UNIVERSITY OF KWAZULU- NATAL

FORECASTING THE NATURAL GAS CONSUMPTION FOR THE PERIOD OF 2018 TO 2028, THE SOUTH OF KZN PROVINCE

By

Mabuti Vilakazi

217077489

A dissertation submitted in partial fulfillment of the requirements for the degree of
Masters of Business Administration

Graduate School of Business and Leadership
College of Law and Management Studies

Supervisor: Doctor Bibi Zaheenah Chummun

2018

DECLARATION

I Mabuti Vilakazi declare that:

- The research reported in this thesis, except where otherwise indicated, is my original work.
- This thesis has not been submitted for any degree or examination at any other university.
- This thesis does not contain other persons' data, pictures, graphs or other information unless specifically acknowledged as being sourced from other persons.
- This thesis does not contain other persons' writing unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written but the general information attributed to them has been referenced;

b) where their exact words have been used, their writing has been placed inside quotation marks and referenced.

c) Where I have reproduced a publication of which I am an author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.

d) This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed: [redacted]

LIST OF ACRONYMS AND ABBREVIATIONS

ADF	-	Augmented Dicky Fuller
AR	-	Autoregressive
BAU	-	Business as Usual
bcm	-	billion cubic metre
BP	-	British Petroleum
ARIMA	-	Auto-Regressive Moving average
CUSUM	-	Cumulative Sum
CUSUMQ	-	Cumulative Sum Squared
CNG	-	Compressed Natural Gas
CO_2	-	Carbon Di-Oxide
DoE	-	Department of Energy
DMR	-	Department of Mineral Resource
GHG	-	Green House Gas
GUMP	-	Gas Users Master Plan
IEA	-	International Energy Agency
IEP	-	Integrated Energy Plan
IEO	-	International Energy Outlook
IRP	-	Integrated Resource Plan
LNCON	-	Natural log of gas consumption
LNPMI	-	Natural log of Purchasing Management Index
LNG	-	Liquefied Natural Gas
LPG	-	Liquefied Petroleum Gas
MA	-	Moving Average
MRG	-	Methane Rich Gas
MSP	-	Mozambique to Secunda Pipe
NERSA	-	National Energy Regulator of South Africa
NG	-	Natural Gas
PJ	-	Peta Joules (Quadrillion Joules)
PMI	-	Purchasing Management Index
ROMPCO	-	Republic of Mozambique Pipe Company
TCM	-	trillion cubic meter
VECM	-	Vector Error Correction Model

ABSTRACT

An upward trajectory in the consumption of natural gas around the globe as a source of energy to the industrial, residential, and commercial sector is evident. Gas consumption increase or reduction is closely correlated to the energy demand. Growth in the manufacturing industry can have a similar effect on energy and gas consumption demand. Gas industry in South Africa is relatively new and growing. The current gas consumption growth has led to the formation of the gas market structure which, comprises of exploration, transmission, distribution, and reticulation. The continued gas consumption growth requires additional investment since a rise in the capacity of the gas infrastructure would be expected. End-user gas consumption influences the increase in gas consumption which yields the imbalances in the forecast of natural gas and if actual gas consumption is lower than 80% of the initial forecasted gas, contractual penalties apply. The study is motivated by the value of the penalty that comes with the 80% take or pay contract condition. Spring Lights Gas and other traders are contractually bound by the same contract condition and its critical for the company to submit correct gas quantity to the gas producers to avoid penalties. The study focuses on the South African gas industry, particularly the southern region of KwaZulu Natal. The gas consumption time series is inherently non-stationary due to the different production plans and unplanned equipment repair which can cause a swing in the gas consumption. The study aims to forecast future gas consumption for Spring Lights Gas with minimum error. The actual monthly consumption time series data of the KZN southern region for the period of June 2005 to July 2018 was collected from Spring Lights Gas and used to model and estimate the univariate Autoregressive Integrated Moving Average (ARIMA) method which was compared to the bivariate Vector Error Correction Model (VECM) method. The performance of both models was determined by forecasting the in-sample data from 2016 to 2018. The univariate ARIMA method was chosen because it generated low Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) compared to the bivariate VECM method. Consequently, the ARIMA model was used to forecast the future gas consumption for Spring Lights Gas and produced RMSE of seven percent compared to the official future gas plans of the company. The results further revealed that gas consumption in the south region is projected to grow by one percent for the next decade and it was recommended that the company revise down its future gas plans.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

Introduction

1.1 Introduction

An upward trajectory in the consumption of natural gas in Europe (Dilaver, Dilaver & Hunt, 2015) and Turkey (Akpinar and Yumusak, 2016b), as a source of energy in the industrial, residential, and commercial sector is evident. Non or conventional technology can harvest natural gas from the well (Paylor, 2017). There is a considerable number of countries that have proven conventional resource of natural gas namely: Mozambique, Tanzania, and Angola among others (DBSA, 2017). Trade between nations that have gas as an abundant natural resource can be made using transmission pipelines a system which is considered safe and cheap to transport gas across borders (Sasol, 2014).

For trades where gas lines are not available, modern technology has facilitated the virtual forms of transporting gas, namely: Compressed Natural Gas (CNG) and Liquified Natural Gas (LNG) (SLG, 2018). However, substantial initial investments are required to construct the transportation pipe, CNG compression and LNG facilities. Factors that influence the investment in the transportation infrastructure among many include an adequate gas consumption to justify expenditure or the importance and commitment of nations to reduce the Green House Gases (Esso and Keho, 2016), which is undertaken when long-term contracts are concluded.

Take or Pay agreements (Akpinar and Yumusak, 2016b) are common when massive investments are made in developing gas infrastructure to ensure that projected revenue in line with the forecasted gas usage is guaranteed. These contracts are usually concluded on the bases of the forecasted gas usage. Take or Pay agreements are structured such that when the forecasted quantity of gas is not consumed, the consumer will be required to pay the difference between the actual gas usage and the initially forecasted gas usage at the agreed gas price (Akpinar and Yumusak, 2016b). However, should the consumer use more gas than expected, the benefit could be a reduced gas price or restriction of gas flow. To avoid punitive measures in these agreements, gas demand forecast modeling with minimum error is critical.

The primary objective of the study is to forecast future gas consumption for Spring Lights Gas (SLG). SLG, situated in Kwa Zulu Natal, is the second largest trader of natural gas and compressed natural gas in South Africa with Sasol being the largest. The company enjoys a considerable market share in KZN followed by Sasol. The univariate Autoregressive Intergtergrated Moving Average (ARIMA) and bivariate Vector Error Correction Model (VECM) were used to estimate and modeled to determine which method best fit the actual gas consumption.

SLG granted access to actual gas consumption from June 2005 to July 2018 which was used to estimate and model both methods. Purchasing manager index (PMI) was used as the second variable to model the VECM method. In comparing the two models, ARIMA was chosen because it produced low Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) and used to forecast the future gas consumption for SLG.

1.2 Study outline

The study comprises of Chapter one, the overview of the study. It further provides details on the focus of the study, research objective, questions, methodology, sample data source, and limitation. Chapter two gives a detailed background on the study, gas industry in South Africa, sources of gas and literature review on univariate Autoregressive Integrated Moving Average (ARIMA) and the bivariate Vector Error Correction Model (VECM). Chapter three outline model estimation, formulae, sampling, data collection and the validity of the data. Chapter 4 discusses and presents the models result and lastly, Chapter five which concludes and recommends future research topic.

1.3 Background to the study

Natural gas, methane-rich (92% content) gas is a gaseous fuel extracted underground. The formation of this gas resulted from a prolonged reaction of organism deposited million years ago (Paylor, 2017). Natural gas is a flammable mixture of hydrocarbons with different composition and heat value depending on the area of harvest (Paylor, 2017). Compared to other fuels such as coal, woodchips, and fuel oil amongst others, natural gas has significantly reduced Green House Gases (GHG) emissions (Bo, 2017). Figure 1.1 on the next page shows various applications of natural gas and the latest transportation technologies.

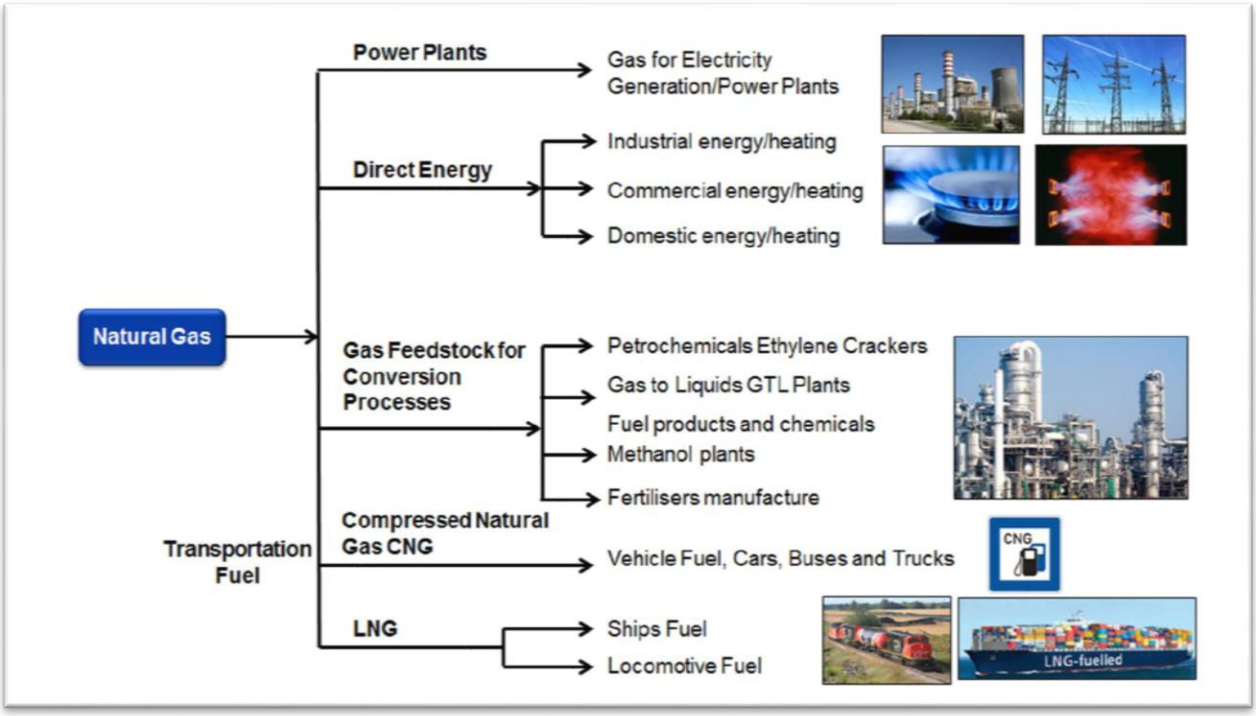


Figure 1.1: Application of natural gas.

Source: Transnet SOC, 2016, p.374.

Globally, natural gas has the fastest growing consumption rate compared to other fossil fuels (Akpinar and Yumusak, 2016b). The increased demand is attributable to factors such as cost per cubic metre, level of emission compared to coal and heavy fuel oil, on tap availability, operational flexibility and high efficiency (Ackah, 2014).

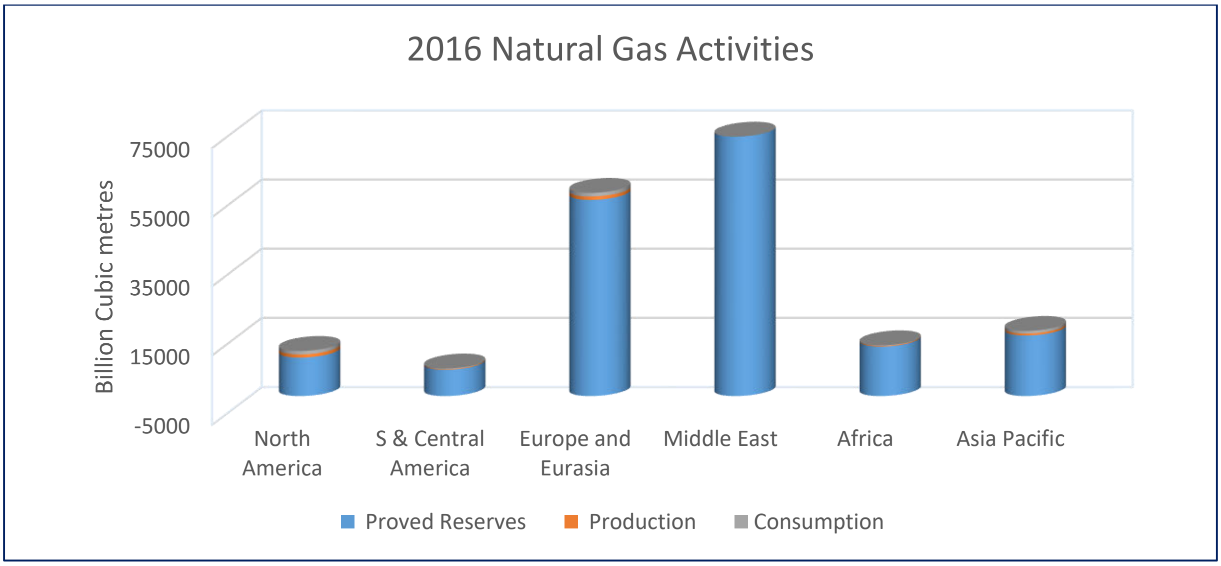


Figure 1.2: Gas consumption, production and reserves of natural gas around the globe.

Sources: BP, 2017.

Globally, natural gas accounts for 24% of the total energy mix (IEA., 2017). Consumption growth in the past decade has been flat on average and projected to be hovering around the 2.4% mark in the next decade (IEA., 2017). Based on the natural gas activities in Figure 1.2, gas reserves are more than current production and consumption trends. Global gas availability is evident at 186 tfc reserves, production of 3551 bcm and consumption of 3542 bcm (IEA., 2017). Consumption is forecasted to increase from 3542 bcm in 2016 to 177 tcf in 2040 (IEA., 2017). Natural gas consumption has risen substantially compared to other fossil fuels, it now holds the second place in the category of the most consumed renewable energies and remains the preferred fuel in the generation of power and heating fuel to the industrial sector (IEO, 2017).

The low Green House Gas emission of natural gas makes it a more attractive fuel option around the world. In 2016, the emissions of Green House Gas (GHG), mainly CO_2 from energy consumption increased by only 0.1% - which remains the lowest recorded in the period 2014 to 2016 compared to the previous decade (IEA., 2017). GHG emissions contribute immensely to global warming. To curtail the continued air pollution and its consequences, countries signed the Kyoto Protocol held in Paris (Esso and Keho, 2016). In 2015, the 21st United Nations conference session was held in Paris. This conference was considered a significant milestone in the commitment of nations to the reduction GHG gases since it identified the precise source of gases and how to find suitable replacements for energy sources (Esso and Keho, 2016).

Natural gas was introduced in South Africa in 2004 by Sasol Energy replacing hydrogen gas (Sasol, 2014). Hydrogen gas is a byproduct of coal processing in Secunda. As a result of this product's tendency to burn and generate heat, Sasol decided to develop a gas market (Sasol, 2014). Delivering gas to customers required a massive investment in the gas infrastructure which covered provinces such as Gauteng, Free State, KZN, and Mpumalanga provinces (Sasol, 2014).

The justification for the initial gas network investment was based on forecasted potential gas customers and their consumption in these provinces. Equally, significant investment was made to construct a pipeline to import natural gas from Mozambique to South Africa. The availability of gas usage information steered the decision to invest in the 865km ROMPCO (Republic of Mozambique Pipeline Company) gas line. The construction of this system was easy because of the already established gas market and quantifiable offtake. Consequently, natural gas consumption in South Africa gradually increased from 2 to 4.9 bcm from 2004 to 2014 (Sasol, 2014). Sasol Energy is the sole supplier of natural gas in the country and imports about 4.6 bcm. The balance is synthetic Methane Reach Gas, which is produced from coal and is similar to natural gas (Sasol, 2014). As an employee of the case study company, the researcher has noted that Sasol is failing to meet the industry gas demand.

Natural gas in South Africa is still at its development stage, and as such there are five active gas traders namely: Spring Lights Gas, Reatile, Novo Energy, Natural Gas Vehicles and Sasol Energy. Sasol Energy is the bulk supplier of gas to other industry players and the most prominent trader (NERSA, 2015b).

Natural gas demand for a country or a province is determined by analysing the expected energy needs of the industrial, commercial, and residential sectors. Szoplik (2015) highlighted that industrial customers consume gas at a constant rate throughout the year. However, commercial and residential consumption is seasonal. Variation in gas consumption in the sectors mentioned above depends on different factors such as the time of the day, weather, gross domestic production, gas price, demographics, building type, operating facility for production, and product needs among others (Szoplik, 2015).

Residential consumption is usually higher in the evenings for cooking purposes and in winter for space heating. Gas consumption in the industry increases with a decrease in price, an increase in the cost of the alternative fuel, product demand, and consumer disposable income. On a monthly and yearly basis, Sasol Energy is required to submit forecasted gas usage to ROMPCO. A bottom-up gas demand forecast is done to ascertain the necessary quantity of gas and should the actual gas usage be lower compared to the amount forecasted, penalties apply (Akpinar et al., 2017).

The gas pipeline is the most preferred form of transportation because of safety and on tap availability of gas. To access customers that are situated far from the network, liquefied and compressed natural gas becomes a feasible option because of the flexibility of delivery. Take or Pay contracts are common when massive investments are made in developing all forms of gas infrastructure to ensure that projected revenue in line with the forecasted gas usage is guaranteed (Akpinar and Yumusak, 2016a).

These contracts are concluded on the bases of longterm forecasted gas usage. Take or Pay agreements are structured such that when the forecasted quantity of gas is not consumed, the consumer will be required to pay the difference between the actual gas usage and the estimated gas usage at the agreed gas price (Akpinar and Yumusak, 2016b). To avoid punitive measures in these agreements, gas demand forecast modeling with minimum error is critical.

1.4 Motivation of the study

The bottom-up forecasting of gas quantity involves a customer, trader, producer and gas transportation companies. It is a common practice for the gas value chain which includes producers, traders, and users to be contractually bound by a take or pay contract condition. The ability to forecast required gas with minimum error will assist the customers, traders and producers to optimise their input cost which will positively influence the bottom line (Akpinar and Yumusak, 2016a).

The identified statistical model will ensure that the natural gas supply value chain is empowered to forecast gas with minimal error. The gas industry is still growing in South Africa, and additional infrastructure is required to roll out gas nationally, and a reasonably accurate forecasting model is essential in increasing the investment appetite in the industry because of quantifiable returns. The motivation of the study is to present a statistical model that can reasonably forecast the amount of gas which will help avoid take or pay penalties which could amount to millions of rands and to quantify the profit margin for the investors in the gas infrastructure (Akpinar and Yumusak, 2016b).

1.5 Focus of the Study

Sasol Energy is the sole supplier of natural gas in the country, and currently, the company is unable to meet the required industry gas needs. One of the options available to producers and traders is to optimise the limited available gas to improve profit margins (SLG, 2018). The current study focuses on closing the research gap by estimating and modeling a statistical method that could be used by Spring Lights Gas to forecast future natural gas consumption. The study did not focus on the gas price elasticity, which plays a role in the growth of the industry and consumption.

1.6 Research Objectives

a) To use the univariate actual gas time series variable to forecast gas usage for Spring Lights Gas (SLG).

b) To establish a bivariate causal relationship between actual gas consumption and purchasing management index (PMI) time series in forecasting gas for SLG.

c) To identify which method between the univariate and bivariate has a smaller error in forecasting gas for SLG.

d) To use the method with smaller error to forecast future gas usage for SLG.

1.7 Research questions

a) How can the univariate actual gas time series be used to forecast gas for SLG?

b) How can the bivariate causal relationship between gas consumption and PMI be used to forecast gas for SLG?

c) Which method has a smaller error?

d) How does the smaller error method forecast future gas consumption for SLG?

1.8 Methodology

The time series regression methodology is the blueprint for modeling and forecasting the quantitative variables of interest because of the inherent unbiased forecasting. Thus, everything required to continue the historical pattern is included in the variable. In this approach, attention is given to the recent historical data (Makananisa and Erero, 2018). Two popular time series regression methods, ARIMA and Vector Error Correction models were used to forecast gas consumption.

The type of evidence required to answer the research questions is actual historical gas consumption data and Purchasing Management Index (PMI). PMI measures the business cycle movements of the manufacturing sector and its performance (Bernstein, Hedeberg, Helu, and Feeney, 2018). Historical consumption data of the southern region (Mobeni, Jacobs, Prospecton, and Umbongintwini) was collected from Spring Lights Gas based in KZN and historical PMI data collected from the Bureau of Economic Research. The rationale was to use the historical univariate and bivariate time series variables to predict future gas demand and casual relationship between gas consumption and PMI respectively. The statistical models were analysed using the EViews statistical data analysis package. The advantages of the data collected from the trader is that it is readily available and with certain validity and reliability as detailed in Chapter 3 of the study.

1.9 Limitations

As a senior employee of the case study company, the researcher has noted that historical gas consumption is competitive information that can be adversely used by the competitor in the market, thus the challenge in accessing gas consumption from other traders which would have enabled a study of the whole province. Equally, the study will not be appending the actual customer's consumption.

1.10 Sampling and data source

The KwaZulu Natal south region has a population of 28 gas customers procuring gas from Spring Lights Gas. The firm has granted access to a monthly total volume of customers to model and estimate future gas consumption. The PMI data used in this study was sourced from the Bureau of Economic Research. As an employee of the case study company, the researcher has noted that SLG aligns its future gas consumption budget with PMI movement, hence the decision to use it.

1.11 Conclusion

Chapter one has discussed the study background, natural gas, South African gas industry and presented an overview of the study. It outlined the motivation for the study; the focus of the study; objectives, research questions. Study limitations have been highlighted. Chapter two presents the literature review on study objectives covered on the subject and detailed background of the gas industry.

CHAPTER TWO
Literature Review

2.1 Introduction

Chapter one enables Chapter two to review and build a literature review in line with the research objectives identified. This Chapter further provides a comprehensive background on the energy scenario in South Africa, the gas industry, identifies a required policy to enable the industry, different sources of natural gas, and the relevant literature regarding forecasting techniques.

The study seeks to test and recommend the suitable statistical method to forecast natural gas for Spring Lights Gas. Natural gas has a weighting of 3% of the total energy mix in South Africa (DoE, 2013b). The overarching objective of the White Energy Paper is to promote the well-diversified energy mix (DoE, 2010). The revised IRP 10 further presents avenues where gas could displace diesel in the generation of power using an open circle turbine (DoE, 2010).

As a senior employee of a case study company, the researcher has noted that the limited available resources of gas and the related required infrastructure to deliver natural gas further puts a strain in archiving the intentions of the White Energy Paper. The natural gas industry is capital intensive and guaranteed healthy profit margin could increase the appetite of investors (Akpinar and Yumusak, 2016a). Before undertaking the investment in the industry, a significant supply resource should be identified, followed by how much gas will be used by the anchor customer, from which the required infrastructure will be established.

2.2 Energy scenario in South Africa

In 2008 South Africans experienced power shortages due to Eskom's compromised generating capacity and the old power stations (Goldberg, 2015) which necessitated a different look at possible energy sources. The following subsections below discuss the South African energy mix and balance.

2.2.1 Energy Mix in South Africa

Energy mix in South Africa is primarily based on the abundance of domestic coal, imported oil with limited renewable energy (solar & wind) and natural gas (DMR, 2015). Coal enables power generation; which is supplied to the industrial, residential and commercial markets. Figure 2.1 shows the weighting of the various fuels in the country. As shown in the graph below, coal is a dominant Energy source while natural gas and nuclear have a weighting of three-percent. Oil at 23% and its derivatives command a greater share of the weighting (DMR, 2017).

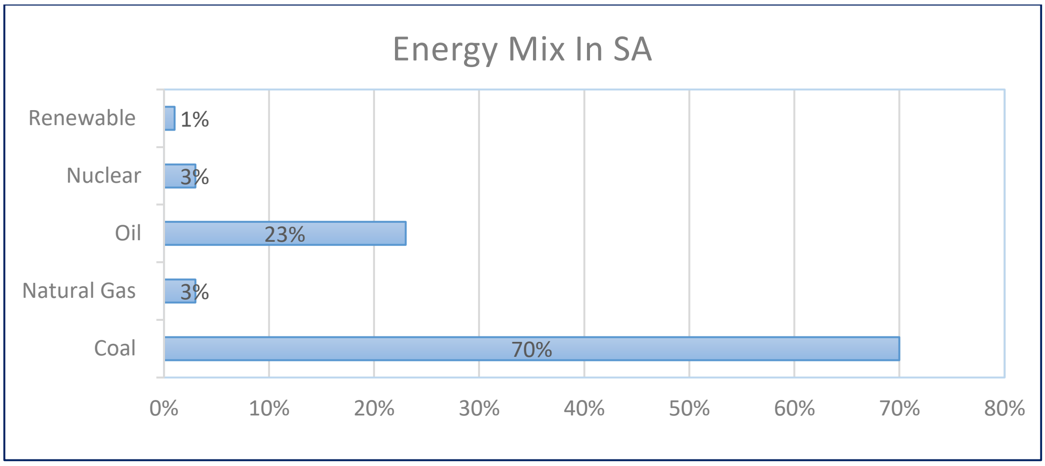


Figure 2.1: Energy mix is South Africa.

Source: IEA, 2015.

2.2.2 Energy balance in South Africa

The South African energy system is currently self-sufficient with less than 20% of import (Wright, 2016). Sources of energy imports comprise of crude oil feedstock refined into liquid fuel, a small amount of gas, and the balance of energy is supplied domestically from coal. Approximately 33% of the liquid fuel is supplied from gas to liquid fuel process (Wright, 2016). This process provides around 4-6% of the fuel (DOE, 2016). Moss Gas fields provide PetroSa with natural gas as feedstock for gas to liquid fuel production. The procurement of renewable electricity through the Independent Power Producers is on the rise. Towards the end of 2015, solar and wind power generation, together accounted for 4,650 Gwatt per hour (DOE, 2016).

Figure 2.2 shows coal as the dominant domestic energy source at 6101 petajoules (PJ), and 28 PJ imported. Most of the oil is imported at 1280PJ with local production of 9 PJ. Similarly, 128 PJ of gas is imported and 43PJ produced locally by PetroSa (IEA 2013; CSIR, 2016). Wind and solar contribute 25PJ the energy mix (Wright, 2016).

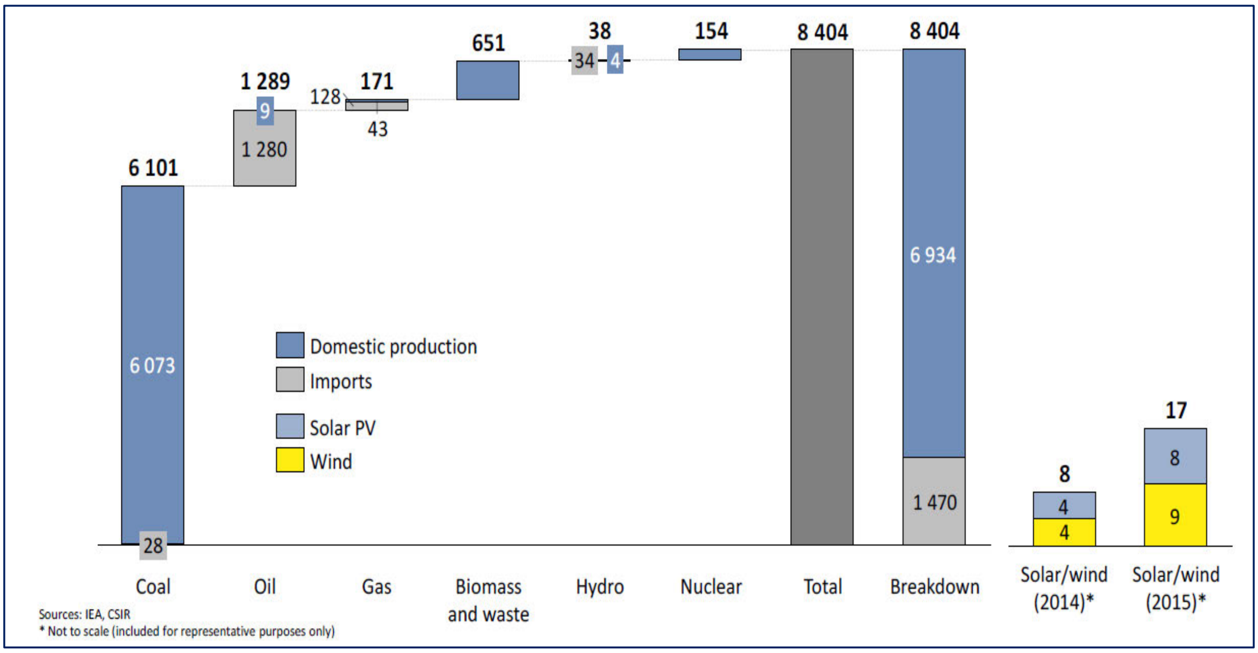


Figure 2.2: Energy balance in South Africa.

Source: CSIR, 2016, p.11.

2.3 History of natural gas

Cape Gas situated in Cape Town began in earnest as a gas industry in South African in 1847. The industry expanded to Port Elizabeth, Kimberly, Grahamston and Johannesburg (Ethekwini Municipality, 2015). Sasol Gas reincarnated the South African gas industry in 1966 (Van Basten, 2007).

The abundance of hydrogen gas from coal gasification and the methane gas synthesis process triggered the re-establishment of the gas industry in South Africa (Collings, 2002). At the time, the regulatory framework was non-existent; a situation which made it easy for Sasol Gas to construct underground pipelines to supply customers (Van Basten, 2007). As a result, Sasol owns all the transmission and distribution networks used to transport gas to commercial and industrial customers in Gauteng, Free State, KZN and Mpumalanga provinces (NERSA, 2015a). In 2004, Sasol replaced hydrogen gas with natural gas imported from Mozambique using an 865 km pipeline. In the south of the country, PetroSa discovered about 28 bcm of natural gas offshore in 1984 and the first gas was delivered in 2000 (PetroSa, 2016).

The current gas fields have reached the maturity stage reducing the production capacity by 40% (PetroSa, 2016). To mitigate the gas supply shortages, PetroSa undertook to drill a new field (Ikhwezi gas field project) to harvest additional gas. Figure 2.3 below shows the current status of natural gas supply. Business As Usual (BAU), represents the current gas field supply and the F-O filed represents the Ikhwezi gas field which was expected to produce 6.8 bcm billion but instead, produced only 0.7 bcm. The diagram shows that the field had a lifespan ending in 2013.

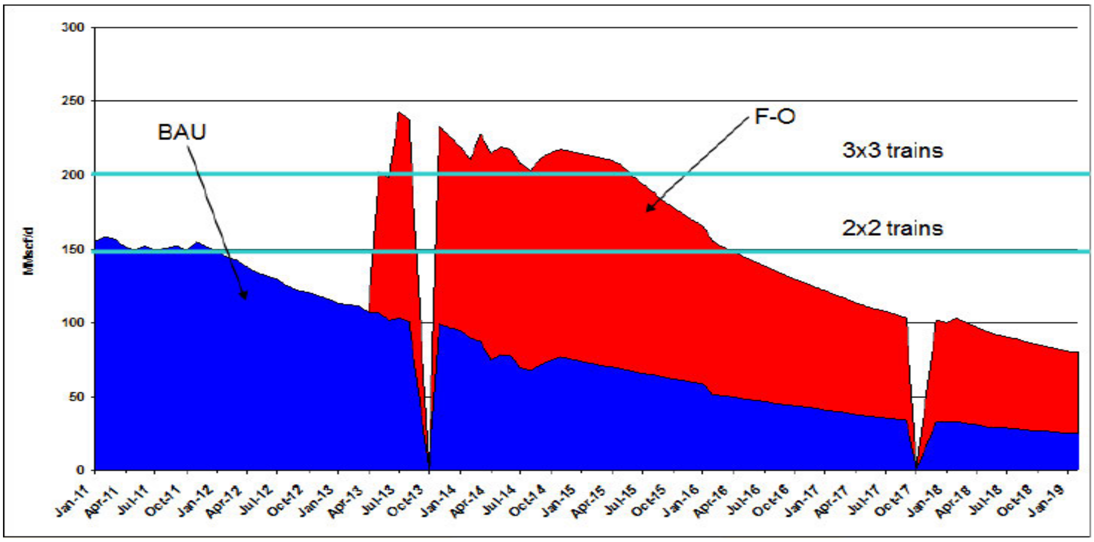


Figure 2.3: PetroSa operation capacity.

Source: PetroSa, 2016, p.7.

According to PetroSa (2016), the current matured gas field lifespan was initially estimated to have been 20 years worth of natural gas supply and was later revised down. The decision to proceed with the Moss Gas project was taken against the uncertainty of a long-term supply for strategic reasons.

2.4 Gas industry in South Africa

As an employee of the case study company, the researcher has noted that the gas industry is growing at a slow pace due to the limited gas supply. The industry is structured by activities ranging from production to supply, the role played by the regulator and the identified policy which anchorages further development of the industry.

2.4.1 Market structure

The market structure of the South African gas industry chain comprises of various components namely: upstream - where production and exploration activities take place; midstream – involving gas transportation activities which include transmission and distribution as well as the downstream - which comprises of trading, reticulation, and end-user (NERSA, 2015a). The following market activities are described in detail in the next sub-sections. Figure 2.4 depicts the gas industry market structure.

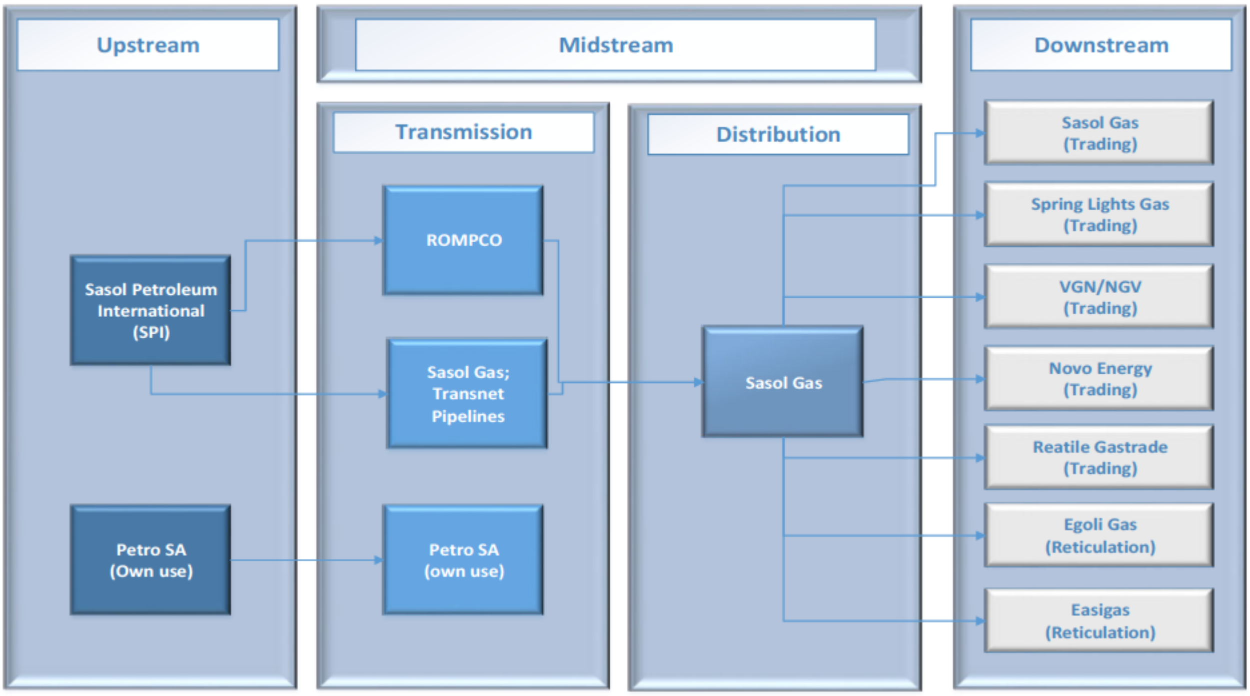


Figure 2.4: South African Gas Market structure.

Source: NERSA, 2015, p.3.

2.4.1.1 Upstream

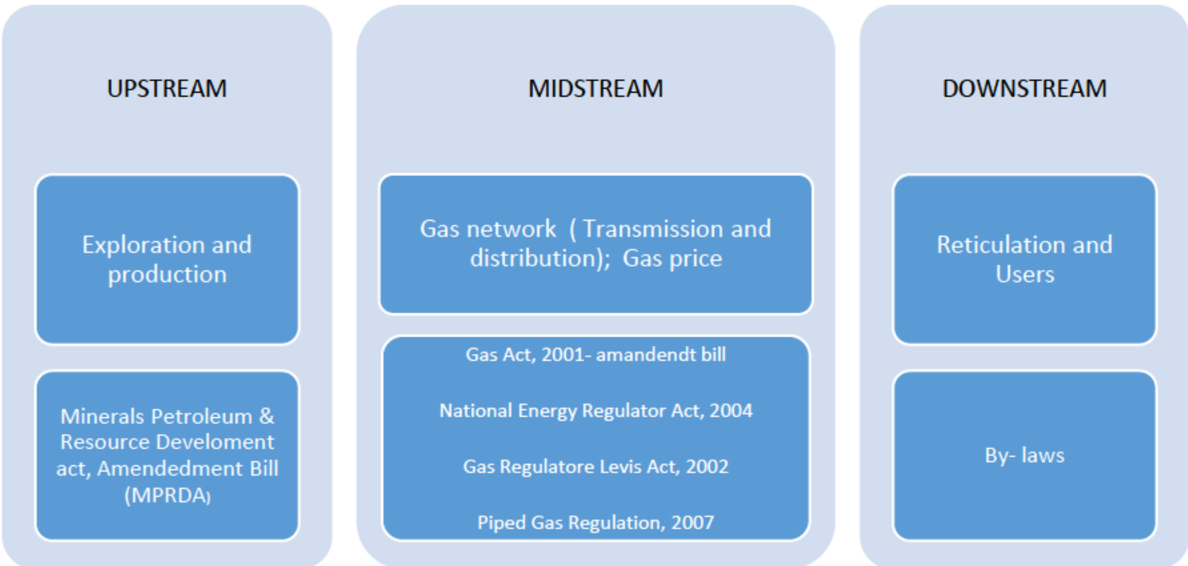


Figure 2.5: South African policies related to the gas industry.

Source: EThekwini Municipality, 2015, p.86.

In 1994, a new administration was elected in South Africa, which required policy review such as fiscal and energy among others. The new era opened the economy to international trade, and the Department of Energy had to ensure appropriate supply and use of energy. This action gave birth to the 1998 Energy White Paper on energy policy.

2.5.2 Summary of the applicable policies to the gas industry

Policy	**Summary**
Energy White Paper of 1998	The core of the policy is to promote various sources of energy in the South African energy mix with the following objectives: • Increased accessibility of reliable and affordable energy services • Promote credible energy governance per province • Stimulate competition in the industry for economic growth • Secured supply diversification • Ensure and manage environmental friendly energy use. The policy promotes natural gas in the country, and it recognises it as the best option for the future. The policy further provided the foundation of the development of the National Integrated Energy Plan (DoE, 2013b).
The national Integrated Energy Plan	The IEP provided a base for a long time energy planning and guidance of the required future infrastructure and policy development. The plan addresses several South African energy objectives one of which is to promote environmentally friendly energy resources. Gas is earmarked to play a crucial role in the reduction of emissions. The IEP presents energy options that could be pursued by RSA for current and future energy demand, promotes and optimises energy sources currently used such as electricity. Natural gas was modeled to assess its impact to facilitate moving the country from a high to a reasonably low carbon economy. The objective mentioned above is echoed by the National Energy Act (DoE, 2013b).

Integrated resource plan (IRP) 2010	The IRP 2010 revised objectives and set out specific targets for retiring different and presenting new energy resources. It is a long-term plan for power generation for the period of 2010 to 2030 indicating various technologies and own timeline for commissioning. The roadmap provided guidelines on the energy mix including nuclear, coal, various forms of renewable energy, gas and power generation capacity for the following two decades (DoE, 2010).
National Development Plan 2012	Promotes gas as an alternative that can play a crucial role in transitioning South Africa to a carbon-free emission economy. It highlights some construction projects for infrastructure development including Liquified Natural Gas to fuel a combined heat power gas turbines. It advocates for the safe and environmentally friendly technology of shale exploration and harvesting. If the government of the day can prove that the Karoo has considerable technical recoverable gas resource and adequately address the concern of the stakeholder, then the development of such resource is encouraged and should be fast-tracked. It further proposes incentives to companies that will use LNG to fuel their fleet (DoE, 2013b).
Gas Act 2001, Act 48 of 2001	The Act provides guidelines on the orderly development of the gas industry in South Africa. It established the National Energy Regulator of South Africa (NERSA) as the custodian of the Act to regulate and enforcers in the industry. The regulator as mandated by the Act regulates all hydrocarbons gases transported by pipe, including natural gas, methane-rich gas, re-gasification of LNG, CNG, produced gas, coal bed methane gas, and LPG (Gas Act, 2001). Provided guideline for NERSA to approve and regulate construction licenses for gas distribution and metering assets, and operation licenses of newly built assets, storage, liquefaction or re-gasification facilities, and trading licenses (Gas Act, 2001).

	Defines registered activities for gas producers, gas importers, transmission of gas for own use, gas reticulation, and LPG activities below two bar. Section 21 of the Act gives NERSA authority to impose license conditions on piped gas activities. Sections 22 of the Act outlines the requirements of pricing and provide guidelines on the development of a pricing model (Gas Act, 2001). The noteworthy exclusion on the Act are: • Upstream sourcing of LNG. • LNG liquefaction. • LNG transportation by sea, rail, and road. • Compressed natural gas. • Gas network below 15 bar. The Gas amendment bill was presented and approved by the cabinet in 2013 and open to public comment, the process is still underway. The bill aims to : • Monitor compliance and to enforce where necessary. • Deliberate on the changes in the gas environment. • To consider new technologies to harvest convention and nonconventional natural gas. • Transportation of natural gas LNG. • Regulate the regasification of LNG and compression rate (GasAct, 2001).
Piped Gas Regulation, Gazzeted 29702, 201 April 2007	Part of the Gas Act, which states that NERSA can not set prices for the gas market but can review wholesale and trader price in line with the maximum pricing methodology. It allows the regulator to review the application for the maximum price on a yearly basis and can request the amendment of adjustment to the maximum price of a trader. The maximum price of the trader is the cap price (DoE, 2013b).

Gas Regulation levis	The purpose of the levies paid by gas users in the industry is to fund the regulatory for administrations and another cost. Levis are revised yearly (GasAct, 2001).
National Energy Regulator Act 40 of 2004	The act gives NERSA the authority to administrate and force the provision of the Gas Act 48 of 2001, Petroleum pipeline Act 60 of 2003. NERSA is also mandated to act out the function stated in section 4 of the Electricity Regulation Act 2006 (NERSA, 2016).
Integrated Resource Plan (IRP) 2010	The IRP 2010 revised objectives set out specific targets for retiring different and presenting new energy resources. It is a long-term plan for power generation for the period of 2010 to 2030 indicating various technologies and own timeline for commissioning. The roadmap provided guidelines on the energy mix including nuclear, coal, various forms of renewable energy, gas and power generation capacity for the following two decades (DoE, 2010).

2.5.3 Gas Utilisation Master Plan (GUMP)

Gas Utilisation Master Plan is a roadmap developed by the DoE to assess the potential and opportunities to develop the gas industry in South Africa and the economic aspect of the resource (DoE, 2016). The plan is expected to fit into other policies and energy plan for the future.

The natural gas account for three percent of the total energy mix in the country and the objective of GUMP is to diversify the resource and gas economy intensely. The GUMP draft has been presented to the parliament in the 3rd quarter of financial year 2016/17. At the time of this report, the document was awaiting approval from the cabinet (DoE, 2016).

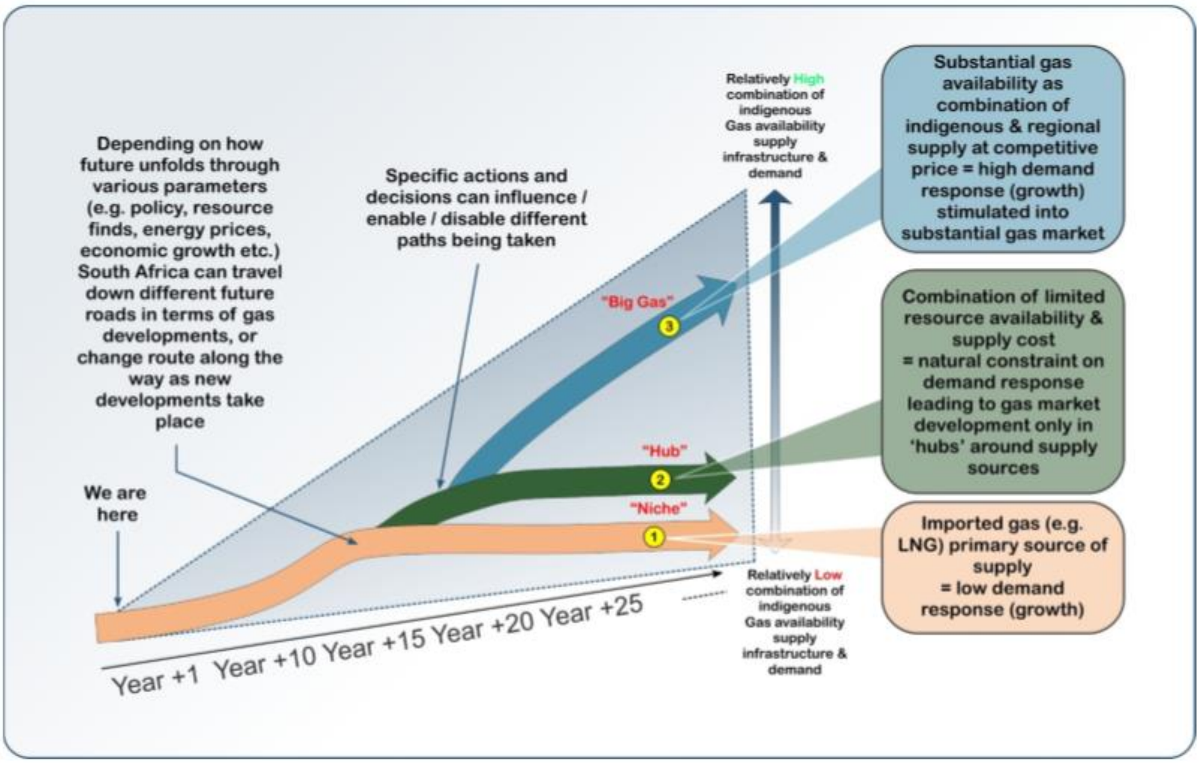


Figure 2.6: Gas Utilisation Master Plan (GUMP) illustration of possible future gas market evolution.

Source: CSIR, 2016, p.22.

Figure 2.6, provides the roadmap to the development of the gas industry. The plan aims to provide guidance and clarity to the traders, producers, and investors about opportunities in the industry. It further gives a view of possible various potential and quantities of sources of energy and time frames.

2.6 National Regulator

National Energy Regulator of South Africa (NERSA) (Gas Act, 2001) derives its mandate to regulate electricity, piped gas, and petroleum piped industry in terms of the National Energy Regulation Act, 2004 (Act No.40 of 2004), in terms of Electricity Regulation Act, 2006 (Act No. 4of 2004), Gas Act, 2001 and Petroleum Pipeline Act, 2003 (Act No.60, 2003).

Section 21(p) of the Gas act prescribes that NERSA should, among other things, regulate and approve maximum prices for wholesales, traders, and reticulators regarding Chapter 2 and 3 of the Competition Act, 1998. NERSA regarding section 4 (g) of the Gas Act and the subsequent section 21 (1) p is mandated to approve the cap price, meaning that the actual price is determined by the traders and should not exceed the maximum rate. An assessment of whether there is enough competition in a given industry precedes regulation, and as a result, the regulator is required to evaluate the state of the competition in the gas industry bi-annually as contemplated in the Competition Act.

Other functions of NERSA are to approve construction and operation license for transmission, storage, distribution, and the operation of such infrastructure regarding section 15 (1) of the Gas Act (Gas Act, 2001). Trading in gas requires a license which is evaluated and approved by NERSA in terms Section 4(iv) of the Gas Act (Gas Act, 2001). Trades are required to apply for a trading license in a specific area where they would be allowed to trade gas.

In the environment where there is limited gas and one supplier, market abuse is inevitable. Hence section 22 of the Gas Act gives authority to NERSA to regulate and approve the pricing mechanism used by the trader to price gas. The pricing mechanism of all the traders should be modeled such that customers are priced equitably except by objectively justifiable conditions such as the quantity of gas. This implies that the more gas is used, the less the price paid.

2.7 Sources of natural gas

As a senior employee of the case study company, the researcher has noted that Sasol as the sole supplier of natural gas in the industry is unable to meet the demand. Developing a fully functional industry, an additional supply of gas is required as well as the infrastructure to transport it. The following subsections discuss the possible sources of gas that the country could explore.

2.7.1 Importing of natural gas

PetroSa is the only South African company which is involved in the offshore gas extraction activities in Mosel Bay. The harvested gas is used as feedstock in the gas to liquid process. A significant supply of natural gas is imported by Sasol Gas using 865km ROMPCO line from Mozambique. Sasol is the only supplier of gas to the external market, and it supports the South African gas industry. The increased supply to the external market has positively influenced the weighing of gas in the total energy mix from 2 to 3% (Merven et al., 2017). Gas usage is forecasted to grow by seven percent over the next 15 years with Eskom anticipated being the possible anchor customer (eThekwini Municipality, 2015).

Sasol currently sells gas to 550 commercial and industrial customers and has displaced coal as feedstock at the Sasolburg operation in the drive to reduce carbon footprint (Sasol, 2014). The natural gas supply in the country has grown from 1.4 to 4.4 bcm 2014 (Ethekwini Municipality, 2015). For the industry to continue growing, a significant supply of gas and the related infrastructure is required. The natural gas demand continues to grow in South Africa, and ROMPCO undertook to increase the capacity of the Mozambique Secunda Pipeline (MSP) by installing compressor facilities at Komatipoort, Mpumalanga.

The compressor increased the capacity to an additional 0.6 bcm (NERSA, 2017). The second expansion of Mozambique Secunda Pipeline (MSP) was introduced in 2014 with a loopline of 128 Km which runs parallel to the MSP as shown in Figure 2.7. The loopline was installed to supply the growing demand in Mozambique. The third expansion of the MSP was the second loop line which will initially transport an additional 0.2 bcm and ramp up to 0.61bcm (NERSA, 2017). The expansion of the infrastructure mentioned above was justified by quantifiable gas demand.

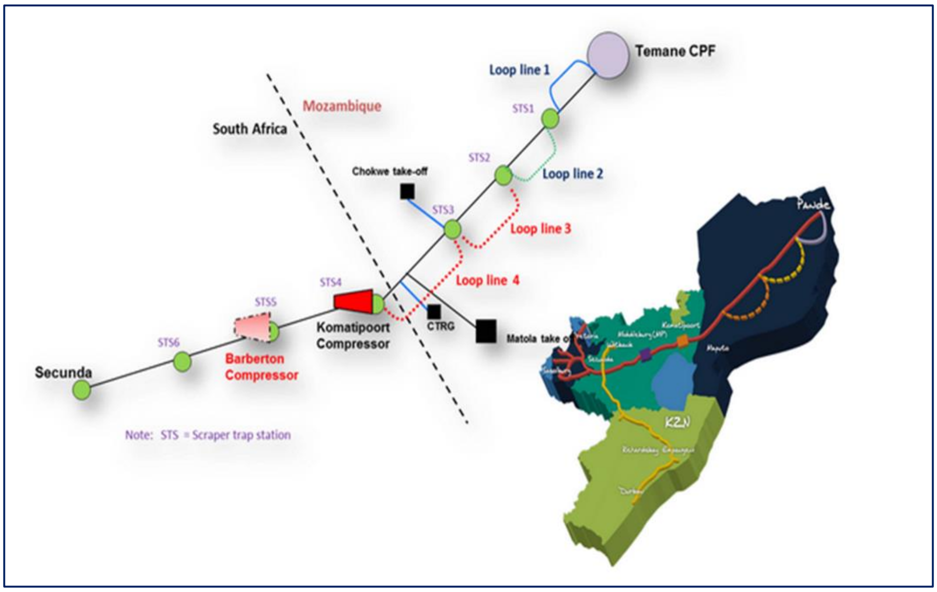


Figure 2.7; ROMPCO gas pipeline.
Source: ROMPCO, 2018.

The logical way of additional sourcing of gas would be to import from other neighboring countries such as Angola, Namibia, and Tanzania among others. Tanzania has a well-developed gas industry and has been growing at an exponential rate. According to DBSA (2017), Tanzania has estimated unproven gas reserves of 1.4tcm. Currently, gas production is quantified at 1,8bcm with the forecasted production of 3.5bcm in 2026 from Songa Songa and Mnazi Bay fields. South Africa is in a favorable position to take advantage of the available proven gas reserve in Tanzania. A trade agreement between the countries could increase natural gas supply and maximise the use of the increased capacity of the MSP line.

In 2002, ROMPCO entered into a Gas Transportation Agreement with Sasol Gas to initially transport 3bcm of natural gas. According to DME (2005), the MSP line was constructed at the cost of $549million US dollars, and the returns for the investments was quantified based on the possible shipping of gas required by the South African gas industry. The agreement comes with the condition of ship or pay of 80% of the annual forecasted 3bcm (NERSA, 2017).

The penalties will apply should Sasol Gas ship below 80% of the 3bcm with the company liable to pay the difference of the actual gas shipped and 80% of the nominated quantity. With the increased supply of gas, the agreement has been amended to transport 4.6bcm in 2014 per annum to South Africa (NERSA, 2017). The initial pay or ship agreement is still applicable to the new gas. The take or pay conditions require consistent forecasting with minimum errors to avoid penalties.

2.7.2 Fracking

South Africa is rated among the top ten countries in the world with 390 tcm technically recoverable shale gas. The safe mining of this gas could be the game changer for the energy landscape in the country. The following subsections discuss the fracking process, Government role and how the locals have received fracking.

2.7.2.1 Global view of shale gas

Globally, the technically recoverable shale gas is estimated at 7112 trillion m^3 (IEA, 2011). This is not the total quantity of gas trapped in shale rock, but the amount that can be extracted using proven technologies. The reserves mentioned above technically increase the global supply of natural gas by an estimated 45% (IEA, 2011). According to the UNDP's (United Nation Development Plan) Human Development Index, highly developed countries hold approximately 43% of the estimated technically recoverable shale gas.

According to Paylor (2017), highly developed countries hold nearly 33 %; medium 21% and less developed roughly 3%. The development index demonstrates that highly developed countries have a better chance of recovering shale gas because of the available required infrastructure (Paylor, 2017). Equally, healthy political landscape, coupled with economic stability and socially coherent environment of a highly developed country allows an environment where exploration can be carried out in determining the country's shale gas reserves. Highly developed countries with a well-balanced energy mix and a large energy market provide a platform that allows a more fluid commercial benefit of exploring shale gas (Paylor, 2017).

Canada and the USA have been commercially extracting shale gas for the past two decades (De Rijke, 2013). However, recent interest in commercial shale extraction has been on the upward trajectory. For example, China had its first commercial shale extraction in 2011(Hu and Xu, 2013); The UK has awarded 176 licenses for onshore fracking (Davies et al., 2014). Algeria adjusted its hydrocarbon law to stimulate the interest in the commercial fracking (Paylor, 2017).
The South African government overturned a moratorium on fracking (Fig and Scholvin, 2015). Equally, some countries have rejected the commercial fracking for different reasons and the primary one being the pollution and exposure to radiation that come with shale gas. France, Germany, Bulgaria, Romania, and Luxemburg are among the countries that have banned hydraulic fracking (Moore et al., 2014).

2.7.2.2 Shale production

Technological advancement has presented two ways which are currently used to mine shale gas. The first one involves directional drilling to extend the well from the well pad, secondly, the hydraulic fracturing which strongly attracts public interest (Striolo, 2018). Stanolind Oil and Gas Operation used the fracking technology for the first time in the 1940s in Kansas, USA. To date, millions of fracking jobs on oil and gas wells have been done throughout the world.

The technique of fracking involves the stimulation of gas flow to a drainage area to upscale the production of the well. The site drainage is generally the reservoir of hydrocarbons trapped in a fractured sub-surface rock formation. Shale is a sedimentary rock punctured by small fissures wherein gas is trapped (Schrag, 2012). In the past, the cost of mining shale gas was high; a situation which eroded profit margins considerably and rendered the extraction unviable.

The recent technological advancement of horizontal drilling which allows maximum productivity from shale rock up to a cross distance of 4 metres from a vertical well pad (Ehrenberg, 2012), with the increase in oil price (global natural gas price index) meant that the profit margin generated by fracking has influenced a decision to undertake the project. Fracking is a three-stage process which involves exploration, production, and abandonment.

2.7.2.3 Exploration

Exploration stage commences with the drilling of two to three wells to test the availability of shale gas, the features of the shale rock, and the complexity of extraction. Fifteen to sixteen more wells may be drilled with the fracturing of a small area of shale rock to determine the features of the rock to determine how to propagate the fissures and the viability of extraction (Paylor, 2017). Further to that, the drilling of additional wells may be required to evaluate the long-term viability of the well fully. The recent trend of the shale gas extraction reveals that in general, profitable well is thousands of metre below the surface of the ground and hundred metre wide (Paylor, 2017).

2.7.2.4 Production phase

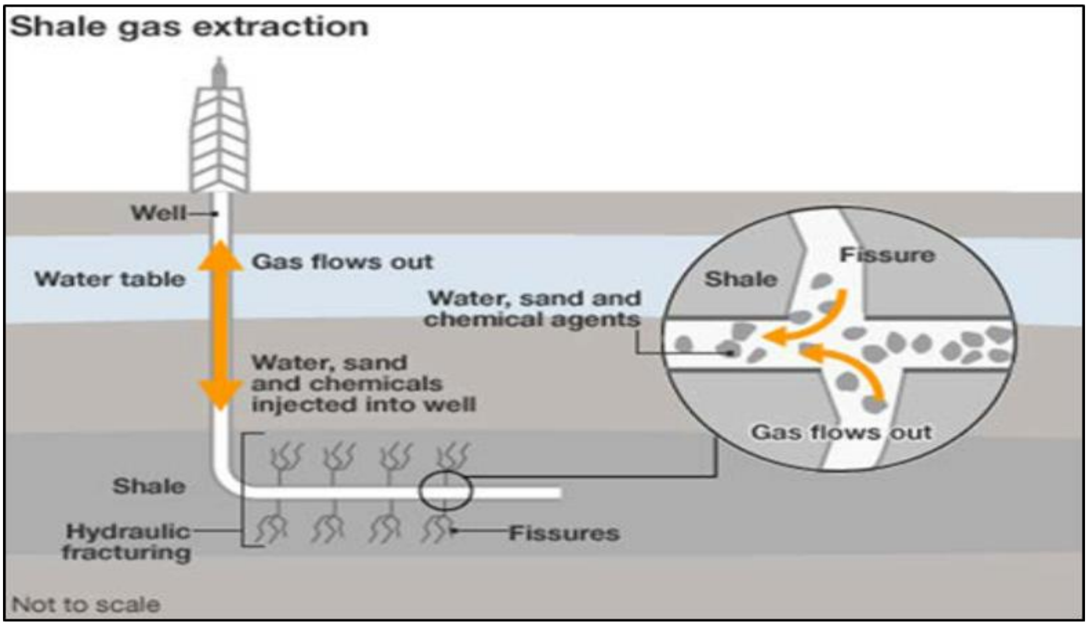


Figure 2.8 Shale gas extraction.

Source: Transnet SOC, 2016, p.373.

Figure 2.8 shows the production of shale gas. Production stage follows the proven viability of the well and its long-term viability. The shale gas shaft takes an L shape with extended vertical drilling followed by a horizontal borrowing to a cross distance of about 4 metre as shown in Figure 2.8.

At the end of the shaft, perforation is done to create a toe which will enable the fracking to commence. The process of hydraulic fracturing encompasses the high pressure propelling of water, sand solution, and additives into an exhumed mine at high pressure. The pressure is used to penetrate the fissure of the rock to fracture and allow trapped gas to escape (Paylor, 2017).

Fracturing is water intensive and thus requires an estimated 7,5 to 15 million litres of water throughout the lifetime of the well (Kondash and Vengosh, 2015). The list of additive used in fracking varies from toxic at a low level of concentration to medium and less. Some of the additives are benzene, ethylbenzene, toluene carbonyl sulfide among others (Weinhold, 2012).
The additives mentioned above are classified as hazardous air pollutants (Paylor, 2017). The objectives of the additives in the fracking process are to prevent bacteria growth, corrosion and lubricate to enable the proponents to permeate the cracks (Ehrenberg, 2012). The fractured shale rock will then release trapped gas together with fluid which contains arsenic, mercury, and other natural radioactive fluids. The escaped gas is captured at the wellhead. When gas is released, thirty to fifty percent of fracking fluid is recovered and recircled, and the rest is disposed of safely.

2.7.2.5 Abandonment phase

Generally, the lifespan of a well ranges from four to five years. At this point, the level of productions has significantly reduced and a process of shutting down commences. The underground freshwater stream is sealed off to the gas and wastewater by filling the well with cement.

2.7.2.6 Shale gas in South Africa

South African Shale gas potential was discovered at the Karoo basin in 1960 (Rowsell, 1976). The Karoo area is a vast geographical setting with a small population and low levels of economic activity. Currently, the area has no environmental management plan, and the quality of air is unknown due to missing quality monitoring systems (Altieri and Stone, 2016). It is conceivable that the production and development of shale gas in the Karoo will have a dramatic impact on the locals, environment, and economy (Steinzor et al., 2013).

The commencement of drilling for shale gas will play out for 24 hours, seven days a week and 365 days a year. Due to the activities in the surrounding area, noise output is high and consistent, pollutants from diesel power generators, 24 hour in and outbound trucking of water and additives.

Daily water consumption is estimated at 1000 m^3 to 5000m^3 litres, and about 30 to 50% of the fracking water mixture is recovered for purification and disposal (Altieri and Stone, 2016).
Shale production comes with significant sources of air pollution which spreads out to a vast geographical area. The primary pollutants are nitrogen oxide, non-methane volatile organic

compounds which could significantly affect the quality of air. Exposure to the pollutants mentioned above can course asthma, decreased lung function, increased hospital admission, chronic respiratory and cardiovascular challenges (Kim et al., 2015).

The other concern that comes with fracking is the contamination of drinking water. According to Darrah et al. (2014), drinking water contamination can happen in two ways, first through the leakage from the well casing and cement challenges. Secondly, through faulty drill casing. The purpose of drilling case during shale gas fracking and gas extraction is to avoid mixing aquifers with fracking fluids and extracted gas; however, if the casting is damaged or improperly installed, aquifers could be exposed and contaminated by fracking fluid (Paylor, 2017).

2.7.2.7 The economics of shale gas

South Africa is rated among the eight-nation with the technically recoverable vast shale gas reserves in the world. Shale gas is situated in the Karoo basin covering an area of 300,000km² of the interior of the country. The technically recoverable resource is estimated at 390 trillion cubic feet. The value of the shale gas is estimated at a low to a higher limit of 3.3 to 10.3 percent of the GDP with estimated job creation of 144 – 7000 jobs (Altieri and Stone, 2016) respectively. If fracking becomes a reality, the growth of the above economic activity will depend on the gas offtake. It had been suggested that fracked gas could be used to generate power.

The infrastructure required to transport gas from the fracking point to the required destination will require massive investment. In generating the revenue to justify the project, future gas offtake is required, and a reliable forecasting model is critical. As mentioned on numerous occasions in this study, contract with the punitive measure is preferred in the industry.

2.7.2.8 Public participation against fracking

The South African government first indicated interest to pursue hydraulic fracturing in 2008, which received a backlash from the public. The apparent rejection from the public forced the government of the day to institute a moratorium on fracking. However, eighteen months later the moratorium was reversed (Feodoroff et al., 2013). The public rejection of shale gas has gained momentum in the past four to five years from various organisations such as Treasure the Karoo Group, Afriforum, and Wildlife environment society of South Africa among others. Lack of Karoo community protection, natural environment and violation of the human right informs the objection

(Temper et al., 2013). These organisations continue to bring awareness to the Karoo community and to lobby for support from the greater South African population against fracking.

2.7.3 Liquefied Natural Gas

It is common practice to transport gas using underground pipes, and the length of the pipe varies from one destination to the other, for example, Sasol uses an 865km pipe to transport gas from Mozambique to Secunda South Africa (NERSA, 2016). Liquefied Natural Gas (LNG) enables the chances of accessing the gas market that is far from the gas pipe. The flexibility provided by the LNG de-regionalise and integrates the natural gas market around the globe (Barnes and Bosworth, 2015). LNG allows for short and long-term supply, and the proportion of such trade has increased from 18.9% in 2010 to 28% in 2015 (Feng et al., 2017). Equally, LNG flexibility brings vulnerability of security of supply to countries that immensely depended on it. Due to its flexibility and short-term agreements, importers face the danger of non-delivery because the supply is not always guaranteed, and for the exporter, sales are not always guarantees which can affect the margin of the trader (Wood, 2012). Trade links with countries that have abundant natural gas to ship are essential for energy mix and to augment energy needs.

2.7.3.1 Production of Liquefied Natural Gas

Natural gas supplied from the conventional well and unconventional fracking process should be first treated and cleaned before liquefaction. The process begins with the scrubbing of hydrogen and carbon dioxide using potassium carbonate (Bridgwood, 2015). The next stage is cooling which allows water to condensate and separate from gas. This stage is concurrently done with the drying of gas to remove remaining small particles of water. Water and carbon dioxide crystalizes below sub-zero temperatures which will cause blockages within the piping system. The composition of natural gas consists of other hydrocarbon chains which are separated from methane gas by using the distillation process at various vapor temperatures below zero (Bridgwood, 2015). The low temperature is achieved by heat exchange where gas losses heat and the coolant gains heat. Further cooling takes place and natural gas condensates and liquefies at -167 ºC (Mokhatab and Poe, 2012).

2.7.3.2 Liquid Natural Gas in South Africa

The LNG option is an optimal choice for South Africa to diversify energy resources and to augment natural gas supply. Currently, South Africa is not trading in LNG, and due consideration has been given to the development of the terminal and other required infrastructure for importing purposes (Wright, 2016). Massive investment is required to facilitate LNG growth in the country. Harnessing partnerships with key stakeholders such as Transnet in erecting imports terminals is crucial to get the business of LNG trading off the ground. All these are the initiatives that the South African government is considering in the growth of the gas industry (Wright, 2016).

The government of Mozambique has approved Anadarko Petroleum project to develop the liquefied natural gas terminal in the north of the country (Anarderko, 2018). The terminal development is costed at 20 billion rands and will facilitate the trade of LNG in the south region of Africa. South Africa has the opportunity to trade and import to PetroSa which is harvesting from mature wells.

Significant investment is required to construct gas infrastructure that will transport gas to all the industrial sites. The current gas infrastructure is limited to four provinces. The accuracy of gas forecasting is essential in ensuring that payback of the investment is realised.

2.7.3.3 Compressed natural gas

Energy demand around the globe increases with time, and the use of fossil energy sources is equally on the upward trajectory. The use of fuels produced from high carbon fossils presents disadvantages such as limited infrastructure to supply, waste management and compromising air quality due to emissions of GHG (Singh et al., 2016). The recent research shows that transportation of fuel contributes to the depletion of air quality, and many scholars have researched alternatives for the fuel economy and emissions. Eco-friendly fuels such as methane gas have been extensively researched, and it is one of the fuels advocated to replace diesel and petrol engines. Compressed natural gas is a promising and futuristic renewable fuel to power diesel and petrol engines (Singh et al., 2016). Pressures to conserve the air quality has promoted and forced a shift from petrol and diesel-powered engines. In the past decade, the number of natural gas vehicles has exponentially increased with the Asia-Pacific and Latin American taking center stage (Singh et al., 2016).

The use of Compressed Natural Gas (CNG) allows industries to access small users that are far from the pipe network and are environmentally sensitive. Trucks and mini-vans transport CNG from a remote production facility to a customer location (SLG, 2018). Another option available to the traders is to truck the filled modules to remote daughter station to CNG-fueled vehicles. SLG is the largest supplier of CNG in the country and an industrial customer with the estimated yearly volume of more than 2 million cubic metres (SLG, 2018). The CNG value chain consists of the following.

- Manufacturing
- Product delivering
- Unloading

2.7.3.4 Production

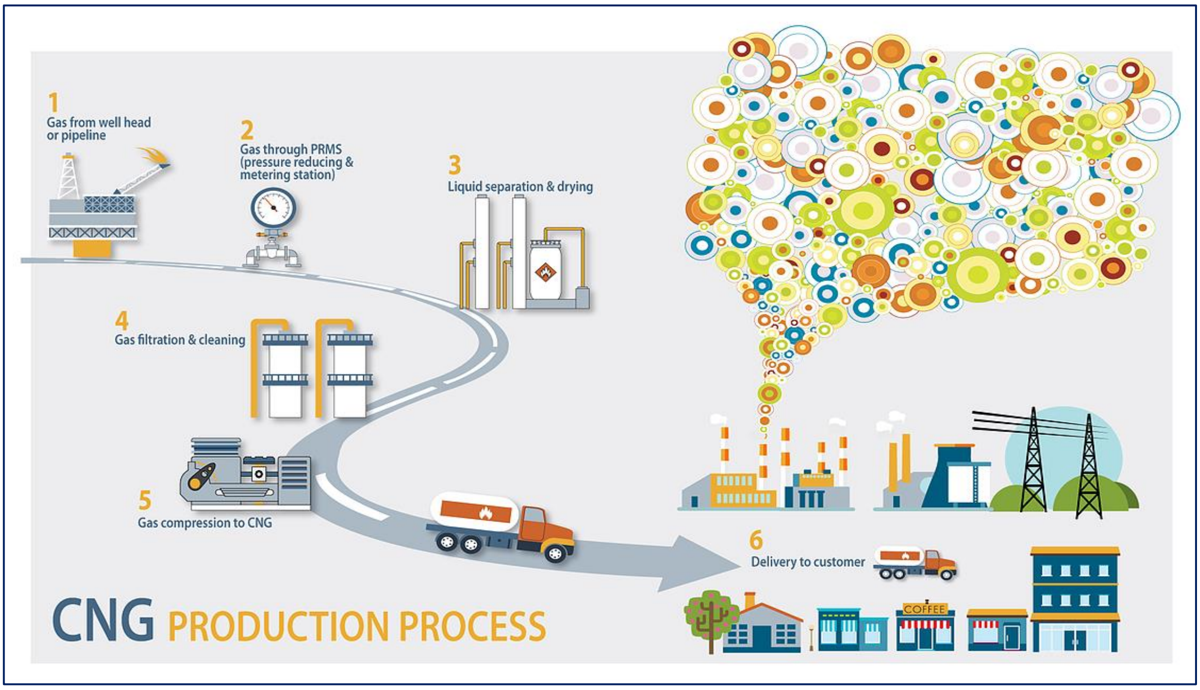


Figure 2.9 CNG value chain.

Source: Power Gas Africa, 2018.

CNG production is more straightforward compared to liquified natural gas. Figure 2.9 graphically shows the CNG value chain. The production consists of drying, cleaning, compression, and delivering. Design parameter of compression and cooling vary for different CNG process. The modified design parameter are influenced by the quality of harvested gas and different reservoir pressures. For example, the calorific value of natural gas in Mozambique has a range of 38- 40 MJ/m^3 while in some European countries ranges from 41 to $42MJ/m^3$ (Sasol, 2017). The process commences with a supply of natural gas from an existing network with a supply pressure of 6 bar. The pressure is increased by compressing it through a multistage compression to a pressure of 250 bar (CNG Holding, 2017). The compressed natural gas is stored in modules called MATTS and transported on trucks to the filling stations or the customer production receiving point. The pressure is reduced to the customer's requirements, and the heating system is supplied to counter the temperature drop due to Joule Thomason effect (SLG,2018).

2.7.3.5 Product delivery

Transportation accounts for a considerable weighing in the total delivered price in the supply of CNG which consists of the containment system, safety and gas control systems. CNG can be transported by the following:

- Cosell CNG carrier
- Volume-optimized transport and storage (VOTRANS)
- Gas transport modules

CNG Holding and Spring Lights Compressed Natural Gas currently use transportation modules to contain and deliver gas to their customers (SLG, 2018). A single trailer can carry up to $6,578m^3$. The central idea behind the gas transport modules is the "flexibility" that maximises the transported gas usage. The loading and unloading is a straightforward process which is executed by truck drivers.

2.7.3.6 Unloading

The CNG unloading process is carried out using flexible steel pipes that are connected to a pressure reducing station. Reduced pressure gas flow is connected to a customer receiving pipe which takes gas to the operating facility. The pressure reducing stations includes a dock with high-pressure pipeline connections and a vaporizer to allow gaseous natural gas into the receiving pipe to the combustion facility (SLG. 2018). The low-pressure trip control system in a module container communicates with the panel controller at the filling station allowing an efficient supply of gas and to ensure that the customer has gas all the times. When the module pressure drops below a low-level trip point, the panel gives an alarm indicating that the module is close to depletion and the following module is released in line with the survival time of remaining gas at the customer site (SLG, 2018).

2.7.3.8 Operation criteria

Transporting CNG could mean a reduced supply due to a temperature increase in the modules. Volume reduction for CNG is influenced by compression conditions which have a range of 7,142 to 8,571m³. For the traders to consider a CNG supply, a potential customer has to procure a minimum of 65,789 m³ monthly within the radius of 150km or less (SLG, 2018). As the distance from the source to the market increases above 150 km, LNG becomes a more favorable option due to its ability to store more liquid gas. The CNG supply is flexible and can accommodate significant gas requirements and can be done by adding more volume modules (CNG Holdings, 2017).

2.7.3.9 Key consideration

CNG technology has been used for decades around the globe but relatively new to the South African gas market. Moreover, as a gas monetization option, infrastructure is lacking and requires massive investment. Companies playing in the CNG space often face difficulties in putting a plan together that is financially attractive and reasonably scaled to justify the cost of a CNG plant. Wholesale gas price is a critical variable that affects the downstream cost of the CNG and transportation equipment (SLG, 2018). As mentioned, massive investment is required to enable CNG as a supply option in the gas market, and the ability to accurately forecast required gas molecule becomes a point of departure for investors and the suppliers of gas. Take or pay contract conditions are common in the industry and correctly sizing the filling equipment at the CNG station and supplying the required gas is critical to keep the business afloat.

2.8 Literature Review on time series analysis

The research objectives highlighted in Chapter one will shape the literature review of the study. Time series approach refers to a process of replicating the pattern of the preceding observations at equal time interval over a period and apply the established pattern to forecast the future movements of the variable of interest (Zhang, 2016). Univariate variable describes a data series which comprise of one observed variable with a single attribute or characteristics. For example, the age of all employees in the company can be visualized using graphs and images (Canova, Cortinovis, and Ambrogi, 2017). The univariate variable can be divided into a discrete or continuous data series. Discrete variable arises from counts, such as the number of student at a university while continuous arises from measurement such as the height of a building (Canova et al., 2017).

2.8.1 Univariate time series

The actual gas consumption is a measure of gas used by a customer and measured at an equal time interval (Akpinar and Yumusak, 2016b). It is a single variable with its attribute and characteristics which can be defined as a time series data. Time series data by nature can replicate the pattern of the preceding observation of equal time interval over a period and apply the established pattern to forecast future observations (Zhang, 2016). It is on this basis that univariate actual gas time series is chosen as a method to forecast future gas usage for Spring Light Gas.

According to Gujarati and Porter (2009), the Box–Jenkins (BJ) methodology technically know as the Autoregressive Integrated Moving Average (ARIMA) time series models allow the variable of interest (*Yt)* to be explained by its preceding, or lagged, values of *Y* itself and stochastic error terms.

Erdogdu (2010), used the ARIMA model to forecast short and long run gas demand in Turkey. Monthly actual gas consumption in million cubic m^3 from 1987 to 2007, which translate to 252 observation was collected from a gas company in Turkey. The research results reveal that gas demand in Turkey is growing at an average of four-percent year on year to 2030 which translate into an increase of 146%. In comparing the shot run to 2020, the official projection varied by 1,3% which is perfectly acceptable. Further to the result, Erdogdu (2010), compared the ARIMA result to the official projection done in 2002 which show that the gas demand was overestimated by 34%. He also noted that the ARIMA model is more of art than science and variation should be expected. The ARIMA model is based on assumptions and the link that exists between the validity of underlying assumptions made and the accuracy of the model. The ARIMA model generally assumes that: the already established pattern in gas consumption will continue in the future.

On the other hand, Wadud, Dey, Kabir, and Khan (2011) argue that using univariate ARIMA method alone to forecast gas may pose challenges with accuracy. Wadud et al. (2011), further, argue that gas consumption is influenced by the related economic variables hence challenging the use of actual gas consumption alone.

Akpinar and Yumusak (2016b), used ARIMA, Holt-Winter exponential smoothing, and scenario decomposition to forecast natural gas usage for the province of Sakarya in Turkey. Actual gas consumption of residential and commercial customer was collected and used to determine the existing pattern which contained a seasonality effect. Monthly data series from 2011 to 2014 was split into two, from 2011 to 2013, which was used to determine the existing pattern and the parametre of the model. The 2011 period was forecast to test the performance of the model. The result show that a complex computation of the model by adding an exogenous variable improves the accuracy of the model. On the other hand, univariate actual consumption generated a satisfactory result and can be recommended for use.

Cabral, Legey, and Freita Cabarl (2017) proposed a hybrid ARIMA model to forecast regional electricity consumption in Brazil with spatial dependency and showing the different patterns among regions. The identification of the spatial patterns suggests that the space dimension must be included in the specific model to ensure consistent, efficient and unbiased estimates. The researcher used the ARIMAS model to forecast electricity, and the results show low error Mean Absolute percentage error of the forecast as compared to a traditional ARIMA model. The result confirms the accuracy of the forecasting spatiotemporal model in forecasting electricity consumption.

Elamin and Fukushige (2018) used the ARIMA model infused with the exogenous variable and the interaction variable to forecast the hourly electricity demand. The main effects of the ARIMA model were determined using the traditional process, and the interaction of the weather variable and seasonality pattern analysed and added to the model. The proposed seasonal ARIMA model with interaction performed better compared to the seasonal ARIMA by a low mean absolute percentage error and the mean absolute error. Thus, including interaction effects can have potential improvements.

Exports and imports sector plays a crucial role in the economic development in Serbia hence the need for accurate forecasting. Mladenovic, Lepojevic, and Jankovoic-Malic (2016) analysed the export trends and variations using data covering the period from 2004 to 2014. The variation was used to forecast the trends from January to December 2015 using the winter holts and ARIMA to analyze the data. Both methods performed well, and both predicted an increase during the summer season and a noticeable drop in October 2015 which is the beginning of the winter period.

2.8.2 Bivariate time series

Bivariate variable means two different variables which are analysed to establish whether there is a relationship between the variables. The two variable are commonly visualised by plotting one variable against another in a scatter plot. If the data fits a curve or a line, then it means there is a relationship or correlation between the variables (Canova et al., 2017).

The study will employ the actual gas consumption and the purchasing management index (PMI) causal relationship to forecast gas usage for SLG. Actual gas consumption is a single variable with its attribute and characteristics which can be defined as a time series data, and PMI has the same features. As an employee of the case study company, the researcher has noted that the company forecasts gas consumption by observing the PMI and it is on this basis that the bivariate method of forecasting is chosen as well as the PMI variable.

According to Gujarati and Porter (2009), Vector Autoregressive (VAR) bivariate method explains the two variables in the context of endogenous setting meaning that its previous or lagged observation explains each variable. VAR measures the causal relationship in the short run, while Vector Error Correction Model (VECM) measure both short and long run (Gujarati and Porter, 2009). Many scholars have used VECM to forecast economic variables such as GDP, inflation, and the money supply among others (Gupta, Kanda, Modise, and Paccagnini, 2015). The application of VECM in this study is relevant because PMI is an economic variable.

Fang and Wolski (2016) evaluated the connecting association between energy consumption, human capital and economic growth in China using data from 1965 to 2014 applying a Granger causality test. They concluded that the causal relationship exists among aggregate energy use and coal, natural gas, and hydroelectricity consumption. However, the unidirectional causality exists from GDP to fuel oil, meaning that future aggregate energy use can be explained by gas usage.

On the other hand, Fang and Chang (2016) examined the causal relationship between energy consumption and economic growth in the Asia Pacific region. The researchers included the control of human capital and the cross country dependence and concluded that the economic growth drives energy consumption in the region, however, vastly different across countries. The research further revealed that physical capital investment and energy are significant contributors to economic growth. The researchers concluded that future energy consumption should be forecasted in conjunction with economic growth and physical capital investment.

Darrah, Vengosh, Jackson, Warner, and Poreda (2014) examined a casual relationship between energy consumption and aggregate GDP at sectoral level using data from 1980 to 2014 in Bangladesh. The research revealed that unidirectional relationship causality from GDP to energy consumption at the sectoral level. However, Vector Error Correction Model (VECM) revealed no causal relationship between energy consumption and GDP at the sectoral and micro level in the short run concluding that GDP is not driven by energy consumption.

Amin and Murshed (2017) examined the relationship between electricity consumption, economic growth and the foreign aid inflow in Bangladesh using data from 1980 – 2013. The researchers used Granger causality to tests for causality between the variables and revealed unidirectional causality relationship running from electricity usage to economic growth in the long run, while the VECM also confirmed that the unidirectional causality is running from electricity consumption to economic growth in the short run. The result concludes that energy consumption in the short run uses its preceding observation to predict the future quantity and in the medium to the long run, the forecast is influenced by GDP.

2.8.3 Smaller Error method

The models mentioned above will be used to forecast gas for SLG, and a need arises to identify the model with a smaller error. Oguz (2013) analysed sectoral energy demand of industry, agriculture, residential energy consumption, transportation, and services sectors in Turkey for the period of 2006 to 2023 using ARIMA, Vector Autoregressive, and Decomposition statistical methods. In his study, he outlines that population has been increasing with 1.3 percent and the GDP at an average of five percent in the past decade meaning that energy demand has been growing at a similar rate. Oguz (2013) used sectoral energy consumption from 1970 to 2006 which was used to estimate and determine the ARIMA parameter and historical pattern.

The VECM method requires a minimum of two variables and changes in GDP data series was added as the second variable (Oguz, 2013). Similar data range was collected to determine the parameters of the method. The outcome of the results suggests that ARIMA performed well when energy, GDP and transport show congruence and outperformed the VECM. On the other hand, VECM performed well compared to ARIMA when GDP, actual energy consumption and agriculture are modeled together. Overall, VECM outperformed ARIMA in other sectors.

Dissanayake and Perera (2015) used ARIMA model to forecast domestic electricity demand and evaluate the dynamic interaction between domestic electricity consumption, GDP, unit electricity cost, average electricity, and the average temperature in Sri Lanka based on a Vector Error Correction model (VECM). The Sri Lanka economy has been on an upward trajectory in the past 15 years, and a need arose to evaluate and forecast electricity consumption. The actual quarterly data from 1997 to 2013 for all the variables of interest was collect to first, determine the ARIMA parametre and to establish the existing pattern.

Secondly, the VECM model was estimated using all the mentioned variables to evaluate the relationship between all the variables. The study revealed that the ARIMA model produced five percent less mean average percentage error (MAPE), while the GDP absorbed the domestic electricity consumption shocks more compared to other variables and produced an error of less than five-percent meaning that the GDP movements can explain the consumption of electricity (Dissanayake and Perera, 2015). In this case, both models produced satisfactory results.

Kinene (2016) compared the ARIMA and seasonal ARIMA to VECM in forecasting Uganda 's inflation. Actual monthly data from 1998 to 2015 was collected for all variables of interest and used inflation as a univariate variable to estimate the ARIMA model. A multivariate approach, the VECM model, employed inflation, exchange rate, and world coffee to determine the parameter of the model. The result show that the ARIMA and season ARIMA produced the lower Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) compared to the VECM for the out and in-sample performance measure. Equally, seasonal ARIMA outperformed the ARIMA and confirming that Uganda's inflation has an element of seasonality (Kinene, 2016).

2.8.4 Forecasting future gas usage for SLG

The literature presented two models that could be used to forecast future gas usage for SLG. Based on the presented literature, both models are capable of forecasting future gas usage in a short and long run. Akpinar and Yumusak (2016b) successfully forecasted natural gas usage for residential and light commercial customers with minimum error using the ARIMA model in Turkey. Erdogdu (2010) forecasted the national natural gas usage with a minimum error when compared to the forecasted consumption done in 2002 in Turkey. The outcome of the results showed that the forecast was overstated by 34%, and he further forecasted gas usage to 2030 which revealed an error of 1,3% compared to the official forecast.

Amin and Murshed (2017) examined the relationship between electricity consumption, economic growth and the foreign aid inflow in Bangladesh using data from 1980 – 2013. The researchers concluded that energy consumption in the short run uses its previous observation to predict the future quantity and in the medium, to the long run, the forecast is influenced by GDP. Fang and Wolski (2016) examined the causal relationship among energy consumption, human capital and economic growth in China using data from 1965 to 2014; the researchers concluded that the unidirectional causality exists from GDP to fuel oil, meaning that future aggregate energy use can be explained by gas usage.

2.9 Conclusion

The Literature Review highlights the various limitations of the gas industry regarding policy, new gas supply challenges and required infrastructure. South Africa has options that could be explored to fully develop the gas industry, ranging from importing gas using the pipeline, fracking, liquefied natural gas, and compressed natural gas. A fully functional policy enables the required activities in the development of the industry. Fracking the Karoo has been termed a game changer for the South African energy landscape, the advantages and disadvantages that come with it have been documented. The literature further reveals that a massive amount of investment is required in developing sources of gas and related infrastructure. A significant gas off-taker would positively steer the business case in securing financing over a period.

The common denominator in this transaction is possible future gas consumption; thus the requirement of the forecasting method that would reasonably forecast future gas usage. The literature review identified Autoregressive Integrated Moving Average (ARIMA) and the Vector Error Correction Model (VECM) to forecast gas for SLG. The method with the lowest error would be used to forecast future gas usage for the company. The identified model would be recommended for the whole region. The next chapter discusses the theoretical estimation of the chosen models, data collection, sampling and the validity of the data.

CHAPTER THREE

Research Methodology

3.1 Introduction

The literature review in Chapter two identified two forecasting methods that could be used to forecast gas for Spring Lights Gas, namely, Autoregressive Integrated Moving Average (ARIMA) and Vector Error Correction Model (VECM). The traders of gas could apply the proposed methods of forecasting, and for quantifying the required gas infrastructure. Chapter three, presents model's methodology, the aim of the study, research paradigm, sampling, study setting and the in-depth discussion of Vector Error Correction Model (VECM) and Autoregressive Integrated Moving Average (ARIMA). The model's formulae, estimation, and application of both models. The chapter closes with a summary.

A statistical model which forms part of a time series quantitative research has been chosen to be the blueprint for study. The study emphasises the scientific empiricist method to generate the outcome that is purely data and uninfluenced by the observer's interpretation. The time series observations used in the study are pure data and suitable to answer the research questions outlined in Chapter two. The outcome allows inference to made about the population of the study. The study takes a view to answer the research questions in Chapter one.

3.2 Aim of the Study

The bottom-up forecasting of gas quantity involves a customer, trader, producer and gas transportation companies. It is good practice for traders to enter into a back to back take or pay agreement with the producers and the customer. The ability to forecast required gas with minimum error could assist the customer, trader and producers to optimise their input cost which will positively influence the bottom line (Akpinar and Yumusak, 2016a). The study aims to identify a suitable statistical model that could be used to forecast gas for Spring Light Gas with a minimum error which would help avoid the take or pay penalties.

3.3 Research paradigm

Sauders, Lewis, and Thornhill (2016) relate the positivism paradigm to a natural investigation which entails working with the observed social reality which could be generalised to the population. It emphasises the scientific empiricist method to generate the outcome that is purely data and uninfluenced by the observer's interpretation (Saunders et al., 2016). A positivist approach is commonly conducted through quantitative research approach and is suitable where a researcher purposes to maintaining an objective stance (Sekaran and Bougie, 2016). Creswell (2014) argues that quantitative research should be used if the researcher aims to discover factors that influence the outcome or to identify the best predictor of the outcome. Positivist paradigm is suitable for this study as it seeks to determine the best scientific method of quantitative forecast which could be generalised to the region.

3.4 Research Approach

The study will employ the quantitative forecasting research approach to answer the research questions. The primary objective of the study is to forecast future gas usage for Spring Lights Gas (SLG). The study will use the actual historical gas consumption data series of SLG customers situated in the southern region of KwaZulu Natal (Mobeni, Jacobs, Prospecton, and uMbongintwini) in the manufacturing sector.

3.5 Research methodology

Time series quantitative research has been chosen as the blueprint for the study. The research questions outlined in Chapter two will shape the research methodology. The following subsections discuss the estimation of the univariate and bivariate forecasting methods, identification, and using a method with small error to forecast future gas usage for SLG.

3.5.1 Methodology

The time series regression methodology is the foundation for modeling and forecasting the quantitative variables of interest because of the inherent unbiased forecasting. Thus, everything required to continue the historical pattern is included in the variable. Here, attention is given to the recent historical data. Two popular time series regression methods, ARIMA Model and Vector Error Correction Model were used to forecast gas consumption for SLG.

3.5.1.1 Stationary data series

Stationary data means that the observed variables in the analysis have a constant mean, and variance, and covariance over a specified time which implies that the metrics mentioned above are equal, independent of any representative period. A stationary data series is essential to model or predict future outcome. However, the opposite of stationary data series results in unpredictable modeling or inaccurate future prediction. Using non-stationary data can be spurious and give the false relationship between variables where it does not exists. The Augmented Dickey-Fuller test is used to test for a stationary data series (Dickey and Fuller, 1979).

3.5.1.2 Univariate forecasting method

The ARIMA model will be used to estimate the univariate forecasting method as outlined in Chapter 2. The following subsection will discuss the origins of the model, component, estimation, diagnostics, and the final forecasting model.

3.5.1.2.1 Autoregressive Integrated Moving Average (ARIMA)

Box and Jenkins (1976) popularised the Autoregressive Integrated Moving Average (ARIMA) which is a combination of autoregression (AR), integration (I) and moving average (MA). The ARIMA model is a powerful univariate statistical model used to forecast different kinds of time series in various sectors, and it models, and analyses the autocorrelation of the variable within its stationary time series. The model is represented by the values of (p, d, q), where p, d, q ≥ 1 where p and q denote the order of autoregression and moving average respectively. The value of 'd' refers to the differencing to the order of 'd', which indicates the size of differencing to make the data series stationary (Butt, 2014).

- **Auto-Regressive (AR)**

The AR process is given by equation (1)

$$(Y_t - \delta) = \alpha_1 (Y_{t-1} - \delta) + \varepsilon_t \qquad (1)$$

Where δ is the mean of Y and ε is uncorrelated random error term with constant variance and zero mean. Equation (1) represents the first order Autoregressive or AR (1) stochastic process. The value of Y at time t is determined by its previous time period and the error term ε_t which means that the Y value is explained by its own mean deviation (Butt, 2014).

- **Moving Average (MA)**

The MA process is defined by equation (2)

$$Y_t = \mu + \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} \qquad (2)$$

Where μ is a constant and the ε_t is white noise error term. The *Y* value at time *t* is determine by a constant plus moving average of a present and past error term; thus the first order moving avarage (Gujarati and Porter, 2009).

- **Autoregressive and Moving Average**

As discussed above, the *Y* value could be generated from the AR and MA mechanism and therefore the Autoregressive Moving Average. *Y* value follows the ARMA (1,1) and is given by equation (3).

$$Yt = \theta + \alpha_1 \mathrm{Y}_{t-1} + \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} \qquad (3)$$

The ARMA process is defined by (p) Autoregressive and (q) Moving average values (Gujarati and Porter, 2009).

- **Autoregressive Integrated Moving Average (ARIMA)**

Generally, a stationary data series is preferred when a model is estimated. When a data series is non-stationary and non-seasonal, differentiating is used to make the data series stationary to proceed in estimating the model (Butt, 2014). If Y_t is changed to $\Delta^d Y_t$ by taking the order of 'd' to differentiate, then the ARMA (p,q) changes to ARIMA (p,d,q) model. This model was used for forecasting and therefore essential that the structure of the model is constant over time and given by the equation (4) (Butt, 2014).

$$\Delta^d Y_t = \theta + \alpha_1 \mathrm{Y}_{t-1} + \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} \qquad (4)$$

- **Box – Jenkins (BJ) Methodology**

The ARIMA model can be estimated by following the steps provided by the Box-Jenkins methodology.

- **Identification of the Order**

The ARIMA require the determination of the values of (p,d,q). Identification can be objective by following the Augmented Dickey-Fuller test for a unit root in the data series. The presence of a unit root means that the data series is not stationary, and the opposite refers to a stationary data series. The subjective approach involves the use of the Autocorrelation function (ACF) and Partial autocorrelation function (PACF) and resulting correlogram which shows a plot of ACF and PACF against the lag length. PACF is a conditional correlation which refers to the correlation between Y_x and $\mathrm{Y_{x-k}}$ after removing the effect of the Y_x values between the two mentioned values. ACF is the correlation between Y_t and $\mathrm{Y_{t-h}}$ where h is 1,2,3.. If the series is not stationary, then differentiating is applied to make the series stationary.

Table 3.1: ACF and PACF theoretical patterns.

Model	**Typical Pattern of ACF**	**Typical Pattern of PACF**
AR(p)	Logarithm decay	Significant spike at lags *p*
MA(q)	Significant spikes at lags *q*	Logarithm decay
ARMA	Logarithm decay	Logarithm decay

Source: Tabachnick and Fidell, 2007, p.45.

Table 3.1 above shows the patterns for the ACF and PACF for Autoregression (AR) and Moving average (MA). ACF predict AR model with an exponential decline and the PACF cuts off after few significant initial lags and reverts to zero. For the MA model, Table 3.1 predicts the opposite of the AR. It also predicts the patterns for the ARMA process where the ACF and PACF predict a similar pattern at a time. The significant number of lags of the ACF gives an estimated value of q and PACF predicts the value of p (Tabachnick and Fidell, 2007).

- **Model estimation**

After determining the estimated values of (p,d,q), the final model is estimated using different combinations of AR and MA. Eviews 10 statistical packages was used to conduct a trial and error to determine the final model. The most suitable model showing superior results is used as the final model (Butt, 2014).

The final model was chosen using the following criteria:

a) The p-value for MA and AR term have to be statically significant, i.e., smaller than α 5% at value.
b) Low parameterized model to archive parsimony (Box et al., 2015).
c) Adjusted R squared value, the higher the value, the better the model fit (Gujarati and Porter, 2009).
d) Then BIC and AIC value, the smaller, the better (Vrieze, 2012).

- **Diagnostics**

A series of random errors evaluate and indicate a good model and the applicability. The residual error forms a set of observation that is evaluated using ACF and PACF correlogram plots. For the chosen model, ACF and PACF correlogram should be flat, meaning that all the lags should be non-significant and must fall within the higher and the lower limit of the correlogram (Tabachnick and Fidell, 2007).

- **Forecasting**

The final model is chosen at 5% statistically significant level, low Akaike and Schwarz Information Criterion, higher adjusted R squared value.

3.5.1.3 Bivariate forecasting method

3.5.1.3.1Vector Error Correction Model (VECM)

The VECM model will be used to estimate the bivariate forecasting method as outlined in Chapter 2. The following subsection will discuss the stationary data series, regression, lag selection, test for cointegration, diagnostic check, model estimation and definition, and decomposition forecast.

- **Stationary test**

The framework employed to examine the dynamic relationship between gas consumption and manufacturing output volume index involves a unit root, cointegration and causality test. This study performs an Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979) to test for a stationary data series. The null hypothesis states a non-stationary data series, and the alternative hypothesis states a stationary data series.

Should the results show a non-stationary data series in both variables, successive differencing transforms the data series from non-stationary to stationary. In general, if the time series is differenced to the size ‘d’ times to make the data size ‘d,’ then the data series is stationary and integrated to the order of ‘d’ (Fang and Wolski, 2016).

- **Regression Lag selection**

The VECM model requires two variables to estimate the parameter of the model. Both variables interact in an endogenous setting and should have the same regression lag (Gupta, 2014). The

Akaike information criterion (AIC) and the Schwarz Information Criterion (BIC) are used to determine the lag size.

The objective of the two criterion is to impose penalties for adding an unnecessary number of regression lag thus enables the selection of a parsimonious model. When comparing multiple models, the low AIC and BIC model is preferred. The over-parameterized model can lead to multicollinearity and loss of a degree of freedom, leading to inaccurate estimates and erroneous out of sample forecasting (Gupta, 2014).

- **Cointegration**

Testing for co-integration between the two variables. The Johansen test is used to test co-integration among two or more variables within the VECM (Johansen and Juselius, 1990). The Johansen maximum likelihood test uses two test statistics to evaluate cointegration, namely: Trace and Maximum Eigenvalue statistic test. The test results show the extent of cointegration between the two non-stationary variables using the likelihood model. The presence of cointegrating equation confirms that there is an existing long-run relationship among the variables under study (Gujarati and Porter, 2009).

- **Model estimation**

The variables that confirm the presence of cointegration equation allows the estimation of the VECM which would be used to determine the long-run causality. The long-run causality between the variables is identified by the coefficient of the error correction term (ECT, $\emptyset$ in equation 5, 6 and 7). A significantly different coefficient from zero allows a conclusion that the causality runs from the independent to the dependent variables.
The primary objective of VECM is to study the dynamic interaction between gas consumption (LNCON) and the Purchasing management index (LNPMI). The establishment of co-integration allows the modeling of long-run causality by adjusting the equilibrium error and the short run using past changes. Many researchers have used the restricted AR model which generates VECM that uses cointegrated non-stationary data series (Gupta et al., 2014). VECM is suitable for short-run adjustment and can incorporate cointegration relationship to restrict the movement of endogenous variables to congregate to their long-run relationship. The error correction term, also cointegration, gradually correct through a series of short-run adjustment.

To demonstrate: assume that Yt includes n time series integrated to the order of one.

Using the VAR model in equation (1) converts into a VECM model as follows:

$$\Delta \mathrm{LNCON_t} = \alpha_o + \sum_{i=1}^{K} \Theta_1 \ \Delta \mathrm{LNCON_{t-1}} + \sum_{i=1}^{K} \phi_1 \ \Delta \mathrm{LNPMI_{t-1}} + \emptyset \ \Delta ECT_{t-1} + \varepsilon_{1,t} \quad (5)$$

$$\Delta \mathrm{LNPMI_t} = \beta_o + \sum_{i=1}^{K} \Theta_2 \ \Delta \mathrm{LNCON_{t-1}} + \sum_{i=1}^{K} \phi_2 \ \Delta \mathrm{LNPMI_{t-1}} + \emptyset 2 \ \Delta ECT_{t-1} + \varepsilon_{2,t} \quad (6)$$

$$ECT_{t-1} = Y_{t-1} \text{ (LNCON)} - X_{t-1}\text{(LNPMI)} + \text{CONSTANT} \quad (7)$$

Where k is the lag length and $\emptyset$ is the coefficient error (speed of adjustment parameter with a negative sign).

- **Diagnostic test**

Diagnostic check is applied to verify the applicability of the model. The following subsections evaluate the model by discussing the autocorrelation, heteroscedasticity and the normality test.

- **Autocorrelation test**

The tests for Auto-correlation in the error of the VECM gives information about the autocorrelation of errors in the VECM. The model is accepted when there is no serial correlation, and the errors are purely random. The null hypothesis of the auto-correlation holds that residual errors are not serially correlated while the alternative hypothesis refers to the presence of autocorrelation (Hussain and Haque, 2017).

- **Normality test**

The normality test for error distribution gives information about the distribution of the error in VECM. The closer the distribution of the sample to the population the better the model and in cases of a large sample, normality test seems to lose its significance. The null hypothesis holds that the errors are normally distributed, and the alternative hypothesis; the errors are not normally distributed. The test is carried out by observing the Jarque Bera measure (Jarque and Bera, 1987).

- **Heteroscedasticity test**

Heteroscedasticity test is done to assess whether the variance of the error in the VECM is constant over time. The Breusch–Pagan Test and White noise test are carried out to evaluate the errors. The null hypothesis of Heteroscedasticity holds that residual error has equal variance while the alternative hypothesis refers that error variance is not constant over time (Hussain and Haque, 2017).

3.5.1.4 Identification of smaller error method

SLG as a trader of gas requires a reliable model to forecast gas usage. A reliable method could be identified by using the industry tradition of 80% take or pay contract condition. Should the actual gas usage of the trader or a customer be lower than the 80% of the initially forecasted gas quantity, penalties apply. As a consequence, the model that generates an error of 20% and below is recommended. For a given time series, there may be several competing models to choose from in finalising the forecast model. Statistical models can be compared against each other for the goodness of data series fit and forecasting using t-1($0 < l \leq t$) observation of the data series and use the fitted model to forecast against the sample data to evaluate the accuracy of the model using the following measuring matrix (Makananisa and Erero, 2018).

$$e_t(t - l + j) = y_{t-l+j} - \hat{y}_{t-l+j}, j = 1,2,...l \qquad (8)$$

Where $J = 1,2...,l$, $\hat{y}_{t-l+j}$ is the forecast for $\hat{y}_{t-l+j}$ using a competing model and compute

$$MAPE = \frac{1}{l}\sum_{j=1}^{1} \frac{e_t(t - l + j)}{y_{t-l+j}} * 100 \qquad (9)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{J-1}^{l}(\hat{y} - y_t)^2}{\text{T}}} \qquad (10)$$

$$\text{MSE} = \frac{1}{l}\sum_{j=1}^{1} e^2{}_t(t - l + j) * 100 \qquad (11)$$

Model performance will be tested by forecasting the in-sample for the period from 2016 to 2018.

3.5.1.5 Forecasting with the smaller error method

The identified smaller error method would be used to forecast future gas usage for SLG. The ability of both methods to forecast future gas usage has been demonstrated by many researchers (Akpinar and Yumusak, 2016a).

3.6 Research strategy

Research could be conducted using either qualitative or quantitative forecasting study or a combination of both. Both approaches were considered and assessed for suitability to ensure that the study produces good results. The qualitative research takes a form of interpretive philosophy (Saunders et al., 2016). Interpretive research in business means looking at an organisation from different groups of people. Organisation executives, managers, blue and white-collar employees see and experience work realities differently (Sanders, 2016:140). A qualitative study is characterised by open-ended questions and interview data to establish themes and patterns (Creswell, 2014).

In a quantitative comparison, the strategy seeks to study the relationship between variables using numerical measures and is commonly conducted using experiments and surveys methods (Sekaran and Bougie, 2016). The expected outcome of the strategy is to generalise the sample to the population for inferences to be made about some characteristics of the attitude or behavior of the population (Babbie, 1990). The objective of the study is to assess and identify the suitable forecast model for Spring Lights Gas which warrants the use of the quantitative strategy and can accommodate large numerical data. For the study, actual historical consumption data of the KZN south region with customers in the manufacturing sector was collected to develop statistical models.

3.7 Study setting

The customers that make up the total gas volume under study are Spring Lights Gas customers procuring a total of 11million m^3/month of natural gas and are situated in the south of KZN. The customers are in the manufacturing sector producing various products with different market share. The south region was chosen because of SLG 's clustered customers and granted access to the actual gas time series data. Due to a significant market share in the region, an inference can be made about the outcome and behavior of the forecast concerning the population of customers.

3.8 Population and Sample of the study

The KZN south region has a population of 33 gas customers procuring gas from all the gas traders in the KZN province. The target population consists of 28 customers that procure gas from Spring Lights Gas. The firm has granted access to a total volume of customers to model and develop a statistical model. Due to limited time for the project and possible repetition in the model process, the south region is chosen to assess and recommend a suitable statistical model.

The south region is selected for two reasons: firstly, because of its significant volume consumption level which could be a good representation of the southern region consumption and; secondly, SLG is investigating to expand its CNG plant in the region and wishes to predict the future gas usage accurately. The study aimed to identify a suitable model for future gas consumption requirements.

3.9 Sampling methods

Sampling design is a process of selecting a sample of a population of interest for purposes of making observation and inference about that population. It comprises the techniques and method used to select the sample. Population defines the total number of elements such as all companies, industries, and markets that are the subject of the study from which the sample is selected (Elfil and Negida, 2017). The sampling frame is the source list of the population elements under investigation (Walliman, 2011). Due to challenges in accessing the total southern region volume from all the traders except for SLG, the sample automatically chooses itself.

3.10 Research Sampling design

According to Gentles, Charles, Ploeg, and Mckibbon (2015) sampling design is a process to extract a representative sample from the population. The process further comprises the techniques used to select the sample. The term "population" denotes a total number of customers in the south region of KZN which their consumption is the subject of the study (Walliman, 2011) and the sample became the population of the study. In this study, the sample frame and the population are the same as all the customer's consumption was used. The two common techniques in sampling design are non-probability and probability sampling; both have its own merits and utility.

- **Non-probability**

Non-probability sampling is a technique where not all the elements in the population are given a fair chance to be selected (Elfil and Negida, 2017). The principal drawbacks are the lack of representativeness of the population and considered biased. It can be sub-classified into convenience, purposive, quota and snowball sampling and depending on the research objectives it could affect the quality of the results (Gentles et al., 2015).

- **Probability sampling**

Probability sampling is a technique where all the elements in the research project have an equal opportunity of inclusion in the sample (Elfil and Negida, 2017) and include simple, systematic, stratified and clustering sampling techniques.

The technique used in collecting data was a simple download of data in the archive of Spring Lights Gas. No human involvement when monthly consumption is collected for invoicing. It is for this reason that the quality of the sample and the data is high.

3.11 Data Collection

Data Collection plays a vital role in all kinds of research projects. Erroneously collected data could result in flawed finding and inaccurate inference about the population under study (Rimando, Brace, Namango-Funa, Parr, Sealy, and Davis, 2015). The quantitative data collection methods commonly use structured data collection instruments which involves a wide range of experiments into predetermined response categories and produced results are easy to summarize, compare, and generalise. The actual historical gas consumption data of the south region was collected from the Spring Lights Gas archives. The data collection is from July 2005 to June 2018 which coincides with the company financial year. The Purchasing Management Index (PMI) was collected from the Bureau of Economic Research. The index is compiled and released monthly. It is based on the monthly survey sent to senior management and executives at more than 400 companies in different sectors. The survey is based on new orders, productions, inventory levels, supplier delivery and employment (Mudgal, 2015).

3.12 Data analysis

The initial step in data analysis was to export the historical data from the company archives in share point system to excel format. The consumption data was checked against the customer archived invoices for verification and correctness. After the data confirmation, the EViews statistical package was used to assess and estimate the model.

The data from 2005 to 2018 was used as a base consumption to determine the ARIMA model, and the identified model was tested by forecasting in sample consumption from 2016 to 2018. The second analysis was to evaluate the causal relationship between the actual gas consumption and PMI for the same period using the Vector Error Correction Model (VECM).

3.13 Reliability and validity

Reliability and validity are crucial to any form of assessment that is dependable and accurate. It is the magnitude to which a model is correctly measured in a quantitative study; meaning, a survey was designed to measure depression but was instead used to measure anxiety, would not be considered valid (Heale and Twycross, 2015). According to Creswell (2014), validity can take three forms namely: content validity, and construct validity, and criterion validity.

- **Content validity**

Content validity is an idea which holds that a data used is relevant and answers the research question. The data used to estimate the model must constitute a content study area, actual gas consumption was used to develop a gas forecasting model which generated high quality of validity (Creswell, 2014).

- **Construct validity**

The construct is the concept that determines whether an inference can be made about the outcome of the test in the related study area (Heale and Twycross, 2015). The south region of KZN is clustered with manufacturing customers that procure gas from Spring Lights Gas (SLG, 2018). The collected actual consumption data and the outcome of the result is a representation of the region which enables a high quality of construct validity.

- **Criterion validity**

Criterion validity refers to testing the used measurement tool against another established measurement or against itself (Heale and Twycross, 2015). At SLG, historical data is used to forecast future consumption. The chosen model also used the historical actual gas consumption which allows the model to measure its outcome against the trader approach in forecasting future gas usage.

- **Reliability**

- Reliability is the measure of consistency. The data used to develop the ARIMA and VECM model should have approximately the same parametre each time the model is estimated and used (Heale and Twycross, 2015). Reliability can take three common forms (Heale and Twycross, 2015).

- The stability which refers to test and retest when the same data is used more than once in a similar setting. The similar outcome provides the quality of reliability.
- Equivalent refers to the allowable error due to different researchers who may undertake the same study and use the same data.
- Internal consistency measures the extent of consistency of responses of questions in the questionnaire or scale to each other.

Lee Cronbach developed the Cronbach alpha in 1951 which shows reliability as an index which is expressed within a range from 0 to 1, where 0 indicates no reliability, and one strongly demonstrates reliability (Tavakol and Dennick, 2011). The Cronbach measures the internal consistency which defines the degree to which all the subject in a test or a scale measure the same paradigm and hence it is associated to the inter-connectedness of the items within the test (Bonett and Wright, 2015). Due to the nature of the time series variable, the Cronbach alpha could not be measured because of a single variable that interacts and regresses its previous observations.

SLG granted access to the study data which was downloaded from the company archives. The consumption data required to invoice the customers is virtually downloaded with no human intervention and verified by paid invoices, the quality of reliability is therefore high (SLG, 2018).

3.14 Elimination of Bias

The presence of bias is experienced when there is a deviation from the legitimacy in data interpretation, collection, analysis, and publication which can lead to incorrect conclusions (Simundic, 2013). The researcher can intentionally, or unintentionally experience bias, either way, the consequence of bias can be harmful.

- **Researcher bias**

Due to its nature and mostly numeric form, quantitative research is susceptible to bias. Researcher bias can be international or unintentional (Simundic, 2013). For example, a researcher can unintentionally use monthly and weekly data and assume a consistent interval. In this study, monthly time series data is used to develop the forecasting models.

- **Selection sampling bias**

According to Simundic (2013), sample selection bias can occur when an incorrect sampling technique is used which could lead to the incorrect and erroneous representation of the population. In this study, the sample consists of all Spring Lights Gas customers in the southern region of KZN which are part of the manufacturing sector (SLG, 2018). For a given time series, using quantitative forecasting methods, and the fact that the data was sourced from Spring Lights Gas in which there is no human intervention, no bias was experienced.

3.15 Ethical Stance and Consideration

The first requirement for ensuring that the study was conducted ethically, was to obtain a gatekeeper letter from Spring Lights Gas group Chief Executive Officer, a trader of gas to the industry. Human Resource department was approached to gain approval for research and to use private company information; to ensure that the study is in line with the researcher's development plans, and to minimise interference with a daily work requirement. Upon receiving study approval from the company, the gatekeeper letter from the company and the ethical clearance application form was sent to the research office at the University of KwaZulu Natal.

The research office accepted the informed concent, storage of the data and research method used in this study and the ethical clearance number was assigned (HSS/0990/018M). Data collection commenced after the reference number was assigned. The Business Development Manager and the custodian of the customer data provided consent to access the historical data of the customers. Spring Lights gas as represented by the Busines Development manager was asked to provide informed consent before the participation which was duly granted.

As an employee of Spring Lights Gas, the verbal discussion was held to provide detailed information about the study. The Business manager was given assurance that the data used and the outcome of the study would be kept confidential as the willing participant in the study.

3.16 Summary

Chapter two presented the two methods that could be used to forecast gas usage for Spring Lights Gas namely, univariate ARIMA and bivariate VECM methods. Chapter 3 presented a theoretical estimation, model formulae, diagnostic checks, and the final model. The presented measure of performance will aid in identifying the model with the low error. The chosen method with the respective model demonstrated the ability to forecast into the future as used by other scholars (Makananisa and Erero, 2018). The chapter also presented the study approach, sampling, data collection, validity and reliability of the study.

The chosen research approach was considered appropriate for answering research questions involving time series survey and yielded satisfactory results (Akpinar and Yumusak, 2016a). Time series data collected from the trader was free from a human intervention which ensured the highest quality and a significantly reduced bias. The next chapter presents the results of the study.

CHAPTER FOUR

Results Presentation and Discussion

4.1 Introduction

Chapter 3 discussed the methodology of univariate and bivariate methods using the Autoregressive Integrated Moving Average (ARIMA) and Vector Error Correction Method (VECM) respectively, performance measurements in choosing the model with a low error, and the final forecasting model. The study aims to use univariate and bivariate methods to forecast gas consumption for Spring Lights Gas (SLG). Chapter four presents and discuss the various steps in estimating the ARIMA and VECM model. The Chapter commences with the estimation of the ARIMA model and the related process followed by the VECM. The model with a low Mean Root Square Error (MRSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE) is presented and chosen.

4.2 Univariate forecasting method

The univariate method uses ARIMA model, and the following subsections present and discuss the model results focusing on estimation, diagnostics, and the final forecasting model.

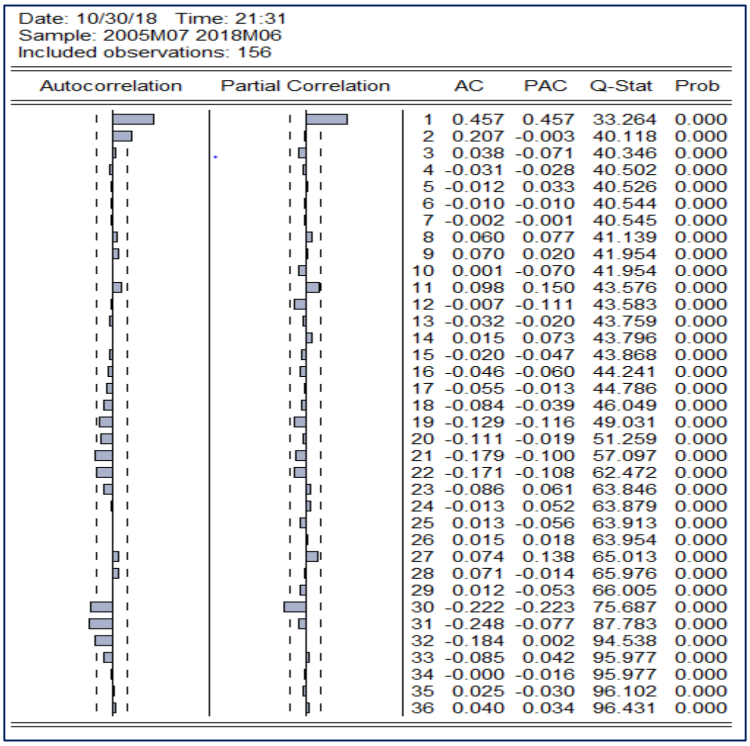
4.2.1 ARIMA model presentation and discussion

Final forecast models AR (1) MA (11) MA (30) at level values has been finalised. The model gave reasonable forecast values when compared to actuals, and the model is considered to be suitable for the corresponding time series. The Box Jenkin approach was followed in analysing and interpreting the actual data to obtain a suitable forecast model (ARIMA), model. Eviews 10 statistical software package was used to analyse the data. The detailed outcome of this analysis is given as follows:

4.2.1.1 Unit Root testing (Stationarity test)

The tradition in the estimation of the time series regression analysis is to test for a stationary data series and to establish the order of integration to avoid spurious regression. The actual gas consumption time series data has no unit root as given by Figure 4.1. Visual evaluation of the correlogram (Figure 4.1) has no recognisable pattern which indicates a stationary data series. The

Augmented Dickey-Fuller (ADF) Table 4.1, shows that p-value is lower than 5% which is an indication that the null hypothesis of Unit Root test, that is; the null hypothesis is rejected.



Date: 10/30/18 Time: 21:31
Sample: 2005M07 2018M06
Included observations: 156

Lag	AC	PAC	Q-Stat	Prob
1	0.457	0.457	33.264	0.000
2	0.207	-0.003	40.118	0.000
3	0.038	-0.071	40.346	0.000
4	-0.031	-0.028	40.502	0.000
5	-0.012	0.033	40.526	0.000
6	-0.010	-0.010	40.544	0.000
7	-0.002	-0.001	40.545	0.000
8	0.060	0.077	41.139	0.000
9	0.070	0.020	41.954	0.000
10	0.001	-0.070	41.954	0.000
11	0.098	0.150	43.576	0.000
12	-0.007	-0.111	43.583	0.000
13	-0.032	-0.020	43.759	0.000
14	0.015	0.073	43.796	0.000
15	-0.020	-0.047	43.868	0.000
16	-0.046	-0.060	44.241	0.000
17	-0.055	-0.013	44.786	0.000
18	-0.084	-0.039	46.049	0.000
19	-0.129	-0.116	49.031	0.000
20	-0.111	-0.019	51.259	0.000
21	-0.179	-0.100	57.097	0.000
22	-0.171	-0.108	62.472	0.000
23	-0.086	0.061	63.846	0.000
24	-0.013	0.052	63.879	0.000
25	0.013	-0.056	63.913	0.000
26	0.015	0.018	63.954	0.000
27	0.074	0.138	65.013	0.000
28	0.071	-0.014	65.976	0.000
29	0.012	-0.053	66.005	0.000
30	-0.222	-0.223	75.687	0.000
31	-0.248	-0.077	87.783	0.000
32	-0.184	0.002	94.538	0.000
33	-0.085	0.042	95.977	0.000
34	-0.000	-0.016	95.977	0.000
35	0.025	-0.030	96.102	0.000
36	0.040	0.034	96.431	0.000

Figure 4.1 Correlogram for the stationary test.

Table 4.1 Augmented Dickey-Fuller unit root test.

Null Hypothesis: LNCON has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on AIC, maxlag=13)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.585663	0.0000
Test critical values:	1% level	-3.472813	
	5% level	-2.880088	
	10% level	-2.576739	

*MacKinnon (1996) one-sided p-values.

4.2.1.2 Model Identification

The autocorrelation and partial autocorrelation correlogram spikes as shown in Figure 4.1 above were used as an initial estimate for the lag size for autoregressive (AR) denoted by (p), and the moving average (MA) denoted by (q). The spikes outside of the boundary indicate information that is not captured by the model. Table 4.2 shows the summary of selected estimates for AR and

MA from the correlogram spikes (selected from ACF and PACF lags) with the corresponding selected models.

Table 4.2 Correlogram model estimates.

Models identified from the Correlogram	**The finalised model selection criterion**
ARIMA (1 , 0 , 0)	ARIMA (1, 0 , 0)
ARIMA (1 , 0 , 1)	
ARIMA (1, 0 , 2)	ARIMA (1, 0, 11)
ARIMA (1 , 0 , 11)	
ARIMA (1, 1, 21)	
ARIMA(1 , 0 , 22)	AR(1) MA(11) MA(30)
AR(1) MA(11) MA(30)	

The final model is chosen based on the following criteria:

- A higher value of the adjusted R squared.
- The low value of the Akaike (AIC) and Schwarz (BIC) information criterion.
- Statistically significant model at 5%.

Table 4.3 below shows the estimated models and their respective selection criteria measurements. Using the least square method and subjecting the models to different processes of diagnostic checks, a model was finalised as AR (1) MA (11) MA (30).

Table 4.3 Model selection criteria.

Model	**Adjusted R^2**	**AIC**	**BIC**	**P values of the model α 5%**	**Selected model**
ARIMA(1 , 0 , 0)	20%	- 0.25	-0.18	α < P	x
ARIMA (1, 0, 11)	22%	- 0.23	-0.18	α < P	x
AR(1) MA(11) MA(30)	26%	- 0.29	- 0.19	α < P	√

4.2.1.3 Diagnostic Checks

The diagnostic check involved the following criteria:

➢ ARMA structure

The inevitability and stationary condition of AR and MA terms are analysed using the ARMA structure. The process is confirmed to be inevitable and stationary if the AR and MA roots for the model are inside a unit circle. In this case, AR and MA roots terms are inside and at the unit circle which confirms the inevitability and stationary state of the model as given in Figure 4.2 below.

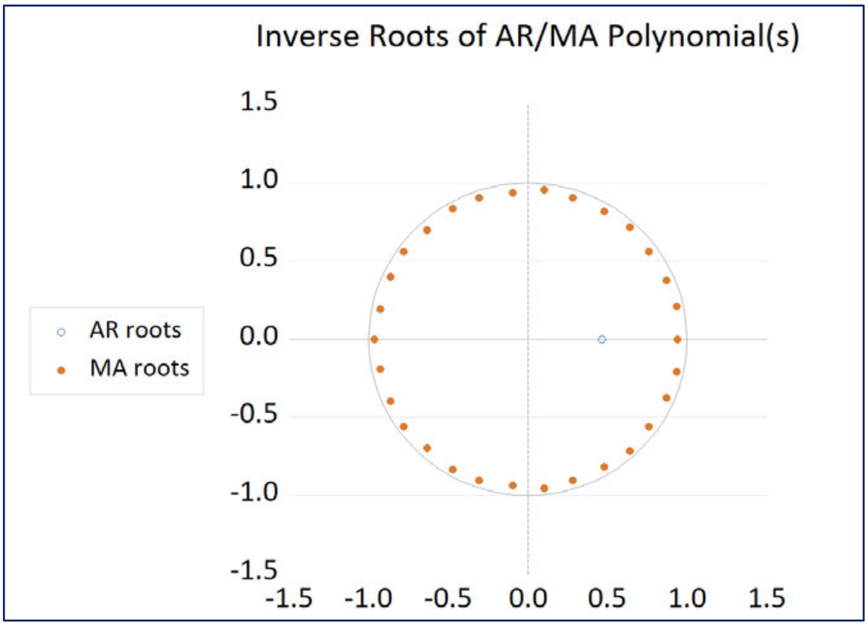
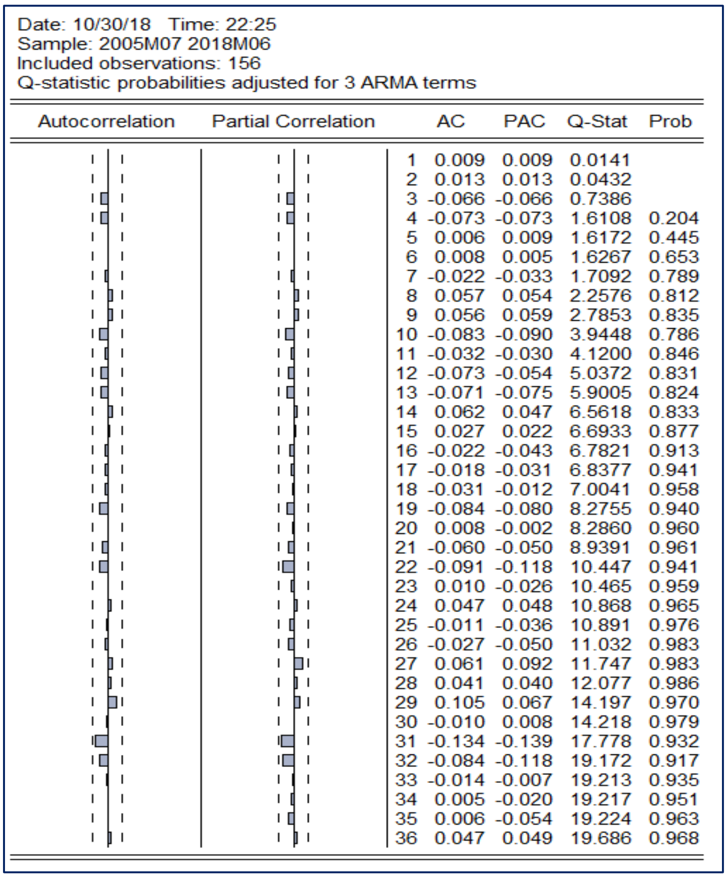


Figure 4.2: Inverse root of AR.

➢ Residual test

The correlogram of the residuals give the information about the chosen model whether all the lag information have been captured and if not, another model should be estimated. Should the residuals of the ACF and PACF lie within the boundary lines, then it means that the chosen model has captured all the lag information. Figure 4.3 shows a flat ACF and PACF which means the model has captured all the information and can be used to forecast future values.



Date: 10/30/18 Time: 22:25
Sample: 2005M07 2018M06
Included observations: 156
Q-statistic probabilities adjusted for 3 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.009	0.009	0.0141	
		2	0.013	0.013	0.0432	
		3	-0.066	-0.066	0.7386	
		4	-0.073	-0.073	1.6108	0.204
		5	0.006	0.009	1.6172	0.445
		6	0.008	0.005	1.6267	0.653
		7	-0.022	-0.033	1.7092	0.789
		8	0.057	0.054	2.2576	0.812
		9	0.056	0.059	2.7853	0.835
		10	-0.083	-0.090	3.9448	0.786
		11	-0.032	-0.030	4.1200	0.846
		12	-0.073	-0.054	5.0372	0.831
		13	-0.071	-0.075	5.9005	0.824
		14	0.062	0.047	6.5618	0.833
		15	0.027	0.022	6.6933	0.877
		16	-0.022	-0.043	6.7821	0.913
		17	-0.018	-0.031	6.8377	0.941
		18	-0.031	-0.012	7.0041	0.958
		19	-0.084	-0.080	8.2755	0.940
		20	0.008	-0.002	8.2860	0.960
		21	-0.060	-0.050	8.9391	0.961
		22	-0.091	-0.118	10.447	0.941
		23	0.010	-0.026	10.465	0.959
		24	0.047	0.048	10.868	0.965
		25	-0.011	-0.036	10.891	0.976
		26	-0.027	-0.050	11.032	0.983
		27	0.061	0.092	11.747	0.983
		28	0.041	0.040	12.077	0.986
		29	0.105	0.067	14.197	0.970
		30	-0.010	0.008	14.218	0.979
		31	-0.134	-0.139	17.778	0.932
		32	-0.084	-0.118	19.172	0.917
		33	-0.014	-0.007	19.213	0.935
		34	0.005	-0.020	19.217	0.951
		35	0.006	-0.054	19.224	0.963
		36	0.047	0.049	19.686	0.968

Figure 4.3: Residual correlogram test.

➢ **The Ljung Box test**

The Ljung Box test for autocorrelation, it is carried out by autocorrelation residuals correlogram which indicates whether the residuals variance is constant and at the same time, analysed the p-values of Q-statistics for autocorrelation. If the spikes are within the boundary lines, then the residual errors are constant and purely random. To further strengthen the argument, the p-values of Q-statistics must be greater than α 5%, which means that no autocorrelation exists, and the opposite will apply for residual errors that are outside the bounds and Q- statistics below α 5%. Figure 4.4 shows a flat autocorrelation correlogram and Q-statistics values which are greater than α 5%. Based on the criterion mentioned above, the model is fit for use and can be used to forecast future values. Model AR (1) MA (11) MA (30) passes the diagnostic checks and has been used to forecast future values.

Date: 10/30/18 Time: 22:41
Sample: 2005M07 2018M06
Included observations: 156

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.069	0.069	0.7576	0.384
		2	0.141	0.136	3.9188	0.141
		3	-0.020	-0.039	3.9850	0.263
		4	-0.075	-0.092	4.8881	0.299
		5	-0.086	-0.070	6.1010	0.297
		6	-0.137	-0.109	9.1906	0.163
		7	-0.117	-0.089	11.437	0.121
		8	-0.105	-0.075	13.279	0.103
		9	0.000	0.019	13.279	0.150
		10	-0.009	-0.016	13.294	0.208
		11	-0.037	-0.080	13.527	0.260
		12	0.020	-0.014	13.596	0.327
		13	0.003	-0.019	13.597	0.403
		14	-0.005	-0.043	13.602	0.480
		15	-0.027	-0.053	13.733	0.546
		16	-0.005	-0.014	13.737	0.618
		17	0.029	0.025	13.891	0.675
		18	-0.066	-0.095	14.671	0.684
		19	-0.035	-0.066	14.888	0.730
		20	0.089	0.108	16.310	0.697
		21	-0.032	-0.052	16.501	0.741
		22	0.105	0.055	18.524	0.675
		23	-0.001	-0.010	18.524	0.728
		24	-0.049	-0.086	18.978	0.753
		25	-0.054	-0.067	19.519	0.772
		26	0.010	0.031	19.539	0.813
		27	-0.014	0.010	19.579	0.848
		28	-0.031	-0.026	19.758	0.873
		29	0.105	0.093	21.909	0.824
		30	-0.105	-0.129	24.084	0.768
		31	-0.094	-0.153	25.815	0.730
		32	-0.005	0.015	25.819	0.772
		33	-0.028	0.005	25.977	0.803
		34	-0.067	-0.105	26.893	0.802
		35	0.034	0.037	27.130	0.827
		36	0.056	0.049	27.765	0.835

Figure 4.4 Correlogram for residual squared and Q- statistics.

4.2.1.4 Final forecast model

The model has been chosen, and it is given by equation (12).

LNCON = 12.83 + [AR (1) = 0.465, MA (11) = 0.197, MA (30) = - 0.224] (12)

4.3 Bivariate forecasting method

The bivariate method used the VECM model, and the following subsections present and discuss the model results focusing on the stationary test, cointegration, estimation, diagnostic check, and final forecasting model**.**

4.3.1 The Vector Error Correction Model

The tradition in the estimation of the time series regression analysis is to test for a stationary data series and to establish the order of integration to avoid spurious regression result. Augmented Dickey-Fuller (ADF) test presented the absence of the unit root in the data series. The null

Hypothesis states that unit root exists in the data series (non-stationary). ADF result presented in Table 4.4a-b shows that actual gas consumption and Purchasing Management Index (PMI) are stationary. The p-value is lower than 5% which is an indication that the null hypothesis of Unit Root test, i.e., the null hypothesis is rejected.

Table 4.4a-b Augmented Dickey-Fuller test for stationary LNPMI and LNCON.

Null Hypothesis: LNPMI has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=13)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.705410	0.0049
Test critical values:	1% level	-3.472813	
	5% level	-2.880088	
	10% level	-2.576739	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: LNCON has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=13)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.585663	0.0000
Test critical values:	1% level	-3.472813	
	5% level	-2.880088	
	10% level	-2.576739	

*MacKinnon (1996) one-sided p-values.

4.3.1.1 Lag selection

The VAR and VECM models employ the same lag length for all the endogenous variables in a given study. The lag selection was determined empirically by using the Akaike and Schwartz Information Criterion. The lower lag selection avoids the over-parameterised model which could lead to multicollinearity and on the other hand, the less parameterised model could lead to statistical errors (Gupta et al., 2014). Table 4.5 shows the criteria used to select the lag for the model. The model with the lowest lag size was selected.

Table 4.5 Lag selection.

VAR Lag Order Selection Criteria
Endogenous variables: LNCON LNPMI
Exogenous variables: C
Date: 11/01/18 Time: 15:33
Sample: 2005M07 2028M06
Included observations: 148

Lag	LogL	LR	FPE	AIC	SC	HQ
0	142.1165	NA	0.000516	-1.893466	-1.852964	-1.877010
1	249.9230	211.2424*	0.000127*	-3.296256*	-3.174748*	-3.246888*
2	251.5248	3.095502	0.000131	-3.263849	-3.061335	-3.181568
3	253.1848	3.162840	0.000135	-3.232227	-2.948707	-3.117033
4	254.9786	3.369492	0.000139	-3.202414	-2.837888	-3.054308
5	256.6010	3.003681	0.000144	-3.170284	-2.724753	-2.989266
6	258.5602	3.574100	0.000148	-3.142705	-2.616168	-2.928774
7	259.8357	2.292451	0.000154	-3.105887	-2.498344	-2.859044
8	264.6096	8.451135	0.000152	-3.116346	-2.427797	-2.836590

* indicates lag order selected by the criterion

4.3.1.2 Stability tests

The Vector Error Correction Model (VECM) is a tool used to evaluate the long-term causal relationship between two endogenous variables hence the importance of long-run stability test coefficient which is carried out using the Cumulative Sum (CUSUM), and Cumulative Sum Squared (CUSUMQ) and AR Inverse Root test. The result of the three analysis is presented in Figure 4.5 a–c. The graphical representation for the CUSUM and CUSUMQ plots falls within five percent critical bounds; meaning that the residual variance is stable over time, while the roots for the AR graph lie within the unit circle suggesting that the model is stable.

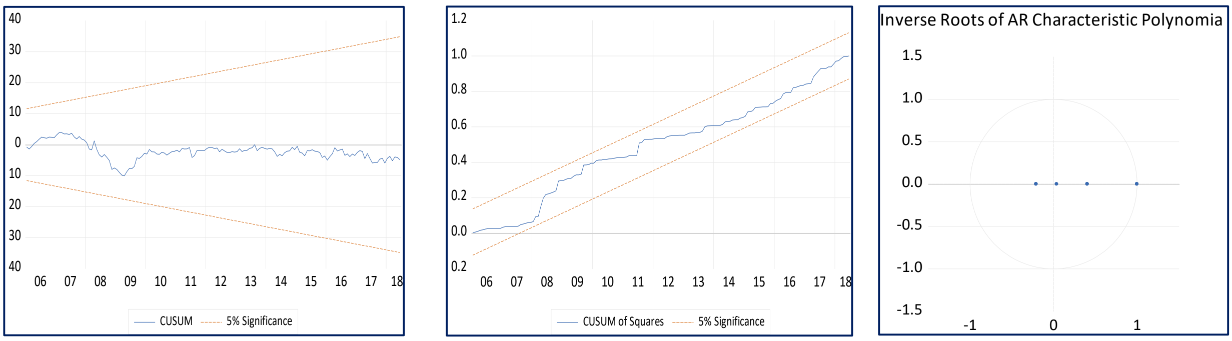


Figure 4.5 a-c CUSUM, CUSUMQ and AR Root test for stationarity.

4.3.1.3 Cointegration test

Table 4.6 shows the result of the Johansen–Juselius cointegration test. The Akaike Information Criterion (AIC) is employed to choose the lag length. The AIC is preferred because of its ability to choose a model that will minimise the Mean Squared Error (MSE) prediction in the absence of the actual model in a list of competing models (Vrieze, 2012). The results confirm the presence of a long-run relationship between the variables.

Table 4.6 Johansen–Juselius Cointegration test.

Date: 11/01/18 Time: 15:49
Sample (adjusted): 2005M09 2018M06
Included observations: 154 after adjustments
Trend assumption: Linear deterministic trend
Series: LNCON LNPMI
Lags interval (in first differences): 1 to 1

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.238137	51.31798	15.49471	0.0000
At most 1 *	0.059408	9.431842	3.841466	0.0021

Trace test indicates 2 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.238137	41.88614	14.26460	0.0000
At most 1 *	0.059408	9.431842	3.841466	0.0021

Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Based on the outcome of cointegration, the short and long-run causal relationship can be estimated using the Vector Error Correction.

4.3.1.4 The model diagnostic checks

The presence of cointegration allows the estimation of the VECM model. In verifying the results, a model diagnostic check is required. The following subsections discuss the diagnostics checks and the acceptance of the model.

- **Autocorrelation test**

The tests for error auto-correlation in the VECM gives information about serial auto-correlation of errors in the VECM. The model is accepted when there is no serial correlation, and the errors are purely random. Table 4.7 shows auto-correlation results. The p-value is higher than five percent which is an indication that the null hypothesis which states that residual errors are not serial correlated, that is, series error not correlated; we fail to reject null hypothesis.

Table 4.7 Serial autocorrelation test.

VEC Residual Serial Correlation LM Tests
Date: 11/01/18 Time: 17:08
Sample: 2005M07 2018M06
Included observations: 154

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	2.507732	4	0.6433	0.627471	(4, 294.0)	0.6433

- **Normality test**

The normality test for error distribution gives information about the distribution of the error in VECM. The closer the distribution of the modeled sample to the population the better the model. In cases of a large sample, normality test seems to lose its significance. Table 4.3 shows distribution results. The p-value is lower than five percent which is an indication that the null hypothesis which states that residuals are normally distributed, i.e., series is normally distributed; null hypothesis is rejected.

Table 4.8 Residual normality test.

VEC Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: Residuals are multivariate normal
Date: 11/01/18 Time: 17:10
Sample: 2005M07 2018M06
Included observations: 154

Component	Skewness	Chi-sq	df	Prob.*
1	-0.167261	0.718060	1	0.3968
2	-0.114892	0.338803	1	0.5605
Joint		1.056863	2	0.5895

Component	Kurtosis	Chi-sq	df	Prob.
1	3.404238	1.048535	1	0.3058
2	4.327469	11.30729	1	0.0008
Joint		12.35582	2	0.0021

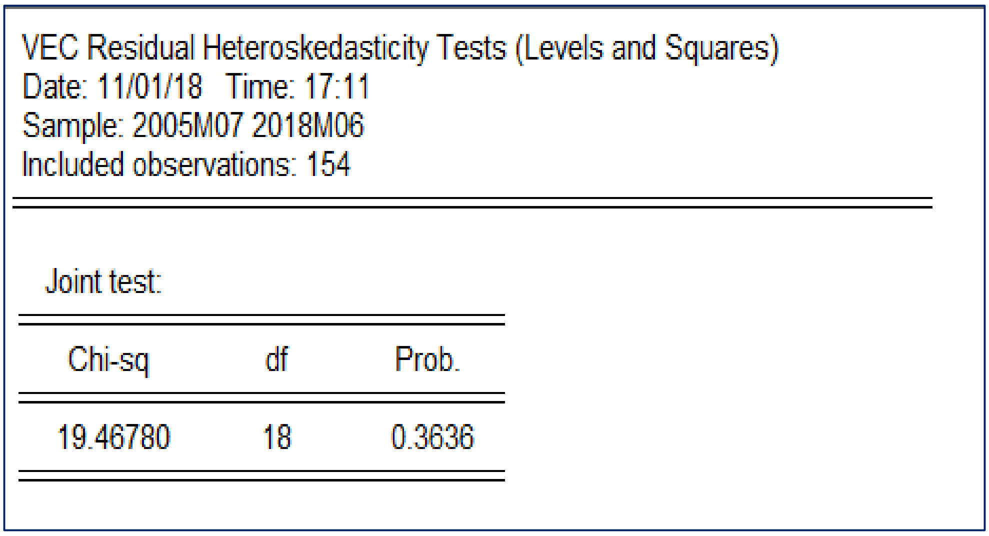
Component	Jarque-B...	df	Prob.
1	1.766595	2	0.4134
2	11.64609	2	0.0030
Joint	13.41269	4	0.0094

*Approximate p-values do not account for coefficient estimation

- **Heteroscedasticity test**

Heteroscedasticity test is done to evaluate whether the variance of the error in the VECM is constant over time. The Breusch–Pagan, and White noise test is carried out to evaluate the error. Table 4.9 shows the heteroscedasticity result of the series. The p-value is higher than five percent which is an indication that the null hypothesis which states that residuals are heteroscedastic, that is, the presence of heteroscedasticity; null hypothesis is rejected.

Table 4.9 Heteroscedasticity test for constant residual variance over time.


VEC Residual Heteroskedasticity Tests (Levels and Squares)
Date: 11/01/18 Time: 17:11
Sample: 2005M07 2018M06
Included observations: 154

Joint test:

Chi-sq	df	Prob.
19.46780	18	0.3636

The model diagnostic checks reveal that the model is acceptable even though the normality test rejected the null hypothesis. According to Ghasemi and Zahediasl (2012), violation of the normality test is not considered an issue when working with a large data series. The data series that has observation above 30 is considered significant, and in this case, 154 observations were considered.

4.3.1.5 The model causality test

The VECM allows the long-run relationship of the endogenous variables to return to their long-run equilibrium behavior while allowing a wide range of short-run dynamics.

The long and short-run causal relationship of the endogenous variables is shown in Table 4.10 below.

Table 4.10 Error correction term Estimates.

Cointegrating Eq:	CointEq1	
LNCON(-1)	1.000000	
LNPMI(-1)	-0.571664 (0.30685) [-1.86298]	
C	-10.60147	
Error Correction:	**D(LNCON)**	**D(LNPMI)**
CointEq1	-0.579240 (0.08461) [-6.84587]	0.008597 (0.02184) [0.39365]
D(LNCON(-1))	0.013092 (0.08022) [0.16319]	0.013705 (0.02071) [0.66189]
D(LNPMI(-1))	0.049633 (0.31072) [0.15973]	-0.193245 (0.08020) [-2.40961]
C	0.003494 (0.01690) [0.20671]	-0.001122 (0.00436) [-0.25716]

Table 4.10 is interpreted as follows: The correction error term of LNCON is negative meaning that it can capture the long-run equilibrium signifying convergence to a long run consumption time series. The previous period deviation of LNCON from a long run is corrected in the current period at an adjustment speed of 0.5792 back to equilibrium and statistically significant at one percent level. The speed of adjustment of 0.5792 is significant implying that consumption of gas closely follows the PMI movement hence one month lag; while LNCON is associated with 0.013 increase in itself. LNPMI is associated with LNCON at an increase of 0.049 in the short run, meaning that PMI does not influence the consumption of gas in the short run. LNPMI, on the other hand, has a positive error term signifying a short run relationship. In the short run, LNCON has associated with LNPMI with an increase of 0.0013 meaning that consumption of gas is unable to explain the PMI movements.

3.3.1.6 Final forecast model

The model has been chosen, and its given by the following equations (13, 14 and 15.

$$\Delta \text{LNCON}t = 0.0130\Delta\text{LNCON}_{t-1} + 0.0496\Delta\text{LNPMI}_{t-1} - 0.5792\Delta ECT_{t-1} + 0.0034_{t} \quad (13)$$

$$\Delta \text{LNPMI}t = 0.0137\Delta\text{LNCON}_{t-1} - 0.1932\Delta\text{LNPMI}_{t-1} + 0.0085\Delta ECT_{t-1} + 0.0011_{t} \quad (14)$$

$$ECT_{t-1} = 1\ \text{LNCON} - 0.571\text{PMI} - 10.6 \quad (15)$$

4.4 Identification of small error method

Identification of small error method could be done scientifically by comparing the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). This section focuses on both methods of forecasting and measures the performance using the in-sample data. The model fit on the actual data series enables the model to be used to forecast with some level of accuracy. The model was tested on the industry standard of 20% take or pay contract (a condition which triggers punitive measures should the user procure gas less than 80% of the initial requirement).

An error of 20% is therefore applicable to determining a good model. The models forecasted gas for the period of 24 months starting from 2016 to 2018. The following subsections present and discuss error measurements of both models to identify a low error method. Table 4.11 below shows the measurement of accuracy for both methods. The univariate ARIMA model shows low error in all measurements. Thus the ARIMA model is selected to forecast gas for Spring Lights Gas (SLG).

Table 4.11: Measure of accuracy.

Measurement of accuracy			
METHOD	**RMSE**	**MAE**	**MAPE**
ARIMA	**19%**	**11%**	**93%**
VECM	**20%**	**13%**	**102%**

4.5 Forecasting future gas consumption using the small error method

The univariate forecast model was identified and used to forecast future gas usage for SLG. The univariate ARIMA forecast model is a process where a variable of concern is a function of its past variables. Future consumption is predicted by observing the pattern of past consumption and replicating the pattern to predict future consumption. This section presents the result of future gas consumption.

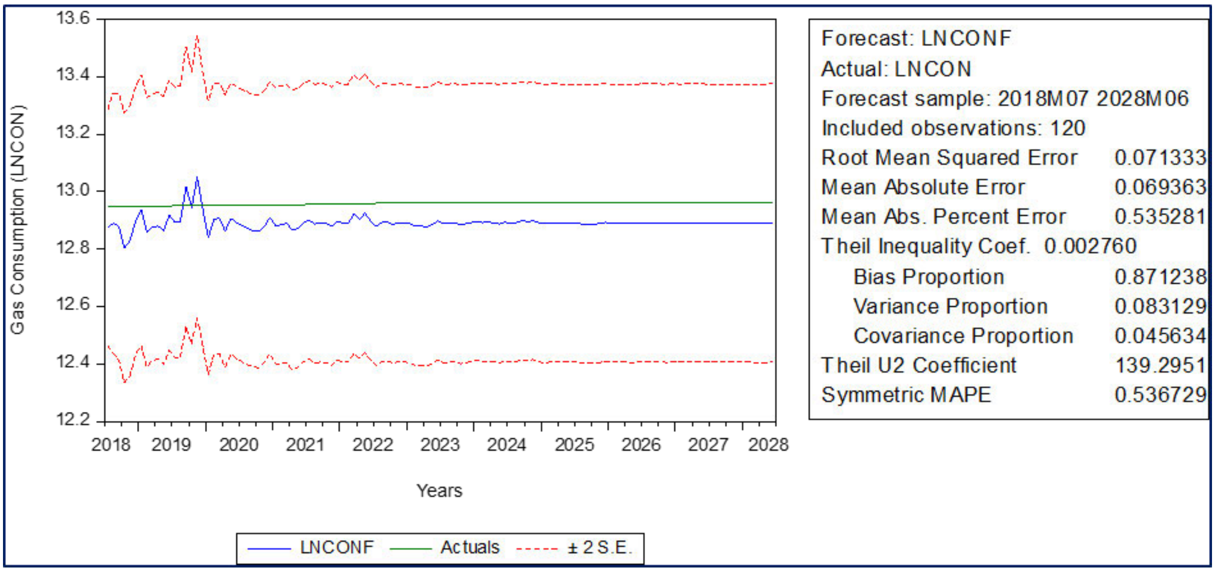


Figure 4.6 ARIMA model future gas consumption.

Figure 4.6 shows the forecasted gas consumption for the period 2018 to 2028. The y-axis shows SLG's ten-year official future gas forecast and gas consumption forecasted by the ARIMA model which fall within the standard error boundaries, while the x-axis shows yearly consumption. The model produced the RMSE of seven-percent when compared to the official forecasted gas consumption which confirms that the univariate method is suitable to forecast future gas consumption.

Table 4.12, monthly future gas usage converted to yearly consumption. The table below shows year on year gas consumption movement in percentages from 2018 to 2028. The model forecast that the growth of gas consumption in the southern region will grow at an average of one percent in the next decade and has possibly reached its maximum level.

Table 4.12: Year on Year consumption movement.

Year	Year on year movement
2019	4%
2020	-2%
2021	-0.09%
2022	3.8%
2023	-2%
2024	0.02%
2025	-0.24%
2026	1%
2027	-0.04%
2028	0.04
Ten-year average growth	1%

Future gas usage predicted by the model was compared to SLG's official future gas consumption. Based on the results above, SLG may have overstated its future gas sales plans. The company forecast its future gas consumption by observing the purchasing management index (PMI) and other econometric variables such as the growth of the South African economy and exchange rate amongst others. The state of factors such as gas availability, transmission and distribution infrastructure, suitable policy, regulation, and completion should be taken into account before gas consumption can be revised down in line with the outcome of the model.

4.6 Summary

Chapter four discussed the results of the univariate and the bivariate method of forecasting using the ARIMA and VECM respectively. Both models were estimated and the diagnostic check applied to validate the applicability of the methods. The ARIMA model generated smaller RMSE and MAE at 19% and 11% respectively, while the VECM generated 20% and 13 %. The ARIMA model was chosen because the error generated is below the industry standard of 20% which is derived from the 80% take or pay contract condition. The outcome of the model indicates that the univariate variable method is suitable to forecast longterm gas usage. Akpinar and Yumusak (2016b) forecasted the day ahead gas consumption in Turkey, and the model produced a Mean Absolute Percentage Error (MAPE) of 12.9% which further proves the applicability of the univariate ARIMA. The seven-percent RMSE further strengthened the argument that the model captured all the information required to reasonably forecast into the future. The ARIMA model was used to predict future gas usage for SLG which was compared to the official company forecast for the next decade and based on the results presented, SLG may wish to lower planned future gas sales. The next chapter is a conclusion to the study.

CHAPTER FIVE

Conclusion and Recommendations

5.1 Introduction

The study results were presented and discussed in Chapter four where the univariate ARIMA model was chosen to forecast future gas consumption. ARIMA was chosen because of the low error generated from the in-sample accuracy measurements. The study has numerously emphasized the importance of accurate forecasting of future gas consumption and the contractual penalties associated with 80% take or pay conditions. Chapter two highlighted industry limitation relating to policy, sources of gas, associated infrastructure, and forecasting techniques. The development of the industry regarding the mentioned limitation hinges entirely on the possible gas offtake which is crucial in making an investment decision, thus the importance of accurate forecasting. Chapter five concludes and briefly discusses the outcome of the study, summarises Chapter four results and the implication of the results, limitations, recommendation on how to solve the problem and the research gap that exist in the region.

5.2 Conclusion

The main reasons for Spring Lights Gas (SLG) customers to use gas is to fuel the furnaces and mostly boilers. Gas consumption increase is associated with infrastructure investment and market growth. For Spring Lights Gas to continue to be relevant in the industry, it needs to maximise its gas allocation by placing it where it will flow on a consistent basis, thereby generate revenue for the company and increase shareholder value. Future gas usage is crucial in budgeting for future revenue, thus, a need for a reasonable future gas consumption method.

The univariate and bivariate method using Autoregressive Integrated Moving Average (ARIMA) and Vector Error Correction Model (VECM) respectively, were employed to forecast gas for Spring Lights Gas. Monthly consumption data from July 2005 to June 2018 was collected from Spring Lights Gas to estimate and model the south of KZN (Mobeni, Jacobs, Prospecton, and Umbongintwini) region customers. The results show that the univariate ARIMA model better captured the historical movement of gas consumption compared to the bivariate VECM. This is shown by the measurements of accuracies such as RMSE, MAE, and MAPE on the forecasting of in-sample data from 2016 to June 2018 (Makananisa and Erero, 2018). Based on the outcome of the measurements of accuracy, the univariate ARIMA method of forecasting was chosen because it generated the lowest error of RMSE at 19%, MAE at 11% and MAPE at 99% compared to VECM at 20%, 13%, and 102% respectively. The chosen model predicted the future gas usage with seven percent error. It showed a drop in gas consumption compared to the company official future gas plans which suggest that SLG considering all applicable factors in the industry such as the growth of South African economy, new gas reserves, regulation, competition, suitable policy, and gas distribution and transmission infrastructure, may wish to revise down its future gas consumption plans.

5.3 The implication of this research

The model forecast one percent growth in the consumption for the next decade in the southern region. The organic growth of gas consumption in the region could be stimulated by customer investment in scaling up operations and procuring gas equipment provided that the current list of customers continue to use gas in the future.

A reasonably low gas price could also stimulate the usage of gas, as an employee of the case study company, the researcher has noted that some customers discontinue or reduce gas usage due to price concerns. The new gas connections could also grow the region which is achievable by improving the current technical offering, and the price to entice customers to choose gas over other forms of fuel. The low RMSE between the model's future gas consumption suggests that the company's future outlook for the southern region is flat and may wish to place gas in other growing regional economy or other provinces. NERSA has granted SLG a license to trade in Gauteng, Mpumalanga and Free State (SLG, 2018). The study affords SLG an opportunity to forecast gas for future use accurately and to develop a strategy that encourages the expansion of its footprint into other markets and to investigate the possibility of moving gas molecules from KwaZulu Natal to the other licensed areas.

The gas industry is capital intensive, and most projects realise the profit margins at an average of 5 to 10 years (SLG, 2018). The required investment to further develop the gas infrastructure and other related virtual gas sources will inevitably be hindered by the lack of growth in the gas economy of the region. Energy White Paper is seen as the overarching policy for the energy in the country, which anchorages the diversification of energy sources. The successful exploration of gas in the coastal line of the region will likely land in other provinces.

5.4 Limitation of the study

Although SLG has a considerable market share, the region has other traders that sell gas. The study could not access the actual gas consumption from other gas traders in the region. Due to SLG 's considerable market share in the region, the outcome of the study can be generalised to the whole region.

5.5 Recommendations to solve the research problem

The primary objective of the study was to forecast natural gas consumption for Spring Lights Gas from 2018 to 2028. The study compared two forecasting method namely, univariate ARIMA and bivariate VECM to predict future gas usage. The study result shows that the univariate method of forecasting is suitable for the company. The method allows the company to use its actual gas consumption to predict future usage. The ARIMA model requires historical actual gas data, and standard mathematical knowledge using spreadsheet software. The statistical computer package containing ARIMA will obtain the best results and requires a moderate understanding of the operation. The EViews business package statistical software could cost the company R27,500 (Quantec, 2018).

5.6 Recommendation of future studies

As a senior employee of the study case company, the researcher has noted that SLG lost gas volumes to customer closure, competitors and other sources of energy, hence the forecasted one percent growth in the next ten years. Natural gas is competing with other sources of energy such as heavy fuel oil, coal, and electricity. Coal is a preferred energy source for boilers, and the cost of generating one-ton steam is low compared to gas (SLG, 2018). Gas price is one of the reasons for volume loss. This is an area of concern that the study could not thoroughly examine and that could be the reason for low consumption growth. However, future studies could consider investigating the natural gas price elasticity (Arora, 2014). The outcome of the investigation could assist Spring Lights Gas and other traders to evaluate their pricing strategy to make better-informed decisions about the future of the company.

5.7 Summary

The study sought to present a statistical method fit for use by Spring Lights Gas to forecast future gas usage. The study concluded that the univariate ARIMA method of forecasting was suitable for the company because it generated a low error of RMSE, MAE, and MAPE compared to the bivariate VECM method. The study forecasted that gas consumption would grow by one percent for the next ten years. Natural gas is competing with other fuels in the region and price is among the reason for slow growth. A full investigation on the price elasticity will assist SLG to revise its pricing strategy to meet the customer expectation and thereby grow its portfolio of customers within and outside the region.

REFERENCES

ACKAH, I. 2014. Determinants of natural gas demand in Ghana. *OPEC Energy Review,* 38**,** 272-295.

AFRICA, P. G. 2018. Available: http://www.powergas.com/ [Accessed].

AKPINAR, M., ADAK, M. F. & YUMUSAK, N. 2017. Day-Ahead Natural Gas Demand Forecasting Using Optimized ABC-Based Neural Network with Sliding Window Technique: The Case Study of Regional Basis in Turkey. *Energies (19961073),* 10**,** 1-20.

AKPINAR, M. & YUMUSAK, N. 2016a. Year ahead demand forecast of city natural gas using seasonal time series methods. *Energies,* 9**,** 727.

AKPINAR, M. & YUMUSAK, N. 2016b. Year Ahead Demand Forecast of City Natural Gas Using Seasonal Time Series Methods. *Energies (19961073),* 9**,** 1-17.

ALTIERI, K. E. & STONE, A. 2016. Prospective air pollutant emissions inventory for the development and production of unconventional natural gas in the Karoo basin, South Africa. *Atmospheric Environment,* 129**,** 34-42.

AMIN, S. B. & MURSHED, M. 2017. AN EMPIRICAL ANALYSIS OF MULTIVARIATE CAUSALITY BETWEEN ELECTRICITY CONSUMPTION, ECONOMIC GROWTH AND FOREIGN AID: EVIDENCE FROM BANGLADESH. *Journal of Developing Areas,* 51**,** 369-380.

BABBIE, E. R. 1990. Survey research methods Wadsworth Pub. *Co Belmont, Calif,* 3.

BARNES, R. & BOSWORTH, R. 2015. LNG is linking regional natural gas markets: Evidence from the gravity model. *Energy Economics,* 47**,** 11-17.

BERNSTEIN, W. Z., HEDBERG JR, T. D., HELU, M. & FEENEY, A. B. 2018. Contextualising manufacturing data for lifecycle decision-making. *International journal of product lifecycle management,* 10**,** 326.

BO, Z. 2017. Forecasting the relation of supply and demand of natural gas in China during 2015-2020 using a novel grey model. *Journal of Intelligent & Fuzzy Systems,* 32**,** 141-155.

BONETT, D. G. & WRIGHT, T. A. 2015. Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior,* 36**,** 3-15.

BOX, G. E. & JENKINS, G. M. 1976. *Time series analysis: forecasting and control, revised ed*, Holden-Day.

BOX, G. E., JENKINS, G. M., REINSEL, G. C. & LJUNG, G. M. 2015. *Time series analysis: forecasting and control*, John Wiley & Sons.

BRIDGWOOD, P. 2015. Method and system for production of liquid natural gas. Google Patents.

BUTT, M. 2014. Forecast Model for Transmission of Natural Gas of Loopline and Mainline: A Time Series Analysis using Box Jenkins Approach. *Journal of Statistics,* 21**,** 41-61.

CANOVA, S., CORTINOVIS, D. L. & AMBROGI, F. 2017. How to describe univariate data. *Journal of thoracic disease,* 9**,** 1741-1743.

COLLINGS, J. 2002. *Mind over matter: The Sasol Story: A half-century of technological innovation*, Sasol.

CRESWELL, J. W. 2014. *A concise introduction to mixed methods research*, Sage Publications.

CSIR 2016. Effects on National Energy Planning and Energy Security

DARRAH, T. H., VENGOSH, A., JACKSON, R. B., WARNER, N. R. & POREDA, R. J. 2014. Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales. *Proceedings of the National Academy of Sciences,* 111**,** 14076-14081.

DAVIES, R. J., ALMOND, S., WARD, R. S., JACKSON, R. B., ADAMS, C., WORRALL, F., HERRINGSHAW, L. G., GLUYAS, J. G. & WHITEHEAD, M. A. 2014. Oil and gas wells and their integrity: Implications for shale and unconventional resource exploitation. *Marine and Petroleum Geology,* 56**,** 239-254.

DBSA 2017. NATURAL GAS IN SOUTHERN AFRICA. *In:* UNIT, K. M. A. I. K. (ed.).

DE RIJKE, K. 2013. Hydraulically fractured: Unconventional gas and anthropology (Respond to this article at http://www. therai. org. uk/at/debate). *Anthropology Today,* 29**,** 13-17.

DICKEY, D. A. & FULLER, W. A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association,* 74**,** 427-431.

DILAVER, Ö., DILAVER, Z. & HUNT, L. C. 2014. What drives natural gas consumption in Europe? Analysis and projections. *Journal of Natural Gas Science and Engineering,* 19**,** 125-136.

DISSANAYAKE, D. & PERERA, S. 2015. An empirical study of domestic electricity demand in Sri Lanka.
DME 2005. Gas Infra-Structure Plan *In:* ENERGY, D. O. M. A. (ed.).
DMR 2015. SOUTH AFRICA'S MINERAL INDUSTRY *In:* AFRICA, M. R. R. O. S. (ed.).
DMR 2017. SOUTH AFRICA'S MINERAL INDUSTRY
DOE 2010. South African Energy Synopsis *In:* ENERGY, D. O. (ed.). Department of Energy.
DOE 2013b. Department of Energy *In:* ENERGY, S. A. D. O. (ed.).
EHRENBERG, R. 2012. The facts behind the frack: Scientists weigh in on the hydraulic fracturing debate. *Science News,* 182**,** 20-25.
ELAMIN, N. & FUKUSHIGE, M. 2018. Modeling and forecasting hourly electricity demand by SARIMAX with interactions. *Energy,* 165**,** 257-268.
ELFIL, M. & NEGIDA, A. 2017. Sampling methods in clinical research; an educational review. *Emergency,* 5.
ERDOGDU, E. 2010. Natural gas demand in Turkey. *Applied Energy,* 87**,** 211-219.
ESSO, L. J. & KEHO, Y. 2016. Energy consumption, economic growth and carbon emissions: Cointegration and causality evidence from selected African countries. *Energy,* 114**,** 492-497.
FANG, Z. & CHANG, Y. 2016. Energy, human capital and economic growth in Asia Pacific countries—Evidence from a panel cointegration and causality analysis. *Energy Economics,* 56**,** 177-184.
FANG, Z. & WOLSKI, M. 2016. Human capital, energy and economic growth in China–Evidence from multivariate nonlinear Granger causality. *European Investment Bank,* 1**,** 1-25.
FENG, S., LI, H., QI, Y., GUAN, Q. & WEN, S. 2017. Who will build new trade relations? Finding potential relations in international liquefied natural gas trade. *Energy,* 141**,** 1226-1238.
FEODOROFF, T., FRANCO, J. & MARTINEZ, A. 2013. Old Story, New Threat: Fracking and the global land grab. *TNI Agrarian Justice Programme Briefing Paper. Transa national Institute*.
FIG, D. & SCHOLVIN, S. 2015. Fracking the Karoo: The Barriers to Shale Gas Extraction in South Africa Based on Experiences from Europe and the US.
GASACT 2001. Gas Act No. 48 OF 2001. *In:* DOE (ed.). South Africa.
GENTLES, S. J., CHARLES, C., PLOEG, J. & MCKIBBON, K. 2015. Sampling in qualitative research: Insights from an overview of the methods literature. *The Qualitative Report,* 20**,** 1772-1789.
GHASEMI, A. & ZAHEDIASL, S. 2012. Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism,* 10**,** 486.
GOLDBERG, A. 2015. *The economic impact of load shedding: the case of South African retailers.* University of Pretoria.
GUJARATI, D. N. & PORTER, D. 2009. Basic Econometrics Mc Graw-Hill International Edition.
GUPTA, R., KABUNDI, A., MILLER, S. M. & UWILINGIYE, J. 2014. Using large data sets to forecast sectoral employment. *Statistical Methods & Applications,* 23**,** 229-264.
GUPTA, R., KANDA, P. T., MODISE, M. P. & PACCAGNINI, A. 2015. DSGE model-based forecasting of modeled and nonmodeled inflation variables in South Africa. *Applied Economics,* 47**,** 207-221.
HEALE, R. & TWYCROSS, A. 2015. Validity and reliability in quantitative studies. *Evidence-Based Nursing,* 18**,** 66-67.
HU, D. & XU, S. 2013. Opportunity, challenges and policy choices for China on the development of shale gas. *Energy Policy,* 60**,** 21-26.
HUSSAIN, M. E. & HAQUE, M. 2017. Fiscal Deficit and Its Impact on Economic Growth: Evidence from Bangladesh. *Economies,* 5**,** 37.
IEA 2015.
IEA, I. 2011. World energy outlook 2011. *Int. Energy Agency***,** 666.
IEA. 2017. *World Energy Outlook 2017*, Organisation for Economic Co-operation and Development, OECD.
JARQUE, C. M. & BERA, A. K. 1987. A test for normality of observations and regression residuals. *International Statistical Review/Revue Internationale de Statistique***,** 163-172.
JOHANSEN, S. & JUSELIUS, K. 1990. Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and statistics,* 52**,** 169-210.

KIM, K.-H., KABIR, E. & KABIR, S. 2015. A review on the human health impact of airborne particulate matter. *Environment international,* 74**,** 136-143.
KINENE, A. 2016. FORECASTING OF THE INFLATION RATES IN UGANDA:: A COMPARISON OF ARIMA, SARIMA AND VECM MODELS.
KONDASH, A. & VENGOSH, A. 2015. Water footprint of hydraulic fracturing. *Environmental Science & Technology Letters,* 2**,** 276-280.
MAKANANISA, M. P. & ERERO, J. L. 2018. Predicting South African personal income tax - using Holt-Winters and SARIMA. *Journal of Economics & Management,* 31**,** 24-49.
MERVEN, B., MARQUARD, A., SENATLA, M., STONE, A., MCCALL, B. & AHJUM, F. 2017. The long-term uptake of natural gas in the South African energy system.
MLADENOVIĆ, J., LEPOJEVIĆ, V. & JANKOVIĆ-MILIĆ, V. 2016. Modeling AND PROGNOSIS OF THE EXPORT OF THE REPUBLIC OF SERBIA BY USING SEASONAL HOLT-WINTERS AND ARIMA METHOD. *MODELIRANJE I PROGNOZIRANJE IZVOZA REPUBLIKE SRBIJE PRIMENOM SEZONSKOG HOLT-WINTERSOVOG I ARIMA METODA.,* 54**,** 233-260.
MOKHATAB, S. & POE, W. A. 2012. *Handbook of natural gas transmission and processing*, Gulf professional publishing.
MOORE, C. W., ZIELINSKA, B., PETRON, G. & JACKSON, R. B. 2014. Air impacts of increased natural gas acquisition, processing, and use: a critical review. *Environmental science & technology,* 48**,** 8349-8359.
MUDGAL, N. 2015. *An empirical analysis of the relationship between the Purchasing Managers' Index (PMI) and share prices in the manufacturing sector of South Africa.*
MUNICIPALITY, E. 2015. Natural Gas Position Paper. *In:* DOE (ed.).
NERSA 2015a. Determination of the inadequate competition in the piped-gas industry as contemplated in chapters 2 and 3 of the Competition Act, 1998 (Act No. 89 of 1998) *In:* ENERGY, D. O. (ed.).
NERSA 2015b. Discussion document – 2015, Determination of the inadequate competition in the piped-gas industry as contemplated in chapters 2 and 3 of the Competition Act, 1998 (Act No. 89 of 1998) *In:* DOE (ed.).
NERSA 2016.
NERSA 2017. REASONS FOR DECISION ON ROMPCO (PTY) LTD'S APPLICATION FOR A MULTI-YEAR TARIFF

In: GAS, N. P. (ed.).
OGUZ, M. E. 2013. Forecasting Turkey's Sectoral Energy Demand
PAYLOR, A. 2017. The social–economic impact of shale gas extraction: a global perspective. *Third World Quarterly,* 38**,** 340-355.
PETROSA 2016. Presentation on the 2014/2015 Impairment Report Presented to the Portfolio Committee on Energy. *In:* ENERGY, D. O. (ed.).
QUANTEC. 2018. Available: https://www.quantec.co.za/software/ [Accessed].
RIMANDO, M., BRACE, A. M., NAMAGEYO-FUNA, A., PARR, T. L., SEALY, D.-A., DAVIS, T. L., MARTINEZ, L. M. & CHRISTIANA, R. W. 2015. Data collection challenges and recommendations for early career researchers. *The Qualitative Report,* 20**,** 2025-2036.
ROMPCO. 2018. Available: www.rompco.co.za/content/gas-pipeline [Accessed 2018].
ROWSELL, A. 1976. Diagenesis in Cape and Karroo sediments, South Africa, and its bearing on their hydrocarbon potential. *South African Journal of Geology,* 79**,** 81-145.
SASOL 2014. Unlocking the potential of SOUTHERN AFRICA's gas resources.
SAUNDERS, M., LEWIS, P. & THORNHILL, A. 2016. *Research methods for business students*, Pearson education.
SCHRAG, D. P. 2012. Is shale gas good for climate change? *Daedalus,* 141**,** 72-80.
SEKARAN, U. & BOUGIE, R. 2016. *Research methods for business: A skill-building approach*, John Wiley & Sons.
SIMUNDIC, A.-M. 2013. Bias in research. *Biochemia medica: Biochemia medica,* 23**,** 12-15.

SINGH, S., MISHRA, S., MATHAI, R., SEHGAL, A. K. & SURESH, R. 2016. Comparative Study of Unregulated Emissions on a Heavy Duty CNG Engine using CNG & Hydrogen Blended CNG as Fuels. *SAE International Journal of Engines,* 9**,** 21-29.

SLG. 2018. Available: https://www.google.co.za/search?q=spring+lights+gas&rlz=1C1PRFI_enZA785ZA785&oq=spring+li&aqs=chrome.2.69i57j0l5.8042j0j8&sourceid=chrome&ie=UTF-8 [Accessed].

SOC, T. 2016. NATURAL GAS INFRASTRUCTURE PLANNING. CHAPTER 6.

STEINZOR, N., SUBRA, W. & SUMI, L. 2013. Investigating links between shale gas development and health impacts through a community survey project in Pennsylvania. *New solutions: a journal of environmental and occupational health policy,* 23**,** 55-83.

STRIOLO, A. 2018. Shale Gas: Friend or Foe?: An overview of the scientific and technical challenges that have surfaced because of shale gas. *TCE: The Chemical Engineer***,** 40-43.

SZOPLIK, J. 2015. Forecasting of natural gas consumption with artificial neural networks. *Energy,* 85**,** 208-220.

TABACHNICK, B. G. & FIDELL, L. S. 2007. *Using multivariate statistics*, Allyn & Bacon/Pearson Education.

TAVAKOL, M. & DENNICK, R. 2011. Making sense of Cronbach's alpha. *International journal of medical education,* 2**,** 53.

TEMPER, L., YÁNEZ, I., SHARIFE, K., OJO, G., MARTINEZ-ALIER, J., CANA, M. C., CORNELISSEN, K., LERKELUND, H., LOUW, M. & MARTÍNEZ, E. 2013. Towards a Post-Oil Civilization. Yasunization and other initiatives to leave fossil fuels in the soil. *Environmental Justice Organizations, Liabilities and Trade (EJOLT) Report*.

VAN BASTEN, C. 2007. Economic regulation of SA's public utilities. Pretoria: TIPS.

VRIEZE, S. I. 2012. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods,* 17**,** 228.

WADUD, Z., DEY, H. S., KABIR, M. A. & KHAN, S. I. 2011. Modeling and forecasting natural gas demand in Bangladesh. *Energy Policy,* 39**,** 7372-7380.

WALLIMAN, N. 2011. *Your research project: Designing and planning your work*, Sage Publications.

WEINHOLD, B. 2012. The future of fracking: new rules target air emissions for cleaner natural gas production. *Environmental health perspectives,* 120**,** a272.

WOOD, D. A. 2012. A review and outlook for the global LNG trade. *Journal of Natural gas science and Engineering,* 9**,** 16-27.

WRIGHT, J. 2016. EFFECTS ON NATIONAL ENERGY PLANNING AND ENERGY SECURITY.

ZHANG, Z. 2016. Univariate description and bivariate statistical inference: the first step delving into data. *Annals of translational medicine,* 4**,** 91-91.