Essays on Examining Financial Markets' Dynamics and Forecasting by Deep Learning and Econometrics Models

Parisa Foroutan

A Thesis
in John Molson School of Business

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Business Administration) at
Concordia University

Montreal, Quebec, Canada

June 2024

© Parisa Foroutan, 2024

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Parisa Foroutan

Entitled: Essays on Examining Financial Markets' Dynamics and Forecasting by Deep Learning and Econometrics Models

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy (Business Administration - Supply Chain and Business Technology Management)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Bianca Grohmann — Chair

Dr. Roger Nkambou — External Examiner

Dr. Rustam Vahidov — Arm's Length Examiner

Dr. Yu Ma — Examiner

Dr. Danielle Morin — Examiner

Dr. Salim Lahmiri — Thesis Supervisor (s)

Approved by

Dr. Tracy Hecht, Graduate Program Director — Chair of Department or Graduate Program Director

06/06/2024
Date of Defence

Dr. Anne-Marie Croteau — Dean, John Molson School of Business

Abstract

Three Essays on Examining Financial Markets' Dynamics and Forecasting by Deep Learning and Econometrics Models

Parisa Foroutan, Ph.D.
Concordia University, 2024

Understanding the dynamics of financial markets specially during financial crises and being able to forecast these markets are crucial for policymakers and investors. This dissertation aims to explore the dynamics of Crude oil, Gold, Silver, and Cryptocurrency markets from various perspectives.

The first topic of the dissertation involves comparing the dynamics of cryptocurrencies, crude oil, and gold markets before and during the COVID-19 pandemic. This topic comprises two research studies: First, we investigated the effect of COVID-19 pandemic on the return-volume and return-volatility relationships of crude oil, gold, and ten-most traded cryptocurrency markets. The findings of the first study enable policymakers and investors to better react to the dynamics of digital currencies, and commodity markets during financial crisis. Then, using statistical and econometrics methods, we examined the interactions between these markets before and after the COVID-19 pandemic and investigated whether gold or crude oil can play a safe-haven role for cryptocurrency markets during the pandemic crisis. This study assists hedge fund managers or individual investors to adapt their risk exposure to crude oil, gold, and cryptocurrency markets during the financial crises.

For the second topic, several deep learning, machine learning, and hybrid models are adapted to improve the forecasting of crude oil, gold, and silver markets. For this purpose, I implemented sixteen different deep learning and machine learning models on historical price data and compared the prediction performance of these models across four different input sequence lengths to find the optimal settings in forecasting each market. The findings of this study assist investors, policymakers, and governmental agencies to effectively anticipate market trends and make informed timely decisions regarding crude oil, gold, and silver markets.

Lastly, I propose three graph-based neural networks models to predict the direction of price movements in crude oil, gold, and silver markets using a comprehensive set of features such as historical price data, global macroeconomic factors, supply and demand-related factors, other financial markets, and technical indicators. The proposed graph-based models consider the relationship among various factors that can affect the direction of price movements in crude oil and precious metal markets and can be considered as a feature extraction module for predicting the future trend of crude oil, gold, and silver markets.

Keywords: Crude oil, Precious metals, Cryptocurrency, Forecasting, Deep Learning, Graph Neural Networks, COVID-19

Acknowledgements

I extend my deepest gratitude to my esteemed supervisor, Dr. Salim Lahmiri, for his invaluable guidance, insightful feedback, and constant encouragement throughout my doctoral journey. His expertise, encouragement, and commitment to my academic growth have been instrumental in the completion of this research.

I would like to thank the faculty of the Supply Chain and Business Technology Management department at John Molson School of Business. The stimulating learning environment they fostered, significantly contributed to my intellectual growth and prepared me for this research endeavor.

My heartfelt thanks go to my beloved family—my mother, Golshohreh, my father, Behrouz, and my sister, Shiva. Mom, your boundless love, and patience have been my pillars of strength miles away from home. Dad, your constant support and understanding have been my lifeline, keeping me going through all the ups and downs. Shiva, you are my rock, always there to celebrate my successes and offer a shoulder to cry on during setbacks. Thank you all for believing in me and giving me the strength to achieve this dream.

Special thanks to my amazing partner, Babak Zamiri, for his unwavering support, patience, and encouragement. His belief in my capabilities has been the spark fueling my resilience and determination.

To my friends in Canada, who have become my family away from home, I extend my gratitude for the shared experiences, and support. I also want to convey my deep affection and appreciation to my friends back in my home country. Although miles apart, your thoughts have always been with me. I miss you all dearly and cherish the memories we share.

Last but not least, I would like to express my sincere appreciation of my mentor, Dr. Reza Dehghan who taught me how to truly appreciate life and the preciousness of time. I am forever grateful for your belief in my potential and for always pushing me to strive for excellence.

A heartfelt thank you to each one of you for being an essential part of both my academic and personal journey. Whether through significant gestures or small acts of kindness, your contributions have carved an unforgettable presence in this chapter of my life.

Dedication

To my wonderful mother, father, and sister for their unconditional love and support.

Contribution of Authors

This thesis consists of four research papers that I have co-authored with my supervisor, Dr. Salim Lahmiri.

1. Foroutan, P., Lahmiri, S., (2022) The effect of COVID-19 pandemic on return-volume and return-volatility relationships in cryptocurrency markets. *Chaos, Solitons & Fractals* 162:112443.
2. Foroutan, P., Lahmiri, S., (2024) Connectedness between cryptocurrency markets and crude oil and gold markets: An analysis of the effect of COVID-19 pandemic. *Financial Innovation* 10, 68 (2024).
3. Foroutan, P., Lahmiri, S., (2024) Deep Learning Approaches for Forecasting the Crude Oil and Precious Metals Prices. *Financial Innovation*.
4. Foroutan, P., Lahmiri, S., (2024) Deep Learning-based Spatial-Temporal Graph Neural Networks for Price Movement Classification in Crude Oil and Precious Metals Markets. *Machine Learning with Applications* 16:100552.

Authors' contributions:

Parisa Foroutan: Conceptualization, methodology, software, coding, validation, formal analysis, investigation, data gathering, data curation, writing original draft, writing review and editing.

Salim Lahmiri: Conceptualization, methodology, review, and editing.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Exploring financial markets and predicting their movements is critical in the fields of economics and investments. These markets serve as the lifeblood of global economies, impacting everything from personal savings to global trade and political ties. Accurate predictions provide crucial insights for investors, businesses, and policymakers, enabling them to make informed decisions, manage risks, and navigate the constantly evolving landscape of financial opportunities and challenges. In a world where financial markets are intricately interconnected, the ability to foresee their dynamics becomes an invaluable tool for achieving financial goals and economic stability.

This study primarily concentrates on financial markets, with a specific focus on crude oil, precious metal, and cryptocurrency markets. Non-renewable commodities that are usually mined in certain countries can strongly impact their economies, policies, currencies, and international or political issues. Energy and precious metals markets, among other commodities, are well-known alternatives to stock markets (Pullen et al., 2014; Hussain Shahzad et al., 2017; Akbar et al., 2019; Adekoya et al., 2022; Phan et al., 2016; Sarwar et al., 2019). Crude oil is considered a strategic market due to its critical role in powering economies, industries, and transportation systems around the world. It serves as a fundamental energy source, making it indispensable for countries' economic stability and growth. Fluctuations in crude oil prices have profound implications for a country's political and economic security. In this regard, understanding the dynamics of such markets and forecasting their evolutions is crucial for portfolio optimization and management.

Likewise, gold is a commodity asset of certain importance for investment portfolios diversification and hedging (ben Khelifa et al., 2021; Reboredo, 2013; Baek, 2019). Gold contributes as a large portion of commodity reserves of major economies. As of September 2023, the official US gold reserve was 8133.46 tons, approximately 68% of total US reserves [1]. In addition, Silver is not only valued for its ornamental uses but also serves as a crucial component in industrial manufacturing.

[1] https://www.gold.org/goldhub/data/gold-reserves-by-country

On the other hand, Cryptocurrencies have emerged as recent financial assets that enable direct, transparent, and secure electronic payments through blockchain technology. The decentralized and secure nature of cryptocurrencies has attracted investors worldwide but holding these assets for longer periods can expose investors to high risks. Therefore, diversification and hedging strategies beyond cryptocurrency markets are necessary to safeguard investments, particularly in the event of economic uncertainty. Contrary to cryptocurrencies that do not have any tangible values, commodity markets are popular for their intrinsic physical value. With the increasing popularity of digital cryptocurrencies among governments, companies, and individuals, there is a growing interest among traders, investors, and scholars to enhance their understanding of the market dynamics and explore methods for forecasting the future potential returns on investments. There is limited knowledge about the behavior of cryptocurrencies during the financial crisis since these digital currencies were developed after the last global recession in 2008. The most recent global distress has been COVID-19 disease, which was declared a global pandemic on March 11, 2020.

During the COVID-19 pandemic, the extensive shutdown of industrial operations and travel restrictions imposed by the lockdown measures had a profound impact on global crude oil demand, leading to a sharp decline in prices and increased financial market risk (Qin et al., 2020; Le et al., 2021). For the first time in history, the price of West Texas Intermediate (WTI) oil dipped below -$37 on April 20, 2020 due to significant abnormal market pressures, geopolitical tensions, and global concerns about the severity of the COVID-19 pandemic[2] (Le et al., 2021). In this research we explore cryptocurrency markets and compare the important financial aspects of these markets with commodity markets such as crude oil and precious metals. We will consider the effect of COVID-19 on these markets and examine whether their behavior varies before and during the COVID-19 pandemic. Moreover, we explore whether there is any connection between these markets considering the effect of the COVID-19 crisis.

This research is being conducted to explore and predict financial markets using various econometrics methods, and predictive models. The research will conclude in four peer-revied papers as follows.

The first paper explores the relationship between volatility and return of cryptocurrency, crude oil, and precious metal markets before and during the COVID-19 crisis. Volatility, in finance literature, refers to the extent of price or return fluctuations of a financial asset and is measured by the variance of the rate of return (Bhowmik & Wang, 2020). Furthermore, research on stock markets show that trading volume is a significant factor in determining the return of a financial asset. Therefore, we will also examine the relationship between cryptocurrency returns and trading volume, both before and during the COVID-19 pandemic.

[2] https://www.investopedia.com/articles/investing/100615/will-oil-prices-go-2017.asp

The second paper extends the first paper by investigating the connectedness between the ten most traded cryptocurrencies and Gold and crude oil markets in pre-COVID-19 and during the COVID-19 periods. A safe haven asset refers to an investment that is expected to retain or increase its value during times of market turbulence, economic downturns, or uncertainty. These assets are considered safe because they have a history of maintaining or appreciating in value when other investments, such as stocks or riskier bonds, might experience significant declines (Baur et al., 2021). The potential safe haven effect of gold and crude oil for these cryptocurrency markets during financial crisis are studied using statistical tests and econometrics models.

Accurate forecasting of crude oil and precious metal prices is critical for informed economic and financial decisions, risk management, economic planning, and geopolitical considerations. Fluctuations impact energy costs, global trade, consumer budgets, and market stability, making reliable forecasts essential for stakeholders ranging from governments to individual investors. Consequently, researchers have dedicated their efforts to developing and improving models that capture the intrinsic behavior and dynamics of financial market time series. In the third paper a predictive analytics approach is employed to forecast the price of crude oil and precious metal markets. Using historical price data, we utilized sixteen deep learning, machine learning, and ensemble models to predict the price of crude oil, gold, and silver markets. The performance of these models is compared across four input sequence lengths to find which window size provides more useful data for each market.

The last paper approaches the prediction of crude oil and precious metals markets from a classification perspective. The primary concern for investors revolves around the pivotal decision of whether to buy or sell an asset before the next trading period. In this regard, predicting the direction-of-price-movement can be more informative than price level forecasting. Predicting the direction of price movements in these markets has been a prevailing challenge, necessitating innovative approaches for accurate forecasts. Our study addresses this challenge by adapting three innovative spatial-temporal graph neural network models to the unique characteristics of crude oil, gold, and silver markets. Spatial-Temporal Graph Convolutional Networks (ST-GCNs) are an extension of Graph Convolutional Networks (GCNs) tailored to handle data with both spatial and temporal characteristics. They address the need to capture intricate dependencies within datasets structured as graphs, which may evolve over time. To accomplish this prediction, we leverage a rich dataset of historical, economic, financial, and supply-demand features. These models are adaptable to each market because their hyperparameters are tailored to the specific characteristics of each market. Moreover, we show the effectiveness of attention mechanisms in improving the accuracy of models.

A summary of our contributions can be outlined as follows:

Paper 1 makes significant contributions by examining the behavior of cryptocurrencies during the COVID-19 pandemic, a novel global crisis. It analyzes and compares cryptocurrency dynamics

before and during the pandemic, investigating the effect of volatilities on market returns, and compares return-volatility relationships with other markets. Moreover, the relationship between return and trading volume of cryptocurrencies is examined. The study evaluates the ten most traded cryptocurrency markets, ensuring a comprehensive analysis and unbiased estimation. The results provide valuable insights for investors to understand risk-reward dynamics and develop investment strategies during the pandemic.

Paper 2 contributes to the literature by employing various analyses to explore the relationship between cryptocurrency and commodity markets, focusing on Gold and Crude Oil. It examines the hedging and safe-haven roles of these commodities during the COVID-19 pandemic on cryptocurrency dynamics. The comprehensive analysis covers ten cryptocurrency markets, enhancing understanding of interconnectedness and contributing to a more nuanced understanding of safe-haven effects.

Paper 3 addresses the gap in literature on deep learning models for commodity market forecasting. It implements and compares state-of-the-art deep learning models for predicting crude oil, gold, and silver prices. The study is unique in forecasting both gold and silver prices, using advanced models such as TCN, Time2Vector embedding module, and hybrid TCN-BiLSTM and TCN-BiGRU. Each model is meticulously tailored and optimized to provide accurate predictions for its respective market. The results provide valuable insights for players and investors in crude oil and precious metal markets.

Paper 4 introduces novel advancements in spatial-temporal graph neural network models for predicting price movement directions in crude oil and precious metal markets. It customizes Spatial-Temporal Graph Convolutional Network models, examining a comprehensive set of 25 variables. The study leverages the potential of ST-GCNs in financial time series classification tasks, improving predictive performance with attention mechanisms and temporal dilated convolution networks. The research is pioneering in utilizing graph neural networks for price movement prediction and contributes to the understanding of complex market dynamics.

The subsequent chapters of this dissertation unfold as follows: Chapter 2 includes the details of our first paper, while Chapters 3, 4, and 5 thoroughly explain the content and insights derived from Papers 2, 3, and 4, respectively. A conclusive summary of the dissertation is presented in Chapter 6.

Chapter 2

The effect of COVID-19 pandemic on return-volume and return-volatility relationships in cryptocurrency markets

2.1 Introduction and Literature Review

There is limited knowledge about the behavior of cryptocurrencies during the financial crisis since these digital currencies were developed after the last global recession in 2008. The most recent global distress has been COVID-19 disease, which was first detected in Wuhan, China, on December 31, 2019, and was subsequently declared a global pandemic by the World Health Organization (WHO) on March 11, 2020 (Neslihanoglu, 2021). Governments enforced many immediate measures such as quarantines, lockdowns, and social distancing to reduce the number of confirmed and death cases due to the pandemic. COVID-19 outbreak was a severe threat that dramatically affected the world economy as many companies were shut down, sales and productions fell, and unemployment rates surged, leading to downward movement in the majority of the industries.

Considering that the COVID-19 pandemic is an unforeseen crisis, many researchers scrutinized the effect of this pandemic on financial markets' properties and relationships (Chaudhary et al., 2020; Ozili, 2020; Şenol & Zeren, 2020; Sharif et al., 2020). Among these markets, cryptocurrencies are the new digital currencies with many unrecognized characteristics that need to be investigated. For instance, the authors in (Naeem et al., 2021) studied the asymmetric efficiency of four cryptocurrencies and found that significant amounts of market inefficiency can appear in periods of a global health crisis. The implication of COVID-19 confirmed and death cases on cryptocurrency market prices are examined in (Sarkodie et al., 2021), showing that the number of daily COVID-19 confirmed and death cases directly affect these markets' prices.

Volatility is a measure of uncertainty and risk, reflecting the potential for rapid and significant price changes (S. Baek et al., 2020). Volatility is a key factor in investment decision-making, risk assessment, and financial modeling, as it can influence asset pricing, portfolio management, and overall market stability. In finance literature, several studies on the return-volatility relationship have been conducted (Black, 1976; Caporale et al., 2016; Christie, 1982) where evidence of a negative and asymmetric relationship was reported. In this context, volatility is often considered a measure reflecting investor sentiment (Whaley, 2000). A growing body of literature has empirically investigated the volatility of cryptocurrencies. An examination of 45 cryptocurrencies reveals heightened instability and irregularities during the COVID-19 pandemic compared to international stock markets (Lahmiri & Bekiros, 2020). Additionally, findings from an EGARCH model highlight that the leverage effect is significant for Litecoin, Ripple, and Ethereum, but not for Bitcoin (Yousuf Khan et al., 2021). Other research suggests that Bitcoin's volatility experiences notable fluctuations between speculative and stable periods (Kumar & Anandarao, 2019; López-Cabarcos et al., 2021). Furthermore, the impact of news on predicting return volatility in the cryptocurrency market during the COVID-19 pandemic is explored using a GARCH-MIDAS framework, revealing increased risk in the return volatility of digital currencies during this period (Salisu & Ogbonna, 2021).

Besides, understanding the relationship between price and volume of financial markets has been a prominent subject of study among researchers. It provides insights into the structure of markets and is essential for event studies that use a combination of stock returns and trading volume data to make inferences (Karpoff, 1987). Trading volume is linked to investors' attention and reveals how investors react to news about a firm or an asset (Hou et al., 2009). Moreover, trading volume describes investors' learning curve that causes overconfidence and further alters future stock returns (W. Liu, 2006; Statman et al., 2006). The Sequential Arrival of Information (SAI) model (Copeland, 1976) states that information is spread sequentially, and trading volume is a proxy for the information flow rate, implying a positive correlation between volume and the absolute value of price changes which is supported by the mixture of distributions model (Epps & Epps, 1976).

Existing works provide various analyses and findings regarding the dynamic relationship between return and trading volume. For instance, a positive correlation between stock market trading volume and the absolute value of return was found in (Lee & Rui, 2002; Smirlock & Starks, 1988), while it was shown that trading volume does not Granger-cause stock returns (Lee & Rui, 2002). In a more recent study (Behrendt & Schmidt, 2021), the information transfer between stock prices and trading volume is investigated using Shannon transfer entropy which confirms a significant nonlinear information transfer from stock returns to trading volume changes.

Although the causal relations between trading volume and stock returns have been widely investigated in the literature, there is limited empirical research to examine these relationships in the cryptocurrency markets. Most recently, cryptocurrency returns are found to significantly impact the volume changes before the COVID-19 outbreak (Corbet et al., 2021). However, in

(Corbet et al., 2021), only the short-term effect of the outbreak is considered as the sample only covers the data up to May 27, 2020. Likewise, Leirvik, 2021 found a significant but time-varying correlation between cryptocurrency liquidity volatility and returns (Leirvik, 2021).

Despite all efforts by scholars, there is still a lack of knowledge about cryptocurrencies' volatility-return relationship during the COVID-19 pandemic. In this study, we fill this gap in the literature by utilizing an ARMA-EGARCH model to examine the effect of return volatilities on the ten most traded cryptocurrency market returns before and during the COVID-19 outbreak. More particularly, we will test whether there is a difference in cryptocurrencies' returns due to their return volatility in pre-COVID-19 and during COVID-19 periods. Since cryptocurrency markets are highly volatile in essence, this effect will be compared with the return and volatility of crude oil and gold commodities for the same periods. The COVID-19 risk is perceived differently over the short and the long run and may be regarded as an economic crisis in the early stages of its emergence. In this regard, a time frame of one year before the COVID-19 pandemic and one year during this pandemic is considered to capture both short and relatively long-term effects.

In addition, we study the unidirectional and bidirectional Granger causality relationship between the ten most traded cryptocurrency returns and trading volume changes and further test the effect of COVID-19 pandemic on these relations. This analysis aims to explain these relationships in the pre-COVID-19 and during the COVID-19 pandemic periods and provides more comprehensive understanding of movements in the digital currency market. To the best of our knowledge, none of the preceding studies have undertaken a comparative examination of the relationships between returns and volume changes before and during the COVID-19 pandemic. Hence, the current study seeks to remedy this situation.

Given the literature review, our research questions are as follows:

1. Does the volatility of cryptocurrency, crude oil, and gold markets affect their returns?
2. Is the relationship between volatility and return of cryptocurrency, crude oil, and gold markets affected by COVID-19 crisis?
3. Which market is less risky during the COVID-19 pandemic?
4. Do changes in the return of cryptocurrencies cause any changes in their trading volume? How about the reverse causal relationship?

In summary, our study makes the following important contributions to the existing literature on the effect of the COVID-19 pandemic on the dynamics of cryptocurrencies and commodity markets:

1. As the COVID-19 pandemic is the first global crisis after the advent of cryptocurrencies, it is essential to examine the behavior of digital currencies during this distress. This study analyzes and compares cryptocurrency dynamics in the pre-COVID-19 and during COVID-19 pandemic periods, which has received limited attention compared to the conventional financial markets.

2. We examine the effect of return volatilities on the ten most traded cryptocurrency market returns before and during the COVID-19 outbreak. Such investigation would help understand the risk-reward dynamics in cryptocurrency markets before and during the pandemic.
3. We compare the return and volatility relationship of cryptocurrencies with the same relationship in other markets such as Crude oil and Gold for pre-COVID-19 and during-COVID-19 periods. This comparison will give some insights to investors for distinguishing the associated risk with cryptocurrencies and commodity markets.
4. In contrast to the existing literature that limits its analysis to a few cryptocurrency markets, mainly Bitcoin and Ether, our study evaluates the ten most traded cryptocurrency markets to have a better and generalized idea of digital currency markets and investigate crude oil and gold markets.
5. We consider almost the same sample size for both the before and during COVID-19 pandemic periods to ensure that there will not be any estimation bias related to sample size. Moreover, our empirical study will cover a more extended period during the COVID-19 pandemic compared to other similar studies.
6. The results from our study would assist investors in comparing the potential risk of investing in crypto markets with other commodity markets during the COVID-19 pandemic and in developing investment strategies by considering the return-volatility and return-volume relationships.

2.2 Methodology

In this study, return and volume change series are defined as follows:

$$R_t = 100 \times \log(\frac{P_{t+1}}{P_t}) \tag{2-1}$$

$$V_t = \log(\frac{v_{t+1}}{v_t}) \tag{2-2}$$

where P_t and v_t are, respectively, the price and trading volume of the asset at time t. From now on, the *volume change series* will be called *volume series* for simplicity.

2.2.1 Return-Volatility relationships

Many financial markets exhibit asymmetric responses to positive and negative shocks, with negative shocks at time t-1 having a greater effect on variance at time t than positive shocks of equal magnitude (Black, 1976; Christie, 1982). Such an asymmetric relationship can be explained by the leverage effect. The exponential autoregressive conditional heteroscedasticity (EGARCH) model uses logged conditional variances to relax the positiveness restriction of GARCH model coefficients and allows for asymmetric impacts between positive and negative asset returns. The main advantage of the EGARCH model is that it does not require any parameter constraints

because the equation is built on log variance rather than variance itself; hence the nonnegativity is irrelevant. In general, the likelihood maximization with no restrictions provides faster and more reliable optimizations. Furthermore, the EGARCH-M addresses the potential heteroskedasticity problem that would lead to inefficient estimators and possibly incorrect inferences.

To this end, we employ the Exponential Generalized Autoregressive Conditional Heteroskedasticity with Mean (EGARCH-M) model proposed by (Nelson, 1991) to examine the impact of the COVID-19 pandemic on the relationship between returns and volatility. The estimation of mean returns is carried out using the Autoregressive Moving Average (ARMA) model, as outlined by (Box George E.P. et al., 2015). The formulation of the EGARCH(1,1) in mean model is:

$$r_t = c + \sum_p \varphi_p r_{t-p} + \sum_q \theta_q \varepsilon_{t-q} + \lambda\sigma_t^2 + \varepsilon_t \quad , \quad \varepsilon_t = z_t\sigma_t, \ and \tag{2-3}$$

$$\ln(\sigma_t^2) = \omega + \alpha_1\big(|z_{t-1}| - E(|z_{t-1}|)\big) + \alpha_2 z_{t-1} + \beta \ln(\sigma_{t-1}^2) \tag{2-4}$$

In Eq. (2-3), c is the constant intercept, ε_t is the error term, σ_t^2 is the conditional variance, z_t is the standardized shock, and φ_p and θ_q are the parameters of autoregressive and moving average terms, respectively.

The structure of ARMA models for each market are determined according to the Ljung-Box Q-test for autocorrelations (Ljung & Box, 1978) and the Akaike Information Criterion (AIC) (Akaike, 1974). Similarly, EGRACH model structure is selected by using Bayesian Information Criterion (BIC), also known as Schwarz Information Criterion (SIC) (Schwarz, 1978). EGARCH-in-mean parameter (λ) captures the impact of return volatility on cryptocurrency returns. Similarly, the EGARCH model in Eq.(2-4) estimates the current conditional variance by the sum of these linear components:

- Past logged conditional variances (the GARCH component β)
- Magnitudes of past standardized innovations (the ARCH component α_1)
- Past standardized innovations (the leverage component α_2)

The maximum likelihood estimation (MLE) routine (Fisher, 1925) is employed to estimate all parameters of the EGARCH process.

2.2.2 Return-Volume relationships

Causality in time series has been explored through diverse approaches, including Granger causality (Granger, 1969), Lasso-Granger (Hlaváčková-Schindler & Pereverzyev, 2015), causal diagrams (Shojaie & Michailidis, 2010), neural networks (Rosoł et al., 2022), and network centrality methods (Dablander & Hinne, 2019). In this study, linear causality tests are opted for investigating the causal relationship between cryptocurrency return and trading volume for the several reasons. First, both return and trading volume change are stationary series with a linear relationship. Second, the sample size is large enough to allow for robust inferences which makes it irrelevant to

use other causality methods such as Restricted variants of the VAR model (CGCI) (Siggiridou & Kugiumtzis, 2016), which are more suitable for smaller samples. Lastly, given the bivariate causality test is required by the variable setting of this study, Granger causality is considered more appropriate compared to high-dimensional methods such as Lasso-Granger, causal diagrams, or network centrality methods.

In order to find the return-volume change relationships, first a vector autoregression (VAR) model (Sims, 1980a) is created and then a Granger causality test is performed on the estimated coefficients for the VAR model. This model can be expressed as:

$$R_t = a_r + \sum_{i=1}^{p} b_{r,i} R_{t-i} + \sum_{i=1}^{p} c_{r,i} V_{t-i} + u_{r,t} \quad (2\text{-}5)$$

$$V_t = a_v + \sum_{i=1}^{p} b_{v,i} R_{t-i} + \sum_{i=1}^{p} c_{v,i} V_{t-i} + u_{v,t} \quad (2\text{-}6)$$

in which R_t represents returns, V_t denotes volume, $u_{r,t}, u_{v,t}$ are error terms and p denotes the autoregressive lag length. The optimal lag structure in Eq.(2-5) and Eq.(2-6) is chosen according to the corresponding AIC. Recall that Eq.(2-5) and Eq.(2-6) have been estimated by using MLE method.

The variables in a VAR model should be stationary so that the VAR estimates are reliable. For this, the Augmented Dicky Fuller (ADF) test (Said & Dickey, 1984) is performed in which the null hypothesis is that series have a unit root. Rejecting the null hypothesis signifies that series are stationary.

The estimated regression coefficients from Eq.(2-5) and Eq.(2-6) are then used to carry out pairwise Granger causality tests to find whether an endogenous variable can be treated as exogenous. The Wald Chi-Square test for the joint hypothesis of $c_{r,i} = 0, (\mathrm{i} = 1, .., \mathrm{p})$ is used to examine the Granger causality effect from volume to return. Similarly, to test the Granger causality effect from return to volume, Wald Chi-Square test for the joint hypothesis of $b_{v,i} = 0, (\mathrm{i} = 1, .., \mathrm{p})$ is performed. The null hypothesis indicates that there is no Granger causality effect between these market series. A bidirectional Granger causality exists between variables if both null hypotheses are rejected.

To get a common understanding of the aforementioned markets' behavior, descriptive analyses of return, volatility, and volume series for pre-COVID-19 and during COVID-19 periods are performed. Then, results of statistical tests involving simple Pearson correlations between the returns and volume for both periods at the 5% significance level will be presented. Additionally, Granger causality tests are applied to investigate any lead-lag relation between volume and return time-series. The optimal number of lags has been determined by minimizing the AIC. Hence, we test whether the cryptocurrency returns "Granger cause" its trading volume and vice-versa.

2.3 Data and Empirical Results

2.3.1 Data Description

This study seeks to comprehensively explore the interplay between the volatility and returns of cryptocurrencies, along with the relationship between their trading volume and returns. A thorough examination of the literature revealed a predominant focus on a limited set of cryptocurrencies, notably Bitcoin and Ether, given their substantial share of the cryptocurrency market capitalization. However, with the emergence of numerous new cryptocurrencies in recent years, coupled with increased trading activity, there is a growing interest in understanding their dynamics. To address this, we have strategically chosen to analyze the ten most traded cryptocurrencies based on their average trading volume in the fourth quarter of 2020.

This paper studies thirteen markets, including ten cryptocurrencies, Gold, West Texas Intermediate (WTI), and BRENT Crude Oil. The cryptocurrencies studied in this paper are Tether, Bitcoin, Ethereum, Ripple, Litecoin, Bitcoin Cash, EOS, Chainlink, Cardano, and Monero. Daily closing prices of the ten most traded cryptocurrencies (as of Oct-Dec 2020), and daily spot prices of Gold, WTI, and BRENT crude oils are collected from January 01, 2019, to December 31, 2020. The choice of sample size is mainly affected by our objective to compare the volatility and volume dynamics of cryptocurrencies before and during the COVID-19 pandemic. The first reported COVID-19 cases on December 31, 2019, followed by WHO's declaration of a global pandemic in March 2020 prompted emergency measures such as quarantines, lockdowns, and business closures which dramatically affected the world economy. Although COVID-19 pandemic is still ongoing in many countries, the most dramatic financial impacts of this pandemic has occurred in 2020. By this virtue, the entire sample is split into two subsamples: daily trading data from January 01, 2019, to December 31, 2019, as the pre-COVID-19 sample, and from January 01, 2020, to December 31, 2020, as the during COVID-19 sample. We have selected one year of daily trading data for each period to provide a large amount of data for the robust estimation of econometrics models. Our sample selection assures almost equal sample sizes in the two periods, and our comparisons will not be biased by sample size. It is worth mentioning that cryptocurrency prices are available seven days a week, while the WTI, BRENT, and Gold prices are only available five days a week. The sample size and dataset sources are listed in Table 2-1.

Table 2-1: Sample data

Market	2019 sample (prior-Covid19)	2020 sample (during-Covid19)	Data Source
Cryptocurrencies	365	361	Yahoo Finance
WTI and BRENT Crude Oil	248	246	Thomson Reuters from the U.S. Energy Information Administration
Gold	248	246	World Gold Council

Table 2-2: Descriptive statistics of returns

Markets	Mean		Std. Dev.		Minimum		Maximum		Skewness		Kurtosis	
	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19
Tether	-0.0017	0.0001	0.172	0.243	-0.619	-2.283	0.653	2.319	0.18	0.29	1.93	46.01
Bitcoin	0.0777	0.1673	1.532	1.748	-6.593	-20.183	6.951	7.257	0.23	-4.05	4.39	50.29
Ether	-0.0034	0.2086	1.795	2.283	-7.959	-23.918	6.297	7.533	-0.45	-3.17	3.60	34.08
Ripple	-0.0718	0.0157	1.588	2.690	-5.827	-23.908	9.927	14.523	0.51	-1.64	5.82	25.76
Litecoin	0.0363	0.1324	2.067	2.292	-7.830	-19.502	11.671	8.292	0.68	-1.68	4.86	15.32
Bitcoin Cash	0.0361	0.0620	2.262	2.497	-11.993	-24.379	14.881	11.703	0.62	-2.33	9.56	26.53
EOS	0.0006	0.0008	2.178	2.329	-11.670	-21.898	8.052	9.047	-0.27	-2.37	4.28	22.87
Chainlink	0.2154	0.2221	2.804	2.996	-9.396	-26.691	20.873	10.701	1.64	-1.71	9.19	18.15
Cardano	-0.0266	0.2050	2.003	2.585	-9.037	-21.873	7.377	7.986	-0.05	-1.50	2.11	14.46
Monero	-0.0044	0.1507	1.823	2.118	-8.248	-21.465	6.126	6.115	-0.13	-3.00	2.44	29.05
Gold	0.067	0.089	0.723	1.282	-2.048	-5.265	2.746	5.133	0.499	-0.59	4.635	6.474
WTI	0.112	-0.095	2.167	24.27	-8.724	-290.74	14.17	218.71	0.58	-3.99	10.658	111.36
BRENT	0.091	-0.114	2.082	7.311	-6.337	-64.37	11.07	41.202	0.247	-2.16	6.03	32.57

Table 2-3: Descriptive statistics of volume changes.

Markets	Mean		Std. Dev.		Minimum		Maximum		Skewness		Kurtosis	
	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19
Tether	0.0023	0.0011	0.077	0.085	-0.201	-0.313	0.358	0.338	0.59	0.18	1.35	0.76
Bitcoin	0.0018	0.001	0.072	0.086	-0.305	-0.312	0.322	0.353	0.49	0.26	2.34	1.31
Ether	0.0016	0.0005	0.069	0.09	-0.207	-0.34	0.329	0.334	0.61	0.21	1.57	1.53
Ripple	0.001	0.0019	0.119	0.107	-0.852	-0.499	0.507	0.411	-0.09	0.1	8.6	2.74
Litecoin	0.0025	0.0009	0.074	0.084	-0.205	-0.363	0.422	0.304	1.08	0.28	4.57	2.28
Bitcoin Cash	0.0027	0.001	0.105	0.118	-0.226	-0.461	0.549	0.69	1.21	0.76	3.7	4.71
EOS	0.0009	0.0004	0.099	0.11	-0.251	-0.412	0.395	0.542	0.62	0.52	1.36	2.53
Chainlink	0.0044	0.0036	0.183	0.117	-0.488	-0.465	1.063	0.52	1.3	0.37	4.92	1.71
Cardano	0.0003	0.0046	0.138	0.131	-0.404	-0.493	0.42	0.501	0.18	0.26	0.05	0.79
Monero	0.002	0.0035	0.109	0.233	-0.564	-1.999	0.679	2.143	0.93	1.32	9.74	60.27

Table 2-2 presents a summary of descriptive statistics for the return time-series in periods of pre-COVI19 and during COVID-19 pandemic. Before COVID-19, cryptocurrencies average returns ranged from -0.0718 to 0.2154 percent, with Ripple having the least and Chainlink having the most return. While during COVID-19, the average returns ranged from 0.0001 to 0.2221 percent, with Tether having the least and Chainlink having the most return. In both periods, Chainlink has been the most volatile cryptocurrency. Furthermore, the standard deviations of all analyzed markets are higher during the pandemic, indicating that, in general, returns are more volatile during the pandemic, with Chainlink being the most volatile cryptocurrency both before and during COVID-19. Moreover, except for Tether, all market return distributions are negatively skewed throughout the pandemic, indicating that the return series are not normally distributed. When the maximum and minimum returns are compared in both periods, it can be inferred that the return range of all markets is higher during the pandemic compared to the pre-pandemic period. Table 2-1 also reports the kurtosis and skewness of the return distributions in 2019 and 2020. These findings reveal that the return distributions are non-normal in both periods, indicating excess kurtosis and negative skewness before and during the pandemic. During the pandemic, the excess kurtosis and negative skewness of the return distributions are more extreme. This confirms the presence of volatility and GARCH structure for return series in both periods.

The descriptive statistics related to trading volume changes in Table 2-3 show that the average trading volumes for all cryptocurrencies during the pandemic are lower than the pre-pandemic period, while the standard deviation of trading volume is higher during the pandemic. The evidence from Table 2-3 does not confirm any significant skewness in trading volume series in pre-COVID-19 and during COVID-19 periods.

Several statistical tests are conducted to test the equality of returns' means, variances, and distributions before and during the COVID-19 pandemic, and the findings are reported in Table 2-4. It is evident that the mean returns of all markets in the pre-COVID-19 period are not significantly different from the mean returns during the COVID-19 pandemic. However, except for Bitcoin Cash, EOS, and Chainlink, the variances of other markets are significantly different between the two periods. The Kolmogorov-Smirnov test (Massey, 1951) examines whether the return distributions before and during the COVID-19 pandemic are statistically equal and the results demonstrate that only Tether, Ether, Monero, Gold, WTI, and BRENT have statistically different return distributions.

Table 2-4: Reported p-value of statistical tests for return and volatility

Market	Equality of Means*		Kolmogorov-Smirnov test**		Equality of Variances***
	Return	Volatility	Return	Volatility	Return
Tether	0.9097	**0.004**	**0.0039**	**0.000**	**0.0000**
Bitcoin	0.4635	**0.000**	0.2923	**0.000**	**0.0125**
Ether	0.1647	**0.000**	**0.0314**	**0.000**	**0.0000**
Ripple	0.5936	**0.000**	0.0740	**0.000**	**0.0000**
Litecoin	0.5533	**0.000**	0.3848	**0.000**	**0.0492**
Bitcoin Cash	0.8838	**0.000**	0.5546	**0.000**	0.0602
EOS	0.9992	**0.000**	0.8267	**0.000**	0.2012
Chainlink	0.9752	0.066	0.1225	**0.000**	0.2091
Cardano	0.1777	**0.000**	0.1327	**0.000**	**0.0000**
Monero	0.2907	**0.000**	**0.0217**	**0.000**	**0.0044**
Gold	0.8080	**0.000**	**0.0013**	**0.000**	**0.0000**
WTI	0.8935	0.315	**0.0411**	**0.000**	**0.0000**
BRENT	0.6714	**0.000**	**0.0338**	**0.000**	**0.0000**

H0: Average return (volatility) in pre-COVID-19 period is equal to the average return (volatility) during COVID-19. ** H0: Return (volatility) distributions are equal in pre-COVID-19 and during COVID-19 periods. *** H0: Returns' variance in pre-COVID-19 period is equal to the variance during COVID-19. Values in bold show that the null hypothesis is rejected at the 5% significance level.

Table 2-4 summarizes the results of statistical tests comparing market volatilities before and during the COVID-19. These findings reveal that, except for Chainlink and WTI, the average volatility of all other markets differs significantly between the two periods, as does the volatility distribution of all assets. Figure 2-1 depicts the volatility of all thirteen markets before and during the pandemic. As displayed, there is a large jump in volatility across all markets in March 2020 due to the market crash following the COVID-19 pandemic declaration by WHO. Moreover, on April 20, 2020, the May 2020 contract futures price for WTI plunged to around -$37 per barrel, causing even another leap in volatility. The scale of volatility magnitudes in 2020 is considerably larger than in 2019. The impact of these volatilities on market returns will be examined in the following section.

In order to explore potential linear associations between return and volume, we conducted Pearson correlation analyses for each cryptocurrency, and the outcomes are detailed in Table 2-5. Prior to the COVID-19 pandemic, all cryptocurrencies, except for Tether, demonstrated a significant correlation between return and trading volume at the 5% significance level. In contrast, during the pandemic, Tether, EOS, and Monero did not support a significant correlation. Consequently, the likelihood of a causal association between return and trading volume for Tether, EOS, and Monero appears improbable during the pandemic period.

Table 2-5: Pearson correlation between returns and volume changes

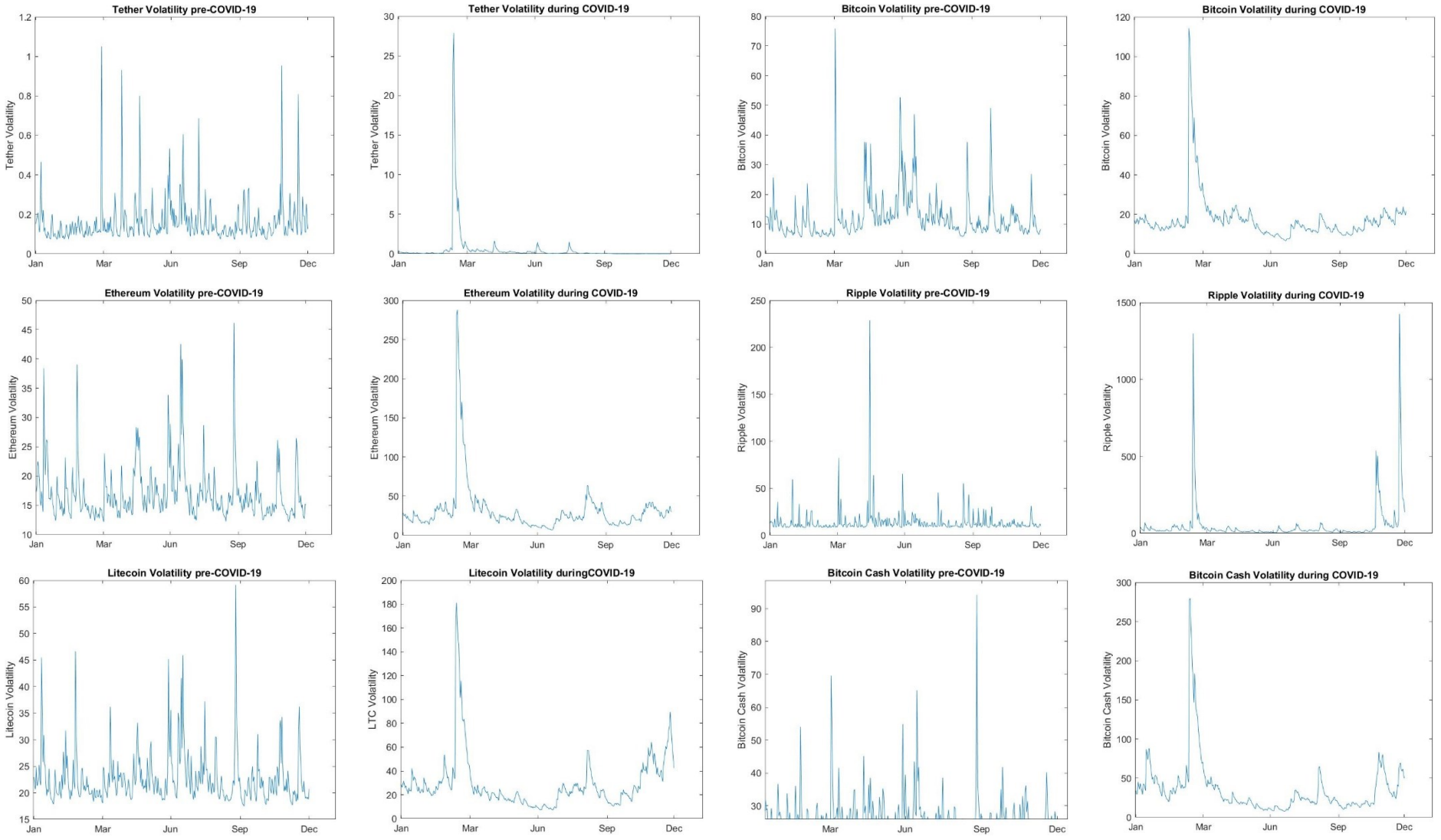
	Pre-COVID-19 (2019) 2019/01/01-2019/12/31		During COVID-19 (2020) 2020/01/01-2020/12/31	
	Correlation	p-value	Correlation	p-value
Tether	0.007	0.8929	0.000	0.9988
Bitcoin	0.212	0.0001	0.141	0.0072
Ether	0.195	0.0002	0.124	0.0180
Ripple	0.164	0.0017	0.135	0.0098
Litecoin	0.360	0.0001	0.198	0.0001
Bitcoin Cash	0.284	0.0001	0.200	0.0001
EOS	0.124	0.0177	0.058	0.2639
Chainlink	0.465	0.0001	0.185	0.0004
Cardano	0.217	0.0001	0.201	0.0001
Monero	0.199	0.0001	0.036	0.4871

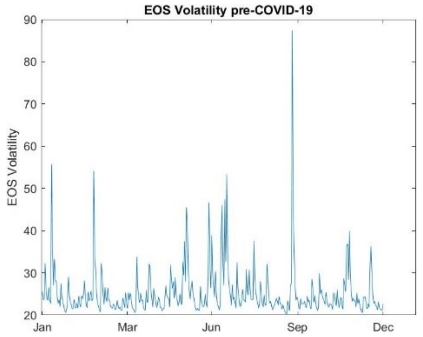
2.3.2 Return and volatility of return relationships (EGARCH-M)

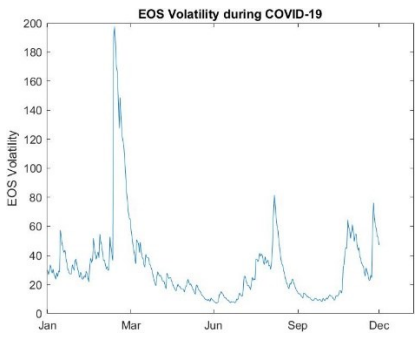
This section presents the outcomes of the Return-Volatility relationship analysis. The stationarity of all return series is scrutinized using the Augmented Dickey-Fuller test, confirming their stationary nature. Considering the kurtosis and skewness of the return distributions in Table 2-2, the possibility of a GARCH effect influencing the volatility of returns is acknowledged. To assess the presence of GARCH effects, the Jarque-Bera normality test (Jarque & Bera, 1980) is conducted on the return series. As detailed in Table 2-6, the GARCH structure is evident across all thirteen markets, both before and during the COVID-19 pandemic.

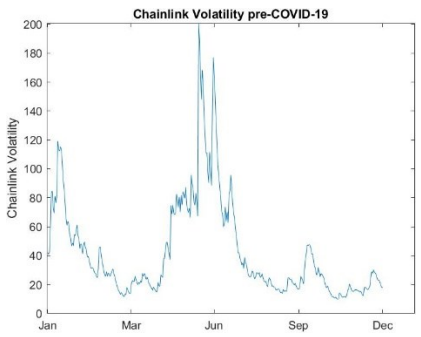
For each sample, the selection of the ARMA model structure in each market is guided by the Ljung-Box Q-test for autocorrelations and the AIC method. Subsequently, an EGARCH-M model is employed on the return series to investigate the effect of the pandemic on the return-volatility relationship.

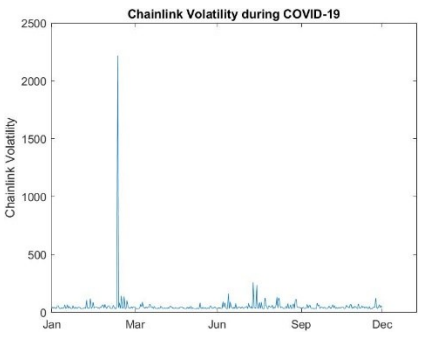
Figure 2-1: Market volatilities in Pre-COVID-19 and during COVID-19 pandemic periods

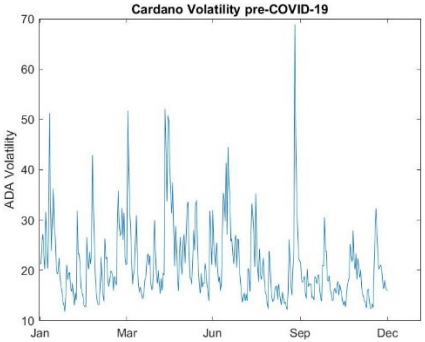


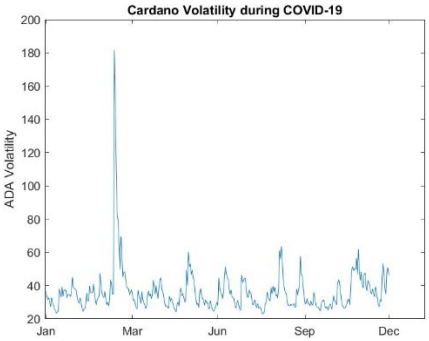

EOS Volatility pre-COVID-19
EOS Volatility
Jan
Mar
Jun
Sep
Dec

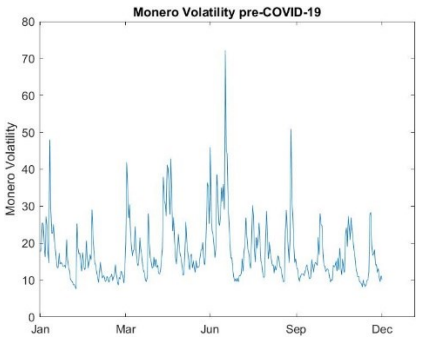

EOS Volatility during COVID-19
EOS Volatility
Jan
Mar
Jun
Sep
Dec

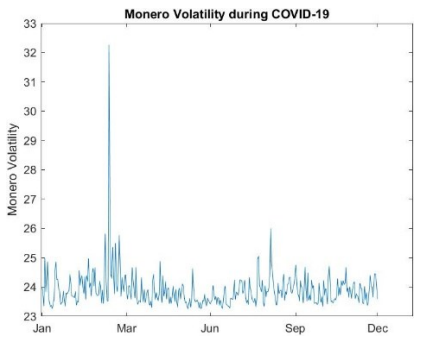

Chainlink Volatility pre-COVID-19
Chainlink Volatility
Jan
Mar
Jun
Sep
Dec

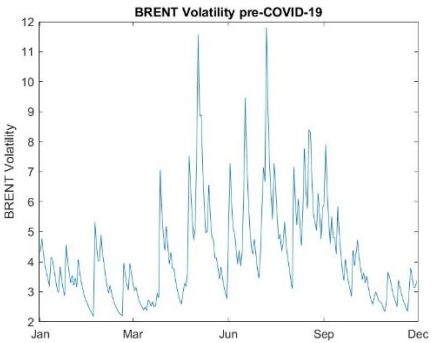

Chainlink Volatility during COVID-19
Chainlink Volatility
Jan
Mar
Jun
Sep
Dec

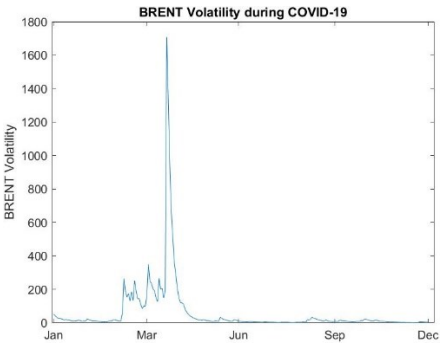

Cardano Volatility pre-COVID-19
ADA Volatility
Jan
Mar
Jun
Sep
Dec

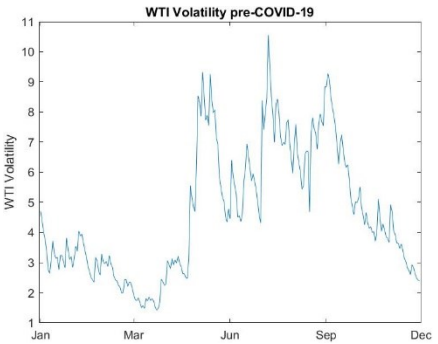

Cardano Volatility during COVID-19
ADA Volatility
Jan
Mar
Jun
Sep
Dec

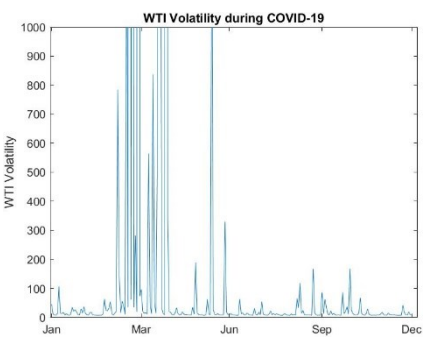

Monero Volatility pre-COVID-19
Monero Volatility
Jan
Mar
Jun
Sep
Dec

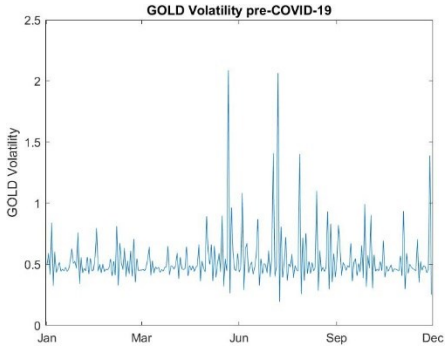

Monero Volatility during COVID-19
Monero Volatility
Jan
Mar
Jun
Sep
Dec


BRENT Volatility pre-COVID-19
BRENT Volatility
Jan
Mar
Jun
Sep
Dec


BRENT Volatility during COVID-19
BRENT Volatility
Jan
Mar
Jun
Sep
Dec


WTI Volatility pre-COVID-19
WTI Volatility
Jan
Mar
Jun
Sep
Dec


WTI Volatility during COVID-19
WTI Volatility
Jan
Mar
Jun
Sep
Dec


GOLD Volatility pre-COVID-19
GOLD Volatility
2.5
2
1.5
1
0.5
0
Jan
Mar
Jun
Sep
Dec

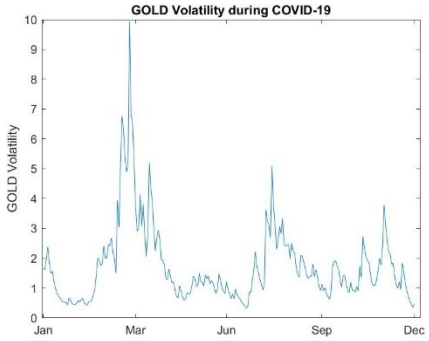

GOLD Volatility during COVID-19
GOLD Volatility
10
9
8
7
6
5
4
3
2
1
0
Jan
Mar
Jun
Sep
Dec

Table 2-6: Statistical tests for return series' stationarity and normality of return distributions

	Pre-COVID-19 (2019)			During COVID-19 (2020)		
	Return*	Volume Changes*	Jarque-Bera**	Return*	Volume Changes*	Jarque-Bera**
Tether	-13.049 (0.0000)	-13.049 (0.0000)	58.882 (0.0000)	-8.725 (0.0000)	-8.725 (0.0000)	31941 (0.0000)
Bitcoin	-19.699 (0.0000)	-7.854 (0.0000)	296.45 (0.0000)	-21.486 (0.0000)	-9.242 (0.0000)	39136 (0.0000)
Ether	-20.232 (0.0000)	-13.304 (0.0000)	208.99 (0.0000)	-8.642 (0.0000)	-15.322 (0.0000)	18120 (0.0000)
Ripple	-15.154 (0.0000)	-13.2957 (0.0000)	531.24 (0.0000)	-12.700 (0.0000)	-7.871 (0.0000)	10174 (0.0000)
Litecoin	-18.459 (0.0000)	-11.362 (0.0000)	386.60 (0.0000)	-21.095 (0.0000)	-17.789 (0.0000)	3708 (0.0000)
Bitcoin Cash	-19.362 (0.0000)	-15.775 (0.0000)	1413 (0.0000)	-9.020 (0.0000)	-11.574 (0.0000)	10947 (0.0000)
EOS	-20.761 (0.0000)	-13.472 (0.0000)	283.51 (0.0000)	-8.952 (0.0000)	-15.001 (0.0000)	8227 (0.0000)
Chainlink	-20.086 (0.0000)	-8.220 (0.0000)	1448 (0.0000)	-20.565 (0.0000)	-11.168 (0.0000)	5145 (0.0000)
Cardano	-20.655 (0.0000)	-14.285 (0.0000)	67.772 (0.0000)	-13.017 (0.0000)	-7.893 (0.0000)	3288 (0.0000)
Monero	-20.535 (0.0000)	-14.054 (0.0000)	91.624 (0.0000)	-7.207 (0.0000)	-10.512 (0.0000)	13274 (0.0000)
Gold	-15.568 (0.0000)		38.047 (0.0000)	-16.249 (0.0000)		138.12 (0.0000)
WTI	-17.027 (0.0000)		622.33 (0.0000)	-22.794 (0.0000)		121006 (0.0000)
BRENT	-16.497 (0.0000)		97.797 (0.0000)	-16.450 (0.0000)		9151 (0.0000)

*Augmented Dicky Fuller Unit Root Test (H0: Series have a unit root; maximum lag =20, Intercept Only). The first value in each cell is the t-statistics and the second value in the parentheses is the associated *p*-value. All Series are stationary as they are significant at the 1% level. ** The null hypothesis is that return distributions are normal.

Table 2-7 presents the magnitude and the direction of the volatility effect on cryptocurrencies, Gold, WTI, and BRENT crude oil returns in both pre-COVID-19 and during COVID-19 periods. The EGARCH-M effects are investigated by three different error distribution assumptions: Normal Distribution, Student's t Distribution, and Generalized Error Distribution (GED).

According to the observed skewness and kurtosis values for return series in Table 2-2, all of our studied markets, excluding Tether, show excess kurtosis before and during the COVID-19 pandemic, with higher values during the pandemic. Besides, throughout the pandemic, these Leptokurtic return distributions are negatively skewed, indicating that the return series are not normally distributed. Additionally, prior to the COVID-19 pandemic, we can observe negative skewness in the return distributions of Ether, EOS, Cardano, and Monero and positive skewness

in Tether, Bitcoin, Ripple, Litecoin, Bitcoin Cash, Chainlink, Gold, WTI, and BRENT return distributions.

This evidence suggests that return series do not exhibit a Gaussian distribution, and therefore the EGARCH model's general assumption of normality does not apply to the return series investigated in this paper. In order to represent leptokurtosis more accurately, student's t-distribution and Generalized Error Distribution (GED) are adopted as heavy tail alternative distributions (H. Chen et al., 2019). The parameters of the GED distribution are estimated by the maximum likelihood estimation method (MLE) (Purczyński & Bednarz-Okrzyńska, 2014).

Our empirical results show that EGARCH-M estimations for cryptocurrencies are more statistically significant and robust with the postulated GED distribution for errors.

Table 2-7: EGARCH in Mean effects with three different residual distribution assumptions

Market	Pre-COVID-19			During COVID-19		
	Normal	*t*-student	GED	Normal	*t*-student	GED
Tether	0.1818 (0.46)	-0.0185 (0.92)	-0.0150 (0.94)	-0.0504 (0.54)	-0.0445* (0.06)	**-0.0792** **(0.016)**
Bitcoin	0.0295 (0.27)	0.0310 (0.25)	0.0102 (0.51)	0.0253 (0.38)	0.0130 (0.45)	-0.0022 (0.88)
Ether	0.0115 (0.84)	0.9684 (0.88)	-0.3721 (0.44)	0.0079 (0.75)	-0.1122 (0.54)	**0.1160** **(0.00)**
Ripple	0.0175 (0.56)	-0.0073 (0.49)	0.0162 (0.20)	-0.0036 (0.68)	-0.0050 (0.44)	**-0.0136** **(0.03)**
Litecoin	-0.0140 (0.83)	0.3650 (0.34)	0.0355 (0.23)	-0.0081 (0.37)	0.0100 (0.47)	0.023 **(0.11)**
Bitcoin Cash	0.0770 (0.47)	0.1809 (0.64)	0.1368* (0.08)	**-0.1047** **(0.015)**	0.0059 (0.61)	**0.036** **(0.020)**
EOS	0.0268 (0.72)	0.0435 (0.69)	0.4931 (0.62)	0.0705 (0.52)	0.0000 (0.96)	**0.0272** **(0.002)**
Chainlink	0.0284 (0.45)	0.0113 (0.45)	0.0094 (0.48)	0.0006 (0.96)	0.0134 (0.46)	-0.7572 **(0.24)**
Cardano	0.0360 (0.37)	0.0508 (0.13)	0.0538 (0.13)	0.0177 (0.54)	0.3054 (0.11)	0.0920 **(0.12)**
Monero	-0.0056 (0.81)	-0.0044 (0.85)	0.0075 (0.75)	0.0196 (0.44)	0.0585 (0.53)	**0.8008** **(0.001)**
GOLD	-0.1281 (0.11)	-0.3429 (0.24)	-0.4773 (0.17)	-0.0412 (0.63)	- 0.0621 (0.36)	-0.0532 **(0.42)**
WTI	**-0.7605** **(0.00)**	**-4.311** **(0.00)**	-1.492 (0.12)	0.0000 (0.99)	0.0019 (0.45)	0.0022 **(0.14)**
BRENT	0.0772 (0.41)	**-0.2990** **(0.0003)**	**-0.3002** **(0.0002)**	-0.0059 (0.14)	0.0072 (0.32)	0.0036 (0.62)

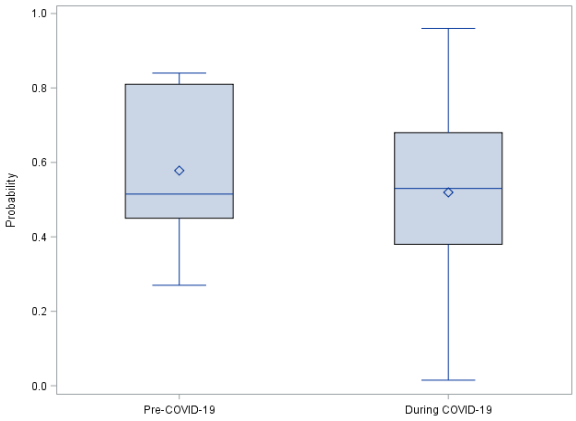
This table presents the value of λ from Eq.(2-3). Values in the parentheses are associated p-values. Significant coefficients at the 5% level are in bold. Values with (*) are significant at the 10% level.

Referring to the results in Table 2-7, there are no significant relationships between volatilities and returns of all cryptocurrencies prior to the COVID-19 pandemic with any of the error distributions considered. During the COVID-19, this relationship is significant for Tether, Ether, Ripple, Bitcoin Cash, EOS, and Monero with GED error distribution. However, there is no significant return-volatility relationship for Bitcoin, Litecoin, Chainlink, and Cardano in this period.

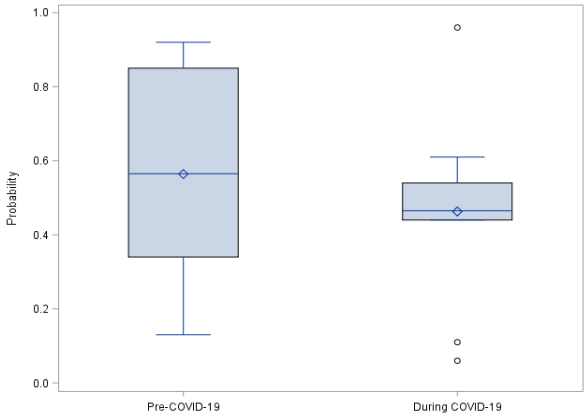
By considering volatility as a proxy for market risk, these results suggest that investors' returns on cryptocurrencies are not significantly affected by the risk of these markets during stable economic conditions. However, due to COVID-related economic hassles, the escalated risks of investing in Tether, Ether, Ripple, Bitcoin Cash, EOS, and Monero have significantly affected the return on these assets. Ceteris paribus, the higher risk of investing in Bitcoin Cash, EOS, and Monero is associated with higher capital returns during financially stable periods, while higher risk of investing in Tether, Ether, and Ripple is resulted in lower average return on these assets. To mitigate the risk of investing in cryptocurrencies during a financial crisis, cryptocurrency market agents can direct their capitals towards safer assets such as Bitcoin, Litecoin, Chainlink, and Cardano.

Regarding the effect of COVID-19 on commodity markets such as Gold, WTI, and BRENT crude oil, it can be concluded that Gold is a less volatile asset than cryptocurrencies, and the effect of Gold volatility on its return is not significant in both pre-CVOID-19 and during the COVID-19 periods. Hence, our results suggest that Gold is resilient to the corresponding risks of financial recession related to the COVID-19 pandemic, and it could be a potential candidate instrument for hedging cryptocurrency portfolios. However, a thorough study of safe-haven or hedging properties of Gold and crude oil for cryptocurrency assets is considered for our future studies.

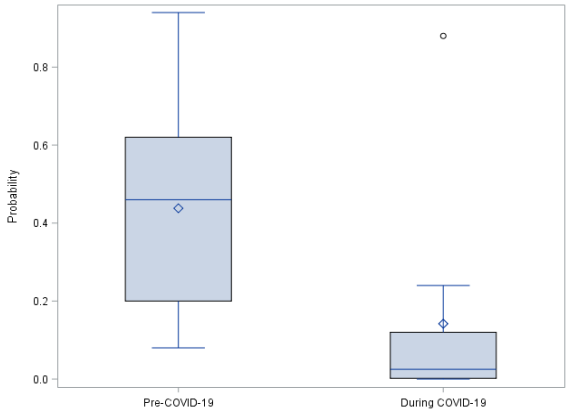
In the pre-COVID-19 period, the return-volatility relationship for WTI and BRENT crude oil appears significant, and market volatilities have decreased crude oil returns. However, from Table 2-7, it can be inferred that the return-volatility relationships for these oil markets are not significant during the COVID-19 pandemic. These results indicate that crude oil markets do not behave in the same way as the majority of the cryptocurrency markets examined here. Furthermore, the effect of risk on crude oil return seems to be only significant during the stable economic conditions compared to the financial distress of COVID-19. Thus, investors in crude oil markets may not expect any impact from COVID-19 pandemic on their long-term returns.



(a) Normal Distribution



(b) Student t Distribution



(c) GED Distribution

Figure 2-2: Distribution of p-values for the significance of EGARCH-M parameter under three different residual distribution assumptions.

Figure 2-2 presents boxplots of the p-values for the significance of EGARCH-M parameter across pre-COVID-19 and during the COVID-19 periods. The analysis suggests that the median p-value of EGARCH-M parameters in cryptocurrency markets differs only under the GED distribution assumption between pre-COVID-19 and during the COVID-19 periods.

2.3.3 Return and volume change relationships by Granger causality tests

This section examines the unidirectional Granger causality from return to trading volume and from trading volume to return for all cryptocurrencies. The ADF unit root test is applied to return and trading volume time series to verify these markets' stationarity before applying the VAR model and Granger causality tests. As presented in Table 2-6, the null hypothesis of having a unit root in the ADF test for all return and trading volume time series is rejected at a 1% significant level, confirming the stationarity of these time series.

Results from Table 2-8 indicate that only returns of Chainlink and Monero Granger cause their volume prior to the COVID-19 pandemic. However, during the COVID-19, there is a

significant Granger causality relationship from the return to the trading volume in Tether, Ether, Ripple, Litecoin, EOS, and Cardano at the 5% level. This analysis could not find any significant causal relationship from return to the trading volume in Bitcoin, Bitcoin Cash, Chainlink, and Monero cryptocurrencies during the COVID-19 pandemic. It can be concluded that fluctuation in the price of these cryptocurrencies is a significant drive for changes in their trading volume. We could not find an overall effect of COVID-19 pandemic on all cryptocurrencies. However, in comparison to the pre-pandemic period, this crisis has forced the return of more cryptocurrencies to have predictive power for their trading volume. Interestingly, neither Bitcoin's return nor its trading volume has any predictive power for one another. Therefore, Bitcoin investors may not rely on trading volume changes to forecast their return in any of the time periods examined in this paper.

Similarly, the Granger causal relationship from trading volume to each cryptocurrency's return is explored. The results confirm that, except for Litecoin, there is no significant evidence of causal relations from trading volume to the return in the pre-COVID-19 period, indicating that volume does not contain predictive power for cryptocurrency return. The Granger causality relations from volume to returns are only present in Tether and Chainlink during the pandemic. Our analyses could not discover any significant unidirectional or bidirectional return-volume relationships for Bitcoin or Bitcoin Cash in the pre-COVID-19 or during the COVID-19 periods. In general, our findings are consistent with the efficient market hypothesis, which argues that returns should not be forecast by publicly available information, like trading volume. Our conclusions regarding cryptocurrencies extend previous research on stock markets, such as studies of Jarque & Bera, 1980; and Lee & Rui, 2002, that state trading volume cannot forecast the return. Our results are consistent with those studies and the efficient markets hypothesis (Fama, 1970), which states that returns should not be predicted by publicly available information, like trading volume.

Table 2-8: Return-volume Granger causality tests (p-values)

	H0: Changes in cryptocurrency price (Return) Granger causes changes in the volume		H0: Changes in the cryptocurrency volume Granger cause changes in the price (return)	
	Pre-COVID-19 (2019)	During COVID-19 (2020)	Pre-COVID-19 (2019)	During COVID-19 (2020)
Tether	0.9478	**0.0005**	0.8425	**0.033**
Bitcoin	0.1816	0.3675	0.3006	0.6024
Ether	0.052*	**0.0083**	0.3021	0.5428
Ripple	0.1246	**0.0272**	0.4859	0.3998
Litecoin	0.2531	**0.0005**	**0.0355**	0.2865
Bitcoin Cash	0.2561	0.4694	0.1154	0.0574*
EOS	0.3051	**0.0033**	0.0654*	0.9452
Chainlink	**0.008**	0.2939	0.2979	**0.0142**
Cardano	0.0935*	**0.0007**	0.4047	0.3489
Monero	**0.0155**	0.6966	0.211	0.9402

Values in bold are significant at the 5% level and values with (*) are significant at the 10% level.

Table 2-9 and Table 2-10 present the estimated parameters and p-values for the VAR models used to carry out the Granger causality tests between cryptocurrencies' return and trading volume in pre-COVID-19 and during COVID-19 periods, respectively.

Table 2-9: VAR model estimations for Return-Volume relationships in the pre-COVID-19 period

Markets	Response	Features									
		R_{t-1}	R_{t-2}	R_{t-3}	R_{t-4}	R_{t-5}	ΔV_{t-1}	ΔV_{t-2}	ΔV_{t-3}	ΔV_{t-4}	ΔV_{t-5}
Tether	R_t	**-0.526 (0.000)**	**-0.398 (0.000)**	**-0.248 (0.000)**	**-0.158 (0.003)**		0.075 (0.492)	0.088 (0.422)	0.001 (0.990)	-0.042 (0.695)	
	ΔV_t	0.005 (0.841)	-0.001 (0.953)	0.005 (0.866)	-0.016 (0.521)		**-0.209 (0.000)**	**-0.259 (0.000)**	**-0.182 (0.000)**	-0.098 (0.063)	
Bitcoin	R_t	-0.028 (0.603)	-0.026 (0.634)	0.035 (0.519)	-0.049 (0.366)		-0.171 (0.888)	1.752 (0.155)	0.337 (0.785)	2.250 (0.064)	
	ΔV_t	0.005 (0.056)	0.004 (0.101)	0.001 (0.586)	0.000 (0.847)		**-0.255 (0.000)**	**-0.276 (0.000)**	**-0.187 (0.000)**	**-0.131 (0.015)**	
Ether	R_t	-0.055 (0.308)	0.019 (0.721)	0.039 (0.468)	-0.049 (0.365)	0.084 (0.110)	-0.775 (0.608)	1.837 (0.328)	0.081 (0.965)	1.985 (0.286)	**-3.337 (0.025)**
	ΔV_t	0.003 (0.122)	**0.005 (0.010)**	0.001 (0.605)	0.001 (0.490)	0.002 (0.213)	**0.719 (0.000)**	-0.089 (0.181)	**0.164 (0.013)**	0.029 (0.659)	**0.130 (0.015)**
Ripple	R_t	-0.039 (0.471)	**-0.116 (0.033)**	-0.018 (0.745)	-0.052 (0.337)	0.077 (0.151)	0.302 (0.694)	0.790 (0.318)	0.715 (0.383)	**1.428 (0.071)**	-0.231 (0.764)
	ΔV_t	**0.009 (0.016)**	0.005 (0.218)	0.005 (0.220)	-0.001 (0.815)	0.004 (0.294)	**-0.313 (0.000)**	**-0.337 (0.000)**	**-0.189 (0.001)**	**-0.178 (0.001)**	**-0.108 (0.048)**
Litecoin	R_t	-0.001 (0.979)	-0.081 (0.159)	-0.032 (0.566)	-0.017 (0.761)		2.126 (0.200)	**3.375 (0.045)**	1.634 (0.332)	**4.693 (0.005)**	
	ΔV_t	**0.004 (0.044)**	0.002 (0.439)	-0.001 (0.518)	0.001 (0.679)		**-0.212 (0.000)**	**-0.274 (0.000)**	**-0.126 (0.031)**	**-0.120 (0.036)**	
Bitcoin Cash	R_t	-0.005 (0.927)	-0.016 (0.766)	0.079 (0.151)			-0.297 (0.808)	1.404 (0.245)	-1.743 (0.154)		
	ΔV_t	**0.006 (0.024)**	0.000 (0.953)	0.002 (0.343)			**-0.211 (0.000)**	**-0.228 (0.000)**	**-0.242 (0.000)**		
EOS	R_t	-0.065 (0.223)	0.036 (0.503)	0.024 (0.653)	-0.042 (0.427)		0.293 (0.809)	1.564 (0.202)	-1.676 (0.172)	**2.438 (0.045)**	
	ΔV_t	0.003 (0.219)	0.002 (0.277)	0.003 (0.226)	0.002 (0.299)		**-0.244 (0.000)**	**-0.283 (0.000)**	**-0.203 (0.000)**	**-0.145 (0.007)**	
Chainlink	R_t	**-0.132 (0.030)**	-0.054 (0.386)	0.088 (0.150)	0.048 (0.425)	0.109 (0.067)	1.932 (0.067)	0.865 (0.394)	1.035 (0.309)	0.259 (0.790)	-0.973 (0.305)
	ΔV_t	**0.008 (0.023)**	-0.002 (0.540)	**0.010 (0.006)**	-0.001 (0.740)	0.000 (0.908)	**-0.213 (0.000)**	**-0.280 (0.000)**	**-0.189 (0.002)**	-0.111 (0.062)	**-0.155 (0.007)**
Cardano	R_t	-0.098 (0.069)	-0.038 (0.485)				0.671 (0.400)	0.946 (0.234)			
	ΔV_t	**0.008 (0.035)**	0.003 (0.454)				**-0.224 (0.000)**	**-0.219 (0.000)**			
Monero	R_t	-0.060 (0.270)	-0.024 (0.653)	0.027 (0.621)	0.019 (0.724)		-1.535 (0.117)	-0.385 (0.694)	-1.317 (0.174)	0.794 (0.408)	
	ΔV_t	0.003 (0.282)	-0.003 (0.278)	**0.009 (0.002)**	0.000 (0.855)		**-0.279 (0.000)**	**-0.231 (0.000)**	**-0.268 (0.000)**	**-0.144 (0.006)**	

Values in the parentheses are associated probabilities. Significant coefficients at the 0.05 level are in **bold**.

Table 2-10: VAR model estimations for Return-Volume relationships during the COVID-19 period

Markets	Response	Features													
		R_{t-1}	R_{t-2}	R_{t-3}	R_{t-4}	R_{t-5}	R_{t-6}	R_{t-7}	ΔV_{t-1}	ΔV_{t-2}	ΔV_{t-3}	ΔV_{t-4}	ΔV_{t-5}	ΔV_{t-6}	ΔV_{t-7}
Tether	R_t	**-0.658 (0.000)**	**-0.446 (0.000)**	**-0.261 (0.000)**	-0.028 (0.833)				-0.028 (0.833)	0.166 (0.212)	**0.350 (0.008)**				
	ΔV_t	**0.041 (0.039)**	-0.034 (0.122)	**-0.048 (0.016)**					**-0.279 (0.000)**	**-0.300 (0.000)**	**-0.160 (0.002)**				
Bitcoin	R_t	**-0.106 (0.048)**	0.076 (0.162)	-0.036 (0.510)	**0.140 (0.010)**	0.055 (0.313)	-0.003 (0.961)	**-0.116 (0.032)**	0.354 (0.768)	-1.007 (0.429)	-1.269 (0.347)	-1.798 (0.197)	-1.198 (0.377)	-0.690 (0.590)	1.298 (0.283)
	ΔV_t	0.003 (0.179)	**0.005 (0.026)**	0.001 (0.777)	0.000 (0.959)	0.000 (0.967)	**0.002 (0.449)**	0.001 (0.725)	**-0.350 (0.000)**	**-0.422 (0.000)**	**-0.314 (0.000)**	**-0.147 (0.019)**	**-0.198 (0.001)**	-0.080 (0.164)	0.096 (0.077)
Ether	R_t	**-0.110 (0.040)**	0.053 (0.331)	-0.063 (0.248)	**0.128 (0.017)**				2.166 (0.137)	0.934 (0.539)	-0.365 (0.810)	-0.098 (0.946)			
	ΔV_t	**0.004 (0.036)**	0.006 (0.001)	0.002 (0.223)	0.000 (0.809)				**-0.313 (0.000)**	**-0.379 (0.000)**	**-0.174 (0.002)**	-0.059 (0.274)			
Ripple	R_t	-0.078 (0.145)	0.069 (0.197)	0.004 (0.937)					1.154 (0.413)	1.914 (0.170)	2.528 (0.076)				
	ΔV_t	**0.005 (0.013)**	0.004 (0.070)	**0.004 (0.027)**					**-0.240 (0.000)**	**-0.309 (0.000)**	**-0.164 (0.002)**				
Litecoin	R_t	**-0.115 (0.033)**	0.058 (0.278)						-1.920 (0.200)	1.786 (0.235)					
	ΔV_t	0.003 (0.065)	**0.007 (0.000)**						**-0.271 (0.000)**	**-0.240 (0.000)**					
Bitcoin Cash	R_t	**-0.128 (0.019)**	0.019 (0.728)	-0.064 (0.243)	0.102 (0.065)	-0.029 (0.594)			-0.115 (0.923)	-0.234 (0.848)	2.241 (0.074)	-0.051 (0.966)	-0.311 (0.794)		
	ΔV_t	**0.005 (0.039)**	**0.006 (0.013)**	0.003 (0.220)	0.000 (0.896)	-0.002 (0.347)			**-0.246 (0.000)**	**-0.319 (0.000)**	**-0.182 (0.000)**	-0.096 (0.085)	**-0.133 (0.015)**		
EOS	R_t	**-0.1634 (0.002)**	0.036 (0.502)	-0.032 (0.558)	**0.110 (0.041)**				-0.565 (0.639)	0.110 (0.929)	0.571 (0.643)	-0.177 (0.882)			
	ΔV_t	0.003 (0.178)	**0.007 (0.002)**	**0.006 (0.011)**	-0.001 (0.585)				**-0.288 (0.000)**	**-0.331 (0.000)**	-0.097 (0.075)	0.024 (0.650)			
Chainlink	R_t	**-0.127 (0.021)**	0.019 (0.737)	0.013 (0.811)	**0.137 (0.013)**	0.045 (0.402)			**4.849 (0.001)**	1.941 (0.198)	0.529 (0.737)	0.127 (0.933)	-1.857 (0.206)		
	ΔV_t	**0.005 (0.030)**	0.003 (0.200)	0.001 (0.645)	0.000 (0.917)	0.001 (0.769)			**-0.209 (0.000)**	**-0.337 (0.000)**	**-0.145 (0.015)**	-0.099 (0.082)	**-0.111 (0.046)**		

Markets	Response	Features													
		R_{t-1}	R_{t-2}	R_{t-3}	R_{t-4}	R_{t-5}	R_{t-6}	R_{t-7}	ΔV_{t-1}	ΔV_{t-2}	ΔV_{t-3}	ΔV_{t-4}	ΔV_{t-5}	ΔV_{t-6}	ΔV_{t-7}
Cardano	R_t	**-0.136**	0.063	-0.033					1.906	1.097	0.665				
		(0.013)	(0.255)	(0.550)					(0.086)	(0.328)	(0.548)				
	ΔV_t	**0.008**	**0.009**	0.004					**-0.268**	**-0.258**	**-0.155**				
		(0.004)	**(0.001)**	(0.147)					**(0.000)**	**(0.000)**	**(0.004)**				
Monero	R_t	**-0.168**	0.054	-0.026	**0.160**	0.082	-0.040		0.291	0.190	0.048	-0.035	-0.651	-0.292	
		(0.002)	(0.324)	(0.625)	**(0.003)**	(0.134)	(0.468)		(0.619)	(0.783)	(0.947)	(0.961)	(0.344)	(0.617)	
	ΔV_t	0.002	0.007	0.000	0.001	-0.006	0.000		**-0.628**	**-0.397**	-0.021	**-0.150**	-0.116	**-0.116**	
		(0.621)	(0.180)	(0.977)	(0.912)	(0.199)	(0.956)		**(0.000)**	**(0.000)**	(0.746)	**(0.022)**	(0.064)	**(0.029)**	

Values in the parentheses are associated probabilities. Significant coefficients at the 0.05 level are in **bold**.

Another important motivation of this research is to investigate the causality effect between absolute returns and trading volume changes. As presented in Table 2-11, with the exception of Monero and Tether, absolute returns have a significant causal impact on the volume of all cryptocurrencies both before and during the COVID-19 pandemic periods. The significant $\chi 2$ statistic indicates that, except for Monero, absolute cryptocurrency return provides useful predictive information for trading volume in both periods. Conversely, a significant predictive power of trading volume for absolute returns could not be found for most of the cryptocurrencies in this research. However, Bitcoin's volume Granger causes its absolute return prior to the COVID-19 period (p-value is 0.0228), while Litecoin shows a similar effect during the COVID-19 period (p-value is 0.0141). These results comply with the sequential arrival of information theory, confirming that as the price values change more extremely, more investors will buy or sell their cryptocurrency assets.

Table 2-11: Absolute return-volume Granger causality tests (p-values)

	H0: Absolute changes in cryptocurrency price (absolute return) Granger cause changes in the volume		H0: Changes in the cryptocurrency volume Granger cause absolute changes in the price (absolute return)	
	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19
Tether	0.2068	**0.0044**	0.2262	0.1372
Bitcoin	**0.0491**	**0.0086**	**0.0228**	0.4515
Ether	**0.0013**	**0.0038**	0.1625	0.0563*
Ripple	**0.0004**	**0.0000**	0.6997	0.7871
Litecoin	**0.0033**	**0.0006**	0.7548	**0.0141**
Bitcoin Cash	**0.0008**	**0.0274**	0.1916	0.5929
EOS	**0.0002**	**0.0001**	0.2654	0.0868*
Chainlink	**0.0223**	**0.0000**	0.2299	0.1365
Cardano	**0.0202**	**0.0001**	0.2024	0.3196
Monero	0.6797	0.5347	0.3831	0.5219

Values in bold are significant at the 5% level of significance and values with (*) are significant at the 10% level of significance.

Table 2-12 and Table 2-13 present the estimated parameters and p-values for the VAR models used to carry out the Granger causality tests between cryptocurrencies' absolute return and trading volume in pre-COVID-19 and during COVID-19 periods, respectively.

Table 2-12: VAR model estimations for absolute Return-Volume relationships in the pre-COVID-19 period

Markets	Response	Features									
		$\lvert R_{t-1}\rvert$	$\lvert R_{t-2}\rvert$	$\lvert R_{t-3}\rvert$	$\lvert R_{t-4}\rvert$	$\lvert R_{t-5}\rvert$	ΔV_{t-1}	ΔV_{t-2}	ΔV_{t-3}	ΔV_{t-4}	ΔV_{t-5}
Tether	$\lvert R_t\rvert$	**0.247** **(0.000)**	0.008 (0.885)	-0.047 (0.389)	**0.134** **(0.012)**		0.076 (0.360)	**0.166** **(0.046)**	0.030 (0.721)	-0.051 (0.533)	
	ΔV_t	-0.037 (0.285)	-0.045 (0.953)	-0.035 (0.333)	0.0316 (0.363)		**-0.212** **(0.000)**	**-0.247** **(0.000)**	**-0.159** **(0.003)**	-0.086 (0.103)	
Bitcoin	$\lvert R_t\rvert$	**0.139** **(0.036)**	**0.204** **(0.000)**	**0.147** **(0.021)**	0.069 (0.295)		0.065 (0.951)	-1.376 (0.199)	**-3.01** **(0.002)**	-1.480 (0.118)	
	ΔV_t	0.006 (0.135)	-0.006 (0.139)	**-0.009** **(0.029)**	0.001 (0.756)		**-0.291** **(0.000)**	**-0.224** **(0.001)**	-0.101 (0.100)	-0.077 (0.189)	
Ether	$\lvert R_t\rvert$	0.061 (0.344)	0.118 (0.068)	**0.158** **(0.015)**	0.017 (0.790)	-0.067 (0.258)	0.446 (0.737)	-2.807 (0.080)	-0.181 (0.909)	1.486 (0.347)	1.324 (0.278)
	ΔV_t	0.004 (0.193)	-0.006 (0.054)	-0.004 (0.163)	-0.002 (0.491)	**-0.006** **(0.031)**	**0.679** **(0.000)**	0.007 (0.923)	0.139 (0.069)	0.044 (0.565)	0.102 (0.084)
Ripple	$\lvert R_t\rvert$	**0.220** **(0.001)**	0.011 (0.872)	**0.163** **(0.014)**	0.012 (0.860)	0.002 (0.972)	0.774 (0.266)	-0.029 (0.969)	0.132 (0.856)	-0.383 (0.562)	0.447 (0.440)
	ΔV_t	0.010 (0.114)	-0.011 (0.087)	0.005 (0.437)	**-0.015** **(0.021)**	-0.011 (0.073)	**-0.386** **(0.000)**	**-0.337** **(0.000)**	**-0.177** **(0.011)**	-0.108 (0.090)	0.005 (0.932)
Litecoin	$\lvert R_t\rvert$	0.098 (0.158)	0.063 (0.355)	0.101 (0.142)	0.012 (0.859)		0.495 (0.738)	-0.949 (0.518)	-1.597 (0.245)	-0.061 (0.963)	
	ΔV_t	0.002 (0.641)	**-0.007** **(0.027)**	**-0.006** **(0.049)**	-0.005 (0.123)		**-0.221** **(0.001)**	**-0.192** **(0.005)**	-0.043 (0.507)	-0.021 (0.736)	
Bitcoin Cash	$\lvert R_t\rvert$	0.138 (0.051)	0.042 (0.543)	0.064 (0.362)			0.338 (0.782)	-1.570 (0.181)	-2.024 (0.072)		
	ΔV_t	**0.009** **(0.018)**	**-0.011** **(0.004)**	-0.008 (0.052)			**-0.290** **(0.000)**	**-0.131** **(0.049)**	-0.115 (0.072)		
EOS	$\lvert R_t\rvert$	-0.038 (0.581)	0.008 (0.902)	0.119 (0.080)	0.125 (0.075)		**2.574** **(0.035)**	0.242 (0.844)	-0.558 (0.621)	-0.472 (0.653)	
	ΔV_t	-0.003 (0.430)	**-0.013** **(0.001)**	-0.008 (0.052)	-0.005 (0.258)		**-0.244** **(0.000)**	**-0.169** **(0.015)**	-0.089 (0.164)	-0.063 (0.290)	
Chainlink	$\lvert R_t\rvert$	0.086 (0.178)	0.104 (0.097)	0.052 (0.421)	**0.141** **(0.025)**	0.095 (0.130)	0.298 (0.688)	-0.858 (0.254)	0.123 (0.867)	0.089 (0.896)	**-1.423** **(0.033)**
	ΔV_t	**0.013** **(0.021)**	**-0.013** **(0.016)**	0.000 (0.934)	0.002 (0.695)	-0.008 (0.126)	**-0.238** **(0.000)**	**-0.211** **(0.001)**	-0.112 (0.076)	-0.104 (0.074)	-0.101 (0.076)
Cardano	$\lvert R_t\rvert$	**0.151** **(0.013)**	0.079 (0.189)				0.344 (0.569)	-0.929 (0.108)			
	ΔV_t	-0.006 (0.286)	**-0.013** **(0.023)**				**-0.179** **(0.003)**	**-0.137** **(0.016)**			
Monero	$\lvert R_t\rvert$	**0.136** **(0.010)**	0.094 (0.08)	0.057 (0.291)	**0.137** **(0.01)**		0.264 (0.684)	-0.302 (0.647)	-1.164 (0.074)	0.108 (0.866)	
	ΔV_t	0.001 (0.816)	-0.004 (0.398)	0.004 (0.329)	-0.004 (0.348)		**-0.281** **(0.000)**	**-0.237** **(0.000)**	**-0.239** **(0.000)**	**-0.154** **(0.003)**	

Values in the parentheses are associated probabilities. Significant coefficients at the 0.05 level are in **bold**.

Table 2-13: VAR model estimations for absolute Return-Volume relationships during the COVID-19 period

Markets	Response	Features													
		$\lvert R_{t-1}\rvert$	$\lvert R_{t-2}\rvert$	$\lvert R_{t-3}\rvert$	$\lvert R_{t-4}\rvert$	$\lvert R_{t-5}\rvert$	$\lvert R_{t-6}\rvert$	$\lvert R_{t-7}\rvert$	ΔV_{t-1}	ΔV_{t-2}	ΔV_{t-3}	ΔV_{t-4}	ΔV_{t-5}	ΔV_{t-6}	ΔV_{t-7}
Tether	$\lvert R_t\rvert$	**0.478** **(0.000)**	**-0.129** **(0.028)**	**0.11** **(0.041)**	0.015 (0.906)				**-0.287** **(0.026)**	-0.024 (0.853)	**0.062** **(0.000)**				
	ΔV_t	0.014 (0.521)	-0.081 (**0.000**)	**0.051** **(0.02)**					**-0.277** **(0.000)**	**-0.297** **(0.000)**	**-0.138** **(0.008)**				
Bitcoin	$\lvert R_t\rvert$	0.112 (0.057)	0.031 (0.604)	-0.049 (0.408)	**0.163** **(0.006)**	0.073 (0.222)	-0.021 (0.728)	**0.147** **(0.014)**	-0.634 (0.555)	-1.031 (0.367)	0.990 (0.399)	-0.391 (0.745)	-1.736 (0.136)	0.503 (0.639)	0.414 (0.681)
	ΔV_t	**0.007** **(0.031)**	**-0.01** **(0.002)**	-0.006 (0.082)	0.004 (0.268)	-0.002 (0.507)	-0.003 (0.3)	0.001 (0.741)	**-0.382** **(0.000)**	**-0.325** **(0.000)**	**-0.256** **(0.000)**	**-0.158** **(0.016)**	**-0.162** **(0.011)**	-0.026 (0.651)	0.084 (0.129)
Ether	$\lvert R_t\rvert$	**0.158** **(0.006)**	0.041 (0.474)	-0.081 (0.169)	**0.217** **(0.000)**				-1.490 (0.223)	-1.49 (0.236)	**2.495** **(0.037)**	0.04 (0.972)			
	ΔV_t	0.003 (0.328)	**-0.01** **(0.000)**	-0.003 (0.299)	0.000 (0.87)				**-0.304** **(0.000)**	**-0.254** **(0.000)**	-0.096 (0.096)	-0.034 (0.545)			
Ripple	$\lvert R_t\rvert$	**0.303** **(0.000)**	**0.152** **(0.009)**	0.046 (0.431)					-1.217 (0.315)	-0.298 (0.795)	-0.471 (0.68)				
	ΔV_t	**0.011** **(0.000)**	**-0.008** **(0.003)**	**-0.007** **(0.013)**					**-0.297** **(0.000)**	**-0.234** **(0.000)**	-0.099 (0.068)				
Litecoin	$\lvert R_t\rvert$	**0.252** **(0.000)**	0.068 (0.243)						**-2.756** **(0.021)**	**-2.513** **(0.034)**					
	ΔV_t	0.005 (0.099)	**-0.01** **(0.000)**						**-0.276** **(0.000)**	**-0.111** **(0.047)**					
Bitcoin Cash	$\lvert R_t\rvert$	**0.15** **(0.037)**	0.071 (0.325)	0.108 (0.136)	0.063 (0.386)	**-0.197** **(0.006)**			0.041 (0.974)	-2.245 (0.088)	**-3.28** **(0.011)**	-1.643 (0.164)	1.809 (0.111)		
	ΔV_t	**0.011** **(0.008)**	**-0.009** **(0.021)**	-0.006 (0.177)	-0.001 (0.759)	**-0.013** **(0.002)**			**-0.308** **(0.000)**	**-0.175** **(0.019)**	**-0.18** **(0.013)**	0.015 (0.822)	0.124 (0.054)		
EOS	$\lvert R_t\rvert$	**0.167** **(0.008)**	0.026 (0.676)	-0.067 (0.296)	**0.169** **(0.008)**				-0.453 (0.695)	-0.542 (0.649)	**2.654** **(0.016)**	1.116 (0.283)			
	ΔV_t	0.006 (0.074)	**-0.012** **(0.001)**	**-0.009** **(0.007)**	0.000 (0.942)				**-0.342** **(0.000)**	**-0.223** **(0.001)**	0.014 (0.817)	0.048 (0.409)			
Chainlink	$\lvert R_t\rvert$	**0.18** **(0.002)**	0.024 (0.675)	0.032 (0.594)	0.046 (0.438)	-0.006 (0.925)			-1.845 (0.118)	-0.856 (0.473)	-0.251 (0.835)	2.157 (0.055)	-1.146 (0.308)		
	ΔV_t	**0.008** **(0.004)**	**-0.013** **(0.000)**	-0.002 (0.519)	0.003 (0.312)	-0.004 (0.159)			**-0.22** **(0.000)**	**-0.19** **(0.001)**	-0.115 (0.051)	-0.093 (0.089)	-0.087 (0.112)		
Cardano	$\lvert R_t\rvert$	**0.166** **(0.005)**	0.036 (0.537)	-0.005 (0.932)					-1.4 (0.118)	-0.275 (0.752)	0.768 (0.365)				
	ΔV_t	0.004 (0.362)	**-0.015** **(0.000)**	-0.006 (0.147)					**-0.247** **(0.000)**	-0.106 (0.064)	-0.07 (0.21)				

Markets	Response	Features													
		$\lvert R_{t-1}\rvert$	$\lvert R_{t-2}\rvert$	$\lvert R_{t-3}\rvert$	$\lvert R_{t-4}\rvert$	$\lvert R_{t-5}\rvert$	$\lvert R_{t-6}\rvert$	$\lvert R_{t-7}\rvert$	ΔV_{t-1}	ΔV_{t-2}	ΔV_{t-3}	ΔV_{t-4}	ΔV_{t-5}	ΔV_{t-6}	ΔV_{t-7}
Monero	$\lvert R_t\rvert$	**0.112**	-0.059	0.05	**0.216**	0.01	0.001		0.06	0.053	0.371	0.375	-0.215	0.455	
		(0.039)	(0.278)	(0.354)	**(0.000)**	(0.855)	(0.988)		(0.888)	(0.917)	(0.481)	(0.475)	(0.668)	(0.286)	
	ΔV_t	0.003	-0.008	-0.01	0.000	0.005	0.003		**-0.634**	**-0.38**	0.001	**-0.138**	-0.121	**-0.115**	
		(0.622)	(0.217)	(0.148)	(0.96)	(0.441)	(0.679)		**(0.000)**	**(0.000)**	(0.991)	**(0.036)**	(0.054)	**(0.031)**	

Values in the parentheses are associated probabilities. Significant coefficients at the 0.05 level are in **bold**.

To further investigate the general return-volume relationships for the cryptocurrency market, Student t-tests are applied to the sample of resulting p-values from Granger causality tests. Table 2-14 provides the associated probabilities for testing the significance of bidirectional and unidirectional relationships between cryptocurrency returns and their trading volume changes for both pre-COVID-19 and during COVID-19 periods.

Table 2-14: Statistical *t*-tests for the causality between cryptocurrency returns (or absolute returns) and changes in volume

Pre-COVID-19 Period	***p*-value**
H0: The mean of probabilities for bidirectional Granger causality tests between return and change in volume = 0	0.0002
H0: The mean of probabilities for unidirectional Granger causality tests from returns to changes in volume = 0	0.0301*
H0: The mean of probabilities for unidirectional Granger causality tests from changes in volume to returns = 0	0.0028
H0: The mean of probabilities for bidirectional Granger causality tests between absolute return and volume changes = 0	0.0013
H0: The mean of probabilities for unidirectional Granger causality tests from absolute returns to volume changes= 0	0.1794
H0: The mean of probabilities for unidirectional Granger causality tests from volume changes to absolute returns = 0	0.0023
During the COVID-19 Period	***p*-value**
H0: The mean of probabilities for bidirectional Granger causality tests between returns and volume changes = 0	0.0004
H0: The mean of probabilities for unidirectional Granger causality tests from returns to volume changes = 0	0.0448*
H0: The mean of probabilities for unidirectional Granger causality tests from volume changes to returns = 0	0.0040
H0: The mean of probabilities for bidirectional Granger causality tests between absolute returns and volume changes = 0	0.0040
H0: The mean of probabilities for unidirectional Granger causality tests from absolute returns to volume changes = 0	0.3028
H0: The mean of probabilities for unidirectional Granger causality tests from volume changes to absolute returns = 0	0.0049

*The null hypothesis is not rejected at the 1% significant level. The null hypothesis is not rejected at the 5% significant level for values in bold.

As indicated in Table 2-14, the average p-value for bidirectional Granger causality tests between returns (or absolute returns) and volume changes of all ten cryptocurrencies significantly differs from zero at the 1% level. Hence, the bidirectional return (or absolute return)-volume relationship is not supported in cryptocurrency markets for the pre-COVID-19 and during COVID-19 periods. Similarly, the results of these tests do not confirm any Granger causality effect from cryptocurrencies' trading volume to their returns (or absolute returns). However, in both pre-COVID19 and during COVID19 periods, the average p-value for the unidirectional Granger causality effect from returns to trading volume changes of all ten cryptocurrencies is not

significantly different from zero at the 1% level. Thus, cryptocurrency returns show a predictive power for trading volume. Besides, the average p-value for Granger causality tests from absolute returns to volume changes of all ten cryptocurrencies is not significant at the 5% level. This denotes the existence of a significant causality relation from cryptocurrency absolute returns to volume changes for both periods. Therefore, a general significant predictive power of cryptocurrency returns for trading volume changes is confirmed by this study.

2.4 Conclusion

Causal relationships between returns and the trading volume of financial markets have been a popular subject among researchers for a long time. Moreover, since investors need to evaluate the associated risk with any investment strategy before making decisions, it is crucial to understand the effect of price volatility on financial asset returns. These relationships have not been adequately investigated for the recently emerged cryptocurrency markets. In this study, the effect of the COVID-19 pandemic on the return-volatility and return-volume relationships for the ten most traded cryptocurrencies, namely Tether, Bitcoin, Ether, Ripple, Litecoin, Bitcoin Cash, EOS, Chainlink, Cardano, and Monero, is investigated.

This effect is also investigated for Gold, WTI, and BRENT crude oil markets to compare the behaviour of cryptocurrencies with less volatile markets such as commodity markets. Evidence from the EGARCH-M model suggests that the return-volatility relationships for Tether, Ether, Ripple, Bitcoin Cash, EOS, and Monero are significant during the COVID-19 pandemic. Although, the same relationship is not significant, for any of the studied cryptocurrencies, in the pre-pandemic period. Moreover, it is concluded that the COVID-19 pandemic does not play an essential role in the relationship between returns and volatilities of GOLD, WTI, and BRENT crude oil markets.

Our findings of the return-volume relationship support the availability of causal relations from the return to trading volume for Chainlink and Monero in the pre-COVID-19 period and for Ether, Ripple, Litecoin, EOS, and Cardano during the COVID-19 period. Except for Litecoin, there is no significant evidence of causal effects from trading volume to the return of cryptocurrencies prior to the COVID-19, while during the COVID-19 period, trading volume of Tether and Chainlink Granger cause their returns.

As a further investigation, the general return-volume relation for cryptocurrency markets is tested, and the results did not show any significant relationship. However, considering the absolute return values, we found a significant causal effect from cryptocurrencies' absolute returns to trading volume for both the prior and during COVID-19 periods.

Our analyses have several implications for policymakers. Even though cryptocurrencies are not yet effectively backed by all governments, understanding the effect of financial crisis, such as the one followed by the COVID-19, on these markets is essential. Our findings enable policymakers to better react to the dynamics of digital currencies, and the potential effects of these markets on other

financial and commodity markets. Accordingly, they can adjust their monetary policy decisions promptly. Additionally, this study will provide investors with practical insights to distinguish the associated risk with cryptocurrencies and commodity markets, allowing them to make more effective decisions about their trading positions during the COVID-19 pandemic. In this regard, Gold can be considered a suitable asset for portfolio hedging during the pandemic period. Besides, it was found that cryptocurrency traders tend to trade in high volumes while prices vary extremely and this behaviour is not significantly affected by the COVID-19 crisis. Our findings about the trading volume can help traders and investors identify the effect of momentum and potential trend in cryptocurrencies on their investments.

Chapter 3

Connectedness Between Cryptocurrency, Crude Oil and Gold Markets: An Analysis of the Effect of COVID-19 Pandemic

3.1 Introduction and Literature Review

The COVID-19 crisis had far-reaching effects on various financial markets, including stocks, commodities, exchange rates, and cryptocurrency markets (Ozili, 2020; Naeem et al., 2021; Le et al., 2021; Drake, 2022). The changes in economic and financial market dynamics due to the COVID-19 pandemic are expected to have significant implications for hedge fund managers, cryptocurrency market investors, and policymakers (Khelifa et al., 2021; Corbet et al., 2021). Financial market players and investors need to account for the changes in market dynamics due to the COVID-19 pandemic and diversify their investment portfolios with suitable hedge and safe-haven assets to reduce risk.

The common notion that investors shift to gold during economic market crises remains unverified for many cryptocurrency markets. This paper investigates the connectedness between the ten most traded cryptocurrencies and Gold and crude oil markets in pre-COVID-19 and during the COVID-19 periods.

A safe haven asset is either uncorrelated or negatively correlated with another asset, specially during negative market conditions. A hedge asset demonstrates no correlation or a negative correlation with another asset, and a diversifier displays a positive but not perfect correlation with another asset (Nedved & Kristoufek, 2023). A safe haven investment is an asset that investors turn to during times of economic and financial uncertainty or market turbulence (Baur & Lucey, 2010; Baur & McDermott, 2010). The primary purpose of a safe haven is to preserve capital and reduce risk exposure during periods of market stress. Baur & McDermott (2010) define a strong safe haven as an asset that is negatively correlated with equities, whereas a weak safe haven asset is uncorrelated with equities during corresponding periods.

Due to the intrinsic value of commodities, they are less vulnerable to inflations or financial crises. Thus, they are considered a potential hedging candidate for other markets (Disli et al., 2021). Crude oil is one of the key commodities traded internationally which holds significant influence over macroeconomic factors such as inflation, exchange rates, and economic growth. Research by Bashiri Behmiri and Pires Manso (2013) demonstrates that crude oil prices are influenced by various economic, political, environmental, financial, and technological factors. On the other hand, Gold has historically been attractive as a financial asset due to its tangible ability to maintain its value, particularly in periods of financial, political, and economic uncertainty (Baur & Lucey, 2010). Thus, numerous financial researchers have suggested that gold can serve as a reliable safe-haven asset for other markets, including stock and bond markets, as well as for investors holding U.S. dollars, providing protection against macroeconomic risks during financial crises (Baur & Lucey, 2010; Reboredo, 2013; Junttila et al., 2018; Baek, 2019). For instance, Akhtaruzzaman et al. (2021) examine the role of gold as a hedge or safe-haven asset during the COVID-19 crisis and find evidence that gold initially served as a safe haven for stock markets at the onset of the pandemic. However, during the second phase of the crisis (March 17 to April 24, 2020), there was no clear evidence of gold's safe-haven properties. Additionally, Nkrumah-Boadu et al. (2022) discover that both gold and cryptocurrency markets provide safe haven characteristics for investors in the African stock market.

The literature extensively examines crude oil's hedging and safe-haven properties for stock markets (Park and Ratti's, 2008; Creti et al., 2013; Śmiech & Papież, 2017; Junttila et al., 2018). However, there are limited studies on the safe-haven effect of crude oil on cryptocurrency markets. Elie et al. (2019) study the potential safe-haven roles of gold and crude oil against extreme downward trends in clean energy stock indices. Their analysis, using daily price data, demonstrated that both crude oil and gold exhibited weak safe-haven characteristics for clean energy indices. Building upon this body of evidence, we aim to investigate whether the hedging or safe-haven effects of gold and crude oil is also evident in cryptocurrency markets.

Recognizing hedge assets that are uncorrelated with cryptocurrencies or safe-haven assets that exhibit negative correlation during cryptocurrency downturns is a plausible approach to mitigate risks and maintain profits (Nedved & Kristoufek, 2023). This has led to a growing interest in exploring the interactions between cryptocurrencies and traditional financial assets, such as gold and crude oil (Khelifa et al., 2021; Lahmiri & Bekiros, 2020; Conlon & McGee, 2020; Mnif et al., 2020; Corbet et al., 2020; Goodell & Goutte, 2021).

In recent studies, the relationship between commodities, such as gold and crude oil, and cryptocurrency markets has gained significant attention. For instance, Owusu Junior et al. (2020) employed an EEMD-based quantile-on-quantile regression model with the data on gold and eight cryptocurrencies (Bitcoin, Ethereum, Dash, Litecoin, Ripple, Stellar, NEM, and Monero) from April 2013 to April 2019. Their findings show that cryptocurrencies and gold can serve as both hedge and diversification assets for each other across various conditional distributions of their returns. González et al. (2021) examined the relationship between gold and cryptocurrency returns,

highlighting an asymmetric connectedness of cryptocurrencies to gold returns during economic disturbances. Similarly, Barson et al. (2022) conducted a causality analysis to study the relationship between gold and cryptocurrencies during the COVID-19 pandemic. Their empirical results suggest that gold exhibits weak safe haven characteristics during the COVID-19 pandemic while acting as a hedge at normal cryptocurrency market conditions.

Several recent studies have examined the interdependence between crude oil and cryptocurrencies. Selmi et al. (2018) compared the hedging and safe haven properties of Bitcoin and Gold for oil markets and concluded that both Gold and Bitcoin could serve as hedges, safe havens, and diversifiers against crude oil. Okorie and Lin (2020) explored the volatility connectedness between crude oil and cryptocurrencies, finding a unidirectional volatility spillover from Ethereum, XRP, and ReddCoin to the crude oil markets. The assessment of the potential safe haven role of gold, crude oil, and cryptocurrency during the COVID-19 pandemic by Disli et al. (2021) rejects the safe haven characteristics of gold, oil, and Bitcoin for the stock market. On the other hand, the study of the hedging effect of cryptocurrencies for Brent crude oil price movements during the COVID-19 pandemic shows that only stablecoins protect against plunging oil price movements. Still, they do not reduce investment volatility (Będowska-Sójka & Kliber, 2022)https://doi.org/10.1016/j.resourpol.2020.101731. The study by Wang et al. (2023) reveals that Bitcoin has a limited safe-haven effect on the crude oil market, whereas gold demonstrates a strong safe-haven ability for crude oil, before and after the COVID-19 pandemic. This finding supports the results from the study of Wen et al. (2022), which concluded that Bitcoin does not serve as a safe haven for oil markets. Nedved and Kristoufek (2023) investigated the safe-haven role of gold, crude oil, and stock markets for Bitcoin and concluded that while Bitcoin movements were positively correlated with that of the stock market, oil and gold demonstrated safe haven characteristics, with gold being a strong safe haven for Bitcoin.

As shown in this section, most research studies on the connectedness between cryptocurrencies and other financial markets are limited to one or two cryptocurrencies, namely Bitcoin and Ether (Katsiampa, 2019; Conlon & McGee, 2020; Goodell & Goutte, 2021). Thus, it is crucial for crypto literature to extend its attention to other prominent cryptocurrencies that have gained traction in recent years. This broader examination will enhance our overall comprehension of the relationship between cryptocurrency markets and various domains, such as commodities. Currently, there exists limited knowledge regarding the safe haven effect of gold or crude oil for cryptocurrency markets during financial crises. In the current study, an attempt has been made to fill this gap in the literature. In this context, our study will examine dynamic short-term and long-term relationships and potential hedging effects of gold and crude oil markets for the ten most traded cryptocurrencies. Specifically, we aim to investigate whether the safe haven potential of gold or crude oil for cryptocurrency remains consistent before and during the COVID-19 pandemic. We find that during the COVID-19 pandemic, Gold plays a significant safe haven role for Bitcoin, Litecoin, and Monero, while crude oil is a weak safe haven for Ether, Bitcoin Cash, EOS, and Monero. Among the crude oil assets, Brent can be a strong safe haven for Bitcoin during the COVID-19 pandemic. Our causality analyses

indicate that a bidirectional relationship exists between Gold and cryptocurrencies such as Bitcoin and Monero during the pandemic.. Indeed, knowing this relationship holds profound implications for hedge fund managers and investors for altering their risk exposure and modifying a cryptocurrency portfolio's return. Our findings are intended to serve as a basis for future discussions about safe havens for cryptocurrencies, as these assets are becoming increasingly popular in individual and institutional investors' portfolios. This study is crucial in foreseeing the behavior of cryptocurrency markets and protecting investors against extreme price movements, especially during global pandemics.

Following the literature review, the following research questions is examined in this paper:

1. Is gold or crude oil suitable a safe haven asset for cryptocurrency markets during the COVID-19 crisis?
2. Can cryptocurrency markets be considered as safe haven assets for gold or crude oil markets during the COVID-19 crisis? If yes, which cryptocurrencies?
3. Does the hedging ability of gold and crude oil markets differ before and during the COVID-19 pandemic?
4. Is there any Granger causality effect between gold, crude oil, and cryptocurrency markets? If yes, what is the direction of causal effect?

Subsequently, this study makes the following important contributions to the existing literature on the connectedness of cryptocurrencies with other markets:

1. This study employs VAR, VECM, ARDL, and Granger causality analyses to examine the relationship between ten cryptocurrency markets with commodity markets such as Gold and Crude Oil and finds whether a hedging or safe-haven role exists among these markets.
2. Examination of the COVID-19 Pandemic: We analyze the effect of the COVID-19 pandemic as the first economic distress after the emergence of cryptocurrencies on the hedging opportunity of Gold and Crude Oil for cryptocurrency markets dynamics prior to and during the COVID-19 pandemic periods, which has received limited attention compared to the conventional financial markets.
3. Comprehensive Analysis of Cryptocurrency Markets: Contrary to the existing literature that restricts their analysis to a limited number of cryptocurrency markets, mainly Bitcoin and Ether, our study seeks to find the hedging effect of Gold and Crude Oil on the ten most traded cryptocurrency markets. Our findings will contribute to a comprehensive understanding of the interconnectedness between Gold and Crude Oil markets and digital currencies.
4. Exploration of Safe Haven Effects on cryptocurrencies: while prior literature primarily investigates the safe haven effect of Bitcoin on the crude oil market, in this study, we aim to examine the contrary safe haven effect of crude oil for cryptocurrency markets. To enhance the robustness of our findings, we investigate the safe haven effect of two widely traded oils, specifically WTI and Brent, on cryptocurrencies. By incorporating the analysis of these oils, we aim to strengthen the validity and reliability of our results.

3.2 Methodology

Our empirical analysis begins by conducting preliminary tests to confirm the presence of nonlinearities and the stationarity properties in the variables used in this study. The Augmented Dickey-Fuller (ADF) test (Cheung & Lai, 1995) is employed on each series in levels to determine whether a series has a unit root. In order to examine the co-movement between cryptocurrency markets and Oil and Gold markets, the Johansen and Bounds cointegration tests are conducted. These tests help identify the appropriate econometric model for analyzing these markets' short-term and long-term relationships. Following the results of the ADF test, one of the following three cases may apply: *(i)* Series are integrated of order 0 (stationary in level): cointegration test is unnecessary; *(ii)* Series are integrated of order 1 (stationary after first difference): a cointegration test is necessary to establish a long-run relationship. Johansen cointegration test (Johansen, 1991) can be performed on the level. If any cointegration equation is found, we estimate both VAR (Sims, 1980b) for short-term relationships and VECM (Engle & Granger, 1987) for long-term relationships. If no cointegration equation is found, only the VAR (for first difference) model should be estimated; and *(iii)* Series are integrated of different orders: in this case, the Bounds test (Pesaran et al., 2001) for cointegration can be performed on the level. After conducting the Bounds cointegration test, we determine the applicable model based on the results. If there is no cointegration, we estimate the short-term relationships using the ARDL (Pesaran & Shin, 1995) model. If cointegration exists, we estimate the long-term VECM model.

This rigorous methodology enhances the reliability and validity of our analysis. Following the cointegration tests, one of the below models is applicable.

The equations for $\mathrm{VAR(p)}$ model for the first difference of series in this study are:

$$\Delta P_t = a_{0,p} + \sum_{i=1}^{p} b_{p,i}\, \Delta P_{t-i} + \sum_{i=1}^{p} c_p\, \Delta x_{t-i} + u_{p,t} \tag{3-1}$$

$$\Delta x_t = a_{0,x} + \sum_{i=1}^{p} b_{x,i}\, \Delta P_{t-i} + \sum_{i=1}^{p} c_x\, \Delta x_{t-i} + u_{x,t} \tag{3-2}$$

Where, $a_{0,p}$ and $a_{0,x}$ are constants, P_t is the log of cryptocurrency price at time t, P_{t-i} is the i^{th} lag of cryptocurrency price, x_{t-j} is the j^{th} lag of the log of Gold, WTI, or Brent price, and $u_{p,t}$ and $u_{x,t}$ are the error terms for the cryptocurrency and crude oil or gold markets, respectively.

The equation for the long-term $\mathrm{VECM(p,q)}$ model in this study is expressed as:

$$\Delta P_t = \alpha_0 + \sum_{i=1}^{p} \varphi_i \Delta P_{t-i} + \sum_{j=1}^{q} \beta_j \Delta x_{t-j} + \lambda\, ECT_{t-1} + \varepsilon_{t,vecm} \tag{3-3-a}$$

$$ECT_{t-1} = c + \gamma P_{t-1} + \theta x_{t-1} \tag{3-3-b}$$

Where, α_0 is a constant, P_t is the log of cryptocurrency price at time t, P_{t-i} is the i^{th} lag of cryptocurrency log price, x_{t-j} is the j^{th} lag of the log of Gold, WTI, or Brent price, and $\varepsilon_{t,vecm}$ is the error term in the VECM model. The long-term effect between markets is captured by the ECT_{t-1} term in equation (3-a), and λ is the adjustment coefficient. This cointegrating term represents the long-run equilibrium relationship between cryptocurrency and other markets, while λ shows the speed of convergence towards the long-run equilibrium. The lag structures of $\mathrm{VAR(p)}$ and

$\mathrm{VECM}(p, q)$ models for each market are determined according to the Akaike Information Criterion (AIC) (Akaike, 1974).

ARDL model is an ordinary least square (OLS) based model which is applicable for both non-stationary time series and times series with mixed order of integration. An $\mathrm{ARDL}(p, q)$ model is shown below:

$$P_t = \alpha_0 + \sum_{i=1}^{p} \varphi_i P_{t-i} + \sum_{j=0}^{q} \beta_j x_{t-j} + \varepsilon_{t,ardl} \quad (3\text{-}4)$$

Where, α_0 is a constant, P_t is the log of cryptocurrency price at time t, P_{t-i} is the i^{th} lag of cryptocurrency log price, x_{t-j} is the j^{th} lag of the log of Gold, WTI, or Brent price, and $\varepsilon_{t,ardl}$ is the error term in the ARDL model. The proper lag structure in ARDL model is selected by Schwarz criterion (SIC) (Schwarz, 1978).

The estimated regression coefficients from Eq.(3-1) to Eq.(3-4) are then used to carry out pairwise Granger causality tests (Granger, 1969) to find whether an endogenous variable can be treated as exogenous. The Wald Chi-Square test for the null hypothesis that Gold, WTI or Gold does not Granger cause the cryptocurrency market is examined. A bidirectional Granger causality exists between variables if both null hypotheses are rejected. Recall that Granger causality analysis adopted in the current work was successful in various financial problems, including the studying causality between stock prices and economic activity (Yilanci et al., 2021), insurance market density and economic growth (Pradhan et al., 2017), dynamic price changes in securities (Virgilio, 2022), banking activities and economic growth (Mushtaq, 2016), Bitcoin market and internet attention (Zhang et al., 2021), pricing dynamics of cryptocurrencies (Kristoufek, 2022), and foreign direct investment and economic growth (Sarker, 2020).

3.3 Data Description

In line with the previous discussions, gold and crude oil are popular hedges and safe haven assets for many financial markets. This study looks into the connections between the ten most traded cryptocurrency markets and the markets for Gold, WTI, and Brent Crude Oil. For this study, we have used the same dataset as in section Data2.3.1. However, since cryptocurrency prices are available seven days a week, while the WTI, Brent, and Gold spot prices are only available five days a week, we have matched the trading days of cryptocurrencies in accordance with the trading days of Gold and Crude oil markets. This resulted in a sample size of 248 and 246 daily prices for the pre-COVID-19 and during the COVID-19 periods, respectively.

3.4 Empirical Results

We first employ the ADF test to ascertain the presence of a unit root in the price time series. The p-values presented in Table 3-1 indicate that the logarithmic price series of Tether during both pre-COVID-19 and during COVID-19 periods, as well as WTI and Brent in pre-COVID-19 period are stationary. Conversely, for all other markets, only the return series, represented by the first difference of the logarithmic price series, demonstrate stationarity. To assess short-term and long-term

relationships between cryptocurrency markets and crude oil and gold markets, we employ the Johansen and Bounds Cointegration Tests. Cointegration implies a long-term equilibrium relationship between variables, indicating their tendency to move together despite short-term deviations (Jansen et al., 1993). Table 3-2 shows the F-values of Bounds test and the p-values of Johansen test for the null hypothesis of no cointegration between market pairs. The F-value for Tether-Gold market pairs in the pre-COVID-19 period and for Tether-Gold, Tether-WTI, and Tether-Brent pairs during the COVID-19 period exceed the I(1) critical value. Consequently, we infer the existence of a cointegration equation between these markets. Accordingly, we will estimate both ARDL for the short-term and VECM for the long-term relationships among Tether and gold, WTI, and Brent crude oil markets. The p-values from the Johansen test verify that cointegration equation only exists among Bitcoin cash-WTI, Bitcoin cash-Brent, and Chainlink-Gold market pairs during COVID-19. Table 3-3 summarizes the suitable models following the results of cointegration tests for each market.

Table 3-1: ADF test results (p-values)

Markets	Pre-COVID-19 (2019) 2019/01/01-2019/12/31		During COVID-19 (2020) 2020/01/01-2020/12/31	
	Return	Log Price	Return	Log Price
Tether	**0.000**	**0.007**	**0.000**	**0.000**
Bitcoin	**0.000**	0.574	**0.000**	0.994
Ether	**0.000**	0.626	**0.000**	0.837
Ripple	**0.000**	0.686	**0.000**	0.169
Litecoin	**0.000**	0.537	**0.000**	0.805
Bitcoin Cash	**0.000**	0.607	**0.000**	0.116
EOS	**0.000**	0.609	**0.000**	0.091
Chainlink	**0.000**	0.427	**0.000**	0.455
Cardano	**0.000**	0.743	**0.000**	0.729
Monero	**0.000**	0.765	**0.000**	0.739
Gold	**0.000**	0.819	**0.000**	0.397
WTI	**0.000**	**0.008**	**0.000**	0.181
Brent	**0.000**	**0.032**	**0.000**	0.133

Table 3-2: Bounds or Johansen cointegration test

Market pairs	Pre-COVID-19 (2019)		During COVID-19 (2020)	
	Bounds Test	Johansen Test	Bounds Test	Johansen Test
Tether-Gold	**6.56**		**60.09**	
Tether-WTI			**56.97**	
Tether-Brent			**56.71**	
Bitcoin-Gold		0.502		0.710
Bitcoin-WTI	1.12			0.393
Bitcoin-Brent	1.52			0.568
Ether-Gold		0.664		0.725
Ether-WTI	0.91			0.440
Ether-Brent	1.70			0.514
Ripple-Gold		0.763		0.231
Ripple-WTI	0.78			0.079
Ripple-Brent	1.08			0.107
Litecoin-Gold		0.254		0.735
Litecoin-WTI	0.76			0.430
Litecoin-Brent	1.36			0.637
Bitcoin cash-Gold		0.725		0.193
Bitcoin cash-WTI	1.61			**0.029**
Bitcoin cash-Brent	1.89			**0.040**
EOS-Gold		0.594		0.174
EOS-WTI	0.84			0.068
EOS-Brent	1.84			0.117
Chainlink-Gold		0.207		**0.029**
Chainlink-WTI	2.04			0.274
Chainlink-Brent	1.65			0.289
Cardano-Gold		0.630		0.552
Cardano-WTI	0.42			0.233
Cardano-Brent	0.99			0.210
Monero -Gold		0.420		0.707
Monero -WTI	0.63			0.259
Monero -Brent	0.85			0.274

Values in bold are significant at a 5% significance level. Values for the Bounds test are F-statistic for the null hypothesis that there is no cointegration in levels. F- critical values for lower and upper bounds are I(0)=3.62, I(1)= 4.16 at a 5% significance level. Values for the Johansen Unrestricted Cointegration Rank Test (Trace) are p-values for the null hypothesis that there is no Cointegration equation.

Table *3-3:* Selected econometrics models to investigate market co-movement

Markets	Models	
	Pre-COVID-19 (2019)	During COVID-19 (2020)
Tether-Gold	VECM	VECM
Bitcoin-Gold Ether-Gold Ripple-Gold Litecoin-Gold Bitcoin Cash-Gold EOS-Gold Cardano-Gold Monero-Gold	VAR-first D	VAR-first D
Chainlink-Gold	VAR-first D	VECM
Tether-WTI/ Brent	VAR- level	VECM
Bitcoin cash-WTI/Brent	ARDL	VECM
Bitcoin-WTI/Brent Ether-WTI/Brent Ripple-WTI/Brent Litecoin-WTI/Brent EOS-WTI/Brent/Brent Chainlink-WTI/Brent Cardano-WTI/Brent Monero-WTI/Brent	ARDL	VAR-first D

Following the cointegration tests, the VECM model can only analyze the potential long-term relationship between Tether-Gold, Tether-WTI, Tether-Brent, chainlink-Gold, Bitcoin Cash-WTI, and Bitcoin Cash-Brent market pairs. Moreover, short-term relationships between cryptocurrency markets and crude oil and gold will be explored by utilizing the ARDL or VAR-in-first-difference models. According to the results of VECM models in Table 3-4, during the COVID-19 period, λ_{Gold}, λ_{WTI}, and λ_{Brent} coefficients for Tether market are significantly negative. Thus, during COVID-19, there is a significant long-term negative relationship between Tether market and Gold, WTI, and Brent markets at a 5% level. The Negative long-term λ coefficient in Eq.(3-3-a) shows that series are convergent in the long run to their long-term equilibrium relationship when deviations occur in the short term. Baur & McDermott (2010) define a strong safe haven as an asset that is negatively correlated with equities, whereas a weak safe haven asset is uncorrelated with equities during corresponding periods. Here, we consider an asset with a significant negative relationship as a strong safe haven and an asset with an insignificant negative coefficient as a weak safe haven. The coefficient of lagged Gold, $\beta_{1,Gold}$, for Tether is negative but not statistically significant, indicating that Gold can be a weak safe haven for Tether during the stable markets before the pandemic. However, the short-term effects of Gold, WTI, and Brent crude oil on Tether, $\beta_{1,Gold}$, $\beta_{1,WTI}$, and $\beta_{1,Brent}$, are positive and insignificant, showing that Gold, WTI, and Brent do not show any safe haven properties for Tether during the pandemic. Similarly, significant long-term negative relationships between Bitcoin Cash and WTI and Bitcoin Cash and Brent crude oil markets are found. This indicates that Bitcoin cash will converge to the long-run equilibrium with WTI and Brent crude oil when deviations occur in the short term. Regarding the safe haven effects, the insignificant negative coefficients of crude oil lag for

Bitcoin Cash, $\beta_{1,WTI}$, and $\beta_{1,Brent}$, show a weak safe haven property during the COVID-19 pandemic. Our results also show a significant long-term negative relationship, λ_{Gold}, between the Chainlink and Gold markets and a weak safe haven characteristic of Gold, $\beta_{1,Gold}$, for Chainlink market during the COVID-19 pandemic.

Comparing the results of the VECM model in the pre-COVID-19 period shows that the long-term relationship, λ_{Gold}, is only available between Tether and Gold markets. Our results suggest that Gold can have a weak safe-haven effect for Tether in the pre-COVID-19 period and for Chainlink during the COVID-19 period. Besides, during the COVID-19 period, crude oil markets such as WTI and Brent can play an investment hedging role for Bitcoin Cash market. Thus, our findings suggest diversifying the Bitcoin Cash portfolios with crude oil assets during COVID-19 to minimize the risk of investment. However, during the stable markets (pre-COVID-19 period), it is suggested that investors in Tether diversify their portfolios with Gold assets.

Table 3-4: Estimated coefficients in VECM models

	Tether		Bitcoin Cash	Chainlink
	Pre-COVID-19	During COVID-19	During COVID-19	
λ_{Gold}	**-0.2623**	**-0.8122**		**-0.0574**
	(-5.365)	**(-9.742)**		**(-2.917)**
θ_{Gold}	0.0173	-0.002		**-8.1508**
	(1.343)	(-0.467)		**(-9.257)**
$\beta_{1,Gold}$	-0.0179	0.0027		-0.3408
	(-0.551)	(0.107)		(-0.800)
λ_{WTI}		**-0.8187**	**-0.0331**	
		(-9.923)	**(-1.919)**	
θ_{WTI}		0.0011	**-0.693**	
		(1.108)	**(-4.716)**	
$\beta_{1,WTI}$		0.0002	-0.0144	
		(0.125)	(-0.764)	
λ_{Brent}		**-0.8241**	**-0.0718**	
		(-10.046)	**(-2.967)**	
θ_{Brent}		0.0016	**-0.465**	
		(1.470)	**(-3.265)**	
$\beta_{1,Brent}$		0.0063	-0.0145	
		(1.475)	(-0.246)	

This table presents the coefficients in Eq.(3-3-a) and Eq.(3-3-b). Values in parentheses are t-statistics. Significant coefficients at a 0.05 level are in bold.

The results of the estimated short-term relationships between cryptocurrency markets and crude oil and gold markets with ARDL and VAR models are presented in Table 3-5 and Table 3-6, respectively. As indicated in Table 3-5, our findings did not suggest any short-term relationship between cryptocurrency markets and crude oil markets before the COVID-19 pandemic at a 5% level of significance. However, considering a 10% level, a significant positive effect of Brent price, represented by β_0 coefficient, on the price of Ether, Bitcoin Cash and EOS is evident before the pandemic. Likewise, WTI price has a significant positive effect on Bitcoin Cash price with a 10% level. These finding show that crude oil

price movements are similar to the price movements of Ether, Bitcoin Cash, and EOS and that crude oil is not a safe haven for these cryptocurrencies during the stable markets. However, the negative insignificant WTI and Brent crude oil coefficients for other assets in Table 3-5 indicate that crude oil can be a weak safe haven asset for Ripple, and Chainlink in short-term.

Table 3-5: Estimated coefficients in ARDL models

	Bitcoin	Ether	Ripple	Litecoin	Bitcoin Cash	EOS	Chainlink	Cardano	Monero
Pre-COVID-19 (2019)									
WTI	0.0302	0.0481	-0.0073	0.0003	**0.1069**	0.0533	-0.0885	0.0158	0.0003
	(0.722)	(0.977)	(-0.169)	(0.005)	**(1.727)**	(0.839)	(-1.132)	(0.272)	(0.006)
Brent	0.0529	**0.0848**	0.0400	0.0737	**0.1129**	**0.1209**	-0.0271	0.0782	0.0380
	(1.312)	**(1.821)**	(0.970)	(1.335)	**(1.950)**	**(1.918)**	(-0.348)	(1.332)	(0.797)

This table presents the β_0 coefficients in ARDL model (Eq.(3-4)). Values in the parentheses are associated t-statistic. Significant coefficients at 0.1 level are in bold.

Besides, the findings presented in Table 3-6 provide evidence that there is no statistically significant short-term connectedness between cryptocurrencies and WTI, Brent, and Gold markets in the pre-COVID-19 period, as their coefficients are not significant at a 5% level. However, with a 10% significance level, we observe a negative relationship between lagged Gold returns have and Litecoin and Monero before the COVID-19 pandemic. Moreover, the first lag of Gold has an insignificant negative relationship with all cryptocurrency markets before the COVID-19 pandemic. These results confirm weak safe haven properties of Gold for all ten cryptocurrencies in the periods of financial stability. Likewise, the insignificant negative lagged WTI and Brent coefficients for Tether market in Table 3-6 suggest that crude oil can serve as a weak safe haven for Tether in stable periods.

On the other hand, during the COVID-19 pandemic, lagged Gold returns have a significant negative relationship with Bitcoin, Litecoin, and Monero at a 5% level. Thus, Gold is a strong safe haven asset for Bitcoin, Litecoin, and Monero during financial crisis during the COVID-19. Meanwhile, the lagged WTI has a significant direct association with Bitcoin return while the lagged Brent crude oil returns have a significant negative relationship with Bitcoin returns, at the 5% level. Therefore, during the COVID-19 pandemic, Brent crude oil can be considered as a strong safe haven for Bitcoin, whereas WTI is not a safe haven for Bitcoin in this period. This study did not find any significant relationship between WTI and Brent crude oil markets and Ripple, Litecoin, Bitcoin Cash, EOS, Chainlink, and Cardano during the COVID-19 pandemic. We neither found any significant relationship between Gold and Ether, Ripple, Bitcoin Cash, EOS, Chainlink, and Cardano markets in any of our sub-sample periods. However, due to insignificant negative coefficients, WTI is a weak safe haven for Ether, Ripple, EOS, Chainlink, and Monero; Brent is a weak safe haven for Ether, Cardano, and Monero; and Gold is a weak safe haven for Bitcoin Cash, EOS, and Cardano during the COVID-19 crisis. Overall, the findings of our study indicate that Gold exhibits stronger safe haven characteristics for cryptocurrencies compared to the crude oil markets, particularly during the COVID-19 pandemic. In addition, the safe haven characteristics of both Gold and Crude oil markets for most of cryptocurrencies

have improved during the COVID-19 pandemic when we compare it with the prior to the COVID-19 pandemic.

Table 3-6: Estimated coefficients in VAR-in-first-difference models

	Bitcoin	Ether	Ripple	Litecoin	Bitcoin Cash	EOS	Chainlink	Cardano	Monero
Pre-COVID-19 (2019)									
DGold (-1)#	-0.018	-0.241	-0.466	-0.928*	-0.72-	-0.594	-0.839	-0.611	-0.146
	(-0.048)	(-0.550)	(-1.215)	(-1.862)	(-1.342)	(-1.118)	(-1.215)	(-1.259)	(-0.326)
DGold (-2)	0.2184								-0.336
	(0.579)								(-0.758)
DGold (-3)	-0.453								-0.762*
	(-1.188)								(-1.697)
DGold (-4)	-0.053								-0.214
	(-0.138)								(-0.475)
Tether in level									
WTI (-1)	0.001	Brent (-1)	0.003						
	(0.066)		(0.273)						
WTI (-2)	0.003	Brent (-2)	-0.002						
	(0.177)		(-0.12)						
WTI (-3)	-0.007	Brent (-3)	-0.003						
	(-0.670)		(-0.261)						
During COVID-19 (2020)									
DGold (-1)	-0.196	0.069	0.125	0.024	-0.235	-0.084		-0.107	0.122
	(-0.822)	(0.220)	(0.313)	(0.078)	(-0.689)	(-0.268)		(-0.298)	(0.417)
DGold (-2)	0.206	0.367	0.286	0.230					0.283
	(0.864)	(1.162)	(0.718)	(0.756)					(0.969)
DGold (-3)	**-0.657**			**-0.936**					**-0.731**
	(-2.813)			**(-3.087)**					**(-2.552)**
DGold (-4)	**-0.491**								-0.539*
	(-2.078)								(-1.865)
DWTI (-1)	0.014	-0.001	0.002	0.011		-0.000	0.001	0.011	0.007
	(1.005)	(-0.029)	(0.086)	(0.633)		(-0.024)	(0.021)	(0.611)	(0.402)
DWTI (-2)	0.005	-0.005	-0.005	0.007			-0.004		-0.000
	(0.317)	(-0.256)	(-0.222)	(0.357)			(-0.149)		(-0.016)
DWTI (-3)	0.014			0.001			0.005		0.002
	(0.919)			(0.077)			(0.197)		(0.098)
DWTI (-4)	0.021						0.005		0.004
	(1.356)						(0.203)		(0.230)
DWTI (-5)	**0.027**								
	(1.977)								
DBrent (-1)	0.008	0.021	0.009	0.017		0.003	0.002	-0.019	0.031
	(0.204)	(0.374)	(0.122)	(0.315)		(0.045)	(0.025)	(-0.308)	(0.620)
DBrent (-2)	0.004	0.058	0.036	0.028			0.084		-0.020
	(0.096)	(1.056)	(0.509)	(0.519)			(1.148)		(-0.406)
DBrent (-3)	0.051	0.035		0.049			0.105		0.044
	(1.231)	(0.621)		(0.911)			(1.423)		(0.880)
DBrent (-4)	0.029	0.008					0.075		0.012
	(0.708)	(0.152)					(1.023)		(0.246)
DBrent (-5)	**-0.085**	-0.104*							-0.085*
	(-2.077)	(-1.898)							(-1.723)

This table presents the coefficients in VAR model (equation (1)). Values in the parentheses are associated t-statistic. Significant coefficients at a 0.05 level are in bold and values with (*) are significant at the 10% level. # DGold (-1) shows the first difference of gold at the first lag.

Table 3-7 presents the p-values for unidirectional Granger causality tests, examining the causal effect of Gold, WTI, and Brent return series towards cryptocurrency returns. The results reveal that during the COVID-19 pandemic, Gold returns significantly Granger caused Bitcoin, Litecoin, and Monero returns at a 5% significance level. However, no significant Granger causality effect of WTI and Brent crude oil returns on cryptocurrency returns was found in the same period. In the pre-COVID-19 period, Gold returns Granger caused Litecoin returns at a 10% significant level. Similarly, WTI and Brent crude oil prices show a significant causal effect on Bitcoin Cash at a 10% level in this period. This study could not find any significant causality relationship from Gold, WTI, and Brent crude oil towards other cryptocurrencies in pre-COVID-19 and during COVID-19 periods. These results from the Granger causality test also confirm our findings about safe haven effects discussed previously. More specifically, the significant Granger causality from Gold towards Bitcoin, Litecoin, and Monero and the observed negative effects from Table 3-6 confirm that Gold can serve as a strong safe haven asset for these markets during the COVID-19 crisis.

Table 3-7: Granger causality from Gold, WTI, and Brent towards cryptocurrencies (p-values)

Markets	H0: Gold Granger causes Cryptocurrency		H0: WTI Granger causes Cryptocurrency		H0: Brent Granger causes Cryptocurrency	
	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19
Tether	0.5819	0.9146	0.7183	0.9008	0.942	0.1403
Bitcoin	0.7788	**0.008**	0.2475	0.4312	0.166	0.2192
Ether	0.5819	0.5017	0.3141	0.9655	0.1389	0.3732
Ripple	0.2244	0.7421	0.8814	0.96	0.384	0.8745
Litecoin	0.0626*	**0.0152**	0.7477	0.9351	0.2312	0.7816
Bitcoin Cash	0.1795	0.4904	0.0616*	0.4446	0.0771*	0.8053
EOS	0.2636	0.7883	0.3327	0.9806	0.1424	0.9639
Chainlink	0.2245	0.4236	0.6113	0.9976	0.871	0.4251
Cardano	0.208	0.7656	0.7163	0.541	0.3177	0.758
Monero	0.4355	**0.0313**	0.6177	0.9924	0.3675	0.4436

Values in bold are significant at the 5% level and values with (*) are significant at the 10% level.

The analyses in this study suggest that, in general, Gold is a better safe haven asset than crude oil for cryptocurrencies in both periods of pre-COVID-19 and during COVID-19. Even though crude oil is generally found to be a weak safe haven for cryptocurrencies, the safe haven properties of crude oil for Bitcoin, Ether, Bitcoin Cash, EOS, and Monero have risen during the COVID-19 pandemic. During the COVID-19 crisis, hedge fund managers can reduce the risk of investing in Bitcoin, Litecoin, and Monero by assigning a portion of their portfolio investments to the Gold market. Besides, Bitcoin investors can also reduce their investment risks during the COVID-19 pandemic by investing in Brent crude oil. However, as the absolute value of the coefficients for the effect of Gold on Bitcoin is larger than the effect of Brent crude oil on Bitcoin, gold would be a better hedging asset for Bitcoin during the COVID-19 pandemic. These results are consistent with prior literature (Owusu Junior et al., 2020; González et al., 2021; Barson et al., 2022; Nedved & Kristoufek, 2023), which provide evidence of

Gold's safe haven properties for cryptocurrency markets. Our results corroborate the findings of Nedved & Kristoufek (2023) regarding the strong safe haven nature of Gold and the weaker safe haven role of crude oil for the Bitcoin market. However, we were unable to compare our results on crude oil's safe haven properties for other cryptocurrencies with prior studies, as we did not find any existing research on this specific topic. Table 3-8 summarizes the existence and the intensity of safe haven properties of Gold, WTI, and Brent crude oil for ten cryptocurrencies that are analyzed in this study.

Table 3-8: Safe haven properties of Gold, WTI, and Brent crude oil

	Gold		WTI		Brent	
	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19	Pre-COVID-19	During COVID-19
Tether	Weak	None	Weak	None	Weak	None
Bitcoin	Weak	**Strong**	None	None	None	**Weak**
Ether	Weak	None	None	Weak	None	Weak
Ripple	Weak	None	Weak	Weak	None	Weak
Litecoin	**Strong**	**Strong**	None	None	None	None
Bitcoin Cash	Weak	Weak	None	Weak	None	Weak
EOS	Weak	Weak	None	Weak	None	None
Chainlink	Weak	Weak	Weak	Weak	Weak	None
Cardano	Weak	Weak	None	None	None	Weak
Monero	**Strong**	**Strong**	None	Weak	None	Weak

To investigate the possible causal relationships from cryptocurrency markets toward the Gold, WTI, and Brent crude oil markets, Table 3-9 presents the p-values of Granger causality tests. These tests consider the cryptocurrency returns as exogenous variables to estimate Gold, WTI, and Brent market returns. Our findings suggest that before the COVID-19 pandemic, only Bitcoin and Chainlink prices have a significant causal effect on Brent crude oil price at 0.1 level. However, during the COVID-19 period, Tether, Bitcoin, and Ether returns significantly Granger caused the Brent returns. Likewise, Bitcoin return Granger causes the WTI return at this period.

Table 3-9: Granger causality effect from cryptocurrencies towards Gold, WTI, and Brent (p-values)

	Tether	Bitcoin	Ether	Ripple	Litecoin	Bitcoin Cash	EOS	Chainlink	Cardano	Monero
Pre-COVID-19 (2019)										
Gold	0.979	0.213	0.349	0.330	0.967	0.213	0.605	0.597	0.354	0.351
WTI	0.303	0.357	0.395	0.862	0.868	0.703	0.771	0.335	0.497	0.799
Brent	0.472	0.084*	0.181	0.963	0.48	0.406	0.918	0.059*	0.543	0.419
During COVID-19 (2020)										
Gold	**0.011**	**0.015**	**0.003**	0.391	0.173	0.646	0.593	0.968	0.970	**0.027**
WTI	0.963	0.074*	0.419	0.927	0.508	0.264	0.474	0.511	0.629	0.113
Brent	0.077*	**0.040**	0.052*	0.538	0.523	0.488	0.373	0.453	0.181	0.140

Values in bold are significant at the 5% level and values with (*) are significant at the 10% level.

It is evident from our empirical studies that Tether, Bitcoin, Ether, and Monero returns have a significant causal effect on Gold market return, at a 5% level, during the COVID-19 period. Referring to Table 3-7, it can be concluded that the Granger causality relationship between Bitcoin-Gold and Monero-Gold markets are bi-directional. A summary of discovered significant long-term and short-term relationships between the cryptocurrency markets and Gold, and Crude oil markets is presented in Table 3-10. Finally, Table 3-11 summarizes our findings about significant Granger causal effects between the cryptocurrency markets and Gold, WTI, and Brent crude oil markets.

Table 3-10: Summary of significant long-term and short-term relationships

	Pre-COVID-19 (2019)		During COVID-19 (2020)	
	Gold	Crude Oil	Gold	Crude Oil
Tether	Long-term	-	Long-term	Long-term
Bitcoin	-	-	Short-term	Short-term
Ether	-	-	-	-
Ripple	-	-	-	-
Litecoin	-	-	Short-term	-
Bitcoin Cash	-	Short-term	-	Long-term
EOS	-	-	-	-
Chainlink	-	-	Long-term	-
Cardano	-	-	-	-
Monero	-	-	Short-term	-

Table 3-11: Summary of significant Granger causality effects

	Pre-COVID-19 (2019)			During COVID-19 (2020)		
	Gold	WTI	Brent	Gold	WTI	Brent
Tether	-	-	-	Tether GC Gold		Tether GC Brent*
Bitcoin	-	-	-	Bidirectional GC	Bitcoin GC WTI*	Bitcoin GC Brent
Ether	-	-	-	Ether GC Gold	-	Ether GC Brent*
Ripple	-	-	-	-	-	-
Litecoin	Gold GC Litecoin*	-	-	Gold GC Litecoin	-	-
Bitcoin Cash	-	WTI GC Bitcoin Cash*	Brent GC Bitcoin Cash*	-	-	-
EOS	-	-	-	-	-	-
Chainlink	-	-	Chainlink GC Brent*	-	-	-
Cardano	-	-	-	-	-	-
Monero	-	-	-	Bidirectional GC	-	-

This table shows the significance at the 5% level. Rows with * are significant at the 10% level.

Robustness of empirical findings

The robustness of our findings was assessed through several methods to ensure the reliability and validity of the results. Firstly, the analysis was conducted using different statistical models, including cointegration tests, VAR, VECM, and ARDL models, and Granger causality tests which provided consistent and converging outcomes. This cross-validation approach strengthened the confidence in the findings, as they were not reliant on a single model. Additionally, to enhance the generalizability of the

findings, this study employed a large and diverse dataset, encompassing ten cryptocurrency markets, two crude oil markets, and gold market in a sample of 494 daily instances covering the pre-COVID-19 and during COVID-19 periods. The inclusion of this comprehensive dataset allowed for a more comprehensive analysis, reducing potential biases and increasing the reliability of our findings.

3.5 Conclusion

Numerous studies have highlighted the lower volatility of gold markets compared to high-frequency traded markets such as stocks and digital currencies (Klein et al., 2018; Dyhrberg, 2016; Maghyereh & Abdoh, 2022). By employing various statistical and econometrics models, this study investigated the connectedness between ten most traded cryptocurrencies and gold and crude oil markets. Our findings suggest that, during the COVID-19 pandemic, Gold acts as a superior safe haven asset compared to WTI, and Brent crude oil for minimizing the risk of cryptocurrency investments. Similarly, the significant long-term negative relationships between Bitcoin Cash and WTI, as well as Bitcoin Cash and Brent crude oil markets during the COVID-19 period, indicate that crude oil can be a safe haven for Bitcoin Cash investments. Prior to the COVID-19 pandemic, we did not find any significant short-term connectedness between any of the cryptocurrency markets and Crude Oil and Gold markets. However, during the COVID-19 pandemic, lagged WTI and Brent crude oil returns demonstrate a significant relationship with Bitcoin returns. Overall, our findings suggest that the emergence of the COVID-19 pandemic has strengthened the safe haven effect of Gold and Crude Oil markets for portfolios primarily invested in Bitcoin, Litecoin, Monero, Chainlink, and Bitcoin Cash. However, the magnitude of the effect of Gold on these markets is more prominent than the effect of Crude Oil. We did not find any significant connection between Gold and Crude Oil markets with Ether, Ripple, EOS, and Cardano, at a 5% level of significance, during any of the study periods. Moreover, Granger causality tests are conducted on cryptocurrency, Gold, and WTI, and Brent crude oil to determine the direction of causal relationships between these markets. Our empirical results show that the direction of significant causal relations is predominantly from cryptocurrencies toward Gold or crude oil markets. However, during the COVID-19 pandemic, Gold significantly Granger caused Bitcoin, Litecoin, and Monero. Results from this study shed some light on the availability of the safe haven effect of gold and crude oil markets for cryptocurrency markets prior to and during the financial crisis due to the COVID-19 pandemic. This will help hedge fund managers and investors in digital currencies balance their risk exposures and maximize their returns in the event of a financial crisis. More specifically, the results indicate that incorporating Gold as a safe haven asset in Bitcoin, Litecoin, Monero, Bitcoin cash, Chainlink, EOS, and Cardano portfolios can provide risk mitigation benefits, particularly during periods of market turbulence such as the COVID-19 pandemic.

Gold provides diversification benefits in investment portfolios due to its low correlation with stocks, bonds, and other financial assets. This diversification helps reduce overall portfolio risk and enhances stability during turbulent market conditions. Unlike crude oil, which is consumed and has a finite lifespan as an asset, gold retains its value over extended periods. Its enduring appeal as a wealth-preserving asset makes it attractive for long-term investment strategies, particularly in

uncertain economic environments. In contrast, crude oil's price volatility and its dependence on global economic growth, geopolitical factors, and supply-demand dynamics make it less suitable as a safe haven asset. While oil can provide opportunities for profit during periods of economic expansion and rising demand, its price fluctuations can be unpredictable and subject to sudden shifts based on geopolitical tensions or supply disruptions.

Therefore, from an economic perspective, gold's stability, historical role as a store of value, liquidity, and perceived safety during economic turmoil make it a preferred safe haven asset compared to crude oil, which is more closely tied to industrial demand, geopolitical factors, and global economic conditions. The findings of this study indicate that policymakers, investors, and portfolio managers could enhance portfolio resilience by diversifying their cryptocurrency holdings to include gold. Additionally, this study shows that the safe haven characteristics of both Gold and Crude oil markets have improved during the COVID-19 pandemic compared to the period before the pandemic. This suggests that market participants should take a long-term perspective when evaluating the safe haven potential of these assets.

Chapter 4

Deep Learning Approaches for Forecasting the Crude Oil and Precious Metals Prices

4.1 Introduction and Literature Review

Non-renewable commodities that are usually mined in certain countries can strongly impact their economies, policies, currencies, and international or political issues. Energy and precious metals markets, among other commodities, are not only critical indicators of economic health but also crucial determinants for financial planning and decision-making. Given the multifaceted nature of these markets, forecasting the trajectories of these commodities is crucial in financial markets, serving as an essential tool for investors, policymakers, and analysts. For investors, the ability to anticipate price movements in crude oil and precious metals provides a strategic advantage in optimizing portfolio performance and risk management. A comprehensive understanding of potential price fluctuations allows investors to make informed decisions, allocate resources optimally, and, ultimately, enhance their overall financial returns (Bhowmik & Wang, 2020). Policymakers, on the other hand, rely on accurate market forecasts to formulate effective economic policies and mitigate the potential impact of market volatility on national economies. Fluctuations in crude oil prices, for instance, can have cascading effects on inflation, trade balances, and overall economic stability (Uzo-Peters et al., 2018; Xiuzhen et al., 2022; Periwal, 2023). Similarly, the prices of precious metals are often indicative of broader economic sentiments and can influence monetary policies and international trade relationships.

In this context, accurate forecasting of financial markets is a critical guide in determining economic policies. Consequently, researchers have dedicated their efforts to developing and improving models that capture the intrinsic behavior and dynamics of financial market time series. The prediction methods used in these studies are generally comprised of statistical or econometrics, machine learning, and deep learning methods. Recently, several forecasting modeling approaches have been applied to crude oil and precious metal commodities. For instance, Szarek et al. (2020) propose a new stochastic distribution, skewed Student's t-distribution, for silver, copper, and gold time series

estimation, which accounts for the time-dependent parameters and the non-Gaussian behavior of time series data. Drachal (2022) employs Bayesian Symbolic Regression method to address variable uncertainty in monthly crude oil price forecasting. A recent review paper (Mohamed & Messaadia, 2023) highlights that artificial neural networks (ANN) and support vector machines (SVMs) are the most popular Artificial Intelligence techniques to forecast the crude oil price

By virtue of ongoing improvements in Natural Language Processing (NLP) tasks, some recent studies have incorporated news text and google trends features as inputs to their forecasting models (X. Li et al., 2019; Salisu et al., 2020; Tang et al., 2020; Bai et al., 2022; Kertlly de Medeiros et al., 2022; Fang et al., 2023). These approaches leverage the valuable information contained in textual data to enhance the accuracy of predictions.

Considering the non-linear, non-stationary, noisy, and heteroscedastic structure of crude oil and precious metals markets, capturing their behavior precisely remains significantly challenging and leads to difficulties in forecasting (Dutta et al., 2019). Traditional statistical forecasting models like vector autoregressive (VAR), ARIMA, ARDL, etc., encounter challenges in achieving robust performance in forecasting tasks. This is primarily due to their reliance on assumptions about the normality and stationarity of price data, assumptions that frequently do not align with the characteristics of many commodity market time series data. As a result, the inclination of recent studies is towards using machine-learning and deep-learning models, which excel in handling non-linear data and do not rely on the normality assumption for accurate price predictions. In the literature, three main types of deep neural networks are used for sequence modeling, and they can be applied for time series forecasting (Lim & Zohren, 2021): (i). recurrent neural networks (RNNs) and their variants such as long short-term memory (LSTM) (Hochreiter & Schmidhuber, 1997) and gated recurrent units (GRUs) (Cho et al., 2014); (ii). Convolutional Neural Networks (CNN) (Lecun et al., 1998) and their recent variant, temporal convolutional networks (TCN) (Lea et al., 2016) , and (iii). Transformer (Vaswani et al., 2017) and its variants (Devlin et al., 2018; P. He et al., 2020; Y. Liu et al., 2019).

Considering the importance of gold price forecasting, several studies utilized statistical, machine learning and deep learning models. Alameer et al. (2019) use a Multi-Layer Perceptron model with a whale optimization algorithm (WOA) for gold next-month price forecasting. Using lagged gold prices, gold demand, and treasury bills rates as predictors, Madziwa et al. (2022) employ an ARDL model to forecast annual gold price. In another study, P. Zhang & Ci (2020) use U.S. Consumer Price Index (CPI), Crude oil price, exchange rate, Dow Jones Industrial Price Index in a deep belief network to predict monthly gold price. Risse (2019) predicts gold excess returns to risk free rate of return by a SVR model. SVR finds the non-linear relationship in the data by mapping a linear function into a high dimensional feature space. Tree-based ensemble models have demonstrated promising performance in forecasting gold prices. Yuan (2023) leverages XGBoost (T. Chen & Guestrin, 2016) and LightGBM (Ke et al., 2017) models for gold and bitcoin price forecasting. Additionally, there is an increasing trend of using deep learning methods for gold price prediction.

For instance, using the association rules and LSTM mode, Boongasame et al. (2022) predict the gold price. Vidal & Kristjanpoller (2020) develop a hybrid of convolutional neural networks and long short-term memory models (CNN-LSTM) which incorporates historical log-return series and time series data in an image format to predict the volatility of gold spot prices. Likewise, deep learning models are used for crude oil price forecasting in various studies. Orojo et al. (2019) employ a Multi-recurrent Network (MRN) to forecast a one-month ahead WTI crude oil price. Lin et al. (2022) forecast crude oil futures prices by utilizing a BiLSTM-Attention-CNN model with Wavelet transform. Swamy & Lagesh (2023) explore the effectiveness of investor sentiments from Twitter in predicting the daily gold price by a wavelet analysis method and unveil a strong correlation between Twitter sentiments and gold price.

However, the literature on forecasting other precious metal markets is relatively limited. Sroka (2022) utilizes block bootstrap methods to forecast daily silver prices, while Salisu et al. (2020) test the impact of Google Trends on forecasting the prices of four precious metal markets using an ARDL model. Y. Zhang et al. (2022) introduce a new objective function to forecast commodity markets including the silver price. To the best of our knowledge, there exists no precedent study to forecast the silver price by machine learning and deep learning models. We will attempt to fill this void in literature.

Some studies achieve improved forecasting performances by developing ensemble models. Zhao et al. (2017) combine the advantages of stacked denoising autoencoders (SDAE) and bootstrap aggregation (bagging) techniques to model the nonlinear and complex relationships of oil price factors. J. Wang et al. (2020) propose an ensemble of five linear and non-linear submodels to produce the prediction intervals of crude oil spot price. S. Zhang et al. (2021) developed an ensemble deep-learning model for electricity price series prediction. Jiang et al. (2022) combine a decomposition-ensemble approach, optimized by the seagull algorithm, with a sentiment analysis to forecast crude oil future prices. Su et al. (2022) propose a hybrid forecasting model using support vector machines (SVM), Extreme learning machines (ELM), XGBoost, and LSTM models to prediction of the crude oil futures series. Sun et al. (2022) propose a secondary decomposition-reconstruction-ensemble approach for crude oil price forecasting.

The gradient boosting methods are powerful predicting models for many tasks. Borisov et al. (2021) compare the performance of tree-based ensembles such as XGBoost, LightGBM, and CatBoost (Prokhorenkova et al., 2018) with some deep learning models, including but not limited to multilayer perceptron (MLP), regularization learning networks, Neural oblivious decision ensembles, and transformers. They assert that machine learning tree-based models outperform deep learning models in several prediction tasks with tabular data. However, their study does not include the deep learning models for sequential data and is silent about forecasting the financial market price. To address this shortfall, in the current study, we will use tree-based ensemble models such as Random Forest and LightGBM in comparison with twelve deep learning models and two other machine learning models (KNN and SVR) to forecast daily crude oil and precious metals market prices.

Recently Temporal Convolutional Networks (Lea et al., 2016) draw more attention from scholars and are applied to a wide range of time series data. For instance, Lara-Benítez et al. (2020) utilize a TCN model to forecast the electricity demand and price in Spain. In environmental milieu, Yan et al. (2020) predict the El Niño-Southern Oscillation (ENSO), which is an index to measure the earth's climate variability, by applying an ensemble empirical mode decomposition-temporal convolutional network (EEMD-TCN) model. This model shows improved prediction performance compared to a LSTM model.

Considering temporal patterns in predicting a time series data is a significant challenge for many models. Some recent works tried to introduce learnable time representations to account for the temporal patterns in sequential data (Xu et al., 2019; Xu et al., 2021; Y. Li et al., 2017). Among these studies, Kazemi et al. (2019) introduced the Time2Vector method to represent sequential data as periodic and non-periodic vectors that can capture complex temporal patterns in data. M. Yang et al. (2021) improve the performance of an attention neural network for nonintrusive load monitoring by applying the Time2Vector method. In the current study, we will apply Time2Vector embedding to input series and incorporate the resulted periodic and non-periodic features to several deep learning models to forecast the crude oil, gold, and silver prices.

Given the recent innovations in deep learning models for time series forecasting, the forecasting literature about crude oil and precious metals has not sufficiently utilized the deep learning models for price prediction. In this study, we attempt to fill this gap in forecasting literature by applying several deep learning and machine learning models to predict the daily closing price of crude oil, gold, and silver prices. Firstly, the time series data of daily spot prices of two prominent crude oils, WTI and Brent, and two precious metal markets, Gold, and Silver are gathered and normalized. Then, several input sequences are prepared by the sliding window method with four different window lengths. Next, the dataset is split into training, validation and test sets with a time-based splitting approach. Finally, a comprehensive set of sixteen forecasting models, consisting of twelve deep learning models, two baseline ensemble models, and two baseline machine learning models are implemented to predict the next-day market price. The deep learning models in the current study include long short-term memory (LSTM), bidirectional LSTM (BiLSTM), gated recurrent units (GRU), bidirectional GRU (BiGRU), time2vector BiLSTM (T2V-BiLSTM), time2vector BiGRU (T2V-BiGRU), convolutional neural networks (CNN), hybrid CNN-BiLSTM, hybrid CNN-BiGRU, temporal convolutional networks (TCN), hybrid TCN-BiLSTM, and hybrid TCN-BiGRU models. Two baseline ensemble models are Random Forest and LightGBM gradient boosting models, and two baseline machine learning models are support vector regression (SVR) and k-nearest neighborhood (KNN) models.

Each of the employed models has its own strengths and limitations. LSTM models are a type of RNNs that are popular for their ability to capture long-term dependencies, overcoming the gradient vanishing problem, and handle variable-length sequences. However, LSTMs can be computationally expensive and prone to overfitting, requiring regularization techniques (Y. Yu et al., 2019). GRU models, another type of RNN, have a simpler architecture, resulting in faster

training and inference times. However, they may have limitations in capturing complex patterns compared to LSTM models. Bidirectional models, such as BiLSTM or BiGRU, consider both forward and backward information, making them more robust to variations in input sequence order. However, they are computationally complex and require more memory resources (Khan et al., 2021). CNNs are effective at capturing local patterns and features within time series data. They learn filters to detect specific temporal patterns and are translation invariant, meaning they can detect patterns regardless of their position in the input sequence. CNNs, however, have limitations such as the requirement for fixed-length inputs, limited consideration of temporal ordering, and the ability to capture long-term dependencies. Hybrid CNN-LSTM models combine the strengths of both CNNs and LSTMs, capturing both spatial and temporal features. They are suitable for tasks that require capturing complex patterns in time series data. However, they can be less interpretable compared to standalone models (Gharghory, 2021). TCNs are designed to capture long-term dependencies efficiently. They use dilated convolutions to capture information from a wide range of past time steps. TCNs are adaptable to different time series lengths without the need for padding or truncation. However, they can be complex to design and tune, and they are sensitive to input scaling (Gopali et al., 2021). Ensemble machine learning models like Random Forest and LightGBM are also used in time series analysis. Random Forest combines multiple decision trees and offers high prediction accuracy and robustness to outliers. LightGBM is an efficient gradient boosting framework that handles large datasets effectively. Both models have their strengths in terms of accuracy and generalization, but they lack the ability to explicitly capture temporal dependencies (Ke et al., 2017). SVR is a flexible model that can capture linear and nonlinear relationships. It focuses on support vectors, which have the most influence on the model's decision boundary. SVR can handle high-dimensional datasets and complex relationships between variables. However, SVR's performance depends on selecting appropriate hyperparameters, and it does not explicitly model temporal dependencies. KNN is an instance-based algorithm that makes predictions based on the similarity of training instances. It requires no training phase but suffers from the curse of dimensionality and lacks the ability to capture temporal dependencies.

Our paper compares the forecasting performance of these models by mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE) error functions. Mainly, the objectives of this paper are to answer the following questions through empirical experiments: (1) What is the best deep learning model that can reliably and precisely predict the crude oil, gold, and silver spot prices? (2) Supposing the response to the first question, whether there exists a particular model that outperforms other models for both crude oil and precious metals prices? (3) Which input sequence length is more informative for each market’s price prediction? (4) Are hybrid models effective in crude oil, gold, and silver spot price forecasting? (5) what conclusions about the properties of each deep learning model can be drawn in the context of crude oil and precious metals time series forecasting?

A summary of the literature on crude oil and precious metal forecasting is presented in Table 4-1.

Table 4-1: Literature review of crude oil and precious metal forecasting

Method type	Author(s)	Method(s)	Features		Data			
			Financial market	Google trend or sentiment	WTI	Brent	Gold	Silver
Statistical and Econometric methods	L. T. Zhao et al. (2018)	VTFM	ü			ü		
	Szarek et al. (2020)	SGT	ü				ü	ü
	Drachal (2022)	BSR	ü		ü	ü		
	Kertlly de Medeiros et al. (2022)	MIDAS	ü	ü		ü		
	Salisu et al. (2020)	ARDL	ü	ü			ü	ü
	Tang et al. (2020)	MEMD	ü	ü		ü		
	Madziwa et al. (2022)	ARDL	ü				ü	
	Swamy & Lagesh (2023)	wavelet analysis method	ü	ü			ü	
	Sroka (2022)	block bootstrap-ARIMA	ü					ü
Machine Learning methods	Zhao et al. (2017)	Ensemble (SDAE-bagging)	ü		ü			
	Bai et al. (2022)	AdaBoost.RT	ü	ü	ü			
	Risse (2019)	SVR	ü				ü	
	Yuan (2023)	XGBoost & LightGBM	ü				ü	
	J. Wang et al. (2020)	EPPA	ü			ü		
	Su et al. (2022)	ensemble	ü		OPEC crude oil			
	Sun et al. (2022)	ensemble	ü		ü	ü		
Deep Learning methods	X. Li et al. (2019)	CNN	ü	ü	ü			
	Fang et al. (2023)	FineBERT-VMD-Att-BiGRU	ü	ü	ü			
	Fang et al. (2023)	ISBN-EMD-FNN	ü			ü		
	Liang et al. (2023)	Deepreinforcement learning	ü		ü	ü		
	Alameer et al. (2019)	MLP	ü				ü	
	P. Zhang & Ci (2020)	Deep belief network	ü				ü	
	Boongasame et al. (2022)	LSTM	ü				ü	
	Vidal & Kristjanpoller (2020)	CNN-LSTM	ü				ü	
	Orojo et al. (2019)	MRN	ü		ü			
	Lin et al. (2022)	BiLSTM-Attention-CNN	ü		ü	ü		

In this chapter, the following research questions are answered through empirical experiments.

1. What is the best deep learning model that can reliably and precisely predict the crude oil, gold, and silver spot prices?
2. Supposing the response to the first question, whether there exists a particular model that outperforms other models for both crude oil and precious metals prices?
3. Which input sequence length is more informative for each market's price prediction?
4. Are hybrid models effective in crude oil, gold, and silver spot price forecasting?
5. What conclusions about the properties of each deep learning model can be drawn in the context of crude oil and precious metals time series forecasting?

This study makes the following important contributions to the existing literature on the forecasting of commodity markets prices:

- Considering that there is limited literature on the use of deep learning models to forecast the price of commodity markets, this study implements and compares various types of the state-of-the-art deep learning models for crude oil and precious metal spot price forecasting. Hence, our study will encompass a wide range of forecasting results that provides comprehensive insights for crude oil, gold, and silver market players and investors.
- Most of the studies on precious metals, only focus on gold price predictions. However, in this study we forecast the price of both gold and silver to maintain a more general understanding of the precious metal markets.
- To the best of our knowledge, this study is the first in forecasting literature that applies the TCN model, Time2Vector embedding module, and the hybrid TCN-BiLSTM, and TCN-BiGRU models to forecast the spot price of WTI, Brent, Gold, and Silver time series.

4.2 Methodology

4.2.1 LSTM and BiLSTM

LSTM and BiLSTM are structural variants of RNN models that can remember the important information from the time series sequences (Y. Lin et al., 2022). Particularly, BiLSTM concatenates two LSTM layers in opposite direction. The interior structure of a common LSTM cell is shown in Figure 4-1-a. An LSTM unit consists of an input gate, a forget gate, and an output gate. These gates facilitate the information flow and help the cell to forget unnecessary information. Firstly, the forgetting gate decides what information from the inputs and previous hidden states to discard. Secondly, the input gate decides what information from the inputs and previous cell states to keep and updates the cell state. Finally, the output gate obtains the output h_t by multiplying the o_t of the input information processed by the sigmoid activation function and the cell state vector transformed by the tanh activation function. The equations of a forward pass in an LSTM unit are as follows:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{4-1}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{4-2}$$

$$c'_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{4-3}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot c'_t \tag{4-4}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{4-5}$$

$$h_t = o_t \odot \tanh(c_t)\ , \tag{4-6}$$

where, $x_t \in \mathbb{R}^d$ is the input vector, $h_t \in \mathbb{R}^h$ hidden state vector, f_t forget gate vector, i_t input gate vector, o_t output gate vector, c'_t temporary cell state vector, $c_t \in \mathbb{R}^h$ cell state vector, $W \in \mathbb{R}^{h \times d}, U \in \mathbb{R}^{h \times h}, b \in \mathbb{R}^h$ parameter matrices and vectors.

In a BiLSTM model, h_t from opposite directions are concatenated to construct the bidirectional hidden state. The formula of bidirectional h_t are as following:

$$\vec{h}_t = LSTM\left(x_t, \vec{h}_{t-1}\right) \tag{4-7}$$

$$\overleftarrow{h}_t = LSTM\left(x_t, \overleftarrow{h}_{t+1}\right) \tag{4-8}$$

$$h_t = \left[\vec{h}_t, \overleftarrow{h}_t\right] \tag{4-9}$$

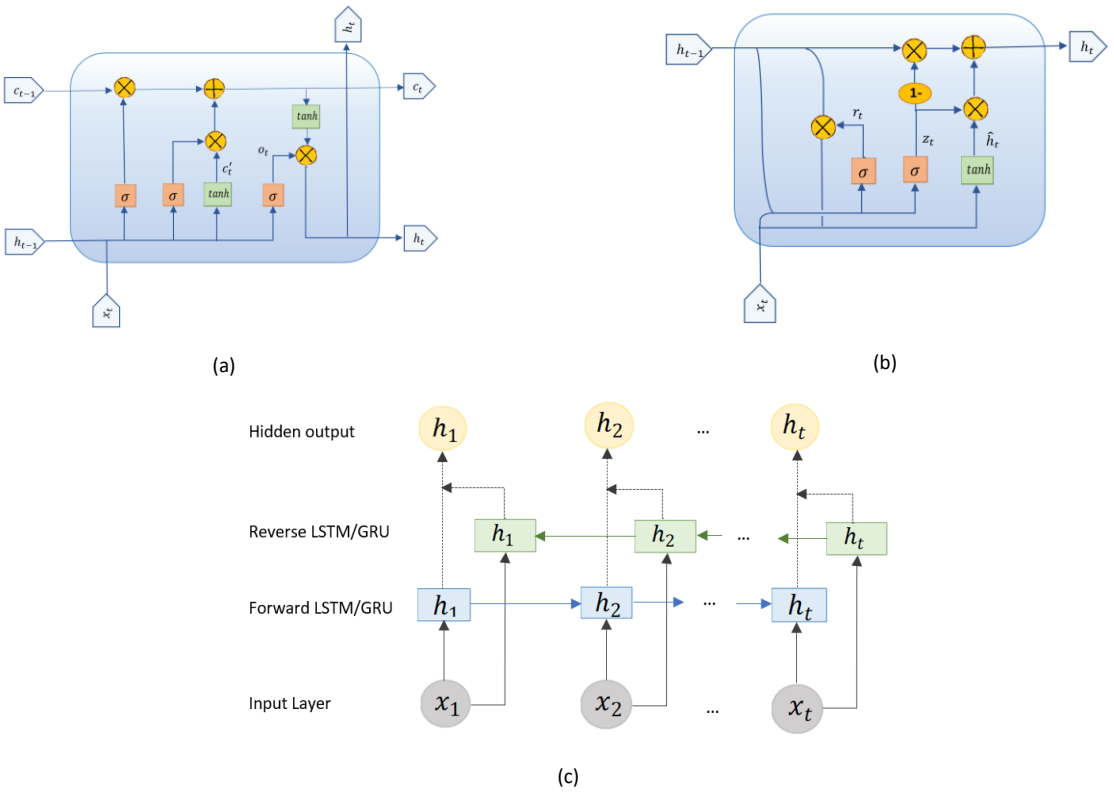


Figure 4-1: (a) LSTM internal cell structure, (b) GRU internal cell structure, (c) A single layer BiLSTM or BiGRU model.

4.2.2 GRU and BiGRU

Like the LSTM, gated recurrent unit (GRU) is a variant of RNN cells that can forget the insignificant information and help the model to utilize longer sequences of data. GRU has fewer parameters than LSTM, as it eliminates the output gate.

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (4\text{-}10)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (4\text{-}11)$$

$$\hat{h}_t = \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \quad (4\text{-}12)$$

$$h_t = z_t \odot \hat{h}_t + (1 - z_t) \odot h_{t-1} \quad (4\text{-}13)$$

where, $x_t \in \mathbb{R}^d$ is the input vector, $h_t \in \mathbb{R}^h$ hidden state vector, z_t forget gate vector, r_t reset gate vector, $\hat{h}_t$ candidate activation vector, $W \in \mathbb{R}^{h \times d}, U \in \mathbb{R}^{h \times h}, b \in \mathbb{R}^h$ parameter matrices and vectors, and σ is the sigmoid activation function. For certain sequential datasets, GRUs have outperform LSTM models (Chung et al., 2014; Gruber & Jockisch, 2020). The internal structure of a GRU cell is depicted in Figure 4-1-b.

For a bidirectional GRU model, hidden state vectors from two opposite direction are concatenated as follows:

$$\vec{h}_t = GRU(x_t, \vec{h}_{t-1}) \quad (4\text{-}14)$$

$$\overleftarrow{h}_t = GRU(x_t, \overleftarrow{h}_{t+1}) \quad (4\text{-}15)$$

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (4\text{-}16)$$

Figure 4-1-c shows the architecture of a single layer bi-directional LSTM (BiLSTM) or bi-directional GRU (BiGRU) model.

4.2.3 CNN

Convolutional neural network is a kind of feedforward neural network model proposed by Lecun et al. (1998). CNNs are very popular in computer vision applications such as Facial recognition systems, Object Localization, Object Detection, Semantic Segmentation, etc. CNNS are effective at capturing local patterns and features within a time series. The convolutional layers learn filters that can detect specific temporal patterns. This makes CNNs well-suited for capturing local dependencies and short-term patterns in time series data. CNNs are inherently translation invariant, meaning they can detect patterns regardless of their position in the input sequence. This property is useful for time series analysis since the same patterns may occur at different time steps. The local perception and weight sharing of CNN can greatly reduce the number of parameters, thus improving the efficiency of model learning (W. Lu et al., 2020). They, however, suffer from limitations such as requirement for fixed-length inputs, lack of considering the temporal ordering, and limited ability to catch long-term temporal dependencies.

The architecture of this model is generally constructed from two layers, namely convolution layer and pooling layer. The convolution layer will extract the useful features from the input series by applying several convolution kernels on inputs, as indicated in Eq. 4-17), which will downsample the input for final forecasting. Then, a pooling layer will apply to the output of convolution layer to further reduce the dimensionality of the model.

$$l_t = \sigma(x_t * k_t + b_t) \quad 4\text{-}17)$$

where l_t is the output of convolution layer, σ is the activation function, $x_t \in \mathbb{R}^d$ is the input vector, $k_t \in \mathbb{R}^d$ is the parameter vector of the convolution kernel, and b_t is the bias term.

4.2.4 TCN

The intrinsic weaknesses of CNN, including the fixed-size inputs and mismatched input and output dimensions, restrict its application in time-series forecasting. The Temporal Convolutional Networks (Lea et al., 2016) are variants of the CNN models that, employ casual convolutions and dilations to predict sequential data with temporality and large receptive fields. Causal means there is no information leakage from future to past and the receptive field means the set of sample elements of the original input that affects a specific element of the output. By setting a proper dilated factor and kernel size, a TCN model can show a full coverage of input history. A simple convolution is only able to look-back at a fixed timing window, whereas TCN uses dilated convolutions to achieve a large receptive field with fewer convolutional layers. TCNs capture long-term patterns using a hierarchy of temporal convolutional filters and in that manner, they tend to outperform Bidirectional LSTM models and are over a magnitude faster to train (J. Yan et al., 2020). TCN was first developed for action detection in video data settings to account for both spatial and temporal features of input (Lea et al., 2016). However, recently TCNs draw more attention from scholars and are applied to a wide range of time series data. Figure 4-2 shows a general representation of our TCN model with dilated causal convolutions. The architecture of this model consists of the following:

Dilated convolution layer: The dilated convolution architecture modifies Kronecker-factored convolutional filters, it allows to achieve larger receptive field with fewer parameters and fewer layers (Zhou et al., 2015). For a sequence of $x_t \in \mathbb{R}^d$ and a filter $f: \{0, \dots, k-1\} \rightarrow \mathbb{R}$, the dilated convolution operation $*_D$ on entries s of the sequence is defined as follows:

$$\mathrm{F}(s) = (x_t *_D f)(s) = \sum_{i=1}^{k-1} f(i).\, x_{s-D.i} \quad 4\text{-}18)$$

where, D is the dilation factor, k is the filter size, and $s - D.i$ assures that only past data are convoluted. The output of dilated causal convolution layer is transformed by a tanh function.

Dropout layer: a dropout layer with the probability of 0.2 is applied after each dilated convolution layer to regularize the model and eliminate the overfitting problem.

Residual block: We have used a stack of two dilated causal convolution layers together, and the results from the final convolution are added back to the inputs to obtain the outputs of the block. The residual connection avoids the vanishing and/or exploding gradient problem in deep learning models.

Fully connected layer: the output of residual block is then inputted into a fully connected layer to predict the next day price.

In Figure 4-2, the TCN model has a stack of two layers, a residual connection, and a fully connected layer. Each layer in the stack has a dilated causal convolution, tanh activation function, and a dropout for regularization. The dilation factors for the dilated convolution layer are $D = 1, 2, 4$ and a filter size of $k = 2$. It is evident that when $D = 1$, the dilated convolution becomes a basic convolution.

In recurrent-type neural networks, operations apply sequentially while in a TCN model all sequences are convolved at the same time in each dilated convolutional layer; hence, the training of TCN is much faster than LSTM or GRU models (Lea et al., 2016).

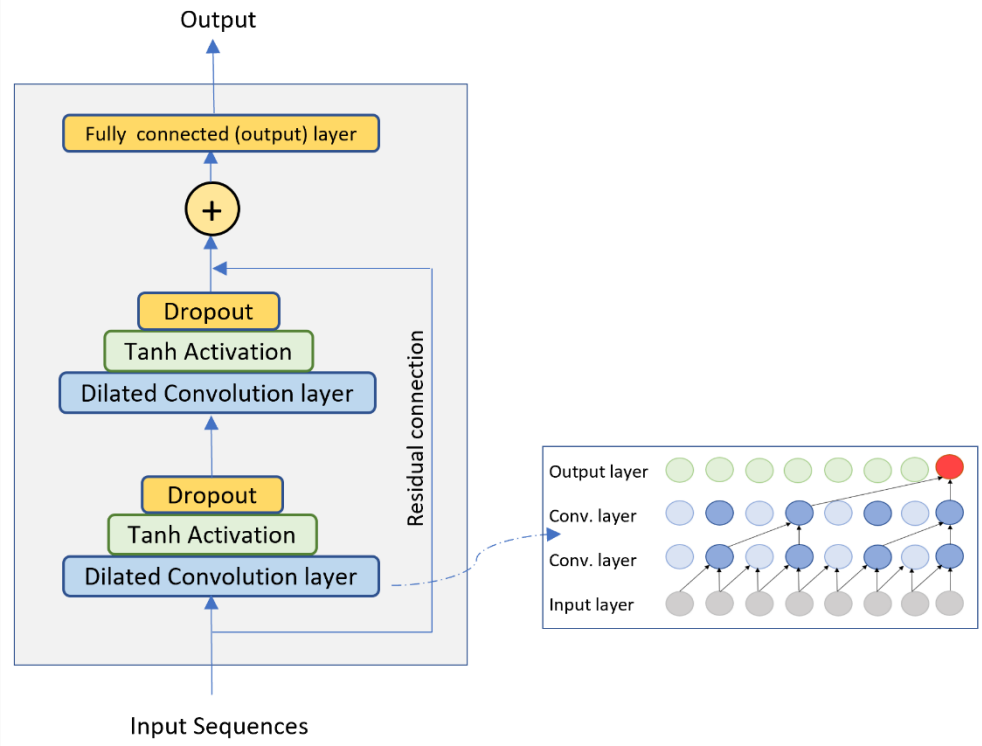


Figure 4-2: (left) The architecture of a TCN model with a stack of two dilated causal convolutional layers and a residual connection. (right) a dilated causal convolution layer with dilated factors D = {1, 2, 4} and kernel size k = 2.

4.2.5 Time2Vector (T2V-BiLSTM and T2V-BiGRU)

Time series input can be considered as a sequence where, rather than being identically and independently distributed (i.i.d), a dependency across time exists among the sample data. For this,

it is important to account for time features while developing a time series forecasting model. Vector embedding has been formerly successfully utilized in many natural language processing tasks (Pennington et al., 2014; Mikolov et al., 2013; Almeida & Xexéo, 2019). Similarly, Time2Vector (Kazemi et al., 2019) is a learnable vector embedding for time which can be easily combined with many deep learning models. Time2Vector is a decomposition technique that encodes a temporal signal into a set of periodic and non-periodic patterns, allowing the model to better understand and learn from the time-dependent patterns. Time2Vector eliminates the need for explicit feature engineering when dealing with time-related features. By incorporating temporal information in a meaningful way, Time2Vector can improve the performance of time series models.

For a given scalar notion of time τ, Time2Vec of τ is a vector of size k + 1 defined as follows:

$$T2V(\tau)[i] = \begin{cases} w_i\tau + b_i, & if\ i = 0. \\ \mathcal{F}(w_i\tau + b_i), & if\ 1 \leq i \leq k. \end{cases} \quad 4\text{-}19)$$

where $T2V(\tau)[i]$ is the i[th] element of $T2V(\tau)$, $\mathcal{F}$ is a periodic activation function, and w and b are learnable weight and bias parameters. Following the suggested activation function in the original T2V paper (Kazemi et al., 2019), we use sine function as $\mathcal{F}$. Time2Vector (T2V) assures that the learned periodic and nonperiodic time features will not be affected by the time scale (M. Yang et al., 2021).

To construct the T2V-BiLSTM and T2V-BiGRU models, firstly, the input sequences are transformed by Time2Vector embeddings, then the embedded input vectors are entered a single layer BiLSTM or BiGRU models, and finally the output is predicted through a fully connected layer. Figure 4-3 presents a schematic of T2V-BiLSTM or T2V-BiGRU model.

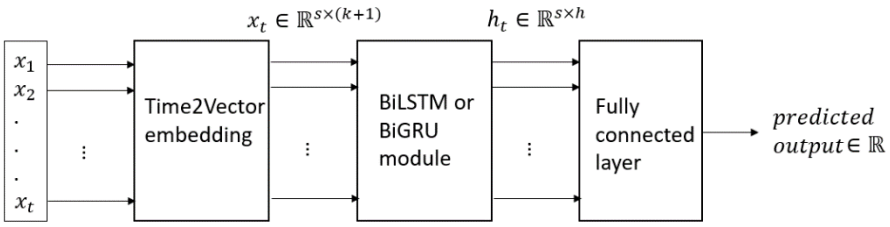


Figure 4-3: T2V-BiLSTM or T2V-BiGRU models. ***s*** is the input sequence length, ***k*** is the T2V output size, ***h*** is the recurrent hidden size.

Figure 4-4 summarizes the full process of data preprocessing, model training, and predicting for the test set in this study.

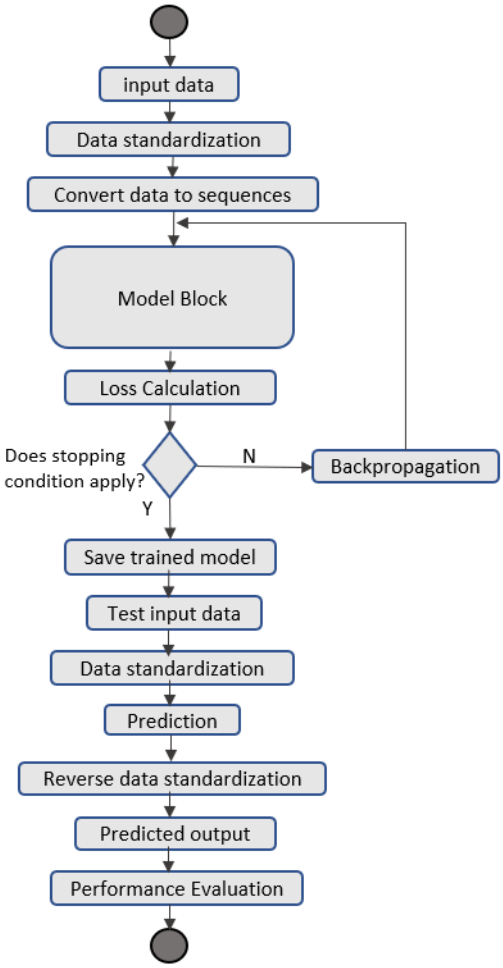


Figure 4-4: The price time-series forecasting flow chart

4.2.6 Hybrid Models

To verify the applicability of hybrid models in forecasting the crude oil, gold, and Silver daily prices, we have used CNN-BiLSTM, CNN-BiGRU, TCN-BiLSTM, and TCN-BiGRU models. CNNs in the initial layers of the hybrid model can learn low-level spatial features, such as local patterns, while the BiLSTM layers can learn high-level temporal dependencies. This hierarchical representation learning allows the model to capture both local and global dependencies in the time series data. CNNs and TCNs are well-suited for feature extraction from raw data, including time series. They can automatically learn relevant features and reduce the dimensionality of the input, which can be beneficial for downstream BiLSTM or BiGRU layers to learn more meaningful representations.

The explanation of each model structure is as follows:

CNN-BiLSTM and CNN-BiGRU models: Firstly, in the CNN module, a one-dimensional convolution layer is applied to input sequences and then a max pooling layer is applied to the output

of convolution layer to extract the important features. Next, the output of pooling layer is entered into a single layer BiLSTM or BiGRU module, and then the final output is predicted through a fully connected layer.

TCN-BiLSTM, and TCN-BiGRU models: Firstly, a TCN module will receive the input sequences, next the output of the TCN is introduced into a single layer BiLSTM or BiGRU module, and then the final output is predicted through a fully connected layer.

4.2.7 Ensemble and Machine Learning Models

Among the ensemble machine learning models, Random Forest and LightGBM, which is a type of gradient boosting technique, are utilized in this study. Random Forest generally provides high prediction accuracy due to the aggregation of multiple decision trees. It is less prone to overfitting compared to individual decision trees. By combining multiple trees and using techniques such as bagging and random feature selection, Random Forest reduces the variance and improves the model's ability to generalize. It is also robust to outliers and missing values. However, since Random Forest treats each data point independently and do not explicitly consider the temporal dependencies between consecutive observations in the time series, it lacks autocorrelation modeling. Random Forest is not well-suited for extrapolation, especially for long-term forecasts. It may struggle to capture and project future trends that extend beyond the range of the observed data. While Random Forest is generally robust to overfitting, it can still be sensitive to noisy data. If the dataset contains a substantial amount of noise or irrelevant features, Random Forest may still overfit to the noise, leading to degraded performance.

LightGBM is a powerful and efficient gradient boosting framework that offers excellent performance in various machine learning tasks. LightGBM is designed to be highly efficient and can handle large datasets with millions of instances and features. It uses a histogram-based algorithm to achieve faster training and prediction times compared to traditional gradient boosting implementations. The main advantage of LightGBM is low memory usage due to using a compact data structure for representing the dataset during training. Like other gradient boosting algorithms, LightGBM can be prone to overfitting if not properly regularized or tuned. LightGBM may struggle to capture complex feature interactions compared to deep learning models.

SVR is a machine learning model that can capture both linear and nonlinear relationships between variables and it can handle high-dimensional datasets and capture complex relationships between variables. The algorithm focuses on the support vectors, which are the data points that have the most influence on the model's decision boundary. Outliers have less impact on this model due to the use of a margin. SVR allows the use of different kernel functions, such as linear, polynomial, radial basis function (RBF), and sigmoid. This flexibility enables modeling various types of relationships between the input variables and the target variable. However, SVR performance is highly dependent on selecting appropriate hyperparameters, such as the kernel type, regularization parameter, and kernel-specific parameters. Training an SVR model can be computationally expensive, especially

when dealing with large datasets or complex kernel functions. For time series datasets, SVR lacks to account for the temporal dependencies among observations.

KNN is an instance-based, non-parametric algorithm that uses different distance metrics, such as Euclidean distance, Manhattan distance, or cosine similarity to make predictions. KNN does not explicitly learn a model from the training data. Instead, it stores the entire training dataset and uses it during the prediction phase. This eliminates the need for a time-consuming training phase. As the number of training instances increases, the algorithm's prediction time can be significant since it requires calculating distances to all training samples. Some limitations of KNN models are curse of dimensionality, sensitivity to the scale of features, intensive memory requirement, time-consuming predictions with large datasets, and lack of capturing temporal dependencies.

4.2.8 Evaluation Criteria

To evaluate the prediction performance, this paper adopts the following three metrics to calculate the forecasting error: mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE). MAE is a measure of the magnitude of difference between two continuous variables and treats all errors equally without emphasizing outliers. MAE is particularly useful for a clear understanding of the average magnitude of errors in our predictions. MAPE is a scaleless error value that measures the relative forecasting error and provides insights into how well the model predicts in percentage terms. RMSE represents the standard deviation of the residual error between the predicted value and the observed value which is important when we want to penalize larger errors more significantly. The prediction performance of the models increases with the decrease of these error measures. The formula of the above evaluation criteria is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i| \quad 4\text{-}20)$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} \left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad 4\text{-}21)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad 4\text{-}22)$$

Where, n is the sample size, and y_i and $\hat{y}_i$ are the true value and the predicted value for sample i, respectively.

4.3 Empirical Analysis and Results

4.3.1 Data description and preprocessing

The daily closing prices of West Texas Intermediate (WTI) and Brent crude oil, Gold, and Silver are collected from 2000-01-04 to 2022-03-25. The original spot price data of WTI and Brent crude oil are derived from the U.S. Energy Information Administration (https://www.eia.gov), while the spot price of Gold and Silver are derived from KITCO (https://www.kitco.com). We have used data from the same trading days across all four markets to have identical sample size for all time-series.

To find the best hyperparameters and evaluate the models' real-world performances, it is essential to evaluate them on a separate validation set and a test set that represents future unseen data. Splitting the time series datasets is challenging due to temporal dependencies, and the existence of seasonality and trends. If we split the data randomly, it breaks the temporal order, and the model may be trained on future data, leading to data leakage and overfitting. Moreover, if the training set does not capture the full range of seasonality or fails to include representative trend patterns, the model's ability to generalize to unseen data may be compromised. It is crucial to ensure that the training set contains consecutive past observations to predict future observations, includes multiple seasonal cycles, and captures the underlying trends adequately. To address these challenges, time-based splitting and rolling window approaches can be used in time series analysis. In time-based splitting, we split the data based on a specific date or time point, ensuring that the training set only contains past observations, and the test set contains future observations, and in rolling window approach, a sliding window is used to create samples in training, validation, and test sets, where each sample includes past observations, and the corresponding future target observation. Thus, for each market, the whole dataset is split in three parts: 65% training data (from 2000-01-04 to 2014-06-15), 25% validation data (from 2014-06-16 to 2020-01-02), and 10% test data (from 2020-01-03 to 2022-03-25). It is worth mentioning that the test data period includes the financial crisis due to the COVID-19 pandemic and the sharp decline in crude oil prices in April 2020. Therefore, test data includes highly volatile price data that make the forecasting even more challenging.

Since deep learning models are sensitive to the scale of data, we have normalized each dataset into [0,1] interval to limit the effect of noise, speed up the updating of neural network parameters and enhance the training performance of the model. The formula to standardize data is as follows:

$$x_t' = \frac{x_t - \min(x_t)}{\max(x_t) - \min(x_t)} \tag{4-23}$$

where, x_t and x_t' denote the data before and after standardization, respectively. Table 4-2 presents a summary of descriptive statistics and statistical tests for WTI and Brent crude oil, Gold, and Silver in the whole sample. The total sample size for all markets is 5426. All four market spot prices show the significant characteristics of skewness, while WTI, Brent, and Gold also represent significant leptokurtic properties at a 5% significance level. Besides, the significant Jarque-Bera test statistics at a 1% significance level show the WTI, Brent, Gold, and Silver price time-series do not comply with the normal distribution. Hence, these markets can be treated as non-stationary signals.

Table 4-2: Descriptive statistics

Index	WTI	Brent	Gold	Silver
Count	5426	5426	5426	5426
Mean	61.47	64.27	1020.98	15.73
Standard deviation	25.79	29.33	510.85	8.51
Min	-36.98	9.12	255.95	4.07
Max	145.31	143.95	2067.15	48.7
Skewness[1]	0.39*	0.41*	-0.07**	0.61*
Kurtosis[2]	-0.598*	-0.817*	-1.249*	0.084
Jarque-Bera[3]	218.903*	301.375*	357.452*	342.924*

[1] null hypothesis is that the series are skewed. [2] null hypothesis is that the series show normal kurtosis. [3] null hypothesis is that the series are normally distributed. * , ** denote the rejection of the null hypothesis at the 1% and 5% significance level, respectively.

For these forecasting tasks, $x_t = \{x_1, x_2, \dots, x_s\}$ is the input vector, where x_i is the price data at day i and s is the sequence length (sliding window length), and $y_t = \{x_{s+1}\}$ is the target. Before sending series into the model, we have created inputs of different sequences. In this study, we train sixteen deep learning and machine learning models with four different sliding window lengths of 5, 30, 60, and 90 days to predict the next day WTI, Brent, Gold, and Silver prices. We have considered 5 as a relatively short sliding window length and 30, 60, and 90 as relatively long sliding window lengths to capture any seasonality or trend in data. We will compare various deep learning and machine learning models to find how they perform on forecasting commodity price time-series with longer input sequences.

4.3.2 Empirical Results

Crude oil and precious metals are highly important commodities in financial markets. The purpose of this study is to forecast the daily price of WTI and Brent crude oil, Gold, and Silver through deep learning models and compare the prediction performance of deep learning models with Random Forest, LightGBM, SVR, and KNN models as the baseline machine learning models. Hence, our results will suggest the best deep learning model to forecast crude oil, gold, and silver daily prices. We will experiment the performance of all models across four sliding window lengths of 5, 30, 60, 90 days to suggest the suitable input length for a superior performance with each model. The deep learning models used in this study are LSTM, BiLSTM, GRU, BiGRU, T2V-BiLSTM, T2V-BiGRU, CNN, CNN-BiLSTM, CNN-BiGRU, TCN, TCN-BiLSTM, TCN-BiGRU models.

We have used Grid Search on the validation dataset to tune and select the optimal hyperparameters of each model. The common hyperparameters among all models are number of epochs, batch size, dropout rate, and learning rate which are equal to 50, 32, 0.2, and 0.001, respectively. Table 4-3 presents the selected hyperparameters of four best performing models in this study. Due to the large scale of the study and space limitations we only presented the selected hyperparameters of BiGRU,

T2-BiGRU, TCN, and TCN-BiGRU models for each market. The hyperparameters of other models are available upon request from the corresponding author.

Table 4-3: Selected hyperparameters of models

models	markets	Hidden size				T2V size				Num. filters				Kernel size			
		5	30	60	90	5	30	60	90	5	30	60	90	5	30	60	90
BiGRU	WTI	64	128	128	128												
	Brent	64	128	128	128												
	Gold	128	128	128	128												
	Silver	64	128	128	128												
T2V-BiGRU	WTI	64	128	128	128	32	90	32	128								
	Brent	64	128	150	128	64	128	128	128								
	Gold	128	150	128	128	90	64	128	128								
	Silver	256	256	256	128	128	128	128	32								
TCN	WTI									32	16	16	16	2	4	2	4
	Brent									32	16	16	16	2	4	2	4
	Gold									64	64	64	32	2	16	4	4
	Silver									32	64	64	64	2	4	2	4
TCN-BiGRU	WTI	64	64	64	64					32	16	16	16	2	4	2	4
	Brent	64	64	64	64					32	16	16	16	2	4	2	4
	Gold	128	128	128	128					64	64	64	32	2	16	16	4
	Silver	64	128	64	128					32	64	64	32	2	16	4	4

Other common hyperparameters among all models and all four markets are: epochs=50, batch size=32, dropout rate=0.2, and initial learning rate=0.001.

After each training step, the weights of models are updated by Adam optimizer with a scheduled learning rate (lr) as follows:

$$lr = \begin{cases} lr_0 & if\ epochs < 5 \\ lr\ * e^{(-0.1)} & otherwise \end{cases} \tag{4-24}$$

the initial learning rate (lr_0) is 0.001 which is applied on epoch one through epoch five and then is exponentially decreased for each epoch after epoch five. In this study, the models are trained to minimize the mean squared error (MSE) loss function. The objective function of training process is as follows:

$$\text{Objective function } = \text{ Minimize } MSE = \text{ Minimize } \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{4-25}$$

Where $\hat{y}_i$ is the predicted price and y_i is the true target price for sample i.

Overfitting in financial market price forecasting experiments can lead to misleading and unreliable results. Overfitting occurs when a model is too complex, and it is able to capture the noise in the data, rather than the underlying patterns. The consequences of overfitting in financial market price forecasting can be severe. Traders who rely on the overfitted model may make poor investment

decisions, leading to significant losses. In addition, the overfitted model may be highly sensitive to changes in the market, making it difficult to use in real-world situations. To mitigate the risk of overfitting in crude oil and precious metals market price forecasting, in this study, techniques such as cross-validation, dropout, early stopping, and pruning (for Random Forest and LightGBM) are employed. Cross-validation involves partitioning the data into training and validation sets and evaluating the model on the validation set to assess its generalization performance. Model regularization in this study is achieved through dropout layer in models' architectures and early stopping after ten epochs during the training. Early stopping will end the training process if there is no improvement in validation error. To further assure the robustness of forecasting results, all reported errors and predicted values are the average outputs from ten runs of each model.

All deep learning models are implemented by using Tensorflow Keras and machine learning models are created using Sklearn. The experiments are conducted by using Python 3.8 and run on a computing system with a 70 W Tesla T4 NVIDIA-SMI GPU, CUDA version 11.2, and 16GB RAM.

4.3.2.1 WTI price forecasting

To show the computational performance of our deep learning models for WTI next-day spot price forecasting, we draw the forecasting performance of LSTM, BiLSTM, GRU, BiGRU, T2V-BiLSTM, T2V-BiGRU, CNN, CNN-BiLSTM, CNN-BiGRU, TCN, TCN-BiLSTM, TCN-BiGRU models and compared them with the baseline models, i.e., Random Forest, LightGBM, KNN and SVR models. Each model was executed 10 times to reduce randomness and improve robustness in results. Table 4-4 presents the MAE, MAPE, and RMSE values for the forecasted next-day WTI prices in the test dataset across all models. It is observed that among the evaluated models and considering two out of three performance criteria, the TCN model consistently achieves the lowest MAE and MAPE for WTI price forecasting across all input sliding window sizes. However, when considering the RMSE metric, the BiGRU model outperforms other models for input sequences of length 5 and 30. On the other hand, for input sequences of length 60 and 90, the TCN-BiGRU and T2V-BiGRU models demonstrate superior performance, respectively. In addition to the superior prediction performance, the forecasting error of TCN model is not significantly affected by the input sequence length as we get the MAE of 1.510, 1.455, 1.444, and 1.472 with sequence lengths of 5, 30, 60, and 90, respectively. Comparing this with other models, we can see that the performance of most of the models is more sensitive to the input sequence length. Using bidirectional models was proved effective in natural language processing tasks (Arbane et al., 2023; Y. Huang et al., 2023; G. Liu & Guo, 2019; Raza & Schwartz, 2023). However, little attention has been given to using these models for price time series forecasting. In this study, all three performance criteria from Table 4-4 show that bidirectional recurrent models such as BiLSTM and BiGRU perform better than unidirectional models such as LSTM and GRU for all sequence lengths. Bidirectional RNNs exploit the network memory used in processing the information from both backward and forward directions. So, interdependency among data samples is learned better compared to unidirectional models that

only use the forward direction information processing. Our findings are in compliance with the studies of M. Yang & Wang (2022) and Siami-Namini et al. (2019) which found BiLSTM model outperforming the LSTM model for time series prediction. Besides, it is evident from Table 4-4 that GRU-type models such as GRU, BiGRU, T2V-BiGRU, CNN-BiGRU, and TCN-BiGRU perform better than LSTM-type models such as LSTM, Bi LSTM, T2V-Bi LSTM, CNN-Bi LSTM, and TCN-Bi LSTM in WTI price forecasting.

To evaluate the effectiveness of Time2Vector embedding in WTI price forecasting, we compare the MAE, MAPE, and RMSE of BiLSTM and BiGRU models with T2V-BiLSTM and T2V-BiGRU models, respectively. Using the T2V input embedding, the MAE of BiLSTM and BiGRU models with input sequence 5 increases from 1.821 and 1.570 to 1.985 and 1.889, respectively. On the contrary, the MAE of BiLSTM and BiGRU models with input sequence 90 decreases from 1.904 and 1.699 to 1.670 and 1.523, respectively. It can be argued that Time2Vector embedding does not have an improving effect on forecasting with smaller input sequences, 5 and 30, while it has improved the WTI price forecasting performance for longer sequences of 60 and 90. To study the impact of using hybrid models such as CNN-BiLSTM and CNN-BiGRU we compare their performance with a single BiLSTM and BiGRU models, respectively. It can be seen that the combination of the CNN model with recurrent-type models has a detrimental effect on the forecasting performance of WTI prices, as evidenced by an increase in the MAE across all sequence lengths. The reason for this would be that CNN module downsamples the input sequence and some information that might be useful for BiLSTM or BiGRU models will be lost, hence resulting in higher forecasting errors. Similarly, a single TCN model outperforms the hybrid TCN-BiLSTM and TCN-BiGRU models. TCN model is able to see the full sequence in its receptive field and use the best temporal features to forecast WTI price. Therefore, combining it with a recurrent-type model will only increase the complexity of model and causes overfitting problem without any significant improvements in forecasting performance.

Upon examining the forecasting errors of ensemble tree-based models, i.e., Random Forest and LightGBM, it becomes apparent that Random Forest performs poorly in predicting WTI prices, whereas LightGBM demonstrates exceptional forecasting capabilities. The MAPE and RMSE values of LightGBM across sequence lengths of 5, 30, and 90 days are consistently the lowest among all sixteen forecasting models. Consequently, LightGBM can be considered on par with the TCN model as the top-performing method for WTI price forecasting. Moreover, it is worth noting that the performance of LightGBM exhibits a slight decline as the input sequence lengths increase. However, this decrease in performance is not significant, indicating that LightGBM is relatively insensitive to variations in input sequence length. On the other hand, using SVR and KNN models, it becomes apparent that the performance of conventional machine learning models tends to deteriorate as the input sequences grow larger. In contrast, deep learning models are less affected by larger input sequences, showcasing their robustness. For larger input sequences, all deep learning models outperform the SVR and KNN models. However, for smaller sequences, such as those with

a length of 5, it is observed that the KNN model performs better than the deep learning models, except for the BiGRU and TCN models. This discrepancy can be attributed to the fact that the data within each sequence serves as input features for the KNN model. As the sequence length increases, the KNN model faces greater challenges in identifying the nearest neighbors required for predicting the target price accurately.

Table 4-4: WTI price forecasting performance

Models	MAE				MAPE				RMSE			
	5	30	60	90	5	30	60	90	5	30	60	90
LSTM	1.950	1.872	1.875	1.873	7.190	8.005	8.605	8.284	3.787	3.719	3.720	3.720
GRU	1.893	1.787	1.765	1.631	8.300	6.092	5.672	4.965	3.682	3.584	3.580	3.539
BiLSTM	1.821	1.654	1.679	1.904	7.532	5.966	5.720	5.286	3.588	3.511	3.496	3.747
BiGRU	1.570	1.559	1.575	1.699	6.868	5.094	5.230	6.725	3.415	3.432	3.439	3.488
T2V-BiLSTM	1.985	2.196	1.668	1.670	5.594	5.033	6.274	6.204	3.770	3.986	3.461	3.458
T2V-BiGRU	1.889	1.606	1.511	1.523	5.107	5.987	3.853	5.060	3.691	3.451	3.392	3.405
CNN	1.887	2.068	3.201	3.201	7.277	8.213	8.957	9.357	4.000	6.064	5.085	5.071
CNN-BiLSTM	1.972	2.581	2.563	3.650	6.733	6.158	6.566	8.207	4.496	4.845	4.444	5.695
CNN-BiGRU	1.851	2.406	2.244	2.885	8.133	5.850	7.793	7.196	4.243	4.650	4.161	4.815
TCN	1.510	**1.455**	**1.444**	**1.472**	3.829	3.663	**3.530**	3.882	3.550	3.495	3.559	3.552
TCN-BiLSTM	2.063	2.354	1.704	2.494	5.374	5.566	4.301	5.377	3.857	4.267	3.419	4.456
TCN-BiGRU	1.787	2.015	1.672	1.911	5.437	5.471	7.023	5.033	3.520	3.872	**3.372**	3.720
Random Forest	3.209	3.151	1.887	1.868	7.209	7.131	4.307	4.286	4.869	4.801	3.574	3.569
LightGBM	**1.487**	1.551	1.554	1.558	**3.490**	**3.622**	3.624	**3.643**	**3.213**	**3.338**	3.392	**3.386**
SVR	2.231	4.199	3.861	5.586	5.591	9.224	8.560	16.12	4.947	5.733	5.687	7.821
KNN	1.772	3.358	5.049	6.559	4.075	7.599	10.68	12.83	3.440	5.586	7.792	10.03

To find the best sliding window length for each forecasting model, Figure 4-5 presents the RMSE of WTI next-day spot price forecasting models. RMSE is particularly selected for comparing models because it effectively highlights accurate predictions and mitigates the influence of large errors. Our experiments with WTI price forecasting show that using only recurrent-type models such as LSTM, GRU, BiLSTM, BiGRU, T2V-BiLSTM, and T2V-BiGRU, we get better prediction performance compared to using only CNN or a hybrid of CNN with Recurrent-type models such as CNN-BiLSTM, and CNN-BiGRU. It can be noted that recurrent-type models, are not very sensitive to the input sequence length and they even perform slightly better with relatively longer input sequences as longer sequences enable the model to learn more upward, downward, and complex patterns and generalize better in predicting the unseen data. However, since the CNN models are not able to memorize the important information from the past data points, the forecasting error of CNN-type models such as a single CNN, CNN-BiLSTM, and CNN-BiGRU increases by the input sequence length. As in general the RMSE of TCN-BiLSTM and TCN-BiGRU is smaller than the RMSE of CNN-BiLSTM and CNN-BiGRU models, it can be concluded that among the hybrid models, TCN module performs better than CNN module in extracting the important temporal features. From Figure

4-5, in general, input sequence of 60 days of lagged data points is better than other sliding window lengths such as 5, 30, or 90 days for WTI daily price forecasting. However, CNN, CNN-BiLSTM, and CNN-BiGRU models perform better with input sequence of 5 days compare to the other sequence lengths for WTI price prediction. Among the machine learning models, Ensemble tree-based models emerge as the leading models for forecasting WTI prices. Notably, the Random Forest model exhibits subpar performance with shorter input sequences. Conversely, LightGBM consistently performs well across all input sequences, showcasing its robust forecasting capabilities. In contrast, the forecasting performance of the SVR and KNN models deteriorates as the input sequence length increases. This suggests that these models struggle to effectively capture the complex patterns and relationships within longer sequences of data.

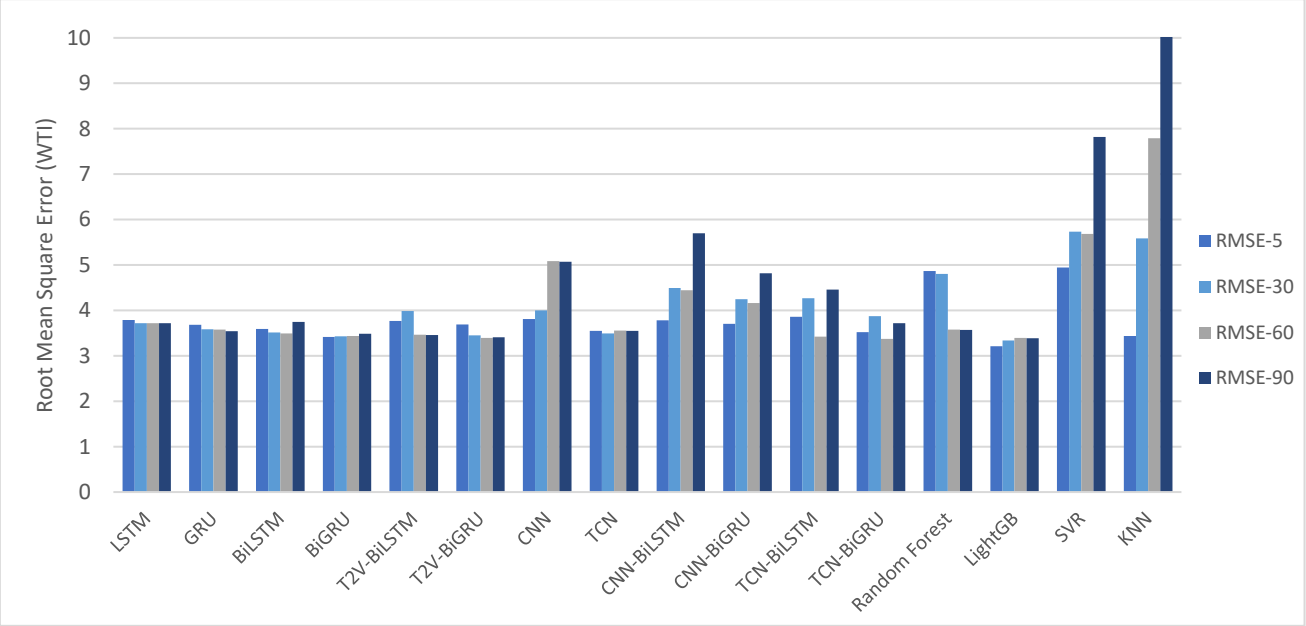


Figure 4-5: RMSE of WTI crude oil next-day price forecasting models

Our observations regarding WTI forecasting align with the outcomes presented in the study by Qin et al. (2023), where the GRU model demonstrated superior performance compared to random forest, SVR, and LSTM models, achieving a lower MAPE value. Similarly, our results corroborate the conclusions drawn by J. Yuan et al. (2023), highlighting that LightGBM exhibited significantly better performance than LSTM and SVR models.

Figure 4-6 compares the line chart of predicted WTI prices in the test dataset with the actual WTI price value in the period from 2020-01-03 to 2022-03-25. Upon looking at the predicted values at the end of April 2020, it becomes apparent that the TCN model surpasses the LightGBM model in accurately capturing sharp changes in WTI price. The TCN model demonstrates superior performance in detecting and predicting abrupt fluctuations in price, showcasing its ability to capture and respond to sudden market dynamics with greater precision than the LightGBM model.

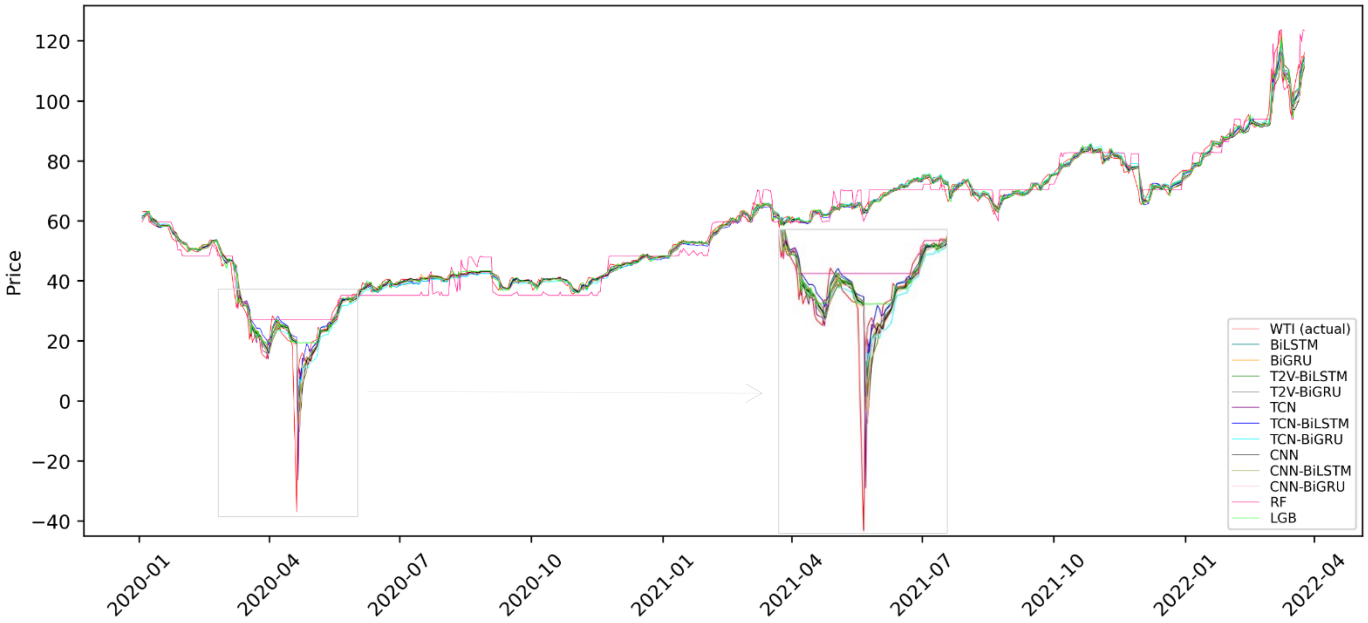


Figure 4-6: Comparison of WTI crude oil price forecasting models on the test dataset

4.3.2.2 Brent price forecasting

Table 4-5 shows the errors, MAE, MAPE, and RMSE, of our forecasting models for Brent next-day spot price forecasting. We compare the forecasting performance of LSTM, BiLSTM, GRU, BiGRU, T2V-BiLSTM, T2V-BiGRU, CNN, CNN-BiLSTM, CNN-BiGRU, TCN, TCN-BiLSTM, TCN-BiGRU models with the baseline models, Random Forest, LightGBM, KNN and SVR models. According to the lowest values of the MAE and RMSE measures for all input sequence lengths, 5, 30, 60, and 90, the TCN is the best performing model in predicting the Brent crude oil price in the test dataset. Considering the MAPE, although, for input sequences with 5 lagged data points the TCN model has the best Brent price prediction performance, for input sequences of length 30, 60 and 90, the T2V-BiGRU model outperforms other models. Additionally, the TCN model is not particularly sensitive to the input sequence length. The TCN achieves a robust and stable forecasting performance for all input sequence lengths as the MAE with sequence lengths of 5, 30, 60, and 90 are 1.295, 1.353, 1.315, and 1.301, respectively. Evidently, the performance of the majority of the other models exhibits higher sensitivity to changes in the input sequence length for Brent crude oil. For instance, the MAEs of the CNN model grows with increasing the sequence length as it gets the MAE of 1.542, 1.879, 2.818, and 5.194 with sequence lengths of 5, 30, 60, and 90, respectively. Similar to our findings for the WTI crude oil price forecasting, we found that, in general, BiLSTM and BiGRU models outperform the unidirectional LSTM and GRU models in forecasting the Brent crude oil price. By juxtaposing the MAE, MAPE, and RMSE of the GRU-type models such as GRU, BiGRU, T2V-BiGRU, CNN-BiGRU, and TCN-BiGRU with those of the LSTM-type models such as LSTM, BiLSTM, T2V-BiLSTM, CNN-BiLSTM, and TCN-BiLSTM, we found that a GRU unit is a more appropriate recurrent unit for Brent crude oil price forecasting.

The impact of Time2Vector embedding in Brent crude oil price forecasting is assessed through the comparison of MAE, MAPE, and RMSE of T2V-BiLSTM and T2V-BiGRU models with BiLSTM and BiGRU models, respectively. Table 4-5 shows that the T2V embedding is improving the forecasting performance of BiLSTM model for input sequences of 60 and 90, while it stimulates the performance of BiGRU model for input sequences of 30, 60, and 90. The results of Brent crude oil price forecasting confirms that T2V embedding has a favorable influence on forecasting with longer input sequences. For the hybrid models, our results suggest that combining the CNN model with recurrent-type models has an adverse effect on the performance of BiLSTM and BiGRU models for the Brent crude oil price forecasting. The same pattern is visible in comparing the forecasting performance of a single TCN model with the TCN-BiLSTM and TCN-BiGRU hybrid models in predicting the Brent daily prices. The TCN model outperforms the hybrid models.

By comparing the forecasting errors of Random Forest, LightGBN, SVR and KNN models with our deep learning models, we conclude that basically the forecasting performance of deep learning models is superior to that of machine learning models. However, it is worth noting that the ensemble LightGBM model stands as an exception, demonstrating remarkable performance as the second-best model among all sixteen models for forecasting Brent crude oil price across all input sequence lengths. This exceptional performance sets LightGBM apart from the other models, emphasizing its robustness and effectiveness in accurately predicting Brent crude oil price, regardless of the length of the input sequence. Yet, for the short sequence length of 5, the KNN performs better than the deep learning models, with the exception of the BiGRU, CNN, and TCN models.

Table 4-5: Brent price forecasting performance

Models	MAE				MAPE				RMSE			
	5	30	60	90	5	30	60	90	5	30	60	90
LSTM	1.856	1.806	1.789	1.821	3.689	3.563	3.509	3.630	2.860	2.826	2.800	2.830
GRU	1.760	1.721	1.436	1.412	3.343	3.277	2.752	2.715	2.760	2.670	2.289	2.257
BiLSTM	1.703	1.492	1.556	1.562	3.328	2.895	2.974	3.013	2.643	2.364	2.446	2.449
BiGRU	1.421	1.428	1.439	1.425	2.888	2.861	2.895	2.833	2.222	2.217	2.237	2.228
T2V-BiLSTM	1.893	1.503	1.417	1.403	3.738	2.922	2.788	2.683	2.922	2.371	2.211	2.200
T2V-BiGRU	1.721	1.386	1.372	1.403	3.345	**2.669**	**2.641**	**2.671**	2.671	2.189	2.171	2.204
CNN	1.542	1.879	2.818	5.194	2.653	3.679	5.773	5.150	2.459	2.934	4.335	4.032
CNN-BiLSTM	1.857	2.456	2.680	2.641	3.618	4.887	5.241	5.121	2.844	3.807	4.126	4.096
CNN-BiGRU	1.769	2.414	2.265	2.279	2.136	5.026	4.684	4.233	2.747	3.694	3.478	3.248
TCN	**1.295**	**1.353**	**1.315**	**1.301**	**2.637**	2.905	2.658	2.697	**2.035**	**2.059**	**2.052**	**2.045**
TCN-BiLSTM	2.063	2.320	1.653	1.879	4.062	4.620	3.215	3.592	3.171	3.658	2.467	2.923
TCN-BiGRU	1.768	2.035	1.363	2.061	3.440	4.061	2.639	4.121	2.746	3.236	2.129	3.217
Random Forest	1.989	2.681	2.717	2.667	3.811	5.425	5.475	5.275	2.909	3.585	3.565	3.501
LightGBM	1.375	1.426	1.452	1.478	2.719	2.910	2.963	3.063	2.130	2.174	2.190	2.210
SVR	2.130	2.787	4.480	4.530	4.685	5.192	9.343	8.417	3.284	4.102	6.284	6.448
KNN	1.594	3.224	5.440	6.421	3.197	6.737	10.72	12.94	2.483	4.984	8.379	10.14

To assure the robustness of models' performances, the average of errors in ten runs of the models are reported here.

Figure 4-7 represents the RMSE of the forecasting models that are implemented in this study to predict the next-day Brent crude oil price in the test dataset. Our results denote that the recurrent-type models such as LSTM, GRU, BiLSTM, BiGRU, T2V-BiLSTM, and T2V-BiGRU, outperform the CNN and hybrid models such as CNN-BiLSTM, CNN-BiGRU, TCN-BiLSTM, and TCN-BiGRU in term of Brent price forecasting. From Figure 4-7, we realize that, in general, the efficacity of recurrent-type models in predicting the Brent price enhances with relatively longer input sequences. However, the CNN and hybrid models do not perform well with longer input sequences. As the RMSE of TCN-BiLSTM and TCN-BiGRU are mainly lower than the RMSE of CNN-BiLSTM and CNN-BiGRU models, it can be inferred that the TCN module performs better than CNN module in extracting the important temporal features of Brent crude oil price. When examining the ensemble and conventional machine learning models, namely Random Forest, LightGBM, SVR, and KNN, it becomes evident that the optimal forecasting input sequence for Brent price prediction is 5 days. LightGBM model achieve a superior forecasting across all input sequences and thus is not significantly affected by the changes is the input sequence length. As a general observation, the forecasting performance of these baseline models declines as the input sequence length increases. This suggests that shorter input sequences provide more accurate and reliable predictions compared to longer sequences when utilizing these models for forecasting Brent price. Regardless of machine learning-type models, CNN, CNN-BiLSTM, and CNN-BiGRU models that perform better with shorter input sequences, our experiments suggest that, in general, the best input sequence length for Brent crude oil forecasting is 60 days of past data. Hence, the lowest RMSE values across most of the deep learning models in this study, are achieved for input sequence length of 60 for Brent crude oil price forecasting.

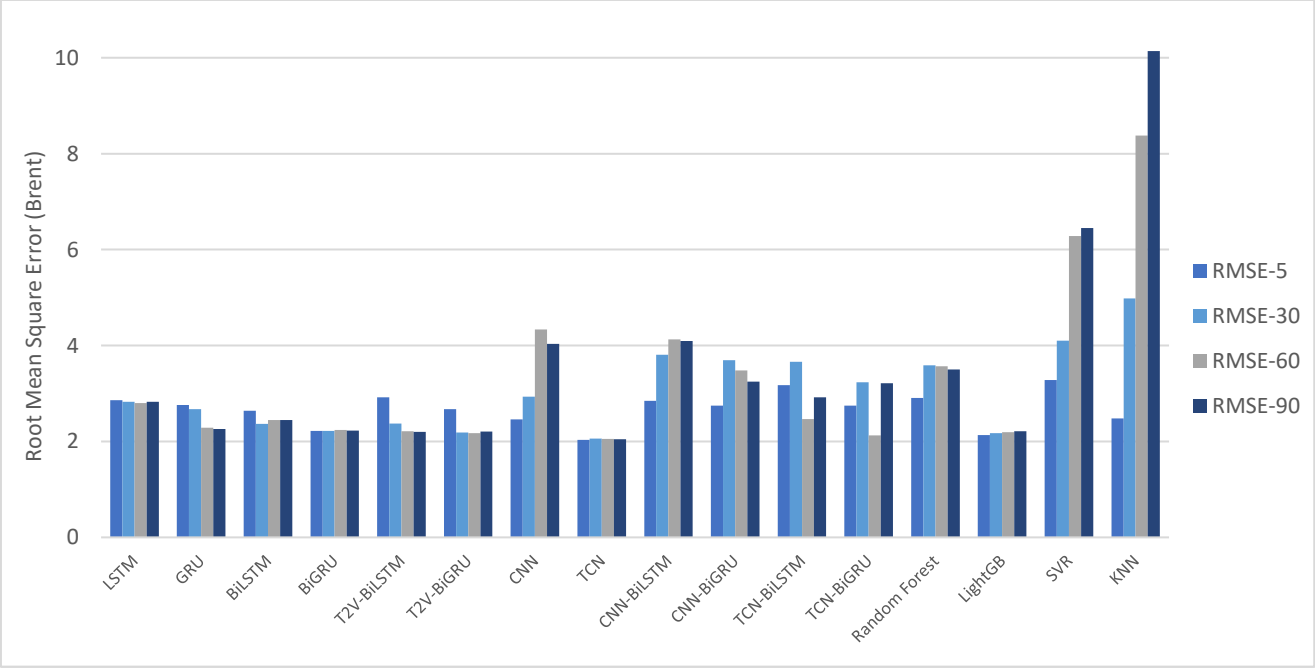


Figure 4-7: RMSE of Brent next-day price forecasting models

Our results validate the conclusions drawn by Y. Zhao et al. (2017), indicating that deep learning models outperform machine learning models, such as SVR, in forecasting crude oil prices.

Figure 4-8 compares the line chart of predicted Brent crude oil prices in the test dataset with the actual Brent price values in the period from 2020-01-03 to 2022-03-25. Upon analyzing the predicted value during the abrupt Brent price change periods, it becomes apparent that the TCN model outperform the LightGBM model in accurately capturing sharp changes in Brent price. Thus, TCN is a more reliable model for predicting the sudden changes in Brent price.

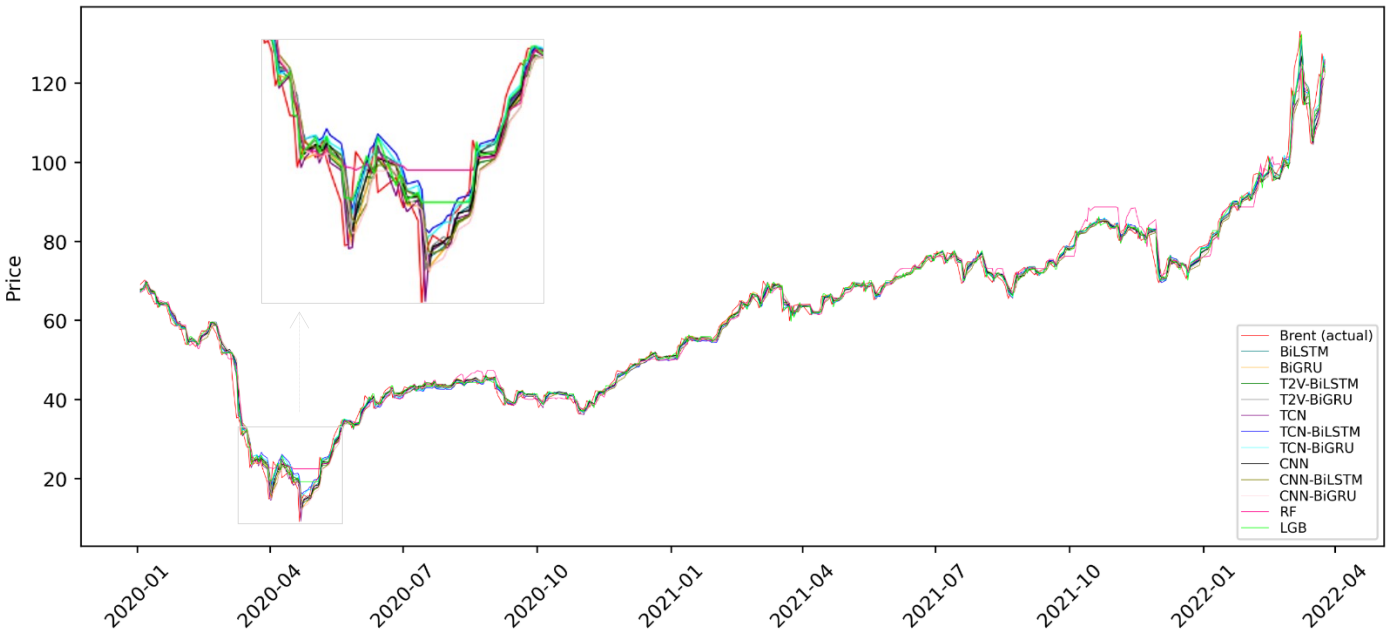


Figure 4-8*:* Comparison of Brent crude oil price forecasting models on the test dataset

4.3.2.3 Gold price forecasting

The forecasting errors of Gold price prediction with sixteen deep learning and machine learning models are presented in Table 4-6. Considering the resulted MAE, MAPE, and RMSE of the models, the TCN model has the best Gold price prediction performance for input sequences of 5 and 90 day. Meanwhile, for Gold price predictions with input sequences of 30 and 60, the BiGRU and GRU models show the superior performance among all models. Our results show that in most of the cases, the deep learning models performed remarkably better than the baseline Random Forest, LightGBM, SVR and KNN models in predicting the price of Gold market. Compared to the CNN-BiLSTM, TCN-BiLSTM, and TCN-BiGRU, the SVR model achieved lower MAE, MAPE, and RMSE values. The prediction with Gold price data show that bidirectional LSTM models perform better than the unidirectional LSTM models for all input sequences. In the meantime, the BiGRU model outperformed the GRU model exclusively for input sequences of 5, and 60 days. Comparing

the Gold price forecasting errors of the GRU-type models such as GRU, BiGRU, T2V-BiGRU, CNN-BiGRU, and TCN-BiGRU with those of the LSTM-type models such as LSTM, Bi LSTM, T2V-Bi LSTM, CNN-Bi LSTM, and TCN-Bi LSTM, we found that GRU-type models are more proper than the LSTM-type models for Gold price forecasting.

Our deep learning models were able to predict the gold price for the test data relatively well. In contrast to its performance in WTI and Brent price forecasting, the LightGBM model surprisingly did not exhibit strong generalization capabilities when predicting the Gold price during the test data period. Despite its success in other forecasting tasks, the LightGBM model failed to provide accurate and reliable predictions for Gold price, indicating that the underlying dynamics and patterns of Gold price data might differ significantly from those of WTI and Brent. Table 4-8 shows the coefficient of variation for the resulted MAEs of all forecasting models. The coefficient of variation is a scaleless value calculated via dividing the standard deviation of model MAEs through various input sequence lengths by the mean of those MAEs. Comparing the forecasting results of Gold market with the results of WTI and Brent crude oil markets, from Table 4-8, we notice that the models are more sensitive to the input sequence lengths of Gold market as the MAE forecasting error of each model vary markedly across the sequence lengths.

Table 4-6: Gold price forecasting performance

Models	MAE				MAPE				RMSE			
	5	30	60	90	5	30	60	90	5	30	60	90
LSTM	26.92	29.53	33.95	35.39	1.52	1.67	1.92	2.00	33.26	35.91	40.37	41.67
GRU	18.21	17.17	**21.50**	24.50	1.01	0.95	**1.20**	1.35	25.26	23.36	**28.57**	32.88
BiLSTM	22.33	26.17	32.76	36.39	1.25	1.48	1.85	2.05	29.26	33.66	42.27	42.41
BiGRU	19.07	**15.19**	27.64	22.15	1.05	**0.85**	1.56	1.22	26.13	**20.85**	35.12	28.84
T2V-BiLSTM	25.08	33.58	36.57	32.41	1.41	1.89	2.05	1.81	31.37	42.02	45.55	69.55
T2V-BiGRU	19.93	24.35	24.49	35.06	1.11	1.36	1.38	2.00	28.05	30.96	31.43	60.80
CNN	23.28	34.05	53.34	64.35	1.30	1.91	2.99	3.70	31.05	45.08	65.78	81.94
CNN-BiLSTM	30.39	68.59	68.62	66.56	1.71	3.98	4.01	3.84	36.97	79.67	85.07	77.20
CNN-BiGRU	24.44	36.16	46.48	42.36	1.37	2.05	2.67	2.31	31.48	47.93	58.03	53.65
TCN	**15.56**	17.21	27.86	**19.89**	**0.87**	0.96	1.57	**1.12**	**20.83**	22.58	33.73	**24.82**
TCN-BiLSTM	35.67	96.56	81.28	59.71	2.01	5.66	4.69	3.40	43.68	112.97	94.1	39.86
TCN-BiGRU	42.83	78.34	66.33	55.97	2.42	4.55	3.83	3.20	51.53	93.99	81.01	42.96
Random Forest	96.26	73.14	156.5	73.43	5.61	4.19	9.48	4.209	121.33	97.42	177.28	97.52
LightGBM	43.41	69.40	70.27	79.67	2.43	3.97	4.02	4.58	64.39	94.06	95.05	104.55
SVR	87.02	49.64	69.90	48.11	5.07	2.83	4.01	2.76	92.06	60.12	80.39	63.037
KNN	45.45	96.48	109.76	135.61	2.55	5.62	6.45	8.10	65.26	116.41	131.78	156.72

To assure the robustness of models' performances, the average of errors in ten runs of the models are reported here.

Figure 4-9 depicts the RMSE of our forecasting models to predict the Gold next-day price in the test dataset. It can be seen that the recurrent-type models such as LSTM, GRU, BiLSTM, BiGRU, T2V-BiLSTM, and T2V-BiGRU, in general, have lower RMSE values compared to the CNN and hybrid models such as CNN-BiLSTM, CNN-BiGRU, TCN-BiLSTM, and TCN-BiGRU. This result aligns with the research conducted by Z. He et al. (2019) on gold price prediction, demonstrating that a

hybrid CNN-LSTM model did not exhibit superior performance compared to the individual CNN or LSTM models.

We notice that, in general, shorter input sequence of 5 days price data is more useful in Gold price predictions with both deep leaning and machine learning models. Besides, in general, the Gold price forecasting performance deteriorates by increasing the input sequence length. The best prediction performance across all models and all sequences is achieve through the BiGRU model using 30 days of Gold price data. Based on the findings presented in Table 4-8, it is evident that LightGBM exhibits a higher coefficient of variation for MAE in Gold price forecasting compared to WTI and Brent crude oil. This indicates that LightGBM is considerably sensitive to changes in the input sequence length when predicting Gold price. The higher coefficient of variation suggests that the performance of LightGBM may vary significantly when the input sequence length is altered, underscoring the need for careful consideration and optimization of the input sequence length specifically for Gold price forecasting with LightGBM. Figure 4-10 compares the line chart of predicted Gold prices in the test dataset with the actual Gold price values in the period from 2020-01-03 to 2022-03-25. These results suggest that Random Forest and LightGBM models do not generalize well in forecasting the Gold price.

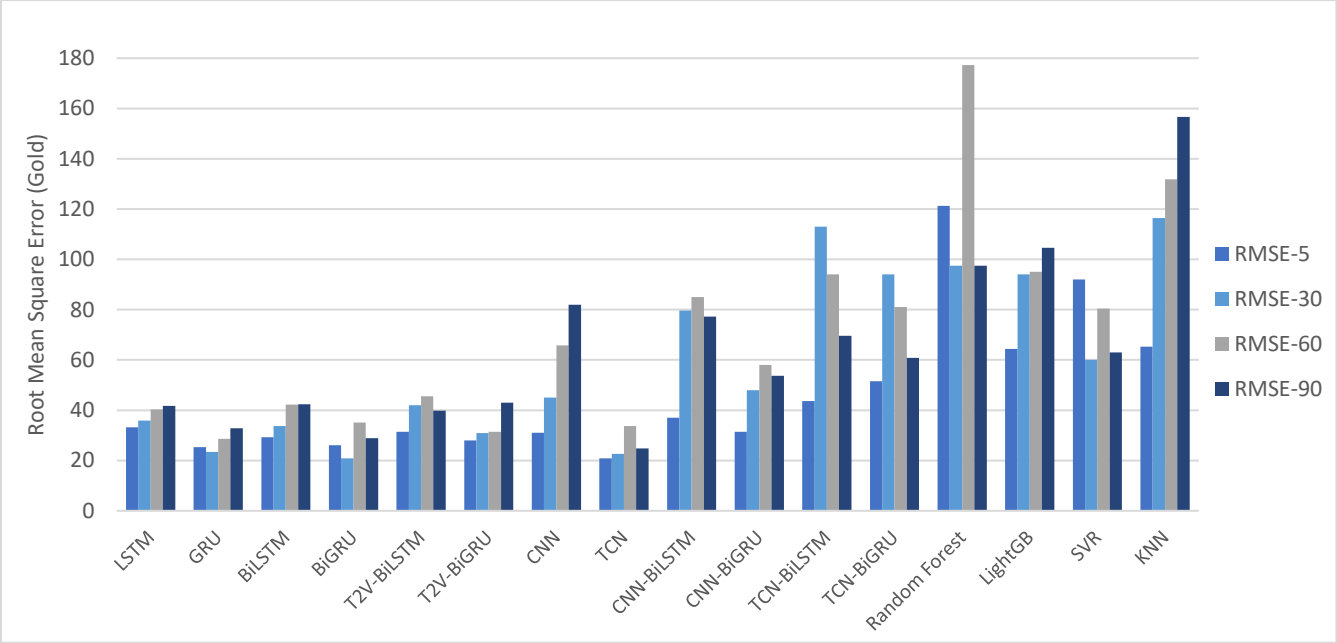


Figure 4-9: RMSE of Gold next-day price forecasting models

Comparing the performance of LightGBM and KNN models in predicting gold prices, our results demonstrate the superiority of LightGBM, a conclusion also supported by the study conducted by Z. Yuan (2023).

Figure 4-10: Comparison of Gold price forecasting models on the test dataset

4.3.2.4 Silver price forecasting

As a kind of precious metal, the daily spot price of Silver is forecasted through the deep learning models in this study and compared with the Random Forest, LightGBM, SVR and KNN forecasts. Table 4-7 shows the MAE, MAPE, and RMSE of Silver price predictions. The TCN model is the best performing model across all input sequence lengths to forecast the daily silver price as it scores the lowest MAE, MAPE, and RMSE amidst all models. Besides the TCN's superior ability in forecasting the Silver price, this model is the least susceptible to the input sequence length as the MAE coefficient of variation from Table 4-8 shows. The coefficient of MAE variation across all sequence lengths is 0.015 for the TCN model which is the lowest value amongst all models. Results of this study suggest that, except for the TCN-BiLSTM and TCN-BiGRU models with input sequence of 5 days, our deep learning models are superior to the SVR and KNN models in predicting the price of Silver market. For Silver price forecasting, providing the bidirectional information seems promising with the BiLSTM model as it reached lower MAE, MAPE, and RMSE values compared to the unidirectional LSTM. However, the bidirectional information did not improve the forecasting performance of the GRU model for Silver price prediction. Furthermore, the results from Table 4-7 propose that GRU-type models have a relatively better forecasting performance than the LSTM-type models, for Silver price prediction.

Using the ensemble (Random Forest, and LightGBM) or conventional (SVR, and KNN) machine learning models, only LightGBM could outperform some of the deep learning models, namely, CNN, CNN-BiLSTM, CNN-BiGRU, TCN-BiLSTM, and TCN-BiGRU in Silver price forecasting. LightGBM is found to be the best machine learning model in Silver price forecasting across all sequence lengths.

Upon comparing the MAE coefficient of variations between the Silver and Gold markets in Table 4-8, it can be concluded that the performance of our forecasting models is relatively less affected by changes in the input sequence length when predicting the Silver market. This suggests that the forecasting models exhibit greater stability and consistency in their predictions for the Silver market, regardless of variations in the input sequence length. In contrast to the Gold market, where the models show higher sensitivity to changes in the input sequence length, the Silver market demonstrates a more robust and reliable forecasting performance across different input sequence lengths.

Table 4-7: Silver price forecasting performance

Models	MAE				MAPE				RMSE			
	5	30	60	90	5	30	60	90	5	30	60	90
LSTM	0.490	0.463	0.497	0.491	2.186	2.086	2.239	2.221	0.721	0.694	0.741	0.710
GRU	0.440	0.394	0.381	0.385	1.967	1.753	1.667	1.684	0.658	0.590	0.549	0.556
BiLSTM	0.471	0.387	0.499	0.476	2.095	1.716	2.237	2.137	0.702	0.575	0.733	0.710
BiGRU	0.425	0.442	0.388	0.364	1.897	1.973	1.721	1.609	0.639	0.649	0.571	0.543
T2V-BiLSTM	0.478	0.589	0.491	0.434	2.138	2.666	2.206	1.933	0.717	0.850	0.730	0.653
T2V-BiGRU	0.472	0.548	0.462	0.370	2.125	2.483	2.065	1.641	0.703	0.794	0.679	0.553
CNN	0.551	0.730	0.752	0.768	2.454	3.273	3.400	3.468	0.786	1.030	1.090	1.140
CNN-BiLSTM	0.571	0.747	0.670	0.775	2.538	3.307	3.022	3.478	0.800	1.103	0.980	1.108
CNN-BiGRU	0.528	0.611	0.656	0.635	2.358	2.737	2.952	2.843	0.779	0.926	0.953	0.893
TCN	**0.355**	**0.349**	**0.346**	**0.357**	**1.573**	**1.547**	**1.525**	**1.575**	**0.529**	**0.520**	**0.513**	**0.525**
TCN-BiLSTM	0.647	0.942	0.746	0.603	2.881	4.269	3.345	2.668	0.960	1.327	1.013	0.867
TCN-BiGRU	0.714	0.928	0.553	0.467	3.186	4.155	2.476	2.086	1.058	1.307	0.780	0.651
Random Forest	1.831	0.831	0.848	2.476	8.076	3.562	3.646	11.30	2.149	1.059	1.075	2.764
LightGBM	0.509	0.475	0.486	0.475	2.154	2.047	2.070	2.061	0.710	0.665	0.684	0.646
SVR	0.602	1.164	1.077	1.130	2.656	5.006	4.609	5.034	0.775	1.321	1.359	1.513
KNN	0.667	1.170	1.847	2.686	2.805	5.078	8.022	11.47	0.913	1.501	2.295	3.300

To assure the robustness of models' performances, the average of errors in ten runs of the models are reported here.

Figure 4-11 presents the RMSE of our deep learning models to forecast the Silver next-day price in the test dataset. Similar to the results of WTI, Brent, and Gold markets, it can be observed that the Silver price forecasting error of the recurrent-type models such as LSTM, GRU, BiLSTM, BiGRU, T2V-BiLSTM, and T2V-BiGRU are in general lower than the forecasting error of the CNN and hybrid models such as CNN-BiLSTM, CNN-BiGRU, TCN-BiLSTM, and TCN-BiGRU. The best performing model to predict the Silver price is the TCN model which demonstrates a robust forecasting performance across all input sequence lengths. Our results show that the recurrent-type models generally perform better with longer input sequence of 90 days to predict the next-day Silver price. The best prediction performance across all models and all sequences is achieve through the TCN model using 60 days of past Silver price data. Moreover, in the hybrid models such as CNN-BiLSTm, CNN-BiGRU, TCN-BiLSTM, and TCN-BiGRU, the TCN module performs better than the CNN module in extracting the temporal features of Silver market price.

Figure 4-12 illustrates the line chart of the best predicted Silver prices in the test dataset with the actual Silver price values from 2020-01-03 to 2022-03-25. It shows that the TCN and Random Forest models are best and least generalizing models in Silver price forecasting.

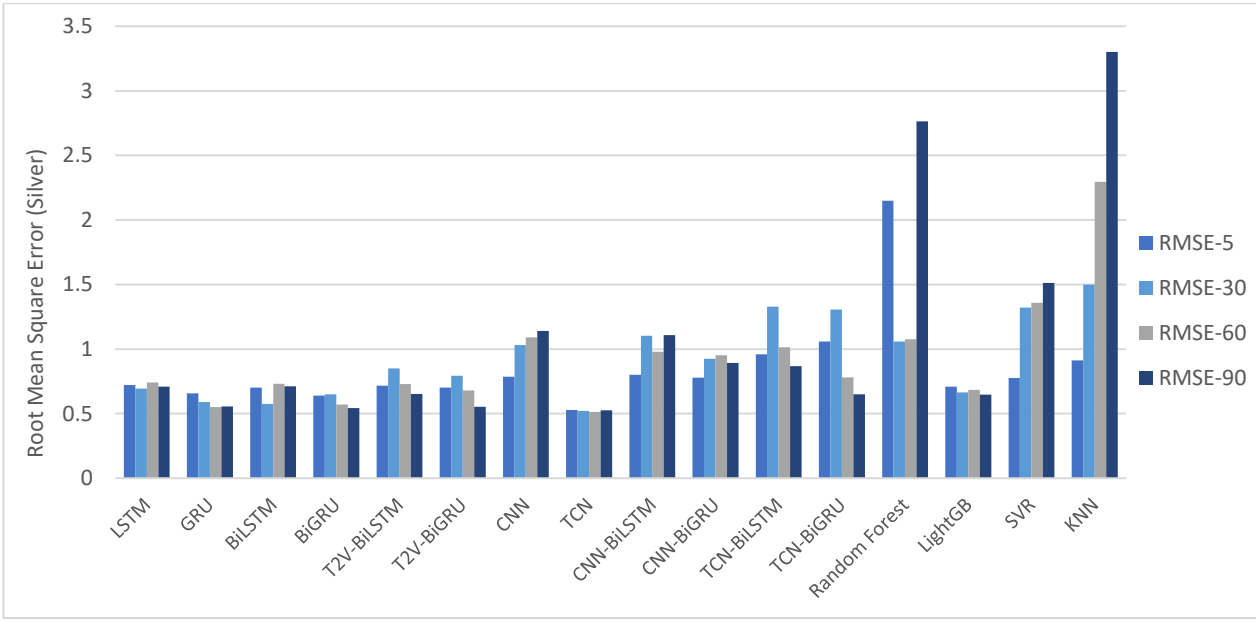


Figure 4-11: RMSE of Silver next-day price forecasting models

Using MAPE as the metric, our silver price prediction results surpass those of a comparable study conducted by Gono et al., (2023), which employed Random Forest and XGBoost methods. Our best MAPE for silver price prediction, 1.52%, significantly outperforms the best MAPE of 5.98% achieved in the study of Gono et al., (2023).

Figure 4-12: Comparison of Silver price forecasting models on the test dataset

Table 4-8: Coefficient of variation (CoV) for the MAE of forecasting models

	WTI	Brent	Gold	Silver
LSTM	0.020	0.016	**0.124**	0.031
GRU	0.061	0.116	0.164	0.068
BiLSTM	0.067	0.056	0.216	0.107
BiGRU	0.041	**0.005**	0.250	0.087
T2V-BiLSTM	0.137	0.148	0.153	0.131
T2V-BiGRU	0.108	0.114	0.248	0.158
CNN	0.274	0.577	0.423	0.144
CNN-BiLSTM	0.260	0.158	0.321	0.132
CNN-BiGRU	0.182	0.130	0.257	0.092
TCN	0.020	0.020	0.271	**0.015**
TCN-BiLSTM	0.162	0.143	0.388	0.205
TCN-BiGRU	0.081	0.179	0.248	0.305
Random Forest	0.258	0.121	0.341	0.465
LightGBM	**0.019**	0.027	0.201	0.028
SVR	0.301	0.301	0.251	0.229
KNN	0.429	0.452	0.339	0.476
H0: the mean CoV of Gold and Silver markets are equal (p-value = 0.0074)				
H0: the mean CoV of WTI and Brent markets are equal (p-value = 0.7142)				

The coefficient of variation is calculated by $\frac{standard\ deviation(MAE_i)}{mean(MAE_i)}$,

$i = s_5, s_{30}, s_{60}, s_{90}$. Models with values in bold are least sensitive to the input sequence lengths for each market's price predictions.

4.4 Conclusion

Crude oil, in particular WTI and Brent, perform a crucial role in the global financial markets and market economics. In recent years, the price of crude oil has been more vulnerable to geopolitical and macroeconomic factors. Thus, understanding the dynamics of crude oil markets seems inevitable. Besides, precious metals such as gold and silver are key commodities that are mined in particular countries which make the economics of these countries highly rely on precious metal markets. Moreover, gold is a substitute asset for stock markets and plays an indispensable role in financial investment portfolios. Therefore, developing an accurate forecasting model for crude oil, gold and silver price movements is a vital key for policymakers, business owners, investors, and other stakeholders to mobilize timely political movements, foresee the market trends and properly design their investment strategies to mitigate investment risks. In this study, we implement 12 deep learning models, namely, LSTM, BiLSTM, GRU, BiGRU, T2V-BiLSTM, T2V-BiGRU, CNN, CNN-BiLSTM, CNN-BiGRU, TCN, TCN-BiLSTM, and TCN-BiGRU, to forecast the WTI, Brent, Gold, and Silver markets prices, and compare their forecasting performance with four baseline models, namely, Random Forest, LightGBM, SVR and KNN models. For this, we use each market's historical price information and apply 4 different sliding window lengths of 5, 30, 60, and 90 days. To assess the forecasting power of each model, the MAE, MAPE, and RMSE evaluation metrics are employed. Specifically, we chose RMSE for comparing models because it effectively highlights accurate predictions and mitigates the influence of large errors, which could otherwise substantially impact returns in our markets. We compared the forecasting performance of these models across various input sequence lengths and found that the TCN model is the best performing model to forecast the price of Brent, Gold, and Silver. LightGBM exhibits comparable forecasting performance to the TCN model in accurately predicting WTI and Brent crude oil prices, while it outperforms TCN in predicting the WTI prices. Additionally, our results indicate that the BiGRU and GRU models are the best models to predict the Gold spot prices with input sequence of 30 and 60, respectively. The best forecasting performance for each market is WTI through a LigthtGBM model with input sequence 5, RMSE 3.213, Brent through a TCN model with input sequence 5, RMSE 2.035, Gold through a TCN model with input sequence 5, RMSE 20.83, and Silver through a TCN model with input sequence 60, RMSE 0.513. Eventually, our study suggests utilizing the TCN model for superior financial time series price predictions. From the empirical results, we figure out that the bidirectional LSTM and GRU models outperform the unidirectional LSTM and GRU models, respectively. Moreover, in general, the GRU-type models such as GRU, BiGRU, T2V-BiGRU, CNN-BiGRU, and TCN-BiGRU outperformed their LSTM-type peers in predicting the WTI, Brent, Gold, and Silver prices.

Our study has a number of implications for policymakers and investors. First, the results of this paper can assist investors and decision makers to promptly anticipate crude oil, gold, and silver market prices and adjust their investment portfolios. Additionally, stakeholders can execute risk hedging methods and lowering their losses with timely predictions. Particularly, gold is considered

a suitable safe-haven asset for the stock and cryptocurrencies markets (Junttila et al., 2018a). Therefore, timely prediction of gold market price will help stock market investors to hedge their portfolios. In terms of the organizational-level and country level relationships, organizations such as Organization of the Petroleum Exporting Countries (OPEC), World Petroleum Council (WPC) and International Energy Agency (IEA), and government agencies can further apply the suggested method, i.e., the TCN model, to devise profitable policies related to global crude oil price. Finally, our study would be particularly valuable to forecast crude oil, gold, and silver prices in case of some extreme events such as the COVID-19 pandemic and the recent conflict between Russia and Ukraine which were covered in the time period that was considered in this study.

In our research focusing on the forecasting of crude oil and precious metal prices, several limitations need to be acknowledged. Firstly, the volatile and non-linear nature of these markets poses difficulties in capturing all the intricate patterns and sudden changes in prices. Additionally, external factors such as natural disasters, geopolitical events and supply-demand dynamics can significantly influence commodity prices, and accurately incorporating these factors into forecasting models remains a complex task. Finally, it is essential to acknowledge the inherent uncertainty in forecasting and implement appropriate risk management strategies. Addressing these limitations will enhance the robustness and reliability of our research findings.

For future improvement of the crude oil and precious metals price forecasting, there are some possible directions. Firstly, rather than only using the historical price data, other features such as technical indicators, macroeconomic features, supply and demand data, production rate, and the interconnections with other financial markets can be utilized to predict the crude oil and precious metals prices. Secondly, incorporating the stakeholders' sentiments which can be derived from the news articles and social media platforms, might improve the forecasting performance of our suggested method. Finally, as an alternative to using the sequential data, other data structure and learning methods such as temporal graph neural networks can be implemented to forecast price time series data.

Chapter 5

Spatial-Temporal Graph Neural Networks for Price Movement Classification in Crude Oil and Precious Metals Markets

5.1 Introduction

Crude oil, gold and silver markets are considerably attractive for individuals, institutional investors, and governments both domestically and internationally. Drawing upon the precision of price trend predictions in these markets, stakeholders such as suppliers, producers, and consumers along the entire supply chain can adjust their strategies, ensuring optimal resource allocation, enhanced operational efficiency and resource utilization.

Investors face a crucial decision-making point centered on whether to buy or sell an asset ahead of the upcoming trading periods. In this context, forecasting the direction of price movement holds significant. As financial time series are characterized by inherent volatility, non-linearity and non-stationarity, forecasting price movement directions within such markets is a complex and challenging task. This paper investigates the direction-of-price-movement predictability in crude oil and precious metals, utilizing diverse analytical approaches.

Many researchers have dedicated their studies to find the variables that affect crude oil or precious metal markets. For instance, Y.-J. Zhang (2013) found that speculators' positions significantly impact WTI crude oil futures returns. Shin et al. (2013) utilized supply and demand related variables, and economic indicators such as producer price indices and US exchange rates to predict the direction of crude oil price movements. Kia et al. (2018) predicted the Brent crude oil price by employing various stock markets and gold market as features. P. Zhang & Ci (2020) used inflation rate, exchange rates, and stock markets indices to forecast gold price. Deng et al. (2023) employed the Japanese technical indicator Ichimoku Kinko Hyo to extract features from high-frequency crude oil price data. With the exponential growth of web-based information, accessing data about financial markets from various sources has become increasingly convenient. Notably, researchers pay more attention to unstructured data such as textual content from news articles, Google Trends, and Twitter, to enhance their predictions for the crude oil and precious metal market (X. Zhang et al.,

2012; X. Li et al., 2015; J. Wang et al., 2018; Y. Yang et al., 2021; B. Wu et al., 2021; Qin et al., 2023).

In forecasting literature, many studies on crude oil and precious metals markets have predominantly focused on price level or return predictions, often neglecting the critical aspect of classifying price trends. Our study seeks to address this notable gap in the existing literature by specifically focusing on predicting the direction of price movement in the crude oil, gold, and silver markets. The methods employed for forecasting in the crude oil and precious metals markets can be broadly classified into three main categories: traditional statistical models (J. Wang et al., 2020; Mittal, 2023; Kumar et al., 2023), machine learning models, and deep learning models. Machine learning approaches such as support vector machines (SVM), random forest (RF), linear regression (LR), decision tree (DT) have been widely used to forecast crude oil and precious metals markets (Abdullah Ahmed & Bin Shabri, 2014; Abraham et al., 2022; Das et al., 2022; Ongsritrakul & Soonthornphisaj, 2003; Pierdzioch & Risse, 2017; Weng et al., 2020; J. Yang et al., 2022; L. Zhao et al., 2015). More recently, deep learning methods such as LSTM, GRU, and CNN models for crude oil and precious metals forecasting have drawn researchers' interest (Z. He et al., 2019; Livieris et al., 2020; Q. Lu et al., 2021; Y.-X. Wu et al., 2019; K. Zhang & Hong, 2022). In a study of Fang et al. (2023), textual information with historical price data is employed in a hybrid FinBERT-VMD-Att-BiGRU model which combines integrates FinBERT, variational mode decomposition (VMD), an attention mechanism, and the BiGRU deep-learning model to predict WTI crude oil price. Yuan et al. (2023) forecast WTI price by constructing an ensemble model of statistical, machine learning and deep learning models. To forecast the future price of gold market, P. Zhang & Ci (2020) proposed a deep belief network (DBN) model that utilized a restricted Boltzmann machines (RBM) for pre-training and a layer of supervised back-propagation (BP) for fine-tuning.

The studies on predicting the direction of crude oil price movements are getting more popular in recent years. Deng et al. (2019) proposed a hybrid method to predict the one-week ahead direction changes of crude oil price using Multiple Dynamic Time Wrapping (MDTW) optimized by genetic algorithms. J. Liu et al. (2021) explored the predictability of price direction changes in commodity futures markets by employing binary probabilistic techniques, including Variable Length Markov Chain (VLMC). Nayak et al. (2023) proposed an Artificial Electric Field Algorithm-based Artificial Neural Networks (AEFA-ANN) model to optimize the parameters of a neural network for modeling and predicting crude oil price movements. Deng et al. (2023) forecast the crude oil futures in the Chinese market by a hybrid model of Fuzzy Rough Set (FRS), Non-dominated Sorting Genetic Algorithm-II (NSGA-II) and Sliding Window (SW) approaches. To the best of our knowledge, prior research has not focused on predicting price trends in precious metals markets. This study aims to fill this gap in the existing literature.

In this study, we approach the challenge of forecasting crude oil and precious metals prices by framing it as a Multivariate Time Series (MTS) classification task aimed at predicting one-day ahead movement direction in timeseries. Particularly, we are interested in classifying the direction of price

movements in WTI, Brent, Gold, and Silver markets. Investment decisions based solely on historical trends from a few predictors are more prone to risk. Thus, using a wide range of economic, financial, and supply and demand related predictors leads to more informed and resilient investment strategies. Achieving precise forecasts in MTS is a complex task as it demands simultaneous attention to temporal patterns within each variable and the interplay among various variables. To address this complexity, traditional methods such as vector autoregression (VAR), Vector Error Correction Model (VECM), vector auto-regression moving average (VARMA), and gaussian process (GP), often rely on the strict stationary assumption and suffer to capture the non-linearity among variables. Deep neural networks have shown better performance than traditional econometrics models on modeling non-stationary and non-linear dependencies. However, they still struggle in effectively capturing the interdependencies among predictors, particularly in cases involving a large number of variables. Recently, spatial-temporal graph neural networks (ST-GNNs) have gained significant attention as a powerful framework for modeling and analyzing data with both spatial and temporal dependencies. ST-GNNs are mainly examined in traffic speed forecasting (L. Bai et al., 2020; Y. Li, Yu, et al., 2017), skeleton-based forecasting, synthetic video prediction task (Y. Seo et al., 2018), and sign language recognition (Parelli et al., 2022) applications. In this study we adapt three state-of-the-art spatial-temporal graph neural networks, namely Spatial-Temporal Graph Attention Networks (ST-GAT) (Song et al., 2022), Multivariate Time-series Graph Neural Networks (MTGNN) (Z. Wu et al., 2020), and Attention Spatial-Temporal Graph Convolutional Networks (ASTGCN) (S. Guo et al., 2019), to predict the next-day direction of price movements in crude oil and precious metals' markets. We construct a graph of multivariate time series where each variable is a node, and the edges show the connection between each pair of variables.

The introduction of attention mechanism (Vaswani et al., 2017) has significantly improved the performance of deep learning models across many different domains and tasks (Brauwers & Frasincar, 2023). Attention mechanisms are particularly useful in tasks involving sequential data modeling such as neural machine translation (Peng et al., 2019), natural language processing (Galassi et al., 2021), and financial timeseries forecasting (Jiali, 2021; Uddin et al., 2021). At its core, an attention mechanism allows a model to selectively focus on specific parts of its input data while ignoring irrelevant information. Veličković et al. (2017) integrated the attention mechanism with graph neural networks which showed significant improvements on Citation network, and Protein interaction benchmark datasets. In this study, we examine the effectiveness of spatial and temporal attention incorporation in spatial-temporal neural networks for multivariate timeseries trends classification tasks.

Within this research, we attempt to highlight the capabilities of ST-GNN models, enhanced by attention mechanisms, for modeling multivariate time series data in the context of financial market predictions. More specifically, the following research questions will be explored.

1. Do spatial-temporal graph neural networks outperform deep learning models in predicting the direction of price movements in crude oil and precious metal markets?

2. How the interaction between various variables affecting the price of crude oil and precious metal markets can be modeled?
3. Does an attention mechanism improve the accuracy of ST-GNN models in predicting the market trends?
4. What is the best performing model in accurately predicting the trend in each market?
5. Since each market has its own unique specifications and price dynamics, how can we make our predictive models adaptive to each market?

In summary, our study offers the following important contributions to the existing literature on predicting the price trend in crude oil and precious metal markets:

- We customize and enhance advanced spatial-temporal graph neural network models, including MTGNN (Z. Wu et al., 2020), ST-GAT (Song et al., 2022), and ASTGCN (S. Guo et al., 2019), for the purpose of classifying price movement directions in crude oil and precious metal markets.
- To the best of our knowledge, this study marks the pioneering attempt within the crude oil and precious metals literature to utilize the potential of graph neural networks for predicting price movement directions in these markets.
- Considering the complex interplay of factors influencing crude oil and precious metal markets, this study examines the impact of a comprehensive set of 25 variables spanning commodity markets, prominent stock indices, exchange rates, global macroeconomic factors, as well as supply and demand dynamics. We aim to leverage the predictive potential of these variables in determining the directional shifts in prices within WTI, Brent, Gold, and Silver markets.
- The majority of spatial-temporal graph neural network models have been predominantly tailored for applications such as traffic flow forecasting or tasks centered around skeleton-based data, where the dataset inherently assumes a graph-like structure. However, this study endeavors to unveil the power of ST-GNNs in financial time series classification tasks.
- We improve the predictive performance of MTGNN model by incorporating an attention mechanism in the temporal module of this model. We call the new architecture multivariate timeseries graph neural networks with temporal attention and learnable adjacency matrix (MTGNN-TAttLA). Furthermore, we revamp the temporal module of ST-GAT by implementing temporal dilated convolution networks, constructing a new model named spatial attention graph with temporal convolutional networks (SAG-TCN).

5.2 Related Works

In this section, we present an overview of related works to crude oil and precious metals price trend predictions, and spatial-temporal graph neural networks. Given the scarcity of studies focused on predicting price movement directions in our target markets, namely, WTI, Brent, Gold, and Silver, we conducted a thorough literature review within the broader domain of financial markets.

5.2.1 Price trend classification in financial markets

In recent years, many methods have been developed to address financial markets' trend prediction problem. Initially traditional statistical approaches were used (Lauren & Harlili, 2014; Xiao et al., 2022), but with the rapid advancement of artificial intelligence technology, deep learning techniques have gained popularity in this domain. Deng et al. (2023) forecast the crude oil futures in the Chinese market by a hybrid model of Fuzzy Rough Set (FRS), Non-dominated Sorting Genetic Algorithm-II (NSGA-II) and Sliding Window (SW) approaches. Various deep learning methods have been employed for predicting market trends, including RNN, RBM, LSTM and CNN (Buczkowski, 2017; Haq et al., 2021; Ma et al., 2022; Thakkar & Chaudhari, 2021). The capacity of these models to capture spatial dependencies among multiple variables that influence financial market trends is constrained. The intricate relationships that exist between financial markets and other factors, whether economic, financial, geopolitical, or supply and demand-related, can be effectively represented as a graph structure. Traditional deep learning models struggle to handle such graph-structured data. Consequently, there is a growing trend among researchers to turn to spatial-temporal neural networks to better capture the interdependencies among variables across various time steps. Remarkably, the application of ST-GNN methods to address the prediction of trends in crude oil and precious metals markets remains relatively unexplored in the existing literature.

Shin et al. (2013) proposed a graph network representation that connected WTI crude oil with a range of global and domestic economic factors, including variables associated with global supply and demand, producer price indices, and US dollar exchange rates. Within this graph representation, the nodes represent variables, and their connections are determined by their similarities, which are computed using a weighted K-Nearest Neighbors method. By utilizing a semi-supervised learning (SSL) algorithm which combines Nonlinear Principal Component Analysis (NLPCA) with an Auto-Associative Neural Network (AANN), the authors predicted the next month's upward or downward changes oil prices. Notably, this SSL model did not process the data sequentially but rather leveraged technical indicators to transform the time-series data into a vector-based format. In a similar manner, Kia et al. (2018) designed a hybrid supervised semi-supervised (HyS3) graph-based model to predict one-day ahead movement of global stock markets, Brent crude oil and gold prices. This hybrid model combined a supervised Support Vector Machine (SVM) component with a graph-based semi-supervised learning (GSSL) component, which involved iterative label spreading. To construct the graph structure, the authors utilized the maximum spanning tree derived from the correlation network of markets as its base structure. They further enhanced this structure by adding supplementary edges that contributed to the overall prediction accuracy of the model. The label propagation method was then employed to forecast the unknown labels for markets.

Most of the direction of price trend prediction literature focus on stock market movements. M. Kim & Sayama (2017) predicted the future movements of Standard & Poor's 500 Index (S&P 500) by constructing a network of S&P 500 constituent companies. Each pair of companies within this network is linked by weighted mutual information. The authors add network centrality

measurements to an Autoregressive Integrated Moving Average (ARIMA) model to enhance the prediction accuracy. In another study, Y. Chen et al. (2018) constructed a network of companies based on knowledge graphs to predict stock prices using a graph convolutional network (GCN). In finance, a knowledge graph serves as a tool for uncovering intricate connections between entities such as companies, management, news events, and user preferences. These knowledge graphs essentially function as databases that facilitate semantic search by preserving intricate relationships among multiple entities (Q. Wang et al., 2017). Y. Liu et al. (2018) extracted features from financial news using a hybrid knowledge graph embedding and a Convolutional Neural Network (CNN) model. These extracted features, in conjunction with daily S&P500 price data and various technical indicators, are then integrated into a Long Short-term Memory (LSTM) model to predict Apple's stock price movement. In a separate research investigation focused on the FOREX market, Rundo (2019) introduced a joint deep learning and reinforcement learning (RL) algorithm to predict the short-term trend in the EUR/USD exchange rate. In their proposed pipeline, first, a block of LSTM layers predicts the market trend. Then, these predicted market signals undergo a verification or correction process through a reinforcement learning (RL) module, designed to maximize the return on investment. More recently, researchers noticed the importance of modeling the interconnection between stock market predictors and employed graph neural networks for capturing a representation that shows these relationships. For instance, Matsunaga et al. (2019) extracted a representation of a company knowledge graph using a graph convolutional network (GCN) and integrated this graph representation into a recurrent neural network (RNN) model to predict stock prices in the Japanese Nikkei 225 market. W. Li et al. (2020) introduced a LSTM Relational Graph Convolutional Network (LSTM-RGCN) model to predict the overnight stock price movement in Tokyo Stock Exchange. To achieve this, the model leveraged input features like news vectors and historical price embeddings of companies in a GCN layer to find the relationship among various stocks. The connection among stocks (i.e. edge weights) are determined by a correlation matrix. W. Chen et al. (2021) used a model which combined graph convolutional networks with convolutional neural network (GC–CNN) to predict Chinese stock market trends. They created a network representation of the stock market, where each market is represented as a node, and the correlations between markets are edge weights. Using daily trading data and technical indicators as node features, the stock network then transformed into images that were further processed by a CNN model to make final trend predictions.

5.2.2 Spatial-Temporal Graph Neural Networks

Graph Neural Networks (GNNs) have emerged as a powerful paradigm in machine learning, revolutionizing how we analyze and interpret data with complex relationships. In the past decade, a growing body of literature focused on the utilization of Graph Convolutional Networks (GCNs) (Bruna et al., 2013; Kipf & Welling, 2016) to operate convolutional filters on data structured as graphs. Graph data is formed by a collection of nodes interconnected by edges, and the relationships

among these edges are defined within an adjacency matrix. GCNs are typically categorized in two key paradigms: the spatial and spectral approaches. Spatial techniques directly execute convolution operations on graph vertices and their adjacent nodes. The convolution operation enables the central node to assimilate information from its neighbors, thereby capturing and encoding local patterns and relationships within the graph. In contrast, spectral perspective methods leverage graph Laplacian matrix eigenvalues and eigenvectors. They conduct convolution operations in the frequency domain through graph Fourier transforms, eliminating the necessity to extract locally connected regions from graphs during each convolutional step (Shi et al., 2018). Kipf & Welling (2016) introduced a first approximation of Chebyshev spectral filter (Defferrard et al., 2016).

Spatial-Temporal Graph Convolutional Networks (ST-GCNs) can handle data with both spatial (inter-variable relationships) and temporal characteristics. Current research on spatial-temporal graph modeling can be broadly categorized into two main directions. One approach integrates graph convolution networks (GCN) within recurrent neural networks (RNN) such as LSTM and Gated Recurrent Units (GRU) (Andreoletti et al., 2019; Jain et al., 2016; C. Li et al., 2022; Y. Seo et al., 2018), while the other incorporates them into convolution neural networks (CNN) (S. Yan et al., 2018; B. Yu et al., 2018). RNN-based approaches suffer from time-consuming iterative propagation and gradient explosion or vanishing problems when capturing long-term sequences (M. Seo et al., 2017; X. Zhang et al., 2018). In contrast, CNN-based approaches offer benefits such as efficient parallel computing, and low memory demands. However, CNNs require multiple layers to effectively capture extended sequences due to their use of standard 1D convolution, where the receptive field size increases linearly with the number of hidden layers. To address this challenge, Z. Wu et al. (2019) introduced the Graph WaveNet model which combined GCNs with stacked dilated casual convolution networks to capture temporal dependencies. Since the receptive field size of stacked dilated casual convolution networks grows exponentially with an increase in the number of hidden layers, Graph WaveNet is capable of handling very long sequences.

The majority of ST-GCN models are utilized in multivariate timeseries tasks involving traffic flow forecasting (L. Bai et al., 2020; S. Guo et al., 2019; Z. Wu et al., 2019; B. Yu et al., 2018; Zheng et al., 2019), and body movement estimation (Jain et al., 2016; Parelli et al., 2022; Shi et al., 2018). More recently, researchers have started unveiling the potential of ST-GCN model for stock market forecasting. For instance, Feng et al. (2018) developed a Relational Stock Ranking (RSR) method including a Temporal Graph Convolution (TGC) for stock ranking prediction. The TGC model combines a LSTM and a GCN with Chebyshev polynomials filters for capturing the temporal and spatial relationships, respectively, in stock market. The stock market graph was constructed by sector-industry relations and Wikidata company-based relations. Z. Wu et al. (2020) leveraged the advantage of dilated convolutional filters in a ST-GCN model to enhance prediction tasks involving diverse datasets, including solar energy, traffic, electricity, and exchange rate data. Their approach integrates a GCN featuring a mix-hop propagation method, effectively capturing spatial relationships among variables. Additionally, a temporal graph convolution module was employed to identify intra-variable temporal dependencies. Since for most multivariate timeseries the

relationship among variables is not predefined, they proposed a graph learning module to extract the unidirectional relations among variables. In another study, Xiao et al. (2022) introduced the Adaptive Fused Spatial-Temporal Graph Convolutional (AFSTGC) model to fuse disordered temporal, spatial, and spatial-temporal dependencies into structured data. By constructing an adaptive fused adjacency matrix that dynamically and concurrently captured potential temporal, spatial, and spatial-temporal relations, a GCN module forecast the NASDAQ100 index value without requiring any separate temporal module. L. Chen et al. (2022) developed a multi-scale adaptive graph neural network (MAGNN) to capture inter-variable and temporal dependencies at different time scales. Within the MAGNN framework, they incorporated an adaptive graph learning module capable of automatically deducing scale-specific graph structures. The model's effectiveness was demonstrated through its application to diverse forecasting tasks, encompassing datasets related to solar energy, traffic, electricity, exchange rates, and NASDAQ.

Attention Mechanisms have gained widespread popularity across various domains due to their high efficiency, flexibility, and adaptability in capturing dependencies within variable-length sequences (M.-H. Guo et al., 2022; J. Lin et al., 2022; Vaswani et al., 2017). The core idea of attention mechanisms is to adaptively focus on the most relevant features according to the input data. Recently, researchers applied attention mechanisms to graph-structured data (Brody et al., 2021; Veličković et al., 2017) to model spatial correlations. They introduce graph multi-head attention networks to ascertain how much attention should be allocated to neighboring node j in order to determine the representation of each central node i. J. Zhang et al. (2018) constructed Gated Attention Networks (GaAN) by using gated attention aggregators to collect crucial information from neighboring nodes for a central node in a traffic network graph. They replaced the fully connected layer within GRU units with GaAN modules to form graph gated recurrent units (Graph GRU), allowing them to capture both spatial and temporal dependencies. Later, studies of J. Bai et al. (2021), Song et al. (2022), and Zheng et al. (2019) extended the attention mechanism to spatial-temporal graph data for traffic prediction. They employed the temporal attention mechanism to modulate the significance of various time points and gather comprehensive temporal information, ultimately enhancing prediction accuracy. S. Guo et al. (2019) employed the attention mechanism in both the spatial and temporal component of their proposed attention based spatial-temporal graph convolutional network (ASTGCN) model to predict traffic speed. The spatial module of ASTGCN is constructed of a spectral GCN while the temporal module is a standard convolutional neural network. The graph attention network is also used for stock movement prediction. R. Kim et al. (2019) used an LSTM model to extract temporal representations for each company listed in the S&P500 market index. These temporal representations were then employed as node features within a graph representing the relationships among companies. To identify spatial dependencies among these companies, the authors introduced a hierarchical graph attention network that selectively aggregates information from neighboring companies.

To construct the graph network and define the adjacency matrix, various approaches have been proposed in prior literature. One common approach is to use domain-specific knowledge or pre-

existing relationships to directly define the edges and weights between nodes. This method is useful when the graph structure is well-defined and known in advance. Examples of pre-defined graph structures are road networks for traffic flow forecasting and joint connections in skeleton-based tasks (Zheng et al., 2019). In financial time series analysis, knowledge graphs such as sector-industry relations and corporate relations data are used to find the relationship among entities (Feng et al., 2018; R. Kim et al., 2019; Y. Yang et al., 2019). Moreover, Correlation analysis is used to find the existence of relationship among nodes and connect them together (W.-Q. Huang et al., 2009; Namaki et al., 2011). M. Kim & Sayama (2017) constructed a stock market graph with non-linear correlation among variables measured by mutual information. In financial domains, it's often challenging to have predefined relationships between factors that affect a financial market as ground truth. Additionally, the spatial correlations between variables can be dynamic and evolve over time. Graph neural networks can employ unsupervised learning techniques to automatically learn the adjacency matrix from the data itself. This can involve methods like k-nearest neighbors (KNN) to establish connections based on proximity or similarity measures (Jain et al., 2016; Y. Seo et al., 2018). Lately, supervised graph learning approaches are proposed to tackle this issue within the Spatial-temporal graph convolutional networks that learns the adjacency matrix of graphs in an end-to-end manner (L. Bai et al., 2020; L. Chen et al., 2022; K. Guo et al., 2022; Shi et al., 2018; Z. Wu et al., 2019, 2020; Xiao et al., 2022).

In this study, we utilize a KNN approach, and the graph learning module proposed in the study of Z. Wu et al. (2020) to construct the multivariate graph network for predicting the trend of crude oil and precious metals markets.

5.3 Methodology

5.3.1 Constructing the Markets' Graph

Unlike conventional price prediction models that estimate future price levels, models for predicting the direction of price movement predominantly center around evaluating the likelihood of an event (i.e. occurrence of upward or downward price movements). In the subsequent section, we delve into diverse methodologies employed for classifying movements in multivariate time series data. Despite variations in their architectural aspects, these approaches share a common objective: estimating the probability of a positive price movement in the next period. Relying on this predicted probability, we are able to predict the future class for the direction-of-price-movement.

Definition 2.1 (Problem): Let $x_t \epsilon \boldsymbol{R}^N$ denote the value of a multivariate variable of dimension N at time step t, where $x_t[i] \epsilon \boldsymbol{R}$ denote the value of the i^{th} variable at time step t. Given a sequence of historical S time steps of multivariate observations, $\mathbf{X} = \{x_{t-S+1}, x_{t-S+2}, \dots, x_t\}$, our goal is to predict the direction of price movements in one-day-ahead future, denoted by $\mathbf{Y} = \{C_{t+1}\}$, where $C_t \epsilon \{0,1\}^N$ is the binary class of 1 or 0 for the future x value moving up or down, respectively.

We aim to build a mapping f (·) from **X** to **Y** by minimizing the binary cross entropy loss with $\boldsymbol{L2}$ regularization as follows.

$$Loss = -\frac{1}{n}\sum_{i=1}^{n} y_i.\log\big(\sigma(\hat{y}_i)\big) + (1-y_i).\log\big(1-\sigma(\hat{y}_i)\big) + \frac{\lambda}{2n}\sum_w w^2 \quad (5\text{-}1)$$

Where, n represents the batch size, y_i is the actual class, $\sigma(\hat{y}_i)$ is the predicted probability of that class, $\hat{y}_i$ is the model output, $\sigma(.)$ is the sigmoid function, λ is the regularization parameter, and w is the model parameter.

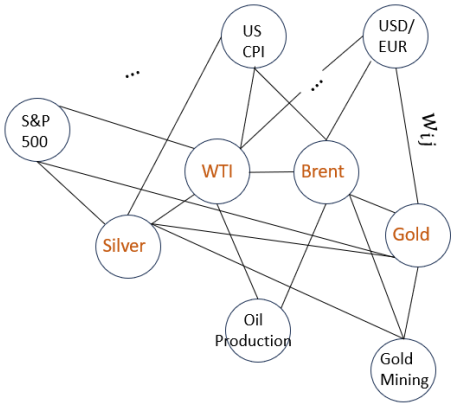
In this study, we use three different spatial-temporal graph neural networks as the mapping function f (·) to predict the direction of crude oil and precious metals price movements. The relationship among our features is conceptualized as a graph network defined below.

Definition 2.2 (Graph). A graph is denoted by $G = (V, E)$ where V is the set of nodes, E is the set of edges, and $|V| = N$ (number of nodes is N).

Definition 2.3 (Edges). Let $v_i \in V$, and $v_j \in V$ to denote two connected nodes in graph. An edge pointing from v_j to v_i is defined as $e = (v_i, v_j) \in E$.

Definition 2.4 (Adjacency Matrix). The adjacency matrix is a mathematical representation that indicates whether pairs of nodes are adjacent or not. It is denoted as $A \in \boldsymbol{R}^{N\times N}$ with $A_{ij} = r > 0$ if $(v_i, v_j) \in E$ and $A_{ij} = 0$ if $(v_i, v_j) \notin E$.

To construct the graph neural networks, we consider the variables present in our multivariate time series dataset as nodes, outlining connections between these nodes through the utilization of a graph adjacency matrix. Figure 5-1-a presents a schematic of our features graph. The collaboration of a graph convolutional module with a temporal module in a ST-GNN model is demonstrated in Figure 5-1-b. A temporal module filters the inputs by sliding a 1D window over the time and node axes, as denoted by the green. A graph convolution module filters the inputs at each step, denoted by the red.



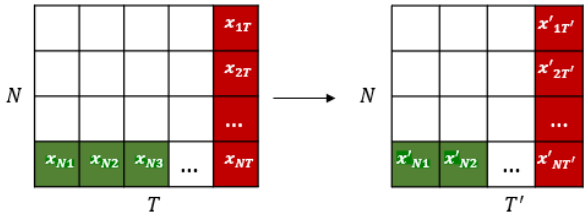


Figure 5-1: (a) Schematic of features graph **(b)** Representation of spatial and temporal module collaboration

Construct Graph Adjacency Matrix: Unlike road traffic or skeleton movement time series data, multivariate financial time series data are not inherently structured in a graph format. Therefore, it

becomes necessary to construct a graph that captures the potential relationships among variables within our dataset. In SGA-TCN and ASTGCN graph neural networks, the adjacency matrices need to be predefined. For this purpose, we have utilized the KNN graph from the PyTorch Geometric library to identify the k-nearest nodes (representing features in our dataset) to a given node. To assess the similarities between nodes, we have employed the Euclidean distance metric across all time steps. The KNN method keeps the K largest weights in the network and eliminates the rest. The justification of filtering a network and eliminating some edges with smaller weights is noise reduction (Kia et al., 2018).

5.3.2 Spatial Graph Attention -Temporal Convolution Network (SGA-TCN)

Spatial Graph Attention -Temporal Convolution Network (SGA-TCN) is designed to effectively capture the complex interplay between spatial and temporal dependencies in graph-structured data. We created the SGA-TCN model for our classification task by adapting the ST-GAT (Song et al., 2022) model. Specifically, in this study we utilize blocks of temporal dilated convolutional networks as the temporal module while in ST-GAT, LSTM layers are employed. Several studies have proven that TCNs are more powerful than LSTM models in timeseries forecasting. Figure 5-2 depicts the architecture of SGA-TCN model.

Spatial Graph Attention: Graph attention is an extension of GCN that applies an explicit attention mechanism (Vaswani et al., 2017) to learn the hidden state of each node by iteratively using node feature for similarity computation. The fundamental distinction between GAT and GCN lies in their approach to gathering and aggregating feature representations from neighboring nodes.

Let h^l denote the input features of layer l in the spatial graph attention module: $h^l = \{h_1^l, h_2^l, \dots, h_N^l\}, h_i^l \in R^S$, where S is the historical window size for each node. The attention score between Nodes i and j is calculated using Eq. (5-2) and Eq. (5-3).

$$e_{ij} = SelfAttention(\Phi h_i^l, \Phi h_j^l) \tag{5-2}$$

Where $\Phi \in R^{S \times S'}$ is the weight matrix that transforms node features into higher-level states.

$$\alpha_{ij} = softmax(LeakyReLU(e_{ij})) \tag{5-3}$$

These attention coefficients, α_{ij}, are used to update node representations for the next layer:

$$h_i^{l+1} = Concat_K\left(\sigma\left(\sum_{j \in N(i)} \alpha_{ij} \Phi^l h_j^l\right)\right) \tag{5-4}$$

Where h_i^{l+1} represents the updated hidden states of node i, $N(i)$ is the set of immediate neighbor nodes of node i, σ is a non-linear activation function, and K is number of attention heads.

Temporal TCN module: The updated node representations will be utilized as inputs of a TCN module. The TCN module includes B number of TCN blocks. Each block is constructed of two Dilated Convolution layers followed by batch normalization, ReLU activation, and dropout layers. A dilated convolution layer is a type of convolutional layer for processing input sequences, with an expanded receptive field. Unlike traditional convolution where a filter is applied directly to adjacent elements, dilated convolution introduces gaps in the filter, allowing it to capture information from

a wider range without increasing the number of parameters (L. Chen et al., 2014). Batch normalization is applied to improve training stability, convergence speed, and generalization performance. The purpose of dropout layer is to regularize the model for reducing the potential overfitting problem. In each block, a residual connection is used to address the vanishing gradient problem, improve gradient flow, and help in training deeper networks more effectively.

The output of TCN layer i, $H^i \in R^{C_{tcn_i} \times N \times S_{new}}$, is:

$$H^i = dropout(ReLU(BatchNorm(H^{i-1} *_D W_i + b_i))) \quad (5\text{-}5)$$

Where, C_{tcn_i} represents output channels of the TCN layer i, S_{new} is a new sequence number, $*_D$ is the dilated convolution operation.

The final predictions are generated through a fully connected convolutional neural network in the output layer.

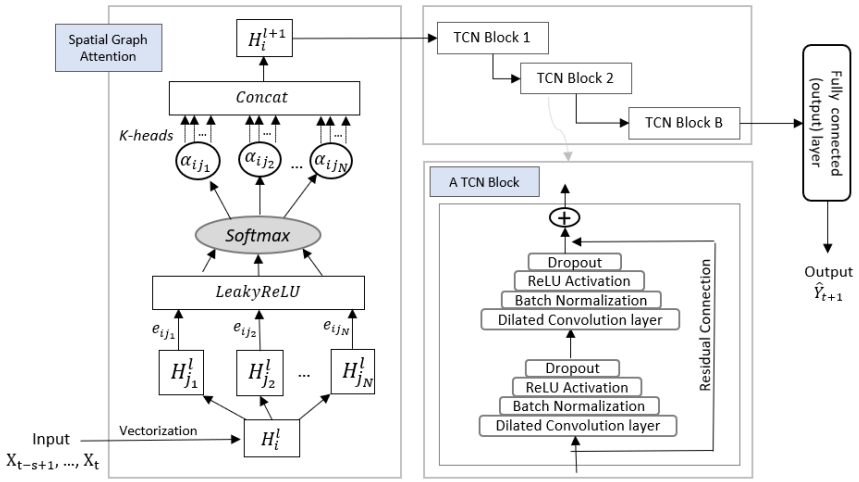


Figure 5-2: SAG-TCN framework

5.3.3 Multivariate Timeseries Graph Neural Network with Temporal Attention and Learnable Adjacency Matrix (MTGNN-TAttLA)

The second graph neural network for the classifying the direction-of-price-movement of our target markets is Multivariate Timeseries Graph Neural Network with Temporal Attention and Learnable Adjacency Matrix (MTGNN -TAttLA). The attention mechanism is argued to improve prediction performance in many prior studies (Peng et al., 2019; Galassi et al., 2021; Jiali, 2021; Uddin et al., 2021). In this work, MTGNN -TAttLA is adapted from the MTGNN (Z. Wu et al., 2020) model by incorporating a temporal attention mechanism to the TCN modules of MTGNN. According to our knowledge gained during the literature review process, we found MTGNN to be the most suitable graph neural networks for multivariate time series forecasting. The reason for this is the existence of a graph learning layer in MTGNN that learns the adjacency matrix. In MTGNN, the graph

captures the relationships and interactions between different variables (nodes) over time. Figure 5-3-(a) illustrates the MTGNN-TAttLA framework.

The four core components of MTGNN-TAttLA model are: graph construction module, graph convolution networks, attention-based temporal convolution networks, and an output module. The graph construction module finds the relationship between variables in an adjacency matrix, which then is utilized as inputs to graph convolution modules. This module also acts as a feature importance approach that determines which variables in our dataset are most useful in the classification of WTI, Brent, Gold, and Silver's direction-of-price-movement. Each spatial-temporal layer (ST-layer) includes a graph convolution and an attention-based temporal convolution module. Graph convolution and temporal convolution modules work together to unveil the spatial and temporal dependencies in our dataset. To bypass the potential gradient vanishing problem, residual connections are applied from the inputs to the outputs of each ST-layer. Following each temporal convolution module, skip connections are added to help maintain the original information in the data throughout the network layers. The detailed explanation of each module of our MTGNN-TAttLA model follows:

Graph Construction Module: This module is responsible for constructing the graph representation of the data. We have used the proposed method in (Z. Wu et al., 2020) to extract uni-directional connections between nodes as follows:

$$M_1 = \tanh(\alpha E_1 \theta_1) \tag{5-6}$$

$$M_2 = \tanh(\alpha E_2 \theta_2) \tag{5-7}$$

$$A = ReLU(\tanh(\alpha(M_1 M_2^T - M_2 M_1^T)) \tag{5-8}$$

where $E_1 \in R^{N \times D}$, $E_2 \in R^{N \times D}$ represent randomly initialized node embeddings, which are learnable during training, θ_1, θ_2 are model parameters, α is a hyper-parameter for controlling the saturation rate of the activation function. For every individual node, we opt for its nearest k nodes to serve as its neighboring nodes. While maintaining the weights for linked nodes, we assign a weight of zero to nodes that are not connected as below.

$$index = argtopk(A[i,:]), \qquad for\ i = 1,2,\dots,N \tag{5-9}$$

$$A[i, -inex] = 0 \tag{5-10}$$

Graph Convolution Module: The graph convolution module is designed to blend the information of a node with that of its neighbors, effectively managing spatial interdependencies within a graph. As shown in Figure 5-3-(b), this module includes two mix-hop propagation layers to process the flow of information across nodes with spatial dependencies. This mix-hop propagation layer operates through information propagation and information selection steps as follows:

The information propagation step is defined as:

$$H^k = \beta H_{in} + (1-\beta)\tilde{A}H^{k-1}, \tag{5-11}$$

where β serves as a hyperparameter that regulates the ratio of maintaining the root node's original states. A fraction of nodes' initial states is conserved throughout the propagation process, allowing the node states to uphold locality while also exploring a broader neighborhood. The output of the information selection step is defined as:

$$H_{out} = \sum_{k=0}^{K} H^k W^k \tag{5-12}$$

where K represents the depth of propagation, H_{in} is the input hidden states outputted by the previous layer, $H^0 = H_{in}$, $\tilde{A} = \widetilde{D}^{-1}(A+I)$, and $\widetilde{D}_{ii} = 1 + \sum_j A_{ij}$. Figure 5-3-(c) illustrates that first, the information is propagated horizontally to depth of K and then is selected vertically. The parameter W^k functions as a feature selection matrix.

Attention Temporal Convolution Module: This module comprises two dilated convolution layers, each followed by a temporal attention mechanism.

The dilated convolution layer captures temporal patterns within timeseries data through a set of dilated 1D convolution filters. We employ diverse dilated filter sizes of $1 \times 2, 1 \times 3, 1 \times 6, 1 \times 7$ as outlined in (Z. Wu et al., 2020). The incorporation of these various filter sizes enables the model to recognize temporal patterns across different ranges, while the utilization of dilated convolution filters empowers the model to handle very long sequences effectively.

The output of each dilated convolution layer is then processed through a batch normalization layer, a ReLU activation function, and a masked k-head self-attention mechanism (Vaswani et al., 2017). The masked attention layer is used to find the attention weights of the temporal hidden states as follows:

$$H^l = ReLU(BatchNorm(H^l_{DC})) \tag{5-13}$$

$$Attention_i = softmax\left(\frac{H^l W_1 (H^l W_2)^T}{\sqrt{d}}\right) H^l W_3 \tag{5-14}$$

$$H^{l_t} = Concat_k(Attention_i) \tag{5-15}$$

Where $H^l_{DC} \in \boldsymbol{R}^{S_l \times N \times C_l}$ denotes the output of l^{th} dilated convolution layer, $H^l \in \boldsymbol{R}^{N \times S_l \times d}$ is the input of self-attention layer, $H^{l_t} \in \boldsymbol{R}^{N \times S_l \times C_l}$ is the output of attention layer, C_l is the number of channels in the l^{th} temporal convolution output, S_l is the number of sequences in the temporal convolution output, d is equal to $\frac{C_l}{k}$, and $W_1, W_2,$ and W_3 are the query, key, and value weights in the attention mechanism, respectively.

We used the $tanh$ activation function on the output of one attention dilated convolution layer and $sigmoid$ activation on the other layer. These functions act as a filter and a gate, respectively. The gate controls the amount of information that the filter can transmit to the subsequent module. Figure 5-3-(d) and Figure 5-3-(e) demonstrate the mechanics of the attention temporal convolution module and the dilated convolution layer.

Output module: The output layer of MTGNN-TAttLA comprises of two 1×1 standard convolutional layers, which effectively reshape the channel dimension of the input data to match the desired output dimension. In this study the output dimension is 1 as we predict the price direction class for one-day ahead.

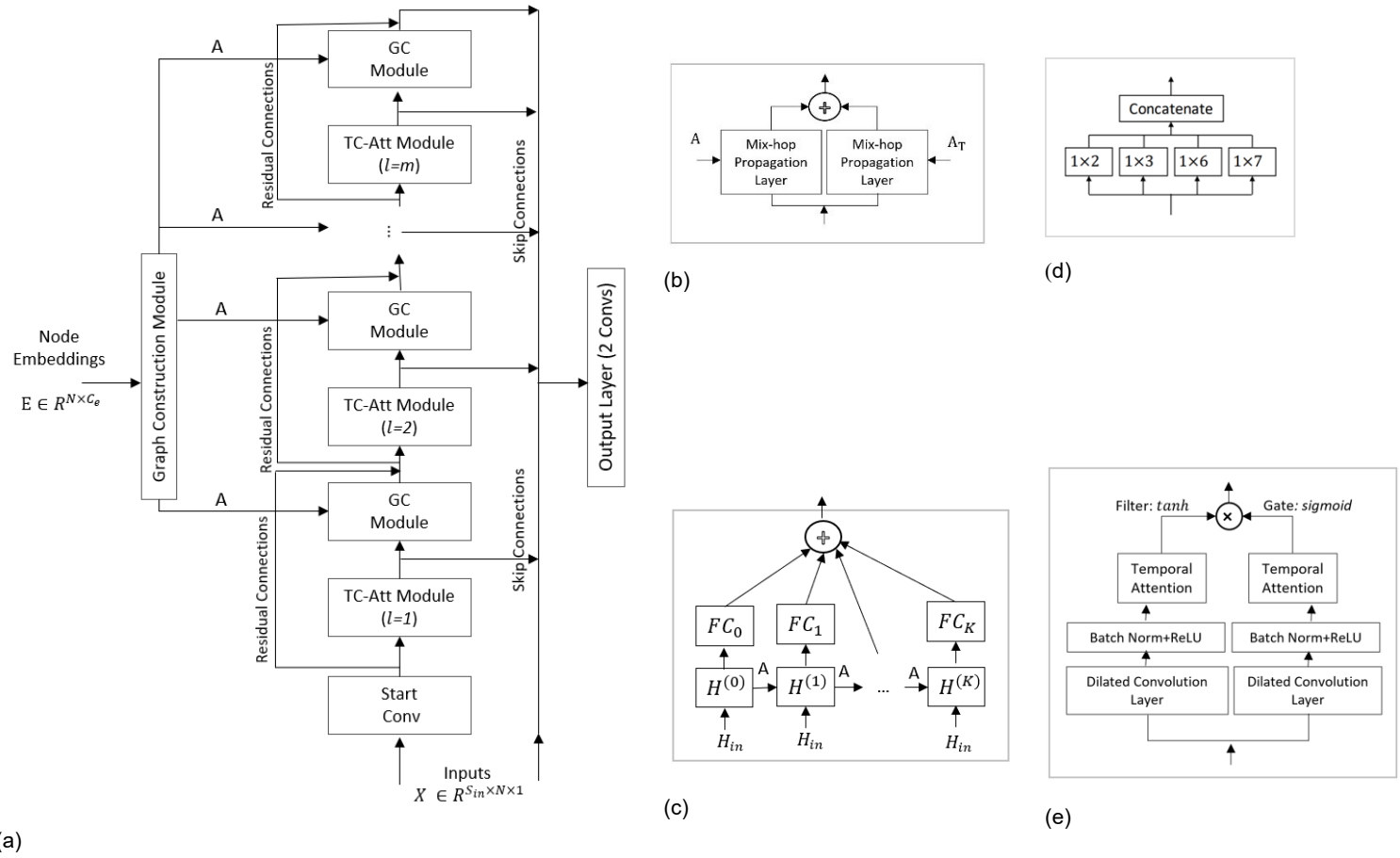


Figure 5-3: (a) MTGNN-TAttLA framework, **(b)** Graph convolution module, **(c)** a mix-hop propagation layer, **(d)** a dilated convolution layer with different kernel sizes, **(e)** a gated attention-based temporal module

5.3.4 Attention-based Spatial and Temporal Graph Convolution Networks (ASTGCN)

The attention-based spatial and temporal graph convolution networks (ASTGCN) (S. Guo et al., 2019) was first introduced for traffic flow forecasting. Similar to MTGNN, ASTGCN employs graph convolutional layers to capture spatial dependencies within a graph structure and temporal convolutional layers to uncover sequential patterns over time. However, this model differs with MTGNN in following aspects: (a) the adjacency matrix is predetermined and not learnable, (b) the temporal module utilizes standard convolution operations instead of dilated convolution operations and filters, (c) in the graph convolution module, it does not include the information selection step,

(d) there is no skip connection in the ASTGCN model architecture. In addition, ASTGN applies self-attention mechanisms to spatial and temporal dimensions of the input features to dynamically weigh the importance of information from different nodes and time steps, enhancing the model's ability to focus on relevant data. Figure 5-4 Shows the ASTGCN model framework.

Spatial attention:

$$S = V_s\,\sigma\left(\left(H^{l-1}W_1\right)W_2\left(W_3H^{l-1}\right)^T + b_s\right) \tag{5-16-a}$$

$$S'_{ij} = softmax(S_{ij})\ ,\ S' \in \boldsymbol{R}^{N\times N} \tag{5-16-b}$$

Where $H^{l-1} \in \boldsymbol{R}^{N\times C_{l-1}\times S_{l-1}}$ is the input of l^{th} spatial-temporal layer, C_{l-1} is the number of input channels, and S_{l-1} is the length of temporal dimension in l^{th} layer, respectively. $V_s, b_s \in \boldsymbol{R}^{N\times N}$, $W_1 \in \boldsymbol{R}^{S_{l-1}}, W_2 \in \boldsymbol{R}^{C_{l-1}\times S_{l-1}}, W_3 \in \boldsymbol{R}^{C_{l-1}}$ are learnable parameters and σ is sigmoid activation function.

Temporal attention:

$$E = V_e\,\sigma((H^{l-1}U_1)U_2(U_3H^{l-1})^T + b_e) \tag{5-17-a}$$

$$E'_{ij} = softmax(E_{ij})\ \ ,\ E' \in \boldsymbol{R}^{N\times C_{l-1}\times S_{l-1}} \tag{5-17-b}$$

Where $V_e, b_e \in \boldsymbol{R}^{S_{l-1}\times S_{l-1}}$, $U_1 \in \boldsymbol{R}^{N}, U_2 \in \boldsymbol{R}^{C_{l-1}\times N}, U_3 \in \boldsymbol{R}^{C_{l-1}}$ are learnable parameters. The value of an element E_{ij} indicates the strength of relationship between time instances i and j.

Graph convolution module:

$$g_\theta *_G x = \ g_\theta(\boldsymbol{L})x \tag{5-18}$$

where $*_G$ represents a graph convolution operation, $\boldsymbol{L}$ is the normalized graph Laplacian matrix, $\boldsymbol{L} = I_N - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \in \boldsymbol{R}^{N\times N}$, D is the graph degree matrix with $D_{ii} = \sum_j A_{ij}$. Chebyshev polynomials of the graph Laplacian matrix are adopted to approximate the spectral filters g_θ:

$$g_\theta *_G x = \ g_\theta(\boldsymbol{L})x = \sum_{k=0}^{K-1}\theta_k T_k\left(\tilde{\boldsymbol{L}}\right)x \tag{5-19}$$

where $\tilde{\boldsymbol{L}} = \frac{2}{\lambda_{max}}\boldsymbol{L} - \boldsymbol{I_N}$, λ_{max} is the maximum eigenvalue of the Laplacian matrix.

Temporal convolution module:

$$H^l = ReLU(\Phi * ReLU(g_\theta *_G H^{l-1}))) \in \boldsymbol{R}^{C_l\times N\times S_l} \tag{5-20}$$

where $*$ is a standard convolution operation, Φ is the parameters of the temporal dimension convolution kernel.

The ASTGCN is composed of several spatial-temporal blocks (ST-blocks) aimed at capturing a broader spectrum of dynamic spatial-temporal interdependencies. To avoid gradient vanishing problem, a residual connection is applied from the input to the output of each ST-block. A fully

connected layer is added after the last ST-block to ensure the output dimension matches the target dimensions.

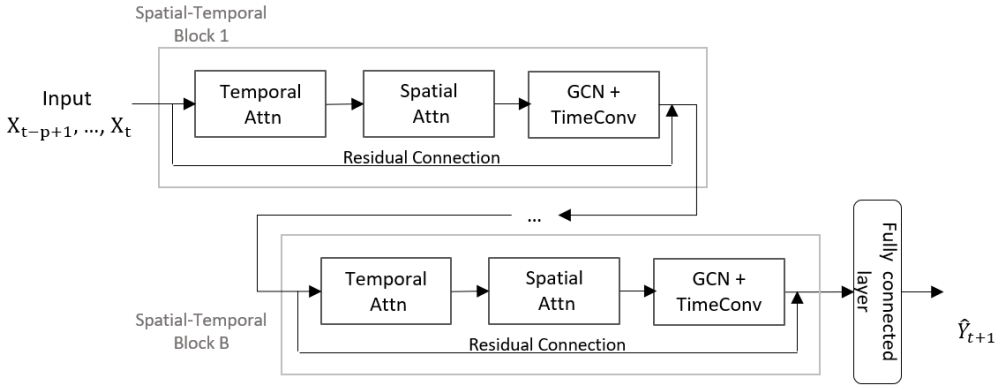


Figure 5-4: The ASTGCN framework

5.3.5 Baseline Models

In order to compare the efficiency and performance of graph neural networks in the classification of price movement directions, following non-graph deep learning models are utilized as baselines:

- Long Short-term Memory Networks (LSTM) Model: LSTM is a modified form of RNN architectures designed to retain crucial details from time series sequences (Lin et al., 2022). An LSTM unit is composed of an input gate, a forget gate, and an output gate, which regulate the flow of information and aid in discarding redundant data.
- Convolutional Neural Network (CNN) Model: The convolutional neural network (CNN) is a type of feedforward neural network architecture introduced by Lecun et al. (1998). CNNS excel at capturing localized patterns and characteristics within time series data. The convolutional layers are adept at acquiring filters that can identify distinct temporal patterns. The local perspective and parameter sharing in CNN contribute to parameter reduction, thereby enhancing the model's learning efficiency (Lu et al., 2020). However, they encounter challenges like the need for inputs of fixed lengths, disregard for temporal sequencing, and a restricted capability to capture long-term temporal relationships.
- Temporal Convolutional Network (TCN) Model: The temporal convolutional network (TCN) (Lea et al., 2016) represents an adaptation of the CNN architecture. TCN utilizes causal and dilated convolutions, which are well-suited for sequential data containing temporal patterns and requiring extensive receptive fields. The term "causal" implies that no data flows from the future to the past.

5.3.6 Evaluation Criteria

To evaluate the classification performance, this paper adopts the following two metrics to calculate the prediction error: accuracy and F1 score. Accuracy is a metric in classification tasks to measure the effectiveness of a model in correctly predicting the classes of a dataset. It measures the proportion of correctly predicted instances among the total instances in the dataset. The F1 score combines both precision and recall into a single value and provides a balanced measure of a model's performance. Precision represents how many of the predicted positive instances were actually positive. Recall indicates the proportion of correctly predicted actual positive instances. The formula of the above evaluation criteria is as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{5-21}$$

$$Precision = \frac{TP}{TP+FP}, \qquad Recall = \frac{TP}{TP+FN} \tag{5-22), (5-23}$$

$$F1\ Score = 2 * \frac{Recall*Precision}{Recall+Precision} \tag{5-24}$$

Where, TP is the number of true positive ($y = 1$) predictions, TN is the number of true negative ($y = 0$) predictions, FP is the number of False positive predictions, and FN is the number of False negative predictions.

5.4 Data Description

The dataset used in this study encompasses a comprehensive range of financial and economic variables spanning multiple markets and sectors. The data covers a period from July 12, 2001, to December 28, 2022, allowing for a thorough investigation of trends and patterns over more than two decades. The dataset includes a diverse set of variables, consisting of commodity markets prices, stock indices, exchange rates, global macroeconomic indicators, supply and demand metrics, and a variety of technical indicators. We have eliminated dates from our dataset where certain variables lacked values, retaining only those dates where all variables had recorded values. The description of each feature in dataset is as following:

Commodity markets: the daily spot price of five prominent commodity markets, namely WTI, Brent, Natural Gas, Gold, and Silver, are included in the dataset. These crude oil, natural gas, and precious metal markets are particularly important for governments, economies, and industries due to their intrinsic value and their impact on other markets, trade, and global relations. In this study, we aim to predict the direction of future price movements of our target markets, WTI, Brent, Gold, and Silver.

Global stock indices: The dataset encompasses the daily closing price data from nine key global stock indices, reflecting the performance of diverse segments of the stock market. These indices include Standard and Poor's 500 (S&P500), Dow Jones Industrial Average (DJIA), National Association of Securities Dealers Automated Quotations (NASDAQ), S&P Energy, Standard and Poor's Goldman

Sachs Commodity Index (S&P GSCI), NYSE Arca Oil and Gas Index (XOI), London stock exchange index (FTSE100), Barrick Gold Corporation stock, and Tadawul All Shares Index.

Exchange rates: Five major daily exchange rates are incorporated into the dataset, tracking the fluctuations between USD and other currencies: USD/EUR, USD/CNY, USD/JPY, USD/SAR, and USD/QAR.

Global macroeconomics: The dataset features three critical global macroeconomic indicators: US Federal Funds Effective Rate, US unemployment rate, and US Consumer Price Index (CPI). Given that the initial frequency of US unemployment rate and Consumer Price Index was on a monthly basis, we employed linear interpolation techniques to generate daily data points. These indicators offer insights into the broader economic conditions that impact financial markets.

Supply and demand factors: Three indicators related to crude oil, gold, and silver supply and demand dynamics are included: world petroleum production rate, world petroleum consumption rate, and US Gold and Silver Mining Index. These indicators shed light on the availability and utilization of key resources.

Technical Indicators: To investigate the effectiveness of technical indicators in predicting the direction of price movements, this study defines some technical indicators related to our target markets. Our choices of technical indicators are based on prior studies on the effect of technical indicators in predicting the crude oil or precious metals prices (Shin et al., 2013).We define seven momentum and trend technical indicators for WTI, Brent, Gold, and Silver markets. These comprehensive set of 28 technical indicators, distributed across the four target markets represent various aspects of price movements and trading patterns. The technical indicators in this study are Rate of Change (ROC), Relative Strength Index (RSI), three Exponential Moving Averages (EMA), Moving Average Convergence/Divergence (MACD), and Triple Exponential Moving Average (TEMA). ROC is Calculated over a 5-day look-back period and provides historical performance context that aids in target labeling. RSI is computed over a 14-day period, RSI gauges the speed and magnitude of recent price changes. Three EMAs are incorporated with look-back periods of 5 days, 50 days, and 100 days, offering insights into varying degrees of short-term, medium-term, and long-term trends. MACD is derived from the 26-period and 12-period EMAs, MACD facilitates the identification of potential trend changes. TEMA is computed over a 20-day period, TEMA enhances trend analysis by providing a smoothed moving average.

Table 5-1 presents the origins of data for each feature in our dataset.

Table 5-1: Data Sources

Feature	Data Source
Daily price index of S&P500, DJIA, NASDAQ, S&P Energy, S&P GSCI, FTSE100, XOI, Barrick Gold Corporation stock, Tadawul All Shares Index	Yahoo Finance
Daily spot price of WTI, Brent, and Natural Gas; monthly world petroleum production and consumption rates	http://www.eia.gov
Daily spot price of Gold, and Silver	https://www.kitco.com

Data Partitioning

To ensure a robust evaluation of model performance, the dataset is divided into three distinct sets: training, validation, and test. The training data which includes the first 65% of the dataset, covering the period from July 12, 2001, to February 4, 2016, is allocated for model training, enabling the models to learn patterns and relationships present in the historical data. The validation data contains 15% of the dataset from February 5, 2016, January 17, 2019, and is reserved for validation, enabling model tuning and hyperparameter optimization. The test data is the remaining 20% of the dataset, covering the period from January 18, 2019, to December 28, 2022, serves as the test set. This subset is utilized to assess the models' generalization capabilities and predictive accuracy on unseen data.

Utilizing the collected data and applying the methodologies outlined in Section 3, we conduct an empirical analysis aimed at addressing the issue outlined in Definition 3.1. This analysis is elaborated upon in Section 5.

5.5 Empirical Analysis and Results

In this section, we delve into data preprocessing techniques, the experimental setup, and the presentation of the achieved results. By meticulously analyzing the outcomes, we aim to provide valuable insights into the model's performance and its potential implications for the classification of the direction-of-price-movement in WTI, Brent, Gold, and Silver markets.

Data Preprocessing: Input normalization in time series prediction tasks is a vital preprocessing step that involves scaling the input data to a common range. By normalizing the data, the model can better capture patterns and relationships, mitigate the effects of varying scales, and ensure stable convergence during training. In this study, the input data is normalized into [0,1] using Eq. (4-23).

where, x_t is the original-scale sample and x_t' denote the scaled data.

For each target market, the label is determined according to Eq. (5-25(5-25), where Y_{t+1} is the actual label at time t+1, X_{t+1} is the closing price at time t+1, EMA_t is the 9-day exponential moving average trend, label 1 indicates the up trend and label 0 represents the down trend. The labeling method applied here is similar to the approach utilized in prior studies, such as (Shin et al., 2013) and (Kia et al., 2018).

$$Y_{t+1} = \begin{cases} 1(up), & if\ X_{t+1} \geq EMA_t\ , \\ 0\ (down) & otherwise \end{cases} \tag{5-25}$$

$$EMA_t = \alpha X_t + (1-\alpha) EMA_{t-1} \tag{5-26}$$

$$\alpha = 2/(span+1), \quad \text{for } span \geq 1 \tag{5-27}$$

This labeling method based on the Exponential Moving Average (EMA) relates to the EMA Crossover trading strategy which is a type of trend-following strategy used in technical analysis. It suggests buying when the price of an asset is above the EMA and selling when the price is below the EMA. By employing this labeling approach, we have successfully generated an almost balanced dataset, as illustrated in Table 5-2.

Table 5-2: Balanced training set

	WTI	Brent	Gold	Silver
Class 0 (down-trend) in training samples	1059 (45.5%)	1071 (46%)	1027 (44.1%)	1087 (46.7%)
Class 1 (up-trend) in training samples	1270 (54.5%)	1258 (54%)	1302 (55.9%)	1242 (53.3%)

The final preprocessing step is to create input sequences using a specific window size. Given the inherent reliance of time series data on past observations, capturing this interdependency is important for precise predictions. Furthermore, the input sequences should properly reflect temporal trends and patterns present within the data. Hence, the selection of an appropriate input sequence length requires careful consideration. This is accomplished through the hyperparameter tuning process.

Experimental setup: The hyperparameters of our classification graph neural networks is optimized throughout three stages: graph construction, model architecture creation, and the training stage. In this study, we have used the Neural Network Intelligence (NNI) toolkit[3] from Microsoft with Tree-structured Parzen Estimator (TPE) tuning algorithm to optimize the hyperparameters of our GNN models on the validation dataset. The TPE tuner belongs to the Bayesian optimization family and is particularly effective for optimizing hyperparameters when the search space is complex and nonlinear. TPE works by modeling the relationship between hyperparameters and the model's performance using probability distributions. It divides the search space into two parts: one for exploring (trying out new parameter values) and the other for exploiting (focusing on regions that have shown promising results). This division allows TPE to iteratively refine its search to converge towards optimal hyperparameters efficiently (Bergstra et al., 2011).

Since we utilize the KNN technique to construct the adjacency matrix of SGA-TCN and ASTGCN models, it becomes crucial to ascertain the optimal value of number of neighbors (k). To achieve this, we conducted a hyperparameter tuning process and found that the most suitable value for k is 18 when N=25 (i.e. when technical indicators are not included in input features). In the case of N=53 (i.e. when 28 technical indicators of our target markets are integrated to the input features), the optimal value of $k = 30$. For neighboring nodes, the adjacency value A_{ij} is set to one, while non-neighbor nodes are assigned a value of A_{ij} equal to 0.

[3] https://github.com/microsoft/nni

In model construction stage, we have used NNI toolkit to find the best hyperparameters for each ST-GNN architecture and baseline models. Moreover, we found the input sequence window size is a significant factor in the classification performance of SGA-TCN, MTGNN-TAttLA, and ASTGCN models. Thus, we included input sequence length as a hyperparameter of our models and search for an optimal window size. Table 5-3 present the value of optimal hyperparameters in our search space for each ST-GNN model.

For the other hyperparameters of MTGNN-TAttLA and MTGNN-LA models, we have used the values of $\alpha = 3$, and $\beta = 0.05$, and $k = 18$ (subgraph size) in Eq.(5-8) and Eq.(5-11), respectively, same as the original MTGNN paper.

Table 5-3: ST-GNNs hyperparameters

Models	Window size	layers	attention heads	Node embedding size	Output layer channel size	Convolutional channel size	GCN depth	Dilation factor
MTGNN-TAttLA	20	2	4	32	64	64	2	2
MTGNN-LA	10	2	-	16	128	32	1	2
SGA-TCN	Window size	TCN layers	attention heads	$[C_{tcn_1}, C_{tcn_2}, C_{tcn_3}]$	TCN blocks	TCN kernel size	GATConv layers	Dilation factor
	20	3	8	[16, 64, 64]	4	5	1	2
ASTGCN	Window size	blocks	Temp. conv. channels	Graph conv. channels	GCN depth			
	30	2	64	16	5			

All ST-GNN models in this study are implemented using Pytorch. The experiments are conducted by using Python 3.10.12 and run on a computing system with a 70 W Tesla T4 NVIDIA-SMI GPU, CUDA version 12.0, and 16GB RAM.

We have trained all ST-GNN models for 40 epochs and using a batch size of 32. We have utilized Adam optimizer (Kingma & Ba, 2014) for minimizing the loss function in Eq. (5-1), with a learning rate equal to 0.005, and a regularization penalty equal to 0.0001. To avoid any potential gradient exploding, we have clipped the gradient values to 5.

This study incorporates model regularization through the integration of dropout layers within the models' architectures and the implementation of $\boldsymbol{L2}$ regularization. Using a validation dataset also enhances regularization by allowing us to fine-tune model parameters and hyperparameters based on its performance. This helps in preventing overfitting, as the model's generalization ability is assessed on data it hasn't seen during training. Regularization techniques such as dropout layers and $\boldsymbol{L2}$ regularization work in conjunction with the validation dataset to ensure the model doesn't become overly specialized to the training data. To further enhance the reliability of the classification outcomes, we repeat the experiments 10 times and report the average value of accuracies and F1 scores for each model.

Results and Discussions: using the above experimental setups, we predicted the one-day ahead up and down price trends by SAG-TCN, MTGNN-TAttLA, and ASTGCN models and compared their

classification ability with non-graph deep learning models for sequential data such as TCN, CNN, and LSTM models. To find the effectiveness of the temporal attention mechanism in enhancing the accuracy of classification, we also compare the MTGNN-TAttLA model with the MTGNN-LA model which does not include any attention layer. Table 5-4 presents the accuracy and F1 score of each model in classification of one-day ahead price movements in WTI, Brent, Gold, and Silver markets.

Table 5-4: Classification performance using 25 features (N=25)

Models	WTI		Brent		Silver		Gold	
	Acc	F1	Acc	F1	Acc	F1	Acc	F1
MTGNN-LA	0.8317	0.8642	0.8022	0.8307	**0.7924**	0.7944	0.7727	0.7923
SGA-TCN	0.8081	0.8473	0.8052	0.8405	0.7621	0.7764	0.7607	0.7937
ASTGCN	0.8368	0.8712	0.8089	0.8431	0.7838	**0.803**	**0.7922**	**0.8265**
MTGNN-TAttLA	**0.8492**	**0.8798**	**0.8179**	**0.8479**	0.7837	0.794	0.7439	0.7619
TCN	0.7858	0.8242	0.7719	0.8066	0.7454	0.7601	0.7315	0.7643
CNN	0.6764	0.7149	0.615	0.6123	0.5495	0.5405	0.5844	0.6534
LSTM	0.6073	0.7444	0.4993	0.5452	0.5718	0.6446	0.5619	0.705

Results from Table 5-4 show that MTGNN-TAttLA is the best performing model in classifying the price movement direction of WTI and Brent Crude oil with an accuracy of 84.92% and 81.79% on the test set, respectively. However, this model does not outperform other models in precious metal markets. One reason for this could be that, as shown in Figure 5-5, the training set of WTI and Brent markets demonstrate more dynamics and complex patterns compared to the training set of Silver and Gold. Thus, adding an attention mechanism to the model assists the model to focus more on learning from important time steps, while for the Silver and Gold markets adding attention mechanism does not make any significant improvement and rather increases the complexity and number of parameters in model.

For classification of price movement direction in Silver and Gold markets, the MTGNN-LA and the ASTGCN are the best ones with accuracies of 79.24% and 79.22%, respectively.

Table 5-5 summarizes whether the ST-GNN models are able to learn the adjacency matrix or they require a predefined adjacency matrix and whether the ST-GNN model uses temporal or spatial attention mechanism in its architecture.

Table 5-5: ST-GNN architecture properties

Models	Learns matrix A	Has Spatial Attention	Has Temporal Attention	Graph model
MTGNN-TAttLA	ü	û	ü	ü
MTGNN-LA	ü	û	û	ü
SGA-TCN	û	ü	û	ü
ASTGCN	û	ü	ü	ü

To understand the effectiveness of adding a temporal attention module to a ST-GNN model, we compare the accuracy of SGA-TCN and ASTGCN. Both these models use graph convolutional networks with spatial attention as their spatial module and temporal convolutional networks as their temporal module. However, the ASTGCN model also incorporates a temporal attention mechanism to its architecture. As the accuracies of predicted direction-of-price-movement from Table 5-4 show, ASTGCN is outperforming the SGA-TCN model across all four target markets (i.e., WTI, Brent, Silver, and Gold). This demonstrates that a temporal attention mechanism can improve the classification performance of a spatial-temporal graph neural network.

Comparing the accuracies of non-graph TCN, CNN, and LSTM models with those of graph-type, MATGNN-LA, SGA-TCN, MTGNN-TAttLA, and ASTGCN, our results show that spatial-temporal graph neural networks significantly outperform traditional deep learning models in classification of WTI, Brent, Gold, and Silver price movement directions. This showcases the importance of finding the spatial relationship between features that can affect the crude oil and precious metals price dynamics. The advantages of defining a multivariate time series input in a graph-structured data is that we can include large number of variables that can affect the target markets and let the graph convolution networks to find the significant relationship and patterns among variables. However, while non-graph TCN, CNN, and LSTM models can extract the interdependencies among variables to some extent, they are not very powerful in discovering the complex relationship among multivariate timeseries. TCN is the best non-graph model for the classification task in hand reaching an accuracy of 78.58%, 77.19%, 74.54%, and 73.15% on WTI, Brent, Silver, and Gold, respectively. The TCN is outperforming the CNN model due to the fact that it uses dilated convolutions, which have a larger receptive field than traditional convolutions. This allows TCNs to capture information over longer spans of time, making them better equipped to model sequences with varying temporal patterns, while CNNs are only able to extract local patterns and dependencies. Similarly, the ability of TCN to capture long-term dependencies in sequences requires less iterative training process and effectively overcomes the vanishing gradient problem often encountered in LSTMs. This makes TCN better at capturing relationships over extended time spans. Moreover, TCN's convolutional filters can capture interactions between variables at different time steps, allowing it to capture complex temporal patterns and relationships.

Figure 5-5: WTI, Brent, Gold and Silver price movements from 2001-04-07 to 2022-12- 28.

To investigate whether adding technical indicators of our target markets will improve the classification performance of our ST-GNNs, we repeated our experiments by including seven technical indicators for each target market into our input feature set represented as a graph. The technical indicators are ROC, RSI, MACD, TEMA, 5-day EMA, 50-day EMA, 100-day EMA. Thus, 28 new features have been added to our data se, totaling to 53 features. Since each feature is a node in our markets' graph, the number of nodes in the new graph will equal to 53. Table 5-6 shows the performance of ST-GNN and non-graph deep learning models in predicting the direction-of-price-movement using 53 features.

Upon comparing the results from Table 5-5 and Table 5-6, it becomes evident that incorporating technical indicator features into our input data significantly enhances the accuracy and F1 scores of nearly all ST-GNN models, particularly in predicting trends for WTI and Brent crude oil markets. Similarly, we arrive at a similar inference regarding the favorable impact of incorporating technical indicators on the classification of Gold price trends, though ASTGCN's performance gain appears more nuanced. Thus, our results suggest that technical indicators perform an important role in predicting the direction-of-price-movement within WTI, Brent, and Gold markets. Moreover, in a broader context, we infer that graph-oriented models exhibit remarkable performance in scenarios with high-dimensional inputs, setting them apart from traditional deep learning models. Nevertheless, even though these models exhibited favorable results for crude oil and gold markets, their effectiveness in predicting Silver price trends did not consistently improve upon the inclusion of technical indicators in our feature dimension. This discrepancy can be attributed to the intricate dynamics of the Silver market, which is influenced by a complex interplay of economic indicators,

industrial demand, and technological advancements. Consequently, the relationship between technical indicators and Silver prices becomes less straightforward, potentially introducing noise and reducing data quality upon integration. Moreover, our observations indicate that the addition of technical indicators to our features set didn't consistently improve the accuracy of CNN and LSTM models. This implies deep learning models' limitations in comprehending complex relationships between variables, especially in cases characterized by a high input dimension.

As highlighted in Table 5-6, using comprehensive set of 53 input features, MTGNN-TAttLA is the top-performing model for classifying the WTI, Brent, and Silver price movement directions. Meanwhile, ASTGCN maintains its superiority in predicting Gold price trend directions among other models. These results confirm that the temporal attention comes handy in classification of timeseries trends even when the input feature size is large.

Table 5-6: Classification performance using 53 features (N = 53)

Models	WTI		Brent		Silver		Gold	
	Acc	F1	Acc	F1	Acc	F1	Acc	F1
MTGNN-LA -TI	0.8387	0.8697	0.8078	0.8355	0.7798	0.7792	0.784	0.8094
SGA-TCN -TI	0.8122	0.8505	0.8081	0.8435	0.7719	0.7859	0.7608	0.7908
ASTGCN-TI	0.8396	0.8718	0.8131	**0.8498**	0.7755	0.7873	**0.7922**	**0.822**
MTGNN-TAttLA-TI	**0.852**	**0.8844**	**0.8165**	0.8455	**0.7838**	**0.7895**	0.7738	0.8
TCN-TI	0.7899	0.8278	0.7774	0.8126	0.7399	0.7581	0.7343	0.7696
CNN-TI	0.6248	0.6258	0.6192	0.6245	0.537	0.4924	0.5941	0.6643
LSTM-TI	0.6401	0.6903	0.6216	0.6624	0.5358	0.5154	0.5064	0.4718

Figure 5-6 presents the heatmap of the adjacency matrix learned by MTGNN-TAttLA using 25 features in the model. The matrix values signify the level of spatial connections among variables, and the importance of each variable for the classification tasks undertaken in this study. For a variable in row i, A_{ij} shows the importance of the j^{th} column's variable in classifying the movement direction of the i^{th} variable. In this context, the most relevant variables for our designated targets – WTI, Brent, Silver, and Gold – are detailed in Table 5-7.

Table 5-7: Connected variables to target markets

Market	Connected variables
WTI	Brent, DJIA, FTSE100, Gas, NASDAQ, S&P500, Silver, Tadawul, US CPI, USDJPY, WPetCons, WPetPro, XOI
Brent	DJIA, Gold, GS mining, S&P Energy, S&P GSCI, S&P500, Silver, Tadawul, UnempRate, US CPI, USDCNY, USDJPY, USDQAR
Silver	DJIA, Gold, GS mining, S&P Energy, S&P GSCI, S&P500, Tadawul, UnempRate, US CPI, USDCNY, USDJPY, USDQAR
Gold	Barrick, FTSE100, Gas, GS mining, NASDAQ, S&P Energy, S&P GSCI, USDCNY, USDEUR, USDQAR, USDSAR, WTI

These relationships suggest that these markets are interconnected and influenced by diverse economic, supply and demand, financial, and commodity-related factors. As an example, the relationship between NASDAQ, which represents technology and growth-oriented companies, and WTI and Gold could indicate that during periods of market uncertainty or economic instability, investors might allocate funds to WTI or Gold as a protective measure against potential losses in the equity markets. Likewise, London Stock Exchange index (FTSE100) demonstrates stronger correlations with WTI and Gold markets in comparison to Brent or Silver markets. The relationship among the variables in Table 5-7 are complicated and cannot be interpreted within the scope of the current study. One needs more through analysis for this purpose which we leave for future studies. The variables in Table 5-7 offer valuable insights into the potential factors that could influence the price dynamics of WTI, Brent, Silver, and Gold markets. Hence, these relationships should be carefully taken into account when conducting analyses or formulating predictions regarding these markets.

Figure 5-6: Heat map of the learned adjacency matrix

Figure 5-7 illustrates box plots for the distribution of accuracies across all spatial-temporal neural networks employed for classifying the short-term price trend directions in WTI, Brent, Silver and Gold markets. On the x-axis of the plots, we have denoted the classification performance of models incorporating technical indicators alongside 25 other features as "Market-TI". For example, in the box plot for Gold, the "Gold-TI" notation signifies the accuracy of predictions for Gold price movement direction using 53 factors, including 28 technical indicators in addition to other 25 inputs. As it is evident from Figure 5-7, the inclusion of technical indicators for the respective targets led to higher accuracies in the case of Brent, Gold, and WTI. Additionally, the dispersion of accuracy values across all markets becomes narrower with the incorporation of technical indicators. This highlights the enhancing influence of technical indicators on the precision of price movement direction classification through ST-GNNs. In general, we have achieved higher accuracies for crude oil markets compared to the precious metals markets. This divergence might be attributed to various underlying factors. The set of features used for prediction might have a stronger predictive power

for crude oil markets compared to precious metals markets. Given that the ST-GNN models are fine-tuned to classify labels across all multivariate features, the model complexities might be better suited to capturing the underlying patterns of crude oil rather than precious metals markets. Finally, as presented in Figure 5-5, the training dataset within the crude oil markets encompasses numerous diverse temporal patterns, making the predictive models more equipped to learn and generalize from historical data. In contrast, the temporal patterns in precious metals typically adhere to long-term trends, which might deviate from the trends present in the test set. This characteristic make them less suitable for effectively classifying short-term, one-day ahead price movements in precious metals markets.

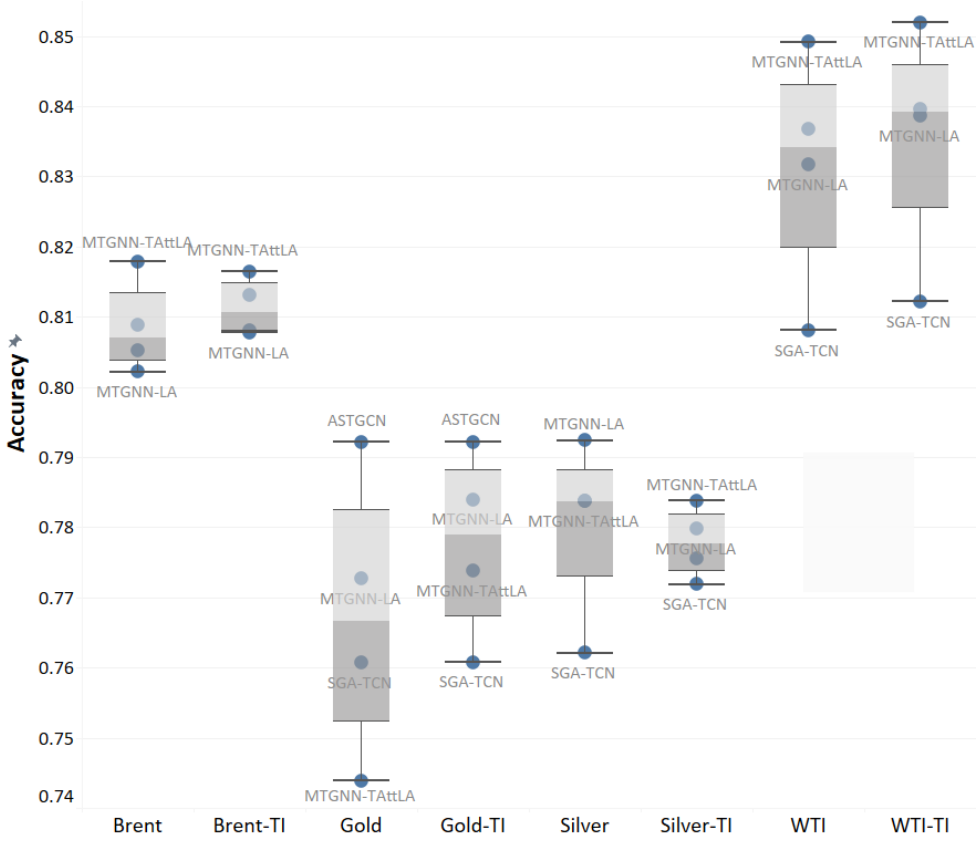


Figure 5-7: Accuracy distributions

Figure 5-8 shows the training and validation losses for all four spatial-temporal neural networks across forty epochs. Notably, MTGNN-LA, MTGNN-TAttLA, and ASTGCN exhibit quicker loss convergence and achieve lower training losses compared to SGA-TCN. Moreover, MTGNN-LA and MTGNN-TAttLA have reached the lowest validation error, underscoring their superior performance in time series classification tasks.

Figure 5-8: Training and validation losses

5.6 Conclusion

The crude oil market has historically been susceptible to the influence of geopolitical and macroeconomic factors. Precious metals, on the other hand, are investment assets that have significant implications in various industries. Additionally, gold serves as a substitute asset for traditional stock markets and is a fundamental component of financial investment portfolios. Given these factors, the development of accurate predictive models for determining the future direction of crude oil, gold, and silver prices is important. This study addressed the formidable challenge of predicting price movement directions in the crude oil, gold, and silver markets. To tackle this challenge, we employ three spatial-temporal graph neural network models, namely MTGNN-TAttLA, SAG-TCN, and ASTGCN. These models use several variables such as historical prices of crude oil, silver, and gold markets, global stock market indices, exchange rates, supply and demand-related factors, global macroeconomic factors, and technical indicators as predictors. Our findings underscore the efficacy of these models in capturing the intricate interplay of spatial and temporal dependencies inherent in market data. The incorporation of an attention mechanism, particularly in the MTGNN-TAttLA model, yielded exceptional prediction accuracy. Notably, all three models surpassed conventional deep learning approaches like TCN, LSTM and CNN in forecasting accuracy. Particularly, MTGNN-TAttLA demonstrated outstanding performance in predicting price movement directions for WTI, Brent, and silver markets, while ASTGCN excelled in forecasting gold market trends. These results carry substantial implications for investors, financial institutions, and policymakers, as they empower informed decision-making, effective risk management, and portfolio optimization.

In this research, several limitations need to be acknowledged. First, our analysis relies on historical data, and market conditions are subject to change. Future research should consider incorporating real-time data and exploring the adaptability of these models to evolving market dynamics. Additionally, further investigation into the generalization of these models to other commodities and

financial assets is warranted. Future research can explore additional optimization techniques for refining the network structure and parameter settings. Finally, other variables that might affect the direction of crude oil and precious metals market movements ought to be considered in the graph network construction. For example, investors' sentiments from social media or geopolitical events can be incorporated into the predictors' set.

Chapter 6

Conclusion

Crude oil, in particular WTI and Brent, perform a crucial role in the global financial markets and market economics. In recent years, the price of crude oil has been more vulnerable to geopolitical and macroeconomic factors. Thus, understanding the dynamics of crude oil markets seems inevitable. Besides, precious metals such as gold and silver are key commodities that are mined in particular countries which make the economy of these countries highly rely on precious metal markets. Moreover, gold is a substitute asset for stock markets and plays an indispensable role in financial investment portfolios. On the other hand, cryptocurrencies are modern digital assets that have been traded for approximately fifteen years to date and still need more investigation under various circumstances.

This research has examined the characteristics and dynamics of cryptocurrency, crude oil, and precious metal markets and developed several timeseries prediction models to forecast the price level, and the direction-of-price-movement in crude oil and precious metal markets. Our study is concluded into four peer-reviewed research articles, three of which are published in prestigious Q1 business journals, and the last one is accepted for publication in Financial Innovation journal.

Our first paper explored the return-volatility relationships in cryptocurrencies, crude oil, and gold markets before and during the COVID-19 pandemic. Moreover, the cryptocurrencies return-volume relations are investigated across both periods. For this, we have implemented econometrics and statistical models such as EGRACH-in-Mean, VAR, and Granger causality tests. Our empirical analysis indicates no significant return-volatility relationship in any of the cryptocurrencies prior to the COVID-19 pandemic. However, during the pandemic, this relationship became significant for Tether, Ether, Ripple, Bitcoin Cash, EOS, and Monero. The effect of volatility on returns of Bitcoin, Litecoin, Chainlink, and Cardano is negligible during the pandemic. Thus, traditional risk-return relationships (i.e. higher risks are associated with higher returns) may not hold for these markets. In this situation, investors may need to reassess how they evaluate and manage risk in Bitcoin, Litecoin, Chainlink, and Cardano during financial crisis and adjust their investment strategies, considering other factors that influence returns, such as investment sentiments, or external events. In addition, the return-volatility relationship for WTI and Brent crude oil was significantly negative before the

pandemic but became non-significant during the crisis. Our results from the granger causality tests on return-volume relationships suggest the absolute returns of cryptocurrencies significantly influence their trading volumes, indicating that traders engage in higher volumes during periods of significant price fluctuations, a behavior unaffected by the COVID-19 crisis.

The second paper approaches these markets from a different angle and studies the relationship among cryptocurrencies and commodity markets including crude oil and gold. This study searches for the potential safe haven effects of gold or crude oil on cryptocurrency markets under stable and crisis financial market conditions, especially before and during the COVID-19 pandemic. Our results show that, during the COVID-19 pandemic, Gold demonstrated stronger safe haven characteristics compared to crude oil, offering a more effective risk mitigation for cryptocurrency investments. However, crude oil emerged as a robust safe haven for Bitcoin Cash. This study could not find any significant relationship between gold and crude oil markets and Ether, Ripple, EOS, and Cardano across the examined periods. Unlike crude oil, which is a commodity with significant price volatility influenced by geopolitical factors, supply disruptions, and demand fluctuations, gold's price movements are relatively more stable and less influenced by short-term supply-demand dynamics. This stability enhances its attractiveness as a safe haven during uncertain times. The findings from Granger causality tests affirm that, overall, causal relationships predominantly exist from cryptocurrencies toward gold and crude oil markets in both periods. The implications of our findings extend to hedge fund managers and digital currency investors, providing valuable insights to help them strategically balance risk exposures and optimize returns during financial crises.

In the third essay, we approach crude oil and precious metal markets from a predictive perspective. More specifically, we tailor various deep learning models to improve the prediction of the future daily price of these markets and compare the performance of deep learning models with ensemble, and individual machine learning models. Following an extensive literature review, it became evident that proposing a universally applicable model capable of effectively forecasting all markets is impractical due to the unique characteristics of each market. Thus, we examined sixteen deep learning, hybrid, ensemble, and machine learning models with various hyperparameter settings and input lengths to forecast the price of each market. In addition, we implemented a novel Time2Vector embedding to enhance forecasting performance with longer input sequences. Our results indicate that the Temporal Convolutional Networks are the best model among the sixteen models in generalizing and forecasting the commodity market prices. LightGBM is the best machine learning type model in forecasting the commodity market prices. However, compared to the TCN model, it performs poorly in capturing and responding to sharp market dynamics. Regarding the type of deep learning models to forecast crude oil and precious metal prices, our results show that GRU-type RNNs perform better than LSTM-types, and bidirectional RNNS are superior to unidirectional ones. Furthermore, among WTI, Brent, Gold, and Silver, Gold exhibits the highest sensitivity to the input sequence length in price forecasting. Consequently, investors and analysts should exercise greater caution and precision in selecting the appropriate input sequence window when forecasting gold prices. Our findings

provide valuable insights for analysts seeking to improve the accuracy of commodity market price forecasts. It offers guidance for selecting the most suitable models and input parameters for forecasting commodity market prices. This knowledge equips governments, energy sector managers, and crude oil and precious metals investors to make sensible decisions. In a governmental context, crude oil and precious metal price forecasting helps governments in fiscal planning, economic policy decisions, resource allocation, revenue management, international trade negotiations, socioeconomic development, environmental policies, and geopolitical considerations. Accurate forecasts can provide a competitive advantage by enabling managers to make timely and informed decisions. They can anticipate market trends, respond quickly to price fluctuations, and maintain a competitive advantage in terms of pricing, supply chain management, and customer satisfaction.

Lastly, the fourth paper expands the forecasting scope to predict future direction-of-price-movements in crude oil, gold, and silver markets, framing it as a multivariable timeseries classification problem. To address this challenge comprehensively, we incorporate a wide array of historical, economic, financial, supply and demand factors known to influence movements in crude oil and precious metal markets. We establish a feature graph to model relationships among these factors and introduce three innovative spatial-temporal graph neural networks, namely MATGNN-TAttLA, ASTGCN, and SGA-TCN, to classify price movements in the mentioned markets. Moreover, we show the effectiveness of the attention mechanism in improving the accuracy of these models. Our research presents significant empirical findings that provide crucial insights into robust models and influential factors shaping short-term price trend classification in commodity markets. MATGNN-TAttLA emerges as the top-performing model for accurately classifying the direction of one-day ahead price movement in WTI, Brent, and Gold markets, while ASTGCN outperforms other ST-GNNs in the Silver market. The incorporation of technical indicators and a diverse set of features enhances classification accuracy across all ST-GNN models for WTI, Brent, and Gold markets. The inclusion of a temporal attention mechanism further proves advantageous, enhancing classification performance by emphasizing critical patterns and significant input sequences. Notably, spatial-temporal graph neural networks consistently outperform non-graph deep learning models such as TCN, CNN, and LSTM in classifying the direction-of-price-movement across diverse commodity markets. The ability of ST-GNNs to learn spatial relationships among variables makes them particularly well-suited for timeseries prediction applications involving a substantial number of input features. Our findings with improved accuracy of price movement forecasts empower investors, financial institutions, and policymakers to make informed decisions, and optimize their portfolios. Beyond practical applications, these models contribute to a deeper understanding of the dynamics in crude oil and precious metals markets and provide insights into the factors influencing price trends.

A summary of research papers in this work is presented in Table 6-1.

Table 6-1: Summary of four papers

Paper	Methodology	Input data	Data period	Evaluation metrics
Return-Volume, Return-Volatility relationships	ARMA, EGARCH in Mean, Granger Causality	Daily Historical returns, volumes, and volatilities	2019-01-01 to 2020-12-31	Statistical analysis
Connectedness between cryptocurrency, crude oil, and gold	VAR, VECM, ARDL, Franger Causality	Daily Historical Prices	2019-01-01 to 2020-12-31	Statistical analysis
Forecasting Crude oil and precious metals prices	Sixteen deep learning and machine learning models	Historical spot prices	2021-01-01 to 2021-12-31	MSE, MAE, RMSE
Price movement prediction in Crude oil and precious metals markets	Three Spatial-Temporal Graph Neural Networks, and Deep learning models such as LSTM, CNN, and TCN	Historical spot prices, global economic factors, supply-demand factors, other financial markets, technical indicators	2001-07-12 to 2022-12-28	Accuracy, F1-score

Bibliography

Abdullah Ahmed, R., & Bin Shabri, A. (2014). Daily Crude Oil Price Forecasting Model Using Arima, Generalized Autoregressive Conditional Heteroscedastic and Support Vector Machines. *American Journal of Applied Sciences*, *11*(3), 425–432.

Abraham, R., Samad, M. El, Bakhach, A. M., El-Chaarani, H., Sardouk, A., Nemar, S. El, & Jaber, D. (2022). Forecasting a Stock Trend Using Genetic Algorithm and Random Forest. *Journal of Risk and Financial Management*, *15*(5), 188.

Adekoya, O. B., Akinseye, A. B., Antonakakis, N., Chatziantoniou, I., Gabauer, D., & Oliyide, J. (2022). Crude oil and Islamic sectoral stocks: Asymmetric TVP-VAR connectedness and investment strategies. *Resources Policy*, *78*, 102877.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Akbar, M., Iqbal, F., & Noor, F. (2019). Bayesian analysis of dynamic linkages among gold price, stock prices, exchange rate and interest rate in Pakistan. *Resources Policy*, *62*, 154–164.

Akhtaruzzaman, M., Boubaker, S., Lucey, B. M., & Sensoy, A. (2021). Is gold a hedge or a safe-haven asset in the COVID–19 crisis? *Economic Modelling*, *102*, 105588.

Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. *Resources Policy*, *61*, 250–260.

Almeida, F., & Xexéo, G. (2019). Word Embeddings: A Survey. *ArXiv*.

Andreoletti, D., Troia, S., Musumeci, F., Giordano, S., Maier, G., & Tornatore, M. (2019). Network Traffic Prediction based on Diffusion Convolutional Recurrent Neural Networks. *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 246–251.

Arbane, M., Benlamri, R., Brik, Y., & Alahmar, A. D. (2023). Social media-based COVID-19 sentiment classification model using Bi-LSTM. *Expert Systems with Applications*, *212*, 118710.

Baek, C. (2019). How are gold returns related to stock or bond returns in the U.S. market? Evidence from the past 10-year gold market. *Applied Economics*, *51*(50), 5490–5497.

Baek, S., Mohanty, S. K., & Glambosky, M. (2020). COVID-19 and stock market volatility: An industry level analysis. *Finance Research Letters*, *37*, 101748.

Bai, J., Zhu, J., Song, Y., Zhao, L., Hou, Z., Du, R., & Li, H. (2021). A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting. *ISPRS International Journal of Geo-Information*, *10*(7), 485.

Bai, L., Yao, L., Li, C., Wang, X., & Wang, C. (2020). Adaptive Graph Convolutional Recurrent Network for Traffic Forecasting. *ArXiv*.

Bai, Y., Li, X., Yu, H., & Jia, S. (2022). Crude oil price forecasting incorporating news text. *International Journal of Forecasting*, *38*(1), 367–383.

Barson, Z., Junior, P. O., Adam, A. M., & Asafo-Adjei, E. (2022). Connectedness between Gold and Cryptocurrencies in COVID-19 Pandemic: A Frequency-Dependent Asymmetric and Causality Analysis. *Complexity*, *2022*, 1–17.

Bashiri Behmiri, N., & Pires Manso, J. R. (2013). Crude Oil Price Movements and Determinant Factors: A Historical Overview. *SSRN Electronic Journal*.

Baur, D. G., Dimpfl, T., & Kuck, K. (2021). Safe Haven Assets - The Bigger Picture. *SSRN Electronic Journal*.

Baur, D. G., & Lucey, B. M. (2010). Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. *Financial Review*, *45*(2), 217–229.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance*, *34*(8), 1886–1898.

Będowska-Sójka, B., & Kliber, A. (2022). Can cryptocurrencies hedge oil price fluctuations? A pandemic perspective. *Energy Economics*, *115*, 106360.

Behrendt, S., & Schmidt, A. (2021). Nonlinearity matters: The stock price – trading volume relation revisited. *Economic Modelling*, *98*, 371–385.

Ben Khelifa, S., Guesmi, K., & Urom, C. (2021). Exploring the relationship between cryptocurrencies and hedge funds during COVID-19 crisis. *International Review of Financial Analysis*, *76*.

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for Hyper-Parameter Optimization. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 24). Curran Associates, Inc.

Bhowmik, R., & Wang, S. (2020). Stock Market Volatility and Return Analysis: A Systematic Literature Review. *Entropy*, *22*(5), 522.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economic Statistics Section* , 177–181.

Boongasame, L., Viriyaphol, P., Tassanavipas, K., & Temdee, P. (2022). Gold-Price Forecasting Method Using Long Short-Term Memory and the Association Rule. *Journal of Mobile Multimedia*.

Borisov, V., Leemann, T., Seßler, K., Haug, J., Pawelczyk, M., & Kasneci, G. (2021). *Deep Neural Networks and Tabular Data: A Survey*.

Box George E.P., Jenkins Gwilym M., Reinsel Gregory C., & Ljung Greta M. (2015). *Time series analysis: forecasting and control* (fifth). John Wiley & Sons.

Brauwers, G., & Frasincar, F. (2023). A General Survey on Attention Mechanisms in Deep Learning. *IEEE Transactions on Knowledge and Data Engineering*, *35*(4), 3279–3298.

Brody, S., Alon, U., & Yahav, E. (2021). How Attentive are Graph Attention Networks? *ArXiv*.

Bruna, J., Zaremba, W., Szlam, A., & LeCun, Y. (2013). Spectral Networks and Locally Connected Networks on Graphs. *ArXiv*.

Buczkowski, P. (2017). *Predicting Stock Trends Based on Expert Recommendations Using GRU/LSTM Neural Networks* (pp. 708–717).

Caporale, G. M., Spagnolo, F., & Spagnolo, N. (2016). Macro news and stock returns in the Euro area: A VAR-GARCH-in-mean analysis. *International Review of Financial Analysis*, *45*, 180–188.

Chaudhary, M., Sodani, P. R., & Das, S. (2020). Effect of COVID-19 on Economy in India: Some Reflections for Policy and Programme. *Journal of Health Management*, *22*(2), 169–180.

Chen, H., Zhang, J., Tao, Y., & Tan, F. (2019). Asymmetric GARCH type models for asymmetric volatility characteristics analysis and wind power forecasting. *Protection and Control of Modern Power Systems*, *4*(1), 29.

Chen, L., Chen, D., Shang, Z., Wu, B., Zheng, C., Wen, B., & Zhang, W. (2022). Multi-Scale Adaptive Graph Neural Network for Multivariate Time Series Forecasting. *ArXiv*.

Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2014). Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. *ArXiv*.

Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System*.

Chen, W., Jiang, M., Zhang, W. G., & Chen, Z. (2021). A novel graph convolutional feature based convolutional neural network for stock trend prediction. *Information Sciences*, *556*, 67–94.

Chen, Y., Wei, Z., & Huang, X. (2018). Incorporating Corporation Relationship via Graph Convolutional Neural Networks for Stock Price Prediction. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 1655–1658.

Cheung, Y.-W., & Lai, K. S. (1995). Lag Order and Critical Values of the Augmented Dickey-Fuller Test. *Journal of Business & Economic Statistics*, *13*(3), 277.

Cho, K., van Merrienboer, B., Bahdanau, D., & Bengio, Y. (2014). *On the Properties of Neural Machine Translation: Encoder-Decoder Approaches*.

Christie, A. (1982). The stochastic behavior of common stock variances Value, leverage and interest rate effects. *Journal of Financial Economics*, *10*(4), 407–432.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*.

Conlon, T., & McGee, R. (2020). Safe haven or risky hazard? Bitcoin during the Covid-19 bear market. *Finance Research Letters*, *35*, 101607.

Copeland, T. E. (1976). A Model of Asset Trading Under the Assumption of Sequential Information Arrival. In *Source: The Journal of Finance* (Vol. 31, Issue 4).

Corbet, S., Hou, Y., Hu, Y., Larkin, C., Lucey, B., & Oxley, L. (2021). Cryptocurrency liquidity and volatility interrelationships during the COVID-19 pandemic. *Finance Research Letters*.

Corbet, S., Larkin, C., & Lucey, B. (2020). The contagion effects of the COVID-19 pandemic: Evidence from gold and cryptocurrencies. *Finance Research Letters*, *35*, 101554.

Creti, A., Joëts, M., & Mignon, V. (2013). On the links between stock and commodity markets' volatility. *Energy Economics*, *37*, 16–28.

Dablander, F., & Hinne, M. (2019). Node centrality measures are a poor substitute for causal inference. *Scientific Reports*, *9*(1), 6846.

Das, S., Nayak, J., Kamesh Rao, B., Vakula, K., & Ranjan Routray, A. (2022). *Gold Price Forecasting Using Machine Learning Techniques: Review of a Decade* (pp. 679–695).

Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. *ArXiv*.

Deng, S., Xiang, Y., Fu, Z., Wang, M., & Wang, Y. (2019). A hybrid method for crude oil price direction forecasting using multiple timeframes dynamic time wrapping and genetic algorithm. *Applied Soft Computing Journal*, *82*.

Deng, S., Xiao, C., Zhu, Y., Peng, J., Li, J., & Liu, Z. (2023). High-frequency direction forecasting and simulation trading of the crude oil futures using Ichimoku KinkoHyo and Fuzzy Rough Set. *Expert Systems with Applications*, *215*, 119326.

Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv*.

Disli, M., Nagayev, R., Salim, K., Rizkiah, S. K., & Aysan, A. F. (2021). In search of safe haven assets during COVID-19 pandemic: An empirical analysis of different investor types. *Research in International Business and Finance*, *58*, 101461.

Drachal, K. (2022). Forecasting the Crude Oil Spot Price with Bayesian Symbolic Regression. *Energies*, *16*(1), 4.

Drake, P. P. (2022). The gold-stock market relationship during COVID-19. *Finance Research Letters*, *44*, 102111.

Dutta, A., Bouri, E., & Roubaud, D. (2019). Nonlinear relationships amongst the implied volatilities of crude oil and precious metals. *Resources Policy*, *61*, 473–478.

Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar – A GARCH volatility analysis. *Finance Research Letters*, *16*, 85–92.

Elie, B., Naji, J., Dutta, A., & Uddin, G. S. (2019). Gold and crude oil as safe-haven assets for clean energy stock indices: Blended copulas approach. *Energy*, *178*, 544–553.

Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, *55*(2), 251.

Epps, T. W., & Epps, M. L. (1976). *The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis* 44(2).

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, *25*(2), 383.

Fang, Y., Wang, W., Wu, P., & Zhao, Y. (2023). A sentiment-enhanced hybrid model for crude oil price forecasting. *Expert Systems with Applications*, *215*, 119329.

Feng, F., He, X., Wang, X., Luo, C., Liu, Y., & Chua, T.-S. (2018). Temporal Relational Ranking for Stock Prediction. *ArXiv*.

Fisher, R. A. (1925). Theory of Statistical Estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, *22*(5), 700–725.

Galassi, A., Lippi, M., & Torroni, P. (2021). Attention in Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(10), 4291–4308.

Gharghory, S. M. (2021). A Hybrid Model of Bidirectional Long-Short Term Memory and CNN for Multivariate Time Series Classification of Remote Sensing Data. *Journal of Computer Science*, *17*(9), 789–802.

Gono, D. N., Napitupulu, H., & Firdaniza. (2023). Silver Price Forecasting Using Extreme Gradient Boosting (XGBoost) Method. *Mathematics*, *11*(18), 3813.

González, M. de la O., Jareño, F., & Skinner, F. S. (2021). Asymmetric interdependencies between large capital cryptocurrency and Gold returns during the COVID-19 pandemic crisis. *International Review of Financial Analysis*, *76*, 101773.

Goodell, J. W., & Goutte, S. (2021). Co-movement of COVID-19 and Bitcoin: Evidence from wavelet coherence analysis. *Finance Research Letters*, *38*, 101625.

Gopali, S., Abri, F., Siami-Namini, S., & Namin, A. S. (2021). A Comparison of TCN and LSTM Models in Detecting Anomalies in Time Series Data. *2021 IEEE International Conference on Big Data (Big Data)*, 2415–2420.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424.

Gruber, N., & Jockisch, A. (2020). Are GRU Cells More Specific and LSTM Cells More Sensitive in Motive Classification of Text? *Frontiers in Artificial Intelligence*, *3*.

Guo, K., Hu, Y., Qian, Z., Sun, Y., Gao, J., & Yin, B. (2022). Dynamic Graph Convolution Network for Traffic Forecasting Based on Latent Network of Laplace Matrix Estimation. *IEEE Transactions on Intelligent Transportation Systems*, *23*(2), 1009–1018.

Guo, M.-H., Xu, T.-X., Liu, J.-J., Liu, Z.-N., Jiang, P.-T., Mu, T.-J., Zhang, S.-H., Martin, R. R., Cheng, M.-M., & Hu, S.-M. (2022). Attention mechanisms in computer vision: A survey. *Computational Visual Media*, *8*(3), 331–368.

Guo, S., Lin, Y., Feng, N., Song, C., & Wan, H. (2019). Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 922–929.

Haq, A. U., Zeb, A., Lei, Z., & Zhang, D. (2021). Forecasting daily stock trend using multi-filter feature selection and deep learning. *Expert Systems with Applications*, *168*, 114444.

He, P., Liu, X., Gao, J., & Chen, W. (2020, June 5). DeBERTa: Decoding-enhanced BERT with Disentangled Attention. *International Conference on Learning Representations*.

He, Z., Zhou, J., Dai, H.-N., & Wang, H. (2019). Gold Price Forecast Based on LSTM-CNN Model. *2019 IEEE Intl Conf (DASC/PiCom/CBDCom/CyberSciTech)*, 1046–1053.

Hlaváčková-Schindler, K., & Pereverzyev, S. (2015). *Lasso Granger Causal Models: Some Strategies and Their Efficiency for Gene Expression Regulatory Networks* (pp. 91–117).

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780.

Hou, K., Peng, L., Xiong, W., Barberis, N., Brandt, M., Cohen, L., Daniel, K., Vigna, S. della, Hirshleifer, D., Hong, H., Seasholes, M., Simin, T., & Vorkink, K. (2009). A Tale of Two Anomalies: The Implications of Investor Attention for Price and Earnings Momentum.

Huang, W.-Q., Zhuang, X.-T., & Yao, S. (2009). A network analysis of the Chinese stock market. *Physica A: Statistical Mechanics and Its Applications*, *388*(14), 2956–2964.

Huang, Y., Liu, Q., Peng, H., Wang, J., Yang, Q., & Orellana-Martín, D. (2023). Sentiment classification using bidirectional LSTM-SNP model and attention mechanism. *Expert Systems with Applications*, *221*, 119730.

Hussain Shahzad, S. J., Raza, N., Shahbaz, M., & Ali, A. (2017). Dependence of stock markets with gold and bonds under bullish and bearish market states. *Resources Policy*, *52*, 308–319.

Jain, A., Zamir, A. R., Savarese, S., & Saxena, A. (2016). Structural-RNN: Deep Learning on Spatio-Temporal Graphs. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5308–5317.

Jansen, D. W., Engle, R. F., & Granger, C. W. J. (1993). Long-Run Economic Relationships: Readings in Cointegrating. *Southern Economic Journal*, *59*(3), 548.

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, *6*(3), 255–259.

Jiali, X. (2021). Financial Time Series Prediction Based on Adversarial Network Generated by Attention Mechanism. *2021 International Conference on Public Management and Intelligent Society (PMIS)*, 246–249.

Jiang, H., Hu, W., Xiao, L., & Dong, Y. (2022). A decomposition ensemble based deep learning approach for crude oil price forecasting. *Resources Policy*, *78*, 102855.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, *59*(6), 1551

Junttila, J., Pesonen, J., & Raatikainen, J. (2018a). Commodity market based hedging against stock market risk in times of financial crisis: The case of crude oil and gold. *Journal of International Financial Markets, Institutions and Money*, *56*, 255–280.

Junttila, J., Pesonen, J., & Raatikainen, J. (2018b). Commodity market based hedging against stock market risk in times of financial crisis: The case of crude oil and gold. *Journal of International Financial Markets, Institutions and Money*, *56*, 255–280.

Karpoff, J. M. (1987). The Relation Between Price Changes and Trading Volume: A Survey. In *Source: The Journal of Financial and Quantitative Analysis* (Vol. 22, Issue 1).

Katsiampa, P. (2019). Volatility co-movement between Bitcoin and Ether. *Finance Research Letters*, *30*, 221–227.

Kazemi, S. M., Goel, R., Eghbali, S., Ramanan, J., Sahota, J., Thakur, S., Wu, S., Smyth, C., Poupart, P., & Brubaker, M. (2019). Time2Vec: Learning a Vector Representation of Time. *ArXiv*.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc.

Kertlly de Medeiros, R., da Nóbrega Besarria, C., Pitta de Jesus, D., & Phillipe de Albuquerquemello, V. (2022). Forecasting oil prices: New approaches. *Energy*, *238*, 121968.

Khan, M., Wang, H., Riaz, A., Elfatyany, A., & Karim, S. (2021). Bidirectional LSTM-RNN-based hybrid deep learning frameworks for univariate time series classification. *The Journal of Supercomputing*, *77*(7), 7021–7045.

Kia, A. N., Haratizadeh, S., & Shouraki, S. B. (2018). A hybrid supervised semi-supervised graph-based model to predict one-day ahead movement of global stock markets and commodity prices. *Expert Systems with Applications*, *105*, 159–173.

Kim, M., & Sayama, H. (2017). Predicting stock market movements using network science: an information theoretic approach. *Applied Network Science*, *2*(1), 35.

Kim, R., So, C. H., Jeong, M., Lee, S., Kim, J., & Kang, J. (2019). HATS: A Hierarchical Graph Attention Network for Stock Movement Prediction. *ArXiv*.

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *ArXiv*.

Kipf, T. N., & Welling, M. (2016). Semi-Supervised Classification with Graph Convolutional Networks. *ArXiv*.

Klein, T., Pham Thu, H., & Walther, T. (2018). Bitcoin is not the New Gold – A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis*, *59*, 105–116.

Kumar, A. S., & Anandarao, S. (2019). Volatility spillover in crypto-currency markets: Some evidences from GARCH and wavelet analysis. *Physica A: Statistical Mechanics and Its Applications*, *524*, 448–458.

Lahmiri, S., & Bekiros, S. (2020). The impact of COVID-19 pandemic upon stability and sequential irregularity of equity and cryptocurrency markets. *Chaos, Solitons & Fractals*, *138*, 109936.

Lara-Benítez, P., Carranza-García, M., Luna-Romera, J. M., & Riquelme, J. C. (2020). Temporal Convolutional Networks Applied to Energy-Related Time Series Forecasting. *Applied Sciences*, *10*(7), 2322.

Lauren, S., & Harlili, S. Dra. (2014). Stock trend prediction using simple moving average supported by news classification. *2014 International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA)*, 135–139.

Le, T.-H., Le, A. T., & Le, H.-C. (2021). The historic oil price fluctuation during the Covid-19 pandemic: What are the causes? *Research in International Business and Finance*, *58*, 101489.

Lea, C., Flynn, M. D., Vidal, R., Reiter, A., & Hager, G. D. (2016). Temporal Convolutional Networks for Action Segmentation and Detection. *ArXiv*.

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Lee, B.-S., & Rui, O. M. (2002). The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence. In *Journal of Banking & Finance* (Vol. 26).

Leirvik, T. (2021). Cryptocurrency returns and the volatility of liquidity. *Finance Research Letters*, 102031.

Li, C., Zhang, H., Wang, Z., Wu, Y., & Yang, F. (2022). Spatial-Temporal Attention Mechanism and Graph Convolutional Networks for Destination Prediction. *Frontiers in Neurorobotics*, *16*.

Li, W., Bao, R., Harimoto, K., Chen, D., Xu, J., & Su, Q. (2020). Modeling the Stock Relation with Graph Network for Overnight Stock Movement Prediction. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 4541–4547.

Li, X., Ma, J., Wang, S., & Zhang, X. (2015). How does Google search affect trader positions and crude oil prices? *Economic Modelling*, *49*, 162–171.

Li, X., Shang, W., & Wang, S. (2019). Text-based crude oil price forecasting: A deep learning approach. *International Journal of Forecasting*, *35*(4), 1548–1560.

Li, Y., Du, N., & Bengio, S. (2017). Time-Dependent Representation for Neural Event Sequence Prediction. *ArXiv*.

Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2017). Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. *ArXiv*.

Liang, X., Luo, P., Li, X., Wang, X., & Shu, L. (2023). Crude oil price prediction using deep reinforcement learning. *Resources Policy*, *81*, 103363.

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *379*(2194), 20200209.

Lin, J., Ma, J., Zhu, J., & Cui, Y. (2022). Short-term load forecasting based on LSTM networks considering attention mechanism. *International Journal of Electrical Power & Energy Systems*, *137*, 107818.

Lin, Y., Chen, K., Zhang, X., Tan, B., & Lu, Q. (2022). Forecasting crude oil futures prices using BiLSTM-Attention-CNN model with Wavelet transform. *Applied Soft Computing*, *130*, 109723.

Liu, G., & Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, *337*, 325–338.

Liu, J., Papailias, F., & Quinn, B. (2021). Direction-of-change forecasting in commodity futures markets. *International Review of Financial Analysis*, *74*.

Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics*, *82*(3), 631–671.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *ArXiv*.

Liu, Y., Zeng, Q., Yang, H., & Carrio, A. (2018). Stock Price Movement Prediction from Financial News with Deep Learning and Knowledge Graph Embedding. In *Lecture Notes in Computer Science* (Vol. 11016, pp. 102–113). Springer.

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Computing and Applications*, *32*(23), 17351–17360.

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297–303.

López-Cabarcos, M. Á., Pérez-Pico, A. M., Piñeiro-Chousa, J., & Šević, A. (2021). Bitcoin volatility, stock market and investor sentiment. Are they connected? *Finance Research Letters*, *38*, 101399.

Lu, Q., Sun, S., Duan, H., & Wang, S. (2021). Analysis and forecasting of crude oil price based on the variable selection-LSTM integrated model. *Energy Informatics*, *4*(S2), 47.

Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN-LSTM-Based Model to Forecast Stock Prices. *Complexity*, *2020*, 1–10.

Ma, G., Chen, P., Liu, Z., & Liu, J. (2022). The Prediction of Enterprise Stock Change Trend by Deep Neural Network Model. *Computational Intelligence and Neuroscience*, *2022*, 1–9.

Madziwa, L., Pillalamarry, M., & Chatterjee, S. (2022). Gold price forecasting using multivariate stochastic model. *Resources Policy*, *76*, 102544.

Maghyereh, A. I., & Abdoh, H. A. (2022). COVID-19 pandemic and volatility interdependence between gold and financial assets. *Applied Economics*, *54*(13), 1473–1486.

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, *46*(253), 68.

Matsunaga, D., Suzumura, T., & Takahashi, T. (2019). Exploring Graph Neural Networks for Stock Market Predictions with Rolling Window Analysis. *ArXiv*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *ArXiv*.

Mnif, E., Jarboui, A., & Mouakhar, K. (2020). How the cryptocurrency market has performed during COVID 19? A multifractal analysis. *Finance Research Letters*, *36*, 101647.

Mohamed, N. A., & Messaadia, M. (2023). Artificial Intelligence Techniques for the Forecasting of Crude Oil Price: A Literature Review. *2023 International Conference On Cyber Management And Engineering (CyMaEn)*, 340–343.

Naeem, M. A., Bouri, E., Peng, Z., Shahzad, S. J. H., & Vo, X. V. (2021). Asymmetric efficiency of cryptocurrencies during COVID19. *Physica A: Statistical Mechanics and Its Applications*, *565*.

Namaki, A., Shirazi, A. H., Raei, R., & Jafari, G. R. (2011). Network analysis of a financial market based on genuine correlation and threshold method. *Physica A: Statistical Mechanics and Its Applications*, *390*(21–22), 3835–3841.

Nayak, S. C., Das, S., Sahoo, B., & Satyanarayana, B. (2023). *AEFA-ANN: Artificial Electric Field Algorithm-Based Artificial Neural Networks for Forecasting Crude Oil Prices* (87–94).

Nedved, M., & Kristoufek, L. (2023). Safe havens for Bitcoin. *Finance Research Letters*, *51*, 103436.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, *59*(2), 347.

Neslihanoglu, S. (2021). Linearity extensions of the market model: a case of the top 10 cryptocurrency prices during the pre-COVID-19 and COVID-19 periods. *Financial Innovation*, *7*(1).

Nkrumah-Boadu, B., Owusu Junior, P., Adam, A., & Asafo-Adjei, E. (2022). Safe haven, hedge and diversification for African stocks: cryptocurrencies versus gold in time-frequency perspective. *Cogent Economics & Finance*, *10*(1).

Okorie, D. I., & Lin, B. (2020). Crude oil price and cryptocurrencies: Evidence of volatility connectedness and hedging strategy. *Energy Economics*, *87*, 104703.

Ongsritrakul, P., & Soonthornphisaj, N. (2003). Apply decision tree and support vector regression to predict the gold price. *Proceedings of the International Joint Conference on Neural Networks*, 2488–2492.

Orojo, O., Tepper, J., McGinnity, T. M., & Mahmud, M. (2019). A Multi-recurrent Network for Crude Oil Price Prediction. *IEEE Symposium Series on Computational Intelligence (SSCI)*.

Owusu Junior, P., Adam, A. M., & Tweneboah, G. (2020). Connectedness of cryptocurrencies and gold returns: Evidence from frequency-dependent quantile regressions. *Cogent Economics & Finance*, *8*(1), 1804037.

Ozili, P. (2020). *Spillover of COVID-19: impact on the Global Economy*.

Parelli, M., Papadimitriou, K., Potamianos, G., Pavlakos, G., & Maragos, P. (2022). Spatio-Temporal Graph Convolutional Networks for Continuous Sign Language Recognition. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 8457–8461.

Park, J., & Ratti, R. A. (2008). Oil price shocks and stock markets in the U.S. and 13 European countries. *Energy Economics*, *30*(5), 2587–2608.

Peng, R., Chen, Z., Hao, T., & Fang, Y. (2019). Neural Machine Translation with Attention Based on a New Syntactic Branch Distance. In *Communications in Computer and Information Science* (Vol. 1104, pp. 47–57). Springer.

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543.

Periwal, A. (2023). The Impact of Crude Oil Price Fluctuations on Indian Economy. *International Journal for Research in Applied Science and Engineering Technology*, *11*(4), 3173–3202.

Pesaran, M. H., & Shin, Y. (n.d.). An Autoregressive Distributed-Lag Modelling Approach to Cointegration Analysis. In S. Strom (Ed.), *Econometrics and Economic Theory in the 20th Century* (pp. 371–413). Cambridge University Press.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289–326.

Phan, D. H. B., Sharma, S. S., & Narayan, P. K. (2016). Intraday volatility interaction between the crude oil and equity markets. *Journal of International Financial Markets, Institutions and Money*, *40*, 1–13.

Pierdzioch, C., & Risse, M. (2017). Forecasting Precious Metal Returns with Multivariate Random Forests. *SSRN Electronic Journal*.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 31). Curran Associates, Inc.

Pullen, T., Benson, K., & Faff, R. (2014). A Comparative Analysis of the Investment Characteristics of Alternative Gold Assets. *Abacus*, *50*(1), 76–92.

Purczyński, J., & Bednarz-Okrzyńska, K. (2014). Estimation of the Shape Parameter of Ged Distribution for a Small Sample Size. *Folia Oeconomica Stetinensia*, *14*(1), 35–46.

Qin, M., Zhang, Y.-C., & Su, C.-W. (2020). The Essential Role of Pandemics: A Fresh Insight Into the Oil Market. *Energy RESEARCH LETTERS*, *1*(1).

Qin, Q., Huang, Z., Zhou, Z., Chen, C., & Liu, R. (2023). Crude oil price forecasting with machine learning and Google search data: An accuracy comparison of single-model versus multiple-model. *Engineering Applications of Artificial Intelligence*, *123*, 106266.

Raza, S., & Schwartz, B. (2023). Entity and relation extraction from clinical case reports of COVID-19: a natural language processing approach. *BMC Medical Informatics and Decision Making*, *23*(1), 20.

Reboredo, J. C. (2013). Is gold a safe haven or a hedge for the US dollar? Implications for risk management. *Journal of Banking & Finance*, *37*(8), 2665–2676.

Risse, M. (2019). Combining wavelet decomposition with machine learning to forecast gold returns. *International Journal of Forecasting*, *35*(2), 601–615.

Rosoł, M., Młyńczak, M., & Cybulski, G. (2022). Granger causality test with nonlinear neural-network-based methods: Python package and simulation study. *Computer Methods and Programs in Biomedicine*, *216*, 106669.

Rundo, F. (2019). Deep LSTM with reinforcement learning layer for financial trend prediction in FX high frequency trading systems. *Applied Sciences (Switzerland)*, *9*(20).

Said, S. E., & Dickey, D. A. (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, *71*(3), 599.

Salisu, A. A., & Ogbonna, A. E. (2021). The return volatility of cryptocurrencies during the COVID-19 pandemic: Assessing the news effect. *Global Finance Journal*, 100641.

Salisu, A. A., Ogbonna, A. E., & Adewuyi, A. (2020). Google trends and the predictability of precious metals. *Resources Policy*, *65*.

Sarkodie, S. A., Ahmed, M. Y., & Owusu, P. A. (2021). COVID-19 pandemic improves market signals of cryptocurrencies–evidence from Bitcoin, Bitcoin Cash, Ethereum, and Litecoin. *Finance Research Letters*.

Sarwar, S., Shahbaz, M., Anwar, A., & Tiwari, A. K. (2019). The importance of oil assets for portfolio optimization: The analysis of firm level stocks. *Energy Economics*, *78*, 217–234.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2).

Selmi, R., Mensi, W., Hammoudeh, S., & Bouoiyour, J. (2018). Is Bitcoin a hedge, a safe haven or a diversifier for oil price movements? A comparison with gold. *Energy Economics*, *74*, 787–801.

Şenol, Z., & Zeren, F. (2020). *Coronavirus (COVID-19) and Stock Markets: The Effects of the Pandemic on the Global Economy*

Seo, M., Min, S., Farhadi, A., & Hajishirzi, H. (2017). Neural Speed Reading via Skim-RNN. *ArXiv*.

Seo, Y., Defferrard, M., Vandergheynst, P., & Bresson, X. (2018). *Structured Sequence Modeling with Graph Convolutional Recurrent Networks* (pp. 362–373).

Sharif, A., Aloui, C., & Yarovaya, L. (2020). COVID-19 pandemic, oil prices, stock market, geopolitical risk and policy uncertainty nexus in the US economy: Fresh evidence from the wavelet-based approach. *International Review of Financial Analysis*, *70*.

Shi, L., Zhang, Y., Cheng, J., & Lu, H. (2018). Two-Stream Adaptive Graph Convolutional Networks for Skeleton-Based Action Recognition. *ArXiv*.

Shin, H., Hou, T., Park, K., Park, C. K., & Choi, S. (2013). Prediction of movement direction in crude oil prices based on semi-supervised learning. *Decision Support Systems*, *55*(1)348–358.

Shojaie, A., & Michailidis, G. (2010). Discovering graphical Granger causality using the truncating lasso penalty. *Bioinformatics*, *26*(18), i517–i523.

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). The Performance of LSTM and BiLSTM in Forecasting Time Series. *2019 IEEE International Conference on Big Data*, 3285–3292.

Siggiridou, E., & Kugiumtzis, D. (2016). Granger Causality in Multivariate Time Series Using a Time-Ordered Restricted Vector Autoregressive Model. *IEEE Transactions on Signal Processing*, *64*(7), 1759–1773.

Sims, C. A. (1980a). Macroeconomics and Reality. *Econometrica*, *48*(1), 1.

Sims, C. A. (1980b). Macroeconomics and Reality. *Econometrica*, *48*(1), 1.

Śmiech, S., & Papież, M. (2017). In search of hedges and safe havens: Revisiting the relations between gold and oil in the rolling regression framework. *Finance Research Letters*, *20*, 238–244.

Smirlock, M., & Starks, L. (1988). An empirical analysis of the stock price-volume relationship. *Journal of Banking and Finance*, *12*, 31–41.

Song, J., Son, J., Seo, D., Han, K., Kim, N., & Kim, S.-W. (2022). ST-GAT: A Spatio-Temporal Graph Attention Network for Accurate Traffic Speed Prediction. *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 4500–4504.

Sroka, Ł. (2022). Applying block bootstrap methods in silver prices forecasting. *Econometrics*, *26*(2), 15–29.

Statman, M., Thorley, S., & Vorkink, K. (2006). Investor overconfidence and trading volume. In *Review of Financial Studies* (Vol. 19, Issue 4, pp. 1531–1565).

Su, M., Liu, H., Yu, C., & Duan, Z. (2022). A new crude oil futures forecasting method based on fusing quadratic forecasting with residual forecasting. *Digital Signal Processing*, *130*, 103691.

Sun, J., Zhao, P., & Sun, S. (2022). A new secondary decomposition-reconstruction-ensemble approach for crude oil price forecasting. *Resources Policy*, *77*, 102762.

Swamy, V., & Lagesh, M. A. (2023). Does happy Twitter forecast gold price? *Resources Policy*, *81*, 103299.

Szarek, D., Bielak, Ł., & Wyłomańska, A. (2020). Long-term prediction of the metals' prices using non-Gaussian time-inhomogeneous stochastic process. *Physica A: Statistical Mechanics and Its Applications*, *555*.

Tang, L., Zhang, C., Li, L., & Wang, S. (2020). A multi-scale method for forecasting oil price with multi-factor search engine data. *Applied Energy*, *257*.

Thakkar, A., & Chaudhari, K. (2021). A comprehensive survey on deep neural networks for stock market: The need, challenges, and future directions. *Expert Systems with Applications*, *177*, 114800.

Uddin, A., Tao, X., & Yu, D. (2021). Attention Based Dynamic Graph Learning Framework for Asset Pricing. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 1844–1853.

Uzo-Peters, A., Laniran, T., & Adenikinju, A. (2018). Brent prices and oil stock behaviors: evidence from Nigerian listed oil stocks. *Financial Innovation*, *4*(1), 8.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *Arxiv*.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2017). Graph Attention Networks. *ArXiv*.

Vidal, A., & Kristjanpoller, W. (2020). Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, *157*.

Wang, J., Athanasopoulos, G., Hyndman, R. J., & Wang, S. (2018). Crude oil price forecasting based on internet concern using an extreme learning machine. *International Journal of Forecasting*, *34*(4), 665–677.

Wang, J., Niu, T., Du, P., & Yang, W. (2020). Ensemble probabilistic prediction approach for modeling uncertainty in crude oil price. *Applied Soft Computing Journal*, *95*.

Wang, J., Niu, X., Liu, Z., & Zhang, L. (2020). Analysis of the influence of international benchmark oil price on China's real exchange rate forecasting. *Engineering Applications of Artificial Intelligence*, *94*, 103783.

Wang, Q., Mao, Z., Wang, B., & Guo, L. (2017). Knowledge Graph Embedding: A Survey of Approaches and Applications. *IEEE Transactions on Knowledge and Data Engineering*, *29*(12), 2724–2743.

Wang, Q., Wei, Y., Zhang, Y., & Liu, Y. (2023). Evaluating the Safe-Haven Abilities of Bitcoin and Gold for Crude Oil Market: Evidence During the COVID-19 Pandemic. *Evaluation Review*, *47*(3), 391–432.

Wen, F., Tong, X., & Ren, X. (2022). Gold or Bitcoin, which is the safe haven during the COVID-19 pandemic? *International Review of Financial Analysis*, *81*, 102121.

Weng, F., Chen, Y., Wang, Z., Hou, M., Luo, J., & Tian, Z. (2020). Gold price forecasting research based on an improved online extreme learning machine algorithm. *Journal of Ambient Intelligence and Humanized Computing*, *11*(10), 4101–4111.

Whaley, R. E. (2000). The Investor Fear Gauge. *The Journal of Portfolio Management*, *26*(3) 12–17.

Wu, B., Wang, L., Wang, S., & Zeng, Y.-R. (2021). Forecasting the U.S. oil markets based on social media information during the COVID-19 pandemic. *Energy*, *226*, 120403.

Wu, Y.-X., Wu, Q.-B., & Zhu, J.-Q. (2019). Improved EEMD-based crude oil price forecasting using LSTM networks. *Physica A: Statistical Mechanics and Its Applications*, *516*, 114–124.

Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 753–763.

Wu, Z., Pan, S., Long, G., Jiang, J., & Zhang, C. (2019). Graph WaveNet for Deep Spatial-Temporal Graph Modeling. *ArXiv*.

Xiao, Y., Xia, K., Yin, H., Zhang, Y.-D., Qian, Z., Liu, Z., Liang, Y., & Li, X. (2022). AFSTGCN: Prediction for multivariate time series using an adaptive fused spatial-temporal graph convolutional network. *Digital Communications and Networks*.

Xiuzhen, X., Zheng, W., & Umair, M. (2022). Testing the Fluctuations of Oil Resource Price Volatility: A Hurdle for Economic Recovery. *SSRN Electronic Journal*.

Xu, D., Ruan, C., Korpeoglu, E., Kumar, S., & Achan, K. (2021). A Temporal Kernel Approach for Deep Learning with Continuous-time Information. *ArXiv*.

Xu, D., Ruan, C., Kumar, S., Korpeoglu, E., & Achan, K. (2019). Self-attention with Functional Time Representation Learning. *ArXiv*.

Yan, J., Mu, L., Wang, L., Ranjan, R., & Zomaya, A. Y. (2020). Temporal Convolutional Networks for the Advance Prediction of ENSO. *Scientific Reports*, *10*(1), 8055.

Yan, S., Xiong, Y., & Lin, D. (2018). Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition. *Proceedings of the AAAI Conference on Artificial Intelligence*, *32*(1).

Yang, J., De Montigny, D., & Treleaven, P. (2022). ANN, LSTM, and SVR for Gold Price Forecasting. *2022 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFEr)*, 1–7.

Yang, M., Li, X., & Liu, Y. (2021). Sequence to Point Learning Based on an Attention Neural Network for Nonintrusive Load Decomposition. *Electronics*, *10*(14), 1657.

Yang, M., & Wang, J. (2022). Adaptability of Financial Time Series Prediction Based on BiLSTM. *Procedia Computer Science*, *199*, 18–25.

Yang, Y., Guo, J., Sun, S., & Li, Y. (2021). Forecasting crude oil price with a new hybrid approach and multi-source data. *Engineering Applications of Artificial Intelligence*, *101*, 104217.

Yang, Y., Wei, Z., Chen, Q., & Wu, L. (2019). Using External Knowledge for Financial Event Prediction Based on Graph Neural Networks. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2161–2164.

Yousuf Khan, S., Amir, A., Khawaja, A., Kidwai, O. K., & Gheblehzadeh, M. (2021). Article ID: IJM_12_03_018 Impact of Covid-19 on Price Volitality of Cryptocurrency. *International Journal of Management (IJM)*, *12*(3), 193–205.

Yu, B., Yin, H., & Zhu, Z. (2018). Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 3634–3640.

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.

Yuan, J., Li, J., & Hao, J. (2023). A dynamic clustering ensemble learning approach for crude oil price forecasting. *Engineering Applications of Artificial Intelligence*, *123*, 106408.

Yuan, Z. (2023). Gold and Bitcoin Price Prediction based on KNN, XGBoost and LightGBM Model. *Highlights in Science, Engineering and Technology*, *39*, 720–725.

Zhang, J., Shi, X., Xie, J., Ma, H., King, I., & Yeung, D.-Y. (2018). GaAN: Gated Attention Networks for Learning on Large and Spatiotemporal Graphs. *ArXiv*.

Zhang, K., & Hong, M. (2022). Forecasting crude oil price using LSTM neural networks. *Data Science in Finance and Economics*, *2*(3), 163–180.

Zhang, P., & Ci, B. (2020). Deep belief network for gold price forecasting. *Resources Policy*, *69*.

Zhang, S., Chen, Y., Zhang, W., & Feng, R. (2021). A novel ensemble deep learning model with dynamic error correction and multi-objective ensemble pruning for time series forecasting. *Information Sciences*, *544*, 427–445.

Zhang, X., Chen, F., & Huang, R. (2018). A Combination of RNN and CNN for Attention-based Relation Classification. *Procedia Computer Science*, *131*, 911–917.

Zhang, X., Fuehres, H., & Gloor, P. A. (2012). Predicting Asset Value through Twitter Buzz. In *Advances in Intelligent and Soft Computing* (pp. 23–34). Springer.

Zhang, Y., Wang, J., Yu, L., & Wang, S. (2022). An extreme bias-penalized forecast combination approach to commodity price forecasting. *Information Sciences*, *615*, 774–793.

Zhang, Y.-J. (2013). Speculative trading and WTI crude oil futures price movement: An empirical analysis. *Applied Energy*, *107*, 394–402.

Zhao, L., Cheng, L., Wan, Y., Zhang, H., & Zhang, Z. (2015). A VAR-SVM model for crude oil price forecasting. *International Journal of Global Energy Issues*, *38*(1/2/3), 126.

Zhao, L., Wang, Y., Guo, S. Q., & Zeng, G. R. (2018). A novel method based on numerical fitting for oil price trend forecasting. *Applied Energy*, *220*, 154–163.

Zhao, Y., Li, J., & Yu, L. (2017). A deep learning ensemble approach for crude oil price forecasting. *Energy Economics*, *66*, 9–16.

Zheng, C., Fan, X., Wang, C., & Qi, J. (2019). *GMAN: A Graph Multi-Attention Network for Traffic Prediction*.

Zhou, S., Wu, J.-N., Wu, Y., & Zhou, X. (2015). Exploiting Local Structures with the Kronecker Layer in Convolutional Networks. *ArXiv*.