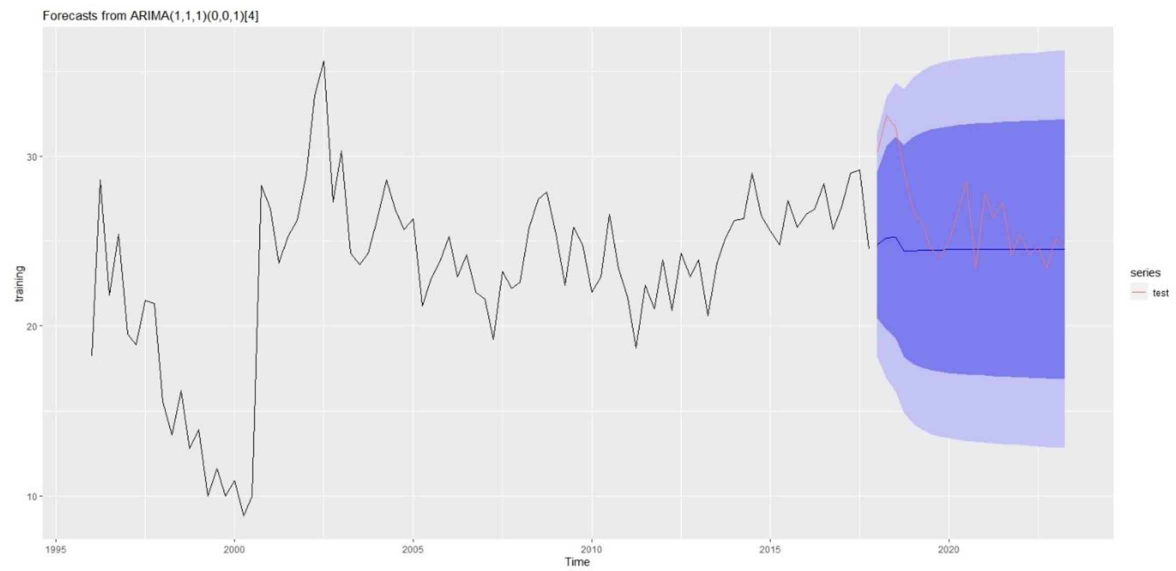

Figure 4.10 Forecasts of unemployment rate based on the results of ARIMA(1,1,1)(0,0,1)[4]

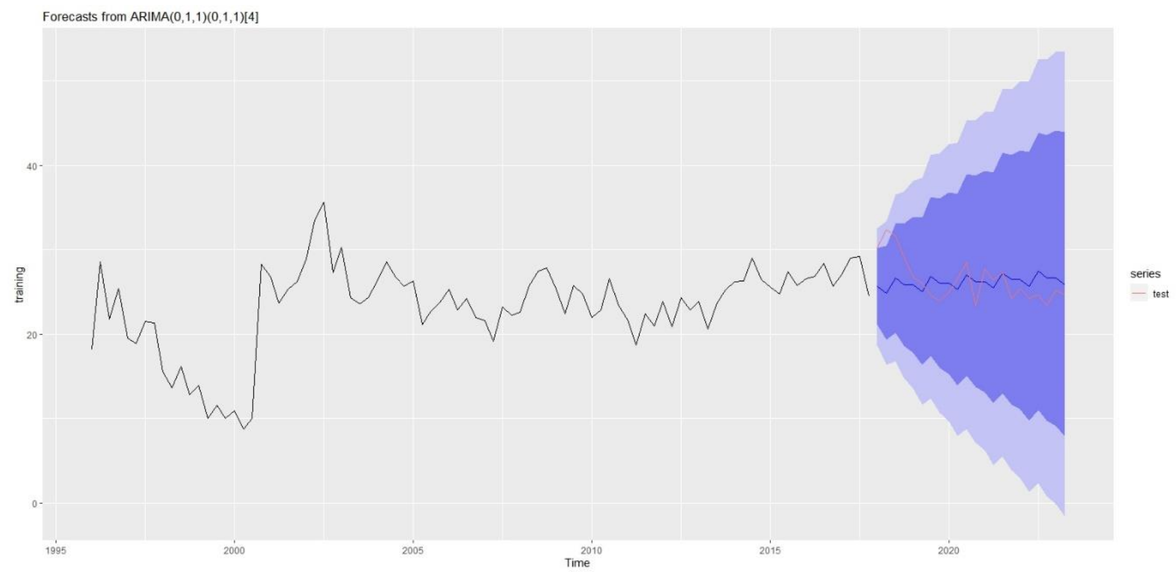

Figure 4.11 Forecasts of unemployment rate based on the results of ARIMA(0,1,1)(0,1,1)[4]

4.4 NNAR Model

For fitting the NNAR model, the data were divided into two parts, training dataset covering the period 1996Q1 to 2017Q4 and testing dataset covering the period from 2018Q1 to 2023Q2. This division is based on the recommendation from the literature that training data set must comprise of 80% of the whole dataset (Joseph, 2022).

Given that the graphical representation of the ACF and PACF of the original series of unemployment rate in Palestine exhibited a slight autoregressive pattern of order 1; AR(1) as well as the number of seasonal lags was set at a value of 1 since there is a seasonality in the data, the current study modeled the unemployment rate in Palestine using NNAR(1,1,k)[4]. The NNAR model has been estimated to identify the optimal value of k, whereby the in-sample (i.e., training) and out-sample (i.e., testing) RMSE, MAE, MASE, and MAPE and the findings are displayed in Table 4.10. The findings suggest that the optimal value of k is ten nodes in the hidden layer using both training(in-sample) and testing (out-of-sample) datasets. Hence, this study reveals that an NNAR(1,1,10)[4] model is the optimal model for predicting the unemployment rate in Palestine. Figure 4.12 shows the prediction of the unemployment rate in Palestine, which exhibits a slight decreasing trend with values fluctuating around 24.7 to 26.7% and exhibiting strong pattern of seasonality.

Table 4.10. Forecasting Performance of the NNAR(1,1,10)[4].

Accuracy Measure	Training dataset	Testing dataset
RMSE	2.208	2.492
MAE	1.568	1.872
MAPE	7.577	6.675
MASE	0.442	0.527

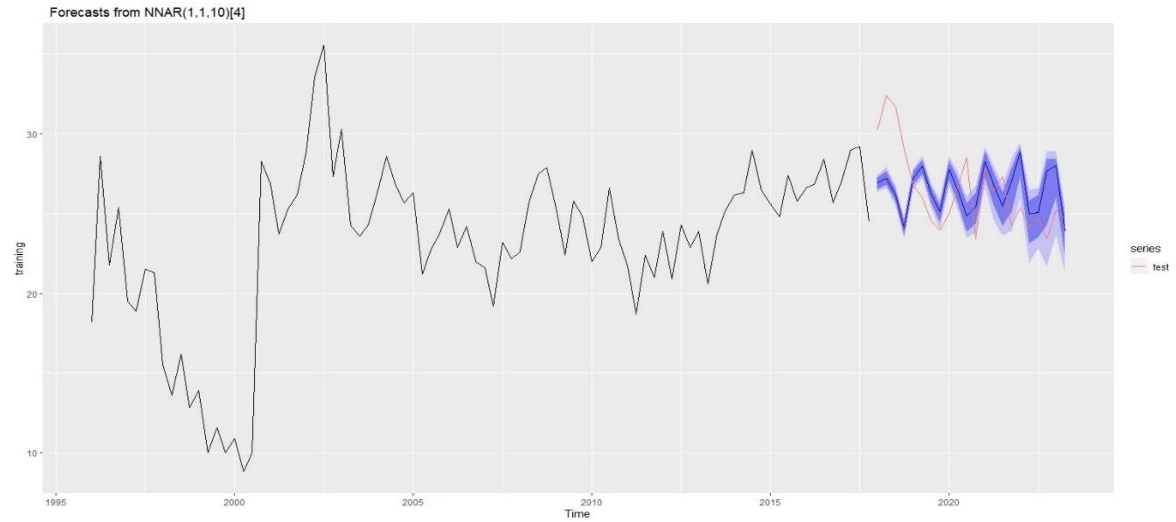

Figure 4.12 Predictions for the unemployment rate derived from the outcomes of the NNAR(1,1,10)[4] model.

4.4.1 Model Diagnostics

To ensure the accuracy of the estimated NNAR model, the residuals of this model were tested for various assumptions including serial autocorrelation and normality. The results of the Ljung-Box test indicate that the p-values associated with the test statistic exceed the 5% significance level for all lag orders as detailed in Table 4.11. This suggests that there is no statistically significant autocorrelation observed in the residuals. (see Figure 4.13). Furthermore, Figure 4.13 shows that the residuals follow an approximate normal distribution and as it evident by Shapiro-Wilk test whereby the p-value is 0.6612, which is larger than 5%.

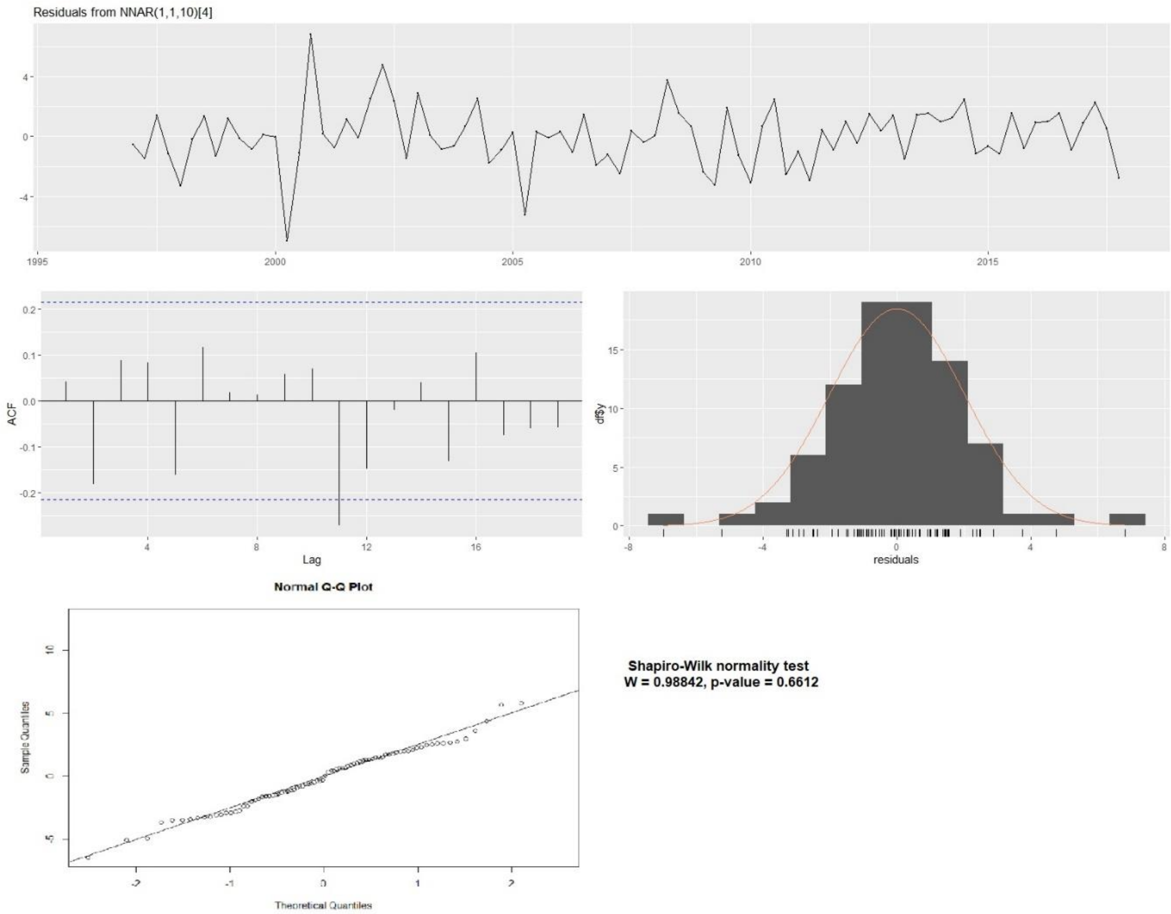

Figure 4.13 Residuals Diagnostic Plots for NNAR(1,1,10)[4]

Table 4.11. Residuals Diagnostics using Ljung-Box test.

Lag	Ljung-Box	p-value
4	7.244	0.124
8	9.225	0.324
12	9.956	0.620
16	10.358	0.847

4.5 Comparing Forecasting Performance for SARIMA and NNAR models

As it can be seen from Table 4.12, the measures of accuracy obtained from NNAR model for both training and testing datasets were lower than those obtained from SARIMA model. However, the focus is particularly should be placed on the MAPE and MASE values. These measures are not influenced by the scale of the data and serve as suitable measures when comparing forecast accuracy across series with varying scales. They are particularly useful when dealing with the testing sample data that does not match the length of the training sample data.

Table 4.12. Comparing Forecasting Performance for SARIMA and NNAR models

	ARIMA(0,1,1)(0,1,1)[4]		**ARIMA(1,1,1)(0,0,1)[4]**		**NNAR(1,1,10)[4].**	
Accuracy Measure	**Training dataset**	**Testing dataset**	**Training dataset**	**Testing dataset**	**Training dataset**	**Testing dataset**
RMSE	3.371	2.787	3.296	3.015	2.208	2.492
MAE	2.322	2.236	2.377	2.365	1.568	1.872
MAPE	10.804	8.237	10.907	8.582	7.577	6.675
MASE	0.654	0.629	0.669	0.640	0.442	0.527

Figure 4.14 displays the values of the unemployment rate in Palestine for the testing dataset, predicted values obtained from SARIMA and NNAR models. The figure illustrates that the forecasted Palestinian unemployment rate values obtained from the NNAR model are more closely align with those in the testing dataset, whereas the forecasted values from the SARIMA model show a less close alignment.

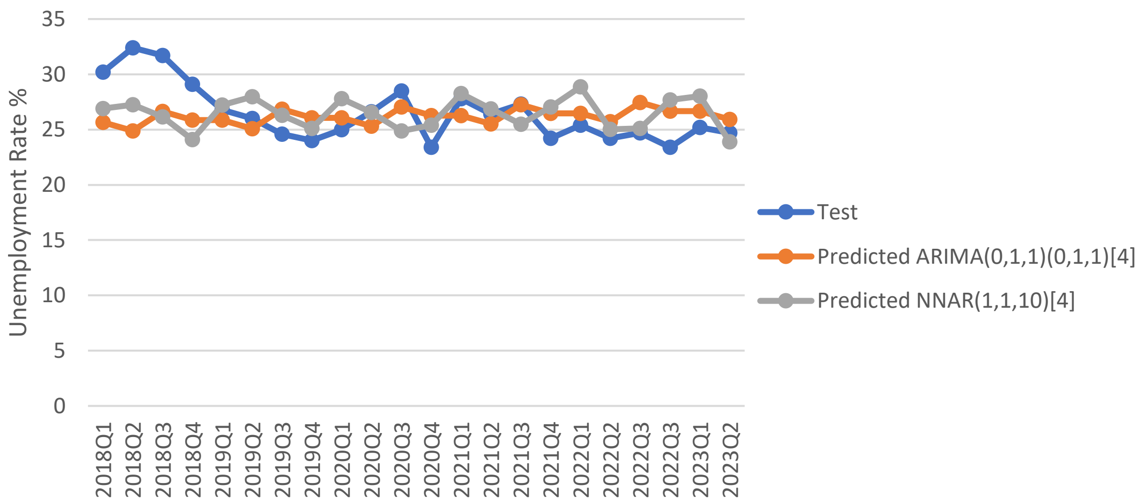

Figure 4.14. Palestinian Unemployment Rate for Testing Dataset and Forecasted Values from SARIMA and NNAR.

This study also performed Diebold-Mariano (DM) test to assess and compare predictive accuracy obtained from both models, namely, SARIMA and NNAR models. The empirical results of this test reveal that NNAR(1,1,10)[4] model performed better than ARIMA(0,1,1)(0,1,1)[4] in terms of forecasting accuracy and performance for both in-sample dataset (p-value = 0.011) and out-of-sample dataset (p-value = 0.010) as displayed in Table 4.12. Therefore, we reject the null hypothesis of this study that indicates that SARIMA and NNAR models exhibit comparable forecasting performance to forecast unemployment rate in Palestine. Hence, it can be concluded that NNAR(1,1,10)[4] model has better forecasting performance than ARIMA(0,1,1)(0,1,1)[4] model to forecast unemployment rate in Palestine.

Table 4.13. Results of Diebold-Mariano Test for Predictive Accuracy

	DM Test Statistic	**P-value**
Training dataset	2.356	0.011
Testing dataset	2.357	0.010
DM: Diebold-Mariano		

4.6 Forecasting Unemployment Rate in Palestine (2023Q3 – 2025Q2)

Table 4.14 illustrates the forecasted unemployment rate in Palestine for the next eight quarters (i.e., 2023Q3 to 2025Q2). Based on the emerged results of the present study about the best models using NNAR(1,1,10)[4] model, the predicted Palestinian unemployment rate in the third quarter of 2023 is 24.1% compared to 24.7% in the second quarter of 2023, while for the overall forecasting period it is expected to range from 23.2 to 27.5% indicating evidence for seasonality. For SARIMA model, on the other hand, the predicted value of the Palestinian unemployment rate for the third quarter of 2023 is expected to increase to 27.7% as compared with the second quarter of 2023 and for the overall forecasting period it expected to fluctuate between 26.3 to 27.8%. This suggests that the Palestinian unemployment rate is expected to remain chronic problem and encounter an oscillating pattern using both NNAR and SARIMA models.

In the meantime, combining both results obtained from these two methods using a hybrid approach could be useful as an alternative approach to improve forecasting accuracy and performance of the Palestinian unemployment rate.

Table 4.14. Forecasted Unemployment Rate in Palestine (2023Q3 – 2025Q2).

Year (Quarter)	ARIMA(0,1,1)(0,1,1)[4]	NNAR(1,1,10)[4]
2023 (Q3)	27.7	24.1
2023 (Q4)	26.9	25.3
2024 (Q1)	26.7	26.6
2024 (Q2)	26.1	25.9
2024 (Q3)	27.8	27.5
2024 (Q4)	27.1	26.0
2025 (Q1)	27.1	25.1
2025 (Q2)	26.3	23.2

4.7 Discussion

Various studies have been conducted to forecast the unemployment rate in several countries worldwide, given its importance as a key macroeconomic indicator for each country. Forecasting such a vital indicator necessitates the use of accurate statistical methods to assist decision-makers in formulating relevant policies aimed at addressing this economic challenge. Accordingly, numerous studies in the existing literature have focused on forecasting the unemployment rate through classical time series modeling techniques, such as ARIMA and Holt-Winter methods (Dritsaki, 2016; Dritsaki & Klazoglou, 2018; Mladenovic, 2017; Nor et al., 2018), machine learning and artificial neural networks (ANN) (Gogas et al., 2022; Husin et al., 2023), and hybrid methods (Ahmad et al., 2021; Chakraborty et al., 2021; Katris, 2019). These studies represent the various approaches employed to tackle the significant challenge of forecasting unemployment rates. Each approach has its unique strengths and applicability within different economic contexts, timeframes, and geographical variations.

The key findings of this study highlight that univariate forecasting of the Palestinian unemployment rate using NNAR methods outperforms the results obtained from the SARIMA models. This is evident from the DM test and the accuracy measures for training sample and testing sample forecasts. These findings are in line with similar studies, such as

Davidescu et al. (2021), who found that forecasting Romania's unemployment rate using NNAR was superior to the SARIMA model, particularly for out-of-sample forecasts. However, it is worth noting that there are variations in results based on different economic contexts. For instance, a study conducted in Turkey compared the forecasting performance of unemployment rates using SARIMA and NNAR models during the economic downturn triggered by the COVID-19 pandemic. The findings indicated that the perfect model for predicting the monthly unemployment rate in Turkey varied based on the crisis's impact. Before the crisis, ARMA (2,1) performed better, while during the crisis, NNAR was more effective (Yamacli & Yamacli, 2023). These results are not consistent with the findings of our study. Additionally, Chakraborty et al. (2020) suggested that employing a hybrid approach for forecasting the unemployment rate provided superior results compared to traditional ARIMA models, which differs from the findings of this study. These variations in results indicate that the choice of forecasting method can be influenced by specific economic conditions and the presence of external factors like economic crises.

In the existing literature, It is evident that there is no universally acknowledged superior method for predicting the unemployment rate, and diverse sample periods necessitate distinct forecasting approaches. (Chakraborty et al., 2020; Katris, 2019). For instance, Chakraborty et al. (2020) pointed out that hybrid approaches, which combine ARIMA and ANN models, can enhance the forecasting accuracy of unemployment rates in seven developed countries. Katris (2019) emphasized that there is no universally recognized single model, and the choice of approach depends on factors such as the forecasting horizon and geographic location. He demonstrated that for 1-step ahead forecasts, fractional autoregressive integrated moving average (FARIMA) models showed the best forecasting performance, whereas for longer periods ($h \geq 12$), neural network approaches yielded results comparable to FARIMA-based models. Moreover, when $h = 3$, the Holt-Winters model proved to be more suitable.

In this study, it was evident that the ARIMA model exhibited non-normal errors and showed a borderline significance of heteroscedasticity. These findings might explain our results that indicate that the NNAR model outperform the SARIMA model, as indicated in Katris (2019). Given the asymmetries in the Palestinian unemployment rate stemming from its complex political and economic context, non-linear methods have proven to be more effective than traditional linear time series models. Ahmad et al. (2017) have also suggested that non-linear time series models offer superior forecasting for the Canadian unemployment rate, especially in capturing the asymmetry across both short and long-forecasting horizons when compared to linear time series models. Therefore, additional research is needed to identify the most suitable forecasting methods for unemployment rate considering different forecasting periods.

4.8 Chapter Summary

In this chapter, the unemployment rate in Palestine over the period from 1996Q1 to 2023Q2 has been described. Furthermore, the findings indicated that the best ARIMA model was ARIMA(0,1,1)(0,1,1)[4] while the best NNAR model was NNAR(1,1,10)[4]. The forecasting performance using NNAR model was better than that obtained from SARIMA whereby all accuracy measures of both training and testing datasets for NNAR model was lower than those for SARIMA model as well as providing the predict of the unemployment rate in the next few quarters. Furthermore, the chapter provided discussions of the study findings and contrast them with the literature showing that there is inconclusive evidence on the statistical methods of forecasting unemployment rate. The next chapter will present the conclusion and recommendation of this study.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter is devoted to show the conclusion and recommendations of this study. The Second Section shows the conclusion of this study. The Third Section provides the recommendations for policy implications and future work.

5.2 Conclusion

Considering the significance of predicting the unemployment rate in Palestine as a crucial macroeconomic indicator for the country's economy. The current study aimed to assess and compare the accuracy of the forecasted Palestinian unemployment rate using SARIMA and NNAR models. This is important for policy implications that are usually require accurate statistical results.

The examinations indicated that the unemployment rate in Palestine demonstrates a non-stationary, nonlinear, and seasonal pattern over the study period. This study found that the SARIMA(0,1,1)(0,1,1)[4] is the best classical linear time series model to forecast the out-of-sample unemployment rate in Palestine. However, its residuals of the errors were not normally distributed which affect the accuracy and predictive performance of the model. On the other hand, the results of this study unveiled that the best nonlinear model using neural networks was NNAR(1,1,10)[4], which exhibits the best forecasting performance metrics including RMSE, MAE, and MAPE, and MASE using 10 nodes input hidden layer as compared to other hidden layers.

In terms of comparing forecasting performance, this study showed that NNAR(1,1,10)[4] outperformed SARIMA(0,1,1)(0,1,1)[4] in terms of forecasting accuracy measures (i.e., RMSE, MAE, MAPE, and MASE), predictive values, and DM test for predictive accuracy for both training sample and testing sample datasets.

Finally, unemployment in Palestine persisted as a significant persistent problem, with rates oscillating between approximately 24.7% and 26.7% based on the NNAR model while it fluctuated around the rate of 24.9 to 27.5% based on the SARIMA model for the testing dataset. Hence, it exhibited a pronounced and consistent seasonal pattern.

5.3 Recommendations

Based on the conclusions emerged from this study, a several of potential recommendations can be forwarded for both policy implications and future research, which are mentioned below:

1. It appears that the unemployment issue in Palestine is chronic and proves challenging to predict using both linear and non-linear models. In light of this, policymakers and the government of Palestine should implement early interventions for long-term and sustainable labor market regulations to manage and eradicate unemployment effectively. Indeed, this could encompass a range of strategies such as job training programs, apprenticeship schemes, and educational initiatives aimed at equipping individuals with the skills demanded by the job market.

2. Further work is needed to compare more forecasting methods and assess the predictive accuracy of unemployment rate in Palestine including Holt-Winters, self-exciting threshold auto-regressive models, machine learning and hybrid methods.
3. Given that ARIMA models exhibited non-normal and heteroscedastic of residuals. it is recommended to employ a Generalized Autoregressive Conditional Heteroscedasticity (GARCH) can be used to account for nonmorality and heteroscedasticity in the error residuals.

4. To increase the accuracy of the forecasting models, it is important to include important factors of unemployment in the future work such as inflation rate, gross domestic product, interest rate, and percentage of tertiary education.

5. In light of our findings, confirms the importance of considering spatial and gender-specific variations in policy formulation. Future research and policy implications should take into account the nuanced impacts across diverse geographic regions and gender demographics, ensuring a more comprehensive and equitable approach to addressing the identified trends.

References

Abugamea, G. (2018). Determinants of Unemployment: Empirical Evidence from Palestine. *Munich Personal RePEc Archive*, Paper No. 89424.

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics, 2014*, 1-7.

Bădulescu, A. (2006). Unemployment in Romania. A retrospective study. *Theor. Appl. Econ., 2*, 71–76.

Barnichon, R., & Garda, P. (2016). Forecasting unemployment across countries: The ins and outs. *European Economic Review*, 84, 165–183.

Barrow, D., & Kourentzes, N. (2018). The impact of special days in call arrivals forecasting: A neural network approach to modelling special days. *European Journal of Operational Research*, 264(3):967–977.

Box, G., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* San Francisco: John Wiley & Sons.

Bussiere, M., & Fratzscher, M. (2006). Towards a new early warning system of financial crises. *J. Int. Money Financ., 25*, 953–973.

Cai, F., & Wang, M. (2010). Growth and structural changes in employment in transition China. *J. Comp. Econ., 38*, 71–81.

Chakraborty, T., Chakraborty, A. K., Biswas, M., Banerjee, S., & Bhattacharya , S. (2021). Unemployment Rate Forecasting: A Hybrid Approach. *Computational Economics , 57*, pages183–201.

Charemza, W., & Deadman, D. (1992). *New Directions in Econometric Practice.* 1st Edition, Aldershot:Edward Elgar.

Ciaburro, G., & Venkateswaran, B. (2017). *Neural Networks with R.* Packt Publishing.

Cryer, J., & Chan, K.-S. (2008). *Time Series Analysis in R Second Edition.* LLC, 233 Spring Street, New York, NY 10013, USA: Springer Science+Business Media.

Daoud, Y. (2006). Short Term Measures to alleviate unemployment in Palestine. *Prceeding of MAS's Annual Conference: Unemployment in the Palestinian Territroy "Relaity and Stratigies to Alleviate it"* (pp. 23 - 36). Rammalh: MAS.

Davidescu, A., Apostu, S., & 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), 325.

Diebold , F. (2015). Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold–Mariano Tests. *Journal of Business & Economic Statistics*, 33:1, 1-1.

Dritsaki, C. (2016). Forecast of SARIMA models: An application to unemployment rates of Greece. *Am. J. Appl. Math. Stat., 4*, 136–148.

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.

Dritsakis, N., Athianos, S., Stylianou, T., & Samaras, I. (2018). Forecasting unemployment rates in Greece. *Int. J. Sci. Basic Appl. Res., 37*, 43–55.

Dumičić, K., Čeh Časni, A., & Žmuk, B. (2015). Forecasting unemployment rate in selected european countries using smoothing methods. *International Journal of Social, Education, Economics and Management Engineering, 9*(4), 867–872.

ERDAL, G., DOĞAN, H. G., & KARAKAŞ, G. (2015). The Analysis of The Relationship Between Unemployment And Inflation In Turkey By Var Model. *journal of new results in science*, Number:8, Pages: 22-29.

Feuerriegel, S., & Gordon, J. (2019). News-based forecasts of macroeconomic indicators: A semantic path model for interpretable predictions. *European Journal of Operational Research, 272*(1), 162–175.

Firmino, P. R., de Mattos Neto, P. S., & Ferreira, T. A. (2014). Correcting and combining time series forecasters. *Neural Networks, 50*, 1 - 11.

Funke, M. (1992). Time-series forecasting of the German unemployment rate. *J. Forecast., 11*, 111-125.

Gagea, M., & Balan, C. (2008). Prognosis of monthly unemployment rate in the European Union through methods based on econometric models. *Ann. Fac. Econ., 2*, 848–853.

Gogas, P., Papadimitriou , T., & Sofianos, E. (2022). Forecasting unemployment in the euro area with machine learning. *Journal of Forecasting, 41*(3), 551-566.

Gostkowski, M., & Rokicki, T. (2021). Forecasting the Unemployment Rate: Application of Selected. *European Research Studies Journal, Volume XXIV, Issue 3*, pp. 985-1000.

Hill, T., O'Connor, M., & Remus, W. (1996). Neural network models for time series forecasts. *forecasts. Management science*, 42(7):1082–1092.

Husin, W., Abdullah, N., & Rockie, N. (2023). Neural Network Model in Forecasting Malaysia's Unemployment Rates.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice.* OTexts.

International Labor Organisation. (2013). *Global Employment Trends for Youth*. Retrieved from https://www.ilo.org/wcmsp5

Joseph, V. (2022). Optimal ratio for data splitting. Statistical Analysis and Data Mining. *The ASA Data Science Journal, 15(4)*, 531-538.

Kavkler, A., Dncic, D., Babucea, A., Biani, I., Bohm, B., Tevdoski, D., . . . Borsi, D. (2009). Cox regression models for unemployment duration in Romania, Austria, Slovenia, Croatia and Macedonia. *Rom. J. Econ. Forecast*, 2, 81–104.

Khan Jafur, Z. R., Sookia, N. U., Nunkoo Gonpot, P., & Seetanah, B. (2017). Out-of-sample forecasting of the Canadian unemployment rates using univariate models. *Applied Economics Letters, 24*(5), 1097–1101.

Khashei, M., & Bijari, M. (2010, January). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with Applications, 37*(1), 479-489.

Khashei, M., Bijari, M., & Ardali, R. G. (2009). Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and Artificial Neural Networks (ANNs). *Neurocomputing, 72*(4-6), 956-967.

Kozuch, A., Cywicka, D., & Adamowicz, K. (2023). A Comparison of Artificial Neural Network and Time Series Models for Timber Price Forecasting. *Forests*, 14, 177.

Lo Duca, A. (2021, july 13). Understanding the Seasonal Order of the SARIMA Model. Towards Data Science.

Meyler, A., Kenny, G., & Quinn, T. (1998, December 2). Forecasting irish inflation using ARIMA models. *Central Bank and Financial Services Authority of Ireland Technical Paper Series*, pp. 1-48.

Mishra, A. K., & Desai, V. R. (2005). Spatial and temporal drought analysis in the Kansabati River Basin, India. *Int. J. River Basin Manage*, 3 (1), pp. 31-41.

Mladenovic, J., Ilic, I., & Kostic, Z. (2017). Modeling the unemployment rate at the EU level by using box-jenkins methodology. *KnE Soc. Sci.*

Nagao, S., Takeda, F., & Tanaka, R. (2019). Nowcasting of the us unemployment rate using google trends. *Finance Research Letters, 30*, 103–109.

Nikolaos, D., Stergios, A., Tasos, S., & Ioannis, S. (2016). Forecasting Unemployment Rates in Greece. *International Journal of Sciences: Basic and Applied Research*, 4531, 43–55.

Nkwatoh, L. (2012). Forecasting Unemployment Rates in Nigeria Using Univariate time series Models. *International Journal of Business and Commerce*, 1(12), 33–46.

Nor, M., Saharan, S., Lin, L., & Salleh, R. (2018). Forecasting of Unemployment Rate in Malaysia Using Exponential Smoothing Methods. International Journal of Engineering and Technology. 7(4.30), 451.

PCBS. (2022). *Labor Force Survey Report.* Ramallah: Palestinian Central Bureau of Statistics.

Peláez , R. F. (2006). Using neural nets to forecast the unemployment rate: a promising application of an emerging quantitative method. *Business Economics*, 37-44.

Rublikova, E., & Lubyova, M. (2013). Estimating ARIMA-ARCH model rate of unemployment in Slovakia. *Forecast. Pap. Progn. Pr., 5*, 275–289.

Schanne, N., Wapler, R., & Weyh, A. (2010). Regional unemployment forecasts with spatial interdependencies. *Int. J. Forecast., 26*, 908–926.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics (Vol. 6, pp. 497-516). Boston, MA: pearson.* (4th edition ed.). NJ, USA: Pearson Education, Upper Saddle River, NJ, USA.

Tansel, I., Yang, S., Venkataraman, G., Sasirathsiri, A., Bao, W., & Mahendrakar, N. (1999). *Modeling time series data by using neural networks and genetic algorithms.* New York, NY, USA, 1999: ASME Press.

Taylor, J. W. (2008). A comparison of univariate time series methods for forecasting intraday. *Management Science*, 54(2):253–265.

Teräsvirta, T., Van Dijk, D., & Medeiros, M. C. (2005). Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting, 21*(4), 755–774.

Vicente, M. R., López-Menéndez, A. J., & 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.

Vladi, E., & Eglantina, H. (2019). *The Impact of Macroeconomic Indicators on Unemployment Rate: Western Balkan Countries.*

Yamacli, D. S., & Yamacli, S. (2023). Estimation of the unemployment rate in Turkey: A comparison of the ARIMA and machine learning models including Covid-19 pandemic periods. *Heliyon, 9*, e12796.

Appendix A

R code for ARIMA and NNAR models

R Packages

```
library(forecast)
library(tidyverse)
library(readxl)
library(TSstudio)
library(lmtest)
library(Metrics)
library(uroot)
library(urca)
library(aTSA)
library(portes)
library(FinTS)
library(TSA)
library(tseries)
library(gt)
```

Import Data

```
unemployment_rate =
read_excel("C:/Users/Downloads/thesis/Entropy/unemployment_rate.xlsx")
y <- ts(unemployment_rate, start=1996, frequency = 4)
autoplot(y)
ggsubseriesplot(y)
ggseasonplot(y, polar=TRUE) +  ylab("%")
ggseasonplot(y, polar=TRUE)
ggAcf(y, lag=48,main="Autocorrelation")
ggPacf(y,lag=48,main ="Partial Autocorrelation")
```

Splitting data (training; test)

```
training <- window(y, start=1996, end=c(2017,4))
test <- tail(y, 5.5*4)
```

Fitting SARIMA Model

```
training%>% diff(lag=1) %>% ggtsdisplay()
```

Stationary Test

```
adf.test(training)
pp.test(training, type = c("Z_rho", "Z_tau"), lag.short = TRUE, output = TRUE)
```

Structural Breaks - Zivot Andrew Test

```
za_trend <- ur.za(training, model =  "trend")
summary(za_trend)
za_both <- ur.za(training, model =  "both")
summary(za_both)
```

Hegy Test

```
hegy.test(training)
```

Fitting The Best SARIMA Models

```
sarima1=Arima(training, order=c(1,1,1),seasonal=list(order=c(0,0,1)))
summary(arima1)
sarima2=Arima(training, order=c(0,1,1),seasonal=list(order=c(0,1,1)))
summary(arima2)
```

Checking Residuals For ARIMA Model

```
checkresiduals(arima1, lag = 48)
tsdiag(arima1)
shapiro.test(arima1$residuals)
qqnorm(arima1$residuals)
qqline(arima1$residuals)
fcast<- arima1 %>% forecast::forecast(h=22)
forecast::accuracy(fcast, test)
autoplot(fcast)+autolayer(test)
checkresiduals(fcast, lag = 48)
Box.test(fcast, lag = 4)
Box.test(fcast, lag = 8)
Box.test(fcast, lag = 12)
Box.test(fcast, lag = 16)
```

```
arch.test(arima1)
checkresiduals(arima2, lag =48)
tsdiag(arima2)
shapiro.test(arima2$residuals)
qqnorm(arima2$residuals)
qqline(arima2$residuals)
fcast1<- arima2 %>% forecast::forecast(h=22)
forecast::accuracy(fcast1, test)
autoplot(fcast1)+autolayer(test)
checkresiduals(fcast1, lag=48)
Box.test(fcast1, lag = 4)
Box.test(fcast1, lag = 8)
Box.test(fcast1, lag = 12)
Box.test(fcast1, lag = 16)
arch.test(arima2)
```

Fitting NNAR Model

```
nnar = nnetar(training, repeats = 20,p= 1, P = 1, size =10)
fcastnnar = forecast::forecast(nnar, h = 22)
forecast::accuracy(fcastnnar, test)
shapiro.test(nnar$residuals)
qqnorm(nnar$residuals)
qqline(nnar$residuals)
autoplot(fcast)+autolayer(test)
```

Checking Residuals For NNAR Model

```
checkresiduals(fcastnnar)
Box.test(fcastnnar, lag = 4)
Box.test(fcastnnar, lag = 8)
Box.test(fcastnnar, lag = 12)
Box.test(fcastnnar, lag = 16)
```

DM Test To Compare Forecasting Accuracy Between SARIMA And NNAR

```
e1 = residuals(fcast1)
```

```
e2 = residuals(fcastnnar)
dm.test(e1, e2, alternative = c("greater"),  h = 1,varestimator = c("acf", "bartlett"))
e11 = residuals(arima2)
e22 = residuals(nnar)
dm.test(e11, e22, alternative = c("greater"),  h = 1,varestimator = c("acf", "bartlett"))
```