2026-04

Hybrid Econometric - Deep Learning Models for Volatility and Correlation Forecasting in Stock and Agricultural Commodity Markets

Elias, Dessie

Hybrid Econometric - Deep Learning Models for Volatility and Correlation Forecasting in Stock and Agricultural Commodity Markets

By

ELIAS DESSIE JEMBERIE

Department of Mathematics

BAHIR DAR UNIVERSITY

A dissertation submitted to the Department of Mathematics, College of Science, Bahir Dar University in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Financial Mathematics.

Supervisor: Tesfahun Berhane (PhD, Associate Professor), Department of Mathematics, College of Science, Bahir Dar University

April, 2026

BAHIR DAR UNIVERSITY

DEPARTMENT OF MATHEMATICS

Thesis Approval

This is to certify that the thesis entitled **Hybrid Econometric - Deep Learning Models for Volatility and Correlation Forecasting in Stock and Agricultural Commodity Markets** , which is summited to Department of Mathematics, College of Science, Bahir Dar University for the award of degree of Doctor of Philosophy, is the research work by **Elias Dessie Jemberie**.

External Examiner 1: **Professor Jimbo Henri Claver** Signature ________________

External Examiner 2: **Professor Jules Sadefo Kamdem** Signature ________________

Internal Examiner: **Professor Elias Rabson Offen** Signature ________________

Chair Person : **Hunegnaw Dessie (Asso. Professor)** Signature ________________

April, 2026

Declaration of Authorship

I declare that the work embodied in this thesis entitled *Hybrid Econometric - Deep Learning Models for Volatility and Correlation Forecasting in Stock and Agricultural Commodity Markets* is my own bona fide work carried out by me under the supervision of Dr.Tesfahun Berhane in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes. The matter embodied in this dissertation has not been submitted for any other academic award. I certify that the work is the candidate's own work except where explicit reference is made to the contribution of others. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

Abstract

Predicting volatility and correlation in financial assets is crucial for assessing financial risk and achieving effective portfolio diversification. However, conventional econometric models used for forecasting volatility and dependence structures have frequently been criticized for their limited predictive performance. Consequently, the adoption of deep learning techniques for volatility and correlation forecasting has emerged as a logical alternative. Nevertheless, although the application of machine learning models in volatility prediction has advanced considerably, these approaches still face challenges in effectively capturing market noise and key stylized facts, such as volatility clustering and leverage effects. These considerations highlight the potential for integrating traditional econometric models with deep learning techniques, particularly convolutional neural networks and long short-term memory networks, an area that still requires further exploration. In this thesis, we investigate the integration of GARCH-type models with a CNN–LSTM hybrid framework to enhance the accuracy of volatility prediction in financial markets. Convolutional neural networks (CNNs) are effective in identifying local spatial patterns in time-series data, whereas long short-term memory (LSTM) networks are capable of retaining crucial information over extended temporal horizons. In addition, GARCH models are well known for their ability to capture key stylized facts of financial returns, including volatility clustering and excess kurtosis.

For GARCH-type model selection, several model adequacy and selection criteria are considered, including the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), maximum log-likelihood values, and the statistical significance of model parameters. Furthermore, the mean-reversion property of volatility dynamics is evaluated under alternative distributional assumptions, such as the normal, Student's-t, and generalized error distributions. The CNN and LSTM models are subsequently selected due to

their complementary capabilities in volatility prediction. The forecasting performance of the proposed hybrid models is evaluated using loss functions computed from daily out-of-sample test data for the S&P 500 index. Specifically, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) are employed as evaluation metrics. The empirical results indicate that the proposed GARCH–CNN–LSTM model outperforms alternative hybrid specifications, achieving up to a 13% improvement in forecasting accuracy over the study period.

Furthermore, this study proposes a DCC–EGARCH model integrated with a Long Short-Term Memory (LSTM) network to model multivariate volatility and to obtain more precise estimates of dynamic correlations. The DCC–EGARCH framework effectively captures asymmetric effects in standardized residuals within the conditional covariance structure, while the LSTM network is capable of learning long-term dependencies and complex temporal patterns in sequential data. Consequently, the combined framework is expected to enhance forecasting accuracy for dynamic correlations across stock market indices as well as selected agricultural commodities.

In particular, the volatility dynamics and time-varying correlations among the S&P 500, the Shanghai Stock Exchange Composite Index (SHCOMP), the NIFTY 50 of the National Stock Exchange of India, and the S&P/TSX Composite Index of Canada are empirically examined. The experimental results indicate that incorporating LSTM into the DCC–EGARCH framework substantially improves forecasting performance, as evidenced by lower Mean Squared Error (MSE) and Mean Absolute Error (MAE) values. Overall, the findings demonstrate that the proposed DCC–EGARCH–LSTM hybrid model outperforms alternative DCC specifications, highlighting the effectiveness of integrating deep learning techniques to achieve more reliable and accurate predictions of dynamic correlations in financial markets.

In the nutshell, we integrated econometric ARCH -type models with hybrid deep learning models for more accurate univariate and multivariate volatility and correlation prediction in selected agricultural commodities and stock market indices.

Publications

- Dessie, E., Birhane, T., Mohammed, A., & Walelign, A. (2025). A hybrid GARCH, convolutional neural network and long short term memory methods for volatility prediction in stock market. J. COMBIN. MATH. COMBIN. COMPUT, 124(461), 476.

- Dessie, E., Brhane, T., Mohammed, A., & Walelign, A. (2023). Agricultural Commodity Price Volatility Models and Prediction Performance of GARCH Family Models. Horn of African Journal of Business and Economics (HAJBE), 6(2), 150-171.

Acknowledgements

I would like to thank my advisor, Dr. Tesfahun Berhane (Associate Prof.) for his unconditional support. His guidance has helped me every step of the way in researching and writing this thesis. I would also like to thank to my co-advisors Dr. Abdu Mohammed and Dr. Assaye Walelign for their consistent feedback, encouragement and invaluable advice. I wish to extend my special thanks to all staff members of the department of Mathematics for their valuable support. Finally, I would also like to thank my family for their endless support.

Bahir Dar, Ethiopia

Elias Dessie Jemberie

April, 2026

Contents

List of Figures

List of Tables

List of Abbreviations

ACF	**A**uto**c**orrelation **F**unction
ADF	**A**ugmented **D**ucky **F**uller
ADCC	**A**symmetric **D**ynamic **C**onditional **C**orrelation
ANN	**A**rtificial **N**eural **N**etwork
BEKK	**B**aba **E**ngle **K**raft **K**roner
CCC	**C**onstant **C**onditional **C**orrelation
CNN	**C**onvolutional **N**eural **N**etwork
DCC	**D**ynamic **C**onditional **C**orrelation
DNN	**D**eep **N**eural **N**etwork
EGARCH	**E**xponential **G**eneralized **A**utoregressive **C**onditional **H**eteroscedasticity
GARCH	**G**eneralized **A**utoregressive **C**onditional **H**eteroscedasticity
GED	**G**eneralized **E**rror **D**istribution
LM	**L**agrange **M**ultiplier
LSTM	**L**ong **S**hort **T**erm **M**emory
MGARCH	**M**ultivariate **G**eneralized **A**utoregressive **C**onditional **H**eteroscedasticity
MLE	**M**aximum **L**ikelihood **E**stimation
MSE	**M**ean **S**quare **E**rror
MAE	**M**ean **A**bsolute **E**rror
ND	**N**ormal **D**istribution
NIFTY 50	**N**ational **S**tock **E**xchange **F**ifty
PACF	**P**artial **A**uto**c**orrelation **F**unction
ReLU	**Re**ctified **L**inear **U**nit

RMSE	**R**oot **M**ean **S**quare **E**rror
RWC	**R**olling **W**indow **C**orrelation
SHCOMP	**S**hanghai **C**omposite **I**ndex
STD	**S**tudent's **T** **D**istribution
S&P 500	**S**tandard & **P**oor's 500
TGARCH	**T**hreshold **G**eneralized **A**utoregressive **C**onditional **H**eteroscedasticity
TSX	**T**oronto **S**tock **E**xchange
VEC	**V**ector **E**rror **C**orrection

Chapter 1

Introduction

1.1 Background of the Study

Forecasting volatility and correlation is a key part of making financial decisions, managing risk, and deciding how to distribute assets. Thus, precise prediction of these two market traits gives an opportunity for investors, portfolio managers, and policymakers to make better choices and deal with uncertainty more effectively. In financial markets, volatility reflects the degree of variability in asset prices over time, while correlation measures the degree to which different financial assets move together.

Volatility has three stylized facts. The first is volatility clustering which the empirical evidence dates back to the well-known pioneering studies of Mandelbrot (1963) [120] and Fama (1965) [121] demonstrated that large price (small price) changes tend to be followed by large price (small price) changes. The second asymmetric behavior of volatility in financial market is the leverage effect, in which depreciation (a downward movement) is always followed by higher volatility (volatility is high after negative shocks) while appreciation (upward movement) is followed by low volatility (volatility is low after positive shocks). There is evidence that bad news has higher volatility than good news which indicates that volatility is higher after negative shocks than after positive shocks of the same magnitude [68]. The third crucial stylized fact of volatility is the tail behavior or kurtosis of a continuous random variable. The shifting of the distribution either to the right or to the left from the normal distribution.

Together, these fundamental measures are essential for evaluating market risk, determining optimal portfolio diversification strategies, and assessing the stability of the financial system. Given their importance, the development of accurate and reliable models for forecasting volatility and correlation has become a central focus in the fields of financial econometrics and quantitative finance.

Conventional econometric models have been extensively employed for several decades to forecast financial market volatility and dependence structures. In particular, the Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models are among the most widely used and influential frameworks in financial econometrics. Their popularity stems from their ability to capture key stylized facts observed in financial time series, including volatility clustering, leptokurtosis (fat-tailed distributions), and time-varying conditional variance.

Furthermore, numerous extensions of the GARCH family have been proposed to improve model flexibility and empirical performance. For instance, asymmetric specifications such as the Exponential GARCH (EGARCH) and the Glosten–Jagannathan–Runkle GARCH (GJR-GARCH) models are designed to account for leverage effects and the asymmetric response of volatility to positive and negative market shocks. In addition, multivariate frameworks, particularly Dynamic Conditional Correlation (DCC) models, have been widely applied to capture the evolution of time-varying correlations across different financial assets and markets.

Although, these traditional econometric models have been used extensively their predictive accuracy and capacity to adjust to the intricacies of contemporary financial markets have come under growing criticism. Their parametric nature and linear structure frequently restrict their ability to model complex and nonlinear dynamics[134]. These models are unable to adequately capture the abrupt regime shifts, structural breaks, and evolving patterns that are common in financial time series data. Because of this, their forecasting abilities may suffer in erratic and unpredictable market circumstances, underscoring the necessity for more advanced and flexible modeling techniques.

In recent years, there have been new opportunities to improve financial forecasting due to the rapid development of deep learning and machine learning techniques. Deep learning models have shown remarkable success in domains such as speech processing, natural language understanding, and fraud detection because they can automatically extract complex nonlinear patterns from large datasets. Two of these methods that have attracted increased attention in the financial sector are convolutional neural networks (CNNs) and Long Short-Term Memory (LSTM) networks. CNNs are particularly adept at identifying local and spatial patterns in time series data, whereas LSTMs excel at capturing long-term dependencies and temporal relationships. Their complementary benefits make them useful tools for financial time series forecasting.

Despite their ability to extract non-linear features, deep learning methods alone face challenges such as noise, volatility clustering, and the leverage effect when applied to financial data [135]. Unlike physical or engineering systems, financial markets are inherently noisy, influenced by behavioral factors, and characterized by stylized facts such as volatility clustering and leverage effects. Purely data-driven deep learning models may fail to capture these characteristics effectively, leading to instability or reduced interpretability. This limitation highlights the potential benefits of hybrid modeling approaches that combine the strengths of econometric models—known for their capacity to handle stylized facts and provide statistical interpretability—with the flexibility and predictive power of deep learning techniques.

The need for accurate volatility and correlation forecasting is particularly critical in an increasingly interconnected global financial system. Shocks originating in one market can rapidly transmit to others through various channels, amplifying systemic risk. Modeling dynamic correlations between key stock market indices and other asset classes, such as agricultural commodities, is therefore essential for portfolio diversification, hedging strategies, and financial stability monitoring. While DCC and EGARCH models provide useful frameworks for modeling time-varying correlations, they can be further enhanced by incorporating deep learning components that capture nonlinear temporal dynamics more effectively. The integration of LSTM networks with DCC-EGARCH models can improve their ability to predict dynamic correlations under changing market conditions.

This thesis explores these opportunities by investigating the predictive performance of hybrid econometric and deep learning models for both univariate and multivariate volatility and correlation forecasting. Specifically, the study develops and tests integrated models such as GARCH–CNN–LSTM for volatility forecasting and DCC-EGARCHC–LSTM for dynamic correlation prediction. The empirical analysis focuses on major global stock indices, including S&P 500, Shanghai Stock Exchange Composite Index (SHCOMP), National Stock Exchange of India (NIFTY 50), and Toronto Stock Exchange (TSX), as well as selected agricultural commodities. By employing evaluation metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE), the study assesses the extent to which hybrid models outperform single model approaches.

In conclusion, this thesis is motivated by the need to improve the accuracy and reliability of volatility and correlation forecasts in financial markets. By Combining the statistical rigor of econometric models with the predictive strength of deep learning techniques, this study aims to contribute to both academic literature and practical financial applications. The findings are expected to support better risk management strategies, improve portfolio allocation, and provide valuable insights for financial institutions, investors, and policy-makers in managing uncertainty in increasingly complex and interconnected markets.

.

1.2 Statement of the Problem

Accurately predicting volatility and correlation in financial markets remains one of the most challenging and essential tasks in modern finance. Volatility reflects the level of uncertainty and risk associated with asset price movements, while correlation indicates the degree of co-movement between different financial assets. Together, they play a central role in portfolio classification, asset pricing, and financial risk management. Inaccurate forecasts can lead to poor investment decisions, mispricing of financial instruments, and inadequate hedging strategies—ultimately exposing investors and institutions to significant financial losses. Therefore, the development of reliable volatility and correlation forecasting models is a critical priority in financial research and practice.

Conventional econometric models such as Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) and Dynamic Conditional Correlation (DCC) have been extensively used to model volatility and dependence structures. Although these models effectively capture some stylized facts of financial time series, including volatility clustering, leverage effects and time-varying correlations, they exhibit limited flexibility and adaptabilty which restrict their predictive power when dealing with nonlinear and complex market dynamics [68, 60]. Their forecasting accuracy often deteriorates in highly volatile or rapidly changing market conditions. Moreover, their parametric structure restricts their ability to adapt to new patterns in financial data, reducing their effectiveness for modern risk management applications.

On the other hand, recent advancements in deep learning techniques particularly Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) have shown strong potential in handling complex, nonlinear relationships in time series data. CNNs can efficiently extract local and spatial patterns, while LSTMs are capable of learning long-term temporal dependencies. However, these models alone struggle to account for certain financial characteristics such as leverage effects, volatility clustering, and fat-tailed distributions, which are essential for realistic volatility modeling [138]. Additionally, both CNN and LSTM models lack providing statistically interpretable parameters (like persistence, asymmetry), interpretability and structured assumptions that econometric models provide, such as ensuring positive semi-definite correlation matrices or quantifying volatility persistence. This creates a methodological gap between the statistical interpretability of econometric models and the predictive power of deep learning approaches.

Furthermore, the increasing interconnectedness of global financial markets requires not only accurate univariate volatility forecasting but also precise modeling of dynamic correlations across multiple assets and markets. This is particularly relevant for major stock market indices and agricultural commodities, which are influenced by global shocks, policy changes, and macroeconomic uncertainty. Traditional correlation models often fail to capture nonlinear and evolving dependence structures, leading to less reliable risk assessments and portfolio strategies.

Given these challenges, there is a need for hybrid modeling frameworks that can integrate the strengths of econometric models and deep learning techniques. Such models should be capable of capturing stylized facts, learning complex nonlinear patterns, and adapting to dynamic market environments. This study addresses this gap by developing and empirically testing hybrid models that combine GARCH-type models with CNN and LSTM architectures for volatility forecasting, as well as integrating LSTM into DCC-EGARCH frameworks for dynamic correlation prediction. The focus is on improving predictive accuracy and robustness in both univariate and multivariate settings across major stock market indices and selected agricultural commodities.

1.3 Objective of the Study

1.3.1 General Objective

The general objective of this thesis is to develop hybrid deep learnig integrated with asymmetric autoregresive heteroscedastic type models to predict volatility and covariance in financial markets.

1.3.2 Specific Objectives

This thesis has the following specific objectives:

- To develop and identify appropriate ARCH-family models for forecasting the price returns of selected agricultural commodities, capturing volatility dynamics and stylized facts such as clustering and leverage effects.

- To design a hybrid econometric–deep learning model for univariate volatility forecasting in stock market indice, improving predictive accuracy and robustness compared to conventional models.

- To construct a multivariate DCC-EGARCH-LSTM hybrid model for dynamic conditional correlation estimation among major stock market indices, enhancing the accuracy of correlation forecasts for improved risk management and portfolio diversification.

1.4 Significance of the Study

This study is significant, because it provides practical, theoretical, and policy-relevant contributions to financial markets, agricultural commodity management, and risk mitigation. By integrating deep learning techniques with autoregressive volatility and correlation models, the study offers robust forecasting tools that can enhance decision-making for multiple stakeholders.

The specific benefits and significance of this research include:

1. **Supporting Farmers, Traders, and Policymakers**: The models developed in this study can help stakeholders manage price risk by providing reliable forecasts of commodity price volatility. This enables informed decisions regarding production planning, storage strategies, and optimal timing for sales, minimizing losses and maximizing returns.

2. **Guiding Market Participants Along the Value Chain**: By offering actionable insights into price dynamics, this research assists all actors in the market chain—from producers to distributors and retailers—in making informed strategic decisions that align with market conditions.

3. **Informing Investment and Portfolio Strategies**: The hybrid models provide early information to investors, enabling them to monitor and manage both local and international portfolios. The ability to forecast volatility and correlations across countries allows for more effective portfolio diversification and risk reduction.

4. **Enhancing Market Stability**: By anticipating sudden price fluctuations caused by weather events, supply disruptions, or policy changes, the study contributes to stabilizing financial and commodity markets. This reduces the adverse effects of unexpected shocks on both domestic and global markets.

5. **Facilitating Risk Management for Traders and Investors**: The research helps traders, investors, and risk managers understand and forecast price movements in stock indices and commodities. This supports risk management strategies and helps identify profitable trading opportunities while mitigating exposure to extreme market fluctuations.

6. **Understanding Macroeconomic Implications**: By analyzing the effects of currency movements, exchange rates, and capital flows across countries, the study provides insights into how global financial interconnections influence local and international markets. This information is valuable for policymakers, international investors, and central banks in decision-making processes.

Overall, this research bridges the gap between advanced quantitative modeling and practical financial and commodity market applications. By integrating econometric models with deep learning approaches, the study delivers tools that improve forecasting accuracy, support strategic decision-making, and enhance market efficiency and stability in increasingly complex and interconnected financial systems.

1.5 Limitation of the Study

High frequency data for stock market index and daily price data of agricultural commodies are crutial in modeling realized volatility and correlation for checking the performance of the model. The basic limitation of this study is availability of quality and adequate amount of daily and high frequency financial asset data. Thus, we have used daily closed stock market index and weekly retail price data of selected agricultural commodities in the empirical data analysis. Moreover due to the availability of daily and high frequency data in Ethiopia, particularly for training deep learning approaches, we were forced to consider stock market indices for model building.

1.6 Structure of the Thesis

This thesis is organized as follows: In chapter one, we present the background of the study, statement of the problem, objectives of the thesis and significance. chapter two explores the theoretical and supportive literature review related to the work. This includes theoretical and empirical findings on volatility and correlation models of asset returns. Chapter three covers basic preliminary concepts, model validation tests,and types of error distributions used in model building. Additionally, it presents maximum log-likelihood estimation, performance metrics, fundamental concepts in deep learning neural networks, and activation functions, which serve as the foundation for the subsequent chapters. In chapter four, we present univariate GARCH type volatility and deep learning hybrid models in forecasting the volatility of selected agricultural commodities and stock market. In this chapter empircal data analysis in asymmetric GARCH extention models under different error distributions for the retail price of teff, maize, wheat and barley and identifies appropriate volatility predictive models for each commodity. Chapter five deals with modeling and predicting the volatility and time varying correlation of stock market indices and some selected agricultural commodities by multivariate GARCH type models with Long Short-Term memory hybrid model was presented. Chapter six finally presents the conclusion and future work direction of the thesis.

Chapter 2

Literature Review

In this chapter, an overview of research work with the relevant background related to the work presented in the thesis is provided. We have tried to give a comprehensive overview of related work. Thus, this chapter clearly concentrates mainly on previously investigated volatility and correlation researchs closely related to the work presented later.

The first popular model for forecasting volatility is autoregressive conditional heteroscedasticity (ARCH) pioneered by Robert Engle (1982) [44] in which the forecast of the volatility is derived from historical time series data. The time varying volatility of asset return is a function of squared shocks of the past q lags. However, there are some limitations of ARCH (q) model. Among the drawbacks, how to decide the appropriate number of lags of the squared residual in the model; it is symmetric to the bad news and good news; the large value of q may induce a non- parsimonious conditional variance model. To overcome some of these problems Bollerslev (1986) [39] developed generalized autoregressive conditional heteroscedasticity (GARCH) that is efficient model to estimate and analyze time-varying volatility of asset returns as the combination of p lags in previous square shocks and q lags in previous variance.

GARCH (p,q) model is a more parsimonious model than ARCH (q) model and provides better framework for time-varying volatility estimation. Additionally, GARCH (p,q) model can effectively remove the excess q lags of the shocks. Particularly, GARCH (1,1) model is widely accepted as a benchmark framework for modeling volatilities of many financial time series. However, the standard symmetric GARCH (p,q) model also has three underlying limitations. The first is the requirement that the conditional variance is positive and to keep this the coefficient parameters forced to be positive. The only way to avoid this problem is to place the constraints for coefficients to force them to be positive. The second limitation is that it cannot explain the leverage effect, although it has good performance for explaining volatility clustering and leptokurtosis in a time series. Thirdly, the direct feedback between the conditional mean and conditional variance is not allowed by the standard GARCH (p,q) model. Small shocks may have little effect, while large shocks can causes unproportional volatility change.

In order to overcome the limitations of the standard symmetric GARCH (p,q) model, a number of extensions have been introduced, such as the asymmetric TGARCH(p, q, r), GJR-GARCH (p, q, r), EGARCH (p, q) and power GARCH (p, q) models which can better capture the leverage effects and make the model more efficient in modeling and forecasting the dynamics of financial time series.

Many scholars explored that the asymmetric GARCH models EGARCH, TGARCH, GJR-GARCH highly provide significantly better forecast the volaility of assets more than symmetric GARCH models which exhibit leverage effects [58, 59, 60, 61]. [131] studied the volatility of Real Estate Investment Trusts (REITs) using standard GARCH models with different error distributions, and investigated that the sGARCH with a generalized error distribution outperforms other models in forecasting volatility.

The approach of machine learning algorithms has been growing to enhance the accuracy and reliability of stock market volatility forecasting [75, 82]. Numerous deep learning models have been used in the prediction of stock market indices and correlation of the price volatilities [90, 89, 88]. LSTM is a type of artificial neural network architecture used in deep learning, specifically designed to address the issue of long-term dependencies in sequential data as the traditional recurrent neural network tends to forget important information from earlier parts of the sequence because of issues like vanishing gradient [86, 87, 84, 85].

The combination of multi GARCH type models and the LSTM has better prediction performance in the stock price index volatility of COSPI 200 [52]. The hybrid model of EGARCH and the feed forward artificial neural network demonstrated better volatility forecast in S&P 500 return index [45]. The hybrid model of ARIMA GARCH and LSTM model outperformes the single alone models in the prediction of stock market time series data [40].

Kim, H. Y., & Won, C. H. (2018) [62] used Long Short-Term Memory combined with various generalized autoregressive conditional heteroscedasticity (GARCH)-type models in modeling and forecasting the volatility of COSPI 200 and the combination of more than one single econometric models, hybrid with LSTM model has better performance than the other heteroscedastic GARCH extension models as well as the deep forward neural (DFN) model.

1D CNN recently, has been used in time series data for forecasting purpose.1D-CNN framework for stock price forecasting demonstrates CNN's ability to capture nonlinear temporal dependencies and achieve superior prediction performance compared to traditional statistical baselines (ARIMA) and comparability with LSTM, with additional interpretability insights [139]. Additionally,electricity load forecasting and it gives accurate forecasts in multistep temporal data [140],and other time series predictions [141, 142, 144, 143].

Vidal, A., & Kristjanpoller, W. (2020) [71] proposed two deep learning networks: Convolutional Neural Network (CNN) hybrid with Long Short-Term Memory to predict gold price volatility for risk hedge and portfolio diversification, which significantly improves the forecasting performance of GARCH and LSTM.

Kristjanpoller, W., & Minutolo, M. C. (2016) [63] used the integration of Artificial Neural Network (ANN) that enables to capture the non linear features of the time series with the autoregressive econometric GARCH model that handles the conditional heteroscedasticity to predict time varying oil price volatility and the model out performs other statistical models. The implication of the model is not only providing oil volatility future market dynamic information for trade participants but also motivations for reseachers in hybrid deep learning architectures.

Jardine, A.K.S., Carver, R.B., & Yu, D.L.K.C. (2020) [64] explored autoregressive integrated moving average (ARIMA) combined GARCH volatility dynamic hybridized with Long Short-Term Memory (LSTM) leveraging GARCH for linear structure, while LSTM enables to capture long term and non linear relationships of financial time series data in predicting volatility forecasting.

Amirshahi, B., & Lahmiri, S. (2023) [65] used the combination of deep learning and GARCH extension models with different error distributions in the prediction of cryptocurrencies volatility and the result showed that deep feed forward neural network and LSTM improves the prediction power of the GARCH models in all error distributions.
Dynamic conditional correlation model is an econometric parametric method that used to model and forecast the time varying volatility and conditional correlation of multivariate asset retuns from covariance matrices of each univariate standardized residuals in the autoregressive conditional volatility and Cholesky decomposition proces [77].

Aielli, G. P. (2013) [78] improved the dynamc conditional correlation (DCC) by recursive updating the conditional correlaton in the covariance matrix and this enables the model parameter consistent estimator and better statistical properties.

Bauwens, L., & Otranto, E.(2020) [66] introduced dynamic conditional correlation of stock markets with regime switching that provides a specific relation for each and they imply a nonlinear autoregressive form of dependence on lagged correlations and are based on properties of the Hadamard exponential matrix. Moreover, the model with GARCH dynamics out performs the other conditional correlation models in forecasting performance.

Wang, H., Liu, C., Tran, M.-N., & Wang, C. (2025) [79] used enhanced deep learning multivariate GARCH model which is LSTM-BEEK that leverages the structural strengths of BEKK while enhancing it with LSTM's ability to capture nonlinear and long-term dynamics. The empirical result demonstrates that the combined model out performs both BEEK and DCC in equities covariance and correlation prediction accuracy interms of negative log-likelihood metrics.

There have been many studies on the dynamic correlation and spillover effects on the S&P 500, the Shanghai Composite index, and NIFTY 50 [91, 83, 92, 93]. However, improving the prediction accuracy of stock market index correlation is a crucial part of financial economics, as it plays a fundamental role in investment risk hedging, portfolio management, and diversification of trading strategies.

Many researchers have employed DCC-GARCH model to predict the correlations between stock indices and other assets. Specifically it is favored for its balance of flexibility and parsimony in modeling time-varying correlations and it decouples univariate volatility dynamics from the correlation structure, enabling dynamic co-movements to be tracked with a manageable parameter count as the asset set grows, which underpins its widespread use in forecasting covariances, risk management, and portfolio optimization. In addition, the framework benefits from computational efficiency relative to fully parameterized MGARCH models, robust interpretability through a clear separation of volatility and correlation components, and broad applicability across markets, where it often captures contagion and evolving linkages while remaining a strong benchmarking tool against more complex specifications [94, 97, 95, 96, 98, 99, 100]. However, the model has limitations in capturing complex, nonlinear temporal dependencies of financial data series. The model leads to instability in highly volatile data that gives incorrect risk estimation [80]. The other draw back of DCC-GARCH is that the covariance parameters a and b are scalars that ensure a positive definite matrix for the time series which leads the same dynamics. In addtition, it is limitted in capturing tail dependence and also path dependent [133, 81].

The DCC-EGARCH model better captures the leverage effects and volatility clusters and adapts to market changes that describes the conditional relations between asset returns. Many research has been utilized this model to predict the covariance and correlation multivariate assets [101, 102, 103].

On the other hand, LSTM networks have significant merits in covariance and correlation prediction in time series. Their ability to capture the long-term dependence of the time series and non-linear series and forecast dynamic relationships over time makes them particularly useful [104]. The advantages of DCC asymmetric GARCH with the combination long short-term memory enhances the performance of modeling and predicting the correlation of stock market indices as well as agricultural commodities.

Chapter 3

Preliminary Concepts

3.1 Basic Tests in Return Data

3.1.1 Augmented Dickey-Fuller (ADF) test

To model, predict, and make decisions using statistical methods, the return series should be stationary so that past behavior can serve as an adequate predictor of future trends. The mathematical definition of stationarity depends on the type under consideration: strict stationarity or weak stationarity.

A time series X_t is said to be strictly stationary if $P(X_{t_1}, X_{t_2}, ..., X_{t_k}) = P(X_{t_1+h}, X_{t_2+h}, ..., X_{t_k+h})$ that is k time series data and the same number of time series k having lag h have the same probability distribution. While,Weak stationarity requires that the first two moments (mean and variance) of the time series are time-invariant. Mathematically:

$E[X_t] = \mu$, constant mean

$Var(X_t) = \sigma^2$, constant variance

$Cov(X_t, X_{t+h})$=$\gamma(h)$, the covariance between X_t and X_{t+h} is a function of the time difference h not on time t.

$$\Delta y_t = \alpha + \beta_t + \gamma y_{t-1} + \sum_{i=1}^{p} (\sigma_i \Delta y_{t-i} + \varepsilon_t, \tag{3.1.1}$$

where α is constant (drift), β_t is the trend intercept, γ is the coefficient of the lag of the series , σ_i the coefficient of the lagged difference to higher order and ε_t is the white noise error term. And the ADF test statistic is expressed as:

$$\tau = \frac{\hat{\gamma}}{SE(\hat{\gamma})}, \tag{3.1.2}$$

where SE is the standard error and when the magnitude of τ greater than critical value (at levels) it is stationary otherwise non stationary.

The distribution of returns in financial time series data is not usually characterized by normality but fat-tails, high peakedness and skewness. The kurtosis and skewness are defined as:

$$K(R) = E[\frac{(R-\mu)^4}{\sigma^4}] \tag{3.1.3}$$

$$S(R) = E[\frac{(R-\mu)^3}{\sigma^3}] \tag{3.1.4}$$

where μ and σ are the mean and standard deviation of the return price. When $K(R)$ is greater than three it indicates that the distribution has heavy tail properties. Negative values of $S(R)$ shows skewed to the left and positive value indicates the skewed to the right.

3.1.2 Breusch–Godfrey test

The Breusch–Godfrey test based on the Lagrange Multiplier (LM) is a statistical test used to detect the presence of serial correlation (autocorrelation) in the residuals of a regression model. The test statistic is:

$$\begin{aligned} LM &= (n-p)R^2 \\ &= (n-p)\left(1 - \frac{RSS}{TSS}\right) \\ &= (n-p)\left(1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}\right), \end{aligned} \tag{3.1.5}$$

where n is the sample size, p is the number of lags tested, R^2 is the coefficient of determination from the auxiliary regression, $\bar{y}$ is the mean of observed values, y_i is the observed value, $\hat{y}_i$ is the predicted value.

If the computed LM statistic is greater than the critical value from the chi-square distribution ($LM > X^2_{(p,\alpha)}$) reject the null hypothesis (there is serial auto correlation of residuals). If $LM \leq X^2_{(p,\alpha)}$, fail to reject the null hypothesis (there is no evidence of serial auto correlation of residuals).

3.1.3 Jarque-Bera test

Jarque-Bera test is a normality test based on skewness and kurtosis. The test statistic is defined [72]

$$JB = \frac{n}{6}\left(\frac{(K-3)^2}{4} + S^2\right) \tag{3.1.6}$$

The assumption of normality under H_0 is rejected when the p-value of JB test is less than the significance level. (normal distribution when it is less than 5%)

3.1.4 ARCH-LM test

The presence of heteroscedasticity in the time series data can be checked using ARCH-LM test (ARCH effect). Engle (1984) [15] proposed Lagrange Multiplier tests to verify whether the standardized residuals exhibit ARCH effect or not in the time series data.

There are two ways to test the ARCH effect:

Ljung-Box Q statistic or Lagrange Multiplier (LM) which are computed as:

$$Q = T(T+2)\sum_{k=1}^{h}\frac{r_k^2}{T-k} \sim X^2(h), \tag{3.1.7}$$

Where T number of observations, h number of lags tested, r_k the autocorrelation at lag k.

$$LM = T.R^2 \sim X^2(h), \tag{3.1.8}$$

where R^2 is the coefficient of determination from the square of the residual regression. From these tests when the square of residuals auto correlated with the previous square of residuals in lag k or p value is significant reject the null hypothesis and there is an ARCH effect otherwise fail to reject the null hypothesis which indicates that no ARCH effect.

3.1.5 Diebo-Mariano test

The Diebold-Mariano (DM) test is a statistical test used to compare the predictive accuracy of two competing forecasting models. It tests the null hypothesis that both models have equal forecast accuracy based on the loss differential of their forecast errors. The test accounts for autocorrelation in forecast errors using robust variance estimators [136]. And the statistical computation is expression as:

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t, \quad where \quad d_t = L(e_{1,t}) - L(e_{2,t}) \tag{3.1.9}$$

$\bar{d}$ is error difference of the models, T is the number of samples.

$$DM = \frac{\bar{d}}{\sqrt{\hat{S}_d/T}}, \quad where \quad \hat{S}_d = \gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k \quad \& \quad \gamma_k = \frac{1}{T}\sum_{t=k+1}^{T}(d_t - \bar{d})(d_{t-k} - \bar{d}) \tag{3.1.10}$$

3.1.6 Information Criterion

The determination of ARCH and GARCH orders to build the the conditional variance equation requires a conventional criteria. The two most widely used information criteria are Akaike information criterion (AIC) and Bayesian information criterion (BIC), and their respective algebraic expressions are:

$$AIC = \ln\sigma^2 + \frac{2K}{T} \tag{3.1.11}$$

$$BIC = \ln\sigma^2 + \frac{K}{T}\ln T, \tag{3.1.12}$$

where σ^2 is the variance of the residuals, T is the sample size, and K is the total number of parameters. Particularly $K = p+q+1$ for GARCH (p,q) models. Based on this the model with smallest values of AIC and BIC is considred to be the best one among the candidate models.

3.2 Evaluation Metrics

The forecasting performance of the model is measured by comparing the forecasted with Actual or proxy for actual in the out sample data using the metrics mean square error (MSE), root mean square error (RMSE) and the mean absolute error (MAE). MSE describes the average of the difference squares in the observed and predicted volatility or correlation. RMSE indicates the magnitude of the error in square root of the average square of the predicted and observed values, where as MAE shows the magnitude of absolute difference average of predicted and observed value. These equations are as follows:

Mean square error

$$\begin{aligned} MSE &= \frac{1}{n}\sum_{n=1}^{n}(\hat{\sigma}_t - \sigma_t)^2 \\ MSE &= \frac{1}{n}\sum_{n=1}^{n}(\hat{\rho}_t - \rho_t)^2 \end{aligned} \tag{3.2.1}$$

Root mean square error

$$
\begin{aligned}
RMSE &= \sqrt{\frac{1}{n}\sum_{n=1}^{n}(\hat{\sigma}_t - \sigma_t)^2} \\
RMSE &= \sqrt{\frac{1}{n}\sum_{n=1}^{n}(\hat{\rho}_t - \rho_t)^2}
\end{aligned} \tag{3.2.2}
$$

Mean absolute error

$$
\begin{aligned}
MAE &= \frac{1}{n}\sum_{n=1}^{n} \mid \hat{\sigma}_t - \sigma_t \mid \\
MAE &= \frac{1}{n}\sum_{n=1}^{n} \mid \hat{\rho}_t - \rho_t \mid
\end{aligned} \tag{3.2.3}
$$

Where $\hat{\sigma}_t$ and $\hat{\rho}_t$ are the forecasted volatility and forecasted correlation respectively, σ_t and ρ_t are the observed (actual) volatility and correlation value in time period t and n is number of out sample periods.

3.3 Distributions

3.3.1 Univariate distribution

For maximum loglikelihood estimation the log return follows three conditional error distribution functions, namely normal (N), standardized student-t (ST) and generalized errror distribution (GED), and their Mathemaical expressions are:

$$
f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} exp^{-}\frac{(x-\mu)^2}{2\sigma^2} \tag{3.3.1}
$$

And the log-likelihood expression for each x_i for $i = 1, 2, ..., n$

$$
\begin{aligned}
\ell(\mu, \sigma^2) &= \sum_{i=1}^{n} \log f(x_i \mid \mu, \sigma^2) \\
&= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2
\end{aligned} \tag{3.3.2}
$$

Taking the derivative with respect to μ and set to zero:

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i - n\mu\right) \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.3.3}$$

Next, taking the derivative with respect to σ^2

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \mu)^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 \tag{3.3.4}$$

hence the maximum log-likelihood estimation of the mean and variance are the sample mean and sample variance respectively for normal distribution.

$$f(x \mid \mu, \sigma^2, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi\sigma^2}}(1 + \frac{(x-\mu)^2}{\nu\sigma^2})^{-\frac{\nu+1}{2}}, \tag{3.3.5}$$

where Γ is the usual gamma function and $\nu > 2$ is the number of degree of freedom for student-t distribution.
The log-likelihood with $\theta = \mu, \sigma^2, \nu$

$$\ell(\theta) = n\log\Gamma\left(\frac{\nu+1}{2}\right) - n\log\Gamma\left(\frac{\nu}{2}\right) - \frac{n}{2}\log\left(\nu\pi\sigma^2\right) - \frac{\nu+1}{2}\sum_{i=1}^{n}\log\left(1 + \frac{(x_i-\mu)^2}{\nu\sigma^2}\right) \tag{3.3.6}$$

For generalized error distribution, we have

$$f(x; \mu, \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(\frac{1}{\beta})} exp^{-}(\frac{\mid x - \mu \mid}{\alpha})^{\beta}, \tag{3.3.7}$$

where x is the random variable. μ is the location parameter, indicating the mean of the distribution. α is the scale parameter, controlling the spread of the distribution. β is the shape parameter, controlling the shape of the tails. When $\beta = 2$ the distribution reduces to the normal distribution. This distribution allows for skewness and heavier or lighter tails than the normal distribution, making it more flexible for modeling various data distributions [22, 44, 39].

3.3.2 Multivariate distribution

Multivariate Gaussian Distribution and its loglikelihood is expressed as:

$$f(\mathbf{x} \mid \mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\right), \tag{3.3.8}$$

where: Σ is the covariance matrix and positive definite, μ is the mean vector of the assets, d is the dimension of the random vector x.

The likelihood function for the sample is:

$$\begin{aligned} L(\mu, \Sigma) &= \prod_{i=1}^{n} f(x_i \mid \mu, \Sigma) \\ &= (2\pi)^{-\frac{nd}{2}} |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)\right) \end{aligned} \tag{3.3.9}$$

The corresponding log-likelihood of this distribution is:

$$\log \ell(\mu, \Sigma) = -\frac{nd}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i-\mu)'\Sigma^{-1}(\mathbf{x}_i-\mu) \tag{3.3.10}$$

To find the maximum log-likelihood estimations, differentiate first with respect to μ and set to zero:

$$\frac{\partial \ell}{\partial \mu} = \sum_{i=1}^{n} \Sigma^{-1}(x_i-\mu) = 0 \implies \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.3.11}$$

Which implies the MLE of μ is the sample mean.

Next, differentiate with respect to Σ^{-1} (using matrix calculus):

$$\frac{\partial \ell}{\partial \Sigma^{-1}} = \frac{n}{2}\Sigma - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)(x_i-\mu)' = 0 \tag{3.3.12}$$

After rearrangement it becomes:

$$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})(x_i - \hat{\mu})' \tag{3.3.13}$$

which is the maximum log-likelihood covariance matrix is the sample covariance matrices.

Multivariate Student's t-Distribution:

$$f(\mathbf{x} \mid \mu, \Sigma, \nu) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\nu^{d/2}\pi^{d/2}|\Sigma|^{1/2}}\left[1 + \frac{1}{\nu}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\right]^{-\frac{\nu+d}{2}}, \tag{3.3.14}$$

where: ν is the degree of freedom. As $\nu \to \infty$, it reduced to multivariate Gaussian distribution. The log likelihood of this distribution is expressed as:

$$\log \mathrm{L}(\mu, \Sigma, \nu) = \sum_{i=1}^{N}\left[\log\Gamma\left(\frac{\nu+d}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{d}{2}\log(\nu\pi) - \frac{1}{2}\log|\Sigma| - \frac{\nu+d}{2}\log\left(1 + \frac{1}{\nu}\delta_i\right)\right] \tag{3.3.15}$$

where:

$$\delta_i = (\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)$$

Because the maximum log-likelihood equations are nonlinear and involve the gamma and digamma functions, the parameters are estimated iteratively using an Expectation-Maximization (EM) algorithm. Generalized Gaussian Distribution (MGGD)

$$f(\mathbf{x}; \mu, \Sigma.\beta) = \frac{\beta}{2^{d/2}\Gamma\left(\frac{d}{2\beta}\right)|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}\left[(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\right]^{\beta}\right) \tag{3.3.16}$$

when $\beta = 1 \implies f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\right)$, Since $\Gamma(\frac{1}{2}) = \pi^{\frac{1}{2}}$

The log likelihood description of this generalized multivariate distribution is:

$$\log \mathrm{L}(\mu, \Sigma, \beta) = N \log \left(\frac{\beta}{2^{d/2}\Gamma\left(\frac{d}{2\beta}\right)|\Sigma|^{1/2}} \right) - \frac{1}{2}\sum_{i=1}^{N} \left[(\mathbf{x}_i-\mu)'\Sigma^{-1}(\mathbf{x}_i-\mu)\right]^{\beta} \quad (3.3.17)$$

3.4 Deep Learning Neural Network

Artificial neural network (ANN) is one of the best known machine learning algorithms. It is mimc to the neuron in the human brain. Artificial neurons are connected to each other by weights and solve problems by learning.

Deep Neural Networks (DNNs) are a class of machine learning models based on artificial neural networks (ANNs) that consist of multiple layers between the input and output layers. These networks are designed to automatically learn patterns and representations from data by utilizing many layers of interconnected nodes (or neurons). The term “deep” refers to the presence of multiple hidden layers in the network, which enables DNNs to learn hierarchical features and complex patterns in data [125].

3.4.1 Structure of Deep Neural Networks

A typical deep neural network consists of several key components:

Input Layer: This layer takes in the raw data, which could be in the form of images, text, time series data, etc. Each neuron in the input layer represents one feature of the input data.

Hidden Layers: These are the layers that exist between the input and output layers. A deep neural network usually has more than one hidden layer. Each layer in a DNN learns progressively more complex features. Neurons in these layers are connected to neurons in the previous and next layers. The number of hidden layers is a key factor in the "depth" of the network, and it determines the complexity of the model [125].

Neurons (Nodes): Each neuron in the hidden layers performs a simple computation. It receives input from the previous layer, processes it with a weighted sum, applies an activation function, and then passes the result to the next layer. The weight and bias associated with each neuron are learned during the training process. The mathematical expression of neural network in the function f is defines as:

$$f : \mathrm{R}^n \to \mathrm{R}$$
$$f(x_1, x_2, \ldots, x_n) = \rho\left(\sum_{i=1}^{n} x_i w_i + b\right), \tag{3.4.1}$$

where $w_1, w_2, \ldots w_n \in \mathrm{R}$ are the weights, $x_1, x_2, \ldots x_n$ are inputs, $b \in \mathrm{R}$ is the bias and ρ is the activation function. The activation functions support the neuron to train the non linear and complex patterns. Some of the most common activation functions are:

- Sigmoid or logistic function that maps any input value to a range between 0 and 1. It is often used in binary classification tasks and mathematically:

  $$\sigma(x) = \frac{1}{1 + \exp(-x)}, \tag{3.4.2}$$

  where x is the input.

- Rectified linear unit (ReLU) function that converts all negative values to zero and keeps non negative values unchanged. It helps to prevent the vanishing gradient problem in deep networks.

  $$ReLU(x) = max\{0, x\} \tag{3.4.3}$$

- Tangent hyperbolic (Tanh) function that maps inputs to a range between -1 and 1. It is similar to the Sigmoid function but with a wider output range.

  $$tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}, \tag{3.4.4}$$

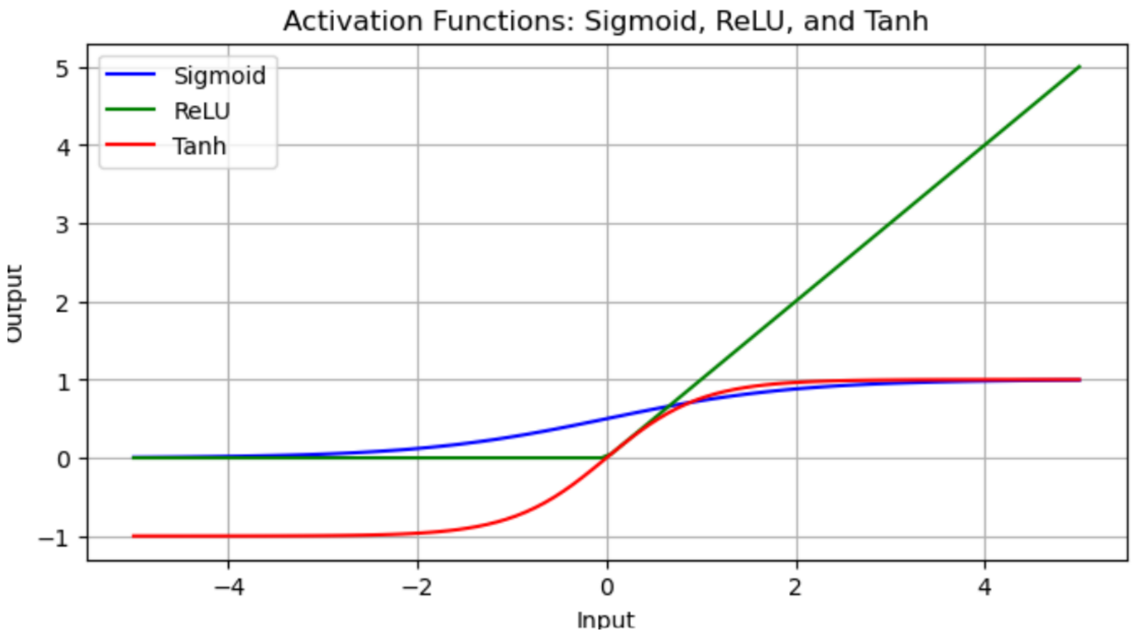

FIGURE 3.1: Activation functions

The weights and biases in each layer are connected to neurons in the adjacent layers, and each connection of the weights are adjusted during training. Each neuron also has a bias term in which the unobservable factor that are not considred as an input. The training process involves optimizing these weights and biases to minimize a loss function.

The loss Function (cost function) is used to quantify the difference between the predicted outputs and the actual target values. The goal of training is to minimize the loss function and its optimization is described as:

$$z = \arg\min_{w \in W} \sum_{i=1}^{N} \ell\big(F(x_i; w), y_i\big), \tag{3.4.5}$$

where z is is the model parameter in W that minimize the loss function l, F is the predictive function and y is the actual or the target value. Depending on the kind of neuron and purpose the model optimization varies. In the case of time series forecasting it is error metrics.

3.4.2 Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are a class of deep learning models capable of capturing sequential dependencies in time series data as an input to future prediction purpose.

The Mathematical description of input to hidden layer equation is given as:

$$h_t = f(Bx_t + Ch_{t-1} + b) \tag{3.4.6}$$

$$y_t = f(Ah_t + b_1) \tag{3.4.7}$$

where h_t is hidden layer at time t, f is activation function, B is the weight of input to hidden layer, x_t is input at time t, h_{t-1} is hidden layer at time $t-1$, b bias for hidden layers,y_t out put vector at time t, A is the weight to output layer and b_1 bias for the out put layer [126].

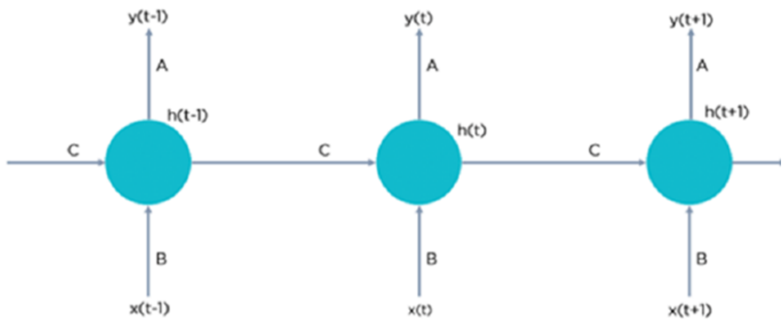

FIGURE 3.2: Structure of RNN

3.4.3 Back propagation

Backpropagation is a fundamental algorithm used to train deep learning neural networks by minimizing the error between predicted outputs $\hat{y}$ and actual targets y by the loss or cost function. It consists of two main phases: forward error propagation and backward error propagation.

Forward propagation is the first phase where the input data x is fed through the network layer by layer. Each neuron in a layer computes a weighted sum of inputs from the previous layer, applies an activation function (such as sigmoid, ReLU, or tanh), and passes the output to the next layer. This process continues until the network produces an output at the final layer. The output is then compared to the target (true value) to compute the error, typically by using a loss or cost function that quantifies the difference.

Backward propagation error is the second phase where the error computed at the output is propagated backward through the network. Using the chain rule, the algorithm calculates gradients of the loss function with respect to each weight in the network. This gradient tells how much each weight contributed to the overall error and how it should be adjusted to reduce the error.

Weights in each layer are updated typically using gradient descent or its variants. The error propagates layer-by-layer in reverse order from the output back to the input layer, and weights get adjusted in proportion to their contribution to the error. This enables the network to improve its predictions iteratively during training by minimizing the loss function. One complete forward and backward phase for each weight update is called an epoch.

Using the gradient descent method at each iteration the weight are updated as:

$$w_j^{(t+1)} = w_j^{(t)} - \lambda \frac{\partial l}{\partial (w_j)}, \tag{3.4.8}$$

where $\frac{\partial l}{\partial (w_j)}$ is the gradient, and λ is called the learning rate which contains the rate at which the model learns. Thus, the error minimized by $\frac{\partial l}{\partial w_j}$. in particular, the back propagation error $\frac{\partial l}{\partial (w_j)}$ for every weight w_j at every connection/node the chain rule of weight up date from hidden to out put (W_{ho}), hidden to hidden (W_{hh}) and input to hidden layer (W_{xh})

$$\frac{\partial l}{\partial W_{ho}} = \sum_{t=1}^{T} \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot \frac{\partial \phi_o}{\partial W_{ho}} \tag{3.4.9}$$

$$\frac{\partial l}{\partial W_{hh}} = \sum_{t=1}^{T} \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot \frac{\partial \phi_o}{\partial H_t} \cdot \frac{\partial H_t}{\partial \phi_h} \cdot \frac{\partial \phi_h}{\partial W_{hh}} \tag{3.4.10}$$

$$\frac{\partial l}{\partial W_{xh}} = \sum_{t=1}^{T} \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot \frac{\partial \phi_o}{\partial H_t} \cdot \frac{\partial H_t}{\partial \phi_h} \cdot \frac{\partial \phi_h}{\partial W_{xh}}, \tag{3.4.11}$$

where ϕ_0 is the final state activation function, O_t is pre-activation out put, H_t the hidden state at time t, ϕ_h is the hidden state activation function usually either $tanh$ or $ReLU$.

3.4.4 Adam Optimizer

The Adam (Adaptive Moment Estimation) optimizer is an adaptive gradient descent optimization algorithm widely used in training recurrent neural networks due to its efficiency, low memory usage, and ability to adaptively adjust learning rates during training [127].

Chapter 4

Econometric and Deep Learning Hybrid Models for Univariate Volatility Forecasting

4.1 Introduction

This chapter has two sections: the first section focuses on modeling and forecasting the volatility of selected staple crops using asymmetric GARCH-type models, while the second section addresses the prediction of single stock return volatility through a hybrid model combining GARCH and deep learning techniques. In the first part we took teff, wheat, maize and barley retail price volatility in asymmetric GARCH models against univariate alternative error distributions and compare which model describes the best predictive model using selection criteria and error metrics. The inherent agricultural commodity price volatility is one of the major concern in countries whose dominant economy depends on agricultural products like Ethiopia.

To capture the potential asymmetries in price volatility, particularly the differential impact of positive and negative shocks, the GARCH, TGARCH (treshold GARCH), EGARCH (Exponential GARCH) or PARCH (power ARCH) models were employed. These models were estimated under normal (Gaussian), student's t and generalized distributional assumptions for the error terms. The empirecal analysis evaluates leverage effects, kurtosis and peristence in the volatility of staple crops that better fit models using AIC, BIC and log-likelihood comparison.

Part two presents results of modeling and forecasting volatility of a single stock market index return data analysis in GARCH extension and deep learning hybrid approaches. The main objective of this analysis is to assess the comparative forecasting performance of asymmetric GARCH models integrated with deep learning architectures particularly, Convolutional Neural Network (CNN), Long Short-Term Memory network (LSTM), and their hybrid configurations, namely CNN-LSTM and LSTM-CNN.

The asymmetric GARCH model has an ability to capture the stylzed volatility features like leverage effect and volatility clusters, but rely on the parametric assumptions and limited capacity to model nonlinear sequential temporal dependencies, which may restrict their effectiveness in capturing the complex patterns present in financial return series.

CNN has a feature of extracting special temporal patterns from the return series; where as, LSTM is leveraged for its strength in learning long-term dependencies and memory of past events. So the combined effect of these two enables to extract both the short term and long term patterns support the accuracy of forecasting the volatility of stock returns. The comparative analysis is based on performance evaluation metrics and information criteria. Historical daily closing index return is used for training, validation and testing the models.

4.2 Source and Description of Data

The data used for modeling price volatility of teff, wheat, maize and barely is taken from the Ethiopian Statistical Agency (ESA) record as the average retail price of Debre Markos market which is found in the most productive zone of teff, maize, wheat and barely in Ethiopia from September 2010 to August 2022. The quality of these staple crops varies as white, mixed and black (red), we take mixed one , and since the retail price is not the same for each retailer we use average weekly observations. The reason for selecting these staple crops has three folds: The first is their dominant in daily consumption in Ethiopia and the second is agricultural economic value and the third is worldwide demand.

To understand the price dynamic behavior and the price return of the time series it is necessary to plot the original series and its log return of teff, wheat, barely and maize price and their log returns. The price dynamics and returns are depicted in the Figure 4.1 and Figure 4.2. For better description of volatility we use return of retail price at time t and it is calculated as:

$$R_t = ln(\frac{P_t}{P_{t-1}}) \tag{4.2.1}$$

Figure 4.1, The retail price of teff, wheat, barely and maize in which all of them are non stationary while the return series of teff, wheat, barely and maize on Figure 4.2, shows stationary. However the plot of the figure is not sufficient but it give us a clue and we check by unit root test.

Moreover, from Figure 4.2, the variability of the changes, small changes are followed by small changes and large changes are followed by large changes which shows that the existence of volatility clustering in all of the crops. The results in Table 4.1, shows that all the return prices are stationary since the absolute value of the test statistic exceeds the critical value at the 5% significance level; thus, it is mean reverting.

The other data used in this chapter is stock price data of S&P 500 extracted from Yahoo

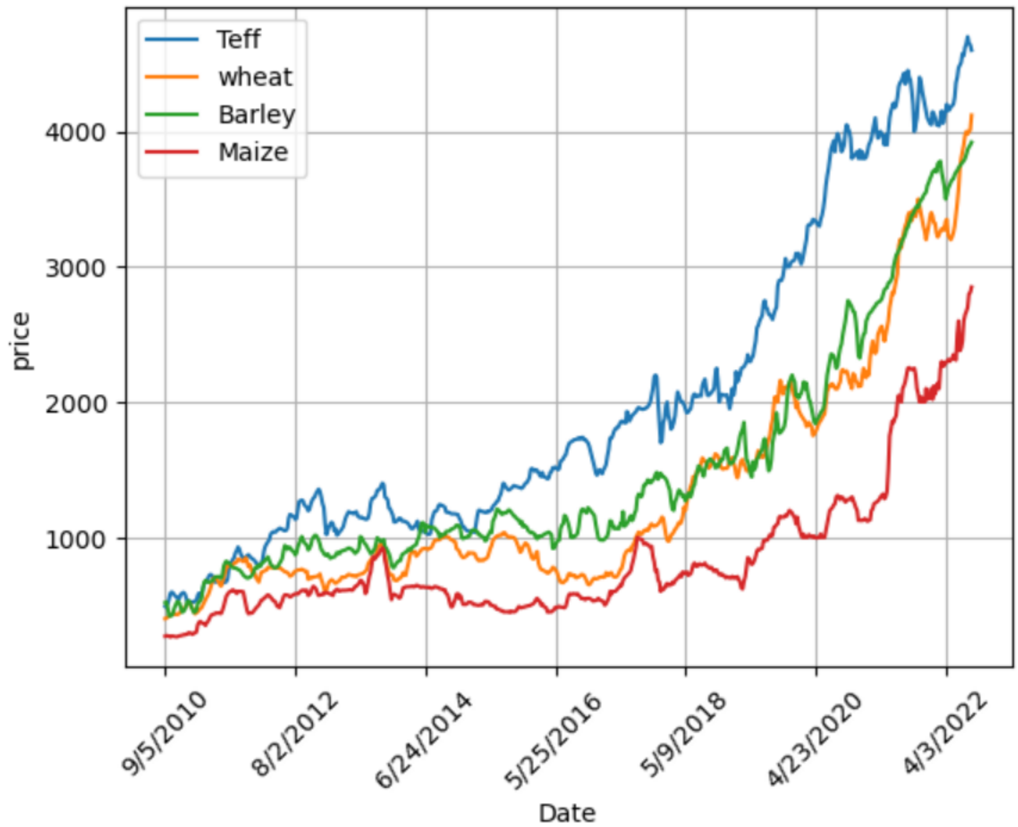

FIGURE 4.1: Retail price of crops

finance (GSPC) database that covers from 6 May 2019 to 6 May 2024 trading days. The time series plots of S&P 500 stock price is not stationary but its logarithmic return is stationary which are depicted in Figure 4.3 and Figure 4.4 which is justified using unit root test.

TABLE 4.1: Unit root tests for Log-return series at level of selected cereals

		ADF test critical values			
Variables	t-statistics	1%	5 %	10 %	p-value
Teff return	-16.2573	-3.4415	-2.8664	-2.5694	0.0000
Wheat return	-19.3078	-3.4415	-2.8664	-2.5694	0.0000
Barely return	-15.9849	-3.4415	-2.8664	-2.5694	0.0000
Maize return	-12.9560	-3.4415	-2.8664	-2.5694	0.0000

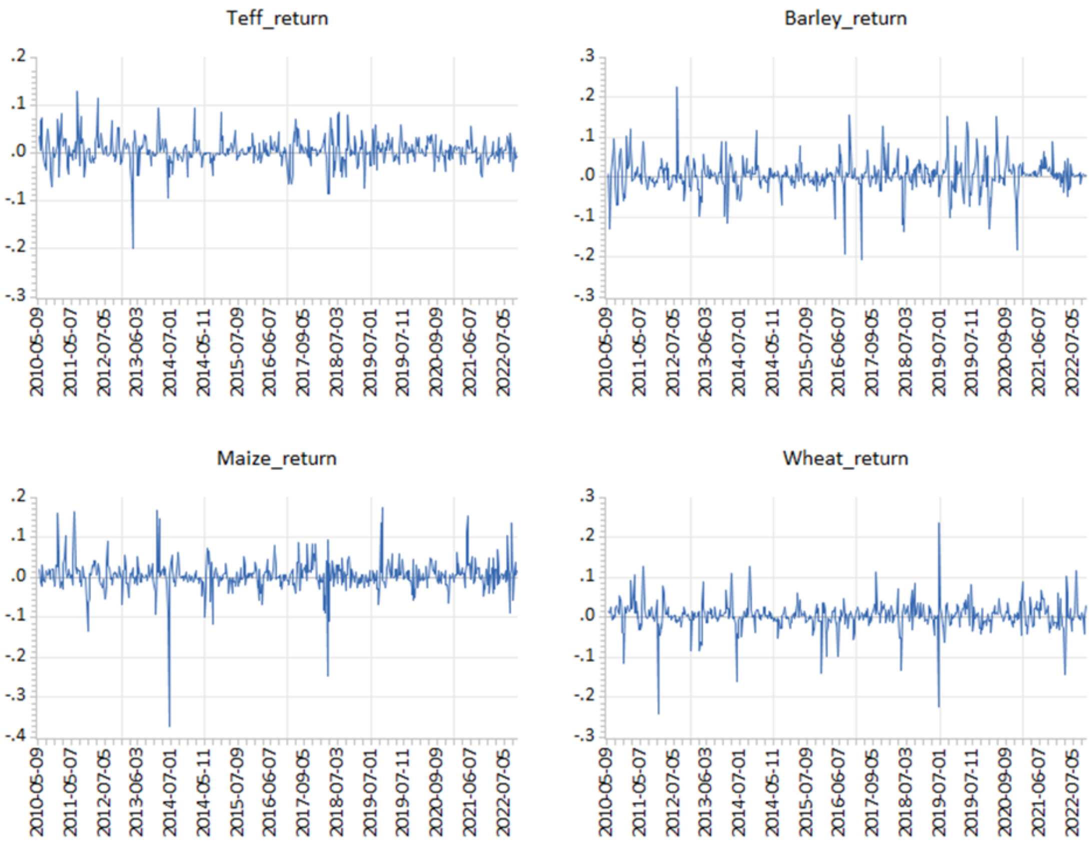

FIGURE 4.2: Log price return of crops in the study period

TABLE 4.2: Unit root tests for S&P 500 return

ADF test critical values					
Variables	t-statistics	1%	5 %	10 %	p-value
S&P500	-24.6854	-3.4353	-2.8636	-2.5694	0.0000

From, Table 4.2 clearly shows that the magnitude of S&P 500 stock return in t-statistics is greater than the magnitude at 5 % level which indicates that the return is stationary; thus, it is mean reverting.

From Table 4.3 the LM - statistic of teff, wheat, maize and barely returns respectively are 24.96, 36.35, 6.85 and 10.05 respectively where as the p- values are 0.0000, 0.0000, 0.0088 and 0.0015 which are statistically significant. This shows that the null hypothesis

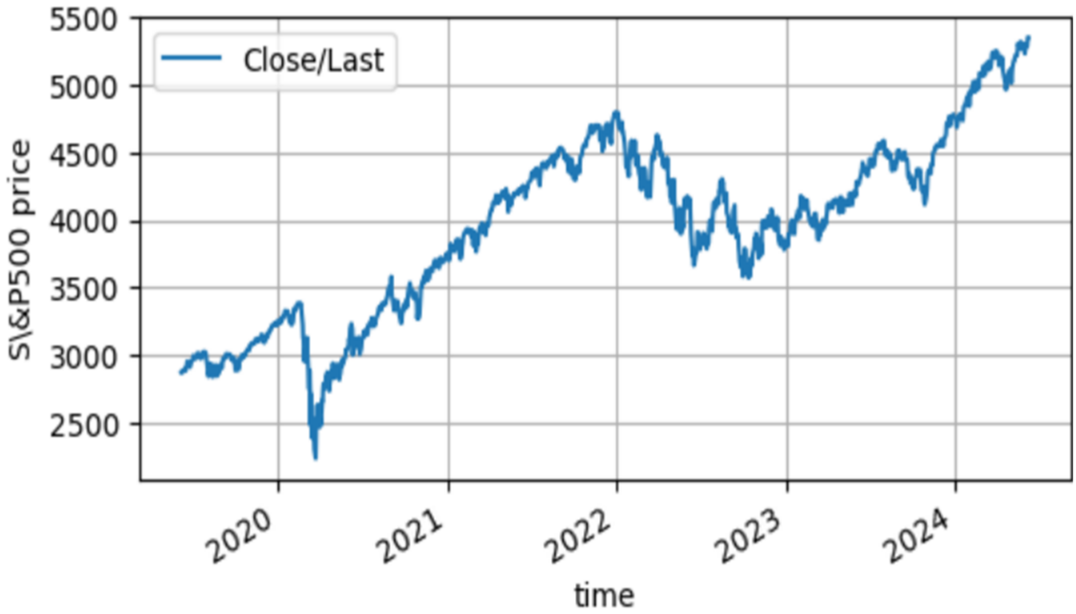

FIGURE 4.3: Daily S&P 500 stock price

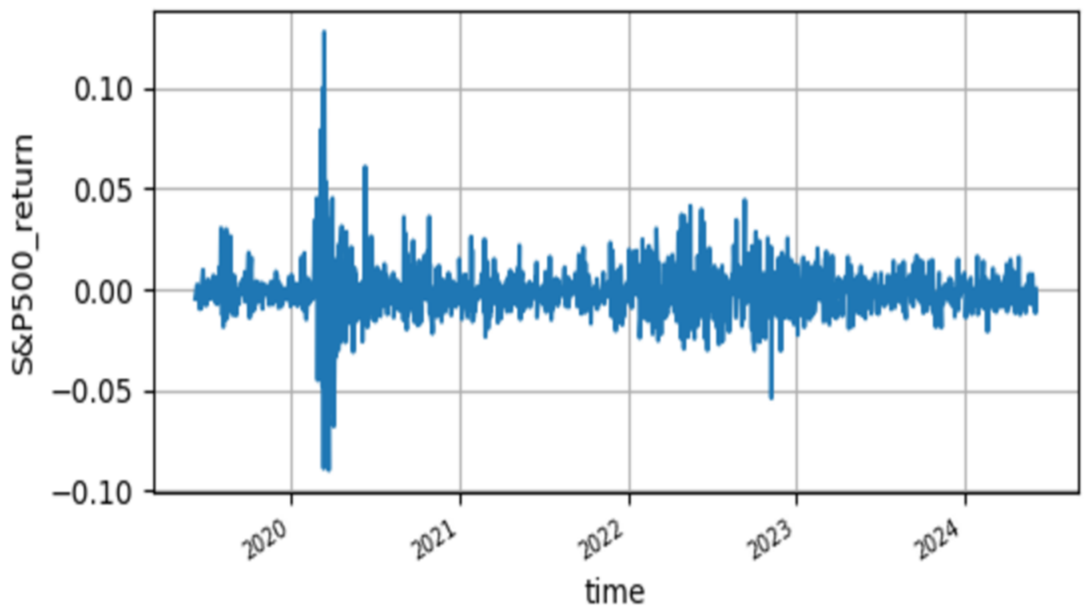

FIGURE 4.4: S&P 500 stock daily return

(Homoskedasticity) is rejected and reveals the presence of heteroskedasticity (time varying volatility). Therefore it is better to estimate GARCH family models. To determine

TABLE 4.3: Heteroskedasticity test of return price for selected cereals

	Teff-return	Wheat-return	Maize-return	Barley-return
F-statistic	26.0142(**0.0000**)	38.6794(**0.0000**)	6.9055(**0.0088**)	10.1938(**0.0015**)
Obs*R-squared	24.9677(**0.0000**)	36.3524(**0.0000**)	6.8469(**0.0089**)	10.0501(**0.0015**)

Values in parenthesis are p-values.

TABLE 4.4: Heteroskedasticity test of S&P 500 return

F-statistic	P-value	Obs*R-squared	P-value
46.0142	0.0000	38.6794	0.0000

the order of GARCH type family models Akaike information criterion (AIC), Schwarz information criterion (SIC) and error distribution are used for each models.

Similarly from Table 4.4 the LM - statistics of S&P 500 stock price return is 46.014 and p-value 0 which is statistically significant. This shows that the null hypothesis (homoskedasticity) is rejected and reveals the presence of heteroskedasticity (time varying volatility). Therefore, GARCH family models are better fitting for forecasting the volatility of S&P 500 stock price .

4.3 Realized Volatility

This is the actual observed volatility over a specific interval, calculated using high-frequency trading data. It is model-free and an aggregates actual return movements within the day. It is often used to assess the accuracy of volatility forecasts [116, 124]. The realized variance (RV) is mathematically expressed as:

$$RV = \sum_{i=1}^{m} r_{t,i}^2, \tag{4.3.1}$$

where r is the return, m is the number of intraday return during day t. A 5 minute intraday RV is the most standard and most popular in practical applications [119, 117, 118]. As volatility is the standard deviation of the return (σ), the daily realized volatility is expressed as:

$$\sigma = \sqrt{RV} \tag{4.3.2}$$

Whenever there is only daily historical data, the proxy for daily realized volatility is the magnitude or absolute value of the return. However, it is less precise compared to realized volatility measures calculated with intraday high-frequency data.

Realized correlation measures the time-varying correlation between two asset returns using high-frequency data. It is formulated by using the realized covariance and the individual realized volatilities of two assets. The mathematical expression of this realized covariance is expressed as:

$$RCov_t = \sum_{i=1}^{M} r_{1,t,i} \cdot r_{2,t,i} \tag{4.3.3}$$

The **realized correlation** is described as:

$$\rho_t^R = \frac{RCov_t}{\sqrt{RV_{1,t}} \cdot \sqrt{RV_{2,t}}} = \frac{\sum_{i=1}^{M} r_{1,t,i} r_{2,t,i}}{\sqrt{\sum_{i=1}^{M} r_{1,t,i}^2} \cdot \sqrt{\sum_{i=1}^{M} r_{2,t,i}^2}}, \tag{4.3.4}$$

where $RCov_t$ is the realized covariance of two assets during day t, $r_{1,t}$ and $r_{2,t}$ are the returns of the assets at day t and ρ_t^R is the realized correlation at day t. This correlation has limitation that it is noisy and the assets may not have the same intraday frequency data due to the variation of opening and closing time in different countries .

4.4 Univariate GARCH family Models

4.4.1 Autoregressive Conditional Heteroskedasticity(ARCH model)

An ARCH model is one of the classical models in time series data for analyzing and forecasting volatility. It is originally proposed by Engle (1982) [15], where the prediction accuracy of conditional variance depends on the previous historical shocks and this is expressed as:

$$r_t = \mu + \varepsilon_t; \quad \varepsilon_t = \sigma_t z_t \tag{4.4.1}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, \tag{4.4.2}$$

where r_t is the return, μ is the mean, $z_t \sim i.i.d(0,1), \alpha_0 > 0, \alpha_i > 0$, σ_t^2 is the conditional variance, ε_t is the innovation of an asset, σ_t is the volatility of the series at time t and, z_t is a white-noise. The equation

$$\sum_{i=1}^{q} \alpha_i < 1 \tag{4.4.3}$$

ensures the conditional variance process is covariance stationary (mean reversion) and the measure of persistance depend on the closer to unity. Equation 4.4.2 often requires many parameters and a high order of the ARCH term q to capture the dynamic behavior of conditional variance that leads to over predict the volatility and fails to capture the leverage effect due to its symmetric variance specification [28].

4.4.2 Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model.

In GARCH the the current conditional variance is expressed not only the previous shocks but also the the previous conditional variance and mathematically expressed as:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{4.4.4}$$

$\alpha_0 > 0, \alpha_i > 0, q > 0, p > 0$

when p = 1 and q = 1, it is called GARCH (1,1) and the conditional volatility of Equation 4.4.4 becomes:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{4.4.5}$$

The condition that covariance stationary for GARCH is

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1. \tag{4.4.6}$$

The limitation of GARCH is that it still fails to capture the leverage effect due to its symmetric distribution. This model is proposed by[12]

4.4.3 Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH)

This model is proposed by Nelson (1991) to take into account the leverage effects of price fluctuation on conditional variance. This means that a negative shock (bad news) can have a greater impact on volatility than a positive shock (good news) of the same magnitude. [55]. The conditional variance in EGARCH is given as:

$$log\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \frac{\alpha_i \mid \varepsilon_{t-i} \mid}{\sigma_{t-i}} + \sum_{i=1}^{p} \frac{\gamma_i \varepsilon_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{q} \beta_j log\sigma_{t-j}^2 \tag{4.4.7}$$

Where p, and q are the ARCH and GARCH orders respectively. As log of the variance σ_t^2 makes the leverage effect exponential. When $\gamma_i = 0$, regardless of the sign of ε_{t-i} the model is symmetric; thus no leverage effect. When ε_{t-i} is positive (good news) the total effect of ε_{t-i} is $(\alpha_i + \gamma_i)\varepsilon_{t-i}$ on the contrary, ε_{t-i} is negative (bad news) the total effect of ε_{t-i} is $(\alpha_i - \gamma_i)\varepsilon_{t-i}$. Bad news can have a larger impact on volatility, and the value of γ_i would be expected to be negative. Another advantage of the EGARCH model over the basic GARCH model is that the conditional variance σ_t^2 is guaranteed to be positive regardless of the values of the coefficients in the above Equation 4.4.7 because the logarithm of σ_t^2 instead of σ_t^2 itself is modeled [35].

The Covariance-stationarity of EGARCH log-variance process requires that the AR part be stable. Particularly, when p =1, $\mid \beta_1 \mid < 1$ and for a common assumption;

$$\sum_{j=1}^{p} \beta_j < 1, \tag{4.4.8}$$

for every $\beta_j \geq 0$.

4.4.4 Threshold Generalized Autoregressive Conditional Heteroskedastic (TGARGH)

The other alternative for modeling the conditional variance with leverage effect is the threshold GARCH model which was proposed by [34] that allows asymmetric shocks to volatility with positive and negative shocks of equal size to have different impacts on volatility.

$$\sigma_t^2 = \alpha_0 + \sum_{i}^{p} (\alpha_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-i}^2 I) + \sum_{j}^{q} \beta_j \sigma_{t-j}^2 \tag{4.4.9}$$

where

$$I = \begin{cases} 1 & if \varepsilon_{t-i} < 0 \\ 0 & otherwise \end{cases}$$

In this model if γ is zero it become GARCH, if γ is negative then bad news decrease volatility which is not likely and it is expected to be positive and good news decrease volatility. The mean reversion or the covariance stationary condition for TGARCH is given as:

$$\sum_{i=1}^{p}(\alpha_i + 0.5\,\gamma_i) + \sum_{j=1}^{q}\beta_j < 1 \tag{4.4.10}$$

4.4.5 An asymmetric power ARCH model

An other extension of GARCH model is the power ARCH expressed as:

$$\sigma_t^{\delta} = \alpha_0 + \sum_{i}^{p}\beta_i\sigma_{t-i}^{\delta} + \sum_{j}^{q}\alpha_j(\mid \varepsilon_{t-j} \mid - \gamma\varepsilon_{t-j})^{\delta} \tag{4.4.11}$$

where γ is the asymmetric parameter and delta is the parameter of the power term.

4.5 Long Short- Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks have become increasingly important in applied financial econometrics due to their ability to model nonlinear temporal dependencies and long-memory structures commonly observed in financial time-series data. Financial volatility and cross-asset correlations exhibit complex dynamics such as persistence, clustering, and regime shifts, which are often difficult to capture using traditional econometric models such as GARCH or DCC-GARCH that rely on restrictive parametric assumptions. LSTM networks address these limitations through a gated recurrent architecture that allows the model to retain relevant historical information over long sequences while filtering out irrelevant noise, thereby enabling the learning of complex temporal relationships in financial markets.

Recent empirical studies demonstrate that LSTM-based approaches can improve volatility forecasting performance relative to classical models, particularly in environments characterized by nonlinear dynamics and structural changes. For instance, studies show that LSTM models produce more accurate realized volatility forecasts than conventional econometric benchmarks such as GARCH and HAR models in many financial settings.

Furthermore, hybrid models that integrate LSTM with traditional time-series frameworks or other deep learning architectures have shown strong predictive performance in modeling market volatility and financial risk during turbulent periods, including the COVID-19 crisis. These improvements arise because LSTM networks can capture volatility clustering, nonlinear interactions, and evolving dependence structures across assets, which are critical for accurate dynamic correlation forecasting and portfolio risk management. Consequently, LSTM-based models have become an important methodological tool in modern financial forecasting, particularly for Volatility Forecasting, Financial Market Prediction, forecasts of time-varying correlations, and Credit risk modeling. These capabilities make LSTM one of the most important deep learning models for financial time-series analysis.

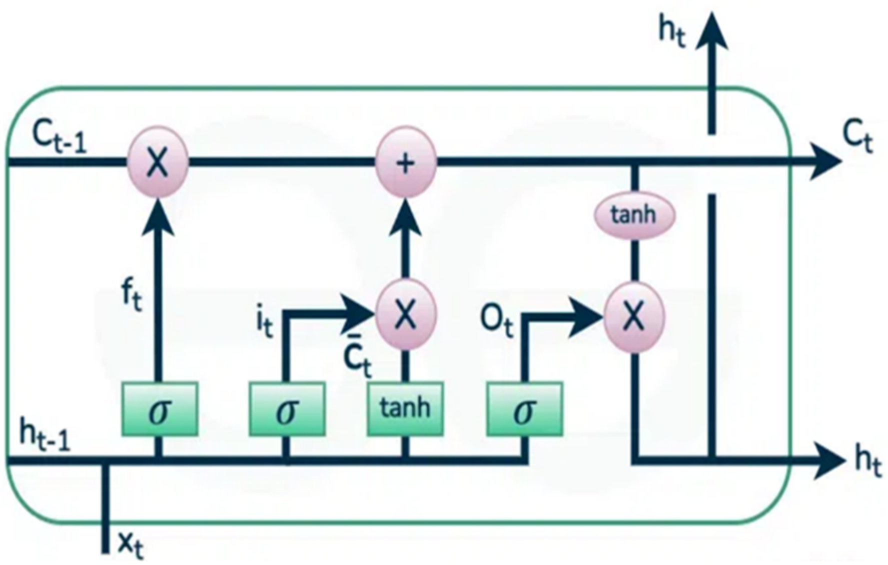

FIGURE 4.5: Architecture of LSTM

Each LSTM has four chain interacting communicative gates. LSTM networks address the problem of vanishing gradients of RNN by splitting in three inner-cell gates and build memory cells to store information in a long range context [106]. As we can see from Figure 4.5, a typical LSTM cell is configured mainly by four gates:

forget gate schedules when to forget the output data and thus selects the optimal time lag for the input sequence. The forget gate f_t deletes unimportant information from the previous time step and has an equation:

$$f_t = \sigma(w_f h_{t-1} + u_f x_t + b_f) \tag{4.5.1}$$

The Memory cell C_{t-1} takes input from the output of LSTM cell in the last iteration.

input and input modulation gates determine which information is significant through logistic (sigmoid) and tangent hyperbolic activation functions respectively and have equations:

$$\bar{i}_t = \sigma(w_i h_{t-1} + u_i x_t + b_i) \tag{4.5.2}$$

$$\bar{c}_t = tanh(w_c h_{t-1} + u_c x_t + b_c) \tag{4.5.3}$$

The point multiplication of forget gate with previous memory cell of the unit sum up with input gate forms memory cell of the unit and goes to tangent activation function and has equations:

$$i_t = \bar{i}_t * \bar{c}_t \tag{4.5.4}$$

$$C_t = C_{t-1} * f_t + i_t \tag{4.5.5}$$

output gate takes through sigmoid activation function and all results processed and generate the output.

$$o_t = \sigma(w_o h_{t-1} + u_o x_t + b_o) \tag{4.5.6}$$

finally,

$$h_t = o_t * tanh(C_t), \tag{4.5.7}$$

where w_f, u_f, w_i, u_i, w_c, u_c, w_o, u_o are weights of the forgate, input, input modulation and out put gates respectively. b the bias term. σ and $tanh$ are the sigmoid (logistic) and hyperbolic tangent activation functions respectively and h_t is the hidden state.

4.6 Convolutional Neural Network (CNN)

Recent advances in deep learning have demonstrated the effectiveness of one-dimensional Convolutional Neural Networks (1D-CNN) in time-series forecasting. Unlike traditional econometric models that rely on predefined lag structures, 1D-CNN applies convolutional filters over ordered temporal sequences to automatically extract local dependency patterns and nonlinear features. Empirical evidence shows that 1D-CNN performs competitively in electricity load forecasting [140],and it has an ability to capture nonlinear temporal dependencies and achieve superior prediction performance compared to traditional statistical baselines like ARIMA [139, 142]. Furthermore, hybrid architectures such as CNN-LSTM combine local feature extraction with long-term dependency learning, significantly improving forecasting accuracy in complex temporal environments [143, 145]. More recent studies incorporate attention mechanisms and interpretable CNN frameworks to enhance transparency and robustness in applied financial forecasting. Collectively, these findings provide strong methodological justification for employing 1D-CNN as either a standalone or hybrid component in volatility prediction model. The structure of CNN has input layer which accepts the inputs, convolutional layers that are kernels of the CNN which works in convolution operation (Hadmard's product), pooling layers which reduce the dimension by retaining the important imformation, fully connected layers (Dense) and flatten layers which changes into one dimension to the data and out put layers.

As we see from Figure 4.6 at convolutional layer the input data operates the hadmard product with the kernel or filter and gives several results and these sent to RelU activation

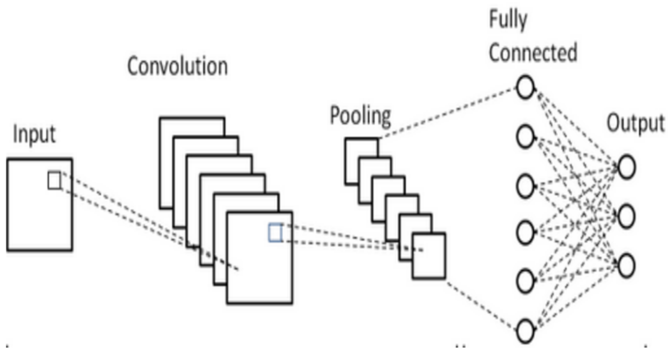

FIGURE 4.6: Architecture of CNN [108]

function and mathematically these expressed as;

$$y_t = (f * g)(t) = \sum_{a=1}^{k} f(a) * g(t-a) \quad \forall t \in [1,2,3,...,L], \tag{4.6.1}$$

where, y_t is the output of the convolution operation at position t, $f(a)$ is the kernel convolutioned at position a, k the kernel size, L is the length of the out put sequence, $*$ is convolution operation (Hadamard operation). To keep the number of out put as the number of input in the convolution operation we use padding. Similarly in in pooling we use unit stride. The operation of ReLU (σ) activation function as:

$$z_t = \sigma(y_t + b), \tag{4.6.2}$$

where z_t is the out put and b is the bias term. The next process of the ReLU out put value goes to the maximum pooling or average pooling and finally changed to flatten data at fully connected layer.

4.7 The Hybrid Models

The purpose of the hybrid models is to improve the volatility prediction performance of the traditional GARCH family models.

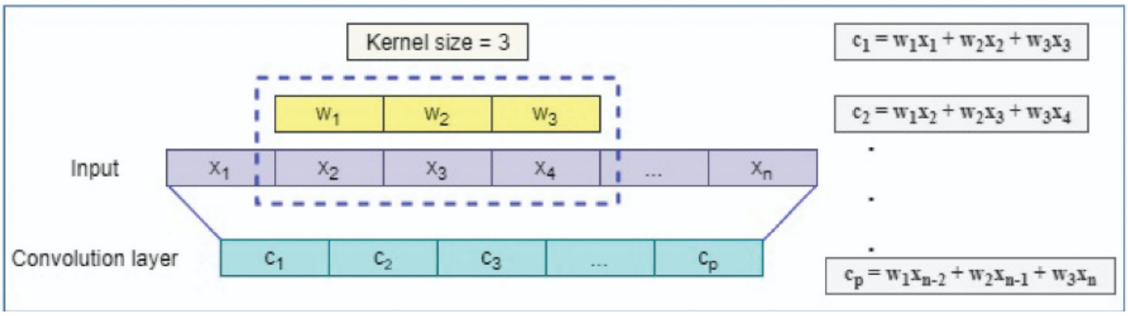

FIGURE 4.7: Convolution operation (1D) [114]

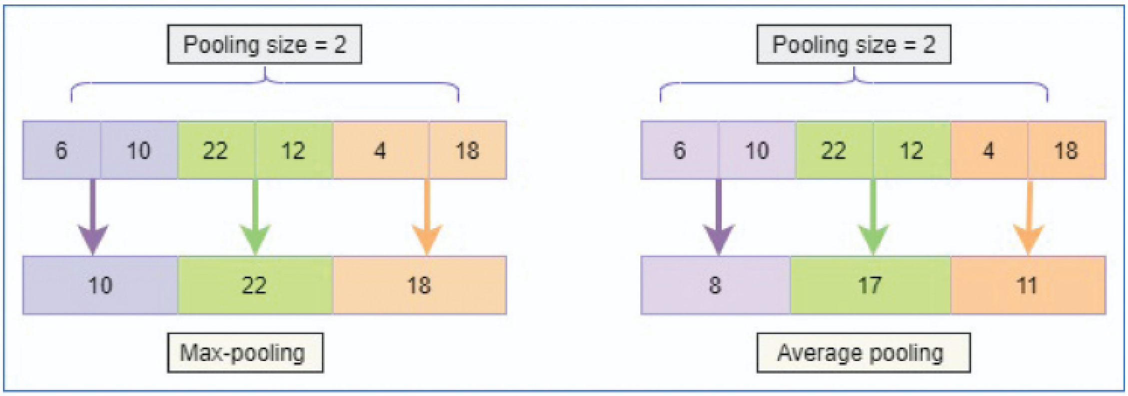

FIGURE 4.8: Pooling (1D) [115]

Univariate Hybribrid Models

The features of volatility like clusters makes the inputs of deep leaning models increase the prediction accuracy on result and estimated parameters [49, 69, 54, 50], [51, 52]. As the result of this we design GARCH and CNN-LSTM, GARCH and LSTM, CNN and LSTM & GARCH and LSTM-CNN hybrid models.

The basic steps in the hybrid Models

1. Data and preprocessing: define the return series and compute realized measures and split data into training, validation, and test sets.

2. Secondly, develop GARCH-type model using statistical criteria and estimate the parameters that indicates the volatility features.

3. Use the residuals from mean equation (the conditional volatility of the log returns) to the input for deep learning models.

4. Based on this GARCH-CNN-LSTM, GARCH-LSTM, CNN-LSTM and GARCH-LSTM-CNN hybrid models are designed in the volatility prediction of stock index return.

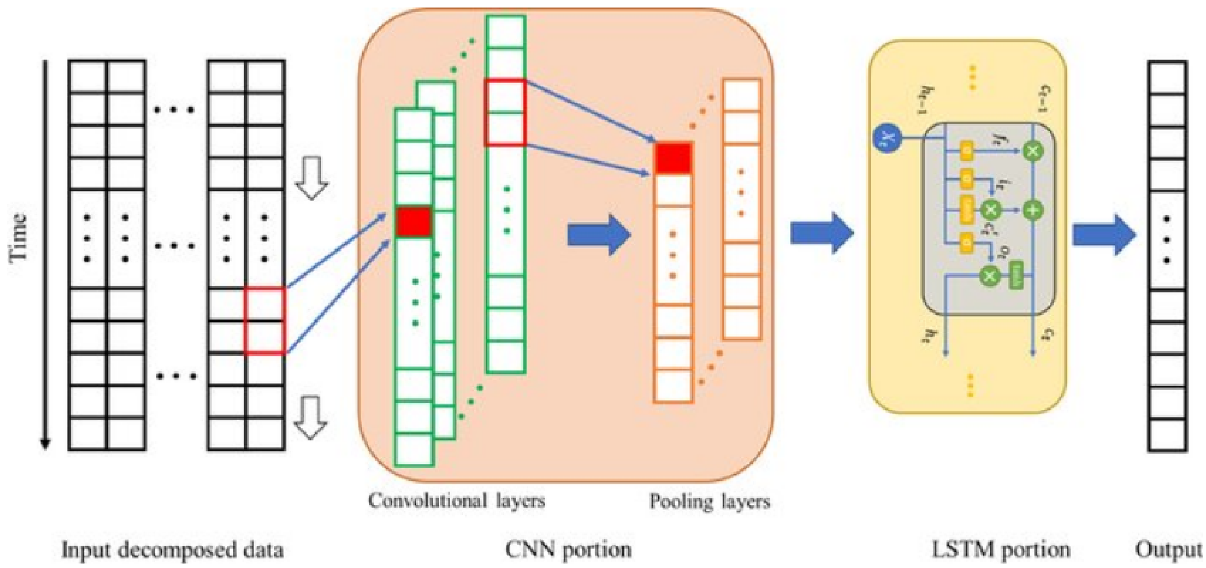

FIGURE 4.9: Architecture of sequential CNN-LSTM model

In case of GARCH-CNN-LSTM after selection of conditional volatility from GARCH type models using the criterias and tests it joins CNN keras tensor flow of kernel 3, pool size 2, number of filter 64 , ReLU activation function, batch size of 32 and padding the same to keep the number of input is the same as number of out puts and again this out put as an input for the LSTM with 50 cell units of batch size 32 in adam optimizer and 120-200 epoch. In CNN-LSTM the input starts directly the retun of the stock into CNN and then into LSTM taking the parameters and hyperparameters.

As shown in Figure 4.9 the conditional volatility result from GARCH (1,1) as a sequential data enters into convolutional layer and after convolutional operations passes to the flatten layer and this out put as an input of the gates in LSTM.

4.8 Result and Discussion

4.8.1 Analysis on selected agricultural commodities

TABLE 4.5: Return statistics for selected staple crops

Parameter	teff-return	wheat-return	maize-return	barley-return
Mean	0.003895	0.004056	0.004099	0.003915
Median	0.004158	0.004488	0.003752	0.004073
Maximum	0.128121	0.234647	0.173019	0.225738
Minimum	-0.198851	-0.243662	-0.373966	-0.205590
Std.Dev	0.026947	0.034557	0.038920	0.039305
Skewness	-0.461898	-0.840425	-1.438414	-0.185334
Kurtosis	10.72215	16.800029	23.77059	9.744343
Jarque-Bera	1449.119	4630.504	10534.32	1093.065
Probability	0.000000	0.0000000	0.000000	0.000000

From the descriptive statistics on Table 4.5 the mean log return prices of teff, wheat, barely and maizes are 0.0039, 0.0041, 0.0041 and 0.0039 respectively. The maximum values are 0.2346, 0.2257,0.1730 and 0.1281 for wheat, barely, maize and teff respectively that indicates high return deviations from the mean. Similarly the minimum values -0.3739, -0.2437, -0.2056 and -0.1989 for maize, wheat, barely and teff respectively which shows high weekly variability of log returns.

The skewness of the normal distribution is 0. However, Table 4.5 reveals that the skewness of maize is -1.4384 that is negatively skewed and all the remaining three are also negatively skewed. The kurtosis of maize, wheat, teff and barely price returns of the series are 23.7706, 16.8000,10.7222 and 9.7443 respectively that are greater than 3 which reveals leptokurtic. That has fat-tails and high peakedness distributions for each crops.

The value of Jarque-Bera for each staple crop in Table 4.5, are 1449.119, 4630.504, 10534.32 and 1093.065 teff-return, wheat-return, maize-return and barely-return series respectively significantly greater than the p-value 0. Therefore reject the null hypothesis (normality) and the distributions are not normally distributed.

TABLE 4.6: Teff-return GARCH extension model candidate

Model	Error distribution	AIC	SIC	Log likelihood
GARCH(1,1)*	ND	-4.6605	-4.6226	1342.56
GARCH(1,1)*	STD	-4.9015	-4.8560	1412.73
GARCH(1,1)*	GED	-4.9002	-4.8547	1412.35
TGARCH(1,1)*	ND	-4.6576	-4.6121	1342.74
TGARCG(1,1)	STD	-4.9009	-4.8478	1413.55
TGARCH(1,1)	GED	-4.8977	-4.8446	1412.64
TGARCH(2,1)*	ND	-4.6559	-4.6129	1340.99
TGARCH(2,1)	STD	-4.9081	-4.8474	1416.16
TGARCH(2,1)	GED	-4.9046	-4.8439	1415.61
EGARCH(1,1)*	ND	-4.6514	-4.6059	1340.96
EGARCH(1,1)	STD	-4.8932	-4.8401	1411.34
EGARCH(1,1)	GED	-4.8923	-4.8393	1411.10
PARCH(1,1)	ND	-4.6588	-4.6057	1344.08
PARCH(1,1)	STD	-4.9074	-4.8467	1412.14
PARCH(1,1)	GED	-4.9057	-4.8374	1416.93

* all parameters are significant(p-value <0.05)

From Table 4.6 we observed that GARCH(1,1), TGARCH(1,1), TGARCH(2,1), EGARCH(1,1) and PARCH(1,1) models under normal distribution, student's t-distribution and generalized error distribution assumptions of teff-return were selected as the candidate models using minimum Akaike Information Criteria(AIC), Schwarz Information Criterion(SIC) and maximum likelihood with significant p value of parameters(p-value less than 5%). Thus, based on the minimum information criteria, maximum loglikelihood and significant coefficients, GARCH(1,1) with generalzed error distribution assumption for teff-return model is identified as the best performing model among the selected candidate models. The distribution is GED that reveals it is leptokurtic.

The conditional variance equation of teff-return in GARCH(1,1) GED is:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{4.8.1}$$

$$\sigma_t^2 = 0.00020 + 0.54583\varepsilon_{t-1}^2 + 0.25463\sigma_{t-1}^2 \tag{4.8.2}$$

From Equation 4.8.2 above indicates that the volatility of returns is persistent, with the sum of α_1 and β_1 equal to 0.80046 (close to unity). Although the volatility of teff price return has long memory, it is still mean reverting. Moreover stationarity condition of $\alpha_1 + \beta_1 < 1$ is satisfied and it shows that the conditional variance process of teff log returns series is stable and predictable.

TABLE 4.7: Wheat-return GARCH extension candidate models

Model	Error distribution	AIC	SIC	Log likelihood
GARCH(1,1)*	ND	-4.0514	-4.0135	1167.76
GARCH(1,1)*	STD	-4.6662	-4.6112	1345.34
GARCH(1,1)*	GED	-4.6131	-4.5676	1327.96
TGARCH(1,1)*	ND	-3.9825	-3.9377	1148.98
TGARCG(1,1)	STD	-4.6648	-4.6117	1345.80
TGARCH(1,1)*	GED	-4.6101	-4.5570	1330.11
EGARCH(1,1)*	ND	-4.0592	-4.0137	1170.99
EGARCH(1,1)*	STD	-4.6733	-4.6202	1348.24
EGARCH(1,1)*	GED	-4.6139	-4.5608	1331.19
PARCH(1,1)*	ND	-4.0580	-4.0049	1171.65
PARCH(1,1)	STD	-4.6510	-4.5904	1342.860
PARCH(1,1)	GED	-4.6072	-4.5465	1330.28

* all parameters are significant(p-value <0.05)

Using minimum AIC and SIC and maximum loglikelihood having significant parametric values from Table 4.7, EGARCH (1,1) in student's t error distribution in single asymmetric order is identified as the best model for the conditional variance of wheat price return. The distribution is STD that shows it is leptokurtic (has fat tails).

The conditional variance of the wheat price return is presented as:

$$log\sigma_t^2 = \alpha_0 + \frac{\alpha_1 \mid \varepsilon_{t-1} \mid}{\sigma_{t-1}} + \frac{\gamma_1 \varepsilon_{t-1}}{\sigma_{t-1}} + \beta_1 log\sigma_{t-1}^2 \quad (4.8.3)$$

$$log\sigma_t^2 = 4.88688 + \frac{0.9095 \mid \varepsilon_{t-1} \mid}{\sigma_{t-1}} + \frac{-0.2151\varepsilon_{t-1}}{\sigma_{t-1}} + 0.4160 log\sigma_{t-1}^2 \quad (4.8.4)$$

The coefficient of the asymmetric term γ_1 is negative (-0.21516) and statistically significant at 1% level which reveals that negative return (bad news) has larger volatility than positive return (good news). The total effect of bad news on $log\sigma_t^2$ is $(1 + 0.2151) \mid \varepsilon_{t-1} \mid$

where as the total effect of good news on $log\sigma_t^2$ is (1- 0.2151)ε_{t-1}. The GARCH term β_1 = 0.4160 which shows weak persistence to die out as far from unity.

TABLE 4.8: Barely-return GARCH family model selection

Model	Error distribution	AIC	SIC	Log likelihood
GARCH(1,1)*	ND	-4.5404	–4.5029	1308.116
GARCH(1,1)*	STD	-4.6589	-4.6134	1343.129
GARCH(1,1)*	GED	-4.6576	-4.6121	1342.747
TGARCH(1,1)	ND	-4.5382	-4.4927	1308.483
TGARCH(1,1)	STD	-4.6556	-4.6025	1343.177
TGARCH(1,1)	GED	-4.6545	-4.6014	1342.841
EGARCH(1,1)	ND	-4.5446	-4.4991	1310.308
EGARCH(1,1)	STD	-4.6608	-4.6074	1344.584
EGARCH(1,1)	GED	-4.6625	-4.6094	1345.149
PARCH(1,1)	ND	-4.5385	-4.4854	1309.562
PARCH(1,1)	STD	-4.6612	-4.6006	1345.773
PARCH(1,1)	GED	-4.6569	-4.5933	1344.559

* All parameters are significant

Using in the same way, minimum AIC and SIC and maximum log-likelihood having significant parametric values from candidate models of Table 4.8 many asymmetric models are not significant which at least one of the parametrs are not significant (p-value greater than 5%). The conditional variance of barley is symmetrical to bad news and good news. Although GARCH (1,1) in STD shows minimum in AIC and BIC but it is not best predictive and not mean reverting as $\alpha + \beta > 1$. Thus GARCH (1,1) GED is identified as the best model for the conditional variance of barley price return.

The conditional variance of the barley price return is presented as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{4.8.5}$$

$$\sigma_t^2 = 0.000064 + 0.2855\varepsilon_{t-1}^2 + 0.6722\sigma_{t-1}^2 \tag{4.8.6}$$

From Equation 4.8.6 above indicates that the volatility of returns is persistent, with the sum of α_1 and β_1 being 0.9577 which shows that price return has long memory but it is still mean reverting. Moreover stationarity condition of $\alpha_1 + \beta_1 < 1$ is satisfied and it indicates that the conditional variance process of barley log returns series is stable and predictable.

TABLE 4.9: Maize-return GARCH extension models

Model	Error distribution	AIC	SIC	Log likelihood
GARCH(1,1)*	ND	-4.1469	-4.1090	1195.16
GARCH(1,1)*	STD	-4.3588	-4.3133	1256.97
GARCH(1,1*)	GED	-4.3520	-4.3065	1255.05
TGARCH(1,1)*	ND	-4.1485	-4.1030	1196.62
TGARCG(1,1)	STD	-4.3594	-4.3063	1258.16
TGARCH(1,1)	GED	-4.3531	-4.3000	1256.35
EGARCH(1,1)*	ND	-4.1498	-4.1043	1196.99
EGARCH(1,1)*	STD	-4.3745	-4.3214	1262.48
EGARCH(1,1)*	GED	-4.3635	-4.3104	1259.32
PARCH(1,1)	ND	-4.1546	-4.1015	1199.34
PARCH(1,1)*	STD	-4.3733	-4.3127	1262.16
PARCH(1,1)	GED	-4.3633	-4.3027	1260.29

* All parameters are significant (p-value<0.05)

From Table 4.9 having all significant parameters and the criteria of minimum AIC, SIC and maximum likelihood EGARCH (1,1) with student's t distribution is the best conditional variance of maize price return and the equation is:

$$log\sigma_t^2 = \alpha_0 + \alpha_1 \frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} + \frac{\gamma_1 \varepsilon_{t-1}}{\sigma_{t-1}} + \beta_1 log\sigma_{t-1}^2 \quad (4.8.7)$$

$$log\sigma_t^2 = -2.39994 + 0.87989 \frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} + \frac{-0.21545\varepsilon_{t-1}}{\sigma_{t-1}} + 0.71129 log\sigma_{t-1}^2 \quad (4.8.8)$$

From Equation 4.8.8 γ_1 is negatives that shows bad news has high volatility than good news. That is the current volatility depends not only on one lag of the magnitude of the return but also the sign of one lag log returns of maize. The result is synonyms with [23]. As GARCH term $\beta_1 = 0.71129$ its volatility persistence is moderate and it satisfies stationary condition.

The empirical result have emphasized that for these agricultural commodities, a negative return increases conditional variance more than a positive return of the same magnitude does which is the leverage effect. The pivotal point implies that negative (shocks to) commodity returns ought to be followed by an increase in conditional variance, or at least that negative returns ought to affect subsequent conditional variance more than positive returns do.

TABLE 4.10: Breusch-Godfrey test for standardized residuals of fitted models

Teff	F-statistics	1.5448	Prob.F(1,571)	0.2144
	Obs*R-squared	1.5460	Prob.Chi-square(1)	0.2137
Wheat	F-statistics	0.0498	Prob.F(1,571)	0.8235
	Obs*R-squared	0.0499	Prob.Chi-square(1)	0.8231
Barely	F-statistics	0.4210	Prob.F(1,571)	0.5167
	Obs*R-squared	0.4222	Prob.Chi-square(1)	0.5158
Maize	F-statistics	0.7915	Prob.F(1,571)	0.3740
	Obs*R-squared	0.7931	Prob.Chi-square(1)	0.3731

From Table 4.10, the standardized residuals of the fitted models did not exhibit any additional ARCH effects across all series, as both the F-statistics and the observed R-squared values were insignificant. Furthermore, the correlogram of the squared residuals shows that the p-values of both the autocorrelation function (ACF) and the partial autocorrelation function (PACF) exceed 5% (0.05) or lie within the corresponding confidence intervals. This indicates the absence of serial correlation in the residuals. Therefore, the selected models for price volatility are statistically well justified.

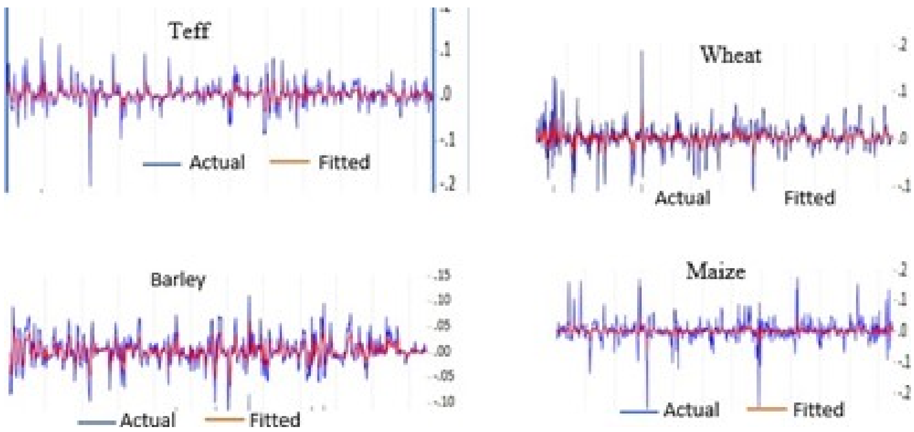

FIGURE 4.10: Actual-fitted log return for selected models

TABLE 4.11: Forecasting performance of selected models

Staple crops	Model	Error dis.	RMSE	MAE
Teff	GARCH(1,1)	GED	0.0101	0.0082
Wheat	EGARCH(1,1)	STD	0.0078	0.0066
Barely	GARCH(1,1)	GED	0.0078	0.0065
Maize	EGARCH(1,1)	STD	0.0180	0.0144

From Table 4.11 we see that the four weeks ahead forecasting performance of GARCH (1,1) GED for teff-return, EGARCH (1,1) STD for wheat-return, GARCH (1,1) GED for barely-return and EGARCH (1,1) STD for maize-return in mean absolute error are 0.0082, 0.0066, 0.0065 and 0.0144 respectively and it indicates that they better perform.

4.8.2 Analysis of S&P 500 stock indices

TABLE 4.12: descriptive statistics of S&P 500 log-return

Mean	Max.	Min	Std.Dev	Skewness	Kurtosis	Jarque-Bera	P-value
0.0005	0.08968	-0.12765	0.01343	-0.839736	17.501	11170.50	0.00000

From the descriptive statistics in Table 4.12, S&P 500 stock return has maximum and minimum values 0.089 and -0.1276 respectively which are highly deviated from the mean 0.0005. Moreover, the skewness and kurtosis of S&P 500 stock price return distiribution indicates negatively skewed and leptokurtic. The value of Jarque-Bera in Table 4.12 is 11170 significantly greater than the p-value 0. Therefore, reject the null hypothesis (normality) and the distribution is not normally distributed.

TABLE 4.13: S&P 500 stock price-return GARCH extension candidate models

Model	Error distribution	AIC	SIC	Log likelihood	p-vale α	p-value. β
GARCH(1,1)	ND	-4.3414	-4.3210	3990.59	0.0000	0.00000
GARCH(1,1)	STD	-4.3741	-4.3496	4012.23	0.0000	0.1020
GARCH(1,1)	GED	-4.3684	-4.3439	4008.56	0.0165	0.0250
TGARCH(1,1)	ND	-4.3548	-4.3302	4000.02	0.0012	0.0030
TGARCG(1,1)	STD	-4.3909	-4.3784	4017.55	0.0516	0.4270
TGARCH(1,1)	GED	-4.3850	-4.3564	4020.01	0.0064	0.2850
EGARCH(1,1)	ND	-4.3570	-4.3325	4001.39	0.0186	0.0251
EGARCH(1,1)	STD	-4.4027	-4.3741	4031.13	0.1951	0.6350
EGARCH(1,1)	GED	-4.3865	-4.3579	4020.97	0.0265	0.6673

From Table 4.13, we observed that GARCH (1,1), TGARCH (1,1) and EGARCH (1,1) models under normal distribution (ND), student's t-distribution (STD) and generalized error distribution (GED) assumptions of S&P 500-return were selected as the candidate models using minimum Akaike Information Criteria (AIC), minimum Schwarz Information Criterion (SIC) and maximum log likelihood with significant p value of parameters (p-value less than 5%). Thus, based on the minimum information criteria, maximum log likelihood and significant coefficients, GARCH (1,1) with GED assumption for S&P 500-return is identified as the best performing model among the selected candidate models. The distribution is GED that reveals it is leptokurtic.

The conditional variance equation of S&P 500-return in GARCH (1,1) GED is:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{4.8.9}$$

$$\sigma_t^2 = 0.0000036 + 0.1668\varepsilon_{t-1}^2 + 0.8145\sigma_{t-1}^2 \tag{4.8.10}$$

From Equation 4.8.10 above indicates that the volatility of returns is persistent, with $\alpha_1 + \beta_1$ equal to 0.97 (close to unity) which is mean reverting. Moreover, stationarity condition the magnitude of $\alpha_1 + \beta_1 < 1$ is satisfied and it shows that the conditional variance of S&P 500 return series is stable and predictable.

After the identification of GARCH -type, we estimated the conditional volatility of S&P 500 stock return and the forecasting performance is 0.0000286, 0.005352 and 0.004516 in mean squared error, root mean squared error and mean absolute error respectively in the out sample test data.

The fitness of the actual volatility which is the logarithmic return of S&P 500 with conditional volatility obtained from GARCH (1,1)-GED in the training data and the test data with forecasted volatility are depicted in Figure 4.11.

S&P 500 results in hybrid models

We imported the units of CNN and LSTM from tensorflow and keras library in python software. From estimated conditional volatility splits into 2019-06-10 to 2023-03-08(944)

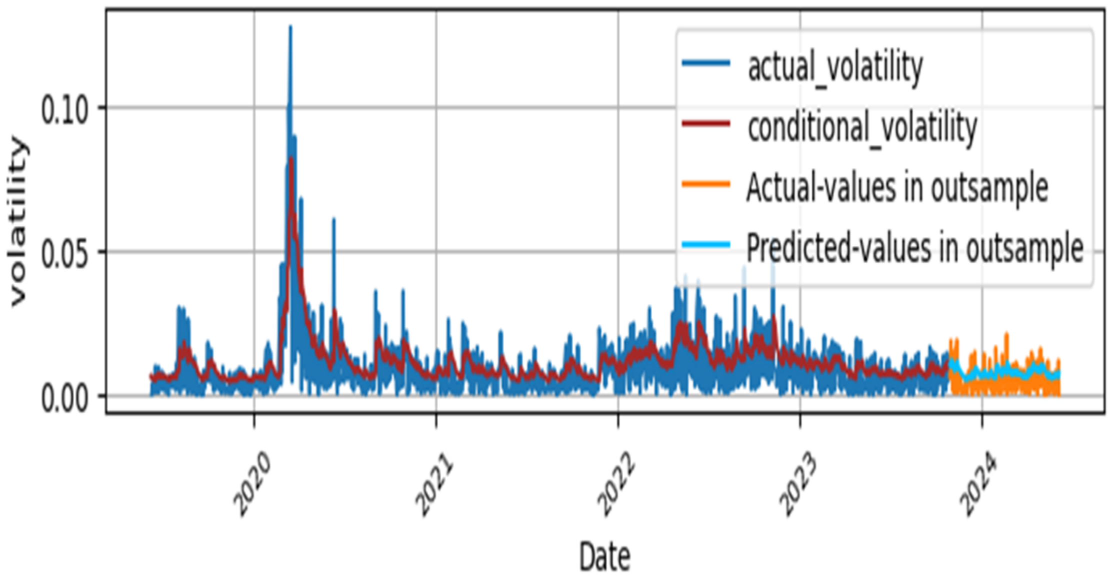

FIGURE 4.11: Actual vs conditional volatility by GARCH

for training, 2023-03-09 to 2023-10-30 (163) for validation and 2023-10-31 to 2024-06-06 (151) for test (outsamples).

Based on this partition, we trained and test different hybrid models. In the first part, we trained this output conditional volatility into the input CNN -LSTM hybrid model with 20 time steps (lookbacks),75% time splits, 11 number of folds and hyperparameters in Table 4.14. Hyperparameter tuning is based on a trial-and-error (random search method) in minimum metrics.

TABLE 4.14: CNN and LSTM hyperparametrs

Model	Activation function	Pool size	kernel	Number of filters/units	Optimizer	epoch	Batch size
CNN	ReLU	2	3	64	Adam(lr=0.0001-0.0005)	120-200	32
LSTM	sigmoid, tanh			50	Adam(lr=0.0001-0.0005)	120-200	32

The loss value in the first fold of GARCH-CNN-LSTM training process is depicted in Figure 4.12 and almost similar figures are observed in the process which validate the training.

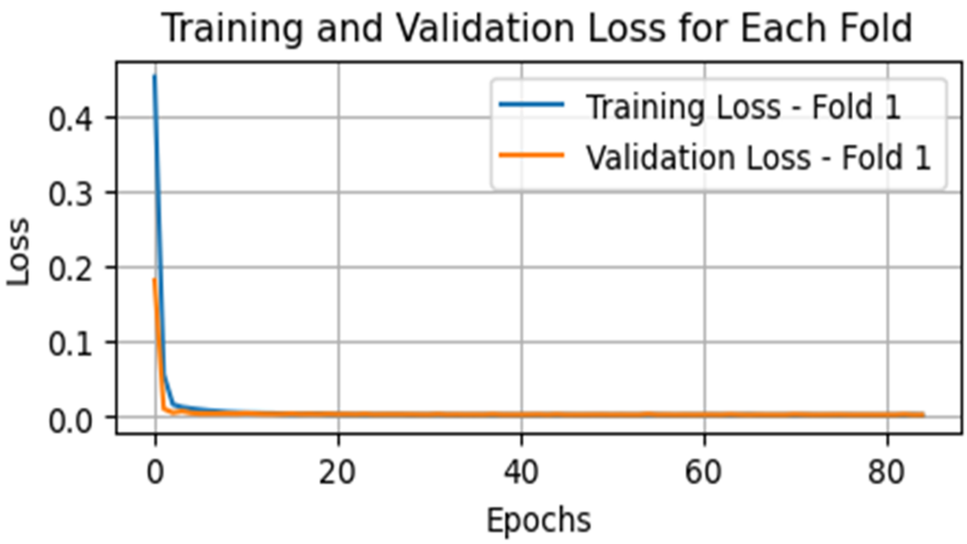

FIGURE 4.12: loss value in the process of training

Furthermore, the fittness of actual, predicted from GARCH and predicted volatility from GARCH-CNN-LSTM of out-sample data set and the comparison of the predicted and actual test data (realized volatility of the out-sample set) is depicted in Figure 4.13. The forecasting performance of GARCH-CNN-LSTM hybrid model in outsample data set by the loss functions MSE, RMSE and MAE is depicted in Table 4.15.

The experiment of GARCH-LSTM is processed by taking the output of GARCH (1,1) as the input for LSTM with the same training, validation and test set data size as GARCH-CNN-LSTM and the same kind and number of parameters and hyperparameters. The fittness of predicted versus test data is plotted at Figure 4.14.

Keeping the parameters and hyperparamers as it is and taking S&P 500 stock return as the input for CNN integration into LSTM with the same size of training, validation and test data is carried out for CNN-LSTM hybrid model. Comparatively the performance of this model is stated at Table 4.15. Usng the conditional volatility of GARCH (1,1) with the same size validation, test data and hyper parametres as GARCH-CNN-LSTM only

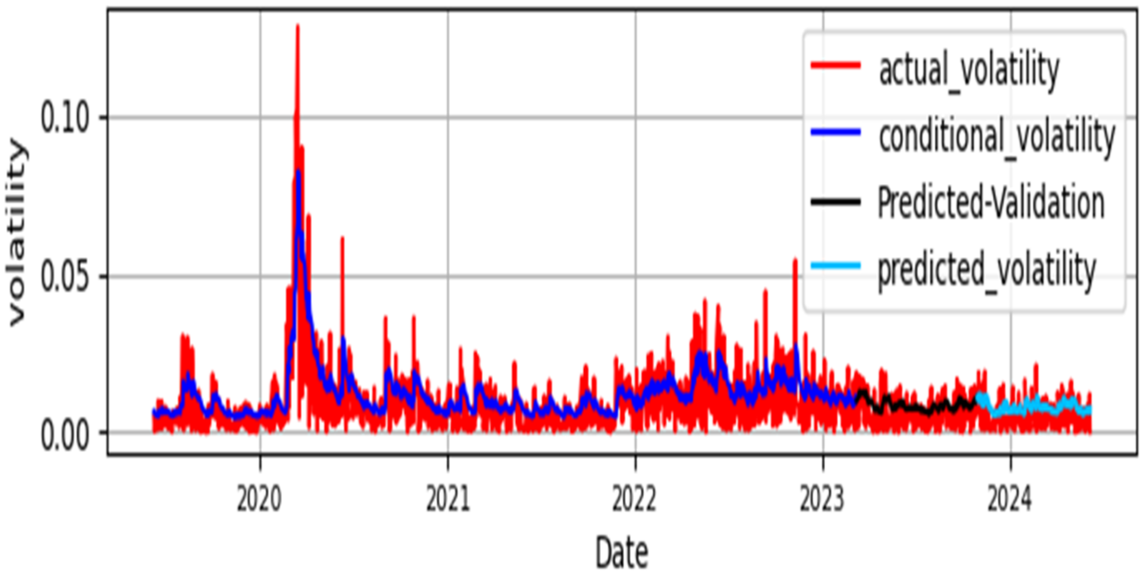

FIGURE 4.13: Actual vs predicted in-sample outsample plot for GARCH-CNN-LSTM

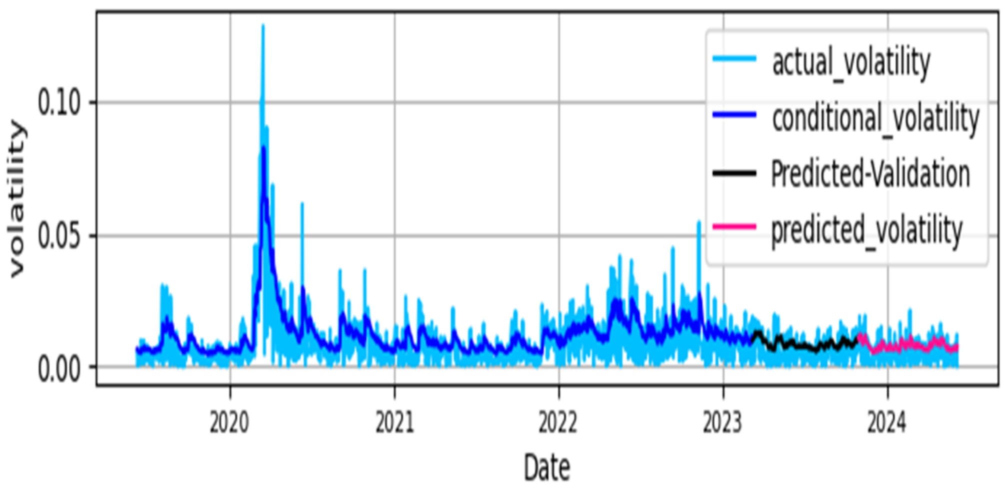

FIGURE 4.14: Actual vs predicted insample outsample plot for GARCH-LSTM

changng the the order CNN and LSTM deep learning models the hybrid GARCH-LSTM-CNN is experimented and volatility forecasting performance in mean square error, root mean square error and mean absolute error is depicted at Table 4.15.

The loss value during the training process as well as the fitness of predictions with the actual data in training set, validation with actual data in validation set and with predictions in test data are depicted in Figure 4.15 and Figure 4.16 for CNN-LSTM and GARCH-LSTM-CNN models respectively.

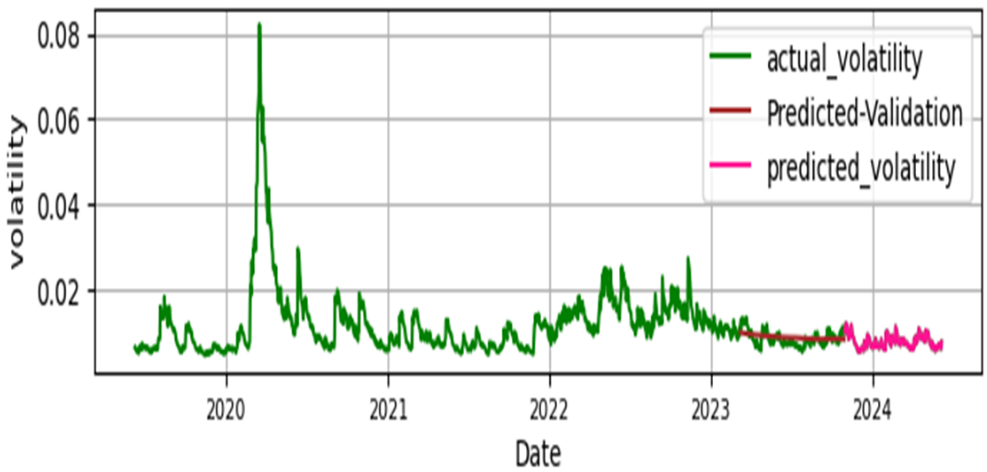

FIGURE 4.15: Actual vs predited insample and outsample plot for GARCH-LSTM-CNN

TABLE 4.15: Forecasting performance of the models in out-sample data

Metrics	GARCH	GARCH CNN-LSTM	GARCH LSTM	CNN-LSTM	GARCH LSTM-CNN
MSE	0.00002865	**0.00002524**	0.00002741	0.00002813	0.000027631
RMSE	0.005352	**0.005024**	0.005235	0.005304	0.005256
MAE	0.004516	**0.004218**	0.004515	0.004685	0.004553

From Table 4.15, the minimum values in mean square error, root mean square error and mean absolute errors we observe that GARCH-CNN-LSTM has the lowest values in the three metrics and GARCH-LSTM has the second lowest values, where as GARCH has the highest values which indicates that GARCH-CNN-LSTM is the most accurate model in forecasting the volatility of S&P 500 stock price and GARCH-LSTM is the second one. GARCH performs worst among the compared models in the volatility predicting performance of the S&P 500 stock price. The S&P 500 stock price volatility prediction power of GARCH-LSTM-CNN less than GARCH-CNN-LSTM the sequential input and output matters.

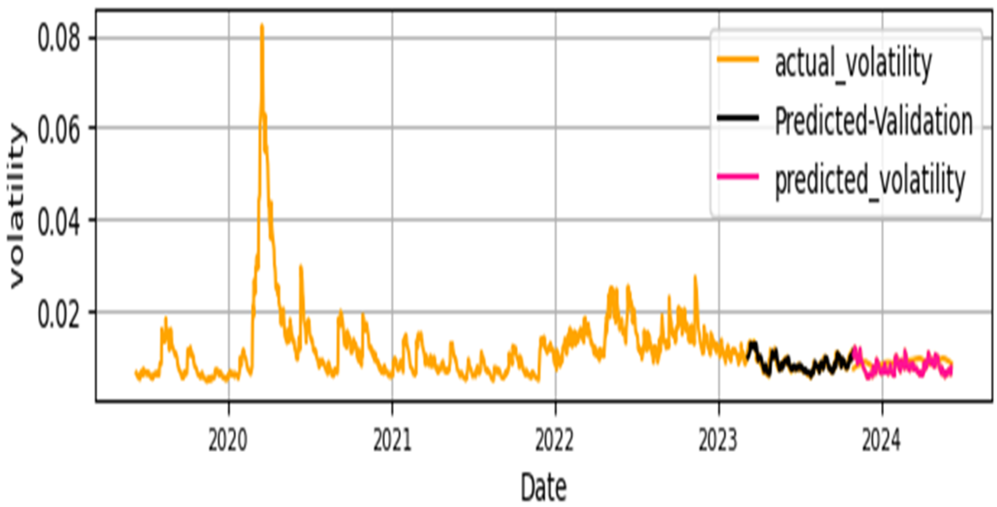

FIGURE 4.16: CNN-LSTM actual and predicted plot

Table 4.16 shows the lower the statistical value and upper triangular the p-value of Diebold Mariano test results for the same number of outsample days. The models have comparable accuracy as all the p-values are significant at 5% significant lablel. This indicates that each model has different volatility prediction performance in S&P 500 stock market dynamics.

TABLE 4.16: Diebold Mariano (DM) test results

	GARCH	GARCH CNN-LSTM	GARCH LSTM	CNN-LSTM	GARCH LSTM-CNN
GARCH		0	0	0	0.03
GARCH-CNN-LSTM	3.65		0	0	0
GARCH-LSTM	3.62	2.10		0	0
CNN-LSTM	3.09	2.41	3.61		0
GARCH-LSTM-CNN	1.8	3.56	3.64	3.61	

Although deep learning algorithms have an ability of better forecasting in time series, they have drawbacks in parameter setting as tunning is based on trial and error depending on the nature of data [110].

4.9 Conclusions

The weekly log-returns of mixed teff, wheat, barley and maize in the study period and area using Akaike Information Criterion (AIC), Schwarz Information Criteria (SIC), Log-likelihood and significance of the parameters with alternative error distributions, we conclude that GARCH family models, specifically GARCH(1,1) with Generalized Error Distribution (GED) for teff and barley, and EGARCH (1,1) with Student's t Distribution (STD) for wheat and maize, effectively model and forecast of the price volatility of these agricultural commodities. The study found evidence of asymmetry in volatility, where negative news has a stronger impact on price variance than positive news for wheat and maize. These models provide valuable insights for producers, consumers, and traders to anticipate market price fluctuations and make informed decisions to reduce price volatility in these staple crops.

The volatility of S&P 500 stock returns has been analyzed using symmetric GARCH, asymmetric TGARCH, and EGARCH models, evaluated under normal, Student's t, and generalized error distributions. Based on criteria such as the minimum Akaike Information Criterion (AIC) and maximum log-likelihood with significant statistical coefficients, the GARCH (1,1) model with a generalized error distribution outperforms other GARCH extensions in predicting S&P 500 stock price volatility. Additionally, we computed the conditional volatility of S&P 500 stock prices during the study period, providing input for hybrid CNN and LSTM neural network models.

The integration of GARCH with LSTM in forecasting S&P 500 stock price volatility performs effectively, evidenced by low MSE, RMSE, and MAE values in out-of-sample data. GARCH captures spatial volatility features such as volatility clustering, while LSTM captures long-term memory persistence. Similarly, the hybrid model GARCH-CNN-LSTM slightly outperforms in volatility forecasting, leveraging CNN's ability to recognize spatial patterns in the time series data, as indicated by the same metrics and out-of-sample data. The combination of CNN and LSTM shows the lowest performance in terms of high MSE, RMSE, and MAE values on the same out-of-sample data. When forecasting the volatility of S&P 500 stock prices, the GARCH-LSTM-CNN hybrid model performs better than CNN-LSTM but not as well as either GARCH-LSTM or GARCH-CNN-LSTM hybrid models.

In this thesis, we focus only on the closing price of the stock. However, incorporating additional factors such as volume, trading activity, and optimizing parameters and hyperparameters selection methods will enhance the accuracy of predicting S&P 500 stock price volatility.

In summary GARCH-CNN-LSTM is best out performs in the forecasting volatility of

S&P 500 stock price upto 13% accuracy improvement in the study period and GARCH-LSTM is the second most out performs well. Improved predictive abilities of forecasting the volatility of S&p 500 stock price is pivotal for the decision of finance and risk management in stock market trading.

Chapter 5

Volatility and Correlation of Stock Markets and Some Agricultural Commodities based on DCC-EGARCH and LSTM

5.1 Introduction

Accurate modeling and forecasting of volatility and spillover effects among financial assets play a pivotal role in understanding and managing market risks, particularly within the increasingly interconnected landscape of international stock markets and major agricultural commodity markets. Price fluctuations in global stock indices and staple agricultural commodities not only affect investors and policymakers but also have profound implications for economic stability and consumer food security.

Multivariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH)-type models have long served as a cornerstone for capturing time-varying volatility clustering and dynamic correlations among asset returns. However, symmetric multivariate GARCH specifications often encounter limitations related to high dimensional, nonlinear dependence structures, and limited flexibility in adapting to structural market changes.

Recent advances in deep learning, particularly recurrent architectures such as Long Short-Term Memory (LSTM) networks, have demonstrated strong capabilities in capturing complex nonlinear temporal patterns and long-range dependencies in financial time series. Consequently, hybrid modeling frameworks that integrate multivariate EGARCH-type structures with LSTM networks have emerged as promising approaches for improving volatility and correlation forecasting performance..

Multivariate econometric correlation models like DCC has drawbacks or limitation in forecasting the time varying relation ship of assets such as conditional correlations sometimes appearing limited flexibility for tail dependence and nonlinearities that is almost constant or indistinguishable and suffers from the curse of dimensionality, and parameter estimates tend to have severe negative biases, causing the estimated correlations to become unnaturally smooth or nearly static [137]. Moreover among these, the number of variables increases, estimating parameters becomes more complex and less reliable in high frequency financial data. Moreover they rely on parametric assumptions like stationarity and linearity which may not be practical and hence struggle to capture non linear dependency. Other multivariate models like BEKK requires large number of parameters in the matrices that leads overfitting and instability [80]. In addition, these multivariate covariance models requires a significant computational material particularly for large number of variables. To reduce such limitations deep learning models like Long Short-Term memory has been recently applied in finance and economics in forecasting stock market correlations and other asset prices. The DCC-LSTM hybrid model is a novel methodology which contains DCC-EGARCH and LSTM designed to forecast the conditional correlation between two assets and their volatilities by combining the strengths of both EGARCH (exponential generalized autoregressive conditional heteroskedasticity) models in asymmetric shocks for volatility estimation and LSTM networks for capturing complex temporal patterns of DCC.

5.2 Data description and Methods

5.2.1 Source and Description of Data

We collected historical closing price time series data for the S&P 500, Shanghai Composite Index (SHCOMP), NIFTY 50, and Toronto Composite Index (TSX) from 4 January, 2016, to 25 October, 2024, sourced from Yahoo Finance. A total of 2,298 trading days were analyzed. Missing values were imputed using two average values. We took also weekly retail price data of teff, wheat and maize from 1 September 2010 to 23 June 2022 from centeral statistical agency of Ethiopia. To simplify statistical analysis and ensure stationarity, the data was transformed into log returns. Additionally, the data was divided into two portions: the training set, which comprised 80% of the data, and the test set, which made up of the remaining 20 % to evaluate the model's performance. The time series plots of S&P 500, SHCOMP, NIFTY 50 and TSX of stock prices and teff, wheat and maize prices are non-stationary; however, their logarithmic returns exhibit stationarity, as illustrated in Figure 5.1 and Figure 5.2. Additionally, the ADF test results in Table 5.2 clearly shows that the t-statistics for the returns of S&P 500, SHCOMP, NIFTY 50, and TSX stocks as well as teff, wheat and maize are all greater than the critical value at the 5% significance level, with significant p-values. This confirms that the returns are stationary, indicating that they are mean-reverting.

From Table 5.1, it indicates the ARCH -LM statistics of the returns. The p-values show rejecting the null hypothesis of "no ARCH effect" at a 5% significant level, which suggests each of the log returns presents heteroskedasticity (time-varying volatility). Therefore, GARCH family models are better fitting for forecasting the volatility.

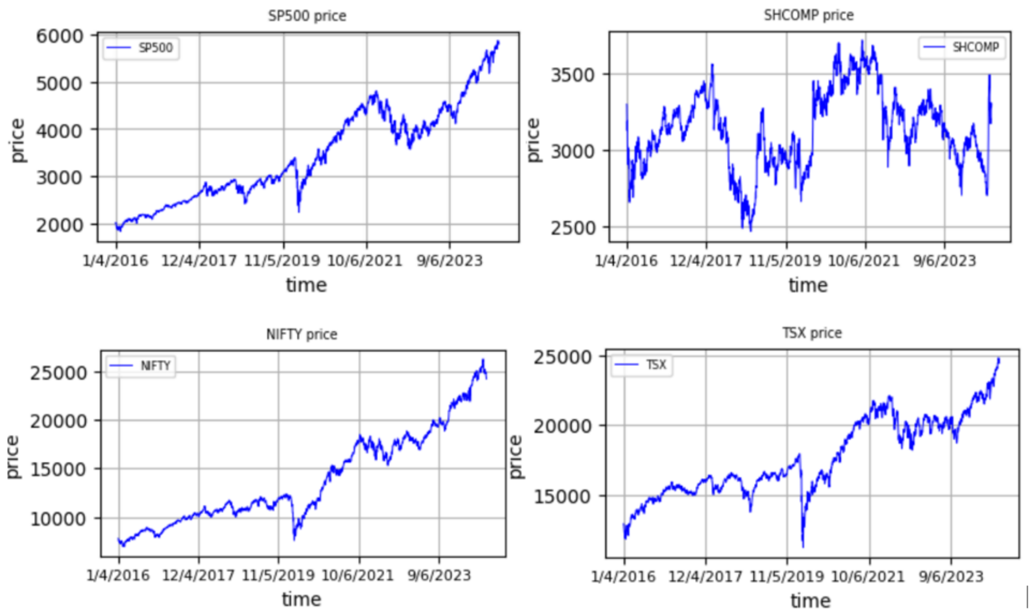

FIGURE 5.1: stock indices

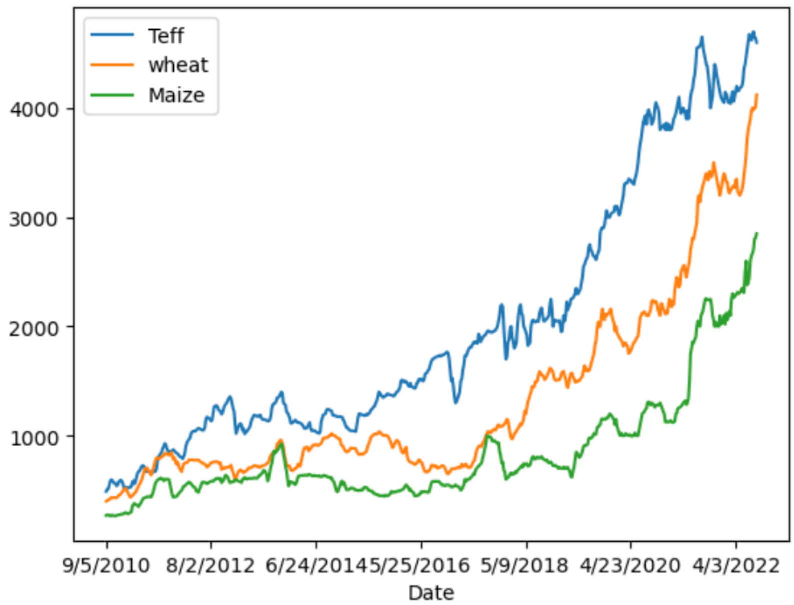

FIGURE 5.2: retail price of staple crops

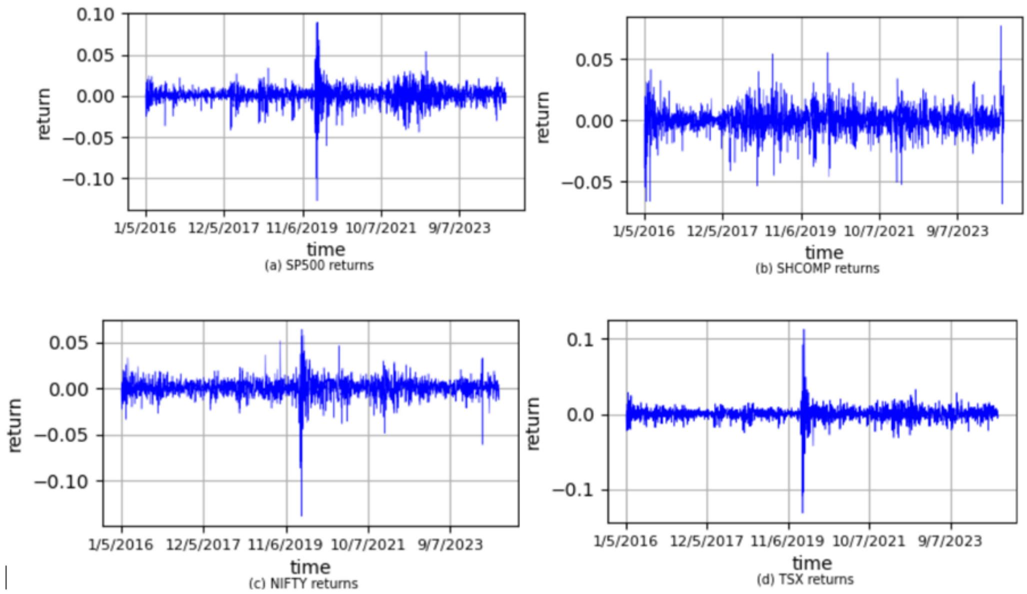

FIGURE 5.3: Daily stock returns

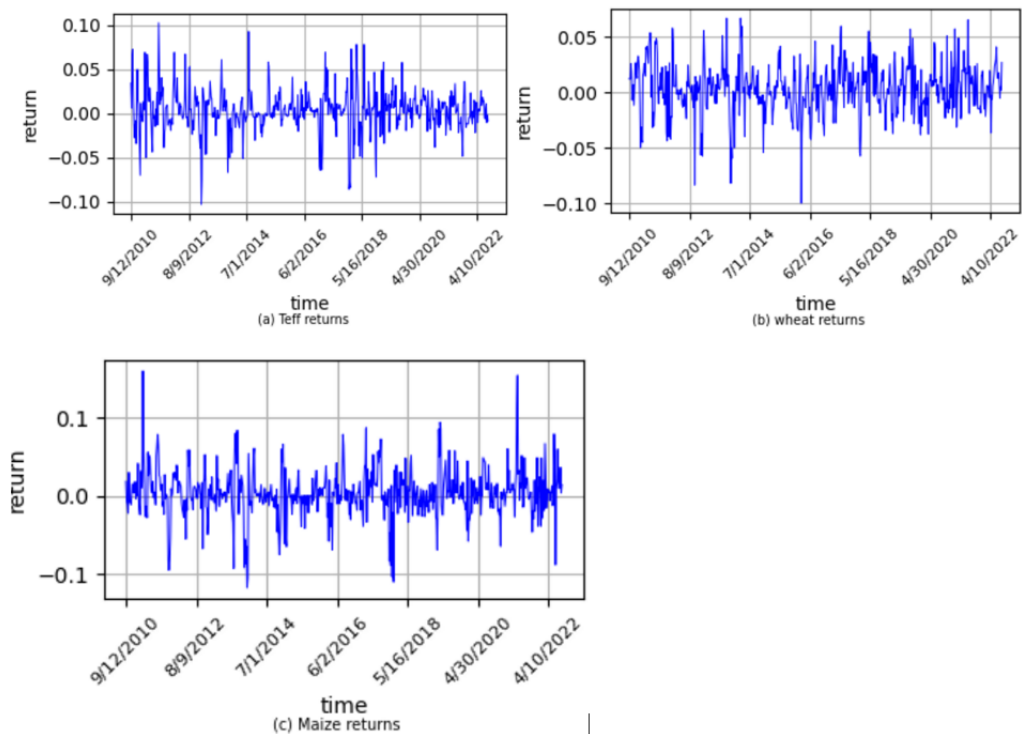

FIGURE 5.4: log return of staple crops

5.2.2 Multivariate GARCH family Models

Multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) is an extension of univariate GARCH that has a vector of n returns at time t and described as:

$$\begin{aligned} r_t &= \mu_t + \varepsilon_t \\ &= \mu_t + H_t^{\frac{1}{2}} z_t, \quad z_t \sim N(0, I) \end{aligned} \tag{5.2.1}$$

TABLE 5.1: ARCH LM test of Stock returns an staple crops

Returns	F-statistic	P-value
S&P 500	116.03	0.0000
SHCOMP	20.36	0.0000
NIFTY 50	63.73	0.0000
TSX	96.04	0.0000
Teff	26.01	0.0000
Wheat	38.67	0.0000
Maize	6.90	0.0088

TABLE 5.2: Unit root tests for Stock indice and Some Commodity returns

		ADF test critical values			
Variables	t-statistics	1%	5 %	10 %	p-value
S&P500	-61.5885	-3.4330	-2.8626	-2.5674	0.0001
SHCOMP	-31.5642	-3.4330	-2.8626	-2.5674	0.0001
NIFTY 50	-20.5903	-3.4330	-2.8626	-2.5674	0.0001
TSX	-14.6294	-3.4330	-2.8626	-2.5674	0.0001
Teff	-16.2573	-3.4415	-2.8664	-2.5694	0.0000
Wheat	-19.3078	-3.4415	-2.8664	-2.5694	0.0000
Maize	-12.9560	-3.4415	-2.8664	-2.5694	0.0000

where μ_t is the conditional mean vector. From Equation 5.2.1 when the mean is zero, the equation is reduced to:

$$r_t = \varepsilon_t \tag{5.2.2}$$

In this case, covariance of two asset expressed as:

$$\begin{aligned} H_t =& \alpha_0 + \sum_{i=1}^{q} \alpha_i r_{t-i} r'_{t-i} + \sum_{i=1}^{p} \beta_i H_{t-i} \\ =& \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i} \varepsilon'_{t-i} + \sum_{i=1}^{p} \beta_i H_{t-i} \end{aligned} \tag{5.2.3}$$

The general multivariate GARCH (p, q) conditional covariance matrix H_t is expressed as:

$$H_t = C + \sum_{i=1}^{q} A_i \varepsilon_{t-i} \varepsilon'_{t-i} A'_i + \sum_{j=1}^{p} B_j H_{t-j} B'_j \tag{5.2.4}$$

where:

H_t is the $n \times n$ conditional covariance matrix at time t, ε_t is the $n \times 1$ vector of residuals (innovations), C is a symmetric positive definite $n \times n$ matrix (constant component), A_i and B_j are $n \times n$ coefficient matrices that captures lag of residuals and covariance respectively, p and q are the GARCH and ARCH orders, respectively. The other parametrization form of multivariate model is vector error correction (VEC) using the lower part of previous lag residual and covariance matrices that is described as:

$$\text{vech}(H_t) = c + \sum_{i=1}^{q} A_i \,\text{vech}(\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{j=1}^{p} B_j \,\text{vech}(H_{t-j}), \tag{5.2.5}$$

where H_t is the conditional covariance matrix of the residuals at time t, ε_t is the vector of residuals, vech$(\cdot)$ is the operator that stacks the lower triangular part of a matrix into a vector, c is a vector of $\frac{n(n+1)}{2}$, A_i and B_j are matrices of $\frac{n(n+1)}{2} x \frac{n(n+1)}{2}$, n number of variables. when the orders p, q = 1, this equation is reduced to:

$$\text{vech}(H_t) = c + A \,\text{vech}(\varepsilon_{t-1} \varepsilon'_{t-1}) + B \,\text{vech}(H_{t-1}) \tag{5.2.6}$$

In this model H_t is positive definite for each t and the number of parameters is $\frac{n(n+1)}{2} + (p+q)(\frac{n(n+1)}{2})^2$ which is difficult to estimate due to large parametrization.

To confirm the positive definiteness , Baba, Engle, Kraft and Kroner (1990) (**BEKK**) proposed the model:

$$H_t = CC' + \sum_{i=1}^{q} \sum_{k=1}^{N} A_{ik} \varepsilon_{t-i,k} \varepsilon'_{t-i,k} A'_{ik} + \sum_{j=1}^{p} \sum_{k=1}^{N} B_{jk} H_{t-j} B'_{jk}, \tag{5.2.7}$$

where A_{ik}, B_{jk} and nxn matrices, C is a lower triangular matrix that guarantee the positive definiteness of the covariance H_t. The order one for this model is expressed as:

$$H_t = CC' + A\varepsilon_{t-1}\varepsilon'_{t-1}A' + BH_{t-1}B' \tag{5.2.8}$$

with each term is positive definite. The number of parameters is still large computationally complex.

A modification was made by Engles (2002) [77] called the dynamic conditional correlation (DCC) of time varying covariance matrix H_t decomposed into conditional standard deviation D_t and a dynamic correlation matrix R_t and mathematically expressed as:

$$H_t = D_t R_t D_t, \tag{5.2.9}$$

where D_t is the diagonal matrix of size n and given by:

$$D_t = diag(\sqrt{h_t}) \tag{5.2.10}$$

Or in matrix form $$D_t = \begin{pmatrix} \sqrt{h_{1,t}} & 0 & 0 & ... & 0 \\ 0 & \sqrt{h_{2,t}} & 0 & ... & 0 \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ 0 & 0 & ... & 0 & \sqrt{h_{n,t}} \end{pmatrix}.$$

Each element $h_{i,t}$ is the conditional volatility of each asset i and can be expressed in univariate EGARCH(1,1) as:

$$\log h_{i,t} = \omega + \alpha \frac{|\varepsilon_{t-1}|}{\sqrt{h_{i,t-1}}} + \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{i,t-1}}} + \beta \log h_{i,t-1}. \tag{5.2.11}$$

The correlation matrix has a symmetric form as: $R_t = \begin{pmatrix} 1 & \rho_{12,t} & \rho_{13,t} & \cdots & \rho_{1n,t} \\ \rho_{12,t} & 1 & \rho_{23,t} & \cdots & \rho_{2n,t} \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ \rho_{1n,t} & \rho_{2n,t} & \rho_{3n,t} & \cdots & 1 \end{pmatrix}$

.

This symmetric matrix is a positive definite as the covariance matrix H_t positive definite and the magnitude of each entry is less than or equal to one.

For constant conditional correlation (CCC) Equation 5.2.9 is expressed as

$$H_t = D_t R D_t \tag{5.2.12}$$

The multivariate GARCH model introduced by Engle (2002) expressed in Equation 5.2.12 shows the dynamic relationship of two time varying stock market process. The dynamic conditional correlation model R_t with GARCH(1,1) covariance of standard residuals is described into

$$R_t = diag(Q_t^{-1/2}) Q_t diag(Q_t^{-1/2}) \quad , \tag{5.2.13}$$

where

$$Q_t = \bar{Q}(1-a-b) + a z_{t-1} z'_{t-1} + b Q_{t-1}, \tag{5.2.14}$$

where $\bar{Q}$ is unconditional covariance , Q_t is the conditional covariance at time t, a and b are the joint shock and covariance parameters respectively and z & z' are standard residuals. The condition $a > 0$, $b > 0$ and $a+b < 1$ which is the mean reversion process. When $a+b = 1$, Equation 5.2.14 is reduced into:

$$\begin{aligned} Q_t &= a(z_{t-1} z'_{t-1}) + b Q_{t-1} \\ &= (1-\lambda)(z_{t-1} z'_{t-1}) + \lambda Q_{t-1}, \end{aligned} \tag{5.2.15}$$

where $\lambda = b$.

The process follows exponential smoother matrix of the standard residuals and unconditional covariance matrix $\bar{Q}$ described as.

$$\bar{Q} = cov(z_t z_t') \tag{5.2.16}$$

particularly, In DCC-GARCH model the conditional correlation is expressed as:

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{h_{i,t}}\sqrt{h_{j,t}}}, \tag{5.2.17}$$

where $\rho_{ij,t}$ is the conditional correlation between two stocks , $q_{ij,t}$ is the covariance of the two stocks and $\sqrt{h_{i,t}}$ & $\sqrt{h_{j,t}}$ are the standard deviations of each single stocks in the GARCH(1,1) process.

5.2.3 Multivariate Hybribrid Models

Steps in multivariate hybrid Models.

1. Fit the asymmetric univariate GARCH models to the individual asset return variances and then use the DCC model to estimate time-varying dynamic correlations between assets based on the standardized residuals from the GARCH models.

2. Extract the conditional variance estimates from EGARCH and the conditional correlation estimates from DCC as time series features. The long run correlations, the residuals from the dynamic conditional correlation (DCC) and estmated rolling window of time varying correlation serves as the input to LSTM. This hybrid model integrates these two powerful techniques to effectively handle the prediction of dynamic conditional correlation of two assets, and the asymmetric GARCH (EGARCH) captures the volatility clustering and leverage effect in financial time series data.

.

3. Feed these EGARCH/DCC time-varying variance and correlation sequences into LSTM network as input features to model complex nonlinear dependencies and temporal dynamics that traditional GARCH and DCC might not capture fully.

4. Train the LSTM network and Compare the forecast with econometric correlation models using the performance error metrics.

5.3 Results and discussion

As previously stated, this analysis is based on the daily closing price returns of four stock market indices and the weekly retail price returns of three selected staple crops, all of which are stationary and heteroscedastic.

TABLE 5.3: Descriptive statistics of asset returns

Stock	Mean	Max.	Min	Std.Dev	Skewness	Kurtosis	Jarque-Bera	P-value
S&P 500	0.0005	0.2593	-0.2634	0.0136	-0.682	129.3	1287	0.0001
SHCOMP	0.0000	0.0775	-0.0730	0.0104	-0.4571	9.9561	4712	0.0000
NIFTY 50	0.0005	0.0641	-0.1390	0.0101	-1.6710	26.723	54934	0.0000
TSX	0.0003	0.1129	-0.1318	0.0094	-1.7843	49.2872	20629	0.0000
Teff	0.0038	0.1281	-0.1988	0.02694	-0.4618	10.72	1449.1	0.0000
Wheat	0.0040	0.2346	-0.2436	0.0345	-0.8404	16.80	4630.5	0.0000
Maize	0.0040	0.1730	-0.3739	0.0389	-1.438	23.77	10534	0.0000

From the descriptive statistics in Table 5.3, S&P 500 stock return has maximum and minimum values of 0.2593 and -0.2634, respectively, which are highly deviated from the mean of 0.0005. Moreover, the skewness and kurtosis of the S&P 500 stock return distribution indicate negatively skewed and leptokurtic. Similarly, SHCOMP, NIFTY 50, and TSX returns have maximum values of 0.0775, 0.0641, and 0.1129, and minimum values of -0.0730, -0.1390, and -0.1318, respectively, that depart from their mean of 0.

For agricultural commodities maximum values are 0.1281, 0.2346 and 0.1739 and minimum values are -0.1988, -0.2436, and -0.3739 for teff, wheat and maize respectively. The means of these teff, wheat and maize are 0.0038, 0.0040 and 0.0040 repectively which is almost the same. Additionally, all are negatively skewed. Furthermore, Figure 5.3 daily return series and Figure 5.4 weekly returns revealed that high volatility was followed by a high volatility period and small volatility stylized to small volatility.

Among the univriate candidate models, GARCH(1,1), TGARCH(1,1), and EGARCH(1,1), the best-performing model is selected based on the Minimum Akaike Information Criterion (AIC), Minimum Schwarz Information Criterion (SIC), and maximum log-likelihood, with parameter p-values less than 5% and from alternative univariate distributions. Based on these criteria, the EGARCH (1,1) model is chosen as the best model for forecasting the volatility of stocks. Where as, GARCH(1,1), EGARCH(1,1) and EGARCH(1,1) are selected for teff, wheat and maize respectively.

From Table 5.4, the coefficient of lag volatility β is significant (p-value is zero) for all stocks and 0.97, which is near to 1, that confirms strong persistence and the lag volatility has a high impact on the current volatility. The leverage effect γ is significant for each stock indices that indicates the volatility is asymmetric. But, its magnitude is very small that shows the existence of little leverage effect. The magnitude of β is less than 1 which satiesfies the stationary condition and strongly persistent as it is closer to 1.

Granger causality is a statistical concept used to test whether one time series can help in the prediction of another. Volatility causality or spillover effect is a test that indicates the wether the past conditional variance or volatility of one asset causes the variance or volatility of the other asset and the vice versa.

Table 5.7 explores bi-directional Granger causality (spillover effects) between pairs of

TABLE 5.4: parameter estimates of DCC-EGARCH stock markets

Stocks	parameters	values	SE	p-value
	ω	-0.2927	0.0094	0.0000
S&P 500	α	0.1880	0.0233	0.0000
	γ	-0.1648	0.0168	0.0000
	β	0.9700	0.0009	0.0000
	ω	-0.2507	0.0066	0.0000
SHCOMP	α	0.1848	0.0250	0.0000
	γ	-0.0436	0.0164	0.0078
	β	0.9730	0.0007	0.0000
	ω	-0.2667	0.0108	0.0000
NIFTY 50	α	0.1246	0.0183	0.0000
	γ	-0.1036	0.0145	0.0000
	β	0.9725	0.0011	0.0000
	ω	-0.2507	0.0174	0.0000
TSX	α	0.1637	0.0436	0.0000
	γ	-0.1343	0.0157	0.0000
	β	0.9756	0.0019	0.0000
S&P 500 - SHCOMP	DCC a	0.0038	0.0040	0.3440
	DCC b	0.9841	0.0098	0.0000
S&P 500 - NIFTY 50	DCC a	0.0092	0.0134	0.4831
	DCC b	0.7720	0.4216	0.0467
S&P 500 - TSX	DCC a	0.0387	0.0171	0.0239
	DCC b	0.8006	0.0429	0.0000
SHCOMP - NIFTY 50	DCC a	0.0178	0.0122	0.1451
	DCC b	0.9084	0.0557	0.0000
SHCOMP - TSX	DCC a	0.0000	0.0090	0.9914
	DCC b	0.9155	0.3165	0.0038
NIFTY 50 - TSX	DCC a	0.0117	0.0111	0.2903
	DCC b	0.9158	0.0628	0.0000

TABLE 5.5: Parameter estimates of DCC-GARCH type Teff,Wheat and Maize

Cereals	parameters	values	SE	p-value
Teff	ω	0.0002	0.00005	0.0000
	α	0.5474	0.1672	0.0011
	β	0.2524	0.1205	0.0363
Wheat	ω	-0.4886	0.0225	0.0000
	α	0.9083	0.0329	0.0000
	γ	-0.2151	0.0197	0.8281
	β	0.4160	0.0023	0.0000
Maize	ω	-0.3344	0.0219	0.0000
	α	0.1778	0.0514	0.0005
	γ	-0.0951	0.0225	0.0000
	β	0.9642	0.0025	0.0000
Teff - Wheat	DCC a	0.0050	0.0036	0.1648
	DCC b	0.9852	0.0087	0.0000
Teff - Maize	DCC a	0.0275	0.0198	0.1648
	DCC b	0.6986	0.0941	0.0000
Wheat - Maize	DCC a	0.0244	0.0090	0.0065
	DCC b	0.8946	0.0349	0.0000

TABLE 5.6: Actual vs Conditional volatility of Stocks in out sample data

Metrics	S&P 500	SHCOMP	NIFTY 50	TSX
MSE	0.00002865	0.00005609	0.00003294	0.00002307
RMSE	0.005352	0.007489	0.005740	0.004804
MAE	0.004516	0.005363	0.004267	0.004217

stocks. In the case of NIFTY 50 and S&P 500, there is no risk transmision from NIFTY 50 to S&P 500 as the as the F-statistic is not significant. But, the volatility (variance) of S&P 500 has an effect on NIFTY 50 as F-statistics is significant. In case of TSX and NIFTY 50 both have directional feed back effects that the variance of one indices has

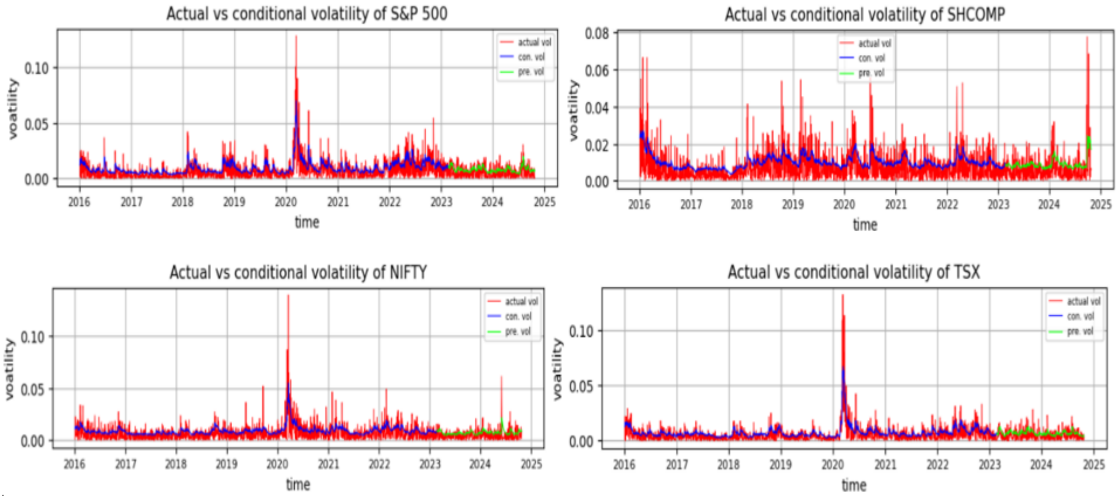

FIGURE 5.5: Conditional vs actual volatility of stocks

TABLE 5.7: Causality test of Stock returns

Returns	F-statistic	P-value
NIFTY 50 does not Granger cause S&P 500	0.635	0.4256
S&P 500 does not Granger cause NIFTY 50	116.3	0.0000
SHCOMP does not Granger cause S&P 500	1.766	0.1830
S&P 500 does not Granger cause SHCOMP	11.68	0.0006
TSX does not Granger cause S&P 500	2.60	0.1066
S&P 500 does not Granger cause TSX	12.35	0.0004
SHCOMP does not Granger cause NIFTY 50	2.66	0.1013
NIFTY 50 does not Granger cause SHCOMP	2.09	0.1420
TSX does not Granger cause NIFTY 50	111.6	0.0000
NIFTY 50 does not Granger cause TSX	11.56	0.0007
TSX does not Granger cause SHCOMP	3.04	0.0813
SHCOMP does not Granger cause TSX	1.185	0.2763

an effect on the variation of the other indice. The F-statistics of Shanghai composite index (SHCOMP) is not significant for S&P 50, NIFTY 50 and TSX which indicates that volatility transmition almost none. On the contrary volatility of S&P 500 has effect on for all stock indices as, the p-value is significant. Overall, the results show a range of predictive relationships among these major stock indices, with volatility some markets influence the others, and some being primarily followers.

In the same Table 5.4, the coefficient of standardized residuals (DCC-a) in the dynamic conditional covariance of S&P 500 with SHCOMP, S&P 500 with NIFTY 50 , and SHCOMP with TSX is not significant; which show the conditional correlations between these markets are persistent and tends to stable for longer period of days. New shocks or temporay shocks in returns do not change the correlation of the stocks. On the other hand DCC-b is significant implies that the existing correlation pattern persists, possibly reflecting stable financial or economic linkages over time rather than short-term contagion. However, DCC-a in TSX with S&P 500 is significant and DCC-b is also significant reveals that temporary shocks in the lag stocks alter the dyamic correlation of the stocks or has an effect at covariance matrix. Not only the significance of coefficient, but their magnitude of the parameters influence the covariance matrix.

For agricultural commodities or staple crops in Table 5.5 the asymmeteric coefficient γ is significant and negative for wheat and maize which indicates that bad shocks increases the volatility. The size of β for teff and wheat are 0.2524 and 0.4160 respectively ,which demonstrates that the lag volatility has low impact on the current volatility and has high impact from the previuos shocks. But, for Maize β is 0.96 that indicadtes the variance highly persistent and α is 0.1778 the temporay shocks has low impact on the current variance. The joint correlation coefficients DCC-a is not significant for both teff with wheat and teff with maize that show the standardized residuals has no short term effect on the correlation of assets. In case of wheat with maize eventhough DCC-a is significant the impact is low as the value is 0.0275. The coefficient of lag covariance DCC-b which are 0.98 and 0.89 in teff with wheat and wheat with maize are significant and strongly influences on the current covariance respectively. Regarding to teff with maize DCC-b is moderate.

So, the weekly correlation between these staple crops is persistent and temporay shock of one has little effect on the other.

TABLE 5.8: Descriptive statistics of DCC correlation

Stock	Mean	Max.	Min	Std.Dev	Skewness	Kurtosis
S&P 500 - SHCOMP	0.1185	0.3112	-0.0798	0.0429	-0.1062	2.71
S&P 500 - NIFTY 50	0.2060	0.5625	-0.0923	0.0490	0.3841	6.84
S&P 500 - TSX	0.6745	0.8727	0.2786	0.0344	-2.499	25.26
SHCOMP - NIFTY 50	0.1976	0.5145	-0.1917	0.0512	0.0184	5.52
SHCOMP - TSX	0.1568	0.4051	-0.1284	0.0442	-0.0614	3.31
NIFTY 50 - TSX	0.2496	0.5913	-0.0161	0.0523	0.2145	5.20
Teff - Wheat	0.0815	0.3509	-0.1564	0.0575	0.3308	2.49
Teff - Maize	0.1598	0.6441	-0.0922	0.0746	0.8953	5.94
Wheat - Maize	0.1800	0.5453	-0.0589	0.0669	0.6621	3.54

During the study period the mean of dynamic conditional correlation (DCC) between pairs of stocks, as shown in Table 5.8, reveals that the S&P 500 has a moderate positive correlation - TSX, indicating a fairly strong positive relationship between the S&P 500 and TSX stock market indices. However, the correlations between S&P 500 - SHCOMP, S&P 500 with NIFTY 50, SHCOMP - NIFTY 50, and NIFTY 50 - TSX are weak positive correlations, suggesting a slight tendency for these variables to move together, but not with a strong relationship. The mean of the dynamic correlations of returns in agricultural commodities in Table 5.10 indicates that there is almost no correlation among the pairs of assets, as the maximum value of 0.2358 occurs between wheat and maize..

The numerical values in Table 5.9 demonstrate the constant relationship between the stocks in the study period.

Figure 5.7 shows the dynamic conditional correlations of four stock returns from January

TABLE 5.9: Long-run (unconditional) correlation of the stocks

Stock	S&P 500	SHCOMP	NIFTY 50	TSX
S&P 500	1	0.1393	0.2990	0.8054
SHCOMP	0.1393	1	0.2232	0.1741
NIFTY 50	0.2990	0.2232	1	0.3870
TSX	0.8054	0.1741	0.3870	1

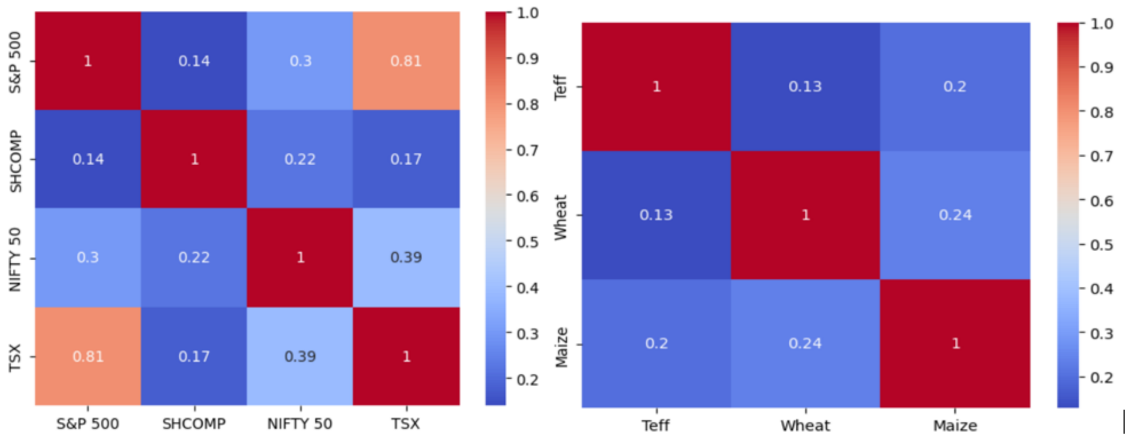

FIGURE 5.6: Long-run (unconditional) correlation heatmap

TABLE 5.10: Long-run (unconditional) correlation of agricultural commodities

Commodity	Teff	Wheat	Maize
Teff	1	0.1283	0.1970
Wheat	0.1283	1	0.2358
Maize	0.1970	0.2358	1

2016 to October 2024 (using a rolling window of 120 days) and the corresponding out-of-sample predictions. Moreover, it demonstrates that the dynamic correlations among all pairs of stocks rise significantly in a positive direction during the COVID-19 pandemic, reflecting periods of market crises and heightened uncertainty that created abnormal market conditions.

Regarding staple crops, Figure 5.8 presents the time-varying conditional correlations of teff, wheat, and maize from 1 September 2010 to 23 June 2022, using a window size of 60 days. The prediction results from the DCC-EGARCH and DCC- EGARCH-LSTM models are also illustrated in the figures.

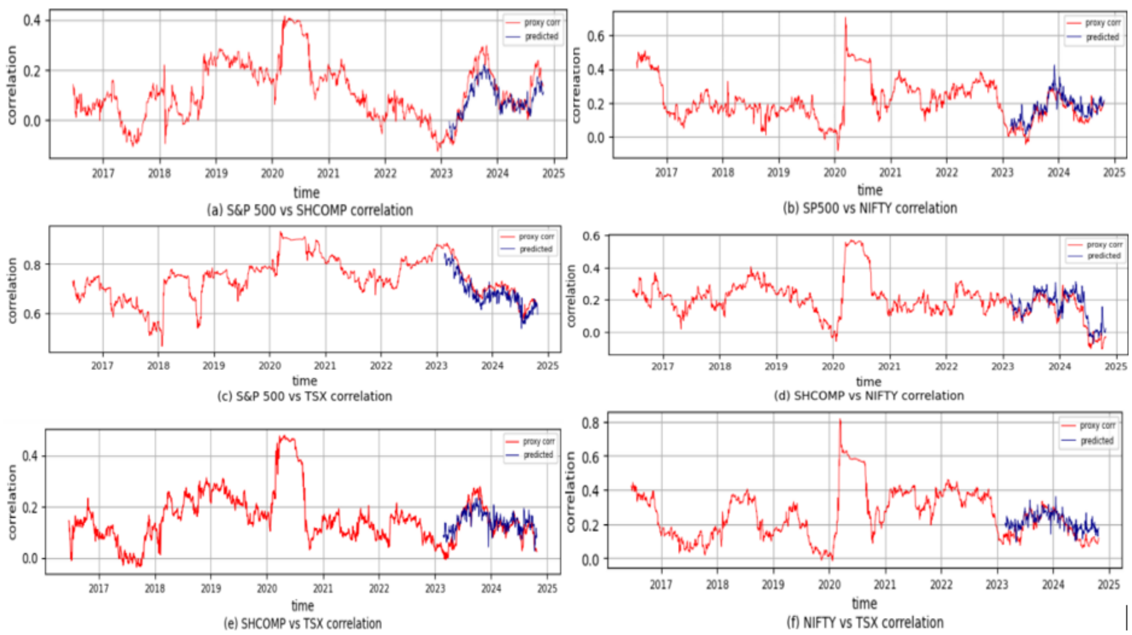

FIGURE 5.7: proxy vs prediction correlations in the out sample data

Additionally, the correlation between S&P 500 - wheat in the same time period and frequency (weekly) price data of the returns it indicates no co movement relation. In particular the Pearson's correlation coeffient is -0.009. Moreover, the mean of the time varying correlation between them in window size 30 and 90 are -0.040 and -0.02 respectively. Additionally the mean, maximum and minimum of the DCC are the same and it is -0.0308 which confirms the variation of S&P 500 return does not have any effect on the wheat retail price return and the vice versa in the study period. This is reasonabliy accepted that the variation of S&P 500 weekly index in US has no any correlation with the weekly retail price of wheat in Ethiopia.

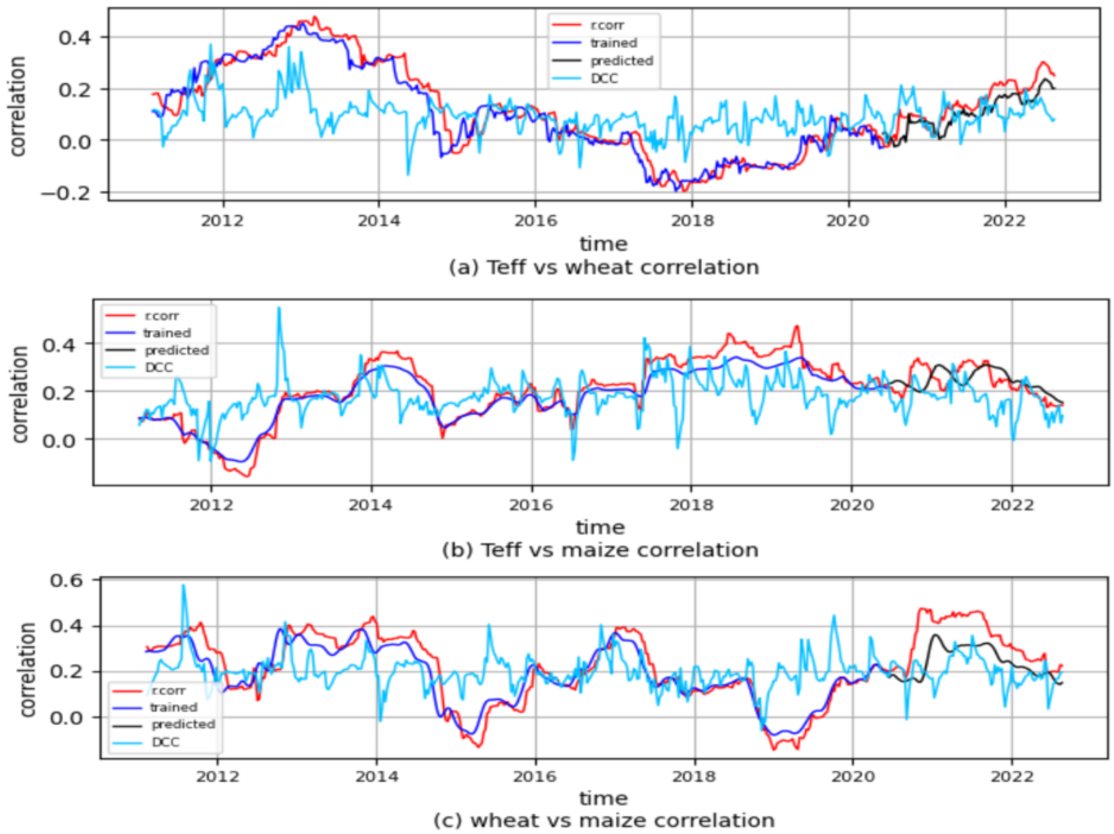

FIGURE 5.8: Time varying correlations of teff,wheat and maize

FIGURE 5.9: Time varying correlations of S&P 500 - wheat

Finally, testing the applicability of integrating LSTM in DCC-EGARCH, we split the data into train and test. Among these 06-20-2016 to 02-22-2023 (1742 trading days) for training and from 02-23-2023 to 10-25-2024 (436 trading days) for test samples. Similar perentage of partition for staple crops from 02-11-2011 to 06-12-2020 for training and 06-19-2020 to 8-23-2022 for test samples. we take 20 lookbacks (time steps), 80% time splits, 11 number of folds, 3 layers, Dense (1), dropout (0.2) and hyperparameters in

Table 5.11. The selection of tuning or hyperparameters is random error (random search method) in minimum metrics. Finally we compare using mean square error and mean absolute errors.

TABLE 5.11: LSTM hyperparametrs

Model	Activation function	Number of units	Optimizer	Epoch	Batch size
LSTM	Sigmoid,tanh	50	Adam(lr=0.0001-0.0005	120-200 loss=MSE)	32

TABLE 5.12: Forecasting performance in outsample data with window size of 120

Pairs of Stocks	DCC-EGARCH		DCC-EGARCH-LSTM	
	MSE	MAE	MSE	MAE
S&P 500 - SHCOMP	0.0110	0.0846	0.0089	0.0802
S&P 500 - NIFTY 50	0.0116	0.0881	0.0107	0.0836
S&P 500 - TSX	0.0081	0.0672	0.0072	0.0629
SHCOMP - NIFTY 50	0.0161	0.0933	0.0152	0.0872
SHCOMP - TSX	0.0069	0.0675	0.0052	0.0600
NIFTY 50 - TSX	0.0114	0.0899	0.0100	0.0848

TABLE 5.13: Forecasting performance in outsample data with rolling window 60

Pairs of Cereals	DCC- EGARCH		DCC-EGARCH-LSTM	
	MSE	MAE	MSE	MAE
teff - wheat	0.0153	0.1046	0.0030	0.0438
teff - maize	0.0162	0.1107	0.0118	0.0956
wheat - maize	0.0552	0.2071	0.0075	0.0715

Table 5.12 presents the mean square error (MSE) and mean absolute error (MAE) for two models taking rolling window size of 120 days as proxy of actual dynamic correlation: DCC-EGARCH and DCC-EGARCH-LSTM, applied to the dynamic conditional correlations of six stock market pairs . S&P 500 with SHCOMP, DCC-EGARCH has MSE of 0.0110 and MAE of 0.0846, while DCC-EGARCH-LSTM performs better with MSE of 0.0089 and MAE of 0.0802. This suggests that adding LSTM to the model significantly reduces both the MSE and MAE, enhancing predictive accuracy for this stock pair. Similarly, for S&P 500 with NIFTY 50 and S&P 500 with TSX, DCC-EGARCH-LSTM outperforms DCC-EGARCH in both metrics (lower MSE and MAE). The MSE for DCC-EGARCH is 0.0116 and 0.0081, reducing to 0.0107 and 0.0072 with LSTM, respectively, similarly by the metrics MAE, which highlights the model's superior ability. In the case of SHCOMP with NIFTY 50, DCC-EGARCH-LSTM achieves lower MSE (0.0152 vs. 0.0161) and slightly lower MAE (0.0872 vs. 0.0933), although the difference is not as pronounced as in some other pairs. This suggests that still DCC-EGARCH-LSTM provides a better fit overall. The DCC-EGARCH-LSTM model performs noticeably better with lower MSE (0.0052 vs. 0.0069) and MAE (0.0600 vs. 0.0675), which indicates its ability to more accurately capture correlation patterns for SHCOMP with TSX. Lastly, in the prediction of dynamic correlation between NIFTY 50 and TSX, DCC-EGARCH-LSTM performs better than DCC-EGARCH.

The performance of DCC-EGARCH and DCC-EGARCH-LSTM in out sample data demonstrates that the values of MSE and MAE for DCC-EGARCH are greater than the values of DCC-EGARCH-LSTM for each pair of stock indices as well as the pairs of staple crops. According to MSE and MAE metrics the performance range from 10 % to 24 % improvement. Based on this, DCC-EGARCH-LSTM performs better than DCC-EGARCH in general, showing improved predictive power and an increase in accuracy. The performance of the models as shown in Table 5.13 there is a significant reduction of MSE and MAE from DCC-EGARCH to DCC-EGARCH combined with LSTM model in each of the three pairs.

TABLE 5.14: Diebold Mariano (DM)test results

	DCC-EGARCH	DCC-EGARCH LSTM
DCC-EGARCH		5.17(2.3e-07)
DCC-EGARCH-LSTM	5.17(2.3e-07)	

The Diebold Mariano test results in Table 5.14 with DM statistics 5.17 and the p-value is almost zero at 5% significant level that confirms in outsample days, the models have comparable accuracy and have different prediction performance in the correlation of stock market dynamics.

The limitation of this part of the study is availability in free source high frequency data , sufficient recent agricultural commodity data and parameter setting as tunning is based on trial and error depending on the nature of data [110]. Additionally what conditions whether economic or political or other that influence the dynamic correlation of the results. These drawbacks create opportunities for future research to explore the relationship and co movement between the volatility of S&P 500, SHCOMP, NIFTY 50, and TSX stock markets and selected commodities in the future time. Additionally, they pave the way for the development of new models that can assist policymakers, investors, and portfolio risk managers in making informed investment decisions in these markets.

5.4 Conclusions

Future market information about stock index, some agricultural commodity price correlation and volatility considerably benefits shareholders, portfolio managers, consumers and financiers to diversify and manage the risk of stock markets and retail asset prices. Understanding the relationship and co-movement of stock markets enables the financial participants for risk alleviation and portfolio diversification. Additionally future information about the volatilty and relationship of teff, wheat and maize helps market decision in trading sterategies.

The estimation of the dynamic conditional correlation (DCC) model with exponential generalized autoregressive conditional heteroskedasticity (EGARCH) plays a significant role in capturing the time-varying relationships between financial markets. In this context, the dynamic correlation of the S&P 500 with other indices, such as the Shanghai Composite Index, the NIFTY 50 of India, and Tronto Composite Index (TSX) of Canada were analyzed to understand the volatility and how stock market interdependencies evolve over time across stocks. The relationship between S&P 500 - SHCOMP, S&P 500 - NIFTY 50, S&P 500 with TSX,SHCOMP - NIFTY 50, and NIFTY 50 - TSX correlation movement were analyzed. Moreover, teff with wheat, teff with maize and wheat with maize with weekly time series data in window size of 60 were discussed.

The combination of LSTM with the DCC-EGARCH model significantly improves the predictive accuracy for forecasting the time varying correlation of stock market indice and commodity asset returns. The performance of the models are compared by the metrics mean square error (MSE) and mean absolute error (MAE) between DCC-EGARCH and DCC-EGARCH-LSTM and explicitly DCC-EGARCH-LSTM performs better than other models for each pairs of stock market indice or argricultural commodoty returns. For example, in the case of S&P 500 with SHCOMP, the DCC-EGARCH-LSTM model both MSE and MAE decreases significantly, indicating its improved ability to capture dynamic correlations more accurately. This measurement is consistent in other stock pairs and the retail price return of teff, wheat and maize correlations, specifically S&P 500 - NIFTY 50, S&P 500 - TSX, and SHCOMP - TSX, where the DCC-EGARCH-LSTM model consistently outperforms the DCC-EGARCH model, achieving lower error metrics. The performance of the model in various dynamic correlation has a range of 10% to

24% accuracy improvements. We believe that, the proposed model is powerful because it leverages the strengths of both statistical and machine learning techniques to handle the complexities of financial data, such as non-linearities, long-range dependencies, and volatility clustering generally.

Chapter 6

Conclusion and Future Works

6.1 Conclusion

Volatility and correlation forecasting are crucial endeavors in the financial assets, serving as the mainstay for risk assessment, portfolio diversification and investment decision. This thesis has aimed to predict the volatility and correlation of financial assets using conventional econometric models hybrid with more recently, deep learning technique approaches that leverage the strengths of both to achieve superior predictive performance.

First, a variety of symmetric and asymmetric GARCH-type models were estimated using different error distributions. The analysis of the selected staple crops shows that the GARCH(1,1)–GED model provided the best fit for teff and barley price volatility, while the EGARCH(1,1)–STD model was most appropriate for wheat and maize. Evidence of asymmetry- where negative shocks exert a stronger influence on volatility than positive shocks-was observed for both wheat and maize. Then, we examined hybrid deep learning models to identify a better volatility forecasting approach for the S&P 500. In particular, the GARCH(1,1)-GED model integrated with CNN and LSTM significantly improved the accuracy of S&P 500 volatility forecasts compared with both the individual models and the other hybrid models tested

Secondly, the DCC-EGARC-LSTM model is the culmination of our thesis. It offers a specialized LSTM recurrent neural network in exponential GARCH with dynamic conditional correlation process. This multivariate model outperforms several established models and demonstrates versatility across diverse conventional DCC in financial data sets. This hybrid approach is a reliable and robust method for time varying volatility and correlation forecasting.

The main focus of this thesis is the complementary relationship between traditional econometric models and modern machine-learning techniques. While each has merits and limitations, their integration can lead to more accurate, interpretable, and versatile volatility models. This is particularly important in the returns of financial markets, where dynamics avenue are not simple.

The overall contribution of the research findings of this thesis are:

- Introduce novel hybrid econometric-deep learning GARCH-CNN-LSTM volatility predictive model.
- We developed new DCC-EGARCH-LSTM model for better accurate and flexible time varying correlation forecasting.
- It provides pre-information for stalkhoders like farmers, investors, traders and policy makers about the risk associated with market fluctuation of assets.
- Accurate prediction of volatility and correlation enables to reduce market risk, portfolio diversificaion and efficient policy decision analysis.
- It provides the motivational information for researchers in the related thematic area.

6.2 Future Works

In future, we will consider the following:

- **Incorporation of Additional Market Factors**

  Future research could aim to include more comprehensive data inputs beyond close price alone, such as trading volume, market sentiment indicators, macroeconomic variables, and financial stress indices and other market areas. These factors could enhance the robustness and accuracy of volatility and correlation forecasting models.

- **Model Optimization and Hybridization**

  Further refinement of hybrid models that combine econometric methods (like ARFIMA) with advanced machine learning techniques (like GRU, CNN, Transformers) considering jump components are essential. This includes exploring better hyperparameter tuning, model architecture optimization, and integration of emerging AI algorithms for improved predictive performance.

- **Broader Asset Classes and Market Conditions**

  Extending the application of forecasting models to a wider variety of asset classes (including Gold, exchange rate index, consumer price index, cryptocurrencies, bonds, and commodities beyond staples like teff, wheat, and maize) and testing in diverse market conditions such as crises like civil war or high volatility periods will provide greater insight into model generalizability and resilience.

- **Real-time and High-Frequency Data Utilization**

  Utilizing high-frequency or intraday data could improve the granularity and real-time forecasting capabilities of volatility and correlation models, aiding timely market decisions and risk management. These directions align with ongoing challenges in volatility forecasting, such as handling non-linearity, adapting to regime changes, and incorporating a richer set of predictive variables, aiming to empower market participants with timely and accurate insights for better decision-making.

These directions align with ongoing challenges in volatility forecasting, such as handling non-linearity, adapting to regime changes, and incorporating a richer set of predictive variables, aiming to empower market participants with timely and accurate insights for better decision-making.

Bibliography

[1] Bouchaud, J.-P., Matacz, A., and Potters, M. (2001). Leverage effect in financial markets: The retarded volatility model. Physical review letters, 87(22):228701.

[2] Chen, Y.-S. and Wang, S.-T. (2004). The empirical evidence of the leverage effect on volatility in international bulk shipping market. Maritime Policy & Management, 31(2):109-124.

[3] Su, C. (2010). Application of egarch model to estimate financial volatility of daily returns:The empirical case of china.

[4] Miron, D. and Tudor, C. (2010). Asymmetric conditional volatility models: Empirical estimation and comparison of forecasting accuracy. Romanian Journal of Economic Forecasting, 13(3):74-92.

[5] Gokcan, S. (2000). Forecasting volatility of emerging stock markets: linear versus non-linear garch models. Journal of forecasting, 19(6):499-504.

[6] Gabriel, A. S. (2012). Evaluating the forecasting performance of garch models. evidence from romania. Procedia-Social and Behavioral Sciences, 62:1006-1010.

[7] Aizenmen, J., & Marion,N.P. (1993). Policy uncertainity, persistence and growth.Review of international Economics, 145-163.

[8] Alentorn, A. (2004). Modeling the implied volatility surface :an emperical study for ftse options. Retrieved from www.theponytail.net/CCFEA.

[9] AL-Najjar. (2016). Modeling and Estimation of volatility using ARCH/GARCH models in Jordan's Stock market. Asian journal of Finance and Acounting, 152-167.

[10] Atoi, N. (2014). Testing voltility in Nigeria stock market Using GARCH models. CBN Journal of Applied Statistics(JAS), 4.

[11] Birau, R., Trivedi,J, & Antonescu. (2015). Modelig S &P bomby stockexchange bankex volatility paterns using garch model. Procedia Economics and Finnance, 520-525.

[12] Bollerslev, T., Engle,R.F, & Nelson,D.B. (1994). ARCH models. Hand book of Econometrics, 2959-3038.

[13] Cont, R. (2007). Volatiity Clusteringin financial markets:emperical facts andagent based models. Long memory in economics, 289-309.

[14] Dinku, T. (2021). Modeling price volatility for selected Agricultural commodities in Ethiopia: Application of GARCH modls. International Journal of Environmental Science, 6.

[15] Engle, R., Granger,C.W, & Kraft,D. (1984). Combining competing forecasts of inflation using a bivariate arch model. Journal of Economic Dynamics and control, 151-165.

[16] Frimpong, J., & Oteng-Abayie. (2006). Modeling and Forecasting volatility of returns on Gana stock exchange using Garch models.

[17] Hailegebrial., Z. (2015). Determinants of food price inflation in Ethiopia . Food Science and Quality management, 30-43.

[18] Haji, J., Gelaw,F., Bekele,W., & Tesfaye,G. (2011). The 'black-box' of Ethiopian Agricultural produces price formation and its determinants within the current liberized market policy. Drylands Coordination group,Mijihuset G, 9.

[19] Huchet-Bourdon. (2011). Agricultural price volatility:an overview.

[20] Kose, M., Prasad,E.S., & Terrones,M.E. (2006). How do trade and financial integration affct the relation ship between growth and volatility. Journal of International Economis, 176-202.

[21] Koutmos, G., & Booth,G.G. (1995). Asymmetric volatility transmission in international stock markets. Journal of international Money and finance, 747-762.

[22] Kuhe, D. (2019). Modeling volatility of agricultural commodity food price indexin Nigeria using ARIMA-GARCH models. Mediterranean Jornal of Basic and Applied Sciences, 55-75.

[23] Musunuru, N., Yu,M, & Larson,A. (2013). Forecasting Volatility of reurnsfor corn Using GARCH models. Texas Journal of Agriculture and Natural Resources, 42-55.

[24] Nelson, D., & Cao,C.Q. (1992). Inequality constraints in the univariate GRCH model. Journal of Business & Economic Statistics, 229-235.

[25] Niu, H. a. (2013). Volatility clusterng and long memory of financial time series and financial price model. Digital signal processing, 489-498.

[26] Niyitegka, O., & Tewar,D. (2013). Volatility clustering at Johannesburg stock exchange: Investigation and Analysis. Mediterranean Journal of Social Sciences, 621-642.

[27] Santeramo, F. (2018). Emerging trends in european food ,diets and food industry. Food research international, 39-47.

[28] Schmidt, L. (2021). Volatility forecasting performance of garch models: Astudy on nordic indices during Covid-19.

[29] Staugaitis, A., & Vaznonis, B. (2022). Short-Term Speculation Effects on Agricultural Commodity Returns and Volatility in the European Market Prior to and during the pandemic. Agriculture (MDPI), 1-26.

[30] Tseng, J. a. (2012). Quantifying Volatility clustering in Financial time series. International Review of Financial Analysis , 11-19.

[31] Wang, C. (2004). Emperical evidence of the leverage effect in international bulk shiping market. Maritime policy & management, 109-124.

[32] Yip, P., Brooks,R, & Nguyen. (2020). Dynamic volatility spillover effects between oil and agricultural products. International Review of Finanial Analysis.

[33] Yu, J. (2005). On leverage in stochastic volatility model. Journal of Econometrics, 165-17.

[34] Zakoian, J. (1994). Threshold heteroskedastic models. Journal of Economic Dynamics and control, 931-955.

[35] Zivot, & Wang,G. (2006). Modeling financial time series withS-PLUS. Springer

[36] Ahmed, K. (2021). Ethiopian Agricultural Market Price Data Analysis and Prediction using Deep Learning Models. PhD thesis.

[37] Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. Econometrica, 71(2):579–625.

[38] Ardabili, S., Mosavi, A., and Várkonyi-Kóczy, A. R. (2019). Advances in machine learning modeling reviewing hybrid and ensemble methods. In International Conference on Global Research and Education, pages 215–227. Springer.

[39] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of econometrics, 31(3):307–327.

[40] Bulut, C. and Hudaverdi, B. (2022). Hybrid approaches in financial time series forecasting: a stock market application. Ekoist Journal of Econometrics and Statistics, (37):53–68.

[41] Cao, J. and Wang, J. (2019). Stock price forecasting model based on modified convolution neural network and financial time series analysis. International Journal of Communication Systems, 32(12):e3987.

[42] Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pages 785–794.

[43] Chen, Y., Wang, Y., Dong, Z., Su, J., Han, Z., Zhou, D., Zhao, Y., and Bao, Y. (2021). 2-d regional short-term wind speed forecast based on cnn-lstm deep learning model. Energy Conversion and Management, 244:114451.

[44] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: Journal of the econometric society, pages 987–1007.

[45] Hajizadeh, E., Seifi, A., Zarandi, M. F., and Turksen, I. (2012). A hybrid modeling approach for forecasting the volatility of s&p 500 index return. Expert Systems with Applications, 39(1):431–436.

[46] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8):1735–1780.

[47] Hu, Y., Ni, J., and Wen, L. (2020). A hybrid deep learning approach by integrating lstm-ann networks with garch model for copper price volatility prediction. Physica A: Statistical Mechanics and its Applications, 557:124907.

[48] Huang, C.-J. and Kuo, P.-H. (2018). A deep cnn-lstm model for particulate matter (pm2. 5) forecasting in smart cities. Sensors, 18(7):2220.

[49] Jumoorty, A. F., Thoplan, R., and Narsoo, J. (2023). High frequency volatility forecasting: A new approach using a hybrid ann-mc-garch model. International Journal of Finance & Economics, 28(4):4156–4175.

[50] Kakade, K., Jain, I., and Mishra, A. K. (2022a). Value-at-risk forecasting: A hybrid ensemble learning garch-lstm based approach. Resources Policy, 78:102903.

[51] Kakade, K., Mishra, A. K., Ghate, K., and Gupta, S. (2022b). Forecasting commodity market returns volatility: A hybrid ensemble learning garch-lstm based approach. Intelligent Systems in Accounting, Finance and Management, 29(2):103–117.

[52] Kim, H. Y. and Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. Expert Systems with Applications, 103:25–37.

[53] Kim, T.-Y. and Cho, S.-B. (2019). Predicting residential energy consumption using cnn-lstm neural networks. Energy, 182:72–81.

[54] Koo, E. and Kim, G. (2022). A hybrid prediction model integrating garch models with a distribution manipulation strategy based on lstm networks for stock market volatility. IEEE Access, 10:34743–34754.

[55] Koutmos, G. and Booth, G. G. (1995). Asymmetric volatility transmission in international stock markets. Journal of international Money and Finance, 14(6):747–762.

[56] Kristjanpoller, W. and Hernández, E. (2017). Volatility of main metals forecasted by a hybrid ann-garch model with regressors. Expert Systems with Applications, 84:290–300.

[57] Lago, J., De Ridder, F., and De Schutter, B. (2018). Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. Applied Energy, 221:386–405.

[58] McAleer, M., Chan, C. A. I., & Tong, D. A. D.(1996). Asymmetric GARCH models in financial econometrics. Journal of Econometrics, 72(1–2), 219–229.

[59] Nelson, D. B., & Foster, M. R. (1995). Asymmetric effects in stock returns and volatility. Journal of Finance, 50(2), 403–426.

[60] Glosten, L. R., Jagannathan,R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance, 48(5), 1779–1801.

[61] Almeida, R. M. F. (2008). Volatility forecasting with asymmetric GARCH models. Journal of Financial Markets, 11(2), 177–203.

[62] Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems with Applications, 103, 25–37.

[63] Kristjanpoller, W., & Minutolo, M. C. (2016). Forecasting volatility of oil price using an artificial neural network–GARCH model. Expert Systems with Applications, 65, 233–241.

[64] Jardine, A.K.S., Carver, R.B.,& Yu, D.L.K.C. (2020). Forecasting volatility using a hybrid ARIMA-GARCH and deep learning model. International Journal of Forecasting, 36(6), 1971–1987.

[65] Amirshahi, B., & Lahmiri, S. (2023). Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. Machine Learning with Applications, 12, 100465

[66] Bauwens, L., & Otranto, E. (2020). Nonlinearities and regimes in conditional correlations with different dynamics. Journal of Econometrics, 217(2), 496–522.

[67] Lu, W., Li, J., Li, Y., Sun, A., and Wang, J. (2020). A cnn-lstm-based model to forecast stock prices. Complexity, 2020(1):6622927.

[68] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the econometric society, pages 347–370.

[69] Song, H. and Choi, H. (2023). Forecasting stock market indices using the recurrent neural network based hybrid models: Cnn-lstm, gru-cnn, and ensemble models. Applied Sciences, 13(7):4644.

[70] Tapia, S. and Kristjanpoller, W. (2022). Framework based on multiplicative error and residual analysis to forecast bitcoin intraday-volatility. Physica A: Statistical Mechanics and its Applications, 589:126613.

[71] Vidal, A. and Kristjanpoller, W. (2020). Gold volatility prediction using a cnn-lstm approach.Expert Systems with Applications, 157:113481.

[72] Tsay, R. S. (2005). Analysis of financial time series. John wiley & sons.

[73] Widodo, A. and Yang, B.-S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. Mechanical systems and signal processing, 21(6):2560–2574.

[74] Wu, J. M.-T., Li, Z., Herencsar, N., Vo, B., and Lin, J. C.-W. (2023). A graph-based cnn-lstm stock price prediction algorithm with leading indicators. Multimedia Systems, 29(3):1751–1770.

[75] Zhang, J. (2023). Some application of machine learning in quantitative finance: model calibration,volatility formation mechanism and news screening. PhD thesis, Institut Polytechnique de Paris.

[76] Zhang, S., Li, C., Qiu, S., Gao, C., Zhang, F., Du, Z., and Liu, R. (2019). Emm-cnn: An etps-based multi-scale and multi-feature method using cnn for high spatial resolution image land-cover classification. Remote Sensing, 12(1):66.

[77] Engle, R.F. (2002) Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models. Journal of Business and Economic Statistics, 20, 339-350.

[78] Aielli, G. P. (2013). Dynamic Conditional Correlation: On Properties and Estimation. Journal of Business & Economic Statistics, 31(3), 282–299.

[79] Wang, H., Liu, C., Tran, M.-N., & Wang, C. (2025, June). Deep Learning Enhanced Multivariate GARCH. arXiv.

[80] Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: A survey. Journal of Applied Econometrics, 21(1), 79–109. https://doi.org/10.1002/jae.842

[81] Lee, H.-T. (2010). Regime switching correlation hedging. Journal of Banking & Finance, 34(11), 2728–2741. https://doi.org/10.1016/j.jbankfin.2010.05.010

[82] Dessie, E., Birhane, T., Mohammed, A., and Walelign, A. (2025). A hybrid garch, convolutional neural network and long short term memory methods for volatility prediction in stock market. J. COMBIN. MATH. COMBIN. COMPUT, 124(461):476.

[83] Singh, P., Kumar, B., and Pandey, A. (2009). Price and volatility spillovers across north american, european and asian stock markets: With special focus on indian stock market. European and Asian Stock Markets: With Special Focus on Indian Stock Market (January 7, 2009).

[84] Mahmood, T., Ahmad, I., Ansar, M. M. Z., Darwish, J. A., & Sherwani, R. A. K. (2024). Comparative study of Long Short-Term Memory (LSTM) and Quantum Long Short-Term Memory (QLSTM): Prediction of stock market movement. arXiv preprint arXiv:2409.08297. https://arxiv.org/abs/2409.08297

[85] Petrozziello, A., Troiano, L., Serra, A., Jordanov, I., Storti, G., Tagliaferri, R., & La Rocca, M. (2022). Deep learning for volatility forecasting in asset management. Soft Computing, 26(17), 8553–8574. https://doi.org/10.1007/s00500-021-06304-3

[86] Alparslan, Y. and Kim, E. (2021). Extreme volatility prediction in stock market: When gamestop meets long short-term memory networks.

[87] Liu, Y. (2019). Novel volatility forecasting using deep learning–long short term memory recurrent neural networks. Expert Systems with Applications, 132, 99–109. https://doi.org/10.1016/j.eswa.2019.04.048

[88] Beniwal, M., Singh, A., & Kumar, N. (2024). Forecasting multistep daily stock prices for long-term investment decisions: A study of deep learning models on global indices. Engineering Applications of Artificial Intelligence, 129, 107617. https://doi.org/10.1016/j.engappai.2023.107617

[89] Samitas, A., Kampouris, E., and Kenourgios, D. (2020). Machine learning as an early warning system to predict financial crisis. International Review of Financial Analysis,71:101507.

[90] Moews, B. and Ibikunle, G. (2020). Predictive intraday correlations in stable and volatile market environments: Evidence from deep learning. Physica A: Statistical Mechanics and its Applications, 547:124392.

[91] Sui, L. and Sun, L. (2016) . Spillover effects between exchange rates and stock prices: Evidence from brics around the recent global financial crisis. Research in International Business and Finance, 36:459–471.

[92] Meher, B. K., Kumar, S., Birau, R., Anand, A., Kumar, S., Lupu, G. A. M., Simion, M. L., and Cirjan, N. T. (2024). Comparative volatility analysis of usa and china stock market indices using garch family models. Revista de Stiinte Politice, (83):108–128.

[93] Kaur, S., Aggarwal, S., and Arora, V. (2024). Co movement of stock market of brics with g7 stock market. Asia-Pacific Financial Markets, pages 1–30.

[94] Ni, J., & Xu, Y. (2023). Forecasting the dynamic correlation of stock indices based on deep learning method. Computational Economics, 61(1), 35–55.

[95] Chen, M., Li, N., Zheng, L., Huang, D., and Wu, B. (2022). Dynamic correlation of market connectivity, risk spillover and abnormal volatility in stock price. Physica A: Statistical Mechanics and Its Applications, 587:126506.

[96] Ceylan, Ö. (2021). Dynamics of global stock market correlations: the vix and attention allocation. Journal of Applied Economics, 24(1):392–400.

[97] Cai, X. J., Tian, S., & Hamori, S. (2016). Dynamic correlation and equicorrelation analysis of global financial turmoil: Evidence from emerging East Asian stock markets. Applied Economics, 48(40), 3789–3803. https://doi.org/10.1080/00036846.2015.1136394

[98] Yin, K., Liu, Z., and Liu, P. (2017). Trend analysis of global stock market linkage based on a dynamic conditional correlation network. Journal of Business Economics and Management, 18(4):779–800.

[99] Ji, X., Wang, S., Xiao, H., Bu, N., & Lin, X. (2022). Contagion effect of financial markets in crisis: An analysis based on the DCC–MGARCH model. Mathematics, 10(11), 1819.

[100] Sclip, A., Dreassi, A., Miani, S., & Paltrinieri, A. (2016). Dynamic correlations and volatility linkages between stocks and sukuk: Evidence from international markets. Review of Financial Economics, 31, 34–44.

[101] Asai, M. and McAleer, M. (2012). Dynamic conditional correlations for asymmetric processes. Journal of the Japan Statistical Society, 41(2):143–157.

[102] Tang, C. and Aruga, K. (2021). Relationships among the fossil fuel and financial markets during the covid-19 pandemic: evidence from bayesian dcc-mgarch models. Sustainability,14(1):51.

[103] Ottosson, I. and Rauf, N. (2024). Cryptocurrency correlation modeling with multivariate garch. Master's Thesis in Mathematical Sciences.

[104] Cook, A. A., Mısırlı, G., and Fan, Z. (2019). Anomaly detection for iot time-series data: A survey. IEEE Internet of Things Journal, 7(7):6481–6494.

[105] Dessie, E., Brhane, T., Mohammed, A., and Walelign, A. (2023). Agricultural commodity price volatility models and prediction performance of garch family models. Horn of African Journal of Business and Economics (HAJBE), 6(2):150–171.

[106] Ardabili, S., Mosavi, A., & Várkonyi-Kóczy, A. R. (2019). Advances in machine learning modeling: Reviewing hybrid and ensemble methods. In International Conference on Global Research and Education (pp. 215–227). Springer

[107] Zhang, S., Li, C., Qiu, S., Gao, C., Zhang, F., Du, Z., & Liu, R. (2019). EMMCNN: An ETPS-based multi-scale and multi-feature method using CNN for high spatial resolution image land-cover classification. Remote Sensing, 12(1), 66. https://doi.org/10.3390/rs12010066 arxiv.org+14

[108] Phung, V. H., & Rhee, E. J. (2018). A deep learning approach for classification of cloud image patches on small datasets. Journal of Information and Communication Convergence Engineering, 16(3), 173–178

[109] Zhou, R., & Zhang, Y. (2022). Reconstruction of missing spring discharge by using deep learning models with ensemble empirical mode decomposition of precipitation. Environmental Science and Pollution Research, 29(54), 82451-82466.

[110] Lago, J., De Ridder, F., and De Schutter, B. (2018). Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. Applied Energy, 221:386-405.

[111] Cappiello, L., Engle, R. F., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. Journal of Financial Econometrics, 4(4), 537–572.

[112] Bekaert, G., Engle, R., & Wu, M. (2007). Volatility and correlation in emerging markets. Journal of Emerging Market Finance, 6(3), 207–228.

[113] K.-C. Ke and M.-S. Huang, "Quality prediction for injection molding by using a multilayer perceptron neural network," Polymers, vol. 12, no. 8, p. 1812, 2020

[114] Aurangzeb, K., M. Alhussein, K. Javaid, and S. I. Haider. 2021. A pyramid-cnn based deep learning model for power load forecasting of similar-profile energy customers based on clustering. IEEE Access 9:14992–15003.

[115] Kuo, P.-H., and C.-J. Huang. 2018. A green energy application in energy management systems by an artificial intelligence-based solar radiation forecasting model. Energies 11 (4):819.

[116] Gupta, Anshu, and Radhika Mishra. "A Comprehensive Study of Stock Market Volatility: Types, Determinants, and Measurement Methods." International Journal of Finance & Management Research, vol. 6, no. 6, November-December 2024.

[117] Jeon, Byung-Cheol, Manki Kim, and Seok-Ho Yoon. "Comparison of volatility measures: high-frequency-based versus daily-based measures." Journal of Forecasting 39, no. 1 (2020): 49-63. Wiley.

[118] Liu, Jun, John M. Maheu, and Tianyang Yang. "Which realized volatility estimator is best? Evidence from large-scale simulation and real data." Journal of Financial Econometrics 13, no. 1 (2015): 34-56. Oxford University Press.

[119] Zhang, Lei, Per-Anton Mykland, and Yacine Aït-Sahalia. "High-frequency financial econometrics." Annual Review of Financial Economics 12 (2020): 111-135. Annual Reviews.

[120] Mandelbrot, B. B. (1963). The variation of certain speculative prices. The Journal of Business, 36(4), 394-419.

[121] Fama, E. F. (1965). The behavior of stock-market prices. The Journal of Business, 38(1), 34-105.

[122] Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. The Economic Journal, 119(534), 158-171.

[123] Bekaert, G., Harvey, C. R., Lundblad, C., & Siegel, S. (2014). Political risk and international valuation.

[124] Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64(2), 253–280.

[125] Goodfellow, A. (2016). Ian; Bengio, Yoshua; Courville. Deep Learning.

[126] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. Neural Computation, 31(7), 1235–1270.

[127] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[128] Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.

[129] Yang, J., & Chen, L. (2021). Volatility forecasting using convolutional neural networks. Applied Soft Computing, 105, 107272.

[130] Bai, Z., & Wang, W. (2020). Volatility forecasting with LSTM model. Finance Research Letters, 35, 101315.

[131] Lee, Y. H., & Pai, T. Y. (2010). REIT volatility prediction for skew-GED distribution of the GARCH model. Expert Systems with Applications, 37(7), 4737-4741.

[132] Alshenawy, F., & Abdo, D. A. (2023). Using Multivariate Dynamic Conditional Correlation GARCH model to analysis financial market data. , 45(4), 34

[133] De Almeida, D., Hotta, L. K., & Ruiz, E. (2018). MGARCH models: Trade-off between feasibility and flexibility. International Journal of Forecasting, 34(1), 45-63.

[134] Lahmiri, S. (2017). Modeling and predicting historical volatility in exchange rates: A comparison of GARCH and nonlinear approaches. Physica A: Statistical Mechanics and its Applications, 483, 405–418.

[135] López de Prado, M. (2018). Advances in Financial Machine Learning. Hoboken, NJ: John Wiley & Sons.

[136] Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. Journal of Business & Economic Statistics, 13(3), 253-263.

[137] Patton, A. J. (2009). Copula–based models for financial time series. In Handbook of financial time series (pp. 767-785). Berlin, Heidelberg: Springer Berlin Heidelberg.

[138] Chung, V., Espinoza, J., & Quispe, R. (2025). Forecasting Financial Volatility Under Structural Breaks: A Comparative Study of GARCH Models and Deep Learning Techniques. Journal of Risk and Financial Management, 18(9), 494.

[139] Ranjan, P., Itani, R., & Faccia, A. (2025). An Interpretable 1D-CNN Framework for Stock Price Forecasting: A Comparative Study with LSTM and ARIMA. FinTech, 4(4), 63.

[140] Lang, C., Steinborn, F., Steffens, O., & Lang, E. W. (2019). Electricity Load Forecasting–An Evaluation of Simple 1D-CNN Network Structures. arXiv preprint arXiv:1911.11536.

[141] Du, G., & Chen, X. (2024). Time Series Prediction Model Based on Convolutional Neural Networks with Attention. Academic Journal of Computing & Information Science, 7(6), 81-87.

[142] Wibawa, A. P., Utama, A. B. P., Elmunsyah, H., Pujianto, U., Dwiyanto, F. A., & Hernandez, L. (2022). Time-series analysis with smoothed Convolutional Neural Network. Journal of big Data, 9(1), 44.

[143] Ishida, K., Ercan, A., Nagasato, T., Kiyama, M., & Amagasaki, M. (2024). Use of one-dimensional CNN for input data size reduction in LSTM for improved computational efficiency and accuracy in hourly rainfall-runoff modeling. Journal of Environmental Management, 359, 120931.

[144] Guessoum, S., Belda, S., Ferrandiz, J. M., Modiri, S., Raut, S., Dhar, S., ... & Schuh, H. (2022). The short-term prediction of length of day using 1D convolutional neural networks (1D CNN). Sensors, 22(23), 9517

[145] Kim, J., Oh, S., Kim, H., & Choi, W. (2023). Tutorial on time series prediction using 1D-CNN and BiLSTM: A case example of peak electricity demand and system marginal price prediction. Engineering Applications of Artificial Intelligence, 126, 106817.

[146] Karmakar, M., Dutta, A., & Roy, S. (2022). A hybrid deep learning approach for volatility forecasting: Integrating GARCH and BiLSTM models. Applied Soft Computing, 127, 109354.