COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI

MÁSTER EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER

DESIGN OF QUANTITATIVE MODELS FOR DEVELOPING AUTOMATIC TRADING STRATEGIES IN ENERGY MARKETS

Autor: Carlos Jesús Pretel Parejo-Merino

Director: Antonio Bello Morales

Co-Director: Javier Reneses Guillén

Madrid

Enero de 2020

AUTORIZACIÓN PARA LA DIGITALIZACIÓN, DEPÓSITO Y DIVULGACIÓN EN RED DE PROYECTOS FIN DE GRADO, FIN DE MÁSTER, TESINAS O MEMORIAS DE BACHILLERATO

1º. Declaración de la autoría y acreditación de la misma.

El autor D. Carlos Jesús Pretel Parejo-Merino

DECLARA ser el titular de los derechos de propiedad intelectual de la obra: 'Design of quantitative models for developing automatic trading strategies in energy markets', que ésta es una obra original, y que ostenta la condición de autor en el sentido que otorga la Ley de Propiedad Intelectual.

2º. Objeto y fines de la cesión.

Con el fin de dar la máxima difusión a la obra citada a través del Repositorio institucional de la Universidad, el autor **CEDE** a la Universidad Pontificia Comillas, de forma gratuita y no exclusiva, por el máximo plazo legal y con ámbito universal, los derechos de digitalización, de archivo, de reproducción, de distribución y de comunicación pública, incluido el derecho de puesta a disposición electrónica, tal y como se describen en la Ley de Propiedad Intelectual. El derecho de transformación se cede a los únicos efectos de lo dispuesto en la letra a) del apartado siguiente.

3º. Condiciones de la cesión y acceso

Sin perjuicio de la titularidad de la obra, que sigue correspondiendo a su autor, la cesión de derechos contemplada en esta licencia habilita para:

a) Transformarla con el fin de adaptarla a cualquier tecnología que permita incorporarla a internet y hacerla accesible; incorporar metadatos para realizar el registro de la obra e incorporar "marcas de agua" o cualquier otro sistema de seguridad o de protección.
b) Reproducirla en un soporte digital para su incorporación a una base de datos electrónica, incluyendo el derecho de reproducir y almacenar la obra en servidores, a los efectos de garantizar su seguridad, conservación y preservar el formato.
c) Comunicarla, por defecto, a través de un archivo institucional abierto, accesible de modo libre y gratuito a través de internet.
d) Cualquier otra forma de acceso (restringido, embargado, cerrado) deberá solicitarse expresamente y obedecer a causas justificadas.
e) Asignar por defecto a estos trabajos una licencia Creative Commons.
f) Asignar por defecto a estos trabajos un HANDLE (URL *persistente)*.

4º. Derechos del autor.

El autor, en tanto que titular de una obra tiene derecho a:

a) Que la Universidad identifique claramente su nombre como autor de la misma
b) Comunicar y dar publicidad a la obra en la versión que ceda y en otras posteriores a través de cualquier medio.
c) Solicitar la retirada de la obra del repositorio por causa justificada.
d) Recibir notificación fehaciente de cualquier reclamación que puedan formular terceras personas en relación con la obra y, en particular, de reclamaciones relativas a los derechos de propiedad intelectual sobre ella.

5º. Deberes del autor.

El autor se compromete a:

a) Garantizar que el compromiso que adquiere mediante el presente escrito no infringe ningún derecho de terceros, ya sean de propiedad industrial, intelectual o cualquier otro.
b) Garantizar que el contenido de las obras no atenta contra los derechos al honor, a la intimidad y a la imagen de terceros.
c) Asumir toda reclamación o responsabilidad, incluyendo las indemnizaciones por daños, que

pudieran ejercitarse contra la Universidad por terceros que vieran infringidos sus derechos e intereses a causa de la cesión.

d) Asumir la responsabilidad en el caso de que las instituciones fueran condenadas por infracción de derechos derivada de las obras objeto de la cesión.

6º. Fines y funcionamiento del Repositorio Institucional.

La obra se pondrá a disposición de los usuarios para que hagan de ella un uso justo y respetuoso con los derechos del autor, según lo permitido por la legislación aplicable, y con fines de estudio, investigación, o cualquier otro fin lícito. Con dicha finalidad, la Universidad asume los siguientes deberes y se reserva las siguientes facultades:

- ➢ La Universidad informará a los usuarios del archivo sobre los usos permitidos, y no garantiza ni asume responsabilidad alguna por otras formas en que los usuarios hagan un uso posterior de las obras no conforme con la legislación vigente. El uso posterior, más allá de la copia privada, requerirá que se cite la fuente y se reconozca la autoría, que no se obtenga beneficio comercial, y que no se realicen obras derivadas.
- ➢ La Universidad no revisará el contenido de las obras, que en todo caso permanecerá bajo la responsabilidad exclusive del autor y no estará obligada a ejercitar acciones legales en nombre del autor en el supuesto de infracciones a derechos de propiedad intelectual derivados del depósito y archivo de las obras. El autor renuncia a cualquier reclamación frente a la Universidad por las formas no ajustadas a la legislación vigente en que los usuarios hagan uso de las obras.
- ➢ La Universidad adoptará las medidas necesarias para la preservación de la obra en un futuro.
- ➢ La Universidad se reserva la facultad de retirar la obra, previa notificación al autor, en supuestos suficientemente justificados, o en caso de reclamaciones de terceros.

Madrid, a 14 de enero de 2020

ACEPTA

Fdo. Carlos Jesús Pretel Parejo-Merino

Motivos para solicitar el acceso restringido, cerrado o embargado del trabajo en el Repositorio Institucional:

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

'Design of quantitative models for developing automatic trading strategies in energy markets'

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el curso académico 2019/20 es de mi autoría, original e inédito y no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido tomada de otros documentos está debidamente referenciada.

Fdo.: Carlos Jesús Pretel Parejo - Merino

Fecha: 10 / 01 / 2020

Autorizada la entrega del proyecto

LOS DIRECTORES DEL PROYECTO

Fdo.: Antonio Bello Morales

Javier Reneses Guillén

Fecha: 10/ 01/ 2020

COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI

MÁSTER EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER

DESIGN OF QUANTITATIVE MODELS FOR DEVELOPING AUTOMATIC TRADING STRATEGIES IN ENERGY MARKETS

Autor: Carlos Jesús Pretel Parejo-Merino

Director: Antonio Bello Morales

Co-Director: Javier Reneses Guillén

Madrid

Enero de 2020

DESIGN OF QUANTITATIVE MODELS FOR DEVELOPING AUTOMATIC TRADING STRATEGIES IN ENERGY MARKETS

Author: Pretel Parejo-Merino, Carlos Jesús

Director: Bello Morales, Antonio

Co-Director: Reneses Guillén, Javier

Collaborating Entity: ICAI – Universidad Pontificia Comillas

Energy markets are complex and difficult to predict. In particular, commodity markets usually present a great amount of structural changes and are affected by political factors, international strategies, etc. A commodity can be defined as a product that is characterized for being commercialized in global markets and by a low level of differentiation among producers (examples of energy commodities are oil, gas and coal). In addition, commodity markets affect volatility and uncertainty of electricity markets, for instance. Therefore, it is essential for those energy companies that need to purchase this type of products for generating electricity to elaborate trading strategies concerning these markets, as well as for other agents and companies that usually operate in them.

In this context, quantitative forecasting models are adequate for developing tools with the ability of carrying out automatic trading strategies. However, these tools are commonly developed for being applied in stock markets and other types of markets due to the complexity energy markets present as it is not easy to replicate them for this application. This is one of the innovations this model constitutes, although it is one of the main difficulties that were required to be faced too.

Regarding the concepts existing in the title of the project, is important to define the principal ones for understanding the methodology that is proposed. Firstly, we can define a trading strategy as those plans and actions that are carried out for achieving a given benefit (in this case, a financial one). These strategies are based on quantitative models, which are mathematical models that are used for adjusting variables and making predictions based on these previous adjustments of the data. Finally, this will be implemented in energy markets, and in particular commodity markets, whose main characteristics have been mentioned above.

Among the forecasting techniques that can be commonly found in the literature, there is a lot of different criteria for classifying them. In this project, they will be classified according to the prediction horizon in which the output forecasts are reasonably accurate. Among the short-term forecasting techniques, the most relevant ones are those based on artificial intelligence techniques and time series analysis. The first ones are models with the ability of correctly interpreting the data, learning from them and using this learning for achieving an accurate adjustment. On the other hand, time series analysis tries to identify patterns in historical data for autocorrelated variables or in those cases in which is not easy to adjust explanatory variables.

Concerning long-term forecasting techniques, it is necessary to remark the importance of error correction models. These are based on assuming the existence of a statistical relationship between two or more variables: cointegration, and they cannot be developed for non-cointegrated variables. Cointegration is a more powerful statistical relationship than correlation. Cointegration analysis consists in identifying for those variables that appear to be correlated in a certain period of time that this correlation is stable and truly a causal relationship in the long term. For proving this, several statistical tests are commonly used. They are known as cointegration tests. Once the existence of a correlation relationship between the variables, error correction models are useful for adjusting these variables and capturing the dynamics they present together taking into account the equilibrium relationship estimated in the long term.

In order to establish a generic methodology for the development of advanced quantitative techniques to build a tool for executing trading strategies automatically and especially concerning their application in energy commodity markets, some steps are proposed. Firstly, the statistical relationship between those commodities that are considered for the implementation of trading strategies is studied. In this first step, it is studied whether the variables have significant correlation relationships and cointegration is also studied, in order to determine for which variables error correction models can be applied. Subsequently, the structural changes that may exist in the series considered are studied. As these types of markets are very complex and with frequent structural changes, this step is important to identify the optimal number of structural changes that exist in these series to take them into account in the predictions. The quantitative models are then developed to adjust the training data. These models are then statistically validated to verify that the adjustments they give rise to are accurate and that the predictions are also reasonably close to the actual values in the validation horizon considered, in order to avoid falling into an overfitting of the models. After this, the type of data on which the user wants to operate and the type of model on which to base the automatic trading tool are selected. In this context, the possibility of developing strategies for monthly or daily prices is considered, depending on whether the computational effort that supposes a greater number of daily price data can be assumed. Finally, the trading strategy is implemented automatically based on these previously developed quantitative models.

A efectos de establecer una metodología genérica para el desarrollo de técnicas cuantitativas avanzadas para construir una herramienta de ejecución de estrategias de trading de manera automática y en especial en relación a su aplicación en mercados de commodities de energía, se proponen una serie de pasos. En primer lugar, se estudia la relación estadística entre aquellas commodities que se consideran para la implementación de estrategias de trading. En este primer paso se estudia si las variables presentan relaciones de correlación significativas y también se estudia la cointegración, a efectos de determinar para qué variables se pueden aplicar modelos de corrección del error. Posteriormente, se estudian los cambios estructurales que puedan existir en las series consideradas. Al ser este tipo de mercados muy complejos y con frecuentes cambios estructurales, es importante este paso para identificar el número óptimo de cambios estructurales que existen en estas series para tenerlos en cuenta en las predicciones.

Thus, the first statistical analysis that is carried out as an initial step of the proposed methodology is to study the correlation and cointegration relationships that may exist between the commodities that are considered. It is interesting to know which commodities present significant and stable correlations. In this context, stable correlations imply they do not depend on the length of the calibration window used. In order to determine this, the existing correlation is studied using different lengths of the calibration windows. Subsequently, it is analyzed which of the significant correlation relationships found give rise to true long-term cointegration relationships, through the implementation of statistical cointegration tests. This analysis will be used to determine for which set of commodities error correction models can be developed. On the other hand, for those commodities that are not cointegrated, other types of quantitative models and forecasting techniques that are based on the past values of the variable itself are proposed for adjusting them.

Once the existing significant statistical relationships have been identified, the quantitative models are built and statistically validated. Then, the logic of the trading algorithm is developed. This logic is based on the outputs that the previously developed forecasting techniques provide. If the prediction in the following instant is greater than the current value of the series, the conclusion is that the quantitative model predicts that the price will rise on the next day (in the case of dealing with daily prices). In this case, the algorithm makes the decision to buy a unit in the market today and sell it the next day, when it predicts that the price will be higher and therefore an economic benefit of this transaction would be obtained. On the contrary, if the prediction in the future instant is lower than the price at the current instant, the algorithm will make the decision to sell a unit today and buy at the moment in which the price is predicted to be cheaper.

The tool that is developed, in addition, is an automated tool that allows the user to enter as inputs a set of variables to be decided. Among these inputs, the user can choose the level of risk he wishes to assume in his trading decisions. This functionality is important, as it allows an energy company to decide the level of risk assumed based on its policies. However, a proposed methodology is also included in this project as a step to follow to find an optimal value for the inputs that the tool allows the user to decide, notwithstanding that the criteria proposed in this methodology can be followed or not depending on the decision made by the user.

After developing the tool, a comparison is made between the forecasting techniques that are usually used in the literature for trading strategies in other types of markets. Validations and simulations are carried out with real commodity data from several years. Based on these simulations, the financial performance of trading strategies is evaluated through financial ratios commonly used in these applications in the literature. In this comparison, it is concluded that error correction models are very powerful for trading applications in energy markets by capturing the joint dynamics of the series, giving rise to excellent financial returns in the simulations.

Based on the previous comparison, a real implementation of the tool for automatic trading strategies based on error correction models is finally carried out. This implementation is carried out for two oil indexes that turn out to be cointegrated after executing the statistical tests: BRENT and WTI. The strategy is implemented with daily spot prices from 2001 to 2019 and an initial investment slightly higher than $40. The profits obtained with respect to the evolution of market prices for that period is higher than $250 with an annual mean financial return of 13,6%.

Overall, this project aims to fulfil the dual objective of serving as a proposal for a generic methodology for the development of automatic trading strategies in energy markets and, at the same time, to develop a tool and provide a potential user an illustrative real simulation as a demonstration of what it can achieved. Thus, this methodology is proposed for developing a tool that aims to improve the decision-making process of energy companies in this type of markets, and that bases this type of decisions on the statistical analyses performed with quantitative models and forecasting techniques.

In addition, some guidelines are mentioned as future work to complement the proposed models. Among them, a technique that is expected to have great applications in the future is mentioned: sentiment analysis. This technique is based on classifying financial news to complement the market predictions of the models used in this project through artificial intelligence techniques.

DISEÑO DE MÉTODOS CUANTITATIVOS PARA EL DESARROLLO DE ESTRATEGIAS AUTOMÁTICAS DE TRADING EN MERCADOS ENERGÉTICOS

Autor: Pretel Parejo-Merino, Carlos Jesús

Director: Bello Morales, Antonio

Co-Director: Reneses Guillén, Javier

Entidad Colaboradora: ICAI – Universidad Pontificia Comillas

Los mercados energéticos son complejos y difíciles de predecir. En especial, los mercados de commodities suelen presentar numerosos cambios estructurales y son afectados por factores políticos, estrategias internacionales, etc. Este tipo de mercados, entendiendo por commodity aquella que se caracteriza por ser comercializada en mercados y mundiales y por la existencia de un bajo grado de diferenciación entre productores (en energía, ejemplos de commodities son el petróleo, el gas natural o el carbón), tienen además un gran impacto en la volatilidad e incertidumbre de, por ejemplo, el mercado eléctrico. Es por ello que para las compañías energéticas que necesiten adquirir este tipo de productos para la generación de electricidad es fundamental elaborar estrategias de inversión en estos mercados, así como para otros agentes o compañías que operen en este tipo de mercados.

De este modo, los modelos cuantitativos de predicción son adecuados para el desarrollo de herramientas que lleven a cabo estrategias automáticas de trading. Sin embargo, la complejidad que estos mercados presentan provoca que este tipo de herramientas se desarrollen comúnmente para su aplicación en mercados de valores y otros, pero no es sencillo su desarrollo en relación a los mercados energéticos. Esta es una de las novedades de este proyecto, y a su vez una de sus mayores dificultades.

En referencia a los conceptos que aparecen en el título del proyecto, es importante definir los principales que aparecen para entender la metodología que se propone. En primer lugar, podemos definir una estrategia de trading como aquellos planes y actuaciones que se realizan para conseguir un determinado beneficio (en este caso, económico). Estas estrategias se pretenden basar en modelos cuantitativos, entendiendo por esto modelos matemáticos que se utilizan para ajustar las variables y establecer predicciones en base al ajuste previo de los datos. Por último, todo ello se implementará en mercados energéticos de commodities, cuya característica principal ha sido definida anteriormente.

Entre las técnicas más habituales de predicción que se pueden hallar en la literatura, existen numerosos criterios para clasificarlas. En este proyecto, se clasifican en función del horizonte de predicción en el cual suelen ser utilizadas y dan lugar a predicciones razonablemente precisas. Así, entre las técnicas de predicción a corto plazo

se pueden mencionar fundamentalmente aquellas que están basadas en modelos de inteligencia artificial y los modelos de series temporales. En cuanto a los primeros, son modelos que tiene la capacidad de interpretar correctamente los datos, aprender de estos y utilizar este aprendizaje para conseguir un buen ajuste. Por su parte, el análisis de series temporales trata de identificar patrones en datos históricos de una variable para aquellas que pueden presentar autocorrelación con respecto a los datos de instantes anteriores, o en aquellos casos en los que no es sencillo recurrir a modelos que tengan en cuenta variables explicativas.

Por otra parte, entre las técnicas de predicción habitualmente empleadas para un horizonte a más largo plazo es necesario destacar la importancia de los modelos de corrección del error. Estos se basan en la existencia de una relación estadística entre dos o más variables: la cointegración, ya que este tipo de modelos no se pueden desarrollar para variables que no estén cointegradas. La cointegración es una relación estadística más potente que la correlación. El análisis de cointegración consiste en identificar para aquellas variables que estén correladas en un periodo de tiempo considerado que esta relación es una relación estable y verdaderamente causal a largo plazo. Para comprobar esto, se emplean tests estadísticos que se denominan tests de cointegración. Una vez que se prueba la existencia de cointegración entre variables, los modelos de corrección del error son útiles para ajustar estas variables y capturar además la dinámica conjunta teniendo en cuenta la relación de equilibrio identificada a largo plazo.

A efectos de establecer una metodología genérica para el desarrollo de técnicas cuantitativas avanzadas para construir una herramienta de ejecución de estrategias de trading de manera automática y en especial en relación a su aplicación en mercados de commodities de energía, se proponen una serie de pasos. En primer lugar, se estudia la relación estadística entre aquellas commodities que se consideran para la implementación de estrategias de trading. En este primer paso se estudia si las variables presentan relaciones de correlación significativas y también se estudia la cointegración, a efectos de determinar para qué variables se pueden aplicar modelos de corrección del error. Posteriormente, se estudian los cambios estructurales que puedan existir en las series consideradas. Al ser este tipo de mercados muy complejos y con frecuentes cambios estructurales, es importante este paso para identificar el número óptimo de cambios estructurales que existen en estas series para tenerlos en cuenta en las predicciones. Seguidamente se desarrollan los modelos cuantitativos para ajustar los datos de entrenamiento. Estos modelos se validan luego estadísticamente para comprobar que los ajustes a los que dan lugar son precisos y que también las predicciones están razonablemente cercanas en los valores reales en el horizonte de validación que se considere, a efectos de evitar caer en un sobreajuste de los modelos. Tras esto, se selecciona el tipo de datos sobre los que se desea operar y el tipo de modelo sobre el que basar la herramienta automática de trading. En este sentido, se considera la posibilidad de desarrollar las estrategias para precios mensuales o diarios, en función de si la carga computacional que supone un mayor número de datos de precios diarios puede asumirse. Por último, se implementa la estrategia de trading de manera automática en base a estos modelos cuantitativos previamente desarrollados.

Una vez propuesta esta metodología, se implementa para el caso particular en el que se centra este proyecto. Utilizando datos reales de los principales mercados de commodities en el contexto internacional, tanto de petróleo como de gas natural, se desea desarrollar los pasos enumerados previamente y realizar una simulación real a modo ilustrativo de lo que se puede lograr con la herramienta que se desarrolla en este documento.

Así, el primer análisis estadístico que se realiza como paso inicial de la metodología propuesta es estudiar las relaciones de correlación y cointegración que pueden existir entre las commodities que se consideran. Es interesante conocer en primer lugar qué commodities presentan correlaciones significativas y estables, es decir, que no dependen de la longitud de la ventana de calibración que se emplee. Para ello, se estudia la correlación existente utilizando diferentes longitudes de las ventanas de calibración. Posteriormente, se analiza cuáles de las relaciones significativas de correlación que se encuentran dan lugar a relaciones verdaderas de cointegración a largo plazo, mediante la implementación de tests estadísticos de cointegración. Este análisis servirá para determinar para qué conjunto de commodities podrán desarrollarse modelos de corrección del error. Por su parte, para aquellas commodities que no resultan cointegradas se propone ajustarlas a través de otro tipo de modelos cuantitativos y técnicas de predicción que se basan en los valores pasados de la propia variable.

Una vez se han identificado las relaciones estadísticas significativas existentes, se construyen los modelos cuantitativos y se validan estadísticamente. Seguidamente, se desarrolla la lógica del algoritmo de trading. Esta lógica se basa en las salidas que proporcionan las técnicas de predicción previamente desarrolladas. Si la predicción en el instante posterior al actual es mayor que el valor actual de la serie, la conclusión es que el modelo cuantitativo predice que el precio subirá en el día siguiente (en el caso de tratar con precios diarios). En este caso, el algoritmo toma la decisión de comprar una unidad en el mercado en el instante actual y venderla al día siguiente, cuando predice que el precio va a ser mayor y por lo tanto se obtendría un beneficio económico de esta transacción. Por el contrario, si la predicción en el instante futuro es inferior al precio en el instante actual, el algoritmo tomará la decisión de vender una unidad en el instante actual y comprar en el instante en el cual se predice que el precio es más barato.

La herramienta que se desarrolla, además, es una herramienta automatizada que permite al usuario introducir como entradas una serie de variables a decidir. Entre estas entradas, el usuario puede elegir el nivel de riesgo que desea asumir en sus decisiones de trading. Esta funcionalidad es importante, ya que permite a una compañía energética decidir en función de sus políticas el nivel de riesgo asumido. No obstante, se incluye también en este proyecto una metodología propuesta a modo de pasos a seguir para encontrar un valor óptimo para las entradas que la herramienta permite decidir al usuario, sin perjuicio de que se pueda seguir el criterio propuesto en esta metodología o diferir de él con vistas a su utilización por una compañía energética.

Tras desarrollar la herramienta, se realiza una comparación entre las técnicas de predicción que habitualmente se utilizan en la literatura para estrategias de trading en otro tipo de mercados. Las validaciones y simulaciones se realizan con datos reales de commodities con datos de varios años. En base a estas simulaciones, se evalúa el rendimiento financiero de las estrategias de trading a través de ratios financieros comúnmente utilizados en estas aplicaciones en la literatura. En esta comparativa, se concluye que los modelos de corrección del error son muy potentes para las aplicaciones de trading en mercados energéticos al capturar la dinámica conjunta de las series, dando lugar en las simulaciones a excelentes rendimientos financieros.

En base a la comparativa anterior, se realiza finalmente una implementación real de la herramienta para estrategias automáticas de trading basadas en modelos de corrección del error. Esta implementación se realiza para dos índices de petróleo que resultan estar cointegrados tras ejecutar los tests estadísticos: BRENT y WTI. La estrategia se implementa para datos de precios diarios reales desde 2001 hasta 2019, con una inversión inicial cercana a las 40 dólares y un beneficio con respecto al mercado superior a 250 dólares para dicho periodo, obteniéndose un retorno financiero medio anual del 13,6%.

En definitiva, se pretende en este proyecto cumplir el doble objetivo de servir como propuesta de una metodología genérica para el desarrollo de estrategias automáticas de trading en mercados energéticos y, al mismo tiempo, desarrollar una herramienta y proporcionar a un potencial usuario de la misma una simulación real a modo de demostración de lo que esta puede conseguir. Así, se propone el desarrollo de una herramienta que pretende mejorar el proceso de decisión de compañías energéticas en este tipo de mercados, y que basa este tipo de decisiones en los análisis estadísticos realizados con modelos cuantitativos y técnicas de predicción.

Asimismo, se mencionan algunas líneas de actuación futuras para complementar los modelos que se proponen. Entre ellas, se menciona una técnica que se prevé que tenga grandes aplicaciones en el futuro: el análisis de sentimiento. Esta técnica se basa en clasificar noticias financieras para complementar las predicciones de mercado de los modelos empleados en este proyecto por medio de técnicas de inteligencia artificial.

Keywords

Commodity, energy markets, quantitative models, foreast, statistics.

Palabras clave

Commodity, mercados energéticos, modelos cuantitativos, predicción, estadística.

Contents

List of Figures

List of Tables

1 Introduction

Energy is an essential input in the production of goods and services. A proof of it is the fact that electricity is considered as an essential good, not only for supplying customers with light and warmth, but it is a basic element for the industry too. On the other hand, fossil fuels are present in our daily lives, not only due to the transportation sector but also creating materials such as computer, plastics or medicines. Because of it, security of supply and reliability associated with both electricity and fossil fuels are essential not only for developed countries, but also for ensuring economic growth in developing regions.

The importance of energy sources and power management strategies will continue being the topic of conversation for many decades in fields such as politics and economics. There is not a clearly predictable future concerning the evolution of global energy consumption. Instead, main energy companies frequently analyse several scenarios as it is uncertain how the countries with highest energy demand will behave with respect to promotion of renewable energy sources and energy transition.

There are widely known advantages of renewable energy in terms of reduction of greenhouse emissions and their alignment with policies against climate change. Nevertheless, problems such as intermittency and energy storage often require technological solutions that are still under research. As a consequence, transportation and electricity sectors are highly dependent on non-renewable sources globally nowadays. Electric vehicles and batteries have a huge potential in the future of the transportation sector, but there is still a lot of progress to be achieved in terms of increasing their autonomy.

Talking about energy sources, we need to distinguish between renewable and nonrenewable resources. A nonrenewable resource is a natural substance that is not replenished with the speed at which it is consumed [1]. Oil, coal and natural gas are examples of nonrenewable energy sources. Meanwhile, renewable resources are infinite resources and are replenished naturally: for example, the sun light and the wind.

Another relevant concept is the definition of commodities. Their basic characteristic is that there is not a significant difference between a commodity supplied by one producer and the same commodity coming from another producer. Commodities need to meet some minimum standards and they are normally traded in spot markets. Oil, coal and natural gas are traditional examples of commodities in the energy sector. However, they are not the only type of commodities as there are agricultural products and metal ones such as gold or silver that accomplish with the requirements for being a commodity too. Nevertheless, the technical part of this project will focus only on energy commodities.

Global primary energy consumption has increased during last years. In 2018, global primary energy consumption was 2.9% higher than in the previous year (the fastest growth since 2010). The main contributors to this increase were in 2018 China, USA and India.

According to the report published by BP *2018 Statistical Review of World Energy*, carbon emissions were higher too (grew at a rate of 2%). Consumption of all fuels increased, especially natural gas, which contributed more than 40% of the increase.

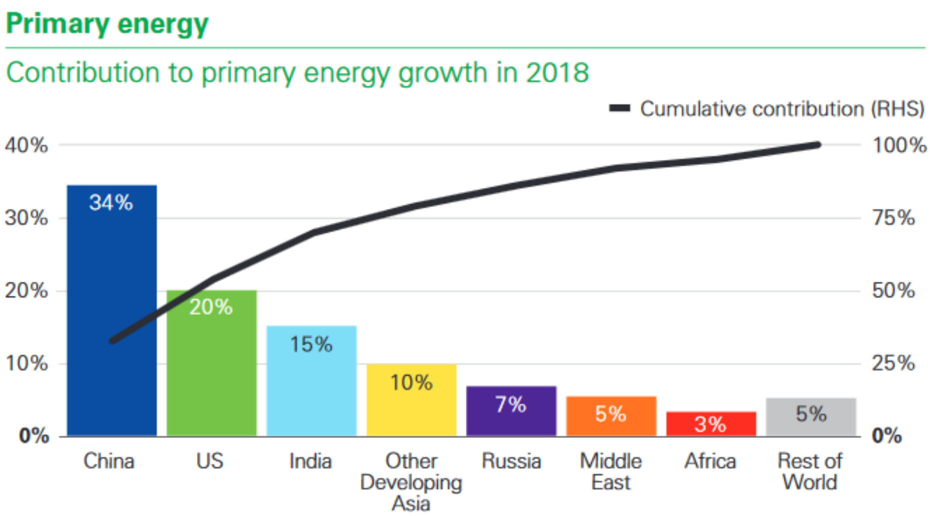

Figure 1.1: *Contribution to primary energy growth in 2018 [2]*

Oil, coal and natural gas continue being the main primary energy sources. In addition, concerning the expected evolution of primary energy consumption published by the International Energy Agency (IEA) in its annual report *World Energy Outlook 2018*, this composition of global primary energy consumption will continue in the next decades.

Overall, commodities markets will continue being the main primary energy markets in the future in terms of the amount of primary energy involved in transactions. However, prices in commodity markets are affected by many factors, such as agents' decisions, stationality or fuel prices. All of them together with the uncertainty caused by energy transition and the increase of renewable production make very likely future prices to be more volatile. As prices in power markets are highly influenced by prices of commodities, developing trading strategies related to them is essential for utilities and energy companies nowadays.

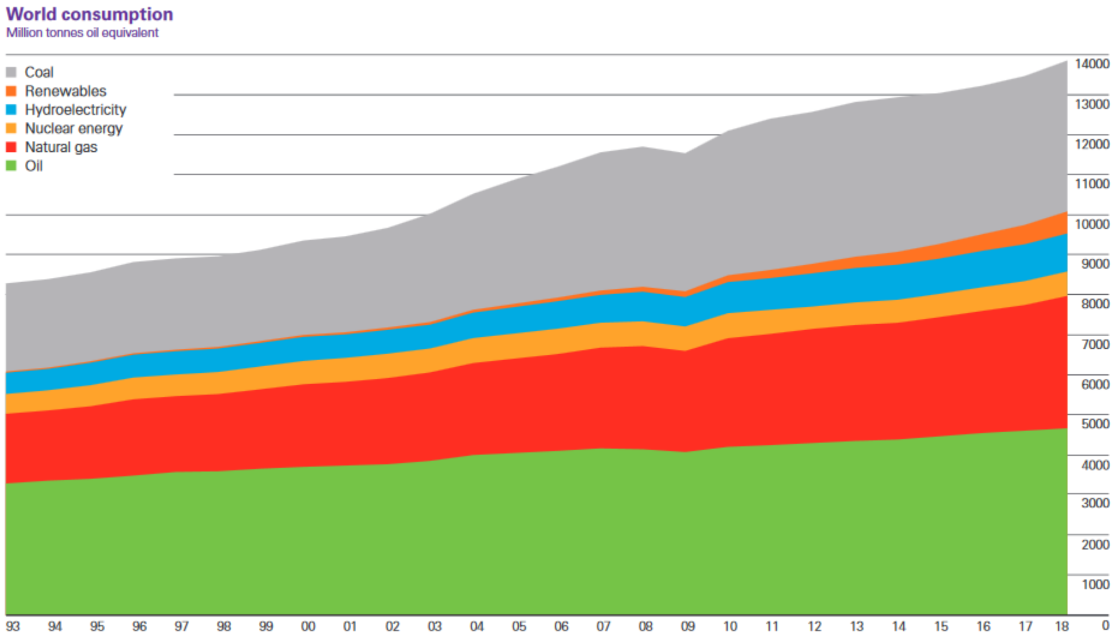

Figure 1.2: *World consumption [2]*

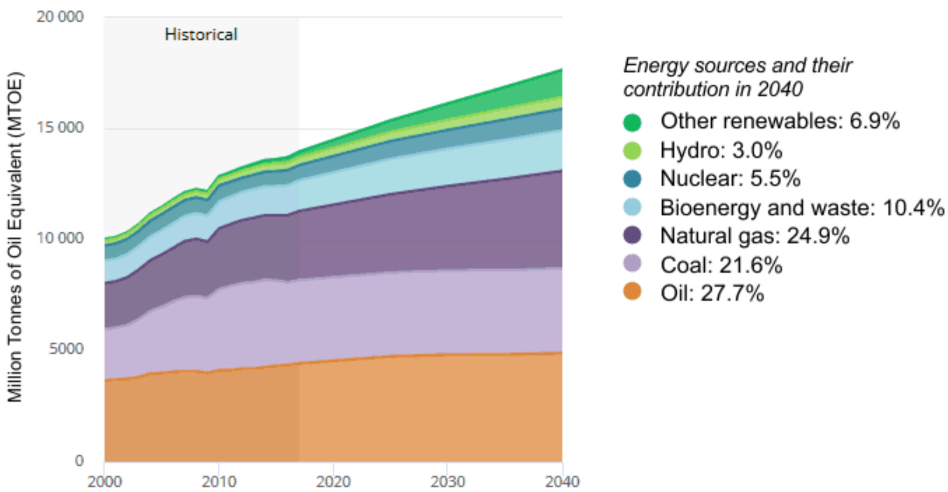

Figure 1.3: *Global primary energy demand by 2040 [3]*

However, the way commodities prices will evolve in the future is not easily predictable. They can be affected by externalities such as political decisions and economic variables together with the global demand of commodities. This is translated into risk and uncertainty for utilities.

To illustrate how externalities can affect commodities prices and how complex energy markets can be, the evolution of WTI Crude Oil Spot Price during the last 35 years is shown in Figure 1.4 [4].

Up to 2008, prices tended to go up due to the exponential increase of demand specially in emerging economies such as China and India. However, after this year oil prices decreased as a consequence of the global economic crisis that took place after 2008 and the decrease in global demand. Once this decrease ended oil prices tended to increase again as global demand started to rise again and emerging economies recovered their rate of growth too. Then, in 2014 oil prices decreased significantly again. This time it happenned because unconventional production in United States hugely increased and South Arabia managed their reserves to affect oil prices and make them decrease with the purpose of reducing the income for unconventional producers and trying to make them bankrupt. All these factors affected the evolution of oil prices and gave rise to structural changes that can even break the relationship between oil and natural gas prices, which reflects how complex and volatile these markets can be depending on political decisions.

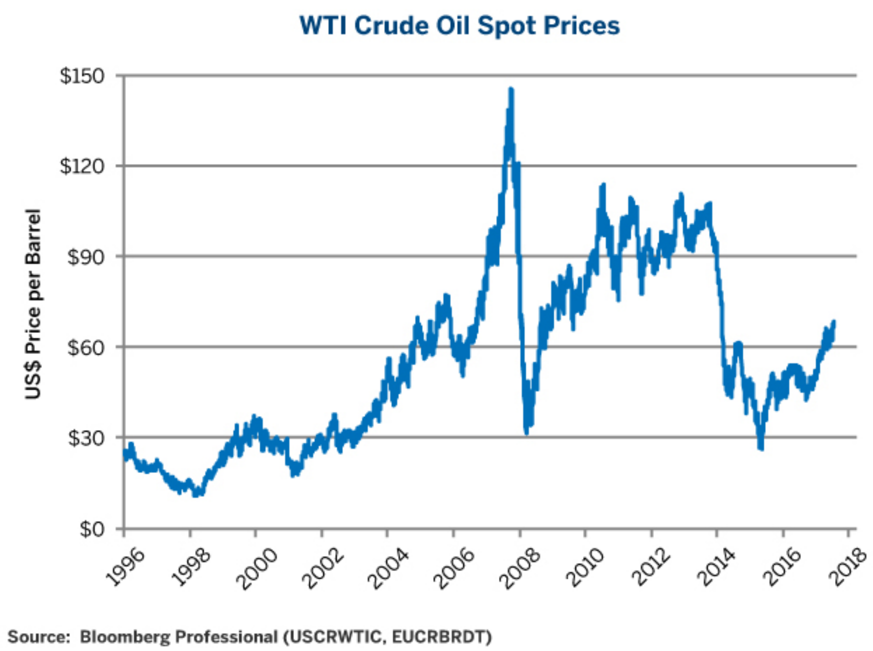

Figure 1.4: *Evolution of WTI Crude Oil Spot Price during the last 22 years*

It can be deduced from those events that there are a lot of aspects of very different nature such as international politics, strategy adopted by countries, etc. that can significantly affect prices of commodities like oil. As a result, commodities are variables which are extremely difficult to forecast. This is translated into risk associated with trading decisions made by utilities.

To mitigate this risk and improve decision-making processes, complex and advanced models are required to capture the dynamics of the evolution of these prices. Achieving this would be very useful for utilities as these models can have lots of applications. The one in which this project will focus on is developing automatic trading strategies.

Companies usually need to take decisions related to trading strategies and risk management. One type of these strategies is based on using financial instruments to reduce the risk associated with market exposure. These tools are a way for utilities to hedge their position and protect themselves against volatility in energy markets. For example, some of the most relevant financial instruments used in energy markets are forward contracts and options. These financial instruments are called derivatives and consist in contracts that establish a right or an obligation to the parties to execute a transaction in the future at an agreed price.

These mechanisms allow companies to hedge their positions and be less exposed to volatility in markets in exchange of assuming the risk of losing the premium paid to the seller of the option. However, these instruments are just tools for mitigating this exposure, but they do not eliminate it without using robust, reliable and accurate forecasting methods to support companies in making trading decisions. This is how building quantitative models for developing trading strategies in energy markets could help utilities to make decisions following a statistical criteria. In addition, tools that can be implemented in an automatic way increase time response and become more reliable as they do not take decisions depending on human psychology.

The aim of this project is to define a methodology to build accurate and reliable models that could be potentially used by utilities for developing automatic trading strategies. However, the complexity of capturing the behaviour of energy markets is not the only difficulty that needs to be faced in this project. There are multiple decisions required to be made to develop a robust forecasting model.

Although a great amount of forecasting methods have been developed along last decades, the decision of choosing the most accurate for a specific application is highly dependent on how the model will be applied in practice. For example, a critical aspect to be considered is the prediction horizon.

The majority of the existing models obtain good results only in the short term. However, there is a lower amount of already developed models robust enough to obtain accurate results in long prediction horizons. In this document, in those cases in which hourly or daily data is available forecasting techniques will be considered as **short-term** techniques if they have a prediction horizon of several hours or days. Meanwhile, prediction horizons of several months or years would mean **medium-term** or **long-term** predictions.

Concerning long-term forecasting mtehods, they became relevant for applications related to studying statistical relationships in the long run. These methods had the objective of distinguishing between spurious correlations and long-term true relationships. Regarding commodities markets, studying relationships between variables in the long run makes sense. Identiifying related variables in the long run can help to build robust explanatory models. In addition, these techniques can be especially relevant for energy companies which are exposed to several international markets not only for capturing the dynamics of each market individually, but especially for finding the potential dependency relationships between them.

As it has been exposed above, there are some variables that have been traditionally correlated. However, some of them are changing over time too. For example, the relationship between the evolution of oil and gas prices is becoming lower recently. As a result, the proportion of gas contracts indexed to oil prices has been decreasing since some years ago.

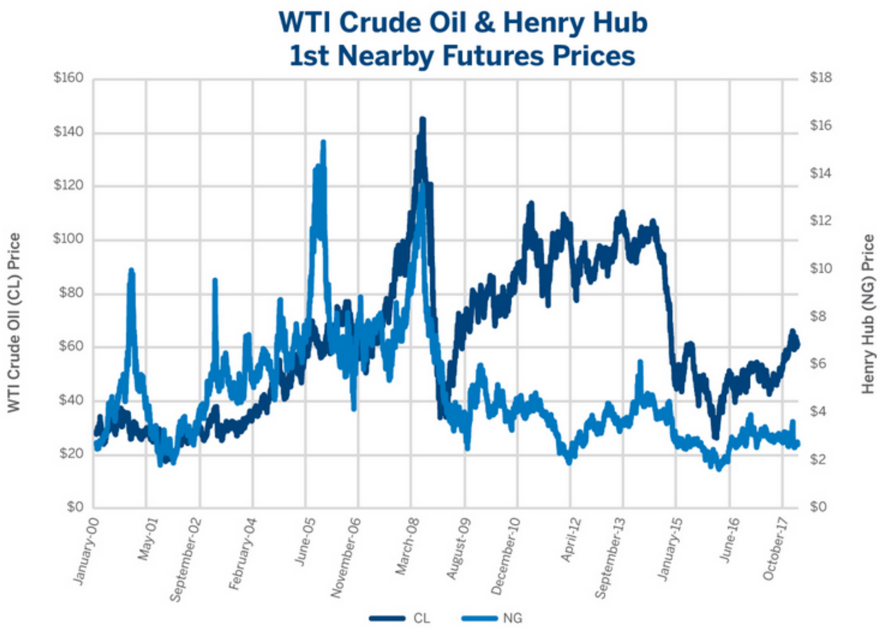

Figure 1.5: *Evolution of the relationship between oil and natural gas prices [5]*

Therefore, statistical analysis able to prove those relationships that are not spurios are necessary to ensurethat only explanatory variables are considered when building mathematical models.

However, most of the references with existing quantitative models developed with trading purposes are focused on other financial markets different than commodities markets. As it has been explained, this is mainly because of the high amount of external factors that can affect their prices and which make it a very challenging task to estimate them only looking at what happened in the past. As a consequence, there are not many examples in which mathematical models are applied in the literature to develop automatic trading strategies. The few examples that exist are normally based on fundamental models that try to predict the behaviour of markets considering only changes in general economic indicators such as demand, interest rates... Because of this reason, the development of complex statistical forecasting models in this project truly represents a big progress with respect to the way trading strategies are carried out in most cases.

The objective of this project is not only to provide trustful models to forecast prices in energy markets in several prediction horizons, but also to complement these models with improvements such as studying structural changes in time series. The reason for this is that it is necessary to anticipate external factors before they take place that can affect commodities prices and that cannot be anticipated only focusing on historical data.

Overall, the technical scope of this project is to develop real quantitative models that can be validated in terms of trading strategies in energy markets. In order to do so, the following additional objectives will be covered in the different sections of this document:

- To develop a **state-of-the-art review** concerning the most relevant techniques to model the behaviour of energy markets that are included in research papers and also which trading strategies have been actually validated in the literature with the aim of developing trading strategies.
- To expose the **current trends** and the most **innovative techniques** that either have been recently developed or are currently under research in the development of automatic trading strategies in energy markets.
- To study **statistical relationships** between commodities prices in energy markets.
- To write a detailed **methodology** with all the steps followed in this project to create, program, train and test quantiative models considering all the difficulties related to using data from real energy markets.
- To propose and implement **quantitative forecasting techniques** for capturing complex dynamics in energy markets.
- To build an **automatic tool** from which the final user can choose the level of risk to be assumed in automatic trading decisions.
- To illustrate the **results of the models** built in this project and simulate its **financial performance in international energy markets** using real and recent data.

- To analyse the results obtained and determine if the models and strategies implemented in this project could be helpful for an utility to carry out trading strategies.
- To determine which ones of the current trends in energy markets exposed in the state-of-the-art review could be most relevant to be used to complement the models in the future.

To fulfil all the objectives described above, the state-of-the-art review will be covered in chapter 2 of this document. The main forecasting techniques used in time series analysis will be presented in sections 2.1 and 2.2, distinguishing between short-term forecasting techniques (included in section 2.1) and those used to predict in the long term (included in section 2.2). Following it, section 2.3 is about the current trends and recent innovative techniques with potential applications in energy markets. The methodology followed to develop, program, train and test the models implemented in the project will be described in section 3.

The development of the models is described in sections 4 and 5. Concerning section 4, statistical analysis are provided using monthly data from energy markets. On the other hand, models proposed in section 5 are trained using daily data. The different techniques explained along the previous sections of the documents are compared and the suitability for being used to develop automatic trading strategies in energy markets is discussed.

2 State of the art

A review of the most relevant forecasting techniques concerning times series analysis is provided in this chapter. As the aim of this project is to develop automatic trading strategies by building forecasting quantitative models, an overview of the different techniques that are commonly used to model time series is necessary. Once these techniques are introduced, a further analysis about their advantages and disadvantages concerning their specific application to fit financial series from energy markets will be carried out.

This chapter aims to provide a study of the main characteristics of principal forecasting methods that are suitable for developing automatic trading strategies in energy markets. In addition, the study is also based on which methods have been traditionally used for forecasting purposes in both financial and energy markets. Most commonly used short-term forecasting methods are neural network regression and some univariate models, while in the long term cointegration and error correction methods have been proved to be successful in forecasting economic variables such as inflation or rates of exchange.

Therefore, a theoretical overview is provided regarding a lot of different forecasting techniques that have been traditionally used for times series analysis. Examples in which some of them have been used in energy markets will be provided too.

Technically speaking, a time series based on a stochastic process is a sequence of chronologically equidistant random variables, referred to one or several characteristics of an observed unit at different moments [6]. Lots of methods for forecasting time series have been researched by many authors due to the importance of anticipating the future in financial decisions.

Forecasting quantitative methods can be defined as statistical techniques used for making predictions about the future which uses numerical measures and prior effects to predict future events [7]. These techniques are based on mathematical models and statistical analysis.

Although there are lots of different ways to classify forecasting methods, a criteria that will be used in this chapter is the prediction horizon. As it has been introduced above, forecasting techniques will be considered as **short-term** techniques if they have a prediction horizon of several hours or days. Prediction horizons of several months or years would correspond to medium-term or long-term predictions.

Both univariate and multivariate models will be divided into short-term or long-term forecasting methods. Univariate models try to predict future values of the series considering just historical data from the forecasted series. On the other hand, multivariate models predict it depending on other variables. They try to identify causal relationships between the forecasted series and explanatory variables.

The suitability of each technique needs to be analysed independently for each specific case, as the decision of using one technique or another is complex and depends on a lot of aspects such as the statistical behaviour of the forecasted series. In addition, the way these models will be applied by the final user in practice should also be considered when studying their advantages and disadvantages. For example, some techniques are more accurate than others for short-term forecasting purposes. The same happens if the user wants to develop models to predict series in the long term. All these aspects increase the complexity the decision of choosing one technique instead of another has.

This chapter includes in its first two subsections a review of the main short-term and long-term forecasting techniques that have been developed based on statistical models, as well as some real applications that can be found in the literature. Sometimes, hybrid models are built by combining some of these techniques. The objective of these hybrid models is to combine the advantages of each technique and obtain a final model with better characteristics than simple techniques individually.

Finally, an overview of some other techniques that are currently under research will also be provided. These techniques will be presented as future alternatives to improve the model that is proposed in this project.

2.1 Short-term forecasting techniques

Short-term forecasting techniques will be presented in this subsection. Among all of them, the fundamentals of the most relevant ones concerning trading strategies applications in energy markets will be provided.

Short-term forecasting techniques can be classified according to different criteria. For example, univariate models can be distinguished from multivariate models. The first ones are built just looking at historical values while multivariate models include different explanatory variables. In this section, they will be divided in artificial intelligence and time series analysis.

2.1.1 Artificial intelligence

Artificial Intelligence (AI) can be defined as a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation [8]. They are also called Machine Learning methods.

Investigations related to machine learning techniques have increased significantly during last years. This can be observed in the increasing number of papers that represent new opportunities explored shown in the following graph. In addition, many of these techniques are explored in energy areas.

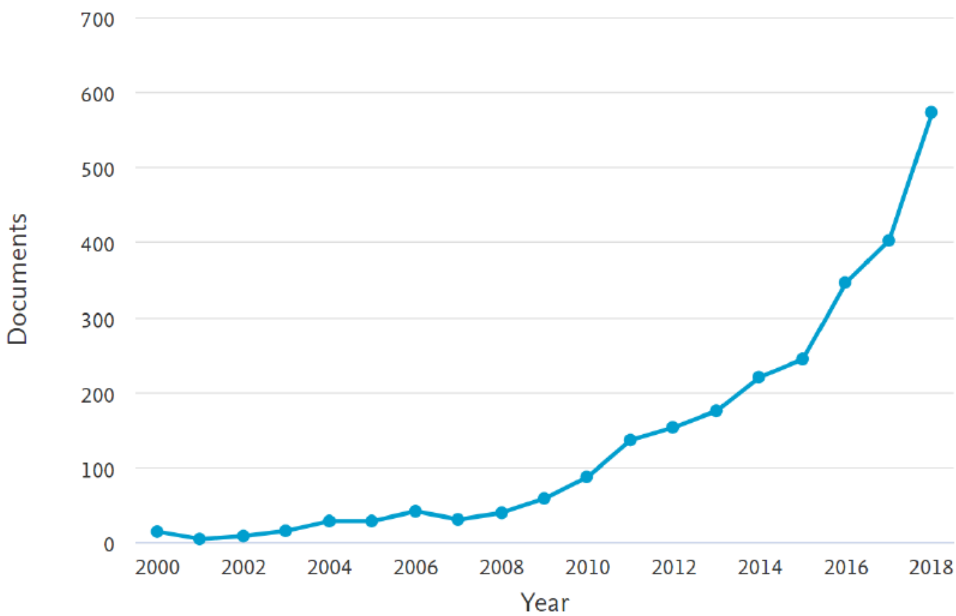

Figure 2.1: *The growth in the number of articles during the past two decades [9]*

Neural networks constitute one of the most popular techniques based on artificial intelligence with forecasting applications in energy markets. Because of this, most of the techniques described in this section will be based on neural networks. In addition, advanced versions of artificial neural networks and hybrid models will also be presented.

Artificial neural networks

In this context, Artificial Neural Networks (ANN) try to predict the existing relationship between the objective variable and those input variables. To achieve this, ANN predict the response of an objective variable through neurons that transform input data into output values through a transfer function. These transfer functions are normally

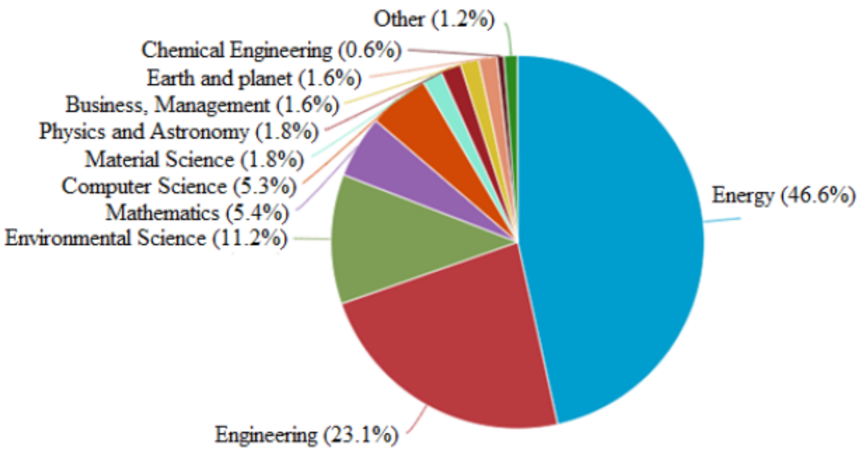

Figure 2.2: *Subject areas of ML in energy systems*

non-linear functions such as the density function of multivariate normal distributions or sigmoid functions. In the training process, weights associated with input data are modified.

ANN whose structure consists of more than one layer of neurons are called Multilayer Perceptrons (MLP). In this case, a necessary steps before training the network is deciding the number of layers. This should be done according to the complexity of the model. The more complex the relationship between input data and the objective variable is the higher number of layers it is required in the structure of the network.

The initial weights are chosen randomly. Then, the iterative learning process is carried out through modifying weights associated to input data according to the error function between the response obtained with current weights and the desired response for the objective variable multiplied by a certain factor. As the initial weights can be randomly chosen, several simulations should be developed in order to create box plots of the different responses obtained after the network training [10].

Advantages of ANN and MLP include high tolerance to the existence of noise in data. Furthermore, these models are able to capture almost any dynamic even with very noisy data. They are good at capturing nonlinearities too. However, ANN lose accuracy when increasing the prediction horizon and overfitting problems are frequent too.

Nevertheless, neural networks do not always learn successfully. This can happen when the input data do not contain necessary information from which the desired output comes or when available data is not enough for completing successfully the training process. There can be some cases in which the structure of network is not properly identified too.

The efficiency of ANN based methods is highly dependent on appropriate tuning of their adjustable parameters, such as the number of hidden layers, nodes, weights and transfer functions. Most of ANN based methods use gradient-based learning algorithms, such as back propagation neural network.

Some of the problems related to ANN models are over-tuning and long computation time. These disadvantages gave rise to the development of advanced versions of ANN.

Recently, a new learning algorithm for single-hidden-layer feedforward neural networks called extreme learning machine (ELM) has been proposed. In ELM, input weights and hidden biases are randomly introduced and output weights are calculated analytically using the Moore-Penrose generalized inverse. Higher learning speed is achieved by using ELM instead of traditional gradient-based learning algorithms, and difficulties such as stopping criteria and over-tuning problems are avoided.

Support Vector Machines

Support Vector Machines (SVM) consitute a machine learning methodology which is widely used for text classification purposes. Given a training dataset, each point belonging to one of two or more categories, an SVM training algorithm builds a model to assign new data to one category or the other. This is achieved by learning linear decision rules given by the following formulation of h, dependent on a weight vector and a threshold b.

$$h(\vec{x}) = sign\{\vec{w} \cdot \vec{x} + b\} \tag{2.1}$$

Where h represents a vector of categories to which points are assigned, $\vec{x}$ is the input data point to be classified, $\vec{w}$ is a weight vector that determines the importance of each element of the input data point and b corresponds to a threshold.

The SVM algorithm can be formulated as an optimization problem [11]. Input is a sample of n training examples, S_n, for which the SVM methodology tries to find the hyperplane with maximum Euclidean distance to the closest training examples. Therefore, the overall optimization problem to be solved can be formulated as follows:

$$S_n = ((\vec{x}_1, y_1), ..., (\vec{x}_n, y_n)), \quad \vec{x}_i \epsilon \mathbb{R}^n, \quad y_i \epsilon \{-1, +1\} \tag{2.2}$$

$$minimize: \quad V(\vec{w}, b, \vec{\xi}) = \frac{1}{2}\vec{w} \cdot \vec{w} + C \sum_{i=1}^{n} \xi_i \tag{2.3}$$

$$subj.to: \quad \forall_{i=1}^{n} \quad y_i(\vec{w}\cdot\vec{x}+b) \geq 1-\xi_i \tag{2.4}$$

$$\forall_{i=1}^{n} \quad \xi_i \geq 0 \tag{2.5}$$

Note that all training examples are classified correctly up to an error represented by ξ_i. If a training example is not classified correctly, the corresponding ξ_i is greater or equal to 1. Meanwhile, the margin of the resulting hyperplane is represented by δ can be defined as:

$$\delta = \frac{1}{\|\vec{w}\|} \tag{2.6}$$

Taking this into consideration, the sum of both the margin of the hyperplane and the sum of all "errors" in classifying training examples are considered in the objective function that has to be minimized.

Figure 2.3 illustrates how the explained SVM algorithm should result in an optimum hyperplane for a binary classification problem. The examples closest to the hyperplane are called support vectors (marked with circles).

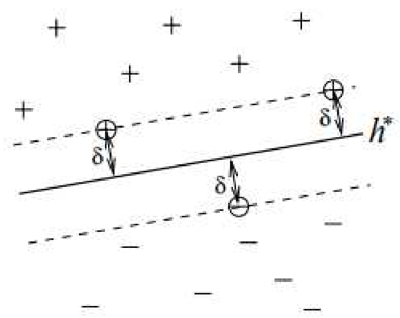

Figure 2.3: *Example of a binary classification problem (+ vs. -) in two dimensions using SVM algorithm*

Artificial Intelligence applications in commodity markets

Some relevant references with applications of artificial intelligence techniques explained in this section for capturing the dynamics of commodity markets are included in Table 2.1.

Table 2.1: *Artificial intelligence in commodity markets*

Year	Reference	Absract
1996	Kohzadi et al. [12]	ARIMA and artificial neural network price forecasting performances are compared in commodity markets.
2009	Kulkarni et al. [13]	This paper presents a model based on multilayer feedforward neural network to forecast crude oil spot price direction in the short-term, up to three days ahead.
2008	Yu et al. [14]	In this study, an empirical mode decomposition (EMD) based neural network ensemble learning paradigm is proposed for world crude oil spot price forecasting.
2019	Li et al. [15]	In this paper, hybrid models are proposed for monthly crude oil price forecasting using variational mode decomposition and artificial intelligence (AI) techniques (support vector machine optimized by genetic algorithm (GASVM) and back propagation neural network optimized by genetic algorithm (GABP) are employed for analyzing).
2018	Prasad Das et al. [16]	A hybrid method that combines a support vector machine (SVM) with teaching–learning-based optimization (TLBO) is developed for forecasting the commodities futures contract index.

2.1.2 Time series analysis

Time series analysis tries to identify patterns from a set of historic data of a random variable. Identifying these patterns through univariate models is suitable for situations in which historic data present autocorrelation, or when it is not easy to generate a model using explanatory variables.

There are several types of univariate time series models depending on patterns presented by historic data. In this context, structural models, exponential smooting and ARIMA techniques are some examples of these univariate models.

Structural models

Structural models are techniques that find time series patterns in the data, and then choose a method that is able to fit the patterns properly. Firstly, these patterns can be identified by decomposing the time series in three components: trend, seasonality and irregularity. These components can be considered in an additive or multiplicative way.

The first step would be to identify the three components. For example, a trend is commonly identified in a time series through simple mathematic models such as Minimum Squares Method when this trend is clearly observable. Then, in case of using an additive model the trend would be subtracted from the model and seasonality could be determined through several techniques. One of them is moving average. Then, a structural model can be built by multiplying or adding all the components.

Nevertheless, one of the main disadvantages moving average has is the fact that all the observations considered on the moving window are equally weighted, regardless of how recent they are. Instead, it is often reasonable to consider the more recent an observation is the more relevant information it contains. Exponential smoothing techniques are based on this principle.

Because of the high amount of external factors and uncertainties that characterize energy commodity markets, there are not many papers in the literature proposing this type of models for capturing the dynamics of these markets. Due to this high uncertainty, it is more frequent to find instead studies about energy commodity prices volatility.

Exponential smoothing

Simple exponential smoothing considers time series is composed of a weighted moving average and an irregular component impossible to be predicted. The weighted moving average gives higher weights to most recent data. This way the more recent an observation is the higher its associated weight is considered in the weighted moving average [17].

$$m_T = \alpha \sum_{j=0}^{T-1} (1-\alpha)^j Y_{T-j} \tag{2.7}$$

Weights depend on the value of parameter α, which needs to be lower or equal to 1. Modifying (2.7) a new expression can be reached in which the weighted moving average can be calculated depending only on the estimated weighted moving average from the last instant and the last observation. This gives rise to one of the main advantages of this method, as it is not required to store the values of all past observations.

$$m_T = \alpha Y_T + \alpha \sum_{j=0}^{t-1} (1-\alpha)^j Y_{t-j} = \alpha Y_T + \alpha(1-\alpha) \sum_{j=0}^{t-2} (1-\alpha)^j Y_{t-1-j}$$

$$m_t = \alpha Y_t + (1-\alpha) m_{t-1} \tag{2.8}$$

Adjusting the value of α between 0 and 1 changes the degree in which the series is smoothed. Higher values imply lots of variance in the smoothed series while low values of α give rise to a much more smoothed series. However, simple exponential smoothing is not effective concerning short-term forecasting of time series presenting an observable trend component, as this would require high values of α. As it can be deduced from the expression shown above, the closer α is to 1 new prediction the closer predictions are to the value adopted by the previous observation.

On the other hand, Holt-Winters method is based on exponential smoothing. This technique includes a trend component in the prediction function. In addition to the estimation of the trend of the series, Holt-Winters method can be modified to include a seasonal component in the prediction function too.

Holt-Winters methodology is attractive in terms of less computational time required in comparison with ARIMA. However, the use of these models gives rise to precise forecasts only in time series with both strong trend and seasonality.

These techniques became popular due to simple model formulation and good forecasting results. Because of these reasons, they constitute an attractive alternative compared with ANN based models for integration in short-term forecasting hybrid models. They avoid some problems such as over-fitting issue and the requirement of a huge amount of data for training in order to obtain accurate enough predictions.

However, the simple formulation of these models triggers some disadvantages too. Complex models with lots of structural changes are not easy to be captured just looking at past values.

ARIMA and SARIMA models

Autoregressive Integrated Moving Average (ARIMA) models are widely used in forecasting models because of including both autoregressive and moving average terms.They usually give rise to precise short-term forecasting. These models inherit the good properties of ANN in terms of accuracy in short-term forecasting, but avoid disadvantages such as costing lots of calculating resources.

These models include three components: autoregressive, integration and moving average. However, the integration term is only required for time series that cannot be assumed to be stationary. ARMA models are designed for stationary processes. The methodology and formulation of these models will be described in the next chapter.

Autoregressive models are function of past values of the own time series considering the error as a white noise with a mean value of zero and constant variance. On the other hand, moving average models are linearly related to past errors and cannot be adjusted by Minimum Squared Method requiring non-linear iterative process, as past errors are not observable. The optimization of the orders of autoregressive (p) and moving average (q) terms in ARMA models can be made for instance attending to ACF[1] and PACF[2] following general Box and Jenkins methodology [18].

The integration term (d) is required for non-stationary processes. For determining the order of integration unit root tests are commonly carried out. Although lots of different types of unit root tests have been developed to study the stationarity of a given stochastic process the augmented Dickey-Fuller test (ADF) is widely spread for this purpose in time series analysis [19]. The theoretical fundamentals of these tests are described in the methodology chapter in this document.

The uses for which ARIMA(p,d,q) models become interesting in terms of short-term forecasting are considerably different. Concerning commodities, it is frequent to find ARIMA models applications for forecasting markets in the short term.For example, Rana Abdullah Ahmed and Ani Bin Shabri [20] proposed in 2014 ARIMA forecasting models for capturing the dynamics of crude oil price West Texas Intermediate (WTI). In this paper, the accuracy of forecasts obtained by ARIMA models are compared with the techniques of SVM and GARCH models[3], whose theoretical fundamentals will be described in the following subsection.

In addition, ARIMA models can also be extended to Seasonal Autoregressive Moving Average (SARIMA). A SARIMA(p,d,q)(P,D,Q)[m] adds three new hyperparameters to include the autoregression (P), differencing (D) and moving average (Q) of the seasonal component with a given period (m). These modesl allow the user to add a seasonal component to an ARIMA model without requiring to previously modify the time series by subtracting seasonality with an empirical decomposition. Therefore, these models are very useful for time series with a strong seasonality as finding the optimal values for p, d, q, P, D, Q and m that best fit the behaviour of the series would allow the user to build the model without previously decomposing the series.

[1]ACF: Autocorrelation Function
[2]PACF: Partial Autocorrelation Function
[3]GARCH: Generalized Autoregressive Conditional Heteroscedastic

However, one of the main disadvantages of both ARIMA ad improved ARIMA techniques is that they usually obtain accurate forecasting results for irregular and non-linear series only in short prediction horizons.

Autoregressive conditional heteroscedasticity

To understand the thoretical fundamentals of autoregressive conditional heteroscedasticity, it is relevant to consider the difference between conditional and unconditional variance. In statistics, while conditional variance is the variance given the value or values of one or more other random variables, unconditional variance can be defined as the squared estimated volatility of a random variable.

Time series representing prices of commodities are mainly characterised by high irregularity and non-linearity and high volatility too. None of the techniques mentioned above considered conditional volatility as relevant for modelling the series. However, Autoregressive conditional heteroscedastic (ARCH) models consider volatility changes over time depending on available information. They try to estimate it to improve forecasting results in the medium or long term [21] [22].

ARCH models appeared after econometrics realised that their ability to predict the future varied from one period to another. McNees suggested in 1979 that the inherent uncertainty or randomness associated with different forecast periods seems to vary widely over time, and that large and small errors tend to cluster together in contiguous time periods [23].

Financial markets were the main field of application for ARCH models. Portfolios of financial assets are held as functions of the expected means and variances of the rates of return. Any shifts in asset demand must be associated with changes in both expected means and variances of the rates of return. Formulation of ARCH(q) models proposed by Engel considered time-varying volatility [24].

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 \tag{2.9}$$

Where $\sigma_t{}^2$ represents time-varying volatility and y_t is the value of time series y at instant t. Parameters $\omega$and $\alpha$are to be determined to model the relationship between $\sigma_t{}^2$ and y_t.

Time series is modelled as a sequence of random variables, independent and with unit variance, multiplied by a factor representing volatility. Volatility is not constant and varies with time as a function of past squared values of the objective variable contained in the time interval between *t-q* and *t-1*.

GARCH(*p*,*q*) models were proposed by Bollerslev in 1986 as a generalization of ARCH(q) models. In this generalization volatility is considered as a function of both past squared values of the objective variable contained in the time interval between *t-q* and *t-1* and of past values of the own volatility in the time interval *t-p* and *t-1* [25].

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{2.10}$$

To ensure conditional volatility is always positive, all parameters ω, α_i and β_j must fulfil the condition of being greater than zero in Bollerslev's formulation.

Due to the existence of high volatility in time series representing prices of commodities, several types of ARCH models have been proposed for modelling and forecasting them with longer prediction horizons. Some of them are hybrid models that combine this technique with other univariate ones explained in this section such as ARIMA. Several relevant references that can be found in the literature concerning time series analysis in energy commodity markets are shown in Table 2.2

Table 2.2: *Time series analysis in commodity markets*

Year	**Reference**	**Absract**
2010	Mohammadi et al. [26]	The usefulness of several ARIMA-GARCH models for modeling and forecasting the conditional mean and volatility of weekly crude oil spot prices in eleven international markets over the 1/2/1997–10/3/2009 period is examined.
2009	Chin Wen Cheong [27]	This study investigates the time-varying volatility of two major crude oil markets, the West Texas Intermediate (WTI) and Europe Brent using ARCH-type models with prediction horizons up to 100 days.
2013	Efimova et al. [28]	This thesis investigates the empirical properties of oil, natural gas and electricity price volatilities using a range of univariate and multivariate GARCH models and daily data from U.S. wholesale markets for the period from 2000 to 2012.

2.1.3 Multivariate models

Among the methodologies described above as short-term forecasting techniques, no multivariate models can be found. Multivariate models try to capture the dynamics of a variable by estimating its relationships with other explanatory variables. To illustrate these type of models with a simple example, simple regression would be a kind of multivariate model.

However, the complexity of capturing the dynamics of energy commodity markets makes this type of models not adequate for this purpose. As training data is limited, it is difficult to find accurate explanatory models to capture these markets. In addition, the probability of these models to be based on spurious relationships even if these relationships are found in the training dataset is very high. Because of this, cointegration analysis becomes interesting in this type of markets as they try to distinguish spurious relationships from true ones. Nevertheless, cointegration is mainly studied in the long term so it will be described in the next section.

2.2 Long-term forecasting techniques

Utilities might also need complex models with accurate forecasts in a long prediction horizon in terms of trading and risk management applications. As the aim of the models proposed in the methodology exposed in the following subsections of this chapter is to validate their financial performance by applyng automatic trading strategies in energy markets, alternatives of techniques which are theoretically more accurate in long-term forecasting purposes need to be considered and compared with the ones explained above. Cointegration and error correction models are studied in this subsection.

Cointegration and error models are well recognized techniques in this context. They have two main advantages that make them suitable for capturing the complex behaviour of energy markets. Firstly, they study the relationships between time series eliminating false or spurious relationships and considering the true ones between contemporary time series. This is aligned with one of the objectives of this project: to capture the behaviour of several financial series in energy markets at the same time to evaluate the implementation of a portfolio strategy. All the methods described above would allow to build models to predict the behaviour of each series individually, but studying cointegration between them means to analyse which variables are related in the long term to develop intertwined predictions.

On the other hand, error correction models forecast the behaviour of cointegrated variables focusing on their long term relationship. This allows this type of models to be used for developing trading strategies assuming cointegrated variables converge to their long term relationship over time. This way simple financial strategies can be developed by buying assets when their price is below the expected one considering long term relationships with the rest of variables and selling them when this relationship is recovered over time.

2.2.1 Theory of cointegration and error correction models

Time series analysis through traditional quantitative methods such as regression can establish false causality relationships between variables, giving rise to the called spurious relationships.

In econometrics, a spurious relationship is a mathematical relationship in which two or more variables are statistically related, but are not in fact causally linked, usually because the statistical relation is caused by a third variable without which this relationship would not exist [29].The objective of cointegration analysis is to determine in which cases statistical relationships coincides with truly causal relationships.

Cointegration analysis is the most used technique for studying long-term relationships between variables. These relationships are found using the statistical tests that are described in the following chapters and whose aim is to prove the error terms of regression between analysed variables are stationary.

According to Granger, cointegration between two or more variables means an error correction (EC) model can be applied [30]. An EC model between two variables (Y_t and Z_t) can be formulated as follows. Y_t and Z_t are assumed to have an order of integration of 1.

$$Y_t = \alpha Y_{t-1} + \beta_0 Z_t + \beta_1 Z_{t-1} + \epsilon_t \tag{2.11}$$

The term ε_t is a white noise and follows a normal distribution with zero mean and constant variance. The expression above can be modified subtracting $Y_{t\text{-}1}$ in both sides of the equation and adding and subtracting $\beta_1 Z_{t\text{-}1}$ in the right side.

$$\Delta Y_t = \beta_0 \Delta Z_t + (\alpha - 1)[Y_{t-1} - K Z_{t-1}] \tag{2.12}$$

where

$$\Delta Y_t = Y_t - Y_{t-1}$$

$$\Delta Z_t = Z_t - Z_{t-1}$$

$$K = \frac{(\beta_0 + \beta_1)}{(1-\alpha)}$$

The parameter β_0 measures how a change in Z_t affects Y_t in the short-term, while K reflects the same in the long-term. This EC model is valid only if all terms are stationary. As Y_t and Z_t are first-order integrated terms corresponding to first differences are stationary, while Y_{t-1}-KZ_{t-1} is also stationary when Y_t and Z_t are cointegrated.

In EC models, changes in Y_t depend on both changes in Z_t and past imbalances. If Y_{t-1} and KZ_{t-1} are different, EC model tries to correct this difference multiplying it by $(\alpha-1)$.

2.2.2 Cointegration analysis in commodity markets

Cointegration analysis between spot and futures prices of commodities have been developed since the beggining of this century to evaluate long-run relationship between spot prices and futures forecasts in commodity markets [31]. Many of its applications consist in studying relationships between energy markets and economic indexes such as interest rates, Gross Domestic Product...

A survey on papers about cointegration analysis and error correction models in energy markets published during last years will be provided in the following table. The purpose of each paper will be summarized, as well as the main conclusions obtained from each analysis. The data used for carrying out the analysis will also be detailed in terms of date of first and last observation and whether they correspond to daily or monthly prices.

Table 2.3: *Cointegration analysis and error correction models in commodity markets*

Reference	**Objective and data**	**Conclusions**
The Relationship between Oil, Exchange Rates, and Commodity Prices (2009) [32]	To examine the price relationship through time of the primary agricultural commodities, exchange rates, and oil prices. Data are monthly observations for the period January 2000 to September 2008.	Commodity prices are linked to oil for corn, cotton, and soybeans, but not for wheat. In addition, exchange rates do play a role in the linkage of prices over time.

Global economic activity and crude oil prices: A cointegration analysis (2009) [33]	To investigate the cointegrating relationship between crude oil prices and global economic activity. Data are monthly observations for the period January 1988 to December 2007.	Real futures prices of crude oil are cointegrated with the Kilian economic index and a trade weighted US dollar index, and crude oil prices are influenced significantly by fluctuations in the Kilian economic index.
Oil price, agricultural commodity prices, and the dollar: A panel cointegration and causality analysis (2012) [34]	This study examines the dynamic relationship between world oil prices and twenty four world agricultural commodity prices accounting for changes in the relative strength of US dollar in a panel setting.The authors employ panel cointegration and Granger causality methods for a panel of twenty four agricultural products based on monthly prices ranging from January 1980 to February 2010.	The empirical results provide strong evidence on the impact of world oil price changes on agricultural commodity prices
Productivity, commodity prices and the real exchange rate: The long-run behaviour of the Canada–US exchange rate (2013) [35]	The paper examines the Canada–US real exchange rate since the early 1970s to test two popular explanations of the long-run real exchange rate based on the influence of sectoral productivities and commodity prices.	The empirical analysis finds that both variables exert a significant long-run effect. In addition, the effect of each variable has become stronger and a positive trend is present since 1990.
Long-run determinants of the Brazilian real: a closer look at commodities (2014) [36]	Cointegration analysis is used to show that the long-run behaviour of the Brazilian Real effective exchange rate between January 1999 and September 2012 can largely be explained by the price variation of a basket of five commodities.	It is estimated that a 10% variation in the real price of these five commodities moves the fundamental long-run real exchange rate by almost 5%.
Understanding dynamic conditional correlations between commodities futures markets (2016) [37]	To estimate dynamic conditional correlations between 10 commodities futures return in energy, metals and agriculture markets over the period 1998-2014 with a Dynamic Conditional Correlation (DCC-) GARCH model.	Correlations between commodities started increasing in the years preceding the 2008 crisis, display a peak during that year and subsequently decreased.

Electricity Consumption and Economic Activity in Malaysia: cointegration, Causality and Assessing the Forecasting Ability of the Vector Error Correction Model (2016) [38]	To empirically investigate the cointegration and causality relationship between electricity consumption (EC) and income based on Malaysia's data over the period 1980-2014 using VECM's framework.	The results from variance decomposition analysis suggested that economic activity is an important variable in explaining future variation in EC of the country.
Multivariate cointegration analysis of the Kaya factors in Ghana (2016) [39]	To investigate the long run relationships between the Kaya factors namely carbon dioxide, total primary energy consumption, population and GDP in Ghana with data spanning from 1980 to 2012.	Research results show an existence of long-run causality running from population, GDP and total primary energy consumption to carbon dioxide emissions.
A Vector Error Correction Approach on Nigeria (2017) [40]	To examine the linkages between oil price shocks and exchange rate volatility in Nigeria using monthly data from January 1996 to December 2015.	Oil price is negatively related to exchange rate in the short-run. A long-run causation is also estimated which reveals that when oil price rises by 1%, exchange rate depreciates by 58%.
Long-Run Commodity Prices, Economic Growth, and Interest Rates: 17th Century to the Present Day (2017) [41]	To examine commodity prices, income, and interest rates over the long-run.	Periods of falling commodity prices will support GDP growth for commodity importers like the US but depress growth for commodity exporters such as Chile.
Cointegration Between Energy Commodities and the South African Financial Market (2017) [42]	The long run relationships between the three energy commodities, namely crude oil, jet kerosene and natural gas, and firstly the FTSE/JSE Top 40 Index and secondly the South African Rand (versus the US Dollar) are examined.	The results indicate that there is a cointegrating relationship between the both relationships investigated and that fluctuations in the three energy commodities studied have an impact on the ZAR.

The majority of these studies carry out cointegration analysis in energy markets.

Some of them include the development of forecasting techniques too.

However, there are no complete studies that complement the analysis of cointegration in energy markets with the development of automatic buying and selling strategies in these markets based on quantitative methods. The most likely reason for this is the already mentioned complexity of capturing energy commodity markets, which undergo structural changes more frequently than, for example, securities markets and company shares.

Therefore, the innovative aspect of this thesis is to provide not only a statistical analysis between different commodity series, but also to build robust models with the ability of capturing the dynamics of these markets in which automatic trading strategies can be based on. These strategies will then be validated by implementing them in real simulations with test datasets which will determine their financial performance under different levels of risk assumed.

3 Methodology to develop automatic trading strategies

The aim of this chapter is to provide an overview of the steps to be followed in order to develop automatic trading strategies with different forecasting techniques. Some techniques will be based on univariate models and will adjust each time series depedning only on past values of the series, but also advanced multivariate models will be proposed: correction error models, introduced in the previous section. Once several models are built, techniques will be compared according to different statistical criteria.

Proposed models need to be tested once they are built concerning their applications for developing automatic trading strategies. Considered techniques not only need to be robust in statistical terms avoiding overfitting and underfitting and giving rise to accurate adjustments in training and test sets, but are also required to validate financial performace in real simulations to test the profitability strategies would have in energy markets.

The objective of this analysis is not only to find an unique solution concerning choosing just one technique as the most suitable for being used for developing automatic trading strategies. Instead, what an energy company would find more useful is to identify those situations and conditions in which each strategy presents more advantages than the others, as variability and complexity of commodities markets will give rise to different scenarios for which sometimes some techniques will perform better and other times in which other alternatives would be more appropiate. Because of these reasons, the comparison between proposed models will be based both on statistical and validated financial performance. In addition, these analysis will be made both on a monthly and a daily basis, using monthly and daily spot prices of different relevant markets.

3.1 Analysis of statistical relationships between time series

The first step to be carried out before building forecasting models is to study co-movement relationships between times series of commodities.

There are lots of factors of different nature (economical indexes, political decisions, etc.) than can affect evolution of prices in energy markets too. Times series are likely to present structural changes in which correlations with other commodities vary depending on the time interval in which they are compared. Linear correlation is the most commonly used method and is required to be studied as a first step, as if some commodities presented high linear correlations simple multivariate models would be enough to give rise to accurate adjustments.

Another alternative to study correlation relationships between commodities is to analyse residuals from ARIMA adjustments after implementing the Box-Jenkins method. The steps for this methodology will be further explained in the next subsection.

In this context, this analysis is also useful for determining if correlations between time series can be found with lagged variables instead of simple contemporary correlations. For example, in case contemporary correlation found between BRENT and TTF markets is not significant, this analysis is able to determine if despite that there is significant correlation between the price of BRENT for today and price of TTF gas for one month ago.

However, as it has been exposed, most of commodities will not present high linear correlations when considering long time intervals. Furthermore, it is also likely to find spurious correlations between commodities. Even if correlation is high it is not certain that this relationship is not based only on a particular period of time in which time series are strongly related instead of representing a true correlation in the long term.

Therefore, more advanced statistical studies are required to find true statistical relationships between commodities. With this purpose, cointegration analysis is a more powerful tool that will be carried out to identify true long term relationships between commodities. This is achieved by implementing some cointegration tests.

The Augmented Dickey-Fuller (ADF) test is the most commonly used among the most relevant unit root tests to determine the order of integration[4] of a given series. Its theoretical principles will be described in the next subsection.

The theoretical fundamentals of the most relevant statistical tests that are commonly used for studying cointegration between variables are described as follows.

Dickey-Fuller test

Dickey and Fuller [43] considered time series as an AR(1) process.

$$y_t = a_1 y_{t-1} + \epsilon_t \tag{3.1}$$

If a_1 is lower than 1, the autoregressive process can be assumed as stationary. However, the process is not stationary and will present an increasing variance if a_1 is greater than 1.

[4]The order of integration of a time series is the number of times it needs to be integrated to become stationary.

The Dickey-Fuller (DF) test consists in contrasting whether the null hypothesis H_0: $a_1 = 1$ can be rejected introducing first differences to the AR(1) process.

$$\Delta y_t = y_t - y_{t-1} = (a_1 - 1)y_{t-1} + \epsilon_t$$

$$\Delta y_t = d_1 y_{t-1} + \epsilon_t \tag{3.2}$$

A trend component and a constant can be optionally added to the AR(1) process for the test.

$$\Delta y_t = d_1 y_{t-1} + d_0 + d_2 t + \epsilon_t \tag{3.3}$$

The value of d_1 is contrasted in the test. It will be zero if a_1 is equal to 1. If this was true stationarity could not be accepted. If the null hypothesis could be rejected no unit root would exist and stationarity could be accepted. For carrying out this test the statistical variable t is calculated subtracting the expected value for d_1 according to the null hypothesis to its estimated value, dividing the result by the standard deviation.

$$t = \frac{(d_1 - 0)}{S_{d_1}} \tag{3.4}$$

The null hypothesis can be rejected in the test if the resulting value of t is higher than the given critical value in absolute terms according to the established level of signification, which is commonly around 5%. Nevertheless, Fuller proved in 1976 that critical value cannot be determined following a t-student distribution. If this is the case, Montecarlo simulation is required instead as t would not follow a t-student distribution under the null hypothesis.

Augmented Dickey-Fuller test

ADF is similar to DF test but including higher orders in the autoregressive process [44]. Therefore, if the null hypothesis cannot be rejected applying first differences for an AR(1) process time series would be considered as non-stationary for having one unit root and would be first order integrated, I(1). Similarly, if the null hypothesis is not rejected applying second differences for an AR(2) process it would be also considered as non-stationary for having two unit roots and would be second order integrated, I(2).

Cointegration Augmented Dickey-Fuller test

In cointegration analysis, the ADF test can be used to determine if the error terms of regression between two variables lack of unit roots and are cointegrated. This application of the ADF test is known as Cointegration ADF (CADF) test [45].

Therefore, the CADF test for studying cointegration between two variables is implemented in two steps. Firstly, a linear regression model is adjusted for the potentially cointegrated variables. Then, the ADF test is implemented over the error terms of the fitted regression model. If these terms can be considered as stationary, both variables are cointegrated and their long-run statistical relationship is given by the slope of the regression.

Two or more variables do not need to have zero order of integration to be cointegrated. If the linear combination of several time series with order of integration greater than zero is stationary and relationships between them do not change over time these time series are cointegrated.

Johansen test

The Johansen test estimates all existing cointegrating vectors when there are more than two variables [46]. If there are n variables which all have unit roots, there are at most *n-1* cointegrating vectors. As forecasting more than two variables corresponding to different commodities using error correction models would previously require to estimate all cointegrating vectors, Johansen test is the methodology that is commonly used in energy markets for developing multivariate cointegration analysis. This is also reflected in Table 2.3, which consists on a survey on main studies that have been developed recently concerning commodities and energy markets.

The theoretical details of the Johansen test are based on the formulation of a general vector autorregresive model, which can be written as:

$$x_t = \mu + A_1 x_{t-1} + ... + A_p x_{t-p} + w_t \tag{3.5}$$

Where μ is the vector-valued mean of the series, A_i are the coefficient matrices for each lag and w_t is a multivariate Gaussian noise term with mean zero [47].

From last equation, the Vector Error Correction Model can be now formulated as:

$$\Delta x_t = \mu + A_1 x_{t-1} + T_1 \Delta x_{t-1} + ... + T_p \Delta x_{t-p} + w_t \tag{3.6}$$

Where $\Delta x_t = x_t - x_{t-1}$. The Johansen test can provide multiple linear stationary combinations of time series. With this purpose, an eigenvalue decomposition of A is carried out. Cointegration is determined by the rank r of the matrix A. The null hypothesis of $r = 0$ means there are no cointegrated variables. A positive rank implies there is cointegration between two or more time series.

3.2 Estimation of structural changes

The structural changes existing in the time series for which the models will be built need to be considered. This is carried out by estimating those points in which either the mean or the variance of the series change significantly. These points are called breakpoints

To estimate the number of breakpoints existing in a time series, a classical regression model is considered. Assuming there are m breakpoints, there are $m+1$ segments in which the regression coeficients are constant. The optimal number of breakpoints is identified by minimizing the total squared error. The algorithm implemented in this thesis concerning this task follows the methodology described by Bai & Perron [48].

3.3 Models development

Once correlation and cointegration between commodities is analysed, the next step is to build statistical models. Both univariate and multivariate models will be built, compared in terms of quality of adjustments and validated statistically and in terms of financial profitability of the implemented strategies.

3.3.1 Univariate models

Univariate models can be built regardless of whether a significant statistical relationship is found or not between time series. Two different univariate models will be tested for being applied to the development of automatic trading strategies in energy markets: ARIMA models adjusted following Box-Jenkins method and ARMA-GARCH models.

The Box-Jenkins Method

ARIMA models have several advantages for being used as forecasting techniques. They are robust models, not very complex to build and give rise to reasonably accurate predictions.

The Box-Jenkins method [18] sets a methodology for identifying, adjusting, checking and using ARIMA models for long time series[5] (at least 50 observations). This methodology is the one that will be followed for adjusting the ARIMA models that will be implemented in this thesis.

Box and Jenkins' methodology can be implemented for stochastic[6] processes that are stationary[7]. This implies that the mean and variance of a stationary process are constant for all time intervals. Therefore, one of the steps required in the Box-Jenkins method is to transform a non-stationary series into a stationary one.

A method for transforming a non-stationary series into a stationary one is to implement a Box Cox transformation. It consists in finding the optimal value for λin the following expression that stabilise data and turn each time series into a modified one closer to follow a perfect normal distribution.

$$y(\lambda) = \begin{cases} \dfrac{y^{\lambda}-1}{\lambda} & if \quad \lambda \neq 0 \\ \\ log(y) & if \quad \lambda = 0 \end{cases} \tag{3.7}$$

Let a time series be denoted by $X_1, X_2, ..., X_t$, where t represents the time instants and X corresponds to the value. ARMA models for stationary processes can be formulated as follows:

$$X_t = \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + a_t - \theta_1 a_{t-1} - ... \theta_q a_{t-q} \tag{3.8}$$

Where the ϕ's are the autoregressive parameters to be estimated, the θ's are the moving average ones and the a's are the residuals from the adjusted autoregressive part.

Box and Jenkins used the backshift operator to make this formulation easier. The backshift operator, B, transforms time period t to time period t-1.

$$BX_t = X_{t-1}\ ,\ B^2 X_t = X_{t-2}\ ,\ ...$$

$$(1 - \phi_1 B - ... - \phi_p B^p) X_t = (1 - \theta_1 B - ... - \theta_q B^q) a_t \tag{3.9}$$

[5] A **time series** can be defined as a set of values observed sequentially through time.

[6] A **stochastic process** is a sequence of random values in which future values can only be described by their probability distribution.

[7] A statistical process is **stationary** if the probability distribution is the same for all starting values of t.

From last expression, we can define the polynomials $\phi_p(B)$ and $\theta_q(B)$ as:

$$\phi_p(B) = (1 - \phi_1 B - ... - \phi_p B^p) \tag{3.10}$$

$$\theta_q(B) = (1 - \theta_1 B - ... - \theta_q B^q) \tag{3.11}$$

These polynomials are important for checking some properties of the time series. The process can be considered as stationary if the roots of $\phi_p(B)$ lie outside the unit circle. On the other hand, it can be considered as an invertible[8] process if the roots of $\theta_q(B)$ lie outside the unit circle.

Other methods that can be used for transforming a non-stationary process into a stationary ones are identifying time series patters (such as trend and seasonality) or introducing an integrated term with order d. Once a time series can be considered as stationary, the Box-Jenkins method can be applied by the iterative implementation of the following three steps:

- **Identification**. Using information such as autocorrelation functions (ACF[9]) and partial autocorrelation functions (PACF[10]), appropiate values for p, d and q can be estimated.
- **Estimation** of appropiate values for ϕ's and θ's through maximum likelihood techniques.
- **Diagnostic checking**, by studying the autocorrelations of the residual series.

The following table shows how ACF and PACF functions can be used for determining the optimal parameters for an ARIMA model following Box and Jenkings methodology.

Table 3.1: *ACF and PACF criteria in ARMA models*

Function	**AR(p)**	**MA(q)**	**ARMA(p,q)**
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

[8] **Invertibility** refers to the fact that the moving average (MA) model can be written as an autoregressive (AR) model.

[9] **ACF** shows the correlation between the values of the series and past values of the series with different lags.

[10] **PACF** represents correlation between real values of the series and past values subtracting the dependency on intermediate lags.

ARMA-GARCH models

As the introduction exposed, models based on autoregressive conditional heteroscedasticity predict future values of the series by fitting the conditional variance. In addition, ARMA-GARCH models use ARMA models for predicting the mean value of the series and GARCH models for predicting its varying volatility depending on past values of the series and past values of its variance.

These models are interesting to be compared with other techniques because they are commonly used for forecasting financial performances in stock markets. Although commodities and energy markets present additional complexities, the objective of building these models is to see if they would perform well when being applied to predict commodities. This way the will be compared with ARIMA models to determine which one of these two univariate models is most suitable for developing automatic strategies in energy markets and maximising profits.

3.3.2 Error correction models

Error correction models can only be developed with cointegrated time series. Taking into account this requirement, the objective of this task is to compare performance of error correction models with univariate models for those variables that are found to be cointegrated.

If after comparing results correction error models are found to give rise to more profitable and realiable trading strategies, the alternative for those variables with no cointegration would be to implement univariate models or to look for structural changes to find out if the no existence of cointegration is due to concrete structural changes that take place along a given time interval.

3.4 Statistical models validation

After the step of models development, statistical validation is essential for determining if models outputs are acceptable compared with real data.This is normally confirmed by separating available data in two sets. The first set is used for training the model and the second set tests it. For the models designed in this project, statistical validation of the different models built using monthly data and daily spot prices will be made in two different ways.

The first method consists in a simple split of the data into two different sets. The second set is used for testing the model. This way models can be proved to be robust and give rise to reasonable outputs in the test set. The length of the test set has chosen to be between 10% and 30% of total amount of data in order to avoid both underfitting and overfitting, as these percentages are frequently selected in the literature. If test set is too short, the model might be overfitted, while if being too long underfitting could be a problem and results may not be accurate. For example, if available data are monthly prices going from 2011 to 2017, data from 2011 to 2016 can be used for training the model and the last year would test it.

On the other hand, statistical models validation will also be carried out through rolling windows. As the application of developed models would hypothetically be implementing automatic trading strategies, these strategies would consist in decisions-making processes of buying or selling in energy markets based on predictions resulted from the model for the next periods after being trained using a sliding window.

3.5 Selection of models and data

Once models are statistically validated, one or several of them are selected for simulating financial performance of trading strategies in real scenarios. With this purpose, the most accurate univariate model will be chosen for being compared with an error correction model based trading strategy. Simulations of the automatic trading strategies are based on sliding windows used for training the models. In these simulations, the breakpoints estimated will be considered and only time intervals in which there are no structural changes will be considered for training the models, removing pre-change observations if there is a sufficient number of points to build the model without considering them.

On the other hand, the type of data to be used in the implementation of chosen models needs to be defined too. There are two types of analysis that will be carried out in this project, one using monthly prices and another one using daily spot prices. The advantage of developing trading strategies using monthly prices is that volatility is expected to be lower than using daily prices, so it would result in more conservative trading strategies. Required computational effort is lower too. However, the main disadvantage of using monthly prices is the amount of data available for training the models is lower, and therefore models can be underfitted.

In addition, higher volatility does not necessarily mean lower profitability. If commodities are found to be cointegrated both with monthly and daily prices, higher volatility also implies higher opportunities for executing transactions based on vector error correction models. If this is the case and a higher computational effort can be assumed, using daily spot prices for implementing final strategies could be a better option.

Overall, monthly prices are firstly used to carry out statical analysis between commodities and monthly forecasting models will be built. Then, the analysis are repeated using daily spot prices to study if the increase in accuracy obtained using daily prices is worthy. If this is true, trading strategies are replicated. Performance of error correction models will be compared with univariate models.

3.6 Development of automatic trading strategies

Trading decisions concerning selling or buying are made depending on what the predictions are for the next periods. These decisions are made with a defined threshold depending on the risk the agent wants to assume when implementing the strategy. A high threshold means decisions to buy or sell assets in markets will only be made if expected benefits from the transactions given by the model are higher than this threshold, assuming low risk. On the contrary, a low threshold implies a higher number of transactions will be made as the condition is relaxed and higher profits can be obtained, but this will also mean higher risk of having economical losses in transactions if predictions are not accurate enough.

Trading strategies based on the selected univariate model and an error correction model will be compared in terms of profits obtained by each one in energy markets with respect a 'Buy and Hold strategy', which basically consists in buying one unit of each commodity and holding it exposed to market price evolution. Simulations will be based on minimum thresholds for the strategies to compare their behaviour assuming the same risk. The strategy with higher profits resulted from the simulation will be selected for being implemented following a full methodology in which inputs of the model are optimised to maximise profits in different scenarios.

One of the techniques commonly used in the literature (although mostly applied to financial but not commodity markets) is pairs trading, which can be based on a previous cointegration analysis. When two price time series are found to be cointegrated, the ratio between both series will vary around a mean and converge to this mean over time. To understand the underlying principle of pairs trading, it can be illustrated using a basic example in which an agent has a pair of securities Y and Z that have some strong economic link, for example two companies that offer the same product like Pepsi and Coca Cola. In this case, the agent expects the ratio between prices of Y and Z, which will represent the cointegrating relationship between Y and Z, to remain constant with time. However, there might be a divergence in this ratio caused by temporary supply/demand changes, reactions for important news about one of the companies, etc. When this happens, one stock moves up while the other moves down relative to each other. Therefore, if the agent expects this divergence to revert back to normal with time, the pairs trade would be to sell the stock that moved up and to buy the stock that moved down hoping to get profit by the time the ratio between prices of Y and Z converge to the expected value [49].

Cointegration analysis are important for developing this type of trading strategies. With this purpose, ADF test can be implemented to ensure time series corresponding to considered stocks are stationary when applying first differences. Then, the null hypothesis of no cointegration among considered time series can be assessed, for instance, by implementing Engel and Granger test. The test one time series on the other and then tests the residuals for a unit root.

In addition, EC models can also be used for implementing pairs trading strategies. Estimating an error correction model of the form of (2.12) will provide information about how a stock Y responds to a divergence from the constant value of the cointegrating relationship. As mentioned above, *(α-1)* in (2.12) is an estimate of the degree to which Y_t responds to imbalances between Y_{t-1} and KZ_{t-1}, which represents the stationary residual from the cointegrating relationship between Y_t and Z_t. Therefore, trading on a pair of stocks Y and Z for which estimated value for *(α-1)* is higher than for other pairs will be preferable. This conclusion is due to the fact that the larger this estimated value is the quicker the price of Y responds to revert to the constant value of the cointegrating relationship between Y and Z, so benefits from applying pairs trading strategy would be reflected in a shorter time frame [50].

Other methods frequently used for developing pairs trading strategies are the distance method and the copula method. In this context, a study was developed in 2016 to compare the performance of the three different pairs trading strategies on the entire US equity market from 1962 to 2014 with time-varying trading costs [51]. In this reference, the distance method is implemented by calculating the spread between the normalized prices of all possible combinations of stock pairs during the formation period, which is chosen to be 12 months. Then, those combinations that have the least sum of squared spreads are selected to form the nominated pairs to trade in the following trading period. Then, when the spread diverges by two or more historical standard deviation (calculated in the formation period), a long and a short position are simultaneously opened in the pair depending on the direction of the divergence. On the other hand, cointegration method uses an EC model for building a quantitative trading strategy that uses deviations from the long-term equilibrium of a cointegrated pair to open long and short positions. Positions are unwound once the equilibrium is restored. Finally, the copula method selects for each pair the marginal distributions that best fit each stock from extreme value, generalized extreme value, logistic and normal distribution. Then, the best fitting copula is selected and parameterized. Once these steps are finished, conditional probabilities are calculated to determine if chances for the stock prices to fall below current price are higher or lower than chanes for them to rise in order to set the trading strategy.

Although the study that has been mentioned above concluded the highest mean monthly excess return was achieved by implementing the distance method, the cointegration method exhibits a higher Sharpe ratio. As a result, in addition to all strategies performing better during periods of significant volatility, the cointegration method is the superior strategy during turbulent market conditions.

Before simulating performance of the strategy in real scenarios in energy markets, two previous steps are needed. This increases the difficulty of the problem. These two steps are the ones that are described as follows:

- **Selection of the optimal length of the sliding window**. Several simulations are required for selecting the length of the sliding window. The longer the sliding window the greater amount of data used for training the model. Therefore, first simulations with short sliding windows will give rise to a high number of transaction with economical losses due to underfitting. Increasing the length will reduce the proportion of transactions with losses up to an optimum size of the window.
- **Determining the optimal level of risk to be assumed**. The level of risk assumed by the trading strategy depends on the definition of the threshold for allowing transactions to be executed when the model detects expected profits from buying or selling. The higher the risk assumed, the higher the profits that can be obtained as more transactions are executed, but the higher the proportion of transactions resulting in economical losses.

These two inputs are optimised by carrying out several simulations with different values of the inputs and selecting the ones that give rise to the optimal level of risk to be assumed when implementing the strategy. After this task is completed, the financial performance of the strategy is evaluated by the use of some ratios. In this project, two main financial ratios are proposed for evaluating the strategy:

- **Sharpe ratio**. This ratio evaluates the mean excess financial returns (profits with respect to the initial investment subtracting the risk-free rate of return) by dividing it by the standard variation of annual returns. This ratio penalises overall volatility of annual returns.
- **Sortino ratio**. It is similar to the Sharpe ratio, but only penalises volatility of negative returns.

4 Statistical analysis of commodities with monthly prices

The objective of the following sections is to implement the different forecasting techniques described in previous chapters. These techniques will be used for building models to capture the behaviour of several commodities in relevant international energy markets.

Firstly, a review of the international energy markets that will be analysed is included. Then, correlation between commodities time series related to those markets will be studied. Finally, the most relevant univariate and multivariate forecasting techniques will be applied to analyse how the behaviour of monthly prices in energy markets can be captured.

The models used in this section will be trained using monthly prices of 7 different years: from January 2011 to December 2017. Advantages and disadvantages of both univariate and multivariate models are also analysed and compared.

4.1 International commodity markets

The time series considered in this first analysis correspond to the following energy commodities and markets. These markets are very liquid as they are some of the most relevant ones in the international context.

- **Brent Crude Oil Prices**. Brent Crude is a major trading classification of sweet light crude oil that is extracted from the North Sea. It is traded through International Petroleum Exchange (IPE).
- **National Balancing Point (NBP)**. It is a virtual trading location for the sale and purchase and exchange of UK natural gas. It is the pricing and delivery point for the ICE Futures Europe (IntercontinentalExchange) natural gas futures contract.
- **NetConnect Germany (NCG)**. Natural Gas at the NetConnect Germany (NCG) Virtual Trading Point.
- **Henry Hub Natural Gas (NGHH)**. Pricing point for natural gas futures contracts traded on the New York Mercantile Exchange (NYMEX) and the OTC swaps traded on Intercontinental Exchange (ICE).
- **PEG Nord (PEGN)**. It is one of 3 virtual trading locations for the sale, purchase and exchange of natural gas and LNG in France. It is one of the pricing and delivery points for Powernext natural gas futures contracts.
- **Punto di Scambio Virtuale (Virtual Trading Point – PSV, Italia)**

- **Title Transfer Facility (TTF)**. It is a virtual trading point for natural gas in the Netherlands.
- **West Texas Intermediate (WTI)**. It is a grade of crude oil used as a benchmark in oil pricing. It is the underlying commodity of New York Mercantile Exchange's oil futures contracts.
- **The Zeebrugge Hub (ZEE)** is the natural gas physical trading point in Zeebrugge, Belgium.

In this context, several alternatives are proposed for fitting both absolute values of the time series and financial performances of the indexes. All the proposed models will be further evaluated within a given validation horizon.

4.2 Correlation between commodities

A first way of studying correlation between commodities is to use residuals from fitted ARMA models. The advantage studying the correlation between residuals and not between time series is that fitting previously an ARMA model erases the autocorrelation component of correlation between considered indexes. Therefore, a "true" value for correlation can be obtained. Another important step that needs to be taken into account is a Box Cox transformation, in case it is required. This step stabilises the data before applying an ARMA model to the series. To evaluate the most accurate approach for studying correlations between residuals the Akaike Information Criterion (AIC) criteria is calculated.

The Akaike information criterion (AIC) is an estimator of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection. Let k be the number of estimated parameters in the model. Let $\hat{L}$ be the maximum value of the likelihood function for the model. Then the AIC value of the model is the following [52] [53]:

$$AIC = 2k - \ln \hat{L} \tag{4.1}$$

Given a set of quantitative models, the best model is the one with the minimum AIC value. Thus, AIC rewards goodness of fit (as assessed by the likelihood function), but it also penalizes the number of estimated parameters. In this case, parameters correspond to p, d and q in ARIMA(p,d,q) models. Table 4.1 shows the parameters selected for the SARMA models. This selection has been carried out according to maximum likelihood criteria and the value for λ obtained in the Box Cox transformations.

Table 4.1: *Fitted ARMA models and values for λ in Box Cox transformations*

Index	**SARIMA(p,d,q)(P,D,Q)[S]**	**λ (Box Cox)**
BRENT	SARIMA(1,0,1)(1,0,0)[12]	0.2673
NBP	ARIMA(1,0,1)	1.4750
NCG	ARIMA(1,0,0)	0.5316
NGHH	SARIMA(1,0,0)(2,0,0)[12]	0.3114
PEGN	ARIMA(1,0,0)	-0.9111
PSV	ARIMA(1,0,0)	0.9360
TFF	ARIMA(1,0,0)	0.4392
WTI	ARIMA(1,0,1)	0.2872
ZEE	SARIMA(2,0,0)(1,0,0)[12]	0.3816

Another important issue that has been considered is the analysis of the residuals derived from these models. Residuals should represent a white noise process, with constant variance and zero mean. A way of checking if this condition is fulfilled is the implementation of the Ljung-Box test [54]. To illustrate the analysis that have been carried out with residuals from fitted SARMA models, the one corresponding to BRENT time series is described as follows. These residuals, ACF and PACF functions are shown in Figure 4.1.

As it has been mentioned above, the Ljung Box test can be implemented to ensure there is no significant autocorrelation in the residuals. The null hypothesis that is tested is that our model does not show lack of fit, and therefore residuals are not autocorrelated. If the null hypothesis is not rejected, it can be assumed residuals are independently distributed.

The test statistic is:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}^2}{n-k} \tag{4.2}$$

Where n represents the number of obervations, $\hat{\rho}$ is the sample autocorrelation at lag k and h is the number of lags being tested. Under the null hypothesis that the residuals are independent, Q is distributed according to a χ^2 with h-p-q degrees of freedom.

The null hypothesis can only be rejected if Q is greater than the 1-α-quantile of the chi-squared distribution with h degrees of freedom (where typical values for α are 5% or 1%).

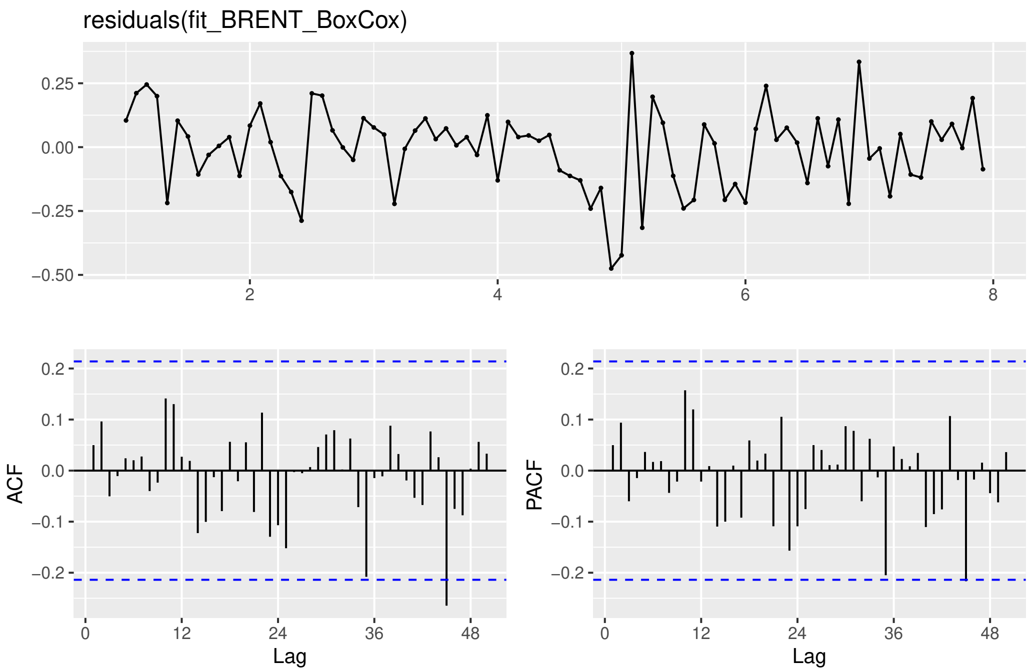

Figure 4.1: *Residuals obtained from fitted SARMA model in the case of BRENT*

Implementing the Ljung Box test to the residuals obtained from fitted SARMA model of BRENT time series, a p-value of 0.8334 is obtained. As this value is not lower than 5%, the null hypothesis that the residuals are independent cannot be rejected.

Once SARMA models for the rest of time series have been proved to generate independent residuals, cross-correlation between them can be assessed. In addition, the existence of cross-correlation between time-lagged residuals needs to be studied too. Correlation functions between lagged residuals provide the answer to whether this is the case. However, these functions reveal there is no significant cross-correlation between lagged variables but only between contemporary residuals for the time series that are considered in this project. One example that illustrate this is the cross-correlation between BRENT and WTI residuals, which is shown in Figure 4.2. It can be observed in this figure there is only one spike out of the boundaries given by the confidence interval of 95% at lag 0.

As only contemporary correlation is significant between some of the series, a correlation matrix between contemporary residuals of fitted ARMA models with a previous Box Cox transformation is shown in Figure 4.3.

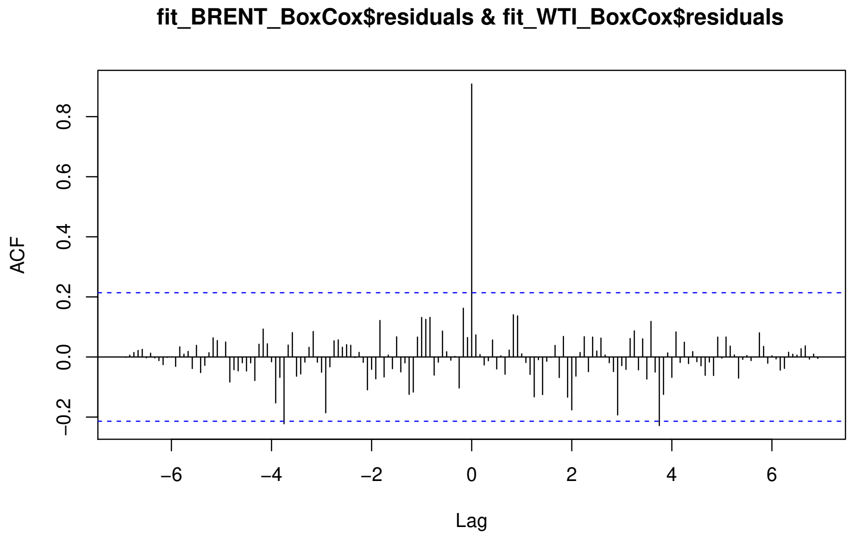

Figure 4.2: *Contemporary and time-lagged cross-correlations (BRENT - WTI)*

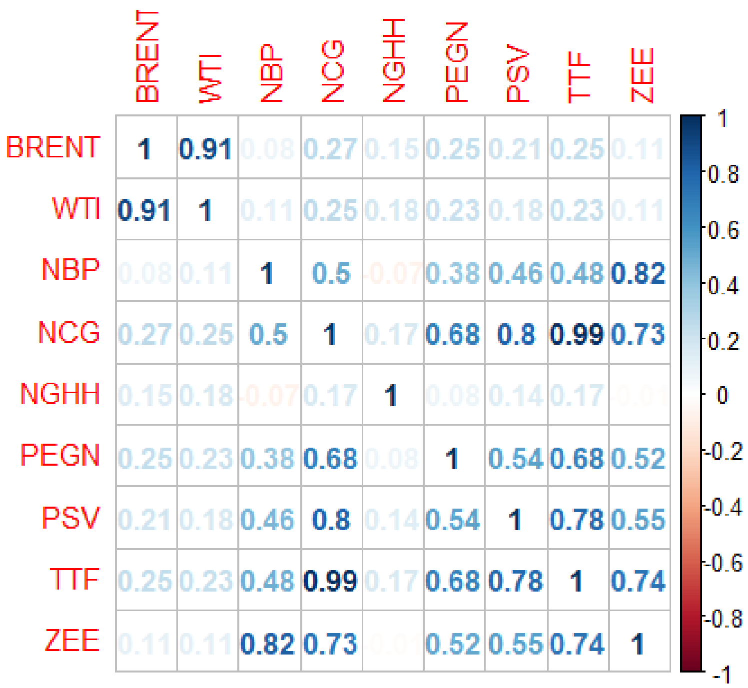

Figure 4.3: *Correlation matrix between residuals of fitted ARMA models for outputs of Box Cox transformation of commodities time series*

In this matrix, significant correlation can be noticed between BRENT and WTI, PSV and NCG, TTF and NCG, ZEE and NBP and ZEE and NCG. Therefore, no significant correlation has been found between oil and gas indexes but only between markets of the same type of commodity. This can be explained by the fact mentioned in the introduction of this thesis about the loss in correlation between oil and gas prices that have taken place during the last years. Nevertheless, the correlation matrix gives us useful information as the indexes that appear to be highly correlated are good candidates for being adjusted by models with the ability to capture these interdependencies.

It is interesting to check if these correlations vary with respect to the number of observations taken into account. To determine this, the significant correlation relationships found above are represented in the following figure with different numbers of years for the length of the calibration window.

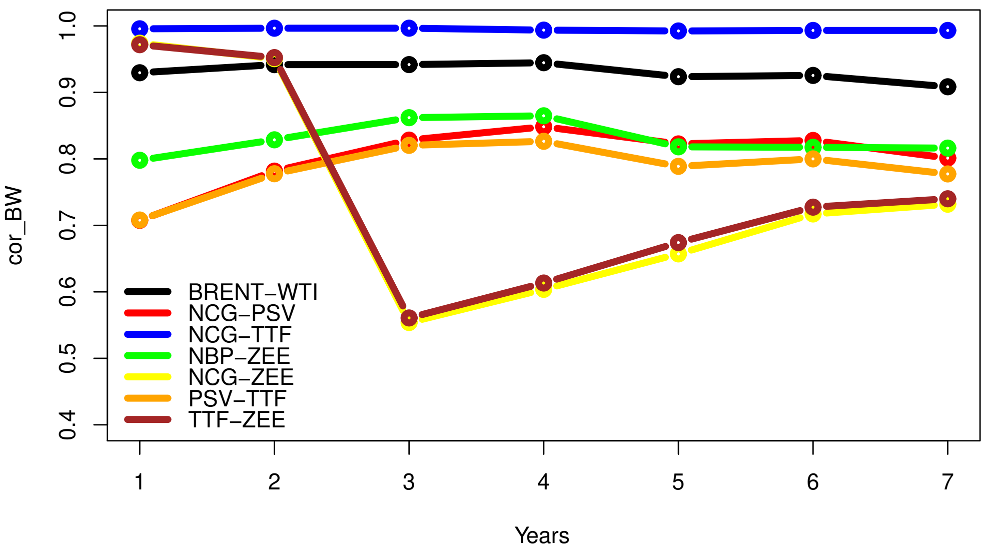

Figure 4.4: *Correlation between residuals for different calibration windows*

According to these lines, the majority of correlated pairs seem to be stable independently of the length of the calibration window considered. Nevertheless, the correlation between NCG-ZEE and TTF-ZEE changes significantly after considering the observations of the third year of the calibration data. We can conclude that those correlation relationships that appear to be more stable are more likely to be true statistical relationships instead of spurious relationships. A very accurate way to classify those two types of relationships is to study cointegration between significantly correlated variables, will be carried out in later sections.

4.3 ARMA-GARCH models

Other types of models commonly used in financial markets are GARCH models. These models are proposed to forecast the behaviour of time series basing on predictions of variance. Therefore, this type of models are useful for forecasting volatility of financial performances.

The purposed model for commodities considered above consists in an ARMA for fitting the mean of the series and a GARCH model for predicting volatility of financial performances of time series of commodities. Logarithmic returns are calculated as follows:

$$FP_{m,i} = \log \frac{P_{m,i}}{P_{m-1,i}} \sqrt{\frac{12}{k}} \tag{4.3}$$

Where i is an index that corresponds to each commodity, m is an index that represents the correspondent month data and k is the number of months between samples (in this case, one month).

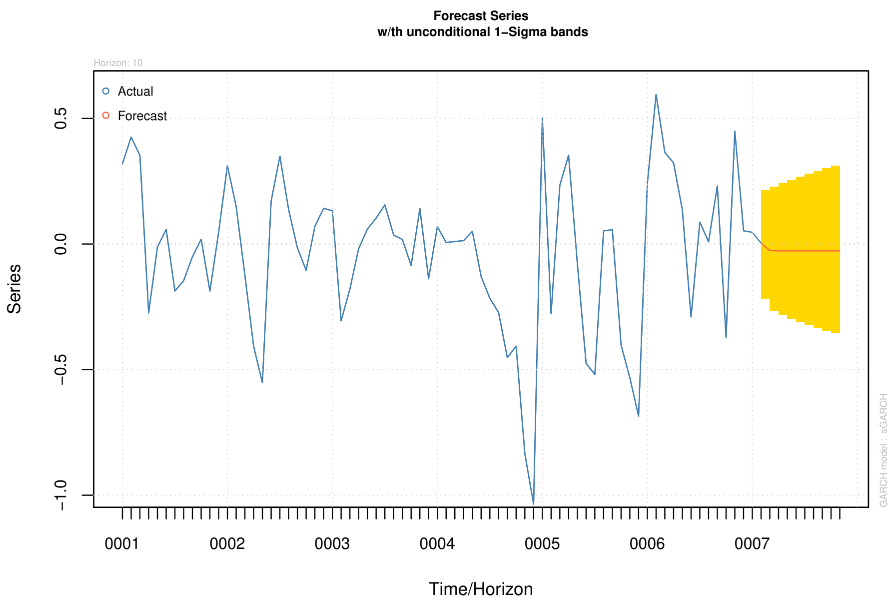

Figure 4.5: *ARMA-GARCH forecasts for BRENT index*

As explained in previous sections, GARCH models fit the variance of the series using not only past values of the own series, but also past values of its variance. An ARMA-GARCH model for BRENT is illustrated in figures 4.5 and 4.6. Figure 4.5 shows the forecast of the time series, while Figure 4.6 illustrates the forecast of the standard deviation.

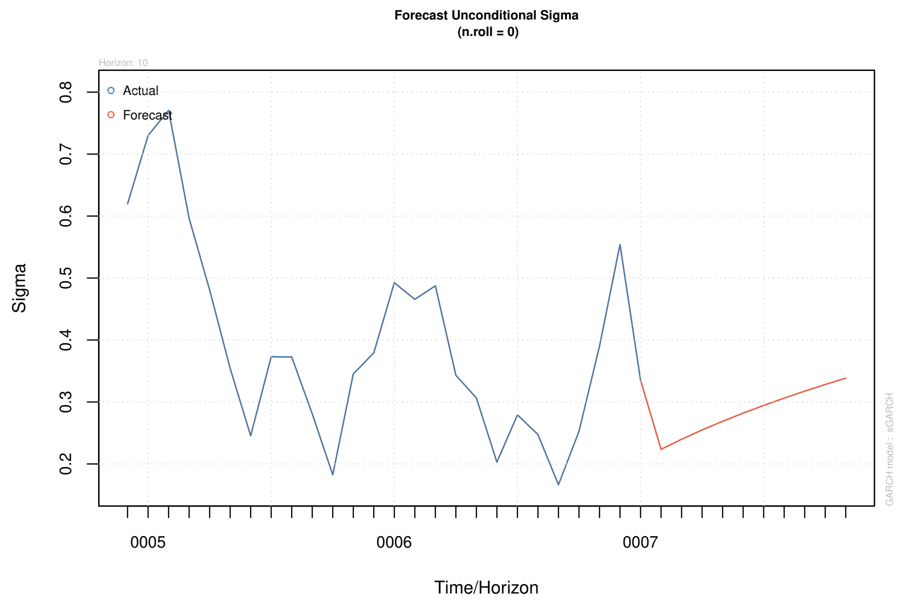

Figure 4.6: *Forecasted sigma in GARCH model for BRENT*

In those figures, the x-axis shows the number of years considered (year 1 corresponds to 2011 and year 7 corresponds to 2017). However, as the model is built using logarithmic performances, a further step consisting in transforming financial performances into price values P from last equation is required to compare its behaviour within a validation horizon with the rest of the models.

4.4 Correlated data simulation based on univariate models

In order to take into account the correlation matrix computed above, ARMA-GARCH models will be used for simulating correlated paths. The idea is to generate a simulation in which stable correlations found in Figure 4.4 is maintained. To achieve this, the following steps have been followed:

1. The time series are forecasted considering the last 10 months of 2017 as the prediction horizon.
2. The stable correlations found in the correlation matrix of rediuals are taken into account, and then a Cholesky decomposition of this matrix is applied.
3. Predictions are stabilized by subtracting the mean and dividing the result by the standard deviation.
4. The Cholesky lower triangular matrix is multiplied by a matrix containing the stabilized paths.

5. The stabilized and correlated paths are multiplied by the standard deviation and the mean value is added again.

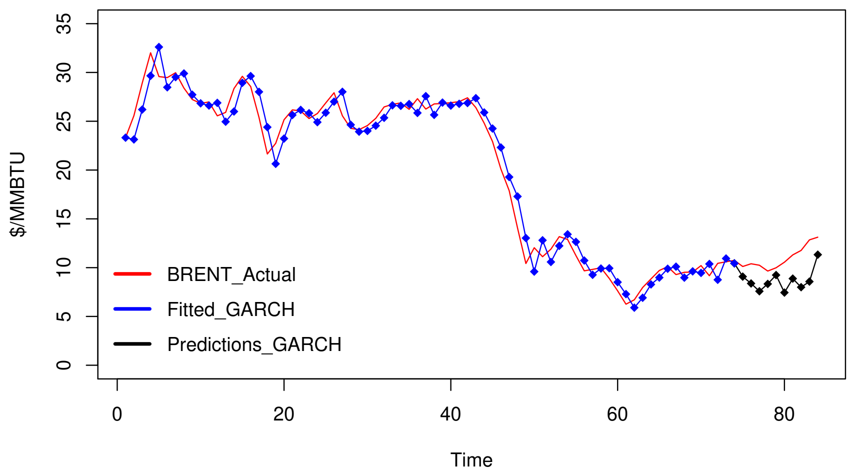

Figure 4.7: *Correlated data simulation based on an ARMA-GARCH model for BRENT*

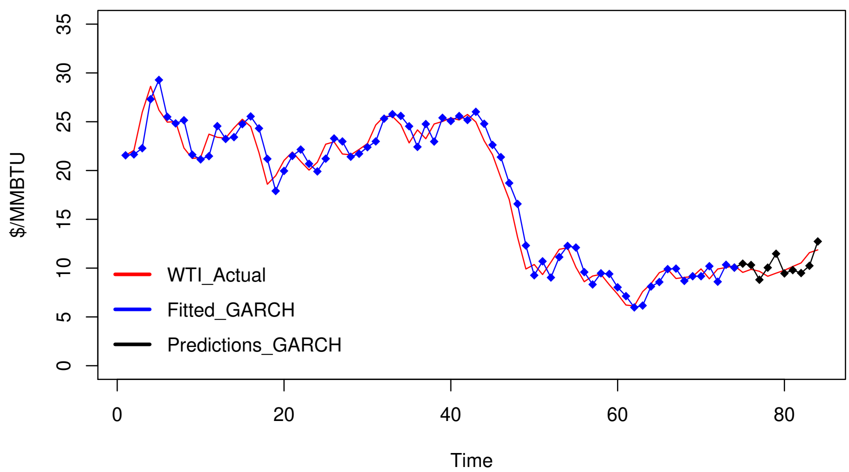

Figure 4.8: *Correlated data simulation based on an ARMA-GARCH model for WTI*

With these steps, the correlation studied in the training dataset is maintained in the predictions. This constitutes a first approach to capture both the individual behaviour of each series with univariate models and the interdependency between those series that present stable correlation relationships. The adjustments of two of the series that are correlated, BRENT and WTI with this approach is shown in figures 4.7 and **??**. Predictions based on correlated data simulation are also included.

4.5 Cointegration analysis of commodities in energy markets

As explained in the state of the art, cointegration analysis is one of the most commonly used for determining true long-term relationships between time series in financial markets. There are several cointegration tests developed with this purpose that make possible to provide statistical evidence of the existence of cointegration between variables. In this section, a cointegration analysis will be carried out focusing on spot prices of energy commodities to determine if some of the time series of commodity prices in the main oil and gas markets are cointegrated.

The following graphs show the available data for the amount of time series considered above. Before executing any cointegration test, it can be expected just by observing the graphs that Brent and WTI are likely to be cointegrated as they are related to crude oil spot prices. Looking at the behaviour of the other time series in the graphs, NCG, PEGN and TTF seem to behave similarly too. BRENT and WTI monthly prices are shown in Figure 4.9, Figure 4.10 shows historical monthly prices of NCG, PEGN and TTF prices, Figure 4.11 shows the ones of NBP and ZEE indexes and Figure 4.12 shows the ones of NGHH and PSV indexes.

This is also reflected in the correlations that appear to be constant with different calibration windows in Figure 4.4. Those series are likely to present true statistical relationships in the long term.

Two different approaches will be described in this thesis to capture this interdependency. The first one has been already presented in the last section, consisting in building univariate models and simulating correlated paths by applying Cholesky decomposition to the correlation matrix of those residuals that appear to have stable correlations.

On the other hand, the second approach will be described in this section. An alternative is to build error correction models for those variables that can be proved to be cointegrated. Nevertheless, for those time series that cannot be proved to be cointegrated these models cannot be proposed. This is why cointegration analysis is necessary before building this type of models.

Figure 4.9: *BRENT and WTI ($/MMBTU)*

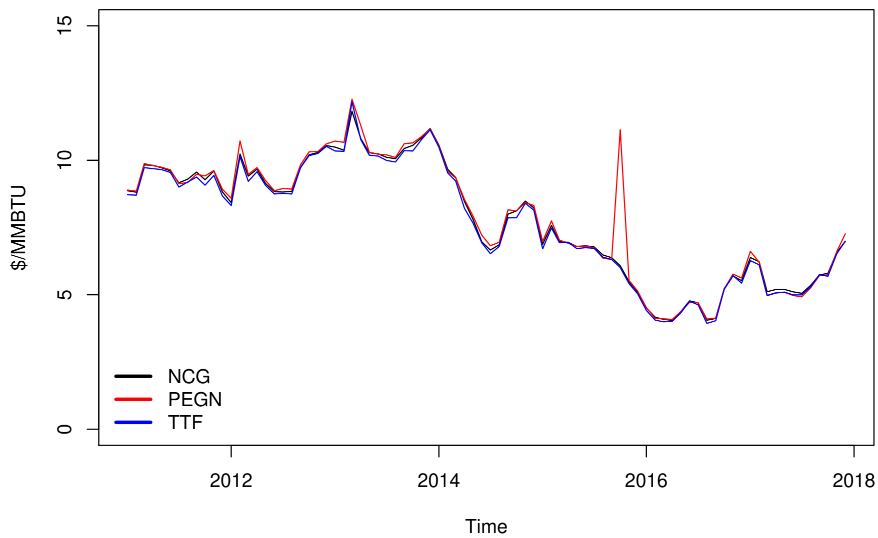

Figure 4.10: *NCG, PEGN and TTF ($/MMBTU)*

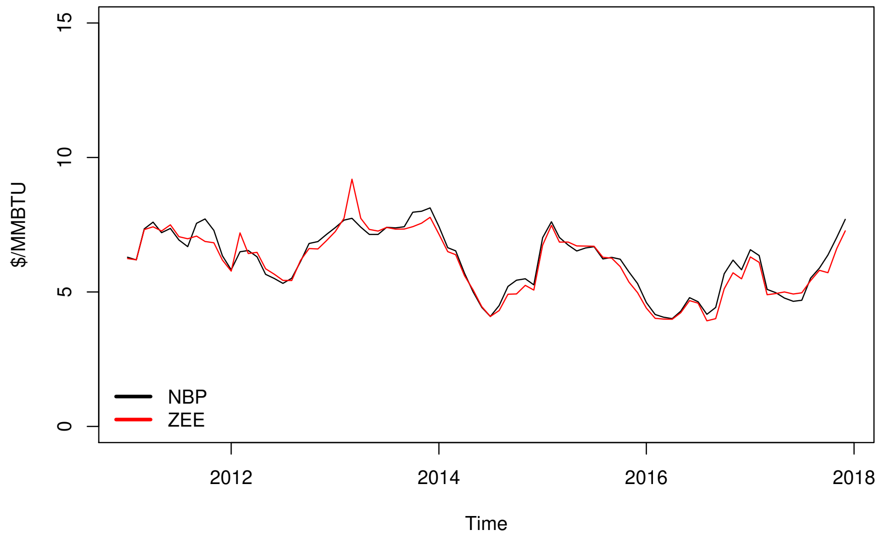

Figure 4.11: *NBP and ZEE ($/MMBTU)*

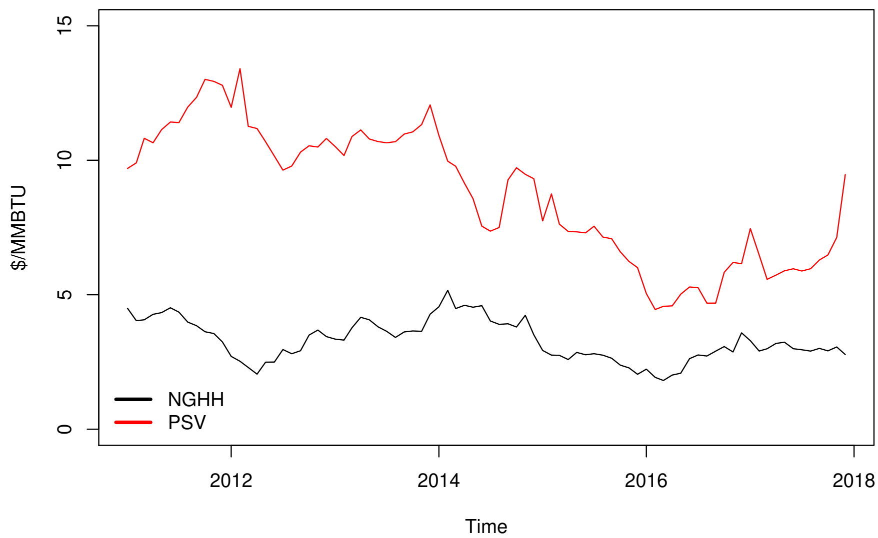

Figure 4.12: *NGHH and PSV ($/MMBTU)*

4.5.1 Cointegration tests

According to how time series evolve, CADF tests are carried out between variables that appear to have a similar behaviour with respect to time. These tests, as it was introduced in the chapter dedicated to explain the methodology to be followed in this analysis, provide evidence of existing cointegration between variables.

The first CADF test is implemented between BRENT and WTI. The test is carried out to determine if residuals from the linear regression model of BRENT on WTI are stationary, which would suggest both variables are cointegrated. A p-value lower than 0.05 (0.01921 for regression of BRENT on WTI) suggests that both variables might be cointegrated, although the obtained p-value is not too low. However, we must be extremely careful when evaluating these results because of the fact that the data used in this analysis consist of monthly prices from 2011 to 2017 (only 84 past measurements).

There are others CADF tests that have been carried out and gave rise to p-values lower than 0.01 as a result, so it can be assumed under a confidence interval of 99% cointegration between tested variables exits. The CADF tests that provided these results test the stationarity of residuals from linear regression between the following variables: NCG and PEGN; NCG and TTF; PEGN and TTF; and NBP and ZEE. Other combinations of commodities did not appear to be cointegrated according to this criteria.

The Johansen test was applied to check if there is a combination of more than two commodities that can be assumed to be cointegrated. However, when applying this test to the basket of all the commodities enumerated above the results obtained suggest there are no cointegrating vectors containing more than two commodities at the same time that can be assumed to be cointegrated (the hypothesis of the rank of the matrix considered in this statistical test being lower or equal to 2 cannot be rejected with a confidence level of 90%). Although NCG is cointegrated with both PEGN and TTF, the combination of the three indexes is not cointegrated according to the Johansen test applied.

4.5.2 Vector Error Correction Models

According to Engle and Granger, if a set of variables are cointegrated, then there exists a valid error correction representation of the data, and viceversa. This section evaluates if VECM can successfully fit some of the pairs of variables that were found to be cointegrated in previous section. Engle-Granger Two-Step approach has been estimated for those pair of variables that appear to be cointegrated according to CADF tests.

The Engle-Granger Two-Step approach for two cointegrated variables y and x consists on the following procedure:

- First step: estimate the long-run equilibrium equation.

$$y_t = \delta_0 + \delta_1 x_t + u_t \tag{4.4}$$

- Second step: estimate the error correction model.

$$\Delta y_t = \phi_0 + \sum_{j=1} \phi_j \Delta y_{t-j} + \sum_{h=0} \theta_h \Delta x_{t-h} + \alpha u_{t-1} + \epsilon_t \tag{4.5}$$

This procedure has been implemented using monthly data for building VECM with 1 lag using cointegrated time series. The first VECM fits BRENT and WTI. Equation (28) shows the estimated parameters of the VECM of BRENT and WTI and p-values corresponding to those parameters between parentheses.

$$\begin{pmatrix} \Delta X_t^1 \\ \Delta X_t^2 \end{pmatrix} = + \begin{pmatrix} 0.0144(0.9074) \\ 0.1675(0.1700) \end{pmatrix} ECT_{-1} \begin{pmatrix} -0.1117(0.4580) \\ -0.0727(0.6223) \end{pmatrix} + \begin{pmatrix} 0.4050(0.1094) & -0.0081(0.9752) \\ 0.1120(0.6492) & 0.2783(0.2777) \end{pmatrix} \begin{pmatrix} \Delta X_{t-1}^1 \\ \Delta X_{t-1}^2 \end{pmatrix} \tag{4.6}$$

Where ${X_t}^1$ represents BRENT prices and ${X_t}^2$ corresponds to WTI index.

On the other hand, equation (29) represents the VECM of NCG and PEGN including p-values of estimated model parameters.

$$\begin{pmatrix} \Delta X_t^1 \\ \Delta X_t^2 \end{pmatrix} = + \begin{pmatrix} 0.1561(0.3333) \\ 1.2058(1.5e\text{-}06) \end{pmatrix} ECT_{-1} \begin{pmatrix} -0.0421(0.5091) \\ -0.0710(0.4400) \end{pmatrix} + \begin{pmatrix} -0.0553(0.7460) & 0.0444(0.6917) \\ -0.0923(0.7071) & 0.0715(0.6574) \end{pmatrix} \begin{pmatrix} \Delta X_{t-1}^1 \\ \Delta X_{t-1}^2 \end{pmatrix} \tag{4.7}$$

Where ${X_t}^1$ represents NCG prices and ${X_t}^2$ corresponds to PEGN index.

The adjusted VECM models are plotted in the following figures. Figure 4.13 and Figure 4.14 represent the adjusted VECM for cointegrated indexes BRENT and WTI, while Figure 4.15 and Figure 4.16 show the adjusted VECM for cointegrated indexes NCG and PEGN.

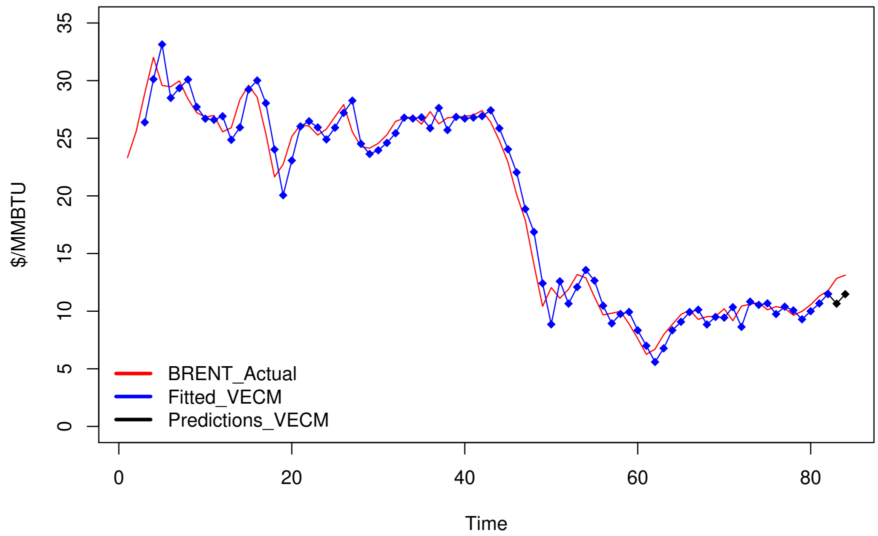

Figure 4.13: *BRENT historical monthly prices and adjusted VECM*

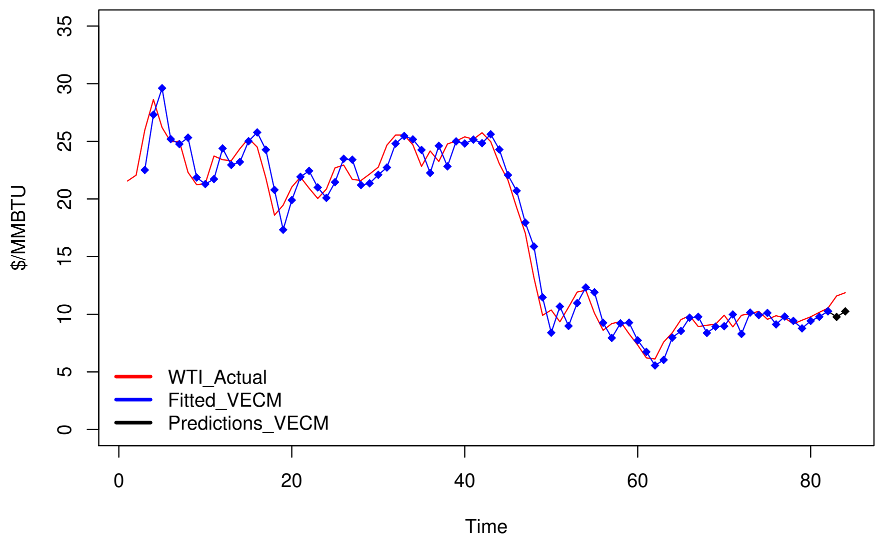

Figure 4.14: *WTI historical monthly prices and adjusted VECM*

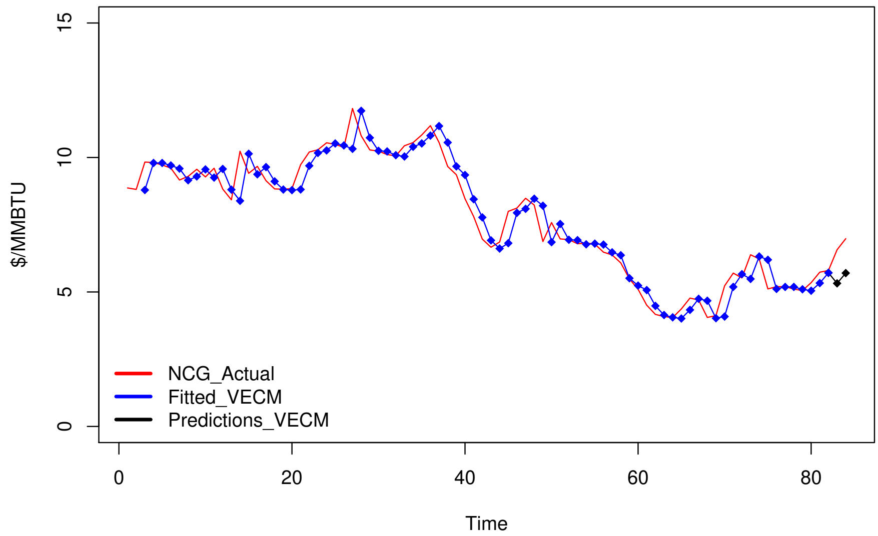

Figure 4.15: *NCG historical monthly prices and adjusted VECM*

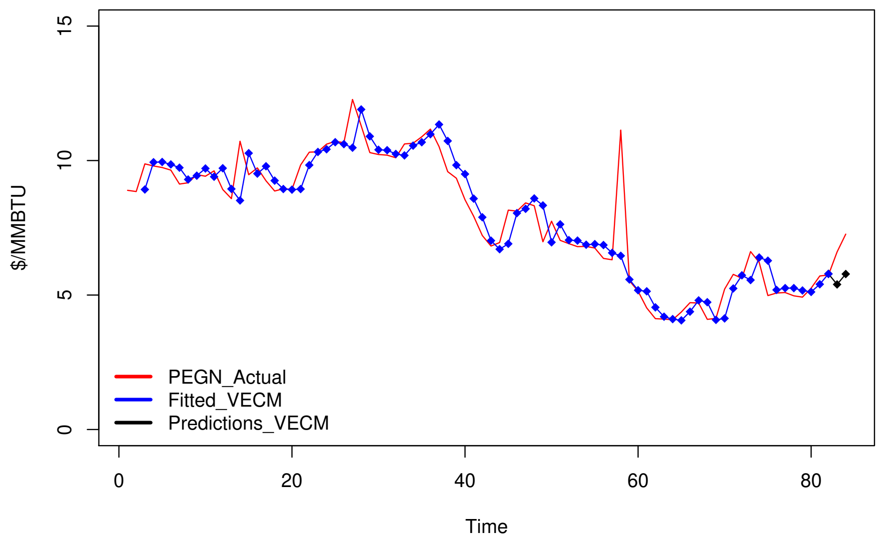

Figure 4.16: *PEGN historical monthly prices and adjusted VECM*

4.6 Statistical validation of the models

As a first approach for testing the models described above, a validation horizon of last 10 months of the monthly prices from 2011 to 2017 is chosen. Figure illustrates predictions from the models for BRENT.

In the case of the VECM, a rolling forecast has been applied. This method allows to check the out-of sample forecasting accuracy by estimating the model on a sub-sample of the original, then making rollingl forcasts of a given horizon, each time by updating the sample.

On the other hand, both the Box and Jenkins methodology and the ARMA-GARCH models are validated by simulating correlated forecasting paths following the steps enumerated in section 4.4. The following figures show the validation carried out for BRENT and WTI following three different approaches. The first two are based on these univariate models and a correlated data simulation. Then, the accuracy of these approaches can be compared with a VECM, as these two series are cointegrated. Only these two commodities will be used for comparing the different approaches, as they represent two cointegrated variables that can be adjusted through a VECM. For non-cointegrated variables, only the first two approaches could be chosen.

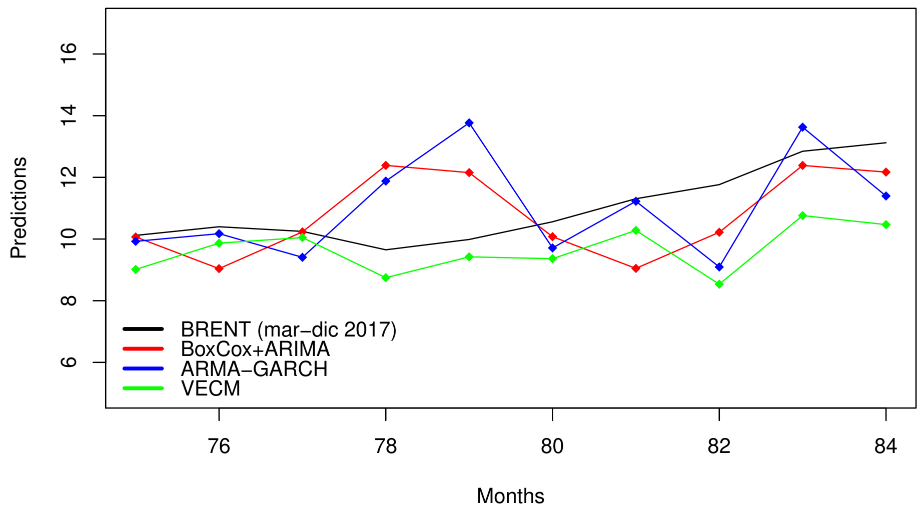

Figure 4.17: *2017 validation of models for BRENT*

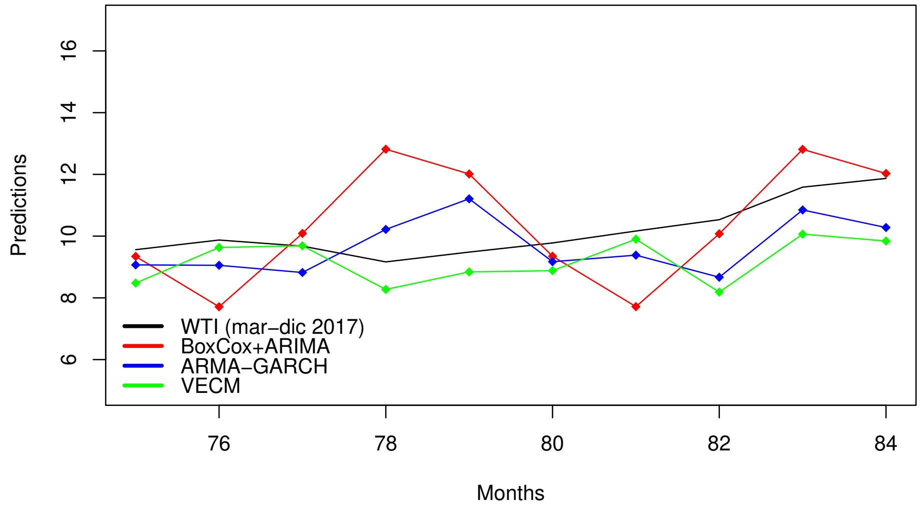

Figure 4.18: *2017 validation of models for WTI*

In the two graphs shown above, it can be observed that predictions are relatively close to real data. Therefore, it can be concluded that overfitting has been avoided in the development of the models.

Although a fixed test set is useful for avoiding statistically overfitted models, a sliding window validation is more commonly used for developing automatic tradint strategies. As ARMA-GARCH and VECM models are commonly used in the literature for this application, a rolling validation will be applied for comparing the performances of both of them in a real context.

For each interval the model is rebuilt with a training window of 48 months and the forecasted value for the next month is computed. To compare the performance of each model, the values for the mean squared error (MSE) are calculated in Table 4.3. Root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) are computed too.

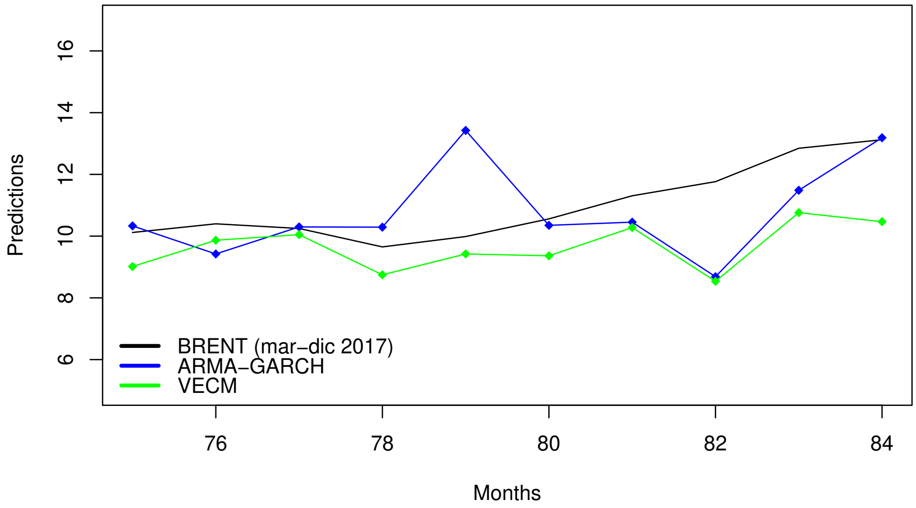

Figure 4.19: *Sliding window validation for models of BRENT*

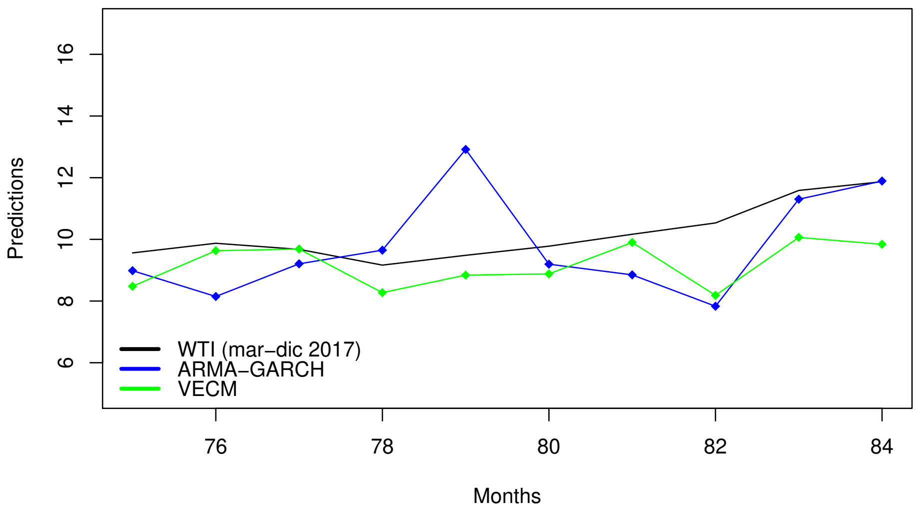

Figure 4.20: *Sliding window validation for models of WTI*

Table 4.2: *Mean Squared Errors in sliding window validations (BRENT)*

Models for BRENT	MSE	RMSE	MAE	MAPE
ARMA-GARCH	1.3657	1.1686	0.8681	9.3516
VECM	1.2877	1.1348	0.8789	9.4135

Table 4.3: *Mean Squared Errors in sliding window validations (WTI)*

Models for WTI	MSE	RMSE	MAE	MAPE
ARMA-GARCH	0.8565	0.9255	0.7447	7.9865
VECM	0.9386	0.9688	0.7765	8.2822

As it can be observed in the errors computed above, ARMA-GARCH models with correlated data simulation and VECM models seem to be equally accurate in both cases. Because of this, trading strategies based on these two approaches will be developed and compared according to other criteria in the next section.

5 Comparison of trading techniques for BRENT and WTI daily spot prices

In this chapter, the different forecasting techniques described above are used for fitting daily spot prices of the energy markets mentioned at the beginning of Chapter 4. Available data go from July 1st 1997 to February 25st 2019.

Using a higher amount of data to build the models requires a higher computational effort, which constitutes a disadvantage with respect to simplified models with monthly prices. However, if this effort can be assumed more accurate results will be obtained too and strategies based on these quantitative models will be more reliable.

The analysis that will be carried out in this chapter include a first step consisting in building the model to capture the evolution of daily prices. Then, advantages and disadvantages of the proposed strategies will be compared. Finally, a risk analysis will be carried out to evaluate the performance under different scenarios of those strategies and determine if they are suitable for being profitable.

With this comparison, the objective is to find the best alternative under which an utility could obtain the best economic results in real applications. Once one of the models is identified as more advantageous in terms of fulfilling the objectives of this project, a full implementation of an automatic trading strategy based on this model will be carried out in Chapter 6. Then, a real simulation of the financial performance of this strategy will be carried out using one of the years with available real data for as the validation horizon.

In order to do so, the analysis carried out in this chapter will use as inputs BRENT and WTI prices, as this pair of variables was one of the few that were found to be cointegrated in Chapter 4. As the purpose of this chapter is to determine if monthly prices are worthy to be replaced by daily spot prices, focusing on just two cointegrated variables is enough for studying the outputs of the different forecasting techniques.

5.1 Development of forecasting models

Using daily spot prices of BRENT and WTI oil from 2013 to 2017, ARMA-GARCH and VECM models are fitted and predictions are computed with a sliding training window of 500 days. The strategies that will be developed in the following sections focus on BRENT and WTI markets because of the proved existing cointegration relationship between them through the CADF test carried out in section 4.4.1. Although this test proved cointegration between monthly prices, it will be assumed in this section that cointegration extends to daily spot prices too, as the objective is to determine which of the proposed techniques is more convenient to use for building a full automatic model. If the VECM strategy appears to be more advantageous, a new cointegration test needs to be carried out before building this model to check that cointegration exists between daily spot prices.

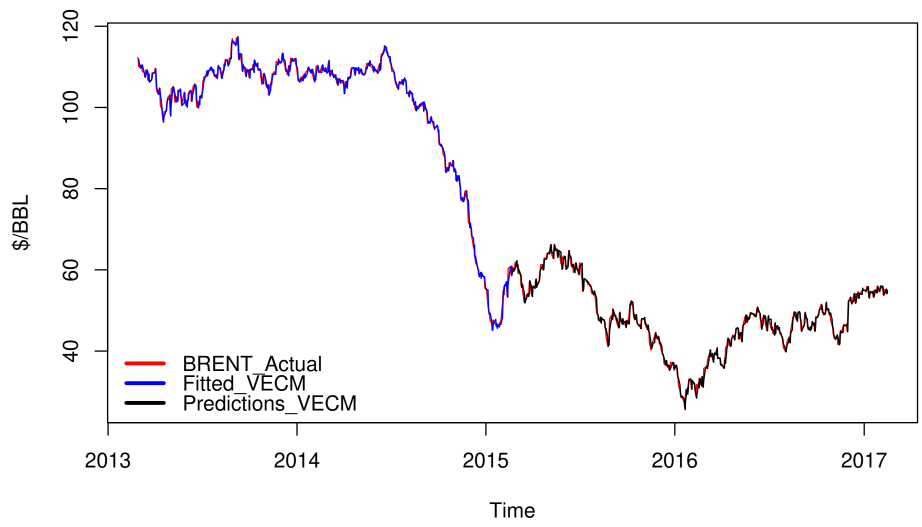

Figure 5.1: *Fitted Error Correction Model for BRENT daily spot prices*

The size of the training window is considered as fixed for a first analysis concerning the profitability of the strategies using daily spot prices, but can be modified if structural changes of the series are considered. With this sliding window, the price for the next day is forecasted by each model. Using a sliding window with a small prediction horizon gives rise to accurate predictions that are illustrated in the case of BRENT forecasted prices using a VECM in Figure 5.1.

On the other hand, it can be observed in the evolution of BRENT daily spot prices the fact that there are several structural changes in the series with high spikes in some instants. This can affect the accuracy of some of the predictions around those periods in which structural changes happen, and therefore the well performance of the trading strategy would be affected too.

Because of this and several facts that affect the profitability of the trading strategy, a process to determine an adequate method for devloping reliable and automatic trading strategies with daily spot prices is needed. It will consist in two steps. Firstly, ARMA-GARCH and VECM models will be compared in terms of volatility and profitability. Once the best alternative among them is chosen, a full model for developing trading strategies will be made with additional features for capturing the complexity of commodities prices including structural changes.

5.2 Definition of safety thresholds for trading strategies

With the models built following the same processes as with monthly prices, similar trading strategies are computed for BRENT and WTI to compare the results. The inputs to be decided in this strategy are associated with risk. For both strategies based on ARMA-GARCH models and VECM models a threshold is defined not to execute any transaction unless this threshold is reached.

Firstly, forecasted financial performance in the next period needs to overcome a threshold (in its positive or negative value) to adopt a short or a long position in ARMA-GARCH strategies.

In the case of VECM models, the threshold is based on a statistical factor z. Once the beta coefficient representing the long-term cointegration between considered variables is computed from the VECM model, the spread between cointegrated variables is calculated for the next period of time as:

$$spread_t = WTI_t - \beta BRENT_t \tag{5.1}$$

Similarly, the spread between both variables for the training data of the sliding window is computed in a vector and its mean and standard deviation are calculated too. Then, a statistical factor z is calculated as follows:

$$z = \frac{spread_t - \mu_{spread,training}}{\sigma_{spread,training}} \tag{5.2}$$

Where $\mu_{spread,training}$ represents the mean value of the spread from the long-term fitted cointegration relationship during the training horizon and $\sigma_{spread,training}$ its standard deviation.

5.3 Comparison of trading strategies

With the models developed in the previous sections, several trading strategies are proposed. The objective of the development of these strategies is to provide a comparative analysis to check if the models built above can be used for developing automatic strategies that can be validated with real data and therefore proved to be economically profitable with respect a Buy and Hold strategy.

In this context, two different trading strategies will be analysed. The first one is based on an ARMA-GARCH model and consists on forecasting the price for the next period of time of the series using a sliding window of 2500 days. The second trading technique is based on a VECM model, and executes a transaction when a spread in the long-term relationship between cointegrated daily prices higher than a given threshold is identified.

The first step is to assess the financial performance of both strategies. To compare them, very risky thresholds are defined to check the outputs and volatility of each strategy. The following graph shows equity curves defining low thresholds. As the ARMA-GARCH strategy is based on financial performances, a threshold of 0.001 is defined for this analysis. This value is equivalent to an increase or decrease in BRENT and WTI prices of 0.023% of the price in the previous day, which is a reasonable value for the purpose of assessing risky versions of both strategies. In the case of the VECM strategy, a threshold of 0.05 times the standard deviation of the spread is defined.

$$INC_{GARCH} = 10^{\frac{0.001}{\sqrt{12}}} - 1 = 0.00023 = 0.023\% \tag{5.3}$$

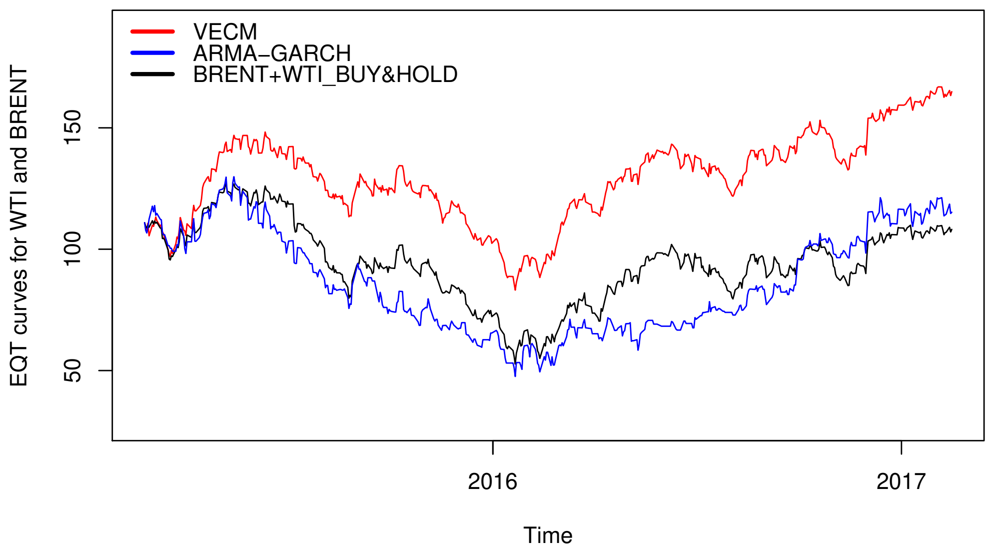

Figure 5.2: *Comparison of trading strategies with risky thresholds*

The financial profits obtained in this simulation with a trading strategy based on a VECM model is higher than the ones corresponding to a Buy and Hold strategy and an ARMA-GARCH model. It can be concluded that, with thresholds that represent a similar level of risk, a VECM trading strategy performs better in this context.

5.4 Risk analysis

As it can be observed, even if defining risky thresholds a VECM strategy is much more reliable and stable with time. Considering the initial investment as the first value of the equity curve of a Buy and Hold strategy, the investment to buy 1 barrel of BRENT oil and 1 barrel of WTI oil at the beginning of 2015 is 111$. After executing a VECM strategy the equity obtained is 164.14$, with a profit with respect the equity that would have been obtained choosing a Buy and Hold strategy (107.68$) of 56.46$. This means that the VECM strategy using a low threshold gives rise to a return of 50.86% after 500 days.

However, to really execute this kind of strategy several criteria must be considered to choose an optimal threshold. Figure 5.3 shows the profits obtained with a VECM strategy with respect to a Buy and Hold strategy for different values of the threshold z. Profits have been computed for values of z between 1e-6 and 2. In this figure, it can be observed that profits tend to decrease when the risk associated with the threshold used decreases too. Nevertheless, some exceptions can be noticed due to the low amount of observations that is used in this comparative analysis as the purpose of it was to assess the performance of univariate and VECM trading strategies with daily spot prices. A more detailed risk analysis will be included in the next chapter.

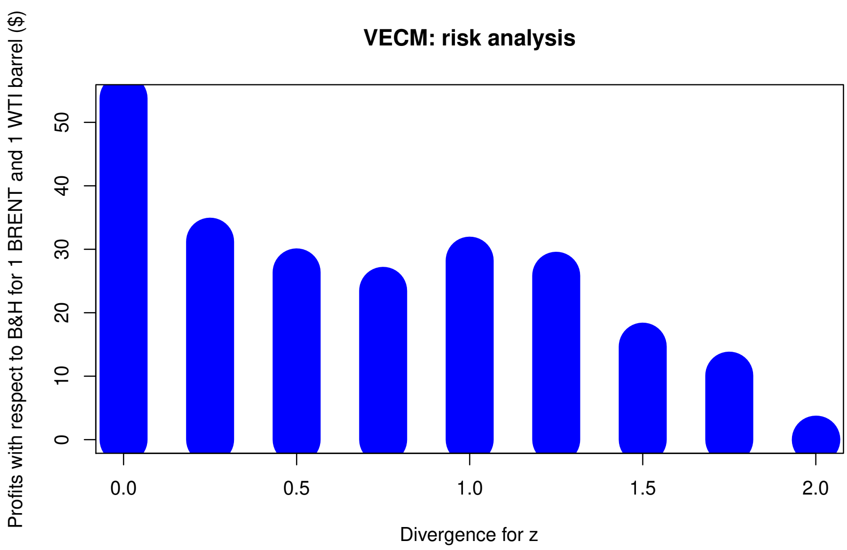

Figure 5.3: *Risk analysis concerning z factor for VECM strategy*

A criteria for chossing a value of *z* must be defined. If z is very low, it is likely to incur in a higher number of transactions that give rise to financial losses, as they would be carried out in a less conservative way whenever the spread overcomes z times the standard deviation. However, with a very high value of z profits will be reduced as less transactions are carried out, but a higher percentage of them will contribute to financial profits. To determine the optimum value of z, a Pareto chart can be used.

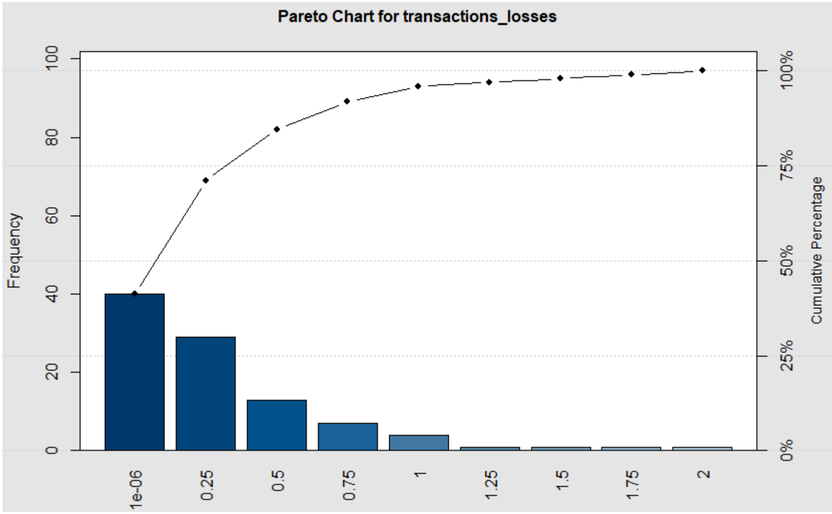

Figure 5.4: *Pareto chart of number of transactions with losses in VECM strategy*

Concerning the Pareto chart of Figure 5.4, it can be observed that defining a higher value of z than 0.25 would allow to avoid more than 70% of the variance of negative transactions. Because of this, an adequate value would be higher than 0.25 and lower than 0.5. Higher values are more conservative, but potential profits of the trading strategy would be reduced as well.

Figure 5.5 focuses on how the number of transactions with losses reduce while increasing the value of z by intervals of 0.05. It can be observed in it that the slope decreases significantly when z reaches 0.4 and decreases again after overcoming values around 0.7. This changes in the slope can be further analysed by structural changes tests.

5.5 Judgement

Once comparisons between techniques and analysis of their advantages and disadvantages have been completed, VECM models seem to be the best alternative for an agent interested in developing automatic trading strategies in energy markets. It is worthy to build the models using daily spot prices instead of monthly prices too, as computational effort could be assumed in the trials carried out. The next section includes a full methodology to develop automatic portfolio trading strategies from the perspective of an utility using data from energy markets mentioned in this chapter.

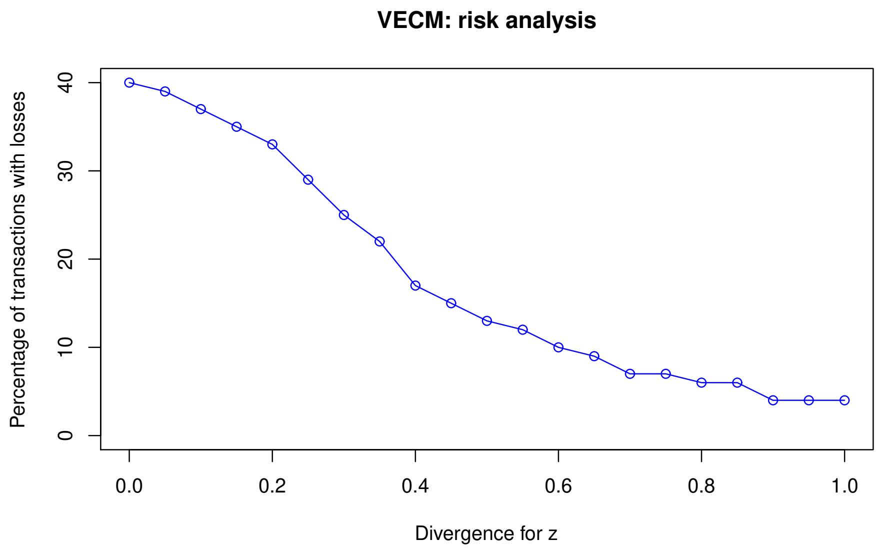

Figure 5.5: *Number of transactions with losses with respect to z*

Although chapter 6 describes all the steps followed to build the VECM strategy in this thesis, several difficulties were identified while doing the tasks of programming and validating results. This means an equilibrium between trying to simplify the models and reaching accurate enough results needed to be obtained, as too complex models would require too much computational effort and on the other hand too simple models would give rise to a poor quality of predictions.

Because of this, several ways of optimising the trading strategy were explored and included in the final model. For example, structural changes had a great impact on results and capturing them was possible without making the problem too big. These improvements are also described in the next section trying to respect the chronological order of the followed steps to provide the reader with a general perspective of how the final model was obtained.

6 Real implementation of a VECM trading strategy

The proposal described in this chapter is to implement a full VECM trading strategy using daily spot prices. After analysing the advantages VECM models have with respect to other forecasting techniques and also after proving the higher computational effort required to train models with daily prices is reasonable to be assumed after the trials that were exeuted in this project described in the previous two chapters, it has been conlcuded error corrections models can give rise to robust and reliable strategies with a higher probability than the rest of the analysed techniques.

Because of this reason, a full methodology to implement a VECM trading strategy is described in this chapter. The steps followed to program and build the strategies are explained. The first step is to study relationships between variables to determine which variables are cointegrated and can be fit through an error correction model. Once the models are built, a validation using part of the available data as a test set is included to prove the financial performance of the proposed strategy would be profitable for an utility.

Firstly, the inputs that the model needs will be identified. These inputs are decisions that need to be made trying to find a equilibrium between not overfitting the variables and obtaining accurate enough predictions. As the models will be implemented using daily spot prices and a sliding window to simulate how an utility would apply this model one day after another in real life, the length of this sliding window is one example of those decisions that need to be made.

On the other hand, trading strategies based on the selected forecast model are implemented according to some safety parameters. This means that decisions concerning when to execute a transaction are made depending on what the model predicts that might occur in the future, but these decisions can be ordered to be made only if a certain security threshold is overcome.

To define the optimum value for this threshold is also an important decision that needs to be made trying again to find an equilibrium between not assuming too much risk and obtaining enough profitability. The purpose of the risk analysis included in this chapter is to describe the methodology that an utility should follow when implementing the proposed strategy to make these decisions, and also to conclude the best parameters for implementing these models using a test dataset for simulation of profits and risk assumed by the trading strategy in real applications.

In the same way as it was done in the previous analysis, the methodology described in this section will implement a full VECM tranding strategy on BRENT and WTI daily spot prices. As these two variables were previously identified to be cointegrated, they can be used to build an example of how a VECM trading strategy should be designed from the perspective of an utility.

6.1 Inputs of the model

Once it has been concluded that a trading strategy based on VECM models is more reliable than one based on ARMA-GARCH models mainly due to volatility triggered by the last one, the methodology followed to build a full VECM strategy within this section. Options already included in the model were:

- The length of the sliding training window.
- The divergence to be overcome to execute transactions, defined as a value times the standard deviation of the spread of long-term relationship between prices of commodities.

Apart from them, a new option is included in the model in this section: the option to take into account or not structural changes in the series. If this option is activated, predictions for each time period are computed based on the most recent data among the sliding window with no structural change.

Input data of the numeric example that is included in this section to explain the methodology to be followed cover daily spot prices of BRENT and WTI oil from January 1997 to February 2019.

6.2 Initial steps for building a VECM

6.2.1 Cointegration test

A VECM of two or more time series can only be built if certain conditions are fulfilled. Firstly, the series considered must have a cointegration relationship, and this needs to be proved by some kind of cointegration test. A CADF test is used for proving cointegration between daily spot prices of BRENT and WTI from 1997 to 2019.

As in previous introductory analysis, this test is carried out for determining if the residuals from a linear combination between the two time series considered with data from an sufficiently long period of time is stationary. In this case, both time series can be considered to be cointegrated over time. This CADF test was executed in R and cointegration was proved with a p-value lower than 0,01.

6.2.2 Determining an adequate size of the training window

Another study that needs to be done is to choose the length of the sliding training window to be used to fit the VECM. A helpful criteria to consider concerning this is the number of transactions with losses that arise as a consequence of the size of the training window, shown in Figure 6.1.

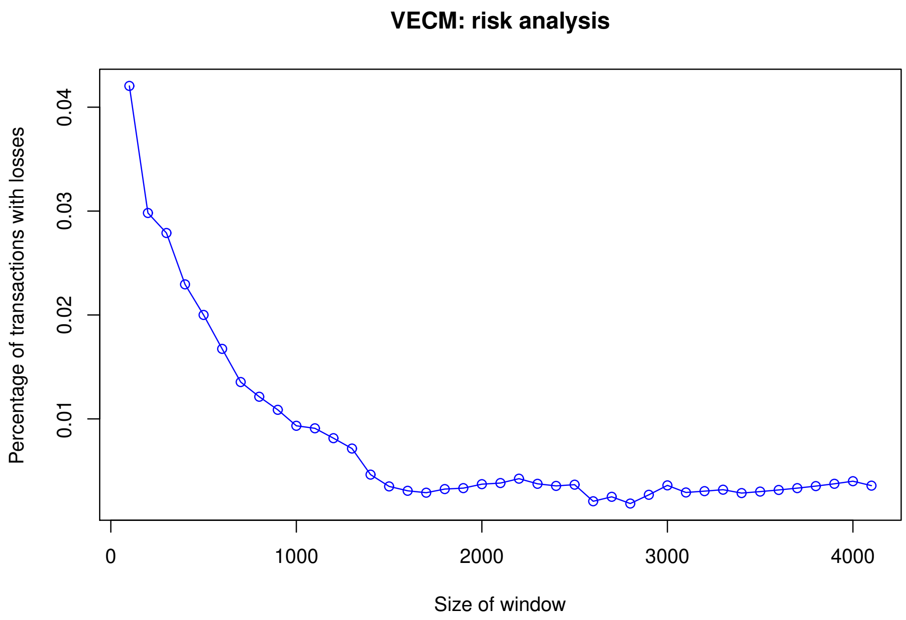

Figure 6.1: *Selection of the size of the training window in a VECM strategy*

The percentage of transactions with economic losses decreases exponentially with the increment of the size of the training window used in the trading strategy. Looking at the evolution shown in the graph, an adequate size of the window would be 1000 days, as the slope with which the percentage of transactions with losses decreases is significantly lower when the size overcomes this value.

6.3 Definition of the automatic trading strategy

Once the beta coefficient representing the long-term cointegration between considered variables is computed from the VECM model, the statistical factor z is calculated followinf the equations defined in the previous chapter:

$$spread_t = WTI_t - \beta BRENT_t \tag{6.1}$$

$$z = \frac{spread_t - \mu_{spread,training}}{\sigma_{spread,training}} \tag{6.2}$$

If z decreases below a defined value, it means that WTI price is below its long-term relationship with BRENT price. The safety factor represented by the statistical factor z determines how many times the standard deviation the difference between the current spread and the mean spread during the training horizon is. If the threshold defined for this difference is overcome, the algorithm sells a barrel of BRENT oil and buys another one of WTI price. Then, when z goes above 0 again, the long-term relationship is recovered and the position is undone.

6.4 Implementation of the trading strategy and improvements

The implentation of the trading strategy can be summarised as follows. If time instants in which structural changes of the series occur are found among the sliding window, only data from the last structural change in the sereis to the instant *t-1* are taken into account for fitting the VECM model and forecasting the next commodities prices.

In order to avoid underfitting, a minimum size of the sample selected once structural changes are found is imposed to be 10% of total length of the sliding window. Furthermore, the maximum length of the sliding window has been reduced to 1000 days instead of 2500 according to the criteria used to define the size of the window explained above.

However, the computational time increases hugely if structural changes are evaluated on a daily basis. To avoid this problem, the model was modified to search structural changes using monthly data. Once the months in which structural changes occur are identified, thay are considered in the execution of the daily strategy by removing data previous to the last structural change in the sliding window. However, a minimum size of 100 days in the sliding window is imposed to avoid underfitting too.

6.5 Results

While profits obtained by the trading strategy has a nearly constant and negative slope with respect to z, the number of transactions decreases exponentially. Both curves are represented in Figure 6.2.

What can be deduced from the graph is that the criteria used for determining the selection of a value for z must be the number of transaction with losses, as the slopes decreases exponentially. Apparently, the maximum difference between the curves happen when **z equals to 0,85**. Selecting this value for validating the trading strategy, its performance is analysed through some of the most commonly used ratios when evaluating automatic trading strategies.

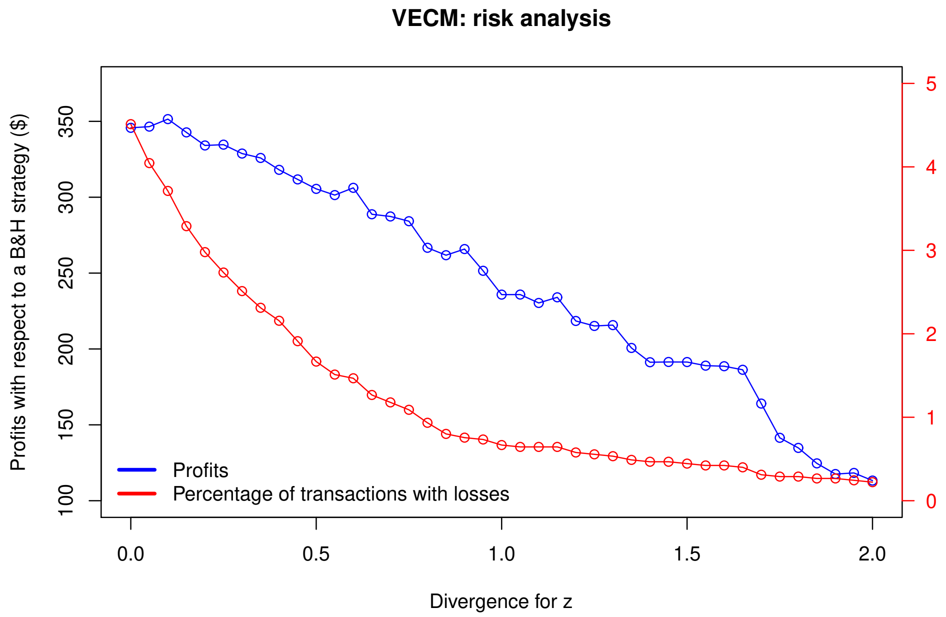

Figure 6.2: *Results for full model implementation with BRENT and WTI*

6.6 Evaluation of the trading strategy

A real simulation has been carried out by implementing the VECM strategy built in this chapter with real BRENT and WTI daily spot prices that go from 2001 to 2017. The profits obtained with respect to a Buy and Hold strategy are illustrated in Figure 6.3.

To analyse the financial performance of the trading strategy, Sharpe Ratio is calculated. In order to do so, yearly excess returns are computed considering a risk-free ratio of 3%. Then, the average excess return and the standard deviation of excess returns are calculated using data from Figure 6.4. The average excess return is 13,62% and the resulting standard deviation is 15,14%. Once these parameters are defined, the **Sharpe Ratio** is calculated.

$$Sharpe = \frac{R - R_f}{\sigma} = \frac{13,62\% - 3\%}{15,14\%} = 0,7013 \tag{6.3}$$

Two important aspects can be observed in Figure 6.4. Firstly, excess returns obtained by applying the VECM strategy are quite high from the beginning until 2009. However, although the value obtained for the sharpe ratio is not bad at all returns decrease after this year except for 2013 and 2018 and the average and the sharpe ratio obtained are significantly reduced because of this fact.

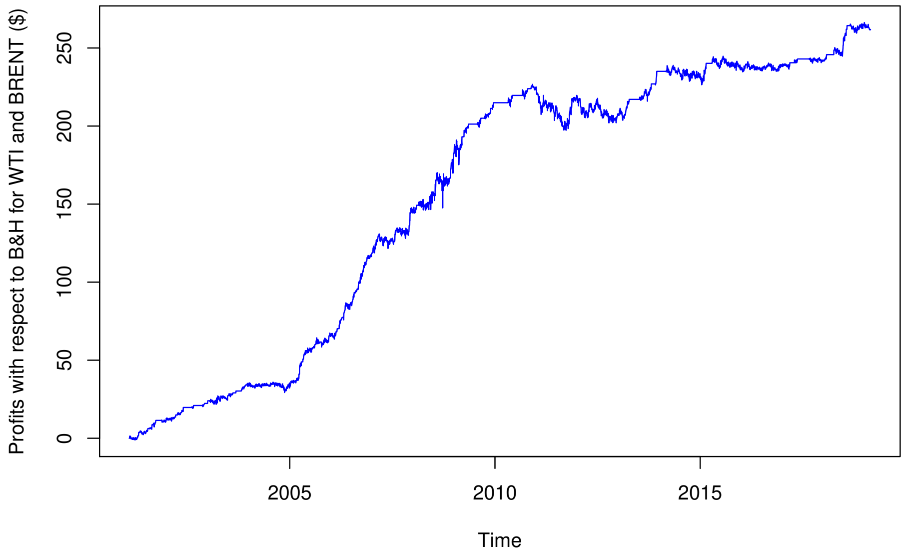

Figure 6.3: *Profits obtained with respect to Buy and Hold*

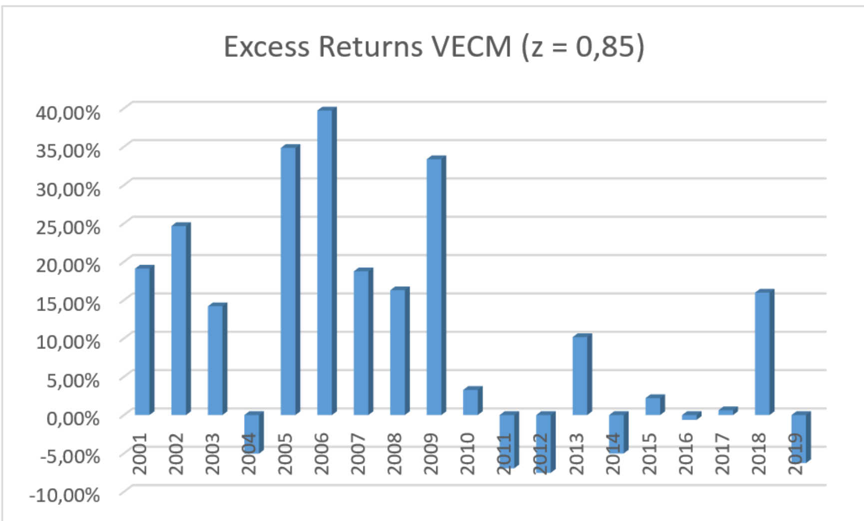

Figure 6.4: *Excess returns, VECM strategy for BRENT and WTI with z = 0,85*

To explain this, Figure 6.5 represents the values calculated for the spread used in the trading strategy for the whole time period considered in the execution compared with the threshold defined as 0,85 for z.

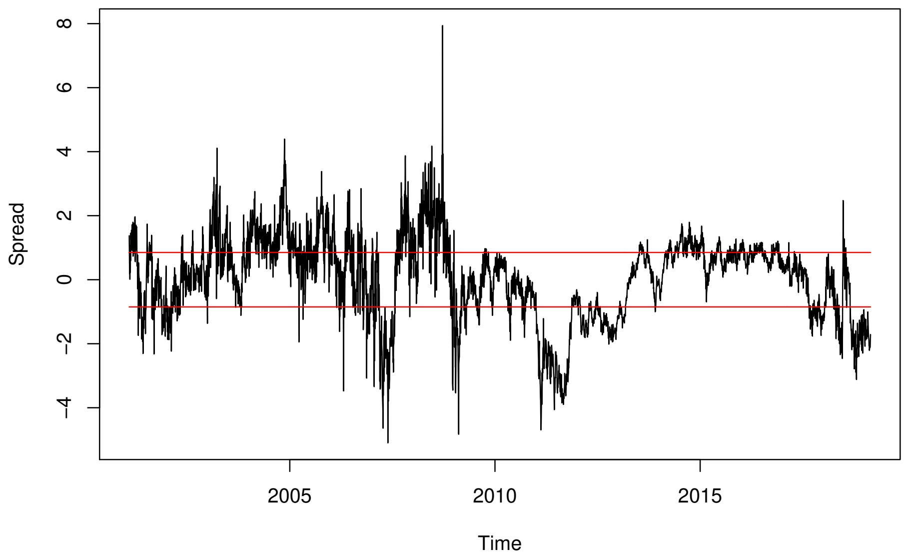

Figure 6.5: *Values obtained for the normalized spread; z = 0,85*

What can be observed in the graph representing values of the spread in the implementation of the VECM strategy is that in the second half of the period the values obtained for the spread are significantly lower than the ones previous to the year 2010, which leads to a lower amount of days in which the spread overcomes the threshold and when it does the values are very close to 0,85 in the majority of cases. As a consequence, a lower number of transactions are executed in this period and those that are executed are less profitable, which is also reflected in Figure 6.3, where the profits obtained in the first 10 years are much higher than after 2010.

Nevertheless, the worse results obtained in the second interval of the trading strategy are not a direct consequence of defects of the model, but they are instead caused by the complexity that time series of commodities present as some intervals may experience high spreads with lots of structural changes while other ones are characterized by low spread with respect to the long-term cointegrated relationship between variables.

The Sharpe Ratio is useful for evaluating the performance of a given trading strategy. However, its main weakness is that it penalizes the overall volatility and therefore both unusually high returns and usually negative returns are equally penalised.

If we consider only downside volatility is bad for an investor, the **Sortino Ratio** is a better option for evaluating the trading strategy. This ratio is similar to the Sharpe Ratio, but considers only the standard deviation of negative returns. In the case of Figure 6.4, the standard deviation of negative returns (below zero) is 2,47%.

$$Sortino = \frac{13,62\% - 3\%}{2,47\%} = 4,2929 \tag{6.4}$$

The result obtained for the Sortino Ratio is much greater than 1, which means a good performance of the trading strategy with much higher returns than harmful volatility. As the above result reflects, the difference between the Sharpe Ratio and the Sortino Ratio can be very high in some cases in which volatility of returns is caused mainly by positive returns. To use one of the above ratios to evaluate the performance of a strategy depend on the preference of the investor in terms of risk and return. If an investor prefers lower risk and constant returns the Sharpe Ratio is a good alternative for evaluating strategies, as it penalizes volatility of both positive and negative returns.

However, there are some markets in which unpredictability and complexity of prices might make it impossible to expect flat curves concerning returns. This is the case of energy markets, with lots of structural changes and complexities. Therefore, in this kind of markets it makes sense to use the Sortino Ratio to penalize only volatility of returns below a defined target, which can be zero or another objective value, although it is common to use a target of zero when considering excess returns with respect to the risk-free rate.

The value obtained for the Sortino Ratio reflects it makes sense in this case to differentiate overall volatility from downside volatility. The extremely high value of the Sortino Ratio means the most of the volatility presented by the excess returns of the VECM strategy is caused by positive returns.

7 Conclusions

Several conclusions can be extracted from the analysis that have been carried out. Firstly, it should be taken into account how complex these markets are, full of structural changes as they can be affected by lots of externalities such as political decisions, economic variables, strategic decisions adopted by powerful regions, etc. Because of this, energy companies require robust and reliable algorithms to maximize their profit when investing in commodity markets. Therefore, in this thesis, a complete methodology for developing automatic trading strategies in energy markets have been proposed.

To achieve this, trading algorithms need to be based on statistical forecasting models. A lot of quantitative models have been studied in this document. The applications that these models can have in building an automatic tool for developing trading strategies can be very interesting either for an energy company or an agent who frequently invests in international energy markets. In this context, it is relevant to study the statistical relationships between commodities. For those time series that appear to be cointegrated, error correction models are usually an accurate technique. On the other hand, univariate models that incorporate the existing correlation relationships between commodities are also an interesting alternative as they have the advantages of implementing a forecasting model which depends on historical values of the series while capturing at the same time the interdependencies of the series considering the correlation obtained in the training dataset.

Concerning commodity markets, a VECM trading strategy seems to represent the best alternative for an agent interested in investing in cointegrated indexes. However, several decisions need to be taken depending on the context. For instance, risk aversion is a relevant aspect. Because of this, the tool that has been developed in this thesis provides the agent with the ability to establish a threshold to find an equilibrium between expected profit and risk aversion.

The statistical analysis carried out in chapters 4 and 5 give rise to several conclusions as well. This analysis led to two types of relevant outcomes. Firstly, some of the commodities that have been studied presented stable correlations for different calibration windows considered. To prove these correlation relationships are due to cointegration between commodities, a cointegration analysis is necessary. The results obtained suggest that the following indexes are cointegrated for the analyzed period with a 95% of confidence level: BRENT and WTI; NCG and PEGN; NCG and TTF; PEGN and TTF; and NBP and ZEE.

On the other hand, other outomes were suggested by the analysis carried out regarding how the different forecasting techniques existing in the literature perform when applied to trading strategies in commodity markets. In chapter 4, three different methods are compared in terms of accuracy for monthly prices forecasts. These three methods consisted in the Box and Jenkins methodology, an ARMA-GARCH model and a VECM. The first two approaches included correlated data simulations for those commodities that presented stable correlations with different calibration windows in the training dataset. From this analysis, it can be concluded that predictions in the test dataset from ARMA-GARCH and VECM models were much more accurate than the first approach. Therefore, these two models were the ones selected for developing automatic trading strategies in these energy markets in the next chapter.

In chapter 5, the selected two approaches were compared when they are applied for building trading strategies algorithms. These strategies consider the forecast outputs from quantitative models to take decisions of buying and selling in commodity markets (in this case, oil barrels). This comparison was based on BRENT and WTI daily spot prices in a prediction horizon of 500 days. The results obtained suggested that a VECM strategy was more reliable and gave rise to a lower amount of transactions with losses and a higher return (around 50% of the initial investment in 500 days). Because of this, the full methodology was applied in chapter 6 to develop real automatic VECM trading strategies. This methodology includes structural changes analysis to ensure investments decisions are based on predictions that consider information provided by past values among which no structural changes take place in the series.

Overall, the methodology that is required to be followed in order to develop a robust algorithm for developing automatic trading strategies is complex and requires a high statistical knowledge. It also necessary to carry out comparative analysis concerning the adequacy of each type of forecasting technique for capturing the dynamics of the markets in which trading strategies are wanted to be developed. However, it is worthy to follow this type of methodology as the results that have been obtained in this thesis suggest it allows us to build powerful tools with the ability of capturing the dynamics of such complex markets. Commodities are essential in the energy sector and investing in this type of markets is crucial for an utility. Therefore, the methodology that is proposed illustrates the steps that can be followed by an energy company for taking advantage of the financial profits that this type of algorithms based on complex statistical models can give rise to even when being applied to indexes that are likely to present structural changes.

8 Future work

An alternative that is suggested as future work for complementing the model purposed in this thesis is introducing a component of **sentiment analysis**. This technique has become relevant since the beginning of this century and it mainly consists in an alternative method to discover future trend of a stock by taking into account news articles about the company and trying to classify positive and negative news. Therefore, this technique has a huge potential and it could also be applied for adding information to forecasts in commodities and energy markets.

Sentiment analysis can be carried out on two levels: document level or sentence level. The first one focuses on classifying the whole document, for instance financial news documents concerning the forecasted variable, as positive, negative or neutral. Meanwhile, on the sentence level analysis each sentence of all analysed documents is classified as positive or negative. Sentence level is becoming in fact very popular because the popularity of twitter, and many research projects carrying out sentiment analysis towards a certain topic have been based on information from this micro-blogging site.

On the other hand, there are mainly two methods to carry out the sentiment analysis: unsupervised learning and supervised learning method. The main difference between them is that unsupervised learning does not require to build a classification model using part of data as training set to classify the rest of the data or test set.

Unsupervised learning methods consist mainly in a dictionary based approach, as a list of words is used to count positive and negative words and determine the sentiment of the document or sentence based on the difference between these two types of words. Instead of this, supervised learning methods use a training dataset to build machine learning models validated by a test dataset to classify the data into different categories. Some of the most successfully tested classifier algorithms are Support Vector Machines, Naïve Bayes and Random Forest.

Other guidelines that can be suggested as future work include to validate the complex models that have been analyzed with all the commodities and to study cointegration relationships for different time periods. In addition, trading strategies with a basket of commodities could also be interesting if more than two commodities could be proved to be cointegrated for different periods.

To sum up, several alternatives can be proposed as future work:

- Firstly, sentiment analysis is a powerful tool that could complement the models described in this document.
- On the other hand, other tasks that could be further developed are related to the

real implementation of a VECM trading strategy that has been applied with BRENT daily spot prices. A future work concerning this issue is to extend the proposed methodology for the rest of cointegrated variables and build correlated data simulation based on univariate models for non-cointegrated ones.

Therefore, a trading agent can find in the different chapters of this document a detailed methodology for developing automatic trading strategies based on the most appropiate quantitative models according to statistical properties of the series in which the agent might be interested. In addition, the guidelines proposed as future work in this section could also serve as potential improvements for complementing the methodology proposed in this thesis.

9 References

[1] *Nonrenewable resources definition*. Available: https://www.investopedia.com/terms/n/nonrenewableresource.asp

[2] *BP Statistical Review of World Energy*, BP, 2019.

[3] *World Energy Outlook 2018*, International Energy Agency (IEA), 2019.

[4] *Oil: How the Markets Dynamics Have Changed*. Available: https://www.cmegroup.com/education/featured-reports/oil-how-the-market-dynamics-have-changed.html

[5] *Are Crude Oil and Natural Gas Prices Linked?*. Available: https://www.cmegroup.com/education/articles-and-reports/are-crude-oil-natural-gas-prices-linked.html

[6] José Alberto Mauricio, *Introducción al análisis de series temporales*, Universidad Complutense de Madrid.

[7] *Definition of Quantitative Forecasting Methods*. Available: https://www.chegg.com/homework-help/definitions/quantitative-forecasting-methods-31

[8] Kaplan, Andreas and Haenlein, Michael; *Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence*, Business Horizons. 62 (1): 15–25. January, 2019.

[9] Amir Mosavi, Mohsen Salimi, Sina Faizollahzadeh Ardabili, Timon Rabczuk, Shahaboddin Shamshirband and Annamaria R. Varkonyi-Koczy *State of the Art of Machine Learning Models in Energy Systems, a Systematic Review*, Energies. 2019.

[10] *Training an Artificial Neural Network - Intro*. March, 2004. Available: https://www.solver.com/training-artificial-neural-network-intro

[11] Thorsten Joachims, *A Statistical Learning Model of Text Classification for Support Vector Machines*, 2001.

[12] Nowrouz Kohzadi; Milton S. Boyd; Bahman Kermanshahi and Iebeling Kaastra *A comparison of artificial neural network and time series models for forecasting commodity prices*, Neurocomputing, Volume 10, Issue 2, March 1996, Pages 169-181.

[13] Siddhivinayak Kulkarni and Imad Haidar *Forecasting Model for Crude Oil Price Using Artificial Neural Networks and Commodity Futures Prices*, International Journal of Computer Science and Information Security (IJCSIS), June 2009 Issue, Vol. 2, No. 1.

[14] Lean Yu; Shouyang Wang and Kin Keung Lai *Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm*, Energy Economics, Volume 30, Issue 5, September 2008, Pages 2623-2635.

[15] Jinchao Li; Shaowen Zhu and Qianqian Wu *Monthly crude oil spot price forecasting using variational mode decomposition*, Energy Economics, Volume 83, September 2019, Pages 240-253.

[16] Shom Prasad Das and Sudarsan Padhy *A novel hybrid model using teaching–learning-based optimization and a support vector machine for commodity futures index forecasting*, International Journal of Machine Learning and Cybernetics, January 2018, Volume 9, Issue 1, pp 97–111.

[17] María Pilar González Casimiro, *Técnicas de predicción económica*, Universidad del País Vasco.

[18] Box, G. E. P. and Jenkins, G. M., *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco, 1976.

[19] M. Carmen Guisán, *Causalidad y cointegración en modelos econométricos. Contrastes DF, ADF, EG y otros*, University of Santiago de Compostela, 2002.

[20] Rana Abdullah Ahmedi and Ani Bin Shabri, *Daily Crude Oil Price Forecasting Model using ARIMA, Generalized Autoregressive Conditional Heteroscedastic and Support Vector Machines*, American Journal of Applied Sciences 11 (3): 425-432, 2014.

[21] Juan Gabriel Rodríguez Hernández, *Una introducción a los modelos de series temporales no lineales*, Universidad Rey Juan Carlos. May, 2001.

[22] Alfonso Novales, *Modelos ARCH univariantes y multivariantes*, Universidad Complutense de Madrid. September, 2013.

[23] S.S. McNees, *The Forecasting Record for the 1970's*, New England Economic Review. October, 1979.

[24] Robert F. Engle, *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation*, The Econometric Society. July, 1982.

[25] Tim Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics. February, 1986.

[26] Hassan Mohammadi and Lixian Su, *International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models*, Energy Economics, 2010, vol. 32, issue 5, 1001-1008

[27] Chin WenCheong, *Modeling and forecasting crude oil markets using ARCH-type models*, Research Centre of Mathematical Sciences, Faculty of Information Technology, Multimedia University, 63100 Cyberjaya, Selangor, Malaysia. March, 2009.

[28] Olga Efimova and Apostolos Serletis, *Energy markets volatility modelling using GARCH*, Energy Economics Volume 43, May 2014, Pages 264-273.

[29] *Spurious Correlations*. Available: http://www.burns.com/wcbspurcorl.htm

[30] Robert F. Engle and C. W. J. Granger, *Co-Integration and Error Correction: Representation, Estimation, and Testing*, Econometrica, Vol. 55, No. 2., pp. 251-276. March, 1987.

[31] Chakriya Bowman and Aasim M. Husain, *Forecasting Commodity Prices: Futures Versus Judgment*. March, 2004.

[32] Ardian Harri, Lanier Nalle and Darren Hudson, *The Relationship between Oil, Exchange Rates, and Commodity Prices*. Journal of Agricultural and Applied Economics. August, 2009.

[33] Yanan He, Shouyang Wang and Kin Keung Lai, *Global economic activity and crude oil prices: A cointegration analysis*. December, 2009.

[34] Saban Nazlioglu and Ugur Soytas, *Oil price, agricultural commodity prices, and the dollar: A panel cointegration and causality analysis*. Energy Economics 34 (2012) 1098–1104

[35] Ehsan U. Choudhri and Lawrence L. Schembri, *Productivity, commodity prices and the real exchange rate: The long-run behavior of the Canada–US exchange rate*, International Review of Economics and Finance, 2013.

[36] Emanuel Kohlscheen, *Long-run determinants of the Brazilian real: a closer look at commodities*. March, 2014.

[37] Niaz Bashiri Behmiri, Matteo Manera and Marcella Nicolini, *Understanding dynamic conditional correlations between commodities futures markets*. 2016.

[38] Zukarnain Zakaria and Sofian Shamsuddin, *Electricity Consumption and Economic Activity in Malaysia: Co-integration, Causality and Assessing the Forecasting Ability of the Vector Error Correction Model*, International Journal of Energy Economics and Policy, 2016, 6(4), 706-713.

[39] Samuel Asumadu-Sarkodie and Phebe Asantewaa Owusu, *Multivariate co-integration analysis of the Kaya factors in Ghana*, Environmental Science and Pollution Research, Volume 23, Issue 10, pp 9934–9943. May, 2016.

[40] Hillary Chijindu Ezeaku, Nwanneka Judith Modebe, Grace Chinyere Eje and Anthony E. Ageme, *Oil Price and Exchange Rate Nexus: A Vector Error Correction Approach on Nigeria*, International Journal of Energy Economics and Policy, 2017.

[41] David I.Harvey, Neil M.Kellard, Jakob B.Madsen and Mark E.Wohar, *Long-Run Commodity Prices, Economic Growth, and Interest Rates: 17th Century to the Present Day*. January, 2017.

[42] Corlise L. Le Roux, *Cointegration Between Energy Commodities and the South African Financial Market*. June, 2017.

[43] David A. Dickey and Wayne A. Fuller, *Distribution of the Estimators for Autoregressive Time Series With a Unit Root*, American Statistical Association. June, 1979.

[44] *Dickey-Fuller and Augmented Dickey-Fuller tests*. Available:
https://nwfsc-timeseries.github.io/atsa-labs/sec-boxjenkins-aug-dickey-fuller.html

[45] *Introduction to CADF and Johansen statistical tests for Cointegration*. Available:
https://www.quantshare.com/blog-542-introduction-to-cadf-and-johansen-statistical-tests-for-cointegration

[46] SørenJohansen, *Statistical analysis of cointegration vectors*, Journal of Economic Dynamics and Control, Volume 12, Issues 2–3, June–September 1988, Pages 231-254.

[47] Michael L. Halls-Moore, *Advanced Algorithmic Trading*, 2017.

[48] Jushan Bai and Pierre Perron *Computation and analysis of multiple structural change models*, Journal of Applied Econometrics, 2003, vol. 18, issue 1, 1-22.

[49] *Pairs Trading using Data-Driven Techniques: Simple Trading Strategies Part 3*. December, 2017. Available: https://medium.com/auquan/pairs-trading-data-science-7dbedafcfe5a

[50] Isak Aggeborn and Olof Hansson, *Pairs Trading: an Extension to the Cointegration Approach*. Uppsala University, Department of Statistics. Spring term, 2017.

[51] Hossein Rad, Rand Kwong Yew Low and Robert Faff, *The profitability of pairs trading strategies: distance, cointegration and copula methods*. Quantitative Finance, 16:10, 1541-1558. April, 2016.

[52] Burnham, K. P. and Anderson, D. R., *Model Selection and Multimodel Inference: A practical information-theoretic approach (2nd ed.)*, 2002.

[53] Akaike, H., *A new look at the statistical model identification*, 1974.

[54] *Ljung Box Test: Definition*. Available:
https://www.statisticshowto.datasciencecentral.com/ljung-box-test/