LUISS

Economics and Finance

Course of Econometric Theory

Trading strategies on the Italian Future Power Prices

Prof.Paolo Santucci de Magistris

SUPERVISOR

Prof.Federico Carlo Eugenio Carlini

CO-SUPERVISOR

Riccardi Giovanni ID 745571

CANDIDATE

Academic Year 2023/2024

ABSTRACT

The realm of energy trading is at its early stages, marked by the relative novelty of this domain. Despite being an nascent field, recent years have witnessed the formulation of various strategies in the market, encompassing both speculative approaches and the management of physical assets.Traditionally, the implementation of speculative strategies in the energy market has been considered peripheral, often standing alongside financial hedging trading practices. This form of trading involves companies seeking to mitigate risks associated with energy price fluctuations by safeguarding contracts through financial transactions.This thesis diverges from the conventional by developing a wholly speculative strategy, akin to bot trading prevalent in financial markets. The approach employed focuses on pure financial commodity trading to define profits. The research demonstrates that profitable strategies can be cultivated in the energy market through the exclusive utilization of statistical forecasting (employing ARIMA model estimations) and data modeling techniques.The financial instruments under scrutiny in this study, specifically Italian power futures Q121, Q221, Q321, Q421 quoted in 2020, and Q122, Q222, Q322, Q422quoted in 2021.The overarching vision of this thesis postulates that, in the forthcoming years, the energy market will increasingly resemble the stock market. Speculative strategies, detached from reliance on assets or client contracts, are anticipated to become the prevailing norm.

Table of Contents

List of Figures

List of Tables

INTRODUCTION

The objective of this thesis is to develop a profitable trading strategy based on the Italian power futures prices of the 2022 quarters and 2021 quoted in 2021 and 2020; a comparative analysis will be carried out to choose on which quarter to implement the trading strategy and obtain the desired ROI. The data that are going to be used are from 2020 and 2021 because the aim of this thesis is to find a trading strategy that can be used in the future to predict the market. Both data before 2020 and those after 2021 aren't enough good (the first because the market was stationary and the second because the fundamental price of the commodities is lower than the actual level). So the data of 2020 and 2021 are useful to understand the future because we will have a market that is going to be near the fundamental until news and other driver are going to create a bull pattern. The data of this two year are useful because represent this situation a period of stability that is going to be change by the geopolitics. The market in the future is not going to be stable as before and overreacting as it was in the past, its going to be at midpoint.It is important to think about the situation of Houti and Gaza.[1] .These two events would have created great bull movement in the past but since the market is now used to similar situations such as the Russian war, it is not overreacting. Often in energy companies especially in Italy a type of hedging or Market access trading is implemented to hedge contracts, speculative trading is less and less used especially after the numerous bankruptcies of energy companies that used speculative energy trading a lot during the beginning of the Russia Ukraine War. In this thesis, a trading strategy will be developed which, however, does not 100% represents the "quintessential" speculative trading strategy in the energy trading world; an attempt will be made to use a method that can also be reused in other return-based asset classes.The "typical" strategy of energy trading instead is to receive orders for energy contracts (example PSV DA.[2] [3], several quarters etc.) "hold them in panic" and then decide when to execute them. In conventional energy trading, a common practice involves waiting for opportune moments to capitalize on more favorable prices, thereby yielding profit margins. For instance, when an order materializes for 1 GW of power for the first quarter of 2021 (Q121), traders typically opt to delay transactions until prices are deemed advantageous. However, this thesis pioneers an alternative methodology. This innovative approach commences with meticulous data initialization, wherein dates and daily prices are meticulously recorded. Specifically, prices and data pertaining to the quarters of 2021 are quoted in 2020, while those for the quarters of 2022 are quoted in 2021. To validate the efficacy of forecasting models, the thesis employs a rigorous cross-validation methodology. Cross-validation encompasses resampling and sample splitting methodologies, leveraging distinct segments of the dataset for both model training and testing across numerous iterations. This technique is particularly prevalent in predictive contexts, where the primary aim is to evaluate the model's predictive accuracy in real-world scenarios. Typically, the model is trained on a dataset comprising known data, termed the training dataset, and subsequently evaluated on a dataset of unseen or first-seen data, known as the validation dataset or testing set. The overarch-

[1] https://ctc.westpoint.edu/houthi-war-machine-guerrilla-war-state-capture/

[2] The PSV DA is the PSV spot

[3] The acronym PSV stands for Virtual Exchange Point, the place where the exchange or sale of natural gas takes place, the meeting point between gas supply and demand in Italy

ing objective of cross-validation transcends mere model validation; it serves to uncover potential issues such as overfitting or selection bias, thus providing insights into the model's generalizability to unseen data. By subjecting the predictive model to diverse datasets, including those unseen during the training phase, cross-validation offers a robust framework for assessing the model's reliability and performance in practical energy trading scenarios. First, the price returns are going to be estimate over the duration of the sample. After that, the log returns will be estimated. The third step and modeling using the ARIMA as done by Newbold (1983), Box et al. (2015) models, to estimate the different forecasts. There will be the comparisons of the forecasts on the graph and then the mean square deviation using the return that occurred as a reference, from there than the mean square deviation will be the best estimation since the smaller the mean square deviation the closer it is to the reference return.The mean square errors measures the extremely useful dispersion defined as the square root of the sum of the quadratic deviations from the arithmetic mean, divided by N. After that as done by Mincer and Zarnowitz (1969),Byun and Cho (2013),Boos and Grob (2023), we are going to do the Mincer and Zarnowitz regression on the various samples. The final step before trading strategies is the Diebold Mariano Test as done by Chen et al. (2021), which will be used to decree the best forecast. The best forecast will be used for the Trading strategy. The application will be done by defining two confidence intervals one lower and one higher and depending on whether the return will rise above or below the intervals there will be stock decisions. At the end of the process the ROE of the different quarters will be calculated. The best quarters with the best performance is going to be select. Prior to delving into practical applications, this thesis will be structured into five comprehensive chapters, each delving into distinct facets of energy trading and market dynamics. Chapter 1: FUNDAMENTALS OF ENERGY MARKET WITH FOCUS ON ITALIAN POWER MARKET: The inaugural chapter serves as a foundational exploration into the intricacies of power fundamentals, elucidating the mechanisms of supply and demand within the electricity market. Notably, it sheds light on the fundamentals of various European countries, offering insights into their energy infrastructure, regulatory frameworks, and market dynamics. A significant focal point of this chapter will be the discussion surrounding the challenge of market illiquidity, dissecting its underlying causes and potential ramifications on trading strategies and market efficiency. Chapter 2: RELATIONSHIP AND INFLUENCE OF GAS AND ITALIAN POWER MARKET: Chapter two shifts focus towards the domain of gas trading, providing a comprehensive analysis of its fundamentals. Special attention will be devoted to examining the geopolitical landscape, particularly the dynamics surrounding gas supply and distribution, with a keen eye on the situation involving Russia. Furthermore, this chapter will undertake an analysis of prominent gas trading hubs, such as PSV (Punto di Scambio Virtuale) and TTF (Title Transfer Facility). Chapter 3: METODOLOGIES FOR THE TRADING STRATEGIES: The third chapter offers an expansive overview of the statistical and econometric instruments that underpin the analysis and forecasting of energy markets. By delving into established methodologies such as ARIMA models and regression analyses, this chapter equips readers with the necessary tools to navigate the complexities of energy market data. Chapter 4:TRADING STRATEGIES: Chapter four marks a pivotal juncture in the thesis, wherein the focus shifts towards the development of trading strategies aimed at optimizing profitability. Through a synthesis of empirical data and theoretical frameworks, this chapter endeavors to identify the most lucrative quarters for energy trading. By leveraging insights gleaned from preceding chapters, such as market dynamics and statistical analyses, this section aims to formulate data-driven strategies tailored to exploit market inefficiencies and capitalize on emerging trends. Chapter 5: CONCLUSION: The concluding chapter offers a retrospective analysis of the results obtained from the preceding chapters, contextualizing findings within the broader landscape of energy trading. Moreover, it proffers insights into potential avenues for future research and implementation, offering recommendations for refining existing methodologies and exploring untapped opportunities within the energy market. By synthesizing key takeaways and implications, this chapter aims to furnish readers with a comprehensive understanding of the intricacies of energy trading and pave the way for future advancements in the field.

Chapter 1

FUNDAMENTALS OF ENERGY MARKET WITH FOCUS ON ITALIAN POWER MARKET

1.1 Introduction

The Energy Market defines the power price (electricity price) of each country by relating the market's electricity demand to the electricity supply in the same market. In determining the price, a normal is often considered, which represents the different values of the fundamentals of both demand and supply. Some important variations, which may be both negative and positive, below and above the normal, lead to variations in the DA power price since the power price is developed the day before, through the study of the different fundamentals and the market news; analyzing these elements the price of the following day is developed. The Power price of each CWE country.[1] and of the other European countries is volatile and is represented by the supply-demand relationship which is also variable and difficult to predict in the medium to long term. When talking about the energy market it is very important to define that energy cannot be stored and, especially for the transportation of energy, you need specific laws. When they say that energy cannot be stored, it is thus partly true; on the other hand, you had better say that it cannot be stored for a reasonable cost. One way to store energy is to use and generate hydropower reserves. However, this cannot be done in all the countries since, as we will see later, it depends on various fundamentals and, above all, it must be said that to create an industry that produces hydropower energy and makes reserves, must need favorable weather conditions. About the storage depend, for sure the transportation can be done. The transport of electricity occurs through Kirchhoff's law[2]. Basically, this law states that the intensity on all nodes must be zero, the voltage in all loops must be zero. The transfer capacities that are available for the exchange among the states first need to generate the hypothesis and then they can take place. As concerns the NTC net capacity transfers, it is handled by ENTSOE (European Network System Operator for Electricity). In the first year of using European and continental energy markets, an auction mechanism was used. This method was very simple and caused so much inefficiency in the market and adverse inflows (as a matter of fact, there were coordination problems between the markets and the transmission capacities). Over the years, market participants have developed and implemented the possibility of trading among different areas; this method has been used by different players in the continental energy market under the name of market coupling.

1.2 Market Microstructure

In each country the market structure is made up of different purposes and different times scales. Is possible to distinguish three types of markets: 1) The intraday market and/or the balance mechanism: the balance mechanism consists of exchanges intraday and Spot and market players to make sure that there are energy exchanges at the balance between generation consumption of consumers; 2) The day-

[1] Central West Europe (covering Benelux, France and Germany)

[2] Kirchhoff's circuit laws are two equalities that deal with the current and potential difference (commonly known as voltage) in the lumped element model of electrical circuits. They were first described in 1845 by German physicist Gustav Kirchhoff. This generalized the work of Georg Ohm and preceded the work of James Clerk Maxwell. Widely used in electrical engineering, they are also called Kirchhoff's rules or simply Kirchhoff's laws. These laws can be applied in time and frequency domains and form the basis for network analysis. Both of Kirchhoff's laws can be understood as corollaries of Maxwell's equations in the low-frequency limit. They are accurate for DC circuits, and for AC circuits at frequencies where the wavelengths of electromagnetic radiation are very large compared to the circuits.

ahead market: quantities are traded the day before being deliberated for the next 24/48 hours of the following day. 3) The futures market: in this case, companies, individuals, investment banks, those who are part of the market can trade energy for future times, for example, they may decide to trade a defined amount of energy for the coming quarter. Focus on the Intraday market is can discovered that. In the short term, energy market participants will define the price and the balance between generation and consumption of their portfolio. The short time frame taken into consideration will be approximately 12 hours. The TSO will be more focused on making sure that there is a balance between generation and consumption within the system. Thus, within the intraday market, two systems coexist: the first one is designated by the TSO which takes into account the budget mechanism that adjusts the generation and the consumption while the other system is composed by cost cutters who trade and generate to reduce and satisfy their own energy needs. The day ahead structure is about the day ahead market that is based on a fixed supply of trading. Every day before 12 a.m. Market participants begin to insert some Bids which will then determine the price for the following day . In the European Union each country has its own day-ahead market. If there's no coordination, prices among the countries will be too different; for this reason, to make countries have the same prices market coupling is used.

1.3 Electricity Demand Fundamentals

Electricity demand is a very important factor in price setting; it can be defined as the main driver of the daily settlement of a DA spot price in Europe. It depends on various factors such as temperature, spot energy prices, general energy sentiment, geopolitics, news and other less important factors. Of course, each country is influenced by these drivers, in different proportions. In France, for example, an increase in temperature in summer or a decrease in winter can drive up the demand and consequently the price since heating in France is fully electric and therefore temperature is a very important driver. The ratio in the transalpine country is in summer for every degree increase in temperature and a rise in consumption by 2000 MWh/h (2 GWh/h). In summer, above-normal temperature rises lead to proportional increases in demand. Of course, if the other fundamentals led to a bullish view, the price would spike high in the market. In winter, on the other hand, below-normal temperatures will lead to price spikes because electricity demand will be pushed up (by the use of radiators and other types of heaters). In this time of the year, the industrial demand is the main driver in the market demand so much so that it is crucial to define what the normal of industrial consumption is in a country because if the Gaussian were to be overestimated or underestimated, the DA Power price that would be defined would be completely wrong. Industrial demand in recent years has decreased due to the relocation of many European industries: the crisis, the covid and the increase in energy prices following the Russian invasion of Ukraine.[3] have led the electricity market to prices that had never been reached before due to an extremely high speculation and the crisis of fundamentals that had only predicted a bullish market trend (low gas storage, dependence on Russia, sanctions against Russia, intrinsic fear in the market, etc.).

[3] The 2022 Russian invasion of Ukraine is the military offensive initiated by the Armed Forces of the Russian Federation on February 24, 2022, invading Ukrainian territory and thus marking an abrupt escalation of the ongoing Russian-Ukrainian conflict since 2014

1.4 Electricity supply fundamentals

The electricity supply must be able to meet the electricity demand. It may happen that at certain times, the energy required to meet electricity demand is not attainable given non-productive fundamentals in the different European countries. So, the goal for a country is to have much higher installed capacities than demands so that there are no periods where demand cannot be met and the price becomes unmanageable. Here it is also important for a country to study the balance between consumption and generation in real time since an imbalance can lead to blackouts and production stoppages on the side of companies and stops in consumption on the side of consumers. Talking about energy reserves, they are capacity generators that can be used given advance notice. This, however, depends on the response time of the generators. There are various types of reserves: the primary reserve and the secondary reserve, in some cases even a tertiary reserve. As concerns the primary reserve, it can be mobilized in less than 15 minutes and, in this case, the national system operator will place the order. The secondary part is after 15 minutes. The tertiary reserve, on the other hand, consists of two parts: the first one is fast and can be mobilized in 15 minutes, up to an hour and can be a complementary factor to the secondary demand. To give an example, in France, the minimum demand values per hour are 50 gigawatts in no -peak hours up to 100 gigawatts in peak hours in summer. The fundamentals of electricity supply vary among the different European countries and the CWE ones. We can distinguish between two types of supply: one is the renewable supply, which comes from clean, the so-called green, sources; the other supply comes from nuclear or continuous-cycle, non-renewable sources (coal- red power stations, gas- red power stations and waste-to-energy plants). The peculiarities of the two offers are: 1).The main sources of renewable production.[4] are wind, solar and hydropower . The wind supply is divided into onshore and offshore wind. Renewable supply, for example, energy derived from renewable sources, is characterized by its dependence on weather factors. Of course, renewable energy sources are neither a closed-cycle or a continuous cycle; they are dependent on weather conditions. Wind depends on the windiness, so the construction of these types of renewable sources usually takes place in areas characterized by a high presence of constant and strong wind. If the weather conditions do not foresee a high wind, naturally the production will be much lower and non-performing, and therefore there will be a non-continuous production. To give a concrete example, Germany is a country that has focused so much on renewable energy sources, especially on wind power, so when there are weeks of high wind strength in Germany, the power price can even become negative because there is a disproportionate supply compared to a much lower demand. Solar power also depends on weather conditions, especially on the intensity of the sun and of the temperature. It is important for photovoltaic panels to be located in areas where the presence of sunlight is constantly high, for example, areas that are not cloudy but have a high degree of solar penetration. The temperature, on the other hand, can be a bearish factor for the performance of solar panels because, since the semiconductors inside the silicon solar panels are made of silicon, high temperatures may reduce the performance of the solar panels Hydrological production depends on

[4]This is an important driver of the prices too, because when there is a great production by renewable energy the prices are low because the cost of renewable energy are lower than other sources

rainfall, on the level of rivers and the amount of snow in the Alps and the Apennines (always referring to Italy). 2.)Continuous cycle production such as nuclear power and non-renewable energies like gas and coal are certainly more stable and more reliable as sources of production.They are internal cycles ; this means that production takes place entirely within the power plant and is not often influenced by exogenous factors. The only factors that can influence them are the CCS and DSS (clean spark spreads and dark spark spreads), which for gas and coal-fired power stations represent the production costs for gas and the dark represents the production costs for coal-fired power stations.Of the three types of production, nuclear power plants are the least reliable and can lead to very high market spikes, as they are producers that can satisfy the needs of entire municipalities and entire cities and therefore any planned or unplanned maintenance can lead to very high supply spikes, which, especially in situations of sustained demand, both for temperatures well above normal in summer and for temperatures well below normal in winter, lead to increases in demand and in price France is the perfect example for analyzing the nuclear power plant supply situation as it is very dependent on nuclear production so if several nuclear power plants are not in production, the French Power price will certainly rocket upwards creating high market mirrors and a very bullish Power price. In Europe over the past few years there has been a trend, which is being reinforced this year, to increase renewable energy sources in all countries except for France, where there are plans to increase nuclear production and not support renewables, which, on the contrary, in Germany and Spain have now become the country's main production. Interconnections.[5] are a fundamental supply driver that is very often overlooked because we tend to erroneously consider only the supply from renewables and non-renewables, high and low efficiency, oil, etc. Interconnections, on the contrary, are a fundamental supply driver for the energy sector and the definition of the price. An example is given by the relationship between France and Germany, so when France has days when demand is high and nuclear supply is in crisis, Germany, if it has a high level of sustained supply, can support France and balance and support the price so that it doesn't rise much. The interconnections are valid in all European countries; another important example may be given by France and Great Britain, the two countries can indeed support each other if Great Britain has a lower demand in high supply and Germany vice versa has a high demand in low supply, Great Britain will be able to pass energy and support France thus stabilizing the Power price .

[5] Interconnections are also important for trading strategies such as on the JAO platform where one can buy cables for days, weeks and months between one country and another and thus has the right to transfer Power from one country to another and therefore speculate on the price spikes that can occur in one or more countries and on the spread between the two countries sharing cable bought by the trader or the trading company willing to speculate on the spreads between prices.

1.5 Fundamentals: France

France is one of the most important countries within the European Union with the second largest population after Germany. France represents a peculiar situation within the European energy scene as it is one of the main industrial powers in Europe, it has many consumers and is dependent on nuclear energy. Nuclear power is the most important source thanks to which France is able to meet its power consumption needs. This peculiarity may also represent a challenge for the transalpine country because the French demand is a consistent one that depends on the variability of the temperature (indeed all systems , even the domestic ones, are completely electric). This is another characteristic, even compared to Italy, which mostly has domestic gas sources. In summer, for every degree above normal you have 1000 megawatts more of consumption and in winter when it's -1° you'll have 1000 megawatts more of consumption. The other renewable sources in France are not as important as nuclear power, which accounts for almost 60% of production. As concerns wind, this is in any case quite interconnected with German wind in the sense that they very often have the same variations, perhaps being in nearby areas. Another country, which is very much interconnected with France, is Great Britain, which is an equally important country for the production of energy consumption within the European continent.

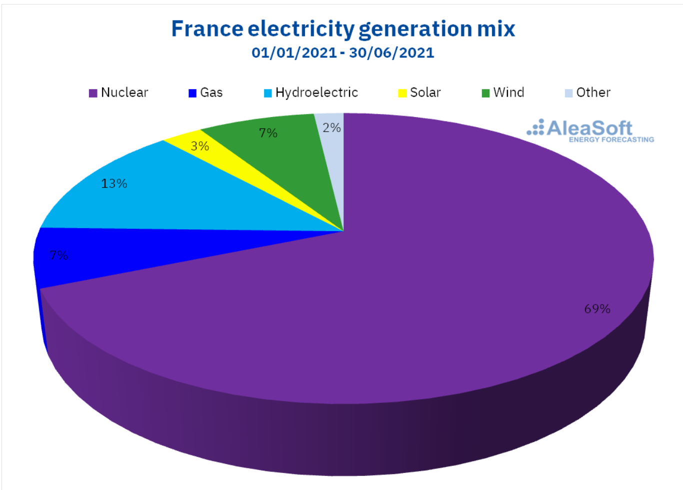

Figure 1.1: In this graph, the sources for electric generation in France are depicted. The violet segment represents nuclear production, which peaked at 69% of total production in 2021. Gas is represented by the blue segment, accounting for 7% of the total. Hydropower is represented by the light blue segment, constituting 13% of the total. Solar energy is depicted by the yellow segment, representing 3% of the total, while wind energy is shown in the green segment, comprising 7%. The remaining 2% is attributed to other sources, represented by the remaining portion of the graph.Graph taken by https://aleasoft.com/it/

The peculiarity of France is that it has an inelastic market; it has macro reactions which sometimes are exaggerated even towards soft movements of the nuclear supply. As a consequence, the main bullish factor in Europe, which can lead to changes in the market, depends on the constant and continuous maintenance of the French power plant, which in any case does not give great stability to the European power price . In recent times, however, the low level of nuclear production has also been helped by a contraction in demand. Demand in France has been fairly high, although it has recovered since last spring, whereas before that it used to be fairly below or close to the norm. This is also due to prices, which two year reached new heights and historical records as bullish prices, especially after the war in Ukraine and the sanctions against Russia, which brought uncertainty and above all a bullish wave to the European Power market. So to cut a long story short, the fundamentals in France are represented by the nuclear offer.

1.6 Fundamentals: Germany

Germany is the locomotive of Europe with the highest industrial production on the continent, the highest industrial demand on the continent and above all the country with the highest population within Europe. The Teutonic country alone has an electricity demand equal to half of the entire European electricity demand, so the German power price manages to influence the entire European power price, being by far the most important country and the country that manages to converge the most in terms of energy production. In recent years, Germany has been known to invest heavily in wind and solar technology and this has led it to cover half of its production with renewable technology in the spring and in the windiest months, when prices are even negative. This is not strange because Germany has a wind production force that can satisfy and lower the power price, even if it is bullish in France and in other interconnections and above all it can bring down the power price in its own country. So the effect of this export benefits other countries within the continent Central Europe when Germany's strong wind power on average has a price at a much lower price than expected because this is also a low-cost energy being renewable. In addition, last year, it has been decided to close the nuclear power plants, which represented about two and a half gigawatts of production per hour, and also to reduce the production of coal-fired power stations (coal used to be an important production tool for Germany).

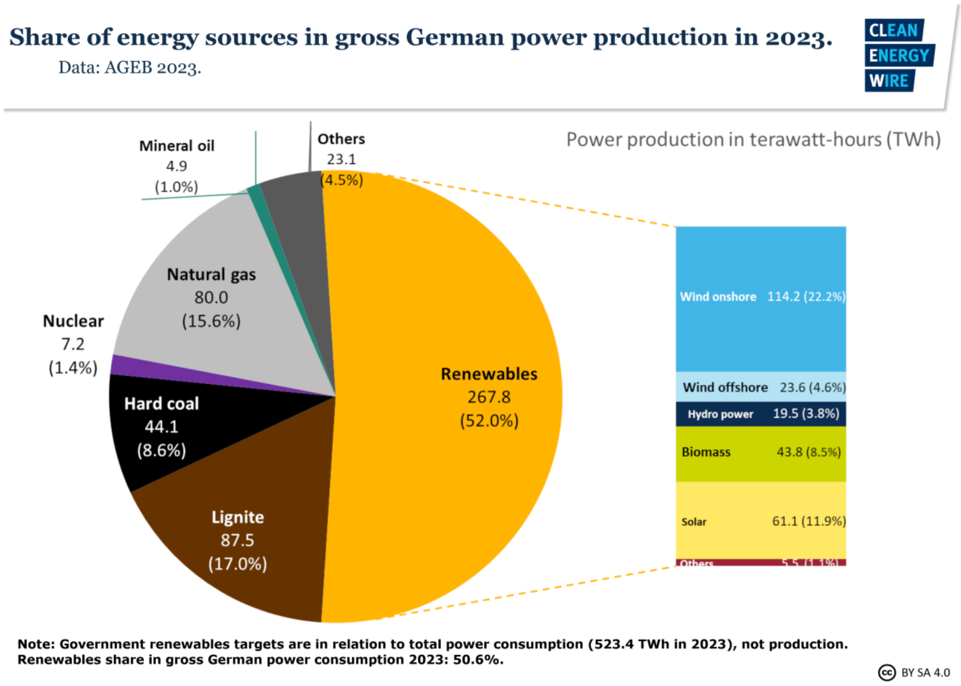

Figure 1.2: In this graph,is defined the shares of energy sources in gross German power production in 2023. The violet segment represents nuclear production, which peaked at 7,2% of total production in 2023. Lignite is represented by the brown segment, accounting for 17% of the total.Hard Coal is represented by the light black segment, constituting 8,6% of the total. Renewables energy is depicted by the orange segment, representing 52% of the total. Gas in grey is 15,6% and Mineral oil is 1% ,other 4,5% .Graph from https://www.cleanenergywire.org/factsheets/germanys-energy-consumption-and-power-mix-charts

The German country has become a point of reference for other European countries, which are trying to shake off old, highly polluting productions and achieve low-cost productions with a very low or even zero environmental impact. However, as mentioned earlier, this means that Germany is also very dependent on weather conditions, even though Germany can reopen coal-fired power stations and repair everything in low-wind conditions if it wishes, but this would lead to heavy pollution. Anyway, as it has just said, summer 2024 and the first three months of 2025 will be crucial, because any shocks caused by low nuclear offers in France could drive the market into a bullish perspective and Germany will not be able to offset the market upturn with wind power production. As is possible to see in the figure Germany has particular fundamentals, which are a strong presence of renewables, but, along with this, a strong presence of coal-fired and gas-fired power station productions.

1.7 Fundamentals: Italy

Consumption in Italy is very similar to that in the UK, both in terms of population and type of consumption. The Italian production share is characterized by a high dependence on the gas source. About half of our country's average annual electricity production is usually satisfied by this source. Before the war in Ukraine, Russia was its main supplier. Currently, a diversification of supply is taking place thanks to countries such as Libya, Algeria and Azerbaijan which suffer from lack of stability. That is why the regulator and the governor are pushing for an increase in renewable energy through incentives and other instruments.

Other less obvious dependencies concern the correlation with hydroelectric production being renewable in general, especially wind, which is mostly present in Southern Italy, but low-cost technologies influence the price by lowering it. Gas dependency is a key factor since the Italian power price depends on the quotation of TTF gas. The PSV which is the gas quoted in Italy is not very liquid so the Dutch one is taken as a reference. Seasonality is reflected in different productions from renewable sources. This has a calming effect on the PUN, which is usually lower between winter-spring, when first wind, then solar and hydro productions are higher. In recent years, the increased volatility of renewable production and especially drought are making the calming effect of renewables less evident and predictable.

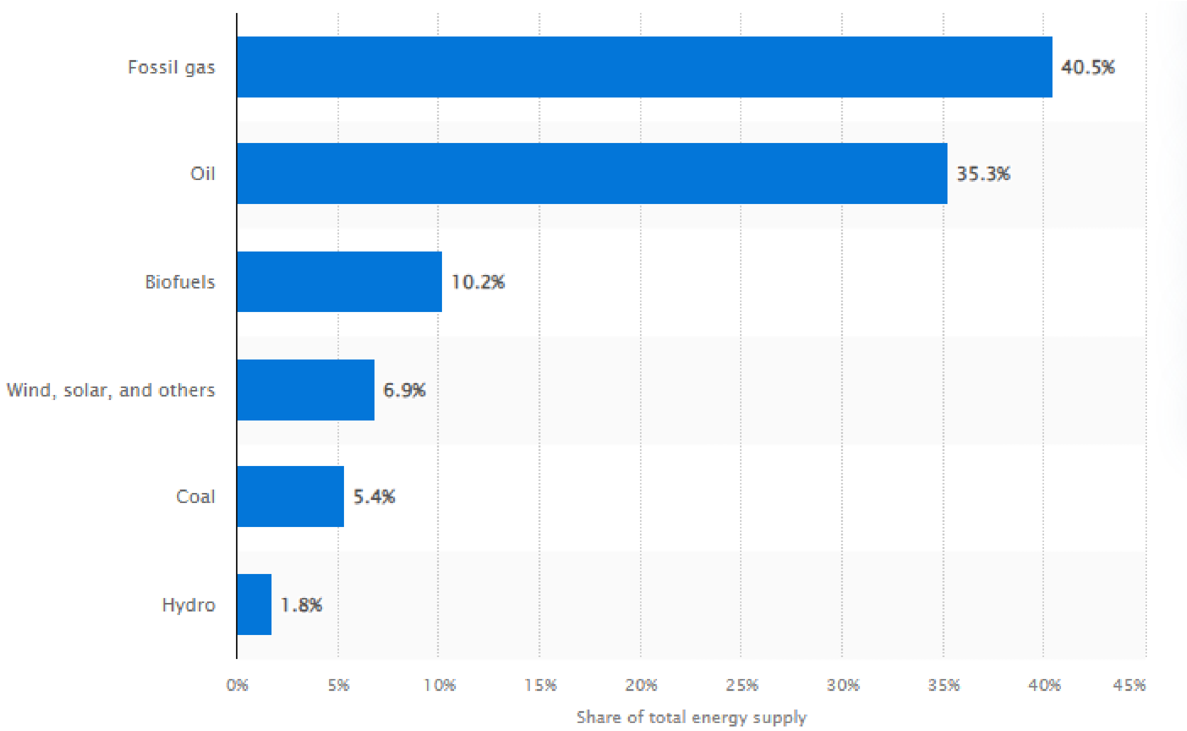

Figure 1.3: The distribution of energy supply in 2022.There is the natural gas that is the most important with 40.5%, after oil 35.3%,Biofuels 10.2%,wind solar and others 6.9%,coal%5,4 and Hydro 1,8%. taken by statista.com.

1.8 The problem of liquidity

It was seen how the TTF Calendar 2024 quotation suddenly rose and the spread reacted promptly, but unexpectedly. As Italy is more sensitive than Germany to the gas price, analysts would have expected to see Italy higher than Germany with a positive spread instead the opposite happened. The more liquid Germany absorbed the increase in gas prices more quickly by raising its price while the less liquid Italy followed the trend late, causing the spread to open in a negative direction. Liquidity is a very important issue for Italy as the Italian power market has very little volatility. The Italian PSV gas market also has low liquidity so the much more liquid Dutch TTF is considered. Low liquidity is not a small problem for a country because when faced with bearish or bullish fundamentals the market does not immediately absorb changes because ,with low liquidity, prices remain stable or take longer to change trend and follow market logic. Germany, Holland, Great Britain and France have much more liquid markets than Italy.

Chapter 2

RELATIONSHIP AND INFLUENCE OF GAS AND ITALIAN POWER MARKET

2.1 GAS demand

Gas is a key element in the European energy market. It was highly volatile last year. To be able to define a possible gas price, it is necessary to study supply and demand. Beyond that, gas is very useful in countries such as Italy, Great Britain, and Sweden. Italy in particular is very dependent on gas, so much that during the crisis, caused by the war in Ukraine, the country faced a supply crisis. Since Italy's main supplier was Russia with the embargoes put in place after the outbreak of the war, it found itself in an 'unhappy' position. Gas is a key element in Italy because many household appliances use gas as an energy source. The same thing happens in England, another country highly dependent on gas. Gas demand.[1] is directly proportional to consumption and inversely proportional to renewable production and to other types of production. This means that gas demand decreases with stable consumption, if renewable energy production increases because renewable resources are cheaper the consumer will then prefer to consume renewable production. The demand for gas therefore depends on a country's total consumption and the production of other sources of energy supply, for example, renewables, nuclear, coal and so on. Demand for gas increases exponentially in summer and winter. In winter, the increase in demand is even more pronounced as many air conditioning and heating systems inside the house need gas to function. In particularly harsh winters, such as we had in past years, gas production is essential for the livelihood of households, consumers and businesses. Fortunately, winter 2023 was a very warm winter so there was no super demand for gas. As previously specified the amount of demand is not stable over time but varies depending on many factors :temperature, seasonality ,other productions and even on the price of the same gas. An example is given by last winter when the demand for gas increased for the very low temperatures and there was a sharp increase in general demand. This increase was not proportional to other years because the price of gas both in the market and in consumer's perception was high. The consumers' perception of gas price isn't very often real but,influenced by the media ,Internet and so on. So, sometimes, the price of gas may not be very high but, if consumers' perception defines a high price, the demand for gas will not be increasing because consumers will not want to pay for gas at very high prices. Last winter, the consumer's perception and the fairly above-average price meant that the growth in demand for gas was not proportional to the growth in demand and the drop in temperatures. Who are the consumers? In Europe we have already mentioned Italy ,Great Britain, etc. In addition to the European states, important players in world demand are Asian countries such as Japan, South Korea and China. China represents a key driver for the so-called 'Asian demand'. It therefore represents the

[1] The basis for this switch is a trivial question, how much does it cost to produce from coal rather than gas? To begin with, it is important to define the inverse correlation with the price of gas; in fact, if the price of gas goes above 130€/MW, energy production costs become too high and therefore there begins to be a preference for operators to produce from coal. Of course in this case the price of coal will rise because there will be a much higher demand than normal. Naturally, the European Union is trying not to let this happen and has introduced new rules for certain production sectors, which will have to respect a maximum emission threshold. In recent years, gas has seen its price rise, first with the war in Ukraine where a real rally began, then with the storage crisis. The short-term marginal cost of power generation of combined cycle plants using gas has become very high, far above the competitive costs of coal or coal-fired combined cycle generation. Since autumn 2022, however, the real change has begun: the price of gas has fallen dramatically from the €300/MWh ,it has peaked to €30/MWh in summer 2023. In order to study this measure of change and the switch between gas and coal, one has to study clean dark spreads and clean spark spreads. This study shows that taking into account the baseload energy contracts, gas reached very efficient competitive costs at the beginning of 2023 while coal was able to defend itself and did not arrive in the full area of the switch. So the costs of the two productions, even though gas is becoming more and more efficient, are quite even. Anyway, the switch between the two sources of energy production has not taken place yet even though a timeline has now emerged that will ensure that energy production will take place mainly through gas and no longer through coal, which has a high pollution coefficient.

demand that comes from outside the Asian continent which has been very volatile in recent times so that the United States preferred to supply Europe which was willing to pay much higher prices than the Asian continent. Right now Chinese demand is not very strong partly because the main industrial country of Asia is not willing and does not want to pay high sums for gas,for that their demand is still much lower than they could be. Market participants know very well that if China were to resume its demand for gas, the price of LNG could spike and increase the price of gas in Europe. Obviously, if there's low gas demand the price of gas decreases but in the face of high demand above normal the price of gas could rise proportionally much more, because, in addition to being driven by fundamentals ,there would also be speculative components that would cause the price to rise proportionally more than it should do in a balanced market. As defined earlier, demand is a very important driver of gas prices. When, last month, China was thought to want to increase its demand, the price of TTF soared in the face of likely higher demand. African demand and South American demand along with demand from countries within the Oceania have no impact on the global gas price because they are much lower and willing to pay similarly lower prices than the European and Asian countries that are the key drivers of the gas price.

2.2 The storage and GAS supply

Gas storage is a key component that determines a country's demand. Indeed, after the crisis in Ukraine, Europe found itself with storage at its lowest level and thus unable to cope with the lack of supply from Russia then blocked by the European embargoes.. In these years the situation has completely changed since European countries, during the fall and winter, filled their storage reaching very high levels above 70/80%; in this way, if there were to be less production and less gas delivery, the countries in question would not suffer much but could still use the reserves until the situation stabilizes. The high levels of gas storage meant that there was a halt to the high speculation carried out by the investment fund market and by various parties who tried to earn as much as possible thanks to price spikes.

The supply of gas depends on many factors, the first one is the production and import by supplier countries.[2]. The gas producing countries are Russia ,Azerbaijan ,Algeria, Libya, Iran and other Middle Eastern and African countries that make up the world gas supply and support the big consumers such as Europe, Asia and North America. The United States is also a very important producer of gas, they use a technique, fracking, that allows them to produce gas and sell it in the market, the American gas is a liquefied one :LNG gas. Now an attempt will be made to analyze the different drivers of both global and European and Italian gas supply. Concepts such as LNG, LNG, SS, CSS, regasification plants, producing countries, gas storage and influence with other factors will be introduced.

2.3 LNG and GNL

LNG gas is much more expensive than conventional gas. It has managed to find space in Europe thanks to the war in Ukraine. Before that, American gas was only in Asia. LNG ships transport liquefied gas all

[2]This is the most important requirement for gas supply. Just think of Russia two years ago , because of the war it suffered an embargo, which in turn led to a huge drop in gas supply throughout Europe.

over the world. It is a very interesting business because it is made up of private individuals. By private individuals, we mean non-governmental companies transporting LNG all over the world by business. The market has grown exponentially over the last year and has become a key factor in gas supply. Of course, the market is based on the demand for gas, so when regasification plants in Europe stop working, the demand grows disproportionately. An example may be the period in early July when maintenance in Norway, which is a major producer and exporter of gas throughout Europe, came to a halt. During this period, the lng market met the production shortfall. Demand is directly proportional to gas demand and is inversely proportional to gas producers, renewables and other energy sources and supply. When regasification plants and gas producing countries export gas through pipelines, demand for LNG decreases proportionally. In the last months LNG.[3]. demand has increased in Europe and Asia, making the cost quite high.

2.4 CSS AND SS

When gas arrives in importing countries, it will be used in gas-fired combined cycle plants for energy production. This driver is called SS and CSS. These two factors SS and CSS tell us nothing more than how much producing energy costs to a combined cycle gas plant. Since gas producing countries such as Russia, Azerbaijan, Libya, Algeria and Burkina Faso produce gas in advance, they then export it through pipelines, ships in the case of LNG, to importing countries. The importing countries will produce energy through gas. The demand of the importing countries will therefore depend on the price of gas since very high prices.[4] will lead to less demand for gas . In this case, importing countries will demand other sources such as renewables, coal, etc. to produce energy The cost of gas depends on the country you are in. In Italy the gas price is represented by the PSV. In Europe, on the other hand, the TTF is used, which is much more liquid than the Italian index. The selling price for Italy is the PUN. The PUN is nothing more than the National Unit Price. Since there are different areas in Italy with different prices, the PUN is the unit price for the whole country.[5].

2.5 Italian Gas Report and the situation with Russia

The Italian production share is characterized by a high dependence on the gas source. About half of the electricity production in our country averages annually is satisfied by this source. This causes a very high dependence of the power price.[6]. on the gas price. Apart from this difference, a correlation exists for both the TTF hub and the PSV. The correlation is present both in the spot market but also in the futures market, offering possibilities for trades. The supply relationship between Italy and Russia. Until the war in Ukraine, Russia was Italy's main energy partner for gas imports; it satisfied 40% of Italian gas demand. This energy source came to the peninsula via the TAG pipeline which was then connected to the Italian

[3] Another Example is the Houti situation in the Red Sea that with their attack are creating a bullish trend in the LNG market

[4] Production Cost=(Gas Cost/Plant Yield)+(C02 factor)+ Fixed Costs
SS=Sales Price-(Gas Cost/Plant Yield)
CSS=SS-Factor C02.

[5] In Italy there are many different zone for elettricity price :Nord,Centro Nord,Centro Sud,Sud,Calabria,Sicilia,Sardegna

[6] On the elettricity market if the price of the Gas increase by one unit the power price is going to increase of two units

network in Tarvisio. This dependence has always been investigated and contested by the Italian secret services. The parliamentary committee for the security of the Republic (Copasir) has pointed out in its report on the consequences of the conflict between Russia and Ukraine in the field of energy security last April stating:

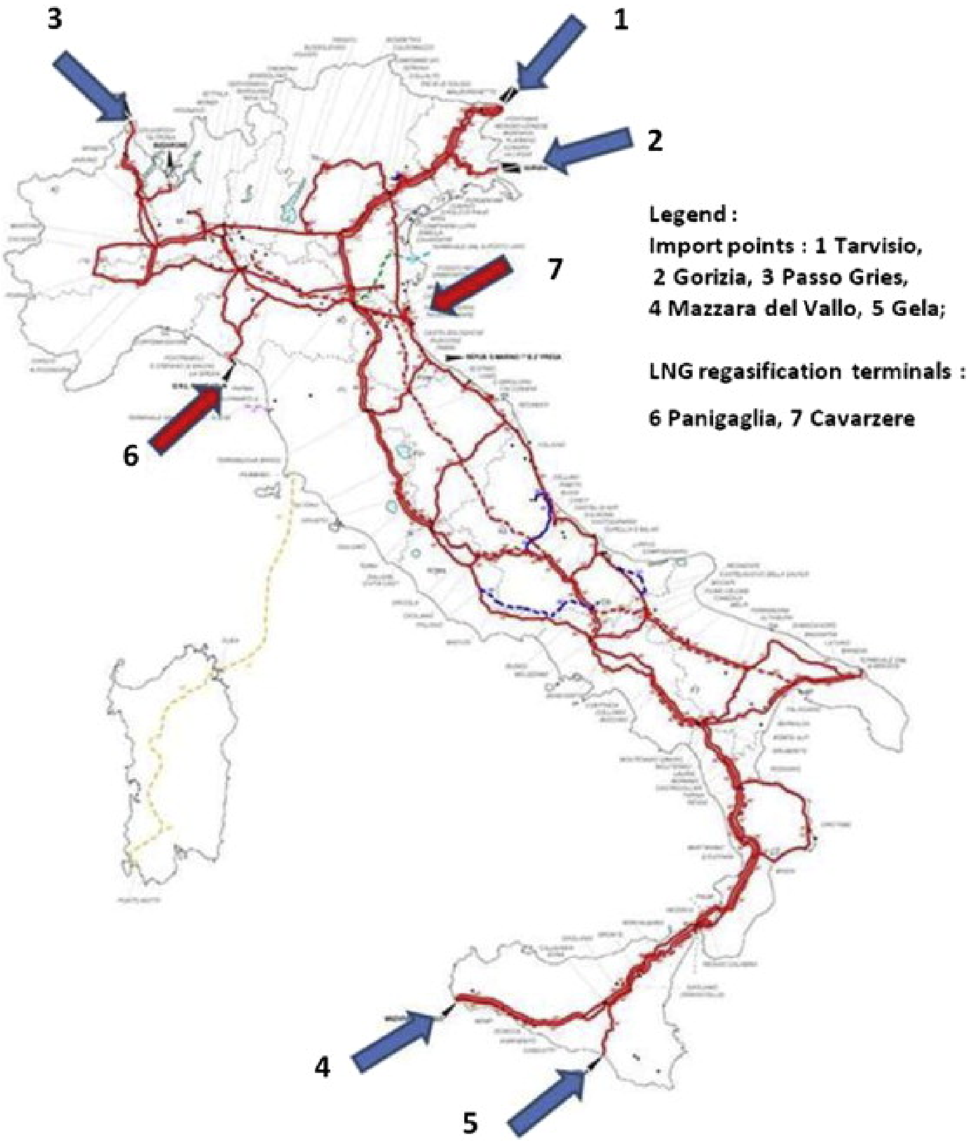

Figure 2.1: Point of entrance for GAS supply in Italy.In the Figure there is the point of entrance for the gas supply .Passo Gries is for the gas that comes from the north of Europe, Tarvisio is the GAS that comes from Russia ,Meledugno is the gas that comes from Azerbajian,Gela is the GAS from Libia and Mazzara del Vallo is from Algeria.This point of entrance is the papeline that connects Italy with the producers .LNG regasification terminals are the point that connect LNG Navy with the pipelines in Italy.These points are Panigaglia and Cavarzere.Graph taken by https://www.spglobal.com/commodityinsights/en/market-insights/blogs/natural-gas/072920-italys-gas-market-an-ecosystem-where-price-takers-thrive.

"Relations between Italy and Russia in the energy sector have deep roots..... Gas was the subject of the first agreements at the end of the 1960s and the volume of supply has grown progressively since then". The reliance on Russia as the main supplier of gas coincided with Putin's rise to power. Today's dramatic events denote how that orientation had objectively underestimated the problem of energy dependence and diversification of supplies, which had therefore become an instrument of pressure used by the Russian power towards European countries, especially Italy". When the price of gas continues to rise, especially in summer, it is also due in part to Russia, which has made sure to use this instrument of political pressure on the states. Before, the gas has increased,because Russia reduced the flow of gas while the supply and demand in the European Union remained constant, causing the price to rise. As in the most classic market law, if demand is the same but supply falls, that good becomes more valuable and therefore the price rises

2.6 ITALY NEW GAS SUPPLIERS

In summer 2022 Italy was in the midest of searching for new strategic gas supply partners and turned its attention to Algeria. As a matter of fact, the North African state, for years Spain's main gas supplier, had suspended its 20-year friendship treaty with Spain in June 2022; this was due to the change of position of President Pedro Sanchez, who had decided to support the Moroccan autonomy plan for the Western Sahara, an area south of Morocco where there is a dispute between Morocco and Algeria over who owns the rights to the area. Spain, Algeria's historical ally in this matter, has therefore lost the favors of its main energy partner. Italy is trying to use this situation to its advantage by strengthening its alliance with Algiers. In June, the Italian Prime Minister, Mario Draghi, visited the North African country and the delegations signed 15 memorandums of understanding including one on gas. Another alternative to Russia is liquefied natural gas. This particular type of gas can be transformed into a liquid state in order to be transported by ship to countries that then request it. Once in the country of destination, this gas then returns to its gaseous state and the process is carried out by the so-called regasification plants. In Italy there are three regasification plants and there is a project to increase their quantity; moreover, there are some methane tankers at disposal as well. The project is to expand the fleet as there are only two ships, one in Ravenna and one in Piombino at the moment. As Copasir reports, Italy is moving to look for new partners to import gas. In this sense, Italy is moving towards Libya and Azerbaijan, through the existing gas pipeline networks, or is sounding out Egypt, Qatar, Congo, Mozambique, Angola and Nigeria for liquid gas. In particular, in this last case, production is also managed by Eni (a historical Italian company in the petrochemical and energy sector in general). Copasir, however, is continuing with the close alliance with Algeria in order to make it our first gas supplier through the TransMed pipeline.

2.7 Analysis on the Italian situation on supplies and possible developments

As Think Thank ECCO reports, it is not really necessary for Italy to equip itself with new infrastructure such as gas pipelines and regasification plants since the response to the energy issue would be unsustainable and uneconomic in the short and long term. For if one thinks the new gas pipelines or LNG carriers will come into service in a few years when the situation will be stabilized and will burden consumers for decades to come. As the International Energy Agency shows in its guidance on the global path to climate neutrality, new investments at the beginning of the fossil energy value chain are inconsistent with climate policies, so investment on regasification plants,LNG ships etc. should not be considered but more should be done towards renewable energies . As Copasir points out, Italy must, therefore, urgently face two situations that have grown over the years and revealed all the fragility and volatility in the energy supply market: the problem of energy dependence and the diversification of supplies: this is the new open game for Italy and all European countries. For the peninsula it can also be a huge gamble since there are not very large renewable facilities within the country yet.

Chapter 3

METODOLOGIES FOR THE TRADING STRATEGIES

3.1 introduction

In the forthcoming methodologies chapter regarding trading strategies, an organized and coherent approach will be adopted to ensure a logical progression in the analysis. The chapter will commence by elucidating the conceptual underpinnings of derivatives, followed by an exhaustive examination of energy derivatives, thereby setting the stage for subsequent discussions. Subsequently, the phenomena of backwardation and contango will be explicated to provide a comprehensive understanding of market dynamics. Furthermore, the chapter will introduce the fundamental statistical components pertinent to trading strategies. Initially, attention will be directed towards the ARIMA model, scrutinizing it both theoretically and practically. Subsequent to model identification, a systematic approach will be undertaken to introduce statistical and econometric tests aimed at assessing forecast quality. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) will be elucidated, providing essential tools for time series analysis. Lastly, the chapter will culminate with the introduction of advanced statistical methodologies, including the Mercier-Zarnowitz regression and the Diebold-Mariano test. These methodologies will serve to evaluate the efficacy of various forecasts generated by the ARIMA model, facilitating the identification of optimal forecasts for utilization within trading strategies. This systematic and structured approach ensures coherence and clarity in the presentation of methodologies, laying a robust foundation for subsequent analysis and application in trading scenarios.

3.2 Derivates

Derivatives represent financial instruments whose value is derived from an underlying market. Predominantly found in equity trading, currency values, and interest rates for loans, these instruments are ideal for both speculation and risk hedging due to their market-based foundation. Their popularity stems from the ability to engage in market speculation even with limited investment capital. In today's market, various energy derivatives serve distinct purposes, ranging from risk management in the energy industry to providing speculative opportunities for investors. The primary energy derivatives in use are futures and options contracts. Futures, traded on exchanges, are forward contracts, while options provide the buyer with the right to buy or sell a specific asset at a predetermined price. Notably, derivatives often carry the classification of securities. This categorization arises because the underlying asset, such as gold or oil, can be held in a trust's bank account. In turn, the trust issues a certificate to the investor, confirming their ownership of a corresponding amount of the underlying asset.

3.3 Future energy derivates

Futures are a type of derivative that is common in the financial industry. They are traded on an exchange and serve as a contract between two parties. One investor must give the other investor a specific amount of the underlying asset (the underlying asset is energy, for example). A futures contract can either be traded on the exchange for cash or for a different futures contract. For example, in the case of energy, the investor can either cash in on the contract or take a contract that has a value that is equivalent to the value of the first contract. Futures contracts are very versatile and can be used for different purposes. They can be used to hedge against risk or for speculation on the market. The two parties involved in the contract have different ways of profiting from it. In a bull market, for example, the buyer can profit from the futures contract by taking the cash value of the contract at the end of the contract. Conversely, a seller can profit when the market is bearish, or the underlying asset's price goes down. This is because the seller will make a profit from the difference in the price of the contract when they sell it.

3.4 Contango and backwardiation

Contango in the energy market occurs when the futures prices of commodities like oil or power are higher than the current spot prices. This situation indicates an anticipation of future price increases or a surplus in the current market. Investors or traders may encounter contango when the cost of storing the energy commodity for future delivery, including associated storage and financing costs, contributes to the higher futures prices. Contango can influence trading strategies and investment decisions in the energy sector. Backwardation[1] in the energy market is characterized by futures prices being lower than the current spot prices of commodities like EUA or natural gas. This scenario suggests an expectation of declining prices or a current shortage in the market. Backwardation may arise when the cost of storing the energy commodity is perceived as lower than the potential future savings in futures prices. Traders and investors often analyze backwardation to make informed decisions, as it may reflect immediate market conditions and impact hedging strategies in the energy industry.

[1] In february 2024 the price of the calendar 2026 is less then the spot prices, because the market is bearish on the future

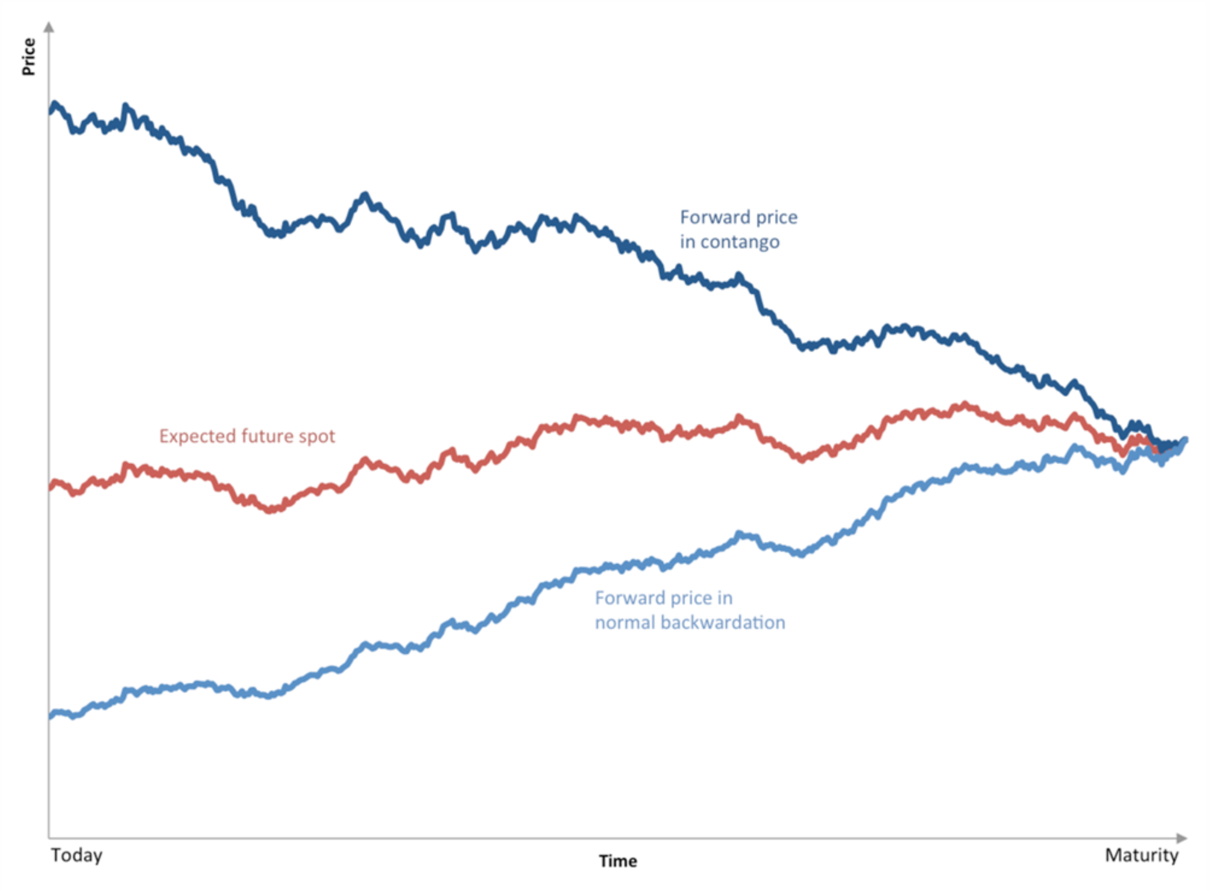

Figure 3.1: The backwardations is represented by the light blue curve.The contango is the Blue curve that is when the future price are higher compared to the spot price,the spot is the red line in the center.The graph shows how contango line will converge on the spot line along with the backwardiation line at time to maturity when the product goes into delivery.

3.5 ARIMA

3.5.1 Introduction ARIMA and applications

In this thesis, the ARIMA model is used for the prediction of time series (future energy prices), because as many papers have shown, the model can make reliable forecasts and has already been used successfully in the past. The ARIMA model was introduced by Box and Jenkins (1970) as a generalization of the ARMA model. To enhance comprehension of the dataset or to project future series points, both ARIMA and seasonal differencing models are employed in the analysis of time series data. ARIMA models are particularly useful when the data exhibit non-stationarity primarily in terms of mean, necessitating initial differencing steps to render the mean function stationary, thereby eliminating trends. Seasonal differencing is employed when there is evidence of seasonality within the time series, aimed at removing the seasonal component. This model has been used in many academic papers in the fields of statistics and econometrics as done by Box et al. (2015) for time series forecasting, or by Brockwell and Davis (2016), Hamilton (1994) for time series analysis or as done by Newbold (1983) for time series analysis and then was compared with the ARMA model and then as a forecasting method in the papers of Hyndman and Athanasopoulos (2018),Makridakis et al. (1998). In the energy trading, the model has been used by Gupta and Kumar (2020) for mid load forecasting,or by Shikhina et al. (2020) , another example is done by Contreras et al. (2003) for the prediction of DA prices. In the paper written by Gao et al. (2017) analyzes how to predict electricity prices in the UK market and places importance on how in today's increasingly stock-like electricity market it has become critical to develop forecasts to have an edge over competitors . An important application of the ARIMA model is done by Karabiber and Xydis (2019) for price forecasting in the energy market in Denmark ,analysis done on the hourly spot price. However, this paper is inapplicable in other European countries as it exploits the unique characteristics of the Scandinavian country.[2] which has extremely different fundamentals than CWE countries and Italy. As described by de Oliveira and Oliveira (2018) the ARIMA model is used for consumption forecasting.This application demonstrates how through the model consumption can be forecasted which is a fondamental element on any trading strategy for both short term and long term. Another possible application, on the other hand, is in hedging strategies for an electric portfolio of a utility operating in the energy sector. The ARIMA model is a very important tool as demonstrated in the antecedent papers. There are numerous applications in the energy market to be able to predict prices and be able to develop trading strategies, or to predict commodity market fundamentals in this case of POWER.

3.5.2 ARIMA Model

In the thesis, when conducting the trading strategy, we will first follow this step to understand which parameters to use in the ARIMA model . We are going to determine whether the series is stationary or not by considering the graph of ACF. If a graph of ACF of the time series values either cuts offfairly

[2] Denmark has an energy demand that is almost completely met by renewable, this makes prices move very differently from other countries that are dependent on continuous cycle production(GAS,COAL,OIL),example Italy has a completely different structure as it is very dependent on the gas price and production as written in chapter two is also based on CSS values at that time

quickly or dies down fairly quickly, then the time series values should be considered stationary. If a graph of ACF dies down extremely slowly,then the time series values should be considered non-stationary. If the series is not stationary, it can often be converted to a stationary series by differencing.Thanks to these passage we can understand if we have to use the parameter D. The ARIMA model developed by Box and Jenkins (1970) combines three main components: AR (AutoRegressive), I (Integrated), and MA (Moving Average). The AR component measures the relationship between an observation and its previous observations within a time interval. The order of AR, denoted as p, specifies how many previous periods are used to predict the current observation. The AR formula is generally expressed as:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t \tag{3.5.1}$$

where:

- y_t is the observation at time t,
- $\phi_1, \phi_2, \ldots, \phi_p$ are the autoregressive coefficients,
- c is a constant,
- ε_t is the white noise error term.

The MA component measures the relationship between an observation and a residual error from a moving average of order q. The MA component formula is generally expressed as:

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_q \varepsilon_{t-q} \tag{3.5.2}$$

where:

- $\theta_1, \theta_2, \ldots, \theta_q$ are the moving average coefficients,
- ε_t is the white noise error term.

ARMA

$$\Phi(B)X_t = \Theta(B)\varepsilon_t \quad \text{with} \quad \varepsilon_t \sim WN(0, \sigma^2). \tag{3.5.3}$$

In the general form, the ARMA(p, q) model writes:

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)X_t = (1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q)\varepsilon_t, \tag{3.5.4}$$

or equivalently:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t. \tag{3.5.5}$$

ARIMA as mentioned before is a generalization of ARMA.The extra part will be introduced in the next paragraph.It is important to note that an ARIMA without part I behaves like an ARMA.As in the case of this thesis after analyzing the ACF it was decided to compile a forecast on the ARIMA model without using part I.

I (Integrated)

The I component refers to differencing the time series to make it stationary. Stationarity is crucial because many statistical models, including ARIMA, assume that the statistical properties of the time series are constant over time. The order of differencing is denoted as d.

The differencing can be represented as:

$$\Delta y_t = y_t - y_{t-1} \tag{3.5.6}$$

where Δ represents the differencing operator. As said as before the overall order of an ARIMA model is represented as ARIMA(p, d, q), where p is the AR order, d is the differencing order, and q is the MA order. The orders can vary depending on the characteristics of the time series under analysis. The goal is to find the optimal order that provides the best model for describing the data.

3.5.3 Forecasting with ARMA

Having used an ARIMA without Part I, the forecast of an ARMA will be analyzed here, which represents the forecast then used by us in the next chapter before implementing the trading strategy. Consider the ARMA(1,1) model given by

$$X_t = \delta + \phi X_{t-1} + \theta \epsilon_{t-1} + \epsilon_t.$$

The one-step ahead forecast is given by

$$\begin{aligned}
\hat{x}_{t+1} &= E_t(X_{t+1}) \\
&= E_t(\delta + \phi X_t + \theta \epsilon_t + \epsilon_{t+1}) \\
&= \delta + \phi x_t + \theta \epsilon_t.
\end{aligned}$$

The forecast error is

$$e_{t+1} = X_{t+1} - \hat{x}_{t+1} = \epsilon_{t+1}.$$

Its variance is

$$\text{Var}(e_{t+1}) = \sigma^2.$$

The forecast for $k = 2$ is

$$\begin{aligned}
\hat{x}_{t+2} &= E_t(X_{t+2}) \\
&= E_t(\delta + \phi X_{t+1} + \theta \epsilon_{t+1} + \epsilon_{t+2}) \\
&= \delta + \phi \hat{x}_{t+1} \\
&= \delta(1 + \phi) + \frac{\phi^2}{2} x_t + \theta \phi \epsilon_t.
\end{aligned}$$

The forecast error is

$$
\begin{aligned}
e_{t+2} &= X_{t+2} - \hat{x}_{t+2} \\
&= \delta + \phi X_{t+1} + \theta\epsilon_{t+1} + \epsilon_{t+2} - (\delta + \phi\hat{x}_{t+1}) \\
&= \phi\epsilon_{t+1} + \theta\epsilon_{t+1} + \epsilon_{t+2} \\
&= (\phi + \delta)\epsilon_{t+1} + \epsilon_{t+2}.
\end{aligned}
$$

Notice that it has zero expected value and variance equal to

$$\text{Var}(e_{t+2}) = \sigma^2((\phi + \delta)^2 + 1).$$

Iterating, the k steps ahead forecast is given by

$$\hat{x}_{t+k} = \delta(1 + \phi + \phi^2 + \cdots + \phi^{k-1}) + \phi^k x_t + \theta\phi^{k-1}\epsilon_t.$$

As k goes to ∞, the variance tends to the unconditional variance of the process.

Notice that after the second step ahead, the predictor resembles that of an AR(1). Indeed, its asymptotic behavior is exactly that of an AR(1), i.e., when $k > 1$, the behavior of the forecast is dominated by the autoregressive part.

For a general ARMA(p, q) model, similar results to those seen for the ARMA(1, 1) can be obtained. When $k > q$, the autoregressive part drives the forecast that converges to the unconditional mean of the ARMA(p, q) as k tends to ∞. Similarly, the variance of the forecast error converges to the unconditional variance.

3.6 ACF

Introduced by Yule (1927) ,the ACF is a statistical measure that evaluates the correlation between an observation in a time series and its previous observations at different time lags. In this thesis is fondamental in the energy trading strategies as said before to understand better how to use ARIMA. The formula of ACF is given by:

$$\rho_k = \frac{\text{Cov}(X_t, X_{t-k})}{\sqrt{\text{Var}(X_t) \cdot \text{Var}(X_{t-k})}} \tag{3.6.1}$$

where:

- ρ_k autocorrelation of the lag k.
- $X_t$time t.
- Cov(X_t, X_{t-k})covariance between observation X_t e X_{t-k}.
- Var(X_t) e Var(X_{t-k}) are the variance X_t e X_{t-k}.

The function $\rho_k$indicates how much the observations to a delay k are correlated with the current observation. A value near 1 indicates strong positive correlation, a value near -1 indicates strong negative correlation, and a value near 0 indicates poor correlation. The coefficient of correlation between two values in a time series is called the autocorrelation function (ACF). As done by Shikhina et al. (2020) after we are going to use in the trading strategies.

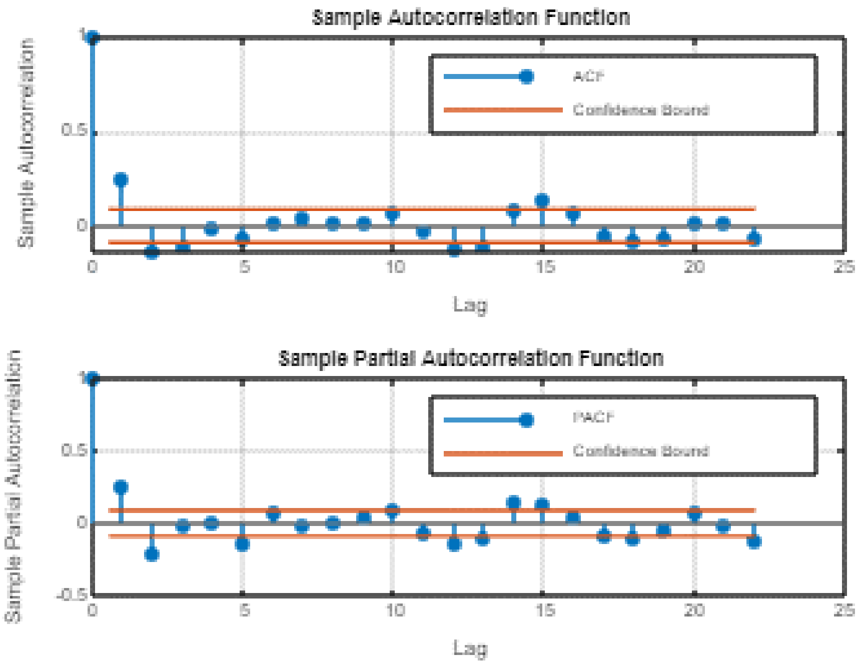

Figure 3.2: Thi is an example of AXF and PACF used in this thesis.The analisys of autocorrelation shows an autocorrelations at lag 1 that is significative.

3.7 PACF

The PACF measures the autocorrelation between X_t and X_{t+h} after removing their linear dependence with the other intermediate variables (recall the partial correlation coefficient in multiple regression). It takes to compute

$$\phi_{hh} = \text{Corr}(X_t, X_{t+h}|X_{t+1}, X_{t+2}, \ldots, X_{t+h-1}). \tag{3.7.1}$$

PACF can be derived as follows: consider a regression model where the dependent variable X_{t+h} is regressed against $X_{t+h-1}, X_{t+h-2}, \ldots, X_t$, i.e.,

$$X_{t+h} = \phi_{h1}X_{t+h-1} + \phi_{h2}X_{t+h-2} + \ldots + \phi_{hh}X_t + e_{t+h}, \tag{3.7.2}$$

where ϕ_{hj} represents the parameter of the regression of X_{t+h} with respect to the variable X_{t+h-j}, and e_{t+h} is the shock uncorrelated with X_{t+h-j} for $j \geq 1$.

We are considering a zero-mean process. Multiplying both sides by X_{t+h-j} and taking the expected values, we get:

$$\gamma(j) = \phi_{h1}\gamma(j-1) + \phi_{h2}\gamma(j-2) + \ldots + \phi_{hh}\gamma(j-h), \tag{3.7.3}$$

thus,

$$\rho(j) = \phi_{h1}\rho(j-1) + \phi_{h2}\rho(j-2) + \ldots + \phi_{hh}\rho(j-h). \tag{3.7.4}$$

For $j = 1, 2, \ldots, h$, we obtain the following system of equations, known as Yule-Walker equations:

$$\rho(1) = \phi_{h1}\rho(0) + \phi_{h2}\rho(1) + \ldots + \phi_{hh}\rho(h-1) \tag{3.7.5}$$

$$\rho(2) = \phi_{h1}\rho(1) + \phi_{h2}\rho(0) + \ldots + \phi_{hh}\rho(h-2) \tag{3.7.6}$$

$$\vdots$$

$$\rho(h) = \phi_{h1}\rho(h-1) + \phi_{h2}\rho(h-2) + \ldots + \phi_{hh}\rho(0). \tag{3.7.7}$$

It can be shown that, after some computations for $h = 1, 2, \ldots$, we get:

$$\phi_{11} = \rho(1) \tag{3.7.8}$$

$$\phi_{22} = \frac{1}{\rho(1)} \begin{pmatrix} \rho(1) & \rho(2) \\ \rho(2) & \rho(1) \end{pmatrix} \tag{3.7.9}$$

$$\phi_{33} = \frac{1}{\rho(1)^2 - \rho(2)^2} \begin{pmatrix} \rho(1) & \rho(2) & \rho(3) \\ \rho(2) & \rho(1) & \rho(2) \\ \rho(3) & \rho(2) & \rho(1) \end{pmatrix}. \tag{3.7.10}$$

3.8 Mincer and Zarnowitz Regression

After estimating the different forecasts using the ARIMA model, as demonstrated by Guler et al. (2017), we will apply the Mincer and Zarnowitz regression within a model to assess forecast accuracy. The Mincer and Zarnowitz regression, examines the relationship between a dependent variable Y and an independent variable, typically representing time or a specific predictor variable. This method is particularly useful in economic research and forecasting.

3.8.1 Regression Model: Mincer and Zarnowitz

The Mincer and Zarnowitz regression can be represented as:

$$Y_{t+h} = \beta_0 + \beta_1 \hat{Y}_{t+h|t} + \varepsilon_{t+h|h} \tag{3.8.1}$$

where:

- Y_{t+h} denotes the dependent variable at time $t + h$.
- $\hat{Y}_{t+h|t}$ represents the independent variable at time $t+h$, often representing a predictor or explanatory variable.

- β_0 is the intercept term.
- β_1 is the coefficient associated with the independent variable.
- $\varepsilon_{t+h|h}$ represents the error term, capturing unobserved factors affecting Y_{t+h}.

In our case h is egual to 1 because we are using daily returns.

3.8.2 Hypothesis Test

The unbiasedness and efficiency of the forecast can be evaluated by testing the intercept and slope through the joint hypothesis:

$$H_0 : \beta_0 = 0, \beta_1 = 1 \tag{3.8.2}$$

Optimal forecast is characterized by the upholding of H_0.

3.8.3 Results and Implications

The results of the Mincer and Zarnowitz regression can be presented in a table, detailing the estimated coefficients and their statistical significance. Researchers can then interpret these results to draw conclusions about the relationship between the independent variable and the dependent variable.We are going to have conclusion about the different forecast.

3.8.4 Conclusion

The Mincer and Zarnowitz regression, when applied provides a valuable tool for understanding the impact of a factor on a dependent variable over time. Its application can offer insights into economic and social phenomena by examining the relationship between a predictor variable and the outcome of interest.

3.9 The diebold mariano Test

As introduce in Chen et al. (2014)Mariano (2002) the Diebold Mariano Test is a statistical test widely used in the field of forecasting to assess whether one predictive model outperforms another in terms of forecast accuracy. Named after Francis X. Diebold and Robert S. Mariano, this test is particularly useful when comparing the forecast performance of two competing models.In this thesis is going to be used as last step before the trading strategies.

3.9.1 Hypotheses

Consider two forecasting models, Model A and Model B. The null hypothesis and alternative hypothesis for the Diebold-Mariano test are defined as follows:

- **Null Hypothesis (H_0):** There is no significant difference in forecast accuracy between Model A and Model B.

- **Alternative Hypothesis (H_1):** Model A significantly outperforms Model B in terms of forecast accuracy.

3.9.2 Test Statistic

The Diebold-Mariano test statistic is computed as:

$$DM = \frac{\bar{e}_A - \bar{e}_B}{\sqrt{\frac{s_d^2}{T}}} \tag{3.9.1}$$

where:

- $\bar{e}_A$ and $\bar{e}_B$ are the mean forecast errors for Model A and Model B, respectively.
- s_d^2 is the variance of the differences in forecast errors.
- T is the number of forecast observations.

3.9.3 Decision Rule

Under the null hypothesis, the test statistic DM follows a standard normal distribution. Researchers typically compare the computed test statistic to critical values from the standard normal distribution to make a decision.

- If $|DM| > Z_{\alpha/2}$, where $Z_{\alpha/2}$ is the critical value at significance level $\alpha/2$, the null hypothesis is rejected in favor of the alternative hypothesis.
- If $|DM| \leq Z_{\alpha/2}$, there is insufficient evidence to reject the null hypothesis.

3.9.4 Conclusion

The Diebold-Mariano test provides a robust method for comparing the forecast accuracy of two models. It is widely applied in empirical research to determine whether a new forecasting model offers a statistically significant improvement over an existing one.

Chapter 4

TRADING STRATEGIES

4.1 Structure of Trading strategies

In the fourth chapter, our analytical focus centers on the POWER futures, specifically targeting the 2021 quarters (quoted in 2020) and 2022 quarters (quoted in 2021). The objective is to formulate forecasts for the ensuing 2022 quarters (quoted in 2021), and subsequently, to devise trading strategies based on these forecasts. At that time prices were correlated to the difficult geopolitical context[1]. To initiate this process, we employ cross-validation techniques to rigorously develop forecasts. Subsequently, we subject these forecasts to comprehensive evaluations utilizing statistical tests, including the Mincer-Zarnowitz regression and the Diebold-Mariano test. These assessments are paramount for appraising the forecasting accuracy, facilitating an informed selection of the most robust models. Our approach maintains consistency across all quarters under scrutiny. The foundational steps include a meticulous plot representation, log returns calculation and an exploration of autocorrelation and partial autocorrelation to discern series properties. The subsequent modeling phase involves the application of diverse configurations, ranging from 0,0,0 to 2,0,2.There will be a model selections. The model performances are pitted against each other using mean square root comparisons. The Mincer-Zarnowitz regression is then applied to further evaluate the forecasting accuracy. The Diebold-Mariano test emerges as the decisive factor in selecting the optimal forecast for subsequent implementation in the development of trading strategies. The strategies commence with the definition of confidence intervals, manifested as upper and lower barriers. The trading strategy hinges on strategic buy and sell decisions, executed when prices breach below or above the confidence interval, respectively. Subsequent to these transactions, meticulous calcula-

[1]The prices taken in the analysis that run from the beginning of 2020 to the end of 2021 are stable for almost the entire time frame, but at the end we see the beginning of a dizzying bullish trend in prices. This disruptive trend so unprecedented in energy ("a financial asset class considered by many to be dead") was made possible by a violent geopolitical crisis caused by the Russia of "czar" Vladimir Putin. Russia for years after the Euromaidan crisis has been trying to impose its own "puppet" in Ukraine, but after the Crimean crisis and the Donbass war the former Soviet country has only one card left to play. A mobilization of more than 100,000 soldiers of the Russian Armed Forces was observed on Ukraine's borders in March and April 2021; representing the largest mobilization since the end of the Cold War, this triggered a serious international crisis, which was later contained after a meeting in Geneva between U.S. President Joe Biden and Russian President Vladimir Putin. The crisis had a serious moment when a Russian Navy ship and two aircraft opened fire at a British Navy ship, firing warning shots and dropping bombs a few dozen meters from the NATO ship. The crisis re-exploded violently when a large Russian mobilization of more than 100,000 troops was observed again in October 2021. By January 2022, Russian forces on the Ukrainian border numbered 150,000-180,000 men. Such a massive exercise only presaged the worst. Markets reacted by sending Energy commodity prices soaring, especially power Italy.

tions of returns and percentage returns are performed, illuminating the Return on Equity (ROE) to gauge the investment's performance. This chapter encapsulates a systematic and statistically sound approach, where forecasting prowess converges with trading strategy development, ensuring a robust and informed foray into the dynamics of POWER futures.

4.2 Trading Strategies Applications and Result

4.2.1 Future Q1

Primarily, data is sourced from the Excel file, previously obtained through Reuters DataStream. The columns containing price and date information are extracted. Subsequently, the data that is the daily Italian future power price[2] Q1 2021 quoted in 2020 and Q1 2022 quoted in 2021 is taken[3]. Following this, leveraging the cross-validation method, the forecast for the year 2021 is conducted.

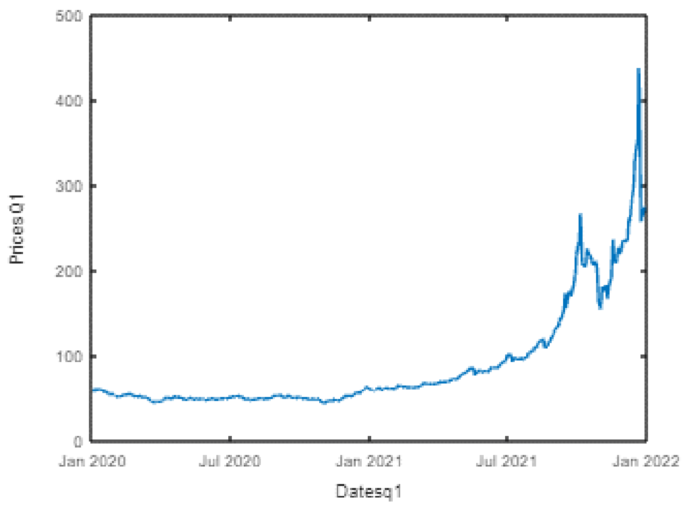

Figure 4.1: There is the energy future power price Q121 and Q122 quoted in the data 2020 2021 Q1.The prices are on the y-axis, and the dates are on the x-axis.

Following the data download, the subsequent steps involve the execution and calculation of yields. To ensure the attainment of stationary data , logarithmic returns are computed.

[2] ITALY BASELOAD QUARTERLY EEX

[3] The exchange in this case is the EEX

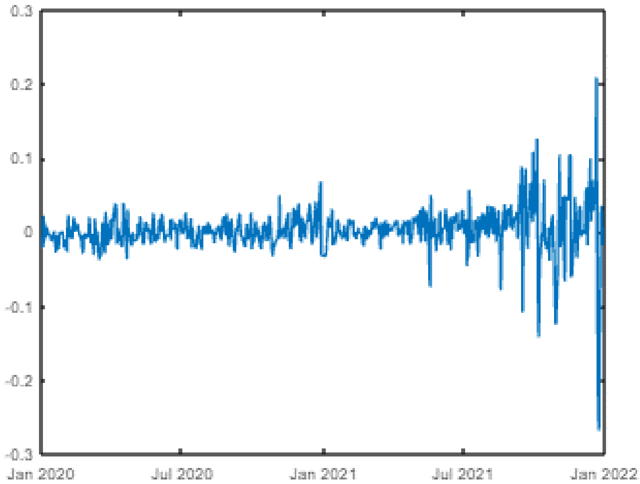

Figure 4.2: There are the Log returns of the 2021Q1 quoted in 2020 and 2022Q1 quoted in 2021.The dates are on the x-axis and the logreturns on the y-axis.

The next step is doing ACF on a sample of 20 Lags. This is important to understand the property of the series and if is useable the ARIMA model.

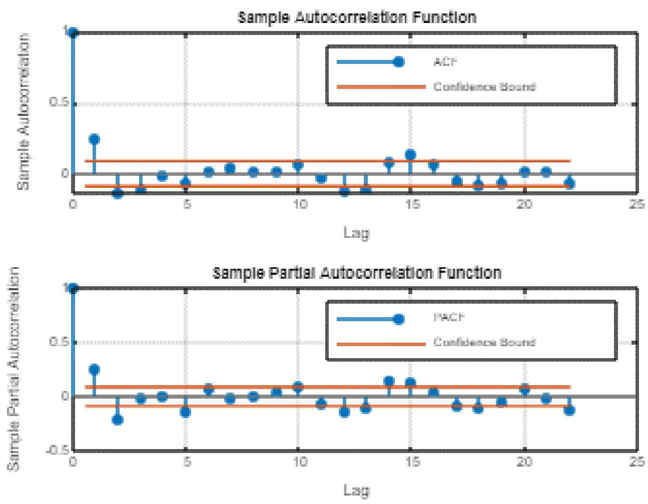

Figure 4.3: The ACF and PACF show autocorrelation and partial autocorrelation.The analisys of autocorrelation shows an autocorrelations at lag 1 that is significative, that justify the use of the model.

This aspect is beneficial for the analysis as it signifies the feasibility of conducting future estimation and forecasting processes. We are going to do the Training Sample, which comprises data from the year 2020. This dataset serves as the foundation for training the ARIMA model to predict data for the following year, 2021. Diverse forecasts are generated to correspond with different orders within the ARIMA model. Conversely, the Test_Set comprises the actual outcomes and serves as a benchmark for evaluating the efficacy of the forecasts. This comparison allows for a comprehensive assessment of the forecast performance against the real-world data. As done by Xue et al. (2022) we are going to estimate the forecast through the ARIMA model. Once completed tests will be run so that the various estimations can then be examined. The subsequent step involves assessing predictive accuracy through mean square errors. This entails comparing the forecasts with the test set. In this context, the objective of the analysis is to minimize the mean square errors, aiming for the most accurate prediction. The formula shows that

Forecast	MSE
Fore_000	0.0017370
Fore_100	0.0016809
Fore_001	0.0016066
Fore_101	0.0016435
Fore_200	0.0016013
Fore_201	0.0016607
Fore_002	0.0016484
Fore_102	0.0016388
Fore_202	0.0017047

Table 4.1: In this table there are the different forecast with their MSE; Fore_000=0.0017370, Fore_100=0.0016809, Fore_001=0.0016066, Fore_101=0.0016435, Fore_200=0.0016013, Fore_201=0.0016607, Fore_002=0.0016484, Fore_102=0.0016388, Fore_202=0.0017047

the best MSE is the fore_200 . So we can say that these forecast is the best estimations for now, since is the ones that is closest to the testset. The subsequent stage involves the implementation of the Mincer and Zarnowitz regression, employed to examine and assess the properties inherent in various forecasts. This regression analysis serves as a methodological tool for scrutinizing the accuracy and reliability of the forecasted values, thereby contributing to a comprehensive evaluation of forecasting performance.As done by Byun and Cho (2013) we are going to implement the regression.

Forecast	ALPHA	BETA
fore000	0.004 (0.08)	-0.1 (0.97090)
fore100	0.016 (0.86)	0.89 (0.00090)
fore001	0.001 (0.01)	0.97 (0.00070)
fore101	0.004 (0.41)	1.01 (0.00012)
fore200	0.001 (0.01)	1.07 (0.00002)
fore201	0.024 (0.18)	1.34 (0.00340)
fore002	0.005 (0.45)	0.76 (0.00780)
fore102	0.004 (0.32)	0.90 (0.00145)
fore202	0.006 (0.02)	-0.8 (0.64000)

Table 4.2: Table that shows the alpha and beta of the Mincer and Zarnowitz regression.The value between parentesis are the p-values.fore000=0.004 (0.08)and beta= -0.1 (0.97090);fore100=0.016 (0.86)and beta=0.89 (0.00090);fore001=0.001 (0.01)and beta=0.97 (0.00070);fore101=0.004 (0.41)and beta=1.01 (0.00012);fore200=0.001 (0.01) and beta=1.07 (0.00002);fore201=0.024 (0.18)and beta =1.34 (0.00340);fore002=0.005 (0.45) and beta=0.76 (0.00780);fore102=0.004 (0.32)and beta=0.90 (0.00145);fore202=0.006 (0.02) and beta=-0.8 (0.64000)

As done by Drachal (2021) we are going to develop the Diebold-Mariano test , conducted on the best-estimated forecasts(fore_200 and fore_001 based on MSE), reveals a positive relationship between the two forecasts. So because the MSE of the fore_200 is smaller than MSE of the fore_001 , we are going to use this in the trading strategies. The articulated trading strategy comprises the following key elements: Utilization of the best forecast Definition of two confidence intervals (upper and lower) Execution of a buy order when yields fall below the lower threshold. Execution of a sell order when yields surpass the upper threshold. The final step involves the computation of returns on the investment . The selected company is allocated a budget of 1,000,000 .Upon observing returns falling below the lower band, the trading bot initiates a purchase of 20,000, subsequently selling when returns surpass the upper band. Employing a confidence interval of 70% the profit margin is determined to be 4.71%. This outcome signifies a positive result for the investment strategy, indicating its viability and success.

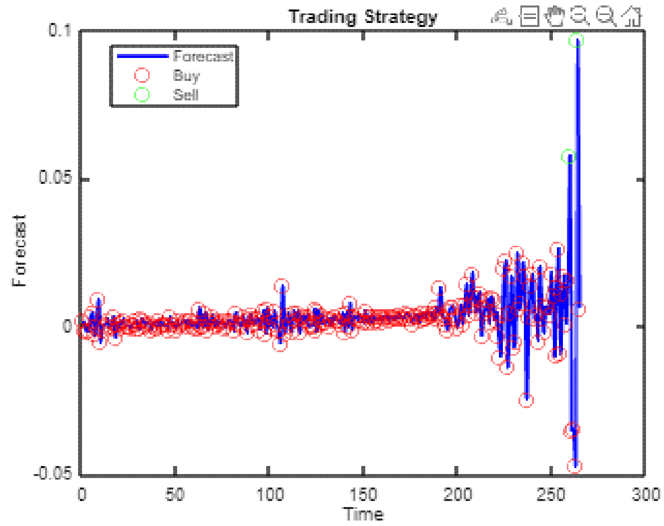

Figure 4.4: Showcases the application of the trading startegies. The green circles represent the selling points , and the red circles represent the buying point.

4.2.2 Future Q2

Initially, data is retrieved from an Excel file, which was originally acquired via Reuters DataStream. The columns containing price and date details are then extracted. Subsequently, the data pertaining to the daily Italian future power prices for Q2 2021 quoted in 2020 and Q2 2022 quoted in 2021 are isolated[4]. Using the cross-validation method, a forecast for the year 2021 is then generated.

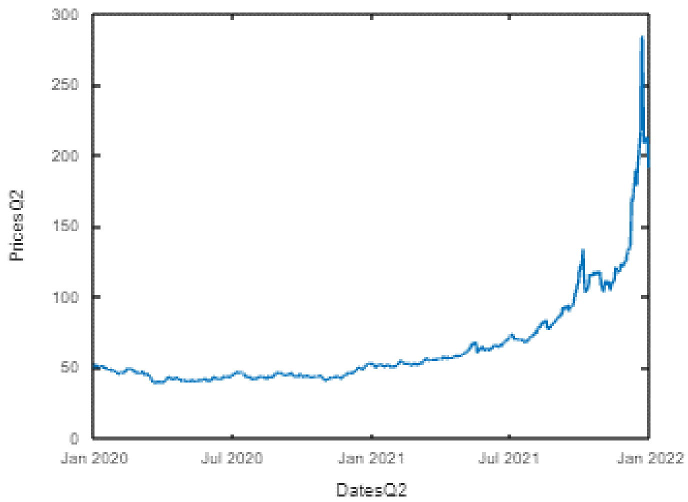

Figure 4.5: There is the power future price Q221 and Q222 quoted in the data 2020 2021 Q1.The prices are on the y-axis, and the dates are on the x-axis.

[4]These prices are sourced from the EEX exchange

On these data, the execution and calculation of yields are realized. Logarithmic returns are computed.

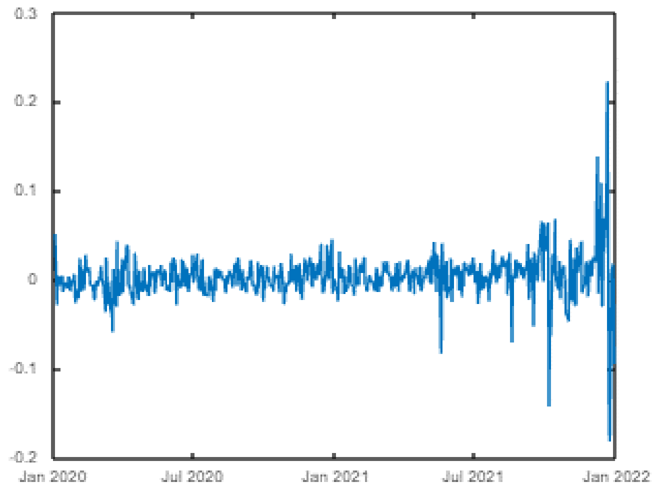

Figure 4.6: There are the Log returns of the 2021 Q2 quoted in 2020 and 2022Q2 quoted in 2021.The dates are on the x-axis and the logreturns on the y-axis.

The subsequent procedure entails conducting an Autocorrelation Function (ACF) analysis spanning a sample of 20 lags.

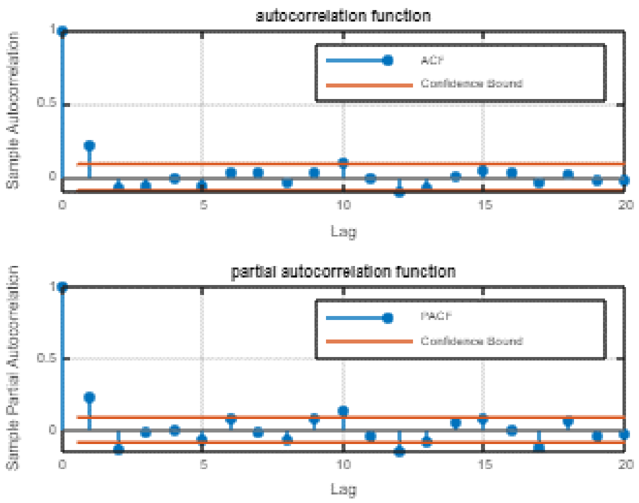

Figure 4.7: The ACF and PACF show autocorrelation and partial autocorrelation.The analisys of autocorrelation shows an autocorrelations at lag 1 that is significative, that justify the use of the model.

This aspect is useful for the analysis as it denotes the utility of conducting future estimation and forecasting processes. The Training Sample, comprising data from the year 2020. This dataset serves as the basis for training the ARIMA model to forecast data for the ensuing year, 2021. Multiple forecasts are generated corresponding to different orders within the ARIMA model. On the other hand , the Test_Set comprises the actual outcomes and serves as a benchmark for evaluating the efficacy of the forecasts. This comparison allows for a comprehensive assessment of the forecast performance against the real-world data. As done by Gao et al. (2017) we are going to carry out the ARIMA model . Different tests will be conducted to scrutinize the various estimations. Subsequently, the process involves evaluating predictive accuracy through mean square errors. This involves comparing the forecasts with the test set. In this context, the aim of the analysis is to minimize the mean square deviation, striving for the utmost precision in prediction. As done by Manchalwar et al. (2023) we are goin to implement and doing different MSE comparing the different forecast to the test_set.

Forecast	MSE
Fore_100	0.0010681
Fore_001	0.0010393
Fore_101	0.0010724
Fore_200	0.0010644
Fore_201	0.0010903
Fore_002	0.0010208
Fore_102	0.0011014
Fore_202	0.0011584
Fore_000	0.0013270

Table 4.3: In this table there are the different forecast with their MSE. Fore_000=0.001270, Fore_100=0.0010681, Fore_001=0.0010393, Fore_101=0.0010724, Fore_200=0.0010644, Fore_201=0.0010903, Fore_002=0.0010208, Fore_102=0.0011014, Fore_202=0.0011584

The formula shows that the best MSE is the one for forecast fore_002 . So we can say that these one is the best estimations for now, since this one is the closest to the testset. The next step is the application of the Mincer & Zarnowitz regression, used to test and carp the properties of the different forecasts. The result of Mincer and Zarowitz regression is: The Parameter Values are quite good since the first beta

Forecast	ALPHA	BETA
fore000	0.005 (0.18)	-0.1 (0.97090)
fore100	0.003 (0.21)	0.83 (0.00400)
fore001	0.003 (0.21)	0.84 (0.00400)
fore101	0.002 (0.31)	1.04 (0.00019)
fore200	0.002 (0.33)	1.07 (0.00002)
fore201	0.003 (0.25)	1.07 (0.00020)
fore002	0.003 (0.25)	0.96 (0.00670)
fore102	0.003 (0.02)	1.01 (0.00012)
fore202	0.006 (0.02)	-0.8 (0.64000)

Table 4.4: Table that shows the alpha and beta of the Mincer and Zarnowitz regression.The value between parentesis are the p-values.fore000=0.005 (0.18)and beta= -0.1 (0.97090);fore100=0.003 (0.21)and beta=0.83 (0.00400);fore001=0.003 (0.21)and beta=0.84 (0.0400);fore101=0.002 (0.31)and beta=1.04 (0.00019);fore200=0.002(0.33) and beta=1.07 (0.00002);fore201=0.003(0.25)and beta =1.07 (0.00020);fore002=0.003 (0.25) and beta=0.96 (0.00670);fore102=0.003 (0.02)and beta=1.01 (0.00012);fore202=0.006 (0.02) and beta=-0.8 (0.64000)

coefficient must tend to one while the second one which is alpha must tend to 0. The next test that will be done to evaluate the best forecast for the investment strategy as done by Chen et al. (2021)is the Diebold Mariano. As done by Xu and Zhang (2023) we are going to use the Diebold-Mariano test that is conducted on the best-estimated forecasts(fore_001 and fore_002 based on MSE).The test reveals a negative relationship between the two forecasts. The difference between the two test shows that fore_002 is better. We are going to use the "fore_002" . Subsequently, the focus shifts towards the formulation and implementation of a trading strategy. The articulated trading strategy comprises the following key elements and components: Utilization of the optimal forecast. Establishment of two confidence intervals (upper and lower). Initiation of a buy order when yields decline below the lower threshold. Initiation of a sell order when yields exceed the upper threshold. The concluding step involves computing returns on the investment. The designated company is allocated a budget of 1,000,000. Upon observing returns dipping below the lower band, the trading bot initiates a purchase of 20,000 units, subsequently selling when returns exceed the upper band. Employing a confidence interval of 70%, the profit margin is determined to be 389.97%. This outcome signifies a positive result for the investment strategy, indicating its viability and success

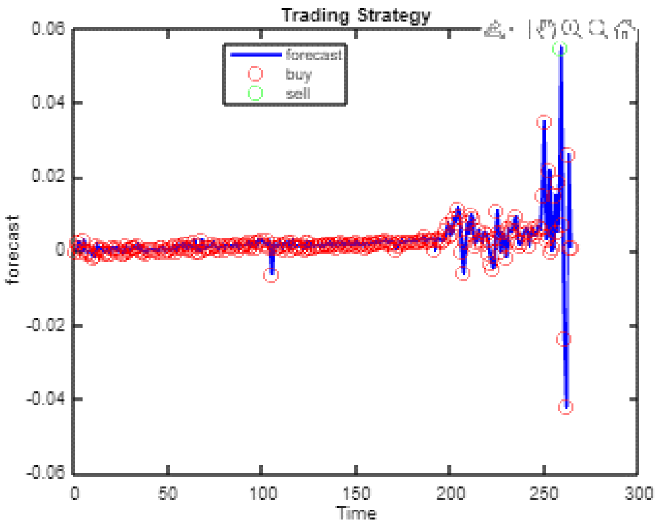

Figure 4.8: Showcases the application of the trading startegies. The green circles represent the selling points , and the red circles represent the buying point.

4.2.3 Future Q3

As always, data is sourced from the Excel file, previously obtained through Reuters DataStream. In this case data is the daily Italian future power price Q3 2021 quoted in 2020 and Q3 2022 quoted in 2021[5]. Following this, leveraging the cross-validation method, the forecast for the year 2021 is conducted.

[5]The exchange in this case is the EEX

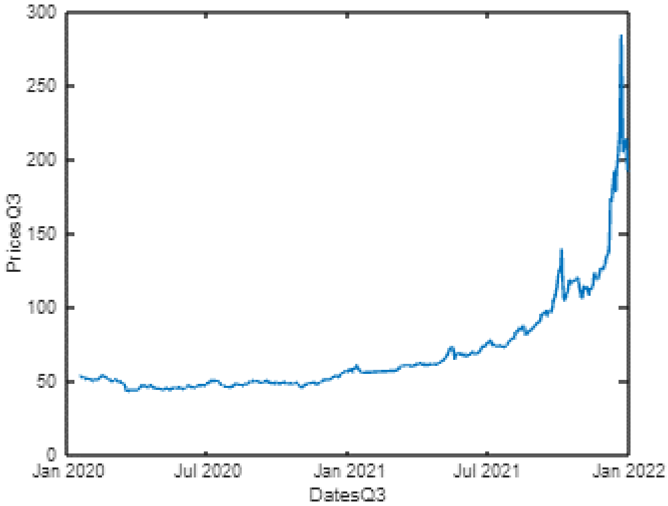

Figure 4.9: There is the power future price Q321 and Q322 quoted in the data 2020 2021 Q3.The prices are on the y-axis, and the dates are on the x-axis.

Following the data download, the subsequent steps involve the execution and calculation of yields. Logarithmic returns are computed.

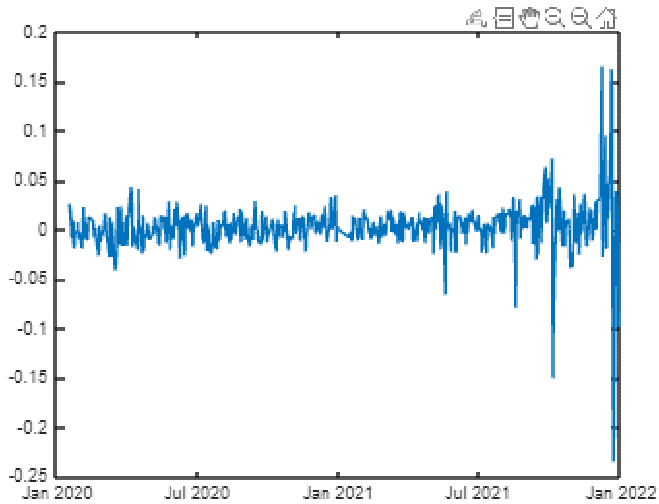

Figure 4.10: There are the Log returns of the 2021 Q3 quoted in 2020 and 2022Q3 quoted in 2021.The dates are on the x-axis and the logreturns on the y-axis.

The next step is doing ACFand PACF on a sample of 20 Lags.

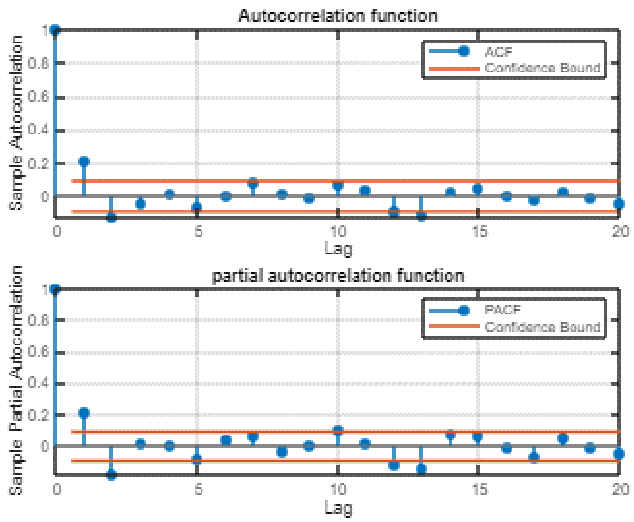

Figure 4.11: The ACF and PACF show autocorrelation and partial autocorrelation.The analisys of autocorrelation shows an autocorrelations at lag 1 that is significative, that justify the use of the model.

This aspect is advantageous for the analysis as it signifies the feasibility of conducting future estimation and forecasting processes. The Training Sample, comprising data from the year 2020. This dataset is employed to train the ARIMA model to predict data for the following year, 2021. Diverse forecasts are producted . Conversely, the Test_Set comprises the actual outcomes and serves as a benchmark for evaluating the efficacy of the forecasts. This comparison allows for a comprehensive assessment of the forecast performance against the real-world data. Subsequently, the procedure entails evaluating predictive accuracy through the utilization of mean square errors. This involves juxtaposing the forecasts against the test set. Within this framework, the primary objective of the analysis is to minimize mean square errors, thereby endeavoring to achieve the utmost precision in prediction. As done by Miseta et al. (2022) we are going to perform the MSE.

Forecast	MSE
Fore_100	0.0012573
Fore_001	0.0012972
Fore_101	0.0012349
Fore_200	0.0012321
Fore_201	0.0012670
Fore_002	0.0012257
Fore_102	0.0012398
Fore_202	0.0013419
Fore_000	0.001423

Table 4.5: In this table there are the different forecast with their MSE. Fore_000=0.001270, Fore_100=0.0012573, Fore_001=0.0012972, Fore_101=0.0012349, Fore_200=0.0012321, Fore_201=0.0012670, Fore_002=0.0012257, Fore_102=0.0012398, Fore_202=0.0013419

The formula shows that the best MSE is the fore_002 . So we can say that these two are the best estimations for now, since they are the ones that are closest to the testset. The next step is the application of the Mincer & Zarnowitz regression, used to test and carp the properties of the different forecasts. As done by Bhattacharya and Gupta (2016) the regression is going to by estimate and the results of Mincer and Zarowitz regression are:

Forecast	ALPHA	BETA
fore000	0.003 (0.13)	0.98 (0.56000)
fore100	0.003 (0.13)	0.59 (0.12757)
fore001	0.003 (0.21)	0.84 (0.00400)
fore101	0.003 (0.20)	0.94 (0.01861)
fore200	0.003 (0.21)	0.82 (0.01515)
fore201	0.004 (0.08)	0.49 (0.12764)
fore002	0.002 (0.29)	1.05 (0.00839)
fore102	0.003 (0.15)	0.83 (0.02589)
fore202	0.005 (0.02)	-0.9 (0.74000)

Table 4.6: Table that shows the alpha and beta of the Mincer and Zarnowitz regression.The value between parentesis are the p-values.fore000=0.003(0.13)and beta=0.98 (0.00060);fore100=0.003 (0.12)and beta=0.59 (0.12575);fore001=0.003 (0.21)and beta=0.84 (0.0400);fore101=0.003 (0.20)and beta=0.94 (0.018);fore200=0.003(0.21) and beta=0.82 (0.01515);fore201=0.004(0.08)and beta =1.05 (0.00839);fore002=0.002 (0.29) and beta=1.05 (0.00839);fore102=0.003 (0.15)and beta=0.83 (0.02589);fore202=0.005(0.02) and beta=-0.9 (0.74000).

The Parameter Values are accetptables. As done as before in the other applications now is the step of the Diebold Mariano. The Diebold-Mariano test reveals a positive relationship between the two forecast(fore_001 and fore_002). The decision of which forecast using is totally subjective. As done in the other squarters now there is the final step of the applivcations of the trading strategies. The process of the strategy is the same as before. Employing a confidence interval of 70%, the profit margin is determined to be 6.56%. This outcome signifies a positive result for the investment strategy, indicating its viability and success

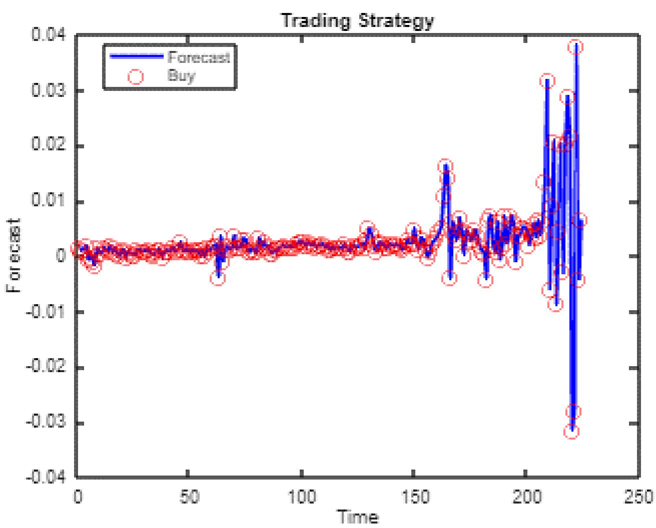

Figure 4.12: Showcases the application of the trading startegies. The green circles represent the selling points,and the red circles represent the buying point.

4.2.4 Future Q4

Initially, data is extracted from the Excel file, which was obtained through Reuters DataStream. Specifically, the columns containing price and date information are retrieved. Next, the data pertaining to the daily Italian future power price for Q4 2021 quoted in 2020 and Q4 2022 quoted in 2021 are selected[6]. Subsequently, utilizing the cross-validation method, forecasting for the year 2021 is performed

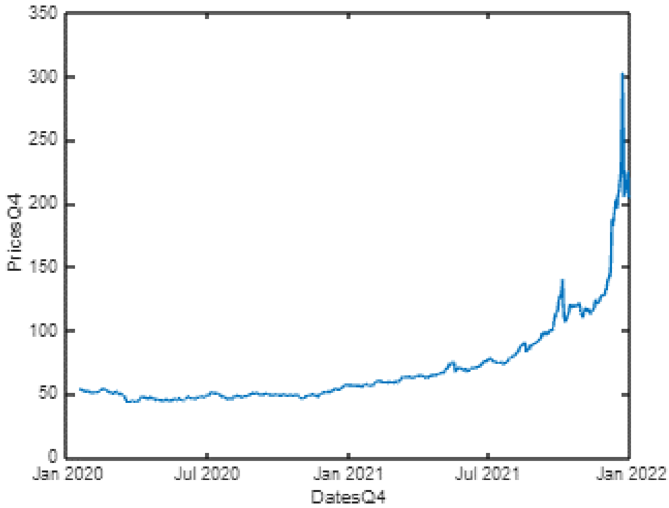

Figure 4.13: There is the power future price Q421 and Q422 quoted in the data 2020 2021 Q4.The prices are on the y-axis, and the dates are on the x-axis.

After the data has been downloaded, the subsequent steps entail executing and computing yields.Logarithmic returns are computed.

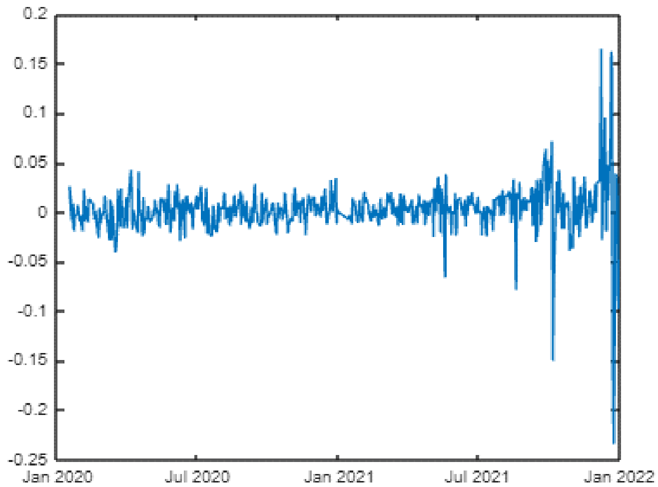

Figure 4.14: There are the Log returns of the 2021Q4 quoted in 2020 and 2022Q4 quoted in 2021.The dates are on the x-axis and the logreturns on the y-axis.

The next step is doing ACF and PACF on a sample of 20 Returns. This is important to understand the property of the series and if is useable the ARIMA model.

[6]The exchange used in this case is the EEX

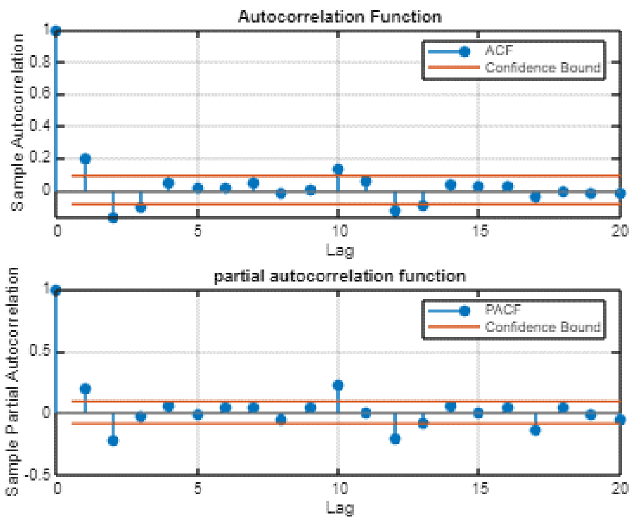

Figure 4.15: The ACF and PACF show autocorrelation and partial autocorrelation.The analisys of autocorrelation shows an autocorrelations at lag 1 that is significative, that justify the use of the model..

This is good for the analysis because it indicates that a future estimation and forecasting process can take place. The Training Sample, encompassing the data from the year 2020. This dataset is utilized to train the ARIMA model for predicting the data in the subsequent year, 2021 . Various forecasts are generated corresponding to distinct order within the ARIMA model. Conversely, the Test Set comprises the actual outcomes and serves as a benchmark for evaluating the efficacy of the forecasts. This comparison allows for a comprehensive assessment of the forecast performance against the real-world data. As done by Gao et al. (2017) we are going to perform the ARIMA model . Following completion, tests will be conducted to scrutinize the various estimations. Subsequently, the procedure involves assessing predictive accuracy through the utilization of mean square errors. This involves meticulous comparison of the forecasts with the test set. Within this framework, the primary objective of the analysis is to minimize mean square errors, with the aim of achieving the utmost precision in prediction.

Forecast	MSE
Fore_100	0.0012238
Fore_001	0.0018319
Fore_101	0.0011887
Fore_200	0.0011742
Fore_201	0.0012136
Fore_002	0.0011753
Fore_102	0.0012163
Fore_202	0.0012648

Table 4.7: In this table there are the different forecasts with their MSE. Fore_000=0.001270, Fore_100=0.0012573, Fore_001=0.0012972, Fore_101=0.0012349, Fore_200=0.0012321, Fore_201=0.0012670, Fore_002=0.0012257, Fore_102=0.0012398, Fore_202=0.001341

The equation indicates that the optimal Mean Squared Error (MSE) corresponds to the forecast fore_200. Therefore, we can assert that this forecast exhibits the most favorable estimations at present, as it closely aligns with the test set. The subsequent stage involves implementing the Mincer & Zarnowitz regression, aimed at examining and elucidating the characteristics of the various forecasts. As done by Bhattacharya and Gupta (2016) the regression is going to by estimate and the outcome of the Mincer and Zarnowitz regression is:

Forecast	ALPHA	BETA
fore000	0.002 (0.04)	1.25 (0.43408)
fore100	0.004 (0.05)	0.26 (0.38222)
fore001	0.003 (0.21)	0.84 (0.00400)
fore101	0.003 (0.12)	0.60 (0.05186)
fore200	0.003 (0.16)	0.65 (0.00858)
fore201	0.003 (0.08)	0.41 (0.00080)
fore002	0.003 (0.18)	0.74 (0.01793)
fore102	0.004 (0.08)	0.35 (0.19262)
fore202	0.004 (0.04)	0.12 (0.60000)

Table 4.8: Table that shows the alpha and beta of the Mincer and Zarnowitz regression.The value between parentesis are the p-values.fore000=0.002(0.04)and beta=1.25 (0.43408);fore100=0.004(0.05)and beta=0.26(0.38222);fore001=0.003 (0.21)and beta=0.84 (0.0400);fore101=0.003 (0.12)and beta=0.60 (0.05186);fore200=0.003(0.16) and beta=0.65 (0.00858);fore201=0.003(0.18)and beta =0.41 (0.00080);fore002=0.003(0.18) and beta=0.74 (0.01793);fore102=0.004(0.08)and beta=0.35(0.19262);fore202=0.004(0.04) and beta=0.12 (0.60000)

The Parameter Values are quite good since the first beta coefficient must tend to one while the second one which is alpha must tend to 0. The next test that will be done to evaluate the best forecast for the investment strategy is the Diebold Mariano as done by Chen et al. (2021). The Diebold-Mariano test , conducted on the best-estimated forecasts(fore_200and fore_002 based on the MSE), reveals a negative relationship between the two forecast. The best forecast come out from the difference is the fore_200. Subsequently, the focus shifts towards the formulation and implementation of a trading strategy as used in the other quarters. Employing a confidence interval of 70%, the profit margin is determined to be 387,87%. This outcome signifies a positive result for the investment strategy, indicating its viability and success

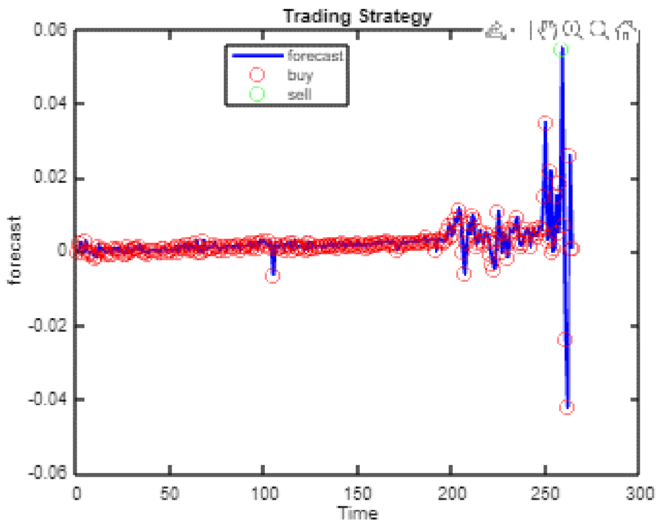

Figure 4.16: Showcases the application of the trading startegies. The green circles represent the selling points , and the red circles represent the buying point.

Chapter 5

CONCLUSION

The primary objective of this thesis is to formulate a profitable trading strategy centered on Italian power futures prices for the 2022 quarters, as quoted in the year 2021.The selection of specific quarters for strategy implementation have been based on a comprehensive comparison of yields. The overarching goal has achieved the desired Return on Investment (ROI). In the realm of energy companies, particularly in Italy, the prevalent practices involve hedging or market access trading to mitigate contract risks. Speculative trading, once widely employed, has diminished, especially in the aftermath of numerous energy company bankruptcies during the early stages of the Russia-Ukraine War. This thesis has aimed at developing a trading strategy that, while not conforming entirely to the conventional speculative energy trading strategy, endeavors to adopt a methodology applicable to other return-based asset classes. In contrast to the typical energy trading strategy, which involves receiving energy contract orders and strategically timing their execution, this approach has started with data initialization. Dates and prices for each quarter over two years (2020 with Q1 2021 and Q1 2022quoted in 2021 as the prediction subject) have been used employing the cross-validation method. The procedural steps have included estimating price returns over the sample duration, calculating log returns and modeling using ARIMA models to generate different forecasts. Comparison of forecasts on graphs, mean square errors computations and Mincer-Zarnowitz regressions have followed. The Diebold-Mariano Test has then been employed to identify the best forecast, subsequently informing the trading strategy. Among the various trading strategies explored, the Q2 strategy emerges with significantly higher returns, making it the preferred choice. This thesis has demonstrated the feasibility of implementing a successful speculative strategy in the energy market. As a prospective extension, constructing an investment portfolio in the commodities sector may be considered. This entails applying Markowitz's Portfolio Theory and studying covariances to ensure non-positive correlations between invested commodities. Diversifying risks may be achieved through different trading strategies within the portfolio, each based on distinct commodities Furthermore, an examination of quantiles would be insightful for integrating risk measures and safeguarding invested capital. Quantile regression, a statistical method, provides detailed insights into the variability of returns distribution tails. Additionally, risk measures such as Value at Risk (VAR)as done by Sadeghi and Shavvalpour (2006) Jackson (2010) Denton et al. (2003)and backtesting applications could be explored for comprehensive risk assessment.

SITOGRAPHY

//ctc.westpoint.edu/houthi-war-machine-guerrilla-war-state-capture/ https://www.mercatoelettrico.org/it/

https://www.sciencedirect.com/science/article/abs/pii/S0306261920316585

https://www.lacimagroup.com/wp-content/uploads/2020/10/energyderivatives_chapter1.pdf

https://www.cmegroup.com/education/courses/introduction-to-ferrous-metals/what-is-contango-and-backwardation.html

https://www.sciencedirect.com/topics/economics-econometrics-and-finance/arma-model

https://www.smallake.kr/wp-content/uploads/2014/04/economics-of-ctfs.pdf

https://dspace.mit.edu/handle/1721.1/61602

https://pubsonline.informs.org/doi/abs/10.1287/mnsc.1090.1049

https://www.sciencedirect.com/science/article/abs/pii/S0301421514006661

https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6419.00172

https://www.eex.com/en/markets/power/power-futures

https://www.nber.org/system/files/chapters/c1214/c1214.pdf

https://academic.oup.com/biomet/article-abstract/77/3/537/253708

https://link.springer.com/chapter/10.1007/978-3-030-60515-5_5

https://www.mdpi.com/2071-1050/10/1/228

https://world-nuclear.org/information-library/country-profiles/countries-a-f/france.aspx

https://www.statista.com/statistics/1267546/france-monthly-wholesale-electricity-price/

https://www.statista.com/statistics/1267541/germany-monthly-wholesale-electricity-price/

https://www.albasoluzioni.com/ https://www.epexspot.com/ https://www.jao.eu/

https://www.jao.eu/resource-center https://yem-energy.it/glossario/eua/

https://www.enjoyenergy.it/blog/notizie/indici-energetici-luglio-2021/

https://luce-gas.it/guida/mercato/pfor https://www.statista.com/statistics/873552/energy-mix-in-italy/

https://www.cleanenergywire.org/factsheets/germanys-energy-consumption-and-power-mix-charts

https://aleasoft.com/it/

https://www.cleanenergywire.org/factsheets/germanys-energy-consumption-and-power-mix-charts

Bibliography

Sujoy Bhattacharya and Pulkit Gupta. Modeling time-varying volatility in indian commodity futures return: Some empirical evidence. *IUP Journal of Financial Risk Management*, 13(4), 2016.

Dominik Boos and Linus Grob. Tracking speculative trading. *Journal of Financial Markets*, 64:100774, 2023.

George E. P. Box and Gwilym M. Jenkins. The estimation of the gradient of a density function, with applications in time series analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 32(2):311–352, 1970.

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

Peter J Brockwell and Richard A Davis. *Introduction to Time Series and Forecasting*. Springer, 2016.

Suk Joon Byun and Hangjun Cho. Forecasting carbon futures volatility using garch models with energy volatilities. *Energy Economics*, 40:207–221, 2013.

Hao Chen, Qiulan Wan, and Yurong Wang. Refined diebold-mariano test methods for the evaluation of wind power forecasting models. *Energies*, 7(7):4185–4198, 2014.

Rongda Chen, Weiwei Bao, and Chenglu Jin. Investor sentiment and predictability for volatility on energy futures markets: Evidence from china. *International Review of Economics & Finance*, 75:112–129, 2021.

Javier Contreras, Rosario Espinola, Francisco J Nogales, and Antonio J Conejo. Arima models to predict next-day electricity prices. *IEEE transactions on power systems*, 18(3):1014–1020, 2003.

Erick Meira de Oliveira and Fernando Luiz Cyrino Oliveira. Forecasting mid-long term electric energy consumption through bagging arima and exponential smoothing methods. *Energy*, 144:776–788, 2018.

Michael Denton, Adrian Palmer, Ralph Masiello, and Petter Skantze. Managing market risk in energy. *IEEE Transactions on Power Systems*, 18(2):494–502, 2003.

Krzysztof Drachal. Forecasting selected energy commodities prices with bayesian dynamic finite mixtures. *Energy Economics*, 99:105283, 2021.

Gao Gao, Kwoklun Lo, and Fulin Fan. Comparison of arima and ann models used in electricity price forecasting for power market. *Energy and Power Engineering*, 9(4B):120–126, 2017.

Kemal Guler, Pin T Ng, and Zhijie Xiao. Mincer–zarnowitz quantile and expectile regressions for forecast evaluations under aysmmetric loss functions. *Journal of Forecasting*, 36(6):651–679, 2017.

Akshita Gupta and Arun Kumar. Mid term daily load forecasting using arima, wavelet-arima and machine learning. In *2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, pages 1–5. IEEE, 2020.

James Douglas Hamilton. *Time Series Analysis*. Princeton University Press, 1994.

Rob J Hyndman and George Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2018.

Jerry Jackson. Promoting energy efficiency investments with risk management decision tools. *Energy policy*, 38(8):3865–3873, 2010.

Orhan Altuğ Karabiber and George Xydis. Electricity price forecasting in the danish day-ahead market using the tbats, ann and arima methods. *Energies*, 12(5):928, 2019.

Spyros Makridakis, Steven C Wheelwright, and Rob J Hyndman. *Forecasting: methods and applications*. John Wiley & Sons, 1998.

Ashwini D Manchalwar, Nita R Patne, BV Surya Vardhan, and Mohan Khedkar. Peer-to-peer energy trading in a distribution network considering the impact of short-term load forecasting. *Electrical Engineering*, pages 1–13, 2023.

Roberto S Mariano. Testing forecast accuracy. *A companion to economic forecasting*, 2:284–298, 2002.

Jacob A Mincer and Victor Zarnowitz. The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, pages 3–46. NBER, 1969.

Tamás Miseta, Attila Fodor, and Ágnes Vathy-Fogarassy. Energy trading strategy for storage-based renewable power plants. *Energy*, 250:123788, 2022.

Paul Newbold. Arima model building and the time series analysis approach to forecasting. *Journal of forecasting*, 2(1):23–35, 1983.

Mehdi Sadeghi and Saeed Shavvalpour. Energy risk management and value at risk modeling. *Energy policy*, 34(18):3367–3373, 2006.

Anna Shikhina, Alexei Kochengin, George Chrysostomou, and Vladimir Shikhin. Investigation of autoregressive forecasting models for market electricity price. In *2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON)*, pages 570–575. IEEE, 2020.

Xiaojie Xu and Yun Zhang. China mainland new energy index price forecasting with the neural network. *Energy Nexus*, page 100210, 2023.

Jian Xue, Jing Ding, Laijun Zhao, Di Zhu, and Lei Li. An option pricing model based on a renewable energy price index. *Energy*, 239:122117, 2022.

George Udny Yule. On a method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226(636-646):267–298, 1927. doi: 10.1098/rsta.1927.0002.

MATLAB CODE

```
% Questo Q1
Data=readtable('DATI␣2021␣tesi.xlsx');
Dataq121=flipud(table2array(Data(1:516,13)));
prezziq121=flipud(table2array(Data(1:516,14)));
plot(Dataq121,prezziq121)


ret=diff(log(prezziq121))
plot(Dataq121(2:516),ret)
subplot(2,1,1);
autocorr(ret, 20);
title('Funzione␣di␣autocorrelazione');

subplot(2,1,2);
parcorr(ret, 20);
title('Funzione␣di␣autocorrelazione␣parziale');

plot(Dataq121(2:516),ret)
xlabel('Date');
ylabel('log␣return');
axis tight;
datetick('x', 'yyyy');

[~,pvalues\_ret]=adftest(ret,'lags',0:2)

%%
subplot(2,1,1)
autocorr(ret,22)
subplot(2,1,2)
parcorr(ret,22)
%%%%%%%%%


wind=250
T=size(ret,1)
fore=zeros(T-wind,1);
test\_set=ret(wind+1:end);


%arima(0,0,0)
```

```
fore\_000=zeros(T-wind,1);
tic
for i=1:(T-wind);

MdL\_000 = arima(0,0,0);
EstMdl = estimate(Mdl\_000,ret(i:wind+i-1));
fore\_000(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc

%%%%%%%%%%arima(1,0,0)%%%%%%%%%%%%%%%%%%%%%
fore\_100=zeros(T-wind,1);
tic
for i=1:(T-wind);

Mdl\_100 = arima(1,0,0);
EstMdl = estimate(Mdl\_100,ret(i:wind+i-1));
fore\_100(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc
fore\_001=zeros(T-wind,1);
tic
for i=1:(T-wind);

Mdl\_001 = arima(0,0,1);
EstMdl = estimate(Mdl\_001,ret(i:wind+i-1));
fore\_001(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc
%%%%%%%%%%arima(1,0,1)%%%%%%%%%%%%%%%%%%%%%
fore\_101=zeros(T-wind,1);
tic
for i=1:(T-wind);

Mdl\_101 = arima(1,0,1);
EstMdl = estimate(Mdl\_101,ret(i:wind+i-1));
fore\_101(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc

fore\_200=zeros(T-wind,1);
tic
```

```
for i=1:(T-wind);

Mdl\_200 = arima(2,0,0);
EstMdl = estimate(Mdl\_200,ret(i:wind+i-1));
fore\_200(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc

fore\_201=zeros(T-wind,1);
tic
for i=1:(T-wind);

Mdl\_201= arima(2,0,1);
EstMdl = estimate(Mdl\_201,ret(i:wind+i-1));
fore\_201(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc

fore\_002=zeros(T-wind,1);
tic
for i=1:(T-wind);

Mdl\_002 = arima(0,0,2);
EstMdl = estimate(Mdl\_002,ret(i:wind+i-1));
fore\_002(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc

fore\_102=zeros(T-wind,1);
tic
for i=1:(T-wind);

Mdl\_102 = arima(1,0,2);
EstMdl = estimate(Mdl\_102,ret(i:wind+i-1));
fore\_102(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc

fore\_202=zeros(T-wind,1);
tic
for i=1:(T-wind);

```

```
Mdl\_202 = arima(2,0,2);
EstMdl = estimate(Mdl\_202,ret(i:wind+i-1));
fore\_202(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
toc


plot(test\_set(1:end),"black")
hold on
plot(fore\_100,'r')
hold on
plot(fore,'cyan')
hold on
plot(fore\_101,'green')
hold on
plot(fore\_200,'red')
hold on
plot(fore\_201,'yellow')
hold on
plot(fore\_002,'magenta')
hold on
plot(fore\_102,'white')
hold on
plot(fore\_202,'blue')

if ~isempty(test\_set)\\\&\& all(isfinite(test\_set(2:end)))\\ \&\& all(isfinite(fore))
MSE = sum((test\_set(1:end)-fore(1:end)).\^2) / numel(test\_set(1:end));
MSE\_100 = sum((test\_set(1:end)-fore\_100).\^2) / numel(test\_set(1:end));
MSE\_101 = sum((test\_set(1:end)-fore\_101).\^2) / numel(test\_set(1:end));
MSE\_200 = sum((test\_set(1:end)-fore\_200).\^2) / numel(test\_set(1:end));
MSE\_201 = sum((test\_set(1:end)-fore\_201).\^2) / numel(test\_set(1:end));
MSE\_002 = sum((test\_set(1:end)-fore\_002).\^2) / numel(test\_set(1:end));
MSE\_102 = sum((test\_set(1:end)-fore\_102).\^2) / numel(test\_set(1:end));
MSE\_001 = sum((test\_set(1:end)-fore\_001).\^2) / numel(test\_set(1:end));
MSE\_202 = sum((test\_set(1:end)-fore\_202).\^2) / numel(test\_set(1:end));
M = [MSE, MSE\_100, MSE\_101, MSE\_200, MSE\_201, MSE\_002, MSE\_102, MSE\_202];
else
disp('Warning: Test set contains NaN or Inf values, MSE calculation skipped.');
end


```

```
param0=fitlm(fore\_002,test\_set(1:end),1)
param2=fitlm(fore\_100,test\_set(1:end),1)
param3=fitlm(fore\_101,test\_set(1:end),1)
param4=fitlm(fore\_102,test\_set(1:end),1)
param6=fitlm(fore\_200,test\_set(1:end),1)
param7=fitlm(fore\_201,test\_set(1:end),1)
param8=fitlm(fore\_000,test\_set(1:end),1)
param9=fitlm(fore\_202,test\_set(1:end),1)
param10=fitlm(fore\_001,test\_set(1:end),1)


%%%DM
DM\_FINAL=dmtest(fore\_200-test\_set,fore\_001-test\_set,1)

FORE\_final=MSE\_200-MSE\_001





capitale\_iniziale = 1000000;
capitale = capitale\_iniziale;
investimento\_fisso = 20000;
posizione = 0;


confidenza = 0.7;


for t = 1:length(fore\_200)

previsione = fore\_200(t);
errore\_std = sqrt(EstMdl.Variance);

previsione\_intervallo\_inferiore = previsione - norminv((1 - confidenza) / 2) *
    errore\_std;
previsione\_intervallo\_superiore = previsione + norminv((1 + confidenza) / 2) *
    errore\_std;

if previsione < previsione\_intervallo\_inferiore \\\&\& posizione <= 0

quantita\_acquistata = investimento\_fisso / previsione;
```

```
capitale = capitale - investimento\_fisso;
posizione = posizione + quantita\_acquistata;
fprintf('Compra␣\%f␣unit␣di␣asset␣a␣t=\%d\n', quantita\_acquistata, t);
elseif previsione > previsione\_intervallo\_superiore \&\& posizione >= 0

quantita\_venduta = min(posizione, investimento\_fisso / previsione);
capitale = capitale + quantita\_venduta * previsione;
posizione = posizione - quantita\_venduta;
fprintf('Vendi␣\%f␣unit␣di␣asset␣a␣t=\%d\n', quantita\_venduta, t);
end
end


valore\_finale = capitale + posizione * previsione;

fprintf('Valore␣finale␣dell␣investimento:␣\%.2f\n', valore\_finale);
profitto=valore\_finale-capitale


ritorno = (valore\_finale - capitale\_iniziale) / capitale\_iniziale;
profitto\_percentuale = (valore\_finale - capitale\_iniziale) / capitale\_iniziale * 100;

fprintf('Ritorno␣sull''investimento␣(ROI):␣\%.2f\%\%\n', ritorno * 100);
fprintf('Profitto␣percentuale:␣\%.2f\%\%\n', profitto\_percentuale);


tempi = 1:length(fore\_200);


figure;
plot(tempi, fore\_200, 'b-', 'LineWidth', 2);
hold on;


compra = fore\_200(fore\_200 < previsione\_intervallo\_inferiore);
vendi = fore\_200(fore\_200 > previsione\_intervallo\_superiore);
plot(find(fore\_200 < previsione\_intervallo\_inferiore), compra, 'ro', 'MarkerSize', 8);
plot(find(fore\_200> previsione\_intervallo\_superiore), vendi, 'go', 'MarkerSize', 8);

```

```
title('Trading␣Strategy');
xlabel('Tempo');
ylabel('Previsioni');
legend('Previsioni', 'Acquisto', 'Vendita', 'Location', 'Best');

hold off;
%Q2
Data=readtable('DATI␣2021␣tesi.xlsx');
Dataq221=flipud(table2array(Data(1:516,16)));
prezziq221=flipud(table2array(Data(1:516,17)));
plot(Dataq221,prezziq221)
xlabel('prezziq221')
ylabel('Dataq221')


ret=diff(log(prezziq221))
plot(Dataq221(2:516),ret)
subplot(2,1,1);
autocorr(ret, 20);
title('Funzione␣di␣autocorrelazione');

subplot(2,1,2);
parcorr(ret, 20);
title('Funzione␣di␣autocorrelazione␣parziale');

plot(Dataq221(2:516),ret)
xlabel('Date');
ylabel('log␣return');
axis tight;
datetick('x', 'yyyy');

[~,pvalues\_ret]=adftest(ret,'lags',0:2)

%%
subplot(2,1,1)
autocorr(ret,22)
subplot(2,1,2)
parcorr(ret,22)
%%%%%%%%%
```

```
wind=250
T=size(ret,1)
fore=zeros(T-wind,1);
test\_set=ret(wind+1:end);


%%%%%%arima(0,0,0)
fore\_000=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_000 = arima(0,0,0);
    EstMdl = estimate(Mdl\_000,ret(i:wind+i-1));
    fore\_000(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

%%%%%%%%%%arima(1,0,0)
fore\_100=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_100 = arima(1,0,0);
    EstMdl = estimate(Mdl\_100,ret(i:wind+i-1));
    fore\_100(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

 fore\_001=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_001 = arima(0,0,1);
    EstMdl = estimate(Mdl\_001,ret(i:wind+i-1));
    fore\_001(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

%%%%%%%%%%arima(1,0,1)
fore\_101=zeros(T-wind,1);
tic
for i=1:(T-wind);
```

```
    Mdl\_101 = arima(1,0,1);
    EstMdl = estimate(Mdl\_101,ret(i:wind+i-1));
    fore\_101(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc


 %%%%%%%%%%arima(2,0,0)%%%%%%%%%%%%%%%%%%%%%%
fore\_200=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_200 = arima(2,0,0);
    EstMdl = estimate(Mdl\_200,ret(i:wind+i-1));
    fore\_200(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc



fore\_201=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_201= arima(2,0,1);
    EstMdl = estimate(Mdl\_201,ret(i:wind+i-1));
    fore\_201(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc


fore\_002=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_002 = arima(0,0,2);
    EstMdl = estimate(Mdl\_002,ret(i:wind+i-1));
    fore\_002(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc
```

```
fore\_102=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_102 = arima(1,0,2);
    EstMdl = estimate(Mdl\_102,ret(i:wind+i-1));
    fore\_102(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

fore\_202=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_202 = arima(2,0,2);
    EstMdl = estimate(Mdl\_202,ret(i:wind+i-1));
    fore\_202(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc


plot(test\_set(1:end),"cyan")
hold on
plot(fore\_100,'r')
hold on
plot(fore,'black')
hold on
plot(fore\_101,'green')
hold on
plot(fore\_200,'red')
hold on
plot(fore\_201,'yellow')
hold on
plot(fore\_002,'magenta')
hold on
plot(fore\_102,'white')
hold on
plot(fore\_202,'blue')
```

```
if ~isempty(test\_set)\ \&\& all(isfinite(test\_set(2:end))) \&\& all(isfinite(fore))
    MSE = sum((test\_set(1:end)-fore(1:end)).\^2) / numel(test\_set(1:end));
    MSE\_100 = sum((test\_set(1:end)-fore\_100).\^2) / numel(test\_set(1:end));
    MSE\_101 = sum((test\_set(1:end)-fore\_101).\^2) / numel(test\_set(1:end));
    MSE\_200 = sum((test\_set(1:end)-fore\_200).\^2) / numel(test\_set(1:end));
    MSE\_201 = sum((test\_set(1:end)-fore\_201).\^2) / numel(test\_set(1:end));
    MSE\_002 = sum((test\_set(1:end)-fore\_002).\^2) / numel(test\_set(1:end));
    MSE\_102 = sum((test\_set(1:end)-fore\_102).\^2) / numel(test\_set(1:end));
    MSE\_202 = sum((test\_set(1:end)-fore\_202).\^2) / numel(test\_set(1:end));
    MSE\_001 = sum((test\_set(1:end)-fore\_001).\^2) / numel(test\_set(1:end));
  M = [MSE, MSE\_100, MSE\_101, MSE\_200, MSE\_201, MSE\_002, MSE\_102, MSE\_202];
else
    disp('Warning: Test set contains NaN or Inf values, MSE calculation skipped.');
end


param0=fitlm(fore\_002,test\_set(1:end),1)
param2=fitlm(fore\_100,test\_set(1:end),1)
param3=fitlm(fore\_101,test\_set(1:end),1)
param4=fitlm(fore\_102,test\_set(1:end),1)
param6=fitlm(fore\_200,test\_set(1:end),1)
param7=fitlm(fore\_201,test\_set(1:end),1)
param8=fitlm(fore\_000,test\_set(1:end),1)
param9=fitlm(fore\_202,test\_set(1:end),1)
param10=fitlm(fore\_001,test\_set(1:end),1)


%%%DM sui migliori
DM\_FINAL=dmtest(fore\_001-test\_set,fore\_002-test\_set,1)

capitale\_iniziale = 1000000;
capitale = capitale\_iniziale;
investimento\_fisso = 20000;
posizione = 0;


confidenza = 0.7;


for t = 1:length(fore\_002)

```

```
    previsione = fore\_002(t);
    errore\_std = sqrt(EstMdl.Variance);


    previsione\_intervallo\_inferiore = previsione - norminv((1 - confidenza) / 2) * errore\_std;
    previsione\_intervallo\_superiore = previsione + norminv((1 + confidenza) / 2) * errore\_std;


    if previsione < previsione\_intervallo\_inferiore\ \&\& posizione <= 0

        quantita\_acquistata = investimento\_fisso / previsione;
        capitale = capitale - investimento\_fisso;
        posizione = posizione + quantita\_acquistata;
        fprintf('Compra␣\%f␣unit␣di␣asset␣a␣t=\%d\n', quantit\_acquistata, t);
    elseif previsione > previsione\_intervallo\_superiore\ \&\& posizione >= 0
         quantita\_venduta = min(posizione, investimento\_fisso / previsione);
        capitale = capitale + quantita\_venduta * previsione;
        posizione = posizione - quantita\_venduta;
        fprintf('Vendi\%f␣unit␣di␣asset␣a␣t=\%d\n', quantita\_venduta, t);
    end
end


valore\_finale = capitale + posizione * previsione;

fprintf('Valore␣finale␣dell␣investimento:␣\%.2f\n', valore\_finale);
profitto=valore\_finale-capitale

ritorno = (valore\_finale - capitale\_iniziale) / capitale\_iniziale;
profitto\_percentuale = (valore\_finale - capitale\_iniziale) / capitale\_iniziale * 100;

fprintf('Ritorno␣sull''investimento␣(ROI):\␣\%.2f\%\%\n', ritorno * 100);
fprintf('Profitto␣percentuale:␣\%.2f\%\\%\n', profitto\_percentuale);


tempi = 1:length(fore\_002);

figure;
plot(tempi, fore\_002, 'b-', 'LineWidth', 2);
```

```
hold on;


compra = fore\_002(fore\_002 < previsione\_intervallo\_inferiore);
vendi = fore\_002(fore\_002 > previsione\_intervallo\_superiore);
plot(find(fore\_002 < previsione\_intervallo\_inferiore), compra, 'ro', 'MarkerSize', 8);
plot(find(fore\_002 > previsione\_intervallo\_superiore), vendi, 'go', 'MarkerSize', 8);


title('Trading␣Strategy');
xlabel('Tempo');
ylabel('Previsioni');
legend('Previsioni', 'Acquisto', 'Vendita', 'Location', 'Best');

hold off;
%Q3
Data=readtable('DATI␣2021␣tesi.xlsx');
Dataq321=flipud(table2array(Data(1:475,19)));
prezziq321=flipud(table2array(Data(1:475,20)));
plot(Dataq321,prezziq321)
xlabel('prezziq321')
ylabel('Dataq321')


ret=diff(log(prezziq321))
plot(Dataq321(2:475),ret)
subplot(2,1,1);
autocorr(ret, 20);
title('Funzione␣di␣autocorrelazione');

subplot(2,1,2);
parcorr(ret, 20);
title('Funzione␣di␣autocorrelazione␣parziale');

plot(Dataq321(2:475),ret)
xlabel('Date');
ylabel('log␣return');
axis tight;
datetick('x', 'yyyy');

```

```
[~,pvalues\_ret]=adftest(ret,'lags',0:2)

%%
subplot(2,1,1)
autocorr(ret,22)
subplot(2,1,2)
parcorr(ret,22)
%%%%%%%%%


wind=250
T=size(ret,1)
fore=zeros(T-wind,1);
test\_set=ret(wind+1:end);


%%%%%%arima(0,0,0)%%%
fore\_000=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_000 = arima(0,0,0);
    EstMdl = estimate(Mdl\_000,ret(i:wind+i-1));
    fore\_000(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

%%%%%%%%%%arima(1,0,0)%%%%%%%%%%%%%%%%%%%%%
fore\_100=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_100 = arima(1,0,0);
    EstMdl = estimate(Mdl\_100,ret(i:wind+i-1));
    fore\_100(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc
 fore\_001=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_001 = arima(0,0,1);
```

```
    EstMdl = estimate(Mdl\_001,ret(i:wind+i-1));
    fore\_001(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc



%%%%%%%%%%arima(1,0,1)%%%%%%%%
fore\_101=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_101 = arima(1,0,1);
    EstMdl = estimate(Mdl\_101,ret(i:wind+i-1));
    fore\_101(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc


 %%%%%%%%%%%arima(2,0,0)%%%%%%%%%%%%%%%%%%%%%%
fore\_200=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_200 = arima(2,0,0);
    EstMdl = estimate(Mdl_200,ret(i:wind+i-1));
    fore\_200(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc



fore\_201=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_201= arima(2,0,1);
    EstMdl = estimate(Mdl\_201,ret(i:wind+i-1));
    fore\_201(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

```

```
fore\_002=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_002 = arima(0,0,2);
    EstMdl = estimate(Mdl\_002,ret(i:wind+i-1));
    fore\_002(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc



fore\_102=zeros(T-wind,1);
tic
for i=1:(T-wind);
    ;
    Md\l\_102 = arima(1,0,2);
    EstMdl = estimate(Mdl\_102,ret(i:wind+i-1));
    fore\_102(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

fore\_202=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_202 = arima(2,0,2);
    EstMdl = estimate(Mdl\_202,ret(i:wind+i-1));
    fore\_202(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc


plot(test\_set(1:end),"cyan")
hold on
plot(fore\_100,'r')
hold on
plot(fore,'black')
hold on
plot(fore\_101,'green')
hold on
```

```
plot(fore\_200,'red')
hold on
plot(fore\_201,'yellow')
hold on
plot(fore\_002,'magenta')
hold on
plot(fore\_102,'white')
hold on
plot(fore\_202,'blue')

if ~isempty(test\_set) \&\& all(isfinite(test\_set(2:end)))\ \&\& all(isfinite(fore))
    MSE = sum((test\_set(1:end)-fore\_000(1:end)).\^2) / numel(test\_set(1:end));
    MSE\_100 = sum((test\_set(1:end)-fore\_100).\^2) / numel(test\_set(1:end));
    MSE\_101 = sum((test\_set(1:end)-fore\_101).\^2) / numel(test\_set(1:end));
    MSE\_200 = sum((test\_set(1:end)-fore\_200).\^2) / numel(test\_set(1:end));
    MSE\_201 = sum((test\_set(1:end)-fore\_201).\^2) / numel(test\_set(1:end));
    MSE\_002 = sum((test\_set(1:end)-fore\_002).\^2) / numel(test\_set(1:end));
    MSE\_102 = sum((test\_set(1:end)-fore\_102).\^2) / numel(test\_set(1:end));
    MSE\_202 = sum((test\_set(1:end)-fore\_202).\^2) / numel(test\_set(1:end));
   MSE\_001 = sum((test\_set(1:end)-fore\_001).\^2) / numel(test\_set(1:end));
    M = [MSE, MSE\_100, MSE\_101, MSE\_200, MSE\_201, MSE\_002, MSE\_102, MSE\_202];
else
    disp('Warning:␣Test␣set␣contains␣NaN␣or␣Inf␣values,␣MSE␣calculation␣skipped.');
end

param0=fitlm(fore\_002,test\_set(1:end),1)
param2=fitlm(fore\_100,test\_set(1:end),1)
param3=fitlm(fore\_101,test\_set(1:end),1)
param4=fitlm(fore\_102,test\_set(1:end),1)
param6=fitlm(fore\_200,test\_set(1:end),1)
param7=fitlm(fore\_201,test\_set(1:end),1)
param8=fitlm(fore\_000,test\_set(1:end),1)
param9=fitlm(fore\_202,test\_set(1:end),1)
param10=fitlm(fore\_001,test\_set(1:end),1)


%%%DM sui migliori
DM\_FINAL=dmtest(fore\_001-test\_set,fore\_002-test\_set,1)
```

```
capitale\_iniziale = 1000000;
capitale = capitale\_iniziale;
investimento\_fisso = 20000;
posizione = 0;


confidenza = 0.9;


for t = 1:length(for\e\_002)

    previsione = fore\_002(t);
    errore\_std = sqrt(EstMdl.Variance);

    previsione\_intervallo\_inferiore = previsione - norminv((1 - confidenza) / 2) * errore\_std;
    previsione\_intervallo\_superiore = previsione + norminv((1 + confidenza) / 2) * errore\_std;

    if previsione < previsione\_intervallo\_inferiore\ \&\& posizione <= 0

        quantita\_acquistata = investimento\_fisso / previsione;
        capitale = capitale - investimento\_fisso;
        posizione = posizione + quantita\_acquistata;
        fprintf('Compra␣\%f␣unit␣di␣asset␣a␣t=\%d\n', quantita\_acquistata, t);
    elseif previsione > previsione\_intervallo\_superiore \&\& posizione >= 0

        quantita\_venduta = min(posizione, investimento\_fisso / previsione);
        capitale = capitale + quantita\_venduta * previsione;
        posizione = posizione - quantita\_venduta;
        fprintf('Vendi␣\%f␣unit␣di␣asset␣a␣t=\%d\n', quantita\_venduta, t);
    end
end


valore\_finale = capitale + posizione * previsione;

fprintf('Valore␣finale␣dell␣investimento:␣\%.2f\n', valore\_finale);
profitto=valore\_finale-capitale
```

```


ritorno = (valore\_finale - capitale\_iniziale) / capitale\_iniziale;
profitto\_percentuale = (valore\_finale - capitale\_iniziale) / capitale\_iniziale * 100;

fprintf('Ritorno␣sull''investimento␣(ROI):␣\%.2f\%\%\n', ritorno * 100);
fprintf('Profitto␣percentuale:␣\%.2f\%\%\n', profitto\_percentuale);

tempi = 1:length(fore\_002);


figure;
plot(tempi, fore\_002, 'b-', 'LineWidth', 2);
hold on;


compra = fore\_002(fore\_002 < previsione\_intervallo\_inferiore);
vendi = fore\_002(fore\_002 > previsione\_intervallo\_superiore);
plot(find(fore\_002 < previsione\_intervallo\_inferiore), compra, 'ro', 'MarkerSize', 8);
plot(find(fore\_002 > previsione\_intervallo\_superiore), vendi, 'go', 'MarkerSize', 8);


title('Trading␣Strategy');
xlabel('Tempo');
ylabel('Previsioni');
legend('Previsioni', 'Acquisto', 'Vendita', 'Location', 'Best');

hold off;
%%%%%
%Q4
Data=readtable('DATI␣2021␣tesi.xlsx');
Dataq421=flipud(table2array(Data(1:502,22)));
prezziq421=flipud(table2array(Data(1:502,23)));
plot(Dataq421,prezziq421)
xlabel('prezziq421')
ylabel('Dataq421')


ret=diff(log(prezziq421))
```

```
plot(Dataq421(2:502),ret)
subplot(2,1,1);
autocorr(ret, 20);
title('Funzione di autocorrelazione');

subplot(2,1,2);
parcorr(ret, 20);
title('Funzione di autocorrelazione parziale');

plot(Dataq421(2:502),ret)
xlabel('Date');
ylabel('log return');
axis tight;
datetick('x', 'yyyy');

[~,pvalues\_ret]=adftest(ret,'lags',0:2)

%%
subplot(2,1,1)
autocorr(ret,22)
subplot(2,1,2)
parcorr(ret,22)
%%%%%%%%%


wind=250
T=size(ret,1)
fore=zeros(T-wind,1);
test_set=ret(wind+1:end);


%%%%%%arima(0,0,0)%%%%
fore\_000=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_000 = arima(0,0,0);
    EstMdl = estimate(Mdl\_000,ret(i:wind+i-1));
    fore\_000(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc
```

```
%%%%%%%%%%arima(1,0,0)%%%%%%%%%%%%%%%%%%%%%
fore\_100=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_100 = arima(1,0,0);
    EstMdl = estimate(Mdl\_100,ret(i:wind+i-1));
    fore\_100(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc



%%%%%%%%%%arima(1,0,1)%%%%%%%%%%%%%%%%%%%%%
fore\_101=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_101 = arima(1,0,1);
    EstMdl = estimate(Mdl\_101,ret(i:wind+i-1));
    fore\_101(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc


 %%%%%%%%%%arima(2,0,0)%%%%%%%%%%%%%%%%%%%%%
fore\_200=zeros(T-wind,1);
tic
for i=1:(T-wind);
;
    Mdl\_200 = arima(2,0,0);
    EstMdl = estimate(Mdl\_200,ret(i:wind+i-1));
    fore\_200(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc



fore\_201=zeros(T-wind,1);
tic
for i=1:(T-wind);

```

```
    Mdl\_201= arima(2,0,1);
    EstMdl = estimate(Mdl\_201,ret(i:wind+i-1));
    fore\_201(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc


fore\_002=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_002 = arima(0,0,2);
    EstMdl = estimate(Mdl\_002,ret(i:wind+i-1));
    fore\_002(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc



fore\_102=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_102 = arima(1,0,2);
    EstMdl = estimate(Mdl\_102,ret(i:wind+i-1));
    fore\_102(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

fore\_202=zeros(T-wind,1);
tic
for i=1:(T-wind);

    Mdl\_202 = arima(2,0,2);
    EstMdl = estimate(Mdl\_202,ret(i:wind+i-1));
    fore\_202(i,1) = forecast(EstMdl,1,'Y0',ret(i:wind+i-1));
end
 toc

plot(test\_set(1:end),"cyan")
hold on
plot(fore\_100,'r')
```

```
hold on
plot(fore,'black')
hold on
plot(fore\_101,'green')
hold on
plot(fore\_200,'red')
hold on
plot(fore\_201,'yellow')
hold on
plot(fore\_002,'magenta')
hold on
plot(fore\_102,'white')
hold on
plot(fore\_202,'blue')

if ~isempty(tes\t\_set)\ \&\& all(isfinite(test\_set(2:end)))\\ \&\& all(isfinite(
    fore))
    MSE = sum((test\_set(1:end)-fore\_000(1:end)).\^2) / numel(test\_set(1:end));
    MSE\_100 = sum((test\_set(1:end)-fore\_100).\^2) / numel(test\_set(1:end));
    MSE\_101 = sum((test\_set(1:end)-fore\_101).\^2) / numel(test\_set(1:end));
    MSE\_200 = sum((test\_set(1:end)-fore\_200).\^2) / numel(test\_set(1:end));
    MSE\_201 = sum((test\_set(1:end)-fore\_201).\^2) / numel(test\_set(1:end));
    MSE\_002 = sum((test\_set(1:end)-fore\_002).\^2) / numel(test\_set(1:end));
    MSE\_102 = sum((test\_set(1:end)-fore\_102).\^2) / numel(test\_set(1:end));
    MSE\_202 = sum((test\_set(1:end)-fore\_202).\^2) / numel(test\_set(1:end));
    M = [MSE, MSE\_100, MSE\_101, MSE\_200, MSE\_201, MSE\_002, MSE\_102, MSE\_202];
else
    disp('Warning: Test set contains NaN or Inf values, MSE calculation skipped.');
end


param0=fitlm(fore\_002,test\_set(1:end),1)
param2=fitlm(fore\_100,test\_set(1:end),1)
param3=fitlm(fore\_101,test\_set(1:end),1)
param4=fitlm(fore\_102,test\_set(1:end),1)
param6=fitlm(fore\_200,test\_set(1:end),1)
param7=fitlm(fore\_201,test\_set(1:end),1)
param8=fitlm(fore\_000,test\_set(1:end),1)
param9=fitlm(fore\_202,test\_set(1:end),1)
param10=fitlm(fore\_001,test\_set(1:end),1)
```

```
%%%DM sui migliori
DM\_FINAL=dmtest(fore\_200-test\_set,fore\_002-test\_set,1)
diff=MSE\_200-MSE\_002

capitale\_iniziale = 1000000;
capitale = capitale\_iniziale;
investimento\_fisso = 20000;
posizione = 0;

confidenza = 0.7;

for t = 1:length(fore\_200)

    previsione = fore\_200(t);
    errore\_std = sqrt(EstMdl.Variance);

    previsione\_intervallo\_inferiore = previsione - norminv((1 - confidenza) / 2) * errore\_std;
    previsione\_intervallo\_superiore = previsione + norminv((1 + confidenza) / 2) * errore\_std;

    if previsione < previsione\_intervallo\_inferiore \&\& posizione <= 0

        quantita\_acquistata = investimento\_fisso / previsione;
        capitale = capitale - investimento\_fisso;
        posizione = posizione + quantita\_acquistata;
        fprintf('Compra␣\%f␣unit␣di␣asset␣a␣t=\%d\n', quantita\_acquistata, t);
    elseif previsione > previsione\_intervallo\_superiore\ \&\& posizione >= 0

        quantita\_venduta = min(posizione, investimento\_fisso / previsione);
        capitale = capitale + quantita\_venduta * previsione;
        posizione = posizione - quantita\_venduta;
        fprintf('Vendi␣\%f␣unit␣di␣asset␣a␣t=\%d\n', quantita\_venduta, t);
    end
```

```
end


valore\_finale = capitale + posizione * previsione;

fprintf('Valore␣finale␣dell␣investimento:␣\%.2f\n', valore\_finale);
profitto=valore\_finale-capitale



ritorno = (valore\_finale - capitale\_iniziale) / capitale\_iniziale;
profitto\_percentuale = (valore\_finale - capitale\_iniziale) / capitale\_iniziale *
     100;

fprintf('Ritorno␣sull''investimento␣(ROI):␣\%.2f\%\%\n', ritorno * 100);
fprintf('Profitto␣percentuale:␣\%.2f\%\%\n', profitto\_percentuale);

tempi = 1:length(fore\_200);


figure;
plot(tempi, fore\_200, 'b-', 'LineWidth', 2);
hold on;
```

Listing 5.1: Matlab codex future