A Hybrid Model for Analyzing the Effect of the Carbon Border Adjustment Mechanism on the Historical Volatility of the EU ETS

EIVIND RYTTER HUSEBY
ALEXANDER ROSÅSEN LONGNES

SUPERVISOR
Jochen Jungeilges

University of Agder, 2024
Faculty of Business and Law
Department of Economics and Finance

Preface

This thesis concludes our master studies in Business and Administration with a specialization in Analytical Finance at the School of Business and Law at the University of Agder (UiA).
We would like to use this opportunity to thank our families, friends and fellow students for interesting discussions, fun moments and support during our academic journey. In particular, to our supervisor Professor Jochen Jungeilges, we wish to show our gratefulness for all input, support and guidance in fulfilling this master thesis.

Abstract

The 1^{st} of October 2023 the Carbon Boarder Adjustment Mechanism (CBAM) transitional phase began. The CBAM is a tool created by the European Union (EU) to reach the target of 55% reduction in greenhouse gas emission levels in 2030 compared with 1990 levels. The CBAM imposes customs on sectors at high risk of carbon leakage. This thesis aims to analyze the historical volatility of the EU Emissions Trading System (ETS) for the end of phase 3 and for phase 4 until the 29^{th} of February 2024. The data consists of the price of indexes related to the CBAM and the price of natural gas. Motivated by a methodology created by Amirshahi and Lahmiri (2023), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), exponential GARCH, Glosten-Jagannathan-Runkle-GARCH models are created based on different assumptions for the distribution of the residuals. In addition, a model is created using Long Short-Term Memory (LSTM). The optimal GARCH-type model for each variable is used to create a hybrid GARCH-LSTM model. Model performance is compared based on the root mean squared error (RMSE) and the mean absolute error (MAE). The results show that the GARCH-LSTM model outperforms the alternatives in terms of RMSE and MAE. In addition, the model shows that the predictability increases in phase 4. The thesis provides the EU with a tool to determine the right policies to ensure a carbon price that does not undermine the main goals of the CBAM.

KEYWORDS: EU ETS, CBAM, GARCH-LSTM, RESIDUAL DISTRIBUTION

Contents

List of Figures

List of Tables

1 Introduction

Climate changes and environmental pollution are major global issues that pose threats to the *"economic, social, political and cultural"* segments of society (Agwu et al., 2021). The global emission levels have increased significantly and pose a threat to the economy, as well as a sustainable future (Flori et al., 2024; Wade and Jennings, 2016). In the nineties, scientific evidence of climate changes began to convince the international community that preventive actions were needed (Massai, 2011, p. 39). In 1997, a UN decision-making body (the Conference of the Parties) decided to adopt the Kyoto Protocol. It is aimed at committing industrialized countries and economies to reduce greenhouse gas (GHG) emissions (Massai, 2011, p. 40; Cary and Stephens, 2024). With the Kyoto Protocol in place, the global effort to reduce climate change and achieve climate neutrality by 2050 began. According to Newell et al. (2013), the debate in the public sector focused on how a global market for trading carbon allowances (or carbon credits) could be designed *"as 'the' vehicle to address global climate change"*. This global market never happened. Eslahi and Mazza (2023) state that carbon allowances represent the right to emit one tonne of carbon dioxide (CO_2) or the equivalent amount of other greenhouse gases in a specific period of time.

In March 2000, the European Commission presented a green paper outlining the initial ideas for a new trading system, called the European Union Emission Trading System (EU ETS). The system was created to serve an important role in reducing the environmental impact of the power sector and heavy industry (Ellerman et al., 2010, p. 22). In 2005, the EU ETS was officially launched (Newell et al., 2013). The EU ETS is the world´s first international trading system and the main idea behind one of the largest initiatives to reduce the European climate impact. The EU ETS forces polluters to pay for their GHG emissions, with the goal of incentivizing a reduction of the European climate impact. Moreover, it will generate revenue for the EU, that can be used in supporting businesses in achieving a greener future (European Commission, n.d.-c). The EU ETS works as a cap-and-trade system where a limit is set on the total amount of allowances distributed (Ellerman et al., 2010, p. 32; European Commission, n.d.-c). Companies are able to buy carbon credits on the market to compensate for emissions that exceed their allocated allowances (Ellerman et al., 2010, p. 131).

1.1 Implementation of the EU ETS

The establishment of the EU ETS has happened gradually, and it is currently in its fourth phase. According to Boungou and Dufau (2024), this gradual introduction of the EU ETS ensures a mitigated loss of competitiveness for firms in regulated sectors. Moreover, this approach offers companies sufficient time to adjust and invest in low-carbon technologies. In 2005, the first phase (2005-2007) of the EU ETS implementation was initiated, mainly focusing on *"the power sector and certain heavy industry"* (Ellerman et al., 2010, p. 18; Newell et al., 2013). The phase was considered as a learning period for the EU, with the main target of data and information gathering. Thereby securing a smooth transition to phase 2 (2008-2012). From phase 1, the EU can take away a successful implementation of a carbon market and free carbon credits trade across the EU (Balietti, 2016). Throughout phase 1, almost all emission allowances are freely allocated to eligible companies (Newell et al., 2013). To ensure the use of allowances, the EU imposed a fee of €40 per tonne of CO_2 as a penalty for companies not complying with the EU ETS (European Commission, n.d.-b). To put this in perspective, the price of emission allowances on the market during this period never exceeded a daily price of €30.450 [1].

The Kyoto Protocol states that the second phase of the implementation of the EU ETS would mark the first commitment period. In other words, agreed emission reduction targets from the Kyoto Protocol needed to be met as of 2008 (Kim, 2021). During phase 2, it was agreed that GHG emission levels should be reduced with up to 8% compared to the emission levels from 1990 (Ellerman et al., 2010, p. 18). The data obtained during phase 1 provided the EU access to annual emissions data. The data gave the EU a helpful tool in adapting the level of cap for the carbon credits (Vollebergh and Brink, 2020). According to Vollebergh and Brink (2020), a good cap system is key for a cap-and-trade policy. This resulted in the EU reducing the cap by 6.5% compared to 2005 levels. In addition, the number of free allowances was reduced to 90% (European Commission, n.d.-b). The global financial crisis of 2008 resulted in significantly greater emission reductions than expected (European Commission, n.d.-b; Zhang and Wei, 2010). Consequently, the market experienced a surplus of emission allowances, which impacted the price of carbon credits during phase 2 (Castagneto-Gissey, 2014; European Commission, n.d.-b; Vollebergh and Brink, 2020).

During phase 3 (2013-2020), the main target for the EU was to reduce the emissions of GHG by an average of 20% compared to the emission levels from 1990 (IEA, 2020). The EU added two features to reach the target. Firstly, the EU introduced a single EU-wide cap on emissions instead of national caps. Secondly, auctioning became the standard method of allocating carbon credits (Newell et al., 2013; Vollebergh and Brink, 2020). In phase 1, all allowances were distributed for free, whereas phase 2 can be recognized by the number of free allowances reducing to around 90%

[1] https://tradingeconomics.com/commodity/carbon

(European Commission, n.d.-b). However, in phase 3, this number would gradually be reduced to 30%, in line with the EU´s climate target (Capros et al., 2019; European Commission, n.d.-a).

The fourth phase (2021-2030) began in 2021 and expands the trading scheme to cover more sectors deemed as high risk for carbon leakage (Boungou and Dufau, 2024). Carbon leakage refers to a situation in which a company relocates its production to a region with lower climate ambitions and fewer penalties of emitting greenhouse gases (Pan and Yu, 2024). Companies that are classified as high risk will be placed on the carbon leakage list and subsequently receive allowances equivalent to 100% of the relevant benchmark for free (European Commission, n.d.-a). The companies that are not on the carbon leakage list will be part of the allocation for free allowances up to 30% until 2026, before the free allowances will be phased out after 2030 (European Commission, n.d.-a). One of the main targets in phase 4 is to reduce the emissions of greenhouse gases by up to 43% within 2030 compared to the emission levels in 2005 (European Commission, n.d.-a). The EU aims to achieve this target by expanding the scope of the EU ETS to include sectors such as maritime and aviation. Additionally, the EU has implemented the Market Stability Reserve (MSR). The purpose of the MSR is to adjust the surplus of allowances and strengthen the system against major market shocks by adjusting the number of allowances available for auction. This tool is useful for the EU to stabilize the market (Osorio et al., 2021). A significant change to the free allowances system is introduced in phase 4, namely the Carbon Border Adjustment Mechanism (CBAM).

1.2 Carbon Border Adjustment Mechanism

In 2019, the European Green Deal was introduced, presenting a plan to reduce GHG emissions by at least 55% by 2030 compared to the 1990 levels. The end goal is to achieve climate neutrality in Europe by 2050 (Pietzcker et al., 2021). Within the framework of the EU Green Deal, the CBAM was introduced as a tool to reduce emissions and the risk of EU based companies moving to countries with lower environmental focus (Schauenberg, 2022). Furthermore, the CBAM seeks to encourage other countries to implement a more sustainable economy (Schauenberg, 2022). In addition, a less carbon-intensive industry is motivated by securing lower exporting costs to the EU and increasing the competitiveness of EU based companies by securing that they are not affected by the CBAM import costs (Cho et al., 2024). The CBAM imposes tariffs on imports of carbon-intensive goods and goods exposed to risk of carbon leakage. In its initial phase, the CBAM will apply to the markets of electricity, iron & steel, fertilizer, hydrogen, cement, and aluminum (Schauenberg, 2022). A phase-out of free allowances in the EU ETS will commence with the CBAM. This phase-out will occur with a yearly reduction until free allowances are completely replaced by CBAM certificates by 2035 (Benson et al., 2023; Cho et al., 2024).

From 2023 to 2026, the CBAM operates in a transitional phase, allowing entities to prepare for its full operationalization. The first reporting period for the CBAM was from 1^{st} of October 2023 to 31^{st} of January 2024 (European Commission, 2024). The reporting period aims to provide insights, data and information on how market mechanisms evolve through the gradual introduction of the CBAM. During this period, importers must prepare CBAM reports that contain information on products and their actual emissions (Cho et al., 2024). With these measures, members need to adapt to comply with new laws, regulations, and guidelines. By the end of the transitional phase in late 2025, the EU aims to have collected sufficient data and information about how the CBAM operates in the market. In 2026, the EU targets a fully operational CBAM (European Commission, 2024).

1.3 The Research Gap Addressed

The carbon market is increasing in size, and according to Segnon et al. (2017) the European Union Allowances (EUAs) market has become the largest worldwide market for CO_2 emissions. The EU ETS has been investigated by many researchers (Chevallier, 2009; Eslahi and Mazza, 2023; Vollebergh and Brink, 2020). However, to the best of our knowledge, there exists no previous research on the effect the chosen variables have on the EU ETS. This thesis seeks to close this gap by providing models for estimating the impact of the CBAM affected sectors on the EU ETS. The CBAM is in a transitional phase, thus the EU is seeking research to secure a good implementation of the CBAM. Previous research on this topic have mainly focused on unfairness in regard of developing countries (Eicke et al., 2021; Gläser and Putaturo, 2022; Lowe, 2021) or legal issues in regard of the CBAM not complying with rules of the World Trade Organization (WTO) (Kaufmann and Weber, 2011; Mehling et al., 2019). This thesis differentiates from the previous research by providing an approach the EU can use to analyze the effect that implementing the CBAM will have on the historical volatility (HV) of the EU ETS. Understanding the pricing of the CBAM certificates is crucial to establish a clear connection between the EU ETS allowances and the CBAM. The European Commission defines the price of the CBAM certificates as a *"weekly average auction price of EU ETS expressed in €/tonne of CO_2 emitted"* (European Commission, 2024). The models developed in this thesis provide a tool for the EU to understand how the historical volatility of the carbon credits will be impacted by the implementation of the CBAM. A successful model will support the EU in determining how they can adjust their policies to avoid carbon leakage by preventing that companies find attractive alternatives. In section 3, plots and tables determining

the historical volatility and fluctuations of prices of EU ETS show why an established model is important. The thesis provides evidence that the sectors implemented in the CBAM are useful to predict the historical volatility of the EU ETS during phase 3 and phase 4. The suggested models are useful to determine how the chosen indexes will affect the EU ETS. Understanding and estimating the prices and the volatility for the CBAM are also of great interest for firms as well as traders. The importance of carbon credits for firms was analyzed by Dewaelheyns et al. (2023), who discovered that when policy-induced carbon risk increased, firms that do not have enough carbon allowances realize value discounts despite carbon prices being low. Oestreich and Tsiakas (2015) support these findings, by arguing that a higher cash flow due to free allowances presents a statistically significant carbon premium for the firms receiving free allowances.
This thesis contributes to the research by introducing hybrid models consisting of Generalized Autoregressive Conditional Heteroskedasticity (GARCH)-type models and a Long Short-Term Memory (LSTM) model to analyze the EU ETS. To the best of our knowledge, this has not been exercised with the CBAM sectors during phase 3 and phase 4. Furthermore, the thesis provides evidence that the GARCH-LSTM model outperforms the chosen GARCH-type model based on the chosen decision criteria. In addition, it outperforms the LSTM model. Moreover, the thesis provides evidence of an increased predictability during phase 4.

This thesis investigates the theoretical foundations that underpin the choice of model. Subsequent to the theoretical exposition, the data employed in our analysis will be introduced in section 3. Thereafter, section 4 will provide a detailed background of the methodology. Furthermore, the empirical results from the application of these methods are presented in section 5. The thesis culminates with a discussion in section 6, wherein the results, the relevance of the results, the implications from the model and future research are discussed. Lastly, section 7 provides a conclusion to summarize the findings.

2 Literature review

2.1 The Theoretical Framework for the CBAM

Researchers have debated the CBAM since before its implementation by the EU. According to many researchers, the aim with the CBAM is to increase the competitiveness of European industry and reduce the carbon leakage (Böning et al., 2023; Magacho et al., 2024; Zhong and Pei, 2024). Many have already argued in favor of the CBAM being an efficient tool in reducing carbon leakage (Mehling et al., 2019; Mörsdorf, 2022; Takeda and Arimura, 2024)). Mörsdorf (2022) analyzed the environmental effects of the CBAM. Their paper showed that implementing the CBAM will have an effect on reducing emissions in Europe and address the problem of carbon leakage. Bellora and Fontagné (2023) used a dynamic general equilibrium model, as well as GHG emissions, input-output relations, and price of emission quota to analyze the economical and environmental impact of the CBAM. They showed that the CBAM is effective in reducing carbon leakage. Their paper further suggests that the CBAM will cause the prices of European carbon credits to increase. The authors argue that the phasing out of the free allowances will cause emission intensive trade exposed (EITE) entities to no longer be able to utilise the free allowances, thus increasing the demand for carbon allowances in the market. Takeda and Arimura (2024) analyzed the effect of the CBAM against the Japanese economy. Their findings showed that the CBAM reduced carbon leakage and resulted in a small positive impact on the gross domestic product (GDP) and welfare in Japan. Furthermore, Sun et al. (2023) analyzed different CBAM schemes to assess the environmental and economic impact. The findings of Sun et al. (2023) show that implementing the CBAM reduces carbon leakage with 19%, but they also found that the CBAM overburdens developing countries. In general, the papers above show that the CBAM overall will have a positive impact on the reduction of carbon leakage. This is in contrast to the findings of Jousseaume et al. (2021). They suggest that the implementation of the CBAM *"significantly increase the risk of carbon leakage"* and argue that this is a result of the CBAMs intention of phasing out the free allowances. Despite Sun et al. (2023) suggesting a reduction of carbon leakage by implementing the CBAM, they argue that the CBAM is inefficient given the price of carbon allowances of $90/tonne, and that the higher the price, the more inefficient the CBAM becomes.

2.2 Research on European Union Allowances

Ever since its launch in 2005, the EUAs have been investigated by researchers. Chevallier (2009) examined the relationship between the returns of carbon futures and changes in macroeconomical conditions. Chevallier (2009) proved that using the premium of junk bonds and equity dividend yields, a weakly forecast of carbon credits return is possible. Vollebergh and Brink (2020) wrote an informative paper investigating the outcome of the EU ETS. They conclude that the EU ETS contributes to the reduction of emissions, even going so far that they claim it *"guarantees reduction in carbon emissions in the long run"*. Other researchers have performed analysis on the relationship between carbon prices and different commodities. Eslahi and Mazza (2023) analyzed the first three phases of the EU ETS to determine if weather and electricity demand can be used to predict the prices of EUAs. Using an extreme gradient boosting technique, they observe that their variables proved to have predictability on the EUAs.

B. Lin and Zhao (2023) analyzed the effects of CBAM on steel and aluminum futures in the Chinese carbon market. In their research, the authors observed that the implementation of the CBAM will have a strong negative impact on steel and aluminum futures. In addition, they proved that for aluminum, the negative effect was larger, and argued that this is mainly due to the emissions being higher in aluminum production. In another paper, B.-Q. Lin and Zhao (2023) analyzed which sectors that the CBAM should cover. In their findings, they suggest that countries such as Russia, United States, China and the United Kingdom are the main source of EITE products to the European market. Moreover, they suggest that the CBAM should be implemented on the plastics, phosphorus fertilisers, aluminum, and copper industries, due to these having a high potential of reducing emissions without having an exorbitant economic impact.

2.3 Forecasting Volatility

Forecasting volatility in the carbon markets has become an increasingly investigated topic. Viteva et al. (2014) use implied volatility (IV) to analyze the carbon options from the European climate exchange (ECX). They observed that the forecasts obtained from IV on the carbon options on the ECX were highly informative. Another model used for forecasting volatility is the GARCH model introduced by Bollerslev (1986) as a generalization of the ARCH model of Engle (1982). Models of GARCH-type have become popular in volatility forecasting. According to Brooks (1997), GARCH-type models are highly important for financial econometric.
Byun and Cho (2013) analyzed the futures market of carbon credits using different GARCH models. They proved that compared to a k-nearest neighbor model and an IV model, the GARCH-type models performs better. They also analyzed the different GARCH models against each other and concluded that the Glosten- Jagannathan-Runkle (GJR)-GARCH with a normal distribution is outperforming the other GARCH-type models. Other researchers have analyzed the EUAs

using different GARCH-type models. Boersen and Scholtens (2014) used a threshold GARCH (tGARCH) to analyze the relationship between the electricity market and the spot price of the EU ETS allowances. They conclude that with the tGARCH, gas and oil amongst others are significant determinants of the carbon futures prices during phase 2. They argue that this is because the price of electricity in the European market is determined by the prices of its fuel inputs. Another GARCH model researchers have used to investigate carbon allowances is the exponential GARCH (eGARCH). Villar-Rubio et al. (2023) used an eGARCH model to predict the volatility in phase 3 of the EU ETS. They found that the eGARCH outperformed both the ARCH and GARCH models. In general, the papers above show that despite the standard GARCH model being useful, other GARCH-type models tend to outperform it when analyzing the carbon markets.

2.4 Long Short-Term Memory Models in Finance

More recently, the use of machine learning algorithms have increased in finance. Since Hochreiter and Schmidhuber (1997) introduced the Long Short-Term Memory (LSTM) algorithm it has become widely used for forecasting financial time series. Bhandari et al. (2022) used an LSTM algorithm to predict the day-ahead closing price of the S&P500 index. They concluded that a single layer LSTM outperformed a multilayer LSTM based on the root mean squared error (RMSE). Their findings are further proved by Fischer and Krauss (2018) who conducted a research on the S&P500 where they tested the LSTM against three different benchmark models: deep nets, random forests, and logistic regression. Their findings suggest that the LSTM model outperforms the benchmarks by showing a higher predictability. This has also been observed by Siami-Namini et al. (2018) who compared an Autoregressive Integrated Moving Average (ARIMA) with an LSTM model. They showed that the LSTM model tends to reduce the obtained error rates with between 84%-87% compared to the ARIMA model.

2.4.1 Challenges of Machine Learning

Despite many great results, machine learning does not come without its challenges. "Black-box" is a term often associated with machine learning models. It describes a phenomenon where a model lacks explanatory power on the output (Sen et al., 2021, p. 6). Output is often readable, and decision makers will be able to interpret the result of the model to some extent. However, arguing for the suitability of the models will become a difficult task (Sen et al., 2021, p. 6). For machine learning to produce great generalized results, the training data is required to be representative. Generalization in machine learning indicates the model´s ability to predict new data out of the training set (X. Y. Huang et al., 2023). If the training data is of low quality, with including errors, as well as extreme values and noise, the model is likely to perform worse due to the machine learning model struggling to identify the underlying patterns (Géron, 2019, p. 25). If the model performs well on training data, but does not generalize well, the training data might be overfitted. When the training data is overfitted, the machine learning model predicts accurately for the training data, but not on new data (Dietterich, 1995). Further, Dietterich (1995) establishes that this is due to the model remembering some of the peculiarities that is observed in the training set instead of the *"predictive rule"*. Moreover, the author claims the goal of a machine learning model is not to perform well on the training set, but rather the test set (Dietterich, 1995). Another challenge with machine learning is underfitting the training data. This is the opposite of overfitting and occurs when the model is too simple to learn the underlying structure of the data and therefore does not generalize well (Géron, 2019, p. 29, Ying, 2019). This results in inaccurate and unreliable predictions.

Hyperparameter tuning poses a challenge in machine learning due to the difficulty and time-consuming nature of setting appropriate values, which often results in inefficiency (Géron, 2019, p. 31). The hyperparameters are important, since they will influence performance and generalization capability of the model (Wojciuk et al., 2024). The values of the hyperparameters that generate the model with the lowest generalization error is preferred. The challenge occurs when the model is used and the output is a generalization error which is higher than expected. This is a result of the original model being adapted to produce the best result for a particular set of data and will therefore be unlikely to perform as well on new data (Géron, 2019, p. 31). Furthermore, too many hyperparameters can be problematic. If the model has too many hyperparameters, the possibility of overfitting the validation set increases. This results in the model performing well on a validation set, but fails to generalize unseen data (Géron, 2019, p. 31).

2.5 Hybrid Models

More recent methodologies combine an econometric model with a machine learning algorithm. These models, often referred to as hybrid models, draw on the advantage from different models to increase predictability. Using an attention mechanism LSTM, Luo et al. (2022) analyze the carbon market and show good results for the model when predicting the carbon prices. Duan et al. (2023) utilized the same method and analyzed the EU carbon futures. Their results support the findings of Luo et al. (2022) showing that the attention mechanism LSTM improves the predicting of carbon futures (Duan et al., 2023, Luo et al., 2022). Y. Huang et al. (2021) analyzed the EU ETS allowances using a combination of a variational mode decompositional model, a GARCH model and an LSTM model. They show that their model performs better than solely using an econometric or a machine learning model. They further suggest that their model is usable for forecasting the phase 3 carbon markets return (Y. Huang et al., 2021).

Liu et al. (2024) analyzed the Chinese carbon market using a GARCH, LSTM, and a hybrid GARCH-LSTM model. Their findings suggest that the GARCH-LSTM model outperforms a single GARCH model and a standalone LSTM model when RMSE is used as performance measure. They conclude that adapting a hybrid GARCH-LSTM model will significantly improve predictions on carbon prices (Liu et al., 2024). This provides a valid argument to utilize a similar method on research of the European carbon market.

2.6 Methodical Background

The methodology used in this thesis is motivated by Amirshahi and Lahmiri (2023). In their paper, the authors forecast the historical volatility on 27 different cryptocurrencies. They use three different GARCH models, the standard GARCH, an eGARCH, and an asymmetric power GARCH (APGARCH). Moreover, they assume the residuals from three different distributions, the normal distribution, the student's t-distribution, and the generalized error distribution. Further, Amirshahi and Lahmiri (2023) consider an LSTM model, before they show that a hybrid GARCH-LSTM model, and specially the APGARCH-LSTM is outperforming the GARCH models or the LSTM model for almost all the cryptocurrencies. The idea of combining different distributions has been adapted by other papers as well (Abdullah et al., 2017; Rahman et al., 2023). Abdullah et al. (2017) compared the student's t-distribution and normal distribution for different GARCH-type models and showed that implying the student's t-distribution made the forecasting of the models better than for the normal distribution. Rahman et al. (2023) analyzed the crude oil price volatility and showed the same results. In addition, they proved that the optimal GARCH model for forecasting the oil price volatility is an eGARCH with an assumption that the residuals are student's t-distributed.

When working with many models, there is a need for performance metrics to determine the optimal model. In terms of machine learning and time series analysis, Karunasingha (2022) presents the RMSE and the MAE as widely used measures for model comparison. The author uses a zero mean symmetric error distribution to analyze the statistical properties of the RMSE and MAE estimators. Based on the findings, it was concluded that it is not possible to compare the RMSE and MAE unambiguously for errors with different distributions. Willmott and Matsuura (2005) examine the abilities of the RMSE and the MAE in describing average model performance. In their findings, they suggest that the MAE has an advantage over the RMSE, due to the MAE being an unambiguous measure. In another paper, Willmott et al. (2009) illustrate the ambiguous nature of sums-of-squares-based measures. Further, Willmott et al. (2009) argue that measures based on sums-of-squares are depending on both the average of the error, as well as the variations in the set of errors. Moreover, they conclude that sum-of-squares-based measures do not have a clear interpretation. Chai and Draxler (2014) have a different view, as they argue against the RMSE being ambiguous. In addition, they claim that when the distribution of the error is expected to be Gaussian, the RMSE is more suitable than the MAE. An important remark in the paper of Chai and Draxler (2014) is that using one metric results in solely emphasizing some *"certain aspect of the error characteristics"*. They conclude that analyzing model performance requires a combination of different metrics.

3 The Data

The variables in this thesis are chosen based on the sectors the CBAM will be implemented in. The data is retrieves from Refinitiv Eikon Datastream[2] and consists of different non-EU indices, in addition to the price of natural gas. The data consists of daily price observations from 28^{th} of February 2017 to the 29^{th} of February 2024. It is split into two periods accommodating for the different phases of the EU ETS. The data from the 28^{th} of February 2017 to the 31^{st} of December 2020 represents the end of phase 3, and the data from the 1^{st} of January 2021 to the 29^{th} of February 2024 represents the beginning of phase 4. In our thesis, we have opted to exclude hydrogen from the analysis due to the unavailability of pertinent datasets.

Figure 1 presents the daily price evolution for the six variables used in this paper. The EU ETS variable is from the intercontinental exchange, and represents the price of the carbon emissions futures. The observed indices for all other variables, except gas, are created by combining different assets operating within each sector. The cement index consists of a combination of Colombian and Australian traded equity and preference shares. For aluminum, there is a combination of an Australian and a Norwegian[3] traded stock. For the other indices, the process remains consistent, wherein assets of non-EU companies are combined to create indices. The variable for natural gas is chosen because of gas being a widely used source of fossil fuels in the EU[4]. In addition, the paper by Boersen and Scholtens (2014) argues that natural gas and oil prices have an impact on the EU ETS during phase 2. Oil is neglected due to the small impact it has on electricity generation in the EU [4]. To retrieve the data used in this thesis, a computer with access to Refinitiv Eikon is needed, and one can search for the text in this footnote [2]. Due to the day of the data fetching, the time span is sat to eight years. This would need to be adjusted if this research is to be replicated.

To work with the data, the prices are turned into logarithmic returns (log returns) using equation 1.

$$r_t = log(P_t) - log(P_{t-1}) \tag{1}$$

The price evolution presented in figure 1 shows all variables from the 28^{th} of February 2017 to the

[2]=@Thomson.Reuters.AFOSpreadsheetFormulas.DSGRID("CFI2Z1C,FRTLZCC,STEELCC,CEMNTCC,ALUMNCC ,NNGCS30";" ";"-8Y";"";"D";"RowHeader=true;ColHeader=true;DispSeriesDescription=false;YearlyTSFormat=false ;QuarterlyTSFormat=false;MonthlyTSFormat=False";"")

[3]Disclaimer: As far as to our knowledge, it is not clear if European Economic Area (EEA) or European Free Trade Association (EFTA) members must pay CBAM customs

[4]https://www.consilium.europa.eu/en/infographics/how-is-eu-electricity-produced-and-sold/

29^{th} of February 2024. The red color represents phase 3 while the green color represents phase 4 of the EU ETS. During both phases, the EU ETS experienced a large increase in price. Furthermore, iron & steel is the only other variable where the price increases during both periods. The cement and aluminum indices experienced a significant price drop during phase 3. Interestingly, aluminum recovered and exceeded phase 3 values in 2022, before a massive price drop in 2022. The price observed the 29^{th} of February 2024 is similar to the observed value at the end of phase 3. Despite a price increase during 2024, cement did not recover before the end of phase 4. This is due to a significant price drop in 2022.

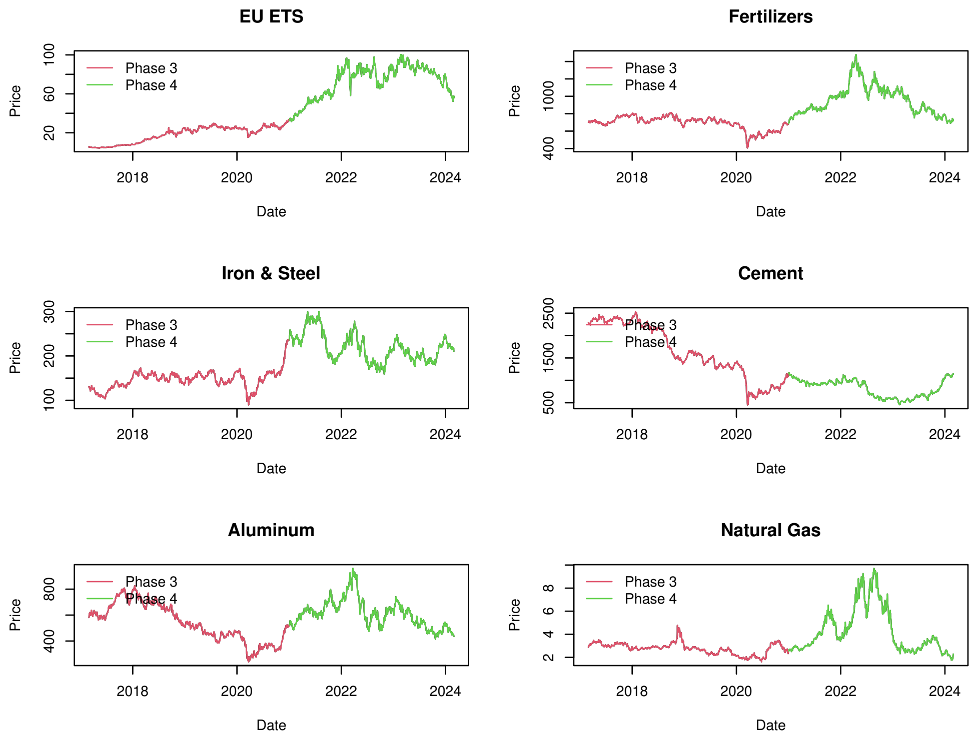


Figure 1: The price evolution of the variables from the 28^{th} of February 2017 to the 29^{th} of February 2024. Phase 3 colored red, phase 4 colored green.

In figure 1, there are especially two dates that are worth noticing due to being special events in the world in the period between the 28^{th} of February 2017 to the 29^{th} of February 2024. With special events there are significantly higher chances for the prices of the variables to be influenced (Balash and Faizliev, 2024). The first special event happened in March 2020 when the COVID-19 pandemic resulted in a worldwide lockdown. With a world in lockdown the demand for raw materials decreased significantly, which furthermore resulted in decreased prices. In figure 1, a decrease in price is observed in the early phase of the pandemic. For the following months the prices of the EU ETS and natural gas experienced a small increase. All price variables experienced an increase in the months following the price drop. The second special event began on the 24^{th} of February 2022 when Russia launched an invasion on Ukraine. The invasion influenced the six price variables and created some interesting observations. Iron & steel, fertilizers, aluminum, and natural gas experienced high price increases. Presented in figure 1, natural gas experienced approximately a 100% increase in price in 2022, while the price of the fertilizers and iron & steel variables increased with around 50%. Due to the invasion, the EU has introduced sanctions against Russia. The supply of natural gas in Europe decreased significantly, which resulted in increased prices. Around 2023, the price of natural gas dropped to the price it was before the invasion. Before the invasion, Russia was one of the biggest importers of pipeline gas to Europe. In 2021, over 40% of the EU's consumption of pipeline gas originated from Russia. In 2023, Russia's share decreased to around 8% (Council of the European Union, 2024).

3.1 Descriptive Statistics

Table 1 presents the descriptive statistics for the log returns of all assets, ranging from the 28^{th} of February 2017 to the 29^{th} of February 2024. In Table 1, the mean in %, standard deviation in %, skewness, kurtosis, and the p-values for four statistical test tests are presented. EU ETS, fertilizers, and iron & steel have positive mean log returns of 0.1297%, 0.0013% and 0.0265%, respectively. Whereas cement, aluminum, and natural gas have negative mean log returns with respectively -0.0380%, -0.0158% and -0.0133%. Further, the standard deviation for the EU ETS is 2.8135%, whereas the standard deviation is 1.6169% for fertilizers, 1.9511% for iron & steel, 2.0898% for cement, 2.1627% for aluminum, and 3.2604% for natural gas. With standard deviations of 2.8135% and 3.2604%, EU ETS and natural gas are more volatile than fertilizers, iron & steel, cement, and aluminum, which in comparison have standard deviations ranging from 1.6169% to 2.1627%.

Furthermore, table 1 shows that all variables have a negative skewness. Skewness defines the shape of the distribution, and measures to which extend it is not symmetric about its mean value (Brooks, 2002, p. 179). Since the skew for all variables are negative, the distributions are skewed to the left, which means that the distributions have a longer/fatter tail on the right-hand side. Moreover, the kurtosis of all variables ranges from 6.0188 to 34.0133, with aluminum having the lowest kurtosis and cement having the highest kurtosis. Kurtosis measures the fatness of the tails of the distribution (Brooks, 2002, p. 179). In the Gaussian distribution, the kurtosis is defined to be equal to three. All variables have kurtosis greater than three, which ensures that extreme values will be observed more often than under a Gaussian distribution. Since the kurtosis of all variables are greater than three, all marginal distributions are leptokurtic.

Table 1: The Descriptive Statistics for the log return of the variables from 28.02.2017 to 29.02.2024

	EU.ETS	Fertilizers	Iron & Steel	Cement	Aluminum	Natural.Gas
Mean in %	0.1297	0.0013	0.0265	-0.0380	-0.0158	-0.0133
Standard Deviation in %	2.8135	1.6169	1.9511	2.0898	2.1627	3.2604
Skewness	-0.4630	-0.6810	-0.3815	-0.6009	-0.3867	-0.0304
Kurtosis	7.7821	9.3481	10.0642	34.0133	6.0188	8.3641
p.value JB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p.value LM	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p.value LB	0.0133	0.0018	0.0009	0.0000	0.0053	0.0151
p.value ADF	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100

Table 1 presents the p-values for four different statistical tests. The Jarque-Bera test (JB) is a statistical test where the null hypothesis states that the price returns are normally distributed, while the alternative hypothesis states that the price returns are not normally distributed. The observed p-values of 0.0000 indicate that the null hypothesis is rejected, providing statistical evidence that all price return series deviates from the normal distribution.

In the Lagrange-Multiplier test for ARCH effects (LM), the null hypothesis states that there is no conditional heteroscedasticity in the price returns. The alternative hypothesis states that conditional heteroscedasticity exists (Catani and Ahlgren, 2017). Table 1 shows that the p-values of the LM test with a max lag of 16 are observed to be equal to zero. Thus, there exists statistical evidence to reject the null hypothesis at a 5% level. There exists statistically significant evidence at a 5% level that until a lag of 16 all variables show the existence of conditional heteroscedasticity.

Moreover, the Ljung-Box (LB) test was performed with a lag of 16. The null hypothesis of the LB test states that the price return series is white noise. The alternative hypothesis states that the price return series is not white noise (Hassani and Yeganegi, 2019). The results from table 1 show that the p-value ranges from 0.0000 to 0.0151. Since all the p-values are below 0.05, the null hypothesis is rejected. Therefore, there exists statistically significant evidence at a 5% level that the price return series are not white noise.

Lastly, the Augmented Dickey-Fuller (ADF) with lag of 16 is performed. The null hypothesis in the ADF is that there exists unit root in the price return series. The alternative hypothesis states a unit root is not present in the price return series. As displayed in table 1, the p-values are below 0.05, which results in the rejection of the null hypothesis. Hence, evidence exists for the price return series being stationary for each variable at the 5% significant level. The results from the statistical tests indicate that the price returns are not independent and identically distributed. This is supported by the rejection of the null hypothesis for the LM test. Moreover, for each variable price return series is stationary and white noise.

The descriptive statistics for phase 3 are presented in table 6 and for phase 4 in table 7, which can be found in Appendix C. For phase 3, the conclusion from the JB, LM, LB and ADF is the same as presented above. The descriptive statistics for phase 4 show equal conclusion for the JB, LM and ADF tests. However, the null hypothesis of the LB test is only rejected for iron & steel and cement.

Figure 2 shows the daily log returns from the 1^{st} of March 2017 to the 29^{th} of February 2024. In Figure 2, there are significant spikes for all variables. In March 2020 all variables experienced a significant negative return shock, except for natural gas. This shock is most likely an affect of the COVID-19 pandemic. It is also worth noticing that natural gas, EU ETS, aluminum, and fertilizers experienced noticeable fluctuations at the start of 2022, most likely related to the Russian invasion of Ukraine. More interestingly is the small impact this had on the log returns of cement. This can be explained by the variable mainly consisting of Colombian assets, and thus the events in Europe are not as relevant on its price. A significant negative log return is observed for the EU ETS, iron & steel, aluminum, and natural gas at the end of 2018/ beginning of 2019. During this period, the EU experienced a shock from the inside, as the deadline for Brexit approached and negotiations between the EU and the British about a deal created uncertainties for many markets.

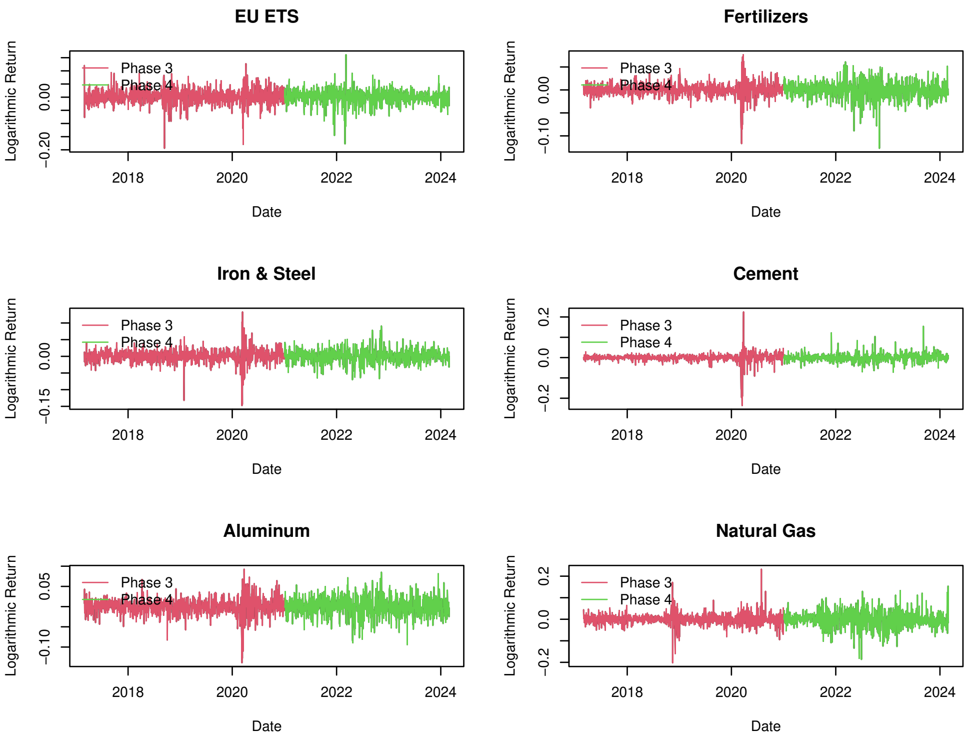


Figure 2: The log return of the variables from the 1^{st} of March 2017 to the 29^{th} of February 2024. Phase 3 colored red, phase 4 colored green.

Figure 3 shows the HV of the variables from the 30^{th} of March 2017 to the 29^{th} of February 2024. This is, the volatility over a time period. In this thesis a period of 22 days is used. Therefore, the figures begin on the 30^{th} of March 2017. The HV is presented in section 4.1. The figure shows that the HV of the price variables experience some significant spikes. For all variables, except natural gas, there is a large spike in the HV in March 2020. The reason behind these spikes is the COVID-19 pandemic that influenced every market in the world. Other than the COVID-19 pandemic, fluctuations are observed for the Russian invasion in 2022 for all variables. The EU ETS and natural gas have some significant spikes in the HV, especially energy commodities have intensified fluctuations the last four years due to special world events. Including the COVID-19 pandemic, natural gas was heavily impacted at the start of 2022 because of the Russian invasion of Ukraine. Before the invasion Europe was heavily dependent on import of Russian gas. After the invasion the EU decided to boycott Russia and the import of Russian gas was therefore cut significantly. An event like this will influence the EU ETS and natural gas significantly. The fluctuations seen on the EU ETS variable shows the importance of developing a model having the possibility to predict the HV.

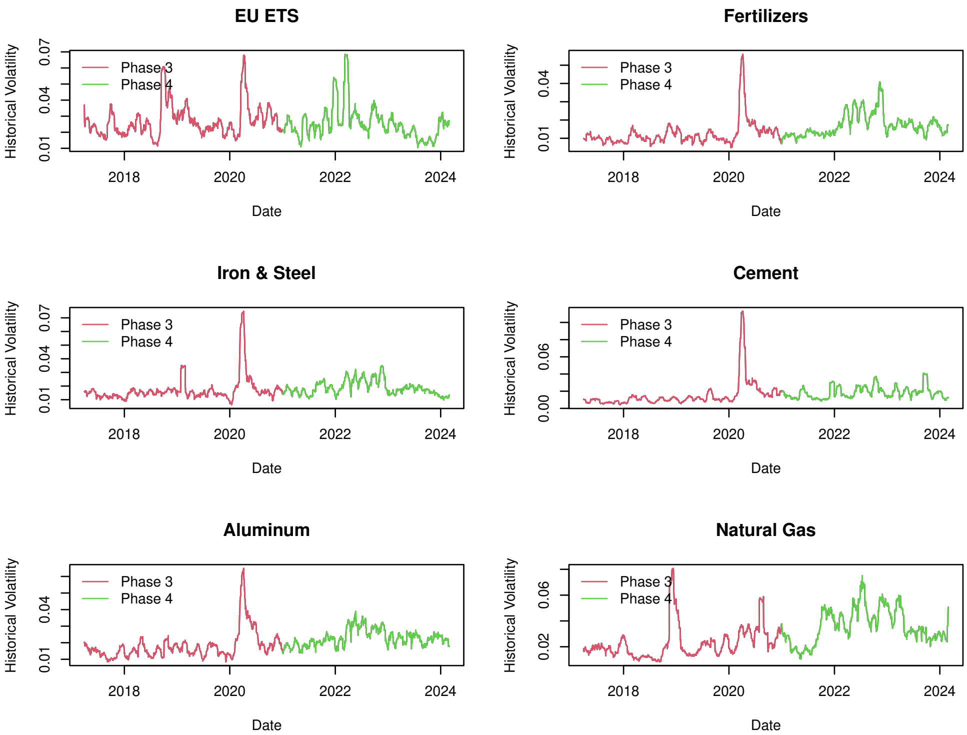


Figure 3: The HV of for all price variables from the 30^{th} of March 2017 to the 29^{th} of February 2024. Phase 3 colored in red and phase 4 colored in green

4 Methodology

In this section we give review of the major methodological components relevant for the thesis project.

4.1 Historical Volatility

The historical volatility is the simplest volatility model, it calculates the standard deviation of the log returns over some historical period (Brooks, 2002, p. 441). The historical volatility therefore is the square root of the variance of the log returns during the T past day(s) (Lahmiri, 2017). The historical volatility at time t is calculated using equation 2 (Ederington and Guan, 2006).

$$HV_t = \sqrt{\frac{1}{T}\sum_{i=0}^{T-1}(r_{t-i} - \overline{r})^2} \tag{2}$$

In equation 2, r_{t-i} is the logarithmic return on day t-i, $\overline{r}$ is the mean log return over the past T trading days. Due to weekends, T is determined to be 22 days resulting in $\overline{r}$ being the mean log return over the past 22 days.

4.2 GARCH

The thesis applies three different GARCH models to conduct the analysis, namely the GARCH (Bollerslev, 1986), the exponential GARCH (Nelson, 1991), and the Glosten-Jagannathan-Runkle-GARCH (Glosten et al., 1993). The goal with the GARCH models is to calculate the conditional variance at time t. By conditional it is meant that the variance is conditioned on some information at a set time (Greene, 2002, p. 661). The different GARCH(p,q) models are used with order (1,1). Previous studies on financial time series have shown the superiority of GARCH(1,1) models when compared with models using other GARCH(p,q) orders (Bollerslev, 1987; Miah and Rahman, 2016). The non-constant parameter for the GARCH is the mean model. The GARCH-type models are estimated with a different mean order, either a zero mean model, Autoregressive Moving Average (ARMA)(1,0) or an ARMA(1,1). Due to observing evidence of stationarity in each price return series in section 3, ARMA is used. For further improvements of the estimations, the error term is estimated with a normal distribution, student's t-distribution (std), or a generalized error distribution (ged). The equations for the different distributions are presented in the paper of Ghalanos (2023).

4.2.1 sGARCH

The GARCH model of Bollerslev (1986) (sGARCH) allows for the interpretation of the fitted variance at time t as a weight of the long term mean value and the fitted variance from a previous period (Brooks, 2019, p. 396). The advantage of the sGARCH model is that the variance can evolve over time in a more general way than the ARCH model (Greene, 2002, p. 662). Compared with the ARCH model, Brooks (2002, p. 453) argues that the sGARCH model will avoid overfitting, thus it will be less likely to breach constraints regarding non-negativity. In addition, the sGARCH(1,1) model has the advantage of being easy to use when forecasting the volatility over a longer period (Campbell et al., 1997, p. 483). The issue with the sGARCH model is that it assumes symmetry, meaning that negative or equally large positive shocks have the same effect on the conditional variance (Zivot, 2008, Brooks, 2002, p. 469). This is due to the conditional variance being calculated based on the lagged squared residuals, resulting in their sign becoming irrelevant (Brooks, 2019, p. 404). Another issue with sGARCH is a leverage issue, meaning that negative news will increase the volatility, while positive news tend to reduce the volatility (Rydberg, 2000, Brooks, 2002, p. 468).

In the explanatory paper describing the "RUGARCH" package, Ghalanos (2023) formulates equation 3 to calculate the conditional variance using the sGARCH model.

$$\sigma_t^2 = (\omega + \sum_{j=1}^{m} \zeta_j v_{jt}) + \sum_{j=1}^{q} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{3}$$

For the sGARCH, and the following eGARCH and GJR-GARCH model, ω is the constant term, m is the amount of external regressors, v_{jt} determines the value of the external regressor j at time t. ϵ_t describes the error term and σ^2 is the conditional variance. The parameters ζ, α and β are estimated by the model, using the maximum likelihood technique (Brooks, 2002, p. 457). Maximum likelihood estimation (MLE) is a search for a set of parameter values that have the highest likelihood of having produced the observed date (Brooks, 2019, p. 400). Amongst the properties of the MLE, the invariance property is especially useful. This is due to the analytical expression of MLE not existing (Martin et al., 2012, p. 71). More results from Martin et al. (2012, p. 71) show that for smaller sample sizes, some maximum likelihood estimators are unbiased.

4.2.2 eGARCH

To address the symmetry issue of sGARCH models, asymmetric GARCH models have been created. In these models various methods have been introduced to ensure disparate effect of equally large, differently signed shocks. One such model, the eGARCH, addresses symmetry by determining a size and a sign of the shock, defined as respectively, γ and α (Ghalanos, 2023). These will either have a negative or a positive sign depending on the input. The advantage of using logarithmic conditional variance renders the eGARCH an appealing model, because it ensures the resulting conditional variance to be positive (Campbell et al., 1997, p. 487). The function used to calculate the conditional variance in the eGARCH model, according to Ghalanos (2023), is presented in equation 4.

$$log_e(\sigma_t^2) = (\omega + \sum_{j=1}^{m} \zeta_j v_{jt}) + \sum_{j=1}^{q}(\alpha_j z_{t-j} + \gamma_j(|z_{t-j}| - E|z_{t-j}|)) + \sum_{j=1}^{p} \beta_j log_e(\sigma_{t-j}^2) \quad (4)$$

In addition to the parameters equal to the sGARCH model. The eGARCH introduces z_{t-j}, which represents a standardization of the residual determined as $\frac{\epsilon_t}{\sqrt{\sigma_t^2}}$, and γ as the weight of the shock. $E|z_{t-j}|$ represents the expected value of z_{t-j}. The parameters in the eGARCH model are estimated using the MLE.

4.2.3 GJR-GARCH

The GJR-GARCH uses an indicator function, I, to model the effects of a shock on the conditional variance. This indicator contributes to asymmetry by being 1 if $\epsilon \leq 0$, and 0 otherwise (Ghalanos, 2023, Brooks, 2002, p. 469-470). The indicator implements the leverage effect described by Rydberg (2000), resulting in a positive shock at previous time t - 1 having a smaller impact on the volatility at time t than an equally large negative shock.

Ghalanos (2023) calculates the conditional variance with the GJR-GARCH model using the formula in equation 5.

$$\sigma_t^2 = (\omega + \sum_{j=1}^{m} \zeta_j v_{jt}) + \sum_{j=1}^{q}(\alpha_j \epsilon_{t-j}^2 + \gamma_t I_{t-j} \epsilon_{t-j}^2) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \quad (5)$$

The GJR-GARCH is created such that if the residuals are larger than 0, the GJR-GARCH is equal to the sGARCH model, else it adds a weight, γ, to the residual. For the GJR-GARCH model, the parameters are estimated using the MLE.

4.3 Distribution of Residuals

Campbell et al. (1997) claim that in reality GARCH models will not show properties of a normal distribution. They argue that the standardized residuals of a GARCH model show signs of excess kurtosis (Campbell et al., 1997, p. 488). Others have shown that despite the literature often assuming normal distribution for financial data this is rarely the case in practical use. Financial data tends to be skewed, show excess kurtosis, and show heavy tails (Adubisi et al., 2022; Cerqueti et al., 2020; Chen and Spokoiny, 2015). Ampadu et al. (2024) argue that a GARCH model assuming a single distribution, often is not suitable to capture these characteristics. Cerqueti et al. (2020) argue that the normal distribution is not suitable, due to the characteristics presented above. As presented in table 1, this is indeed the case for the presented time series. To address these issues, the GARCH-type models are estimated using different expected distribution for the residuals, the normal distribution, std, and ged.

4.4 Information Criteria

The different GARCH-type models are estimated individually for each variable, employing lagged variables as external regressors. The decision of the best model for each variable is done with either Akaike information criterion (AIC) (Akaike, 1974) or Bayesian information criterion (BIC) (Schwarz, 1978), meaning the thesis assumes that a true model exists. The argument is that the best performing model is the model which produces the lowest respective information criteria. The calculation of the AIC and BIC is presented in equation 6 and 7, respectively.

$$AIC(k) = -2ln(likelihood) + 2k \tag{6}$$

$$BIC(k) = -2ln(likelihood) + kln(n) \tag{7}$$

In the equations above, k represent the number of parameters. These criteria do not necessarily present the same model as the best. This is a result of the AIC being developed as a tool to support the finding of the optimal predictor (Akaike, 1974), while the BIC is used to reveal the true model given the data (de-Graft Acquah, 2010). Further, this thesis analyzes models created based on results of either the AIC or the BIC.

4.5 Scaling

Machine learning algorithms can fail to learn from the training data (underfitting), or it can fail to generalize the training data (overfitting) (Rafatirad, 2022, p. 118). This is a result of the algorithm struggling to converge. Amirshahi and Lahmiri (2023) suggest to scale the data to overcome these issues. The scaling is performed by subtracting the mean HV for the variable j from the observed HV for variable j at time t and dividing the standard deviation of the HV of variable j. This is described in equation 8. It is worth mentioning that there exists a vast amount of scalings, and others are presented by Rafatirad (2022, p. 119).

$$Scaled_{HV_t} = \frac{HV_{t,j} - \overline{HV_j}}{\sigma_{HV_j}} \tag{8}$$

Equation 8 represents a standardization scaler, which aims to re-scale the data such that the mean of the data is 0 and its standard deviation is 1. The underlying assumption is that the distribution of the data resembles a normal distribution (Rafatirad, 2022, p. 119).

4.6 Long Short-Term Memory

Long Short-Term Memory (Hochreiter and Schmidhuber, 1997) is a machine learning algorithm created with the intention of addressing the vanishing gradient problem (Houdt et al., 2020). This is a problem occurring in recurrent neural networks where the gradient becomes too small for the weights to be updated (Hull, 2021, p. 264). An LSTM cell is composed of an input gate, output gate, and a forget gate (Géron, 2019, p. 516; Ding et al., 2022). Each of the gates have the possibility to add or delete information from the memory state, which is the model's "information storage" (Ding et al., 2022; Fu et al., 2021). Figure 4 presents a typical LSTM block. In the figure, the forget gate (f), input gate (i), and output gate (o) are displayed. In addition, the block output is presented by z. In the following section, these will be presented.

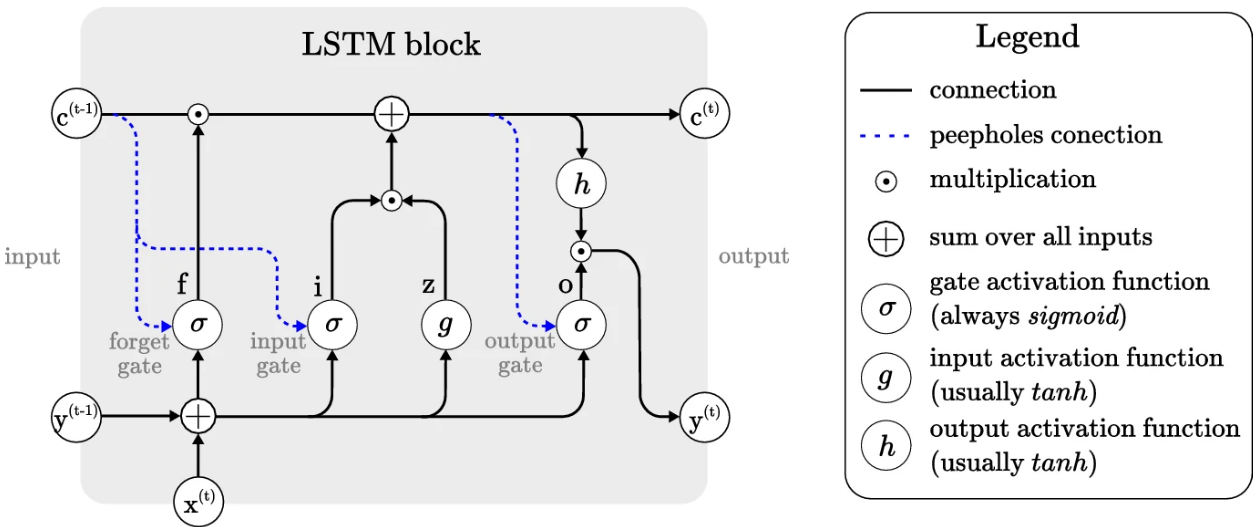


Figure 4: *"Architecture of a typical vanilla LSTM block"*, taken from Houdt et al., 2020

In the forget cell, the previous cell state is retained. It determines what information from the previous memory state to maintain at the current memory state (Ding et al., 2022; Fu et al., 2021). The activation of the forget gate depends on the current input, in addition to the output and state from the previous cell (Houdt et al., 2020). Determining the values of the activation of the forget gate is performed by equation 9, which originates from the paper of Houdt et al. (2020).

$$f^t = \sigma(W_f x^t + R_f y^{t-1} + p_f \odot c^{t-1} + b_f) \tag{9}$$

In equation 9, and for the other equations in section 4.6, W and R represent the weights of the input signal, x, and output, y. The value b represents a bias value and σ represents the sigmoidal activation function. The goal of the activation function is to decide if a signal should progress in the network. If it progresses, the activation function decides to what extent the signal should affect the outcome (Sen et al., 2021, p. 73). The value p represents the weight associated with the cell update c, and $\odot$ is a point-wise multiplicator for two vectors (Houdt et al., 2020).
As information passes between the blocks, the model analyzes a combination of the input with the output from the previous state to determine what information to pass to the cell state (Géron, 2019, p. 517). This procedure is done in the block output, z. Houdt et al. (2020) introduce equation 10 to describe how the update of the block output is performed.

$$z^t = g(W_z x^t + R_z y^{t-1} + +b_z) \tag{10}$$

In the block input, the sigmoid function is replaced with a tangent hyperbolic (tanh) function denoted as g.
The input gate, i, is the gate in the LSTM cell where the input is updated. In this gate the model decides on what information to keep and what information to forget from the current memory state (Ding et al., 2022; Fu et al., 2021; Houdt et al., 2020). Equation 11 presented by Houdt et al. (2020) shows how the input is updated.

$$i^t = \sigma(W_i x^t + R_i y^{t-1} + p_i \odot c^{t-1} + b_i) \tag{11}$$

The values from the input gate, forget gate and block input are used to determine the cell state for the current block. The updating of the current cell state is done by equation 12 (Houdt et al., 2020).

$$c^t = z^t \odot i^t + c^{t-1} \odot f^t \tag{12}$$

The output gate determines what information from the cell the model will feed the specific cells at this time step. The information from this step is fed both to the direct output, but also to the long-term state of the model (Géron, 2019, p. 517; Ding et al., 2022; Fu et al., 2021). The calculation performed in the output gate is presented in equation 13 (Houdt et al., 2020).

$$o^t = \sigma(W_o x^t + R_o y^{t-1} + p_o \odot c^t + b_o) \tag{13}$$

The output from the particular block dictates which information is passed through to the subsequent block and ultimately presented as the predicted value. The computation of the block's output follows equation 14 (Houdt et al., 2020).

$$y^t = g(c^t) \odot \sigma^t \tag{14}$$

4.7 Hybrid Models

The final part of the model creation is estimating the hybrid model. In section 2.5, the paper of Liu et al. (2024) was introduced. The main take away from their paper is the effect of the hybrid GARCH-LSTM model on the Chinese carbon credits market. Their results provide further motivation to investigate the possibility of establishing a GARCH-LSTM model for the EU ETS. The hybrid model introduced in this thesis is a combination of the conditional volatility estimated by the GARCH model chosen by AIC or BIC, with the observed historical volatility of the independent variables. The combination of the observed historical value and the GARCH estimate is then scaled following the formula in equation 8. The lagged scaled values are fed into the GARCH-LSTM model as the input, while the scaled HV value of EU ETS is defined as a dependent variable.

4.8 Forecast Evaluation

In order to analyze the optimal model, the employment of an evaluation method is necessary. The error at time t is determined as the difference between the actual and the predicted HV at the time t. Equation 15 presents how the calculation is performed.

$$\epsilon_t = HV_t - \widehat{HV}_t \tag{15}$$

ϵ_t represents the forecast error at time t, HV_t is the actual historical volatility, and the predicted is represented by $\widehat{HV}_t$.

To measure the performance of the presented models, two different calculations of errors are considered. The root mean squared error presented in equation 16 and the mean absolute error presented in equation 17. Both these are preferred measures as they ensure ϵ_t to be positive, either by squaring or by calculating the absolute value of the error. The rationale for using two performance measures was presented in section 2.6.

$$RMSE = \frac{1}{M}\sqrt{\sum_{i=t}^{M}\epsilon_t^2} \tag{16}$$

$$MAE = \frac{1}{M}\sum_{i=t}^{M}|\epsilon_t| \tag{17}$$

The errors are bootstrapped to provide scientific proof that the model with the lowest RMSE or MAE will outperform the alternative. Bootstrapping involves simulating the errors and determining how many of the simulated errors that outperform the result from a specific test statistic[5]. The null hypothesis of the test is that the analyzed errors are equal. The alternative hypothesis is that the model with the best score outperforms the other. A representation of the hypothesis for the RMSE test is seen below.

$$H_0 : \frac{RMSE_{mod2}}{RMSE_{mod1}} = 1, H_A : \frac{RMSE_{mod2}}{RMSE_{mod1}} > 1$$

The test statistic is calculated similarly for the MAE test, it is obtained by replacing RMSE with MAE in the hypothesis above. The test is performed by simulating the test statistic x amounts of times and calculating the p-value as the amount of simulated statistics that are larger than the test statistic from the original data. If the p-value is smaller than 0.05, the null hypothesis is rejected.

4.9 The Software

The data analysis in the thesis is done using R (R Core Team, 2022). In addition to standard R, different packages are used, which are presented in appendix A. The main packages used in the model creation are also presented below.

4.9.1 RUGARCH

Galanos (2023) created the RUGARCH-package as a tool to aid with the creation of GARCH models in R. The main functions this package provides which are used in this thesis are the *ugarchspec(),ugarchfit()* and *ugarchpath()*. These functions are used to respectively, create the model, fit the model, and perform simulations of future conditional variance.

[5]The function utilized to perform the bootstrap in R is copied from a course in machine learning (course code SE-507) attended at the university.

4.9.2 Keras

The Keras package (Chollet et al., 2015) is a powerful tool to create different machine learning models in R and other programming languages. The package has several different built-in functions to create machine learning models, especially relevant are the *layer_lstm(), keras_model_sequential*[6] and *callback_early_stopping()* functions that aid in creating the LSTM layer and avoiding the model to overfit. The combination of these two packages made it possible to create a model that could combine the results of different GARCH models with the LSTM model to analyze the hybrid models.

[6]Note that in R, a text box will appear asking if a Python environment should be open for the reticulate package. We recommend to press no as the code did not work on our computer when yes was pressed

5 Results

To perform the analysis, the data is split into a training set and a test set. The training set for phase 3 consists of data from the 1^{st} of March 2017 until the 6^{th} of November 2019. The test set in phase 3 consists of data from the the 7^{th} of November 2019 until the the 31^{st} of December 2020.

The training set for phase 4 consists of data from the time period from the 1^{st} of January 2021 to the 21^{st} of March 2023. The test data of phase 4 is the period from 22^{nd} of March 2023 to the 29^{th} of February 2024.

5.1 The Optimal GARCH Model

The GARCH-type models for each asset is predicted using the one day lag of the other commodities as external regressors. For each variable, a model is selected from the GARCH family of models. In addition, three different ARMA orders and three different assumptions for the residual term are used to optimize the model. These were presented in section 4.2. The "optgarch" function presented in appendix B prints the ARMA order, distribution of residual, and which GARCH(1,1)-type model that performs the best according to the AIC and the BIC. The choice of optimal GARCH model for the different variables during phase 3 is presented in table 2, and table 3 presents the choice during phase 4. In addition, the output displays the ARMA order and the distribution of ϵ. The choice of the best model is made either by the AIC or the BIC as explained in section 4.4.

Tables 2 and 3 show that the optimal model for the EU ETS is a sGARCH with an ARMA order (1,0) and with ϵ assumed under a student's t-distribution, regardless of the phase and the information criteria. The same is observed for returns on the iron & steel, with the only difference being that it uses an eGARCH model instead of the sGARCH. For the returns on fertilizers, the BIC's choice of GARCH-type changes from an eGARCH for phase 3 to a sGARCH for phase 4. However, both the ARMA order and the assumed distribution of the residuals remain constant. The AIC chosen GARCH-type model for the returns on fertilizers, shows that the assumed distribution of the residuals changes from ged for phase 3 to std for phase 4, while the eGARCH(1,1) is preferred during both phases. The model selected on the returns on cement is the only to choose GJR-GARCH as the optimal, and this is seen for both phases. For the returns on aluminum, the observation is that for AIC the same model is optimal for both phases, with an eGARCH(1,1) assuming a std and having an ARMA(1,1). Based on the BIC the optimal model for phase 3 is the std

Table 2: The optimal GARCH model, based on the AIC and the BIC, for each variable during phase 3

Variable	Information Criteron	ARMAorder	Residual Distribution	GARCH(1,1)-type Model
EU ETS	AIC	(1,0)	std	sGARCH
EU ETS	BIC	(1,0)	std	sGARCH
Fertilizers	AIC	(1,1)	ged	eGARCH
Fertilizers	BIC	(1,0)	std	eGARCH
Iron &Steel	AIC	(1,0)	std	eGARCH
Iron &Steel	BIC	(1,0)	std	eGARCH
Cement	AIC	(1,1)	ged	GJR-GARCH
Cement	BIC	(1,0)	std	GJR-GARCH
Aluminum	AIC	(1,0)	std	eGARCH
Aluminum	BIC	(1,0)	std	eGARCH
Natural Gas	AIC	(1,1)	ged	sGARCH
Natural Gas	BIC	(1,0)	std	sGARCH

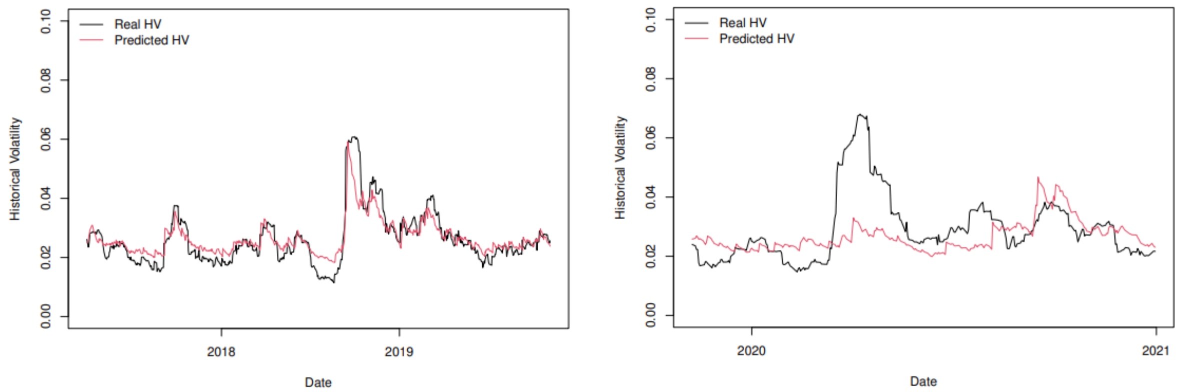
eGARCH with ARMA(1,0). This changes to an ARMA(0,0) sGARCH with a normal distribution for phase 4. The model created on the returns of natural gas has a different optimal GARCH-type model for both phases. This is observed for both the AIC and the BIC.
Another interesting observation is that there seems to be a relationship between the best ARMA order and distribution of residuals in both phases. The models with a std have an ARMA(1,0), the ones with a ged have an ARMA(1,1) and with the norm the ARMA(0,0).

The results from the best GARCH-type model are used to simulate the HV for the respective variables. The simulation is performed by giving the model the coefficients estimated on the training set and using these to simulate the conditional variance for the test period.

Table 3: The optimal GARCH model, based on the AIC and the BIC, for each variable during phase 4

Variable	Information Criteron	ARMAorder	Residual Distribution	GARCH(1,1)-type Model
EU ETS	AIC	(1,0)	std	sGARCH
EU ETS	BIC	(1,0)	std	sGARCH
Fertilizers	AIC	(1,0)	std	eGARCH
Fertilizers	BIC	(1,0)	std	sGARCH
Iron &Steel	AIC	(1,0)	std	eGARCH
Iron &Steel	BIC	(1,0)	std	eGARCH
Cement	AIC	(1,0)	std	GJR-GARCH
Cement	BIC	(1,0)	std	GJR-GARCH
Aluminum	AIC	(1,0)	std	eGARCH
Aluminum	BIC	(0,0)	norm	sGARCH
Natural Gas	AIC	(1,0)	std	eGARCH
Natural Gas	BIC	(0,0)	norm	sGARCH

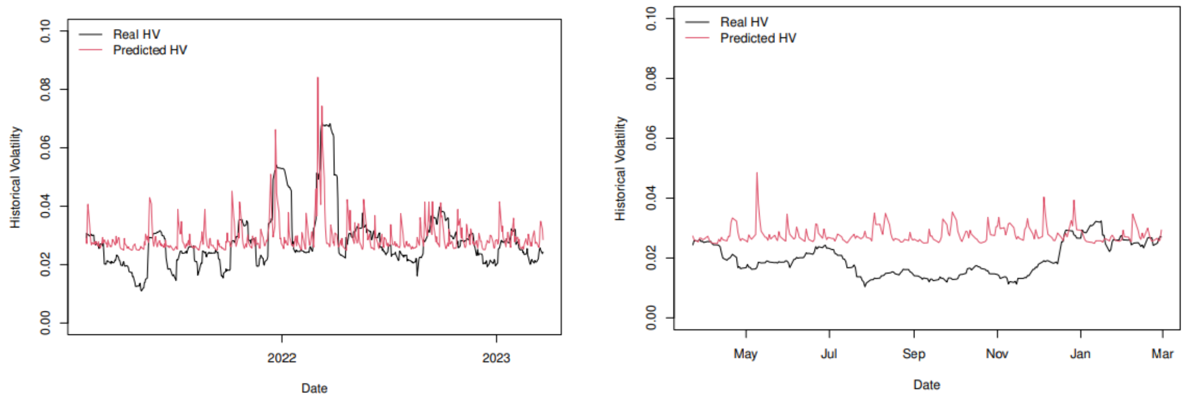
Figure 5 shows the observed and predicted HV of the EU ETS. The predicted HV is based on the result from the sGARCH(1,1) with ARMA(1,0) and residuals assumed to be student's t-distributed. The result on the basis of the training set is presented in figure 5a, and the test set in figure 5b. The predicted HV indicates that the model is generalizing on the training data, which displays that in phase 3 the CBAM variables can be useful to predict the HV of the EU ETS. The performance of the test set is showing that the predictability of the model does not mach the result obtained on the training set. The model is able to capture some trends on the test set, as it fits well with the HV at the end of 2019 and the beginning 2020, as well as at the end of 2020. However, the large increase in the HV in early 2020 is not predicted by the model.



(a) (b)

Figure 5: Plots of the real HV (in black) and the HV predicted by the sGARCH(1,1) model (in red) for the EU ETS on phase 3 (a) training data and (b) test data

Figure 6 shows the real and predicted HV for phase 4, based on the optimal, GARCH-type model presented in table 3. The results on the basis of the training data from figure 6a show that during phase 4, the GARCH model is predicting a more volatile HV than observed. In addition, it is not predicting the lower values. For the max values it predicts higher HV than observed. The same can be seen for the results on the basis of the test set in figure 6b, where almost all predictions are above the observed HV. The model seems to be set on a minimum value of 0.0249 (based on the min() function in R), as this is the lowest number it predicts. There are some positive observations to draw from figure 6. In most instances where the HV increases, the model predicts an increase. This conclusion is more difficult to draw from the model created on the test data. It is worth mentioning that the HV in the test data for phase 4 is showing less fluctuations compared to the previously analyzed period.



(a) (b)

Figure 6: Plots of the real HV (in black) and the HV predicted by the sGARCH(1,1) model (in red) for the EU ETS on phase 4 (a) training data and (b) test data

5.2 The LSTM Model

The creation of the LSTM model is done by creating a simple LSTM cell and adding to this to improve the model. The first block used 32 units, meaning that each LSTM block is composed of 32 cells. After creating one cell another was added using identical parameters. The model contains two dropout layers. These layers remove a random given amount of the data when the model is trained. Further, an early stopping mechanism is included in the model. This allows the model to stop before running through the entire learning process. Early stopping is activated if the performance metric is increasing for a given number of times in a row. According to Ying (2019), this will avoid the model from learning the noise. The model created in this thesis uses a patience level of 100, which is chosen based on the trial and error method. Patience decides how many observations of increasing error the model accepts. All LSTM models are trained using mean squared error (MSE) as performance measure, which is the square of the RMSE presented in section 4.8. This indicates that for every iteration, the model examines if the MSE is reduced.

For representation and later calculation of RMSE and MAE, a de-scaling of the predictions is performed, using equation 18.

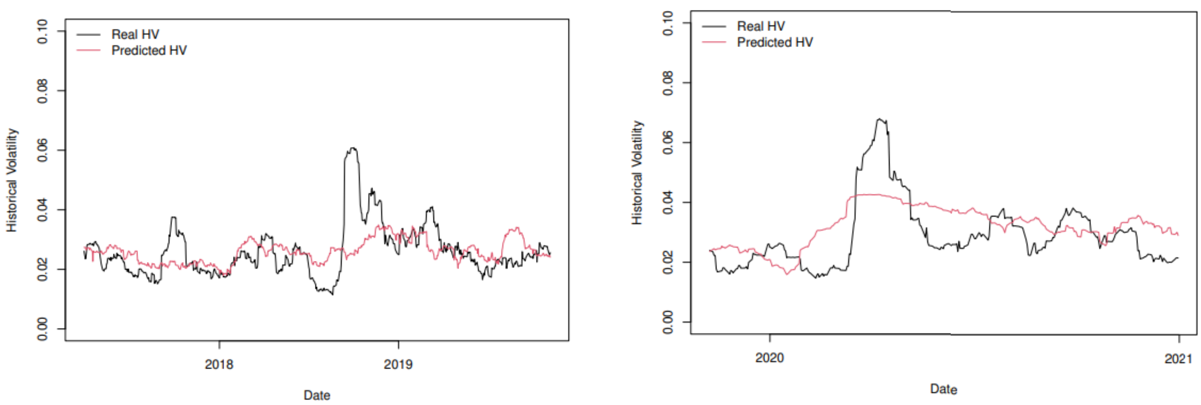
$$\widehat{HV}_{EUETS} = Scaled_{\widehat{HV}_{EUETS}} \times \sigma_{HV_{EUETS}} + \overline{HV}_{EUETS} \tag{18}$$

The de-scaling is performed only based on the HV of the EU ETS, due to this being the value the model is predicting.

The machine learning models are created based on optimizing hyperparameters. The hyperparameters were first tuned on one model and then used on the other models. During this process, many different parameters were investigated. The approach resembles a trial and error strategy. Different parameters, such as the unit size, the input size, the dropout size, the optimizer, the patience and the scaler were changed and tuned to improve the model. The result is the code found in appendix A.

EU ETS at time t constitutes the dependent variable in the LSTM model, while iron & steel, fertilizer, cement, aluminum, and natural gas at lag one function as the independent variables. In figures 7 and 8 the results of the LSTM model, for respectively phase 3 and 4, are displayed. In both figures, panel (a) refers to the training set while panel (b) relates to the test set.

Figure 7a shows evidence that the LSTM model is struggling to learn based on the training set for phase 3. The model predicts less fluctuations in the HV than observed. Diving into figure 5a and 7a, it is observed that the LSTM model is not improving the predictions achieved by the optimal GARCH on the training set during phase 3. This underperformance is further observed in tables 4a and 5a.
From figure 7b, the results on the basis of the test set show that the poor HV predictions on the training set are carried over to the test set. A positive aspect that is worth mentioning is that the model predicts a rise in the HV in early 2020. In addition, there are no extreme predictions. Further, there is a period in the end of 2020 where the predicted HV is close to the observed HV.

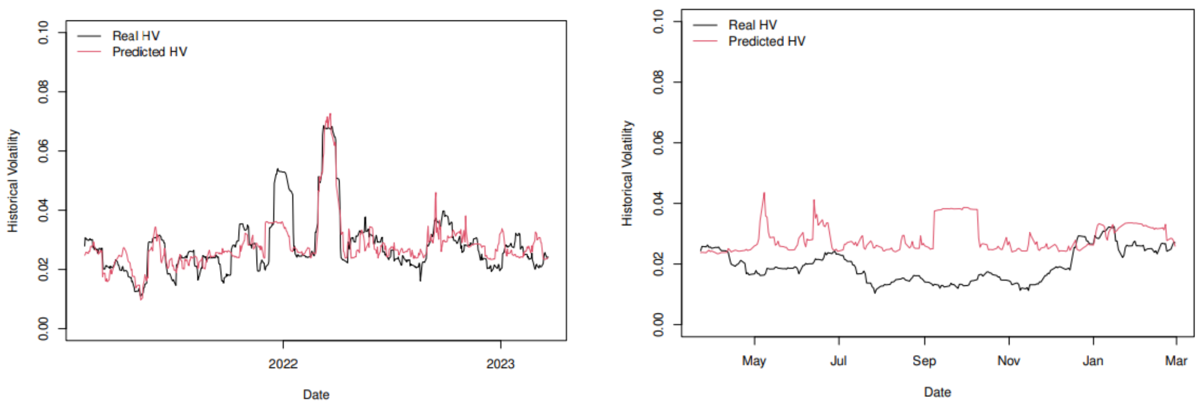


(a) (b)

Figure 7: Plots of the real HV (in black) and the HV predicted by the LSTM mode (in red) for the EU ETS on phase 3 (a) training data and (b) test data

Figure 8a indicates that the LSTM model shows good adaptation on the training set during phase 4, especially when compared with the predictions of the phase 3 model. It predicts especially well the high HV seen in the early 2022. However, poorer predictions are observed on the HV around the year change 2021/2022. After mid-2021, observations in figure 8a indicate that the model is struggling to predict the lower values of the HV. The model's capability of predicting the result on the training set is also observed in tables 4a and 5a. These show that the RMSE and MAE is lower in phase 4, at the same time as the LSTM model outperforms the GARCH-type model.

The results on basis of the test set can be observed in figure 8b. It is seen that the model predicts a spike in the HV around October of 2023, which is the time that the transition period for the CBAM began. However, the large increase in predicted HV is not equal to the actual observations of the HV at this point. Observations in early May and June 2023 indicate that the model predicts some non-observed changes in the HV. The result of these mispredictions is also observed in table 4b where the RMSE of the LSTM model is greater than for the GARCH models. Interestingly, phase 3 and phase 4 show opposite results in terms of the performance. The LSTM model for phase 3 produces higher RMSE and MAE on the result of the training set compared with the GARCH model, and it produces lower RMSE and MAE on the result of the training set for phase 4.



(a) (b)

Figure 8: Plots of the real HV (in black) and the HV predicted by the LSTM model (in red) for the EU ETS on phase 4 (a) training data and (b) test data

Table 4a shows that on the training set, the optimal GARCH model based on both the AIC and the BIC performs better during phase 3 than it does in phase 4. This substantiates the findings in the figures 5a and 6a. This observation is seen on all created models, except the RMSE for the LSTM model.

This observation is also seen on the MAE for the training set in table 5a. Interestingly, this dynamic changes on the test set. From table 4b the LSTM model outperforms the optimal GARCH models in phase 3 for the RMSE. The indication is that despite the learning not being very good on the training set, the predictions of the model are better than for a GARCH model.

Table 4: Observed RMSE on the optimal models

	std-GARCH-AIC	std-GARCH-BIC	LSTM	GARCH-LSTM-AIC	GARCH-LSTM-BIC
Phase 3	0.0042727	0.0042727	0.0085648	0.0028579	0.0026805
Phase 4	0.0094202	0.0094202	0.0057781	0.0040340	0.0039497

(a) Training set RMSE

	std-GARCH-AIC	std-GARCH-BIC	LSTM	GARCH-LSTM-AIC	GARCH-LSTM-BIC
Phase 3	0.0145292	0.0111932	0.0096802	0.0083935	0.0086072
Phase 4	0.0113918	0.0106981	0.0114045	0.0066792	0.0067196

(b) Test set RMSE

Table 5: Observed MAE on the optimal models

	std-GARCH-AIC	std-GARCH-BIC	LSTM	GARCH-LSTM-AIC	GARCH-LSTM-BIC
Phase 3	0.0030935	0.0030935	0.0057632	0.0021615	0.0020281
Phase 4	0.0065409	0.0065409	0.0041949	0.0031483	0.0030800

(a) Training set MAE

	std-GARCH-AIC	std-GARCH-BIC	LSTM	GARCH-LSTM-AIC	GARCH-LSTM-BIC
Phase 3	0.0116590	0.0076144	0.0080518	0.0067246	0.0068996
Phase 4	0.0098134	0.0090074	0.0093926	0.0055683	0.0055260

(b) Test set MAE

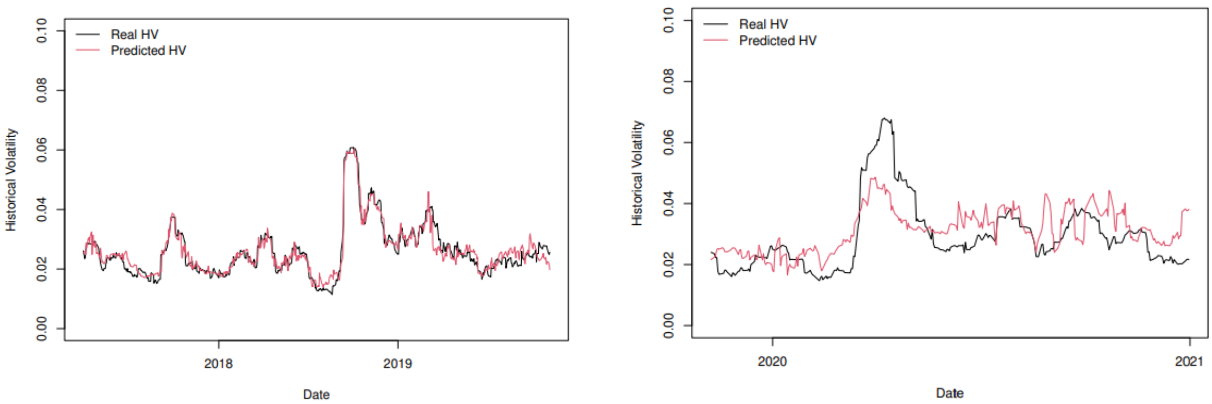
5.3 The GARCH-LSTM Model Chosen Based on BIC

The input data in the GARCH-LSTM model is created by combining the lagged HV for all variables except the EU ETS with the lagged predictions of the optimal GARCH-type models for all variables. The optimal GARCH-type model for each variable was presented in tables 2 and 3. The HV of the EU ETS at time t is therefore predicted using a combination of the HV for the other variables and the predicted HV from the GARCH-LSTM model at time t-1. This dataset is then scaled using equation 8. The same is done for the HV of the EU ETS.

Figure 9 shows the predicted and the real HV on the training set (figure 9a) and the test set (figure 9b). There is clear difference between using only the GARCH model or only the LSTM model, compared to using the hybrid model. For phase 3, the GARCH-LSTM model predicts well on the training set, and often it predicts close to the exact HV. Compared with the predicted HV in figures 5a and 7a, this difference is easily observed. This observation is also seen in both tables 4a and 5a where the RMSE and MAE are lower for the BIC chosen GARCH-LSTM model, than it is for both the GARCH model and the LSTM model.

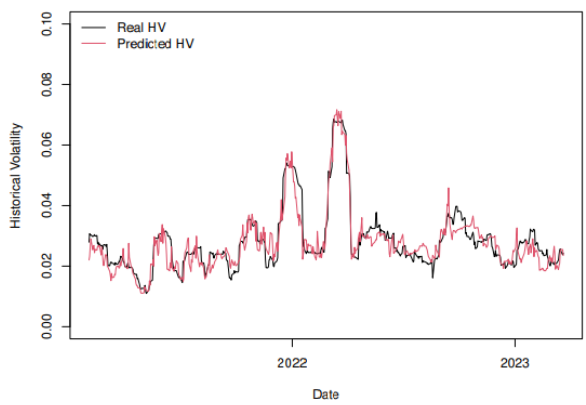
The result on the test set observed in figure 9b shows evidence of the model improving. In addition to better predicting the high increase in HV in early 2020, the other predictions are seemingly better when comparing figures 5b and 7b with figure 9b. The result from table 4b supports that the GARCH-LSTM model does present better prediction on the test set in phase 3. This result additionally holds when looking at the MAE in table 5b.

Analyzing phase 4, the results on the basis of the training set in figure 10a show that the BIC chosen GARCH-LSTM model is learning on the training set. Further, the model shows strong evidence of being able to capture both the higher and the lower values of the HV. This differs strongly from the predicted HV presented for the sGARCH model for phase 4 in section 5.1. Evidence from tables 4a and 5a also show that the GARCH-LSTM model for phase 4 is superior to the sGARCH model or the LSTM model on the training set. From figure 10b the prediction on the test set of the GARCH-LSTM model chosen based on BIC is observed. The model is predicting well until October 2023 where it suffers from the same spike as presented in section 5.2. Despite this spike, tables 4b and 5b show that based on RMSE and MAE, the BIC chosen GARCH-LSTM model is superior on the test data compared with the non-hybrid models.

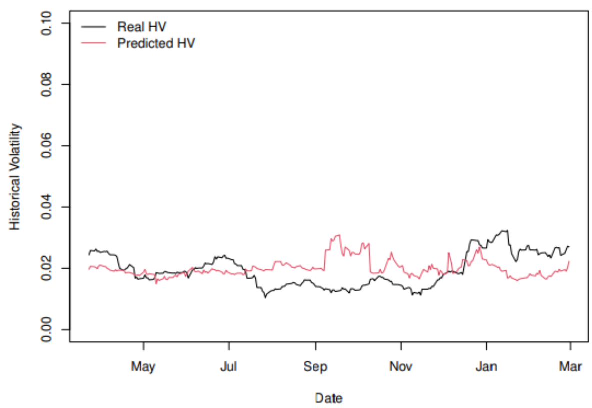


(a) (b)

Figure 9: Plots of the real HV (in black) and the HV predicted due to the BIC chosen GARCH-LSTM model (in red) for the EU ETS on phase 3. (a) training data and (b) test data



(a)



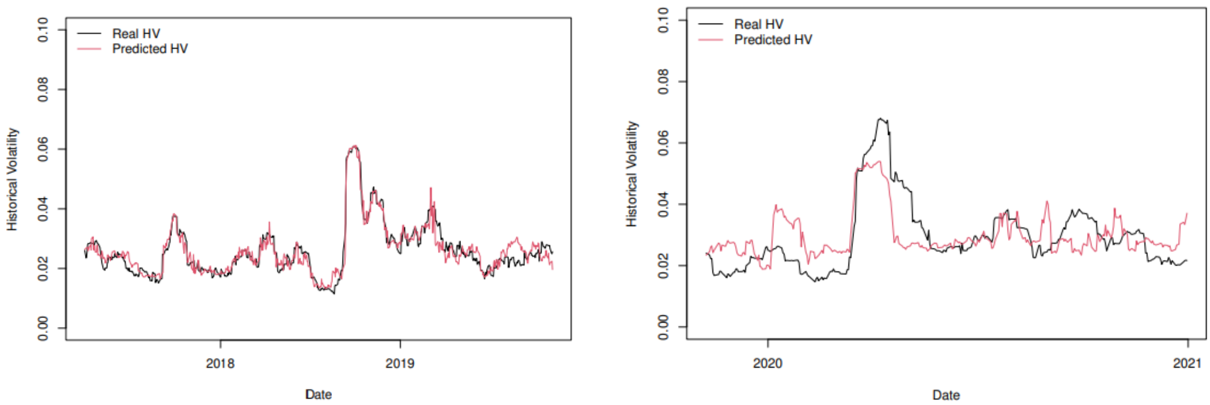
(b)

Figure 10: Plots of the real HV (in black) and the HV predicted due to the BIC chosen GARCH-LSTM model (in red) for the EU ETS on phase 4 (a) training data and (b) test data

5.4 The GARCH-LSTM Model Chosen Based on AIC

Reviewing the results for the AIC chosen GARCH-LSTM model, it is observed that the predictions are good in both phase 3 and 4. This is seen in figures 11a and 12a. In addition, tables 4a and 5a show that the RMSE and MAE is greater for the AIC chosen hybrid models than it is for the optimal AIC chosen GARCH model and the LSTM model during phase 3 and 4.

As expected, the AIC chosen GARCH-LSTM model seems to be superior to the BIC chosen when analyzing the results on basis of the test set both in phase 3 and phase 4. Figures 11b and 12b show that the AIC chosen model does not predict the same increase in the HV in October 2023 that the BIC chosen does. This is observed in tables 4b and 5b to be true for all, except for the MAE in phase 4. This will be further discussed in section 6.



(a) (b)

Figure 11: Plots of the real HV (in black) and the HV predicted due to the AIC based GARCH-LSTM model (in red) for the EU ETS on phase 3 (a) training data , (b) test data

Analyzing figure 12 it is observed that the model learn well on the training set. The results on the test set shows that the model predicts almost no changes in the HV over the period from April 2023 to the 29^{th} of February 2024. This result is similar to the actual observation. However, the model needs to be analyzed with regard to tables 4 and 5 who shows that based on RMSE this is not the optimal model and based, while based on MAE it is. This will be further analyzed in section 6.

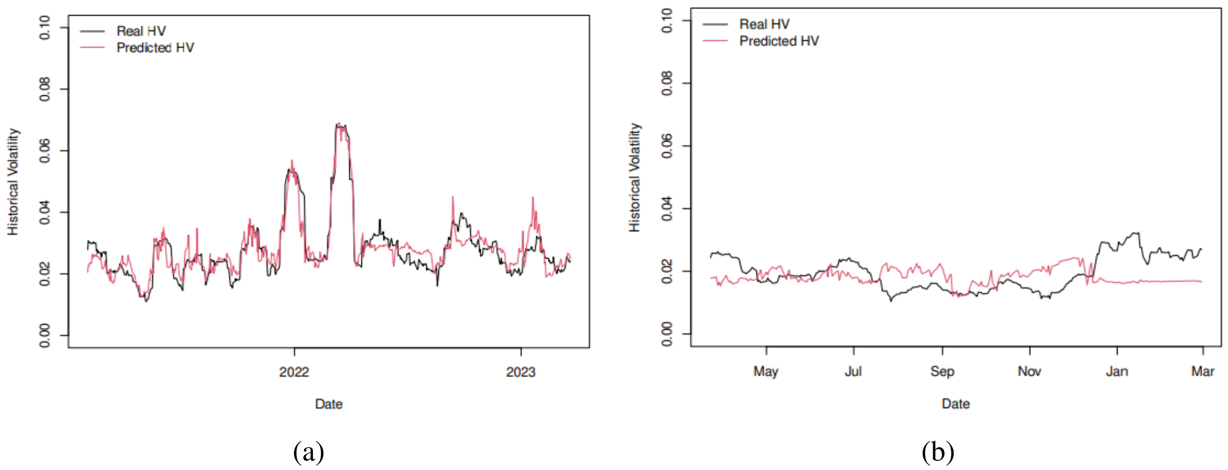


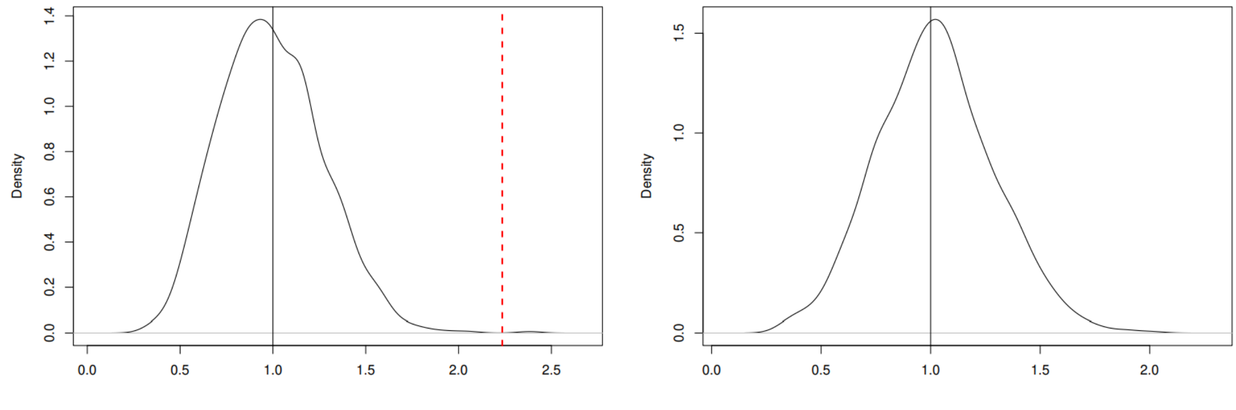
Figure 12: Plots of the real HV (in black) and the HV predicted due to the AIC based GARCH-LSTM model (in red) for the EU ETS on phase 4 (a) training data and (b) test data

5.5 Residual Bootstrap

To address whether the observations in tables 4 and 5 have statistical significance, the observed residuals are bootstrapped. This is performed for the results based on the AIC and the BIC. The null hypothesis as well as the alternative hypothesis were presented in section 4.8.

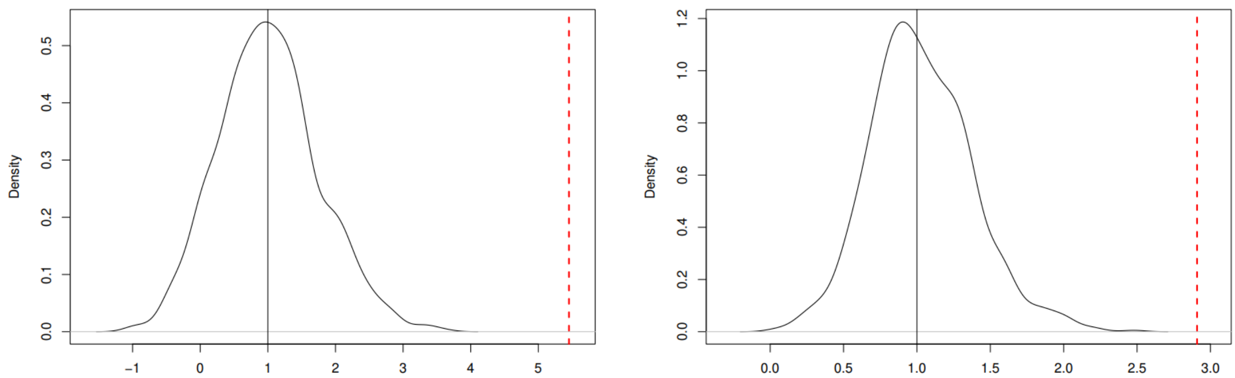
Figures 13 and 14 show the density function and the test statistic of the null hypothesis during phase 3 and 4. The figures show the test conducted where the null hypothesis is that the RMSE of the AIC chosen GARCH-LSTM is equal to the RMSE of the optimal AIC chosen GARCH model. From figure 13 it is clear that the null hypothesis can be rejected for both the training and the test set. This is because the p-value is smaller than 0.05, and the test statistic is seen to be higher than almost all simulated test statistics for the training set. For the test set, the test statistic is higher than all simulated test statistics. This indicates that during phase 3 there exists statistical evidence that the AIC chosen GARCH-LSTM model outperforms the sGARCH model based on RMSE. The result of the bootstrap of the MAE test is seen in appendix C.

The results in figure 14 show clear evidence that the null hypothesis can be rejected for phase 4. The p-value is 0.00, this is seen because the test statistic is far above all simulated test statistics. This result states that during phase 4 there exists statistical evidence that at a 95% confidence level the observation is that the RMSE of the GARCH-LSTM model is lower than that for the GARCH model, when the GARCH models are chosen based on the AIC.



(a) (b)

Figure 13: Plots of the estimated test statistic (in black) and the test statistic (in red) for the bootstrapped RMSE on EU ETS on phase 3 (a) training data and (b) test data. For $H_0 : \frac{RMSE_{Hybrid}}{RMSE_{GARCH}} = 1$



(a) (b)

Figure 14: Plots of the estimated test statistic (in black) and the test statistic (in red) for the bootstrapped RMSE on EU ETS on phase 4 (a) training data and (b) test data. For $H_0 : \frac{RMSE_{Hybrid}}{RMSE_{GARCH}} = 1$

6 Discussion

Since the CBAM currently is in a transitional phase, the EU needs information about the effects of its practical application. This thesis has studied how the historical volatility of the EU ETS are affected by the CBAM during the end of phase 3 and the beginning of phase 4. The results provided in section 5 are valuable for the EU to determine how the CBAM is currently working and if adjustments are needed.

6.1 Relevance of Results for the EU

This thesis can provide the EU with data regarding the market response of implementing the mechanism. In our thesis we decided to work with the prices of natural gas and the EU ETS allowances, as well as price indices on fertilizers, iron & steel, cement, and aluminum as our variables. These were presented in section 1.2 as the sectors required to comply with the CBAM in the transitional phase. By having knowledge about how different markets react, adapt, and evolve during the implementation of the CBAM, the EU can proactively adjust their policies to secure an ideal price for the CBAM certificates, thus avoiding carbon leakage. The hybrid models give the EU a tool to analyze how global financial markets interact with the EU climate policies. The EU can utilize this to ensure that carbon leakage is avoided. Takeda and Arimura (2024) found that the CBAM reduced carbon leakage, in addition to having a small positive impact on the GDP and welfare in Japan. Furthermore, Sun et al. (2023) support the findings that the CBAM reduces carbon leakage, but argues that it also overburdens developing countries. The empirical results from these papers show that the CBAM can influence countries and markets outside of Europe. Despite Sun et al. (2023) further suggesting the inefficiency of the CBAM, the carbon price has been significantly reduced since the publication of their paper. Following from their argument that the higher the price of carbon allowances, the more inefficient the CBAM. The price of carbon allowances at the time of writing this thesis suggests a more efficient CBAM.

Based on figures 10 and 12, in addition to tables 4 and 5, it can be argued that the hybrid model's prediction improves on the test set in phase 4, when comparing with phase 3. As established in section 1.3, the CBAM is fully dependent on the EU ETS, and thus the EU can continue utilizing the model to determine the historical volatility based on indices of the CBAM affected sectors. Liu et al. (2024) showed how the GARCH-LSTM model is superior when predicting price fluctuations on the Chinese carbon market. The model created in this thesis demonstrates that GARCH-LSTM model is indeed superior in predicting the historical volatility on the European carbon market also.

6.2 Relevance of Results for Investors

As mentioned in section 1.3, understanding and estimating the prices and the HV for the CBAM are of great interest for firms as well as investors. The CBAM introduces investment opportunities for firms and investors. The regulations coming with the CBAM require firms to change their productions from carbon-intensive products to low-carbon technology, which can for example consist of renewable energy or carbon capture and storage. Furthermore, the CBAM can result in changes in the market demand for carbon credits. Firms engaged in the production of carbon-intensive commodities and subjects of the CBAM obligations could opt to purchase carbon credits. The demand for the carbon credits would therefore increase and result in investment opportunities. Firms that operate in sectors affected by the CBAM would most likely experience increased costs due to for example exports to countries with carbon pricing or taxes (Cho et al., 2024). Consequently, the demand for the products in these markets can be reduced, which can impact a firm's profitability and influence an investors investment decision (Cho et al., 2024). The opportunities for investing into the CBAM are high, but with any investments there exists an underlying risk. Our hybrid models' usefulness for investors stems from the intentions of the EU. As elaborated by Dewaelheyns et al. (2023) and Oestreich and Tsiakas (2015), the free allowances play an important role in determining the value of a firm due to the extra cash flow they provide. However, a goal with the CBAM is phasing out all free allowances by 2035 (Benson et al., 2023; Cho et al., 2024). During the phase-out period, some sectors may observe a clear advantage due to how the free allowances are distributed. Based on European Commission (n.d.-a), certain sectors will receive free allowances accounting for 100% of their benchmark production of a product. Thus, if the findings of Dewaelheyns et al. (2023) and Oestreich and Tsiakas (2015) hold during phase 4, these firms will benefit from a carbon premium. Our model will aid investors in predicting the HV of the EU ETS based on the CBAM sectors. As mentioned in section 1.2, the sectors are chosen based on the risk of carbon leakage. This argues for the probability of firms operating within these sectors obtaining larger portions of the free allowances. This again results in these being able to generate higher cash flows, as argued by Dewaelheyns et al. (2023) and Oestreich and Tsiakas (2015). Investors can, at their own risk, use the proposed models to determine how the historical volatility of the EU ETS evolves over time.

6.3 Importance of The Distribution of Residuals

The choice of optimal GARCH-type model rises special interest in the assumption of the distribution of the residuals. Especially for phase 4, it is assumed a student's t-distribution in all, but two occasions. For phase 3, it is observed in all, but three occasions. During phase 3, the non-student's t-distribution predictions are only on models based on the AIC, while in phase 4 it is only for models based on the BIC. Both Abdullah et al. (2017) and Rahman et al. (2023) have shown that for their variables the student's t-distribution is preferred. Rahman et al. (2023) also showed that the student's t-distribution provides the lowest AIC and BIC score for sGARCH, eGARCH and GJR-GARCH. Amirshahi and Lahmiri (2023) observed that on average the optimal distribution for the residuals for cryptocurrencies is a generalized error distribution, especially for an asymmetric power GARCH (APGARCH). Further they showed that a student's t-distribution for the APGARCH did not improve the results observed for eGARCh or sGARCH when the residuals were assumed to be generalized error distributed. However, an interesting observation is that for each of the cryptocurrencies, more often than not, the GARCH-type model with a residual distribution assumed as student's t-distribution provides lower RMSE than others on the testing set. The disparity in their observations stems from the generalized error distribution when optimizing is based on RMSE, whereas the selection for each variable is based on choosing according to the AIC. This gives rise to further research, as detailed in section 6.7

6.4 Interpreting The Results

The optimal GARCH model for both phases is observed to be a sGARCH(1,1) under the assumption that the residuals are student's t-distributed. The findings from section 5 show that the models perform differently. Interestingly, the GARCH model on phase 3 shows better adoption to the result on the training set compared with phase 4. This dynamic change on the result of test set for the sGARCH(1,1) model chosen by the AIC, as seen in tables 4 and 5, where both the RMSE and MAE in phase 4 is lower. On the BIC based sGARCH(1,1) this is not observed for the MAE.

When analyzing the two phases there is clear evidence that the hybrid GARCH-LSTM model outperforms the sGARCH(1,1) based on the AIC and the BIC. In addition, they outperform the LSTM model on all occasions. The findings are bootstrapped, and the result shows that there exists statistical evidence that the GARCH-LSTM model is a better option than the GARCH counterpart. This is due to the high test statistic in addition to a p-value being lower than 0.05. On the training set the GARCH-LSTM model is observed to be better for predicting the historical volatility during phase 3. This shows that the model is able to learn on data it has already seen. This is valuable in establishing that the CBAM variables will have an impact on the determination of the HV on the EU ETS allowances. However, as stated by Dietterich (1995), the goal of the model is to perform well on the test set, not the training set. On the test set the RMSE and MAE are lower in phase 4 for all GARCH-LSTM models. This indicates that the model's prediction ability improves as the CBAM launch approaches. In addition, this raises questions about the potential for predictions to improve as the market is exposed to an increasingly amount of information. Furthermore, it is observed that for the models chosen based on either the AIC or the BIC, the GARCH-LSTM outperforms all other models based on either RMSE or MAE. Similar findings have been observed by other researchers who investigated the usefulness of the hybrid GARCH-LSTM model. In Y. Huang et al. (2021), an analysis of phase 3 from 2017 to 2019 showed an advantage in creating a hybrid model for price forecasting of the EU ETS allowances. The findings of Amirshahi and Lahmiri (2023) and Liu et al. (2024) suggest that the implementation of a hybrid GARCH-LSTM model has an advantage compared to GARCH-type model or LSTM models.

From the results on the test set in section 5.3 and 5.4, it is observed that the model chosen by the AIC performs better based on the RMSE for both phases and based on MAE it outperforms in phase 3. The model chosen by the BIC performs best based on the MAE during phase 4. The decision of which that is optimal depends on the model creation process. Intuitively, the most sensible choice is to choose the model with the lowest RMSE, due to the LSTM models being trained to minimize the MSE. However, this is not the case for the GARCH models. As pointed out in section 4.2, GARCH parameters are estimated via the maximum likelihood approach. An interpretation of the difference is that it stems from the predicted spike on the BIC based model. However, figure 12b shows that the AIC based model struggles more in predicting the HV in January and February of 2024. In section 2.6, the paper by Willmott and Matsuura (2005) and Willmott et al. (2009) were presented. They agree that the MAE is a superior performance measure, due to the results only having one interpretation. These findings were argued against by Chai and Draxler (2014). They showed that the RMSE is ambiguous and further proved it to outperform MAE if the distribution of residuals was expected to be Gaussian. Chai and Draxler (2014) concluded that one metric alone is not useful to determine the optimal model.

6.5 Overfitting

Machine learning techniques suffer from the vast amount of hyperparameters that can be optimized, as well as the risk of over- or underfitting. As mentioned in section 2.4.1, these issues can lead to suboptimal prediction of future values. Observations from figures 9a, 10a, 11a and 12a show results that resembles what could be observed in case of overfitting. This means that the prediction is matching the observation to perfect (Dietterich, 1995). Meanwhile, there is evidence against this, since the model is not predicting perfectly for all values, and perhaps in particular for phase 4 for both the AIC and BIC based GARCH-LSTM models. In these cases, the estimated HV is better at predicting trends, and less precise for predicting the actual HV. This is also described in section 5 whereas mentioned, the GARCH-LSTM models obtain a lower RMSE and MAE during phase 3 on the training set. This can also explain why the model is predicting much better on the training set for phase 4 compared with phase 3. This is also presented as a result of overfitting by Géron (2019, p. 28) and Ying (2019), who argued that when the model overfits, the good result obtained on the training data will not be matched on the test data. As presented in section 5.2, early stopping was a measure implemented to prevent the model from overfitting. Early stopping allows the model to stop before running through the entire learning process, if the performance metric is increasing a given number of times in a row. Ying (2019) argues that the model will avoid learning from the noise thus reducing the risk of overfitting.

6.6 The Limitations of The Findings

There are some factors that need to be addressed in regard of the results produced in this thesis. The hydrogen data is not represented in the analysis, due to not finding a source of data that we could access. However, there exist indices, such as the "Solactive World Hydrogen Index", that could be applicable if data were accessible. In addition, as presented in section 1, pricing of the CBAM is based on a weekly average price of the EU ETS. Ideally the data in this thesis should have been a weekly average to give a more precise overview over the effects on the CBAM price. The main reason behind this not being implemented is that the data set would become too small.
Footnote[3] addresses what might be a possible data flaw in this thesis. If the EEA and/or EFTA members are not to pay the CBAM fees, the fertilizers and aluminum indices need to be changed. This is due to them consisting of the Norwegian companies YARA (fertilizers) and Hydro (aluminum). This possible issue can be easily fixed, as there exists vast numbers of different indices to use. The choice in this thesis is based on having the same source of data for all indices.

6.7 Further Research

As mentioned in section 6.3, Amirshahi and Lahmiri (2023) suggest that it is possible to reduce the RMSE of the GARCH model, by optimizing the GARCH models by applying selection criteria other than the AIC or BIC. This will be a possible source of further research. It can be used to analyze whether it is possible to improve the GARCH-LSTM models when the GARCH model is optimized on the model that reduces the RMSE, MAE or another performance measure.

Further research on the EU ETS would also be to include more recent data and monitor what happens the closer we get to full implementation of the CBAM. Further, it can be analyzed whether changing some of the variables, excluding or implementing other variables will affect the EU ETS allowances. An example is to use a combination of the suggested models in this thesis with the variables proposed by B.-Q. Lin and Zhao (2023) to perform an analysis. To build on this, there exists world trade data that will give an indication to what countries the EU are trading with regard to the different variables[7]. Based on this, if data is found, it can be possible to optimize the choice of variables according to the main EU trade partners.

Further research can investigate how the carbon premium evolves with the full implementation of the CBAM and whether the observations of Dewaelheyns et al. (2023) and Oestreich and Tsiakas (2015) still holds. This can be used in accordance with our methodology to develop a model for return forecasts on the EU ETS, and analyze the connection between the CBAM and the carbon premium.

[7]https://wits.worldbank.org/Default.aspx?lang=en

7 Conclusion

This thesis studied how the CBAM impacts the EU ETS. The study utilized daily data on the prices of the EU ETS, cement, aluminum, natural gas, iron & steel, and fertilizer from the 28^{th} of February 2017 to the 29^{th} of February 2024. The data was turned into log returns, as presented in equation 1. Motivated by Amirshahi and Lahmiri (2023), we implemented hybrid models consisting of GARCH-type models and an LSTM model to understand how the historical volatility of carbon credits would be affected by the implementation of the CBAM. Specifically, we created a GARCH, an eGARCH and a GJR-GARCH model based on different assumptions for the distribution of the residuals. The thesis showed that for the EU ETS, sGARCH(1,1) with student's t-distribution as the assumed distribution of the residuals was the best model in terms of both the AIC and BIC during both phases. This model was compared with an LSTM and GARCH-LSTM model in terms of the RMSE and MAE. The results proved that the GARCH-LSTM model outperformed the other models in terms of both measures. We further discussed the results and argued that despite the result based on the training set being better in phase 3, the models' predictions were better in phase 4. This was in line with the literature presented by Dietterich (1995), who argued that the goal of a machine learning model is to perform good on the test set (Dietterich, 1995).

This thesis has contributed to the literature by showing that a GARCH-LSTM model can be useful to analyze the historical volatility of the EU ETS. This was consistent with the findings of previous literature. The baseline paper of Amirshahi and Lahmiri (2023) and the study of Liu et al. (2024) are consistent with the findings in this thesis. The GARCH-LSTM models were improving results and predictability in terms of RMSE and MAE when compared with a GARCH or an LSTM model. Moreover, the thesis provides the EU with a model that they can utilize academically to determine how changes in the CBAM relevant indices and natural gas affect the EU ETS price. This can hopefully assist the EU in the transitional phase and can be used as a tool to determine the correct implementation of CBAM in 2026.

References

Abdullah, S. M., Siddiqua, S., Siddiquee, M. S. H., & Hossain, N. (2017). Modeling and forecasting exchange rate volatility in Bangladesh using GARCH models: a comparison based on normal and Student's t-error distribution. *Financ Innov*, *18*(3). https://doi.org/10.1186/s40854-017-0071-z

Adubisi, O., Abdulkadir, A., Farouk, U., & Chiroma, H. (2022). The exponentiated half logistic skew-t distribution with GARCH-type volatility models. *Scientific African*, *16*, e01253. https://doi.org/https://doi.org/10.1016/j.sciaf.2022.e01253

Agwu, P. C., Emodi, N. V., & Okoye, U. O. (2021). Chapter 3 - Climate Change and Effects: A Qualitative Experience of Selected Older Adults. In T. Chaiechi (Ed.), *Economic effects of natural disasters* (pp. 21–34). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-817465-4.00003-0

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Amirshahi, B., & Lahmiri, S. (2023). Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. *Machine Learning with Applications*, *12*, 100465. https://doi.org/10.1016/j.mlwa.2023.100465

Ampadu, S., Mensah, E. T., Aidoo, E. N., Boateng, A., & Maposa, D. (2024). A comparative study of error distributions in the GARCH model through a Monte Carlo simulation approach. *Scientific African*, *23*, e01988. https://doi.org/10.1016/j.sciaf.2023.e01988

Balietti, A. C. (2016). Trader types and volatility of emission allowance prices. Evidence from EU ETS Phase I. *Energy Policy*, *98*, 607–620. https://doi.org/https://doi.org/10.1016/j.enpol.2016.09.006

Byun, S. J., & Cho, H. (2013). Forecasting carbon futures volatility using GARCH models with energy volatilities. *Energy Economics*, *40*, 207–221. https://doi.org/10.1016/j.eneco.2013.06.017

Boungou, W., & Dufau, B. (2024). EU ETS phase IV and Industrial performance. *Economics Letters*, *236*, 111596. https://doi.org/10.1016/j.econlet.2024.111596

Böning, J., Di Nino, V., & Folger, T. (2023). Benefits and Costs of the ETS in the EU, a Lesson Learned for the CBAM Design. *SSRN Electronic Journal*.

Benson, E., Majkut, J., Reinsch, W. A., & Steinberg, F. (2023). *Analyzing the European Union's Carbon Border Adjustment Mechanism* (tech. rep.). Center for Strategic and International Studies (CSIS). Retrieved May 16, 2024, from http://www.jstor.org/stable/resrep47316

Bellora, C., & Fontagné, L. (2023). EU in search of a Carbon Border Adjustment Mechanism. *Energy Economics*, *123*, 106673. https://doi.org/10.1016/J.ENECO.2023.106673

Balash, V., & Faizliev, A. (2024). Volatility spillovers across Russian oil and gas sector. Evidence of the impact of global markets and extraordinary events. *Energy Economics*, *129*, 107202. https://doi.org/10.1016/j.eneco.2023.107202

Bhandari, H. N., Rimal, B., Pokhrel, N. R., Rimal, R., Dahal, K. R., & Khatri, R. K. (2022). Predicting stock market index using LSTM. *Machine Learning with Applications*, *9*, 100320. https://doi.org/10.1016/j.mlwa.2022.100320

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://ideas.repec.org/a/eee/econom/v31y1986i3p307-327.html

Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, *69*(3), 542–547. Retrieved May 3, 2024, from http://www.jstor.org/stable/1925546

Brooks, C. (2002). *Introductonary econometrics for finance*. Cambridge University Press.

Brooks, C. (2019). *Introductonary econometrics for finance*. Cambridge University Press.

Brooks, C. (1997). GARCH modelling in finance: a review of the software options. *Economic Journal*, *107*(443), 1271–1276.

Boersen, A., & Scholtens, B. (2014). The relationship between European electricity markets and emission allowance futures prices in phase II of the EU (European Union) emission trading scheme. *Energy*, *74*, 585–594. https://doi.org/10.1016/j.energy.2014.07.024

Catani, P., & Ahlgren, N. (2017). Combined Lagrange multiplier test for ARCH in vector autoregressive models. *Econometrics and Statistics*, *1*, 62–84. https://doi.org/https://doi.org/10.1016/j.ecosta.2016.10.006

Capros, P., Zazias, G., Evangelopoulou, S., Kannavou, M., Fotiou, T., Siskos, P., De Vita, A., & Sakellaris, K. (2019). Energy-system modelling of the EU strategy towards climate-neutrality. *Energy Policy*, *134*, 110960. https://doi.org/https://doi.org/10.1016/j.enpol.2019.110960

Castagneto-Gissey, G. (2014). How competitive are EU electricity markets? An assessment of ETS Phase II. *Energy Policy*, *73*, 278–297. https://doi.org/10.1016/j.enpol.2014.06.015

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Cerqueti, R., Giacalone, M., & Mattera, R. (2020). Skewed non-Gaussian GARCH models for cryptocurrencies volatility modelling. *Information Sciences*, *527*, 1–26. https://doi.org/https://doi.org/10.1016/j.ins.2020.03.075

Chevallier, J. (2009). Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics*, *31*, 614–625. https://doi.org/10.1016/J.ENECO.2009.02.008

Chollet, F., et al. (2015). Keras.

Cho, B.-K., Jung, H., Chung, J.-B., & Song, C.-K. (2024). Implications of the Carbon Border Adjustment Mechanism on South Korean industries: Challenges and policy recommendations. *Journal of Cleaner Production*, *444*, 141278. https://doi.org/https://doi.org/10.1016/j.jclepro.2024.141278

Campbell, J. Y., Lo, A. W., & MacKinlay, C. A. (1997). *The Econometrics of Financial Markets* (2nd ed.). Princeton University Press.

Council of the European Union. (2024). Where does the EU's gas come from?

Chen, Y., & Spokoiny, V. (2015). Modeling nonstationary and leptokurtic financial time series. *Econometric Theory*, *31*(4), 703–728. https://doi.org/10.1017/S0266466614000528

Cary, M., & Stephens, H. M. (2024). Economic, environmental, and technical gains from the Kyoto Protocol: Evidence from cement manufacturing. *Resources Policy*, *91*, 104926. https://doi.org/j.resourpol.2024.104926

Dewaelheyns, N., Schoubben, F., Struyfs, K., & Van Hulle, C. (2023). The influence of carbon risk on firm value: Evidence from the European Union Emission Trading Scheme. *Journal of Environmental Management*, *344*, 118293. https://doi.org/10.1016/j.jenvman.2023.118293

de-Graft Acquah, H. (2010). Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. *Journal of Development and Agricultural Economics*, *2*, 1–6.

Dietterich, T. (1995). Overfitting and undercomputing in machine learning. *ACM Comput. Surv.*, *27*(3), 326–327. https://doi.org/10.1145/212094.212114

Ding, S., Zhang, H., Tao, Z., & Li, R. (2022). Integrating data decomposition and machine learning methods: An empirical proposition and analysis for renewable energy generation forecasting. *Expert Systems with Applications*, *204*, 117635. https://doi.org/10.1016/j.eswa.2022.117635

Duan, K., Wang, R., Chen, S., & Ge, L. (2023). Exploring the predictability of attention mechanism with LSTM: Evidence from EU carbon futures prices. *Research in International Business and Finance*, *66*, 102020. https://doi.org/10.1016/j.ribaf.2023.102020

Ederington, L. H., & Guan, W. (2006). Measuring historical volatility. *Journal of Applied Finance*, *16*(1).

Eicke, L., Weko, S., Apergi, M., & Marian, A. (2021). Pulling up the carbon ladder? Decarbonization, dependence, and third-country risks from the European carbon border adjustment mechanism. *Energy Research & Social Science*, *80*, 102240. https://doi.org/10.1016/J.ERSS.2021.102240

Ellerman, A. D., Convery, F. J., de Perthuis, C., Alberola, E., Buchner, B. K., Delbosc, A., Hight, C., Keppler, J. H., & Matthes, F. C. (2010). *Pricing Carbon: The European Union Emissions Trading Scheme*. Cambridge University Press.

Eslahi, M., & Mazza, P. (2023). Can weather variables and electricity demand predict carbon emissions allowances prices? Evidence from the first three phases of the EU ETS. *Ecological Economics*, *214*, 107985. https://doi.org/10.1016/J.ECOLECON.2023.107985

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987–1007. Retrieved March 6, 2024, from http://www.jstor.org/stable/1912773

European Commission. (n.d.-a). Carbon Leakage. https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/free-allocation/carbon-leakage_en

European Commission. (n.d.-b). Development of EU ETS (2005-2020). https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/development-eu-ets-2005-2020_en

European Commission. (n.d.-c). What is the EU ETS? https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/what-eu-ets_en

European Commission. (2024). Carbon Border Adjustment Mechanism. https://taxation-customs.ec.europa.eu/carbon-border-adjustment-mechanism_en

Flori, A., Borghesi, S., & Marin, G. (2024). The environmental-financial performance nexus of EU ETS firms: A quantile regression approach. *Energy Economics*, *131*, 107328. https://doi.org/10.1016/j.eneco.2024.107328

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. https://doi.org/https://doi.org/10.1016/j.ejor.2017.11.054

Fu, X., Yang, B., & Wang, S. (2021). Modeling and control of plasma horizontal displacement for HL-2A tokamak based on LSTM. *Fusion Engineering and Design*, *162*, 112107. https://doi.org/10.1016/j.fusengdes.2020.112107

Galanos, A. (2023). *rugarch: Univariate GARCH models.* [R package version 1.5-1.].

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keara, and Tensorflow*. O'Reilly Media Inc.

Ghalanos, A. (2023). Introduction to the rugarch package.(Version 1.4-3). https://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance*, *48*(5), 1779–1801. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

Gläser, A., & Putaturo, C. (2022). World's poorest should not pay for climate action while EU industry pollutes for free. *Euractiv*. https://www.euractiv.com/section/energy-environment/opinion/worlds-poorest-should-not-pay-for-climate-action-while-eu-industry-pollutes-for-free/

Greene, W. H. (2002). *Econometric Analysis* (5th ed.). Prentice Hall.

Houdt, G. V., Mosquera, C., & Nápoles, G. (2020). A Review on the Long Short-Term Memory Model. *Artif Intell Rev*, *53*, 5929–5955. https://doi.org/10.1007/s10462-020-09838-1

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780.

Huang, Y., Dai, X., Wang, Q., & Zhou, D. (2021). A hybrid model for carbon price forecasting using GARCH and long short-term memory network. *Applied Energy*, *285*, 116485. https://doi.org/10.1016/J.APENERGY.2021.116485

Huang, X. Y., Ao, T.-J., Zhang, X., Li, K., Zhao, X.-Q., Champreda, V., Runguphan, W., Sakdaronnarong, C., Liu, C.-G., & Bai, F.-W. (2023). Developing high-dimensional machine learning models to improve generalization ability and overcome data insufficiency for mixed sugar fermentation simulation. *Bioresource Technology*, *385*, 129375. https://doi.org/https://doi.org/10.1016/j.biortech.2023.129375

Hull, I. (2021). *Machine Learning for Economics and Finance in TensorFlow 2: Deep Learning Models for Research and Industry*. Apress. https://doi.org/10.1007/978-1-4842-6373-0

Hassani, H., & Yeganegi, M. R. (2019). Sum of squared ACF and the Ljung–Box statistics. *Physica A: Statistical Mechanics and its Applications*, *520*, 81–86. https://doi.org/https://doi.org/10.1016/j.physa.2018.12.028

IEA. (2020). EU Emissions Trading System Phase 3 (2013-2020). https://www.iea.org/policies/4790-eu-emissions-trading-system-phase-3-2013-2020

Jousseaume, M., Menner, M., & Reichert, G. (2021). CBAM: Damaging to Climate Protection and EU Export Industries. *Cep-Study*.

Karunasingha, D. S. K. (2022). Root mean square error or mean absolute error? Use their ratio as well. *Information Sciences*, *585*, 609–629. https://doi.org/https://doi.org/10.1016/j.ins.2021.11.036

Kim, Y. (2021). Technological Innovation, the Kyoto Protocol, and Open Innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(3), 198. https://doi.org/10.3390/joitmc7030198

Kaufmann, C., & Weber, R. H. (2011). Carbon-related border tax adjustment: mitigating climate change or restricting international trade? *World Trade Review*, *10*(4), 497–525. https://doi.org/10.1017/S1474745611000292

Lahmiri, S. (2017). Modeling and predicting historical volatility in exchange rate markets. *Physica A: Statistical Mechanics and its Applications*, *471*, 387–395. https://doi.org/https://doi.org/10.1016/j.physa.2016.12.061

Liu, S., Zhang, Y., Wang, J., & Feng, D. (2024). Fluctuations and Forecasting of Carbon Price Based on A Hybrid Ensemble Learning GARCH-LSTM-Based Approach: A Case of Five Carbon Trading Markets in China. *Sustainability*, *16*(4). https://doi.org/10.3390/su16041588

Lowe, S. (2021). *The EU's carbon border adjustment mechanism: How to make it work for developing countries*. Centre For European Reforms. https://www.cer.eu/sites/default/files/pbrief_cbam_sl_21.4.21.pdf

Luo, X., Yu, R., Guo, Y., Jia, H., & Liu, D. (2022). Carbon Price Prediction of LSTM Method Based on Attention Mechanism, 2198–2202. https://doi.org/10.1109/SPIES55999.2022.10082179

Lin, B.-Q., & Zhao, H.-S. (2023). Which sectors should be covered by the EU Carbon Border Adjustment Mechanism? *Advances in Climate Change Research*. https://doi.org/10.1016/J.ACCRE.2023.11.012

Lin, B., & Zhao, H. (2023). Evaluating current effects of upcoming EU Carbon Border Adjustment Mechanism: Evidence from China's futures market. *Energy Policy*, *177*, 113573. https://doi.org/10.1016/J.ENPOL.2023.113573

Massai, L. (2011). *The Kyoto Protocol in the EU : European Community and Member States under International and European Law* (1st ed. 2011.). T.M.C. Asser Press : Imprint: T.M.C. Asser Press.

Magacho, G., Espagne, E., & Godin, A. (2024). Impacts of the CBAM on EU trade partners: consequences for developing countries. *Climate Policy*, *24*(2), 243–259. https://doi.org/10.1080/14693062.2023.2200758

Mehling, M. A., van Asselt, H., Das, K., Droege, S., & Verkuijl, C. (2019). Designing Border Carbon Adjustments for Enhanced Climate Action. *American Journal of International Law*, *113*(3), 433–481. https://doi.org/10.1017/ajil.2019.22

Martin, V., Hurn, S., & Harris, D. (2012). Properties of maximum likelihood estimators. In *Econometric modelling with time series: Specification, estimation and testing* (pp. 33–86). Cambridge University Press.

Mörsdorf, G. (2022). A simple fix for carbon leakage? Assessing the environmental effectiveness of the EU carbon border adjustment. *Energy Policy*, *161*, 112596. https://doi.org/10.1016/j.enpol.2021.112596

Miah, M., & Rahman, A. (2016). Modelling Volatility of Daily Stock Returns: Is GARCH(1,1) Enough? *American Scientific Research Journal for Engineering, Technology, and Sciences*, *18*, 29–39.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, *59*(2), 347–370. Retrieved March 21, 2024, from http://www.jstor.org/stable/2938260

Newell, R. G., Pizer, W. A., & Raimi, D. (2013). Carbon Markets 15 Years after Kyoto: Lessons Learned, New Challenges. *Journal of Economic Perspectives*, *27*(1), 123–46. https://doi.org/10.1257/jep.27.1.123

Osorio, S., Tietjen, O., Pahle, M., Pietzcker, R. C., & Edenhofer, O. (2021). Reviewing the Market Stability Reserve in light of more ambitious EU ETS emission targets. *Energy Policy*, *158*, 112530. https://doi.org/https://doi.org/10.1016/j.enpol.2021.112530

Oestreich, A. M., & Tsiakas, I. (2015). Carbon emissions and stock returns: Evidence from the EU Emissions Trading Scheme. *Journal of Banking & Finance*, *58*, 294–308. https://doi.org/https://doi.org/10.1016/j.jbankfin.2015.05.005

Pietzcker, R. C., Osorio, S., & Rodrigues, R. (2021). Tightening EU ETS targets in line with the European Green Deal: Impacts on the decarbonization of the EU power sector. *Applied Energy*, *293*, 116914. https://doi.org/j.apenergy.2021.116914

Pan, X., & Yu, L. (2024). Do China's pilot emissions trading schemes lead to domestic carbon leakage? Perspective from the firm relocation. *Energy Economics*, *132*, 107334. https://doi.org/10.1016/j.eneco.2024.107334

R Core Team. (2022). *R: A Language and Environment for Statistical Computing (v. 4.2.1 (2022-06-23 ucrt)*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Rafatirad, S. (2022). *Machine learning for computer scientists and data analysts : from an applied perspective*. Springer.

Rahman, N., Jia, G., & Zulkafli, H. (2023). GARCH Models and Distributions Comparison for Nonlinear Time Series with Volatilities. *Malaysian Journal of Fundamental and Applied Sciences*, *19*, 989–1001. https://doi.org/10.11113/mjfas.v19n6.3101

Rydberg, T. H. (2000). Realistic Statistical Modelling of Financial Data. *International Statistical Review / Revue Internationale de Statistique*, *68*(3), 233–258. Retrieved April 3, 2024, from http://www.jstor.org/stable/1403412

Schauenberg, T. (2022). Co$_2$ tax at eu border: revolution or protectionism? *DW NATURE AND ENVIRONMENT EUROPE*. https://p.dw.com/p/43hNl

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. Retrieved April 3, 2024, from http://www.jstor.org/stable/2958889

Sen, J., Sen, R., Dutta, A., Akinbo, R. S., & Daramola, O. A. (2021). *Machine Learning : algorithms, models and applications*. IntechOpen.

Segnon, M., Lux, T., & Gupta, R. (2017). Modeling and forecasting the volatility of carbon dioxide emission allowance prices: A review and comparison of modern volatility models. *Renewable and Sustainable Energy Reviews*, *69*, 692–704. https://doi.org/10.1016/j.rser.2016.11.060

Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series, 1394–1401. https://doi.org/10.1109/ICMLA.2018.00227

Sun, X., Mi, Z., Cheng, L., Coffman, D., & Liu, Y. (2023). The carbon border adjustment mechanism is inefficient in addressing carbon leakage and results in unfair welfare losses. *Fundamental Research*. https://doi.org/10.1016/j.fmre.2023.02.026

Takeda, S., & Arimura, T. H. (2024). A computable general equilibrium analysis of the EU CBAM for the Japanese economy. *Japan and the World Economy*, *70*, 101242. https://doi.org/10.1016/j.japwor.2024.101242

Vollebergh, H. R. J., & Brink, C. (2020). What Can We Learn from EU ETS? (8236). https://www.pbl.nl/sites/default/files/downloads/ifo-dice-2020-1-vollebergh-brink-what-can-we-learn-from-eu-ets-spring.pdf

Villar-Rubio, E., Huete-Morales, M.-D., & Galán-Valdivieso, F. (2023). Using EGARCH models to predict volatility in unconsolidated financial markets: the case of European carbon allowances. *Journal of Environmental Studies and Sciences*, *3*, 500–509. https://doi.org/10.1007/s13412-023-00838-5

Viteva, S., Veld-Merkoulova, Y. V., & Campbell, K. (2014). The forecasting accuracy of implied volatility from ECX carbon options. *Energy Economics*, *45*, 475–484. https://doi.org/10.1016/j.eneco.2014.08.005

Wade, K., & Jennings, M. (2016). The impact of climate change on the global economy. *Schroders Talking Point*.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82. Retrieved May 25, 2024, from http://www.jstor.org/stable/24869236

Willmott, C. J., Matsuura, K., & Robeson, S. M. (2009). Ambiguities inherent in sums-of-squares-based error statistics. *Atmospheric Environment*, *43*(3), 749–752. https://doi.org/https://doi.org/10.1016/j.atmosenv.2008.10.005

Wojciuk, M., Swiderska-Chadaj, Z., Siwek, K., & Gertych, A. (2024). Improving classification accuracy of fine-tuned CNN models: Impact of hyperparameter optimization. *Heliyon*, *10*(5), e26586. https://doi.org/10.1016/j.heliyon.2024.e26586

Ying, X. (2019). An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series*, *1168*(2), 022022. https://doi.org/10.1088/1742-6596/1168/2/022022

Zivot, E. (2008). Practical Issues in the Analysis of Univariate GARCH Models. https://faculty.washington.edu/ezivot/research/practicalgarchfinal.pdf

Zhong, J., & Pei, J. (2024). Carbon border adjustment mechanism: a systematic literature review of the latest developments. *Climate Policy*, *24*(2), 228–242. https://doi.org/10.1080/14693062.2023.2190074

Zhang, Y.-J., & Wei, Y.-M. (2010). An overview of current research on EU ETS: Evidence from its operating mechanism and economic effect. *Applied Energy*, *87*(6), 1804–1814. https://doi.org/10.1016/j.apenergy.2009.12.019

A Datasheet A: Main code

```
rm(list=ls(all=TRUE))
 library("readxl")
  library(moments)
  library("tseries")
  library("rugarch")
  library(keras)
  library(tidyverse)
  library(zoo)
  library(xts)
  library(xtable)
  library("boot")
  library("forecast")

source("functions.r") #See Datasheet B

# Import the data
data <- read_excel("data.xlsx")
data <- data.frame(data)
Date <- as.Date(data[,1])
rownames(data) <- as.character(Date)

# Set the names for the columns
colnames(data) <- c("Data", "EUETS",  "Fertilizers", "Iron Steel",  ...
    "Cement", "Aluminum", "Natural Gas")

# creating the data for Phase 3
phase3 <- with(data, data[(Date >= "2017-02-28" & Date <= "2020-12-31"),] )
Date3 <- phase3[,1]
data3 <- phase3[,-1]

# creating the data for Phase 4
phase4 <- with(data, data[(Date >= "2021-01-01" & Date <= "2024-02-29"),] )
Date4 <- phase4[,1]
data4 <- phase4[,-1]

# Plotting the price evolution of the data
postscript("graph.eps", horizontal=F, width=8, height=6)
par(mfrow = c(3,2))
for(i in 2:(ncol(data))){
```

```
  plot(data[,1], data[,i],type = "l", main = paste("Price evolution ...
     for",colnames(data)[i], "28.02.2017 - 29.02.2024"),
       xlab = "Date", ylab = "Price")
  lines(data[1:nrow(data3),1],data3[,(i-1)], col = 2)
  lines(data[(nrow(data3)+1):nrow(data),1], data4[,(i-1)], col = 3)
  legend("topleft", legend = c("Phase 3", "Phase 4"), col = c(2,3), lty = ...
     1,box.lwd = 0, bty = "n")
}
par(mfrow = c(1,1))
dev.off()
data <- data[,-1]


# Creating logreturns
for(i in 1:ncol(data)){
  ret <- diff(log(data[,i]))
  if(i == 1){r <- ret}
  else{ r <- cbind(r, ret)}
}

rownames(r) <- as.character(Date[-1])
colnames(r) <- colnames(data)
r <- data.frame(r)
r$Date <- Date[-1]
r3 <- with(r, r[(Date >= "2017-02-28" & Date <= "2020-12-31"),] )
r4 <- with(r, r[(Date >= "2021-01-01" & Date <= "2024-02-29"),] )


# Plotting the returns

postscript("return.eps", horizontal=F, width=8, height=6)
par(mfrow = c(3,2))
for(i in 1:(ncol(r)-1)){
  plot(r[,ncol(r)], r[,i],type = "l", main = paste("Log Return ...
     for",colnames(data)[i], "01.03.2017 - 29.02.2024"),
       xlab = "Date", ylab = "Logarithmic Return")
  lines(r[1:nrow(r3),ncol(r)],r3[,i], col = 2)
  lines(r[(nrow(r3)+1):nrow(r),ncol(r)], r4[,i], col = 3)
  legend("topleft", legend = c("Phase 3", "Phase 4"), col = c(2,3), lty = ...
     1,box.lwd = 0, bty = "n")
}
```

```
par(mfrow = c(1,1))
dev.off()

r <- r[-ncol(r)]


# The period the HV is calculated on
lt <- 22


# Creating plotts of the Historical Volatility
HVr <- rollapply(r, FUN = sd, width = lt)
HVr <- data.frame(HVr)
HVr$Date <- Date[-(1:lt)]

HVr3 <- with(HVr, HVr[(Date >= "2017-02-28" & Date <= "2020-12-31"),] )
HVr4 <- with(HVr, HVr[(Date >= "2021-01-01" & Date <= "2024-02-29"),] )

postscript("HistVol.eps", horizontal=F, width=8, height=6)
par(mfrow = c(3,2))
for(i in 1:(ncol(HVr)-1)){
plot(HVr[,ncol(HVr)], HVr[,i],type = "l", main = paste("HV ...
    for",colnames(data)[i], "30.03.2017 - 29.02.2024"),
      xlab = "Date", ylab = "Historical Volatility")
lines(HVr[1:nrow(HVr3),ncol(HVr)],HVr3[,i], col = 2)
lines(HVr[(nrow(HVr3)+1):nrow(HVr),ncol(HVr)], HVr4[,i], col = 3)
legend("topleft", legend = c("Phase 3", "Phase 4"), col = c(2,3), lty = ...
    1,box.lwd = 0, bty = "n")
}
par(mfrow = c(1,1))
dev.off()


#----------------------------------------------------#
#----------------------------------------------------#
# Calculating and creating a table for the descriptive statistics
for( i in 1:ncol(r)){
acf(r[,i])
pacf(r[,i])

mer <- mean(r[,i]);
```

```
std <- sd(r[,i])
skew <- skewness(r[,i])
kurt <- kurtosis(r[,i])
jb <- jarque.bera.test(r[,i])
ADF <- adf.test(r[,i], alternative = "stationary", k = 16)
LM <- Lm.test(r[,i], lag.max = 16)
LB16 <- Box.test(r[,i], lag = 16, type = "Ljung")
desc_stat <- c(mer*100, std*100, skew, kurt, jb[3], LM[3], LB16[3], ADF[4])


if(i == 1){desc.stat <- desc_stat}
else{desc.stat <- rbind(desc.stat, desc_stat)}
}
rownames(desc.stat) <- colnames(r)
colnames(desc.stat) <-  c("Mean in %", "Standard Deviation in %", "Skewness",
                          "Kurtosis", "p.value JB", "p.value LM",  ...
                             "p.value LB",  "p.value ADF")
xtable(desc.stat, digits = 4)

#----------------------------------------------------#
#----------------------------------------------------#
# Phase Selection, choosing data for phase 3

re <- r3[,-(ncol(r3))]
date <- Date3

# Creating the train and test set
insample <- round(nrow(re)*0.7,0)
outofsample <-  (insample+1):nrow(re)
rtrain <- re[1:insample,]
rtest <- re[outofsample,]
colnames(rtrain) <- colnames(data)
rownames(rtrain) <-  as.character(Date[2:(insample+1)])
colnames(rtest) <- colnames(data)
rownames(rtest) <- as.character(Date[(insample+2):length(date)])
#----------------------------------------------------#
#----------------------------------------------------#
# Determining the optimal GARCH model
garchl <- c("sGARCH", "eGARCH", "gjrGARCH")
d <- c("norm",  "std" , "ged")
```

```
garchorder <- c(rep(1,6)) #
armaorder <-c(c(0,0), c(1,0),c(1,1))
lag <- 1
outofsamples <- length(outofsample)


optgarch <- bestgarch(re, garchl, d, garchorder, armaorder, lag, outofsamples)


bestAIC <- data.frame(optgarch$bestAIC)
bestBIC <- data.frame(optgarch$bestBIC)

forAIC <- data.frame(optgarch$forAIC)
forBIC <- data.frame(optgarch$forBIC)

sdAIC <-  data.frame(optgarch$sigAIC)
sdBIC <-  data.frame(optgarch$sigBIC)


HVt <- rollapply(re, FUN = sd, width = lt)
rownames(HVt) <- rownames(re)[lt:nrow(re)]
HVttrain <- HVt[1:(insample-lt+1),]
HVttest <- HVt[(insample-lt+2):(nrow(HVt)),]

# Creating lagged variables of the actual and the simulated optimal GARCH ...
    models
lags = 1
for(i in 1:(ncol(sdAIC))){
  lagsAIC <- make.lags(sdAIC[,i], lags)
  lagsBIC <- make.lags(sdBIC[,i], lags)
  lagsAICt <- make.lags(forAIC[,i], lags)
  lagsBICt <- make.lags(forBIC[,i], lags)
  if(i == 1){HVgarchAICtrain <-lagsAIC$X
  HVgarchBICtrain <- lagsBIC$X
  HVgarchAICtest <-lagsAICt$X
  HVgarchBICtest <- lagsBICt$X}
  else{HVgarchAICtrain <- cbind(HVgarchAICtrain, lagsAIC$X)
  HVgarchBICtrain <- cbind(HVgarchBICtrain, lagsBIC$X)
  HVgarchAICtest <- cbind(HVgarchAICtest, lagsAICt$X)
  HVgarchBICtest <- cbind(HVgarchBICtest, lagsBICt$X)
  }
```

```
}

HVgarchAICtrain <- HVgarchAICtrain[-c(1:(lt-lags-lag)),]
HVgarchBICtrain <-HVgarchBICtrain[-c(1:(lt-lags-lag)),]

for(i in 1:(ncol(HVttest))){
  lagstest <- make.lags(HVttest[,i], lags)
  lagstrain <- make.lags(HVttrain[,i], lags)
  if(i == 1){HVtrain <- lagstrain$X
  HVtest <- lagstest$X}
  else{
    HVtrain <- cbind(HVtrain, lagstrain$X)
    HVtest <- cbind(HVtest, lagstest$X)}
}

rownames(HVtrain) <- rownames(HVttrain)[-lags]
rownames(HVtest) <- rownames(HVttest)[-lags]

# Creating the lagged variable of EUETS
euetstr <- make.lags(HVttrain[,1], lags)
euetste <- make.lags(HVttest[,1], lags)
#Plotting the result of the GARCH models

postscript("Phase3GarchtrainAIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(HVtrain)), euetstr$Y, type = "l", main = paste("Plot ...
    HVtrain for",colnames(bestAIC)[1], "using AIC. For phase 3"),
     ylim = c(-0,0.1),xlab = "Date", ylab = "Historical Volatility")
lines(as.Date(rownames(HVtrain)), HVgarchAICtrain[,1], type = "l", col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()
postscript("Phase3GarchtestAIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(HVtest)), euetste$Y, type = "l", main = paste("Plot ...
    HVtest for",colnames(bestBIC)[1], "using AIC. For phase 3"),
     ylim =  c(-0,0.1),xlab = "Date", ylab = "Historical Volatility")
lines(as.Date(rownames(HVtest)), HVgarchAICtest[,1], type = "l", col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

postscript("Phase3GarchtrainBIC.eps", horizontal=F, width=8, height=6)
```

```
plot(as.Date(rownames(HVtrain)),euetstr$Y, type = "l", main = paste("Plot ...
    HVtrain for",colnames(bestAIC)[1], "using BIC. For phase 3"),
      ylim = c(-0,0.1),xlab = "Date", ylab = "Historical Volatility")
lines(as.Date(rownames(HVtrain)), HVgarchBICtrain[,1], type = "l", col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

postscript("Phase3GarchtestBIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(HVtest)),euetste$Y, type = "l", main = paste("Plot ...
    HVtest for",colnames(bestBIC)[1], "using BIC. For phase 3"),
      ylim =  c(-0,0.1),xlab = "Date", ylab = "Historical Volatility")
lines(as.Date(rownames(HVtest)), HVgarchBICtest[,1], type = "l", col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

# Calculating the residual of the model
ergarchAICtrain <- euetstr$Y - data.frame(HVgarchAICtrain[,1])
ergarchBICtrain <- euetstr$Y - data.frame(HVgarchBICtrain[,1])
ergarcAIChtest <- euetste$Y  - HVgarchAICtest[,1]
ergarcBIChtest <-  euetste$Y  - HVgarchBICtest[,1]

HVwAICtrain <- cbind(HVtrain, HVgarchAICtrain)
HVwAICtest <- cbind(HVtest, HVgarchAICtest)
HVwBICtrain <- cbind(HVtrain, HVgarchBICtrain)
HVwBICtest <- cbind(HVtest, HVgarchBICtest)
#----------------------------------------------------#
#----------------------------------------------------#
#----------------------------------------------------#
#Scaling the data
HVttrain <- euetstr$Y
HVttest <- euetste$Y


for(i in 1:ncol(HVtrain)){
  HVfs <- HVttrain#[,1]
  HVfss <- HVtrain[,i]
  HVtte <- HVttest#[,1]
  HVte <- HVtest[,i]
  sd <-  sd(HVfs)
```

```
  md <- mean(HVfs)
  scaled_HVttr <- scale(HVfs, center = md, scale = sd)#Scaling independant ...
     variable
  scaled_HVtr <- scale(HVfss, center = md, scale = sd)#Scaling independant ...
     variable
  scaled_HVtte <- scale(HVtte,center = md, scale = sd) #Scaling dependant ...
     variable
  scaled_HVte <- scale(HVte,center = md, scale = sd) #Scaling dependant ...
     variable
  if(i == 1){
    s2 <- sd
    m2 <- md
    scaled_HVttrain <- scaled_HVttr
    scaled_HVtrain <- scaled_HVtr
    scaled_HVttest <- scaled_HVtte
    scaled_HVtest <- scaled_HVte
  }
  else{s2 <-cbind(s2, sd)
  m2 <-cbind(m2, md)

  scaled_HVtrain <-cbind(scaled_HVtrain, scaled_HVtr)

  scaled_HVtest <-cbind(scaled_HVtest, scaled_HVte)}
}

for(i in 1:ncol(HVwAICtrain)){
  HVfs <- HVwAICtrain[,i]
  HVtte <- HVwAICtest[,i]
  sd <-  sd(HVfs)
  md <- mean(HVfs)
  scaled_HVttr <- scale(HVfs, center = md, scale = sd)#Scaling independant ...
     variable
  scaled_HVtte <- scale(HVtte,center = md, scale = sd) #Scaling dependant ...
     variable
  if(i == 1){
    s3 <- sd
    m3 <- md
    scaled_HVgarchAICtrain <- scaled_HVttr
    scaled_HVgarchAICtest <- scaled_HVtte
  }
  else{s3 <-cbind(s3, sd)
```

```
  m3 <-cbind(m3, md)
  scaled_HVgarchAICtrain <-cbind(scaled_HVgarchAICtrain, scaled_HVttr)
  scaled_HVgarchAICtest <-cbind(scaled_HVgarchAICtest, scaled_HVtte)}
}



for(i in 1:ncol(HVwBICtrain)){
  HVfs <- HVwBICtrain[,i]
  HVtte <- HVwBICtest[,i]
  sd <-  sd(HVfs) #(max(HVfs)-min(HVfs))
  md <- mean(HVfs) #min(HVfs)
  scaled_HVttr <- scale(HVfs, center = md, scale = sd)#Scaling independant ...
      variable
  scaled_HVtte <- scale(HVtte,center = md, scale = sd) #Scaling dependant ...
      variable
  if(i == 1){
    s4 <- sd
    m4 <- md
    scaled_HVgarchBICtrain <- scaled_HVttr
    scaled_HVgarchBICtest <- scaled_HVtte
  }
  else{s4 <-cbind(s4, sd)
  m4 <-cbind(m4, md)
  scaled_HVgarchBICtrain <-cbind(scaled_HVgarchBICtrain, scaled_HVttr)
  scaled_HVgarchBICtest <-cbind(scaled_HVgarchBICtest, scaled_HVtte)}
}

#----------------------------------------------------#
#----------------------------------------------------#
patience <- 100
# Creating the LSTM model

x.train <- scaled_HVtrain[,-1]
x.train <-as.matrix(x.train)
dim(x.train) <- c(nrow(x.train), ncol(x.train), 1) #
y.train <- scaled_HVttrain



modelLSTM <- keras_model_sequential()
```

```
modelLSTM %>%
  layer_lstm(units = 32, input_shape = c(NULL, ncol(x.train), 1),
             return_sequences = TRUE) %>%
  layer_dropout(rate = 0.2) %>%
  layer_lstm(units = 32, activation = "tanh",   return_sequences = FALSE) %>%
  layer_dropout(rate = 0.2) %>%
  layer_dense(units = 1)
modelLSTM %>% compile(optimizer = optimizer_adam(),
                   loss = "mse")

summary(modelLSTM)

epo <- 200

es = callback_early_stopping(monitor="val_loss", mode='min', verbose=1, ...
    patience=patience)
history <-  modelLSTM %>% fit(x.train, y.train,
                          epochs =  epo,
                          validation_split = 0.1,
                          callback = es,
                          verbose = 1)




y_train_pred <- modelLSTM %>% predict(x.train)

# Unscalling the results

y_train_pred <- y_train_pred*s2[1]+m2[1]
y.train <- scaled_HVttrain[,1]*s2[1]+m2[1]

#Plotting the result on the train set

postscript("Phase3LSTMrain.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.train))),y.train, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
      ylab = "Historical Volatility", main = "The training data with the ...
         LSTM model for Phase 3")
lines(as.Date(rownames(data.frame(y.train))),y_train_pred, col = 2)
```

```r
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

#Analyzing the test set

x.test <- scaled_HVtest[,-1]
x.test <- as.matrix(x.test)
y.test <- scaled_HVttest*s2[1]+m2[1]
dim(x.test) <- c(nrow(x.test), ncol(x.test), 1) #
y_test_pred <- modelLSTM %>% predict(x.test)
y_test_pred <- y_test_pred*s2[1]+m2[1]

#Plotting the result on the test set
postscript("Phase3LSTMtrain.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.test))),y.test, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
      ylab = "Historical Volatility",  main = "The test data with the LSTM ...
          model for Phase 3")
lines(as.Date(rownames(data.frame(y.test))),y_test_pred, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()



# Error of the test set
etrainlstm <-  y.train - y_train_pred
etestlstm <- y.test - y_test_pred
#----------------------------------------------------#
#----------------------------------------------------#
# Creating the GARCH-LSTM AIC

x.train.AIC <- scaled_HVgarchAICtrain[,-1]
x.train.AIC <-as.matrix(x.train.AIC)
dim(x.train.AIC) <- c(nrow(x.train.AIC), ncol(x.train.AIC), 1) #
y.train.AIC <- scaled_HVttrain[,1]

modelAIC <- keras_model_sequential()
modelAIC %>%
  layer_lstm(units = 32, input_shape = c(NULL, ncol(x.train.AIC), 1),
             return_sequences = TRUE) %>%
```

```
  layer_dropout(rate = 0.2) %>%
 layer_lstm(units = 32, activation = "tanh",   return_sequences = FALSE) %>%
  layer_dropout(rate = 0.2) %>%
  layer_dense(units = 1)

modelAIC %>% compile(optimizer = optimizer_adam(),
                     loss = "mse" )

summary(modelAIC)


es = callback_early_stopping(monitor="val_loss", mode='min', verbose=1, ...
    patience=patience)
history <- modelAIC %>% fit(x.train.AIC, y.train.AIC,
                             epochs = epo,
                             validation_split = 0.1,
                             callback = es,
                             verbose = 1)




y_train_predAIC <- modelAIC %>% predict(x.train.AIC)

# Unscalling the results


y_train_predAIC <- y_train_predAIC*s2[1]+m2[1]
y.train.AIC <- scaled_HVttrain[,1]*s2[1]+m2[1]

# Plotting the GARCH-LSTM AIC based model

postscript("Phase3LSTMGARCHtrainAIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.train))),y.train, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
      ylab = "Historical Volatility", main = "The training data with the ...
          GARCH(AIC) - LSTM model for Phase 3")
lines(as.Date(rownames(data.frame(y.train))),y_train_predAIC, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()
```

```
#Analyzing the test set


x.test <- scaled_HVgarchAICtest[,-1]
x.test <-as.matrix(x.test)
y.test <- scaled_HVttest[,1]*s2[1]+m2[1]
dim(x.test) <- c(nrow(x.test), ncol(x.test), 1) #
y_test_predAIC <- modelAIC %>% predict(x.test)
y_test_predAIC <- y_test_predAIC*s2[1]+m2[1]

#Plotting the result on the test set

postscript("Phase3LSTMGARCHtestAIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.test))),y.test, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
      ylab = "Historical Volatility", main = "The test data with the ...
         GARCH(AIC) - LSTM model for Phase 3")
lines(as.Date(rownames(data.frame(y.test))),y_test_predAIC, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

# Calculating the error

etrainAIClstm <- y.train.AIC - y_train_predAIC
etestAIClstm <- y.test - y_test_predAIC
#----------------------------------------------------#
#----------------------------------------------------#
# # Creating the GARCH-LSTM BIC

x.train.BIC <- scaled_HVgarchBICtrain[,-1]
x.train.BIC <- as.matrix(x.train.BIC)
dim(x.train.BIC) <- c(nrow(x.train.BIC), ncol(x.train.BIC), 1) #
y.train.BIC <- scaled_HVttrain[,1]

modelBIC <- keras_model_sequential()
modelBIC %>%
  layer_lstm(units = 32, input_shape = c(NULL, ncol(x.train.BIC), 1),
             return_sequences = TRUE) %>%
  layer_dropout(rate = 0.2) %>%
```

```
  layer_lstm(units = 32, activation = "tanh",   return_sequences = FALSE) %>%
  layer_dropout(rate = 0.2) %>%
  layer_dense(units = 1)#, activation = "tanh")

modelBIC %>% compile(optimizer = optimizer_adam(),
                   loss = "mse")

summary(modelBIC)


es = callback_early_stopping(monitor="val_loss", mode='min', verbose=1, ...
    patience=patience)
history <- modelBIC %>% fit(x.train.BIC, y.train.BIC,
                          epochs = epo,
                          validation_split = 0.1,
                          callbacks = es,
                          verbose = 1)




y_train_predBIC <- modelBIC  %>% predict(x.train.BIC)

# Unscalling the results


y_train_predBIC <- y_train_predBIC*s2[1]+m2[1]
y.train.BIC <- scaled_HVttrain[,1]*s2[1]+m2[1]


# Plotting the GARCH-LSTM AIC based model

postscript("Phase3LSTMGARCHtrainBIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.train))),y.train, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
      ylab = "Historical Volatility", main = "The training data with the ...
        GARCH(BIC) - LSTM model for Phase 3")
lines(as.Date(rownames(data.frame(y.train))),y_train_predBIC, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()
```

```
x.test <- scaled_HVgarchBICtest[,-1]
x.test <-as.matrix(x.test)
y.test <- scaled_HVttest[,1]*s2[1]+m2[1]
dim(x.test) <- c(nrow(x.test), ncol(x.test), 1) #
y_test_predBIC <- modelBIC %>% predict(x.test)
y_test_predBIC <- y_test_predBIC *s2[1]+m2[1]

#Plotting the result on the test set


postscript("Phase3LSTMGARCHtestBIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.test))),y.test, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
     ylab = "Historical Volatility", main = "The test data with the ...
        GARCH(BIC) - LSTM model for Phase 3")
lines(as.Date(rownames(data.frame(y.test))),y_test_predBIC, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

#Calculating the error

etrainBIClstm <- y.train.BIC - y_train_predBIC
etestBIClstm <- y.test - y_test_predBIC
#--------------------------------------------------------------------------------#
#--------------------------------------------------------------------------------#
# Creating the dataframe for the error for phase 3
errtr3 <- cbind(ergarchAICtrain, ergarchBICtrain, etrainlstm, ...
    etrainAIClstm, etrainBIClstm)
errte3 <- cbind(ergarcAIChtest, ergarcBIChtest, etestlstm, etestAIClstm, ...
    etestBIClstm)

# Calculating the RMSE and MAE

for(i in 1:ncol(errtr3)){
errtriRMSE <- sqrt(mean( errtr3[,i]^2 ))
errtriMAE <- mean( abs(errtr3[,i]))
errteiRMSE <- sqrt(mean( errte3[,i]^2 ))
errteiMAE <- mean( abs(errte3[,i] ))
if (i == 1 ){errtrRMSE3 <- errtriRMSE
errteRMSE3 <- errteiRMSE
```

```
errtrMAE3 <- errtriMAE
errteMAE3 <- errteiMAE}
else{errtrRMSE3 <- cbind(errtrRMSE3,errtriRMSE)
errteRMSE3 <- cbind(errteRMSE3,errteiRMSE)
errtrMAE3 <- cbind(errtrMAE3,errtriMAE)
errteMAE3 <- cbind(errteMAE3,errteiMAE)}
}
colnames(errteRMSE3) <- c("GARCH-AIC", "GARCH-BIC", ...
   "LSTM","GARCH-LSTM-AIC", "GARCH-LSTM-BIC" )
colnames(errtrRMSE3) <- c("GARCH-AIC", "GARCH-BIC", ...
   "LSTM","GARCH-LSTM-AIC", "GARCH-LSTM-BIC" )
colnames(errteMAE3) <- c("GARCH-AIC", "GARCH-BIC", ...
   "LSTM","GARCH-LSTM-AIC", "GARCH-LSTM-BIC" )
colnames(errtrMAE3) <- c("GARCH-AIC", "GARCH-BIC", ...
   "LSTM","GARCH-LSTM-AIC", "GARCH-LSTM-BIC" )

#---------------------------------------------------------------------------------#
#Bootstrapping on RMSE based on AIC

rmsetrain3 <- rmse.test(errtr3[,4], errtr3[,1], R=1000)
rmsetrain3$stat
rmsetrain3$pval



den <- density(rmsetrain3$dist)
postscript("Phase3RMSETESTtrainAIC.eps", horizontal=F, width=8, height=6)
plot(den, main="Model: GARCH-LSTM versus Model: GARCH, based on AIC ...
   (training data)", xlim = c((min(den$x)-0.1),(max(den$x)+0.1)))
abline(v=rmsetrain3$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()

rmsetest3 <- rmse.test(errte3[,4], errte3[,1], R=1000)
rmsetest3$stat
rmsetest3$pval



den <- density(rmsetest3$dist)
postscript("Phase3RMSETESTtestAIC.eps", horizontal=F, width=8, height=6)
plot(den, main="Model: GARCH-LSTM versus Model: GARCH, based on AIC (test ...
   data)", xlim = c((min(den$x)-0.1),(max(den$x)+0.1)))
```

```
abline(v=rmsetest3$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()
#----------------------------------------------------------------------------------#
#Bootstrapping on RMSE based on BIC
rmsetrainBIC3 <- rmse.test(errtr3[,5], errtr3[,2], R=1000)
rmsetrainBIC3$stat
rmsetrainBIC3$pval



den <- density(rmsetrainBIC3$dist)
postscript("Phase3RMSETESTtrainBIC.eps", horizontal=F, width=8, height=6)
plot(den, main="Model: GARCH-LSTM versus Model: GARCH, based on BIC ...
    (training data)", xlim = c((min(den$x)-0.1),(max(den$x)+0.1)))
abline(v=rmsetrainBIC3$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()

rmsetest3BIC <- rmse.test(errte3[,5], errte3[,2], R=1000)
rmsetest3BIC$stat
rmsetest3BIC$pval



den <- density(rmsetest3BIC$dist)
postscript("Phase3RMSETESTtestBIC.eps", horizontal=F, width=8, height=6)
plot(den, main="Model: GARCH-LSTM versus Model: GARCH, based on BIC (test ...
    data)", xlim = c((min(den$x)-0.1),(max(den$x)+0.1)))
abline(v=rmsetest3BIC$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()
#-----------------------------------------------------#
#-----------------------------------------------------#
# Phase Selection, choosing phase 4, repeating the above steps

re <- r4[,-(ncol(r4))] #r4
date <- Date4 #Date4

insample <- round(nrow(re)*0.7,0)
outofsample <-  (insample+1):nrow(re)
rtrain <- re[1:insample,]
rtest <- re[outofsample,]
```

```
colnames(rtrain) <- colnames(data)
rownames(rtrain) <-  as.character(date[1:(insample)])
colnames(rtest) <- colnames(data)
rownames(rtest) <- as.character(date[(insample+1):length(date)])
#----------------------------------------------------#
#----------------------------------------------------#
# Finding the optimal GARCH model
garchl <- c("sGARCH", "eGARCH", "gjrGARCH")
d <- c("norm",  "std" , "ged")
garchorder <- c(rep(1,6)) #
armaorder <-c(c(0,0), c(1,0),c(1,1))
lag <- 1
outofsamples <- length(outofsample)#-lt+lag+2


optgarch <- bestgarch(re, garchl, d, garchorder, armaorder, lag, outofsamples)


bestAIC <- data.frame(optgarch$bestAIC)
bestBIC <- data.frame(optgarch$bestBIC)

forAIC <- data.frame(optgarch$forAIC)
forBIC <- data.frame(optgarch$forBIC)

sdAIC <-  data.frame(optgarch$sigAIC)
sdBIC <-  data.frame(optgarch$sigBIC)


HVt <- rollapply(re, FUN = sd, width = lt)
rownames(HVt) <- rownames(re)[lt:nrow(re)]
HVttrain <- HVt[1:(insample-lt+1),]
HVttest <- HVt[(insample-lt+2):(nrow(HVt)),]

lags = 1
for(i in 1:(ncol(sdAIC))){
  lagsAIC <- make.lags(sdAIC[,i], lags)
  lagsBIC <- make.lags(sdBIC[,i], lags)
  lagsAICt <- make.lags(forAIC[,i], lags)
  lagsBICt <- make.lags(forBIC[,i], lags)
  if(i == 1){HVgarchAICtrain <-lagsAIC$X
  HVgarchBICtrain <- lagsBIC$X
```

```
  HVgarchAICtest <-lagsAICt$X
  HVgarchBICtest <- lagsBICt$X}
  else{HVgarchAICtrain <- cbind(HVgarchAICtrain, lagsAIC$X)
  HVgarchBICtrain <- cbind(HVgarchBICtrain, lagsBIC$X)
  HVgarchAICtest <- cbind(HVgarchAICtest, lagsAICt$X)
  HVgarchBICtest <- cbind(HVgarchBICtest, lagsBICt$X)
  }
}

HVgarchAICtrain <- HVgarchAICtrain[-c(1:(lt-lags-lag)),]
HVgarchBICtrain <-HVgarchBICtrain[-c(1:(lt-lags-lag)),]

for(i in 1:(ncol(HVttest))){
  lagstest <- make.lags(HVttest[,i], lags)
  lagstrain <- make.lags(HVttrain[,i], lags)
  if(i == 1){HVtrain <- lagstrain$X
  HVtest <- lagstest$X}
  else{
    HVtrain <- cbind(HVtrain, lagstrain$X)
    HVtest <- cbind(HVtest, lagstest$X)}
}

rownames(HVtrain) <- rownames(HVttrain)[-lags]
rownames(HVtest) <- rownames(HVttest)[-lags]

euetstr <- make.lags(HVttrain[,1], lags)
euetste <- make.lags(HVttest[,1], lags)

postscript("Phase4GarchtrainAIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(HVtrain)), euetstr$Y, type = "l", main = paste("Plot ...
    HVtrain for",colnames(bestAIC)[1], "using AIC. For phase 4"),
      ylim = c(-0,0.1),xlab = "Date", ylab = "Historical Volatility")
lines(as.Date(rownames(HVtrain)), HVgarchAICtrain[,1], type = "l", col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

postscript("Phase4GarchtestAIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(HVtest)), euetste$Y, type = "l", main = paste("Plot ...
    HVtest for",colnames(bestBIC)[1], "using AIC. For phase 4"),
      ylim =  c(-0,0.1),xlab = "Date", ylab = "Historical Volatility")
```

```
lines(as.Date(rownames(HVtest)), HVgarchAICtest[,1], type = "l", col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

postscript("Phase4GarchtrainBIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(HVtrain)),euetstr$Y, type = "l", main = paste("Plot ...
    HVtrain for",colnames(bestAIC)[1], "using BIC. For phase 4"),
     ylim = c(-0,0.1),xlab = "Date", ylab = "Historical Volatility")
lines(as.Date(rownames(HVtrain)), HVgarchBICtrain[,1], type = "l", col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

postscript("Phase4GarchtestBIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(HVtest)),euetste$Y, type = "l", main = paste("Plot ...
    HVtest for",colnames(bestBIC)[1], "using BIC. For phase 4"),
     ylim =  c(-0,0.1),xlab = "Date", ylab = "Historical Volatility")
lines(as.Date(rownames(HVtest)), HVgarchBICtest[,1], type = "l", col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

min(HVgarchBICtrain[,1])

min(HVgarchBICtest[,1])

ergarchAICtrain <- euetstr$Y - data.frame(HVgarchAICtrain[,1])
ergarchBICtrain <- euetstr$Y - data.frame(HVgarchBICtrain[,1])
ergarcAIChtest <- euetste$Y  - HVgarchAICtest[,1]
ergarcBIChtest <-  euetste$Y  - HVgarchBICtest[,1]

HVwAICtrain <- cbind(HVtrain, HVgarchAICtrain)
HVwAICtest <- cbind(HVtest, HVgarchAICtest)
HVwBICtrain <- cbind(HVtrain, HVgarchBICtrain)
HVwBICtest <- cbind(HVtest, HVgarchBICtest)
#----------------------------------------------------#
#----------------------------------------------------#
#Scaling the data
HVttrain <- euetstr$Y
HVttest <- euetste$Y
```

```
for(i in 1:ncol(HVtrain)){
  HVfs <- HVttrain
  HVfss <- HVtrain[,i]
  HVtte <- HVttest
  HVte <- HVtest[,i]
  sd <-  sd(HVfs)
  md <- mean(HVfs)
  scaled_HVttr <- scale(HVfs, center = md, scale = sd)#Scaling independant ...
      variable
  scaled_HVtr <- scale(HVfss, center = md, scale = sd)#Scaling independant ...
      variable
  scaled_HVtte <- scale(HVtte,center = md, scale = sd) #Scaling dependant ...
      variable
  scaled_HVte <- scale(HVte,center = md, scale = sd) #Scaling dependant ...
      variable
  if(i == 1){
    s2 <- sd
    m2 <- md
    scaled_HVttrain <- scaled_HVttr
    scaled_HVtrain <- scaled_HVtr
    scaled_HVttest <- scaled_HVtte
    scaled_HVtest <- scaled_HVte
  }
  else{s2 <-cbind(s2, sd)
  m2 <-cbind(m2, md)
  scaled_HVttrain <-cbind(scaled_HVttrain, scaled_HVttr)
  scaled_HVtrain <-cbind(scaled_HVtrain, scaled_HVtr)
  scaled_HVttest <-cbind(scaled_HVttest, scaled_HVtte)
  scaled_HVtest <-cbind(scaled_HVtest, scaled_HVte)}
}


for(i in 1:ncol(HVwAICtrain)){
  HVfs <- HVwAICtrain[,i]
  HVtte <- HVwAICtest[,i]
  sd <-  sd(HVfs)
  md <- mean(HVfs)
  scaled_HVttr <- scale(HVfs, center = md, scale = sd)#Scaling independant ...
      variable
```

```
  scaled_HVtte <- scale(HVtte,center = md, scale = sd) #Scaling dependant ...
     variable
  if(i == 1){
    s3 <- sd
    m3 <- md
    scaled_HVgarchAICtrain <- scaled_HVttr
    scaled_HVgarchAICtest <- scaled_HVtte
  }
  else{s3 <-cbind(s3, sd)
  m3 <-cbind(m3, md)
  scaled_HVgarchAICtrain <-cbind(scaled_HVgarchAICtrain, scaled_HVttr)
  scaled_HVgarchAICtest <-cbind(scaled_HVgarchAICtest, scaled_HVtte)}
}


for(i in 1:ncol(HVwBICtrain)){
  HVfs <- HVwBICtrain[,i]
  HVtte <- HVwBICtest[,i]
  sd <-  sd(HVfs)
  md <- mean(HVfs)
  scaled_HVttr <- scale(HVfs, center = md, scale = sd)#Scaling independant ...
     variable
  scaled_HVtte <- scale(HVtte,center = md, scale = sd) #Scaling dependant ...
     variable
  if(i == 1){
    s4 <- sd
    m4 <- md
    scaled_HVgarchBICtrain <- scaled_HVttr
    scaled_HVgarchBICtest <- scaled_HVtte
  }
  else{s4 <-cbind(s4, sd)
  m4 <-cbind(m4, md)
  scaled_HVgarchBICtrain <-cbind(scaled_HVgarchBICtrain, scaled_HVttr)
  scaled_HVgarchBICtest <-cbind(scaled_HVgarchBICtest, scaled_HVtte)}
}


#----------------------------------------------------#
#----------------------------------------------------#

# Creating the LSTM train model, repeating as shown above,  for phase 4
```

```
x.train <- scaled_HVtrain[,-1]
x.train <-as.matrix(x.train)
dim(x.train) <- c(nrow(x.train), ncol(x.train), 1) #
y.train <- scaled_HVttrain[,1]



modelLSTM <- keras_model_sequential()
modelLSTM %>%
  layer_lstm(units = 32, input_shape = c(NULL, ncol(x.train), 1),
             return_sequences = TRUE)%>%
  layer_dropout(rate = 0.2) %>%
   layer_lstm(units = 32, activation = "tanh",   return_sequences = FALSE) %>%
  layer_dropout(rate = 0.2) %>%
  layer_dense(units = 1)
modelLSTM %>% compile(optimizer = optimizer_adam(),
                      loss = "mse")

summary(modelLSTM)

epo <- 200

es = callback_early_stopping(monitor="val_loss", mode='min', verbose=1, ...
    patience=patience)
history <-  modelLSTM %>% fit(x.train, y.train,
                               epochs =  epo,
                               validation_split = 0.1,
                              #batch_size = batch.size,
                              callbacks = es,
                              verbose = 1)

y_train_pred <- modelLSTM %>% predict(x.train)
y_train_pred <- y_train_pred*s2[1]+m2[1]
y.train <- scaled_HVttrain[,1]*s2[1]+m2[1]

postscript("Phase4LSTMtrain.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.train))),y.train, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
      ylab = "Historical Volatility", main = "The training data with the ...
          LSTM model for Phase 4")
```

```
lines(as.Date(rownames(data.frame(y.train))),y_train_pred, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

x.test <- scaled_HVtest[,-1]
x.test <- as.matrix(x.test)
y.test <- scaled_HVttest[,1]*s2[1]+m2[1]
dim(x.test) <- c(nrow(x.test), ncol(x.test), 1) #
y_test_pred <- modelLSTM %>% predict(x.test)
y_test_pred <- y_test_pred*s2[1]+m2[1]

postscript("Phase4LSTMtest.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.test))),y.test, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
      ylab = "Historical Volatility", main = "The test data with the LSTM ...
         model for Phase 4")
lines(as.Date(rownames(data.frame(y.test))),y_test_pred, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

etrainlstm <-  y.train - y_train_pred
etestlstm <- y.test - y_test_pred
#----------------------------------------------------#
#----------------------------------------------------#
# Creating the GARCH-LSTM AIC based model, repeating as shown above,  for ...
    phase 4

x.train.AIC <- scaled_HVgarchAICtrain[,-1]
x.train.AIC <-as.matrix(x.train.AIC)
dim(x.train.AIC) <- c(nrow(x.train.AIC), ncol(x.train.AIC), 1) #
y.train.AIC <- scaled_HVttrain[,1]

modelAIC <- keras_model_sequential()
modelAIC %>%
  layer_lstm(units = 32,
             input_shape = c(NULL, ncol(x.train.AIC), 1),
             return_sequences = TRUE) %>%
  layer_dropout(rate = 0.2) %>%
  layer_lstm(units = 32, activation = "tanh",   return_sequences = FALSE) %>%
```

```
  layer_dropout(rate = 0.2) %>%
  layer_dense(units = 1)

modelAIC %>% compile(optimizer = optimizer_adam(),
                     loss = "mse")

summary(modelAIC)


es = callback_early_stopping(monitor="val_loss", mode='min', verbose=1, ...
    patience=patience)
history <- modelAIC %>% fit(x.train.AIC, y.train.AIC,
                            epochs = epo,
                            validation_split = 0.1,
                            callbacks = es,
                            verbose = 1)



y_train_predAIC <- modelAIC %>% predict(x.train.AIC)#, batch_size = ...
    batch.size)
y_train_predAIC <- y_train_predAIC*s2[1]+m2[1]
y.train.AIC <- scaled_HVttrain[,1]*s2[1]+m2[1]

postscript("Phase4GARCHLSTMtrainAIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.train))),y.train.AIC, type = "l", ylim ...
    = c(0,0.1), xlab = "Date",
     ylab = "Historical Volatility", main = "The training data with the ...
        GARCH(AIC) - LSTM model for Phase 4")
lines(as.Date(rownames(data.frame(y.train))),y_train_predAIC, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

x.test <- scaled_HVgarchAICtest[,-1]
x.test <-as.matrix(x.test)
y.test <- scaled_HVttest[,1]*s2[1]+m2[1]
dim(x.test) <- c(nrow(x.test), ncol(x.test), 1) #
y_test_predAIC <- modelAIC %>% predict(x.test)#, batch_size = batch.size)
y_test_predAIC <- y_test_predAIC*s2[1]+m2[1]
```

```
postscript("Phase4GARCHLSTMtestAIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.test))),y.test, type = "l", ylim = ...
    c(0,0.1), xlab = "Date",
      ylab = "Historical Volatility", main = "The test data with the ...
          GARCH(AIC) - LSTM model for Phase 4")
lines(as.Date(rownames(data.frame(y.test))),y_test_predAIC, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
    = 1,box.lwd = 0, bty = "n")
dev.off()

etrainAIClstm <- y.train.AIC - y_train_predAIC
etestAIClstm <- y.test - y_test_predAIC
#----------------------------------------------------#
#----------------------------------------------------#
# Creating the GARCH-LSTM BIC based model, repeating as shown above,  for ...
    phase 4

x.train.BIC <- scaled_HVgarchBICtrain[,-1]
x.train.BIC <- as.matrix(x.train.BIC)
dim(x.train.BIC) <- c(nrow(x.train.BIC), ncol(x.train.BIC), 1) #
y.train.BIC <- scaled_HVttrain[,1]

batch.size <- 1

modelBIC <- keras_model_sequential()
modelBIC %>%
  layer_lstm(units = 32, input_shape = c(NULL, ncol(x.train.BIC), 1),
             return_sequences = TRUE )%>%
  layer_dropout(rate = 0.2) %>%
  layer_lstm(units = 32, activation = "tanh",   return_sequences = FALSE) %>%
  layer_dropout(rate = 0.2) %>%
  layer_dense(units = 1)

modelBIC %>% compile(optimizer = optimizer_adam(),
                     loss = "mse")
summary(modelBIC)



es = callback_early_stopping(monitor="val_loss", mode='min', verbose=1, ...
    patience=patience)
history <- modelBIC %>% fit(x.train.BIC, y.train.BIC,
```

```
                            epochs = epo,
                            validation_split = 0.1,
                            callbacks = es,
                            verbose = 1)



y_train_predBIC <- modelBIC  %>% predict(x.train.BIC)


y_train_predBIC <- y_train_predBIC*s2[1]+m2[1]
y.train.BIC <- scaled_HVttrain[,1]*s2[1]+m2[1]

postscript("Phase4GARCHLSTMtrainBIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.train))),y.train.BIC, type = "l", ylim ...
   = c(0,0.1), xlab = "Date",
     ylab = "Historical Volatility", main = "The training data with the ...
        GARCH(BIC) - LSTM model for Phase 4")
lines(as.Date(rownames(data.frame(y.train))),y_train_predBIC, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
   = 1,box.lwd = 0, bty = "n")
dev.off()

x.test <- scaled_HVgarchBICtest[,-1]
x.test <-as.matrix(x.test)
y.test <- scaled_HVttest[,1]*s2[1]+m2[1]
dim(x.test) <- c(nrow(x.test), ncol(x.test), 1) #
y_test_predBIC <- modelBIC %>% predict(x.test)
y_test_predBIC <- y_test_predBIC *s2[1]+m2[1]

postscript("Phase4GARCHLSTMtestBIC.eps", horizontal=F, width=8, height=6)
plot(as.Date(rownames(data.frame(y.test))),y.test, type = "l", ylim = ...
   c(0,0.1), xlab = "Date",
     ylab = "Historical Volatility", main = "The test data with the ...
        GARCH(BIC) - LSTM model for Phase 4")
lines(as.Date(rownames(data.frame(y.test))),y_test_predBIC, col = 2)
legend("topleft", legend = c("Real HV", "Predicted HV"), col = c(1,2), lty ...
   = 1,box.lwd = 0, bty = "n")
dev.off()

etrainBIClstm <- y.train.BIC - y_train_predBIC
```

```
etestBIClstm <- y.test - y_test_predBIC

#-----------------------------------------------------------------------------#
#-----------------------------------------------------------------------------#
# Creating the dataframe for the error for phase 4

errtr4 <- cbind(ergarchAICtrain, ergarchBICtrain, etrainlstm, ...
    etrainAIClstm, etrainBIClstm)
errte4 <- cbind(ergarcAIChtest, ergarcBIChtest, etestlstm, etestAIClstm, ...
    etestBIClstm)



for(i in 1:ncol(errtr4)){
  errtriRMSE <- sqrt(mean( errtr4[,i]^2 ))
  errtriMAE <- mean( abs(errtr4[,i] ))
  errteiRMSE <- sqrt(mean( errte4[,i]^2 ))
  errteiMAE <- mean( abs(errte4[,i] ))
  if (i == 1 ){errtrRMSE4 <- errtriRMSE
  errteRMSE4 <- errteiRMSE
  errtrMAE4 <- errtriMAE
  errteMAE4 <- errteiMAE}
  else{errtrRMSE4 <- cbind(errtrRMSE4,errtriRMSE)
  errteRMSE4 <- cbind(errteRMSE4,errteiRMSE)
  errtrMAE4 <- cbind(errtrMAE4,errtriMAE)
  errteMAE4 <- cbind(errteMAE4,errteiMAE)}
}
colnames(errteRMSE4) <- c("GARCH-AIC", "GARCH-BIC", ...
    "LSTM","GARCH-LSTM-AIC", "GARCH-LSTM-BIC" )
colnames(errtrRMSE4) <- c("GARCH-AIC", "GARCH-BIC", ...
    "LSTM","GARCH-LSTM-AIC", "GARCH-LSTM-BIC" )
colnames(errteMAE4) <- c("GARCH-AIC", "GARCH-BIC", ...
    "LSTM","GARCH-LSTM-AIC", "GARCH-LSTM-BIC" )
colnames(errtrMAE4) <- c("GARCH-AIC", "GARCH-BIC", ...
    "LSTM","GARCH-LSTM-AIC", "GARCH-LSTM-BIC" )

#-----------------------------------------------------------------------------#
#-----------------------------------------------------------------------------#
#RMSE test for the GARCH-LSTM AIC based model
rmsetest4AIC <- rmse.test(errtr4[,4], errtr4[,1], R=1000)
rmsetest4AIC$stat
```

```
rmsetest4AIC$pval



den <- density(rmsetest4AIC$dist)
postscript("Phase4RMSETESTtrainAIC.eps", horizontal=F, width=8, height=6)
plot(den, main="Model: GARCH-LSTM versus Model: GARCH, based on AIC (train ...
   data)", xlim = c((min(den$x)-0.1),(max(rmsetest4AIC$stat)+0.1)))
abline(v=rmsetest4AIC$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()

rmsetest4AICtest <- rmse.test(errte4[,4], errte4[,1], R=1000)
rmsetest4AICtest$stat
rmsetest4AICtest$pval



den <- density(rmsetest4AICtest$dist)
postscript("Phase4RMSETESTtestAIC.eps", horizontal=F, width=8, height=6)
plot(den, main="Model: GARCH-LSTM versus Model: GARCH, based on AIC (test ...
   data)", xlim = c((min(den$x)-0.1),(max(rmsetest4AICtest$stat)+0.1)))
abline(v=rmsetest4AICtest$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()
#-----------------------------------------------------------------------------#
#-----------------------------------------------------------------------------#
#RMSE test for the GARCH-LSTM BIC based model
rmsetest4BIC <- rmse.test(errtr4[,5], errtr4[,2], R=1000)
rmsetest4BIC$stat
rmsetest4BIC$pval



den <- density(rmsetest4BIC$dist)
postscript("Phase4RMSETESTtrainBIC.eps", horizontal=F, width=8, height=6)
plot(den, main="Model: GARCH-LSTM versus Model: GARCH, based on BIC (train ...
   data)", xlim = c((min(den$x)-0.1),(max(den$x)+0.1)))
abline(v=rmsetest4BIC$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()

rmsetest4BICtest <- rmse.test(errte4[,5], errte4[,2], R=1000)
rmsetest4BICtest$stat
```

```
rmsetest4BICtest$pval



den <- density(rmsetest4BICtest$dist)
postscript("Phase4RMSETESTtestBIC.eps", horizontal=F, width=8, height=6)
plot(den, main="Model: GARCH-LSTM versus Model: GARCH, based on BIC (test ...
    data)", xlim = c((min(den$x)-0.1),(max(rmsetest4BICtest$stat)+0.1)))
abline(v=rmsetest4BICtest$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()
#-------------------------------------------------------------------------------#
#-------------------------------------------------------------------------------#
#MAE test for the GARCH-LSTM AIC based model
maetest4AIC <- mae.test(errtr4[,4], errtr4[,1], R=1000)
maetest4AIC$stat
maetest4AIC$pval



den <- density(maetest4AIC$dist)
postscript("Phase4MAETESTtrainAIC.eps", horizontal=F, width=8, height=6)
plot(den, main="MAE test: Model: GARCH-LSTM versus Model: GARCH, based on ...
    AIC (train data)", xlim = c((min(den$x)-0.1),(max(maetest4AIC$stat)+0.1)))
abline(v=maetest4AIC$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()

maetest4AICtest <- mae.test(errte4[,4], errte4[,1], R=1000)
maetest4AICtest$stat
maetest4AICtest$pval

den <- density(maetest4AICtest$dist)
postscript("Phase4MAETESTtestAIC.eps", horizontal=F, width=8, height=6)
plot(den, main="MAE test: Model: GARCH-LSTM versus Model: GARCH, based on ...
    AIC (test data)", xlim = c((min(den$x)-0.1),(max(den$x)+0.1)))
abline(v=maetest4AICtest$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()
#-------------------------------------------------------------------------------#
#MAE test for the GARCH-LSTM BIC based model
maetest4BIC <- mae.test(errtr4[,5], errtr4[,2], R=1000)
maetest4BIC$stat
```

```
maetest4BIC$pval



den <- density(maetest4BIC$dist)
postscript("Phase4MAETESTtrainBIC.eps", horizontal=F, width=8, height=6)
plot(den, main="MAE test: Model: GARCH-LSTM versus Model: GARCH, based on ...
   BIC (train data)", xlim = c((min(den$x)-0.1),(max(maetest4BIC$stat)+0.1)))
abline(v=maetest4BIC$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()

maetest4BICtest <- mae.test(errte4[,5], errte4[,2], R=1000)
maetest4BICtest$stat
maetest4BICtest$pval



den <- density(maetest4BICtest$dist)
postscript("Phase4MAETESTtestBIC.eps", horizontal=F, width=8, height=6)
plot(den, main="MAE test: Model: GARCH-LSTM versus Model: GARCH, based on ...
   BIC (test data)", xlim = ...
   c((min(den$x)-0.1),(max(maetest4BICtest$stat)+0.1)))
abline(v=maetest4BICtest$stat, col="red", lty=2, lwd=2)
abline(v=1)
dev.off()

#--------------------------------------------------------------------------------#
#--------------------------------------------------------------------------------#

#Creating a table to represent the errors in the different phases

errtrRMSE <- rbind(errtrRMSE3, errtrRMSE4)
errteRMSE <- rbind(errteRMSE3, errteRMSE4)

rownames(errtrRMSE) <- c("Phase 3", "Phase 4")
rownames(errteRMSE) <- c("Phase 3", "Phase 4")

xtable(errtrRMSE, digits = 7)
xtable(errteRMSE, digits = 7)



errtrMAE <- rbind(errtrMAE3, errtrMAE4)
```

```
errteMAE <- rbind(errteMAE3, errteMAE4)

rownames(errtrMAE) <- c("Phase 3", "Phase 4")
rownames(errteMAE) <- c("Phase 3", "Phase 4")
xtable(errtrMAE, digits = 7)
xtable(errteMAE, digits = 7)
```

B Datasheet B: Functions

```
    #This is the functions.r source



## Coppied directly from previous courses, SE-507: Machine learning with ...
   applications in finance:
#Creates lags for a time series

make.lags <- function(x, nlags) {
  # This function prepares the data for estimating the AR(p) model
  # Inputs:
  # x is the time series
  # nlags is the number of required lags

  nobs <- length(x)
  n <- nobs - nlags
  if(n < 2)
    stop("Too short time series for this number of lags!")
  # make the y vector
  Y <- x[(nlags+1):nobs]
  # make the X matrix
  X <- matrix(nrow=n, ncol=nlags)
  for(i in 1:nlags)
    X[,i] <- x[(nlags-i+1):(nobs-i)]
  return(list(Y=Y,X=X))
}

rmse.stat <- function(data, indices) {
  # this function computes the statistic used in the RMSE test

  # access the errors
```

```
  x <- data[indices,]
  errors1 <- x[,1]
  errors2 <- x[,2]
  # compute the RMSEs
  RMSE1 <-mean( abs(errors1^2 )) #sqrt(mean( errors1^2 ))
  RMSE2 <- mean( abs(errors2^2 ))#sqrt(mean( errors2^2 ))
  # compute and return the test statistic
  stat = RMSE2/RMSE1
  return(stat)
}

rmse.test <- function(errors1, errors2, R=1000) {
  # This function conducts the test of Model 1 versus Model 2 using RMSE
  # Null hypothesis is that both models provide equal RMSEs
  # Alternative hypothesis is that Model 1 is better than Model 2 (Model 1 ...
     has smaller RMSE)
  # Inputs:
  # errors1 is the vector of forecast errors of Model 1
  # errors2 is the vector of forecast errors of Model 2
  # R is the number of bootstrap simulations
  # Output is a list that contains:
  # stat is the test statistic
  # pval the p-value of the test
  # dist is the distribution of the test statistic

  n = length(errors1)
  if(length(errors2) != n)
    stop("Different lengths of the errors!\n")

  # bootstrap
  data = cbind(errors1, errors2)
  res = boot(data=data, statistic=rmse.stat, R=R)

  stat = res$t0 # statistic computed using Predicted errors
  stat.sim = res$t # distribution of statistic

  stat.sim = stat.sim - stat + 1 # "null correction" to have mean of 1

  # compute the p-value
  pval <- length(which(stat.sim > stat))/R
```

```
  return(list(stat=stat, pval=pval, dist=stat.sim))
}
#----------------------------------------------------------------------------------#
mae.stat <- function(data, indices) {
  # this function computes the statistic used in the MAE test

  # access the errors
  x <- data[indices,]
  errors1 <- x[,1]
  errors2 <- x[,2]
  # compute the RMSEs
  MAE1 <-mean( abs(errors1^2 ))
  MAE2 <- mean( abs(errors2^2 ))
  # compute and return the test statistic
  stat = MAE2/MAE1
  return(stat)
}
mae.test <- function(errors1, errors2, R=1000) {
  # This function conducts the test of Model 1 versus Model 2 using MAE
  # Null hypothesis is that both models provide equal MAEs
  # Alternative hypothesis is that Model 1 is better than Model 2 (Model 1 ...
     has smaller RMSE)
  # Inputs:
  # errors1 is the vector of forecast errors of Model 1
  # errors2 is the vector of forecast errors of Model 2
  # R is the number of bootstrap simulations
  # Output is a list that contains:
  # stat is the test statistic
  # pval the p-value of the test
  # dist is the distribution of the test statistic

  n = length(errors1)
  if(length(errors2) != n)
    stop("Different lengths of the errors!\n")

  # bootstrap
  data = cbind(errors1, errors2)
  res = boot(data=data, statistic=mae.stat, R=R)

  stat = res$t0 # statistic computed using Predicted errors
  stat.sim = res$t # distribution of statistic
```

```
  stat.sim = stat.sim - stat + 1 # "null correction" to have mean of 1

  # compute the p-value
  pval <- length(which(stat.sim > stat))/R

  return(list(stat=stat, pval=pval, dist=stat.sim))
}



#-------------------------------------------------------------------------------#
#-------------------------------------------------------------------------------#
## Self made code
#------------------------------------------------------------------#
# Best GARCH model
#Returns the optimal garch for each specific return series based on AIC ...
    and BIC


bestgarch <- function(r, garch, d, garchorder, armaorder, lags, outofsample){
  #The function calculates the optimal GARCH model


  for(y in 1:ncol(r)){
    mod.perf <- matrix(nrow = length(garchl)*length(d), ncol = 2)
    rn <- rep(NA, length(d)*length(garchl))
    count <- 0
    counter <- 0



   rg <- r[,-y] #Remove the dependant variable
  rgl = list()
      for(a in 1:ncol(rg)){
    regr <- make.lags(rg[,a], lags) #Create the external regressors
    reg <- regr$X
    rfor <- forecast(ts(rg[-c(1:(nrow(reg)-outofsample+lag)),a])) #create ...
        forecast of the external regressors

    if(a == 1){regressor <- reg
    regfor <- rfor$fitted}
```

```
  else{regressor <- cbind(regressor, reg)
  regfor <- cbind(regfor, rfor$fitted)}
  if(a == 5){rgl[[1]] <- regfor}
    }


for(i in 1:length(d)){
    counter <- counter + 1
    for(t in 1:length(garchl)){

      mod.spec <- ugarchspec(mean.model = ...
          list(armaOrder=armaorder[(counter):(counter+1)], ...
          external.regressors = regressor),
                              variance.model =  list(model = garchl[t], ...
                                  garchOrder = ...
                                  garchorder[(counter):(counter+1)],
                                                    external.regressors = ...
                                                        NULL),
                              distribution.model = d[i]) #Create spec for ...
                                  the given garch, distribution and armaorder



      mod.fit <- ugarchfit( spec = mod.spec, data = r[-(1:lags),y], ...
          solver = "hybrid", out.sample = outofsample) #Fit the model
      count <- count + 1
      mod.perf[count,1] <- infocriteria(mod.fit)[1] # Obtaining the AIC
      mod.perf[count,2] <- infocriteria(mod.fit)[2] # Obtainin the BIC
      arma <- paste(c(armaorder[counter]),c(armaorder[counter+1]), sep=" ...
          , ")
      garchi <- paste(c(garchorder[counter]),c(garchorder[counter+1]), ...
          sep=" , ")
      rn[count] <- paste("Armaorder" ,paste0(c("( "), arma, c(" )")) ...
          ,c(d[i]), c(garchl[t]),paste0(c("( "), garchi, c(" )")))

    }
    counter <- counter + 1
    if(counter == (length(armaorder))){counter <- 0}
  }
  infocrit <- c("AIC", "BIC")
  rownames(mod.perf) <- rn
  colnames(mod.perf) <- infocrit
```

```
#Finding the optimal model based on AIC
bestcolA <- colnames(mod.perf)[ apply(mod.perf, 2, function(x) ...
    any(x==min(mod.perf[,1])))]
bestrowA <- rownames(mod.perf)[ apply(mod.perf, 1, function(x) ...
    any(x==min(mod.perf[,1])))]
print(paste("for",colnames(r)[y], ":" ,bestrowA,"is the best model ...
    based on", infocrit[1]))

#Finding the optimal model based on BIC
bestcolB <- colnames(mod.perf)[ apply(mod.perf, 2, function(x) ...
    any(x==min(mod.perf[,2])))]
bestrowB <- rownames(mod.perf)[ apply(mod.perf, 1, function(x) ...
    any(x==min(mod.perf[,2])))]
print(paste("for",colnames(r)[y], ":", bestrowB,"is the best model ...
    based on", infocrit[2]))

bestAIC <- str_split(bestrowA, pattern = " ", simplify = TRUE)
bestBIC <- str_split(bestrowB, pattern = " ", simplify = TRUE)

# Fetching the optimal specifications for the different models
armA <- as.numeric(bestAIC[3])
armaA <- as.numeric(bestAIC[5])
garA <- as.numeric(bestAIC[10])
garaA <- as.numeric(bestAIC[12])
armB <- as.numeric(bestBIC[3])
armaB <- as.numeric(bestBIC[5])
garB <- as.numeric(bestBIC[10])
garaB <- as.numeric(bestBIC[12])

dA <- bestAIC[7]
dB <- bestBIC[7]

garchA <- bestAIC[8]
garchB <- bestBIC[8]

mod.specAIC <- ugarchspec(mean.model = list(armaOrder=c(armA,armaA), ...
    external.regressors = regressor),
                           variance.model =  list(model = garchA, ...
                              garchOrder = c(garA,garaA),
                                                  external.regressors = ...
                                                     NULL),
```

```
                            distribution.model = dA) #Creating the ...
                               optimal AIC model
mod.fitAIC <- ugarchfit( spec = mod.specAIC, data = r[-(1:lags),y], ...
    solver = "hybrid", out.sample = outofsample)  #Fitting the optimal ...
    AIC model
fit.ret.AIC <- fitted(mod.fitAIC)
setfixed(mod.specAIC) <- as.list(coef(mod.fitAIC))

#Simulating based on the optimal AIC model
sim <- ugarchpath(spec = mod.specAIC,  m.sim = 1, n.sim = 1*outofsample,
                 mexsimdata=rgl, vexsimdata=NULL)
simsAIC <- sim@path$sigmaSim
AICsim <- apply(simsAIC, 1, mean)
fit.sig.AIC <- sigma(mod.fitAIC)
dateoutofsample <- as.Date(tail(rownames(r), outofsample))
forecastAIC <- xts(AICsim, dateoutofsample)



if(y == 1){
retsAIC <- fit.ret.AIC
sigsAIC <- fit.sig.AIC
forAIC <- forecastAIC}
else {retsAIC <- cbind(retsAIC,fit.ret.AIC)
sigsAIC <- cbind(sigsAIC,fit.sig.AIC)
forAIC <- cbind(forAIC, forecastAIC)
}

mod.specBIC <- ugarchspec(mean.model = list(armaOrder=c(armB,armaB), ...
    external.regressors = regressor),
                            variance.model =  list(model = garchB, ...
                               garchOrder = c(garB,garaB),
                                                  external.regressors = ...
                                                     NULL),
                            distribution.model = dB) #Creating the ...
                               optimal BIC model
mod.fitBIC <- ugarchfit(spec = mod.specBIC, data = r[-(1:lags),y], ...
    solver = "hybrid", out.sample = outofsample) #Fitting the optimal ...
    BIC model
fit.ret.BIC <- fitted(mod.fitBIC)
setfixed(mod.specBIC) <- as.list(coef(mod.fitBIC))
```

```
    #Simulating based on the optimal BIC model
    sim <- ugarchpath(spec = mod.specBIC, m.sim = 1, n.sim = 1*outofsample,
                      mexsimdata=rgl, vexsimdata=NULL)
    simsBIC <- sim@path$sigmaSim
    BICsim <- apply(simsBIC, 1, mean)
  fit.sig.BIC <- sigma(mod.fitBIC)
     forecastBIC <- xts(BICsim, dateoutofsample)



    if(y == 1){
      retsBIC <- fit.ret.BIC
      sigsBIC <- fit.sig.BIC
      forBIC <- forecastBIC}
    else {retsBIC <- cbind(retsBIC,fit.ret.BIC)
    sigsBIC <- cbind(sigsBIC,fit.sig.BIC)
    forBIC <- cbind(forBIC, forecastBIC)
    }

  }
  colnames(retsAIC) <-  colnames(r)
  colnames(retsBIC) <-  colnames(r)
  colnames(sigsAIC) <-  colnames(r)
  colnames(sigsBIC) <-  colnames(r)
  colnames(forAIC) <-  colnames(r)
  colnames(forBIC) <-  colnames(r)
  return(list("bestAIC" = retsAIC, "sigAIC" = sigsAIC, "forAIC" = ...
      forAIC,"bestBIC" = retsBIC, "sigBIC" = sigsBIC, "forBIC" = forBIC))

}
```

C Additional Plots and Tables

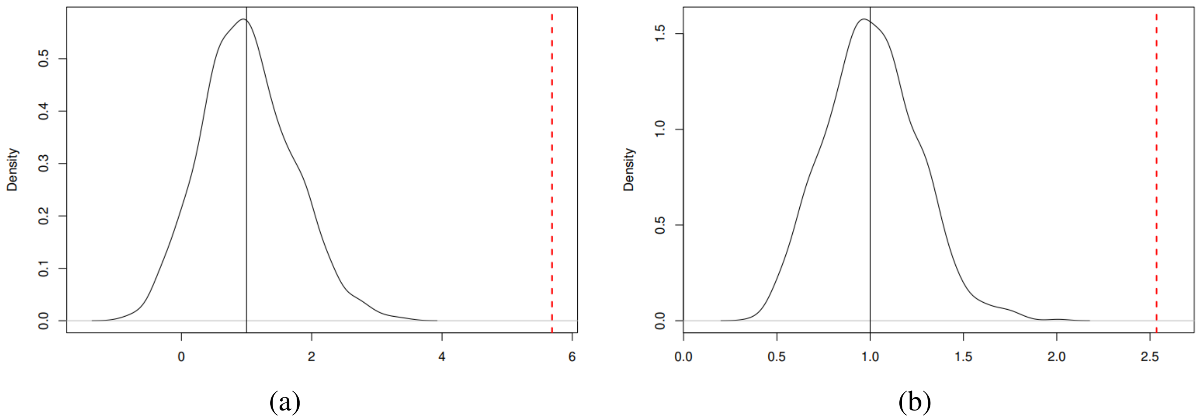


Figure 15: Plots of the estimated test statistic (in black) and the true test statistic (in red) for the bootstrapped MAE on EU ETS on phase 4 (a) training data and (b) test data. For $H_0 : \frac{MAE_{Hybrid}}{MAE_{GARCH}} = 1$

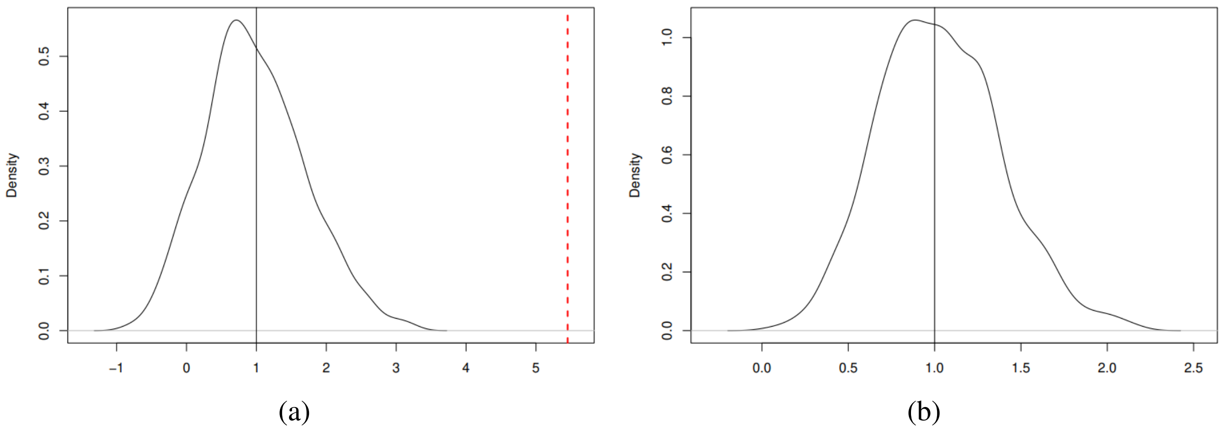


Figure 16: Plots of the estimated test statistic (in black) and the true test statistic (in red) for the bootstrapped MAE on EU ETS on phase 4 (a) training data and (b) test data. For $H_0 : \frac{MAE_{Hybrid}}{MAE_{GARCH}} = 1$

Table 6: Descriptive statistics for phase 3

	EU ETS	Fertilizers	Iron & Steel	Cement	Aluminum	Natural Gas
Mean in %	0.1828	-0.0021	0.0597	-0.0674	-0.0106	-0.0123
Standard Deviation in %	2.8804	1.4103	1.9495	2.1777	1.9974	2.6156
Skewness	-0.4243	-0.6830	-0.7694	-1.6175	-0.6261	0.4356
Kurtosis	7.5498	12.9294	14.7397	46.4632	8.7314	16.8136
p.value JB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p.value LM	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p.value LB	0.0447	0.0000	0.0001	0.0000	0.0202	0.0083
p.value ADF	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100

Table 7: Descriptive statistics for phase 4

	EU ETS	Fertilizers	Iron & Steel	Cement	Aluminum	Natura Gas
Mean in %	0.0651	0.0053	-0.0138	-0.0023	-0.0222	-0.0147
Standard Deviation in %	2.7304	1.8378	1.9535	1.9785	2.3491	3.9047
Skewness	-0.5257	-0.6558	0.0867	1.0621	-0.1990	-0.1978
Kurtosis	8.0848	6.9763	4.5110	10.7830	4.0449	4.9028
p.value JB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p.value LM	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
p.value LB	0.4516	0.0767	0.0000	0.0125	0.0537	0.1233
p.value ADF	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100

D Discussion paper – Alexander Rosåsen Longnes: International

This discussion paper serves as the final remark before completing my master´s degree in Business and Administration with a specialization in analytical finance at the University of Agder. The process of finalizing the master´s thesis has been very exciting and educational. I would like to thank my partner, Eivind, for his invaluable contribution to the thesis and for his assistance when necessary.

In this discussion paper, I will be discussing our thesis in the context of the topic "international". Firstly, I will provide an introduction to our thesis. Following that, I will discuss the relationship between the overarching topic "international" and the findings and relevance of our thesis. Finally, I will the briefly provide a summary and conclusion of the discussion paper.

Introduction

Climate changes and environmental pollution have become one of the biggest global issues. The global emissions level has become a significant threat to both the economy and a sustainable future (Flori et al., 2024; Wade and Jennings, 2016). On this basis changes are needed. In 1997, a UN-decision-making body adopted the Kyoto Protocol as a global effort to reduce greenhouse gases (GHG) in the hope of achieving carbon neutrality by 2050 (Cary and Stephens, 2024). Later, in March 2000, a green paper was introduced and presented the outlining ideas behind the EU ETS. The EU ETS were officially launched in 2005 (Newell et al., 2013) and the main goal was to reduce the impact from power sectors and heavy industries on the environment (Ellerman et al.,2010, p. 22). In 2019, the European Green Deal was introduced. The main goal behind the introduction of the European Green Deal was to achieve climate neutrality by 2050, agreeing with the goal from the Kyoto Protocol (Pietzcker et al., 2021). Within the European Green Deal framework, the CBAM was introduced. The main ideas behind the implementation of the CBAM was to reduce emissions and prevent the risk of firms within EU to relocate their production to countries with lower environmental focus (carbon leakage) (Schauenberg, 2022). With the CBAM, tariffs are imposed on the import of products that are carbon-intensive and exposed to the risk of carbon leakage. The tariffs apply to markets such as electricity, iron & steel, fertilizer, hydrogen, cement, and aluminum (Schauenberg, 2023). CBAM is not expected fully implemented before 2026, but on the 1^{st} of October 2023, the transitional phase of the CBAM began. From the 1^{st} of October

2023 to 31^{st} of January 2024 the first reporting period of the CBAM in the transitional period began (European Commission, 2024). In this period, EU aims to gather information about how different markets and companies evolve and adapt to the gradual implementation of the CBAM.

In our thesis, we decided to work with the prices of aluminum, iron & steel, fertilizers, natural gas, EU ETS and cement. We turned the prices into the logarithmic returns to work with the data. We further analysed the data and applied several statistical tests. Specifically, we carried out the Jarque-Bera, Lagrange-Multiplier, Ljung-Box and Augmented Dickey-Fuller tests. The results from these tests four indicated that the price returns all were all independent and identically distributed and every price return series were stationary. Furthermore, the methodology in our thesis is motivated by the paper by Amirshahi and Lahmiri (2023). In our thesis, we utilized hybrid models consisting of different Generalized Autoregressive Conditional Heteroskedasticity (GARCH) type models and a Long-Short Term Memory (LSTM) model. The GARCH models we decided to use were the standard GARCH model, an exponential GARCH model, and a Glosten-Jagannathan-Runkle GARCH. These were based on different assumptions about the distribution of the residuals, with the most optimal model chosen based on Akaike information criterion or the Bayesian information criterion. Once the optimal GARCH-type model is combined with the true historical volatility, we created a hybrid GARCH-LSTM model to analyze the impact of the CBAM on the EU ETS in phase 3 and phase 4. The results from our model demonstrate that the hybrid GARCH-LSTM model outperforms the other GARCH models and the LSTM in terms of both root mean squared error and mean absolute error. Furthermore, our results show that the model predicts better during phase 4 than phase 3.

Discussion

The topic of our thesis relates to several areas within the overarching theme of “international”. Firstly, our thesis relates to climate changes and environmental policy. Global warming has become one of the world´s biggest threats and has a big impact on the environment (G. Chen, 2023). From The Paris Agreement from 2015, almost all countries in the world agreed to limit the global warming to between 1.5°C and 2°C (United Nations Framework Convention, n.d). The main goal is to reach climate neutrality between 2050 and 2100. Climate neutrality means that we do not emit more GHG than we are able to capture or remove (United Nations Framework Convention, n.d). The EU ETS is one initiative aimed at achieving this goal of climate neutrality. With the

EU ETS being the world´s first trading system and the largest carbon pricing mechanism in the world, the EU ETS becomes one of the largest initiatives and most important policy tools to reduce GHG emissions in Europe (European Commission, n.d.-b). By pricing carbon, it creates an incentive for companies to reduce their GHG emissions (European Commission, n.d.-b). There are several ways for companies to reduce their GHG emissions. For example, companies can shift their focus towards more renewable energy. However, this transition often requires companies to invest more in low-carbon technology and energy efficiency. Different low-carbon technologies include renewable energy sources such as solar power, wind power, hydro power and geothermal energy, transportation technology such as electric and hydrogen vehicles and lastly carbon capture and storage. All these examples will most likely reduce GHG emissions but can at the same time influence other markets. If sectors advance in low-carbon technologies and renewable energies, will most likely influence the non-renewable energy sources like oil, gas and coal.

CBAM was implemented with the intention of reducing GHG emissions and reduce carbon leakage (Schauenberg, 2022), which can result in a level playing field for companies within EU. This was intended to happen by imposing tariffs on carbon intensive products imported to the EU (Cho et al., 2024). Although these tariffs apply to companies within the EU, markets outside of Europe can also be affected. Firstly, the trade dynamics can be influenced. When EU imposes tariffs on imported products, the costs for countries exporting to the EU will increase significantly. This can lead to higher prices being passed on the consumers or exporters seeking new markets. This will either result in increased prices of the products or that the global trade dynamics will shift. Takeda and Arimura (2024) analyzed the effect of CBAM on the Japanese economy. Their findings showed that the CBAM reduced carbon leakage and had a positive impact on the gross domestic product (GDP) and welfare. On the other hand, Sun et al. (2023) analyzed the CBAM and their findings showed that the CBAM reduced carbon leakage, but at the same time overburden developing countries. Developing countries are often rely on export to the EU and often do not have the financial capabilities to invest in the correct and necessary technology. This can lead to economic crises and exacerbate exsisting inequalities.

Our thesis is closely linked to events occurring in the global landscape. Through several analyses, we clearly observe that the price variables in our thesis are being influenced by several international events. The world operates in a way where countries have agreements and policies to follow, and markets follows trends and influences one another. In our thesis there are especially three events that have impacted the price variables in our thesis since the implementation of the CBAM began.

The first event occurred at the end of 2018 and the beginning of 2019, when the United Kingdom was in the final phase of leaving the EU. For a country to leave the EU is a complex and advanced process and require big changes as countries are bound by trade agreements, laws, and regulations within the EU. The decision from the United Kingdom to leave the EU shocked several markets. The EU and United Kingdom had to renegotiate new trade agreements, which created a lot of uncertainty in many countries and markets. In our study, the EU ETS, iron & steel, aluminum and natural gas were negatively impacted in the log returns by the United Kingdom´s decision to leave the EU.

The second international event that impacted the price variables in our thesis was the COVID-19 pandemic, which began in March 2020, and led to a global lockdown. The pandemic influenced and caused shocks in several markets. This led to drastic changes in the demand for various products as the world was in lockdown and many sectors had to pause their productions for some period. From our study, we observed that the COVID-19 pandemic affected the log returns for all our variables.

The most recent and last global event was the Russian invasion of Ukraine, which began in February 2022. Russia has been one of the biggest exporters of natural gas to Europe for several. Following the invasion, the EU decided to sanction Russia. As a result of the sanctions, EU now had to significantly modify the trading patterns of importing natural gas. In 2021, over 40% of the EU´s consumption of pipeline gas originated from Russia, but by 2023, Russia´s share decreased to around 8% (Council of the European Union, 2024). This demonstrates the extend to which the EU had to adapt following a geopolitical event. From our thesis, we can see that the Russian invasion of Ukraine affected natural gas, EU ETS, fertilizers and aluminum. All three events show how international occurrences can impact different markets.

Conclusion

In conclusion, our thesis examines the effect the CBAM on the EU ETS. In this discussion paper, I have discussed this topic with the overarching topic “international”, considering trends and forces, environmental challenges and international market dynamics. Through statistical analysis of the EU ETS, aluminum, fertilizers, iron & steel, cement and natural gas, we observe evidence on how these variables are being affected by international trends and forces. We have seen the variables being affected by Brexit, the COVID-19 pandemic and a geopolitical situation in the Russian invasion of Ukraine.

Furthermore, the implementation of the CBAM influences trade dynamics in international markets outside of EU. In our thesis, we have observed that the CBAM reduces carbon leakage and have had a positive impact on the GDP and welfare in Japan. On the other hand, we have seen that the CBAM overburdens developing countries. By understanding how CBAM, and international markets are being influenced by international trends and forces, we can receive valuable information into how the world can deal with the climate challenges the world is faces. Our thesis can provide policymakers, businesses, researchers, countries and institutions with information on how the implementation of CBAM can affect the EU ETS, and how international trends and forces influence different markets. Understanding these mechanisms can help reduce GHG emissions, achieve climate neutrality and secure economic sustainability for a sustainable future.

Based on our thesis and this discussion paper, it will be very interesting to see how the markets are changing toward the full implementation of the CBAM in 2026. Furthermore, it will be very exciting to see how companies and markets around the world manage to adapt to the changes brought by CBAM, how they are affected by international trends and forces, and whether they can adjust in a way that reduces their GHG emissions to achieve the goal of climate neutrality, economic sustainability, and sustainable future.

References

Amirshahi, B., & Lahmiri, S. (2023). Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. Machine Learning with Applications, 12, 100465. https://doi.org/10.1016/j.mlwa.2023.100465

Cary, M., & Stephens, H. M. (2024). Economic, environmental, and technical gains from the Kyoto Protocol: Evidence from cement manufacturing. Resources Policy, 91, 104926. https://doi.org/10.1016/j.resourpol.2024.104926

Chen, G. (2023). Impact of carbon border adjustment mechanism on China's manufacturing sector: A dynamic recursive CGE model based on an evolutionary game. Journal of Environmental Management, 347, 119029. https://doi.org/10.1016/j.jenvman.2023.119029

Cho, B.-K., Jung, H., Chung, J.-B., & Song, C.-K., (2024). Implications of the Carbon Border Adjustment Mechanism on South Korean Industries: Challenges and policy recommendations. Journals of Cleaner Production, 444, 141278. https://doi.org/10.1016/j.jclepro.2024.141278

Council of the European Union. (2024). Where does the EU's gas come from? Retrieved from https://www.consilium.europa.eu/en/infographics/eu-gas-supply/

Ellerman, A. D., Convery, F. J., de Perthuis, C., Alberola, E., Buchner, B. K., Delbosc, A., Hight, C., Keppler, J. H., & Matthes, F. C. (2010). Pricing carbon: The European Union Emissions Trading Scheme. Cambridge University Press.

European Commission. (2024). Carbon Border Adjustment Mechanism. Retrieved from https://taxation-customs.ec.europa.eu/carbon-border-adjustment-mechanismen

European Commission. (n.d.). Development of EU ETS (2005-2020). Retrieved from https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/development-eu-ets-2005-2020en

Flori, A., Borghesi, S., & Marin, G. (2024). The environmental-financial performance nexus of EU ETS firms: A quantile regression approach. Energy Economics, 131, 107328.

https://doi.org/10.1016/j.eneco.2024.107328

Newell, R. G., Pizer, W. A., & Raimi, D. (2013). Carbon markets 15 years after Kyoto: Lessons learned, new challenges. Journal of Economic Perspectives, 27(1), 123-46. https://doi.org/10.1257/jep.27.1.123

Pietzcker, R. C., Osorio, S., & Rodrigues, R. (2021). Tightening EU ETS targets in line with the European Green Deal: Impacts on the decarbonization of the EU power sector. Applied Energy, 293, 116914. https://doi.org/10.1016/j.apenergy.2021.116914

Schauenberg, T. (2022). CO2 tax at EU border: Revolution or protectionism? DW Nature and Environment Europe. https://p.dw.com/p/43hNl

Sun, X., Mi, Z., Cheng, L., Coffman, D. M., & Liu, Y. (2023). The carbon border adjustment mechanism is inefficient in addressing carbon leakage and results in unfair welfare losses. Fundamental Research. https://doi.org/10.1016/j.fmre.2023.02.026

Takeda, S., & Arimura, T. H. (2024). A computable general equilibrium analysis of the EU CBAM for the Japanese economy. Japan and the World Economy, 70, 101242. https://doi.org/10.1016/j.japwor.2024.101242

United Nations Framework Convention on Climate Change. (n.d). The Paris Agreement. Retrieved from https://unfccc.int/process-and-meetings/the-paris-agreement

Wade, K., & Jennings, M. (2016). The impact of climate change on the global economy. Schroders Talking Point.

E Discussion paper – Eivind Rytter Huseby: Responsible

This master thesis is an analysis of the effect of the Carbon Boarder Adjustment Mechanism (CBAM) on the European Union's Emission Trading System (EU ETS). EU launched the transitional period of the CBAM the 1st of October 2023 as a tool to fight the greenhouse gas (GHG) emissions. During this period business pay no fee, but they need to document their GHG emission. The focus of the CBAM during the transitional phase is the GHG emission that occurs in the production of electricity, iron & steel, fertilizer, hydrogen, cement, and aluminum (Schauenberg, 2022). The data is split into phase 3 and phase 4 based on the directives from EU. The thesis builds on knowledge from the courses in empirical finance and machine learning with application in finance by analyzing different econometric models, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model and a machine learning model, Long Short-Term Memory (LSTM). Based on an idea in the paper of Amirshahi and Lahmiri (2023), the thesis analyzes three different GARCH type models, the standard (Bollerslev, 1986), an exponential GARCH model (Nelson, 1991) and a Glosten-Jagannathan-Runkle GARCH (Glosten et al.,1993). Further the residuals are assumed under three different distributions. For each variable the optimal model is chosen based on Akaike decision criterion or the Bayesian information criterion. The output from the optimal GARCH-type model is combined with the true, lagged historical volatility. This is used as the input data in the hybrid GARCH-LSTM model. To analyze the historical volatility of the EU ETS two different performance measures are used, the root mean squared error (RMSE) and the (MAE). Why two different performance measures are used is elaborated on in the thesis. The results from our model suggests that the hybrid GARCH-LSTM model outperforms the GARCH and the LSTM in terms of root mean squared error and the mean absolute error. This is observed both for models chosen based on AIC and based on BIC. In addition, it is observed for both phases. Further, it is seen that the predictability of the model increased during phase 4. This suggests that implementation of CBAM can have increased the variables effectiveness of forecasting the historical volatility of the EU ETS. The results are bootstrapped to determine that there exists statistical evidence at a 95% level of the GARCH-LSTM model outperforming the alternatives.

CBAM was introduced by EU as a part of the Councils goal of reducing the emissions in the EU by 55% by 2030 when comparing with 1990 levels. From 2026 the EU importers needs to purchase CBAM certificates. These are used to cover the price of GHG emissions of electricity, iron & steel, fertilizer, hydrogen, cement, and aluminum that are not already paid for (Corvino, 2023). However, the CBAM is argued to be unjust. This is a result of more vulnerable economies being

harder affected by climate changes despite them being the least responsible for the GHG emissions (Gläser and Putaturo, 2022). In this regard CBAM have been criticized, Gläser and Putaturo (2022) argues that the CBAM will have particular effects on some of the worlds developing countries. By imposing taxation on goods, some developing countries will lose investment opportunities. In addition, there is a possibility of lost income for exporters of these goods due to European countries choosing alternatives as a cost saving measure (Magacho et al, 2024). This will further enlarge the gap between developing countries and industrialized, European countries (Lowe, 2021). Despite this, the EU have, to the best of my knowledge, yet to propose a solution that will ensure that developing countries are not suffering any financial loss after the CBAM becomes fully operational from 2026.
In the paper of Böhringer et al. (2012) , it is argued that implementing a Boarder Carbon Adjustment (BCA) put the burden of emission reduction on the developing countries. This is neither efficient nor is it responsible. The findings of the above-mentioned papers, suggest that the CBAM need to be carefully addressed. Moreover, the EU need to address these issues before the full launch of CBAM. This to ensure that the CBAM secures a sustainable future for all countries and industries. Moreover, the arguments of the papers show that EU has to take into consideration some of the United Nations (UN) sustainable development goals before fully implementing the CBAM. The arguments of the authors above, can be used to determine that the implementation of the CBAM greatly affects many of the goals. First of all, fighting poverty is difficult if the economy of developing countries is hampered by implementation of the CBAM, this is a result of exporters would seek alternative markets to save costs. This can also greatly affect other goals, such as economic growth and less inequality. However, the CBAM is a tool to reduce the climate changes, and thus a tool to aid the UNs goal of stopping the climate changes. This will be further disgusted below, where the effects the CBAM have to carbon leakage is investigated.

The CBAM is in a transitional phase, thus EU are still able to make changes to the CBAM to ensure the introduction of a system that is more fair. Researchers who have investigated the effects of the CBAM, have also proposed solutions to make the CBAM more fair for developing countries. Gläser and Putaturo (2022) and Lowe (2021) and Magacho et al (2024) suggest that developing countries should be exempted from the CBAM. They differ however in the extent of how to handle the financial aspect. Gläser and Putaturo (2022) suggest that this exemption should be temporarily in addition, they suggest that EU can use the income of the CBAM to aid the developing countries to transform to a decarbonized economy. Lowe (2021 argues that as long as the CBAM only covers the goods implemented in the EU ETS, exempting developing countries will not particularly

be very costly for EU. Thus they do not need to include the developing countries in the CBAM. The argument of Magacho et al (2024) is that the developing countries are not able to adapt their economy to be less carbon emitting, and therefore, they support the findings of Gläser and Putaturo (2022). In their paper, Magacho et al. (2024) argue that returning some of the income from the CBAM will aid these countries to shift their economy to become more sustainable. If the EU decides to use the income to aid the developing countries, this will be a great tool to address some of the UN's sustainability goals, amongst others, those previously mentioned.
Other critics of the CBAM have argued that the CBAM is in conflict with rules presented by the World Trade Organization (WTO) (Kaufmann and Weber, 2011; Mehling et al., 2019). In their paper, Mehling et al. (2019) discuss the design of a BCA. They argue that the conceptual design of a BCA follows a principle of taxing a product at the place of consumption. Further they draw focus to six different key features of a BCA and enlightens on possible compliance these features have with the WTO. The CBAM seeks to increase the competitiveness of the EU based companies, and suggestions are that exporters will receive rebates in the form of free allowances. Thereby large portions of their emissions are excluded from the CBAM. This is among the legal issues WTO members have addresses (Espa et al., 2022). According to Mehling et al. (2019), exporters are crucial in assessing the effectiveness of the CBAM. They claim that excluding them will weaken the CBAMs effect on the carbon leakage.

Arguments by many of the supporters of CBAM and EU is the good effects the CBAM will have on reducing carbon leakage (European Commission, n.d.; Mehling et al., 2019; Mörsdorf, 2022; Takeda and Arimura, 2024). By carbon leakage, it is meant that companies reallocate investment and production to countries with less regulated emission standards (Eicke et al.,2021). In a recent paper, Takeda and Arimura (2024) analyzed the effect CBAM have on Japanese economy, findings in their paper suggests that the implementation of the CBAM will significantly reduce carbon leakage from the EU. They conclude that reduction of carbon leakage shows indications of the CBAM being an effective tool to control carbon leakage. These findings support the literature review of Mehling et al. (2019), who found and argued that a BCA is a promising tool to secure reduction in carbon leakage. In another literature study, Branger and Quirion (2014) perform a meta-analysis on 25 different paper. They find that according to the paper they analyze, there is a difference in implementing a BCA in regard of reduction of carbon leakage. They further conclude that BCAs do reduce leakage, and that implementing BCA will with statistical significance reduce carbon leakage ratio by 6%. Böhringer et al. (2012) discuss the impact of implementing a BCA, they support the previous mentioned findings that carbon leakage can be expected by implementation of the CBAM. They conclude that this will be especially noticeable in the emission-intensive and trade-exposed industries. However, as previously mentioned, the CBAM is being reviewed by the EU and the WTO. These parties need to agree for the CBAM to be launched. Thus, there is still room for changing the CBAM. Espa et al. (2022) claims that final design of the CBAM is crucial in determining the effectiveness of the CBAM. Other researchers have argued that the advantages of the CBAM is that it will force other large emitters to adapt a carbon trading system (Barichella et al., 2021; Espa et al., 2022). As part of the large emitters and economies in the world [1] [2], EU must take a large responsibility in reducing emissions.

In the paper by Barichella et al. (2021) it is argued that the implementation of the CBAM has encouraged the trading partners of the EU to accelerate a system for carbon pricing. The paper concludes that implementing an "international carbon club" is important for the EU, and the world to secure decarbonization of global industries. Shum (2024) support this finding by stating that major EU trade partners are willing to adapt their emissions trading policies in accordance with the EU. This is also argued for by Nordhaus (2020), who use the term "carbon club", suggesting that such a coalition, where members "commit to strong steps to reduce emissions and mechanisms to penalize countries that do not participate". The reality, and as also mentioned by Nordhaus (2020) is that such a club is far from becoming reality. However, if the findings of Barichella et al. (2021) and Shum (2024) hold, the CBAM might be a step in the right direction. It is worth mentioning that the CBAM and such climate club is not the same (Ambec,2022).

As one of the world's largest emitter EU has a responsibility in addressing the climate changes. Further, as one of the world's largest economic forces EU must utilize the influence they have to ensure that others follow. The CBAM is such a tool, and as argued for earlier, there exist evidence that other large emitters are willing to adapt according to the CBAM in some way. As shown above, there exist ethical issues of implementing the CBAM. Some researchers are showing the impact the CBAM will have on developing countries, and as shown above, the research today mainly focuses on the possible negative impacts. In regard of the ethical issues of the CBAM it is important to note that the transitional phase is designed for exactly this purpose. EU seeks aid from affected sectors and academia to ensure that the transition to a fully implemented CBAM goes smoothly and is in accordance with international law. What is important to draw from the research is that many different scenarios are investigated, and different tax systems, prices and other external factors are

[1] https://www.epa.gov/ghgemissions/global-greenhouse-gas-overview

[2] https://www.forbesindia.com/article/explainers/top-10-largest-economies-in-the- world/86159/1

analyzed. It is important that these are analyzed carefully by the EU It is important that these are analyzed carefully by the EU before the full implementation of the CBAM. Other findings presented in this paper discussion note show that the CBAM is an efficient tool at handling carbon leakage. It has been showed that implementing the CBAM reduce carbon leakage. The EU have, through available research, been granted useful insights they need to account for before the full launch of the CBAM in 2026. Despite researchers generally agreeing that the CBAM will have a positive impact on reducing carbon leakage, the question, ”at what cost?” need to be addressed.

References

Amirshahi, B., & Lahmiri, S. (2023). Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. Machine Learning with Applications, 12, 100465. https://doi.org/10.1016/j.mlwa2023.100465

Ambec, S. (2022). The European union's carbon border adjustment mechanism: Challenges and perspectives. 22-1365. https://publications.utcapitole.fr/id/eprint/46339/

Barichella, A., Perthuis, C. d., Mathieu, C., Jianjun Tu, K., & Nagashima, M. (2021). Can the biggest emitters set up a climate club? a review of international carbon pricing debates.

Böhringer, C., Balistreri, E. J., & Rutherford, T. F. (2012). The role of border carbon adjustment in unilateral climate policy: Overview of an energy modeling forum study (emf 29) [The Role of Border Carbon Adjustment in Unilateral Climate Policy: Results from EMF 29]. Energy Economics, 34, S97–S110. https://doi.org/https://doi.org/101016/j.eneco.2012.10.003

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedas ticity. Journal of Econometrics, 31 (3), 307–327. https://ideas.repecorg/a/eee/econom/v31y1986i3p307-327.html

Branger, F., & Quirion, P. (2014). Would border carbon adjustments prevent carbon leakage and heavy industry competitiveness losses? Insights from a meta-analysis of recent economic studies. Ecological Economics, 99, 29–39. https://doi.org / 10. 1016/j.ecolecon.2013.12.010

Corvino, F. (2023). The compound injustice of the EU carbon border adjustment mechanism (CBAM). Ethics, Policy & Environment, 0 (0), 1–20. https://doi.org/10.1080/21550085.2023.2272237

Espa, I., Francois, J., & van Asselt, H. (2022). The EU proposal for a carbon border adjustment mechanism (CBAM): An analysis under WTO and climate change law.

Eicke, L., Weko, S., Apergi, M., & Marian, A. (2021). Pulling up the carbon ladder? decarbonization, dependence, and third-country risks from the European carbon border adjustment mechanism. Energy Research & Social Science, 80, 102240. https://doi.org/10.1016/J.ERSS.2021.102240

European Commission. (n.d.). Carbon leakage. https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/free-allocation/carbon-leakage_en

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance, 48 (5), 1779–1801. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

Gläser, A., & Putaturo, C. (2022). World's poorest should not pay for climate action while eu industry pollutes for free. Euractiv.
https://www.euractiv.com/section/energyenvironment/opinion/worlds-poorest-should- not- pay- for-climate- action- while- eu- industry- pollutes- for-free/

Lowe, S. (2021). The EU's carbon border adjustment mechanism: How to make it work for developing countries. Centre For European Reforms.
https://www.cer.eu/sites/default/files/pbrief cbam sl 21.4.21.pdf

Magacho, G., Espagne, E., & Godin, A. (2024). Impacts of the CBAM on EU trade partners: Consequences for developing countries. Climate Policy, 24 (2), 243–259.
https://doi.org/10.1080/14693062.2023.2200758

Mehling, M. A., van Asselt, H., Das, K., Droege, S., & Verkuijl, C. (2019). Designing border carbon adjustments for enhanced climate action. American Journal of International Law, 113 (3), 433–481.https://doi.org/10.1017/ajil.2019.22

Mörsdorf, G. (2022). A simple fix for carbon leakage? assessing the environmental effectiveness of the eu carbon border adjustment. Energy Policy, 161, 112596.
https://doi.org/10.1016/j.enpol.2021.112596

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica, 59 (2), 347–370. Retrieved March 21, 2024,http://www.jstor.org/stable/2938260

Nordhaus, W. (2020). The climate club: How to fix a failing global effort. Foreign Affairs, 99 (3), pp. 10–17. Retrieved June 1, 2024, https://www.jstor.org/stable/26985603

Schauenberg, T. (2022). Co2 tax at eu border: Revolution or protectionism? DW NATURE AND ENVIRONMENT EUROPE. https://p.dw.com/p/43hNl

Shum, R. Y. (2024). Decarbonizing contending industries: Policy design and the EU's CBAM;. Climate Policy, 24 (2), 274–279. https://doi.org/10.1080/14693062.2023.2194275

Takeda, S., & Arimura, T. H. (2024). A computable general equilibrium analysis of the EU CBAM for the japanese economy. Japan and the World Economy, 70, 101242. https://doi.org/https://doi.org/10.1016/j.japwor.2024.101242