LUND UNIVERSITY
School of Economics and Management

Department of Economics

Forecasting copper price using VAR and the XGBoost model:

an experiment with a relatively small dataset

by

Juanli Hu

August 2023

Master's Programme in Data Analytics and Business Economics

Supervisor: Prof. Joakim Westerlund

Abstract

Given the importance of copper prices to investors, governments, and policymakers, this paper investigates short-term price predictability using VAR and XGBoost models. All models are trained with historical data from November 2021 to December 2022 and using MSE, RMSE and MAE for evaluating the model performance. The results show that the XGBoost model outperforms VAR models, implying that machine learning models are more robust than traditional statistical models. However, specific scenarios with a lag of 1to predict one day ahead(h=1day) show similar performance between XGBoost and VAR, indicating that traditional statistical models can still be competitive in certain situations. Therefore, it is critical not to dismiss traditional statistical models entirely, as they provide benefits in terms of interpretability and computational simplicity. Moreover, we also find that the selection of lag values for models is demonstrated to be empirical, with different lag values resulting in varying model performance. Thus, practitioners are encouraged to experiment with different lag settings in order to find the best model for their specific tasks and dataset sizes.

keywords: Forecasting: copper price; Vector autoregressive model; XGBoost; Time series

Acknowledgments

I'd like to express my appreciation to Joakim Westerlund, not only as my thesis supervisor, but also as an elder, for providing me with great guidance and support throughout this program. Thank you for your support, feedback, and patience, which were crucial in completing this work successfully.

Also, I am grateful for my family's and Chris's unwavering support; they are always there for me unconditionally. Additionally, four parcels full of Italian food from Italian family warm my heart during the wintertime. Of course, I'd like to thank my friends for taking the time to listen to my complaining when I am stressed.

Contents

1. Introduction

Copper is one of the most important minerals in the world, and it is widely traded as a commodity (Donchian, 1960; Choi and Hammoudeh, 2010). The price of copper plays an essential role in various aspects of the economy (Elshkaki et al., 2016; Liu et al., 2017; Mikesell, 1979). Therefore, many businesses and governments are affected by copper price fluctuations and seek to analyse and understand copper price behaviour to make better decisions and choose better policies (Buncic and Moretto, 2015; Alquist et al., 2019; Liu et al., 2022).

From the investor's perspective, copper is one of the most traded commodities on the major futures trading exchanges such as the London Metal Exchange (LME), the New York Commodity Exchange (NYMEX), and the Shanghai Futures Exchange (SHEF) (Sánchez Lasheras et al., 2015). Therefore, accurate copper price forecasts are a crucial point for developing successful investment strategies (Díaz et al., 2020). For traders and investors in particular, copper price forecasting over the short term is vital.

There are numerous factors that influence copper prices. In brief, it can be grouped into two categories: exogenous and endogenous factors (Goldstein and Yang, 2019). According to Guzmán and Silva (2017), all variables related to the copper market are considered endogenous, whereas macroeconomic and financial variables are considered exogenous. Endogenous factors such as physical demand can determine price in the short term, while physical supply determines price in the long term (Guzmán and Silva, 2017). Additionally, different trading markets can have an impact on each other (Li and Zhang, 2013). For example, LME future copper has a greater impact on SHFE copper, and vice versa (Hua and Chen, 2007). Exogenous factors such as the global economic outlook, geopolitical factors, technological factors, and exchange and interest rates can all have an impact on the copper price (Méndez et al., 2019; Vochozka et al., 2021).

Driven by forces of different nature, copper price is highly volatile and accurately forecasting copper price is difficult (Wang et al., 2019). Numerous efforts have been made to provide insight into behaviour of copper prices. Some researchers have used econometric and statistical methods for forecasting copper price. Buncic and Moretto (2015) forecasted monthly copper returns using Autoregressive-Moving-Average (ARMA) models. Similarly, Kriechbaumer et al. (2014) explored a combined approach wavelet-autoregressive integrated moving average (ARIMA), forecasting monthly prices of copper and other metals. Li and Li

(2015) employed generalized autoregressive conditional heteroscedastic (GARCH) models to forecast the volatility of copper futures of LME.

Furthermore, the complexity and nonlinear behaviour of copper prices have convinced researchers to use machine learning methods into their forecasting models to deal with the chaotic movements of copper prices. Their ability to learn complex patterns and relationships from large amounts of data makes them well-suited for long-term and mid-term load forecasting tasks (Baek, 2019). Therefore, advanced machine learning methods such as artificial neural networks (ANNs) and hybrid methods outperform most traditional models, particularly for long-term prediction (Wang et al., 2019). For instance, Lasheras et al. (2015) claimed that the ANN model predicted copper prices more accurately than the ARIMA model, with lower mean forecast error and variance for a 2-year prediction. Wang et al. (2019) provide a hybrid predictive technique that combines a complex network with traditional ANNs and discover that the hybrid method outperforms traditional ANNs methods when forecasting copper prices over a one-year period. Similarly, Liu et al. (2020) proposed a hybrid decision learning method for forecasting copper and other metals. They combine the variational mode decomposition (VMD) technique with long short-term memory network (LSTM) methods. In their study, they found that the VMD-LSTM model outperformed traditional ARIMA models as well as other benchmark models for one-year predictions. However, for short-term forecasting, a simple machine learning method such as random forest can also perform well. Liu et al. (2017) used a decision tree model to forecast copper prices over various time horizons and predictor sets. They concluded that short-term forecast models are often more accurate than long-term forecast models. Similarly, Díaz et al. (2020) used random forests and gradient boosting regression trees to forecast copper prices over various time horizons. Finally, they indicate that the random walk model can forecast daily copper prices with the greatest accuracy.

Although many econometric and artificial intelligence models have been used to forecast copper prices specially for the long-term prediction, little attention has been paid to predicting copper prices with a relatively small data set over a short period. With increasing financial speculation such as hedge funding in the futures markets (Guzmán and Silva, 2017; Yung and Liu, 2009), it is even more important to forecast short-term price movements rather than long-term trends (Lukac, et al., 1988; Gilbert, 2010). For traders, forecasting the price of tomorrow would be more important than knowing the price one year later. Obtaining an approximate accurate estimate of the copper price for the next few days allows traders to make informed decisions (such as identifying potential entry and exit points) that correspond to real-time market conditions (Gilbert, 2010). Thus, they can effectively manage the hedging strategies and risk management accordingly. Several studies have previously investigated the area of immediate prediction. Ni et al. (2022) investigated the use of Recurrent Neural Network (RNN) models to forecast copper prices over one-day, two-day, and three-day forecasting horizons. Consequently, our research will also concentrate on evaluating the performance of our models within these specific one-day, two-day, and three-day forecasting period as well.

Furthermore, the decision to use a relatively small dataset is based on pragmatic reasoning. First, the goal of the task is to provide service for traders with short-term predictions, a smaller dataset may contain more up-to-date information than a 10-year dataset (Díaz et al., 2020). Indeed, a smaller dataset may be more appealing because it can help to reduce noise and improve model accuracy (Díaz et al., 2020). Older and larger datasets, on the other hand, may contain a large number of data points that are no longer relevant or representative of the current situation, which can be unfavourable in these types of tasks.

From this regard, advanced machine learning methods such as ANNs, which are powerful computational tools, would not be appropriate methods for dealing with small datasets since they are easily prone to overfitting (Karimipour et al., 2019). This paper aims to forecast the copper price over a short time horizon using traditional time series methods namely Vector Autoregression (VAR) and classical ensemble methods namely eXtreme Gradient Boosting (XGBoost) on a small dataset. Both methods have been shown that work well with small datasets and are less prone to overfitting (Liang et al., 2020; Fujita et al., 2007; Lütkepohl, 2005). Another advantage of using the VAR model is that it can provide optimal lags. A few studies have used lag features in conjunction with machine learning models to forecast copper prices (Daz, Hansen, and Cabrera, 2020). Lag selection is an important aspect of VAR model specification (Brüggemann and Lütkepohl, 2000). These lag features are then incorporated into the framework of the XGBoost model, improving its predictive capability.

In addition, we include an autoregressive (AR) model as a benchmark for comparison with the VAR and XGBoost models. It acts as a reference point, allowing us to quantify and communicate the progress and effectiveness of more complex models in predicting copper price dynamics.

The research questions are as follows:

RQ1: How do the various lags in the VAR and XGBoost models impact prediction accuracy over different time horizons?

RQ2: Does VAR underperform compared to XGBoost in terms of short-term prediction?

2. Time series forecasting methods

2.1 VAR model

VAR models have a long tradition as tools for multivariate time series (Lütkepohl, 2009; Johansen, 1995). Being linear models, they are relatively easy to work with both in theory and practice. Although the computations are relatively straightforward compared to other AI methods right now, they were sufficient before powerful computers were widely distributed around the world (Lütkepohl, 2009). In VAR models, each variable is a linear function of its own past lags and the past lags of the other variable. The advantage of the VAR model is that it can capture interdependencies between multiple time series (Ramyar and Kianfar, 2017). This is particularly important considering the dependence of the copper price on previous day's prices, which can be identified and analysed in our case using a vector autoregressive model. The VAR model can capture temporal dependencies and provides insights into how past prices influence future copper price movements by incorporating lagged variables and capturing the dynamic relationships between them. A major characteristic of the variables estimated by a VAR model is stationarity. Trends in time series dataset can be a big challenge in accurately predicting future values of a time series, especially if the trend is non-linear or changes direction over time (Stock and Watson, 1988).

All these properties make VAR models particularly useful in explaining the dynamic behaviour of financial and economic time series for forecasting, and it often offers a very rich structure that facilitates capturing more features of the data (Zivot and Wang, 2006; Kadiyala and Karlsson, 1997; Kaura and Rajput, 2021). Therefore, VAR models are commonly used in macroeconomic analysis and forecasting to analyse the relationships between multiple variables over time ((Johansen, 1995; Ramyar and Kianfar, 2017b).

VAR models have been widely used in the commodity market to forecast prices (Akram, 2009; Collier and Goderis, 2012; Kaura and Rajput, 2021), particularly for crude oil. For instance, Park and Ratti (2008) used the VAR model to study oil price shocks in relation to stock markets. Wei et al. (2011) investigated the volatility of West Texas Intermediate (WTI) and Brent crude oil prices using generalized autoregressive conditional heteroskedasticity (GARCH) models. Recently, Ramyar and Kianfar (2017) forecast crude oil by comparing

ANN and VAR models, considering the exhaustible nature of crude oil and the impact of monetary policy.

Given a variety of examples from other fields, VAR shows its capacity to capture the relationships between multiple time series variables. However, there are relatively few studies that have used VAR specifically for copper price prediction. In fact, in comparison to other models, the VAR model is more adaptable since it can be generalized to include any number of variables and has fewer assumptions (Zivot and Wang, 2006). Therefore, forecasts from VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model (Zivot and Wang, 2006). Additionally, the VAR model also shows its ability to capture short-term dynamics of co-integrated systems (Park and Ratti 2008; Salazar and Weale, 1999).

In conclusion, although advanced machine learning methods have gained prevalence and VAR models have been criticized for their inability to capture non-linear relationships (Wang et al., 2019), we believe that VAR models are still considered as a competitive method for forecasting the copper price, especially with datasets such as relatively small.

2.2 XGBoost model

XGBoost is an advanced Gradient Boosting system that was proposed by Chen and He (2015). XGBoost is a tree ensemble model-based algorithm that employs a set of classification and regression trees (CART). The advantage of an ensemble of trees is that it can capture nonlinear relationships between variables and handle complex interactions between features. Furthermore, by employing multiple trees, the algorithm can reduce the risk of overfitting. Therefore, XGBoost has become a popular and effective machine learning algorithm by leveraging both normalization and ensemble techniques, with applications in a wide range of industries.

XGBoost has been extensively used in the financial market. For instance, Basak et al. (2019) used random forest and XGBoost to forecast trends in US and Indian stocks. Chatzis et al. (2018) use XGBoost and Deep Learning to improve classification accuracy while providing a robust method for developing a global systemic early warning tool that is more efficient and risk-sensitive than existing ones, while Wang and Guo (2020b) forecast stock market volatility in time series data using a mixed model of ARIMA and XGBoost.

Indeed, XGBoost models have gained popularity in the commodity market. Jabeur, Mefteh-Wali, and Viviani (2021) show that using XGBoost in conjunction with the SHapley Additive exPlanations (SHAP) approach can significantly improve gold price forecasting performance. Recently, Deng et al. (2023) used Multiple Timeframes eXtreme Gradient Boosting (MTXGBoost) to predict the price of crude oil futures. The experimental findings indicate that MTXGBoost could generate a good profit with minimal trading risk.

However, application of XGBoost in the copper industry is very limited, especially in terms of copper price forecasting. Given the outstanding performance of XGBoost in price prediction, we believe that it is a suitable method to predict copper prices. Additionally, XGBoost is appropriate for structured data, which makes it a suitable choice for time series data. The algorithm is also designed to work with tabular data, where each row represents a single observation, and the columns represent different features or variables.

Finally, we can use the optimal lag determined by the VAR model to feed the appropriate lagged features into the XGBoost model, resulting in more efficient training and predictions.

3. Methodology

The global copper market is predominantly traded on three major commodity markets: the London Metal Exchange (LME), the Commodity Exchange (COMEX), and the Shanghai Futures Exchange (SHFE). In this paper, we chose the LME as the target market for price prediction due to its significant trading volume and influence on price discovery (Watkins and McAleer, 2004).

3.1 Data collection

Copper price highly depends on both exogenous and endogenous factors. Furthermore, Guzmán and Silva (2017) suggest that incorporating both factors is more effective for modelling and forecasting purposes, providing a more comprehensive understanding of price dynamics.

All the variables we've chosen for this paper have received extensive studies and are considered to have a significant impact on copper price. Table 1 shows the variables we involve in this study. For the endogenous factors, we chose copper future price and spot price from LME, copper future price from COMEX, and copper future price and spot price from SHFE (García and Kristjanpoller, 2019). We considered several exogenous factors which have been used by other studies. It includes the prices of gold and silver from COMEX in the mental market, crude oil from NYMEX in the energy market, currency rates (EUR-USD,

USD-CNY), and indexes (NDAQ and S&P 500) (Liu et al., 2017; García and Kristjanpoller, 2019; Luo et al., 2022). Finally, we have a total of 12 variables.

Endogenous factors		**Exogenous factors**			
LME Data	**Copper Data from COMEX & SHFE**	**Metal Industry Data**	**Energy Market**	**Currency rate**	**Stock market index**
LME- Future Copper price	COMEX- Future Copper price	COMEX-GOLD	NYMEX-Oil	EUR-USD	NDAQ Index
LME -Copper Spot price	SHFE- Future Copper price	COMEX-SILVER		USD-CNY	S&P 500 Index
	Copper -Spot price in China				

Table 1: Categories of variables

As we stated before, the aim of this paper is to predict the copper price using relatively small dataset for servicing trading strategies. According to Guzmán and Silva (2017), the effect of financial speculation on commodity prices may exist only during specific periods of time and may not be permanent over time. Additionally, Daz et al. (2020) argue that a longer dataset may contain more noise and irrelevant information, which can interfere with the distillation of useful information from the input data. Therefore, a relatively small dataset might be more appropriate for our task as it can help to reduce noise and increase the model's accuracy for a short-term prediction (Díaz et al., 2020).

Finally, the relevant data was provided by a data service WIND and 12 sets of daily frequency data covering the years 2021.11.30–2023.12.30 including a total of 273 daily observations were collected. The specific data information is shown in Table 2.

No.	**Data name**	**Unit**	**Data type**	**Data frequency**	**Time span**	**Data sources**	**Missing value (%)**
1	LME Future Copper	USD/MT	Price	Daily	2021.11.30 ~2022.12.30	Wind	
2	LME Spot Copper	USD/MT	Price	Daily	2021.11.30 ~2022.12.30	Wind	
3	EUR-USD		Currency	Daily	2021.11.30 ~2022.12.30	Wind	
4	USD-CNY		Currency	Daily	2021.11.30 ~2022.12.30	Wind	5%

5	NDAQ Index		Index	Daily	2021.11.30 ~2022.12.30	Wind	3%
6	S&P 500 Index		Index	Daily	2021.11.30 ~2022.12.30	Wind	
7	NYMEX-Oil	USD/bbl	Price	Daily	2021.11.30 ~2022.12.30	Wind	3%
8	COMEX-GOLD	USD/ Oz	Price	Daily	2021.11.30 ~2022.12.30	Wind	9%
9	COMEX-SILVER	USD/ Oz	Price	Daily	2021.11.30 ~2022.12.30	Wind	13%
10	COMEX Future Copper price	USD/lb	Price	Daily	2021.11.30 ~2022.12.30	Wind	7%
11	SHFE Future Copper price	CNY/ MT	Price	Daily	2021.11.30 ~2022.12.30	Wind	6%
12	Copper Spot price in China	CNY/ MT	Price	Daily	2021.11.30 ~2022.12.30	Wind	6%

Table 2: Raw data information

3.2 Missing value

Table 2 also shows that the dataset has missing values. Overall, the percentage of missing values is small. COMEX silver has the highest percentage (13%), followed by COMEX gold (9%). The NASDAQ index, NYMEX crude oil, and COMEX copper inventory have the lowest (3%). To address this issue, we used a linear interpolation which is commonly used for dealing with missing values in high-frequency data in the economic field (Kohn and Ansley, 1986; Horváth et al., 2020).

3.3 Data correlation

First, we confirmed that there is a correlation between the price of copper and each of the other variables we chose. As shown in Table 3, copper prices from other future markets have a strong positive correlation with LME future copper prices. The correlation coefficients for these variables exceed 0.9, with p-values smaller than 0.001. Additionally, silver and gold from the mental market also show a similarly strong positive correlation. The crude oil from the NYMEX has the weakest positive correlation of the variables chosen. However, the currency pair USD-CNY shows a significant negative correlation with the LME future copper price.

No.	Variables	Correlation coefficient	t-statistic	p-value
1	COMEX Copper price	0.998	285.05	2.20E-16

2	LME Copper Spot price	0.996	180.71	2.20E-16
3	SHFE Copper price	0.932	42.463	2.20E-16
4	Copper Spot price in China	0.927	40.691	2.20E-16
5	COMEX-SILVER	0.874	29.571	2.20E-16
6	COMEX-GOLD	0.872	29.269	2.20E-16
7	EUR-USD	0.869	28.933	2.20E-16
8	S&P 500 Index	0.759	19.186	2.20E-16
9	NDAQ Index	0.710	16.58	2.20E-16
10	NYMEX-Crude Oil	0.294	5.066	7.52E-07
11	USD-CNY	-0.847	-26.272	2.20E-16

Table 3: Correlation analysis of variables

3.4 Train and Test Dataset

Before building models, we need to divide the data into two sets: the training set and the test set. The training set is used to develop models, pre-process predictors, and investigate relationships between predictors and responses, while the test set is used to determine the performance of the predictor model combination.

Time series data differs from other types of data in that it is organized chronologically, and the observations are dependent on previous observations (Cortez, 2010). The temporal order of observations is important in time series data, and there may be correlations or patterns that exist over time. Instead, as illustrated in Figure 1, we use the 12 months of data (253 observations) starting from November 30th, 2021, and ending on November 30th, 2022, as our training set. Additionally, we use the remaining one month of data (20 observations), ranging from December 1st, 2022, to December 30th, 2022, as our test dataset. The decision was made along with the goal of our task—a short-term forecasting. In this regard, we consider our choice to be very practical, efficient, and sufficient.

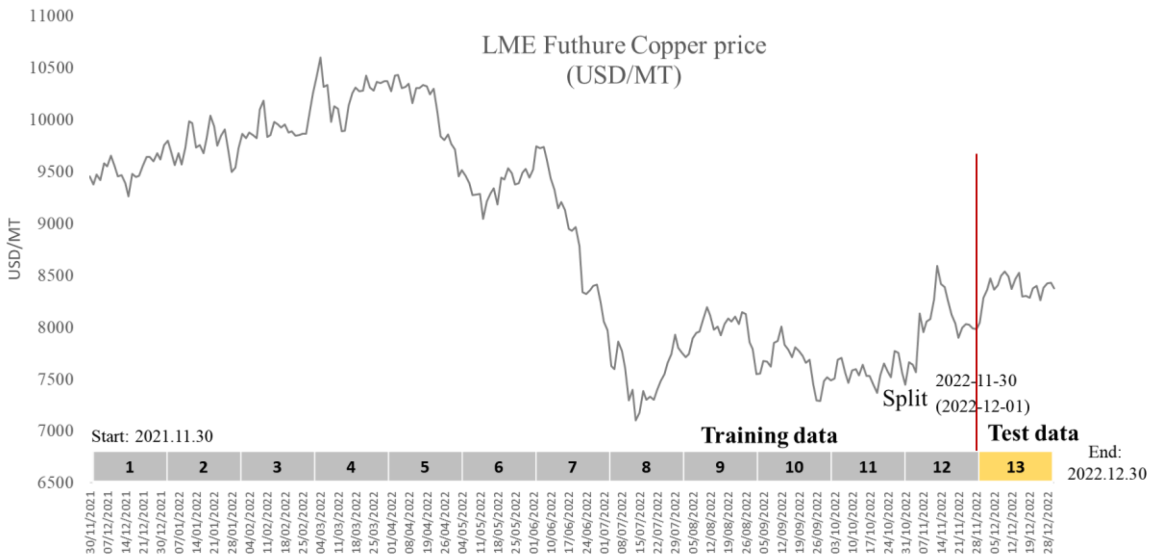
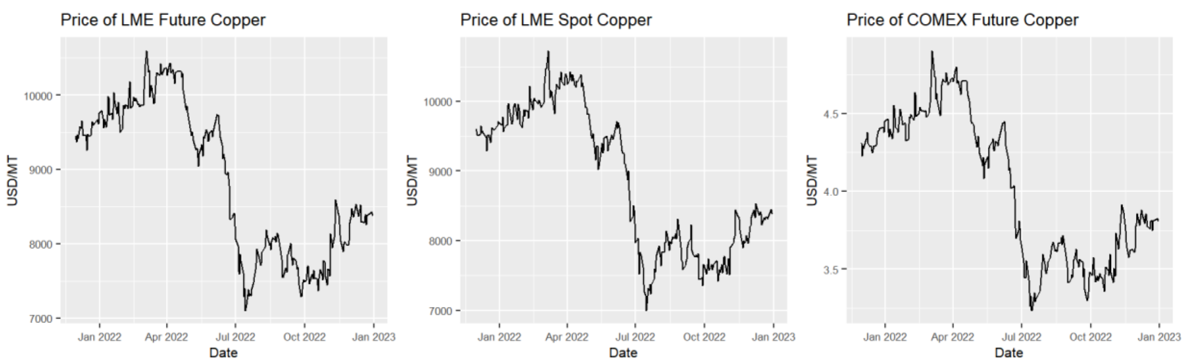


Figure 1: training and test dataset

3.5 Data transformation

However, one of the challenges of time series analysis and forecasting is dealing with trends in the data. Trends can be a significant obstacle in accurately predicting future values of a time series, especially if the trend is non-linear or changes direction over time (Stock and Watson, 1988). Time series data, such as the variables in this study, exhibits non-stationarity at certain levels, as illustrated in Figure 2. All the variables show strong trends.

Therefore, detrending the data is necessary before making certain statistical inferences to estimate its model. A common approach is to remove the trend from the time series by differencing the data or using methods such as exponential smoothing or the Box-Jenkins method. In this study, we use log differencing to remove the trend from our data and get stationary.



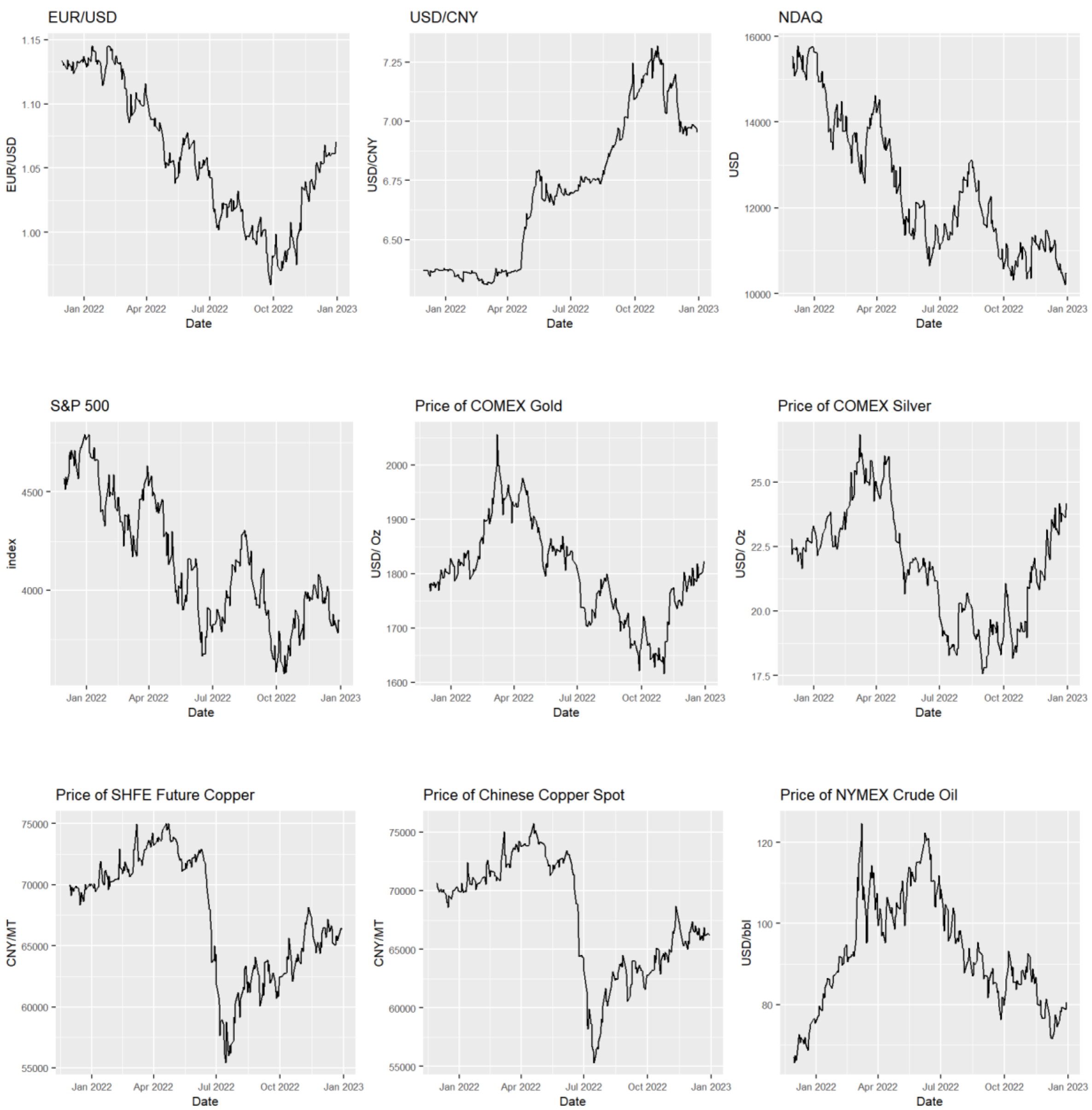


Figure 2: The actual values of copper price from Nov 2021 to Dec 2022

Log difference is a method for making a series of non-constant mean and variance stationary. It is a mathematical transformation applied to a time series to remove its trend and make it stationary. It is commonly used in econometrics and time series analysis to model the behaviour of variables that display trend and seasonality.

The log difference operation transforms the original time series into a new series that represents the percentage change between each pair of consecutive observations. Figure 3 shows logged variables of copper price from LME and COMEX. Taking the log difference helps to eliminate the effect of the trend by normalizing changes in the level of the time series with respect to the previous value. This makes it easier to use the VAR model to model the resulting stationary time series. The remaining factors are listed in Appendix A.

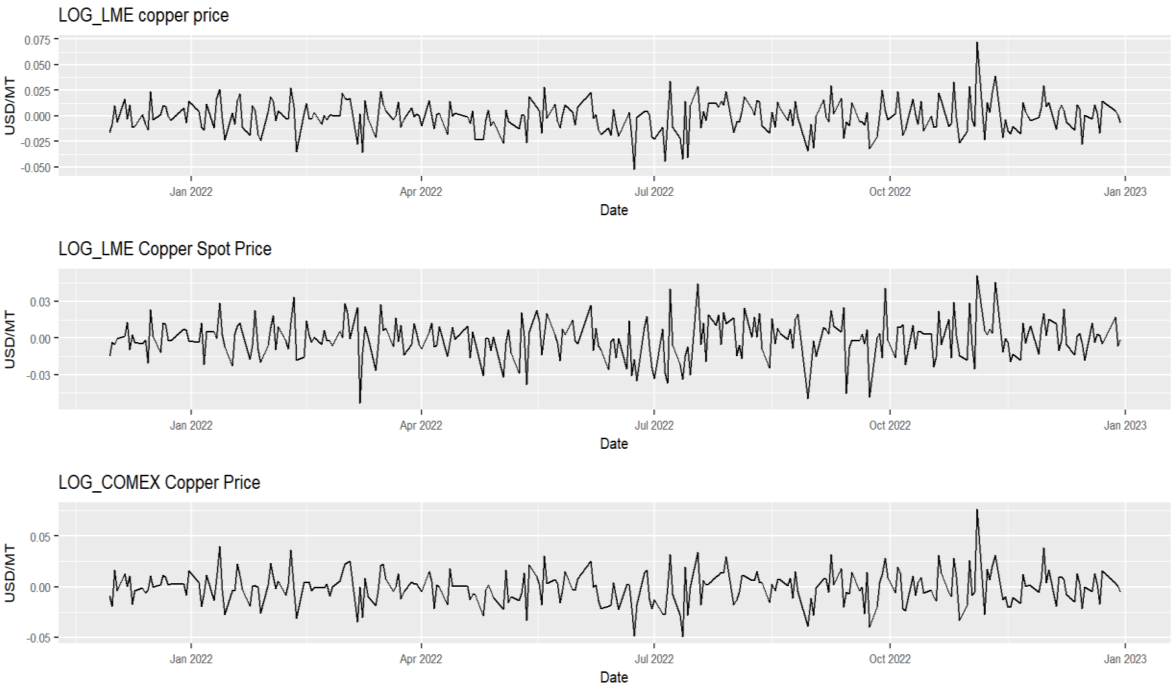


Figure 3: The log difference of copper price (LME and COMEX)

Furthermore, we use the Augmented Dickey-Fuller (ADF) test to validate our dataset; since the p-value for all variables (p-value = 0.01) is less than the chosen significance level (e.g., 0.05), there is sufficient evidence to reject the null hypothesis of non-stationarity. Therefore, based on the ADF test, we can conclude that all data is stationary.

4. Empirical results

The goal of this paper is to forecast copper prices in the short-term using a relatively small dataset, with a focus on very brief intervals to cater to traders and speculators. As we mentioned before, the prediction horizons are recommended by Ni et al. (2022), we will prediction copper prices with our models over one-day(h=1day), two-day (h=2 day), and three-day (h=3 day) forecasting horizons.

Furthermore, in order to compare the predictive technique, a metric to represent the forecasting performance is needed. We use three popular predictability indicators in this study: the mean squared error (MSE), the mean absolute error (MAE), and the root mean squared error (RMSE). The following are the mathematical expressions for these three performance criteria:

$$MSE = \frac{1}{N}\sum_{i=1}^{N} \quad (y_{copper\_true} - y_{copper\_pred})\ \text{^}2$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \quad |y_{copper\_true} - y_{copper\_pred}|$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \quad (y_{copper\_true} - y_{copper\_pred})\ \text{^}2}$$

where N stands for the total number of observations in the copper price database; y_{copper_true} true and y_{copper_pred} are the actual and forecasted copper prices. The smaller the values of MSE, MAE, and $RMSE$, the better the predictive performance of the model being evaluated.

The upcoming section will be structured as follows: initially, we will examine the effects of different orders on individual models, starting with AR, VAR, and XGBoost, to understand how the order influences each model's performance. Following that, we will conduct a model comparison analysis to see how the performance of each model varies across different prediction horizons.

4.1 AR model results

In alignment with previous studies that involving benchmark models for model comparison and evaluation (Hu, Ni, and Wen, 2020; Li and Li, 2015; Garca and Kristjanpoller, 2019), we introduce an AR model as our benchmark for this paper. Indeed, we also believe that the historical copper prices are likely to have a great impact on the copper price (Liu et al., 2017). Thus, we begin by looking at the two fundamental orders of autoregression: AR (1) and AR (2). Additionally, the purpose of using these two simple models as a benchmark is to evaluate the performance of more complexity models later in our case.

The prediction performances of the AR (1) and AR (2) with multiple steps ahead is shown in Table 4. Overall, AR (1) has better performance than AR (2). Particularly, when AR (1) predicts one day ahead has the smallest MSE value (0.00092). Similarly, for the two and three day forecast horizon, AR (1) still outperforms AR (2) in terms of MAE, MSE and RMSE. It indicates that the LME future copper's own most recent lagged variables have a big impact on price forecasting (Hu, Ni, and Wen, 2020). However, as the prediction horizon gets longer, both the AR (1) and AR (2) models perform poorly.

As we discussed before, the copper price can be greatly influenced not only by its own lagged data but also by other variables as the prediction horizon increases (Liu et al., 2017). From this regard, it indicates that the AR model has limited capability to capture a little bit of a long-term pattern due to its limited flexibility.

	AR (1)			**AR (2)**		
	Error criteria					
Horizon	MAE_LAG1	MSE_LAG1	RMSE_LAG1	MAE_LAG2	MSE_LAG2	RMSE_LAG2
h =1day	**0.009566**	**0.000092**	**0.009566**	0.010363	0.000107	0.010363
h = 2 day	0.011728	0.000142	0.011926	0.013494	0.000192	0.013853
h =3 day	0.012061	0.000149	0.012198	0.013259	0.000182	0.013507

Table 4: Prediction performances of AR (1) and AR (2) models for 1,2 and 3 days ahead forecasts

Since AR models only consider the lagged values of a variable for prediction, they ignore other variables that might have an impact on the variable of interest over the time. Particularly in complex and dynamic systems such as commodity markets, this lack of flexibility may have a negative impact on accuracy and predictive power. Therefore, we move to a VAR model that is more adaptable and can handle more variables.

4.2 VAR model results

4.2.1 Lag length selection

It is particularly important to consider the dependence of the copper oil price on previous day's prices, which can be identified and analysed using a vector autoregressive model (Ramyar and Kianfar, 2017). Thus, in order to develop an effective prediction model, we have to identify the appropriate order.

Lag selection is one of the important aspects of VAR model specification (Brüggemann and Lütkepohl, 2000). In practical applications, we generally choose a maximum number of lags, P_{max} and evaluate the performance with $p = 1,2,\ldots,p_{max}$.

The model VAR(p) that minimizes a certain lag selection criterion is then the most suitable one to use. The following are the most typical lag selection criteria:

- Akaike information criterion (AIC)
- Schwarz criterion (SC)
- Hannan-Quinn (HQ)
- Final prediction error (FPE)

The AIC and SC tests are the most commonly used in practice; however, it is difficult to say which is the best (Lütkepohl, 2005). The size of the sample affects the choice of lags required for accurate estimation (Liew et., al 2004). AIC tends to select more complex models than other model selection criteria, while SC tends to select simpler models than other model selection criteria. This can lead to underfitting and poor model performance, particularly when the sample size is large.

However, in terms of economic theory, it often suggests that economic processes are dynamic, it usually cannot be of much help regarding the length of those dynamic processes. Hence, the choice of the lag length in the VAR model is often an empirical issue (Scott Hacker and Hatemi-J, 2008; Pickup, 2014; Liew et al., 2004).

To choose the lag (p) of the model we use the VARselect *()* command from R studio to look at the best possible lag under different criteria. We set lag.max = 10 as suggested by Ni et al. (2022), who discovered that after 10 lags, partial autocorrelation rapidly decreases and becomes relatively low. The lag selection show as follows:

Selected lag lengths under different criteria			
AIC(n)	HQ(n)	SC(n)	FPE(n)
2	1	1	2

Table 5: Selected lag lengths under different criteria

Table 5 shows that different criteria select different lag lengths. HQ and SC select a lag length of 1, AIC and FPE select a lag length of 2. We should consider the practical and logical implications of using that lag before selecting any recommended lag. A lag of 2 means that the current predictions will be based on the previous 2 days of data, whereas a lag of 1 means that the current predictions will be based on the previous 1 days of data. The earlier we can make a prediction; the more time a company can make informed decisions based on that prediction. Such orders of 2 and 1 are all reasonable and acceptable in this case.

Therefore, we train the VAR model using all the recommended lags with different time horizons and analyse how the various lags affect the predictions.

4.2.2 Results with optimal lags

VAR (1) stands for Vector Autoregression of order 1. Each variable in a VAR (1) model is regressed on its own lagged values as well as the lagged values of the other variables in the system. VAR (2), on the other hand, stands for Vector Autoregression of order 2. A VAR (2) model regresses the current values of each variable on its own lagged values as well as the lagged values of all variables in the system up to lag 2. This enables VAR (2) to capture more complex interdependencies and feedback effects between variables.

The prediction performances of the VAR (1) and VAR (2) with multiple steps ahead is shown in Table 6. For the one day forecast horizon, the MSE for the two forecasting models under evaluation, i.e., VAR (1) and VAR (2), are 0.000009 and 0.000073 respectively. Comparing these values, we observe that VAR (1) models provide significantly more accurate forecasts than the VAR (2) model. We obtain similar results when comparing the models using the RMSE. In this case, the RMSE produced by the VAR (1) model is 0.003004, whereas the RMSE produced by the VAR (2) model is 0.008520. These results indicate the VAR (1) model's superior performance when predicting one day ahead. Turning to the MAE metrics, it shows similar results.

	VAR (1) model Lag order (HQ & SC) =1			VAR (2) model Lag order (AIC&FPE) =2		
	Error criteria					
Horizon	MAE	MSE	RMSE	MAE	MSE	RMSE
h =1day	**0.003004**	**0.000009**	**0.003004**	0.008520	0.000073	0.008520
h = 2 day	0.008986	0.000117	0.010795	0.009945	0.000101	0.010046
h =3 day	0.010447	0.000137	0.011716	0.010748	0.000118	0.010870

Table 6: Prediction performances of VAR (1) and VAR (2) models for 1,2 and 3 days ahead forecasts

The outcomes for the two and three day forecast horizons, however, differ from the one day ahead forecast. When the horizons are extended, the model's performance has changed due to the different lag orders in our case. In general, the value of MSE, RMSE and MAE are getting bigger. To better see the impact of the different order of lags on the model performance at different forecast horizons, we make a plot to visualize the criteria. Figure 4 shows how the performance metrics vary across the one, two and three-day forecast horizons.

We found that the MSE and RMSE for the two-day and three-day ahead forecasts using VAR (2) are lower compared to VAR (1). On the other hand, the MAE for the two-day and three-day ahead forecasts using VAR (2) is higher compared to VAR (1). For example, with the VAR (1) model, MSE = 0.000117 for h = 2, MSE = 0.000137 for h = 3, and with VAR (2) model, MSE gets smaller than VAR (1) with h=2 and h=3(MSE=0.000101, MSE =0.000118 respectively). Contrast to RMSE, the values of MAE for the VAR (1) model are slightly smaller than VAR (2) for both forecasting horizons. For instance, MAE =0.008986 for h=1 in VAR (1) while MAE= 0.009945 in VAR (2) model.

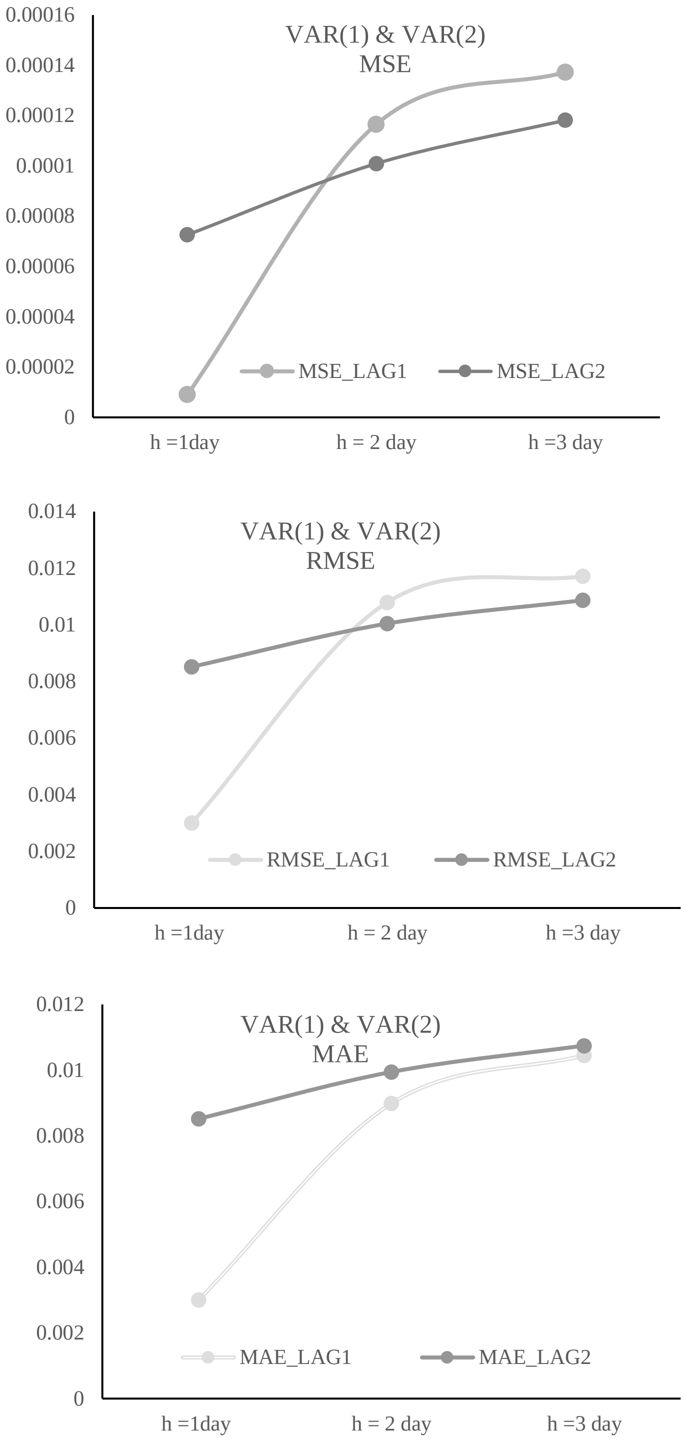


Figure 4: The MSE, RMSE, and MAE trends of the VAR model

Overall, as the task is to forecast the price, we pay more attention to MSE and RMSE. The lower RMSE, MSE indicate better model performance in terms of accuracy and precision. Furthermore, it suggests that the VAR (1) model works well for one-day forecasting, whereas the VAR (2) model performs better for longer time horizons than the VAR (1) model. In this case, the VAR (2) model with additional lagged variables is more robust. However, it is difficult to say which lag is better, it mainly depends on the aim of the task. Anyway, we will see how the lagged variables will impact the prediction in XGBoost.

4.3 XGBoost model results

4.3.1 Tuning Parameters in XGBoost

Before executing XGBoost, several parameters need to be initialized to define the model, which are referred to as hyperparameters. These hyperparameters help address the bias-variance trade-off by balancing the model complexity and predictive ability. The following is a detailed explanation of the XGBoost parameters adapted from the documentation (https://xgboost.readthedocs.io/en/stable/).

objective: we selected “reg:squarederror”, which means that XGBoost will use the mean squared error (MSE) as the loss function. The objective is to minimize the MSE during model training to improve the accuracy of the predictions.

max_depth: This parameter sets the maximum depth of each decision tree in the XGBoost model. A deeper tree can capture more complex patterns in the data but may also overfit. You can experiment with different values to find the optimal depth for your dataset.

learning_rate: This parameter controls the step size that the XGBoost algorithm takes at each iteration. A lower learning rate can result in a more accurate model but may also require more iterations to converge.

gamma: This parameter is used to control the complexity of the model and can be used to avoid overfitting. A higher value of gamma results in fewer splits and a more conservative model, while a lower value of gamma leads to more splits and a more complex model.

subsample: This parameter controls the fraction of observations that are randomly sampled for each tree. A lower subsample can reduce overfitting but may also result in a less accurate model. It can experiment with different values to find the optimal subsample rate for your dataset.

colsample_bytree: This parameter determines how many features are randomly sampled for each tree. Lowering the colsample_bytree can reduce overfitting but may result in a less accurate model. It can experiment with various values to determine the best colsample_bytree rate for your dataset.

Grid SearchCV was used to tune the hyperparameters in this paper. This method volves specifying a grid of hyperparameter values to explore, and then training and evaluating the model using each combination of hyperparameters in the grid. The optimal value was found: colsample_bytree: 1; gamma: 0.0; learning_rate: 0.05; max_depth: 1; n_estimators: 150.

4.3.2 Parameter Tuning Results

Using lag variables as inputs is a common technique in both time series analysis and machine learning to capture temporal dependencies and incorporate historical information into predictive models (Čeperić, Žiković and Čeperić, 2017; Herrera et al., 2019; Antipov and Pokryshevskaya, 2020; Li, Shang and Wang, 2019). Since we can not affirm which lag has better performance in the VAR model, we will test both lags in XGBoost. We denote XGBoost in the order of one (XGBoost (1)) and the order of two (XGBoost (2)).

Table 5 shows the prediction performances of the XGBoost (1)) and XGBoost (2) with multiple steps ahead. Overall, we can see that XGBoost (1) had better performance than XGBoost (2). The MSE for the XGBoost for the one-day forecast horizon and two-day forecast horizon is 0.00000286 and 0.00000287, respectively. Comparing these two values, XGBoost (1) is slightly better than XGBoost (2). When comparing the models using the RMSE and MAE, we get similar results. The smallest MSE, MAE and RMSE given by XGBoost (1) are predicted two days ahead. The MSE produced by the XGBoost (1) model is 0.00000218, whereas the MSE produced by the XGBoost (2) model is 0.00000276. These results show that the XGBoost (1) model outperforms other models when predicting two days ahead in terms of accuracy. Although the MSE of the XGBoost (2) model for h=2 day is greater than that of the XGBoost (1) model for h=2 day, XGBoost (2) model for h=2 day still outperforms h=1 and h = 3 days.

	XGBoost (1) model-optimal Lag order (HQ & SC) =1			XGBoost (2) model-optimal Lag order (AIC&FPE) =2		
	Error criteria					
Horizon	MAE	MSE	RMSE	MAE	MSE	RMSE
h =1day	0.001691	0.00000286	0.001691	0.001694	0.00000287	0.001694
h = 2 day	**0.001457**	**0.00000218**	**0.001475**	**0.001661**	**0.00000276**	**0.001661**
h =3 day	0.002421	0.00000776	0.002785	0.002660	0.00000907	0.003012

Table 5: Prediction performances of XGBoost models for 1,2 and 3 days ahead forecasts

Furthermore, we plot the results in Figure 5 to better see the trend with different lags. The 'U' shape of the line for MAE, MSE, and RMSE indicates that prediction performance increases and then decreases as the horizons extend in general. In this case, the XGBoost (1) model with additional lagged variables is more robust. However, as more lagged variables are added, the model is found to be prone to overfitting.

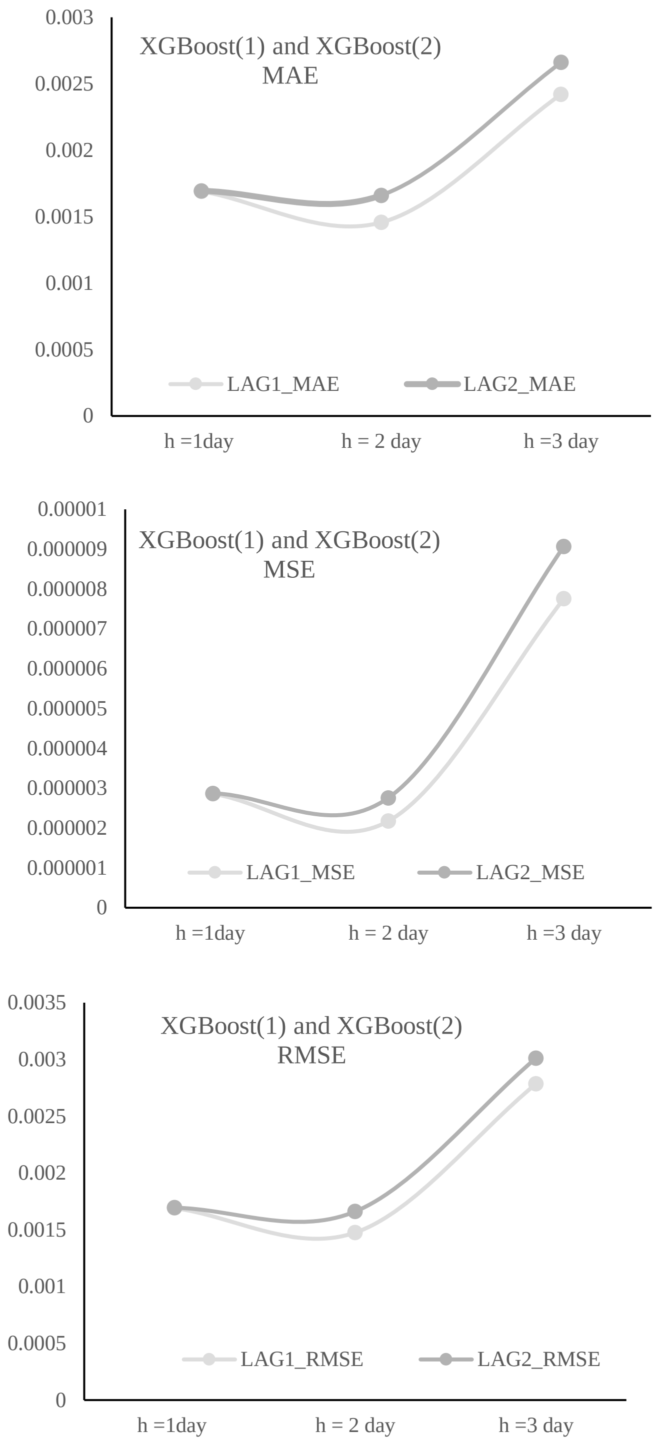


Figure 5: The MSE, RMSE, and MAE trends of the XGBoost model

We conclude that the input of lagged variables is very useful for capturing longer horizon features, but it introduces more noise and increases the model's complexity, making it easier to fit the noise in the training data, especially since the dataset is relatively small in our case. As a result, the model may struggle to generalize to additional lag data points. In this sense, different from the VAR model, VAR (2) has more robust performance than VAR (1), whereas the XGBoost model with one order performs better than two orders.

4.4 Models Comparison

Finally, we can compare three models: AR, VAR, and XGBoost. AR and VAR models are traditional time series models, while XGBoost is a machine learning method. Since the purpose of this paper is to forecast the cooper price, we are primarily concerned with prediction accuracy. Therefore, we choose MSE to evaluate the model performance specifically. The lower the MSE, the better the model's predictive accuracy. Figure 6 shows the MSE values for the three models, allowing for a visual comparison of their performances. The model with the lowest MSE is considered the most accurate in predicting copper prices.

We can immediately recognize that the XGBoost model outperforms the VAR and AR models in terms of prediction accuracy for forecasting copper prices with the lowest MSE. The AR model has the worst performance of the three. These findings support previous studies showing that machine learning models outperform traditional time series models at predicting commodity price (Ramyar and Kianfar, 2017; Sánchez Lasheras et al., 2015; Hu, Ni and Wen, 2020). For instance, Sánchez Lasheras et al. (2015) found that when forecasting the COMEX copper spot price, artificial neural network models outperformed traditional ARIMA models. Similarly, Ramyar and Kianfar (2017) showed that a neural network model outperformed a VAR model in predicting crude oil prices. Both studies show that the machine learning model’s ability to capture nonlinearity and handle dynamic relationships between various factors allowed it to generate more accurate and reliable forecasts than traditional time series models.

Indeed, machine learning methods are not always guaranteed to perform well for every time series forecasting task. The effectiveness of a particular method is often determined by the characteristics of the data and the specific problem being addressed. In some cases, studies have shown that using neural networks to forecast specific time series over an extended period of time may result in poorer performance than traditional models such as ARIMA models (Khashei and Bijari, 2010). In our case, we can see a deterioration in the MSE of XGBoost as the forecasting horizon increases, moving from 0.00000286 (h=1) to 0.00000776 (h=2) in XGBoost (1). This indicates that XGBoost faces challenges in maintaining accurate predictions over longer horizons, resulting in less precise forecasts as the forecasting horizon increases.

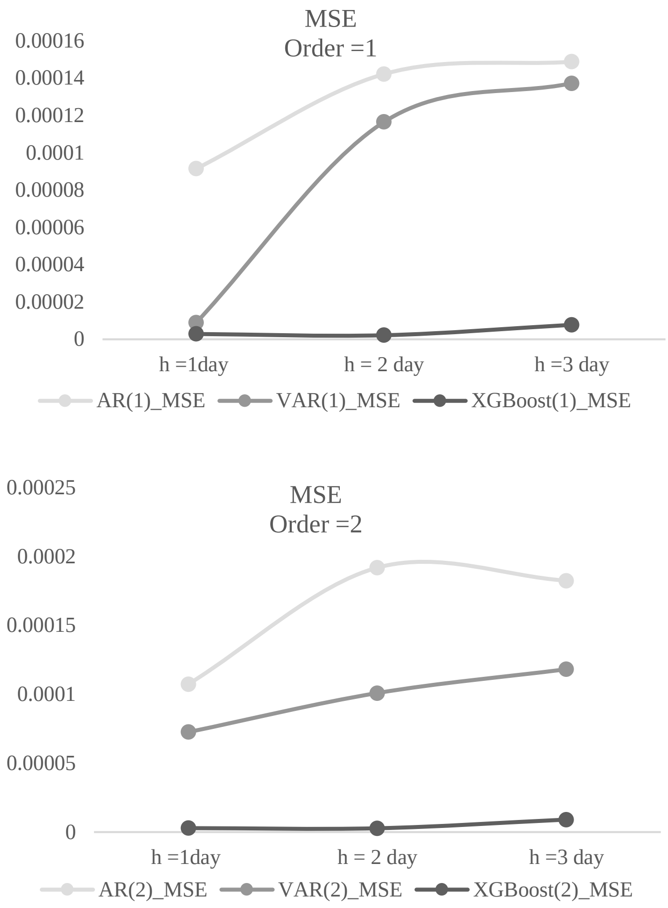


Figure 6: The MSE trends of the AR, VAR and XGBoost model

Similar to XGBoost, VAR models also face the challenge when the prediction period extend. However, the interesting thing is that we can see from Figure 6, where the VAR (1) model performs as well as the XGBoost (1) model for forecasting one day ahead (h=1). This finding is interesting as it demonstrates that, in the context of short-term forecasting one day ahead (h=1), the VAR (1) model and XGBoost (1) model can achieve good forecast in application to real-life problems, yielding comparable MSE of 0.000009 and 0.000003, respectively. It suggests that the VAR (1) model is effective in capturing the short-term dependencies and patterns in the time series data, making it a competitive option for immediate forecasts. It aligns with previous research by Adebiyi et al. (2014) and Sterba and Hilovska (2010), which also highlighted the good performance of the ARIMA model in terms of forecasting when compared to machine learning models such as neural net works. Therefore, from the authors' point of view, in those cases in which the computational capability is a problem and the data analysis will not be performed as a black box for the final user of the forecasting system, we recommend the use of time series models, such as VAR models, even though for the long prediction period, as has been observed in our data, their forecast error would be larger than for XGBoost but always lower than the AR model. Here, the benchmark AR model performs the worst, indicating its limitations in capturing the dynamics of the copper price data. However, it gives a baseline to showcase the improvements achieved by more advanced models.

In sum, the findings indicate that the XGBoost model, by incorporating lagged values as features, outperforms other models in capturing complex relationships and patterns in copper price time series data in this case. Additionally, the VAR model shows its ability to capture short-term predictive power. However, it falls short of capturing the data's non-linear features and complex relationships. Hence, when compared to the XGBoost model, its predictive accuracy may be limited. Based on these results, it can be concluded that the XGBoost model has the lowest MSE, indicating superior performance in predicting copper prices than VAR model. This emphasizes the significance of including lagged variables as features and leveraging the XGBoost algorithm's flexibility to capture non-linear patterns and dependencies within the data.

5. Conclusion

Copper is a widely traded commodity with significant implications for a wide range of industries. Given its importance, accurately forecasting the future price of copper can be beneficial for market participants such as policymakers, shareholders, or traders looking to optimize their investment strategies or make informed policy decisions (Wang et al., 2019; Zhang et al., 2021b). We focused on short-term predictions, as they have been shown to be more accurate compared to long-term forecasts according to previous research (Díaz et al., 2020; Liu et al., 2017). This study attempts to forecast the copper price in a short term in order to provide a reference for traders and speculators when making trading decisions.

To achieve this, we employed classical time series models (VAR and AR), and a classical machine learning model (XGBoost) to capture the complex dynamics of the copper price. By incorporating both endogenous and exogenous variables in our models and applying log-differencing to eliminate trend effects, we aimed to gain a comprehensive understanding of the factors impacting copper price dynamics.

Using daily data from November 2021 to December 2022, our empirical results show that the XGBoost with lagged features outperforms the VAR model for short term prediction and can capture long term features in general. In terms of prediction accuracy as measured by the MSE metric, the MSE values obtained by XGBoost models to predict multiple steps ahead (h=1, 2 and 3 days) were found to be significantly lower than those obtained by the traditional time series models. The RMSE and MAE yield similar results in our case. These findings are consistent with most studies that compare traditional time series models to machine learning models (e.g., Dehghani and Bogdanovic, 2018; García and Kristjanpoller, 2019; Sánchez Lasheras et al., 2015).

In addition, it is worth noting that when forecasting one day ahead, the VAR (1) model and the XGBoost (1) model demonstrate very similar performance. It indicates the VAR model's comparable prediction power for a specific task in this scenario, such as short-term prediction. Furthermore, the VAR model is easier to interpret than most advanced machine learning models. Based on these classical models, researchers can gain the insights from the easily interpretable coefficients of them as a starting point to build more complex models, such as the optimal lags chosen by the VAR model in this case.

Furthermore, it should be emphasized that selecting lags based on different criteria is a very practical decision. In our case, HQ and SC select a lag length of 1, AIC and FPE select a lag length of 2. We find that the VAR (1) model works well for one-day forecasting, whereas the VAR (2) model performs better for longer time horizons than the VAR (1) model. However, different lags have different performance in the XGBoost model. XGBoost (1) has better performance than XGBoost (2). Therefore, the choice of the lag length is often an empirical issue. It emphasizes the importance of selecting optimal lags and meticulously planning the task's objectives and model training.

This study also attempts to combine the strengths of the VAR model's lag order determination and the XGBoost model's ability to learn from informative features can significantly enhance the accuracy and predictive power of copper price forecasting. This approach has the potential to provide valuable support to commodity exchange market participants, including traders and policymakers, in making informed decisions. We can capitalize on the complementary advantages of both models by leveraging the optimal lagged values derived from the VAR model, paving the way for more effective and reliable forecasting in the dynamic world of commodity trading.

Reference

Adebiyi, A.A., Adewumi, A.O. and Ayo, C.K. (2014). *Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction*. [online] Journal of Applied Mathematics. Available at: https://www.hindawi.com/journals/jam/2014/614342/.

Akram, Q.F. (2009). Commodity prices, interest rates and the dollar. *Energy Economics*, 31(6), pp.838–851. doi:https://doi.org/10.1016/j.eneco.2009.05.016.

Alquist, R., Bhattarai, S. and Coibion, O. (2019). Commodity-price comovement and global economic activity. *Journal of Monetary Economics*. doi:https://doi.org/10.1016/j.jmoneco.2019.02.004.

Antipov, E.A. and Pokryshevskaya, E.B. (2020). Interpretable machine learning for demand modeling with high-dimensional data using Gradient Boosting Machines and Shapley values. *Journal of Revenue and Pricing Management*. doi:https://doi.org/10.1057/s41272-020-00236-4.

Baek, S.-M. (2019). Mid-term Load Pattern Forecasting With Recurrent Artificial Neural Network. *IEEE Access*, 7, pp.172830–172838. doi:https://doi.org/10.1109/ACCESS.2019.2957072.

Basak, S., Kar, S., Saha, S., Khaidem, L. and Dey, S.R. (2019). Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance*, 47, pp.552–567. doi:https://doi.org/10.1016/j.najef.2018.06.013.

Brüggemann, R. and Lütkepohl, H. (2000). *Lag Selection in Subset VAR Models with an Application to a U.S. Monetary System*. *edoc.hu-berlin.de*. Humboldt-Universität zu Berlin, Wirtschaftswissenschaftliche Fakultät.

Buncic, D. and Moretto, C. (2015). Forecasting copper prices with dynamic averaging and selection models. *The North American Journal of Economics and Finance,* 33, pp.1–38. doi:https://doi.org/10.1016/j.najef.2015.03.002.

Čeperić, E., Žiković, S. and Čeperić, V. (2017). Short-term forecasting of natural gas prices using machine learning and feature selection algorithms. *Energy*, 140, pp.893–900. doi:https://doi.org/10.1016/j.energy.2017.09.026.

Chatzis, S.P., Siakoulis, V., Petropoulos, A., Stavroulakis, E. and Vlachogiannakis, N. (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Applications*, 112, pp.353–371. doi:https://doi.org/10.1016/j.eswa.2018.06.032.

Chen, T. and He, T. (2015). *xgboost: eXtreme Gradient Boosting*. R package version 0.4-2, 1(4), pp.1-4.

Choi, K. and Hammoudeh, S. (2010). Volatility behavior of oil, industrial commodity and stock markets in a regime-switching environment. *Energy Policy*, 38(8), pp.4388–4399. doi:https://doi.org/10.1016/j.enpol.2010.03.067.

Collier, P. and Goderis, B. (2012). Commodity prices and growth: An empirical investigation. *European Economic Review*, 56(6), pp.1241–1260. doi:https://doi.org/10.1016/j.euroecorev.2012.04.002.

Cortez, P. (2010). *Sensitivity analysis for time lag selection to forecast seasonal time series using Neural Networks and Support Vector Machines*. [online] IEEE Xplore. doi:https://doi.org/10.1109/IJCNN.2010.5596890.

Dehghani, H. and Bogdanovic, D. (2018). Copper price estimation using bat algorithm. *Resources Policy*, 55, pp.55–61. doi:https://doi.org/10.1016/j.resourpol.2017.10.015.

Deng, S., Zhu, Y., Duan, S., Yu, Y., Fu, Z., Liu, J., Yang, X. and Liu, Z. (2023). High-frequency forecasting of the crude oil futures price with multiple timeframe predictions fusion. *Expert Systems with Applications*, 217, p.119580. doi:https://doi.org/10.1016/j.eswa.2023.119580.

Díaz, J.D., Hansen, E. and Cabrera, G. (2020). A random walk through the trees: Forecasting copper prices using decision learning methods. *Resources Policy*, 69, p.101859. doi:https://doi.org/10.1016/j.resourpol.2020.101859.

Donchian, R.D. (1960). Commodities: High Finance in Copper. *Financial Analysts Journal*, 16(6), pp.133–142. doi:https://doi.org/10.2469/faj.v16.n6.133.

Elshkaki, A., Graedel, T.E., Ciacci, L. and Reck, B.K. (2016). Copper demand, supply, and associated energy use to 2050. *Global Environmental Change*, 39, pp.305–315.

doi:https://doi.org/10.1016/j.gloenvcha.2016.06.006.

Fujita, A., João Ricardo Sato, Humberto Miguel Garay-Malpartida, Yamaguchi, R., Satoru Miyano, Mari Cleide Sogayar and Ferreira, C. (2007). Modeling gene expression regulatory networks with the sparse vector autoregressive model. *BMC systems biology*, 1(1). doi:https://doi.org/10.1186/1752-0509-1-39.

García, D. and Kristjanpoller, W. (2019). An adaptive forecasting approach for copper price volatility through hybrid and non-hybrid models. *Applied Soft Computing*, 74, pp.466–478. doi:https://doi.org/10.1016/j.asoc.2018.10.007.

Gilbert, C.L. (2010). Speculative Influences On Commodity Futures Prices 2006-2008. *UNCTAD Discussion Papers*.

Goldstein, I. and Yang, L. (2019). *Commodity Financialization and Information Transmission*. [online] papers.ssrn.com. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2555996 [Accessed 1 May 2023].

Guzmán, J.I. and Silva, E. (2017). Copper price determination: fundamentals versus non-fundamentals. *Mineral Economics*, 31(3), pp.283–300. doi:https://doi.org/10.1007/s13563-017-0130-y.

Herrera, G.P., Constantino, M., Tabak, B.M., Pistori, H., Su, J.-J. and Naranpanawa, A. (2019). Long-term forecast of energy commodities price using machine learning. *Energy*, 179, pp.214–221. doi:https://doi.org/10.1016/j.energy.2019.04.077.

Horváth, L., Liu, Z., Rice, G. and Wang, S. (2020). A functional time series analysis of forward curves derived from commodity futures. *International Journal of Forecasting*, 36(2), pp.646–665. doi:https://doi.org/10.1016/j.ijforecast.2019.08.003.

Hu, Y., Ni, J. and Wen, L. (2020). A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction. *Physica A: Statistical Mechanics and its Applications*, [online] 557(C).

Hua, R. and Chen, B. (2007). International linkages of the Chinese futures markets. *Applied Financial Economics*, 17(16), pp.1275–1287. doi:https://doi.org/10.1080/09603100600735302.

Jabeur, S.B., Mefteh-Wali, S. and Viviani, J.-L. (2021). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Annals of Operations Research*. doi:https://doi.org/10.1007/s10479-021-04187-w.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. *Google Books*. OUP Oxford.

KADIYALA, K.R. and KARLSSON, S. (1997). NUMERICAL METHODS FOR ESTIMATION AND INFERENCE IN BAYESIAN VAR-MODELS. *Journal of Applied Econometrics*, 12(2), pp.99–132. doi:https://doi.org/10.1002/(sici)1099-1255(199703)12:2%3C99::aid-jae429%3E3.0.co;2-a.

Karimipour, A., Bagherzadeh, S.A., Taghipour, A., Abdollahi, A. and Safaei, M.R. (2019). A novel nonlinear regression model of SVR as a substitute for ANN to predict conductivity of MWCNT-CuO/water hybrid nanofluid based on empirical data. *Physica A: Statistical Mechanics and its Applications*, 521, pp.89–97. doi:https://doi.org/10.1016/j.physa.2019.01.055.

Kaura, R. and Rajput, N. (2021). Future–Spot Relationship in Commodity Market: A Comparison Across Commodity Segments in India. *Global Business Review*, p.097215092110172. doi:https://doi.org/10.1177/09721509211017291.

Khalifa, A.A.A., Miao, H. and Ramchander, S. (2010). Return distributions and volatility forecasting in metal futures markets: Evidence from gold, silver, and copper. Journal of Futures Markets, 31(1), pp.55–80. doi: https://doi.org/10.1002/fut.20459.

Kohn, R. and Ansley, C.F. (1986). Estimation, Prediction, and Interpolation for ARIMA Models with Missing Data. *Journal of the American Statistical Association*, 81(395), pp.751–761. doi:https://doi.org/10.1080/01621459.1986.10478332.

Kriechbaumer, T., Angus, A., Parsons, D. and Rivas Casado, M. (2014). An improved wavelet–ARIMA approach for forecasting metal prices. *Resources Policy*, 39, pp.32–41. doi:https://doi.org/10.1016/j.resourpol.2013.10.005.

Li, G. and Li, Y. (2015). Forecasting copper futures volatility under model uncertainty. *Resources Policy*, 46, pp.167–176. doi:https://doi.org/10.1016/j.resourpol.2015.09.009.

Li, X., Shang, W. and Wang, S. (2019). Text-based crude oil price forecasting: A deep learning approach. *International Journal of Forecasting*, 35(4), pp.1548–1560. doi:https://doi.org/10.1016/j.ijforecast.2018.07.006.

Li, Z. and Zhang, L.H. (2013). An Empirical Study of International Linkages of the Shanghai Copper Futures Market. *The Chinese Economy*, 46(3), pp.61–74. doi:https://doi.org/10.2753/ces1097-1475460304.

Liang, W., Luo, S., Zhao, G. and Wu, H. (2020). Predicting Hard Rock Pillar Stability Using GBDT, XGBoost, and LightGBM Algorithms. *Mathematics*, 8(5), p.765. doi:https://doi.org/10.3390/math8050765.

Liew, V., Khim and Sen (2004). Which Lag Length Selection Criteria Should We Employ? *Economics Bulletin*, 3(33), pp.1–9.

Liu, C., Hu, Z., Li, Y. and Liu, S. (2017). Forecasting copper prices by decision tree learning. *Resources Policy*, 52, pp.427–434. doi:https://doi.org/10.1016/j.resourpol.2017.05.007.

Liu, K., Cheng, J. and Yi, J. (2022). Copper price forecasted by hybrid neural network with Bayesian Optimization and wavelet transform. *Resources Policy*, 75, p.102520. doi:https://doi.org/10.1016/j.resourpol.2021.102520.

Liu, Y., Yang, C., Huang, K. and Gui, W. (2020). Non-ferrous metals price forecasting based on variational mode decomposition and LSTM network. *Knowledge-Based Systems*, 188, p.105006. doi:https://doi.org/10.1016/j.knosys.2019.105006.

Lukac, L.P., Brorsen, B.W. and Irwin, S.H. (1988). A test of futures market disequilibrium using twelve different technical trading systems. *Applied Economics*, 20(5), pp.623–639. doi:https://doi.org/10.1080/00036848800000113.

Luo, H., Wang, D., Cheng, J. and Wu, Q. (2022). Multi-step-ahead copper price forecasting using a two-phase architecture based on an improved LSTM with novel input strategy and error correction. *Resources Policy*, 79, p.102962. doi:https://doi.org/10.1016/j.resourpol.2022.102962.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:https://doi.org/10.1007/978-3-540-27752-1.

Lütkepohl, H., 2009. Econometric analysis with vector autoregressive models. *Handbook of computational econometrics*, pp.281-319.

Méndez-Suárez, M., García-Fernández, F. and Gallardo, F. (2019). Artificial Intelligence Modelling Framework for Financial Automated Advising in the Copper Market. *Journal of Open Innovation: Technology, Market, and Complexity*, 5(4), p.81. doi:https://doi.org/10.3390/joitmc5040081.

Ni, J., Xu, Y., Li, Z. and Zhao, J. (2022). Copper price movement prediction using recurrent neural networks and ensemble averaging. *Soft Computing*. doi:https://doi.org/10.1007/s00500-022-07201-w.

Park, J. and Ratti, R.A. (2008). Oil price shocks and stock markets in the U.S. and 13 European countries. *Energy Economics*, 30(5), pp.2587–2608. doi:https://doi.org/10.1016/j.eneco.2008.04.003.

Pickup, M. (2014). *Introduction to Time Series Analysis*. *Google Books*. SAGE Publications.

Ramyar, S. and Kianfar, F. (2017). Forecasting Crude Oil Prices: A Comparison Between Artificial Neural Networks and Vector Autoregressive Models. *Computational Economics*, 53(2), pp.743–761. doi:https://doi.org/10.1007/s10614-017-9764-7.

Resources for the Future and Mikesell, R.F. (1979). *The world copper industry: structure and economic analysis*. Baltimore: Johns Hopkins Univ. Pr.

Salazar, E. and Weale, M. (1999). Monthly data and short-term forecasting: an assessment of monthly data in a VAR model. *Journal of Forecasting*, 18(7), pp.447–462. doi:https://doi.org/10.1002/(sici)1099-131x(199912)18:7%3C447::aid-for726%3E3.0.co;2-t.

Sánchez Lasheras, F., de Cos Juez, F.J., Suárez Sánchez, A., Krzemień, A. and Riesgo Fernández, P. (2015). Forecasting the COMEX copper spot price by means of neural networks and ARIMA models. *Resources Policy*, 45, pp.37–43. doi:https://doi.org/10.1016/j.resourpol.2015.03.004.

Scott Hacker, R. and Hatemi-J, A. (2008). Optimal lag-length choice in stable and unstable VAR models under situations of homoscedasticity and ARCH. *Journal of Applied Statistics*, 35(6), pp.601–615. doi:https://doi.org/10.1080/02664760801920473.

Sterba, J. and Hilovska, K., 2010. The implementation of hybrid ARIMA neural network prediction model for aggregate water consumption prediction. Aplimat—Journal of Applied Mathematics, 3(3), pp.123-131.

Stock, J.H. and Watson, M.W. (1988). Variable Trends in Economic Time Series. *Journal of Economic Perspectives*, 2(3), pp.147–174. doi:https://doi.org/10.1257/jep.2.3.147.

Vochozka, M., Kalinova, E., Gao, P. and Smolikova, L. (2021). Development of copper price from July 1959 and predicted development till the end of year 2022. *Acta Montanistica Slovaca*, pp.262–280.

Wang, C., Zhang, X., Wang, M., Lim, M.K. and Ghadimi, P. (2019). Predictive analytics of the copper spot price by utilizing complex network and artificial neural network techniques. *Resources Policy*, 63, p.101414. doi:https://doi.org/10.1016/j.resourpol.2019.101414.

Wang, Y., Wu, C. and Wei, Y. (2011). Can GARCH-class models capture long memory in WTI crude oil markets? *Economic Modelling*, 28(3), pp.921–927. doi:https://doi.org/10.1016/j.econmod.2010.11.002.

Watkins, C. and McAleer, M. (2002). Cointegration analysis of metals futures. *Mathematics and Computers in Simulation*, 59(1-3), pp.207–221. doi:https://doi.org/10.1016/s0378-4754(01)00409-8.

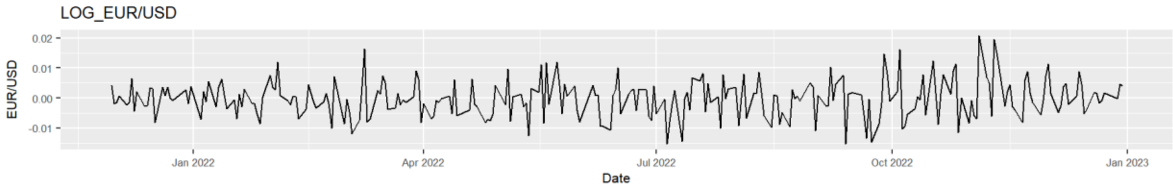
Yung, K. and Liu, Y.-C. (2009). Implications of futures trading volume: Hedgers versus speculators. *Journal of Asset Management*, 10(5), pp.318–337. doi:https://doi.org/10.1057/jam.2009.31.

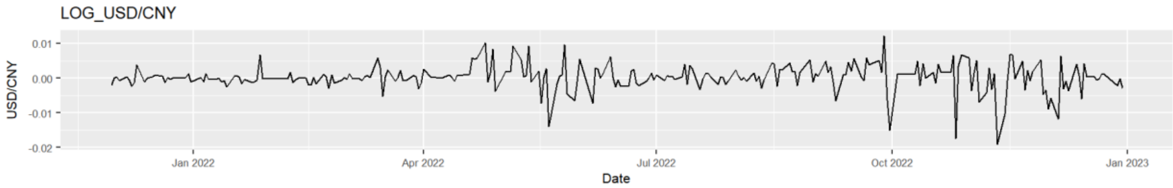
Zhang, H., Nguyen, H., Vu, D.-A., Bui, X.-N. and Pradhan, B. (2021). Forecasting monthly copper price: A comparative study of various machine learning-based methods. *Resources Policy*, 73, p.102189. doi:https://doi.org/10.1016/j.resourpol.2021.102189.

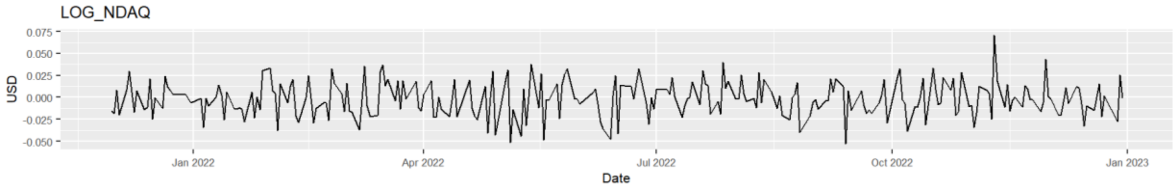
Zivot, E. and Wang, J. (2006). Vector Autoregressive Models for Multivariate Time Series. *Modeling Financial Time Series with S-PLUS®*, pp.385–429. doi:https://doi.org/10.1007/978-0-387-32348-0_11.

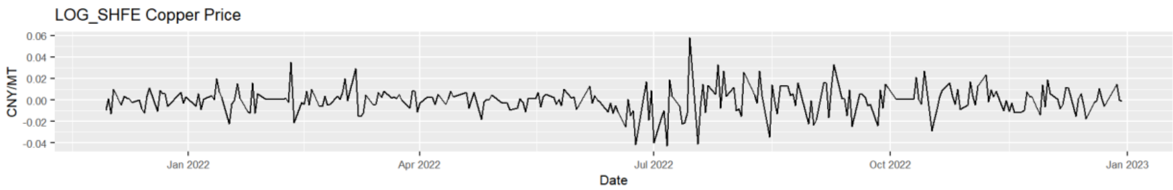
Appendix A

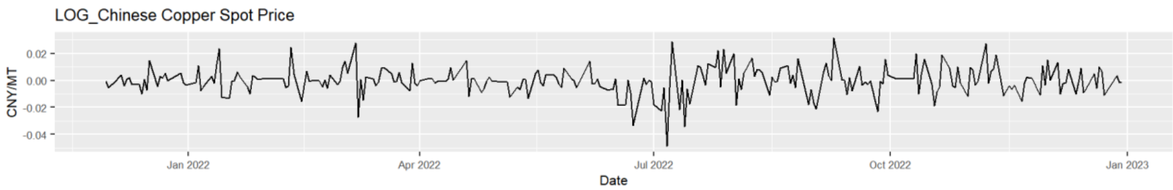
Plots of stationary variables

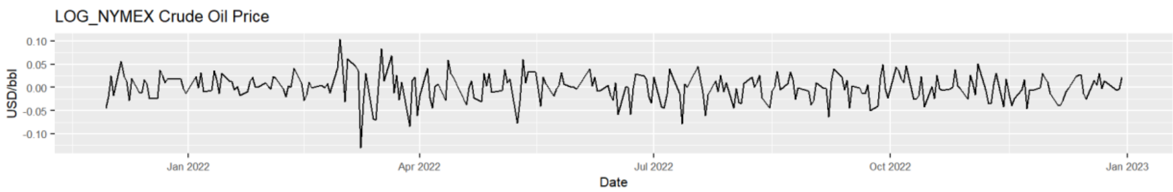


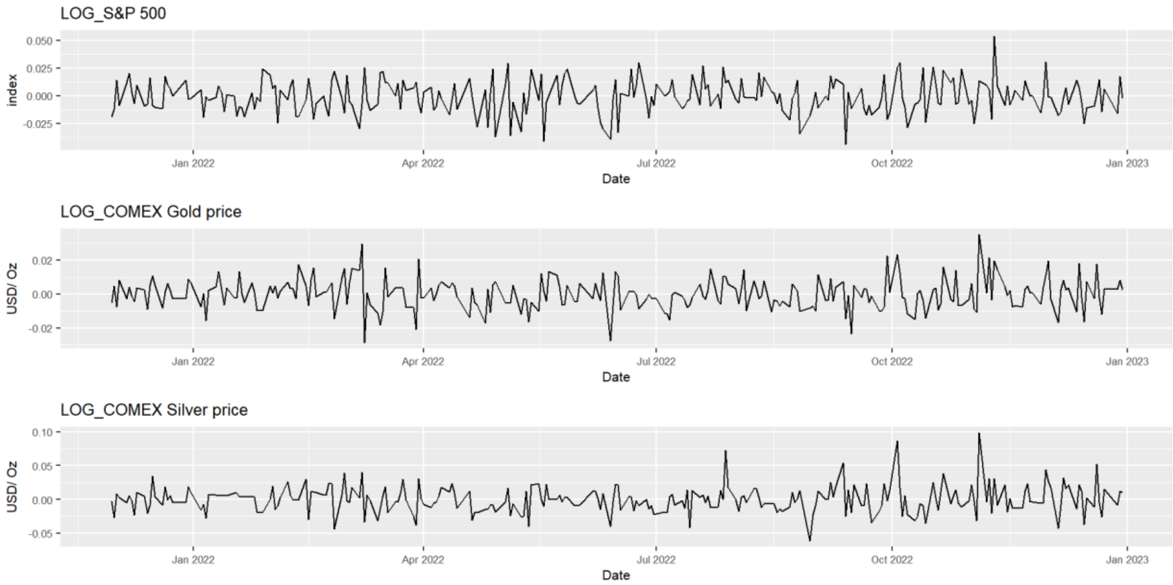











LOG_S&P 500
index
0.050
0.025
0.000
-0.025
Jan 2022
Apr 2022
Jul 2022
Oct 2022
Jan 2023
Date
LOG_COMEX Gold price
USD/ Oz
0.02
0.00
-0.02
Jan 2022
Apr 2022
Jul 2022
Oct 2022
Jan 2023
Date
LOG_COMEX Silver price
USD/ Oz
0.10
0.05
0.00
-0.05
Jan 2022
Apr 2022
Jul 2022
Oct 2022
Jan 2023
Date