School of Social Sciences

Supply Chain Management

(SCM)

Postgraduate Dissertation

“Analyzing the Natural Gas Consumption in Spain’’

ALTIJON RUMANI

Supervisor: Antonios Ntemos

Patras, Greece, July 2024

''Analyzing the Natural Gas Consumption in Spain''

ALTIJON RUMANI

Supervisor: Antonios Ntemos

Patras, Greece, July 2024

This endeavour would not have been possible without the valuable guidance of my supervisor Dr. Antonios Ntemos who generously provided his knowledge and feedback. Additionally, Thanks, should also go to my family for their support during this process, especially my brother.

Abstract

Natural gas (NG) is a fuel energy source that is used for a variety of purposes (heating, power generation etc). The objective of this thesis is to perform a comprehensive analysis of the Spanish consumption of Natural Gas intended for heating, electricity generation and LNG use. We aim to investigate the statistical properties of the sample time series and unveil patterns and trends in NG consumption. Consumption forecasts offer valuable information regarding the expected demand for natural gas over different periods of time. The knowledge of NG demand is important for the efficient allocation of resources like production capacity, storage and transportation infrastructure. For our analysis, we will use a sample of Spanish NG consumption data spanning the years 2004-2024. We will use various statistical/forecasting models and regression models, to address the following research questions:

1) What are the statistical properties of the NG consumption time series? Which characteristics has a major contribution on how gas consumption develops over time?

2) To what extent are consumption patterns predictable? Which model is the most accurate forecasting tool for our data?

3) Have major event (such as the war in Ukraine and the Covid crisis) imposed a structural break in the data-generating process?

Keywords:

Natural Gas, forecasting models, regression analysis, time series

Αναλύοντας την κατανάλωση φυσικού αερίου στην Ισπανία

ΑΛΤΙΓΙΟΝ ΡΟΥΜΑΝΙ

Περίληψη

Το φυσικό αέριο (Natural Gas) είναι μια πηγή ενέργειας καυσίμου που χρησιμοποιείται για διάφορους σκοπούς (θέρμανση, παραγωγή ηλεκτρικής ενέργειας κ.λπ.). Ο στόχος αυτής της διπλωματικής εργασίας είναι να πραγματοποιήσει μια ολοκληρωμένη ανάλυση της ισπανικής κατανάλωσης φυσικού αερίου που προορίζεται για θέρμανση, παραγωγή ηλεκτρικής ενέργειας καθώς και για LNG. Στόχος μας είναι να διερευνήσουμε τις στατιστικές ιδιότητες των δειγματοληπτικών χρονοσειρών και να αποκαλύψουμε μοτίβα και τάσεις στην κατανάλωση φυσικού αερίου. Οι προβλέψεις κατανάλωσης προσφέρουν πολύτιμες πληροφορίες σχετικά με την αναμενόμενη ζήτηση για φυσικό αέριο σε διαφορετικές χρονικές περιόδους. Η γνώση της ζήτησης φυσικού αερίου είναι σημαντική για την αποτελεσματική κατανομή πόρων όπως η παραγωγική ικανότητα, η αποθήκευση και η υποδομή μεταφοράς. Για την ανάλυσή μας, θα χρησιμοποιήσουμε ένα δείγμα ισπανικών δεδομένων κατανάλωσης φυσικού αερίου που καλύπτουν τα έτη 2004-2024. Θα χρησιμοποιήσουμε διάφορα στατιστικά/προγνωστικά μοντέλα και μοντέλα παλινδρόμησης, για να αντιμετωπίσουμε τα ακόλουθα ερευνητικά ερωτήματα:

1) Ποιες είναι οι στατιστικές ιδιότητες της χρονοσειράς κατανάλωσης φυσικού αερίου; Ποια χαρακτηριστικά έχουν σημαντική συμβολή στον τρόπο με τον οποίο αναπτύσσεται η κατανάλωση αερίου με την πάροδο του χρόνου;

2) Σε ποιο βαθμό είναι προβλέψιμα τα πρότυπα κατανάλωσης; Ποιο μοντέλο είναι το πιο ακριβές εργαλείο πρόβλεψης για τα δεδομένα μας;

3) Κάποιο μεγάλο γεγονός (όπως ο πόλεμος στην Ουκρανία και η κρίση του Covid) επέβαλε μια δομική ρήξη στη διαδικασία δημιουργίας δεδομένων;

Λέξεις-Κλειδιά:

Φυσικό Αέριο, μοντέλα πρόβλεψης, ανάλυση παλινδρόμησης, χρονοσειρές

Contents

List of Figures

List of Tables

List of Abbreviations& Acronyms

AR model: Autoregressive model
MAD: Mean Absolute Deviation
MAPE: Mean Absolute Percentage Error
MSE: Mean Squared Error
SMA: Simple Moving Average
NG consumption : Natural Gas consumption

1)INTRODUCTION

Energy availability is one of the most crucial elements in the development of the economy and society since many types of energy are needed for heating, manufacturing, transportation, and the production of electricity. It is widely known that natural gas is one of the energy sources that has permeated practically every aspect of contemporary life and is necessary for both home and commercial applications. The Russia-Ukraine military conflict has caused major supply chain disruptions in the global energy markets for natural gas and crude oil over the past few years, which has led to significant variations in the natural gas markets and sparked interest in examining the security of the natural gas grids.

Furthermore, natural gas has been an essential part of Spain's energy mix in recent decades, greatly assisting with the nation's industrial, power generating, and heating demands. Natural gas is becoming more and more popular due to its advantages over other fossil fuels, such as its lower emissions and higher efficiency, which are in line with international efforts to slow down climate change and lessen its negative effects on the environment. It is becoming more and more important to comprehend and forecast natural gas use as Spain develops its energy infrastructure and policy. For a number of reasons, precise estimates of natural gas use are crucial. They make it possible for energy companies and governments to plan and construct energy infrastructure with knowledge, ensuring that supply and demand are met effectively. Furthermore, accurate consumption forecasts aid in the efficient use of resources throughout the manufacturing, distribution, and transportation phases, improving economic stability and energy security in the process.

With an emphasis on the variables influencing demand and the techniques employed to forecast future consumption trends, this thesis seeks to present a thorough examination of natural gas consumption in Spain. Through the analysis of past consumption data and the application of several statistical and forecasting models, this study aims to identify trends and patterns that may provide valuable insights for improving energy management strategies. The study period, which runs from 2004 to 2024, includes important occurrences and changes that have affected Spain's use of natural gas. Some of them are changes in the global economy, technological development, world events like the COVID-19 epidemic and war in Ukraine and geopolitical unrest. It is essential to comprehend how these factors have influenced consumption patterns in order to create reliable forecasting models that are flexible enough to adjust to changing circumstances.

We will use forecasting models (SMA,EWMA,regression models) that can be used by many experts in logistics market and policy makers in order to address the behavior and patterns of natural gas consumption. This thesis will investigate the dynamics of natural gas consumption by addressing important problems concerning the statistical characteristics of consumption data, the predictability of consumption patterns, and the influence of significant events on the process of data generation. The results will add to the body of knowledge about energy consumption forecasting and give stakeholders in Spain's energy sector important information.

The structure of the thesis is :

Chapter 1=>Introduction

Chapter 2=>Literature review that provides general information about the natural gas(history, transportation)

Chapter 3=>Methodology that refers to the models (forecasting models ,time series analysis) and the data that we will use in our thesis

Chapter 4 =>Empirical study that refers to the use of the different models with our data , accuracy measure ,statistical results

Chapter 5 =>Concluding remarks

2 LITERATURE REVIEW

2.1 HISTORY OF NATURAL GAS

Natural gas is a fossil fuel energy source. Natural gas contains many different compounds. The largest component of natural gas is methane, a compound with one carbon atom and four hydrogen atoms (CH4). Natural gas also contains smaller amounts of natural gas liquids (NGLs, which are also hydrocarbon gas liquids), and nonhydrocarbon gases, such as carbon dioxide and water vapor. We use natural gas as a fuel and to make materials and chemicals.(U.S.Energy Information Administration).

In ancient times, natural gas was observed and used in small quantities. Literature references to flaming bushes, or burning springs imply the occasional usage of natural gas for warmth. There were burning gas springs in Greece and Rome, and in ancient China, brine water was heated by gas springs to extract salt. There are additional documented reports of flaming springs in France, Italy, and Russia. Theologian St. Augustine and philosopher Plutarch both mentioned lights that might have been created by burning natural gas.

Natural gas wasn’t a major source of energy until the 19^{th} century. For example natural gas did not become a major energy source in the United States until 1859, when Colonel Edwin Drake found oil in Titusville, Pennsylvania. Drake discovered natural gas in addition to oil, despite the fact that the two are frequently found in the same geologic levels. The natural gas found in the eastern part of Pennsylvania was sold to local consumers.(Christopher J. Castaneda,2004).There weren't many long-term, successful attempts to use natural gas for commercial or industrial purposes until Drake's discoveries. By the middle of the 19th century, the fuel was limited to use in companies and cities that were close to natural gas wells. Its utility was restricted by the challenges of containing a natural gas spring, storing the gas, and moving it over large distances. Important natural gas discoveries, for instance, like the high-volume well William Tomkins found in 1841 on the Canadaway Creek close to Washington's flaming spring, attracted attention but little business interest. On the other hand, provided coal, or the feedstock, was easily accessible, manufactured coal gas plants could be constructed and run anywhere. As a result, in the 19th century, the business for manufactured coal gas expanded far faster than that of natural gas. Many towns and cities had local distribution systems and built gas plants by the middle of the 19th century, which supplied some coal gas for lighting homes and businesses.

Lack of supply was not the main barrier to the natural gas industry's growth in the middle of the 19th century .It was poor pipeline facilities and technology. Not only could hollow wood pipelines leak and disintegrate, but cast and wrought iron lines had serious inherent flaws as well. During the years 1872–1890, wrought iron lines were usually smaller than 8 inches in diameter, and couplings fastened with screws were used to join the pipe segments. Gas leaks were common. During this time, the majority of gas transported in pipes flowed under the well's natural pressure without the need for extra compression.

2.2 Transportation of natural gas

Ensuring the efficient and reliable distribution of energy from the source to the client is the primary goal of an energy distribution system. A distribution system's operation needs to be resilient enough to handle variations in transportation capacity, which are mostly caused by interruptions in the supply chain. The reliability of the gas supply relative to the needs of the consumer is a matter of risk evaluation and mitigation. All energy supply systems operate at a certain intrinsic level of risk, but it is important to assess whether the level and type of risks are acceptable in the operational context. Furthermore, the reliability of supply in the energy and gas sector is generally more important (e.g., for political and economic reasons) relative to other industries due to the lack of alternative options in the short term. (Sidney Pereira dos Santos,José Eugenio Leal,Fabrício Oliveira,2011).
It is essential to remember that natural gas is explosive; as a result, an advanced network of pipelines makes composed the transportation system in order to accomplish the desired outcome safely. We can separate natural Gas system into 3 stages:

1. 1)Processing
2. 2)Transportation
3. 3)Storage

Natural gas is located, transported to the surface, and prepared for transportation during the production process. If the delivery can be executed with the help of international pipelines, then some components need to be removed from the natural gas before it can be safely delivered to high-pressure, long-distance pipelines.(MET 2020). Hydrocarbon gas liquids (HGL) are the initial element that needs to be eliminated, but other elements including water, oil, and substances like sulfur, helium, nitrogen, hydrogen sulfide, and carbon dioxide are also frequently affected. However, what happens if pipeline transportation is not an option? Then natural gas is usually transported in its liquefied state. For instance, in areas where the requirements for road and water transportation have been established but the locations are too far from the gas extraction sites and there isn't a connecting pipeline. In this case, cargo ships are used to distribute them rather than pipelines because LNG can be carried at sea quite effectively. In typical circumstances, one cubic meter of liquefied LNG yields roughly 600 cubic meters of natural gas after regasification. LNG stands for "liquefied natural gas". It has almost the same composition as natural gas in the traditional sense – only in refrigerated form. This means that it has a temperature of -162 ° C and its density is lower than water's. Several procedures must be followed for natural gas to be transported from production locations to consumers.

- Gathering systems transport raw natural gas from the wellhead to a natural gas processing facility or to an interface with a bigger mainline pipeline. They are mainly composed of small-diameter, low-pressure pipelines.
- Before natural gas is released into a main transmission system, natural gas processing facilities remove the natural gas from water, nonhydrocarbon gases, and hydrocarbon gas liquids.
- Wide-diameter, high-pressure interstate transmission pipelines that cross state boundaries and intrastate transmission pipelines that operate within state boundaries

transport natural gas from the producing and processing areas to storage facilities and distribution centers. Compressor stations (or pumping stations) on the pipeline network keep the natural gas flowing forward through the pipeline system.

- Local distribution companies deliver natural gas to consumers through small-diameter, lower pressure service lines.(Natural gas pipelines ,U.S Energy Information Administration,2022)

2.3)NATURAL GAS IN SPAIN

When Barcelona City Council awarded Charles Lebon the first contract for gas public lighting in 1841, the history of gas use in Spain officially began. A year later, as a consequence of this concession, the first gas plant in Spain was built in Barceloneta. In 1843, Lebon founded the Sociedad Catalana Para el Alumbrado por Gas (SCAG) with the Gil y Serra brothers and other shareholders. Other businesses started to come and compete in the market starting in 1853.

Over the course of the 19th century, as gas became increasingly widely used for public light, the idea emerged to construct a gas plant in the largest municipality and distribute it to the others in nearby areas. And so, piped gas distribution began with the aim of saving on costs and investment. In 1894 gas faced competition in the lighting industry with the introduction of electricity, a war it eventually lost. Gas had to find new applications as a result, mostly in the kitchen. It was heavily promoted via advertising. Spain executed the Development and Stabilization Plans between the late 1950s and the early 1970s, and the peseta began trading on international currency exchange markets. The gas industry's outdated facilities prompted a modernization initiative, the most significant of which was the 1963 transition from producing gas from coal to naphtha .More modern chemical plants were started after this change, increasing the plants' potential for productivity .Gas heating has been heavily promoted as a result of the need to discover a new application for gas in order to find a way to accommodate this increased production.

Supplying natural gas direct without the need for gas plants began in Spain in 1969.

A very important role in this development was played by Catalana de Gas y Electricidad which, with other stakeholders, set up Gas Natural S.A. to import natural gas from Libya, constructing a regasification plant in the Port of Barcelona and the first Spanish LNG tanker: “El Laietà”.(Nedgia Grupo Naturgy). In 1985 a Memorandum of understanding between the Ministry of Industry and Energy (MINER) and the main natural gas distributors and suppliers, led to a long-term expansion strategy involving increased investment and an increase in distributors and consumers. This helped the Spanish industry to catch up with its historical gap behind the rest of Europe. After selling its final shares in conventional electricity, Catalana de Gas y Electricidad changed its name to Catalana de Gas, S.A. in 1987. The last gas plant in the city, located near Sant Martí de Provençals, closed in 1990, marking the end of Barcelona's 1985–1990 transition from city gas to natural gas.

Thanks to the new situation in the Spanish gas sector (excellent possibilities for growth and development), two important shareholders joined the others with a stake in Catalana de Gas:La Caixa and Repsol. Due to favorable circumstances surrounding the Spanish gas industry, which include plenty of possibilities for expansion and improvement, La Caixa and Repsol have become significant investors in Catalana de Gas. A new commercial project including the integration of Spain's gas distribution industry (via the merging of Catalana de Gas, Gas Madrid, and Repsol Butano's piped gas shares) was launched with the backing of these two shareholders. Consequently, Gas Natural SDG, S.A. was established as an entirely new company in 1992. The corporation started a phase of international expansion in Latin America that same year, starting in Argentina and moving on to Brazil, Colombia, and Mexico in the years that followed.

Consequently, Gas Natural SDG, S.A. was established as an entirely new company in 1992. The corporation started a phase of international expansion in Latin America that same year, starting in Argentina and moving on to Brazil, Colombia, and Mexico in the years that followed. The primary source of gas supply between 1985 and 1993 was LNG (liquefied natural gas). The Spanish gas system's first international connection, located at Larrau, France, began operating in 1993. The Maghreb-Europe pipeline officially launched in 1996. This pipeline originates in Algeria, travels through Morocco (a transit nation), and ends in Tarifa, Spain, where it connects to the mainland. With the introduction of the Hydrocarbons Act in 1998, Spain's gas industry started to deregulate. This act made it necessary to divide the responsibilities of System Technical Managers and distribution networks from commercialization and distribution networks.

In 2016, Spain—the seventh-largest natural gas consumer in Europe—burned roughly one trillion cubic feet (Tcf). Approximately 19% of the energy consumed in Spain in 2016 came from natural gas, according to the BP Statistical Review of World Energy 2017. Spain had one of the fastest-growing natural gas markets in Europe before the 2008 financial crisis. Spain saw an approximate 142% increase in natural gas consumption, mostly from the power sector, between 2000 and 2008. Since then, the global financial crisis and competition from alternative fuel sources have caused consumption to fall for several years. Between 2014 and 2016, Spain's natural gas consumption increased by 6% as the country's economy started to revive. Due to its low natural gas production and limited reserves, Spain must import practically all of its natural gas in order to satisfy its needs. Currently the Maghreb-Europe gas pipeline and the Medgaz pipeline are the two undersea pipelines that carry natural gas from Algeria to Spain. In 2016, 42% of all natural gas imports came from pipelines coming from Algeria. It has also to be mentioned that there are very few natural gas pipeline links between the Iberian Peninsula (Spain and Portugal) and the rest of Europe (via a pipeline through France). Liquefied natural gas (LNG) accounted for nearly half (466 Bcf) of Spain's natural gas imports in 2016. Algeria, Qatar, and Nigeria were the top three LNG suppliers to Spain. Approximately 52% of Spain's total imports of natural gas (pipeline and LNG) came from Algeria in 2016, according to the BP Statistical Review of World Energy 2017. El Musel import terminal in Gijon is one of Spain's seven LNG import terminals; it is currently seeking government approval to start operations. At present, Spain possesses the greatest capacity for regasification in Europe, with an approximate annual capacity of 2.4 trillion cubic feet (Tcf/y). The past few years have seen a decrease in the utilization rates of regasification terminals due to capacity increases and a decrease in LNG imports. A total of 134 million cubic feet (MMcf) of LNG storage capacity is available at the ports. Spanish authorities are examining the proposed LNG terminals on the Canary Islands, Tenerife and Gran Canaria, each of which is anticipated to provide 45 billion cubic feet per year of regasification capacity by 2021 (Nedgia Grupo Naturgy).

2.4 Forecasting in Natural Gas sector

Accurate forecasting of natural gas consumption (NGC) plays an important role in energy supply, energy trading, economic effects and environmental sustainability. NG forecasts can be used to adjust production and supply plans to improve gas efficiency and reduce carbon emissions and supply chain waste. (Ning Tian, Bilin Shao, Genqing Bian, Huibin Zeng, Xiaojun Li, Wei Zhao,2024).By applying advanced forecasting techniques, including statistical models, machine learning algorithms, and hybrid model structures, the natural gas industry may more effectively predict variations in demand, manage variations in peak and valley stages, and optimize overall operational efficiency.

One significant benefit of effective forecasting is the ability to optimize energy trading and supply chain management. Energy traders mostly rely on precise forecasts of supply and demand In order to make informed decisions on when and where it is best to purchase or sell natural gas. For example, a company might avoid the volatility of spot market prices by securing contracts at favorable prices in advance if it can foresee an increase in demand during the winter months. This helps businesses control expenses and optimize revenues in addition to guaranteeing constant supplies for customers. Moreover, accurate forecasting allows for strategic storage planning, ensuring sufficient reserves during high-demand periods without unnecessary excess that can lead to storage inefficiencies.

Forecasting also helps to advance environmental sustainability by supporting the transition to cleaner energy sources. Accurate natural gas forecasting is essential for the grid integration of renewable energy sources since natural gas is sometimes regarded as a "bridge energy" to zero-carbon alternatives. Energy firms may optimize their use of natural gas alongside renewables and decrease their reliance on fuels with higher carbon emissions by forecasting demand patterns. This measure is in accordance with international efforts to mitigate climate change and lower greenhouse gas emissions. Accurate forecasts can help governments and politicians create energy policies that reduce environmental impact by encouraging the use of natural gas in a way that minimizes its impact.

In conclusion ,forecasting is crucial to the natural gas market because it helps with the shift to cleaner energy sources, optimizes supply chain management, and offers valuable insight on patterns of demand. Advanced forecasting methods and hybrid model structures allow energy firms to lower costs, increase operational effectiveness, and promote environmental sustainability. Accurate forecasting will continue to be essential for effective decision-making in the energy sector as it develops, and guarantees a consistent and sustainable supply of energy in the future.

3)METHODOLOGY

3.1)Research data

In our research we will use the monthly consumption of natural gas in Spain for the liberalized market. We will use the monthly data of natural gas consumption in Spain per month from January 2004 to January 2024.These data are separated in 3 categories:

1)Conventional Natural gas consumption

20Natural Gas consumption for electricity generation

3)Liquified Natural Gas for direct consumption.

We will retrieve these data from the official site of CORES(Public Law Corporation under the aegis of the Ministry for the Ecological Transition and the Demographic Challenge)(https://www.cores.es/en/cores/quienes-somos, accessed at 21 March 2024). Every member of the hydrocarbons industry in Spain are required by law to submit periodic reports to CORES in order to ensure the security of the supply of natural gas, LPG, and petroleum products. This requirement enables the industry to produce high-quality publications and data. Additionally, CORES is responsible for several of the energy-related data in the National Statistical Plan together with the Ministry for the Ecological Transition and the Demographic Challenge. Additionally, it takes part in the preparation of reports on hydrocarbons that are submitted to various international organizations (such as AIE and EUROSTAT).

Using the data we mentioned previously , we will first estimate the mean and standard deviation, two fundamental statistical measures, to see how distributed the data are and how the annual mean consumption of natural gas changes. Additionally, in order to better understand how the data behave, we will plot the data in graphs. These will allow us to see any trends in the data, such as upward trends when the data are rising, seasonality (when our data rise or fall at particular times of the year), or no patterns at all. We will apply qualitive techniques in order to better understand the data, the outcomes of our calculation, and whether there is a logical connection between them. Additionally, we will use a variety of forecasting methods to see which one best fits the data and measure each one's accuracy level.

3.2)Descriptive statistics

Descriptive statistics help us to analyze, characterize, arrange, or summarize data ,for instance we can detect patterns that may appear in the data, Descriptive statistics, however, do not enable us to draw conclusions about hypotheses we may have had or to draw conclusions from data beyond our analysis. They are just a method of describing our data. Statistical analysis describes the characteristics of the responses in quantitative research, such as the average of one variable or the relationship between two variables, after collecting data. The 3 main categories of descriptive statistics are:

1.The distribution that measures the frequency, the central tendency, which calculates the mean, median, and mode.

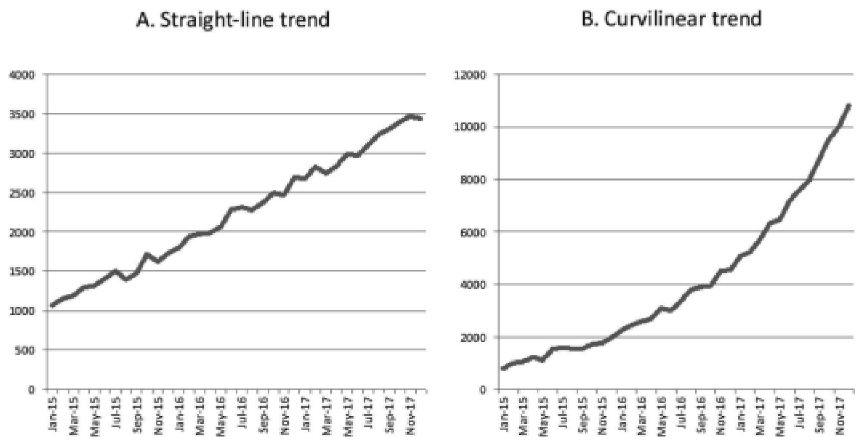
2.The distribution that measures the frequency.

3. The variability, which includes the variance, standard deviation, and range and which quantifies the range of values.
Descriptive statistics were utilized in our study to help us understand the research in a brief overview by allowing us to use the data for preliminary interpretation.

3.3)Time series Analysis

Time-series techniques are a category of algorithms that are designed to identify patterns in historical demand that repeat with time. There are three components of historical demand that these algorithms try to identify and predict: trend, seasonality, and noise.

1.Trend. A trend is a continuing pattern of demand increase or decrease. A trend can be either a straight line (see Figure 1A) or a curve (see Figure 1B).



:
Figure 1 **: Trends in historical Demand**

2. Seasonality. Seasonality is a repeating pattern of demand increases or decreases. Normally, we think of seasonality as occurring within a single year, and cyclicality as occurring over a longer period than a single year. A seasonal demand pattern is illustrated in Figure 2.

3. Noise. Noise represents random demand fluctuation. Noise is that part of the demand history which the other time-series components (trend and seasonality) cannot identify. Figure 3 illustrates a demand pattern that contains no discernable trend or seasonality but is simply noise. Most demand patterns will contain some degree of random fluctuation—the less random fluctuation (i.e., the lower the noise level), the more "forecastable" will be the product or service.

One complicating factor, of course, is that it is often the case that all three of these components can be present in a stream of historical demand. The overall trend might be going up, while at the same time there is repeating seasonal variation, while at the same time there is random variation in the form of noise. The challenge, then, for forecasters who are examining these types of data patterns is to find a time-series algorithm that will do the best possible job of identifying the patterns in the data and projecting those patterns into future time periods. Let's begin with very simple techniques and then move to more sophisticated ones.(Moon, M. A.2018)

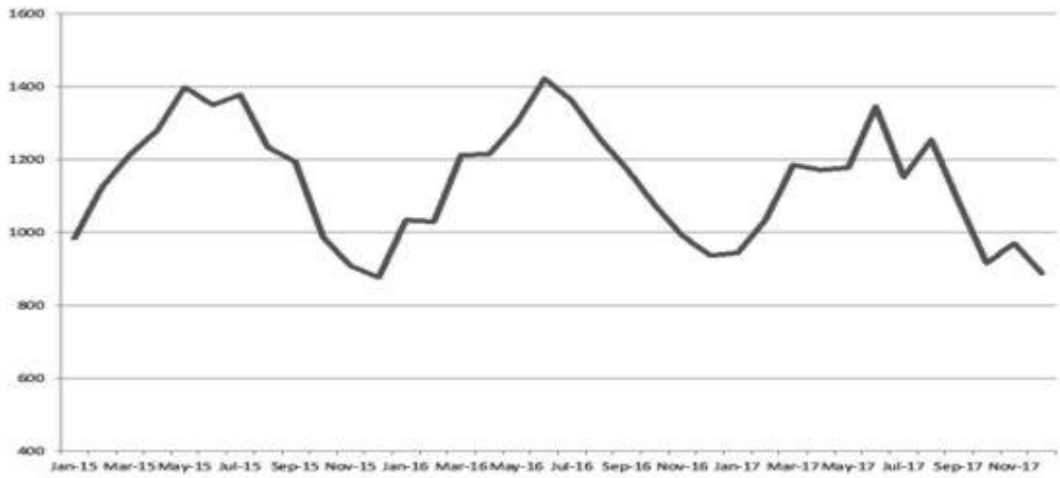


Figure 2: Seasonality in Historical Demand

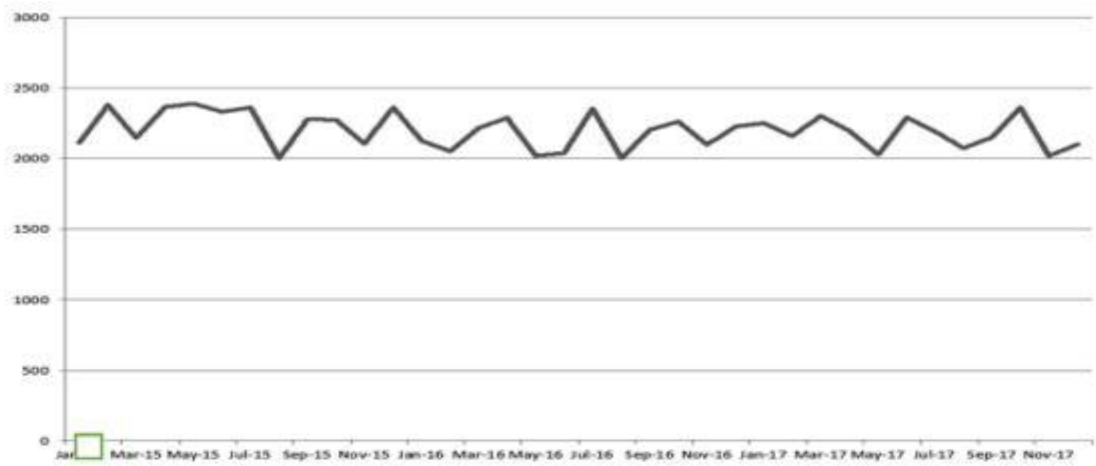


Figure 3**:**Noise in Historical Demand

3.4 SIMPLE MOVING AVERAGE

The simplest form of time-series analysis is a simple average. Simple moving average (SMA) models forecast the value of the time series in a future time period by averaging the most recenFhistorical observations. An average can be expressed arithmetically in the following formula:

$$\boldsymbol{Forecast_{t+1} = Average\ Demand = \sum_{t=1}^{N} D_t/N}$$

Where:

D = Demand

N = Number of periods of demand data

In other words, if using a simple average as a forecasting technique, then next month's forecast, and every future month's forecast, is the average level of demand from all previous months. There is one demand pattern where a simple average is the best forecasting technique that can be used, and that is a pattern of random data, with no detectable pattern of trend or seasonality.

3.5 EXPOTENTIAL WEIGHTED MOVING AVERAGE

The Exponentially Weighted Moving Average (EWMA) is a quantitative or statistical measure used to model or describe a time series. The EWMA is widely used in finance, the main applications being technical analysis and volatility modeling.

The moving average is designed as such that older observations are given lower weights. The weights fall exponentially as the data point gets older – hence the name exponentially weighted.

The only decision a user of the EWMA must make is the parameter alpha. The parameter decides how important the current observation is in the calculation of the EWMA. The higher the value of alpha, the more closely the EWMA tracks the original time series.The EWMA formula is :

$$EWMA = a * x(t) + (1 - a) * EWMA(t - 1)$$

Where,

EWMA(t): moving average at time t

a = degree of mixing parameter value between 0 and 1

x(t) = value of signal x at time x

The main difference between exponential moving averages (EWMA) and simple moving averages is the sensitivity they perform to changes in the data used in their calculations. An exponentially weighted moving average is a means of smoothing random fluctuations that has the following desirable properties (Holt, 2004).

3.6 REGRESSION ANALYSIS

Regression analysis begins with the identification of two categories of variables: dependent variables and independent variables. In the context of demand forecasting, the dependent variable will always be demand(in our case is NG consumption). The independent variable(s) are those factors that the analyst has reason to believe may influence the dependent variable . Thus, once this relationship has been confirmed , we can calculate a formula to help in the prediction of the results of other variables. If there is only one independent variable in our study , we will apply the basic linear regression formula:

The model for simple linear regression is:

Y=$\beta_0 + \beta_1 X_{1+} \beta_2 X_2 + \cdots . + \beta_n X_n$+et , where

β_0 : Y intercept

$\beta_{1\ldots n}$: Coefficient of X

$X_{1\ldots n}$: Independent variable

Y: Dependent variable

et: error term

Regression models that describe the linear relationship between one dependent or response variable and more than one independent or explanatory variables are called multiple linear regression models (Yildiz et al., 2017) .In our study we will use a multiple linear regression with more than one independent variables in order to detect any seasonality in the natural gas consumption in Spain. Besides that we will try to detect if Covid-19 or War in Ukraine has affected the natural gas consumption in Spain as well.

3.7 MEAN ABSOLUTE PERCENTAGE ERROR-MEAN ABSOLUTE ERROR

If we subtract the forecasted value from the actual one and we calculate the average of those differences we will find the Mean Absolute Error:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)$$

The lower the MAE,the better forecasting accuracy we have.

Mean Absolut Percentage Error is E is the percentage difference between the actual value and the predicted one:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_{i-}\hat{Y}_i}{Y_i}$$

In order to have high forecasting accuracy,MAPE has to be as low as possible.

Forecasting accuracy metrics are using the forecast error which is defined as the difference between the actual value and the forecast at time t. If the difference is positive, the forecasting model underestimated the future electricity load while if the difference is negative the forecasting model overestimated the future electricity load(Islam et al., 2020).

4.EMPIRICAL STUDY

4.1 DESCRIPTIVE STATISTICS

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Mean	*15789,77*	17233,37	16370,05	17564,37	19748,47	19718,45	21548,71	21323,65	22273,82	22331,58
Standard Error	426,54	511,55	476,34	582,26	818,98	1118,72	1348,60	1246,97	1391,38	1317,76
Median	15724,14	17181,97	16200,51	17072,14	19139,78	18714,25	20443,30	19846,84	21225,87	22031,66
Mode	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Standard Deviation	1477,57	1772,05	1650,10	2016,99	2837,03	3875,36	4671,70	4319,62	4819,89	4564,86
Sample Variance	2183225	3140148,84	2722831,63	4068255,14	8048764,00	15018391,14	21824763,80	18659138,30	23231323,08	20837917,35
Kurtosis	2,29	0,98	-1,74	-0,42	-0,53	-0,89	-1,37	-1,17	-1,01	-1,29
Skewness	-0,60	-0,93	0,09	0,11	0,33	0,46	0,12	0,32	0,39	0,07
Range	5918,94	5992,17	4349,60	6975,82	9585,32	12121,28	13914,55	13432,59	14745,52	13747,95
Minimum	12365,92	13196,89	14303,05	13953,77	15170,43	14264,46	14313,04	14981,84	15509,79	14979,47
Maximum	18284,85	19189,06	18652,65	20929,59	24755,76	26385,75	28227,59	28414,43	30255,31	28727,42
Sum	189477,19	206800,39	196440,65	210772,49	236981,66	236621,34	258584,56	255883,83	267285,83	267978,97
Count	12	12	12	12	12	12	12	12	12	12

2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
19945,48	20306,90	20306,90	21982,67	22988,64	22988,24	21572,44	22826,06	18200,13	18104,45
1096,72	1137,26	1137,26	1346,69	1229,78	1301,97	1314,11	1218,72	1704,00	969,75
18262,00	18979,13	18979,13	19841,01	22154,05	22167,33	20013,90	22072,46	17516,88	17328,04
#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
3799,13	3939,60	3939,60	4665,05	4260,10	4510,15	4552,22	4221,75	5902,84	3359,31
14433412,83	15520410,82	15520410,82	21762725,97	18148422,55	20341444,08	20722683,95	17823210,97	34843563,84	11284994,16
-0,69	-1,35	-1,35	-0,93	-1,76	-0,44	-1,05	-0,48	-0,21	-1,32
0,70	0,42	0,42	0,68	0,10	0,55	0,63	0,62	0,78	0,43
12267,97	11640,22	11640,22	13608,87	11284,56	14909,16	13101,44	13561,16	18855,12	9903,49
14859,25	15014,84	15014,84	16519,28	17334,76	17135,99	16747,72	17777,56	11190,47	13675,60
27127,21	26655,06	26655,06	30128,16	28619,31	32045,15	29849,15	31338,72	30045,58	23579,10
239345,72	243682,78	243682,78	263792,10	275863,67	275858,83	258869,33	273912,68	218401,56	217253,46
12	12	12	12	12	12	12	12	12	12

Table 1:Descriptive Statistics/conventional reasons

2004-2008.We observe that there is an increasing trend in the average natural gas consumption from 2004 to 2008,starting from around 15789 in 2004 and reaching approximately 19748 in 2008.Also Standard deviation during this period fluctuates. We can assume that distribution approaches normality since skewness is close to 0 these years. Higher standard errors are seen in 2008, which suggests that there was some natural gas consumption volatility in these particular years. The ng consumption range for these years varies, with 2008 having the largest variation, which may point to certain outliers.

2009-2013:The average ng consumption reach a peak in 2013, at around 22331,58.Compared to other years, the standard deviation (SD) values for this time period are comparatively greater, suggesting higher levels of ng consumption volatility. Skewness values are positive, indicating right-skewed distributions. Kurtosis values show leptokurtic distributions with heavier tails.

2014-2018:From 2014 to 2018, average ng consumption was between 19945 and 22988. When compared to the previous period, the standard deviation (SD) values are approximately the same , indicating a stability in ng cunsumption volatility. Skewness remains positive, indicating right-skewed distributions. Leptokurtic distributions are provided by kurtosis values.

2019-2023: From 2019 forward, average ng consumption shows a stability and then a decrease with consumption ranging from 18104 to 22988 in 2023.Leptokurtic distributions are suggested by kurtosis values.The consumption range varies, with the largest range seen in 2022 and the lowest range in 2023, reflecting erratic ng consumption patterns.

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Mean	4802,41694	8201,7211	10908,523	11715,97	15549,0391	13160,83983	11233,07542	9085,1115	7060,175833	4678,393417
Standard Error	255,1892341	367,61848	503,14982	794,68184	487,4786	701,7116021	467,3376227	427,6537451	328,1192011	315,4858636
Median	4742,167806	8495,794	10366,598	12162,996	15537,6518	12944,5805	10963,8015	9347,934	6928,5995	4787,8465
Mode	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Standard Deviation	884,0014379	1273,4678	1742,9621	2752,8586	1688,67541	2430,800294	1618,905014	1481,436029	1136,638254	1092,87509
Sample Variance	781458,5422	1621720,2	3037916,9	7578230,7	2851624,63	5908790,07	2620853,443	2194652,709	1291946,521	1194375,962
Kurtosis	-1,314060524	-0,972804	1,3516958	-1,222429	0,49142713	-1,110388521	1,13467196	-0,122638327	0,214358691	-0,899588137
Skewness	0,039567604	0,2676475	1,2900859	0,0650802	-0,3807271	0,295881357	0,893150289	-0,699517446	0,32883507	-0,260075276
Range	2515,662214	3923,0043	5866,9844	7907,5952	6344,98053	7590,606	5893,542	5004,423	4014,895	3376,09
Minimum	3450,048263	6632,1855	8989,3799	8190,0842	12056,11	9598,667	8965,783	6041,875	5195,727	2778,292
Maximum	5965,710477	10555,19	14856,364	16097,679	18401,0905	17189,273	14859,325	11046,298	9210,622	6154,382
Sum	57629,00328	98420,653	130902,28	140591,65	186588,47	157930,078	134796,905	109021,338	84722,11	56140,721
Count	12	12	12	12	12	12	12	12	12	12

2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
4285,5955	5090,326	4959,379417	6288,060583	5120,129667	9276,921417	7301,741417	7527,5555	11366,64892	8050,869
273,6505023	306,5805278	457,8315861	666,0449519	334,6840032	1026,861188	800,5419196	802,9471261	1092,199007	543,533134
4267,0215	5057,386	4672,3955	6754,6405	5116,771	8565,9205	6790,9915	7460,8225	10162,0775	7660,6755
#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
947,953147	1062,026102	1585,975137	2307,247394	1159,379396	3557,151498	2773,158557	2781,490436	3783,488344	1882,854007
898615,169	1127899,44	2515317,135	5323390,536	1344160,584	12653326,78	7690408,38	7736689,048	14314784,05	3545139,214
-0,850673337	3,392008551	-0,917983157	-1,172981874	-0,149398977	-1,086600591	-0,352732355	-0,295609418	-1,501063945	-1,83228625
0,396652277	1,643308984	0,620527611	-0,191276112	-0,186362475	0,480607857	0,673179781	0,242413692	0,311235138	0,215281964
2971,455	3754,861	4536,203	7039,637	3954,996	10111,834	8808,316	9695,081	10721,278	4822,441
3105,795	4068,094	3045,953	2958,423	3135,872	4919,166	3770,982	3106,837	6350,575	5772,761
6077,25	7822,955	7582,156	9998,06	7090,868	15031	12579,298	12801,918	17071,853	10595,202
51427,146	61083,912	59512,553	75456,727	61441,556	111323,057	87620,897	90330,666	136399,787	96610,428
12	12	12	12	12	12	12	12	12	12

Table 2:Descriptive Statistics/electricity generation

2004-2008: The mean and median values show a continuous upward trend between 2004 and 2008, suggesting a growth in the data's central tendency. This points to a possible upward trajectory or general rise in the measured quantities over the given time frame. Notable increases are also seen in the standard deviation and range, particularly between 2006 and 2008. This suggests that the data throughout these years have become more variable or spread out, suggesting possible fluctuations or dispersion of the observed values around the mean. Over time, skewness stays low and nearly constant, demonstrating symmetry; however, kurtosis varies, pointing to potential shifts in the tail behavior of the distribution.

2009-2013: Between 2009 and 2013, the dataset shows a significant drop in the mean and median values, which suggests an overall decline in the measured quantities' central tendency. The maximum and minimum values decrease in line with this trend, indicating a gradually decreasing range of observations. A more tight distribution of observations around the mean is also implied by the declining standard deviation and range, which point to a more concentrated distribution of values. Variations in kurtosis suggest that the distribution's tail behavior may be changing, even though skewness is still comparatively low.

2014-2018: The dataset shows an ongoing increase in the mean and median values between 2014 and 2017, which is suggestive of an increasing central tendency of the measured variables during that time. .Furthermore, although there are variations in the range and standard deviation, the general trend points to an increasing distribution of observations around the mean, suggesting a less concentrated distribution of values. The distribution's symmetry is shown by the relatively low skewness, while variations in kurtosis suggest that the tail behavior of the distribution may have changed.

2019-2023: The data from 2019 to 2023 show variations in the mean and median values, which point to variations in the measured quantities' central tendency over that time. Although there are occasional increases, especially in some years(e.g. 2022), the tendency is not constantly rising. Standard deviation and range variations show that despite these fluctuations, there are noticeable variations in the data's variability. While fluctuations in kurtosis indicate changes in the tail behavior of the distribution, skewness values stay relatively close to zero, suggesting overall symmetry in the distribution.

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Mean	700,236639	677,13339	629,89874	699,93001	702,306995	833,6253333	911,1658333	1073,929	1111,92	954,7383333
Standard Error	13,6517363	14,008897	20,264166	20,920585	16,6258261	19,32690514	17,070506	27,40719682	22,67192376	18,53397505
Median	704,8324808	675,20947	622,01479	682,42116	706,190044	839,0415	930,5215	1104,7985	1110,3335	955,586
Mode	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Standard Deviation	47,29100177	48,528243	70,19713	72,471031	57,5935512	66,9503633	59,13396741	94,94131477	78,53784772	64,2035729
Sample Variance	2236,438849	2354,9904	4927,6371	5252,0503	3317,01714	4482,351147	3496,826101	9013,85325	6168,193524	4122,098773
Kurtosis	-1,44735898	-0,564741	0,0644158	1,9510772	-1,4765737	-0,266215967	-0,570510234	3,965264123	-0,616738246	0,722738033
Skewness	-0,224956641	0,1868665	0,7485017	1,3829782	-0,218439	-0,446237641	-0,143461547	-1,79216849	-0,035758657	-0,721465308
Range	133,0759793	155,60997	235,08702	242,34341	159,308	218,528	196,313	353,282	256,218	223,142
Minimum	624,9348686	599,4329	541,07248	632,60477	616,931	718,172	818,848	824,696	990,774	814,642
Maximum	758,010848	755,04287	776,1595	874,94818	776,239	936,7	1015,161	1177,978	1246,992	1037,784
Sum	8402,839668	8125,6007	7558,7849	8399,1601	8427,68394	10003,504	10933,99	12887,148	13343,04	11456,86
Count	12	12	12	12	12	12	12	12	12	12

2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
912,32225	786,9066667	827,67875	832,65525	845,0205	913,5546667	999,1584167	1097,857917	756,85575	787,9305833
28,43977408	19,62557845	13,2920979	17,13896143	13,87934236	17,03948994	24,65971725	10,05656929	29,18663116	13,94516442
916,468	793,0945	809,018	810,135	843,287	907,7765	1011,0315	1104,0405	709,3545	791,069
#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
98,51826734	67,98499799	46,04517779	59,37110398	48,0794523	59,02652463	85,42376636	34,83697793	101,1054561	48,3074666
9705,848999	4621,959952	2120,158398	3524,927988	2311,633733	3484,13061	7297,21986	1213,615031	10222,31326	2333,611329
-0,296408662	-0,652387056	-1,339054484	0,391425181	0,158179757	1,73301157	-0,562977122	-0,196941163	-1,012475085	1,148551333
0,460390268	-0,407503154	0,659906351	1,080858899	-0,173327457	-0,025055141	-0,815436721	-0,622299399	0,723572692	-0,15109519
328,166	224,769	127,907	181,699	166,434	241,957	245,937	117,536	293,389	190,363
779,882	663,619	777,262	768,593	764,624	790,952	845,923	1030,449	644,487	692,28
1108,048	888,388	905,169	950,292	931,058	1032,909	1091,86	1147,985	937,876	882,643
10947,867	9442,88	9932,145	9991,863	10140,246	10962,656	11989,901	13174,295	9082,269	9455,167
12	12	12	12	12	12	12	12	12	12

Table 3:Descriptive Statistics/LNG use

2004-2008: The dataset's mean and median values are constant from 2004 and 2008, suggesting a consistent central tendency during that time. The range and standard deviation, on the other hand, show significant variation, indicating variations in the distribution of data points around the mean. Variations in skewness and kurtosis values also point to modifications in the distribution's symmetry and tail behavior

2009-2013: Between approximately 833.63 and 1111.92 are the mean values, while between approximately 839.04 and 1110.33 are the median values. The standard deviation figures show variations in the distribution of data around the mean, ranging from roughly 59.13 to 94.94. Skewness scores are unpredictable, ranging from slightly negative in some years to slightly positive in others. This suggests alterations in the distribution's symmetry. Kurtosis values also fluctuate, with a noteworthy peak in 2011, signifying alterations in the tail behavior of the distribution.

2014-2018: The dataset, which covers the years 2014 to 2018, displays variations in mean and median values, suggesting changing central tendencies over time. The median values range from 793.09 to 916.47, while the mean is about between 786.91 and 912.32. Data dispersion variability is indicated by standard deviation values, which range from 46.05 to 98.52. Skewness values vary, indicating shifts in the symmetry of the distribution; some years have positive skewness and others have negative skewness. Kurtosis values also fluctuate, signifying changes in the tail behavior of the distribution.

2018-2023: The dataset shows mean and median value fluctuations from 2019 to 2023, indicating changes in central tendency over time. The median values span from 709.35 to 1104.04, while the mean is about between 756.86 and 1097.86. Standard deviation measurements show considerable variation in data dispersion, with values ranging from about 34.84 to 101.11. Skewness values vary, suggesting shifts in the symmetry of the distribution; some years have positive skewness and others have negative skewness. Kurtosis values also fluctuate, signifying modifications in the tail behavior of the distribution.

YEAR	NG CONSUMPTION FOR CONVENTIONAL REASONS		NG CONSUMPTION FOR ELECTRICTY GENERATION		NG CONSUMPTION FOR LNG	
	MEAN	VARIANCE	MEAN	VARIANCE	MEAN	VARIANCE
2004	*15789.76585*	2183225.108	4802,417	781458,5	700,2366	2236,439
2005	17233.37	3140148.835	8201,721	1621720,205	677,1334	2354,99
2006	16370.05	2722832	10908,52	3037916,943	629,8987	4927,637
2007	17564.37	4068255	11715,9705	7578230,74	699,93	5252,05
2008	19748.47	8048764.004	15549,03913	2851624,628	702,307	3317,017
2009	19718.45	15018391.14	13160,83983	5908790,07	833,6253	4482,351
2010	21548.71	21824763.8	11233,07542	2620853,443	911,1658	3496,826
2011	21323.65	18659138.3	9085,1115	2194652,709	1073,929	9013,853
2012	22273.82	23231323.08	7060,175833	1291946,521	1111,92	6168,194
2013	22331.58	20837917.35	4678,393	1194375,962	954,7383	4122,099
2014	19945.48	14433412.83	4285,596	898615,169	912,3223	9705,849
2015	20306.9	15520410.82	5090,326	1127899,44	786,9067	4621,96
2016	20876	13437140.57	4959,379	2515317,135	827,6788	2120,158

2017	21982.67	21762725.97	6288,060583	5323390,536	832,6553	3524,928
2018	22988.64	18148422.55	5120,13	1344161	845,0205	2311,634
2019	22988.24	20341444.08	9276,921	12653326,78	913,5547	3484,131
2020	21572.44	20722683.9	7301,741	7690408,38	999,1584	7297,22
2021	22826.06	17823211	7527,556	7736689,048	1097,858	1213,615
2022	18200.13	34843564	11366,65	14314784,05	756,8558	10222,31
2023	18104.45	11284994.16	18104,45	11284994,16	787,9306	2333,611

Table 4:Mean & Variance of Ng consumption

4.2)Time series analysis

Finding the three primary elements that recur throughout time—the trend, the seasonal component, and the irregular or random component—will allow one to completely understand a time series process. It is always possible for all or some of the components to coexist in a given time series. We will try to distill safe conclusions about the seasonal and trend occurrences and identify the key elements that put together the time series of the ng consumption under consideration using the following graphical analysis.

4.2.1)Natural gas consumption for conventional reasons

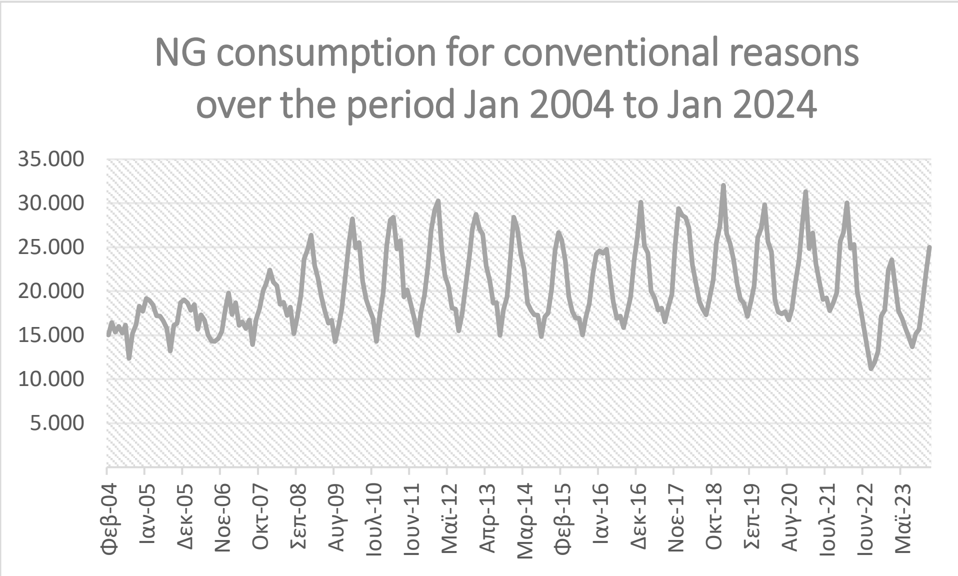


Figure 4:Diagram of Ng consumption

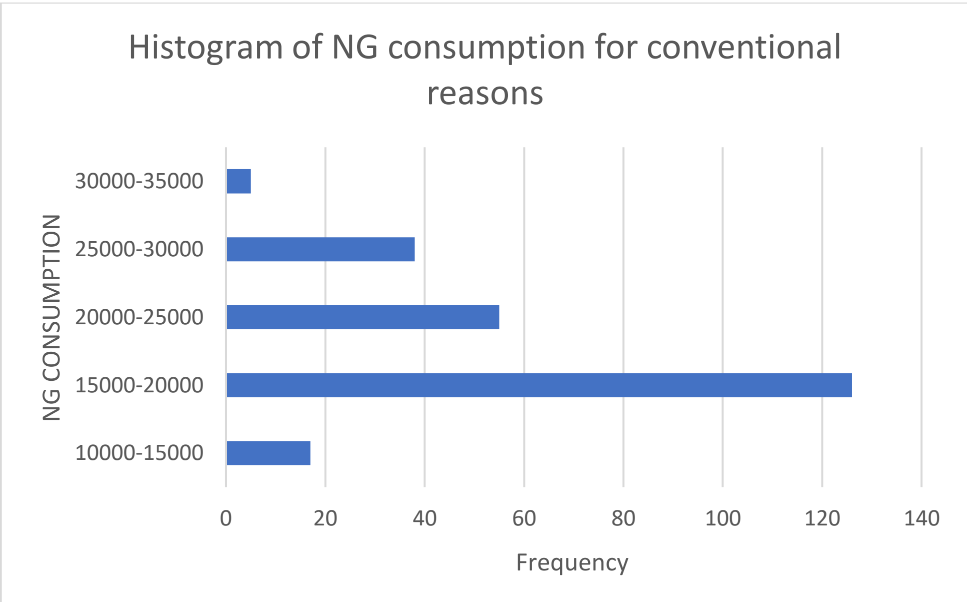


Figure 5:Histogram of ng consumption

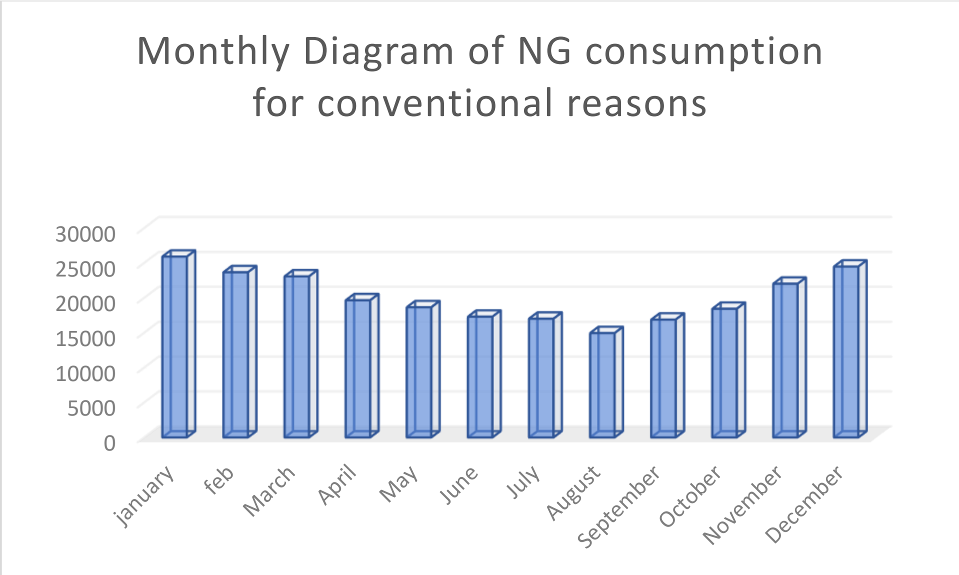


Figure 6 :Monthly diagram of ng consumption

Natural gas consumption in Spain for conventional reasons(heating) has the most increased values among the 3 categories(conventional gas consumption, gas consumption for electricity ,LNG gas consumption).

In the time series plot(Figure 4) we can't observe a stationary process. The evolution of the mean and variance of the ng consumption generated in Table 2 can be used to verify this deduction. Since the mean and the variance are not constant ,we can assume that the series don't exhibit a stationary process. By definition, a stationary process is a time series which its statistical properties are relatively stable over time (Thomaidis, 2021a).

Also the chart displays not only the ng consumption over time, but also the seasonal patterns within a 12 monthly period in the corresponding diagram and the center and variability of the consumption in the histogram plot. A brief analysis of the monthly diagram suggests that there is a strong seasonal impact reflected in the consumption, which is highest in December and January each year, while the lowest demand appears to be in August (People are using more ng for heating on winter due to cold) .The same graphic also shows that there are outliers in the data set and that the monthly sales variance is generally constant.

Over the years, there has been a noticeable upward trend in natural gas consumption for conventional reasons.

4.2.2)Natural gas consumption for electricity generation

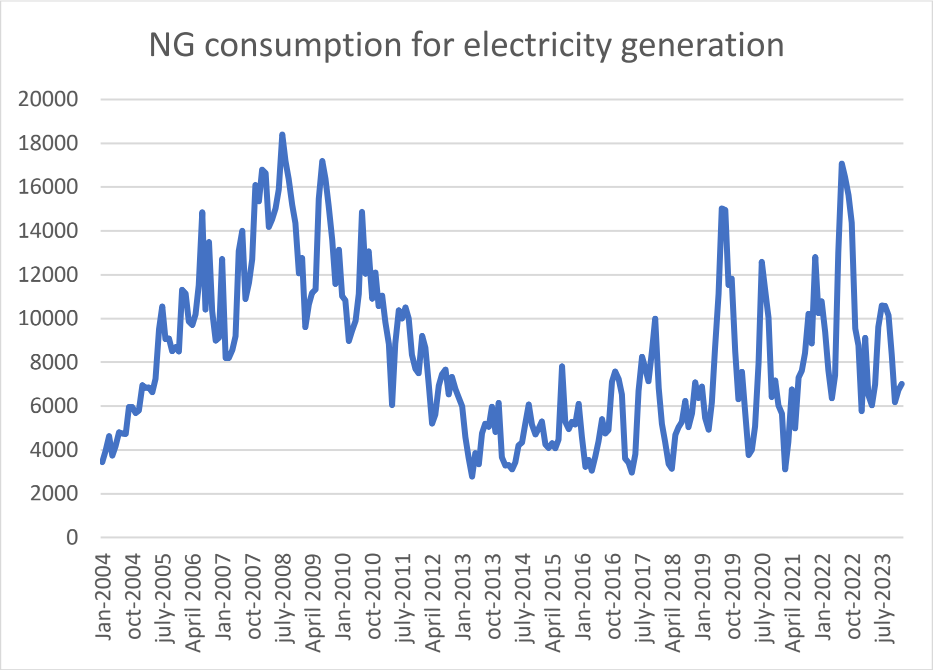


Figure 7: Diagram of Ng consumption

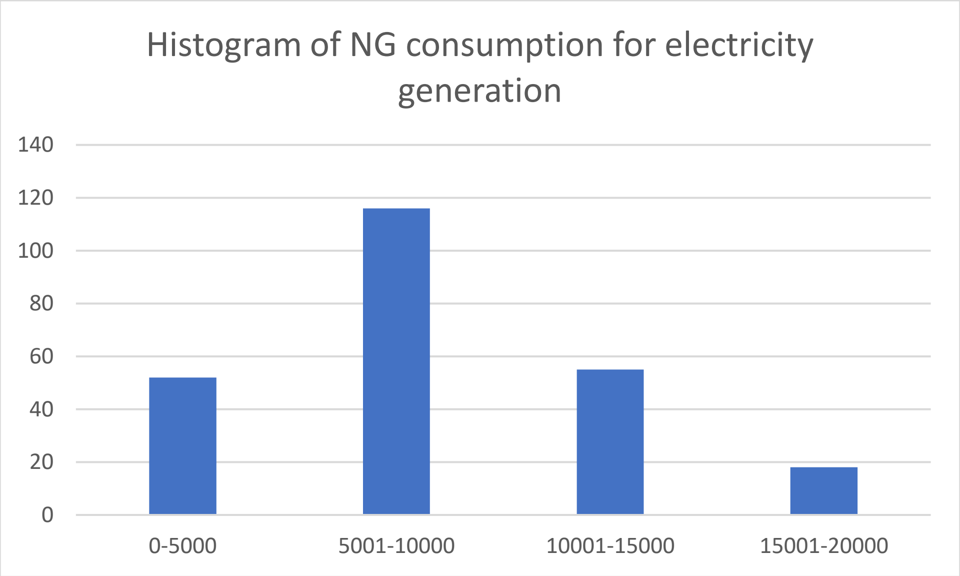


Figure 8: Histogram of ng consumption

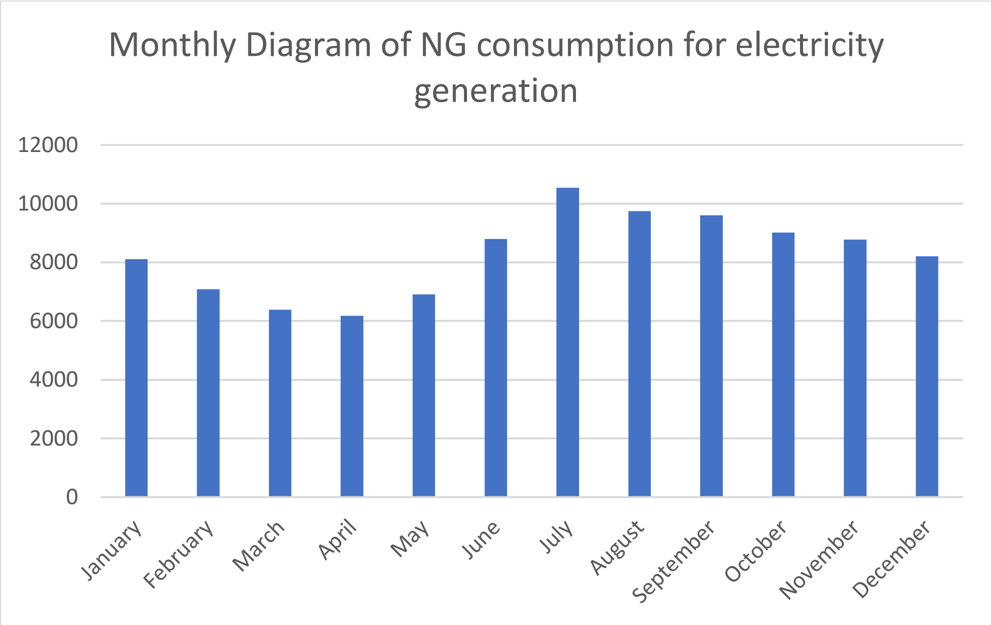


Figure 9: Monthly diagram of ng consumption

Time series data with a constant mean and constant variance are categorized as stationary data . The data under analysis in this instance, however, clearly show a seasonal trend, suggesting that the mean is not stable over time. Additionally, Table 2 demonstrates how the data's variance has evolved over time. Consequently, it can be said that the analysis's data are not stationary. Since the patterns and trends of non-stationary time series data may change over time, it can be more difficult to interpret and make reliable predictions.

Also Figure 7 shows some cyclic behavior. When there are fluctuations in the data that don't occur at a consistent rate, a cycle is present. These variations are typically brought on by the state of the economy and are frequently associated with the “business cycle”. By observing the monthly diagram(Figure 10) we can observe that there is also a seasonal pattern. Ng consumption for electricity generation is very high on July and the lowest level of ng consumption is on April.

4.2.3 NG consumption for LNG

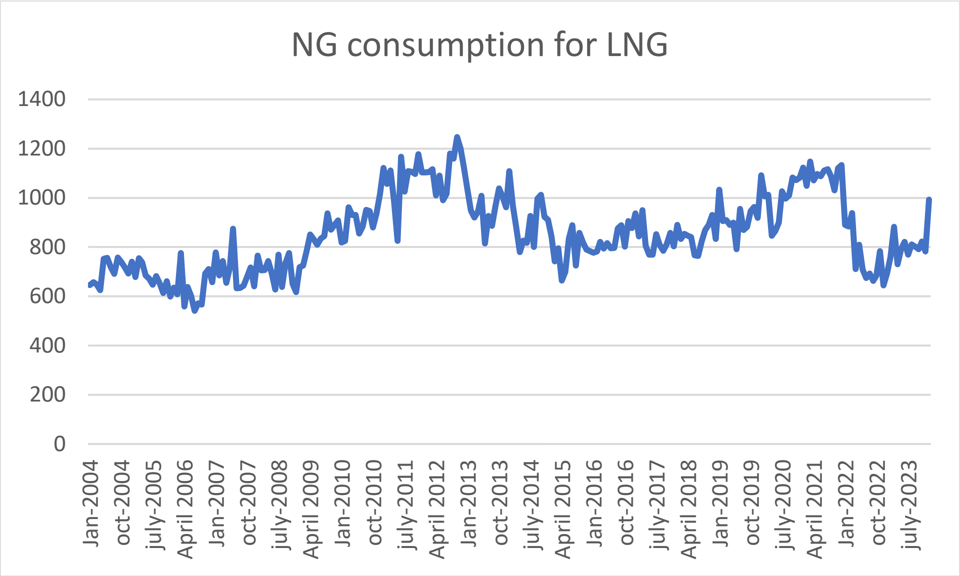


Figure 10: Diagram of Ng consumption

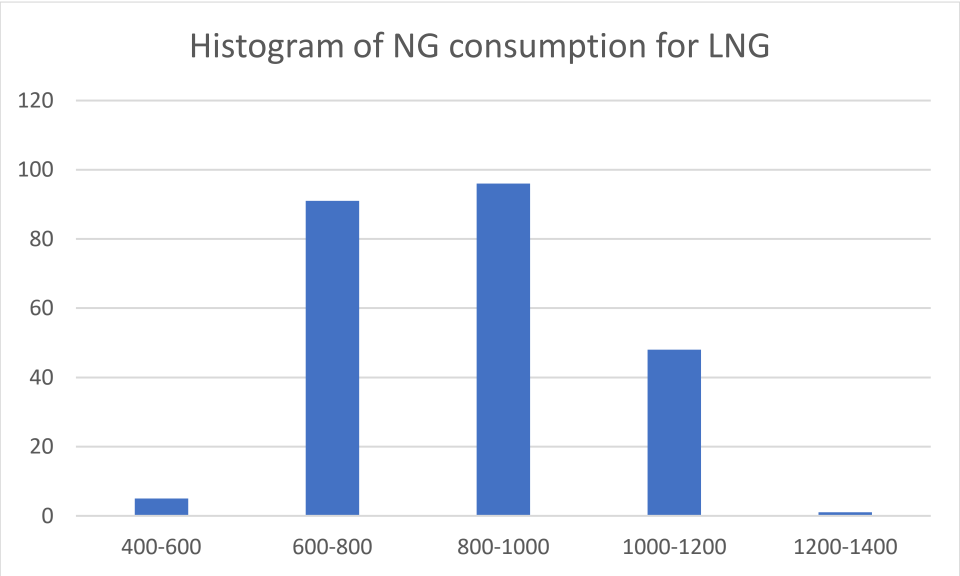


Figure 11: Histogram of ng consumption

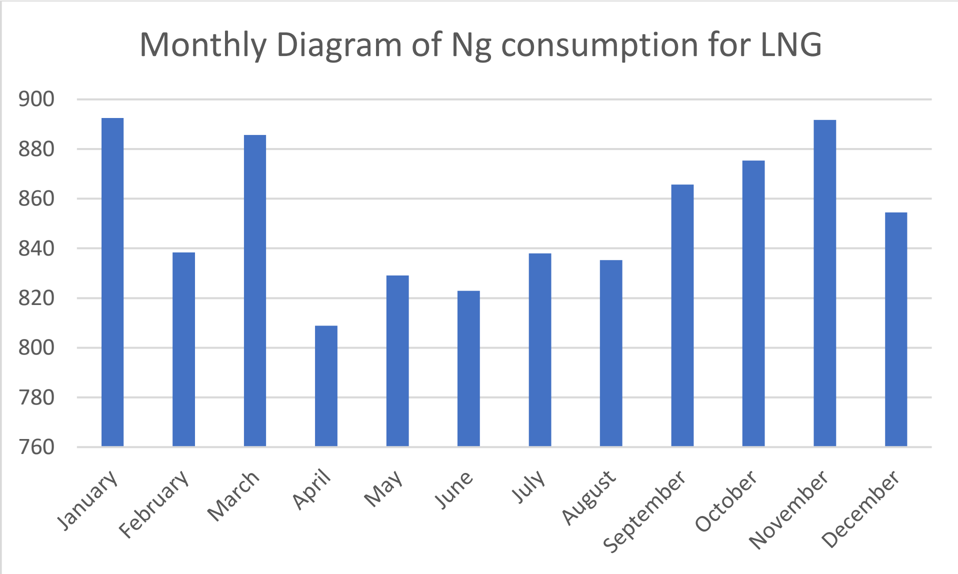


Figure 12:Monthly diagram of ng consumption

There is no obvious pattern or obvious increasing or downward trend to be seen in the time plot that shows the ng consumption for LNG during the whole period under investigation. Furthermore, there are no notable or noticeable variations in the time series that would complicate the analytical process.However mean and variance are not varying so according to Table 2,so we will consider the data as non-stationary.We can also detect from Figure 11 an irregular behavior. According to Figure 14 it seems that there is no seasonality across the specific months and besides that there are no outliers.

4.3 Simple Moving Average

4.3.1 SMA of Ng consumption for Conventional reasons

We will start forecasting with the Simple Moving Average Method which we descripted in chapter **3.2**. When using a simple average as a forecasting technique, the next month's forecast, and every future month's forecast, is the average level of demand from all previous months.We will calculate the display the forecasts from January 2005-January 2024.We will use 723 observations in total from 3 different time series(ng consumption for conventional reasons , electricity generation ,LNG consumption241 per category). So, for forecast one step ahead we will use a window length of 4,6,8,10,12 observations.

For SMA(4) the results are:

The Mean absolute Error is 3587.673531 and the Mean Absolute Percentage error is 0.173642481 =>17,36 %.

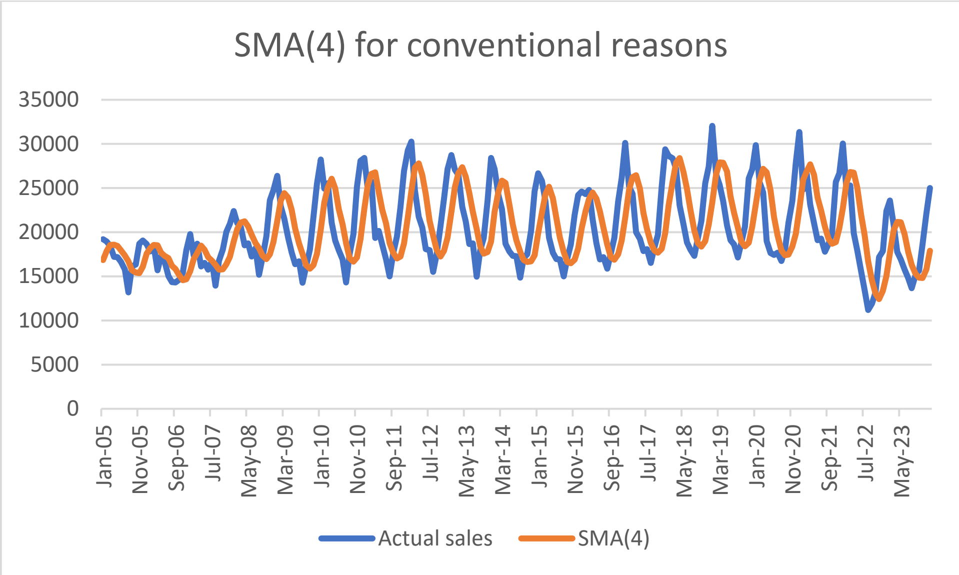


Figure 13 :SMA(4)

As we can see from the Figure 15 ,the forecasting ng consumption follows in the pattern of the actual consumption and can capture the seasonality . SMA line is fast moving and in general terms is approaching the actual ng consumption .There are also some underestimations as well .The MAPE for this method is 17,36 % indicating that the estimation can be considered as accurate.

For SMA(6) the results are :

The Mean absolute Error is 4217,90492 and the Mean Absolute Percentage error is 0,207215179 =>20,72 %.

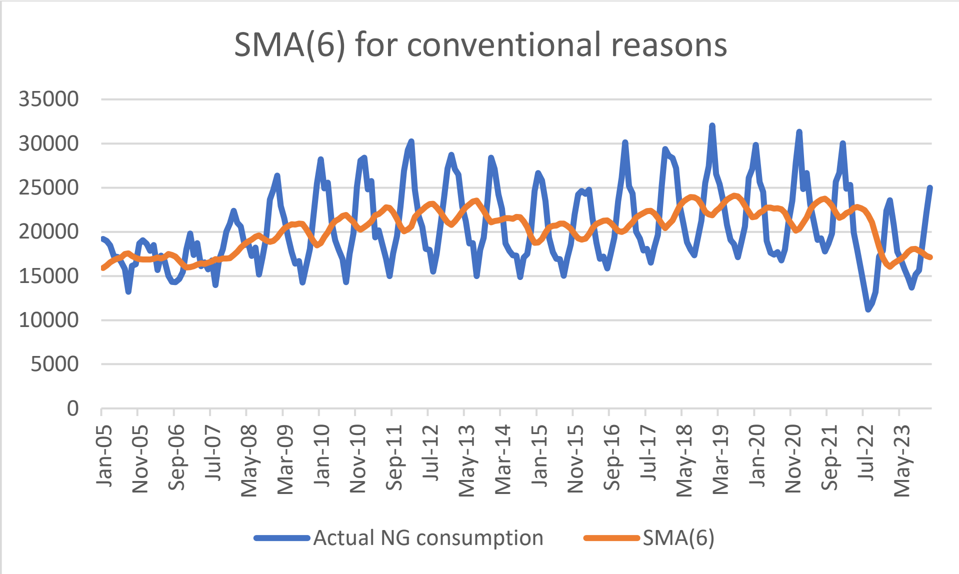


Figure 14:SMA(6)

In some cases the SMA(6) overestimates and in some cases underestimates ng consumption. However. The estimated NG consumption doesn't follow the patterns of the actual ng consumption in many cases. MAPE is 20,72 % and indicates that the forecasting model has a good predictability.

For SMA (8) the results are:

The Mean absolute Error is 4248,045685 and the Mean Absolute Percentage error is 0,210666311 =>21,06 %.

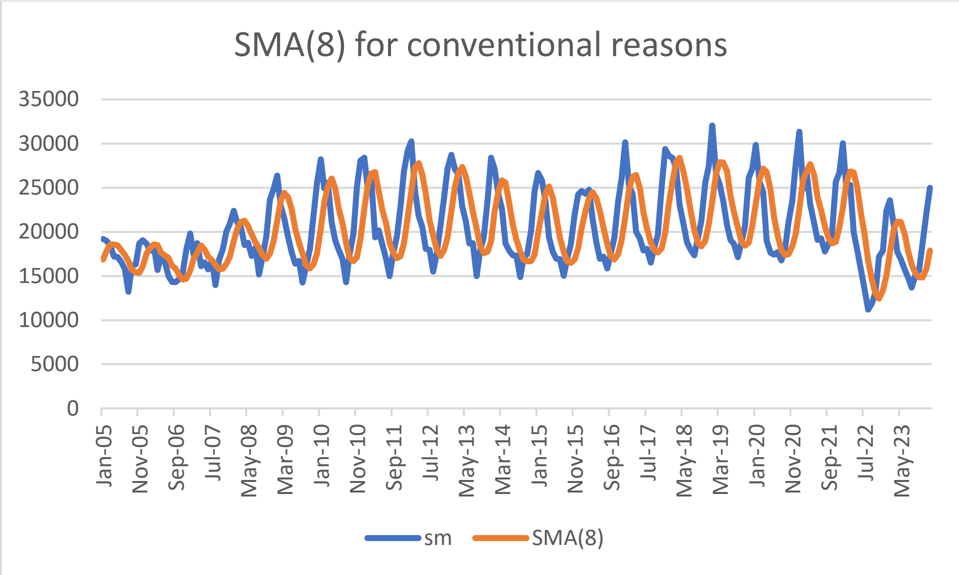


Figure 15: SMA(8)

We can observe in Figure 17 that many times there are underestimations and overestimations of the actual NG consumption. The estimated NG consumption doesn't follow the patterns of the actual ng consumption in many cases(same as SMA(6)). MAPE in this case is 21,06 which cannot be considered so high.

For SMA(10) the results are:

The Mean absolute Error is 3872,845027 and the Mean Absolute Percentage error is 0,192720552=>19,27 %.

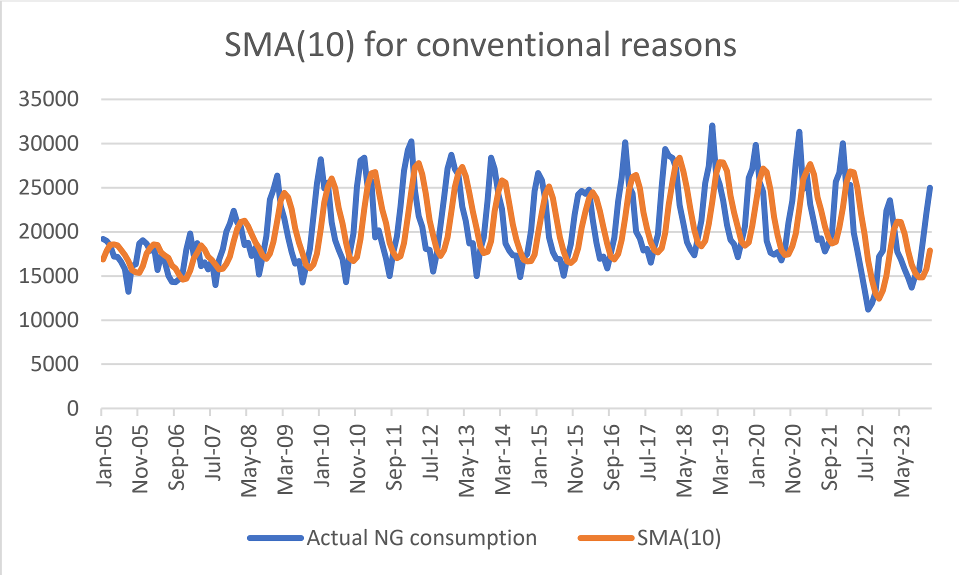


Figure 16:SMA(10)

In this case the forecasting model doesn’t follow the patterns of the actual gas consumption. There are underestimations and overestimations of the actual NG consumption. MAPE is close to 10 %(19,27 %).

For SMA(12) the results are:

The Mean absolute Error is 3350,00824 and the Mean Absolute Percentage error is 0,167123319 =>16,71 %.

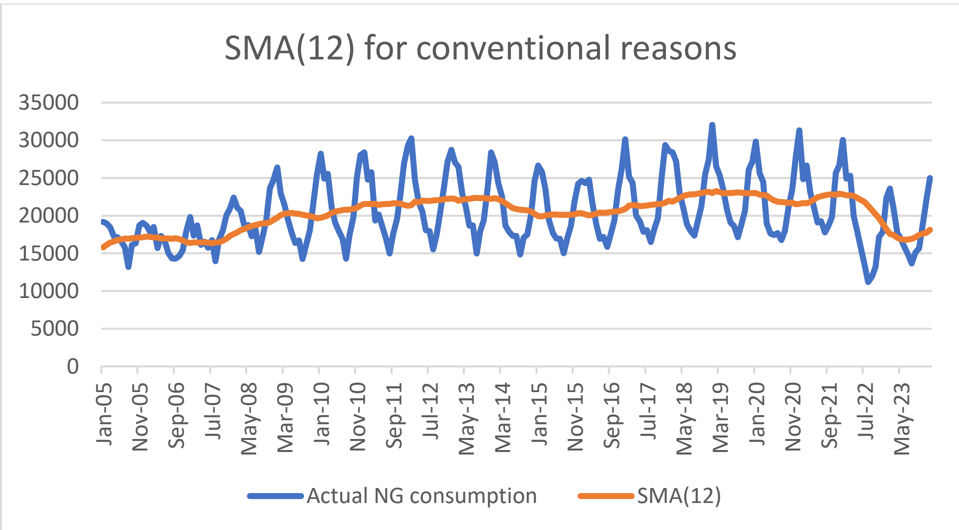


Figure 17:SMA(12)

According to the Figure 18, the forecasting data doesn't follow the pattern of the actual ng consumption and the forecasting line it seems more of a straight trend line. MAPE is low(16,71 %.).

We can observe that as the number of the observations increases in order to calculate the next month's forecast ,the forecasts are smoother(see Figure 19) and are following less the patterns of the actual ng consumption. If we take into consideration only the parameter of MAPE we can assume that Simple Moving Average(12) is the most reliable forecasting model since it has the lowest MAPE(16,71 %).

4.3.2)SMA FOR ELECTRICITY GENERATION

Now we will use the Simple Moving Average method in order to forecast the observations of the ng consumption for electricity generation. We have 241 data.First we will use SMA(4).

The Mean absolute Error is 1797,099074 and the Mean Absolute Percentage error is 0,237121471 =>23,71 %.

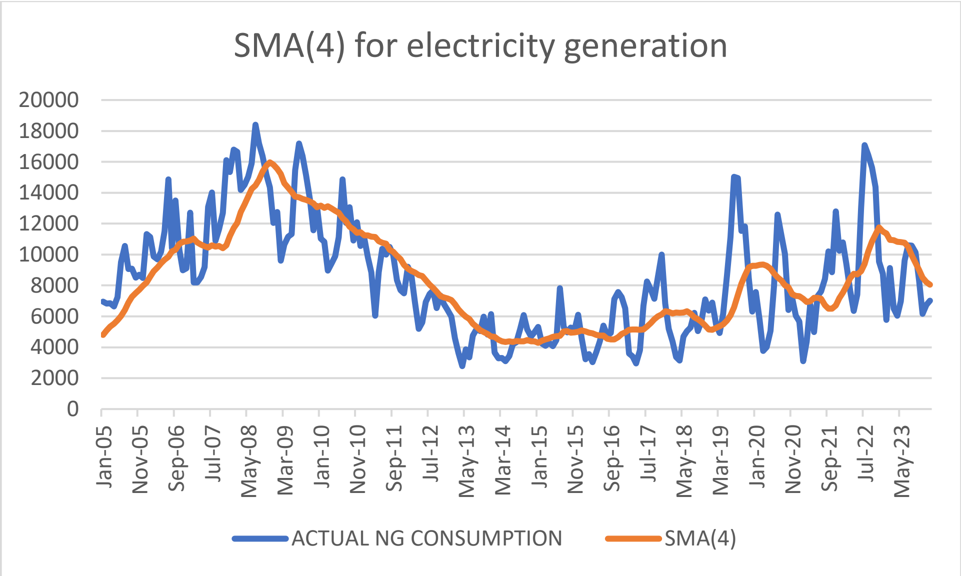


Figure 18: SMA(4)

In the above graph we can observe that the forecasting data are following the patterns of the actual data of ng consumption. If we take into account the forecasting error,it seems that in some cases we have overestimation and in some cases we have underestimation of the actual ng consumption.However we can consider this model as good fit for the data. The SMA line is trying to capture the patterns of the actual data and is quite fast moving.MPE for SMA(4) is 23,71 % which is not so low.

For SMA (6) the results are :

The Mean absolute Error is 2000,004464 and the Mean Absolute Percentage error is 0,269950621=>26,99 %.

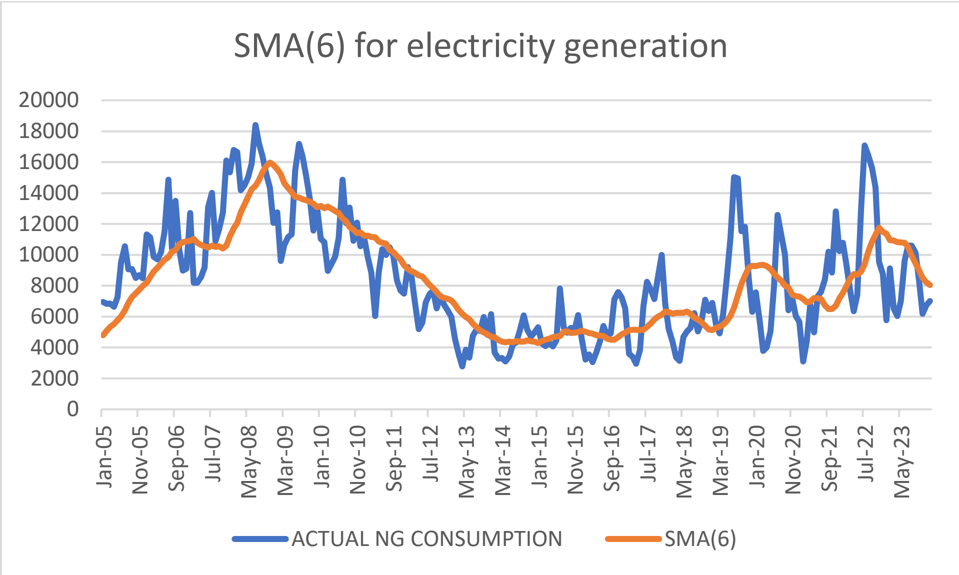


Figure 19:SMA(6)

If we take into account the forecasting error,it seems that in some cases we have overestimation and in some cases we have underestimation of the actual ng consumption. However we can consider this model as good fit for the data. The SMA line is trying to capture the patterns of the actual data and is quite fast moving.We can assume that the lag is not so important.Mape is this occasion is **26,99 %** which cannot be considered neither as high nor as low.

For SMA(8) the results are:

The Mean absolute Error 2230,539201 and the Mean Absolute Percentage error is 0,296490171=>29,64 %.

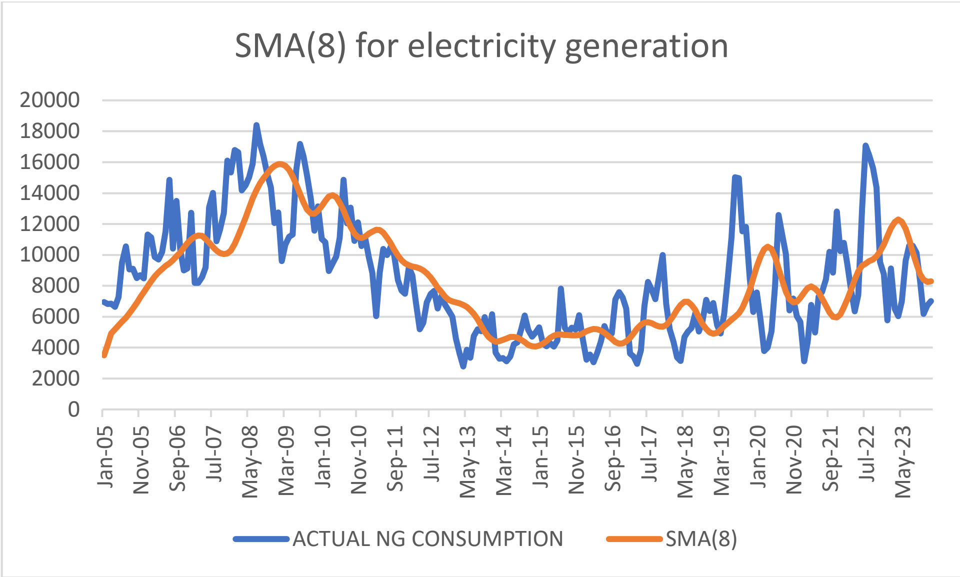


Figure 20:SMA(8)

In the graph we can observe that forecasted ng consumption sometimes underestimates for a long period the actual observations (2011-2013) and sometimes overestimates them (2005-2007).MAPE is 29,64 %. We could say that the forecasted ng consumption doesn't follow successfully the patterns of the actual ng consumption. Also MAPE cannot be considered as low.

For SMA(10) the results are :

The Mean absolute Error 2389,139856 and the Mean Absolute Percentage error is 0,27802441=>27,80 %.

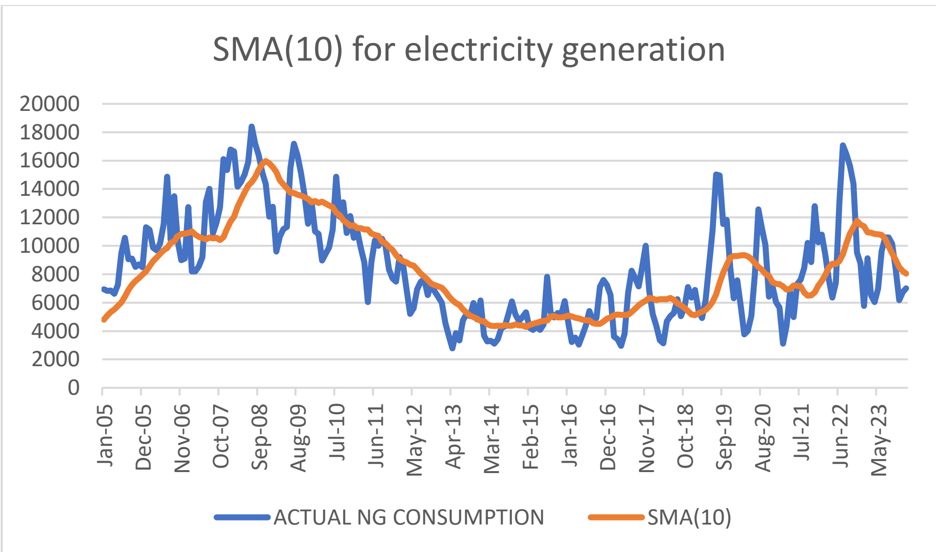


Figure 21:SMA(10)

In many occasions the forecasting model seems to overestimate the actual observations .SMA(10) is trying to follow the patterns of the actual consumption and MAPE is 27,80 %.

For SMA (12) the results are : The Mean absolute Error 1941,4128 and the Mean Absolute Percentage error is 0,257249671=> 25,74 %.

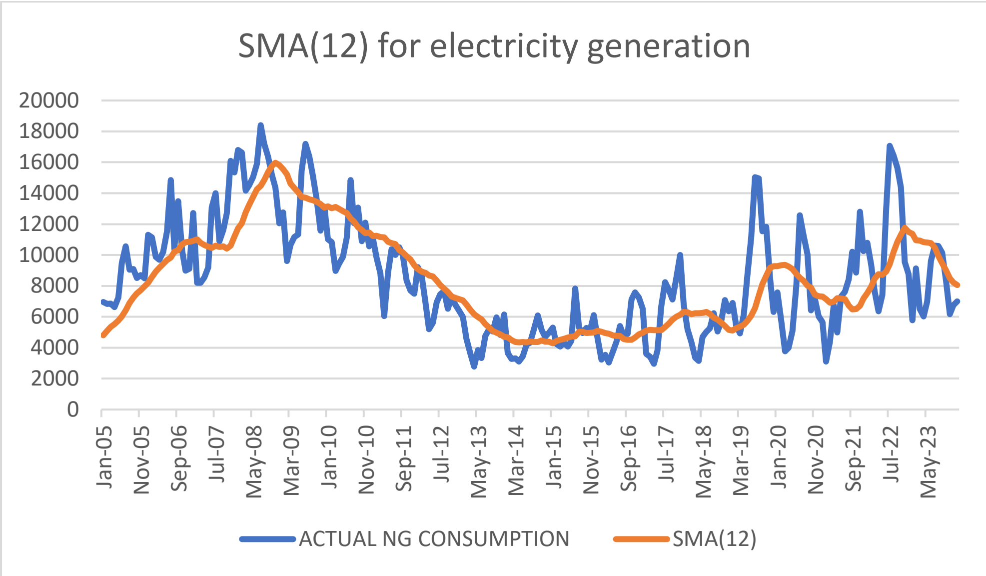


Figure 22:SMA(12)

We can observe both overestimations and underestimations of the actual ng consumption. MAPE is 25,72 %

4.3.3 SMA FOR LNG

The 3rd category of time series that we will use is the monthly ng consumption for LNG.

For SMA (4) the results are : The Mean absolute Error 62,42308286 and the Mean Absolute Percentage error is 0,07355 => 7,35 %.

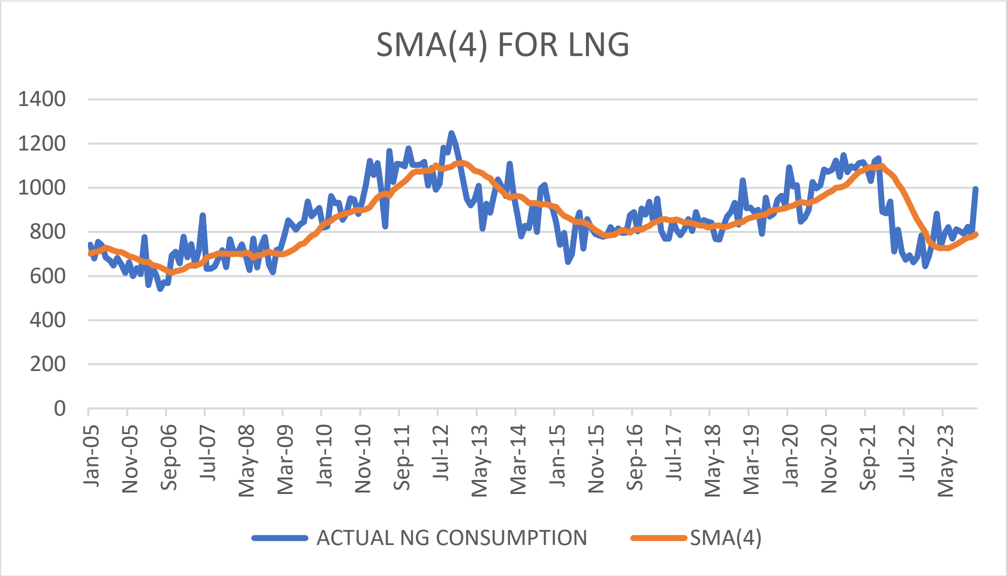


Figure 23:SMA(4)

In the above graph we can observe that the forecasting data are following the patterns of the actual data of ng consumption. We can consider this model as a good fit for the data. The SMA line is trying to capture the patterns of the actual data and is quite fast moving. Also MAPE is very low 7,35 %,so we can assume that forecasting model SMA(4) is very reliable and accurate.

For SMA (6) the results are :

The Mean Absolute Percentage error is 0,078702 => 7,87 %.

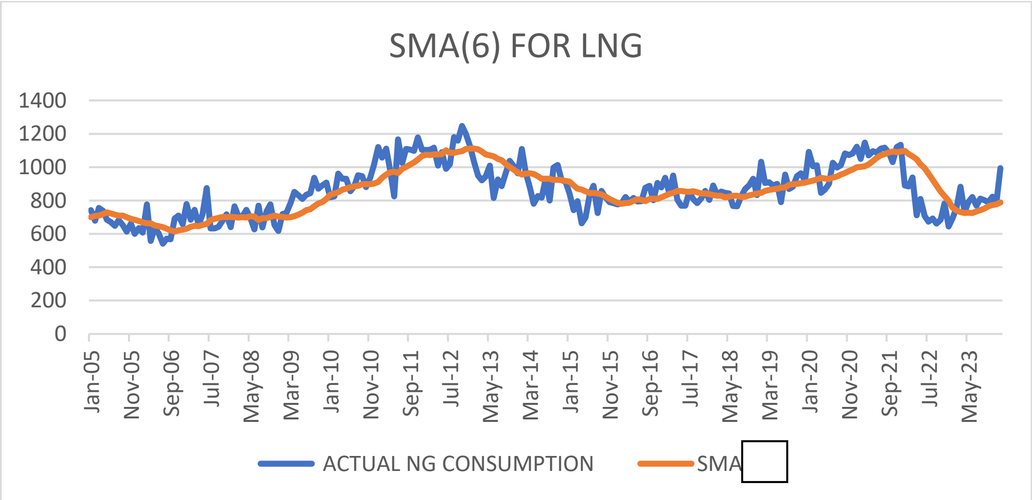


Figure 24:SMA(6)

This forecasting method is the same with the SMA(4).The model is a good fit for the data. Also MAPE is very low(7,87 %.).

For the SMA(8) the results are :

The Mean absolute Error 69,68504096 and the Mean Absolute Percentage error is 0,082778 => 8,27 %.

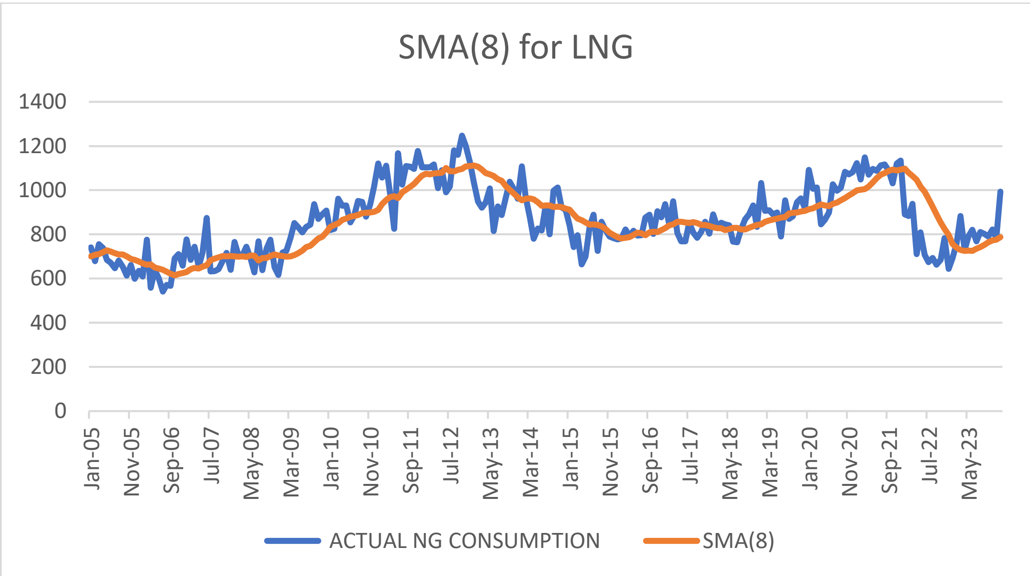


Figure 25:SMA(8)

MAPE is very low (8,27 %.).Also the forecast observations are trying to follow the patterns of the actual observations.There are also some underestimations(2022).

For SMA(10) the results are :

The Mean absolute Error 71,9879 and the Mean Absolute Percentage error is 0,085722

=> 8,57 %.

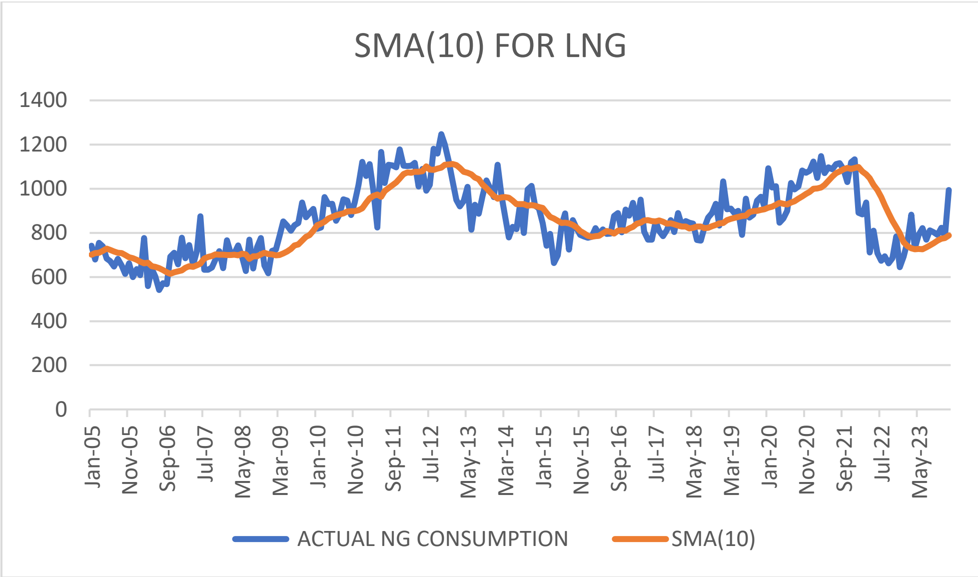


Figure 26:SMA(10)

According to the above graph,the forecasting model is a good fit for the actual data.MAPE is also very low(8,57 %.).

For SMA(12) the results are:

The Mean absolute Error 73,82682305 and the Mean Absolute Percentage error is 0,088021 => 8,8 %.

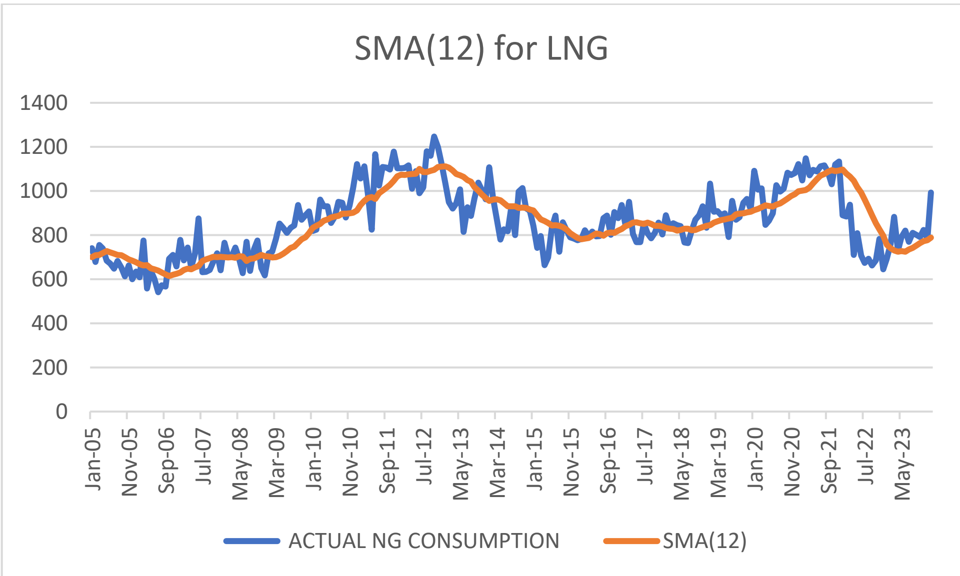


Figure 27:SMA(12)

Forecast observations are trying to follow the patterns of the actual observations. We can also observe some underestimations(2022).MAPE is also low(8,8 %).

4.4)EXPOTENTIAL SMOOTHING AVERAGE

Now we will continue with another forecasting method, the Exponential Weighted Moving Average method.As we already described in chapter **3.5** , EWMA is designed as such that older observations are given lower weights. The weights fall exponentially as the data point gets older – hence the name exponentially weighted. We will display the forecasts from January 2005-January 2024.We will start forecasting from January 2004.We will use 723 observations in total from 3 different time series(ng consumption conventional reasons,electricity generation, LNG consumption/241 observations per category). So, for the forecast method of one step ahead we will use λ =0,1/0,2/0,3/0,4/0.5/0,6/0,7/0,8/0.9.

4.4.1)EWMA FOR CONVENTIONAL REASONS

The results of the EWMA for conventional reasons are the following :

	CONVENTIONAL								
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
MAPE	0,105461	0,11151	0,119286	0,128195	0,137084	0,145721	0,156211	0,165102	0,169244
MAE	2171,648	2300,038	2462,354	2645,405	2824,696	2992,682	3185,071	3347,096	3410,94

Table 5:MAPE/MAE for conventional reasons

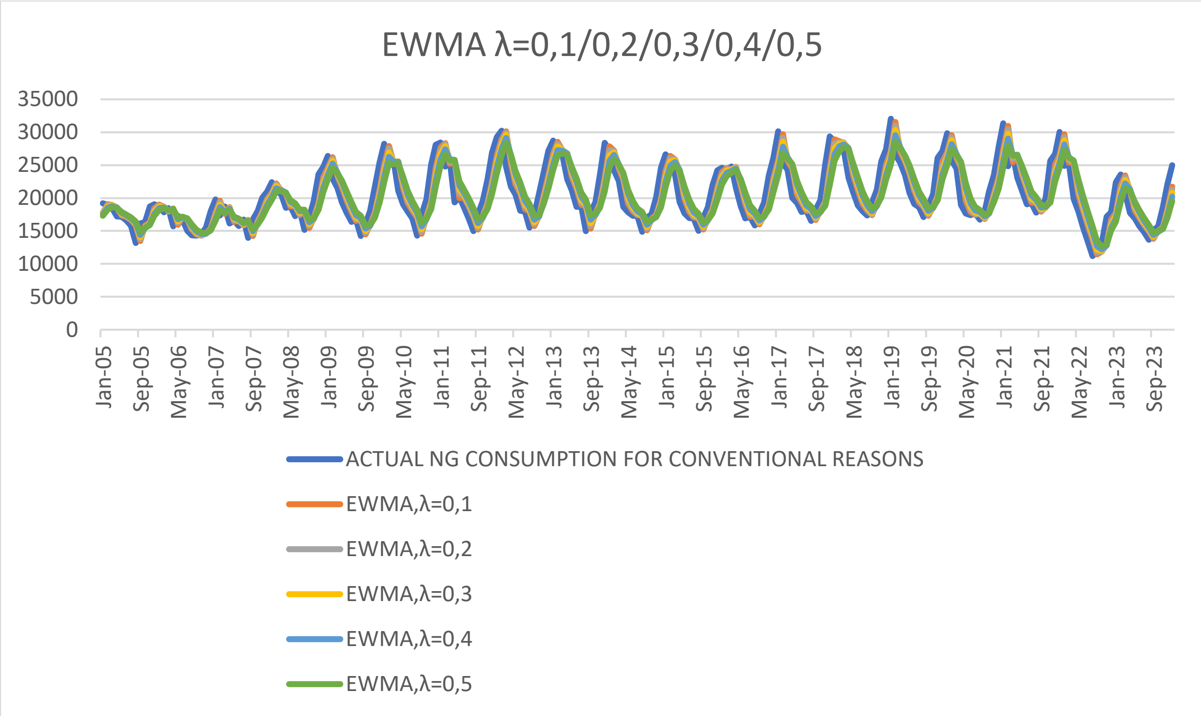


Figure 28:EWMA λ=0,1/0,2/0,3/0,4/0,5

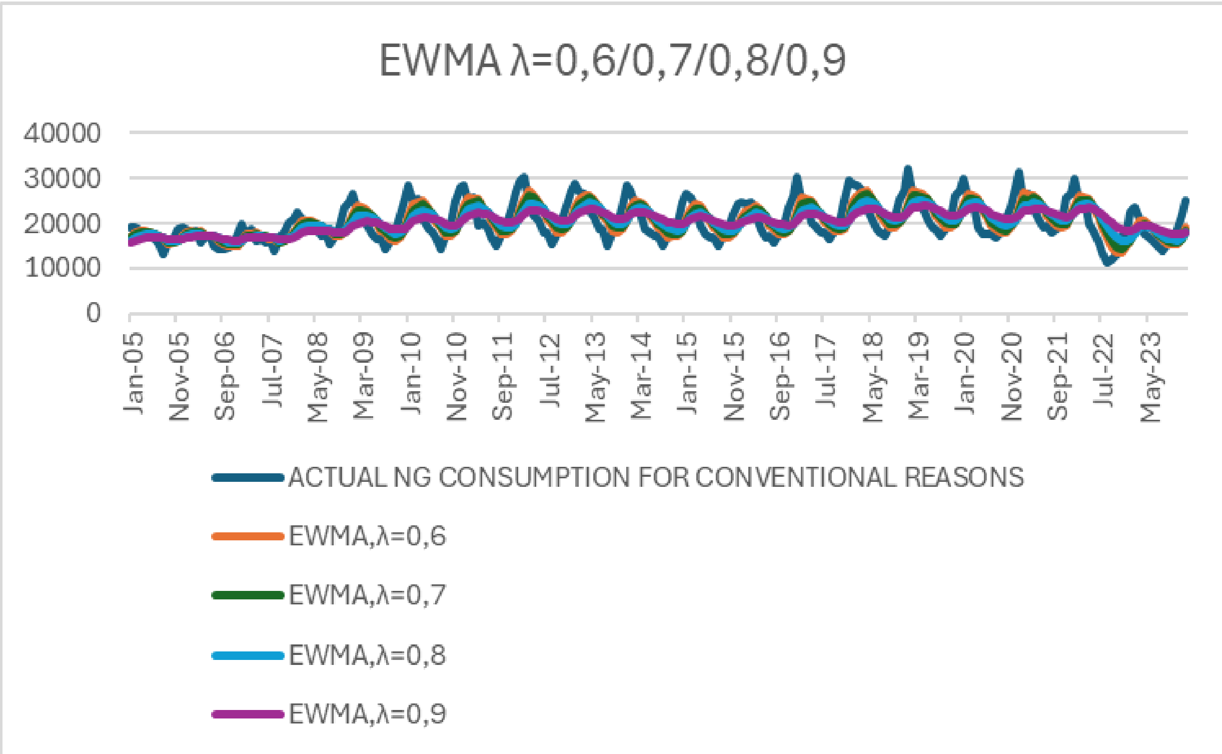


Figure 29:EWMA λ=0,6/0,7/0,8/0,9

The previous graphs offer significant insights into the models' performance. This model demonstrates quick adjustments to sudden variations in the time series data. It gives the most recent observations in the ng consumption data series priority over the older ones.When using EWMA the forecasting error is low and this forecasting model follows the patterns and fluctuations of the actual data of the the ng consumption.We can observe that the smaller the λ,the more accurate forecasts we have. .When λ has a high value,the forecast results are smoother and it’s more difficult to capture ng consumption variations. MAPE is low (from 10,54 % to 16,92 %) .

4.4.2)EWMA FOR ELECTRICITY GENERATION

If we use the monthly data of the ng consumption for electricity generation (241 observations) we have the following results:

NG CONSUMPTION FOR ELECTRICITY									
λ=	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
MAPE	0,170745	0,173011	0,178271	0,185175	0,193893	0,204736	0,219887	0,237273	0,270796
MAE	1326,031	1338,03	1371,309	1418,184	1476,087	1545,036	1645,499	1775,929	2048,453

Table 6:MAPE/MAE for electricity generation

We will also design graphs in order to see the goodness of fit model EWMA.

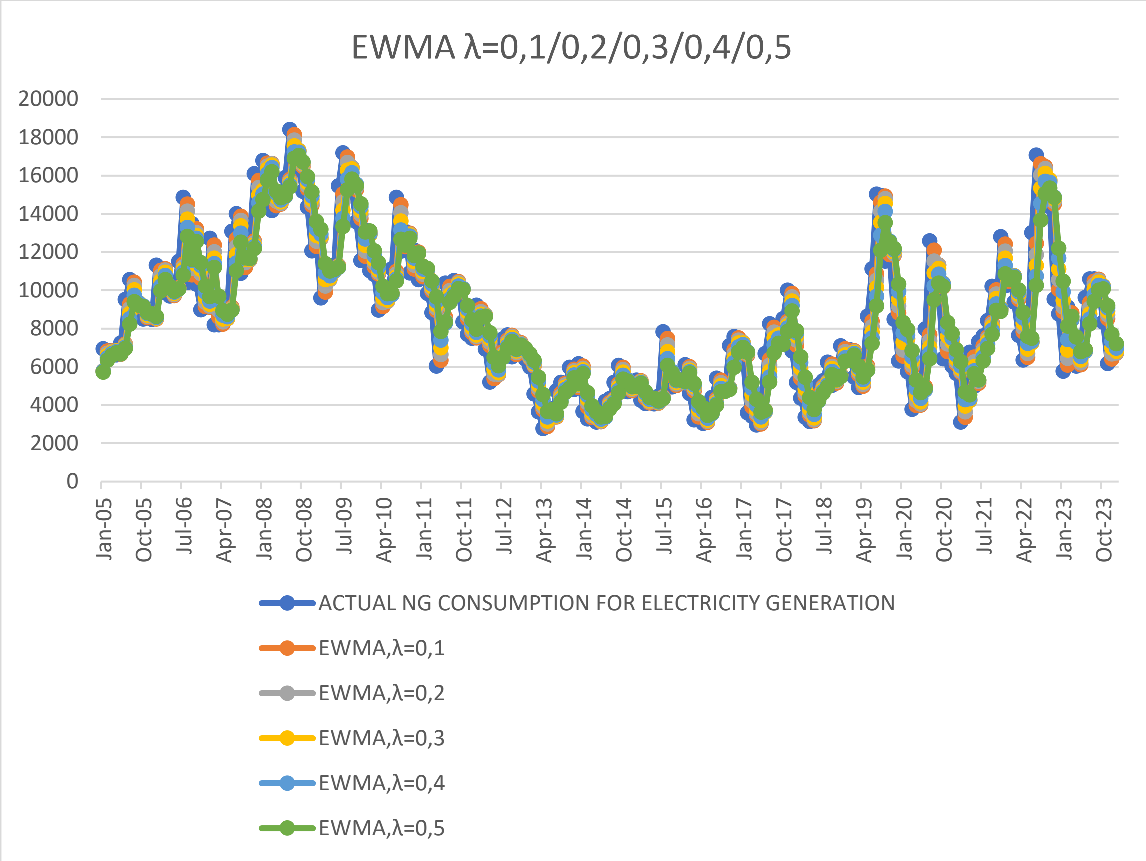


Figure 30:EWMA λ=0,1/0,2/0,3/0,4/0,5

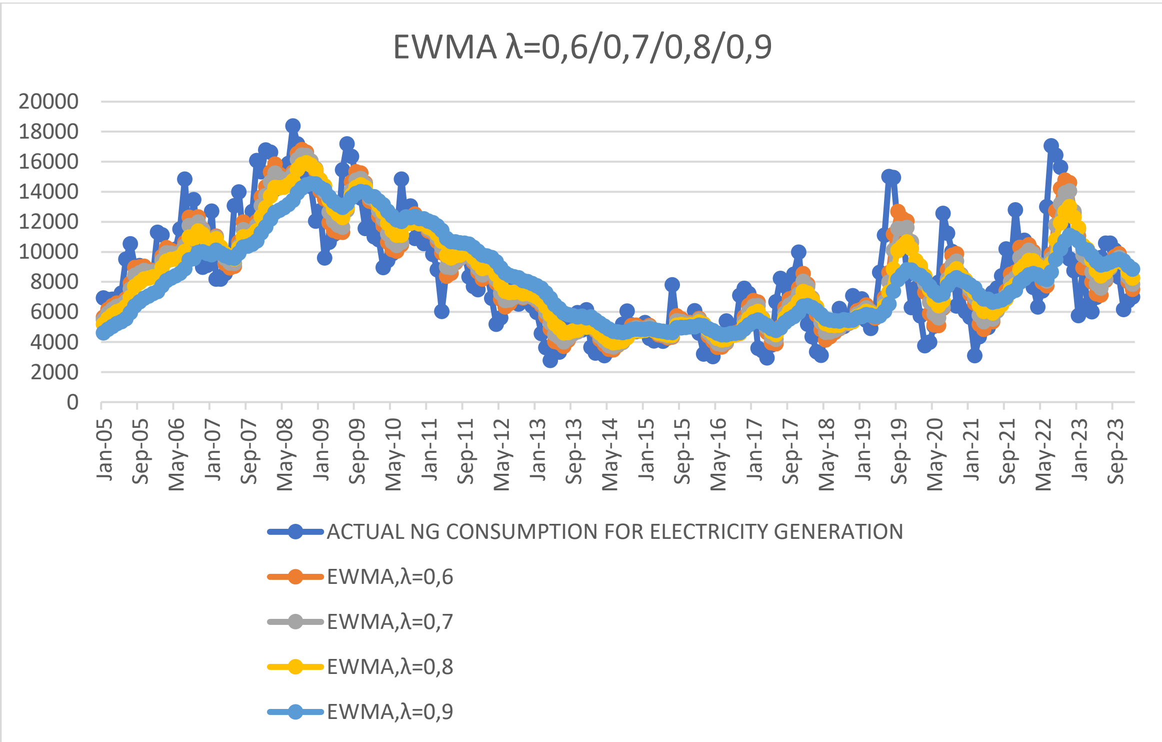


Figure 31:EWMA λ=0,6/0,7/0,8/0,9

The above figures provide important information on how well the models work. This model shows how to quickly adapt to sudden alterations in the time series data. In the ng consumption data series, it prioritizes the most recent observations over the earlier ones.The forecasting error is low when employing EWMA, and the model tracks the trends and variations found in the real data about ng consumption.We can see that our forecasts are more accurate the smaller the λ.When λ has a high value,the forecast results are smoother and it's more difficult to capture ng consumption variations. Although MAPE is 27,07 % for λ=0,9,in general the values of MAPE are low.

4.4.3)EWMA FOR LNG

When we use the monthly data of the ng consumption for LNG (241 observations) we obtain the following results:

NG CONSUMPTION FOR LNG									
λ=	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
MAPE	0,075233	0,072763	0,070852	0,069757	0,0698	0,070936	0,073351	0,078819	0,092629
MAE	63,13292	61,13834	59,62137	58,76485	58,86379	59,86688	61,79632	66,30672	78,51133

Table 7:MAPE/MAE for LNG

We will also design graphs in order to see the goodness of fit model EWMA.

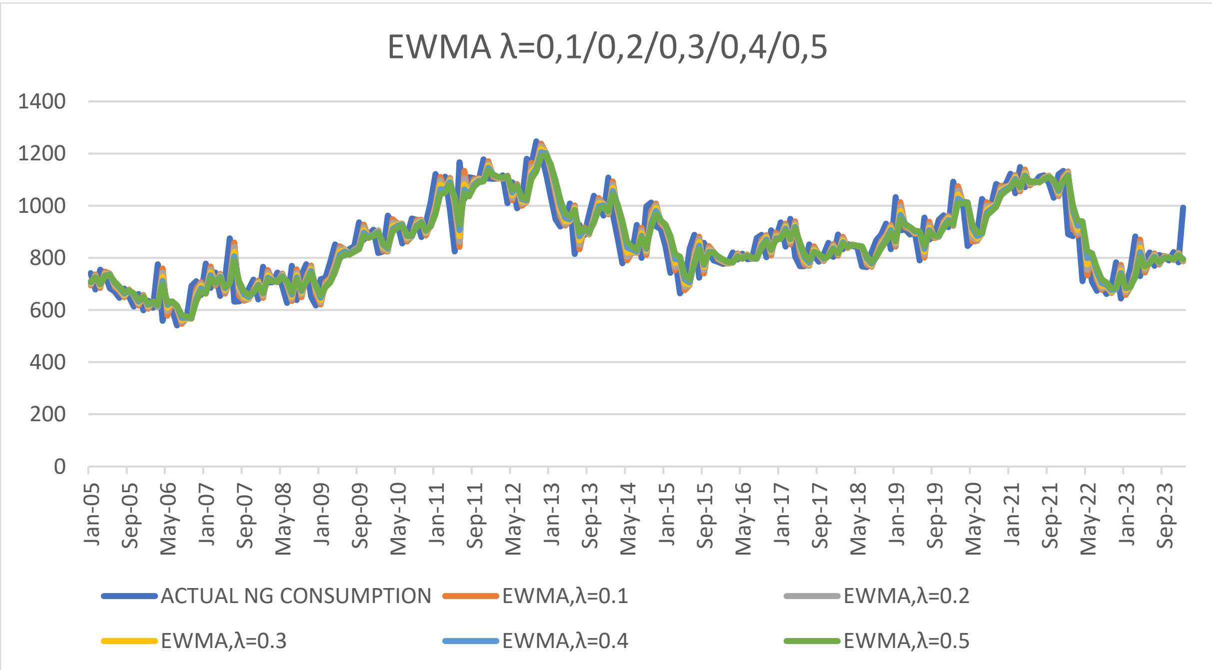


Figure 32:EWMA λ=0,1/0,2/0,3/0,4/0,5

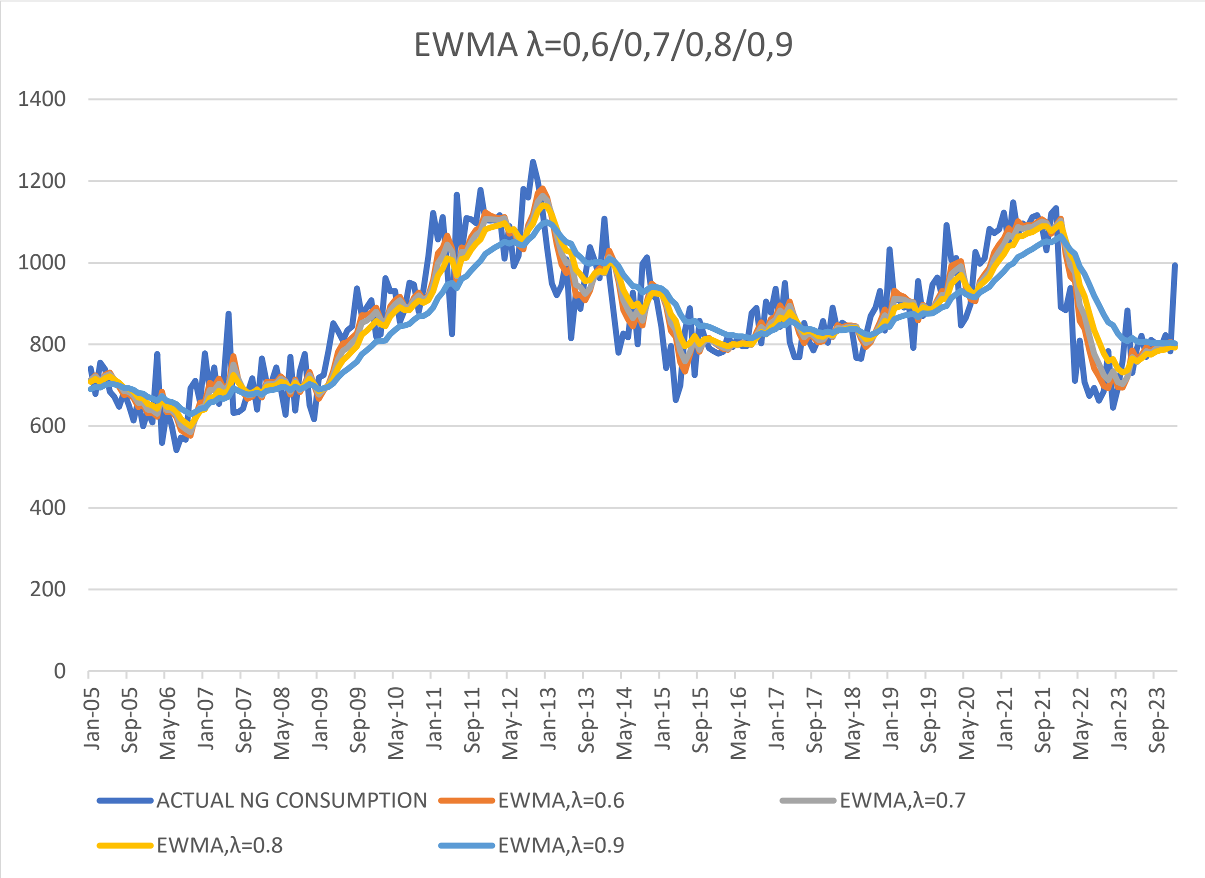


Figure 33:EWMA λ=0,6/0,7/0,8/0,9

As we already saw with the observations of the other 2 categories ,this model shows how to quickly adapt to abrupt changes in the time series data. In the ng consumption data series, it prioritizes the most recent observations over the earlier ones.The forecasting error is minimal when employing EWMA, and the model tracks the trends and variations found in the real data about ng consumption.We can see that our forecasts are more accurate the smaller the λ.A large value for λ makes the forecast results smoother and makes it harder to catch changes in consumption. MAPE doesn’t exceed 10 %.

4.4)REGRESSION ANALYSIS

Regression analysis, according to Moon, is a statistical technique that is helpful when demand is being influenced by a measuring component other than time. (Michele S. Garver, Carlo D. Smith, John T. Mentzer, and Mark A. Moon, 2018).The regression analysis results for the previously mentioned data sets are shown in this chapter. Using Excel's LN function, the sales data are transformed to the natural logarithm (LN.) to reduce heteroscedasticity and deflate exponential increases.

Using the regression analysis, we will use Table 6, the monthly dummy variables, to quantify the seasonality. If the observation is relevant to current month, this variable receives a value of 1, else it receives a value of 0. As per normal, to prevent the negative impacts of multicollinearity, one variable—in this case, the M12 – December—is removed. We will use 241 observations for each category, as we did with the forecasting models in the previous chapters.For the variables Yt-1,2,3 for the first months since we don't have the observations of the previous months ,our observations start from January of 2004) we will assume the observations are equal with the actual observations .We will perform regression analysis in order to build our model.

month	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
1	1	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0
10	0	0	0	0	0	0	0	0	0	1	0
11	0	0	0	0	0	0	0	0	0	0	1
12	0	0	0	0	0	0	0	0	0	0	0

Table 8: Monthly Dummy Variables

4.4.1)Natural gas consumption for conventional reasons

The model that will be examined for all categories has the below formula:

$= \beta o + \beta 1M1 + \beta 2M2 + \beta 3M3 + \beta 4M4 + \beta 5M5 + \beta 6M6 + \beta 7M7$

$+ \beta 8M8 + \beta 9M9 + \beta 10M10 + \beta 11M11$+Yt-1*a1+Yt-2*a2+Yt-3*a3+γτ+et

where:

βo is the intercept

Mi is the month index, for i = 1,2,3,4,5,6,7,8,9,10,11

Yt-1 is the ng consumption of one period ago

a1 is the coefficient of the Yt-1 variable

Yt-2 is the ng consumption of two periods ago

a2 is the coefficient of the Yt-2 variable

Yt-3 is the ng consumption of three periods ago, the lagged variables

a3 is the coefficient of the Yt-3 variable

Global Price of Ng is the global price of natural gas

A4 is the coefficient of Global Price of ng

γt is the coefficient for the linear trend, the slope..

If we run the regressions analysis the results are :

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,965212574
R Square	0,931635313
Adjusted F	0,926752121
Standard I	0,05734246
Observati	241

ANOVA

	df	SS	MS	F	Significance F
Regressio	16	10,03725	0,627328	190,7840875	1,0869E-120
Residual	224	0,736547	0,003288		
Total	240	10,7738			

	Coefficients	andard Err	t Stat	P-value	Lower 95%	Jpper 95%	ower 95,0%	pper 95,0%
Intercept	1,664997245	0,31843	5,228776	3,9083E-07	1,037496326	2,292498	1,037496	2,292498
linear trer	0,000137941	6,43E-05	2,146512	**0,032906908**	1,13039E-05	0,000265	1,13E-05	0,000265
M1	-0,032455625	0,019246	-1,6864	0,093111893	-0,070381138	0,00547	-0,07038	0,00547
M2	-0,177396784	0,020434	-8,68159	8,17994E-16	-0,217663607	-0,13713	-0,21766	-0,13713
M3	-0,133663926	0,025728	-5,1953	4,5907E-07	-0,184363508	-0,08296	-0,18436	-0,08296
M4	-0,255822808	0,02635	-9,70853	7,99788E-19	-0,307749018	-0,2039	-0,30775	-0,2039
M5	-0,187628563	0,02963	-6,33248	1,30366E-09	-0,246016913	-0,12924	-0,24602	-0,12924
M6	-0,199078314	0,028777	-6,91796	4,74664E-11	-0,255786655	-0,14237	-0,25579	-0,14237
M7	-0,165709813	0,026236	-6,31607	1,42658E-09	-0,217411166	-0,11401	-0,21741	-0,11401
M8	-0,273494176	0,024817	-11,0205	7,55599E-23	-0,322398557	-0,22459	-0,3224	-0,22459
M9	-0,054560983	0,028604	-1,90747	0,057738559	-0,110927972	0,001806	-0,11093	0,001806
M10	-0,043662132	0,023558	-1,85336	0,065145602	-0,090086474	0,002762	-0,09009	0,002762
M11	0,03292741	0,01966	1,674815	0,095365956	-0,005815437	0,07167	-0,00582	0,07167
Global Pri	-0,00159523	0,000494	-3,22632	**0,001441489**	-0,002569584	-0,00062	-0,00257	-0,00062
Yt-1	0,767136526	0,065428	11,72495	4,53181E-25	0,638203967	0,896069	0,638204	0,896069
Yt-2	0,171104205	0,081816	2,091326	**0,037625925**	0,00987646	0,332332	0,009876	0,332332
Yt-3	-0,093953545	0,063368	-1,48266	0,139569097	-0,218827304	0,03092	-0,21883	0,03092

Table 9:Regression Analysis for conventional reasons

So, the regression model would be:

Log(s_t)= 1,6649 + 0.00013*γ-0,0324*β1-0,177*β2-0,1336*β3-0,2558*β4-0,187*β5-0,199*β6-0,165*β7-0,273*β8-0,05*β9-0,043*β10+0,032*β11-0,0015*A4+0,767*α1+0,171*α2-0,093*α3

At that point, regression statistics generate an R Square at 0.9316, implying that the model incorporating time trend and seasonality alone is responsible for approximately 93% of ng consumption variation (Albright, Winston, & Zappe, 2011, p. 557).

Now we will use this equation in order to forecast the natural gas consumption for conventional reasons for 2004-2024.At first we will calculate the forecasting error. The forecasting error et is calculated as: $t \square Dt - Ft$, where Dt are the actual consumption and Ft are the forecasted ones.By using excel we calculate MAE=919,59 . Mape has a low rate 4,49 % indicating that this forecasting model has a high accuracy since is below 10 % and close to 0.

From the above results I created the following two graphs. To represent the first graph, I chose the actual ng consumption for conventional reasons and estimated ng consumption of all the 241 observations and from the Excel- 2016 application I chose from the ribbon the line charts(From the tool analysis I also obtained residuals). I ended up the following graph:

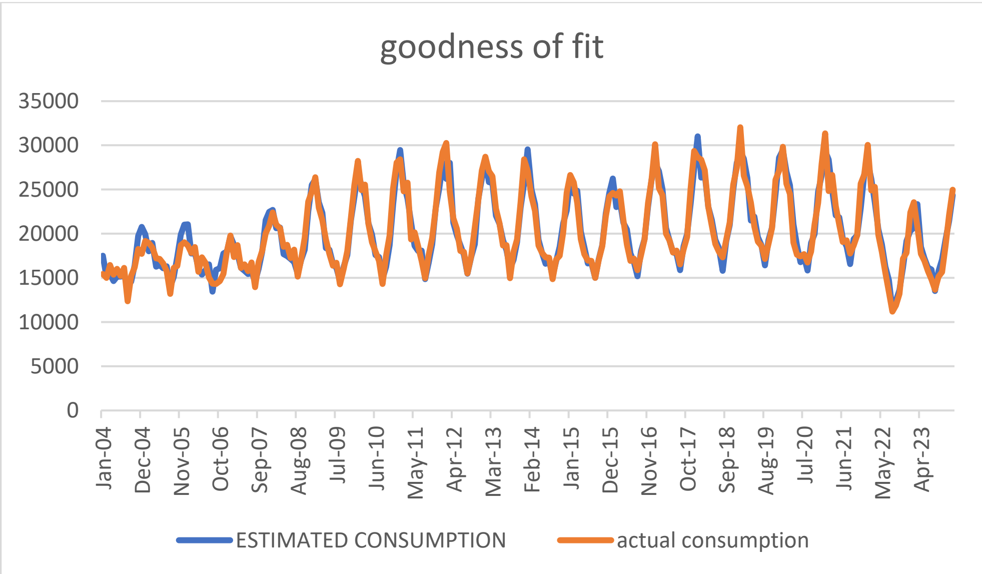


Figure 34:Goodness of Fit -Conventional reasons

We can observe that until 2005 and on 2017 we have an underestimation of the actual consumption. There are also periods that the we have overestimated the actual consumption.In general the model follows the patterns of the actual consumption .And it is almost obviously that there is seasonality (as we can see the pattern remains the same).

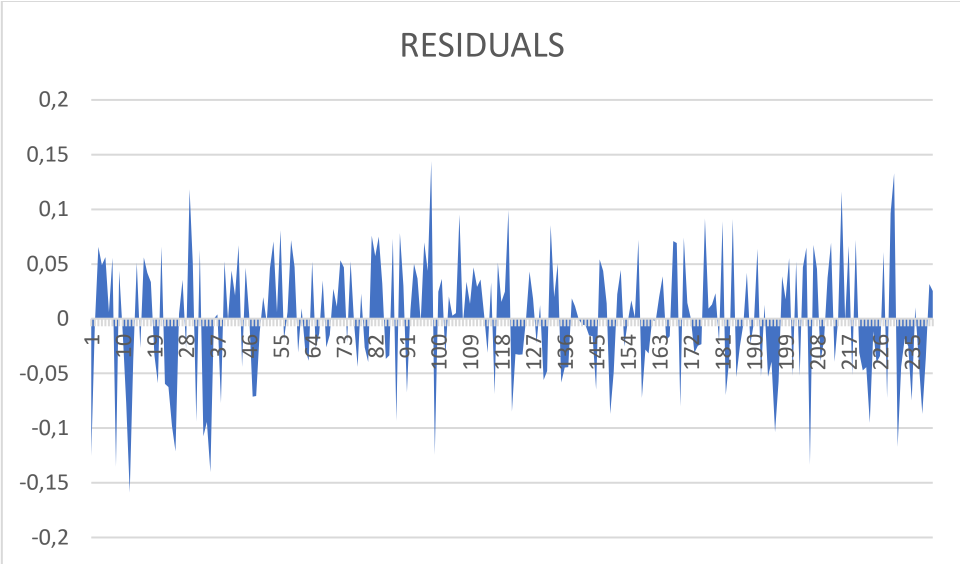


Figure 35:Residuals

We can see that there is no specific pattern and sometimes residuals have negative values and sometimes they remain positive.

4.4.2) Natural Gas consumption for Electricity Generation.

If we run the regressions analysis the results are :

Regression Statistics	
Multiple R	0,927722
R Square	0,860669
Adjusted I	0,850717
Standard	0,176086
Observati	241

ANOVA

	df	*SS*	*MS*	*F*	*gnificance F*
Regressio	16	42,90269	2,681418	86,48016	2,69E-86
Residual	224	6,945381	0,031006		
Total	240	49,84807			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95,0%*	*pper 95,0%*
Intercept	0,854491	0,281534	3,035131	**0,002689**	0,299698	1,409284	0,299698	1,409284
linear trer	-0,0003	0,000182	-1,63732	0,102968	-0,00066	6,08E-05	-0,00066	6,08E-05
M1	0,028893	0,055107	0,524314	0,600579	-0,0797	0,137487	-0,0797	0,137487
M2	-0,12319	0,055823	-2,20686	**0,028339**	-0,2332	-0,01319	-0,2332	-0,01319
M3	-0,08301	0,056716	-1,46367	0,144687	-0,19478	0,028752	-0,19478	0,028752
M4	-0,0439	0,057163	-0,76801	0,44329	-0,15655	0,068744	-0,15655	0,068744
M5	0,143817	0,056888	2,528051	**0,012159**	0,031712	0,255921	0,031712	0,255921
M6	0,289617	0,057627	5,025746	1,03E-06	0,176058	0,403177	0,176058	0,403177
M7	0,281236	0,060464	4,651261	5,64E-06	0,162084	0,400387	0,162084	0,400387
M8	0,053352	0,060903	0,876027	0,381954	-0,06666	0,173368	-0,06666	0,173368
M9	0,057726	0,057397	1,005746	0,315623	-0,05538	0,170833	-0,05538	0,170833
M10	0,025488	0,055801	0,456766	0,648282	-0,08447	0,135449	-0,08447	0,135449
M11	0,02553	0,055746	0,457966	0,647421	-0,08432	0,135383	-0,08432	0,135383
Global Pri	0,00163	0,001588	1,026252	0,30588	-0,0015	0,004759	-0,0015	0,004759
Yt-1	0,731292	0,067091	10,90001	1,8E-22	0,599082	0,863503	0,599082	0,863503
Yt-2	0,123193	0,082019	1,502001	0,134505	-0,03844	0,284822	-0,03844	0,284822
Yt-3	0,046187	0,065799	0,70194	0,483445	-0,08348	0,17585	-0,08348	0,17585

Table 10:Regression Analysis for electricity generation

The model incorporating time trend and seasonality alone is responsible for approximately 86,06% of ng consumption variation.

Log(s_(t))=0,8544-0,0003*γ+0,0288*β1-0,123*β2-0,083*β3-0,043*β4+0,143*β5+0,289*β6+0,281*β7+0,053*β8+0,057*β9+0,002*β10+0,025*β11+0,001 * A4+0,731*α1+0,123*α2+0,046*α3

Now we will use this equation in order to forecast the natural gas consumption for electricity generation for the period 2004-2024.By using excel we calculate MAE= 1016,792936 which is . MAPE is 13,32 % which is good since is near to 10 %.

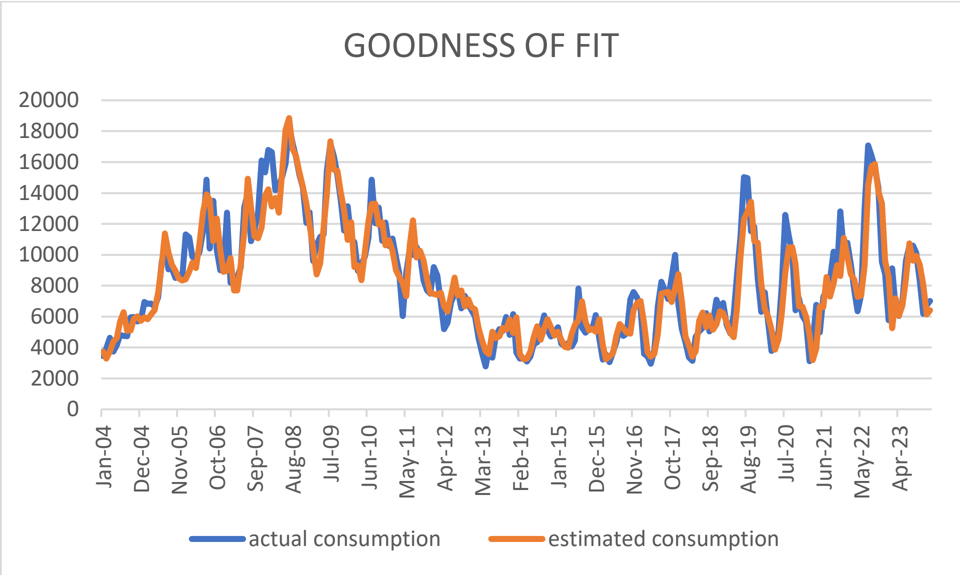


Figure 36:Goodness of Fit -electricity generation

There are both overestimations and underestimations when using this forecasting model.For example on June 2010 we can observe an underestimation of the actual ng consumption. Also on November of 2005 there is an overestimation .After 2014 there are mostly underestimations.

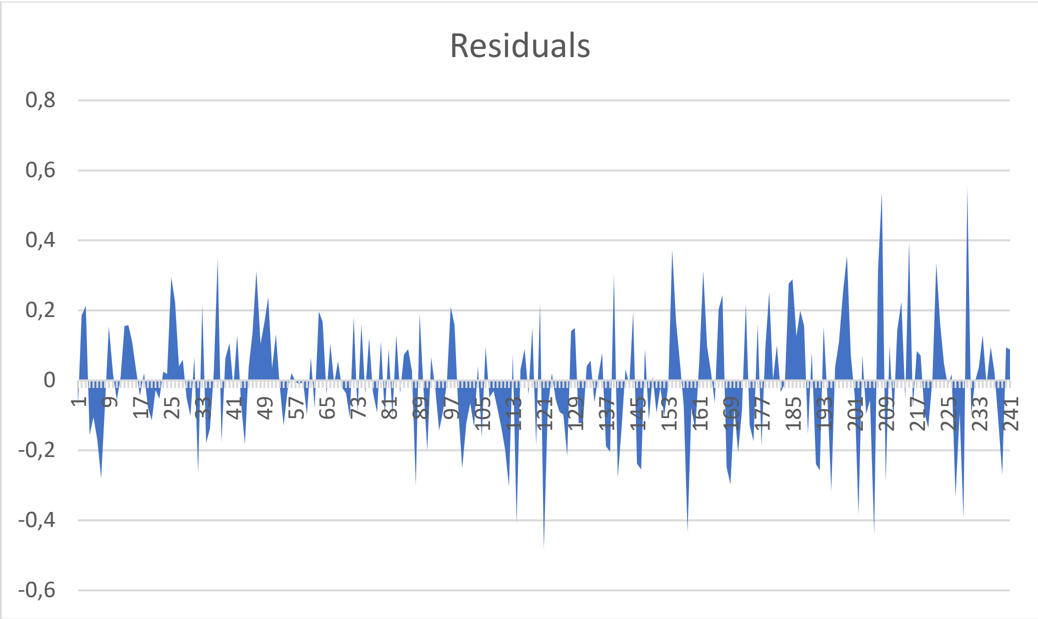


Figure 37:Residuals

There is no specific patterns and sometimes residuals have negative values and sometimes they remain positive.

4.4.3) Natural Gas consumption for LNG

If we run the regressions analysis the results are :

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,900682
R Square	0,811227
Adjusted R Square	0,797744
Standard Error	0,081207
Observations	241

ANOVA

	df	SS	MS	F	gnificance F
Regression	16	6,348054	0,396753	60,16334	1,06E-71
Residual	224	1,477191	0,006595		
Total	240	7,825245			

	Coefficients	andard Err	t Stat	P-value	ower 95%	Jpper 95%	ower 95,0%	pper 95,0%
Intercept	0,596867	0,230961	2,584271	0,010393	0,141732	1,052002	0,141732	1,052002
LINEAR TREND	0,000135	8,94E-05	1,506303	0,133398	-4,2E-05	0,000311	-4,2E-05	0,000311
M1	0,073484	0,025722	2,856872	0,004681	0,022796	0,124172	0,022796	0,124172
M2	-0,01159	0,025956	-0,44647	0,655688	-0,06274	0,03956	-0,06274	0,03956
M3	0,07712	0,026083	2,956671	0,003443	0,02572	0,12852	0,02572	0,12852
M4	-0,05003	0,026046	-1,92103	0,055997	-0,10136	0,001291	-0,10136	0,001291
M5	0,026739	0,026619	1,004507	0,316218	-0,02572	0,079196	-0,02572	0,079196
M6	0,003334	0,026614	0,125259	0,900431	-0,04911	0,055779	-0,04911	0,055779
M7	0,045007	0,025838	1,741921	0,082895	-0,00591	0,095923	-0,00591	0,095923
M8	0,028364	0,025779	1,100297	0,272384	-0,02244	0,079164	-0,02244	0,079164
M9	0,067652	0,025788	2,623415	0,009303	0,016834	0,11847	0,016834	0,11847
M10	0,057263	0,025755	2,223384	0,027189	0,00651	0,108016	0,00651	0,108016
M11	0,067194	0,025735	2,610935	0,009639	0,016479	0,117908	0,016479	0,117908
Global Price of NG	-0,00196	0,000688	-2,84473	0,004857	-0,00331	-0,0006	-0,00331	-0,0006
Yt-1	0,480226	0,064697	7,422687	2,37E-12	0,352734	0,607719	0,352734	0,607719
Yt-2	0,156935	0,0716	2,191843	0,029421	0,01584	0,29803	0,01584	0,29803
Yt-3	0,270214	0,064829	4,168095	4,39E-05	0,142461	0,397967	0,142461	0,397967

Table 11:Regression Analysis for LNG

The model incorporating time trend and seasonality alone is responsible for approximately 81,12 % of ng consumption variation.

Log(s_(t))=0,5968+0,0001*γ+0,073*β1-0,011*β2+0,077*β3-0,050*β4+0,026*β5+0,003*β6+0,045*β7+0,028*β8+0,067*β9+0,057*β10+0,067*β11-0,001* A4+0,48*a1+0,156*a2+0,027*a3

Now we will use this equation in order to forecast the natural gas consumption for LNG reasons for the period 2004-2024.The formula for calculating MAE is the average of the absolute calculated different between actual and forecasted ng consumption . So, by using excel we calculate MAE= 50,8596507**.** Mape has a low rate 6,06 % indicating that this forecasting model has a high accuracy since it below 10 % and can be considered as low.

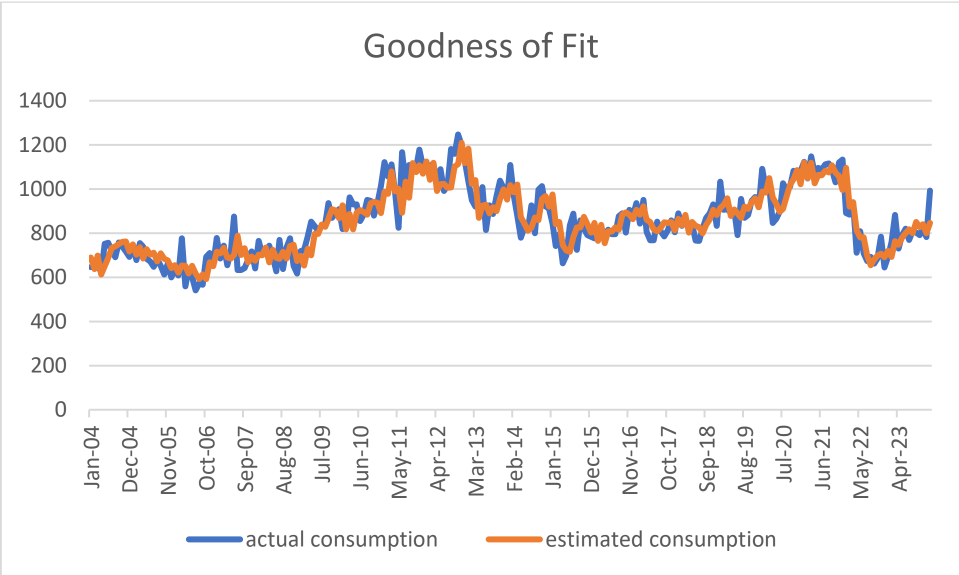


Figure 38:Goodness of Fit

Estimated consumption seems to follow the patterns of the actual ng consumption. Based on Figure 40 and MAPE (6,06 %) we can assume that this forecasting model has high accuracy.

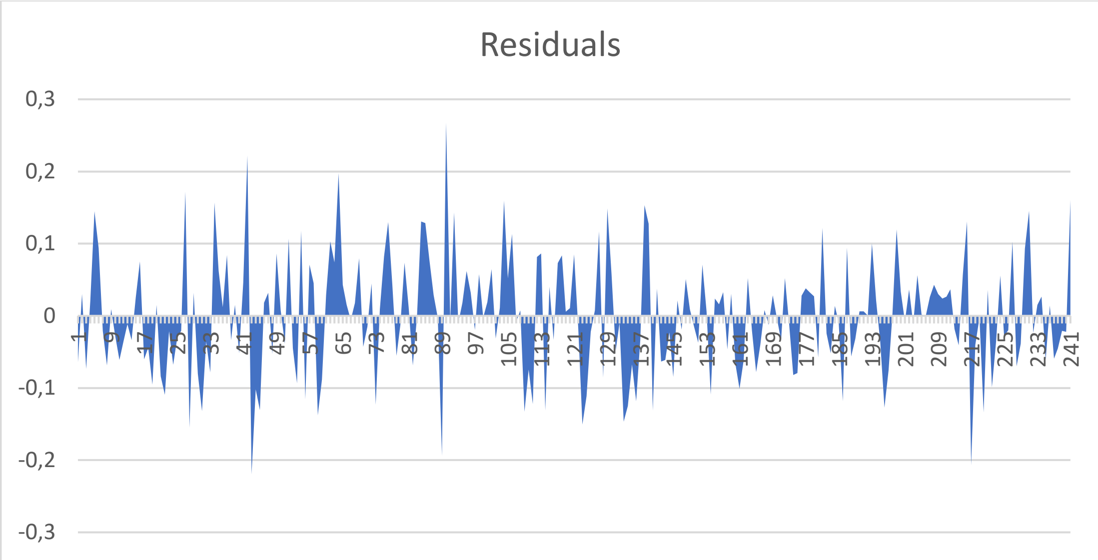


Figure 39: Residuals

We observe a random pattern since residuals go up and down randomly and there are residuals with both positive and negative values.

4.5)Covid Effect

This section makes an effort to determine how the pandemic affected the ng consumption in Spain .First of all we will exclude from our regression analysis the Variable ‘’Global price of NG’’. In order to investigate the impact of COVID-19 on ng consumption , a new dummy variable was added to the model equation. The covid-19 dummy variable separates the times when Spain was under lockdown from the time when pandemic regulations were in place.January 2004 to February 2020 is not considered a COVID-19 timeframe because the first Covid-19 case in Spain was confirmed on February 2020. The nation experienced its first shutdown during observations 195–199, which ran from March 2020 to July 2020. Although there were still some measures in some cities after July 2020,the national lockdown took place between March-July 2020. The aforementioned times are regarded as lockdown months, and the accompanying observations, which range from 195 to 199, have a value of 1 for the dummy variable.Observations 1-194 & 200-241 will take the value 0. The corresponding hypothesis test is:

Null hypothesis H0: **Covid 19 effect** = 0 (the natural gas consumption for conventional reasons is not affected by the pandemic)

Alternative hypothesis H1: **Covid 19 effect** ≠ 0 (the natural gas consumption for conventional reasons is affected by the pandemic)

Below we can see the regression analysis with the new Covid-19 effect variable :

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,963595
R Square	0,928515
Adjusted R Square	0,923409
Standard Error	0,058636
Observations	241

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	16	10,00364	0,625227	181,8462	1,6E-118
Residual	224	0,770161	0,003438		
Total	240	10,7738			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	1,446504	0,318043	4,548146	8,86E-06	0,819766	2,073242	0,819766	2,073242
linear trend	6,73E-05	6,2E-05	1,084956	0,279107	-5,5E-05	0,00019	-5,5E-05	0,00019
M1	-0,03194	0,01968	-1,62307	0,105981	-0,07072	0,00684	-0,07072	0,00684
M2	-0,17678	0,020895	-8,46032	3,48E-15	-0,21795	-0,1356	-0,21795	-0,1356
M3	-0,12821	0,026254	-4,88357	1,98E-06	-0,17995	-0,07648	-0,17995	-0,07648
M4	-0,24799	0,026828	-9,24351	1,92E-17	-0,30085	-0,19512	-0,30085	-0,19512
M5	-0,17446	0,030005	-5,81434	2,08E-08	-0,23358	-0,11533	-0,23358	-0,11533
M6	-0,18535	0,029101	-6,36912	1,07E-09	-0,24269	-0,128	-0,24269	-0,128
M7	-0,15304	0,02652	-5,77057	2,61E-08	-0,2053	-0,10078	-0,2053	-0,10078
M8	-0,26348	0,025217	-10,4487	4,5E-21	-0,31317	-0,21379	-0,31317	-0,21379
M9	-0,04029	0,028935	-1,39252	0,165145	-0,09731	0,016727	-0,09731	0,016727
M10	-0,03289	0,023869	-1,37806	0,169559	-0,07993	0,014144	-0,07993	0,014144
M11	0,03868	0,02002	1,932113	0,054607	-0,00077	0,078131	-0,00077	0,078131
Yt-1	0,803759	0,065972	12,18331	1,56E-26	0,673753	0,933764	0,673753	0,933764
Yt-2	0,167967	0,083663	2,007653	0,045881	0,003099	0,332835	0,003099	0,332835
Yt-3	-0,10686	0,06486	-1,64763	0,100831	-0,23468	0,020948	-0,23468	0,020948
Covid -19 effect	-0,0116	0,027469	-0,42226	0,67324	-0,06573	0,042532	-0,06573	0,042532

Table 12:Covid effect for conventional reasons

The above-mentioned regression model can explain the 92,85 % of variance of the ng consumption for conventional reasons .P value of dummy variable Covid-19 Effect is 67,32% >5 % which means that we don't have enough evidence to reject the null hypothesis in a 95 % significance level test. Thus, we assume that the covid effect is statically insignificant .

I will also add the dummy variable for the covid effect and i will run the regression analysis again for the electricity generation category. The regression analysis output for the natural gas consumption for electricity generation is :

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,927683289
R Square	0,860596285
Adjusted R Squa	0,850638876
Standard Error	0,176131625
Observations	241

ANOVA

	df	SS	MS	F	gnificance F
Regression	16	42,89906	2,681192	86,42774	2,85E-86
Residual	224	6,949006	0,031022		
Total	240	49,84807			

	Coefficients	andard Err	t Stat	P-value	Lower 95%	Upper 95%	ower 95,0%	pper 95,0%
Intercept	0,748453975	0,264762	2,826891	0,005126	0,226711	1,270197	0,226711	1,270197
linear trend	-0,000248704	0,000169	-1,47064	0,142792	-0,00058	8,46E-05	-0,00058	8,46E-05
M1	0,0278541	0,055112	0,505407	0,61377	-0,08075	0,136459	-0,08075	0,136459
M2	-0,124386797	0,055828	-2,22802	0,026873	-0,2344	-0,01437	-0,2344	-0,01437
M3	-0,086036444	0,056873	-1,51278	0,131746	-0,19811	0,026038	-0,19811	0,026038
M4	-0,046450834	0,057304	-0,8106	0,418454	-0,15937	0,066473	-0,15937	0,066473
M5	0,141296671	0,057021	2,477975	0,013952	0,02893	0,253663	0,02893	0,253663
M6	0,286493519	0,057764	4,95975	1,39E-06	0,172664	0,400323	0,172664	0,400323
M7	0,277002865	0,060599	4,571086	8,02E-06	0,157586	0,39642	0,157586	0,39642
M8	0,052464544	0,060911	0,861329	0,389978	-0,06757	0,172497	-0,06757	0,172497
M9	0,056306444	0,057396	0,981014	0,327644	-0,0568	0,169412	-0,0568	0,169412
M10	0,022813396	0,05576	0,409135	0,682832	-0,08707	0,132695	-0,08707	0,132695
M11	0,023833441	0,055737	0,427605	0,66935	-0,086	0,133669	-0,086	0,133669
Yt-1	0,740551349	0,066535	11,13026	3,42E-23	0,609437	0,871666	0,609437	0,871666
Yt-2	0,127332944	0,081953	1,553735	0,121659	-0,03416	0,28883	-0,03416	0,28883
Yt-3	0,04586935	0,065813	0,69697	0,486544	-0,08382	0,17556	-0,08382	0,17556
Covid -19 effect	0,079345959	0,082023	0,967356	0,334409	-0,08229	0,240982	-0,08229	0,240982

Table 13:Covid effect for electricity generation

The above-mentioned regression model can explain 86,05 % of variance of the ng consumption for electricity generation .P value of dummy variable Covid-19 Effect is 33,4 % >5 % which means that we don't have enough evidence to reject the null hypothesis in a 95 % significance level test. Thus, we assume that the covid effect is statically insignificant .

The regression analysis output for the natural gas consumption for LNG is :

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,896892
R Square	0,804415
Adjusted R Squar	0,790444
Standard Error	0,08266
Observations	241

ANOVA

	df	SS	MS	F	gnificance F
Regression	16	6,294742	0,393421	57,58004	5,28E-70
Residual	224	1,530502	0,006833		
Total	240	7,825245			

	Coefficients	andard Err	t Stat	P-value	ower 95%	Jpper 95%	ower 95,0%	pper 95,0%
Intercept	0,515548	0,233608	2,206895	0,028337	0,055198	0,975897	0,055198	0,975897
LINEAR TREND	5,42E-05	8,69E-05	0,623604	0,533522	-0,00012	0,000226	-0,00012	0,000226
M1	0,077317	0,026146	2,957133	0,003438	0,025794	0,128841	0,025794	0,128841
M2	-0,00766	0,026383	-0,29044	0,77175	-0,05965	0,044328	-0,05965	0,044328
M3	0,080263	0,026598	3,017585	0,002843	0,027848	0,132678	0,027848	0,132678
M4	-0,04523	0,026536	-1,70446	0,089681	-0,09752	0,007063	-0,09752	0,007063
M5	0,032708	0,027077	1,207943	0,228343	-0,02065	0,086067	-0,02065	0,086067
M6	0,010495	0,02704	0,388135	0,698285	-0,04279	0,06378	-0,04279	0,06378
M7	0,047937	0,026359	1,818625	0,070304	-0,00401	0,099881	-0,00401	0,099881
M8	0,029222	0,026238	1,113725	0,266591	-0,02248	0,080927	-0,02248	0,080927
M9	0,068809	0,026246	2,621693	0,009349	0,017088	0,120529	0,017088	0,120529
M10	0,059956	0,026198	2,288546	0,023038	0,008329	0,111582	0,008329	0,111582
M11	0,068363	0,026193	2,610024	0,009664	0,016748	0,119978	0,016748	0,119978
Yt-1	0,503783	0,065344	7,709714	4,08E-13	0,375015	0,63255	0,375015	0,63255
Yt-2	0,16758	0,072798	2,301975	0,022255	0,024123	0,311037	0,024123	0,311037
Yt-3	0,246176	0,065457	3,760895	0,000216	0,117186	0,375165	0,117186	0,375165
Covid -19 effect	-0,00347	0,038535	-0,08997	0,92839	-0,0794	0,07247	-0,0794	0,07247

Table 14:Covid effect for LNG

The above-mentioned regression model can explain 80,44 % of variance of the ng consumption for LNG.P value of dummy variable Covid-19 Effect is 92,83 % >5 % which means that we don't have enough evidence to reject the null hypothesis in a 95 % significance level test. Thus, we assume that the covid effect is statically insignificant .

4.6)War In Ukraine

Now we will examine the effect that the ongoing war in Ukraine has in the natural gas consumption of Spain. The war in Ukraine started on January 2022 and it's not over yet. So for our data from January 2022 until January 2024 we will add a dummy variable with a value 1.For the observations January 2004-December 2021 we will add a dummy variable with a value 0. Again we will exclude from our regression analysis the Variable ''Global price of NG''.We run the regression analysis again and we obtain the following results:

Natural Gas consumption for conventional reasons :

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,96602
R Square	0,93319
Adjusted F	0,92841
Standard E	0,05669
Observatio	241

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressio	16	10,054	0,62837	195,537	8E-122
Residual	224	0,71984	0,00321		
Total	240	10,7738			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95,0%*	*pper 95,0%*
Intercept	2,46541	0,40044	6,15676	3,4E-09	1,6763	3,25452	1,6763	3,25452
linear tren	0,00032	8,8E-05	3,67163	0,0003	0,00015	0,0005	0,00015	0,0005
M1	-0,0188	0,01931	-0,9732	0,33149	-0,0569	0,01926	-0,0569	0,01926
M2	-0,1535	0,02102	-7,3055	4,8E-12	-0,195	-0,1121	-0,195	-0,1121
M3	-0,1069	0,02594	-4,1225	5,3E-05	-0,1581	-0,0558	-0,1581	-0,0558
M4	-0,2257	0,02652	-8,5123	2,5E-15	-0,278	-0,1735	-0,278	-0,1735
M5	-0,1659	0,02907	-5,7075	3,6E-08	-0,2232	-0,1086	-0,2232	-0,1086
M6	-0,1814	0,02813	-6,4475	6,9E-10	-0,2368	-0,1259	-0,2368	-0,1259
M7	-0,1605	0,02571	-6,2446	2,1E-09	-0,2112	-0,1099	-0,2112	-0,1099
M8	-0,2739	0,02447	-11,194	2,2E-23	-0,3222	-0,2257	-0,3222	-0,2257
M9	-0,0622	0,02848	-2,1849	0,02993	-0,1183	-0,0061	-0,1183	-0,0061
M10	-0,0487	0,02339	-2,0804	0,03863	-0,0948	-0,0026	-0,0948	-0,0026
M11	0,02287	0,01976	1,15748	0,24831	-0,0161	0,06182	-0,0161	0,06182
Yt-1	0,74213	0,06557	11,3183	8,8E-24	0,61292	0,87134	0,61292	0,87134
Yt-2	0,16507	0,08088	2,0408	0,04244	0,00568	0,32445	0,00568	0,32445
Yt-3	-0,148	0,06322	-2,3403	0,02015	-0,2725	-0,0234	-0,2725	-0,0234
Ukraine W	-0,0737	0,01852	-3,9812	9,3E-05	-0,1102	-0,0372	-0,1102	-0,0372

Table 15:Ukraine War effect for conventional reasons

The above-mentioned regression model can explain 93,31 % of variance of the ng consumption for conventional reasons.War in Ukraine has a negative effect on the ng consumption for conventional reasons(-0,0737).However P value is more than 5 % so we consider the impact of the war in Ukraine as statistically insignificant.

Natural Gas consumption for electricity generation :

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,92757
R Square	0,86038
Adjusted F	0,85041
Standard E	0,17627
Observati	241

ANOVA

	df	SS	MS	F	gnificance F
Regressioi	16	42,8882	2,68051	86,2713	3,4E-86
Residual	224	6,95985	0,03107		
Total	240	49,8481			

	Coefficients	andard Err	t Stat	P-value	Lower 95%	Upper 95%	ower 95,0%	pper 95,0%
Intercept	0,84115	0,28742	2,92659	0,00378	0,27476	1,40753	0,27476	1,40753
linear tren	-0,00032	0,00021	-1,53013	0,12739	-0,00073	9,1E-05	-0,00073	9,1E-05
M1	0,02517	0,05527	0,45537	0,64929	-0,08375	0,13408	-0,08375	0,13408
M2	-0,12605	0,05592	-2,25408	0,02516	-0,23625	-0,01585	-0,23625	-0,01585
M3	-0,08433	0,05685	-1,48343	0,13937	-0,19635	0,02769	-0,19635	0,02769
M4	-0,04547	0,05733	-0,79305	0,42859	-0,15845	0,06751	-0,15845	0,06751
M5	0,14121	0,05715	2,47079	0,01423	0,02859	0,25384	0,02859	0,25384
M6	0,28622	0,05791	4,94226	1,5E-06	0,1721	0,40035	0,1721	0,40035
M7	0,27745	0,06067	4,57274	8E-06	0,15788	0,39701	0,15788	0,39701
M8	0,05068	0,061	0,83078	0,40698	-0,06953	0,17088	-0,06953	0,17088
M9	0,05606	0,05744	0,97587	0,33018	-0,05714	0,16925	-0,05714	0,16925
M10	0,02329	0,05581	0,41735	0,67683	-0,08668	0,13326	-0,08668	0,13326
M11	0,02398	0,05578	0,42993	0,66766	-0,08594	0,1339	-0,08594	0,1339
Yt-1	0,73839	0,06663	11,0815	4,9E-23	0,60708	0,86969	0,60708	0,86969
Yt-2	0,12457	0,08208	1,51758	0,13053	-0,03719	0,28632	-0,03719	0,28632
Yt-3	0,04113	0,06608	0,62241	0,5343	-0,08908	0,17134	-0,08908	0,17134
Ukraine W	0,03612	0,04722	0,76507	0,44503	-0,05692	0,12917	-0,05692	0,12917

Table 16:Ukraine War effect for electricity generation

The above-mentioned regression model can explain 86,03 % of variance of the ng consumption for electricity generation. P value is more than 5 % so we consider the impact of the war in Ukraine as statistically insignificant.

Natural Gas consumption for LNG:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,89988
R Square	0,80978
Adjusted R Squ	0,79619
Standard Error	0,08152
Observations	241

ANOVA

	df	*SS*	*MS*	*F*	*gnificance F*
Regression	16	6,33671	0,39604	59,5982	2,5E-71
Residual	224	1,48853	0,00665		
Total	240	7,82524			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95,0%*	*pper 95,0%*
Intercept	0,81704	0,25928	3,15119	0,00185	0,3061	1,32799	0,3061	1,32799
LINEAR TREND	0,00024	0,00011	2,12643	0,03456	1,8E-05	0,00046	1,8E-05	0,00046
M1	0,07915	0,0258	3,06847	0,00242	0,02832	0,12998	0,02832	0,12998
M2	-0,00621	0,02603	-0,23865	0,8116	-0,0575	0,04507	-0,0575	0,04507
M3	0,08005	0,02616	3,05986	0,00248	0,0285	0,13161	0,0285	0,13161
M4	-0,04421	0,0261	-1,69396	0,09166	-0,09564	0,00722	-0,09564	0,00722
M5	0,03119	0,02665	1,17049	0,24305	-0,02132	0,08371	-0,02132	0,08371
M6	0,00901	0,02661	0,33867	0,73518	-0,04342	0,06144	-0,04342	0,06144
M7	0,04551	0,02593	1,75487	0,08065	-0,0056	0,09662	-0,0056	0,09662
M8	0,02731	0,02589	1,05504	0,29254	-0,0237	0,07832	-0,0237	0,07832
M9	0,06663	0,0259	2,57276	0,01074	0,01559	0,11766	0,01559	0,11766
M10	0,05865	0,02584	2,26962	0,02418	0,00773	0,10957	0,00773	0,10957
M11	0,06757	0,02583	2,6155	0,00952	0,01666	0,11847	0,01666	0,11847
Yt-1	0,47652	0,06532	7,29501	5,1E-12	0,3478	0,60524	0,3478	0,60524
Yt-2	0,15638	0,07191	2,17455	0,03071	0,01467	0,29809	0,01467	0,29809
Yt-3	0,23748	0,0646	3,67611	0,0003	0,11018	0,36478	0,11018	0,36478
Ukraine War	-0,05817	0,02313	-2,51472	0,01261	-0,10376	-0,01259	-0,10376	-0,01259

Table 17:Ukraine War effect for LNG

The above-mentioned regression model can explain 80,97 % of variance of the ng consumption for LNG. The war in Ukraine has a negative effect on the ng consumption for LNG use (-0,05817) .P value is less than 5 % so we consider the impact of the war in Ukraine as statistically significant. In general Russia is a major global supplier of natural gas to numerous nations, especially those in Europe. Any interruptions to Russia's natural gas supplies may have implications for importing nations such as Spain's energy markets.

5 MAIN FINDINGS AND SUMMARY

The outcomes of the aforementioned study will be covered in this chapter, along with a summary of the results of the analysis conducted for each category. The empirical study produced sufficient data to address the research issues posed in this dissertation's first chapter(abstract). The results will be shown below for each question in order to provide an overview of the findings and make a connection to the thesis objectives.

First of all we will compare the forecasting accuracy of the forecasting models that we used in order to estimate the Natural Gas consumption in Spain for the 3 main categories:

1)Natural Gas consumption for Conventional reasons

2)Natural Gas consumption for Electricity Generation

3)Natural Gas consumption for Liquified Natural Gas (LNG)

CONVENTIONAL REASONS		
	MAE	MAPE
SMA 4	3587,67	17,36%
SMA 6	4217,90	20,72%
SMA 8	4248,05	21,06%
SMA 10	3872,85	19,27%
SMA 12	3350,01	16,71%
λ=0,1	2171,65	10,55%
λ=0,2	2300,04	11,15%
λ=0,3	2462,35	11,93%
λ=0,4	2645,40	12,82%
λ=0,5	2824,70	13,71%
λ=0,6	2992,68	14,57%
λ=0,7	3185,07	15,62%
λ=0,8	3347,10	16,51%
λ=0,9	3410,94	16,92%
Regression Analysis	919,59	4,49%

Table 18:MAE-MAPE RESULTS

As we can see from Table 15,the best forecasting method for the natural gas consumption for conventional purposes in Spain ,during the period Jan 2004-Jan 2024 ,is the forecasting model of regression analysis.In this occasion MAE and MAPE values are the lowest.

ELECTRICITY GENERATION		
	MAE	MAPE
SMA 4	1797,10	23,71%
SMA 6	2000,00	26,99%
SMA 8	2230,54	29,64%
SMA 10	2389,14	27,80%
SMA 12	1941,41	25,74%
λ=0,1	1326,03	17,07%
λ=0,2	1338,03	17,30%
λ=0,3	1371,31	17,83%
λ=0,4	1418,18	18,52%
λ=0,5	1476,09	19,39%
λ=0,6	1545,04	20,47%
λ=0,7	1645,50	21,99%
λ=0,8	1775,93	23,73%
λ=0,9	2048,45	27,08%
Regression Analysis	1016,79	**13,32%**

Table 19:MAE-MAPE RESULTS

As we can see from the Table 16,the best forecasting method for the natural gas consumption for electricity generation purposes in Spain ,during the period Jan 2004-Jan 2024, is the forecasting model of regression analysis .In this occasion MAE and MAPE values are the lowest.

LNG		
	MAE	MAPE
SMA 4	62,42	7,35%
SMA 6	66,45	7,87%
SMA 8	69,69	8,27%
SMA 10	71,99	8,57%
SMA 12	73,83	8,80%
λ=0,1	63,13	7,52%
λ=0,2	61,14	7,28%
λ=0,3	59,62	7,09%
λ=0,4	58,76	6,98%
λ=0,5	58,86	6,98%
λ=0,6	59,87	7,09%
λ=0,7	61,80	7,34%
λ=0,8	66,31	7,88%
λ=0,9	78,51	9,26%
Regression Analysis	50,86	**6,06%**

Table 20:MAE-MAPE RESULTS

As we can see from the Table 17,the best forecasting method for the natural gas consumption for electricity generation purposes in Spain ,during the period Jan 2004-Jan 2024, is the forecasting model of regression analysis .In this occasion MAE and MAPE values are the lowest.

What are the statistical properties of the NG consumption time series? Which characteristics has a major contribution on how gas consumption develops over time?

The time series of natural gas (NG) consumption in Spain exhibits statistical characteristics that show dynamic patterns impacted by diverse factors in different categories. First of all, the data show non-stationary behavior, which is defined by variations in the variance and mean across time. All three categories—conventional gas usage (e.g.heating), gas consumption for generating electricity, and LNG consumption—show evidence of this non-stationarity.

Seasonal variations have a big impact on how much conventional gas is used for heating as well as other conventional reasons. Consumption peaks in the winter months of December and January are a reflection of the impact of the weather and the need for heating. The dynamic character of consumption patterns is further enhanced by irregular fluctuations and cyclic behavior, which are probably driven by consumer behavior and economic situations. The statistical characteristics of NG consumption time series are shaped by these elements collectively, which makes them non-stationary and challenging to predict.

Similar to how gas consumption for heating is non-stationary, so is gas consumption for the generation of electricity, although with less obvious seasonal variation. Factors including population increase, industrial activity, and energy policies affect fluctuations in consumption. Long-term consumption patterns are also impacted by the transition to renewable energy sources and technical developments in energy production, underscoring the complex interactions between variables influencing the dynamics of NG consumption.

On the other hand, there are less obvious seasonal patterns or irregular shifts in the consumption of LNG gas over time, indicating generally consistent levels of consumption. The lack of significant outliers indicates a consistent demand for LNG, which may be fueled by elements like transportation, industrial use, and the dynamics of the world energy market.

Seasonal patterns, erratic fluctuations, the state of the economy, technical developments, and energy regulations are all significant elements that have influenced gas usage over time. The statistical characteristics of natural gas consumption time series are determined by these collective characteristics, highlighting the need of understanding the complex processes affecting gas consumption trends in Spain. Policymakers and stakeholders may successfully manage natural gas resources and tackle the nation's energy concerns by taking into account these elements in every aspect.

To what extent are consumption patterns predictable? Which model is the most accurate forecasting tool for our data?

There is an acceptable level of predictability in the natural gas consumption patterns for conventional uses. More complex models, such regression analysis and EWMA, are better at catching fluctuations and seasonal variations and produce more accurate forecasts, even while simpler models, like the SMA, can only follow the overall trend.

Compared to conventional use, the patterns of natural gas consumption in the production of electricity are less predictable due to their increased volatility. The forecasts typically exhibit higher levels of error and unpredictability, which suggests that the underlying consumption drivers are less reliable or more complex.

Out of all the categories, LNG consumption has the highest degree of predictability. Because of the steady consumption patterns and less sensitivity to abrupt shifts, even more basic models like SMA can produce projections that are more accurate. This shows a long-term pattern of constant consumption.

Regarding the forecast accuracy, SMA is good at capturing seasonality and general patterns, but its accuracy decreases with increasing window size. Although it performs very well for short-term forecasts, because it cannot quickly adjust to changes, its accuracy decreases over longer time periods.

By prioritizing recent observations, EWMA increases prediction accuracy. This approach is appropriate for data that changes quickly since it is especially good at catching short-term trends and fluctuations. Nevertheless, the accuracy depends on the smoothing parameter that is selected.

In every category, regression analysis proves to be the most accurate forecasting technique. It produces very accurate projections by skillfully incorporating trends, seasonality, and other outside variables. This strategy is better for long-term projections since it can account for a large amount of the fluctuation in consumption.

In conclusion, regression analysis performs better than other techniques in terms of accuracy and dependability for predicting natural gas consumption across several categories, even though simpler models like SMA offer a solid place to start. Depending on the category, the consumption patterns range from moderately to extremely predictable, with LNG usage being the most predictable.

Have major event (such as the war in Ukraine and the Covid crisis) imposed a structural break in the data-generating process?

Analyzing how significant occurrences like the COVID-19 pandemic and the conflict in Ukraine have affected Spain's natural gas use offers interesting new perspectives on possible shortcomings in the process of gathering data.

COVID-19's Effect on Natural Gas Consumption:

The addition of a COVID-19 dummy variable to the regression model for natural gas use for conventional purposes revealed that the pandemic had no noticeable effect on patterns of consumption. The COVID-19 variable's p-value was 67.32%, far higher than the accepted 5% significance level. This shows that consumption patterns stayed in line with pre-pandemic trends and that the pandemic did not force a fundamental break in the process of obtaining data for conventional natural gas use.

Similarly, with a p-value of 33.4%, the COVID-19 dummy variable did not show a significant impact on natural gas use in the production of electricity. This result suggests that natural gas consumption patterns that are related with the electricity generation were not structurally disrupted by the epidemic and were instead largely unaffected by the disturbances brought on by the pandemic.

With a p-value of 92.83%, the COVID-19 dummy variable also did not demonstrate a significant impact with regard to LNG use. This finding provides more evidence that the consumption patterns of LNG did not change structurally as a result of the pandemic and continued to be consistent with pre-pandemic levels.

Impact of the War in Ukraine on Natural Gas Consumption:

According to the regression model, the war in Ukraine had a negative impact on the use of conventional natural gas. The influence was statistically inconsequential, though, as the p-value for this effect was over the 5% significance level. As a result, the traditional data on natural gas consumption did not undergo a structural disruption due to the conflict in Ukraine.

The fact that the p-value is above 5% suggests that the conflict in Ukraine had no statistically significant effect on the amount of natural gas used to generate power. This result implies that the data-generating process for this category of natural gas usage was not structurally disrupted by the war.

The war in Ukraine, on the other hand, had a statistically significant negative influence on LNG consumption, according to the regression analysis, with a p-value of less than 5%. This suggests that the data on LNG use has a structural break caused by the war.Significant impact likely reflects the broader geopolitical and economic disruptions caused by the conflict, which affected global energy markets and supply chains, thereby influencing LNG consumption patterns in Spain.

In conclusion, the war in Ukraine did cause a structural break in the LNG consumption data, even while the COVID-19 epidemic did not impose one on Spain's natural gas consumption patterns for conventional usage, power generation, or LNG. This demonstrates how these significant events have affected different segments of the natural gas consumption market differently.

Proposal for further research:

Future research on natural gas consumption in Spain could explore several avenues. There are a number of paths that future studies on natural gas consumption in Spain could go. First, it would be advantageous to provide region-specific data for a more detailed examination of consumption trends. Analyzing how the use of renewable energy affects natural gas consumption could put light on the transition dynamics . Prediction accuracy may be improved by using sophisticated forecasting models, such as neural networks or machine learning strategies. More advanced quantitative data analysis methods, such as neural networks or machine learning techniques, could be applied to identify complex (non-linear) characteristics of the sales dynamics and to identify non-linear characteristics of the data set (Thomaidis andDounias, 2012).Furthermore, examining the consequences of policy changes , economic aspects, and adding new variables like the weather may provide a more profound analysis of consumption patterns. Lastly, a more comprehensive understanding of worldwide trends in natural gas use may be obtained through comparative analysis with other countries.

REFERENCES:

1. *Cleveland, C. J. (Ed.). (2004). Natural gas, history of. In Encyclopedia of energy (pp. 207–218). New York: Elsevier. https://doi.org/10.1016/B0-12-176480-X/00042-5*
2. *Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting, 20(1), 5–10. https://doi.org/10.1016/j.ijforecast.2003.09.015*
3. *Santos, S. P., Leal, J. E., & Oliveira, F. (2011). The development of a natural gas transportation logistics management system. Energy Policy, 39(9), 4774–4784. https://doi.org/10.1016/j.enpol.2011.06.047*
4. *Tian, N., Shao, B., Bian, G., Zeng, H., Li, X., & Zhao, W. (2024). Application of forecasting strategies and techniques to natural gas consumption: A comprehensive review and comparative study. Engineering Applications of Artificial Intelligence, 129, 107644. https://doi.org/10.1016/j.engappai.2023.107644*
5. *Yildiz, B., Bilbao, J. I., & Sproul, A. B. (2017). A review and analysis of regression and machine learning models on commercial building electricity load forecasting. Renewable and Sustainable Energy Reviews, 73, 1104–1122. https://doi.org/10.1016/j.rser.2017.02.023*
6. *U.S. Energy Information Administration. (n.d.). Natural gas explained. U.S. Energy Information Administration. Retrieved June 11, 2024, from https://www.eia.gov/energyexplained/natural-gas/*
7. *MET Group. (n.d.). Natural gas transportation. MET Group. Retrieved June 11, 2024, from https://group.met.com/en/media/energy-insight/natural-gas-transportation*
8. *U.S. Energy Information Administration. (n.d.). Natural gas pipelines. U.S. Energy Information Administration. Retrieved June 11, 2024, from https://www.eia.gov/energyexplained/natural-gas/natural-gas-pipelines.php*
9. *Nedgia. (n.d.). History of natural gas and Nedgia. Nedgia. Retrieved June 11, 2024, from https://www.nedgia.es/conocenos/en/natural-gas/history-of-natural-gas-and-nedgia/*
10. *U.S. Energy Information Administration. (n.d.). Spain. U.S. Energy Information Administration. Retrieved June 11, 2024, from https://www.eia.gov/international/analysis/country/ESP*
11. *Cores. (n.d.). Who we are. Cores. Retrieved June 11, 2024, from https://www.cores.es/en/cores/quienes-somos*
12. *ResearchMethod.net. (n.d.). Descriptive statistics. ResearchMethod.net. Retrieved June 11, 2024, from https://researchmethod.net/descriptive-statistics/*
13. *Moon, M. A. (2018). Demand and supply integration: The key to world-class demand forecasting (2nd ed.). Boston: Walter de Gruyter GmbH. Retrieved from https://proxy.eap.gr/login?url=https://www.proquest.com/books/demand-supply-integration/docview/2148548598/se-2*
14. *Corporate Finance Institute. (n.d.). Exponentially weighted moving average (EWMA). Corporate Finance Institute. Retrieved June 11, 2024, from https://corporatefinanceinstitute.com/resources/career-map/sell-side/capital-markets/exponentially-weighted-moving-average-ewma/*
15. *Garver, M. S., Smith, C. D., Mentzer, J. T., & Moon, M. A. (2018). Demand and supply integration: The key to world-class demand forecasting (2nd ed.). Boston: Walter de Gruyter GmbH.*

16. *Thomaidis, N. S., & Dounias, G. D. (2012). A comparison of statistical tests for the adequacy of a neural network regression model. Quantitative Finance, 12(3), 437–449. https://doi.org/10.1080/14697680903426573*
17. *Islam, M. A., Che, H. S., Hasanuzzaman, M., & Rahim, N. A. (2020). Energy demand forecasting. In Energy for sustainable development (pp. 105–123). https://doi.org/10.1016/b978-0-12-814645-3.00005-5*