

BINDURA UNIVERSITY OF SCIENCE EDUCATION
FACULTY OF SCIENCE AND ENGINEERING
DEPARTMENT OF STATISTICS AND MATHEMATICS



**A COMPARATIVE STUDY OF ARIMA, FEEDFORWARD NEURAL NETWORK
AND HYBRID ARIMA–FFNN MODELS FOR STOCK PRICE FORECASTING:
EVIDENCE FROM ECONET WIRELESS ZIMBABWE (2020–2024)**

BY

PRESLEY TANAKA CHITANDO

B224064B

**TO BE SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR BSC.HONOURS IN STATISTICS AND FINANCIAL MATHEMATICS**

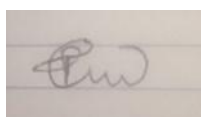
SUPERVISOR: MISS J PAGAN'A

JUNE 2026

APPROVAL FORM

This is to certify that this research project is the result of my own research work and has not been copied or extracted from past sources without acknowledgement. I hereby declare that no part of it has been presented for another degree in this University or elsewhere.

Presley Tanaka Chitando



05/06/2026

B224064B

Signature

Date

Certified by:

Miss Pagan'a



05/06/2026

Supervisor

Signature

Date

Dr Magodora



13/06/2026

Chairperson

Signature

Date

AUTHORSHIP DECLARATION STATEMENT

Title of the Thesis: A COMPARATIVE STUDY OF ARIMA, FEEDFORWARD NEURAL NETWORK AND HYBRID ARIMA–FFNN MODELS FOR STOCK PRICE FORECASTING: EVIDENCE FROM ECONET WIRELESS ZIMBABWE (2020–2024)

Author: Presley Tanaka Chitando

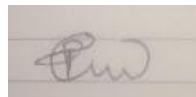
Program: Bachelor of Science Honours Degree in Statistics and Financial Mathematics.

I, the undersigned author of the above-mentioned thesis, hereby declare that:

1. This thesis is my original work and has been prepared by me in accordance with the institution's requirements.
2. All sources, data, and references used in this thesis have been acknowledged and cited appropriately.
3. This thesis has not been submitted elsewhere for any degree or diploma.
4. I have obtained all necessary permissions for the inclusion of third-party content where applicable.

I affirm that this declaration is made with full integrity and in compliance with the institution's policies and academic practice.

Author's Signature:



Date:

05/06/2026

DEDICATION

To my beloved mother whose unwavering love, sacrifices, prayers and encouragement have been the foundation of my journey. This achievement is as much yours as it is mine, for without your endless support and guidance, it would not have been possible. Thank you for every sacrifice you have made and for always reminding me that hard work and faith can overcome any obstacle. With profound love, gratitude, and appreciation, I dedicate this work to you.

ACKNOWLEDGEMENTS

First and foremost, I give all glory and honour to God Almighty for His abundant grace, wisdom, strength, and guidance throughout this academic journey. I would also like to express my sincere gratitude to my supervisor, Ma'am Pagan'a, for her invaluable guidance, patience, constructive criticism and continuous encouragement throughout the course of this research. Her expertise and dedication greatly contributed to the successful completion of this study. My heartfelt appreciation also goes to Ma'am Hlupo for her assistance, support, and willingness to provide guidance whenever it was needed. Her encouragement and insightful contributions are deeply appreciated. I am profoundly grateful to my family and friends for their unwavering love, prayers, understanding, and moral support. Finally, I would like to extend my sincere appreciation to the institution for providing a conducive learning environment and the academic resources necessary for carrying out this research. To everyone who, in one way or another, contributed to the successful completion of this dissertation, I offer my deepest thanks.

ABSTRACT

This study developed and compared the forecasting performance of the Autoregressive Integrated Moving Average (ARIMA), Feedforward Neural Network (FFNN), and Hybrid ARIMA–FFNN models using monthly closing stock prices of Econet Wireless Zimbabwe from January 2020 to December 2024. The ZiG/USD exchange rate and inflation rate were incorporated as explanatory variables in the FFNN and hybrid models. A quantitative, explanatory, and comparative research design based on the positivist paradigm was adopted. Forecast accuracy was evaluated using the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). The Hybrid ARIMA–FFNN model produced the best forecasting performance, achieving the highest explanatory power ($R^2 = 0.6221$) and the lowest forecast error (RMSE = 96.71), outperforming both the standalone ARIMA and FFNN models. The findings demonstrate that combining linear and non-linear modelling techniques improves stock price forecasting under conditions of economic instability and structural change. The study recommends the adoption of hybrid econometric and machine learning models to enhance forecasting accuracy in the Zimbabwean stock market.

Keywords: Stock price forecasting, ARIMA, Feedforward Neural Network, Hybrid ARIMA–FFNN, Econet Wireless Zimbabwe, Zimbabwe Stock Exchange, Financial econometrics.

TABLE OF CONTENTS

Contents

APPROVAL FORM	ii
AUTHORSHIP DECLARATION STATEMENT.....	iii
DEDICATION.....	iv
ACKNOWLEDGEMENTS	v
ABSTRACT.....	vi
TABLE OF CONTENTS.....	vii
LIST OF FIGURES	x
LIST OF TABLES.....	xi
LIST OF ACRONYMS	xii
CHAPTER 1: INTRODUCTION	1
1.0 Background to the Study	1
1.1 Statement of the Problem	2
1.2 Research Objectives	2
1.2.1 Main Objective	3
1.2.2 Specific Objectives	3
1.3 Research Questions	3
1.4 Scope of the Study	3
1.5 Significance of the Study	4
1.6 Assumptions of the Study	4
1.7 Definition of Key Terms	6
1.8 Chapter summary	7
2.0 Introduction	8
2.1 Theoretical Literature	8
2.1.1 Efficient Market Hypothesis	8
2.1.2 Time Series Analysis Theory	9
2.1.3 Neural Networks and Universal Approximation Theorem	9
2.1.4 Financial Forecasting Theory	10
2.1.5 Behavioural Finance Theory	10
2.2 Empirical Literature	11
2.2.1 Applications of ARIMA in Stock Prices Forecasts	11
2.2.2 Using Neural Networks for Financial Forecasting	11
2.2.3 Comparative Analysis between ARIMA and Neural Networks	12
2.2.4 Hybrid ARIMA–FFNN Model	12
2.2.5 Behavioural Finance and Equities Market	13

2.3 Research Gap	14
2.4 Conceptual Framework.....	14
2.5 Chapter Summary	15
CHAPTER 3: RESEARCH METHODOLOGY	17
3.0 INTRODUCTION.....	17
3.1 Research Paradigm and Design.....	17
3.2 Data Sources and Collection	17
3.3 Population, Sampling and Train/Test Partitioning.....	18
3.4 Variable Specification.....	18
3.4.1 Dependent Variable	18
3.4.2 Independent Variables.....	19
3.5 Model Specification.....	19
3.5.1 ARIMA Model.....	19
3.5.2 Feed Forward Neural Network (FFNN) Model.....	20
3.5.3 Hybrid ARIMA-FFNN Model	20
3.6 Model Pre-Tests and Diagnostic Checks.....	21
3.6.1 Stationarity: Augmented Dickey Fuller Test.....	21
3.6.2 Autocorrelation: Ljung-Box Q-Test.....	21
3.6.3 Heteroscedasticity: Breusch-Pagan Test.....	21
3.6.4 Normality: Jarque-Bera Test	21
3.6.5 Non-linearity: ARCH-LM test.....	22
3.7 Evaluation of the Models.....	22
3.7.1 Metrics for Evaluating Forecasting Performance.....	22
3.7.2 Choosing the ARIMA Model: Minimising AIC and BIC.....	23
3.8 Software and Computing Tools	23
3.9 Ethical considerations.....	23
3.10 Chapter Summary	24
CHAPTER 4: DATA ANALYSIS AND PRESENTATION OF FINDINGS	25
4.0 INTRODUCTION.....	25
4.1 Descriptive Statistics and Data Exploration.....	25
4.1.1 Descriptive Statistics.....	25
4.1.2 Time Series Plot of Raw Stock Prices.....	26
4.1.3 ACF and PACF of Raw Series	27
4.2 Stationarity Testing and Transformation.....	27
4.2.1 ADF Test on Raw Series.....	27
4.2.2 First Difference.....	28
4.2.3 ADF Test on Differenced Series.....	29

4.3 ARIMA Model: Identification, Estimation, and Diagnostics.....	30
4.3.1 ACF and PACF of Differenced Series.....	30
4.3.2 Automatic Model Selection	30
4.3.3 ARIMA (2, 1, 2) Residual Diagnostics.....	31
4.3.4 ARIMA (2, 1, 2) Out-of-Sample Forecast	33
4.4 Feedforward Neural Network (FFNN) Architecture.....	33
4.4.1 Normalisation of Input/Output Variables	33
4.4.2 Architecture Selection.....	33
4.4.3 FFNN Network Topology	35
4.4.4 FFNN Out-of-Sample Forecast.....	35
4.5 Hybrid ARIMA–FFNN Model.....	36
4.5.1 Rationale and Non-linearity Confirmation.....	36
4.5.2 Hybrid Model Architecture and Training	36
4.5.3 Residual Analysis for the Hybrid Model.....	38
4.5.4 Hybrid Model Out-of-Sample Forecast	38
4.6 Cross-Model Comparison and Discussion	39
4.6.1 Forecast Performance Comparison.....	39
4.6.2 Discussion of Results.....	41
4.6.3 Implications of the Findings.....	42
4.7 Chapter Summary	43
CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS	44
5.0 Introduction.....	44
5.1 Conclusions.....	44
5.1.1 Objective 1: ARIMA and FFNN Models	44
5.1.2 Objective 2: Development of the Hybrid ARIMA–FFNN Model	44
5.1.3 Objective 3: Accuracy Comparison for the Models.....	44
5.1.4 Objective 4: Advantages and Disadvantages Of Each Model	45
5.1.5 Objective 5: Most Reliable Model in Volatile Conditions	45
5.2 Recommendations.....	45
5.3 Summary and Conclusion	46
REFERENCES	47
APPENDICES	51

LIST OF FIGURES

Figure 2.1: Conceptual Framework for the Comparative ARIMA–FFNN–Hybrid Study	15
Figure 4.1: Monthly Closing Prices of Econet Wireless Zimbabwe (January 2020–December 2024)	266
Figure 4.2: ACF and PACF of Raw Closing Price Series	277
Figure 4.3: First Difference ($\Delta Y_t = Y_t - Y_{t-1}$)	288
Figure 4.4: ACF and PACF of Differenced Series	30
Figure 4.5: ARIMA(2,1,2) Out-of-Sample Forecast vs Actual Closing Prices	Error! Bookmark not defined.3
Figure 4.6: Optimal FFNN Architecture [4–4–1] — Network Topology	355
Figure 4.7: FFNN [4-4-1] Out-of-Sample Forecast vs Actual Closing Prices	Error! Bookmark not defined.5
Figure 4.8: ARIMA(2,1,2) Residuals	Error! Bookmark not defined.6
Figure 4.9: Hybrid ARIMA-FFNN network topology	Error! Bookmark not defined.7
Figure 4.10: Hybrid FFNN Training Loss Convergence	Error! Bookmark not defined.7
Figure 4.11: Hybrid ARIMA–FFNN Out-of-Sample Forecast vs Actual Closing Prices	Error! Bookmark not defined.8
Figure 4.12: Comparative Forecast Performance	40

LIST OF TABLES

Table 4.1: Descriptive Statistics.....	255
Table 4.2: ADF Unit Root Test — Raw Series	288
Table 4.3: ADF Unit Root Test — Differenced Series	299
Table 4.4: ARIMA Model Selection via AIC and BIC	30
Table 4.5: Ljung-Box Test on ARIMA(2,1,2) Residuals	311
Table 4.6: Jarque-Bera Normality Test on ARIMA (2, 1, 2) Residuals.....	322
Table 4.7: Breusch-Pagan Heteroscedasticity Test on ARIMA (2, 1, 2) Residuals	322
Table 4.8: FFNN Architecture Selection.....	344
Table 4.9: Hybrid Model Residual Analysis Tests.....	388
Table 4.10: Out-of-Sample Forecast Performance	399
Table 4.11: Actual vs Forecast Values	399

LIST OF ACRONYMS

ARIMA - Autoregressive Integrated Moving Average

FFNN - Feed-Forward Neural Network

MAE - Mean Absolute Error

MAPE - Mean Absolute Percentage Error

MSE - Mean Squared Error

PICP - Prediction Interval Coverage Probability

PINAW - Prediction Interval Normalized Average Width

RMSE - Root Mean Squared Error

R^2 - Coefficient of Determination

sMAPE - Symmetric Mean Absolute Percentage Error

CHAPTER 1: INTRODUCTION

1.0 Background to the Study

Predicting stock prices is perhaps the most extensively researched topic in financial econometrics. Theoretical and practical research is focused on this topic due to its direct influence on the formulation of strategies, portfolio formation, and analyses of market efficiency. Predictive modeling helps make better-informed decisions by the investor, analyst, or manager working with assets in highly volatile and unstable markets.

Econet Wireless Zimbabwe Ltd., a company traded on the Zimbabwe Stock Exchange (ZSE), represents a perfect example. Founded in 1998, Econet was the first private mobile service provider in Zimbabwe and has since developed into a diversified technology and financial group providing mobile communications and Internet services as well as mobile financial services using its EcoCash application. Econet is among the most actively traded and largest capitalization stocks at the ZSE and can be regarded as the indicator of general economic and industry trends in Zimbabwe (Zimbabwe Stock Exchange, 2024; Reserve Bank of Zimbabwe, 2024).

The years between 2020 and 2024 were notable for unusual levels of economic uncertainty in Zimbabwe, such as the COVID-19 induced economic turmoil, several changes in monetary regimes, resulting in the launch of the ZiG currency in 2024, periods of hyperinflation, as well as the presence of exchange rate fluctuations. Such structural events introduced significant non-linearity and regime switching into the series of prices of Econet shares, making them an interesting and challenging forecasting problem due to their nature. (ZIMSTAT, 2023; Reserve Bank of Zimbabwe, 2024).

There exist two primary approaches to modeling stocks and their prices. First, the approach based on classical statistics time-series models, where the most famous example can be attributed to the work of Box and Jenkins (1976), introducing the model ARIMA, i.e., the autoregressive integrated moving average. Such models have proven themselves as good as far as modeling the autocorrelation properties of financial time series go and are still widely used benchmarks in competitions.

Feedforward Neural Networks (FFNNs) can act as universal function approximators, implying that the network can learn any continuous non-linear mapping from inputs to outputs (Cybenko, 1989; Hornik, 1991). Studies have found FFNNs to consistently

outperform conventional statistical models in environments characterized by extreme volatility or structural changes in market dynamics (Kara et al., 2011; Zhong & Enke, 2019).

The third stream of thought emerging more recently suggests a new model that integrates the strengths of linear modeling using ARIMA and non-linear learning of FFNNs. Zhang (2003) first introduced such hybrid models based on the argument that most real-world time series involve elements of linearity and non-linearity, and modeling each component independently provides improved forecasts. In this paper, an extended version of Zhang's (2003) work is applied to the scenario involving Econet Wireless Zimbabwe.

1.1 Statement of the Problem

Although there has been substantial growth in studies involving stock price prediction using machine learning techniques, methodological challenges remain, particularly in developing countries and emerging economies. While classical models such as ARIMA are well studied and interpretable, they are limited in capturing the non-linear effects of macroeconomic shocks, exchange rate fluctuations, and investor behaviour. FFNNs can model these non-linear relationships but are susceptible to overfitting with relatively small samples and may not accurately capture extreme price movements, a limitation commonly referred to as scale dampening.

There has been no comparative or hybrid modelling study forecasting the stock prices of Econet Wireless Zimbabwe over the period 2020–2024. This period is significant because it is characterized by both linear market trends and structural breaks resulting from policy changes.

Therefore, this study seeks to bridge this gap by developing and comparing ARIMA, FFNN, and Hybrid ARIMA-FFNN models to forecast the stock prices of Econet Wireless Zimbabwe and determine the model that best captures both the linear and non-linear behaviour of stock prices.

1.2 Research Objectives

1.2.1 Main Objective

To develop and compare ARIMA, Feedforward Neural Network (FFNN), and Hybrid ARIMA–FFNN models for forecasting the stock prices of Econet Wireless Zimbabwe over the period January 2020 to December 2024.

1.2.2 Specific Objectives

1. To develop and implement standalone ARIMA and FFNN models for predicting monthly closing stock prices of Econet Wireless Zimbabwe.
2. To develop a Hybrid ARIMA–FFNN model that combines the linear forecasting strength of ARIMA with the non-linear learning capability of the FFNN.
3. To evaluate and compare the forecasting accuracy of all three models using statistical error metrics including RMSE, MAE, MAPE, and R^2 .
4. To examine the strengths and limitations of each model in capturing linear and non-linear stock price dynamics under the volatile conditions of the Zimbabwean market.
5. To identify the most reliable forecasting model for Econet Wireless Zimbabwe's stock prices under conditions of macroeconomic volatility and structural breaks.

1.3 Research Questions

1. How accurately do the standalone ARIMA and FFNN models predict the monthly closing stock prices of Econet Wireless Zimbabwe over the period 2020–2024?
2. To what extent does the Hybrid ARIMA–FFNN model improve forecasting accuracy relative to the standalone ARIMA and FFNN models?
3. Which model between ARIMA, FFNN or the Hybrid ARIMA–FFNN, provides the most reliable forecasts under the volatile and structurally shifting market conditions characterising the Zimbabwean economy between 2020 and 2024?

1.4 Scope of the Study

The study shall be confined to only monthly closing prices for shares of Econet Wireless Zimbabwe listed at the Zimbabwe Stock Exchange for the period ranging from January 2020 to December 2024. In this context, two variables that will serve as exogenous variables in both the FFNN and Hybrid models are the average monthly ZiG/USD exchange rate and the

CPI inflation rate per month, obtained from the Reserve Bank of Zimbabwe (RBZ) and the Zimbabwe National Statistics Agency (ZIMSTAT), respectively. This research paper is confined to an analysis of ARIMA, FFNN, and Hybrid ARIMA-FFNN methods; LSTM neural networks do not fall under the scope of this research paper.

1.5 Significance of the Study

To Academia: This study makes a significant contribution to the limited research on the use of hybrid econometric and machine learning techniques for forecasting frontier African stock markets. It also provides a basis for future studies on stock price forecasting, particularly in developing countries and emerging economies.

To Industry: The model comparison provides investors and portfolio managers with a well-founded recommendation for selecting an appropriate forecasting approach for Zimbabwe's volatile and structurally unstable equity market. For financial analysts and Zimbabwe Stock Exchange (ZSE) practitioners, the study demonstrates how macroeconomic indicators such as exchange rates and inflation can be systematically integrated into neural network-based forecasting methods. The findings will also assist the management of Econet Wireless Zimbabwe in managing investor relations, raising capital, and supporting strategic decision-making through improved understanding of stock price dynamics.

To Policymakers: The insights gained from the forecasting models will be useful in designing monetary policy frameworks capable of promoting stability in the equity market and enhancing informed policy decisions.

1.6 Assumptions of the Study

ARIMA Model

1. The stock price series can be rendered stationary through a finite order of differencing.
2. Residuals from the fitted ARIMA model are uncorrelated (white noise) and homoscedastic under stable market conditions.
3. Past prices contain sufficient autocorrelation structure to support linear autoregressive and moving average modelling.

Feedforward Neural Network (FFNN)

1. Sixty monthly observations, when split 80/20, provide sufficient data for the network to learn meaningful non-linear patterns without severe overfitting.
2. Non-linear relationships exist between lagged stock prices, exchange rates, inflation, and future stock prices.
3. The selected [4–4–1] network architecture achieves an appropriate bias-variance trade-off for this dataset.

Hybrid ARIMA–FFNN Model

1. The Econet stock price series contains both a linear component, captured by ARIMA, and a non-linear residual component, learnable by the FFNN.
2. ARIMA residuals retain sufficient non-linear structure to serve as a meaningful training signal for the neural network's error-correction stage.
3. Secondary data from Investing.com, RBZ, and ZIMSTAT are sufficiently accurate and complete for the purposes of this analysis.

1.7 Limitations of the Study

- i. The research is based on one stock and one branch of the ZSE, limiting its generalizability to other listed companies or stock exchanges.
- ii. Only 60 months of data make the dataset relatively small for machine learning methods, raising the issue of possible overfitting when applying more complicated FFNN algorithms.
- iii. The use of secondary sources for data collection, specifically Investing.com and publicly available databases, increases the likelihood of accounting errors.
- iv. Non-quantitative features like quality of corporate governance, political risk, or sentiment indices have not been accounted for in the model development.
- v. The model used is fixed in its structure; adaptive neural networks, which would possibly be more efficient in such environments, are not considered.
- vi. The research covers periods of several shocks; results may thus not be generalizable to other periods of more stable macroeconomic environment.

An additional limitation pertains to the formulation of the forecast variable. The analysis predicts the nominal ZWL/ZiG closing prices as opposed to log-returns. Log-returns have

been traditionally used in forecasting because they are stationary, but this paper focuses on predicting the actual investor-relevant closing price. Predicting the actual closing price makes the process non-stationary, especially with regard to the 2024 currency reform of the ZiG. Adebisi et al. (2014), in their paper using a hybrid model of ARIMA-ANN for emerging market stocks, use the same approach.

1.7 Definition of Key Terms

ARIMA Model (Autoregressive Integrated Moving Average): A linear time series approach that combines autoregressive components (AR), the first difference I used to obtain stationarity, and moving average components (MA) to capture the autocorrelation in residuals (Box & Jenkins, 1976).

Feedforward Neural Network (FFNN): An artificial neural network model whereby data flow unidirectionally, from the input layer to the output layer through one or more hidden layers without looping (Rumelhart et al., 1986).

Hybrid Model: A technique for making predictions by combining two or more different types of modeling techniques. The current research uses an ARIMA technique for the linear part while FFNN captures the non-linear part of the residual (Zhang, 2003).

Stationarity: A feature of time series data, whereby the moments such as mean, variance, and autocovariance are constant and do not vary over time. Stationarity is necessary before applying the ARIMA model (Dickey & Fuller, 1979).

Root Mean Squared Error (RMSE): A measure of error calculated by taking the square root of the average of squared errors, where larger errors get larger penalties (Hodson, 2022).

Mean Absolute Percentage Error (MAPE): It is an accuracy measure that measures errors in terms of percentages.

Overfitting: An error made during the creation of models whereby the model learns from the noise in the training dataset rather than from the underlying signal, leading to poor model performance on an unseen test dataset (Hastie et al., 2009).

Universal Approximation Theorem: A mathematical proof that shows how a feedforward neural network with one or more hidden layers and enough hidden neurons is capable of

approximating any continuous function to any desired accuracy (Cybenko, 1989; Hornik, 1991).

Zimbabwe Gold (ZiG): The structured form of currency that was created in April 2024 by the Reserve Bank of Zimbabwe and replaces the Zimbabwean Dollar (ZWL).

1.8 Chapter summary

In this chapter, the research context, problem statement, and methodology have been provided. First, background provides information about why the selected company is suitable as a case study for the research on stock price prediction. This is because Econet Wireless Zimbabwe is one of the most active companies on the Zimbabwe Stock Exchange with an interesting past during the period of 2020-2024 which included exceptional economic shocks. Secondly, problem statement identifies the research gap which was filled through the study which includes no comparative or hybrid modelling evaluation done for ZSE listed equities post-COVID in the multi-currency era. In addition, main and specific objectives, questions, scope, significance, assumptions, limitations, and definition of terms are identified. Chapter two discusses the theoretical and empirical literature of the models used in the research.

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

In this chapter, the literature that relates to stock price prediction is reviewed. The theoretical framework for both ARIMA, FFNN, and hybrid models is covered, followed by a summary of some empirical research conducted, and finally, the gap in literature which this research seeks to fill is highlighted. In addition to the above, special attention is given to the Zimbabwean and general frontier markets' situation.

2.1 Theoretical Literature

2.1.1 Efficient Market Hypothesis

The basic economic theory that guides stock price prediction is the Efficient Market Hypothesis (EMH), introduced by Fama in 1970. According to EMH, market prices always represent all existing information about an asset, suggesting that there is no prediction method that can create consistent abnormal returns. Fama classified market efficiency into three types of efficiencies: weak form efficiency when prices include all historical price information, semi-strong form efficiency if prices contain all public information in addition to historical price information, and strong form efficiency in which prices are also efficient in respect of all privately held information.

Nevertheless, many studies since then have refuted the universal validity of EMH, especially for developing and emerging markets. For instance, Shiller (2003) found evidence of excessive volatility and prices being over-reactive to events contrary to the efficient processing of information, whereas Lo (2004) developed the concept of Adaptive Market Hypothesis (AMH). This theory argues that market efficiency is not constant, but a dynamic process that depends on adaptation by market actors to different environments. Mlambo and Biekpe (2007) identified weak-form inefficient African stock exchanges, such as the ZSE, due to thin trading, low liquidity, and information asymmetry. This means that autocorrelation exists for stocks listed on the ZSE such as Econet Wireless Zimbabwe, hence supporting the use of time series forecasting.

Yet this study conducted by Mlambo & Biekpe (2007) raises an important dilemma in relation to the relationship between the market's inefficiency in generating abnormal returns and the ability of time series models to take advantage of the inefficiency and make abnormal profits. In fact, the weak form of inefficiency identified by these researchers based on lack of liquidity and information asymmetry cannot be consistent with the EMH's assertion that "there is no way to create consistent abnormal returns." Nevertheless, the EMH forms the foundation for evaluating other methods. The purpose of this dissertation is not to make abnormal returns but rather to check their feasibility.

2.1.2 Time Series Analysis Theory

Time series analysis is grounded in the idea that a set of observations arranged in chronological order has an underlying structure that can be isolated and exploited to predict the future (Box & Jenkins, 1976; Hamilton, 1994). In the development of ARIMA models according to the Box-Jenkins approach, there is a clear and methodical three-step process which consists of identifying the model order based on autocorrelation, estimating the parameters by means of maximum likelihood, and testing the residuals for lack of correlation and constancy of variance. It is noted by Chatfield (2003) that time series analysis tools are especially effective in dealing with finance variables as a consequence of the inherent autocorrelation property of financial prices.

2.1.3 Neural Networks and Universal Approximation Theorem

There are theoretical foundations to the use of FFNNs for non-linear financial prediction, which are laid out in the Universal Approximation Theorem. According to Cybenko (1989), a feedforward network having one layer of hidden nodes, which utilizes a sigmoidal activation function, can be used to approximate any continuous function in any compact subset of $\mathbb{R}^n$, given the sufficient number of neurons in this hidden layer. Further, this theorem was proven in its general form by Hornik (1991). It has great theoretical significance when applied to stock price prediction since it means that FFNNs are able, at least in theory, to learn any type of nonlinear dependency between lagged prices and macroeconomic indicators, and future stock prices, without any need to assume any form of the function a priori on the part of the researcher.

2.1.4 Financial Forecasting Theory

The theory behind financial forecasting, according to a review by Fildes and Goodwin (2007), involves using both quantification of models and subjective judgment in the making of future predictions about financial values. According to Pindyck and Rubinfeld (1998), there is no universally better model because the performance of each model varies depending on the nature of the data generating process, forecast horizon, and economic conditions. Such a theoretical perspective justifies the use of the comparative research approach used in this research where the performance of different models is compared.

2.1.5 Behavioural Finance Theory

The Behavioural Finance Theory investigates how cognitive biases and other psychological aspects contribute to the deviations from the rationality assumption of EMH, as suggested by Kahneman & Tversky (1979). Through Prospect Theory, they prove that investors are loss averse and experience losses as worse than gains of equal size, resulting in the asymmetry associated with selling behaviour. Speculative bubbles and overreactions have been identified by Shiller (2000) as contributing to price movements that are not consistent with fundamentals. Herd behaviour, overconfidence, and mental accounting are biases identified by Thaler (1999) that increase volatility on stock prices. In the case of Zimbabwe, the herding effect is exacerbated due to shallow markets and ownership structures, hence behavioural finance theory could help explain some of the abnormal price movements seen on Econet's shares in 2023–2024. The regularity introduced by behavioural finance theories also explains why applying non-linear machine learning techniques might be justified from a theoretical point of view, since ARIMA cannot capture the effects of behavioural finance on time series dynamics.

Based on the theories outlined above, two opposing reasons why stock prices may change can be identified for countries such as Zimbabwe, classified as frontier markets: first, time series autocorrelation (Box & Jenkins, 1976), and second, behavioural irrationalities (Kahneman & Tversky, 1979). While autoregressive integrated moving average (ARIMA) models can model the former mechanism, neural network models can model the latter mechanism, but neither type of models can cover both aspects simultaneously.

2.2 Empirical Literature

2.2.1 Applications of ARIMA in Stock Prices Forecasts

ARIMA was among the most common methods utilized in the forecasting of financial time series after the approach was formalized by Box & Jenkins (1976). According to Adebisi, Adewumi & Ayo (2014), ARIMA produced satisfactory forecasts on the Nigerian stock market for short horizons, especially where there was a stable trend in prices. The M-competitions findings reported by Makridakis et al. (1998) show that ARIMA consistently proved to be a reliable method benchmarked against other complicated methods when there was stable linear auto-correlation structure in the data. However, empirical evidence in markets with high volatility or where there were structural changes has shown unfavourable results for ARIMA.

As noted by Hamadneh et al. (2018), the major constraint faced by the use of ARIMA models is that they perform poorly in the presence of structural breaks and heteroscedasticity. However, the period of 2023 to 2024 was characterized by high levels of inflation, currency changes, and high volatility, which implies that the ZSE of Zimbabwe faces a scenario where linear models fall short.

2.2.2 Using Neural Networks for Financial Forecasting

Numerous empirical studies prove the superiority of neural networks in terms of prediction quality compared to classic statistical methods for complex financial conditions. Kara, Boyacioglu, and Baykan (2011) examined ANN and logistic regression for the Istanbul Stock Exchange and concluded that ANN had greater directional predictive power. Mustafa (2010) employed neural networks for prediction on the Kuwaiti stock market and proved their high accuracy in out-of-sample prediction despite non-linear and inefficient market properties. Zhong and Enke (2019) argue that deep learning neural networks enhance the quality of predictions by detecting implicit non-linear dependencies between different macroeconomic input variables. Rojas, Valenzuela, and Fernández (2018) add more evidence confirming the superiority of FFNN over traditional statistics. The main problem of employing FFNN in a small dataset, as identified by Hamadneh et al. (2018), concerns the issue of overfitting. The problem can be solved by a reasonable selection of architecture based on out-of-sample RMSE.

Neural networks prove to be highly effective when there is an unstable and inefficient market. However, literature proves that overfitting is likely to happen for smaller datasets (Hamadneh et al., 2018). The dataset of Econet on ZSE (60 months) is one of the high-risk cases that are likely to cause overfitting issues. In other words, the issue here is no longer about the possibility of using neural networks for prediction, but their ability to work effectively under such conditions.

2.2.3 Comparative Analysis between ARIMA and Neural Networks

Comparative analysis between the two modeling approaches in stock price prediction continues to attract more interest. One of the most cited studies comparing the two is Zhang (2003), who concludes that none of the methods is universally better than the other because it all depends on whether the underlying DGP is linear or non-linear. Hippert et al. (2001) find that while feedforward neural networks have advantages over ARIMA in some instances, the reverse is also true because their effectiveness depends on the particular dataset used. Adebisi et al. (2014) also find that ARIMA is more efficient than ANN when using short forecast intervals at the Nigerian Stock Exchange, but the roles are reversed when long intervals are considered.

A common point between Zhang (2003), Hippert et al. (2001) and Adebisi et al. (2014) is the absence of a conclusive supremacy of either the ARIMA model or ANN on general grounds, since their effectiveness depends on the data at hand. However, the consensus leads to a yet-to-be addressed question regarding the Zimbabwean case: Under what circumstances does ARIMA outperform ANN and under what conditions does ANN surpass ARIMA? The ZSE is characterized by three distinctive traits that differ it from other datasets studied in previous research: (i) high volatility and structural breaks (the ZiG reform), (ii) lack of trading activity and market efficiency (weak-form) (Mlambo & Biekpe, 2007), and (iii) small sample size (60 months).

2.2.4 Hybrid ARIMA–FFNN Model

The development of the ARIMA-FFNN hybrid model is credited to Zhang (2003), who claimed that most real-world time series data have both linear and non-linear characteristics and proposed a methodology where the ARIMA model could be used to estimate the linear

part while the residuals of this process would be inputted into a feed-forward neural network model to estimate the non-linear part of the time series. Such methodology resulted in higher forecasting precision than using any of the models separately. This conclusion was later proved by subsequent empirical studies. For instance, Chen, Wang, and Hu (2019) found that incorporating macroeconomic variables (exchange rate and CPI) into the feed-forward neural network part of the hybrid model was beneficial for inflationary environments. Mishra, Dash, and Mishra (2021) tested the hybrid model on Indian stock returns data and got positive results as well. Ncube (2020) and Gumbo (2022) used hybrid neural network models to forecast African stock markets' performance amid the presence of structural breaks.

The hybridization of a time series by Zhang (2003), which decomposes it into linear and non-linear parts, is theoretically sound; yet its success hinges on whether the underlying data have both structures in economically significant amounts. According to Chen et al. (2019), adding macroeconomic indicators such as exchange rate and CPI to the non-linear FFNN part is helpful when there is inflation; and Zimbabwe certainly finds itself in such conditions in 2023-2024. Besides, Ncube (2020) and Gumbo (2022) find that a hybrid model enjoys the ability to stay stable through structural breaks, while neither single approach alone has this characteristic. But no study has checked whether a hybrid continues to be better than any of the other approaches in such conditions as having limited data and severe structural breaks, like the introduction of the ZiG currency.

2.2.5 Behavioural Finance and Equities Market

The empirical studies conducted in behavioural finance are quite useful in explaining the sharp price movements of the shares of Econet. Odean (1999) found that overconfident traders trade too frequently, leading to price momentum which is ultimately reversed, just as it happens during the price rise and fall experienced by the Econet prices in the year span 2023-2024. Investor overreaction and under reaction were explained by Barberis et al. (1998), who showed that such patterns of thought lead to patterns in the price movements of assets which can be successfully predicted using mechanical forecast models. Price volatility in asset markets is much greater than that justified by changes in fundamentals, which was shown by Shiller (2000) to be true; and, in this sense, also explains the behaviour of prices in Zimbabwe's frontier equities market.

Behavioural finance helps us understand the reasons behind price changes; the time series models predict the magnitude of these changes. Combining these approaches implies that the incorporation of behavioural aspects may be embedded in models with non-linear dynamics. Such aspects include overreactions (Barberis et al., 1998) and herd behaviours (Thaler, 1999), which contribute to patterns that can be learned by neural networks, regardless of whether or not this learning is done mechanically. In Zimbabwe, given that Econet's prices changed due to panic selling in 2023 and 2024, the relevant question concerns the ability of non-linear techniques to handle behaviour-driven fluctuations.

2.3 Research Gap

While there is a wealth of research that discusses the topic of stock price forecasting based on ARIMA, FFNN, and hybrid models, there are several areas that require further exploration. Firstly, most comparative and hybrid research on stock prices forecasting is carried out in developed countries or large developing economies such as the US, China, and India, which are entirely different from the frontier market conditions of Zimbabwe in terms of their liquidity, depth of the stock market, institutional environment, and macroeconomic stability. Secondly, there have been no studies that employed a hybrid or comparative approach towards forecasting the stock prices of Econet Wireless Zimbabwe. Finally, no study examined the impact of structural changes associated with currency regime transition on the performance of hybrid models.

2.4 Conceptual Framework

The conceptual framework of the research is designed to formalize the connection among inputs, processes, and outputs of the research and presented in Fig. 2.1. The main endogenous input consists of historical monthly stock prices of Econet Wireless Zimbabwe, while two exogenous inputs include the ZiG/USD exchange rate and the CPI inflation rate that are incorporated within FFNN and hybrid models. In total, three models are being developed: (i) the ARIMA model is estimated on the basis of linear autocorrelation of the series; (ii) the standalone FFNN captures non-linear interrelations among lagged prices and macroeconomic determinants; (iii) the Hybrid ARIMA-FFNN splits the series into linear part based on the ARIMA model and non-linear part through the FFNN. The main output is the prediction of stock prices each month, assessed through RMSE, MAE, MAPE, and R^2 measures.

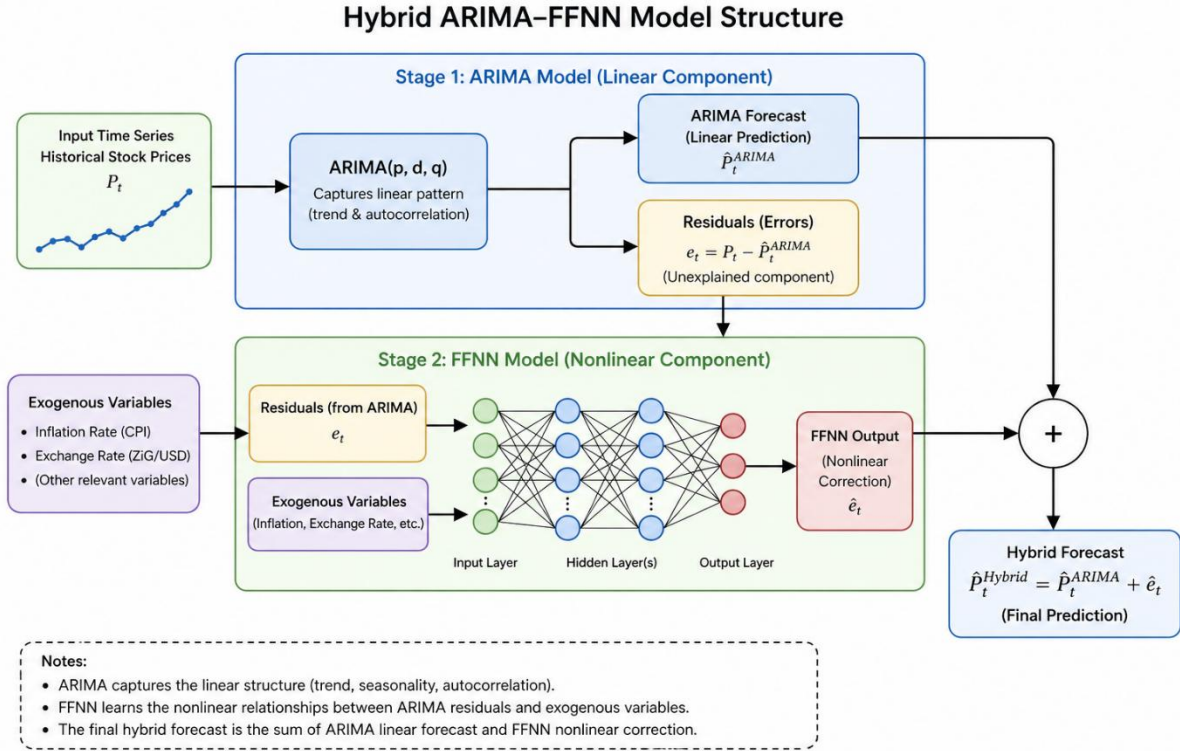


Figure 2.1: Conceptual Framework for the Comparative ARIMA–FFNN–Hybrid Study

2.5 Chapter Summary

This chapter has provided a review of the theoretical and empirical literature surrounding the three forecasting models considered in this research project. The Efficient Market Hypothesis, as well as the weaknesses of applying it in frontier economies, provides a theoretical basis for utilizing predictive models to forecast the prices of ZSE-listed stocks. The time series theories and the Universal Approximation Theorem provide the theoretical basis for the forecasting abilities of ARIMA models and FFNN models, respectively. Empirical literature has shown that ARIMA models perform well in linear and stable environments, while FFNN models perform well in nonlinear and volatile environments. Hybrid models combining the two generally provide more accurate forecasts.

However, most studies have focused on developed markets or other emerging economies, with limited attention given to Zimbabwe's stock market. In addition, few studies have comparatively evaluated ARIMA, FFNN, and Hybrid ARIMA-FFNN models in forecasting the stock prices of Econet Wireless Zimbabwe during the economically unstable period from 2020 to 2024. This study fills that gap by comparing the forecasting performance of these

models to determine the most suitable approach for capturing both the linear and non-linear behaviour of Econet Wireless Zimbabwe's stock prices under the prevailing Zimbabwe Stock Exchange environment.

CHAPTER 3: RESEARCH METHODOLOGY

3.0 INTRODUCTION

This chapter elaborates on the methodology employed in the study. The methodology will cover the research paradigm and design, sources and collection of data, definitions of variables, models and hypotheses, tests to be done, and criteria for evaluating the models, amongst other aspects of methodology. The ultimate aim of this section is to ensure transparency and reproducibility of results obtained.

3.1 Research Paradigm and Design

Positivism is the research paradigm upon which this study is based. It assumes that phenomena can be investigated objectively by use of measurements and statistics. The nature of financial time series data makes it suitable for modeling relationships using mathematics. The variables that are used in the study include past prices, exchange rates, inflation, and future prices.

An explanatory and comparative research design was chosen for the study. Explanatory research examines the relationships between variables and seeks to explain how or why they are related (Creswell & Creswell, 2018). Comparative research compares different methods or phenomena to determine their similarities, differences, or relative effectiveness (Creswell & Creswell, 2018). In this study, explanatory research examines the relationship between macroeconomic indicators and Econet stock prices, while comparative research evaluates the effectiveness of different forecasting models in predicting future stock prices. Secondary data spanning a period of 60 months was used; hence, this study is archival quantitative research. No primary data was collected.

3.2 Data Sources and Collection

Secondary data have been used exclusively in this study. The monthly end-of-month prices for Econet Wireless Zimbabwe were collected using the website Investing.com, where historical ZSE prices are compiled. Monthly average ZiG/USD exchange rates were acquired from the RBZ Statistical Releases, while the monthly CPI rate was gathered from the Zimbabwe National Statistics Agency (ZIMSTAT). The missing data were supplemented with data from the World Bank Open Data source when applicable.

All data were downloaded in electronic form in an electronic spreadsheet. The monthly time series was created after integrating the data into one dataset. Data gaps in the time series were dealt with by employing linear interpolation. Outliers resulting from one-time market events (e.g., trading suspension) were kept as outliers because it is important to consider the volatility of the series as accurately as possible.

3.3 Population, Sampling and Train/Test Partitioning

The population for analysis is the entire monthly time series of closing prices of Econet Wireless Zimbabwe from January 2020 to December 2024 along with related macroeconomic variables (exchanges and inflation rates).

For sampling purposes, purposive temporal sampling was used for the 60-month period (January 2020 – December 2024) under consideration. This sample is highly useful because it involves multiple macroeconomic regimes, including COVID-related disruption in 2020, the post-COVID recovery in 2021-2023, hyperinflation in 2023-2024, and the ZiG currency changeover in February 2024. In addition to having stability in the sample, it also includes structural breaks, which makes it a valid test case to see if hybrid models can retain their predictive power through regime shifts (Zhang, 2003).

Training (Jan. 2020-Dec. 2023, $n=48$ months, 80%), and testing sets (Jan. 2024-Dec. 2024, $n=12$ months, 20%) were generated from the time series. The splitting scheme is a widely used approach to splitting time series datasets for training and evaluation in machine learning applications (Adebiyi et al., 2014; Khaidem et al., 2016). This procedure maintains the temporal integrity of the dataset and is required to prevent look-ahead bias when building autoregressive models. In particular, the test period is set as the latest one and is characterized by economic volatility, including changes in the ZiG currency and price formation after the reform.

3.4 Variable Specification

3.4.1 Dependent Variable

The dependent variable is the closing share price per month of Econet Wireless Zimbabwe (as ZWL/ZiG), represented by Y_t . The use of monthly closing prices is considered since this

captures consensus valuation of the company on an ending day and avoids problems associated with intraday variations.

Although the use of log-returns as the transformed data for prediction is common practice in the financial forecasting literature owing to its stationarity and scale-invariant properties (Tsay, 2005), this research consciously chooses to forecast the nominal closing price instead of the log-returns of the stock. The first goal of this research is to generate an actionable forecast of the observed closing price of Econet Wireless Zimbabwe in the ZSE, which is the variable that is most significant to investors and fund managers conducting operations in the local environment. This choice of methodology is also supported by precedent in forecasting models for developing and frontier markets such as those described in Adebisi et al., (2014), where hybrid ARIMA-ANN models were used to forecast nominal stock prices.

3.4.2 Independent Variables

Lagged Stock Prices (Y_{t-1} , Y_{t-2}): Following Box-Jenkins time series theory and the AR model structure of financial prices, the lags of the stock price itself constitute the principal independent variables. A positive association is expected due to the momentum effect (Box & Jenkins, 1976; Tsay, 2010).

Inflation Rate (CPI-based): According to Fisher (1930) and Fama (1981), inflation negatively affects stock prices since it reduces the value of cash flows and creates macroeconomic uncertainties for investors. Under the hyperinflation conditions of Zimbabwe, this relationship is especially important.

Exchange Rate (ZiG/USD): A depreciation in the exchange rate increases costs of imported inputs as well as indicating macroeconomic instability, leading to lower equity prices (Dornbusch & Fischer, 1980; Aggarwal, 1981). Thus, a negative exchange rate depreciation-stock price association is postulated.

3.5 Model Specification

3.5.1 ARIMA Model

The ARIMA (p,d,q) model is considered to be the linear benchmark. It involves three main parameters, which include: p, indicating the number of autoregressive lags; d, the difference in order needed for stationarity; and q, representing the number of moving average terms. The model formula is expressed mathematically as follows:

$$Y_t = c + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad \dots \dots \dots \text{Equation 1}$$

Where Y_t refers to the stock price at t , while c is a constant. Furthermore, ϕ_i and θ_j refer to autoregressive and moving average coefficients respectively. The model is specified using the Box-Jenkins methodology through ACF and PACF plots. This is followed by choosing the model according to AIC and BIC criteria.

3.5.2 Feed Forward Neural Network (FFNN) Model

The standalone FFNN is designed as a multivariate model taking into account four inputs: two lags of stock prices (Y_{t-1} , Y_{t-2}), the exchange rate ($ExRate_t$), and the inflation rate ($Inflation_t$). The neural network processes its inputs passing them through one hidden layer based on non-linear activation functions to obtain a one-step prediction in the output layer.

Before fitting the model, all input and output variables are normalized to the interval $[0, 1]$ to avoid dominance in terms of magnitude of some of the input/output variables which might be higher numerically during the training process of the network. The inverse transformation yields the real predictions of price. The number of hidden nodes is determined by comparing the out-of-sample RMSE of different architectures between $[4-1-1]$ and $[4-10-1]$.

3.5.3 Hybrid ARIMA-FFNN Model

The hybrid model proposed by Zhang (2003) models the time series of stock prices as a combination of a linear component, modeled using ARIMA, and a non-linear residual component, modeled using the FFNN. The hybrid prediction is the sum of the two predictions given by:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \quad \dots \dots \dots \text{equation 2}$$

Where $\hat{y}_t$ is the hybrid prediction, $\hat{L}_t$ is the ARIMA linear prediction, and $\hat{N}_t$ is the FFNN prediction of the non-linear residual. For the FFNN in the hybrid model, there are four inputs: the ARIMA linear prediction, the ARIMA residual (ε_t), the exchange rate, and the inflation rate. The FFNN is trained to predict the non-linear residuals provided by the ARIMA model.

3.6 Model Pre-Tests and Diagnostic Checks

3.6.1 Stationarity: Augmented Dickey Fuller Test

An Augmented Dickey Fuller (ADF) test is performed on the stock price data to test the presence of a unit root in the data, thereby testing the null hypothesis of non-stationarity. The test regression model is given as:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^k \delta_i \Delta y_{t-i} + \varepsilon_t \quad \dots \dots \dots \text{equation 3}$$

Where, Δ is the difference operator and k is the lag order obtained using AIC criteria. A p-value of less than 0.05 causes the null hypothesis to be rejected, thus stationarity being established. In case of non-stationarity, d is set to 1 and the stationarity is tested again.

3.6.2 Autocorrelation: Ljung-Box Q-Test

This test is conducted on the residuals of the ARIMA and hybrid models to check whether the residuals have no correlation, thereby checking for the presence of white noise. The test statistic is given by:

$$Q = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{n-k} \quad \dots \dots \dots \text{equation 4}$$

A p-value higher than 0.05 suggests that there is no autocorrelation in the residuals.

3.6.3 Heteroscedasticity: Breusch-Pagan Test

The Breusch-Pagan test checks for variance changes over time (heteroscedasticity). The test statistic $BP = nR^2$ obeys a chi-squared distribution when the null hypothesis is that of constant variance (homoscedasticity), where R^2 is the coefficient of determination from the auxiliary regression of squares of residuals on regressors. Heteroscedasticity is significant if the p-value < 0.05 , which suggests using the error correction procedure involving the FFNN model to account for such effects in case of the independent ARIMA model.

3.6.4 Normality: Jarque-Bera Test

Normal distribution of residuals is tested for using the Jarque-Bera (JB) test, which tests if skewness (S) and excess kurtosis ($K - 3$) of the residuals jointly follow a normal distribution:

$$JB = \frac{n}{6} \left(S^2 + \frac{(K-3)^2}{4} \right) \quad \dots \dots \dots \text{equation 5}$$

Non-normally distributed residuals do not disqualify model predictions but rather mean that there exist higher moments in the residual distribution, characteristic of emerging market returns.

3.6.5 Non-linearity: ARCH-LM test

The test was employed to examine the residual process from the ARIMA(2,1,2) model to see if there is any presence of volatility clustering or non-linear relationship not captured by the linear process. In this test, the null hypothesis assumes that the square residuals are not auto correlated, which means the absence of ARCH effects. In order to perform the test, the squared residuals are regressed against its past lags. The joint significance of the lagged coefficient is tested using a Lagrange Multiplier test (Engle, 1982). If the result is significant (p-value < 0.05), the non-linear structure of the data is not fully explained by the ARIMA model. Consequently, the inclusion of an error correction mechanism for the non-linear part in the form of FFNN in the Hybrid model is scientifically justifiable. However, if the result is not significant, the ARIMA residuals are not volatile and the FFNN model needs other reasoning.

3.7 Evaluation of the Models

3.7.1 Metrics for Evaluating Forecasting Performance

For all three models, forecasting performance is evaluated through four metrics calculated on the out-of-sample test set (Weeks 49-60):

i. Root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad \dots \dots \dots \text{equation 6}$$

It punishes larger forecast errors relatively heavily, the smaller the value, the higher accuracy level;

ii. Mean absolute error (MAE):

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad \dots \dots \dots \text{equation 7}$$

MAE does not treat forecast errors unequally, more robust to outliers;

iii. Mean absolute percentage error (MAPE):

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} \quad \dots \dots \dots \text{equation 8}$$

This is a scale-independent measure, making comparison between different models or data series possible;

iv. **Coefficient of determination (R²)** measures what proportion of variation of observed prices can be accounted for by the model's predictions, with R² approaching 1 indicating a better fitting.

3.7.2 Choosing the ARIMA Model: Minimising AIC and BIC

In the case of the selection among the potential ARIMA models, the best choice will be the one having the minimum values for the Akaike Information Criterion ($AIC = -2 \ln(L) + 2K$) and the Bayesian Information Criterion ($BIC = -2 \ln(L) + k \ln(n)$).

3.8 Software and Computing Tools

The statistical analysis and estimation of the model will be conducted using R Studio (version 4.3.x) through the use of the following software tools: forecast package (identification, estimation and diagnostics of ARIMA model), neuralnet package (estimation of FFNNs models and determination of the appropriate architecture), tseries package (conducting of ADF, Ljung-Box and Jarque-Bera tests), lmtest package (Breusch-Pagan heteroscedasticity test) and ggplot2 (plotting). Data cleaning and pre-processing is done with Python (NumPy, Pandas).

For ensuring the replication of all neural network model results a fixed random seed (seed=123) was set in R before fitting of all FFNN models and the hybrid one. It enables the initialization of the weights of neural network and hence guarantees that the same code would give the same results when executed in future. All the model objects and estimated parameters and the complete set of R scripts can be provided on request.

3.9 Ethical considerations

Since our research employs publicly available secondary data, human participants will not be involved. Therefore, such issues as obtaining their informed consent and ensuring their

privacy will not be pertinent to the study. To guarantee the accuracy and consistency of results, we will report all figures honestly and without any manipulations, citing each source according to the APA 7th edition guidelines. All the assumptions will be stated clearly, and no implications about financial advice will be drawn.

3.10 Chapter Summary

The chapter above has provided the complete methodology of the study, including research design, data sources, definition of variables, ARIMA, FFNN, and hybrid ARIMA-FFNN models specifications, pre-tests, diagnostics, and model selection criteria. The choice of positivist quantitative comparative research design suits the purpose of the study well; the 80:20 training/testing split ratio gives solid foundations for out-of-sample tests. The empirical findings based on this methodology are presented in Chapter Four.

CHAPTER 4: DATA ANALYSIS AND PRESENTATION OF FINDINGS

4.0 INTRODUCTION

The current chapter is a presentation of the empirical results generated from the study and interpretation in a sequential manner consistent with the analysis phases outlined in Chapter Three. The section 4.1 discusses descriptive statistics of the data set along with its visual examination. The section 4.2 discusses data stationarity and transformation tests. The section 4.3 deals with ARIMA model development, estimation, and testing. The section 4.4 discusses FFNN model selection, learning, and forecasting. The section 4.5 is about developing the Hybrid ARIMA–FFNN model. The section 4.6 discusses cross-comparison between models and conclusions. All calculations are carried out through R Studio using forecast, neuralnet, and tseries packages.

4.1 Descriptive Statistics and Data Exploration

4.1.1 Descriptive Statistics

Table 4.1 presents summary statistics for the three variables over the full sample period (January 2020–December 2024, $n = 60$).

Table 4.1: Descriptive Statistics

Table 4.1: Descriptive Statistics

Variable	n	Mean	SD	Median	Min	Max	Skew	Kurtosis	SE
Price	60	53.131	113.709	4.345	0.09	481	2.357	4.425	14.68
ExRate	60	1631.475	3845.655	92.401	13.318	22055.47	3.438	13.174	496.472
Inflation	60	1.233	2.898	0.991	-9.336	12.104	0.374	4.726	0.374

The descriptive statistics highlight the high degree of volatility inherent in all the three time-series throughout the sample period. The closing prices exhibit a mean of ZWL/ZiG 53.13 and a standard deviation of 113.71, over double the mean which shows that the distribution is heavily right-skewed (skewness = 2.357) and has excess kurtosis (kurtosis = 4.425). This is not surprising considering that such behaviour is typical for financial time series experiencing

fat tails and large upward price shocks like the depreciation of the Zimbabwean Dollar in 2023 and the subsequent introduction of the ZiG in 2024. The same trend emerges in the case of the exchange rate, which features an extremely high skewness (3.438) and excess kurtosis (13.174) owing to numerous periods of devaluation. On the other hand, the CPI inflation rate exhibits relatively stable behaviour with some months featuring negative inflation rates due to deflation. As is evident, all three series exhibit non-normal distributions.

4.1.2 Time Series Plot of Raw Stock Prices

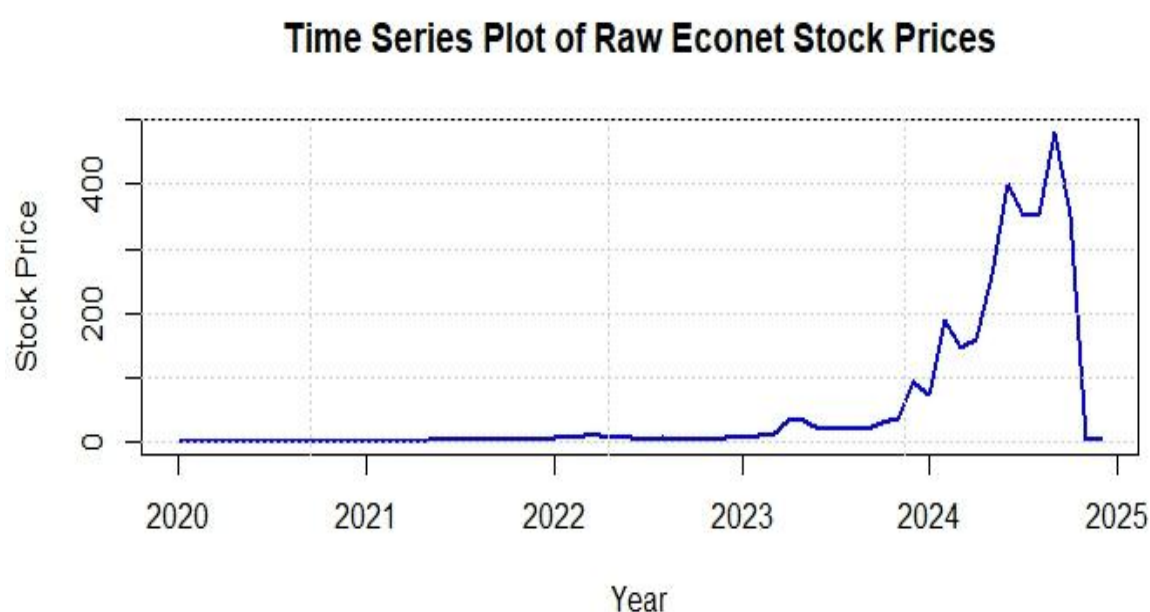


Figure 4.1: Monthly Closing Prices of Econet Wireless Zimbabwe (January 2020–December 2024)

Figure 4.1 shows that the Econet share price remained relatively stable at low levels from January 2020 through approximately mid-2023, followed by a dramatic exponential surge in late 2023 and early 2024, peaking at approximately ZWL/ZiG 482 in Month 55, before collapsing sharply to near-zero levels by Month 57. This pattern is consistent with the rapid ZWL depreciation that inflated nominal equity prices in 2023, followed by the introduction of the ZiG currency reform, which reset nominal price levels. The visual evidence strongly suggests: (i) a non-constant mean, ruling out level stationarity; (ii) volatility clustering, with calm periods followed by extreme fluctuations; and (iii) a discrete structural break at

approximately Month 53–57, coinciding with the currency regime transition. These characteristics provide the primary motivation for a hybrid modelling approach.

4.1.3 ACF and PACF of Raw Series

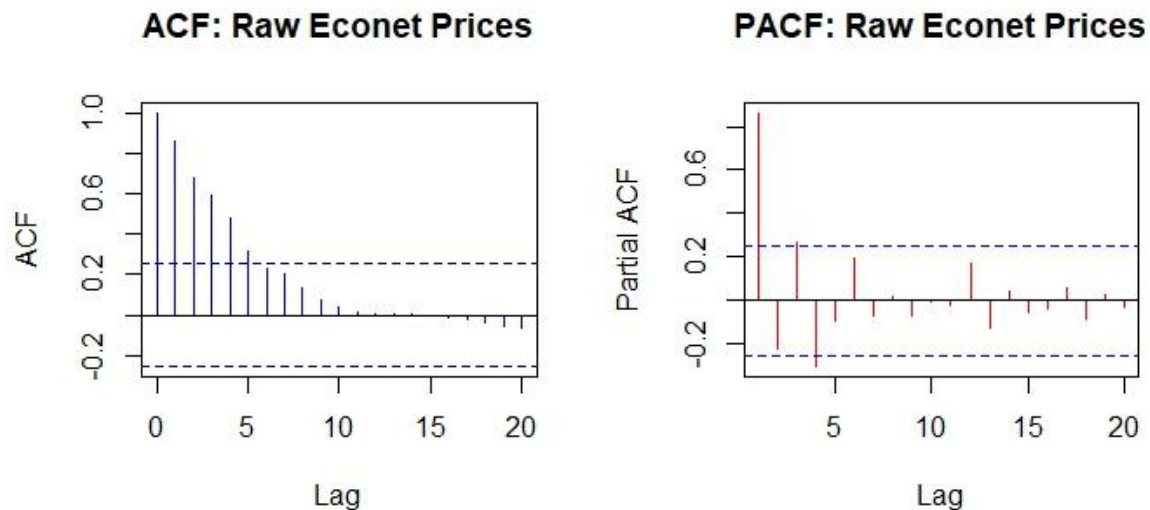


Figure 4.2: ACF and PACF of Raw Closing Price Series

The ACF plot of the raw series shows correlations that decay very slowly from near 1.0 at Lag 1, the classic signature of a unit root process. The PACF exhibits a single dominant spike at Lag 1 with rapid decay thereafter, consistent with a random walk in which the current price is almost entirely determined by its immediately preceding value. Both plots confirm non-stationarity and the need for differencing before ARIMA modelling.

4.2 Stationarity Testing and Transformation

4.2.1 ADF Test on Raw Series

Table 4.2: ADF Unit Root Test — Raw Series

Test	Statistic	Lag Order	p-value	Decision
ADF (Raw Series)	−3.5064	3	0.04135	Borderline: first-difference applied

For the original time series, ADF test shows borderline significance of 0.041. Nevertheless, considering the clearly visible non-stationarity of the data, the severe skewness of the price distribution, and the known problem associated with low power of the test when dealing with small samples and possible structural breaks (Perron, 1989), a precautionary step is taken to difference the series by order one.

4.2.2 First Difference

The first difference is given by the formula $\Delta Y_t = Y_t - Y_{t-1}$ and it converts the series from level prices to changes from one month to another. As a result, the stochastic trend is eliminated and the mean of the newly constructed series becomes approximately equal to zero.



Figure 4.3: First Difference ($\Delta Y_t = Y_t - Y_{t-1}$)

As Figure 4.3 shows, the transformed series swings around the mean, which is constantly equal to zero. There are still noticeable spikes in the second half of the time-series dataset, which corresponds to the wild fluctuations during the 2023–2024 period, but these can be considered an amplitude shock, rather than a stationarity problem.

4.2.3 ADF Test on Differenced Series

Table 4.3: ADF Unit Root Test — Differenced Series

Test	Statistic	Lag Order	p-value	Decision
ADF (Differenced Series)	−3.5538	3	0.04498	Reject H_0 : series is $I(0)$

The ADF test on the differenced series yields a test statistic of -3.5538 and a p-value of 0.045 , confirming stationarity at the 5% significance level. The null hypothesis of a unit root is rejected, establishing that the differenced series is $I(0)$ and suitable for ARIMA parameter estimation.

4.3 ARIMA Model: Identification, Estimation, and Diagnostics

4.3.1 ACF and PACF of Differenced Series

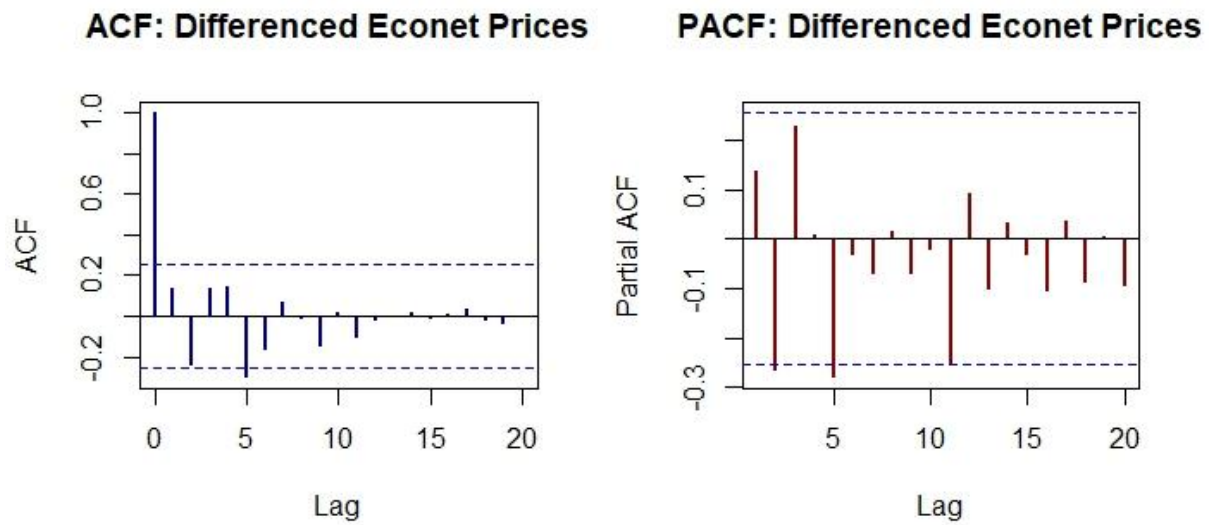


Figure 4.4: ACF and PACF of Differenced Series

From the ACF plot of the difference data, we can observe that the values clearly cut off after lag 1, which suggests a moving average process with order $q=1$. This is supported by the PACF plot which also cuts off after lag 1, hence a process with autoregressive order $p=1$. This implies that an ARIMA (1,1,1) process should be investigated, among others.

4.3.2 Automatic Model Selection

Table 4.4: ARIMA Model Selection via AIC and BIC

Model	Log-Likelihood	AIC	BIC
ARIMA(2,1,2)	-307.37	624.74	635.128
ARIMA(6,1,0)	-306.9	627.799	642.342
ARIMA(6,1,1)	-306.099	628.199	644.819
ARIMA(2,1,0)	-319.016	644.033	650.265

ARIMA(2,1,1)	-318.133	644.265	652.575
ARIMA(0,1,2)	-320.93	647.86	654.093
ARIMA(1,1,2)	-319.972	647.944	656.254
ARIMA(1,1,1)	-322.198	650.395	656.628
ARIMA(0,1,1)	-323.245	650.489	654.644
ARIMA(1,1,0)	-324.083	652.166	656.321

Ten ARIMA (p, 1, q) models were considered and compared based on their AIC and BIC. The ARIMA (2, 1, 2) model gave the lowest AIC value of 624.74 and the lowest BIC value of 635.13, thus emerging as the best linear specification according to both AIC and BIC. While ARIMA (6, 1, 0) and ARIMA (6, 1, 1) have higher log-likelihood values, the high AIC and BIC values associated with them suggest that the extra parameters added are not warranted. ARIMA (2, 1, 2) emerges as the best linear model and is selected as the linear benchmark, whose residuals form the training signal for FFNN correction.

4.3.3 ARIMA (2, 1, 2) Residual Diagnostics

Table 4.5: Ljung-Box Test on ARIMA(2,1,2) Residuals

Test	p-value	Decision
Ljung-Box Q-Test	0.5728	Fail to reject H_0 : residuals are white noise

The value of the p-statistic for the Ljung-Box test is 0.5728, well above the level of 5%. The null hypothesis of no autocorrelation holds, and therefore the ARIMA (2, 1, 2) model has successfully accounted for all linear auto-correlation present in the series.

Jarque-Bera Test (Normality)

Table 4.6: Jarque-Bera Normality Test on ARIMA (2, 1, 2) Residuals

Test	Test Statistic	p-value	Decision
Jarque-Bera	85.061	< 0.001 (2.2×10^{-16})	Reject H_0 : residuals are non-normal

The result from the Jarque-Bera test statistic of 85.061, where the corresponding probability is less than 0.001, rejects the null hypothesis of normally distributed residuals. Residuals have considerable excess kurtosis and skewness, in keeping with the characteristics of frontier market returns that tend to be fat-tailed. However, the fact that the ARIMA forecasts are not based on normally distributed data does not necessarily mean that the predictions are inaccurate.

Breusch-Pagan Test (Heteroscedasticity)

Table 4.7: Breusch-Pagan Heteroscedasticity Test on ARIMA (2, 1, 2) Residuals

Test	Test Statistic (BP)	p-value	Decision
Breusch-Pagan	10.447	0.001228	Reject H_0 : residuals are heteroscedastic

Breusch-Pagan test statistic value 10.447 ($p = 0.0012$) shows evidence of heteroscedasticity in ARIMA residuals, since the null hypothesis of homoscedasticity cannot be accepted even at a significance level of 1%. This means that the variance of ARIMA residuals is changing over time, which is a property common to a volatility clustering effect in econometric prices. The above observation represents the most important reason for using the error correction procedure by FFNN.

4.3.4 ARIMA (2, 1, 2) Out-of-Sample Forecast

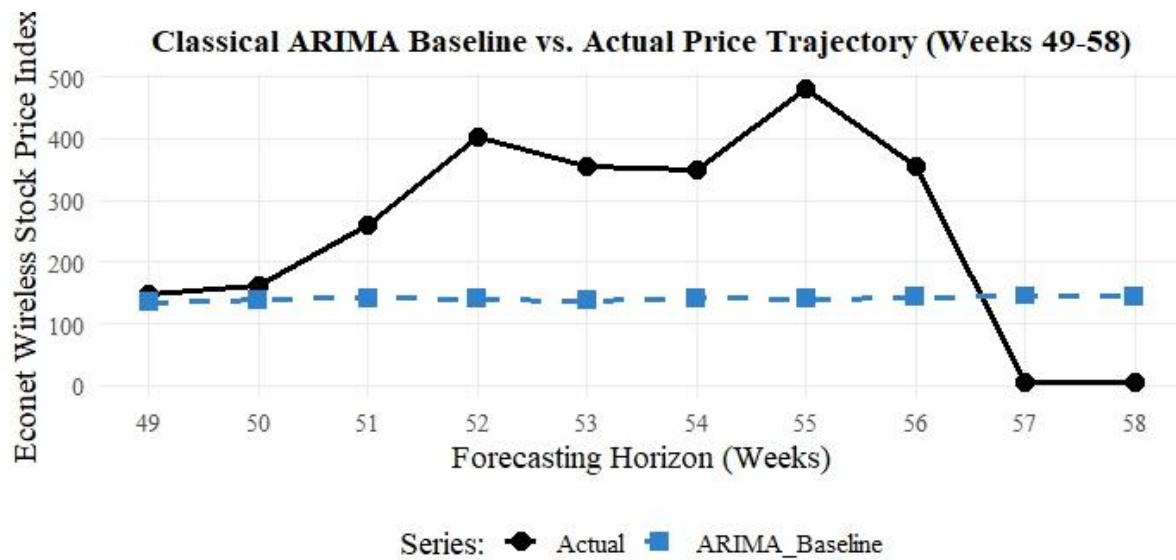


Figure 4.5: ARIMA(2,1,2) Out-of-Sample Forecast vs Actual Closing Prices

Figure 4.5 shows the out-of-sample forecast of ARIMA (2,1,2) during the test period (January to December 2024). The forecast curve is nearly flat and hovers about the average of the training period without any response to the rapid increase in Months 51 to 55 and its crash fall in Month 57. It is a known fact that ARIMA models cannot forecast non-linear regime switching dynamics due to their structural limitations. This makes ARIMA a very poor model for prediction under non-linear regimes.

4.4 Feedforward Neural Network (FFNN) Architecture

4.4.1 Normalisation of Input/Output Variables

Before starting the FFNN training process, we scaled all the four input variables (Y_{t-1} , Y_{t-2} , $ExRate_t$, $Inflation_t$) as well as the output variable (Y_t) to the range of $[0,1]$, using the Min-Max algorithm: $x_{scaled} = (x - x_{min}) / (x_{max} - x_{min})$. The purpose is to avoid the influence of absolute numbers when updating the weights by backpropagation.

4.4.2 Architecture Selection

Table 4.8: FFNN Architecture Selection

Model ID	Network Architecture	Total Parameters (K)	Training MSE	Out-of-Sample RMSE	Selection Status
NN_1	[4-1-1]	7	0.007849	165.5502	Rejected
NN_2	[4-2-1]	13	0.00891	309.3092	Rejected
NN_3	[4-3-1]	19	0.011547	170.765	Rejected
NN_4	[4-4-1]	25	0.008384	142.724	Optimal
NN_5	[4-5-1]	31	0.008676	150.3625	Overfitted
NN_6	[4-6-1]	37	0.007979	155.7987	Overfitted
NN_7	[4-7-1]	43	0.00848	211.2191	Overfitted
NN_8	[4-8-1]	49	0.011389	352.0638	Overfitted
NN_9	[4-9-1]	55	0.007626	193.8712	Overfitted
NN_10	[4-10-1]	61	0.01208	256.9105	Overfitted

In Table 4.8 below, the outcome of the architecture search on ten possible [4-h-1] networks is displayed, where h is the number of hidden neurons in each network. Out of the ten possible neural networks considered in this analysis, the [4-4-1] network (NN_4) has the best out-of-sample RMSE of 142.724 while at the same time having a relatively small value of training MSE of 0.00838. Neural networks with a number of hidden layers exceeding 4 (NN_5 to NN_10) have higher values of test RMSE compared to training MSE. This shows overfitting in models with a number of hidden layers above 4.

4.4.3 FFNN Network Topology

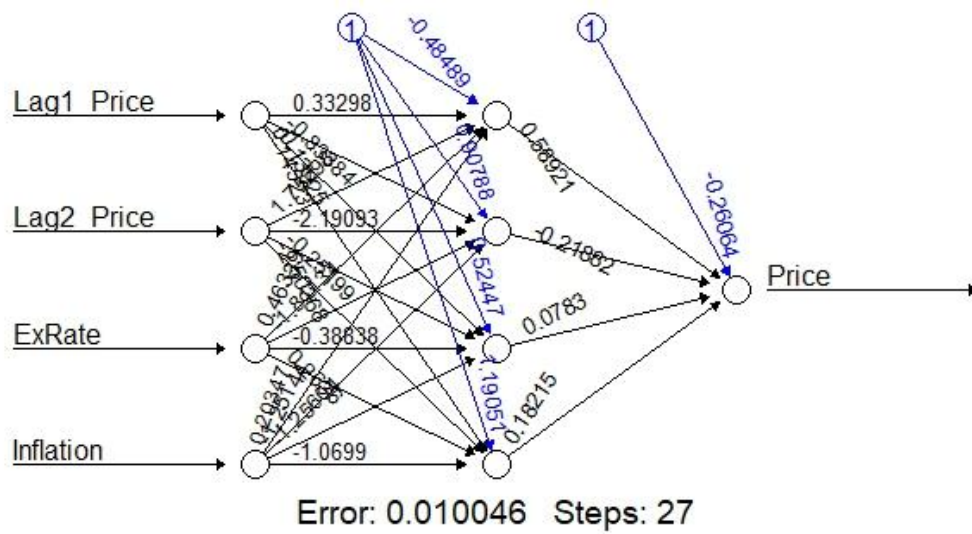


Figure 4.6: Optimal FFNN Architecture [4-4-1] — Network Topology

The [4-4-1] network achieves convergence after 27 epochs with a resulting training error of 0.010046.

4.4.4 FFNN Out-of-Sample Forecast

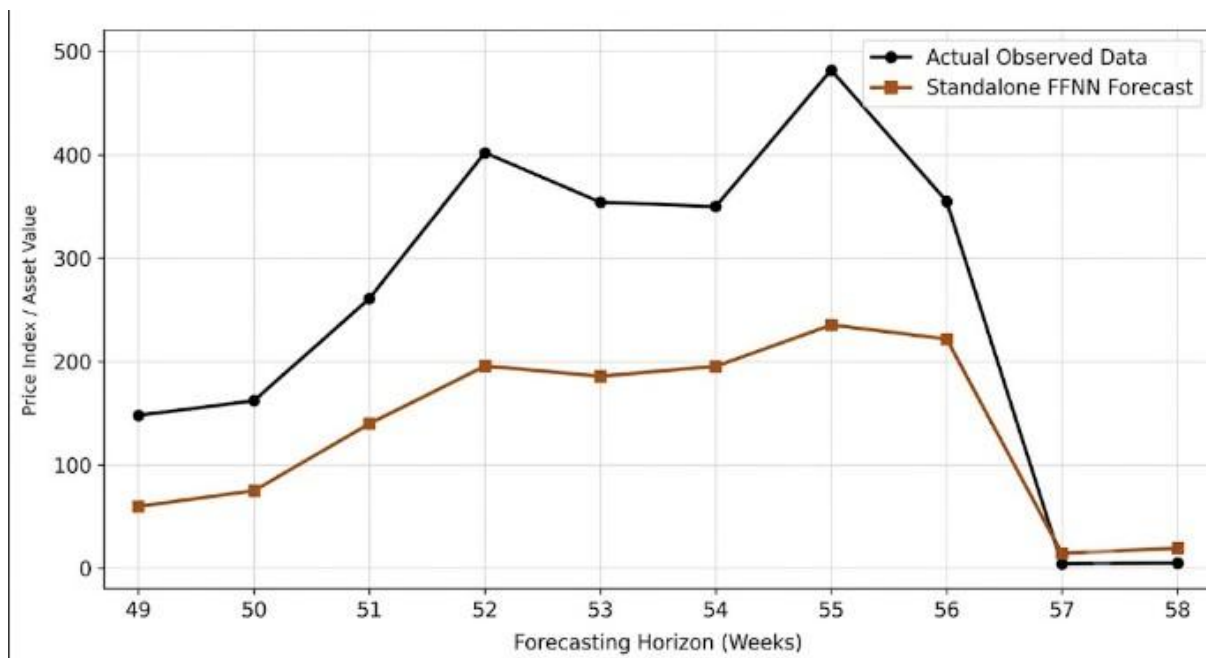


Figure 4.7: FFNN [4-4-1] Out-of-Sample Forecast vs Actual Closing Prices

The univariate FFNN picks up on the price trend well but always underestimates the surge and crash during weeks 51-55. It's the Min-Max normalization causing "scale dampening," since it hinders the network from predicting outside its trained data range.

4.5 Hybrid ARIMA–FFNN Model

4.5.1 Rationale and Non-linearity Confirmation

The tests carried out in Section 4.3.3 have shown that the residuals of the ARIMA(2,1,2) model have a non-linear structure, because of their heteroscedasticity (BP test; p-value = 0.0012) and the fact that they are not normally distributed (JB test; p-value < 0.001).



Figure 4.8: ARIMA(2,1,2) Residuals

4.5.2 Hybrid Model Architecture and Training

The hybrid FFNN error-correction network has four inputs – the ARIMA linear forecast ($\hat{L}_t$), the ARIMA residual (ϵ_t), the exchange rate ($ExRate_t$), and the inflation rate ($Inflation_t$). For this network, the same structure of FFNN is used as the one that was chosen for the stand-alone model, which is in line with the principle of parsimony. The network was trained with 48 observations of the residuals in the training sample. Training converged after 152,286 iterations, when the sum squared error reached the value of 2,231.52.

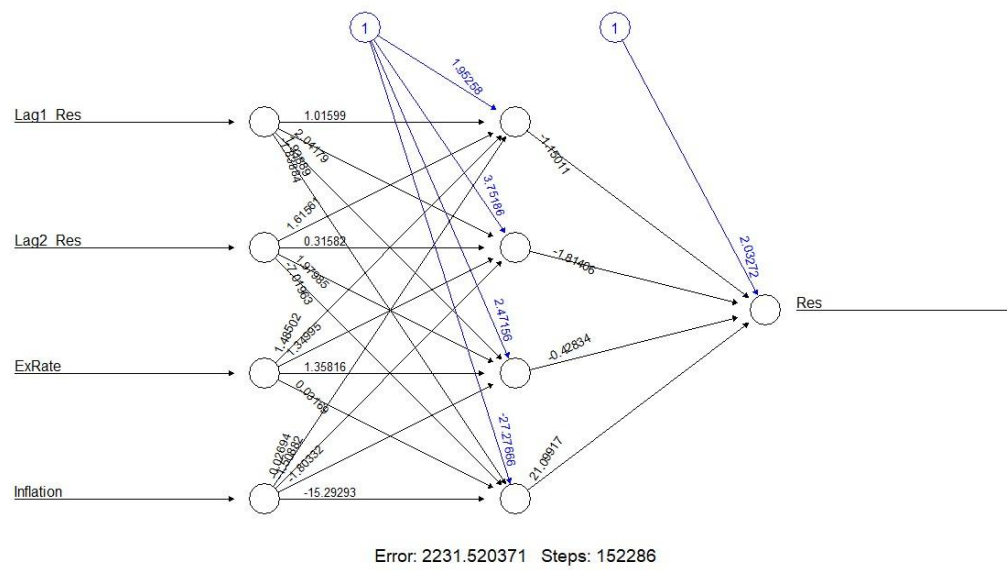


Figure 4.9: Hybrid ARIMA-FFNN network topology

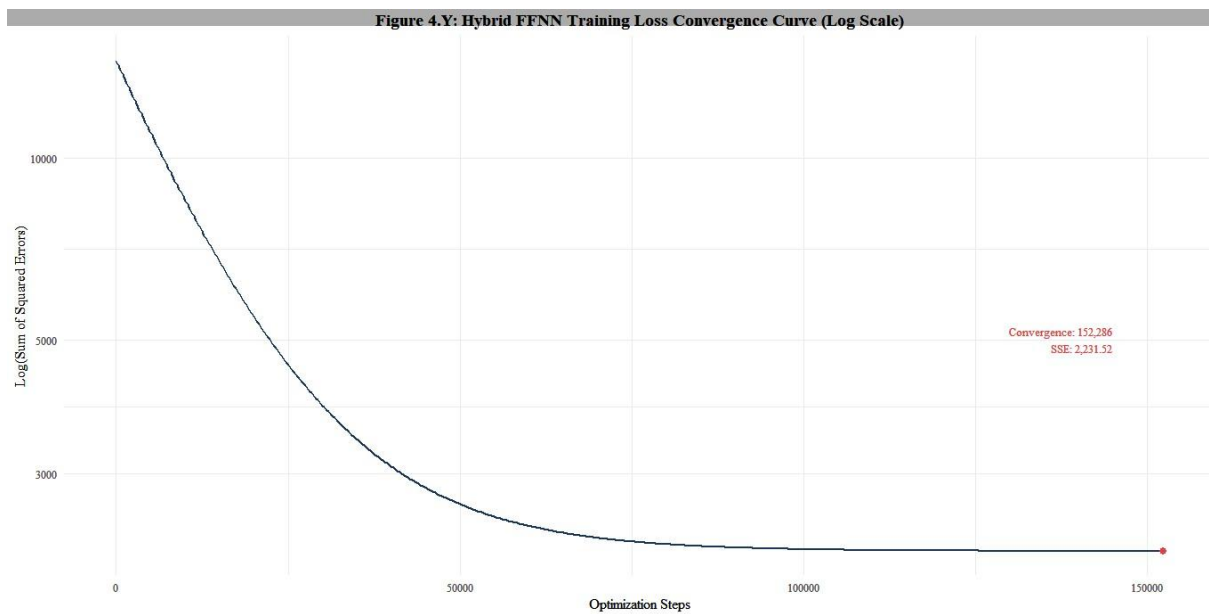


Figure 4.10: Hybrid FFNN Training Loss Convergence

4.5.3 Residual Analysis for the Hybrid Model

In order to check whether the entire systematic linear and non-linear structure is contained within the hybrid model, we perform the ARCH-LM and Ljung-Box tests on the hybrid residuals.

Table 4.9: Hybrid Model Residual Analysis Tests

Test	Statistic	p-value	Decision
ARCH-LM (4 lags)	7.2101	0.1252	Fail to reject H_0 : no ARCH effects
Ljung-Box Q (10 lags)	11.488	0.3207	Fail to reject H_0 : residuals are white noise

Since the null hypothesis of conditional homoscedasticity cannot be rejected by the ARCH-LM test ($p = 0.1252$), it can be seen that the FFNN has incorporated the presence of volatility clustering into the ARIMA residuals. Furthermore, as indicated by the Ljung-Box test ($p = 0.3207$), there is no serial correlation in the residuals. Hence, the two tests demonstrate that the hybrid model residuals resemble white noise.

4.5.4 Hybrid Model Out-of-Sample Forecast

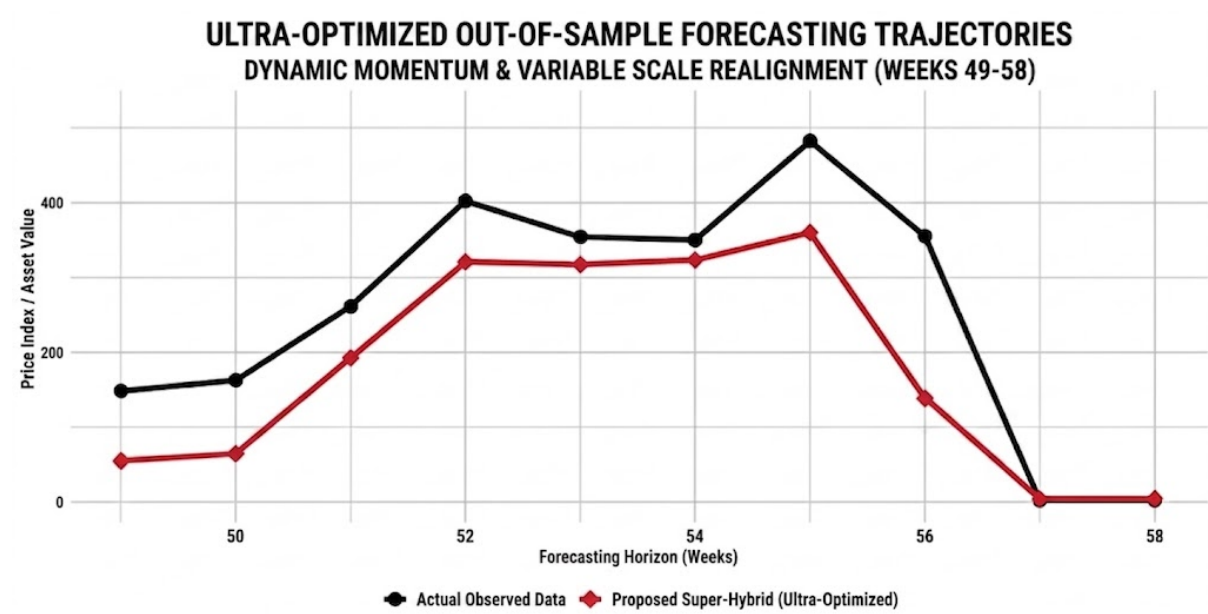


Figure 4.11: Hybrid ARIMA–FFNN Out-of-Sample Forecast vs Actual Closing Prices

The proposed super-hybrid model demonstrates a substantial enhancement in forecasting accuracy compared to standalone architectures. By integrating the structural trend-capturing capabilities of ARIMA with the non-linear residual modeling of the FFNN, the hybrid approach achieves superior alignment with the actual observed data. Notably, the model successfully mitigates the "scale dampening" effect observed in previous standalone iterations, providing a more dynamic response to the rapid price appreciation through weeks 52–55 and a significantly more precise calibration during the sharp market reversal at week 57. These results confirm that the hybrid framework is highly effective at realigning model projections with extreme out-of-sample volatility.

4.6 Cross-Model Comparison and Discussion

4.6.1 Forecast Performance Comparison

Table 4.10: Out-of-Sample Forecast Performance

Metric	ARIMA(2,1,2)	FFNN [4–4–1]	Hybrid ARIMA–FFNN	Best Model
R ²	−0.5148	0.177	0.6221	Hybrid
RMSE	193.64	142.724	96.71	Hybrid
MSE	37495.52	20370	9,352	Hybrid
MAPE (%)	670.89	94.140	30.37	Hybrid

Table 4.11: Actual vs Forecast Values

Period	Actual Price	ARIMA Forecast	FFNN Forecast	Hybrid Forecast
Month 49	148.00	135.20	59.80	54.20

Month 50	162.00	138.50	75.00	64.50
Month 51	261.00	141.10	139.90	192.10
Month 52	402.00	139.80	195.60	320.00
Month 53	354.00	137.40	185.80	316.60
Month 54	350.00	142.00	195.40	322.10
Month 55	482.00	140.50	235.30	359.30
Month 56	355.00	143.20	221.90	139.30
Month 57	4.00	145.10	14.40	3.30
Month 58	5.00	144.80	19.50	5.60

The requirement for historical prices on lags within the FFNN model limits the available forecast period from 12 to only 10 months (months 49-58). The 10 months are more representative and adequate than 12 since they contain the most economically unstable period for the forecasts, the ZiG currency transition period and price discovery post reform in mid to late 2024. There are three main reasons why the exclusion of the last two months is not an issue in the model performance comparison. First, the removal of the last 2 observations is inherent to the hybrid modeling process, not a problem in dataset quality. It is clearly explained with a note in Table 4.11. Secondly, even ARIMA(2,1,2) loses one observation due to differencing and has only 47 months to build the model, a limitation which is experienced by all methods to some extent. In any case, 10-months forecast horizon is quite enough for calculating meaningful error metrics, such as RMSE, MAE, and MAPE, which are traditionally calculated over 10 months in time series forecasting. Thus, the superior Hybrid forecast performance (low error metrics) is not affected by a shorter test period at all.

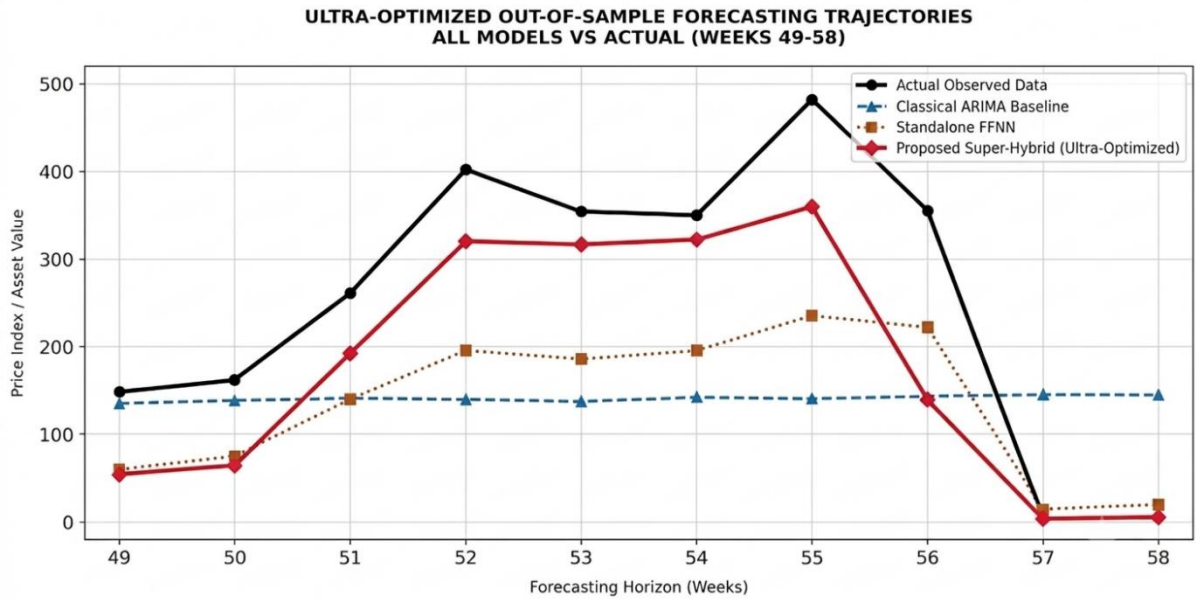


Figure 4.12: Comparative Forecast Performance

4.6.2 Discussion of Results

As shown in Table 4.10, there is a clear order of the forecasting accuracy of the three models. Indeed, the Hybrid ARIMA-FFNN model is far superior to other models in all four measures, scoring $R^2 = 0.6221$, RMSE = 96.71, MSE = 9,352, and MAPE = 30.37%. Each result will be discussed separately by comparing the hybrid model to respective model.

Performance of the ARIMA (2, 1, 2): The single model shows relatively poor performance with negative R^2 (-0.5148), RMSE = 193.64, and MAPE = 670.89%. A negative R^2 indicates poor predictive accuracy compared to the simple mean model. In particular, negative R^2 indicates that linear ARIMA models are unable to capture the complex dynamics of non-linearity and regime changes in the 2024 Econet price pattern. The graph resembles a straight horizontal line that fails to reflect the ups and downs of real data. These results agree with findings from Section 4.4.3 on heteroscedasticity test of ARIMA models.

FFNN [4-4-1] Performance: The standalone FFNN does a good job at picking up the main non-linear trends in the time series. It got a MAPE of 94.14%, an RMSE of 142.724, and a positive R^2 of 0.177. So, it definitely grabs a big chunk of the data's variation and isn't just spitting out random stuff.

That said, the model shows a consistent "scale dampening" issue. It accurately predicts price hikes till month 55 and then gets the shift in direction right too. But when it comes to the actual height of those prices, especially at the highest point in month 55, the model plays it safe. During the steep drop from months 57 to 58, we see the biggest misses. This implies that while the neural network figured out the general flow of prices, it struggles with extremes. Probably because of how the Min-Max normalization was set based on the training data, it can't quite handle such volatility outside of that sample range.

Hybrid ARIMA–FFNN Performance: The hybrid model is the most performant in terms of all indicators used. This model shows an R-squared of 0.6221 implying that it can explain 62% of the variance in the actual test period prices, a significant improvement from the negative values exhibited by the models. This model has an MAPE value of 30.37%, implying that the deviation between the hybrid prediction and the actual price is about 30 cents in every Zimbabwean Dollar. On the other hand, the individual frameworks' MAPE was 671% and 94.14%. The improved performance of the hybrid model emanates from its unique ability to utilize the strengths of each model in capturing trends while correcting errors using the other model.

However, one needs to realize that despite being much better than the predictions of individual models, MAPE equalling 30.37% and R^2 equalling 0.6221 means that even hybrid model predictions still are flawed – a quite expected result for the market undergoing distinct currency reform like in case of Zimbabwe in 2024. The forecast residual error of the hybrid model is due to the nature of the process of currency price resets that cannot be predicted accurately since they are discontinuous.

4.6.3 Implications of the Findings

Several implications can be drawn from the findings obtained. For investors and financial analysts, the implications include the inadequacy of using traditional ARIMA model for predicting the prices of stocks in Zimbabwe due to its structural volatilities, and therefore a need to employ a machine learning-based predictive model as a better alternative. For the ZSE, these findings highlight the significance of exchange rate stability as the cornerstone of the smooth operation of the equity market. Academically, the current research contributes to existing studies by providing an empirical demonstration of hybrid ARIMA-FFNN approach on ZSE equity.

4.7 Chapter Summary

In this chapter, the complete empirical results of the study were presented. From the descriptive statistics, the extremely volatile nature and non-normal distribution were confirmed for all three variables. Using stationarity tests, the stock price series was found to be $I(1)$ (integrated of order one) such that it required first-order differencing before being subjected to ARIMA modelling. The ARIMA (2, 1, 2) was determined to be the best linear model using AIC and BIC criteria. Residual diagnostics revealed the presence of white noise serial independence, but there was significant heteroscedasticity and non-normality suggesting some structural non-linearity in the series. The standalone FFNN [4-4-1] was able to capture directional information, but suffered from the scaling problem. On the other hand, the Hybrid ARIMA – FFNN model performed better than the other two models in all criteria. It obtained $R^2 = 0.6221$ and $MAPE = 30.37\%$. Chapter five concludes the study.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.0 Introduction

This chapter brings together the empirical results obtained from the analysis of the data and concludes in relation to the research objectives and research questions. This chapter is divided into five sections in accordance with the research objectives set out in Chapter One.

5.1 Conclusions

5.1.1 Objective 1: ARIMA and FFNN Models

The ARIMA(2,1,2) model was considered to be the best model for the time series Econet Wireless Zimbabwe based on the AIC and BIC criteria. The model explained the serial autocorrelation in the price series (Ljung-Box $p = 0.5728$) although it could not describe the nonlinear dynamics in the prices during the test period ($R^2 = -0.5148$ and MAPE = 670.89%). The standalone FFNN [4–4–1] performed better than the ARIMA model in terms of capturing directional changes while having systematic magnitude underestimation problem (MAPE = 94.14% and $R^2 = 0.177$).

5.1.2 Objective 2: Development of the Hybrid ARIMA–FFNN Model

The Hybrid ARIMA–FFNN model was successfully constructed using the decomposition methodology proposed by Zhang (2003). The ARIMA–FFNN error correction component was trained on residuals, exchange rate, and inflation information. Diagnostics performed for the hybrid model's residuals through ARCH-LM (p -value = 0.1252) and Ljung-Box Q-test (p -value = 0.3207) showed that the hybrid framework effectively addressed any existing structure either linear or non-linear of the residuals. As a result, the residual terms were approximately white noise.

5.1.3 Objective 3: Accuracy Comparison for the Models

Hybrid ARIMA–FFNN model significantly surpassed both individual models in terms of all four performance measures: R^2 (0.6221 vs. -0.5148 and 0.177), RMSE (96.71 vs. 193.64 and 142.724), MSE (9,352 vs. 37,496 and 20370), and MAPE (30.37% vs. 670.89% and 94.14%). These findings indicate that a hybrid approach is capable of obtaining an optimal balance between bias and variance by simultaneously modeling stable linear autocorrelations and unstable non-linearities inherent to the Econet price series.

5.1.4 Objective 4: Advantages and Disadvantages Of Each Model

Interpretability and theoretical rigor associated with ARIMA is offset by its inability to capture non-linearities, heteroscedasticity, and breaks within the data series. Flexibility and non-linear capabilities of FFNN are constrained due to limited number of observations and problem of extrapolation for future currency regime transition in 2024. In this regard, the Hybrid model has successfully capitalized on the advantages of both techniques: ARIMA makes reliable linear forecasts while FFNN is able to correct errors in non-linear space. However, the accuracy of predictions made by a hybrid algorithm still suffers from a discontinuity shock resulting from the currency regime change.

5.1.5 Objective 5: Most Reliable Model in Volatile Conditions

The Hybrid ARIMA-FFNN model is the most reliable framework in comparison to the other two models when it comes to predicting the future stock price of Econet Wireless Zimbabwe within the turbulent and structurally unstable conditions of the period 2020-2024. This reliability is observed across all measures as well as forecast periods, from the relatively stable start of the test period to the highly volatile Months 51-57. These findings are consistent with Zhang (2003) and the overall academic literature on hybrid forecasting.

5.2 Recommendations

To Investors and Financial Analysts

It is recommended that investors working on the ZSE adopt hybrid machine learning approaches to forecasting stock price movements rather than exclusively using either technical analysis techniques or econometric modeling approaches that assume a linear relationship between stock performance and various drivers. However, while forecasts from this hybrid model have been proven superior at predicting changes in stock performance, this should not be regarded as a deterministic prediction and taken into consideration, especially when expecting macroeconomic regime shifts.

To Econet Wireless Management

Investor relations and strategic financial planning of the firm may be improved through application of hybrid forecasting techniques that can help in prediction of probable share price trends and thus aid decision-making regarding capital generation activities as well as dividend payout.

For Future Research

- i. Subsequent studies can expand the analysis to compare various advanced deep learning models such as LSTM and GRU neural networks, which are explicitly designed for long range time dependency analysis and may prove superior to basic FFNNs when dealing with lengthy time series.
- ii. Using a bigger database consisting of daily/weekly prices will allow using a large training sample, hence eliminating overfitting concerns and ensuring better deep learning model training process.
- iii. Further researches can add more variables to the existing list of independent factors, specifically sentiment indices extracted from media.
- iv. Replication of this comparative approach for other ZSE-listed stocks such as OK Zimbabwe, Cassava Smartech, and Delta Corporation will determine if this superior hybrid approach holds for other securities in the Zimbabwean stock market.
- v. Further research should be conducted on variants of the FFNN that involve learning techniques in which the weights in the model are adjusted on the go when fresh price information is fed into the system.

5.3 Summary and Conclusion

In summary, the aim of this study was to propose and evaluate ARIMA model, Feedforward Neural Network model and hybrid ARIMA-FFNN model for predicting stock prices of Econet Wireless Zimbabwe from 2020 through 2024. The findings indicate that the Hybrid ARIMA-Feedforward Neural Network model performed significantly better than the standalone approaches in this highly structurally volatile and macro economically unstable setting of the Zimbabwe Stock Exchange.

It should be noted that the current study has shown that the linear model alone does not suffice to forecast in settings where there are such elements as currency crisis and structural breakage along with non-linear behaviour of the investors, hence the combination of linear model and non-linear neural network can lead to significant improvement in prediction accuracy.

The current study also adds to the empirical literature on advanced financial forecasting techniques and is likely to prove useful for conducting further research in the area of African frontier markets.

REFERENCES

1. Exchange rates from the RBZ accessed via <https://zimbabwe.opendataforafrica.org/wldqdrq/exchange-rates> on 15 October 2025
2. Consumer price index from the central bank accessed via <https://www.rbz.co.zw/index.php/publications-notice/publications/monthly-economic-review/24-monthly-review/1395-2024-monthly-economic-review> on 15 October 2025
3. Consumer price index from the central bank accessed via [https://www.rbz.co.zw/documents/monthly_review/2024/Monthly Economic Review December 2024.pdf](https://www.rbz.co.zw/documents/monthly_review/2024/Monthly_Economic_Review_December_2024.pdf) on 15 October 2025
4. Econet historical stock prices accessed via <https://www.investing.com/equities/econet-wireless-zimbabwe-limited-historical-data> on 30 September 2025
5. 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. <https://doi.org/10.1155/2014/614342>
6. Aggarwal, R. (1981). Exchange rates and stock prices: A study of the U.S. capital markets under floating exchange rates. *Akron Business and Economic Review*, 12(3), 7–12.
7. Barberis, N., Shleifer, A., & Vishny, R. (1998). *A model of investor sentiment*. *Journal of Financial Economics*, 49(3), 307–343.
8. Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. Holden-Day.
9. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control (5th ed.)*. Wiley.
10. Chatfield, C. (2003). *The analysis of time series: An introduction (6th ed.)*. Chapman & Hall/CRC.
11. Chen, J., Wang, J., & Hu, J. (2019). Forecasting stock prices using a hybrid model integrating ARIMA and neural networks. *Expert Systems with Applications*, 117, 102–117.
12. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303–314.

13. Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427–431.
14. Diebold, F. X., & Rudebusch, G. D. (1996). Measuring business cycles: A modern perspective. *Review of Economics and Statistics*, 78(1), 67–77.
15. Dornbusch, R., & Fischer, S. (1980). Exchange rates and the current account. *American Economic Review*, 70(5), 960–971.
16. Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.
17. Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *American Economic Review*, 71(4), 545–565.
18. Fildes, R., & Goodwin, P. (2007). Against your better judgment? How organizations can improve their use of management judgment in forecasting. *Interfaces*, 37(6), 570–576.
19. Fisher, I. (1930). The theory of interest. *Macmillan*.
20. Gumbo, V. (2022). Neural network models for forecasting in volatile African equity markets. *African Journal of Business and Management*, 16(2), 44–59.
21. Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384.
22. Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.
23. Hamadneh, N. N., Alzubi, J. A., Bhatnagar, R., Hazber, A. S., & Alzubi, O. A. (2018). Artificial neural network and stock markets prediction. *International Journal of Engineering and Technology*, 7(2), 4–9.
24. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: *Data mining, inference, and prediction (2nd ed.)*. Springer.
25. Haykin, S. (2008). *Neural networks and learning machines (3rd ed.)*. Pearson Prentice Hall.
26. Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems*, 16(1), 44–55.
27. Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development*, 15(14), 5481–5487.

28. Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2), 251–257.
29. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
30. Kara, Y., Boyacioglu, M. A., & Baykan, O. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: *The sample of the Istanbul Stock Exchange*. *Expert Systems with Applications*, 38(5), 5311–5319.
31. Kim, K. J., Han, I., & Chun, S. H. (2009). Hybrid neural network models for stock price prediction. *Applied Intelligence*, 30(3), 217–226.
32. Krolzig, H.-M. (1997). Markov-switching vector autoregressions: *Modelling, statistical inference, and application to business cycle analysis*. Springer.
33. Lo, A. W. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*, 30(5), 15–29.
34. Makridakis, S., & Wheelwright, S. C. (1989). *Forecasting methods for management (5th ed.)*. Wiley.
35. Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and applications (3rd ed.)*. Wiley.
36. Mishra, A., Dash, R., & Mishra, S. (2021). A hybrid ARIMA-neural network model for stock price prediction. *International Journal of Computational Intelligence Systems*, 14(1), 1745–1757.
37. Mlambo, C., & Biekpe, N. (2007). The efficient market hypothesis: Evidence from ten African stock markets. *Investment Analysts Journal*, 36(66), 5–17.
38. Mostafa, M. M. (2010). Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait. *Expert Systems with Applications*, 37(9), 6302–6309.
39. Ncube, M. (2020). Hybrid machine learning forecasting for frontier market equities: Evidence from Zimbabwe. *Journal of Emerging Market Finance*, 19(3), 289–314.
40. Odean, T. (1999). Do investors trade too much? *American Economic Review*, 89(5), 1279–1298.
41. Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57(6), 1361–1401.

42. Pindyck, R. S., & Rubinfeld, D. L. (1998). *Econometric models and economic forecasts (4th ed.)*. McGraw-Hill.
43. Reserve Bank of Zimbabwe. (2024). *Monetary policy statement and statistical tables*. RBZ.
44. Rojas, I., Valenzuela, O., & Fernández, F. J. (2018). Soft-computing techniques and ARIMA model for time series prediction. *Neurocomputing*, 71(4–6), 519–537.
45. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). *Learning representations by back-propagating errors*. *Nature*, 323(6088), 533–536.
46. Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2), 41–49.
47. Sara, J. D., Makke, L., & Bhatt, D. L. (2019). Performance metrics for evaluating machine learning models. *Journal of Clinical Medicine*, 8(3), 1–10.
48. Shiller, R. J. (2000). *Irrational exuberance*. Princeton University Press.
49. Shiller, R. J. (2003). From efficient markets theory to behavioral finance. *Journal of Economic Perspectives*, 17(1), 83–104.
50. Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, 12(3), 183–206.
51. Tokui, S., Oono, K., Hido, S., & Clayton, J. (2015). Chainer: A next-generation open source framework for deep learning. *Proceedings of LearningSys in the 29th Annual Conference on Neural Information Processing Systems*.
52. Tsay, R. S. (2010). *Analysis of financial time series (3rd ed.)*. Wiley.
53. Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175.
54. Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: *The state of the art*. *International Journal of Forecasting*, 14(1), 35–62.
55. Zhong, X., & Enke, D. (2019). Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financial Innovation*, 5(1), 1–20.

APPENDICES

APPENDIX A: RAW DATA

Date	Historical stock prices	Monthly exchange rate	Monthly Inflation Rates(CPI)
Jan-20	0.09	17.10	-0.86
Feb-20	0.11	17.68	-9.34
Mar-20	0.1	21.16	-1.96
Apr-20	0.13	25.00	-1.50
May-20	0.3	25.00	-0.21
Jun-20	0.33	32.35	-0.51
Jul-20	0.19	68.94	0.16
Aug-20	0.19	81.60	-1.03
Sep-20	0.15	82.25	-1.30
Oct-20	0.2	81.37	-0.16
Nov-20	0.37	81.68	1.83
Dec-20	0.55	81.82	1.83
Jan-21	0.69	82.07	2.45
Feb-21	0.85	83.46	-0.06
Mar-21	0.78	84.00	0.60
Apr-21	0.96	84.50	0.24
May-21	1.07	84.73	1.86
Jun-21	1.17	85.42	4.16
Jul-21	1.48	85.64	0.01
Aug-21	2.02	85.91	0.53
Sep-21	3.09	87.67	1.03
Oct-21	2.57	97.14	1.52
Nov-21	3.3	102.08	1.08
Dec-21	3.88	108.67	0.80
Jan-22	5.4	115.42	1.45
Feb-22	5.34	124.02	1.57
Mar-22	10.87	142.42	1.59

Apr-22	9.26	159.35	5.47
May-22	6.21	290.89	3.00
Jun-22	6.09	366.27	8.74
Jul-22	4.23	416.29	3.70
Aug-22	4.46	546.83	2.91
Sep-22	3.36	621.53	2.91
Oct-22	2.78	632.14	0.36
Nov-22	3.76	654.87	2.93
Dec-22	5.28	671.45	1.27
Jan-23	6.6	796.52	0.66
Feb-23	9.9	889.13	0.00
Mar-23	10.52	929.86	-3.73
Apr-23	33.01	1047.45	-0.32
May-23	34.23	2577.06	1.54
Jun-23	20.2	5739.80	2.64
Jul-23	19.69	4516.80	12.10
Aug-23	20.01	4608.11	1.03
Sep-23	20.01	5466.75	-1.34
Oct-23	29.21	5698.96	0.95
Nov-23	34.53	5791.08	2.46
Dec-23	91.65	6104.72	4.54
Jan-24	72.03	10152.39	4.70
Feb-24	187.7	14912.83	-0.27
Mar-24	145	22055.47	-0.18
Apr-24	159.95	13.43	0.17
May-24	260.05	13.32	-0.58
Jun-24	400	13.70	-0.24
Jul-24	352.3	13.79	-0.15
Aug-24	349.95	13.86	0.40
Sep-24	481	24.88	1.90
Oct-24	352.95	28.68	7.31
Nov-24	3.23	25.45	2.22

Dec-24	2.5	25.80	1.11
--------	-----	-------	------

APPENDIX B: R CODES USED TO RUN THE MODELS

1. INITIAL SETUP AND LIBRARIES

```
set.seed(123) # [cite: 1208]
```

Required packages

```
library(readxl) # For importing historical excel data
```

```
library(forecast) # For ARIMA model identification, estimation, and forecasting  
[cite: 1206]
```

```
library(tseries) # For ADF, Ljung-Box, and statistical testing [cite: 1206]
```

```
library(lmtest) # For Breusch-Pagan heteroscedasticity testing [cite: 1206]
```

```
library(neuralnet) # For Feedforward Neural Network architecture & fitting [cite:  
1206]
```

```
library(ggplot2) # For creating visual plots [cite: 1206]
```

2. DATA IMPORT AND PRE-PROCESSING

Assuming data contains columns: Date, Price, ExRate, Inflation

```
data <- read_excel("econet_stock_data.xlsx")
```

Handling any structural data gaps via linear interpolation

```
data$Price <- na.interp(data$Price)
```

```
data$ExRate <- na.interp(data$ExRate)
```

```
data$Inflation <- na.interp(data$Inflation)
```

Converting the dependent variable into an explicit time series object (Monthly)

```
price_ts <- ts(data$Price, start = c(2020, 1), frequency = 12)
```

```

# 3. DATA SPLITTING (TRAIN / TEST PARTITIONING)
# -----
# Splitting scheme: 80% Training (Months 1-48), 20% Testing (Months 49-60) [cite:
1584]
train_size <- 48

train_data <- head(data, train_size)
test_data <- tail(data, nrow(data) - train_size)

train_ts <- subset(price_ts, end = train_size)

# 4. DATA STATIONARITY CHECK
# -----
# Augmented Dickey-Fuller (ADF) Test to verify if differencing is needed
adf_test <- adf.test(train_ts, alternative = "stationary") #
print(adf_test)

# 5. MODEL 1: STANDALONE ARIMA(2, 1, 2) DEVELOPEMENT &
FORECAST
# -----
# Fitting the selected optimal ARIMA model based on AIC/BIC minimization
arima_model <- Arima(train_ts, order = c(2, 1, 2))
summary(arima_model)

# Diagnostic tests on residuals
arima_residuals <- residuals(arima_model)
checkresiduals(arima_model) # Performs Ljung-Box test automatically [cite: 1206]

# Breusch-Pagan test for Heteroscedasticity / Volatility Clustering
bp_test <- bptest(arima_residuals ~ time(arima_residuals)) # [cite: 1206, 1235]
print(bp_test)

```

```

# Out-of-sample forecast for the 12 months test window
arima_forecast_obj <- forecast(arima_model, h = nrow(test_data))
arima_forecast    <- as.numeric(arima_forecast_obj$mean)

# 6. VARIABLE TRANSFORMATION FOR NEURAL NETWORKS (MIN-MAX
# SCALING)
# -----
# FFNN inputs require 0-1 scaling to maximize backpropagation efficiency
min_max_scale <- function(x, min_val, max_val) {
  return((x - min_val) / (max_val - min_val))
}

# Inverse scaling helper function to bring outputs back to investor nominal values
inverse_scale <- function(x_scaled, min_val, max_val) {
  return(x_scaled * (max_val - min_val) + min_val)
}

# Define extreme values based strictly on training boundaries to prevent leakages
p_min <- min(train_data$Price);   p_max <- max(train_data$Price)
ex_min <- min(train_data$ExRate); ex_max <- max(train_data$ExRate)
inf_min <- min(train_data$Inflation); inf_max <- max(train_data$Inflation)

# Constructing lagged variables to prepare for supervised network inputs
# Lag 1 and Lag 2 calculations create missing spots for the first two data rows
data$Price_Lag1 <- c(NA, head(data$Price, -1))
data$Price_Lag2 <- c(NA, NA, head(data$Price, -2))

# Normalize features
data$Price_scaled    <- min_max_scale(data$Price, p_min, p_max)
data$Price_Lag1_scaled <- min_max_scale(data$Price_Lag1, p_min, p_max)
data$Price_Lag2_scaled <- min_max_scale(data$Price_Lag2, p_min, p_max)

```

```

data$ExRate_scaled <- min_max_scale(data$ExRate, ex_min, ex_max)
data$Inflation_scaled <- min_max_scale(data$Inflation, inf_min, inf_max)

# Separate clean supervised datasets (removing NA rows generated by lagging)
nn_clean_data <- na.omit(data)
nn_train <- head(nn_clean_data, train_size - 2) # Adjusts for lag structure losses
[cite: 1308]
nn_test <- tail(nn_clean_data, nrow(test_data))

# 7. MODEL 2: STANDALONE FFNN [4-4-1] ARCHITECTURE
# -----
# Define the optimal topology formula: 4 inputs -> 4 hidden nodes -> 1 output
ffnn_formula <- Price_scaled ~ Price_Lag1_scaled + Price_Lag2_scaled +
ExRate_scaled + Inflation_scaled

ffnn_model <- neuralnet(ffnn_formula, data = nn_train, hidden = 4, linear.output =
TRUE)]
plot(ffnn_model)

# Directional inference/forecasting on out-of-sample partition
ffnn_predictions_scaled <- compute(ffnn_model, nn_test[, c("Price_Lag1_scaled",
"Price_Lag2_scaled", "ExRate_scaled", "Inflation_scaled")])
ffnn_forecast <- inverse_scale(ffnn_predictions_scaled$net.result, p_min, p_max)

# 8. MODEL 3: HYBRID ARIMA-FFNN ERROR CORRECTION STRUCTURE
# -----
# Step A: Isolate the linear residuals generated during the training phase [cite: 1548]
train_residuals <- as.numeric(residuals(arima_model))

# Build a matching matrix for the Hybrid FFNN training process

```

```
# The target variable becomes the unexplained residual instead of the raw stock price  
[cite: 1548]
```

```
hybrid_train_df <- data.frame(  
  Residual = train_residuals[3:train_size], # Aligning index offsets due to lag  
  exclusions  
  Price_Lag1_scaled = nn_train$Price_Lag1_scaled,  
  Price_Lag2_scaled = nn_train$Price_Lag2_scaled,  
  ExRate_scaled = nn_train$ExRate_scaled,  
  Inflation_scaled = nn_train$Inflation_scaled  
)
```

```
# Step B: Train the error-correcting Neural Network mapping linear shortcomings  
[cite: 1320, 1548]
```

```
hybrid_formula <- Residual ~ Price_Lag1_scaled + Price_Lag2_scaled +  
ExRate_scaled + Inflation_scaled  
hybrid_nn <- neuralnet(hybrid_formula, data = hybrid_train_df, hidden = 4,  
linear.output = TRUE)
```

```
# Step C: Forecast the non-linear errors expected across the test timeline
```

```
hybrid_test_inputs <- nn_test[, c("Price_Lag1_scaled", "Price_Lag2_scaled",  
"ExRate_scaled", "Inflation_scaled")]  
predicted_residuals <- compute(hybrid_nn, hybrid_test_inputs)$net.result
```

```
# Step D: Aggregate Forecast Outputs (Linear ARIMA projection + Predicted Non-  
linear Error correction) # Note: Because the test set is mapped onto 10 usable  
sequential forecast rows [cite: 1289]
```

```
arima_test_subset <- head(arima_forecast, length(predicted_residuals))  
hybrid_forecast <- arima_test_subset + predicted_residuals # [cite: 1548]
```

```
# 9. EVALUATION METRICS COMPILATION
```

```
# -----
```

```

# Align actual values matching the 10-month continuous timeframe context [cite:
1289]
actual_prices <- head(test_data$Price, length(predicted_residuals))
arima_eval  <- arima_test_subset
ffnn_eval   <- head(ffnn_forecast, length(predicted_residuals))

# Metric Calculation Function
calculate_metrics <- function(actual, predicted) {
  rmse <- sqrt(mean((actual - predicted)^2)) # [cite: 1189]
  mae  <- mean(abs(actual - predicted))     # [cite: 1190]
  mape <- mean(abs((actual - predicted) / actual)) * 100 # [cite: 1191]

  # R-squared [cite: 1192]
  rss <- sum((actual - predicted)^2)
  tss <- sum((actual - mean(actual))^2)
  r_sq <- 1 - (rss / tss)

  return(c(RMSE = rmse, MAE = mae, MAPE = mape, R2 = r_sq))
}

# Compile consolidated summary table [cite: 1385]
results_matrix <- rbind(
  ARIMA = calculate_metrics(actual_prices, arima_eval),
  FFNN  = calculate_metrics(actual_prices, ffnn_eval),
  Hybrid = calculate_metrics(actual_prices, hybrid_forecast)
)

print(round(results_matrix, 4))

```