

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**COMPARATIVE ANALYSIS OF XGBOOST AND LIGHTGBM METHODS
FOR DAY AHEAD SPOT NATURAL GAS PRICE FORECASTING**



M.Sc. THESIS

Doğukan ŞAHİN

Energy Science and Technology Division

Energy Science and Technology Programme

FEBRUARY 2022

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**COMPARATIVE ANALYSIS OF XGBOOST AND LIGHGBM METHODS FOR
DAY AHEAD SPOT NATURAL GAS PRICE FORECASTING**

M.Sc. THESIS

**Doğukan ŞAHİN
(301181010)**

Energy Science and Technology Division

Energy Science and Technology Programme

Thesis Advisor: Dr. Lecturer M. Berker YURTSEVEN

FEBRUARY 2022

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**GÜN ÖNCESİ SPOT DOĞALGAZ FİYATI TAHMİNLEMESİNDE LIGHTGBM
VE XGBOOST METODLARININ KARSILASTIRILMALI ANALİZİ**

YÜKSEK LİSANS TEZİ

**Doğukan ŞAHİN
(301181010)**

Enerji Bilim ve Teknoloji Anabilim Dalı

Enerji Bilim ve Teknoloji Programı

Tez Danışmanı: Dr. Öğr. Üyesi. M.Berker YURTSEVEN

ŞUBAT 2022

Doğukan ŞAHİN, a M.Sc. student of ITU Graduate School student ID 301181010, successfully defended the thesis/dissertation entitled “COMPARATIVE ANALYSIS OF XGBOOST AND LIGHTGBM METHODS FOR DAY AHEAD SPOT NATURAL GAS PRICE FORECASTING”, which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Asst. Prof. Dr. M. Berker YURTSEVEN**
İstanbul Technical University

Jury Members : **Asst. Prof. Dr. Burak BARUTÇU**
İstanbul Technical University

Assoc. Prof. Dr. Cenk YAVUZ
Sakarya University

Date of Submission : 11 January 2022

Date of Defense : 08 February 2022





I would like to thank my mother for continuous support throughout my entire life.



FOREWORD

Natural gas is one of the most important commodities of the Turkish Energy Market. Natural gas is becoming more important with discoveries of new gas fields, increased infrastructure power and market liberalization. No doubt, with the liberalization of the natural gas market, the spot natural gas market of Turkey will be even more important for eligible customers. Price forecasting with machine learning techniques is very popular thanks to its comprehensive usage. The aim of this study is to forecast the price of spot natural gas using machine learning techniques and increase predictability. This study could guide eligible customers who want to trade in the spot gas market and researchers.

December 2021

Doğukan ŞAHİN
(Electrical Engineer)

TABLE OF CONTENTS

	<u>Page</u>
FOREWORD	ix
TABLE OF CONTENTS	xi
ABBREVIATIONS	xiii
SYMBOLS	xv
LIST OF TABLES	xvii
LIST OF FIGURES	xix
SUMMARY	xxi
ÖZET	xxiii
1.INTRODUCTION	1
1.1 Purpose of Thesis.....	2
2. GLOBAL GAS MARKET OUTLOOK	3
2.1 Global Natural Gas Reserves.....	3
2.2 Global Natural Gas Consumption	4
2.3 Global Natural Gas Trades	5
3. TURKISH GAS MARKET	7
3.1 History of Turkish Natural Gas Market.....	7
3.2 Natural Gas Supply of Turkey.....	10
3.2.1 Domestic natural gas sources	10
3.2.2 Natural gas imports of Turkey	13
3.3 Natural gas storage of Turkey	18
3.3.1 Storage prices	19
3.3.2 Natural gas storage facilities of Turkey.....	20
3.4 Natural Gas Consumption of Turkey.....	22
3.5 Natural Gas Export	25
3.6 Market Structure Of Turkey	26
3.6.1 Organized natural gas wholesale market	27
3.6.2 Pricing of natural gas in Turkish Gas Market	28
4.1 Supervised Learning	31
4.2 Unsupervised Learning.....	32
4.3 Ensemble Learning	33
4.4 Bootstrap.....	33
4.5 Bagging.....	34
4.6 Gradient Boosting.....	34
4.6 Random Forests	35
4.7 XGBoost Algorithm.....	36
5. DATA PREPARATION AND MODELLING	37
5.1 Data Preparation	37
5.1.1 Spot day ahead prices and volumes.....	37
5.1.2 BOTAS transactions.....	37
5.1.3 Brent oil prices	38
5.1.4 BOTAS tariffs	38

5.1.5 Dollar.....	38
5.2 Modelling.....	39
5.2.1 Evaluation metrics	39
5.2.2 Standardization	40
5.2.3 Early stopping.....	40
5.3 Results.....	40
5.3.1 XGBoost model results	40
5.3.2 LIGHTGBM results	47
5.3.3 Comparison	52
6. CONCLUSION AND FUTURE RECOMMENDATIONS	53
REFERENCES	55
CURRICULUM VITAE	59



ABBREVIATIONS

ABPP	: Additional Balancing Gas Purchase Price
ABSP	: Additional Balancing Gas Sell Price
Bcm	: Billion Cubic Meter
BGPP	: Balancing Gas Purchase Price
BGSP	: Balancing Gas Sell Price
BOTAS	: Petroleum Pipeline Corporation (Boru Hatları İle Petrol Taşıma Anonim Şirketi)
BP	: British Petroleum
CCGT	: Combined Cycle Power Plant
CNG	: Compressed Natural Gas
EMRA	: Energy Market Regulatory Authority
EPDK	: Enerji Piyasası Düzenleme Kurumu (Energy Market Regulatory Authority)
FSRU	: Floating Storage and Regasification Unit
IGU	: International Gas Union
LightGMB	: Light Gradient Boosting Machine
LNG	: Liquefied Natural Gas
TL	: Turkish Lira
OECD	: Organisation for Economic Co-operation and Development
Tcm	: Trillion Cubic Meter
USSR	: Union of Soviet Socialist Republics
XGBoost	: Extreme Gradient Boosting
XGBRegressor	: Extreme Gradient Boosting Regressor
TANAP	: Trans Anatolian Natural Gas Pipeline (Trans Anadolu Doğalgaz Boru Hattı)



SYMBOLS

°C	: Degrees Celsius
m³	: Cubic Meter
Sm³	: Standard Cubic Meter





LIST OF TABLES

	<u>Page</u>
Table 2.1: Natural Gas Components (Saka, 2003).....	3
Table 5.1: Dataset Summary.	399
Table 5.2: MAE, RMSE and R2 score values according to the estimator numbers. 41	41
Table 5.3: MAE and RMSE and R2 score values according to the learning rate. ..	433
Table 5.4: MAE and RMSE and R2 score values according to the learning rate ...	444
Table 5.5: MAE and RMSE and R2 score values of LIGHTGBM according to the number of estimator	488
Table 5.6: MAE and RMSE and R2 score values of LIGHTGBM according to the number of estimator	499
Table 5.7: MAE and RMSE and R2 score values of LIGHTGBM according to the max depth.	50



LIST OF FIGURES

	Page
Figure 2.1: Proven natural gas reserve map in cubic meters (Url-2).....	4
Figure 2.2: Natural gas consumption map in cubic meters in 2019 (Url-4).....	5
Figure 2.3: Major natural gas trade movements in 2019 (Url-5).....	5
Figure 3.1: Natural gas chronology of Turkey (Url-12).....	9
Figure 3.2: Consumer shares of wholesale licensed production companies in 2019 (%) (EPDK, 2019).....	10
Figure 3.3: Natural gas production rates according to the provinces (%) (EPDK, 2019).....	11
Figure 3.4: Ratio of domestic natural gas production and yearly total natural gas consumption (EPDK, 2019)	11
Figure 3.5: Natural gas production amounts of Turkey by years (Milion Sm ³) (EPDK, 2019).....	12
Figure 3.6: Monthly comparison of natural gas production between 2013-2019 (Milion Sm ³) (EPDK, 2019)	12
Figure 3.7: Long term natural gas purchase contracts of Turkey (Billion Sm ³) (Url-12).....	13
Figure 3.8: Natural gas import amounts by country for 2009-2019 (Billion Sm ³) (EPDK, 2019).....	14
Figure 3.9: Share of imported natural gas by source country in 2019 (%) (EPDK, 2019).....	14
Figure 3.10: Long Term LNG Import Amounts by Source Countries between 2013-2019 (Million Sm ³) (EPDK, 2019)	15
Figure 3.11: Import amounts of natural gas by type (Million Sm ³) (EPDK, 2019)..	16
Figure 3.12: Change in spot LNG import amount between 2008-2019 (Million Sm ³) (EPDK, 2019).....	16
Figure 3.13: Month end natural gas stock amounts by years (Million Sm ³) (EPDK, 2019)	19
Figure 3.14: Comparison of the storage cost of 1 Sm ³ natural gas and the cost of pipeline transmission of 1 Sm ³ in 2021 (TL\ Sm ³) (Url-15)	20
Figure 3.15: Silivri Storage Facility (Url-16).....	20
Figure 3.16: Tuz Lake natural gas storage facility (Url-17).....	21
Figure 3.17: Marmara Ereğlisi LNG terminal (Url-19).....	21
Figure 3.18: Yearly natural gas consumption of Turkey (Billion Sm ³) (EPDK, 2019)	22
Figure 3.19: Annual comparison of sectoral consumption (Million Sm ³) (EPDK, 2019).....	22
Figure 3.20: Effect of temperature (°C) on Istanbul household natural gas consumption (million Sm ³) (EPDK, 2019), (Url-20).....	23
Figure 3.21: Monthly changes of sectoral consumption in 2019 (Million Sm ³) (EPDK, 2019).....	23
Figure 3.22: Monthly industrial natural gas consumption (Million Sm ³) and	

industrial production index (EPDK, 2019) (Url-21).....24



Figure 3.23: Monthly electricity production (MWh) according to the sources in Turkey in 2019 (Url-22).	24
Figure 3.24: Annual natural gas export amounts (Million Sm ³) (EPDK, 2019)	25
Figure 3.25: Turkish natural gas market structure (Turkey's Energy Transition Milestones and Challenges, 2015)	26
Figure 3.26: Wholalase natural gas market contract timelines.	27
Figure 4.1: Machine learning techniques.	32
Figure 4.2: Bootstrap sampling procedure (Rahman, 2019)	33
Figure 4.3: Bagging procedure	34
Figure 4.4: Gradient boosting procedure	35
Figure 5.1: Training accuracy performance of overfit model (Url-26).	40
Figure 5.2: Effect of number of estimators to the model performance.	41
Figure 5.3: Effect of learning rates to model performance.	433
Figure 5.4: Effect of maximum tree depth to model performance	444
Figure 5.5: Working principle of GridsearchCV function. Green blocks show test sets created from dataset, red blocks show training sets (Url-27).	45
Figure 5.6: Predicted and Observed Day Ahead Spot Prices, output of XGBoostRegressor with grid search's best parameters.	466
Figure 5.7: Prediction and Observed day ahead spot gas prices for XGBRegressor model with learning rate 0.1, 6 for max depth with 56 estimator.	46
Figure 5.8: Effect of number of estimators on lightboost.	47
Figure 5.9: Effect of learning rate to LIGHTGBM model performance	499
Figure 5.10: Effect of max depth to LIGHTGBM model performance.	50
Figure 5.11: Prediction and Observed day ahead spot gas prices for LIGHTGBM model with 0.05 learning rate, 4 for max depth with 24 estimator.	51
Figure 5.12: Prediction and Observed day ahead spot gas prices for LIGHTGM model with 0.05 learning rate, 4 for max depth with 24 estimator.	51

COMPARATIVE ANALYSIS OF XGBOOST AND LIGHTGBM METHODS FOR DAY AHEAD SPOT NATURAL GAS PRICE FORECASTING

SUMMARY

Natural gas is an important part of modern life. Advantages such as being storable, being able to be used in many different places as a primary energy source and being transported in the form of LNG or CNG increase the importance of natural gas. The discovery of new natural gas fields and the importance given to infrastructure, the laws adopted for the liberalization of the market indicate that natural gas will be used in Turkey for many years. In addition, studies aimed at using natural gas with hydrogen in Turkey are promising to increase the life of natural gas.

One of the most important criteria for using natural gas in Turkey for many years is predictable prices. Each eligible customer demands predictable pricing and low price. Competition is necessary for liberal markets. Therefore, EPIAS, the market operator, develops the infrastructure of the natural gas market day by day. The weekly product option added to the market after the spot market opened and the opening of the futures gas market are the best examples of this. However, there are very few studies about the spot natural gas market in the literature. This thesis aims increase price predictability in this developing market. In this study, the official reports prepared for the natural gas market were examined first. Later, the dynamics of the spot natural gas market were examined, and a data set was prepared accordingly. The data set is crucial because it is very difficult to get good results without the proper data set.

Machine learning applications are very popular and are very effective in solving various problems. In this study, day ahead spot natural gas price prediction was made with two different machine learning algorithms based on decision trees. XGBoost and LightBoost algorithms were utilized and performances of these two algorithms were compared. First, it was concluded that these algorithms over-learn at which value ranges by iterating the number of trees, tree depth and learning rates. A more detailed price prediction was then made using these ranges together with the GridSearch function of the sklearn library. Then the best features for price prediction are determined by the feature selection function of the XGBoost and LightBoost libraries.



GÜN ÖNCESİ SPOT DOĞALGAZ FİYATI TAHMİNLEMESİNDE LIGHTGBM VE XGBOOST METODLARININ KARSILASTIRILMALI ANALİZİ

ÖZET

Doğal gaz modern yaşamın önemli bir parçasıdır. Depolanabilir olması, birincil enerji kaynağı olarak pek çok farklı yerde kullanılabilmesi, LNG veya CNG formunda taşınabilmesi gibi avantajlar doğal gazın önemini artmaktadır. Yeni yapılan doğal gaz sahalarının keşifleri ve altyapıya verilen önem, marketin liberalleşmesi için çıkarılan yasalar doğalgazın uzun yıllar boyunca Türkiye'de kullanılacağı göstergesidir. Bunun dışında Türkiye'deki doğal gazın hidrojen ile kullanılmasını amaçlayan çalışmalar doğal gazın kullanım ömrünün arttırılması konusunda ümit vericidir.

Türkiye'nin komsuları petrol ve doğalgaz zengini olmasına rağmen Türkiye'nin hidrokarbon rezervleri oldukça düşüktür. Bu yüzden Türkiye enerji ihtiyacının büyük bir kısmını yurtdışından karşılamaktadır. Bu da arz güvenliği problemini beraberinde getirmektedir. Türkiye arz güvenliğini arttırmak için pek çok farklı stratejik yatırım yapmaktadır. Bunlardan biri de doğal gaz altyapısına yapılan iyileştirmelerdir. Altyapı geliştirmeleri sayesinde Türkiye pek çok farklı noktadan doğalgaz ithal edip doğal gazın pek çok farklı tüketim noktasına iletimi sağlamaktadır.

Türkiye'de doğal gaz ticareti ilk kez 1986 yılında Sovyet Rusya ile yapılan gaz alım anlaşması ile başlamaktadır. Ardından kaynakları çeşitlendirmek ve arz güvenliğini sağlamak için 1988 yılında Cezayir ile LNG anlaşması imzalanmış ve 1994 yılında Marmara Ereğlisi LNG Terminali işletmeye başlanılmıştır. Daha sonrasında ise Nijerya, İran ve Azerbeycan ile yapılan anlaşmalar sayesinde doğal gaz arzı çeşitlendirilmiştir.

Gelişen altyapı sayesinde Türkiye doğal gazı pek çok farklı ilde pek çok farklı alanlarda kullanmaktadır. Bu da doğal gazın artık Türkiye için vazgeçilmez olduğunun göstergesidir. Türkiye'nin doğal gaz kullanım alanları ise çok çeşitlidir. Öncelikle evsel kullanımda ve elektrik üretiminde kullanılmaktadır. Bunun yanında yapılan araştırmalarda sanayi üretiminin artışı ile sanayide kullanılan doğal gaz tüketiminin arttığı görülmüştür.

Türkiye'de elektrik arzı için doğalgaz arzı oldukça önemlidir. Çünkü Türkiye'de elektrik üretimi büyük oranda doğal gaza bağlıdır. Aslında elektrik piyasası yapısına göre elektrik enerjisi ilk olarak yenilenebilir enerji kaynaklarından karşılanırsa da bu kaynaklar tüm Türkiye'nin ihtiyaç duyduğu elektrik enerjisini karşılamada yetersizdir. Ayrıca yenilenebilir enerji kaynaklarının kontrolü mümkün değildir. Örneğin son yıllarda yaşanan kuraklık nedeniyle düşen hidroelektrik üretiminden dolayı doğal gazdan elektrik üretimi artış göstermiştir.

Türkiye'deki doğal gaz tüketiminde sezonsallık oldukça fazladır. Bunun en büyük sebebi evsel tüketimdir. Kış aylarında ısınma amaçlı tüketilen doğalgaz miktarı

oldukça yüksektir. Bu sezonluk etki de beraberinde bir depolama ihtiyacı oluşturmuştur. Bu yüzden son yıllarda doğalgaz depolaması için yatırımlar yapılmaktadır.

Doğal gazın fiyatlandırılması konusunda pek çok farklı strateji bulunmaktadır. Yapılan literatür araştırmasında Türkiye'deki doğal gaz fiyatlandırmasının petrol bazlı ikame fiyatlandırma olduğu görülmüştür. Bu bir baz fiyatın üzerine petrol fiyatının belirli stratejilerle eklenmesi sonucu yapılan fiyatlandırma değildir. Bunun dışında doğal gaz fiyatlandırmasındaki en büyük problemlerden biri de kur riskidir. Doların Türk Lirası karşısında değer kazanması dolarla satın alınan doğal gazın fiyatını arttırmaktadır.

BOTAS ilk kurulum amacı petrol taşıma olsa da zamanla şirket yapısı değişmiş ve doğal gaz ticareti ve taşımacılığı da yapmaya başlamıştır. BOTAS halen doğal gaz piyasasında sektör lideridir. Düzenli olarak her ay serbest tüketiciler, serbest olmayan tüketiciler ve elektrik üreticileri için tarife yayınlamaktadır. Bunun yanında BOTAS Türkiye'deki tek doğal gaz iletim şirketidir. Fakat dağıtım konusunda pek çok farklı şirket vardır ve doğal gaz dağıtımını konusunda tekelleşmenin önüne geçilmiştir.

BOTAS'ın tekel yapısını kırmak ve daha liberal bir doğal gaz piyasası oluşturmak için 4646 sayılı yasa çıkartılmıştır. Bu yasa ile BOTAS'ın yapısı değişmiş ve liberalleşmenin yol haritası ortaya konulmuştur. Bu yasa çerçevesinde BOTAS, bazı doğalgaz kontratlarını özel sektöre devretmiştir. Ancak bu liberalleşme adımlarına uyulamamıştır ve bu yasadaki revizyonlara gidileceği öngörülmektedir.

Türkiye'de ki en büyük doğal gaz ithalatçısı BOTAS olmasında rağmen özel sektörün de doğal gaz ithalatında payı vardır. Bunlar genelde BOTAS'ın 4646 numaralı doğalgaz yasası kapsamında özel sektöre devrettikleri kontratlardır. Bunun dışında gerekli lisansa sahip piyasa katılımcıları spot boru gazı ithal edebilmekte ancak bu çok tercih edilmemektedir.

Türkiye, jeopolitik konumu nedeniyle pek çok farklı kaynaktan doğalgaz satın alabilmektedir. Ayrıca pek çok farklı boru hattının kesiştiği bir noktadadır. Bu yüzden Türkiye'nin bir doğalgaz ticaret merkezi olması söz konusudur. Ancak, Türkiye'nin bir doğalgaz ticaret noktası olabilmesi için halen çözmesi gereken bazı sorunları vardır. Örneğin Türkiye'de doğalgaz piyasası halen liberalleşmemiştir. Ancak son yıllarda liberalleşmek için bir çok farklı adım atılmıştır.

Liberalleşmenin en önemli adımlarından biri ise EPIAŞ bünyesinde Sürekli Ticaret Platformu'nun (spot doğalgaz piyasası) kurulmasıdır. Bu platform sayesinde artık piyasa katılımcıları anonim ve kesintisiz olarak doğalgaz ticareti yapabileceklerdir. Piyasa katılımcılarının spot piyasada yapabilecekleri pek çok farklı kontrat seçeneği vardır. Bunlar gün öncesi, gün içi, gün sonu veya haftalık kontratlardır. Fakat en aktif olarak doğal gaz ticareti gün öncesi piyasasında gerçekleşmektedir. Vadeli doğalgaz piyasası ve haftalık kontratlar oldukça yenidir ve şüphesiz ki bu platformların halen zamana ihtiyaçları vardır.

Türkiye'de doğalgazın uzun yıllar kullanılabilir olmasının en önemli kriterlerinden biri de fiyatlandırılması veya fiyatlarının öngörülebilir olmasıdır. Her serbest tüketici

öngörülebilir fiyatlandırma ve düşük fiyat talep eder. Düşük fiyatın oluşabilmesi için rekabet gereklidir. Bu yüzden piyasa operatörü olan EPIAS doğalgaz marketinin altyapısını günden güne geliştirmektedir. Spot piyasa açıldıktan sonra piyasaya eklenen haftalık ürün ticareti opsiyonu ve vadeli gaz piyasasının acılması buna en güzel örneklerdir. Şüphesiz ki var olan bu yeni spot gaz doğalgaz marketi piyasanın diğer dinamiklerinden etkilenmektedir. Bu yüzden özellikle tezin ilk bölümlerinde resmi kurumlarca hazırlanan raporlar incelenmiştir. Yapılan incelemeler sonucunda spot piyasanın bağlı olduğu dinamikler araştırılmıştır. Araştırma sonucunda elde edilen veriler ışığında bir veri seti farklı modelleri eğitmek için derlenmiştir. Bu aşama oldukça önemlidir. Çünkü doğru veri seti hazırlanmadan ve doğru dinamikler ortaya çıkarılmadan iyi tahminler elde edilmesi oldukça güçtür. Literatürde spot doğalgaz piyasası hakkında oldukça az çalışma bulunmaktadır. Bu çalışmanın amacı gelişen bu marketteki fiyat öngörülebilirliğini arttırmaktır.

Makine öğrenmesi uygulamaları günümüzde çok popülerdir. Bu algoritmalar regresyon problemlerinin çözümünde çok etkilidir. Bu çalışmada karar ağacı tabanlı iki farklı makine öğrenmesi algoritması ile gün öncesi spot doğalgaz fiyat tahminlemesi yapılmıştır ve sonuçlar karşılaştırılmıştır. XGBoost ve LightBoost algoritmaları kullanılmıştır. Öncelikle ağaç sayısı, ağaç derinliği ve öğrenme oranları irdelenerek yapılarak hangi değer aralıklarında aşırı öğrenme yaptığı tespit edilmiştir. Ardından bu aralıklar scikit-learn kütüphanesinin GridSearch fonksiyonu ile beraber kullanılarak model parametreleri optimize edilmiştir. Fiyatın hangi parametrelere bağlı olduğu XGBoost ve LightGBM kütüphanelerinin öznelik seçimi fonksiyonu ile belirlenmiştir.

Yapılan çalışmalar sonucunda XGBoost'un LightGBM'den daha iyi sonuç verdiği görülmüştür. Fiyatın bağlı olduğu parametreler ise uzun dönemli fiyat tahminleri için Brent petrol fiyatı, BOTAS tarifeleri ve dolar olduğu belirlenmiştir. Kısa dönemli olarak fiyata etki eden faktörlerin ise BOTAS'ın stok miktarını dengelemek için yaptığı ticaretler olduğu sonucuna ulaşılmıştır.



1.INTRODUCTION

Natural gas is traded in organized and unorganized markets like other energy commodities, but many different features distinguish this commodity from other energy commodities. For example, one of the most important features that distinguish natural gas trades from electricity trade is that natural gas is storable. Apart from that, natural gas can be used as a primary or secondary energy source. In addition, natural gas is used directly as raw material in the industry. The transport of natural gas can be in gas form or liquid form. Such features distinguish natural gas from other energy commodities.

Many factors affect the price of natural gas. Demand, supply, economic growth, oil prices, climate can be given as examples. The price of natural gas in the Turkish market was initially determined by BOTAS. When the law numbered 4646 was enacted in 2001, the Turkish gas market began to be liberalized. Then the spot gas market was opened on 01.09.2018. The spot gas market is an important step for liberalization.

The Turkish Spot Natural Gas Market is developing rapidly. For example, on 01.06.2020, weekly products were also integrated into the spot market, since this date market participants can trade weekly products. On October 1, 2021, natural gas future products are planned to be added to the spot gas market.

Many market participants - with the required license - trade in this new market. Basically, there are 3 different contracts in the spot gas market. These are day-ahead, intraday and day-after contracts. Among these contracts, intraday and day after contracts are generally used for imbalance management. Transactions in the day-ahead market are usually traded according to day-ahead production forecasts. It is important to forecast the day-ahead spot gas price for the consistency of the day-ahead cost-based calculations. In addition to these contracts, weekly trading was added to the market later on. These contracts can be traded on weekends, weekdays or working days only. In addition, natural gas can be traded on a quarterly, semi-annual and annual basis with the future market.

1.1 Purpose of Thesis

It has been observed that there are very few studies about the Turkish Gas Market in the literature. Forecasting natural gas prices is very important for both combined-cycle natural gas power plants and industrial production. A comparative study was conducted using two different machine learning algorithms to understand Turkish Spot Gas Market deeply. Results have been compared with actual prices.



2. GLOBAL GAS MARKET OUTLOOK

Natural gas is one of the most important for modern life. It is one of the cleanest, safest and most used energy sources. It is colorless and odorless in its pure form. It is flammable, and when it burns, a large amount of energy is released. Natural gas consists of various hydrocarbons. These hydrocarbons are primarily methane. Besides methane natural gas contains ethane, propane, butane. The components of natural gas are summarized in the table below (Saka, 2003).

Table 2.1: Natural Gas Components (Saka, 2003).

Natural Gas Components	Percentage of Contents
Methane	70-90%
Etan	0-20%
Carbon Dioxide	0-8%
Oxygen	0-0.2%
Nitrogen	0-5%
Hydrogen Sulfide	0-5%

2.1 Global Natural Gas Reserves

Some key terms are used to classify and report natural gas reserves. These are 1P which means proven reserve, 2P which means proven and probable reserve, 3P which means possible reserve. Many stakeholders demand from natural gas producers to publish their 1P and 2P coded reserves. 1P coded reserves indicate a better view for markets. 1P coded reserves are commercially recoverable resources. 1P reserves are calculated by a deterministic approach. It's close to the actual reserve value with a high probability. In reserve calculations if probabilistic method is used, there is at least 90% accuracy (ACCC, 2019).

The change in the production of fossil fuels over the years was examined in the hypothesis which M.K. Hubbert published in 1949. According to this hypothesis, the

production of fossil resources forms a peak point over the years and then continues to decline. However, this curve can shift with activities such as newly discovered resources and more efficient ways of consumption. Many scientists try to predict the end of oil and natural gas production using this method (Yağız, 2016).

The amount of proven natural gas reserves in the world was calculated as 198.8 Tcm in 2019, with an increase of 1.7 Tcm from 2018. Despite proven reserve growth in China and Azerbaijan, it has been offset by a decline in calculated rezerv amounts in Indonesia. Russia, Iran and Qatar are the countries with the most reserves. When gas reserves and production quantities are considered, there is a 49.8 years gas reserve in the world (Url-1).

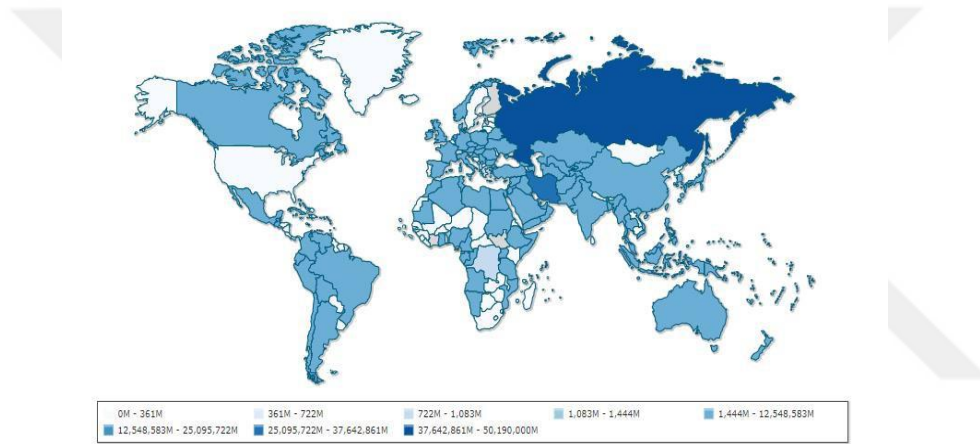


Figure 2.1: Proven natural gas reserve map in cubic meters (Url-2).

2.2 Global Natural Gas Consumption

Natural gas consumption is related to many factors. One of the most important ones is economic development. In a positive manner, a close relationship is observed between the amount of natural gas consumption and economic development. By region, 29.5% of average consumption occurs in Europe and Eurasia 28.3% in North America and an average of 20% in the Asia Pacific region. (Yağız, 2016).

According to the figures published by BP Company, global natural gas consumption increased by 2% on average in 2019. Considering the growth of 5.3% in 2018 and the average consumption growth over the past 10 years, there was a big decrease in the consumption increase ratio in 2019. In 2019, there was an average increase of 78 billion cubic meters. This increase was largely coming from USA (27 Bcm) and China (24 Bcm). However, despite this consumption increase of the USA and China the ratio

of consumption growth of these countries has decreased. In 2019, Mild weather was observed in Russia. Due to mild weather, Russia's natural gas consumption decreased around 10 Bcm. This is the largest natural gas consumption decrease of a country in 2019 (Url-3).

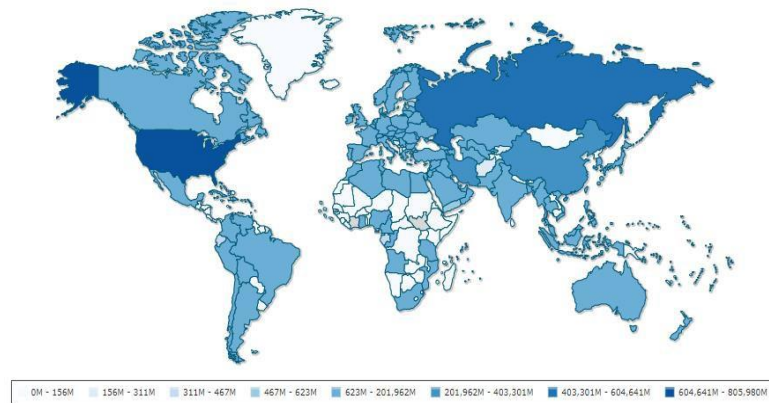


Figure 2.2: Natural gas consumption map in cubic meters in 2019 (Url-4).

2.3 Global Natural Gas Trades

Global energy imports increased by 55.5 Bcm to 1.2 Tcm in 2019. The same trend has been prevailing in the global energy trade since 2017. Gas trade with LNG form was the trend of recent years due to the availability of LNG prices in the world. LNG trades increased by 19.4 Bcm for non-OECD countries and 46.2 Bcm for OECD countries. The rate of LNG trade increased from 34.3% to 38.1% in 2019. Qatar and Australia were the most LNG-exporting countries. Qatar exported 59.1 Bcm LNG and Australia exported 51.2 Bcm LNG in 2019 (Url-5).

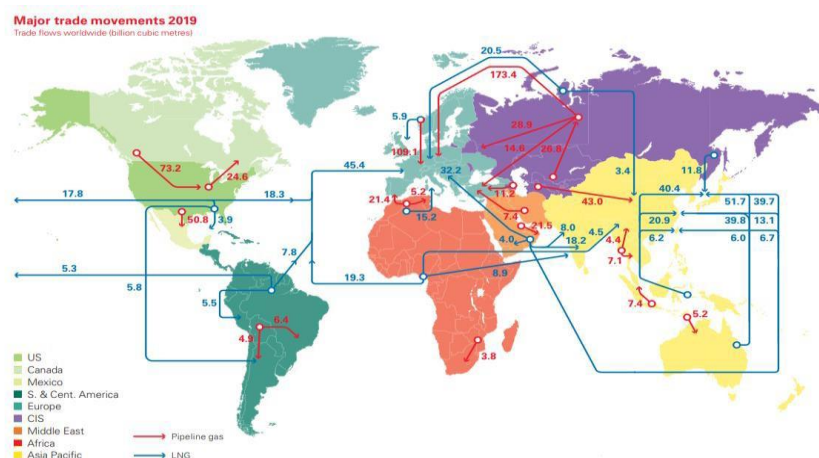


Figure 2.3: Major natural gas trade movements in 2019 (Url-5).



3. TURKISH GAS MARKET

3.1 History of Turkish Natural Gas Market

For domestic heating in Turkey, coal was used during the 1970s. In these years there was air pollution because of the overuse of coal. To protect against air pollution, society began to use their annual leaves during the winter season. Although importing coal has started to be used to reduce air pollution, this has not been a solution. Because of the air pollution, protection of energy supply also energy demand growth and for other reasons Turkey began using natural gas as a source of energy (Engin, 2013).

Turkey's natural gas resources are limited. For example, in 2003, Turkey only provide 2.62% of the domestic consumption with self-resources (Kilic, 2005). After the discovery of limited sources, these resources began to be used in different facilities close to gas fields. For example, natural gas has been discovered for the first time in Kırklareli, Turkey. After 6 years, this gas was started to use in Pınarhisar Cement Factory. In 1975, another gas field was discovered in Mardin. This gas was started to use in Mardin Cement Factory (Yardımcı, 2011).

Due to insufficient domestic supply Turkey began to import natural gas. On September 18, 1984, natural gas transport agreement was signed between USSR and Turkey. After this natural gas transport agreement, a new agreement was made with the Union of Soviet Socialist Republics in 1986 to purchase yearly 6 Bcm of natural gas. The natural gas consumption of Turkey was increasing over the years. Therefore, the security of energy supply has become very important. To secure natural gas supply and meet the increasing demand LNG purchases from Algeria started in 1994. Later, the natural gas supply was diversified by conducting Nigeria and spot LNG trades (TMMOB, 2006).

BOTAS is the biggest participant in the natural gas market of Turkey. It was established in 1974 to transport oil from Iraq. Later, BOTAS started to natural gas transportation and trade in 1987, in order to meet energy demand (Url-6). With a law enacted in 1990, BOTAS was declared as an official institution that sells natural gas.

Botas was deciding on infrastructure investments and determining natural gas prices. In other words, BOTAS became a monopoly under the law (Yardımcı, 2011).

In the 1980s, a liberalization trend started in the world, it also affected the energy markets of Turkey. In this context, Natural Gas Market Law was enacted on 2 May 2001. After the adoption of the Natural Gas Market Law, Energy Market Regulatory Authority started to control the energy markets. According to law, BOTAS will be reorganized in near future. Until the BOTAS import rate decrease to the 20% of national consumption, purchase and sale agreements will be transferred to the private sector (Yardımcı, 2011). In 2005 with a tender a natural gas contract between BOTAS and Russia was transferred to the private sector for the first time (Url-7). The transfer of natural gas contracts has accelerated liberalization in the market. But BOTAS ' transferring its contracts was not enough to establish a liberal market. A platform was required for market participants to buy and sell natural gas anonymously and continuously. Therefore, the Continuous Trade Market (Spot Gas Market) was established and market participants started to trade natural gas (Url-8). Later, the market was developed and the option to do weekly trade option was added on 01.06.2020 to the market. In this way, instead of trading daily, market participants will be able to trade for a week or only on weekdays or only on weekends if they wish. Also, market participants will be able to better manage their portfolios by trading between daily and weekly contracts. (Url-9). The Natural Gas Futures Market is planned to be established in 2021 in to eliminate long-term risks and uncertainties. Natural Gas Futures Market is aimed to increase predictability and create a competitive market. In conjunction with the Natural Gas Futures Market, market participants will be able to trade monthly, quarterly and annually contracts Unlike the future electricity market, this market will be organized by EPIAŞ just like the spot natural gas market. Price formations in this market will be affected by foreign natural gas prices. In other words, it is clear that liberal foreign markets will affect natural gas prices in Turkey, regardless of long-term natural gas purchase contracts. (Url-10).

In 2020, a very important discovery was announced for the Turkish Natural Gas Market. Turkey found 320 billion Sm³ gas reserves in the Black Sea. The government plans to extract and use this gas in 2023. Turkey's dependence on foreign natural gas is expected to decrease (Url-11). This discovered gas will benefit Turkish natural gas consumers in the long run.

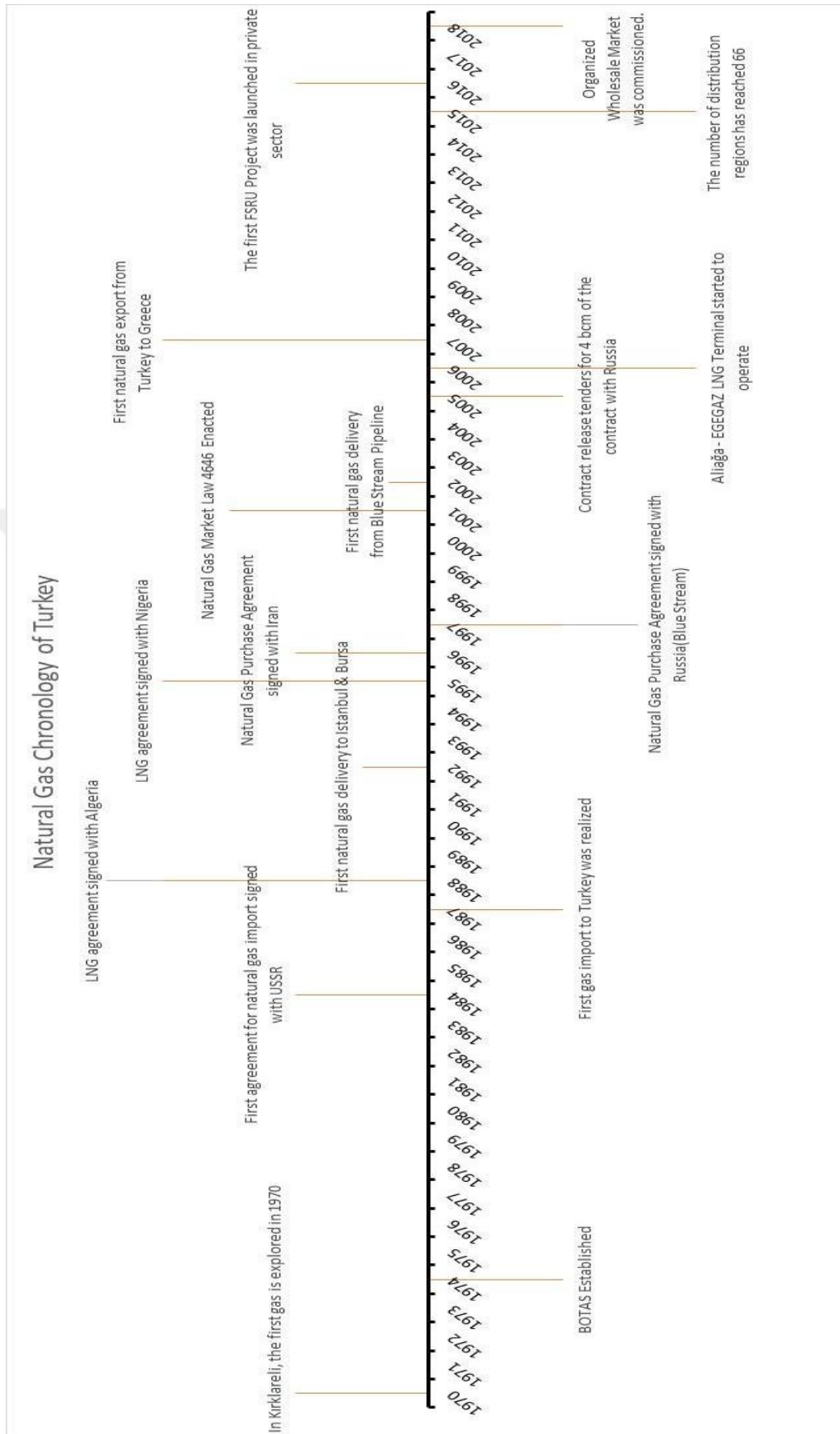


Figure 3.1: Natural gas chronology of Turkey (Url-12).

3.2 Natural Gas Supply of Turkey

3.2.1 Domestic natural gas sources

For companies that want to conduct natural gas production and exploration activities in Turkey, necessary license must be obtained by the General Directorate of Mining and Petroleum Affairs. Natural gas exploration and production companies, if they have the required licenses, can sell the local natural gas they produce to distribution companies, shippers, or CNG companies. In addition, if they have the required licenses, they can sell the natural gas they extract to other countries (EPDK, 2019).

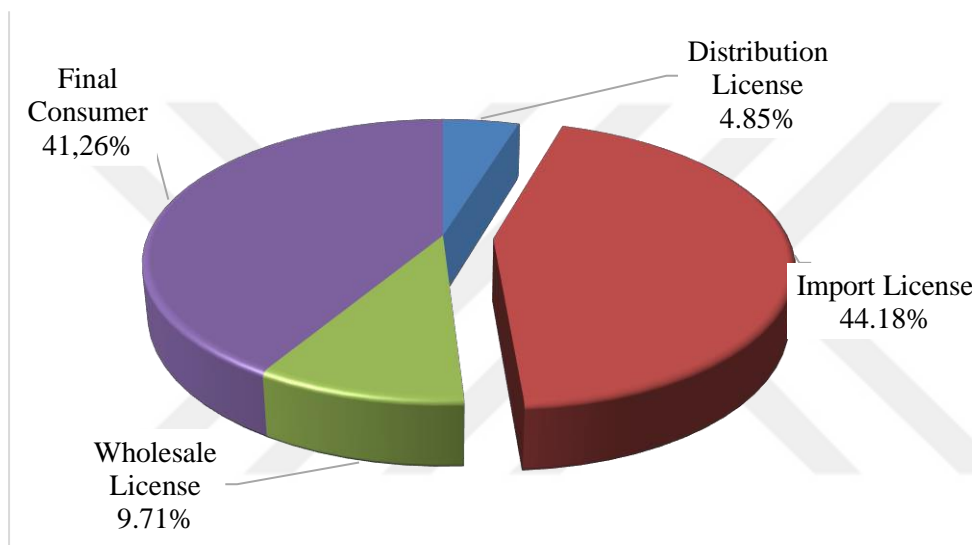


Figure 3.2: Consumer shares of wholesale licensed production companies in 2019 (%) (EPDK, 2019).

Turkey has 13 different natural gas fields. These gas fields located in Adana, Edirne, Kırklareli, İstanbul, Mardin, Adıyaman, Çanakkale and Düzce (Kilic, 2005). Figure 3.3 shows the rate of domestic production of natural gas by provinces in 2019. The maximum amount of natural gas was produced in Istanbul in 2019 with a rate of 44.472% (EPDK, 2019). Produced natural gas is generally used close to gas fields. Local gas is mostly being used in the industry and energy sector. The domestic natural gas production of Turkey is very limited. Figure 3.4 shows the ratio of domestically produced natural gas to Turkey's natural gas consumption by years. Turkey can produce 1.5% of its own average annual natural gas consumption. It's not possible to meet energy demand with local production. Therefore, Turkey imports gas from other countries (Kılıç, 2005).

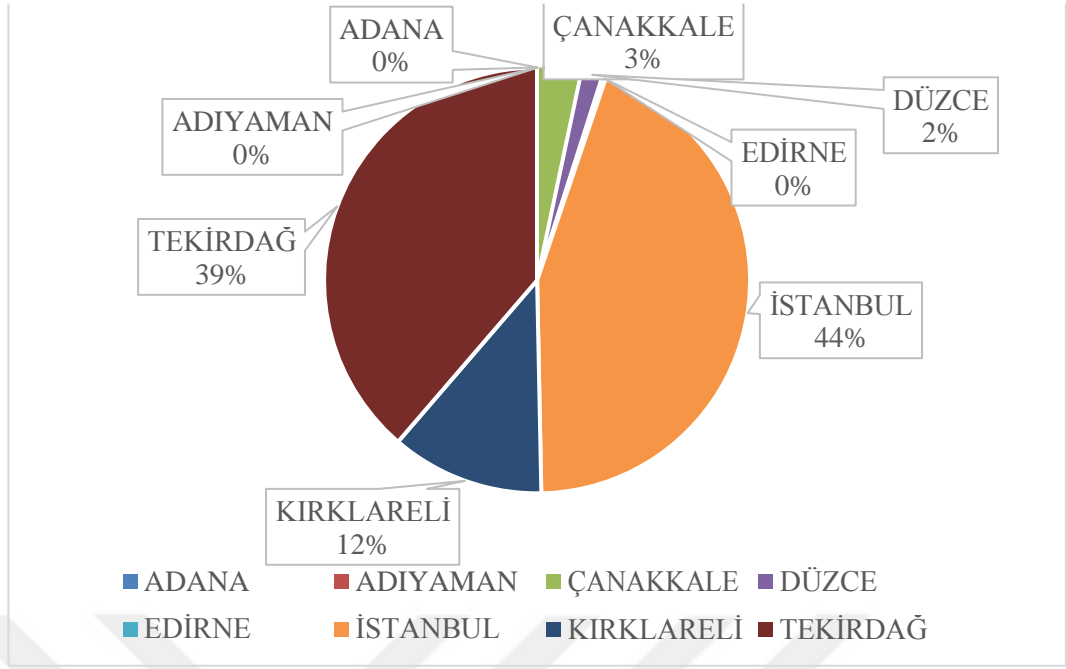


Figure 3.3: Natural gas production rates according to the provinces (%) (EPDK, 2019).

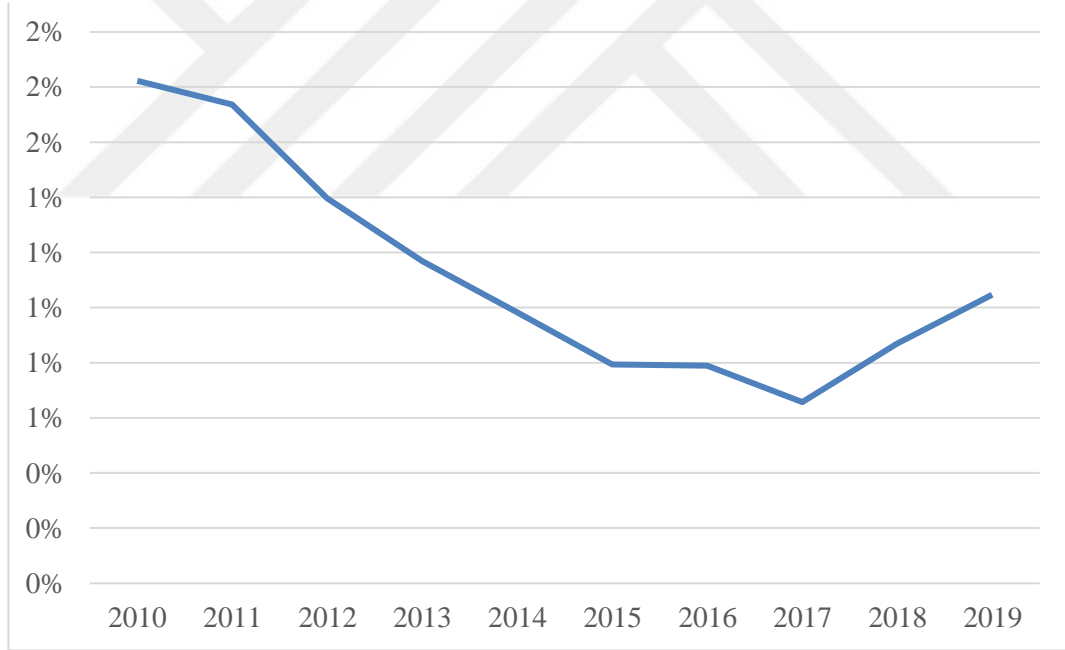


Figure 3.4: Ratio of domestic natural gas production and yearly total natural gas consumption (EPDK, 2019).

Figure 3.5 shows the change in Turkey's local natural gas production between 2011 and 2019. After 2011, there was a decline in Turkey's local natural gas production. But given Turkey's natural gas consumption, this decline is negligible. Figure 3.6 shows the distribution of local natural gas production by month. A seasonal effect on Turkey's

natural gas production by months is not noticeable. It can be said that the reason for this is the use of natural gas produced in industry rather than domestic use.

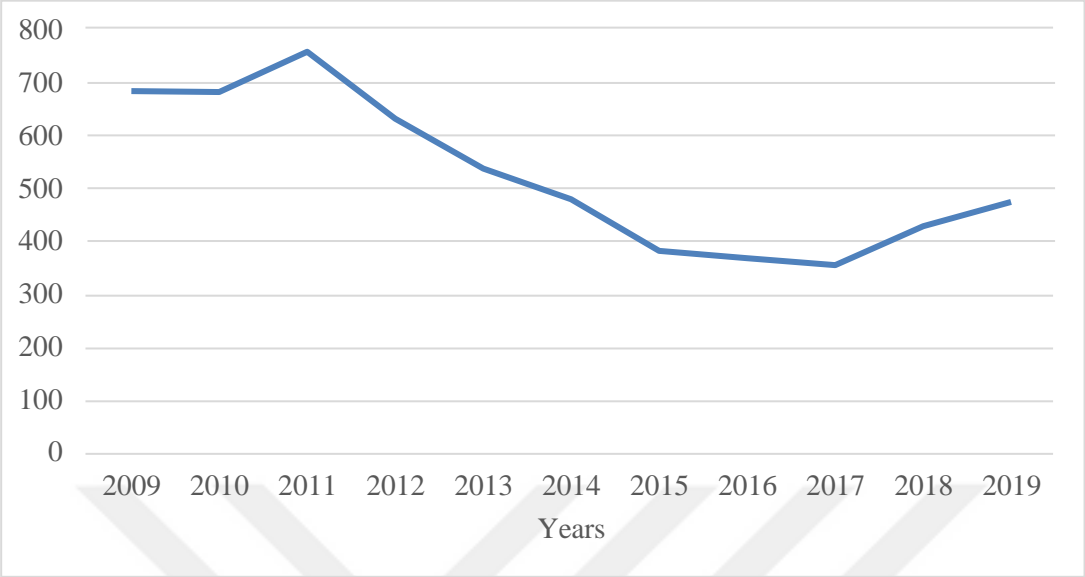


Figure 3.5: Natural gas production of Turkey by years (Milion Sm³) (EPDK, 2019).

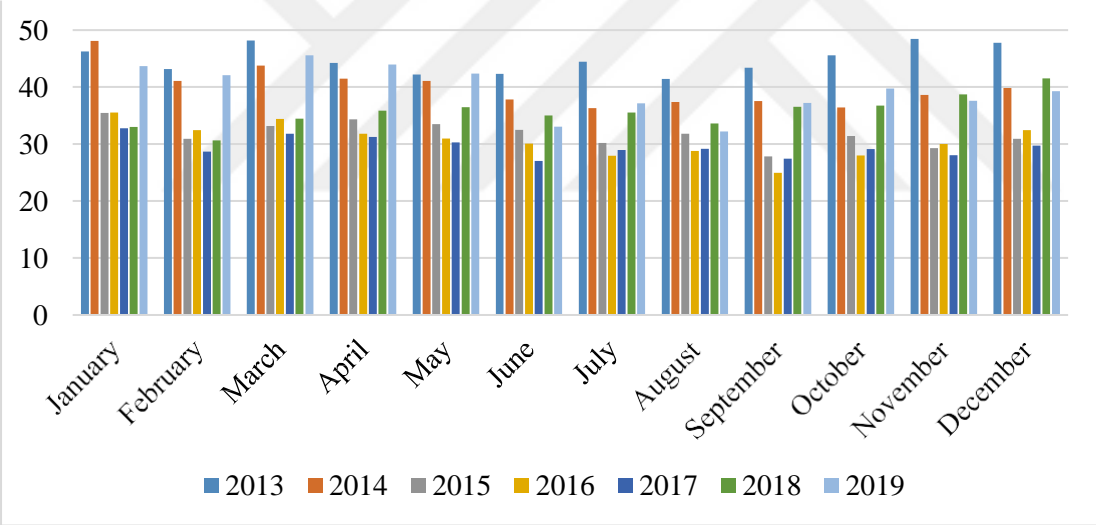


Figure 3.6: Monthly comparison of natural gas production between 2013-2019 (Milion Sm³) (EPDK, 2019).

Turkey has discovered a 320 Bcm of natural gas reserve in the Black Sea in 2020, recently. In addition to the direct economic impact of the local reserve, it is strategically important reducing the dependence on imports in the energy sector. Calorific value, purity of the discovered natural gas and the possibility of connecting to the existing distribution network via pipeline provide economic feasibility to the new reserve. It is predicted that the gas reserve will begin to be used in 2023 with an investment of 3-4 billion dollars (Url-13).

3.2.2 Natural gas imports of Turkey

The insufficient natural gas reserves of Turkey can not meet the natural gas energy demand therefore Turkey imports natural gas from other countries using different ways. Turkey mostly prefers to make long-term contracts to ensure the gas supply. Long term contracts important for energy security. Long term energy contracts mostly have take or pay obligation which guarantees both the source country and the consumer country. Take or pay obligations sometimes provides a flexibility for Turkey. According to the flexibility of these natural gas contracts, Turkey also purchases natural gas from spot markets.

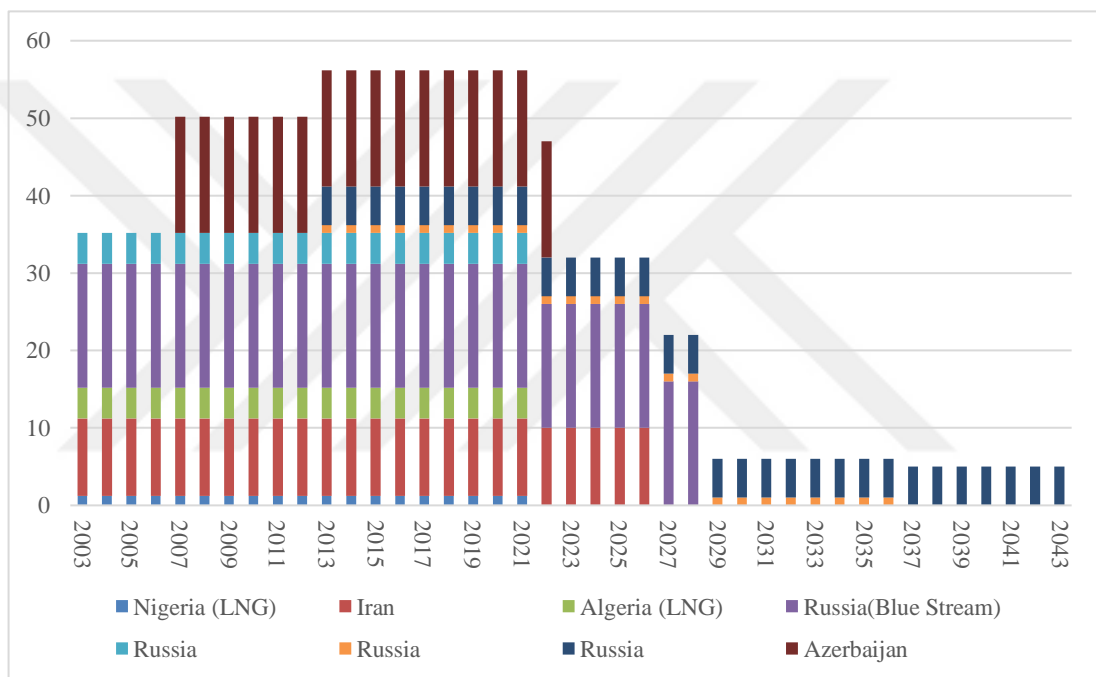


Figure 3.7: Long-term natural gas purchase contracts of Turkey (Billion Sm³) (Url-12).

Turkey's long-term natural gas contracts can be seen at Figure 3.7. Turkey has a long-term natural gas purchase agreement with 5 different countries. Natural gas trades are carried out with Russia, Azerbaijan and Iran with pipelines. There are also long-term LNG contracts with Algeria and Nigeria. Turkey is an oversupplied country in terms of long-term natural gas purchase contracts and consumption between 2009 and 2019. However, some of natural gas purchase contracts of Turkey with Russia and Algeria will expire in 2021. In addition to that, another natural gas purchase contract with Azerbaijan will expire in 2022. These contracts which will expire are a supply threat to Turkey. It is still unknown how natural gas will be supplied after these contracts.

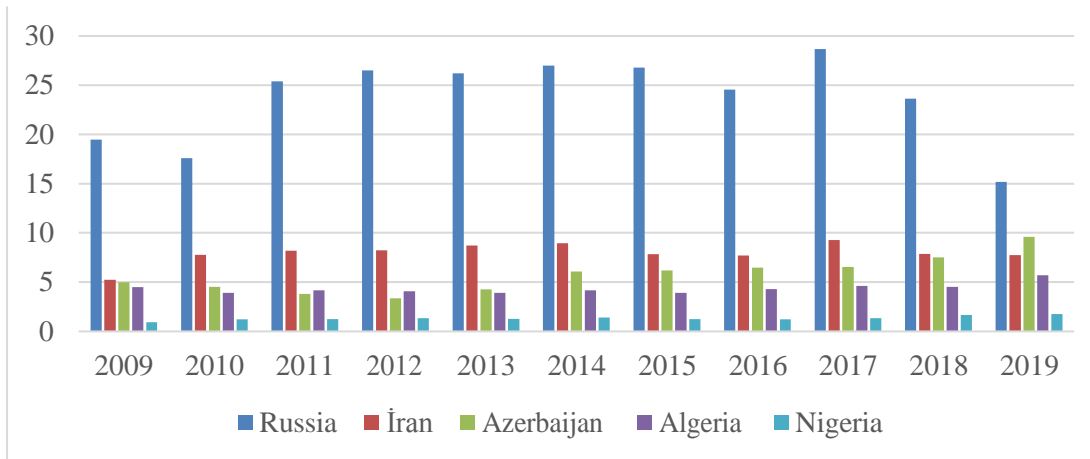


Figure 3.8: Natural gas import amounts by the country for 2009-2019 (Billion Sm³) (EPDK, 2019).

As figure 3.8 shows the natural gas has been imported mostly from Russia via pipelines in recent years. In other words, Turkey's main natural gas supplier is Russia. Turkey began a balanced policy in the long-term natural gas purchase contracts after 2017. In this way Turkey has benefited more from cheaper spot LNG opportunities. Turkey met an average of %50 of its natural gas needs from Russia before 2017, while due to the balanced policy, this ratio decreased to 33%, as shown in figure 3.9.

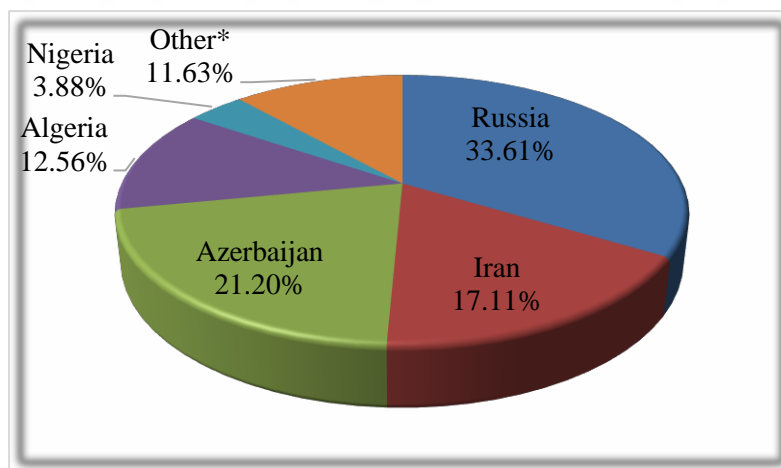


Figure 3.9: Share of imported natural gas by source country in 2019 (%) (EPDK, 2019).

Natural gas is an important primary energy source for Turkey. Therefore, natural gas supply security is highly important. In the past years Turkey has to face the issue of natural gas supply security. For example, due to the crisis between Russia and Ukraine in 2006, the amount of natural gas coming from the West Line decreased. In addition to this, Iran had some natural gas supply problems during winter. Because of such

reasons, LNG infrastructure was developed to ensure supply security. LNG purchases have been started from Algeria since 1994 and from Nigeria since 1999 (EPDK, 2019). Turkey's long-term contracts with Nigeria and Algeria have been actively used since the early 2000s. In addition, it is understood that Turkey has been buying more LNG from time to time than the amount of the contracted volumes since 2013. The annual amount of natural gas that Turkey buys from Algeria and Nigeria is also seen in figure 3.10. On average, 4 billion Sm^3 natural gas is purchased from Algeria and 1.2 billion Sm^3 natural gas is purchased from Nigeria annually. Roughly the amount of LNG taken from these two countries meet 12% of the annual natural gas needs.

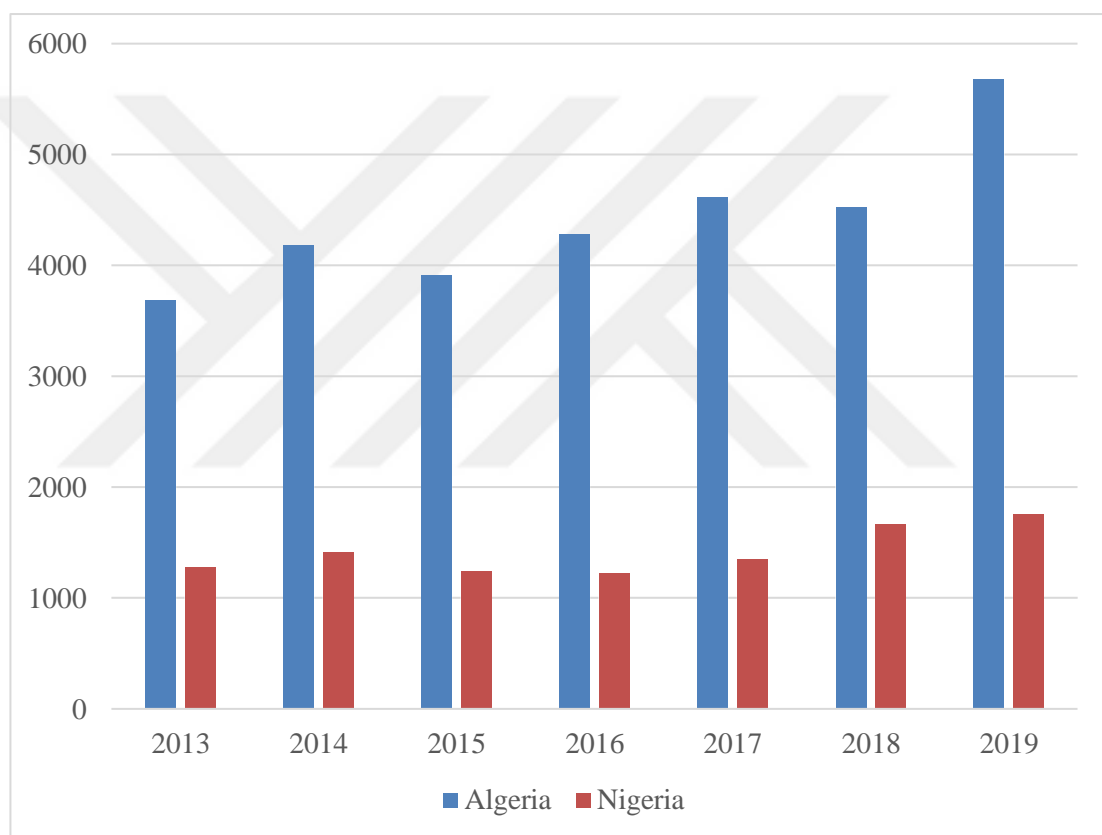


Figure 3.10: Long Term LNG Import Amounts by Source Countries between 2013-2019 (Million Sm^3) (EPDK, 2019).

Turkey's LNG and pipe gas imports between 2013 and 2019 can be seen in Figure 3.11. Since 2013, there has been a regular increase in LNG imports. There are several reasons for this. First of all, Turkey has given importance to LNG and invested in LNG infrastructure in recent years. Currently, there is a surplus of supply in the world on LNG. With these LNG investments, Turkey bought natural gas below the cost of the gas it receives through pipelines. Most analysts interpret the supply surplus in LNG as positive for Turkey. (WIDDERSHOVEN, 2017).

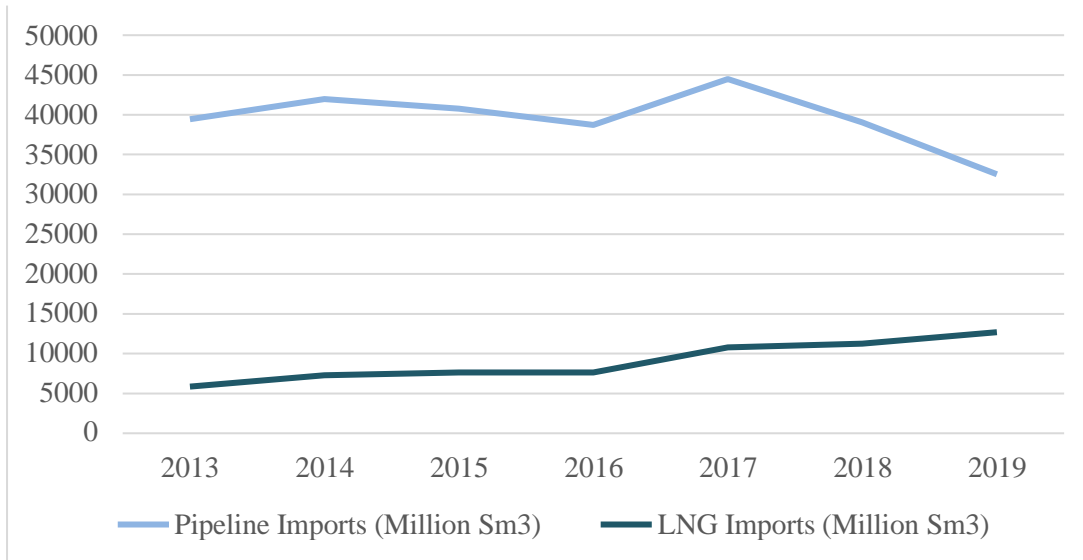


Figure 3.11: Import amounts of natural gas by type (Million Sm³) (EPDK, 2019).

Spot natural gas imports were allowed in 2018 with the amendment to the natural gas law numbered 4646. With this amendment, private companies that will import spot LNG, spot CNG and spot pipe gas are required to obtain a spot license. In other words, it has been possible to import spot gas in many different ways with a single license. In recent years, Turkey has benefited from global spot due to balanced usage of natural gas contracts and infrastructure developments. Spot LNG imports increased by more than 100% as seen in Figure 3.12 after 2016. Spot LNG imports were carried out from many different countries in 2019. Some examples of these countries are USA, France and Qatar (EPDK, 2019).

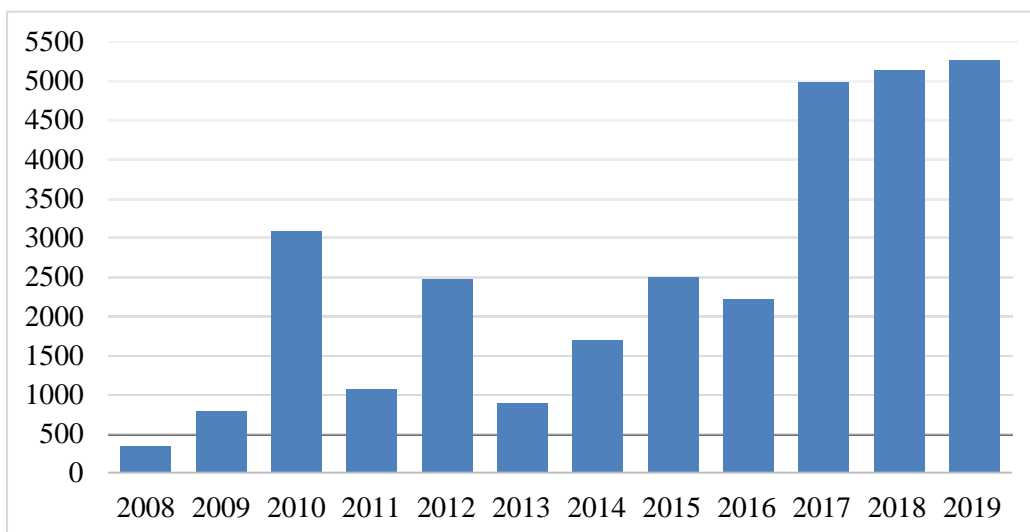


Figure 3.12: Change in spot LNG import amount between 2008-2019 (Million Sm³) (EPDK, 2019).

3.2.2.1 Crossborder pipeline infrastructure of Turkey

Turkey has a very important geopolitical position in terms of natural gas supply security. This is very crucial for Turkey for its goal of becoming a natural gas hub. For this reason, Turkey must first secure its natural gas supply and then secure the natural gas supplies of European countries, which are major natural gas consumers since Turkey is like a bridge between countries having rich natural gas resources in the East and Europe.

Russia - Turkey natural gas pipeline (west line)

Turkey's first natural gas trade took place with Russia over the Russia-Turkey pipeline. However, there were interruptions on this line especially because of the political crisis between Russia and Ukraine. To prevent this problem and secure natural gas supply, the Turkish stream pipeline was constructed. Turkey's long-term contracts on the Russia-Turkey Natural Gas Pipeline were transferred to the Turkish stream pipeline (Url-14).

Blue stream pipeline

Blue stream pipeline was built under the 25-year natural gas purchase agreement signed between BOTAS and Gazexport in 1997. Blue stream pipeline reaches across the Black Sea from Russia to Turkey. According to the purchase agreement, annual 16 Bcm of natural gas was supplied to Turkey via Blue Stream pipeline (Url-14).

Eastern Anatolian natural gas transmission line (Iran - Turkey)

Eastern Anatolian Natural Gas Transmission Line was built after the natural gas purchase and sales agreement with Iran. It was aimed to supply 10 billion sm³ of gas annually. Gas flow from Iran started on this line on December 10, 2001. However, there exists a supply problem during the winter months when the domestic consumption need of Iran increases (Url-14).

Baku-Tbilis-Erzurum natural gas pipeline

Baku-Tbilis-Erzurum natural gas pipeline was built to supply natural gas to Turkey which was produced at the Southern Caspian Sea (Url-14).

TANAP pipeline

On 26 June 2012, an agreement was signed between Azerbaijan and Turkey on the transportation of 10 Bcm of Azerbaijani natural gas to Europe by the TANAP pipeline. This line also supplies natural gas to Turkey. The first gas flow to Turkey started in June 2018. In 2019, this pipeline is connected to Europe (Url-14).

Turkish stream pipeline

Turkish Stream Pipeline Project is a new gas pipeline system consisting of two lines starting from the Russian Federation and coming to Turkey via the Black Sea, each with an annual capacity of 15.75 billion m³. Also this line was built to supply Russian natural gas to Europe via Turkey (Url-14).

3.2.2.2 Crossborder LNG structure

In some cases, there are technical and economic constraints in terms of the transmission of natural gas through pipelines. In these cases, natural gas is cooled down to -162 degrees and liquefies. In this way, the volume of natural gas is reduced up to 600 times. Liquefied natural gas is transported by specially manufactured tankers (EPDK, 2019).

There are four different LNG terminal in Turkey. Firstly, Marmara Ereğlisi LNG terminal started to be used. Then, in 2006, Ege Gaz LNG terminal was started to be used. Then Turkey's first FSRU terminal began to be used in 2016 in Izmir. Finally, Hatay Dörtyol LNG (FSRU) facility received a license in 2017. (EPDK, 2019).

3.3 Natural gas storage of Turkey

Natural gas storage is important for grid management. The storage requirement depends mainly on demand volatility. If a country can meet its natural gas demand with its resources, there may be no need for storage. The demand volatility can be adjusted by the natural gas traction from gas fields. However, when a flow problem occurs in a country that mostly import natural gas or has high demand volatility, natural gas importer countries have to store natural gas in to meet the demand and to manage the network properly. Practices have shown that as the amount of imported gas increases to meet the increasing consumption, storage becomes more necessary (Öztürk, 2007).

The natural gas storage investments of Turkey is increasing every year. As a result of these investments, both natural gas storage capacity and LNG storage capacity have been increasing. In 2019 Turkey's natural gas underground storage capacity was 3,391 million Sm³. The storage capacity of LNG terminals reached 0,964 million Sm³ (EPDK, 2019)

Natural gas storage facilities are used especially in winter. Back natural gas production is carried out from storages to meet the increasing consumption need during winter. In figure 3.13 it can be seen that during winter months usually stored natural gas is used therefore stock levels were decreased. When the data of 2020 were examined, it was seen that 7-9% of the consumption in the winter months was met from storages. Usually, no natural gas was released from the storage facilities in May, August and September. During these months, preparations were made for the winter months and natural gas was injected into the tanks.

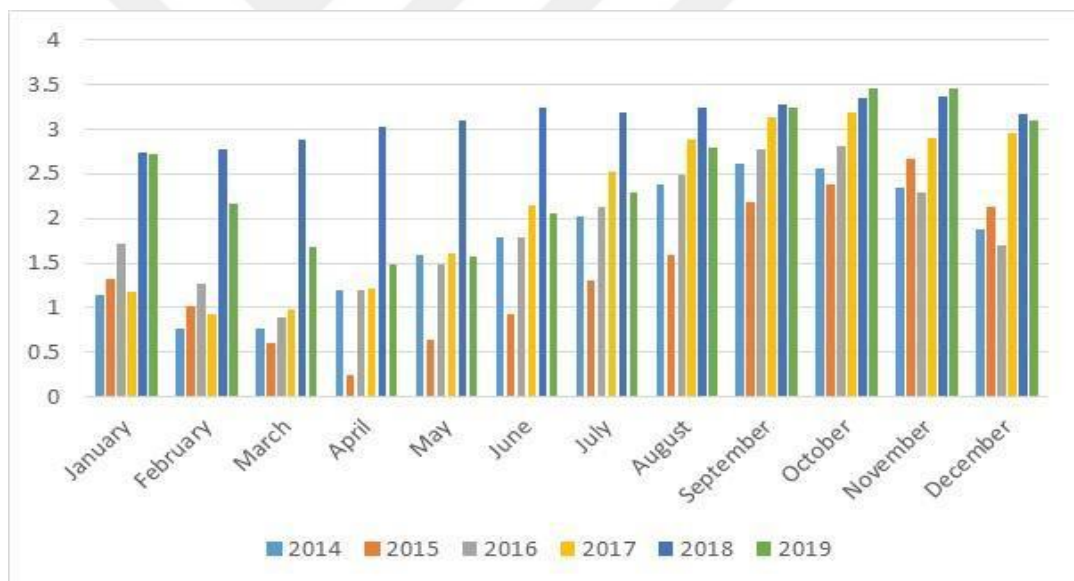


Figure 3.13: Month end natural gas stock amounts by years (Million Sm³) (EPDK, 2019).

3.3.1 Storage prices

Operation costs are very important in terms of using natural gas storages. Low storage costs provides flexibility to portfolio optimizers. A consumer who wants to use the Silivri Storage Facility has to pay different costs such as capacity fee, injection and back production fee, gas exchange fee, service fee etc. Also customers that want to use the storage facility pay daily rent. However there are 3 different costs that a consumer who wants to transport natural gas through pipelines must pay. These are entrance

capacity fee, output capacity fee and transmission service fee. Comparing the costs of pipelines and the costs of the Silivri Storage Facility, the storage cost for 1 Sm³ is 4.3 times more than the cost of transmission with pipelines. While making this calculation, it is assumed that natural gas will be stored for only one day. This ratio will increase as the number of days that natural gas will be stored in storage increases. It is clear that the private sector will not be able to use the storages for optimizing assets and portfolios due to the high storage costs.

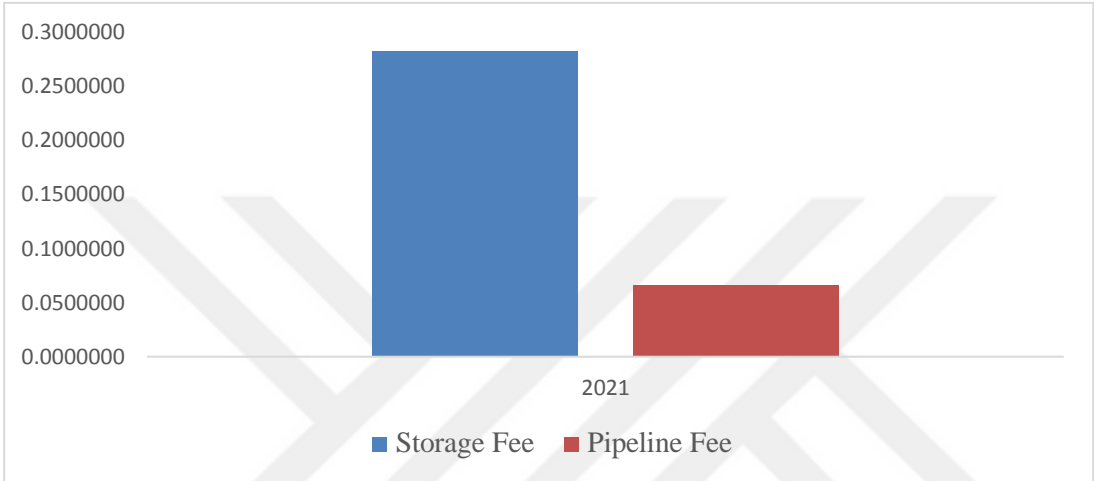


Figure 3.14: Comparison of the storage cost of 1 Sm³ natural gas and the cost of pipeline transmission of 1 Sm³ in 2021 (TL\ Sm³) (Url-15).

3.3.2 Natural gas storage facilities of Turkey

There are four natural gas storage facilities in Turkey. The first of these is the Silivri underground natural gas storage facility. It started to operate in 2007. The total storage capacity is 2,841billion Sm³.The maximum withdrawal amount of the storage facilities is 25 million Sm³/Day and the maximum injection amount is 16 million Sm³/ Day (Url-16).



Figure 3.15: Silivri Storage Facility (Url-16).

There is another natural gas storage facility located in Aksaray. It's located approximately 40 km south of Tuz Lake. Salt layers were used for storing natural gas. It has 20 million Sm³ daily withdrawal capacity (Url-17).



Figure 3.16: Tuz Lake natural gas storage facility (Url-17).

Marmara Ereğlisi LNG Terminal, which was started to be built in 1989 and started to operate in 1994 and has 3 storage tanks. The storage capacity of tanks is 85,000 Sm³. This terminal has three main functions. These are storing imported LNG, gasifying the stored LNG and filling LNG into land tankers (Url-18). Turkey have also an FSRU terminal. FSRU ships performs the functions of storing LNG and transferring LNG to the natural gas network by gasifying LNG. Hatay Dörtyol FSRU terminal started to operate at 07.02.2018. It has 263,000 Sm³ storage capacity and 2 million Sm³ gasification capacity (Url-19).



Figure 3.17: Marmara Ereğlisi LNG terminal (Url-19).

3.4 Natural Gas Consumption of Turkey

Figure 3.18 shows the total natural gas consumption in Turkey between 2010 and 2019. Since 2010, growth has been observed in Turkey's natural gas consumption. However, this growth did not occur continuously. For example, there has been a decrease in natural gas consumption after 2017. Natural gas is used for many different purposes in Turkey. Therefore, natural gas consumption depends on many different factors. For example, domestic natural gas consumption depends on air temperature or industrial consumption depends on industrial production. Figure 3.19 shows the different usage areas of natural gas and the fluctuations according to the sector. Each sector has different dynamics. According to the state of these dynamics, the amount of natural gas consumption varies annually. With the widespread use of natural gas, the parameters that the consumption amount depends on will increase.

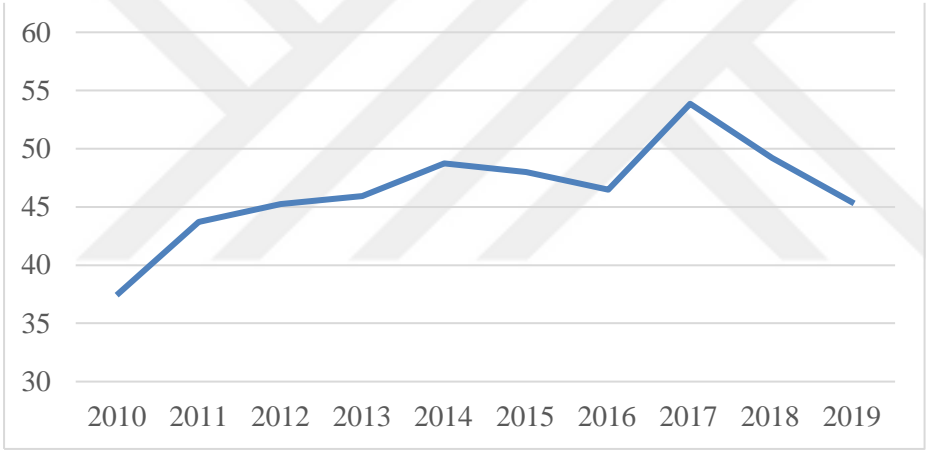


Figure 3.18: Yearly natural gas consumption of Turkey (Billion Sm³) (EPDK, 2019).

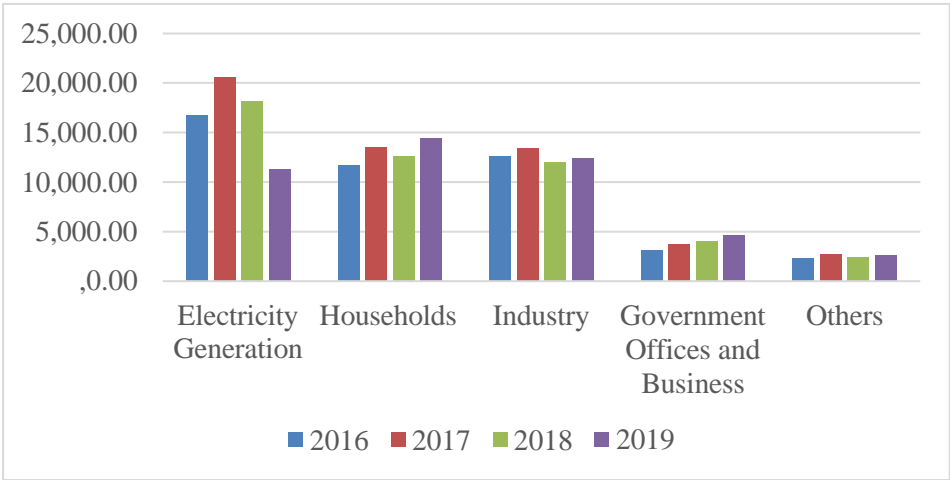


Figure 3.19: Annual comparison of sectoral consumption (Million Sm³) (EPDK, 2019).

One of the most important factors affecting natural gas consumption is temperature. Household consumption increases when weather temperature decreases. Istanbul is the biggest city in Turkey also it's the city with the most household natural gas consumers. Figure 3.20 shows the average temperature of Istanbul in 2019 and the monthly consumed natural gas in 2019. The decrease in the weather temperature increased natural gas consumption. The seasonal effect of temperature on natural gas consumption is clearly seen here. The same consumption profile and seasonality effect can be observed for total natural gas consumption of Turkey. Figure 3.21 shows the 2019 natural gas consumption profile of Turkey. Household consumption in the winter months is considerably higher than in other sectors.

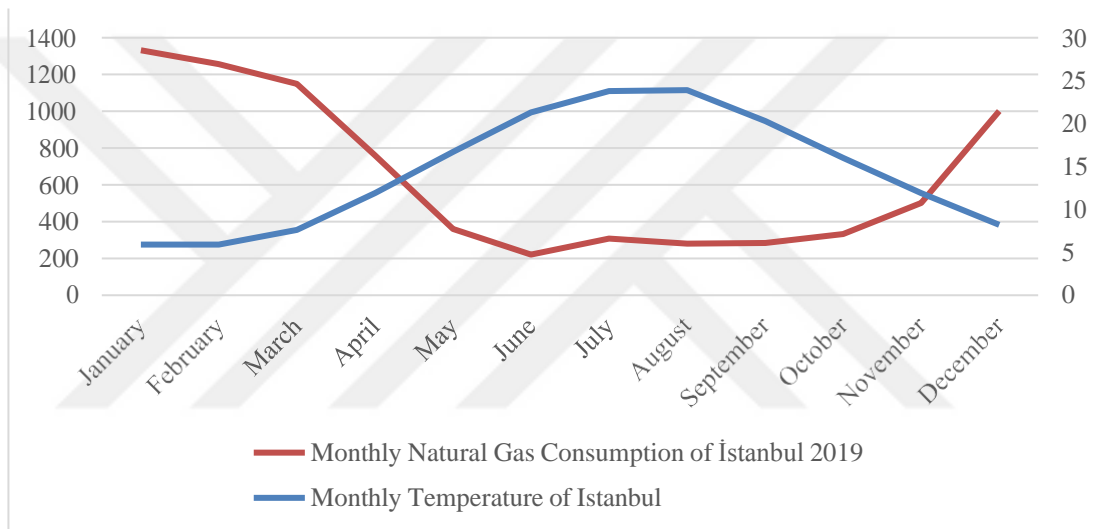


Figure 3.20: Effect of temperature (°C) on Istanbul household natural gas consumption (million Sm³) (EPDK, 2019), (Url-20).

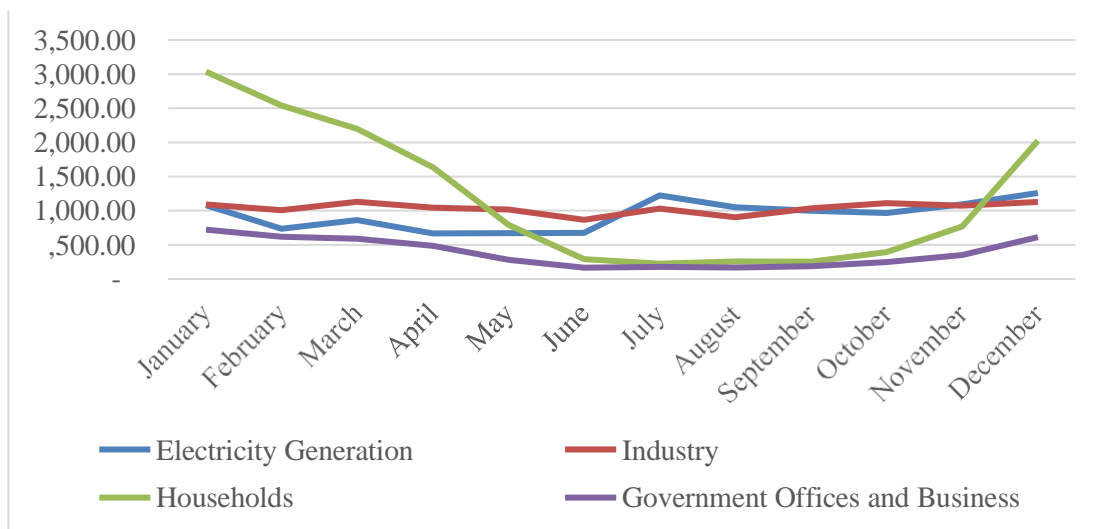


Figure 3.21: Monthly changes of sectoral consumption in 2019 (Million Sm³) (EPDK, 2019).

Natural gas is used in many different sub-sectors in the Turkish industry. Glass, ceramic, cement producers, fertilizer producers or the iron and steel industry can be given as examples. In Figure 3.22, the monthly natural gas consumption and industrial production index of the industry in 2019 are given. As can be seen in Figure 3.22, there is a strong relationship between the natural gas consumption of the Turkish industry and the industrial production index. As the production of Turkish industry increases, natural gas consumption also increases. That's why continuation of gas supplies to the increase of industrial production in Turkey is required.

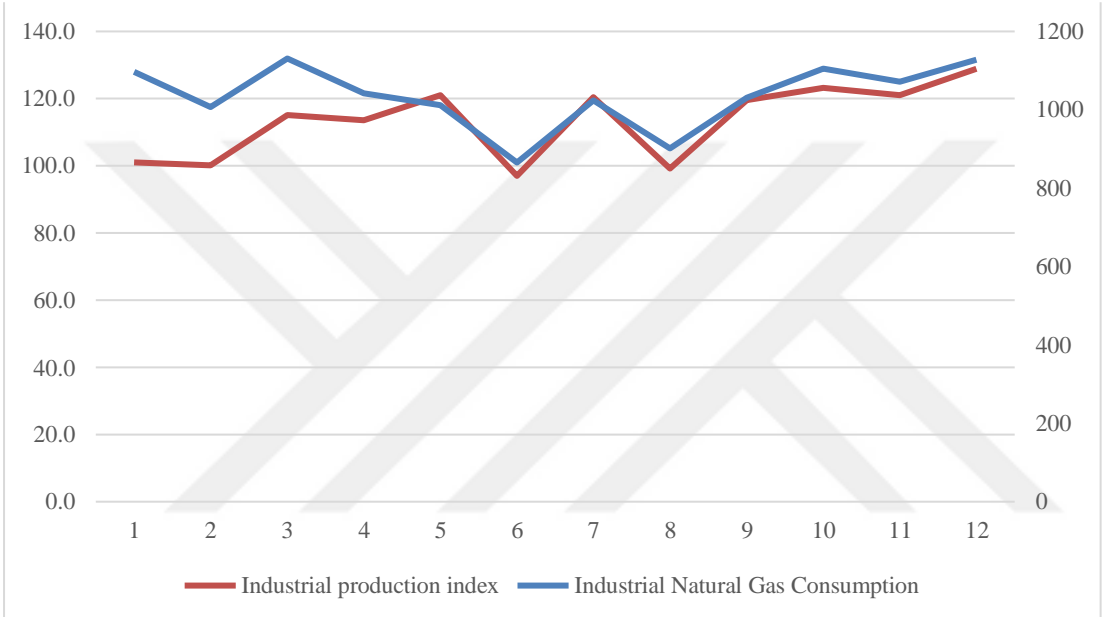


Figure 3.22: Monthly industrial natural gas consumption (Million Sm³) and industrial production index (EPDK, 2019) (Url-21).

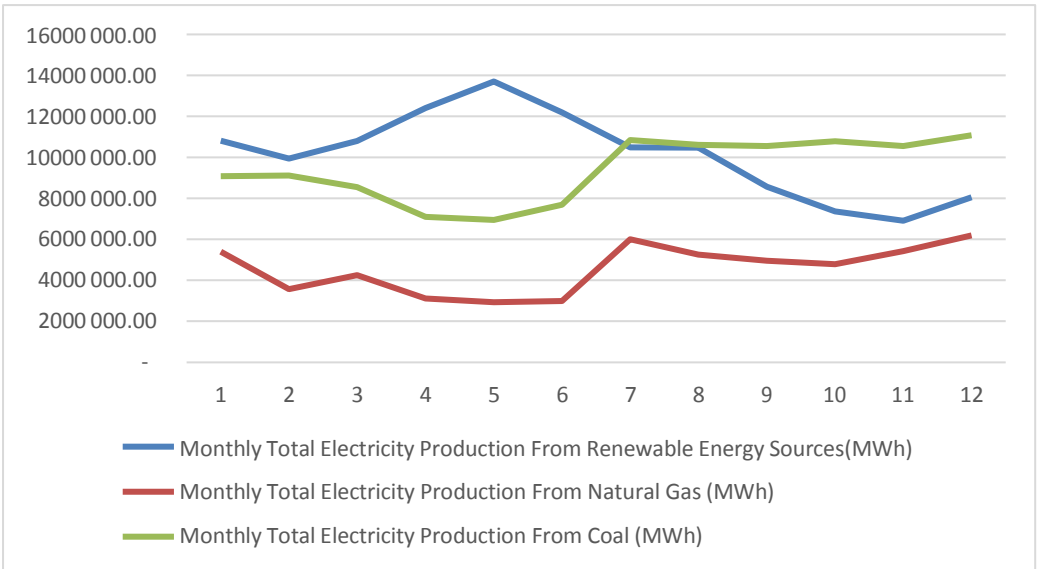


Figure 3.23: Monthly electricity production (MWh) according to the sources in Turkey in 2019 (Url-22).

Figure 3.23 shows the amount of monthly electricity production of Turkey in 2019 according to sources. The use of natural gas for electricity generation in Turkey is quite common. However, the share of natural gas in Turkey's electricity production is constantly changing. Natural gas is not the first choice for electricity generation. Due to low production costs, Turkey's energy needs are met primarily from renewable sources. After renewable energy sources, coal power plants are used to meet the electricity demand. Turkey's coal reserves are greater than its natural gas reserves. However, also coal is imported for electricity generation. If the electricity demand is still not met, natural gas power plants start to operate. As can be seen in Figure 3.23, when electricity production from renewable energy sources decrease, the amount of production from natural gas and coal plants increases. In some cases, the cogeneration facilities of the factories operate as a base load. These cogeneration systems can be given as an example of the on-site consumption strategy.

3.5 Natural Gas Export

The natural gas exports of Turkey is very limited. To export natural gas required licenses must be obtained. Exporter companies can export natural gas to the countries mentioned in their licenses. Turkey usually export natural gas to Greece and Bulgaria. Physical gas flow is used through pipelines for Greece and the gas is exported as LNG to the Bulgaria.

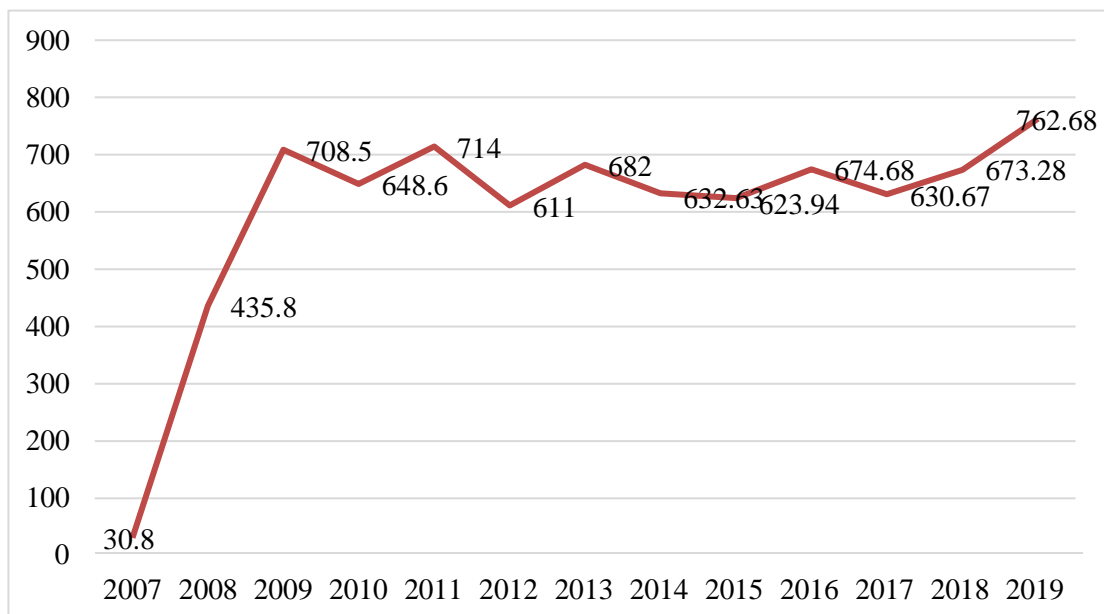


Figure 3.24: Annual natural gas export amounts (Million Sm³) (EPDK, 2019).

3.6 Market Structure Of Turkey

Turkish natural gas market, unlike the electricity market, is liberalizing more slowly and in a controlled way. Liberalization in the Turkish natural gas market began in 2001 with the enactment of the natural gas Law No. 4646. General aims of 4646 Natural Gas Law are as follows:

- Market participants will be licensed by EPDK,
- BOTAS activities will be reorganized
- BOTAS’s trade activities will be reduced.
- Eligible consumer limits will be adjusted.
- New distribution areas will be operated by private sector.

4646 Natural gas market law aims to divide BOTAS’s activities into two groups as trade and infrastructure functions. Also BOTAS's monopoly on natural gas imports, sales and pricing has started to be prevented. Botaş transferred its long-term contract of 4 Bcm to the private sector through the liberalization process in 2005. But despite all these efforts, BOTAS is still a monopoly in the Turkish gas market. By the means of this delay in the liberalization of the market, the impact of the increase in the price of the international gas market, is eased on Turkish industry (WEC, 2018).

The development of the wholesale market continued with the liberalization of spot LNG imports in 2008. Also an amendment to the network operation regulations in 2008 made it possible to trade natural gas in a virtual environment. Despite all these developments, it can be said that the Turkish gas market is still not completely liberal and under the control of BOTAS.

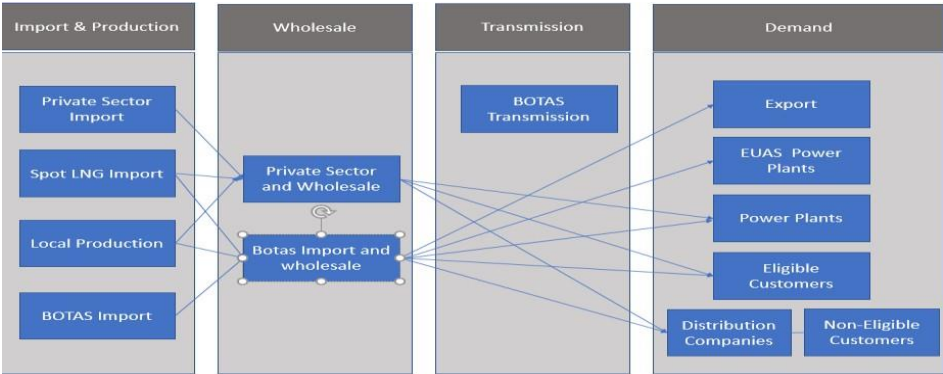
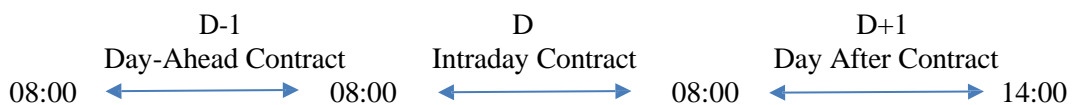


Figure 3.25: Turkish natural gas market structure (Turkey's Energy Transition Milestones and Challenges, 2015).

The general structure of the Turkish gas market is seen in Figure 3.25. Turkey's natural gas supply is provided from 4 different sources. These are local production, spot purchases, private sector imports and BOTAS imports. Wholesale companies and BOTAS transfer physical natural gas to customers through BOTAS Transmission Company. Private sector importers can not sell directly to customers who are not eligible. Only distribution companies can sell natural gas to non-eligible customers. But private sector wholesalers can make transactions with distribution companies. Companies that buy natural gas for electricity generation can buy gas from BOTAS or private sector natural gas wholesalers. Natural gas power plants can purchase natural gas from BOTAS and the private sector at the same time.

3.6.1 Organized natural gas wholesale market

In 2016, it was announced that an organized natural gas market would be established by EMRA. After the announcement in 2017, Organized Natural Gas Wholesale Market Usage Procedures and Principles were published (EPDK, 2019). In this way, the first official step was taken to form an organized natural gas market. This regulation aimed to allow market participants to trade without interruption in an anonymously organized liberal market. The organized wholesale natural gas market started on April 1, 2018. Organized Natural Gas Wholesale Market organized by Energy Exchange (EPIAŞ). The Organized Natural Gas Wholesale Market allowed the transmission company (BOTAS) to balance the number of stocks in pipelines, balance the market participants portfolios, trade natural gas before the day and after the day and create a market-based pricing system (Url-23).



D: Delivery day.

Figure 3.26: Wholasure natural gas market contract timelines.

The gas delivery day starts at 08:00 in the Turkish market. A contract is opened on the spot natural gas market for each natural gas delivery day. A contract on the spot natural gas market can be traded for 54 hours. Contracts are called day ahead contract for the first 24 hours after opening. They are then called intraday contracts for the next 24 hours. For the next 6 hours, contracts are called the day after contract.

Market participants trade natural gas in a day-ahead contract according to the day-ahead planning. In the intraday contract, participants who want to change market position due to a malfunction or change in planning usually make transactions. In addition, a natural gas power plants that receive loading or deloading instructions make transactions in the intraday market. The day after contract is used by suppliers, market participants or natural gas plants to balance their portfolios.

Organized Natural Gas Wholesale has many different prices as an indicator. The day-ahead price is the weighted average of the transactions made in the day ahead contract. Intraday price is the weighted average of transactions of intraday contracts. Day after price is the weighted average of the transactions made on the day after contract. The gas reference price is determined as the weighted average of the transactions made when a contract is a day ahead contract and when it is an intraday contract. BOTAS Transmission Company makes a transaction every day at 12:30 to balance the amount of imbalance in the transmission network. At 12:30, it automatically matches the best offer in the market. In this way, an additional balancing buy or sell price is determined. In other words, the purchase or sale price of BOTAS Transmission Company determines the additional balancing price. Balancing gas purchase and sale prices are used to reconcile the imbalances of wholesale companies or natural gas combined cycle power plants. For each day, that is, for each natural gas contract, the purchase and sale price of the balancing gas is determined by the following formulas.

$$ABPP = \text{Max}((GRF \times 0.93), BGPP) \quad (4.1)$$

$$ABSP = \text{Min}((GRF \times 0.97), BGSP) \quad (4.2)$$

BGPP: Balancing Gas Purchase Price

BGSP: Balancing Gas Sale Price

GRF: Gas Reference price

ABPP: Additional balancing purchase price

ABSP: Additional balancing sale price

3.6.2 Pricing of natural gas in Turkish Gas Market

Natural gas pricing can be considered as domestic pricing and export pricing. Domestic pricing can be considered as specialized pricing for wholesale and

specialized pricing for retail sales. Pricing strategies for wholesale natural gas can be listed as follows:

-Oil price escalation: There is a base price in this pricing. The change in the base price depends on the commodities on which the price of natural gas links. These commodities are competing fuels. For example gasoil, oil and fuel oil (IGU, 2019).

-Gas-on-gas competition: Which is pricing where supply and demand become important. This pricing and competition are seen in physical hubs (IGU, 2019).

-Bilateral Monopoly: These prices occur between a large buyer and seller. For example, governments or state-owned companies. Prices are fixed for certain time intervals (IGU, 2019).

-Netback from the final product: This pricing is generally applied where natural gas is used as raw material in chemical factories (IGU, 2019).

-Regulation based cost of service: The price of natural gas is determined by a regulatory authority, but the price is adjusted to cover the "Cost of Service" and a reasonable rate is set for recovery of the investment (IGU, 2019).

-Regulation based below cost pricing: The price is consciously determined by an authority below its cost, and natural gas prices are usually given to a given population by subsidizing prices (IGU, 2019).

-Regulation based political pricing: Wholesale price is altered irregularly, possibly by a ministry. Natural gas prices are determined by a mixed strategy based on the purpose of increasing income, social needs and covering costs (IGU, 2019).

Natural gas prices in Turkey are still determined by BOTAS, the dominant importer. Given that BOTAS's contract transfer program is progressing slowly, it is expected that prices will remain under BOTAS's influence in the medium term. Since the weighted average cost of gas is the most important factor, BOTAS must reflect these costs within a reasonable period of time. But BOTAS does not always price to reflect its costs. There could be many reasons for this. The most important reason is Turkey's efforts to provide cheap gas supplies to the Turkish people. Because of these subsidies, BOTAS has announced losses in the past (Url-24).

BOTAS tariff includes transmission cost and capacity cost. However, the price of gas is determined independently from transmission, capacity or storage prices. BOTAS

gas prices are also independent of customer consumption profiles and related imbalance costs (Yağız, 2016).

BOTAS's long-term pipeline import contracts are priced based on international oil prices. Spot LNG purchases depend on the international spot market. Factors such as the amount of gas imported under each contract and the exchange rate constitute the average cost of gas. The cost of private import companies depends on contracts between these companies and suppliers

Analysis shows that there is a relationship between oil prices and natural gas prices in Turkey. In natural gas trade agreements where petroleum-based substitution pricing is performed, there is a relationship between the 3-6-9 monthly shifted averages of Brent oil prices and natural gas prices (Yağız, 2016).

4.MACHINE LEARNING

The purpose of machine learning is to learn from previous data, discover patterns and make decisions accordingly. The history of machine learning dates to 1990. With the development of technology, control has become difficult at some points. At the same time, with the increase in the amount of data produced, it is difficult to use these data. But with the development of processor technology and machine learning techniques, the process of using this excess data produced has been facilitated. Machine learning combines a lot of discipline, such as statistics and data analysis. The area of use of machine learning techniques is becoming more and more common. For example, personalized advertisements used on the internet, fraud detection, spam filtering, estimation of maintenance times, etc. (Alpaydm, 2010).

Machine learning can extract patterns from the data that are difficult to find with human effort. From the relationships contained in the data set, future predictions can be made and new relationships can be detected. In other words, machine learning learns from past data and tries to explain today and tomorrow. For example, let's think of a supermarket chain that has lots of products, stores, customers and workers. This market saves huge amounts of data every day. For example, sales volumes, product codes, or the amount of money customers spend. So, what can this supermarket do with this data? It can reveal patterns in this large data using machine learning algorithms. In this way, it can understand the shopping habits and provide better service to the customers locally. Or it can adjust the necessary adjustments for its stocks with machine learning. Machine learning has many different uses. For example, banks use machine learning methods for fraud detection or credit risk analysis. It is used for diagnosis in medicine. In addition, it is also used in image, sound analysis or robotics (Alpaydm, 2010).

4.1 Supervised Learning

Supervised Learning is a machine learning approach. Simply, supervised learning is carried out by revealing the relationship of past data. In supervised learning, a training

dataset and output dataset which are both labeled are provided to the model. The aim is for the supervised learning model to develop a rule or procedure by examining the training data and obtaining the output dataset. Supervised learning is generally used in two different ways. If the result to be predicted is continuous data, such as velocity or length, it is a regression problem. Sometimes the result data can be a category. Like color for example. These kinds of problems are called classification problems (Learned-Miller, 2014).

4.2 Unsupervised Learning

Unsupervised learning is a machine learning approach. Unsupervised learning is used when we do not know how the output from the data should be. Models are created by clustering data based on relationships between input variables. The main purpose here is to reveal the relationship between input data with finding similarities. Clustering or factor analysis are examples of unsupervised machine learning algorithms.

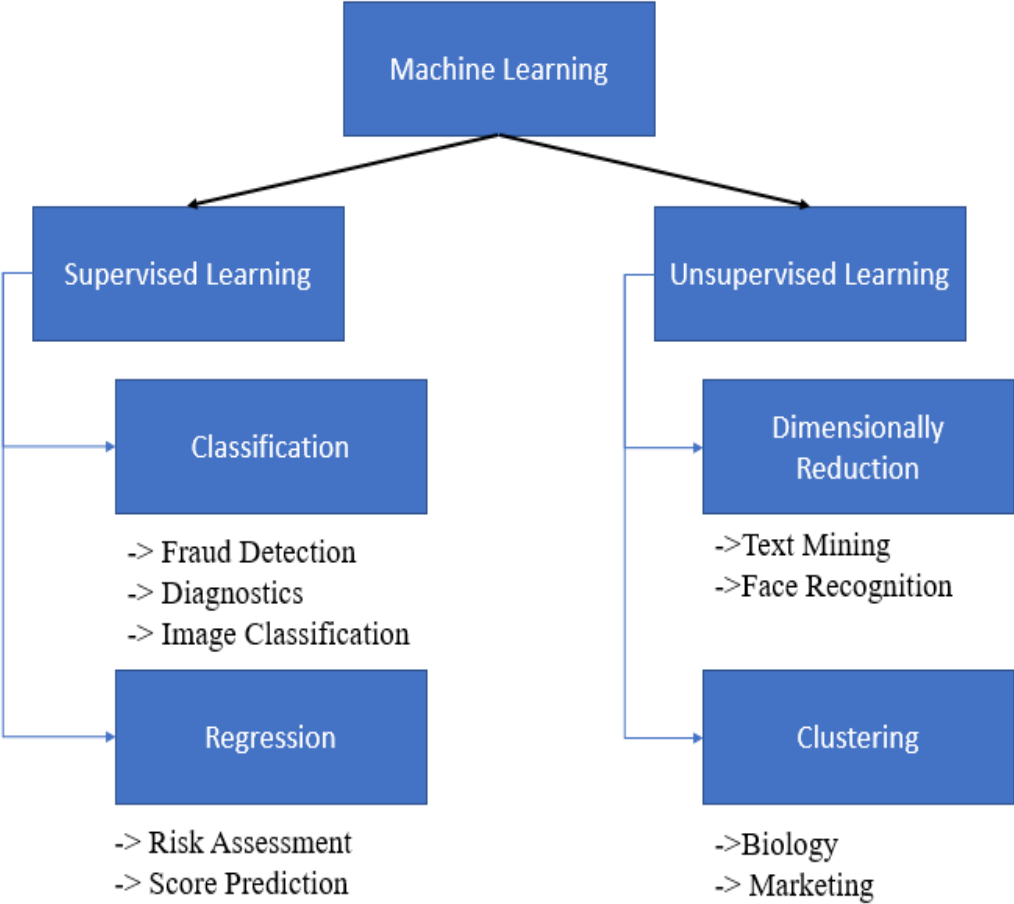


Figure 4.1: Machine learning techniques.

4.3 Ensemble Learning

When using machine learning models, it is often difficult for a single model to be successful. Therefore, solutions in which many different models are used together have been developed. The advantage of community learning algorithms is that more than one models are combined and give results that are close to the most ideal solution. In other words, it is the effort of multiple weak models to produce strong solutions together instead of a single machine learning model. This is just like a patient taking the opinion of more than one medical specialist before undergoing any surgical operation. In other words, it is like evaluating more than one opinion before a decision is accepted (Demiraslan, 2020).

4.4 Bootstrap

Bootstrapping is one of the sampling methods used in statistics and bootstrapping can be applied to many different statistical learning methods easily. Simply, different subsets are randomly generated from an input dataset. Some values in the dataset repeat when creating subsets. Models are developed with these subsets. There are many advantages to bootstrapping. For example, it helps prevent overfitting also by creating different datasets, ensemble learning algorithms are enabled to learn different combinations. In this way, low variance is obtained.

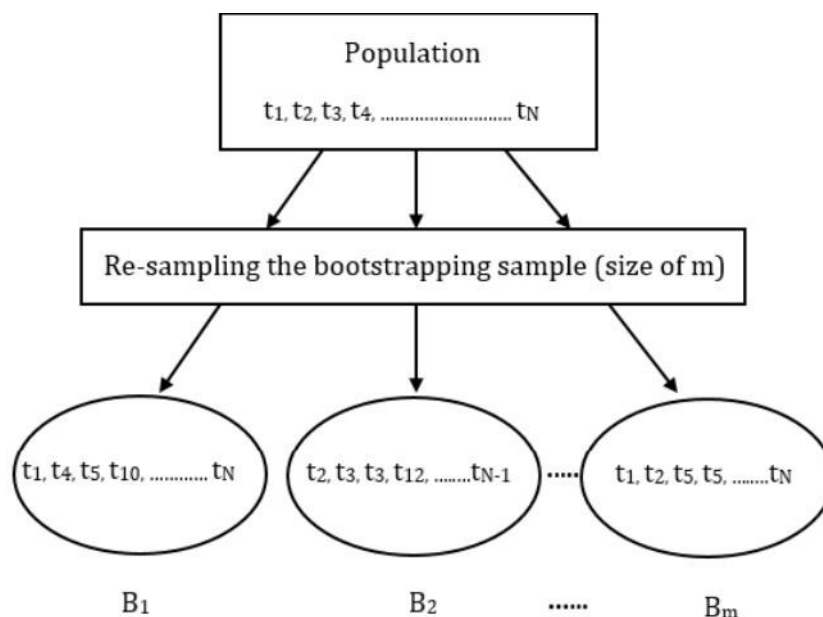


Figure 4.2: Bootstrap sampling procedure (Rahman, 2019).

4.5 Bagging

Bagging is an easy community-based algorithm. Bagging is a shortened form of the words aggregating and bootstrapping. It was discovered by Breiman in 1996. This algorithm aims to train the model with different datasets by creating subsets from the input datasets prepared for training and thus increase the success rate. The size of the created subsets will be the same as the main data set. Therefore, some data cannot be found in a subset as a result of bootstrapping, while some data can be found in subset two or more times. Each subset created from the training data set is used to train a different learner. Then, the decision is made according to the absolute majority of the results of the weak learners trained with different datasets (Breiman, Bagging Predictors, 1996).

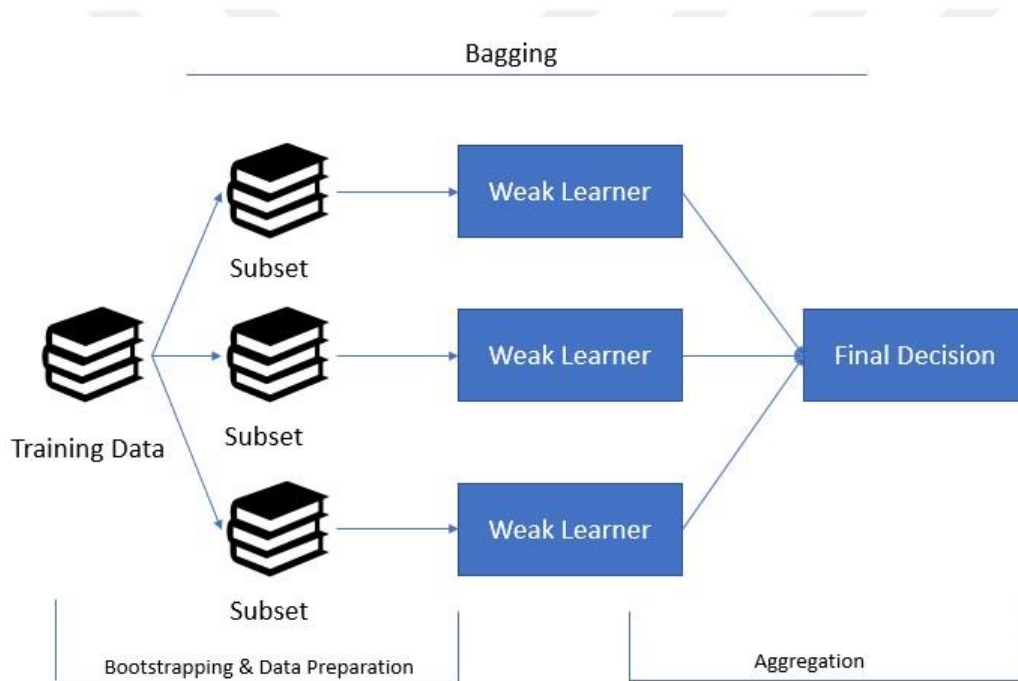


Figure 4.3: Bagging procedure.

4.6 Gradient Boosting

Boosting strategy is very similar to bagging. Just like in bagging, subsets with bootstrapping are used. For regression problems, firstly the results data prepared for training is averaged. Then, the results data prepared for training are extracted from this average and the errors are calculated. Based on these errors, a decision tree is set up. Then, using this decision tree, another result data is calculated. With this calculated result data, error values are recalculated again and another decision tree is set up again.

Better results can be obtained with sequential decision trees that learn from the mistakes of the former tree. In bagging, decision trees are parallel to each other, but in boosting they are cascaded and trained one after the other. In this way, a stronger model is created from poorly learning trees, just like in bagging (Freund & Schapire, 1996).

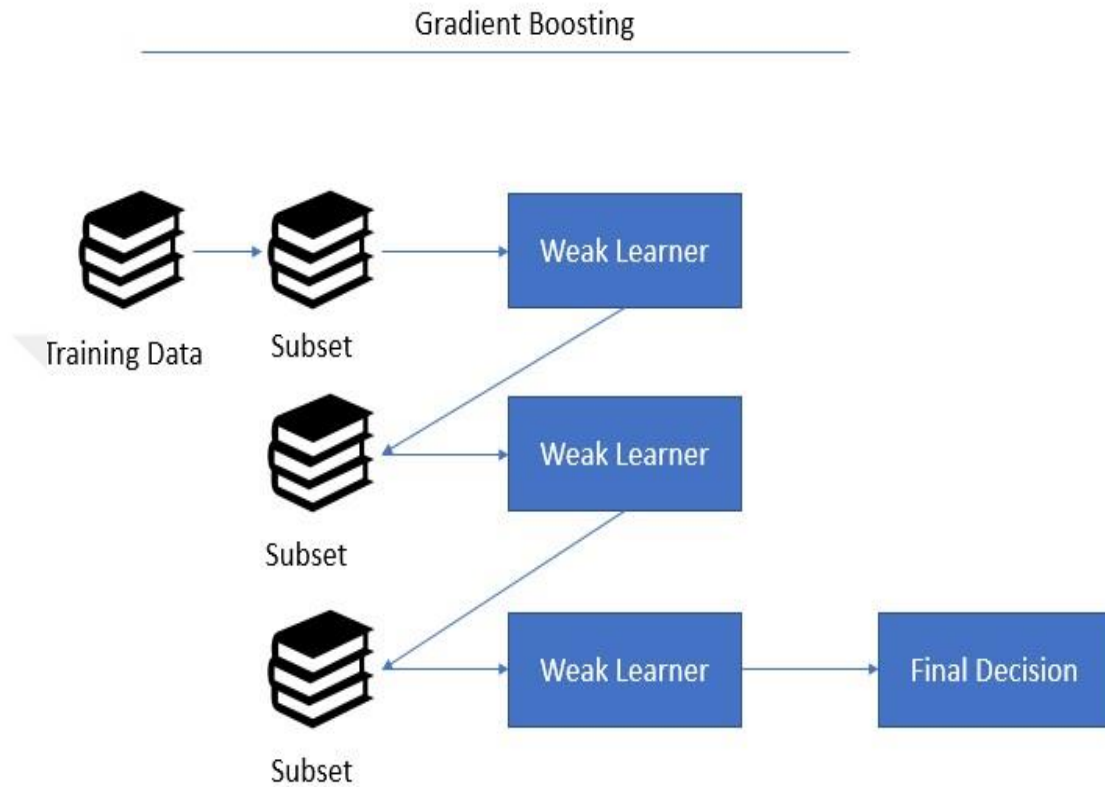


Figure 4.4: Gradient boosting procedure.

4.6 Random Forests

Random forest technique discovered by Breiman in 2001. It is a highly popular community learning algorithm. Because it's relatively fast. It can be used in regression and classification problems. Suitable for multidimensional problems. The logic of the random forest algorithm is similar to bagging. Information acquisition is usually given importance when designing models in decision trees. But in random forests, it makes a random selection instead of a selection of features that will provide the most information gain in a dataset. One of the biggest problems of decision trees is overfitting. The random forest model selects and trains random subsets from both the data set and the feature set to solve this problem. This allows the creation of low variance (Breiman, 2001).

4.7 XGBoost Algorithm

XGBoost stands for extreme gradient boosting. It prevents overfitting by combining gradient boosting, bagging (bootstrap aggregation) and feature randomness strategies. It can be used in both classification and regression problems. It has high flexibility. It performs cross validation. It calculates gain according to the calculated similarity scores and performs a backward prune process accordingly. Besides, it uses a regularization parameter and performs regularization. (Chen & Guestrin, 2016)

4.8 LightGBM Algorithm

LightGBM is a machine learning algorithm aimed at improving the performance of the XGboost algorithm. It was developed by Microsoft. It is a decision tree-based algorithm. It is used in regression and classification problems. LightGBM sometimes performs better than the XGboost algorithm. Just like XGBoost, it does bagging and boosting. LightGBM performs leaf-based growth rather than tree-based like other decision tree algorithms (Url-25).

5. DATA PREPARATION AND MODELLING

The Turkish spot gas market is a fairly new market and has only been active for a few years. In addition, the transactions made by BOTAS Transmission Company, one of the main drivers of the market, to balance the amount of stock, were included in the market after the market opened. Therefore, only data for 2019 and 2020 were used for this study. After the data was prepared, XGBoost and LightGBM algorithms were trained and the results were evaluated.

5.1 Data Preparation

5.1.1 Spot day ahead prices and volumes

The values to be forecasted in the study are the spot day ahead natural gas price. This price is the average of the prices in the first 24 hours of any delivery day contract is open. This price includes balancing transactions made by BOTAS Transmission Company every day at 12:30. The day ahead spot natural gas price is the sum of the special consumption tax and the natural gas price. For this reason, the special consumption tax has been reduced from the day ahead spot natural gas prices. The special consumption tax is 23 Turkish Liras per lot. Prices for 2019 and 2020 were used for this study. When the prices are examined, it is seen that the natural gas prices in the spot markets have changed between 1231-1655 TL in 2 years. On the other hand, 2-year average prices were calculated to be 1411.5 Turkish Liras without special consumption tax. Spot day ahead gas prices are taken from the EPIAS Transparency platform (Url-22).

5.1.2 BOTAS transactions

BOTAS transmission company operates in the spot gas market every day at 12:30 to balance the stock amount. It has been observed that depending on the change in the stock amount, it sometimes makes purchases and sometimes sales. BOTAS automatically matches the best offer available on the market at 12:30. Some days, the day-ahead transaction volume of the BOTAS transmission company reaches 50% of

the total transaction volume. This rate is quite high. For this reason, the transactions made by the BOTAS transmission company for each gas delivery day were added to the established model. In addition, a binary variable has been assigned whether the transactions made by the BOTAS transmission are either selling or buying and it is used in established models. Because the transaction direction of BOTAS transmission affects the market sometimes as excess supply and sometimes as demand. BOTAS transaction volumes are taken from the EPIAS Transparency platform (Url-22).

5.1.3 Brent oil Prices

In the literature review on gas pricing, it was seen that natural gas prices in Turkey are oil-indexed. It was understood from the literature research that natural gas prices in Turkey are formed by adding a formula indexed to oil over a base price. That's why Brent oil prices have been added to the training data. However, again, it is understood from the research that the oil prices are averaged while performing the oil index process. This average is reflected in the price with shifts of 3,6, or 9 month periods. A 9-month scroll was used in the study for this thesis. For each natural gas day, daily Brent oil prices were averaged starting 9 months ago and these obtained data were used for training.

5.1.4 BOTAS tariffs

Botaş tariffs are very important for the spot natural gas market. Because consumers usually decide to buy natural gas from the natural gas tariff or from the spot natural gas market. Sellers try to maximize their profits in this equation. In this study, it has been determined that the BOTAS tariff which applied to power plants has an effect on the spot prices. Also eligible tariff has effect on spot natural gas prices. Tariffs which have been used for this study, excluding transmission charges and capacity charges. In other words only natural gas sale prices of Botas were used in this study.

5.1.5 Dollar

Since most of the natural gas used in Turkey is imported from abroad, all contracts are based on dollars. That's why the day ahead spot dollar rates have been added to the model as training data. The most recent exchange rate is used for public holidays and weekends. Also, indicator prices were used on days when dollar rate were very volatile.

Table 5.1: Dataset Summary.

Input	Number of			
	Data	Min	Max	Average
CCGT Tariff (TL/Lot)	731	1336.1	1611	1490
Eligible Tariff(TL/Lot)	731	1313	1511	1403
Brent Oil Price (\$)	731	37	72	59
Dollar	731	5.16	8.52	6.34
BOTAS Transmission Traded Volume (Lot)	731	10	13714	1277
Spot Day Ahead Price (TL/Lot)	731	1231	1652	1411

5.2 Modelling

5.2.1 Evaluation metrics

Many different methods are used for the accuracy validation of models established in the literature. Accuracy validation is different for regression problems and classification problems. However, mean absolute error and root mean absolute error are relatively more popular for evaluating regression problems. MAE and RMSE were used to evaluate the models established in this study. MAE measures the deviation in error without positive or negative. MAE is the average of the difference between actual results and the results predicted by the model. RMSE is the square root of the squared mean deviation of errors. (Botchkarev, 2019). In addition, R^2 score was considered in evaluating model results

$$e_i = y - \hat{y} \quad (5.1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (5.2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i)^2} \quad (5.3)$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (5.4)$$

Formula 5.1 represents the error. $y - \hat{y}$ denote predicted day-ahead gas price and realized day-ahead gas price. MAE and RMSE are formulized in equations (5.2) and (5.3) respectively. Formula 5.4 represents the R^2 . SS_{res} represents the sum of squares of the residual errors also SS_{tot} represents the total sum of the errors.

5.2.2 Standardization

There is a serious scale difference among the data in the dataset. Therefore, a standardization was performed to prevent any feature from dominating the established model. Python's sklearn library was used for the standardization process. Apart from that, the sklearn library has been used to create train and test datasets. 80% of the data were used for testing and 20% for training.

5.2.3 Early stopping

The purpose of early stopping is to stop training the model before it becomes overfit. Overfit models perform well in the training dataset, while in the test set they perform very poorly compared to the training set. This is something that should be avoided in machine learning based models. However, after a point in models with overfit, there is no improvement in performance (Türkünoğlu, 2019). Therefore, in the case studies conducted for this thesis were carried out by considering the performance of training and test data to prevent overfitting.

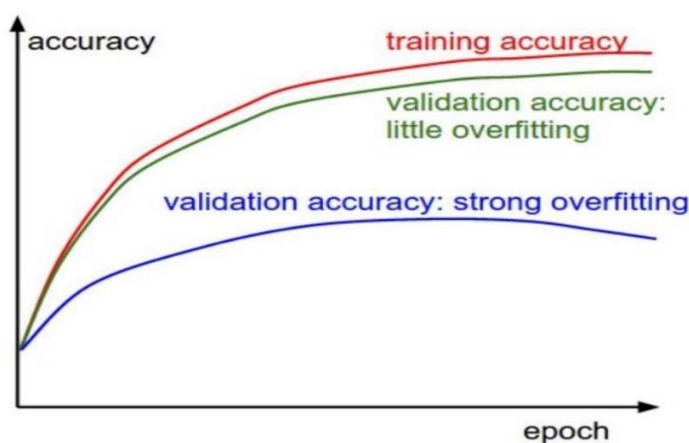


Figure 5.1: Training accuracy performance of overfit model (Url-26).

5.3 Results

5.3.1 XGBoost model results

In this study, XGBRegressor model of python's XGBoost library is used. Sklearn library was also used to create training and test data. With the Sklearn library, the dataset is randomly divided into training and test data also 80% of the dataset used as training data and %20 of the dataset is used as test data. In this study, hyperparameter tuning was performed for the XGBRegressor model. While tuning, the number of estimators, depth of decision trees and the effect of learning coefficient on the model

were examined. During the examinations, these hyperparameters were changed one by one in the first studies and the remaining hyperparameters were fixed to the default values of the library. As a result of the examinations, the independent effect of these three hyperparameters on the model was observed. It was determined after which values the model is overfit. After this, the optimum parameters were determined with the gridsearchCV model of the sklearn library. Then, according to the optimum parameters found, using the feature_importance model of the XGBoost library, input values most affected the model was investigated.

5.3.1.1 Effect of number of estimators

In the study, only the estimator number of the XGBRegressor model was changed. The default values of the library are used in all remaining hyperparameters. The number of estimators was started from 2 and increased by two to reach a maximum of 100. In the observations it was understood that the performance of the model increased with the increase of the number of estimators, and it was also understood that the model was overfitting after 42 estimators. No improvement in the performance of the model was observed after 42 estimators.

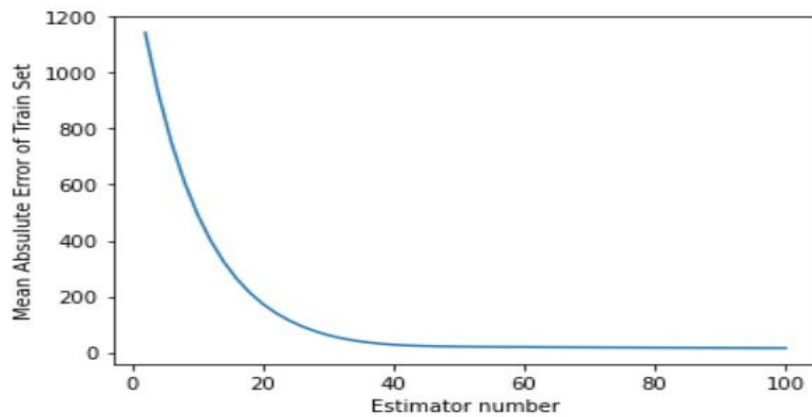


Figure 5.2: Effect of number of estimators to the model performance.

Table 5.2: MAE, RMSE and R² score values according to the estimator numbers.

Number of Estimator	MAE of Test Set(TL/Lot)	RMSE of Test Set (TL/Lot)	R ² score of Test Set (TL/Lot)
42	25.00	33.77	0.80
44	23.23	31.70	0.82
46	22.10	30.24	0.84
48	21.39	29.18	0.85
50	20.94	28.39	0.86

Table 5.3 (continued): MAE, RMSE and R² score values according to the estimator numbers.

Number of Estimator	MAE of Test Set(TL/Lot)	RMSE of Test Set (TL/Lot)	R2 score of Test Set (TL/Lot)
50	20.94	28.39	0.86
52	20.46	27.78	0.86
54	20.16	27.32	0.87
56	19.74	26.85	0.87
58	19.59	26.47	0.88
60	19.30	26.11	0.88
62	19.17	25.91	0.88
64	19.00	25.66	0.88
66	18.85	25.39	0.89
68	18.73	25.24	0.89
70	18.51	25.01	0.89
72	18.39	24.89	0.89
74	18.21	24.69	0.89
76	18.09	24.58	0.89
78	17.95	24.35	0.90
80	17.74	24.12	0.90
82	17.50	23.70	0.90
84	17.27	23.51	0.90
86	17.09	23.31	0.90
88	16.81	22.90	0.91
90	16.63	22.65	0.91
92	16.46	22.42	0.91
94	16.30	22.26	0.91
96	16.08	21.96	0.92
98	15.90	21.72	0.92
100	15.76	21.51	0.92

5.3.1.2 Effect of learning rate

In the second study, only the learning rate of the XGBRegressor model was iterated. The default values of the library are used in all remaining hyperparameters. Learning rate is very critical for model performance and training time. The smaller the learning rates, the longer the training time of the model. Model is iterated for 14 different values. In the observations, it was understood that the performance of the model increased with the increase of the learning rate, but it was understood also that the model was overfitting after 0.08 value. In addition, it has been seen in the study that increasing the learning rate does not affect the performance of the model much after a particular point. By increasing the learning rate, the MAE value was found to be approximately 19-17 liras. This shows the importance of learning rate. In the region where the model is overfit, MAE values are seen in the 14 lira band.

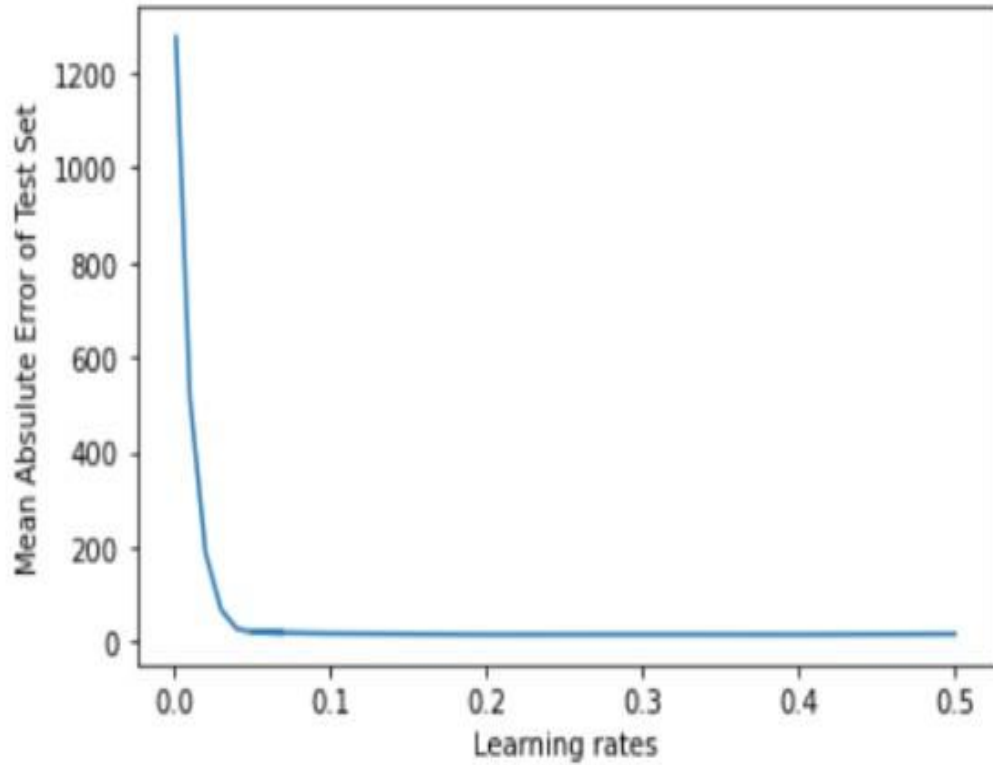


Figure 5. 3: Effect of learning rates to model performance.

Table 5.4: MAE and RMSE and R² score values according to the learning rate.

Learning Rate	MAE of Test Set (TL/Lot)	RMSE of Test Set (TL/Lot)	R ² score of Test Set
0.001	1275.4	1277.3	-292.0
0.01	516.2	518.3	-47.2
0.02	186.6	190.5	-5.5
0.03	66.7	74.0	0.0
0.04	28.5	37.1	0.8
0.05	23.0	29.9	0.8
0.06	21.3	27.2	0.9
0.07	20.5	26.4	0.9
0.08	19.0	24.7	0.9
0.09	17.5	23.1	0.9
0.1	17.2	23.1	0.9
0.2	14.3	20.9	0.9
0.3	14.3	20.7	0.9
0.4	14.5	21.0	0.9
0.5	14.2	21.3	0.9

5.3.1.3 Effect of max depth

In the third study, only the max_depth value of the XGBRegressor model was iterated. Other hyperparameters are kept at default values. In the observations made, it was

understood that the model was overfit when the max depth value exceeded 9. It should be noted here that pruning was not done for this study.

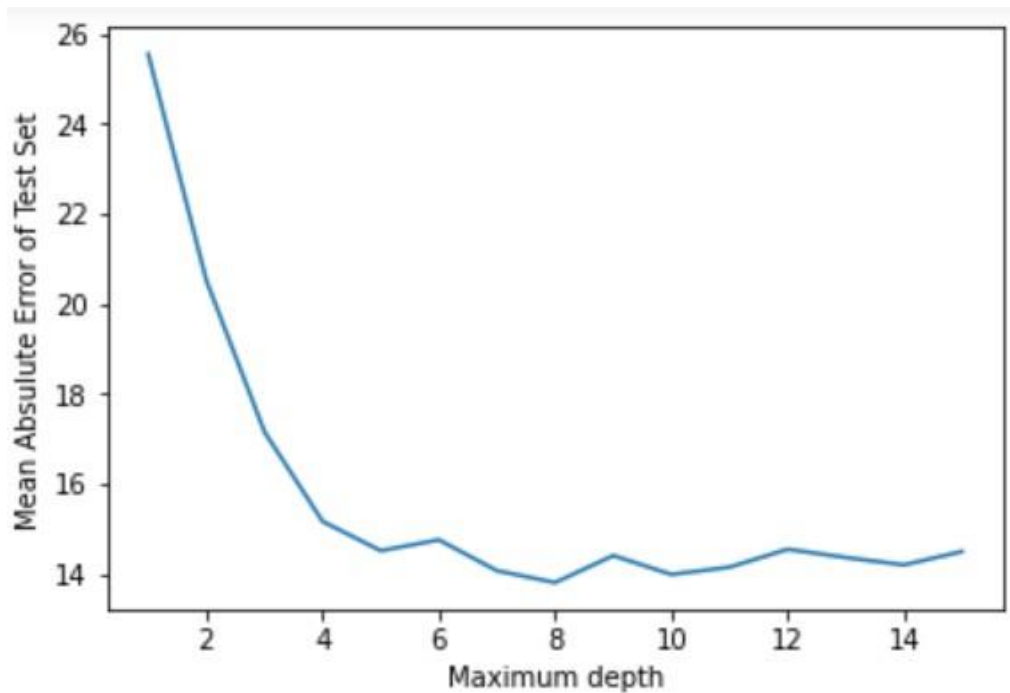


Figure 5. 4: Effect of maximum tree depth to model performance.

Table 5.5: MAE and RMSE and R2 score values according to the learning rate

Max Depth	MAE of Test Set (TL/Lot)	RMSE of Test Set (TL/Lot)	R2 score of Test Set
1	25.559	32.995	0.818
2	20.523	26.757	0.879
3	17.158	21.232	0.904
4	15.167	16.108	0.917
5	14.517	12.887	0.921
6	14.758	9.914	0.916
7	14.071	7.421	0.916
8	13.808	6.172	0.919
9	14.408	5.115	0.914
10	13.988	4.605	0.917
11	14.150	3.147	0.913
12	14.549	2.519	0.911
13	14.365	2.005	0.911
14	14.200	1.197	0.915
15	14.498	1.478	0.909

5.3.1.4 Grid search results

GridsearchCV is a function of a sklearn library that tests the model with all combinations of all values given for hyperparameter tuning. With this method, the

entire dataset is divided into specific parts. In this study, the data set is divided into 10 parts. Each part was used at least once as a test set and 9 times as a training data set. With this method, all dataset can be used as training and test dataset and best parameters can be decided accordingly. The working principle of GridsearchCV is shown in figure 5.6. Calculations have been made for all combinations of previously determined range of hyperparameter values. These values are 0.001, 0.002, 0.004, 0.008, 0.016, 0.032, 0.064, 0.128, 0.256, 0.512, for the learning rate, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42 for the number of estimators and for max_depth value 1, 2, 3, 4 values were trained. Then, the best parameter results are found with the best_params_ function of the sklearn library. These values are 0.512 for learning_rate, 4 for max_depth and 40 for n_estimator. In calculations using these values, the MAE value was found as 14.32 TL/Lot and the RMSE value as 20.99 TL/Lot also R2 score was 0.92. Also mean bias error for gridsearch method is -1.85.

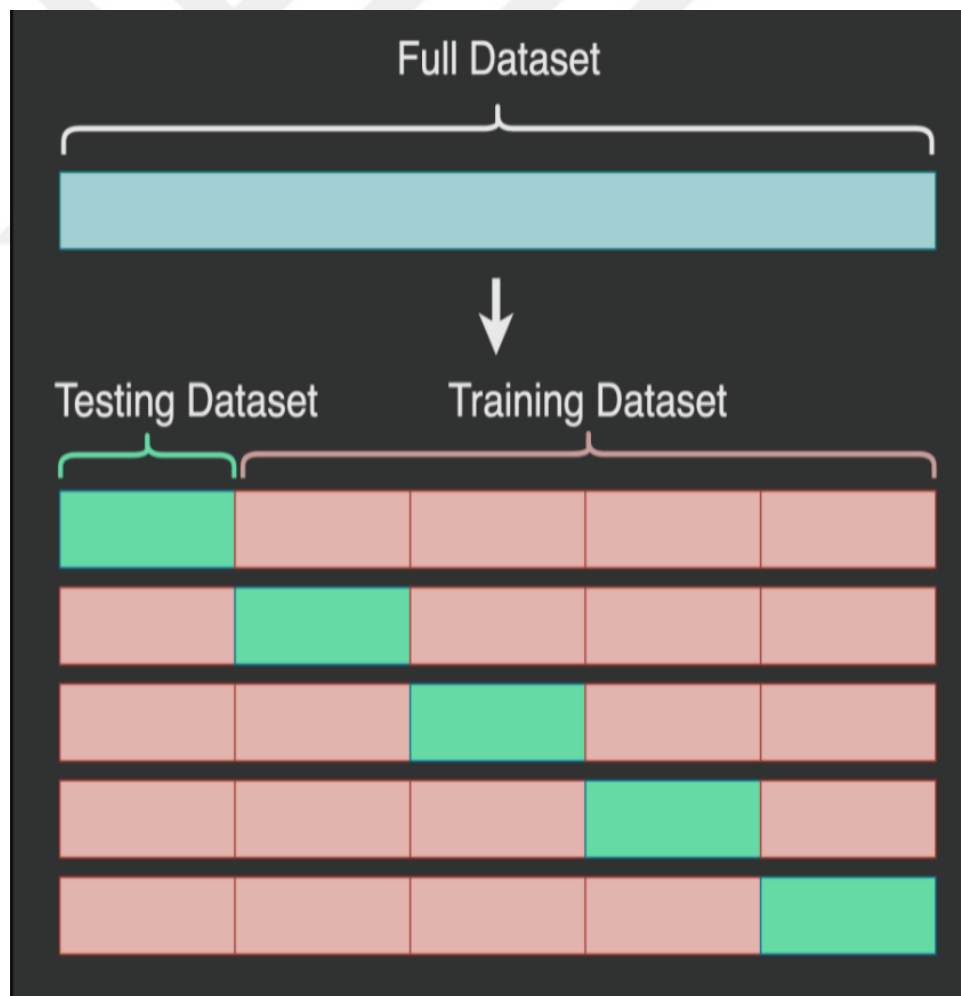


Figure 5. 5: Working principle of GridsearchCV function. Green blocks show test sets created from dataset, red blocks show training sets (Url-27).

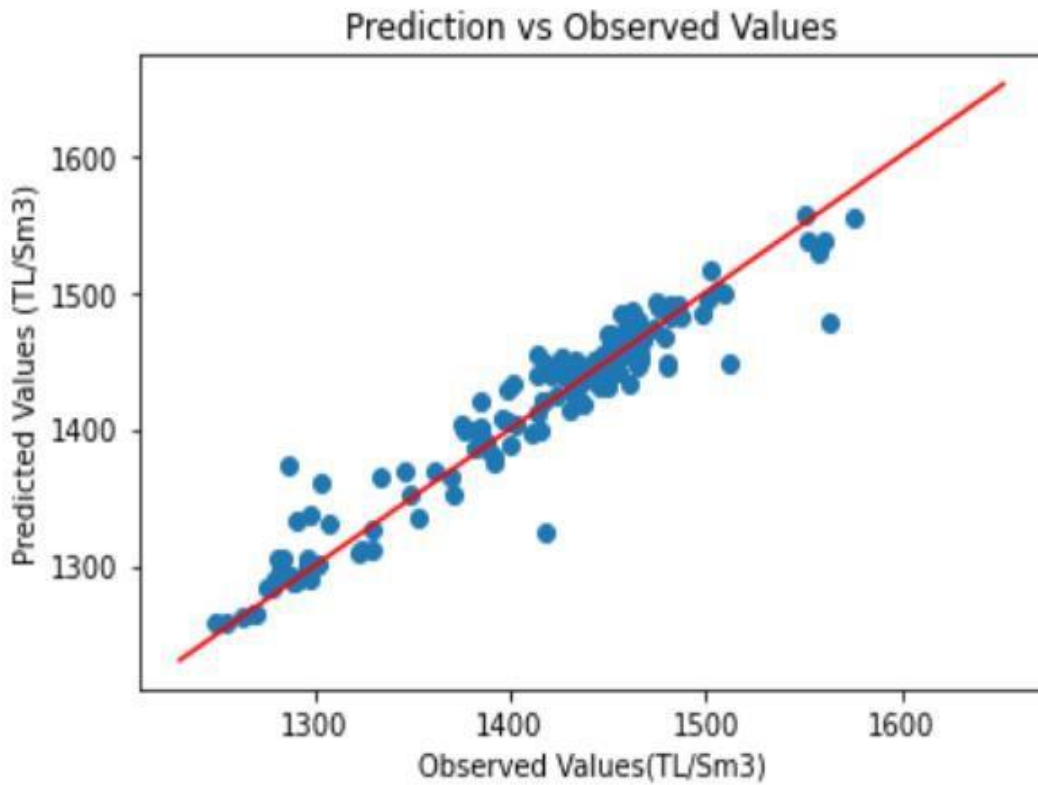


Figure 5. 6: Predicted and Observed Day Ahead Spot Prices, output of XGBoostRegressor with grid search's best parameters.

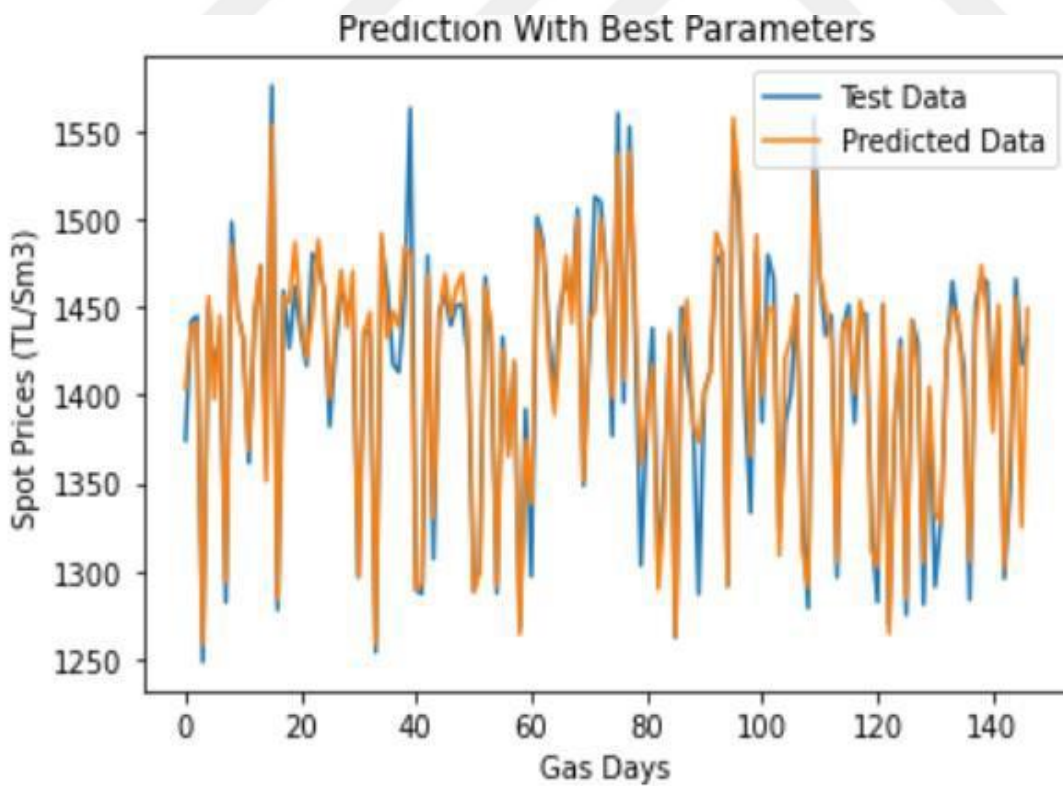


Figure 5. 7: Prediction and Observed day ahead spot gas prices for XGBRegressor model with learning rate 512, 6 for max depth with 56 estimator.

5.3.2 LIGHTGBM results

In this study, LGBMRegressor model of python's LIGHTGBM library is used. Sklearn library was also used to create training and test data. %80 of main dataset used as training, %20 of the main dataset is used as test data. In this study, hyperparameter tuning was performed for the LightBoost model. First the estimator number, then the learning rate and tree depth were optimized, respectively. Then, the best results were found and compared using the GridsearchCV function of the Sklearn library. Then, according to the optimum parameters found, using the feature_importance function of the LightBoost library, it was observed which parameters most affected the model.

5.3.2.1 Effect of number of estimators on LightBoost

In the study, only the estimator number of the LGBMRegressor model was iterated. The default values of the library are used in all remaining hyperparameters. In the study, the number of estimators was started from 2, gradually increased by 2, and estimation was made with a maximum of 100 estimators as the previous study. In the study, it was observed that the performance of the model increased with the increase in the number of estimators, but there was no improvement in the performance of the model after 40 estimators. After 24 estimators, the model was overfit. The biggest difference from the XGBoost study is that, LightBoost gives better results with a lower number of estimators.

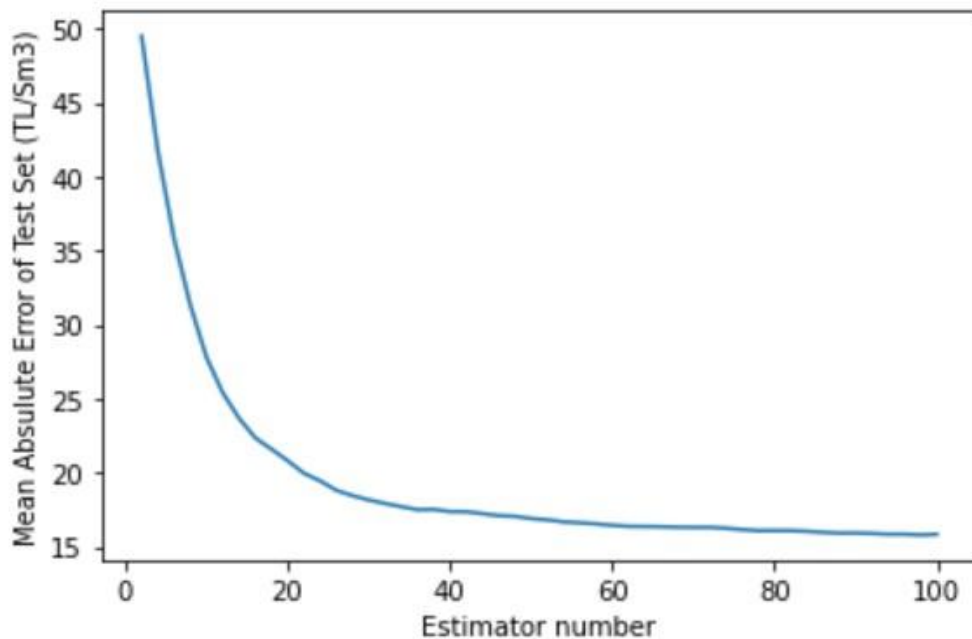


Figure 5. 8: Effect on number of estimators on lightboost.

Table 5.6:MAE and RMSE and R2 score values of LIGHTGBM according to the number of estimators.

Number Of Estimators	MAE of Test Set (TL/Lot)	RMSE of Test Set (TL/Lot)	R2 Score of Test Set
2	49.48	62.42	0.30
4	41.72	52.96	0.50
6	35.80	45.71	0.62
8	31.33	40.45	0.71
10	27.79	36.27	0.76
12	25.39	33.22	0.80
14	23.69	31.01	0.83
16	22.36	29.37	0.85
18	21.59	28.29	0.86
20	20.81	27.33	0.87
22	19.97	26.43	0.87
24	19.45	25.94	0.88
26	18.81	25.28	0.89
28	18.45	24.96	0.89
30	18.16	24.75	0.89
32	17.92	24.58	0.89
34	17.71	24.49	0.89
36	17.51	24.36	0.89
38	17.52	24.32	0.89
40	17.38	24.15	0.90
42	17.37	24.13	0.90
44	17.24	23.92	0.90
46	17.11	23.71	0.90
48	17.06	23.68	0.90
50	16.91	23.55	0.90

5.3.2.2 Effect of learning rate on LightBoost

In the second study with LightBoost, only the learning rate of the LGBMRegressor model was iterated. The default values of the library are used in all remaining hyperparameters. The learning rate value is very important here, just as in XGBOOST. Because large learning rate values can prevent optimization. In addition, small learning rate values can prolong the training time. Also, certain value ranges can operate the model in the overfitting region. It is iterated for 14 different values. In the observations, it was understood that the performance of the model increased with the increase of the learning rate, but it was also understood that the model was overfitting after 0.06 value. By optimizing the learning rate value, MAE values were brought to the band of approximately 16 lira in this study.

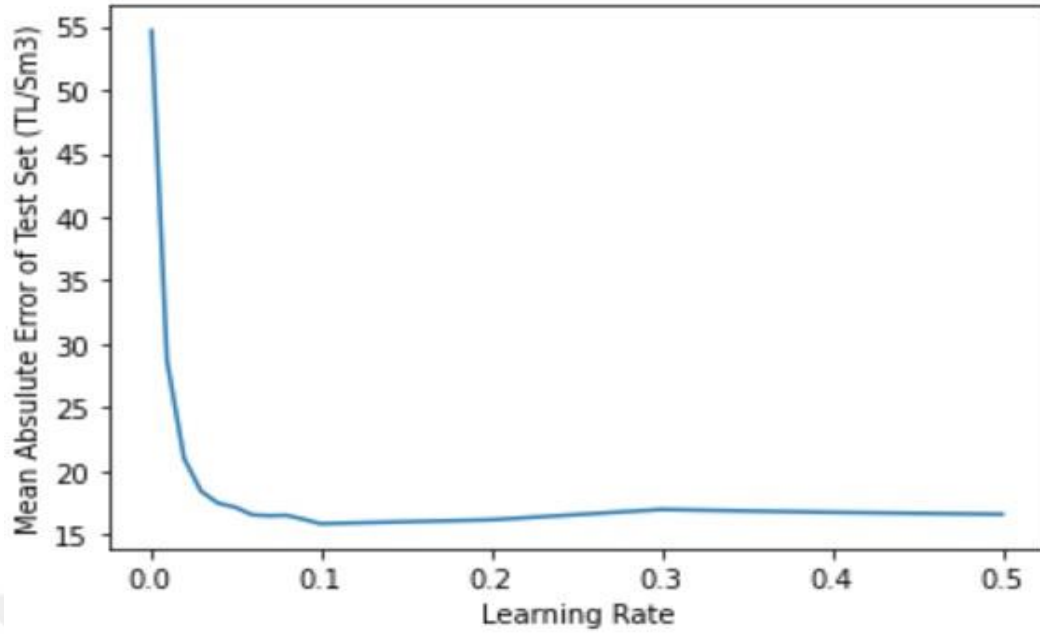


Figure 5. 8: Effect of learning rate to LIGHTGBM model performance.

Table 5.7: MAE and RMSE and R2 score values of LIGHTGBM according to the number of estimator.

Learning Rate	MAE of Test Set (TL/Lot)	RMSE of Test Set (TL/Lot)	R2 Score of Test Set
0.001	54.68	68.47	0.16
0.01	28.69	37.22	0.75
0.02	20.99	27.51	0.86
0.03	18.40	24.94	0.89
0.04	17.46	24.16	0.90
0.05	17.15	23.76	0.90
0.06	16.54	23.08	0.90
0.07	16.48	23.14	0.90
0.08	16.51	23.16	0.90
0.09	16.20	23.06	0.90
0.1	15.85	22.83	0.91
0.2	16.15	23.56	0.90
0.3	16.96	24.66	0.89
0.4	16.74	24.10	0.90
0.5	16.59	23.83	0.90

5.3.2.3 Effect of max_depth parameter on LightBoost Model

In the third study with LightBoost, only the max_depth parameter of the LGBMRegressor model was iterated. The default values of the library are used in all remaining hyperparameters. It is iterated for 15 different values. Iterations started with 1 and end with 15. In the observations, it was understood that the performance of the

model increased with the increase of the max depth value, but it was also understood that the model was overfit when max depth value exceeds 4.

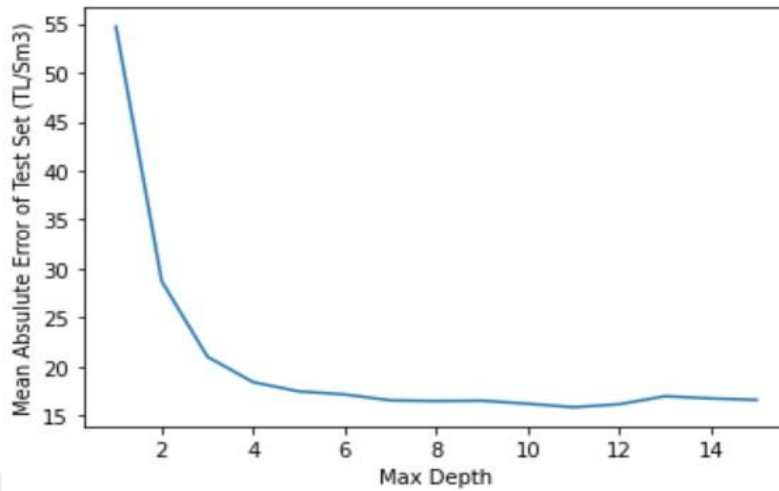


Figure 5. 9: Effect of max depth to LIGHTGBM model performance.

Table 5.8: MAE and RMSE and R2 score values of LIGHTGBM according to the max depth.

Max Depth	MAE of Test Set (TL/Lot)	RMSE of Test Set (TL/Lot)	R2 Score of Test Set
1	54.68	68.47	0.16
2	28.69	37.22	0.75
3	20.99	27.51	0.86
4	18.40	24.94	0.89
5	17.46	24.16	0.90
6	17.15	23.76	0.90
7	16.54	23.08	0.90
8	16.48	23.14	0.90
9	16.51	23.16	0.90
10	16.20	23.06	0.90
11	15.85	22.83	0.91
12	16.15	23.56	0.90
13	16.96	24.66	0.89
14	16.74	24.10	0.90
15	16.59	23.83	0.90

5.3.2.4 GridsearchCV study for LightBoost Model

Considering previous studies with LightBoost hyperparameters are decided. For GridsearchCV function of sklearn library also used for Lightboost model. For learning rates 0.001, 0.01, 0.02, 0.03, 0.04, 0.05 values used. For n_estimator parameter 10, 12, 14, 16, 18, 20, 22, 24 values were used. Finally, for max_depth parameter 2, 3, 4 values

were used. As a result of the iterations, the best parameters found are as follows: learning rate: 0.05, max depth: 4, estimator number 24. In the study conducted with these values, the MAE value was found as 21.54 and the RMSE value as 21.49. Also mean bias error for this study is -1.6. As a result of the study with the best parameters found after the grid search was performed, the parameters that the spot natural gas price depends on were investigated with the feature importance function of LightBoost. As a result of the LIGHTGBM iterations, it has been revealed that the spot gas price is mostly dependent on the dollar, Brent oil price, CCGT tariff, daily trade of the BOTAS transmission company and eligible tariff respectively.

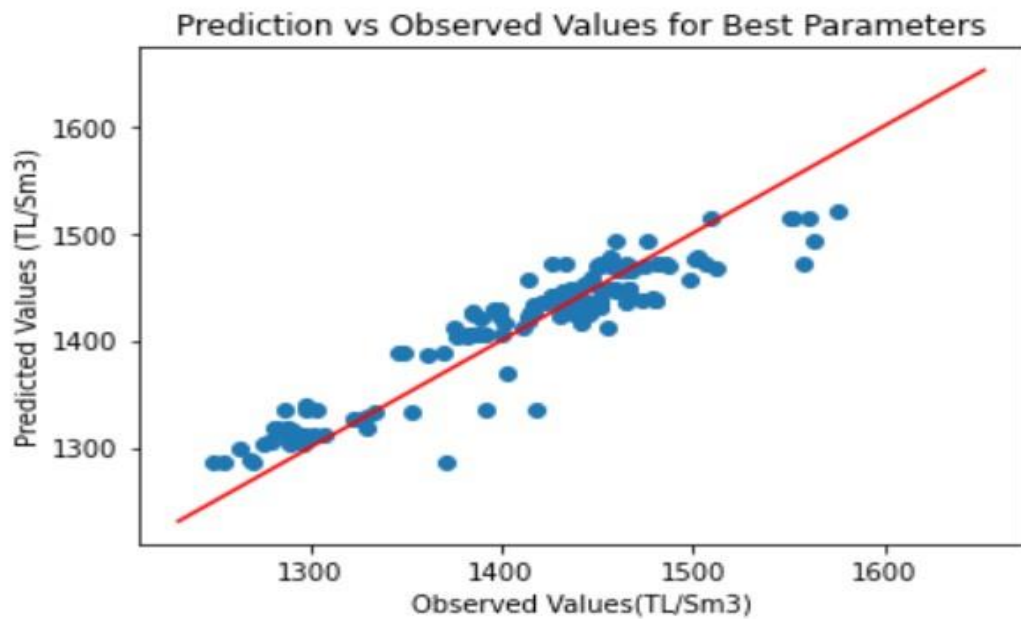


Figure 5.10: Prediction and Observed day ahead spot gas prices for LIGHTGBM model with 0.05 learning rate, 4 for max depth with 24 estimator.

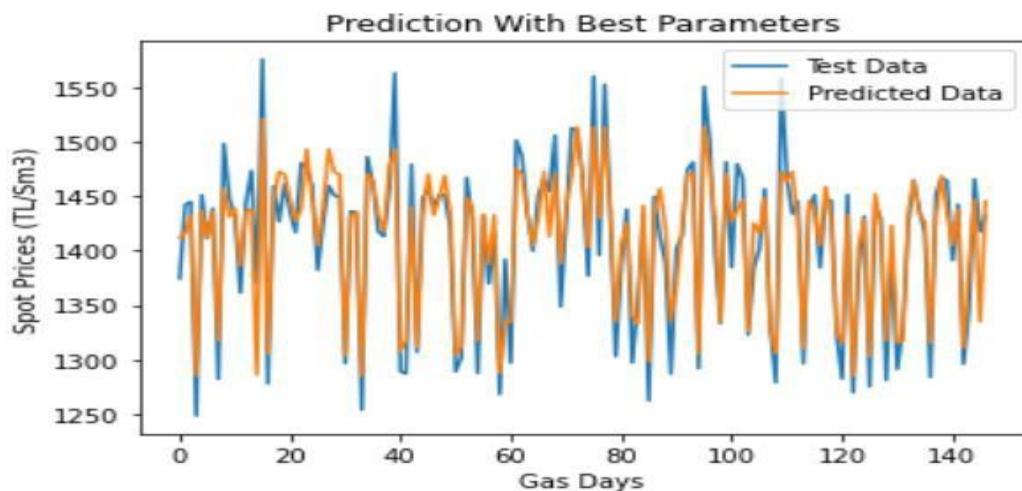


Figure 5.11: Prediction and Observed day ahead spot gas prices for LIGHTGM model with 0.05 learning rate, 4 for max depth with 24 estimator.

5.3.3 Comparison

In the first studies, tree depth, number of estimators and learning rate were iterated one by one, and the models were run again. Only one parameter was changed at a time and the other parameters were left at the default settings of the libraries. These parameters, which were optimized separately, were then used simultaneously with grid search. The most logical and consistent results were found with grid search. In the study, the XGBoost library gave more accurate results. In the grid search study conducted with XGBoost, the optimum values were found to be 0.512 for the learning rate, 4 for the max depth, and 40 for the number of estimators. According to these optimum results, MAE was found as 14.32 TL/Lot. For the optimum values, for LIGHTGBM, the learning rate was found to be 0.05, the maximum depth was 4, and the number of estimators was 24. In the study conducted with these values, the MAE value was found to be 21.4 TL/Lot. As a result, XGBoost found a more accurate result of about 8 TL.

6. CONCLUSION AND FUTURE RECOMMENDATIONS

In this study, first, the Turkish natural gas market has been examined. As a result of the examinations, the parameters on which the natural gas price depends were investigated. According to the literature research, it has been seen that natural gas prices in Turkey are oil-based substitute prices. Apart from this, it is understood that the impact of BOTAS on the natural gas market is quite important. From time to time, it is quite interesting that the prices in the spot natural gas market are lower than the BOTAS price. Because of the take or pay obligations of natural gas contracts, dollar or oil hedges may be the reason for this. The fact that the prices that can compete with BOTAS are at spot prices has undoubtedly attracted the attention of the private sector. Although this interest has increased the amount of trade in the spot natural gas market, total daily transaction volume on spot gas market is quite low when compared to global natural gas markets. However, the Turkish natural gas market is promising. Because new natural gas discoveries, Turkey's ability to purchase natural gas from many different countries and many transit lines through the country give Turkey a great advantage. Apart from that, 2021 is an important year for Turkey in terms of natural gas because some long-term natural gas contracts expire. It is still unclear how the natural gas needs will be met. For future studies, it should not be forgotten that the pricing strategy of the expiring natural gas contracts will undoubtedly affect the spot natural gas market. In addition, it has been seen in the literature research that Turkey's natural gas infrastructure is getting stronger. This strengthening infrastructure has enabled Turkey's natural gas imports to better adapt to changing market conditions. For example, thanks to LNG investments in recent years, cheap natural gas in the spot LNG markets has been used and gas imports from Russia have decreased.

One of the important issues for the Turkish gas market is liberalization. Liberalization of the natural gas market progresses more slowly than the electricity market. BOTAS's market dominance still continues. Apart from that, more data about spot gas market is needed for future studies. The market operator still does not publish the daily amount

of natural gas entering from natural gas import points. If these data are published in the following days and the market becomes more liberal, import prices from different physical entrance points on a daily basis may also affect the spot gas market. However, currently, the daily import amounts are not officially published according to different contracts. Apart from this, in future studies, weekly contracts in the spot market or transactions in the natural gas futures market can be used in price prediction. However, these options in this natural gas market are not used in this study because they are quite new and there is not enough data. In order to obtain better results in future studies, while preparing the data set, the trades made by BOTAS Transmission Company can be deducted from the day-ahead trades and a new day-ahead price can be calculated again without it. In this way, the average of the commercial transactions made by other market participants is found. This may give a better result.

After the literature research and the creation of the data set, the machine learning algorithms to be used for spot gas price prediction started to be researched. It has been seen in the research that decision tree-based algorithms are very popular in price predictions and produce very good results. Therefore, it was decided to use XGBoost and LIGHTGBM algorithms. In addition, python was chosen as the programming language. There is no price prediction made about the Turkish Spot Natural Gas Market in the literature. However, there are a few studies to forecast electricity prices. In the literature study, it has been seen that decision tree-based algorithms give much better results in electricity price prediction.

In this study, the learning rates, tree depths and estimator numbers of the XGBoost and LightBoost libraries were initially iterated one by one. As a result of these iterations, it was seen after which values these parameters overfit the model. Then, with these studies, ranges for hyperparameters were determined and grid search was conducted. Thanks to the grid search, the best parameters were selected. As a result of the study, it was understood that the XGBoost algorithm gave a better result than LightBoost in predicting the spot gas price.

REFERENCES

- ACCC. (2019).** *Framework for the Consistent Reporting of Natural Gas Reserves and Resources.*
- Alpaydın, E. (2010).** *Introduction to Machine Learning.* the MIT Press.
- Botchkarev, A. (2019).** New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms. *Interdisciplinary Journal of Information, Knowledge & Management.*
- Breiman, L. (1996).** Bagging Predictors. *Machine Learning*, 123-140.
- Breiman, L. (2001).** Random forests. *Machine Learning*, 5-32.
- Chen, T., & Guestrin, C. (2016).** XGBoost: A Scalable Tree Boosting System. *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* , (s. 785–794). New York.
- Demiraslan, M. (2020).** *Sağlık Verilerinde Aykırı Gözlemlere Duyarlı ve Hastalık Tanısına Yönelik Yeni Bir Öznitelik Seçim Yöntemi Önerisi* . İzmir .
- Engin, N. (2013).** Enerji Kaynağı Olarak Doğalgaz ve Türkiye. *Marmara Coğrafya Dergisi*, 233-244.
- EPDK. (2019).** *Doğalgaz Piyasası 2019 Yılı Sektör Raporu.* Ankara: T.C. Enerji Piyasası Düzenleme Kurumu Strateji Geliştirme Dairesi Başkanlığı.
- Freund, Y., & Schapire, R. E. (1996).** Experiments with a new boosting algorithm. *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning* , (s. 148-156).
- IGU. (2019).** *Wholesale Gas Price Survey 2019 Edition.*
- Kilic, A. M. (2005).** Major Utilization of Natural Gas in Turkey. *Energy Exploration&Exploitation*, 157-166.
- Kılıç, A. M. (2005).** Major Utilization of Natural Gas in Turkey. *Energy Exploration & Exploitation* , 125-140.
- Learned-Miller, E. G. (2014).** *Introduction to Supervised Learning.*
- Öztürk, S. (2007).** *Doğal gazın depolanması ve Türkiye için önemi (Master's thesis, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü).*
- Rahman, F. (2019).** *Evaluate Probe Speed Data Quality to Improve Transportation Modelling.*
- Saka, B. (2003).** *Doğalgaz Ve Lpg'nin Kullanım Alanları, Dünyadaki Ve Türkiye'deki Rezerv, Üretim Ve Tüketim Miktarı Ve Taşıtlarda Kullanımı.*
- Tianqi Chen, C. G. (2016).** XGBoost: A Scalable Tree Boosting System. *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* , 785–794.
- TMMOB. (2006).** *Türkiye'nin Doğal Gaz Temin ve Tüketim Politikalarının Değerlendirilmesi Raporu.* Makina Mühendisleri Odası.
- The World Bank. (2015).** *Turkey's Energy Transition Milestones and Challenges.*
- Türkunoğlu, A. (2019).** *Short Term Electricity Consumption Forecasting Using Long Short-Term Memory Cells.*

WEC. (2018). *Türkiye Enerji Piyasaları Araştırma Raporu Kısım 1 Mevcut Durum Analizi.*

WIDDERSHOVEN, C. (2017). Turkey's LNG Future & External Threats to Derail Ankara's Strategy. *Energy Policy Turkey*, 62-67.

Yağız, A. (2016). *Türkiye 'de doğal gaz piyasası ve doğal gaz fiyatları üzerinde petrol fiyatları etkisinin incelenmesine yönelik bir araştırma.*

Yardımcı, O. (2011). Türkiye Doğalgaz Piyasası: Geçmiş 25 Yıl, Gelecek 25 Yıl. *Ekonomi Bilimleri Dergisi*, 157-166.

Url-1<<https://www.bp.com>>, date retrieved 03.04.2021.

Url-2<<https://www.indexmundi.com>>, date retrieved 05.03.2021.

Url-3<<https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/natural-gas.html#natural-gas-reserves>>, date retrieved 05.04.201.

Url-4<<https://www.indexmundi.com/map/?t=0&v=137&r=xx&l=en>>, date retrieved 05.04.2021.

Url-5<<https://www.iea.org/reports/natural-gas-information-overview>>, date retrieved 05.04.2021.

Url-6<<https://www.botas.gov.tr>>, date retrieved 02.02.2021.

Url-7<<https://www.etd.org.tr>>, date retrieved 07.04.2021.

Url-8<<https://www.epias.com.tr/tum-duyurular/kurumsal/dogal-gaz-piyasa-sistemi-test-asamasi-1-nisanda-aciliyor-2/>>, date retrieved 07.04.2021.

Url-9<<https://www.epias.com.tr/tum-duyurular/kurumsal/dogal-gaz-piyasasi-haftalik-urunler-isleme-aciliyor-3/>>, date retrieved 07.04.2021.

Url-10<<https://www.epdk.gov.tr>>, date retrieved 07.04.2021.

Url-11<<https://www.bbc.com/turkce/haberler-turkiye-53868045>>, date retrieved 11.04.2021.

Url-12<<https://www.petform.org.tr/dogal-gaz-piyasasi/turkiye-dogal-gaz-piyasasi/>>, date retrieved 11.04.2021.

Url-13<http://rapor.isyatirim.com.tr/2_20200824135204544_1.pdf>, date retrieved 11.04.2021.

Url-14<<https://enerji.gov.tr/bilgi-merkezi-dogal-gaz-boru-hatlari-ve-projeleri>>, date retrieved 11.04.2021.

Url-15<<https://silivriebt.botas.gov.tr/index.php/tr/tarifeler>>, date retrieved 11.04.2021.

Url-16< <https://silivriebt.botas.gov.tr/index.php/tr/>>, date retrieved 11.04.2021.

Url-17<<https://www.botas.gov.tr/Sayfa/tuz-golu-yer-alti-dogal-gaz-depolama/23>>, date retrieved 11.04.2021.

Url-18<<https://www.botas.gov.tr/Sayfa/marmara-ereglisi-lng-terminali/20>>, date retrieved 11.04.2021.

Url-19<<https://www.botas.gov.tr/Sayfa/fsru-dortyol/24>>, date retrieved 11.04.2021.

Url-20<<https://www.mgm.gov.tr/veridegerlendirme/il-ve-ilceler-istatistik.aspx?m=ISTANBUL>>, date retrieved 12.05.2021.

Url-21<<https://data.tuik.gov.tr/Kategori/GetKategori?p=Sanayi-114>>, date retrieved 12.05.2021.

Url-22<<https://seffaflik.epias.com.tr/transparency/>>, date retrieved 12.05.2021.

Url-23<<https://erranet.org>>, date retrieved 12.05.2021.

Url-24<<https://www.botas.gov.tr/uploads/galeri/229130-2018bilanco.pdf>>, date retrieved 12.05.2021.

Url-25<<https://deep-and-shallow.com/2020/02/21/the-gradient-boosters-iii-lightgbm/>>, date retrieved 12.05.2021.

Url-26<<https://deepdatascience.wordpress.com/2016/11/17/how-to-detect-model-overfitting-by-training-accuracy/>>, date retrieved 25.05.2021.

Url-27<<https://mlfromscratch.com/gridsearch-keras-sklearn>>, date retrieved 25.05.2021.



CURRICULUM VITAE



Name Surname: Dođukan Őahin



EDUCATION:

B.Sc.: 2018, Yıldız Technical University, Faculty of Electrical and Electronics Engineering, Electrical Engineering

PROFESSIONAL EXPERIENCE AND REWARDS:

Sade Yazılım DanıŐmanlık, Data Analys since February 2022

ENGIE Global Energy Management, Trading Operation Specialist, Istanbul, June 2019-February 2022

