Torjus Halden

Estimation of Trade Flows With the Use of AIS Data

The Case of LNG Shipping

Master’s thesis in Marine Technology
Supervisor: Bjørn Egil Asbjørnslett
June 2019

Master’s thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Marine Technology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Estimation of Trade Flows With the Use of AIS Data:

The Case of LNG Shipping

Torjus Halden

June 2019

Master Thesis
Department of Marine Technology
Norwegian University of Science and Technology

Supervisor : Bjørn Egil Asbjørnslett

Preface

This thesis marks the final part of my 5 year long integrated Master of Science degree within Marine Systems Design at the Department of Marine Technology (IMT). The work has been carried out at the Norwegian University of Science and Technology (NTNU) during the spring term of 2019, and corresponds to 30 ECTs.

Trondheim, 2019-06-11
Torjus Halden
Torjus Halden

Acknowledgment

I would like to thank the following persons for their great help during my project:

Professor Bjørn Egil Asbjørnslett for supervising me through this project.
Bjørnar Brende Smestad for his guidance and workshops on AIS data.
The Norwegian Coastal Administration for providing me with AIS data.

My office mates, great friends for all the support.

T.H.

Summary

The objective of this thesis is to investigate whether liquefied natural gas (LNG) trade flows can be estimated reliably with the use of vessel positional data from the Automatic Identification System (AIS). Several specialised analytics and consulting firms base their income on providing energy and trade flow analysis. Their data is often based on manual data collection from agents, analysts, scouts and official customs data. With increasing availability and completeness of vessel positions from satellite AIS data, possibilities within automatic and continuous trade flow estimations arrive that can outperform inefficient, expensive and delayed data collection.

In this thesis, a thorough literature review is presented to investigate state-of-the-art within the topic, and an introduction to LNG trade is given. This include historical development, fleet analysis and operational characteristics derived from AIS data. The AIS data used is provided by the Norwegian Coastal Organisation, while the LNG carrier fleet data is provided by IHS Markit.

The created trade flow estimation methodology is based on geo-fencing LNG terminals, detecting unique visits to the terminals and accumulating the visited gas capacity to estimate trade flow. Circular geo-fences of three different sizes were tested on eight LNG exporting countries and six LNG importing countries for both 2017 and 2018. In addition, slow speed and navigational status criteria "moored" or "anchored" where added as a fourth case. All cases were compared to trade flow numbers from the International Gas Union which were used as a benchmark for the import and export values.

Overall, the results from our analysis establish evidence that LNG trade flows can be estimated with a low error by using vessel position data from AIS. Six countries got estimation errors of less than 10% compared to benchmark, and it was found that geographical position, the amount of close by ship traffic, size of geo-fence, and data resolution are all crucial factors for how small the error of the estimation can get. The results are best for 2018 due to higher data resolution, and for terminals placed far north and/or in low vessel density areas. In addition, a large geo-fence decrease the error, as the time frame given to detect the vessel increases. The inclusion of low speed and moored/anchored criteria drastically increase the error and is concluded useless.

To conclude, LNG trade flows estimated from AIS data show good alignment with benchmark values and can hence be estimated with low error. This given that the terminal is within an area with high vessel detection probability. Nevertheless, it is important to state that analysis over a larger time frame is needed to draw a solid conclusion. In addition, implementation of draught analysis are recommended for further work and the methods should if possible be tested on data with higher resolution.

Sammendrag

Målet med denne masteravhandlingen er å undersøke om varestrømmer av flytende naturgass (LNG) kan estimeres pålitelig ved bruk av posisjonsdata for skip fra det automatiske identifikasjonssystemet (AIS). Flere spesialiserte analyse- og konsulentfirmaer baserer sin inntekt på å levere analyser innen energy og varestrøm. Deres data er ofte basert på manuell datainnsamling fra agenter, analytikere, speidere og fra tolldata. Med økende tilgjengelighet og fullstendighet av skipsposisjoner fra satellitt AIS-data, dukker det opp muligheter innen automatisk og kontinuerlig estimering av varestrømmer som vil kunne utkonkurrere den ineffektive, dyre og forsinkede innsamlingen som gjøres i dag.

I denne masteravhandlingen presenteres et grundig litteraturstudie for å sette seg inn i den fremste forskningen innen emnet, og videre blir en introduksjon til frakt av LNG gitt. Dette inkluderer historisk utvikling, flåteanalyse, samt operasjonelle karakteristikker utledet fra AIS data. AIS dataen som er benyttet i denne oppgaven er levert av Kystverket, mens flåte-dataen kommer fra IHS Markit.

Metoden som er laget for å estimere LNG varestrøm baserer seg på å "geofence" LNG terminaler for å detektere unike besøk til terminalen, og akkumulere opp volumet av LNG som har vært i den tilknyttede "geofencen". Tre ulike størrelser av sirkulære "geofencer" blir benyttet på åtte LNG eksportører og seks importører i tidsperiodene 2017 og 2018. I tillegg blir et fjerde "geofence" testet som også inkluderer krav om lav fart, samt at skipet har navigasjonsstatus "ankret" eller "fortøyd". Alle de ulike casene blir sammenlignet med tall fra International Gas Union som blir brukt som referanseverdi.

Totalt sett viser resultatene fra vår analyse klare indikasjoner på at LNG varestrømmer kan estimeres med liten feilmargin. Seks land blir estimert med en feilmargin på under 10% sammenlignet med referanseverdier ved bruk av AIS data, og det blir funnet at resultatene er nært knyttet til deteksjonssansynligheten av skip i det aktuelle området. Geografisk posisjon, mengden av nærgående skipstrafikk, "geofencens" størrelse og dataoppløsningen er alle avgjørende faktorer for hvor liten feilen i estimeringen blir. Resultatene er best for 2018 på grunn av høyere dataoppløsning, og for terminaler plassert langt nord og / eller i områder med lav skipstraffikk. I tillegg reduserer et stort "geofence" feilen ettersom det øker tiden skipet tilbringer i "geofencen". Inkluderingen av krav om lav hastighet og ankret/fortøyd øker feilen dramatisk og kan konkluderes som ubrukelig.

Det kan konkluderes med at LNG varestrømmer estimert fra AIS data er godt tilpasset referanseverdiene og kan derfor estimeres med liten feilmargin. Dette gitt at terminalen ligger innenfor

et område med høy deteksjonssansynlighet. Likevel er det viktig å poengtere at en analyse over en større tidsperiode er nødvendig for å trekke en sikker konklusjon. Som videre arbeid anbefales implementering av dypgangsanalyser i metoden, samt å teste metodikken på enda mer høyoppløselig data.

Contents

Acronyms

AIS Automatic Identification System

BOR Boil-off Rate

COG Course Over Ground

FSRU Floating Storage and Regasification

ETA Estimated Time of Arrival

IMO International Maritime Organisation

LNG Liquefied Natural Gas

MMBTU Million British Termal Units

MMSI Maritime Mobile Service Identity

Mtoe Million tonnes of oil equivalents

MTPA Million Tonnes Per Annum

SQL Structured Query Language

S-AIS Satellite Automatic Identification System

VHF Very High Frequency

WTI Western Texas Intermediate

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The shipping industry has a reputation of being slow and lack the will to innovate. This goes not only for the vessels, but the industry as a complete system. The last couple of years, this trend has started to change. We live in a digitalised world, and the shipping companies have started to understand the commercial potential of utilising the huge amounts of available data in the industry. Norwegian shipowners have for instance established digital labs, while several consulting startups base their income on pure data exploiting and digitalisation in the maritime industry. The Automatic Identification System (AIS) is a big data source that generate live vessel data from more or less all merchant vessels. This big data source gives new and emerging possibilities that can be utilised as a decision support tool for both direct and indirect stakeholders in the maritime industry.

The motivation for working with AIS data is to look into how the huge amount of data can be utilised to give valuable information about shipping and energy markets that potentially can compete with the expensive information and knowledge sold by specialised analytics and consulting firms that base much of their data stream on manual data collection from agents, scouts and reports.

In the context of energy markets, the AIS data can potentially give added value in obtaining and predicting world energy trade through the tracking of ship movements. Extracting and utilising such information can be off high importance for shipowners and investors when making both operational and investment decisions related to the shipping and energy market.

Natural gas is one of the fastest growing energy sources and it is predicted to be an increasingly important energy source due to its cleanness and lower carbon footprint compared to other fossile energy sources. Liquified Natural Gas (LNG) is natural gas predominantly methane, with some mixture of ethane that has been cooled down to liquid form for ease and safety of non-pressurised storage or transport. This commodity introduced in 1964 has enabled natural gas to be transported efficiently world wide with LNG carriers, and the demand increase every year with more and more LNG projects being developed. The LNG market has hardly been subject to academic research. This, combined with the the potential LNG has to reduce emissions world wide are my main motivations for looking into the LNG shipping and energy market with the use of AIS data.

1.2 State-of-the-art AIS Research

Research based on utilising AIS data has gained more and more interest the last ten years, and previous research reveal the wide range of possibilities that lies within AIS data. A large part of the studies are related to shipping networks and patterns. Automatic generation of maritime networks is done in Arguedas et al. (2014), while Kaluza et al. (2010) interpret global cargo ship movements as a complex network, and Spiliopoulos et al. (2018) presents a big data driven approach to extract global trade patterns. Pallotta et al. (2013) purpose a model of anomaly detection and Tu et al. (2018) present methods for collision avoidance and path planning. Ship tracking performance and investigation of the completeness of AIS data is reviewed in Skauen et al. (2013),Skauen and Øystein Olsen (2016) and Eriksen et al. (2018).

Vessel speed analysis from AIS data has among others been done by Leonhardsen (2017) to analyse the fuel saving potential of using rapidly re-configurable bulbous on container ships, and Adland and Jia (2016) who use AIS data to investigate how vessel speed is affected by freight rates and fuel costs. Smestad (2015) show how specific ship types can be identified with the sole use of S-AIS data. Trade flow analysis from AIS data has been done on the crude oil market by Adland et al. (2017) and on the bulk segment by Jia et al. (2019). AIS data has also been used in combination with machine learning methodology by Næss (2018) to predicate freight rates in the LPG market and by Millefiori et al. (2016) to define seaport's operational area. The huge opportunities that lays in using AIS data also bring new security issues to the table in terms of cyber attacks and spoofing. This issue is addressed by Kontopoulos et al. (2018) among few.

There are many investigated user applications for AIS data. Nevertheless, the author see a knowledge gap and large potential of doing research on how to utilise AIS data in order to gain market insight, especially in the LNG shipping business.

1.3 Objectives

The main objective of this thesis is to investigate whether liquefied natural gas (LNG) trade flows can be estimated reliably with the use of vessel positional data from the Automatic Identification System (AIS). Four research tasks have been identified in order to fulfil the objective of the thesis:

1. Obtain key knowledge of the natural gas and LNG industry, including analysis of the fleet composition and basic operational patterns regarding the fleet.
2. Filter and extract relevant LNG vessels from global AIS data.
3. Develop a model that estimates LNG trade volumes based on AIS data.
4. Evaluate the quality and precision of the developed model and identify key weaknesses and improvement areas in the model.

1.4 Scope and Limitations

The scope of this study is first of all limited to LNG shipping and trade volume estimation. However, the methods and models built and used are applicable to other shipping segments with adjustments.

The estimation horizion of this study is yearly import or export volumes of LNG, as the available benchmark volumes are given yearly. A smaller time frame would have given increased information on the reliability and performance of the model, but unfortunately no monthly comparison data has been obtained. The author has been given AIS data from the Norwegian Coastal Organisation for the years 2017 and 2018, and all AIS related analysis are limited to these two years.

1.5 Structure of the Report

The remaining part of the thesis is structured as followed:

Chapter 2 conducts a literature Review on previously done research within the field of AIS data, as well as LNG shipping and trade. The chapter also gives insight to the authors approach when conducting the literature review, as it has been a key part of the literature review.

The data foundation used in this thesis is presented in chapter 3. Here, fundamentals of AIS data are addressed as well as an introduction to the AIS database created by the Norwegian Coastal Administration. Lastly, we introduce the fleet data and trade volume benchmark data used in the thesis.

An introduction to the LNG shipping industry is given in chapter 4 to get expedient background knowledge. The development of the industry, market and pricing is presented, as well as a section analysing the fleet characteristics. The chapter ends with a minor analysis of the LNG fleet operational characteristics based on AIS data.

The methodology and model built to find LNG trade volumes is presented in chapter 5. This includes both the methodology used for data preprocessing and trade flow estimation.

Chapter 6 conducts a case study on LNG trade volumes to evaluate if AIS data can give reliable trade volume information based on the model obtained in chapter 5, and the results from the case study with related discussion can be found in chapter 7.

Chapter 8 gives a discussion on whats done, including the constructed methodology and the data foundation used in the case study, as well as the various choices made in this thesis.

Chapter 9 concludes the work done in this thesis and gives recommendations for further work.

Chapter 2

Literature Review

The literature review has been an important part of this master thesis. Around 40 articles have been reviewed, and a structured approach has been used in the search for articles. In this chapter key articles are presented. The aim of the literature review has been to get an initial understanding of how AIS data has been of use in research and what possibilities the data give. Literature related to LNG trade and energy trade flows have also been an important part of the literature review.

Firstly the approach of the literature review is presented. Then key articles related to the objectives of this thesis are walked trough. The chapter ends with a short conclusion on what remains to be done.

2.1 Approach and Method

Using a structured approach in the literature review has been a focus throughout this master thesis. This section presents the approach and method used in the literature review.

2.1.1 Generally

Scoping studies: Towards a Methodological Framework (Arksey and O'Malley (2005)) has been used as a main framework in this literature review. The framework divide the scoping into five different steps, where three of them are relevant for this project:

- Stage 1: identifying the research question
- Stage 2: identifying relevant studies
- Stage 3: study selection

These three stages give the fundament of the search. However, throughout the master thesis, the author has seen a need to do iterations, as the scope and problem definition has been tuned dynamically as the authors knowledge increased through the work.

2.1.2 Identifying the Research Question

The initial goal of the literature study was to study AIS data and exploit methods used to utilise AIS data. This was primarily done in the authors project thesis written in the autumn semester as a preparation towards this master. Further on the research where narrowed down, as the authors knowledge increased and the master thesis scope where clearly defined. As mentioned, the goal has been iterated throughout the project, and so have the research questions. These two research questions have been key when digging into literature:

1. What are the most researched segments and methods within AIS data and what areas seem to be unexplored?
2. How can AIS data be utilised by vessel management and analysts from a commercial point of view?
3. How are worldwide energy trade flows obtained and predicted?

2.1.3 Identifying Relevant Studies

The primary focus on all the research questions has been on the methods used in the relevant literature. The search tactic has been to start with broad searches before narrowing down the search. Electronic databases have been the primary source in this literature survey. Using multiple databases have broaden the spectre of relevant literature and has reduced the probability of missing relevant literature. Oria, Scopus and Google Scholar have been the key search engines used. Using the bibliography from obtained literature is an other much used and efficient method to find relevant literature. This is also called snowballing. The last search methods used is to be guided directly to relevant sources by people with competence on the field. This can be very efficient.

2.1.4 Study Selection

Although AIS is not the most researched topic out there, there is still countless of literature available. Therefore it is necessary to use a searching technique to select which literature to look deeper into. The search technique is divided into two main steps:

1. The first step is to evaluate literature. This process is based on the thesis guide made by NTNU [1]. The method lets the author review a large amount of literature, and it is of great help to decide where to invest the most time.
2. Step two is to assess the reliability, objectivity, accuracy, and aptitude of the article. This method is called the R-O-A-A principle (T-O-N-E in Norwegian) and it is used to check if the source has high or low quality. Sources with high quality is preferred, but analysing sources with low quality gives the author a better understanding of what remains to be done within the subject of interest.

2.2 Literature reviewed

An important part of research question one has been to look into what "State-of-the-art" AIS research is. The findings from this can be found in the introduction under section 1.2. This section will on the other hand walk through literature that is directly or indirectly relevant to the case study related objectives of this master thesis.

AIS data has attracted more and more researchers the last years. It is important to handle AIS data in a structured manner when conducting research and analysis. However, the diversity of terms and methodologies between different research projects hinders the understanding and communication among them. Wang and Wu (2017) presents how the well known JDL data fusion model can be used to work with AIS data, and they define their own AIS version of the model shown in figure 2.1. The model is entity oriented and separated into four levels (Level 0-3). Level 0 contains the preprocessing of AIS data and the challenges related to this. Level 1 is called the entity assessment. In maritime situation awareness the main entities are vessels, ports, traffic segments, fishing areas etc. Level 2 is the relationship assessment. This level is about how a situation can be defined as relation between situations: Vessel vs vessel, vessel vs pattern etc. Level 3 is the impact assessment, which is about predicting the impact of a state. For example predicting a vessels position 24 hours in advance. This model is an important framework which the author will utilise for working with AIS in this master thesis, and to get an expedient structure.

---

[1] https://www.ntnu.no/viko/

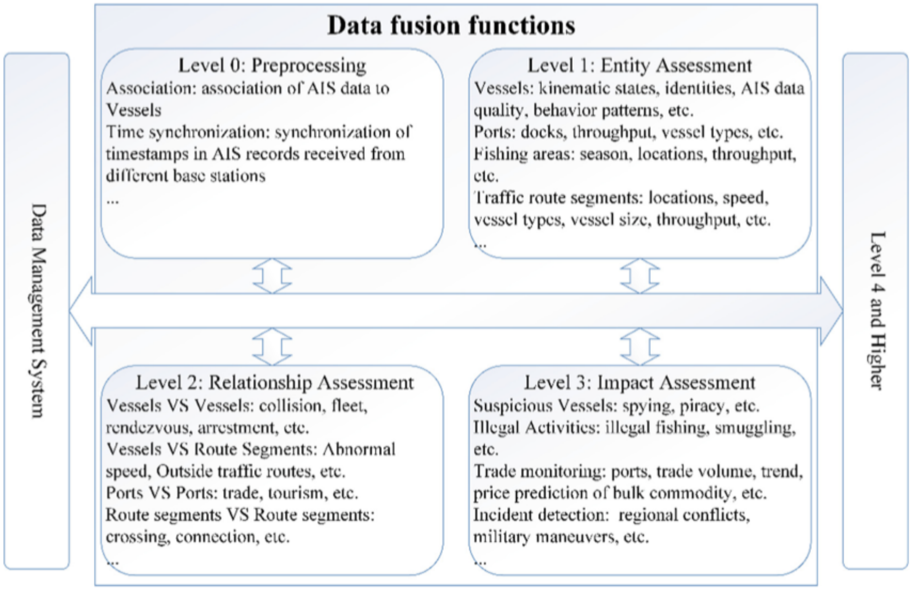


Figure 2.1: JDL AIS data fusion model

LNG freight as a specific shipping market has gotten little research interest compared to other shipping markets such as the oil tanker market, bulk market, and container market. This is also the case when looking at AIS specific research conducted on the LNG freight segment. There has been done research about LNG as ship fuel and safety related research with the use of AIS. However, to the authors knowledge, Axelsen (2018) is the only research conducted on the LNG shipping market with the use of AIS data. Axelsen (2018) obtain a list of all LNG carriers worldwide and preprocess a AIS data set to only contain messages from LNG carriers. He then obtain a list of all LNG carriers trading in the spot market by accessing public information manually, and make a separate database containing AIS messages from spot trading vessels. Further on he analyse the carrier speed distributions from his two databases and compare them to the LNG spot prices. His findings indicate that LNG carriers sail slower than their design speed, and that the spot fraction of the fleet does not adjust speed to high and low rates. His research does not distinguish between vessels sailing fully loaded and vessels sailing a laden return, which could be of interest when analysing speed versus spot prices. Geng et al. (2014) used AIS data to conduct dynamic speed choice analysis on bulk carriers. His findings also indicate that freight rates does not affect speed choice. Instead, vessel-specific variables such as age and design speed, as well as operational factors such as loading conditions, show some explanatory power. Næss (2018) master thesis investigates multivariate freight rate prediction using machine learning on AIS data on the LPG market, which have many similarities with the LNG market. He uses two different models; the Multilayer Perceptron (MLP) and Long Short Term Memory (LSTM) neural network. To the authors knowledge this is the most thorough research done within this field, and thesis concludes that there is evidence that AIS data can be used to predict short-term freight rates, and that machine learning outperforms traditional models. Applying these methods on the LNG market in combination with AIS-derived trade volumes to predicate future trade volumes could give valuable insight.

Estimating and conducting analysis on LNG trade flows is an important part of this thesis. The current literature on trade flows tends to either look at international trade at a macro level with theoretical models, or ship routing and optimisation on a micro level. Few articles focus on modeling seaborne commodity trade. Trade flows can be estimated by discrete choice theory, leading to gravity models formulations. However, Babri et al. point out that trade flows may be affected by long-term contracts or simply trading habits. This leads to a situation where only a fraction of the market is subject to choice. In their article they extend the traditional gravity model such that a fixed amount is separated from observed trade flows and only residuals are subject to choice. The model is applied on seaborne trade of iron, coal and oil, and the results indicate a much better fit than the traditional gravity model. This is highly relevant to the LNG shipping market, as the majority of trade is conducted trough long term contracts. Adland et al. (2017) is to the authors knowledge the only research that compare the accuracy of AIS-derived trade statistics to official customs data. The research use the crude oil market as case and do a breakdown of trade by vessel size over time. They find that while AIS-derived data for seaborne crude exports show good alignment with official export numbers in aggregate, there are substantial temporal and geographical differences across countries and time due to the use of pipelines and transshipment in parts of the supply chain. By conducting a similar analysis on the LNG shipping segment we would expect better results, as all LNG is traded via seaborne trade. Jia et al. (2019) conducts similar research on the bulk segment where the goal was to estimate vessel payloads in bulk shipping using AIS data. Draught from AIS messages combined with a port call dataset is used to estimate the vessels payload on a micro level. The results indicate similar consistency as models relying on principles from physics and navel architecture.

Analysing the natural gas and LNG market development and putting it in relation with other energy commodities is key to understand the LNG market. Geng et al. (2014) has done a thorough analysis on natural gas trade networks. It is found that both the LNG and pipeline gas import and export trade networks display scale-free distributions, while the countries in the LNG trade network are linked more closely than those in the pipeline gas trade network. The research also concludes that the evolution of inter-regional LNG trade will accelerate the integration of the separated international natural gas market. Makholm and Olive (2016) on the other hand argue that the global LNG spot market will not develop in the future, although the trend of increased spot volumes is a fact. A key argument used is the high costs related to LNG transportation.

2.2.1 What Remains to be Done

Previous research has shown that AIS data can give additional insight to shipping and energy trade. Nevertheless, it became clear from the literature review that using AIS data in order to asses commercial intelligence on energy markets has gotten little interest from researchers, especially when looking specifically at the LNG market. There has been some trade flow studies and to the authors knowledge only one study looking at the LNG market from an AIS point of view. However there exists knowledge gaps within obtaining, testing and reviewing methods that can be used to analyse worldwide shipping trade flows, which is non-existing when looking at LNG trade. This master thesis can therefore potentially give high added value both from a research and commercial point of view. For research it will test an unexplored branch within LNG shipping, and commercially the results may be of high value for companies that rely on the work of expensive energy market reports.

Chapter 3

Data Foundation

This chapter presents the data foundation for this master thesis. The main data source is the Automatic Identification System, which we give a short but thorough introduction to. Other data sources used in this thesis are the fleet data and the annual trade volume data. These are also introduced in this chapter.

3.1 Automatic Identification System Data

3.1.1 Introduction to Automatic Identification System

AIS is a communication system introduced by the United Nations maritime organisation IMO to improve safety, environmental impact and to regulate and monitor the shipping traffic. The system communicates through the maritime Very High Frequence (VHF) bands to transmit ship movements and technical data at specific intervals. The AIS regulations was introduced in 2000 by IMO, and came into effect for all ships within the scope of the regulations in 2004. The regulations requires ships to automatically provide data to other ships, aircrafts and appropriately equipped shore stations. The ships are also required to receive data automatically from similarly fitted ships. The data include static data, such as the name of the ship, draught, destination and estimated time of arrival (ETA). The data also include dynamic data such as speed and position. See section 3.1.3 about the message types. A typical use of AIS is to exchange information between vessels within range of each other, to avoid potential collisions.

From around 2008 satellites where also capable of receiving AIS messages, in addition to base stations. This is called S-AIS. The AIS messages can reach around 40-50 nautical miles in horizontal direction at sea level and up to around 200 nautical miles in vertical direction, Skauen et al. (2013). The introduction of S-AIS in combination with the fact that AIS messages are available to anyone with a receiver has given the foundation to AIS used as a commercial tool.

3.1.2 AIS Regulations

The AIS regulations where introduced by IMO, and came into effect in 2004. The regulations requires AIS to be fitted on all ships of 300 gross tonnage and upwards engaged on international voyages, all cargo ships of 500 gross tonnage and upwards and all passenger ships irrespective of size, IMO (2002). The regulations require the ships the send specific messages presented in section 3.1.3 at specific time intervals presented in section 3.1.4. The ships are also required to receive AIS messages.

3.1.3 AIS Message Types and Content

The International Telecommunication (ITU) is the organisation that has defined the AIS message types. ITU (2014) has defined 27 different message types in total, and the 5 most common can be found in table 3.1. The majority of all AIS messages is of type one. Message type one can be seen in table 3.2. Another important message type is type 5, which can be seen in table 3.3. Type 5 include more information, with destination for the vessel as a key element.

Table 3.1: AIS Message type 1-5

ID	Name	Description
1	Position Report	Scheduled position report
2	Position Report	Assigned scheduled position report
3	Position Report	Special position report, response to interrogation
4	Base Station Report	Position, UTC, date and current slot number of base station
5	Static and Voyage Related Report	Scheduled static and voyage related vessel data report

Table 3.2: AIS Message type 1 key features

Information	Description
Unixtime	Number of seconds elapsed since 1 January 1970
Position	Coordinates, longitude and latitude
Speed	Speed over ground (SOG) in knots
Course	Course over ground (COG)
MMSI	Maritime Mobile Service Identity (Vessel ID)
Status	Navigational status

Table 3.3: AIS Message type 5 key features

Information	**Description**
Unixtime	Number of seconds elapsed since 1 January 1970
Vessel specifications	Length and breadth, in meters
Draught	Current draught in meters
IMO Number	International Maritime Organization number
Origin	Origin of current voyage
Destination	Destination of current voyage
ETA	Estimated time of arrival, measured in Unixtime
MMSI	Maritime Mobile Service Identity (Vessel ID)
Vessel type	Vessel type category

Vessels can be identified by two unique numbers included in the AIS messages. The Maritime Mobile Service Identity (MMSI) is connected to the AIS-gear onboard vessels, and it is included in all AIS messages. The MMSI is only changed if the ownership of the vessel is changed. The IMO number included in message type 5 is the other unique identification number. This ship identification number was introduced by IMO in 1987 as a measure aimed at enhancing "maritime safety, and pollution prevention and to facilitate the prevention of maritime fraud" [1]. The number is permanent through a vessels life and does not change upon ownership change. All propelled, sea-going merchant ships of 100 GT and above is assigned an IMO number upon keel laying with the exception of the following: Ships without mechanical means of propulsion, pleasure yachts, ships engaged on special service (e.g. lightships, SAR vessels), hopper barges, hydrofoils, air cushion vehicles, floating docks and structures classified in a similar manner, ships of war and troopships, and wooden ships.

In message type 5 (table 3.3) information about the vessel type is given. This is described by a double digit number between 10 and 99. The first digit represent the ship type which is divided into 9 categorise (see table 3.4), while the second represent whether a cargo is dangerous, hazard or a marine pollutant.

3.1.4 Message Intervals

The different AIS messages are sent at different time intervals. These intervals are defined by 'Guidelines for the onboard operational use of shipborne AIS' by IMO (2002). The guidelines distinguishes between static and dynamic data. The static data is sent every 6 minutes or upon request, while the dynamic data are sent at different time intervals shown in table 3.5.

---

[1] http://www.imo.org/en/OurWork/MSAS/Pages/IMO-identification-number-scheme.aspx

Table 3.4: Ship types defined in the AIS data

First Digit	**Ship Type**
1	Reserved for future use
2	WIG (Wing in Ground)
3	Other vessels
4	High-speed carrier, or vessels<100 GT
5	Special craft
6	Passanger ships>100 GT
7	Cargo Ships
8	Tankers
9	Other types of ships

Table 3.5: AIS Message Intervals

Ship status	**General reporting interval**
Ship at anchor	3 min
Ship at 0-14 knots	12 sec
Ship at 0-14 knots and changing course	4 sec
Ship at 14-23 knots	6 sec
Ship at >23 knots	2 sec
Ship at 14-23 knots and changing course	3 sec
Ship at >23 knots and changing course	2 sec

3.1.5 AIS Data Quality

AIS data is collected by AIS receivers. These are located on board ships, on land , on buoys and satellites. The land based receivers can normally detect messages from ships up to 40-50 nautical miles offshore (Skauen et al. (2013)) and all ships outside this zone will stay undetected by the land receivers. Although the AIS system originally was not intended for space based receivers, several studies indicated that the signals could be picked up from space. As a result, the first Norwegian AIS satellite AISSat-1 was launched in 2010, followed by a second one AISSat-2 in 2013. Today Norway has a total of 4 AIS satellites after Norsat-1 and Norsat-2 was launched in 2017. The use of AIS satellites expands the possibilities of AIS drastically since the whole world in theory can be covered. However, there will still be blindspots and therefore gaps in the data because of the satellites orbits. Interference and messages from ships not detected is another issue that has been a result of the great area covered. Smestad (2015) also discovered that the actual captured data also include thousands of erroneous AIS messages, both in terms of position, dimensions, IMO number etc. Still the majority of AIS data is reliable, but this is important to have in mind when filtering, cleaning and analysing the data.

3.1.6 The AIS Data Used in this Thesis

The AIS data used in this thesis has been provided by the Norwegian Coastal Administration, captured by the four Norwegian satellites and decoded by Bjornar Brende Smestad. The data spans from the first of January 2017 to the end of 2018 and contain more than 30 million AIS messages from more than 500 LNG carriers. Further introduction and preprocessing of the AIS database is done in the case study, chapter 6.

3.2 Fleet Data

The purpose of this master thesis is to conduct analysis on the LNG carrier fleet with the use of AIS data. It is therefore expedient to filter away all AIS messages not containing a LNG carrier. Commercial fleet databases are expensive, hence it is preferable to conduct the filtering process without such a database. Smestad (2015) established heuristics to obtain vessel type with the sole use of AIS data. They found that certain vessel types could be found with very high accuracy. Unfortunately, the obtained heuristics where not able to distinguish between LNG carriers and LPG carriers. Obtaining such heuristics is not part of the scope of this master thesis. Therefore a commercial database called "Seaweb", provided by IHS (2019) is used as foundation for LNG fleet data. The fleet data from Seaweb is used to exclude all vessels that are not LNG carriers from the AIS database, and we we use the Seaweb data to obtain the LNG carrier fleet characteristics. How the AIS LNG carrier database is obtained is described in detail in the methodology section 5.1 and the final database after preprocessing is described in chapter 6. The LNG carrier fleet characteristics are presented in chapter 4.

3.3 Trade Flow Data

The trade flow data used in this master thesis is obtained from the International Gas Union annual report, IGU (2019). The data include yearly LNG imports and exports between all countries world wide. The data is reported in million tonnes per annum (MTPA) and data between 2011-2018 has been used in this thesis. This data is used as a benchmark when conducting trade flow analysis with the use of AIS data.

3.4 LNG Import and Export Terminal Data

A list containing all LNG liquefaction and regasification terminals have been obtained from the International Gas Union annual report, IGU (2019). This data include terminal name, country, capacity and startup year. Sourcewatch[2] has been used as a secondary source when obtaining the coordinates of the various terminals. These coordinates are often only approximate and needs to be tuned further with the use of satellite photos and AIS analysis of the area.

---

[2] https://www.sourcewatch.org/

Chapter 4

LNG Trade

Knowledge about natural gas as an energy source and the market fundamentals is important to be able to conduct analysis related to LNG transportation. Hence, the author has used time to obtain knowledge regarding natural gas and LNG. This chapter will give an introduction to LNG trade and should give the necessary background information in order to fully understand the analysis conducted in this thesis. The chapter ends with a minor study of the operational pattern of LNG carriers based on AIS data.

4.1 Background

4.1.1 Natural Gas

Natural gas is made up by primarily methane and ethane, with varying amounts of heavier hydrocarbons. It is the cleanest fossile energy source used, with carbon dioxide and water vapour as the main combustion products. Natural gas has been used as an energy source since long before the 20th century. Initially it was almost exclusively used as a source of light, and the main source of gas was manufactured from coal. Today natural gas is mainly used for heating and electricity production. It is the third largest energy source, accounting for 23% percent of the worlds energy demand (BP (2019)). Natural gas has about double the reservoir recovery (70-80 percent) than oil (30-40 percent), which lessens the need to continuously find new gas fields. In addition, natural gas requires less processing compared to oil to become "pipeline quality". Lastly, the reserves to production ratio is 50 percent larger than for oil, which means there are plenty of potential for expansions, and the "oil and gas industry" may one day be dubbed the "gas and oil industry", Nersesian (2007).

The main drawback with natural gas has been the logistical challenge of transporting it from production to consumer, as well as storing it. Other energy sources such as oil and coal are liquid and solid at ambient pressure and temperature making transportation on trucks, trains or ships expedient and cost effective. Natural gas, with its low boiling point is in gas form at the same conditions, making it a logistical challenge. Pipelines has been the main transport solution, but these are costly to build. Also the gas could not be stored at the consumer end as with oil products, meaning that the delivery systems had to be designed to handle extreme vagaries in demand. As a consequence, the majority of natural gas had to be consumed close to the production cite. If there were no such demand in the lokal market, the gas was flared (burned) or vented to the atmosphere. These huge amounts of energy waste gave strong incentives to improve pipeline technology, and solve the logistical challenge with natural gas. Another consequence is that the natural gas market has been much more regulated than other energy markets due to the fact that you can only be connected to one pipeline, in the same way that your house only can be connected to one power grid. These factors has been important for how the natural gas market has developed, both technological, and the development of natural gas as a commodity.

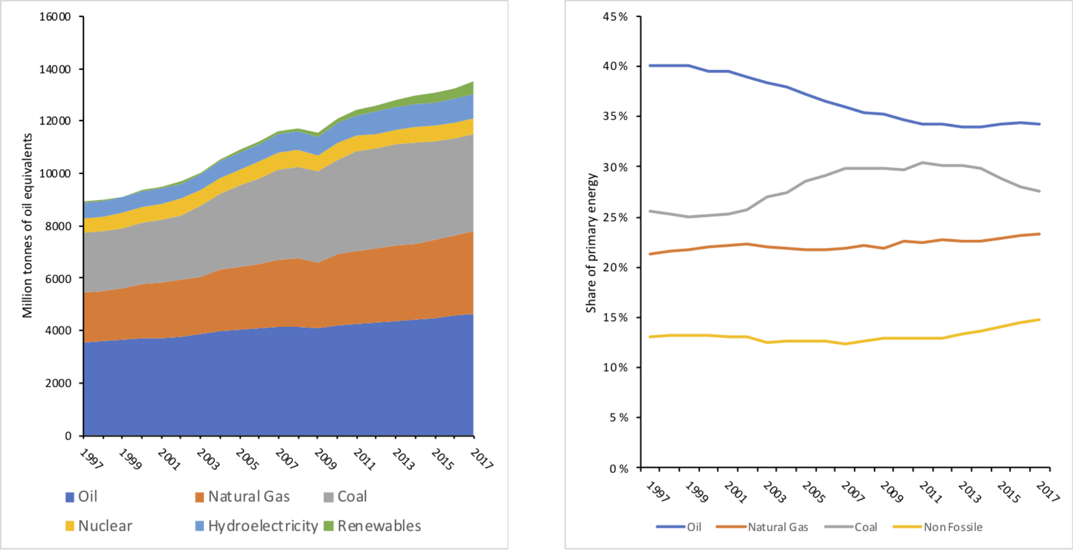
4.1.2 Liquified Natural Gas

Liquified Natural Gas (LNG) is natural gas that has been cooled down to liquid form, which is approximately -162 °C. In its liquid form, LNG occupies about 1/600 the volume of the natural gas feedstock, which gives LNG an obvious advantage in transportation and storing of natural gas. The first patent on natural gas liquefaction was granted in 1914, while the first commercial liquefaction plant was built in Cleveland, Ohio in 1941, Bosma and Nagelvoort (2009). The initial purpose of liquefaction was to liquefy gas when demand were low, and re-gas on demand peaks. This is called peak-shaving and base load. The advantage regarding transportation was first utilised in 1959 with the first commercial LNG carrier, enabling natural gas to be transported over long distances. Today natural gas in the form of LNG is transported world wide with LNG carriers at huge volums, accounting for 11.5% of the total natural gas consumption, IGU (2018).

4.2 Natural Gas and LNG Market Development

4.2.1 Demand

The worlds energy demand has increased by 2.1% yearly on average the last 20 years (Compounded annual growth rate (CAGR)), while the demand for natural gas has increased by 2.6% yearly in the same period. Natural gas has seen the fastest growth in absolute numbers the last



(a) World energy demand by source

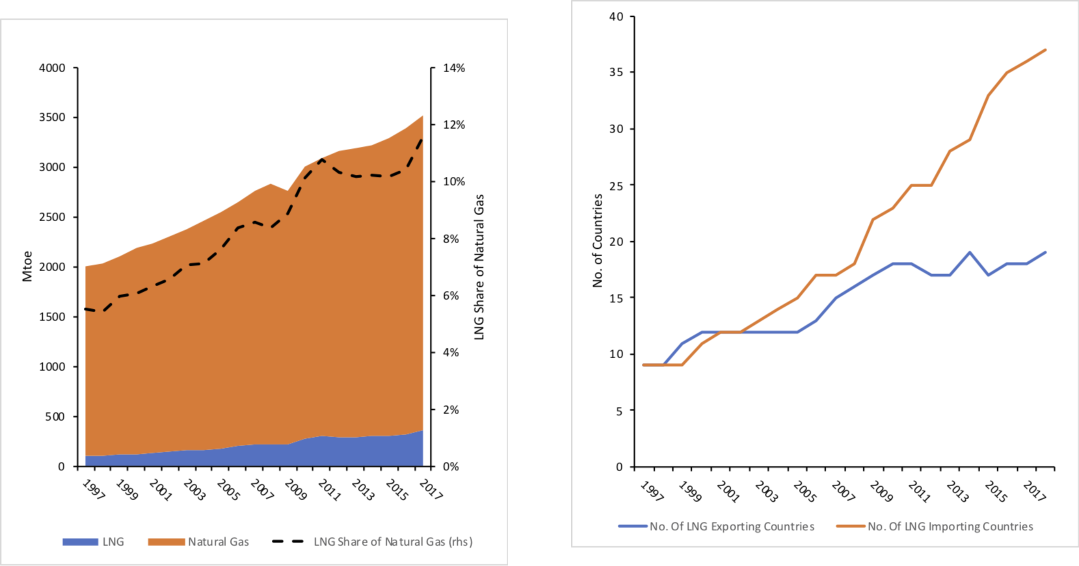
(b) Share of primary energy by source

Figure 4.1: World energy demand development, with data from BP (2019)

ten years, and in 2017 it accounted for approximately 23% of the worlds energy demand, making it the third largest energy source (BP (2019)). Figure 4.1 shows the development of energy demand and mix.

In percentage, World LNG demand has seen the second strongest growth after renewables in the last 20 years with an average annual growth of 6.5%. This gives a total increase of 350% to a total demand of 364 Mtoe in 2017, IGU (2018). As figure 4.2a shows, LNGs share of natural gas demand has doubled the last 20 years, now accounting for close to 12% of the total natural gas demand. Figure 4.2b shows the development in import and export countries of LNG. The number of importing countries has seen a sharp growth to a total of 36 countries in 2017, while the number of export countries has stabilised around 18 for now, IGU (2018). However, several LNG projects are under development and additional countries are expected to both start importing and exporting LNG.

There are several reasons to why natural gas and especially LNG has seen great growth the last two decades. A key reason is related to energy security. Only about one-third of the natural gas reserves are in the politically unstable middle east, while about two-thirds of the oil reserves lie there, Nikhalat-Jahromi et al. (2017). Less risk is thus related to natural gas as an energy source. However, if a country import natural gas through a pipeline system, the country is most likely only connected to one pipeline, as mentioned in section 4.1.1. For instance, several countries in East-Europe have been completely dependent on natural gas imported through a Russian pipeline. The development of LNG terminals has been and is an efficient way of reducing this



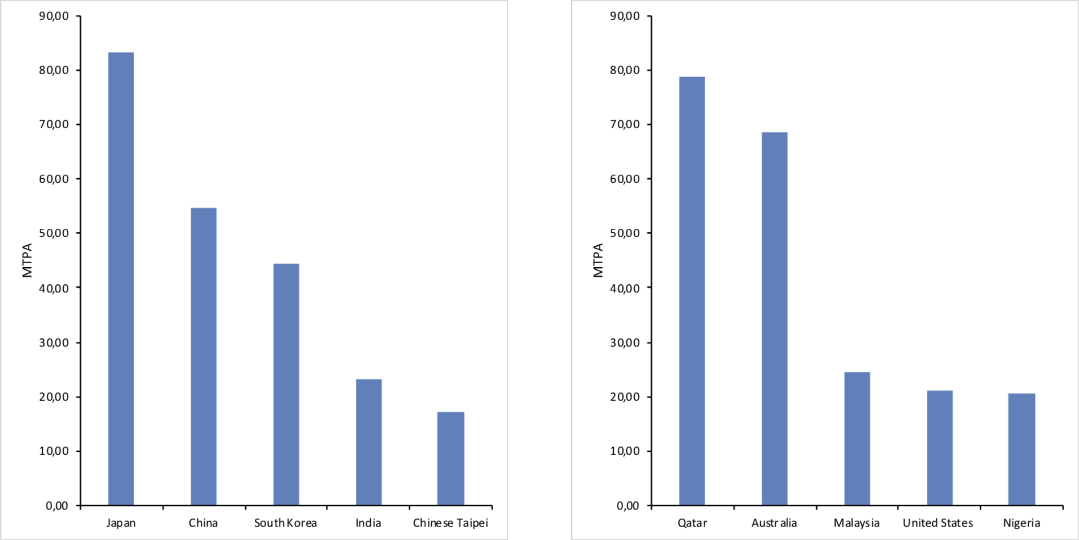
(a) World LNG and natural gas demand

(b) Number of LNG import and export countries

Figure 4.2: World LNG and natural gas market development, with data from IGU (2019) and BP (2019)

dependency and the related energy risk for countries that rely on natural gas, because the LNG can be imported from several different countries, in opposition to a pipeline. Lithuania have among others done this diversification with the chartering of a Floating storage and regasification terminal (FSRU).

Natural gas is also considered the most environmentally friendly fossile fuel, which gives the energy source political support. Last but not least, natural gas and LNG is highly competitive on price, which in the end is extremely important.



(a) Top 5 LNG importing countries

(b) Top 5 LNG exporting countries

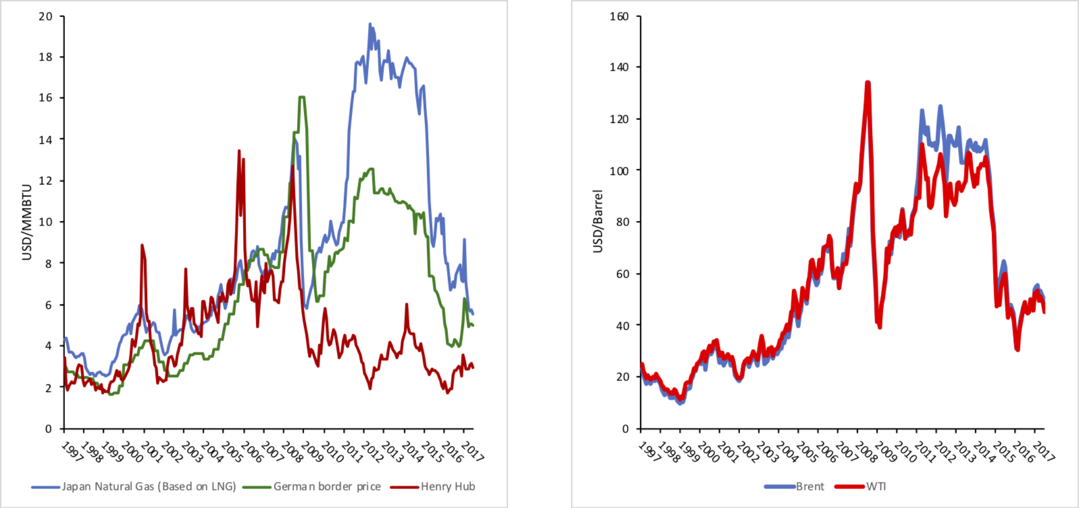
Figure 4.3: Top 5 LNG importers and Exporters of 2017, with data from IGU (2019)

4.2.2 Natural Gas Price

The lack of non-pipeline transport and the high capital cost of LNG facilities means that the kind of well functioning worldwide spot market we have for oil (Brent and Western Texas Intermediate (WTI)), seen in figure 4.4b, does not exist in the natural gas world. Natural gas prices are to some extent liked to the oil price, but historically, natural gas markets have been highly geographically restricted with local market dynamics (supply and demand) and regulations dictating the price. As seen in figure 4.4a the natural gas price has historically been highly volatile with huge geographical price difference, in opposition to oil where we can observe that WTI and Brent are strongly correlated. For instance, the natural gas in Europe (German Border Price) has cost more than the double of US gas (Henry Hub) between 2007 and 2017. In a perfect well functioning market, the price difference should never be higher than the cost of transportation between the various markets. As LNGs share of natural gas increases, the price spread is expected to become more stable, and the link to the oil price should increase.

4.3 LNG Shipping

The first commercial LNG Carrier was built in 1959 with a capacity of 5,500 m^3 to enable natural gas to be transported over long distances. Sixty years later, the industry has changed rapidly both technical and commercial with LNG carriers as large as 264,000 m^3.



(a) Natural gas price 1997-2017

(b) Oil Price 1997-2017

Figure 4.4: Natural gas and Oil historical price development

This section will introduce the LNG value chain with its related costs, fleet development, and market dynamics and characteristics.

4.3.1 LNG Value Chain

The LNG value chain is typically divided into four main steps: Exploration and production, liquefaction, shipping, and regasification, visualised in figure 4.5. The four steps can differ or merge, depending on the methods and facilities used.

Natural gas is typically extracted from onshore gas reservoirs, processed, and sent either to the end user through pipelines in its natural form, or it is sent to a liquefaction facility. However, as the natural gas demand increase, it has been essential to exploit unconventional gas fields. There exist many types of unconventional gas fields, such as shale gas and tight gas fields, but nearly 45% of all natural gas reserves are located offshore, IEA (2018). In order to produce LNG from these type of fields, offshore platforms are commonly used. Further on the gas is pre-treated and cleaned before it is sent by pipelines to land. Here the gas is liquefied and stored in tanks. In 2017 the first offshore liquefaction plant came online which allows cheaper and more remote production of natural gas (LNG), since the production and liquefaction is done at the same site and the need for expensive pipelines are drastically reduced. The Shell Prelude and the Golar Hilli are state of the art examples of offshore liquefaction, but several other projects are under development. After liquefaction, the LNG is ready to get shipped by specialised LNG carriers to its destination. The last step is to regasificate the LNG before it is ready to be used in for example power generation. Regasification is most commonly done on an onshore terminal.

Figure 4.5: LNG value chain

However, floating storage and regasification terminals (FSRU) are getting more and more common.

A similarity with all phases of the LNG value chain is that they are all capital intensive. The cost breakdown between the different phases will vary a lot from project to project, but a rough estimation presented by Maxwell and Zhu (2011), suggest exploration and production of feedstock supplies represent 15-20% of total capital costs, liquefaction comprises 30-45% of costs, shipping accounts for 10-30%, and regasification accounts for the remainder 15-25%.

4.3.2 LNG Carrier Fleet Development

Background knowledge on the LNG carrier fleet is crucial to conduct both operational and trade flow analysis. This section will therefore look at the LNG carrier fleet development, both technical, fleet size and vessel sizes, as well as service speed development.

There has been built approximately 730 LNG carriers since the first was built in 1959, IHS (2019). The first vessel had a capacity of about 5,500 m^3, and it took only ten years before vessels with gas capacity over 70,000 m^3 where built. Figure 4.6 shows the accumulated total gas capacity and number of vessels built. The first big wave of carriers where built in the second half of the 1970s with most vessels having a gas capacity of around 125,000 m^3 and a service speed as high as 19-20 knots. The high service speed was a result of the high boil-off-rate of 0.25%, where the boil-off gas could cover the high fuel consumption at the given speed. A stable fleet growth followed in the 1990s before a new wave where built after entering the 21st century. As shown in figure 4.6, both the total gas capacity and the number of vessels built has exploded since this. As much as 85% of the built gas capacity has been built between 2000-2018, while 78% of the vessels has been built in the same period. Newbuilding orders indicate that the fleet growth will continue at a high pace, driven by the increased demand for LNG and the increasing amount of offshore LNG projects.

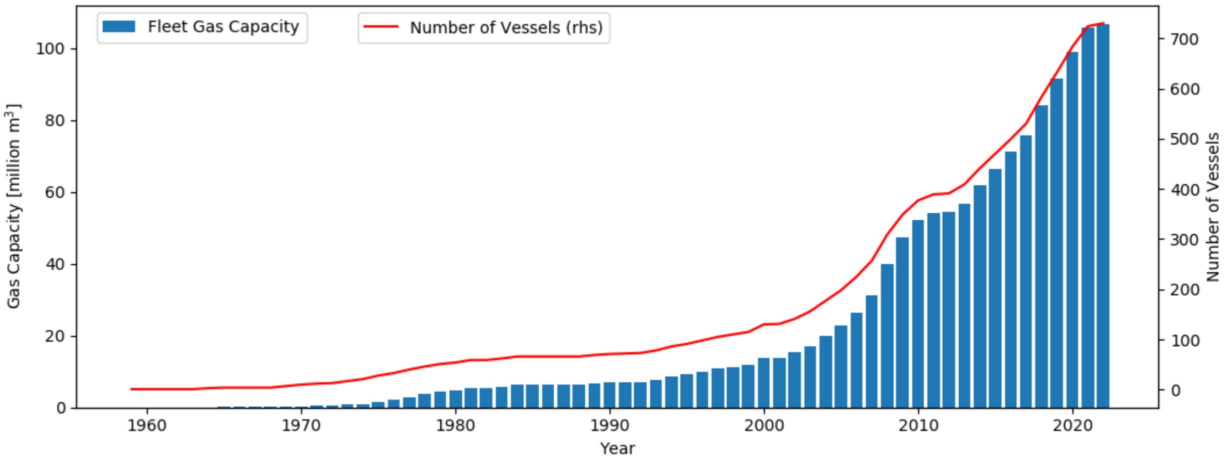


Figure 4.6: LNG total gas capacity built and number of vessels built, with data from IHS (2019)

4.3.3 Current LNG Carrier Fleet

By end of 2018, 503 LNG carriers where in service. These are shown by gas capacity in figure 4.7 and figure 4.8 and by service speed in figure 4.9, also including vessels on order. The fleet has been divided into five categorise based on their gas capacity. The three most common vessels are "Small Gas Carriers" which have a gas capacity under 50,000 m^3, "Conventional Gas Carriers" which have a capacity of 50,001-145,000 m^3 and "New Panamax" (which are the largest type that can pass through the Panama Canal) with a capacity of 145,001-180,000. The two last categorise are the "Q-Flex" and "Q-Max". The Q stands for Qatar, and the vessel design is specifically related to port size in Qatar. Q-Flex has a gas capacity of 205,000 to 215,000 m^3, while the Q-Max is the largest possible vessel that can dock in Qatar with a capacity of 250,000-262,000 m^3.

From 4.7, we can observe that the size of LNG-Small Gas Carriers has been relatively stable historically. These vessels are typically used for short-distance coastal trading. The LNG-Conventional Gas Carriers has on the other hand been built bigger and bigger, creating the "New Panamax" category, which for now has stabilised around 170,000 - 180,000 m^3. The Q-Flex, and Q-Max vessels where all built between 2008-2010, except one vessel. The LNG carrier size trends are more obvious when looking at figure 4.8. Here we can observe that the most popular vessels (deep sea) has increased in size, and that LNG-Small Gas Carriers popularity has increased the last years. We can also observe that the LNG carriers are converging towards mainly two used sizes, which are the LNG-New Panamax Carrier and the LNG-Small Gas Carrier. This indicate that the LNG carrier fleet is solving two different main tasks and that the optimal vessels sizes has been obtained with the current boundaries.

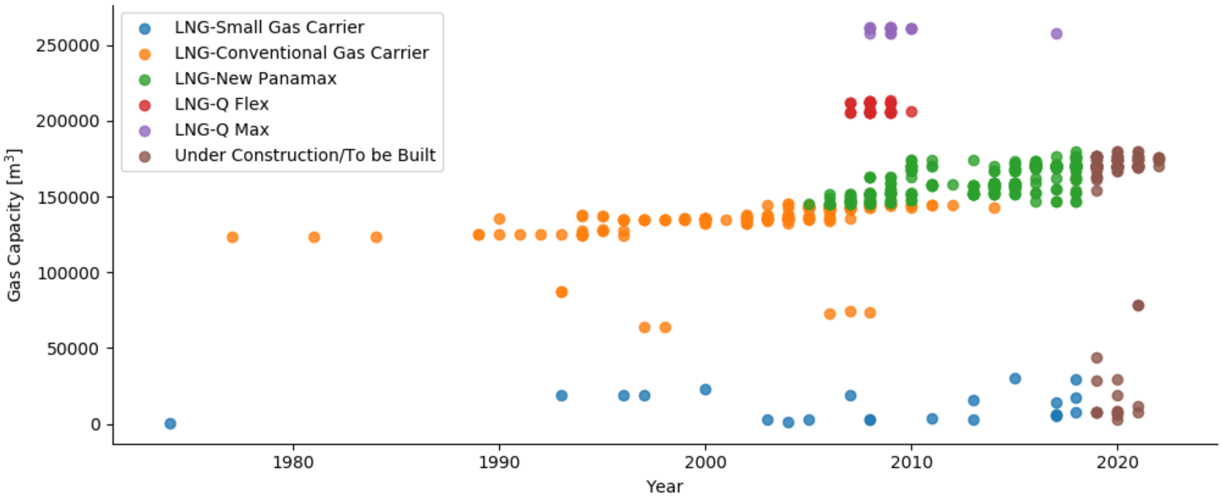


Figure 4.7: LNG carrier fleet in service and on order, by gas capacity and year built, with data from IHS (2019)

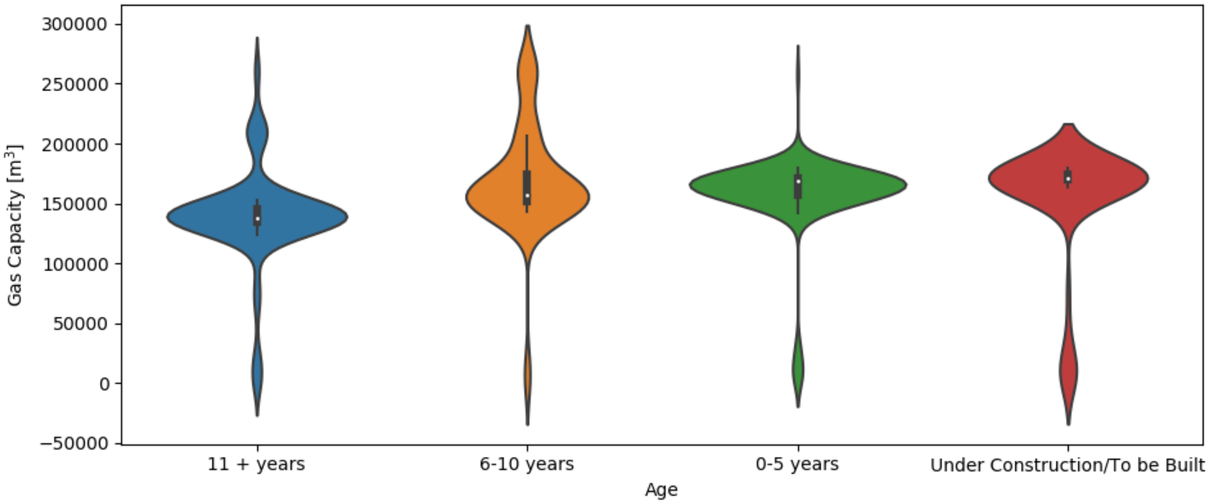


Figure 4.8: LNG carrier fleet in service and on order, by gas capacity and age, with data from IHS (2019)

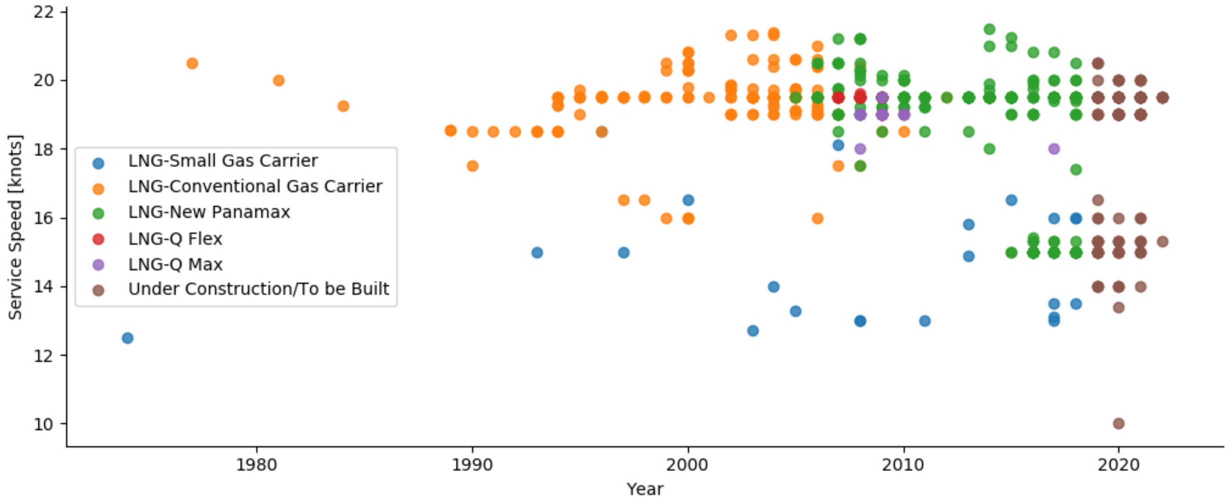


Figure 4.9: LNG carrier fleet in service and on order, by service speed and year built, with data from IHS (2019)

Figure 4.9 shows the development of service speed for the LNG carrier fleet in service and on order. The spread in service speed is huge for the LNG-Small Gas Carriers, while the service speed for the deep sea segment has been stable around 19 knots. The exception is for vessels built after 2015, where we can observe that several deep sea vessels with a service speed of around 15 knots has been built. A difference of 4-5 knots in service speed is quite large when looking at same sized vessels. The actual speeds the vessels are operating at compared to the service speed (design speed) are important factors when looking at the operational patterns in LNG trade. This is investigated in section 4.4.

4.3.4 Technology

LNG carriers are first of all classified by their containment system: Spherical or membrane. The spherical system is a design from the early 1970s where 4-5 spheres carry the LNG cargo. The spheres are made out of aluminium covered by an insulation layer and an outer steel shell. A key advantage with the design is that it limits sloshing of the cargo for improved stability, but a key downside is poor utilisation of the ships hull, due to the usage of spheres. The membrane system have tanks that conform with the ships hull, which give a high volume utilisation. The tank shape makes the system vulnerable to sloshing, thus they have to be completely filled, or almost empty when sailing. While the spherical system where popular in its early days, the membrane containment system has taken more and more over. Today the majority of LNG carriers are built with the membrane containment system.

A key feature and aspect with LNG carriers is the boil-off that occurs continuously. The containment systems job is to keep the boil-off rate (BOR) as low as possible to not waste any cargo. The first vessels in the 1970s had a BOR of 0.25% daily, while the newest technology can limit the BOR to under 0.1% daily. This is naturally an important aspect for charterers and shipowners when chartering or buying a LNG carrier. The boil-off can either be used directly as fuel for the ship, or it can be reliquefied and sent back to the tanks.

The propulsion system is also an important feature on the LNG carriers. Before 2005 almost all LNG carriers where built with a steam propulsion system. This system burn BOG and if necessary heavy fuel to generate steam for the main propulsion turbines and auxiliary systems. From 2005 diesel electric systems with the ability to burn both diesel and BOG took over for the traditional steam systems, due to 25%-30% higher efficiency. Currently almost 50% of the fleet use steam propulsion, while all vessels on order use more modern systems.

4.3.5 Market Characteristics and Dynamics

As mentioned in section 4.3.1, all phases of an LNG project are capital intensive. The projects have a long time frame, and it normally takes more than four years from the investment decision is made before the project start to generate cash (both liquefaction and regasification projects). In addition, a large part of current and potential LNG projects are located in countries with political and/or economical instability. The combination of high capital costs, long time frame (binding capital) and the location of projects gives LNG projects a high risk profile. This has possibly been an important reason to the relatively slow development in the early days of LNG. As mentioned in section 4.2.1, LNG demand and production has increased sharply the last two decades, with expectations of continued high growth looking forward.

By looking at the corresponding shipping market, the majority of LNG trade can be classified as "industrial shipping", Stopford (2009). A main characteristic for LNG shipping has been the large amount of long term contracts. From the start of LNG shipping in the 1960s and until 2000 the amount of long term contracts has always been over 90%, IGU (2019). The contracts where typically 20-25 years of length and included a buyer, a natural gas producer (seller) and an external shipowner if the vessel was not provided by the buyer or seller. A typical long term contract include a clause related to the quality and quantity of LNG to be delivered. Often the buyer is entitled to take a minimum quantity of gas each year and pay for this quantity whether or not it is taken. This is called "Take or pay provision". It is also important to have in mind that vessels with a long term contract also can trade in the spot market if they have available time and capacity.

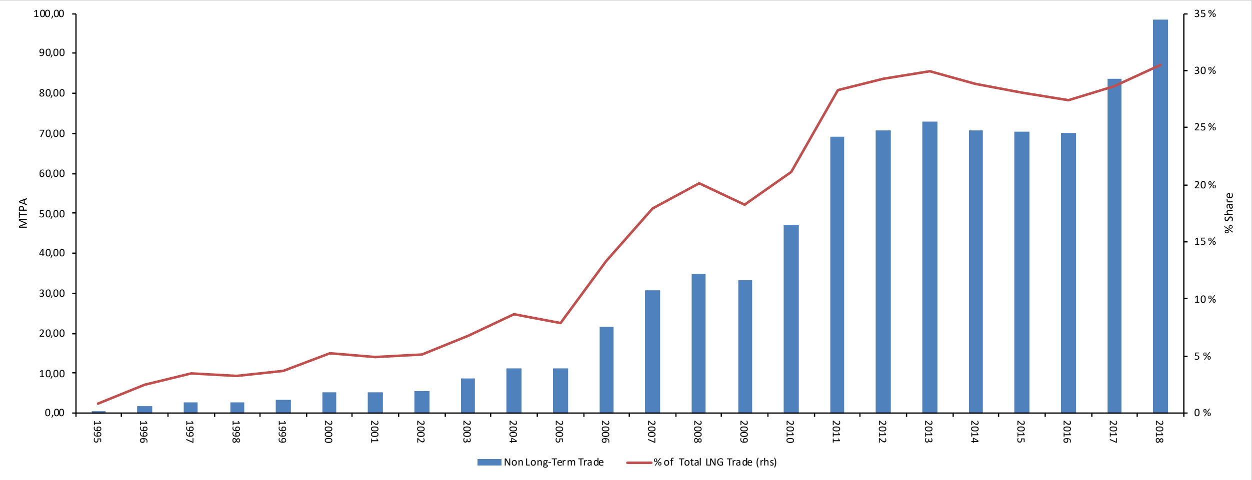


Figure 4.10: LNG short term and spot trade development, with data from IGU (2019)

Since around year 2000, the market has developed fast, and the amount of spot and short term contracts has increased from 10% to now accounting for 30% of total LNG trade (figure 4.10). This is still a very low percentage compared to the oil tanker market, where the majority of tankers are trading in the spot market.

According to IGU (2019), the increased spot and short term trade has been driven by several factors:

- The Growth in LNG contracts with destination flexibility, which has facilitated diversions to higher priced markets.
- The increased number of importers and exporters giving a more complex market.
- The growth of companies with diverse marketing portfolios taking on an aggregator role, allowing long-term offtake contracts to satisfy a variety of short- and long-term buyer commitments.
- Sudden changes in supply or demand dynamics such as the Fukushima disaster in Japan.
- The decline in competitiveness of in LNG interfuel competition such as coal in the power sector (chiefly in Europe) and shale gas (North America) that has freed up volumes to be re-directed elsewhere.
- Periods of large disparity between prices in different basins such as that from 2010 to 2014, which made arbitrage an important and lucrative monetisation strategy.
- The faster development timeline and lower initial capital costs of FSRUs compared to on-shore regasification, which allow new markets to enter the LNG import market.

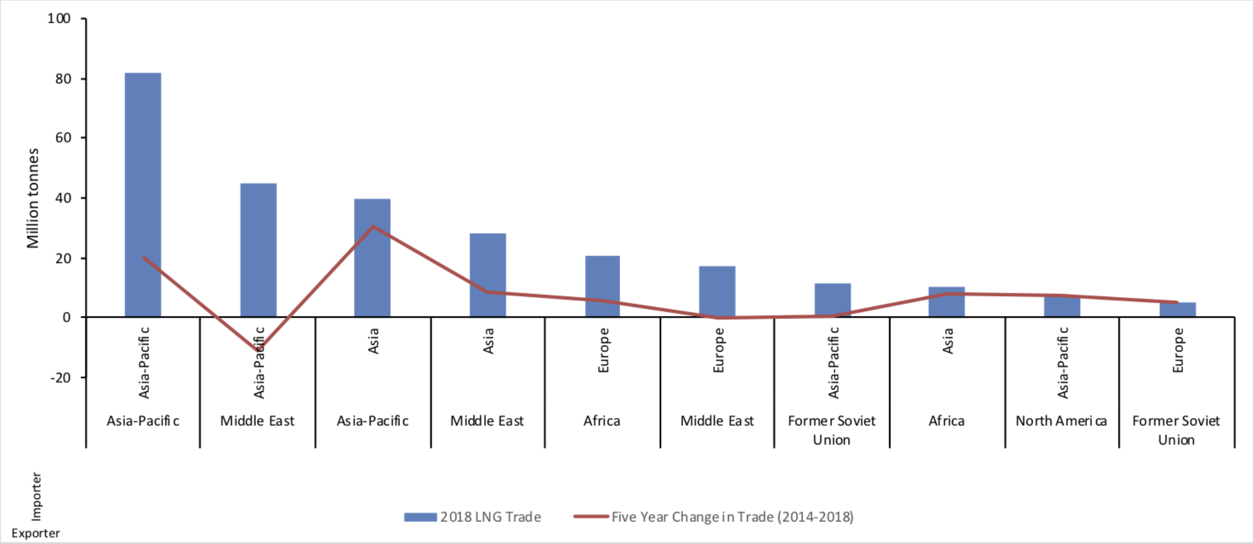


Figure 4.11: LNG trade between regions, data from IGU (2019)

- The large growth in the LNG fleet, especially vessels ordered without a long-term charter, which has at times allowed for low-cost inter-basin deliveries.

The voyage distances and the related ton mile is another important parameter to look at when considering market characteristics and dynamics. Increased ton mile implies increased need of vessel capacity, which can either be solved by building new vessels, or increasing the speed of the current operating fleet. The author do not sit on ton mile data, or voyage distance data. However, IGU (2019) provide data of the LNG trade between different regions, shown in figure 4.11 which can provide important characteristics of the development. The figure show that internal trade in the Asia-Pacific is the largest LNG market followed by exports from the Middle-East to the Asia-Pacific, and exports from the Asia-Pacific to Asia. The largest increase in trade the last five years has occurred from the Asia-Pacific mainly due to increased production in Australia. More interesting is the increase in trade from North America to the Asia-Pacific because of the long voyages that follows.

4.4 LNG Fleet Operational Pattern Obtained from AIS Data

4.4.1 Fleet Operational Areas

With the use of AIS data in combination with vessel specific information from Seaweb, the operational areas for the different types of LNG carriers have been obtained. In figure 4.12 the operational pattern of the Q-Max/Q-Flex vessels can be observed. The vessels mainly operate from the export terminal of Qatar and voyage to Europe and South Asia. The Q-Max/Q-Flex operational pattern is a prime example of industrial shipping with monotone movements and long

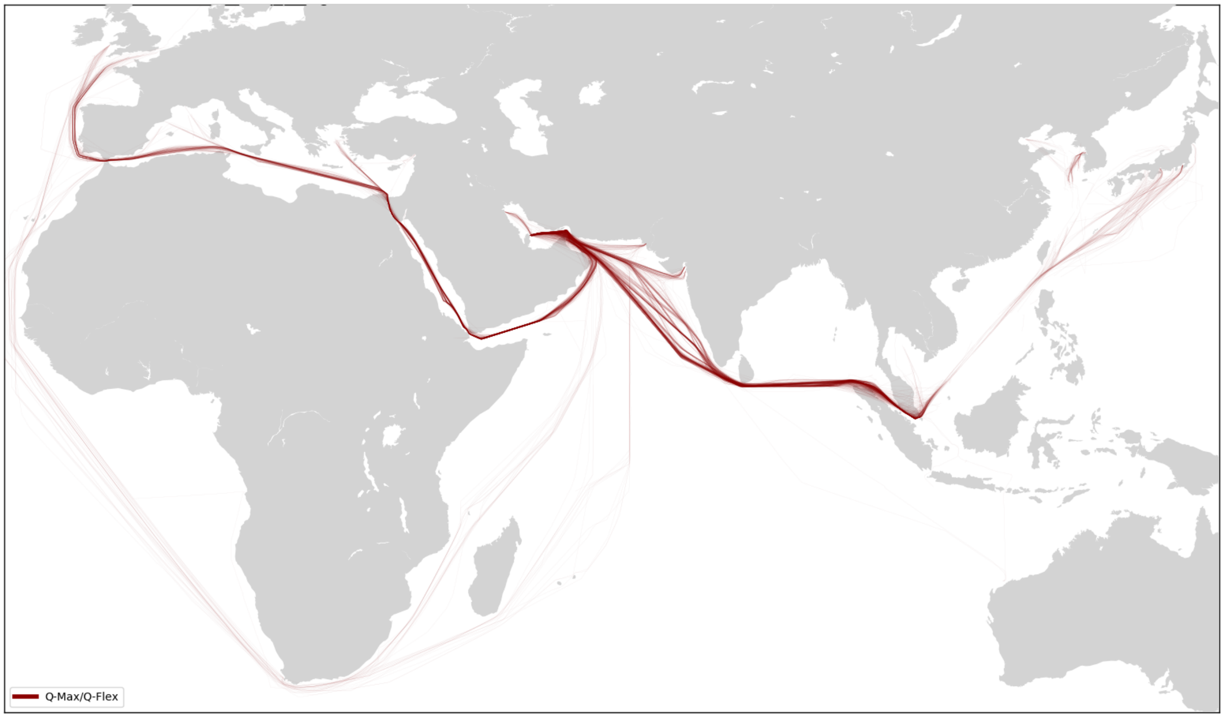


Figure 4.12: Q-Max/Q-Flex carriers operations, 2017-2018

contracts. Figure 4.13 and figure 4.14 show the New-Panamax and Conventional vessels respectively, which are the majority of the LNG carrier fleet. These types of vessels operate between all the main LNG terminals in the world. Lastly, figure 4.15 show the Small-scale LNG carriers. As seen in the figure, these vessels operate on much shorter voyages with three main operational areas. These are Northern Europe, the Caribbean Ocean and East Asia.

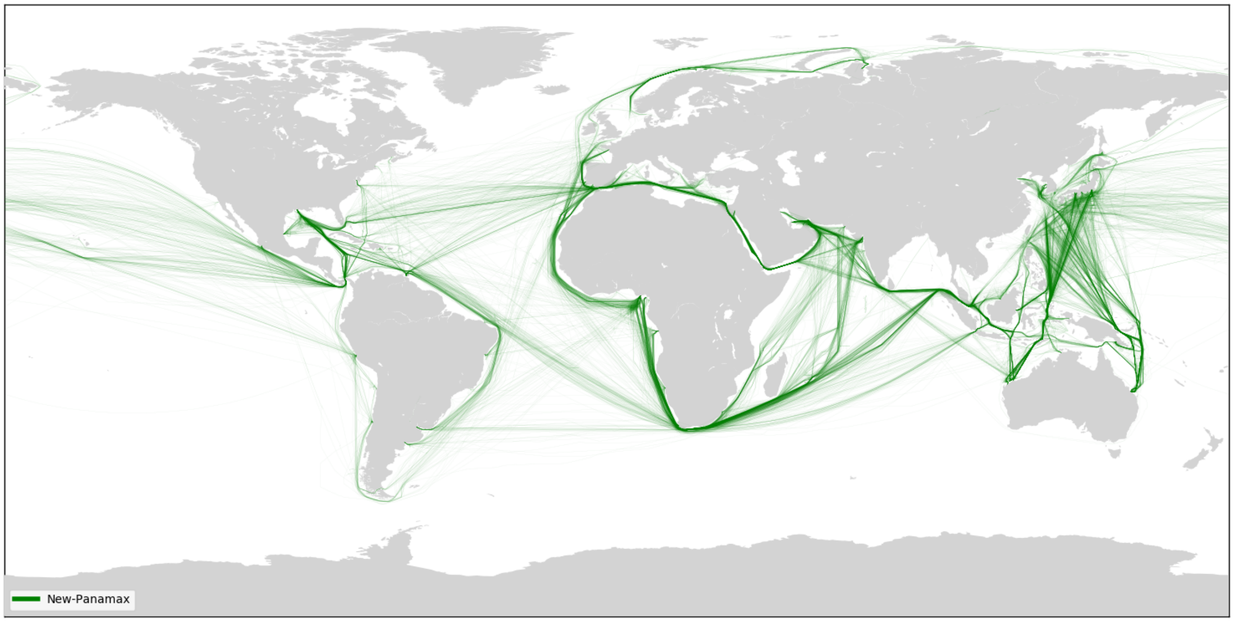


Figure 4.13: New-Panamax carriers operations, 2017-2018

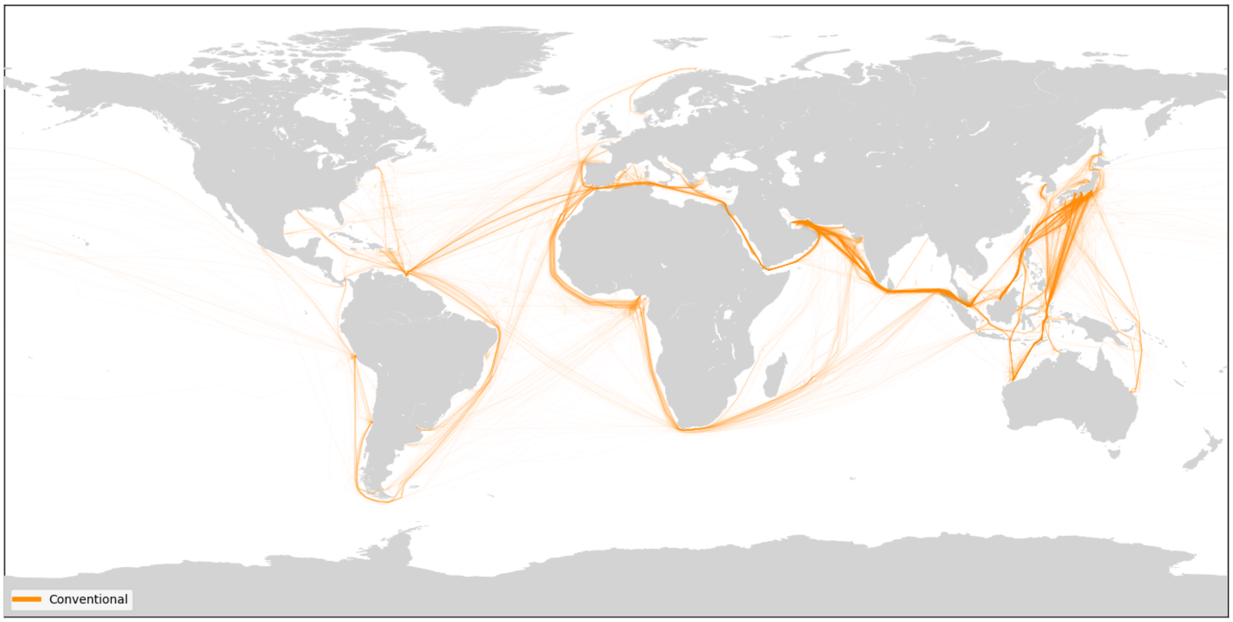


Figure 4.14: Conventional carriers operations, 2017-2018

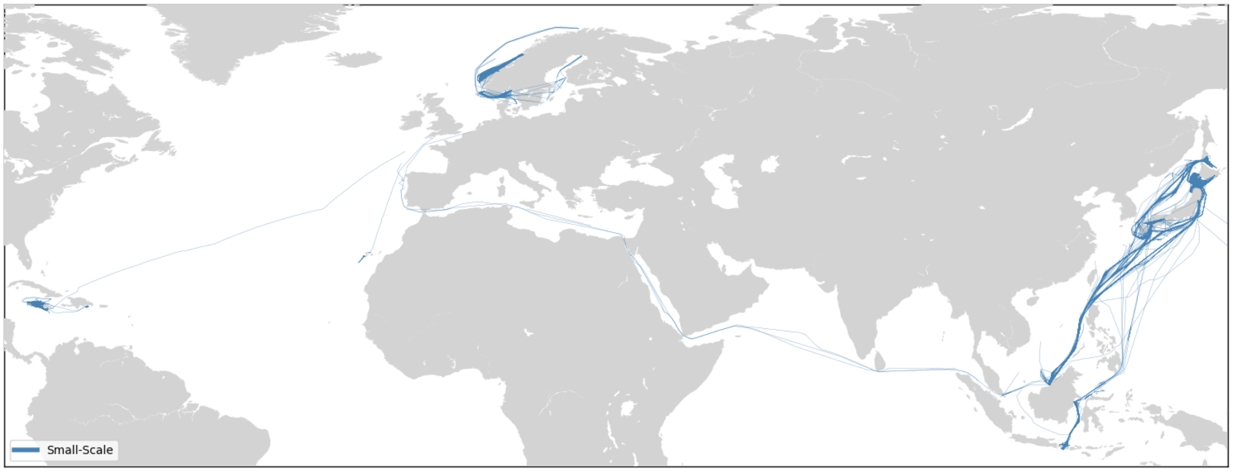


Figure 4.15: Small-scale carriers operations, 2017-2018

4.4.2 Speed Analysis

A brief analysis of the LNG carriers operational speeds have been done with the use of AIS data. In figure 4.16 the development of monthly mean speed is shown categorised by vessel type. The Q-max/Q-Flex vessels generally sail fastest with a mean speed of around 16 knots, while the New-Panamax and Conventional carriers lay relatively stable around 14-15 knots. The small-scale vessels on the other hand have a mean speed of only 10 knots. The speed for all categorise are relatively stable through the period with the drop in April 2018 as an exception. By visual inspection, there is no observable correlation between spot rate and speed, but a visual inspection is of course not sufficient to conclude on the matter. An important notice is that the "Mean Speed" here is the mean SOG from all AIS messages. In other words, the visualisation is not the actual mean speed, rather a good indication.

To recall from section 4.3.3, the majority of Q-Max/Q-Flex, New-Panamax and Conventional carriers have a design speed between 18-20 knots, while Small-scale are in the range of 13-15 knots. Figure 4.17 show a histogram of speed by vessel types for the period of 2017-2018, and the plot gives a lot of information about the operational differences between the different ship types. It can be observed that the mode speed for all vessel types are lower than the design speed, with New-Panamax and Conventional being the closest. It can also be observed that Small-scale vessels stand still over 25% of the time, which is significantly higher than the other vessel categorise. This is both a result of the voyages being much shorter, and the fact that a part of the Small-scale fleet operate as bunkering vessels. A last interesting notice is that while New-Panamax and Conventional operate at a wide range of speeds, the Q-Max/Q-Flex are either standing still, or sailing at 15-18 knots. This indicate that Q-Max/Q-Flex vessels have a stable operational pattern, which is no surprise as they are all on long term contracts. Once again, this is a prime example of industrial shipping.

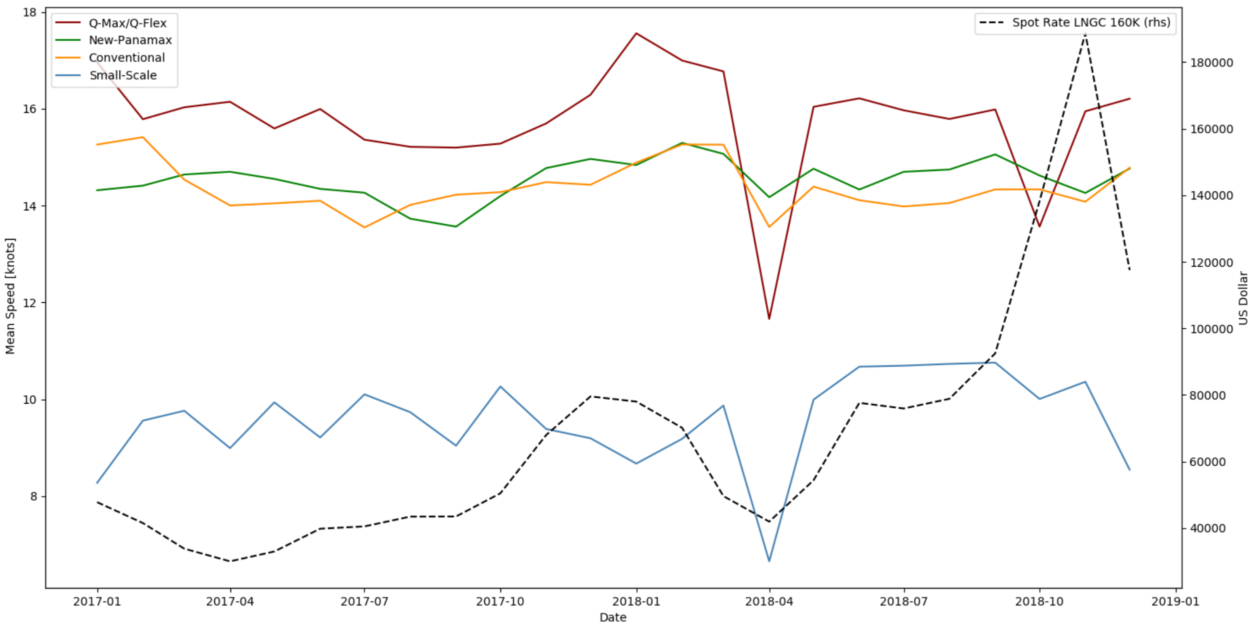


Figure 4.16: Speed monthly by vessel type, 2017-2018. Spot rates obtained from Clarksons Platou

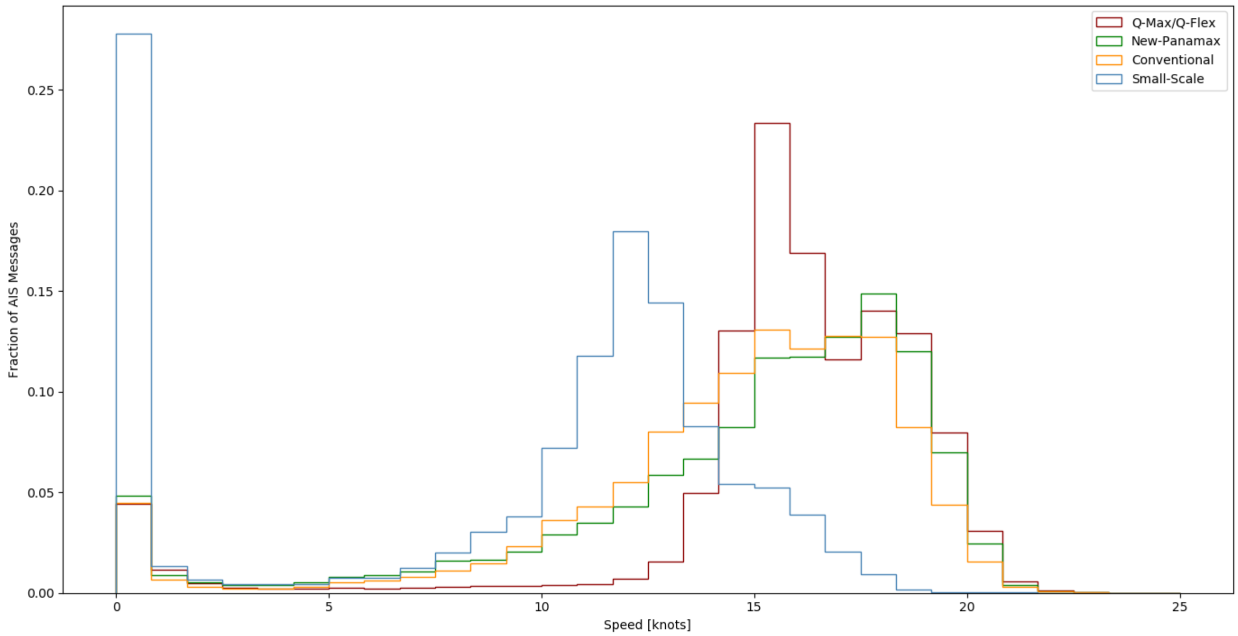


Figure 4.17: Histogram of speed by vessel type, 2017-2018

Chapter 5

Methodology

This chapter presents the main methodology used in order to reach the main objective of this master thesis. The chapter is divided into two main parts, where the first part (section 5.1) describes the methods used to create and prepossess the AIS data. The second part (section 5.2) describes the methodology used to obtain LNG trade flows in or out of terminals. In chapter 6, the methods are put into action and the corresponding results are presented in chapter 7. A methodology for obtaining a complete LNG trade flow network is presented in appendix A, as no case study where conducted on that particular methodology.

5.1 AIS Database Creation

The amount of AIS data is enormous, and the time complexity of every method or algorithm applied on the data will naturally be a function of the amount of data. It is therefore expedient to filter away all unnecessary data that does not meet the purpose of the research/method. In addition, the data contains a significant amounts of erroneous data. In order to conduct reliable analysis in an efficient matter, we want to conduct our analysis on a database containing only data relevant for the study, and with as little erroneous data as possible. A methodology containing five steps, shown in figure 5.1, has been used to create a final database ready for analysis: Data decoding, vessel filtering, message type reduction, data cleaning, and frequency filtering.

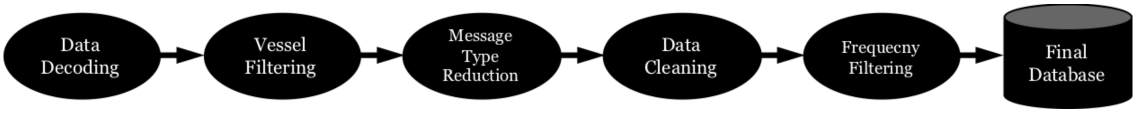


Figure 5.1: Flow chart of LNG Database Creation Methodology

5.1.1 Decoding data

The first step when working with AIS is decoding messages and building a database of the data. A raw AIS message is just a text string containing all the info described in section **??**. It is expedient to get the data on a more readable form in a database, for instance in SQL format. Bjornar Brende Smestad has conducted the decoding of the data used in this thesis. This is done by using an AIS parser and a script utilised in Smestad (2015) and Leonhardsen (2017).

5.1.2 Vessel Filtering

We want to conduct analysis on LNG carriers only, hence we want to filter away all other vessels. By the use of the information in the AIS data alone, we are able to filter away all vessels that are not tankers by filtering on ship type. An external data source is needed to go from all tankers and filter further down to LNG carriers only. We use the Seaweb database from IHS (2019) as data source for LNG carriers. The Seaweb database contain "all" commercial vessels in the world, and allows the user to apply a wide range of filters when searching for vessels. The database contain, but are not limited to IMO number, MMSI, length, breadth, gas capacity, design speed, owner and year built. We want to use the Seaweb database to obtain AIS data containing LNG carriers only. A flowchart of the methodology used to obtain this database can be seen in figure 5.2. From the Seaweb database a complete list of all LNG carriers are obtained. Then a filter based on IMO numbers from the LNG carriers are applied on the AIS message type 5 data, giving us AIS message type 5 data containing only LNG carriers. From this database MMSI numbers from all the LNG carriers are obtained, and a filter based on the MMSI numbers are applied on the dynamic AIS message types 1-3 data. As output we get a AIS database containing only LNG carriers. The final data contain both scrapped, operating and new buildings to hedge us from missing data. Fleet analysis based on the Seaweb database can be found in section 4.3 and the final AIS database based on this method is further introduced in the case study section 6.1.

5.1.3 Message Type Reduction

The AIS Database contains four different message types. These are message type 1,2,3, which are dynamic messages, and 5 which is a static message. The content in type 1-3 is identical, but sent at different criteria. We therefore merge all the dynamic messages to one table, as described in figure 5.3.

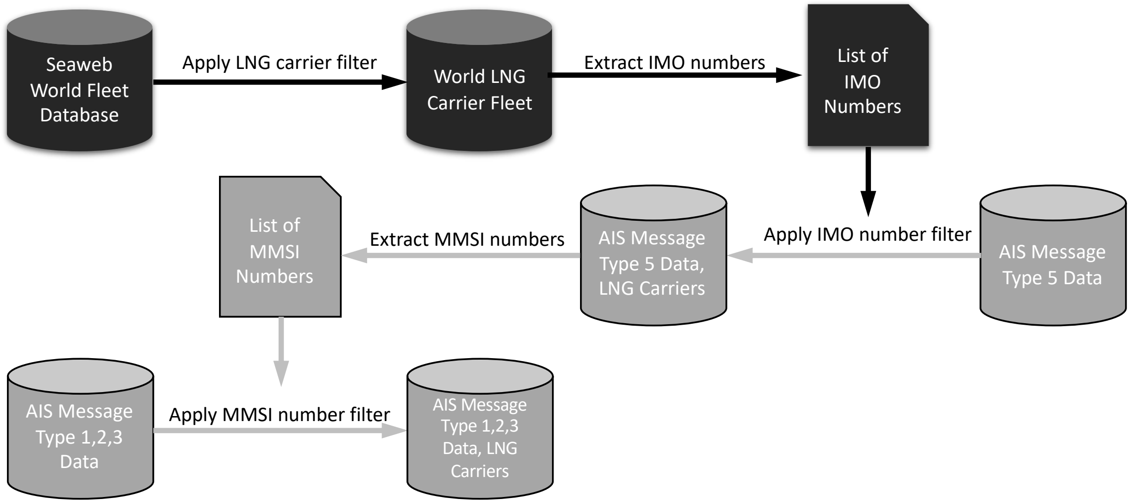


Figure 5.2: Flow chart of LNG carrier search methodology

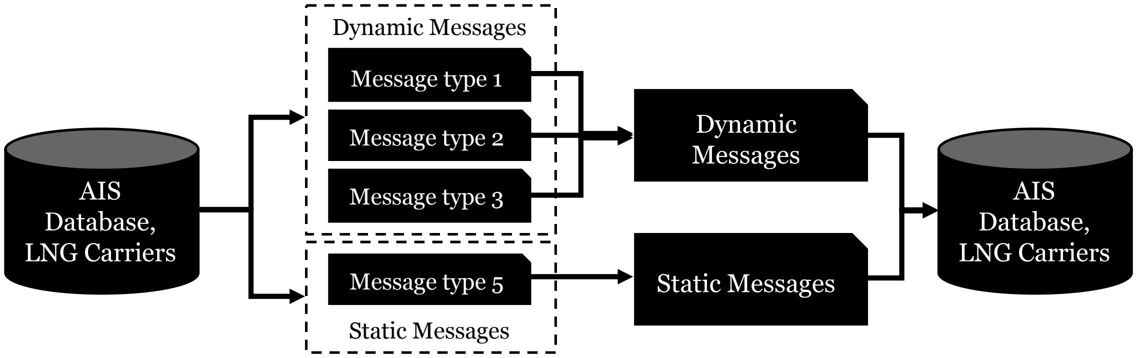


Figure 5.3: Flow chart of message type reduction

5.1.4 Data Cleaning

As described in section 3.1.5, AIS data contain erroneous data. For instance Smestad (2015) identified erroneous IMO numbers, MMSIs, ship dimensions and ship positions. AIS data is also subject to human errors. For instance Næss (2018) revealed that the navigational status, which is set manually is often erroneous. This was checked by comparing navigational status to speed. In this step the database will primarily be checked for erroneous speeds, positions, IMO numbers and MMSI numbers.

5.1.5 Reducing Message Frequency

Vessels send out AIS messages with a frequency based on their speed, change of course and navigational status. High speed and/or fast change of course will give a new message every other second. A message every other second gives us huge amounts of data to process in our analysis, but no additional information gain compared to a frequency of for example 60 seconds, since our main goal is to obtain LNG trade flows in and out of terminals. Therefore the amount of data is reduced by removing all messages sent closer than 60 seconds apart from each other, for each vessel individually. This is primarily done to increase the speed of the analysis conducted, but also to get a more representative picture off for instance the average vessel speed.

5.2 LNG Trade Flow Estimation

Finding LNG trade flow in or out of a port is an important subgoal of this master thesis. We start by looking at pure regasification sites or pure liquefaction sites, meaning port areas with import or export only. Our theory is that trade flow in or out of these areas can be measured by counting the number of unique visits by a vessel with a given gas capacity, and find the sum over all vessels. If the area of interest is a regasification site, we assume import, and export if we are looking at a liquefaction area. The trade flow F in or out of a given port can then be obtained from the following equation:

$$F = \sum_{V} \mu C_v N_v \tag{5.1}$$

Here V is the fleet of vessels v, C_v is the gas capacity of vessel v, N_v is the number of unique visits done by vessel v, and μ is gas capacity utilisation rate. This equation does not fit ports with both large export and import volumes, as it counts the number of unique visits assuming all vessels are either delivering cargo, or collecting cargo. However, the equation can be modified to count number of import visits and export visits if a method to distinguish between vessels unloading cargo and loading cargo is utilised.

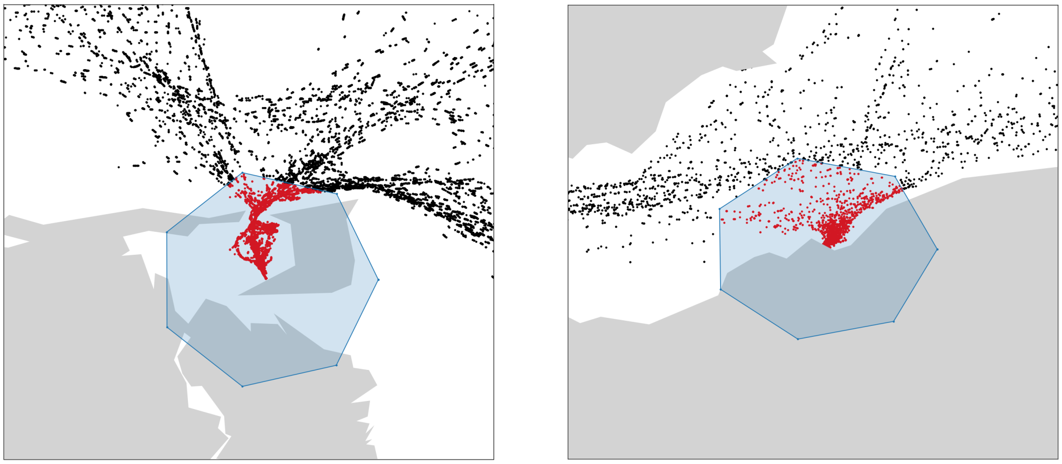
A method containing three main steps has been derived in order to obtain trade flow by using equation 5.1 as origin:

- Step 1: Selecting port area
- Step 2: Deriving unique visits inside port area
- Step 3: Obtaining total LNG import or export volume by summation

5.2.1 Selecting Port Area

The first main step starts with obtaining the coordinates of the port of interest. Then a polygon around the port is defined. The overall methodology is to count the number of unique visits inside this polygon, which is a geo-fence approach. Since we are using S-AIS data, there will be time windows in which the satellites will not get messages from ships within a blind zone.

We therefore face a probability of ships entering and leaving the defined port area without the satellites noticing. A large polygon around the port area will therefore decrease the probability of undetected visits. On the other hand, a large polygon can lead to vessels only passing through the polygon being counted. Especially if the port is located close to a main shipping lane. The optimal polygon size is therefore dependent on the port location as we want the polygon to be large enough to detect all visits, but also small enough to not detect vessels only passing trough. This is exemplified in figure 5.4 where we see the LNG receiving terminal in Trinidad and the LNG export terminal in Algeria with identical polygons surrounding the ports. A visual inspection show that the flow of messages inside the polygon in Trinidad (5.4a) is going towards the port, while a significant part of the vessel flow in Algeria (5.4b) seems to pass by the port. There is also a possibility to add extra criteria to ensure that a vessel is actually loading by limiting maximum speed and demand navigational status to be "anchored" or "moored".



(a) Trinidad LNG receiving terminal

(b) Algeria LNG export terminal

Figure 5.4: Octagon with radius 100 km around port area

5.2.2 Deriving Unique Visits Inside Port Area

The next step after a polygon of expedient size is selected around the port, is to select all unique visits inside the polygon. This step is again divided into two sub-steps which are to tag all AIS messages inside the polygon, and then reduce the selection to only include one message per unique visit. This is called a geo-fence approach.

"Geofencing is a technology that defines a virtual boundary around a real-world geographical area. In doing so, a radius of interest is established that can trigger an action in a geo-enabled

phone or other portable electronic device."[1]

A simple way to check if a message is inside a polygon, is by using the Ray Casting Algorithm, which is also known as the crossing number algorithm or the even-odd rule algorithm. The algorithm is based on two simple observations, and the algorithm is visualised in figure 5.5:

1. Is the location within the x_{max}, x_{min}, y_{max}, y_{min}?
    (a) If yes, continue to step 2.
    (b) If no, the point is not within the geofence (area).
2. Draw a straight line in one direction, and count the number of times the line intersect with the polygon edges.
    (a) If the number of intersections is odd, the location is within the geofence (area).
    (b) If the number of intersections is even, the location is not within the geofence (area).

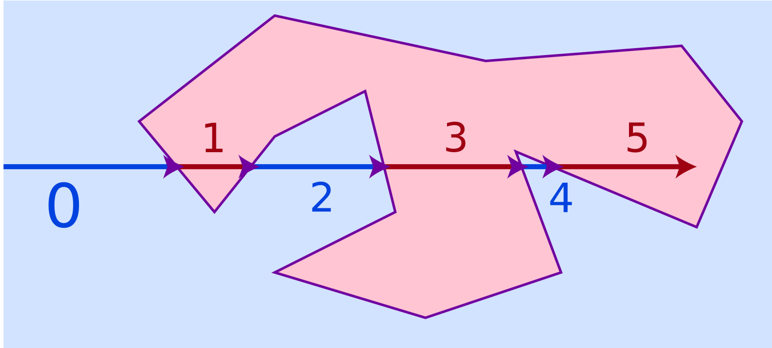


Figure 5.5: Ray Casting Algorithm Visualised[2]

The main strength of this algorithm is that it aborts after step one if the outcome is 1.b. This makes it efficient to use on the large AIS datasets. The algorithm is applied on all AIS messages of type "Dynamic". All messages inside the polygon gets tagged "inside", and all other gets tagged "outside". The tag is saved to a new column in the data frame. Finally, we want to create a port log including only one AIS message per unique visit in the defined polygon. This is done by selecting a data frame of all messages of type "dynamic" for one LNG carrier at a time. Then we look at the "inside/not inside" column and remove all messages where the previous message is of the same inside/not inside category. This gives us only one message categorised as inside per visit. Finally a last criteria is added, saying that unique visits must be at least 24 hours apart

---

[1] https://www.techopedia.com/definition/14937/geofencing

[2] https://en.wikipedia.org/wiki/Point_in_polygon

from each other, as visits any closer is unrealistic. This ensures that we do not double count visits where a vessel goes slightly inside and out again before going in again for loading/unloading. Figure 5.6 show a comparison plot of all messages inside a polygon versus only one per unique visit, and the procedure can be described by the following pseudocode:

Algorithm 1 Obtaining Unique Visits in Port

```
for all LNG carriers do
    for all Messages of type Dynamic except first message do
        if Time since last unique visit by LNG carrier > 24 hours then
            if Message is inside polygon then
                if Previous message is NOT inside polygon then
                    Tag message as new unique visit by LNG carrier
                end if
            end if
        end if
    end for
end for
```

5.2.3 Obtaining Total LNG Import or Export Volume by Summation

A port log containing all unique visits with the corresponding vessel identification MMSI is the final outcome from step 2. The last step is now to add the vessels gas capacity to the port log, making the port log look like table 5.1. The total import or export is now found by taking the sum of the gas capacity column multiplied by a cargo utilisation factor. LNG tanks are normally filled up almost completely (more than 90% and less than 99%) due to sloshing, and emptied keeping a small fraction for heel[3]. In the "Seaweb" database, the gas capacity given is given as the actual gas capacity (97%-99% of available volume), and not the total volume available. Therefore we will set our cargo utilisation factor to 1. As a consequence it is expected that the estimated trade flows can get higher than the actual trade flow and we are per definition measuring the visited gas capacity.

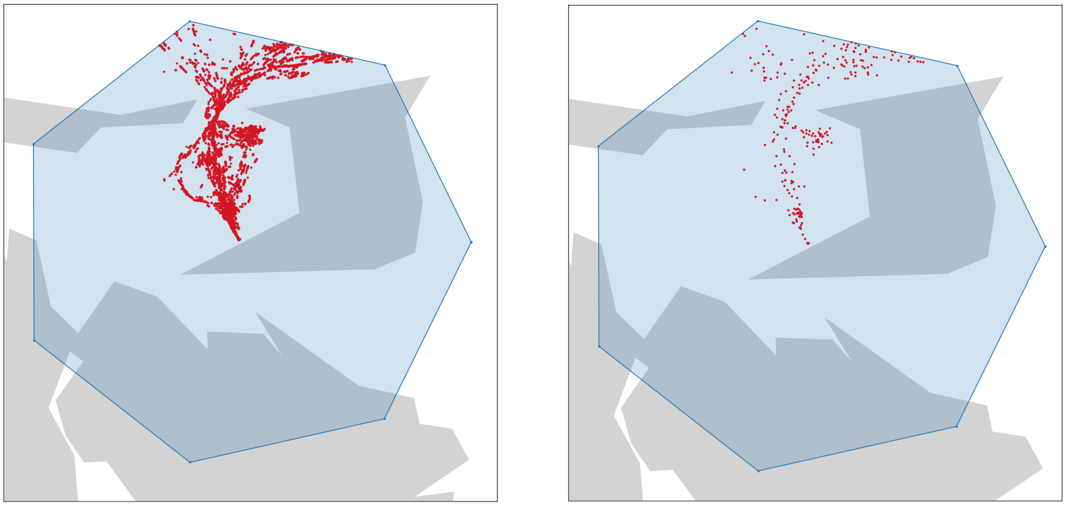
The method is tested in case study 6.2 , with results and corresponding discussion in chapter 7. The methodologies strengths, weaknesses and improvement possibilities are discussed in chapter 8.

---

[3] The small amount of LNG remaining on board a vessel (or storage) after discharge of the regular LNG cargo used to keep tanks cooled and for fuel.

Table 5.1: Port log example from single port analysis

	Date Time	**MMSI**	**Inside**	**Gas Capacity**$[m^3]$
0	2015-01-02 13:38:50	CENSORED	1	135298
1	2015-04-09 01:55:33	CENSORED	1	135298
2	2015-04-24 16:09:05	CENSORED	1	135298
...				
256	2015-11-26 09:05:4	CENSORED	1	152004
257	2015-12-21 13:15:40	CENSORED	1	152004



(a) Trinidad all messages inside polygon

(b) Trinidad unique visits

Figure 5.6: All messages inside polygon versus unique visits

5.2.4 Error Estimation

An error estimation method is used in order to compare results between different methods, different years and between different LNG terminals. To compare the performance between different terminals or between different years absolute percentage error is used:

$$PE = \left| \frac{MV - AV}{AV} \right| \tag{5.2}$$

Where PE is the percentage error, MV is the measured value, and AV is the actual value.

When different methods are compared, the weighted average of PE is taken for all terminals T:

$$WAPE = \frac{\sum_T F_t PE_t}{\sum_T F_t} \tag{5.3}$$

Where WAPE is the weighted absolute percentage error, F_t is the actual tradeflow at terminal t, and PE_t is the percentage error at terminal t.

Chapter 6

Case Study

In this chapter the conducted case study, which consist of two main parts is presented. The first part is the creation and preprocessing of the database, while the second part is the trade flow estimation with the use of AIS data.

6.1 Creating The Database and Initial Analysis

A decoded database consisting LNG carriers in a time-window from the start of 2017 to the end of 2018 was provided by The Norwegian Coastal Administration. Bjørnar Brende Smestad decoded the data and applied the LNG vessel search. This section will walk through the steps described in section 5.1 and visualise the changes made to the database from the initial database to the final database used in the main analysis conducted in this thesis.

Table 6.1 shows the number of messages in the initial database. The amount of messages have approximately doubled from 2017 to 2018.

Table 6.1: Initial database, number of messages

Year	message type 1	message type 2	message type 3	message type 5
2017	10 567 944	1 898	367 241	854 138
2018	19 625 063	5 234	699 388	2 587 973

From the initial database we merge all messages of type 1-3 into one combined message type "Dynamic messages". Next, we analyse the data for erroneous data by looking at speed and position. LNG tankers in general never have a top speed over 25 knots, but we round up and define all messages with speeds over 30 knots ass erroneous. When it comes to erroneous positions, we define all messages with invalid coordinates as erroneous without further inspection.

It is found that over 10% of the messages have a speed over 30 knots, and of these 99% have a speed of 102 knots. This data is obviously error data. Further investigation show that 159 vessels have erroneous data and that 15 vessels have more than 50% of their messages tagged as erroneous. The erroneous messages are sent in packages of 100+ messages with exact same date stamp. Less than 0.005% of the data have invalid position coordinates, and of these, the majority has already been marked erroneous by speed. Figure 6.1 shows a plot of the number of erroneous messages for every month in the time period. The plot shows that the large amount of erroneous data started simultaneously as the Norwegian Costal Authorities launched two additional AIS satellites. The plot also indicate that the fraction of erroneous data increase with increased number of messages.

The trade flow analysis methodology relies on consistent data for all vessels. The high amount of erroneous data could lead to large data gaps which could result in vessels loading or unloading LNG undetected.

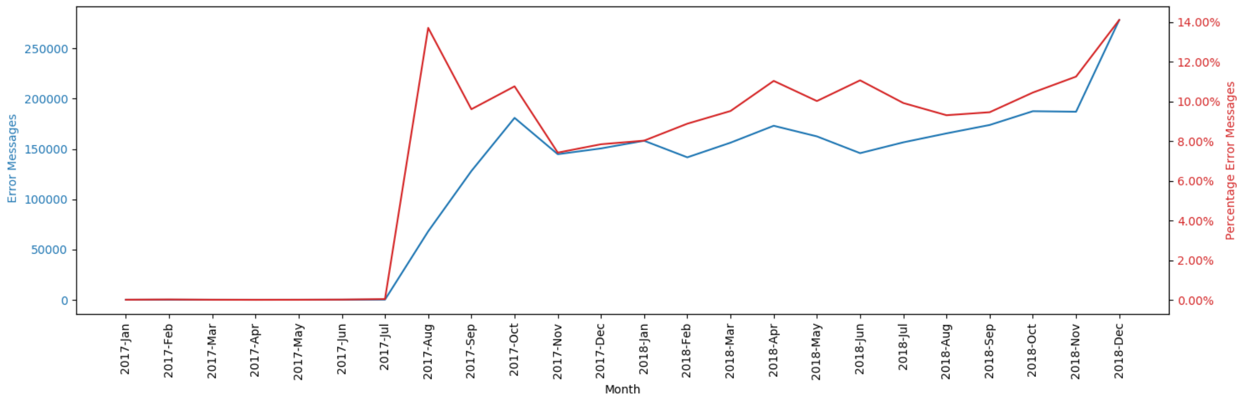


Figure 6.1: Monthly number of erroneous messages

Message frequency reduction is the last step before the database is ready. There are two reasons to why we do this. The first is about speeding up the algorithms by reducing the amount of data to be processed. The second reason is related to the speed analysis. AIS data is sent with a higher frequency when vessels sail fast. Therefore a vessel at a high speed will give more datapoint than a vessel on low speed. This can affect the results in the speed analysis giving unrealistic high average speed values for instance. Our analysis does not depend on knowing the vessels change in position every other second, hence we remove all messages sent with a frequency higher than one message per 60 seconds.

Figure 6.2 shows the amount of data or messages initially, after error removal, and the final data after frequency reduction. We can observe that the amount of monthly data after August 2017 is approximately three times the amount before. This is a result of the two additional satellites launched by the Norwegian Coastal Administration. It is also observable that the final database is less than 50% of the original, which will speed up the analysis significantly. There has obviously also been some startup trouble with the two new satellites. This is clearly shown in figure 6.3, which is a plot of weekly observed MMSIs. It can be seen that the observed MMSIs dropped close to zero around the launch. It is also extremely important to notice that out of the 521 unique MMSIs in our dataset, less than 400 is observed on a weekly basis. This can potentially reduce the performance of the trade flow analysis drastically and it is a key weakness with the database. Figure 6.4 show the number of large message gaps from the vessels. On average each vessel have more than 30 gaps of over one day, and around five gaps of more than seven days. This underlines the weakness further, although there was a small improvement in 2018.

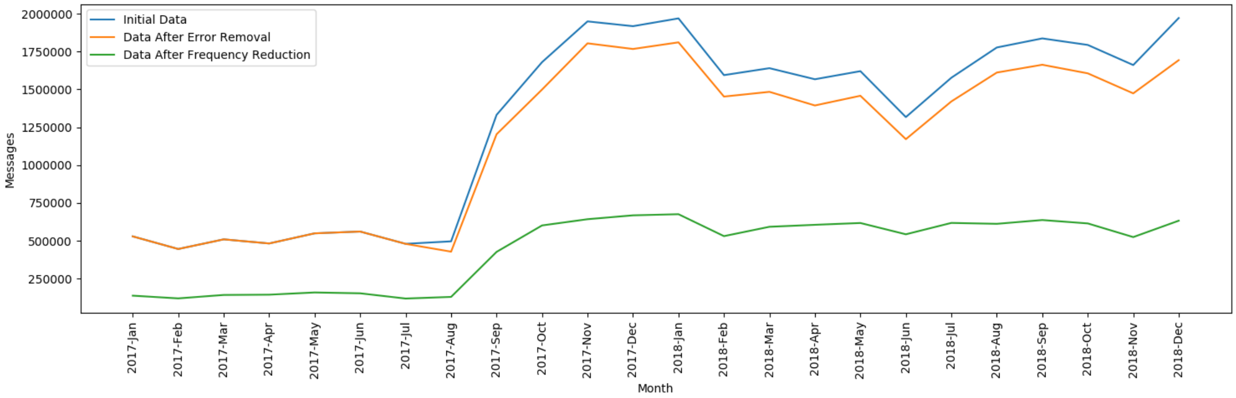


Figure 6.2: Monthly messages

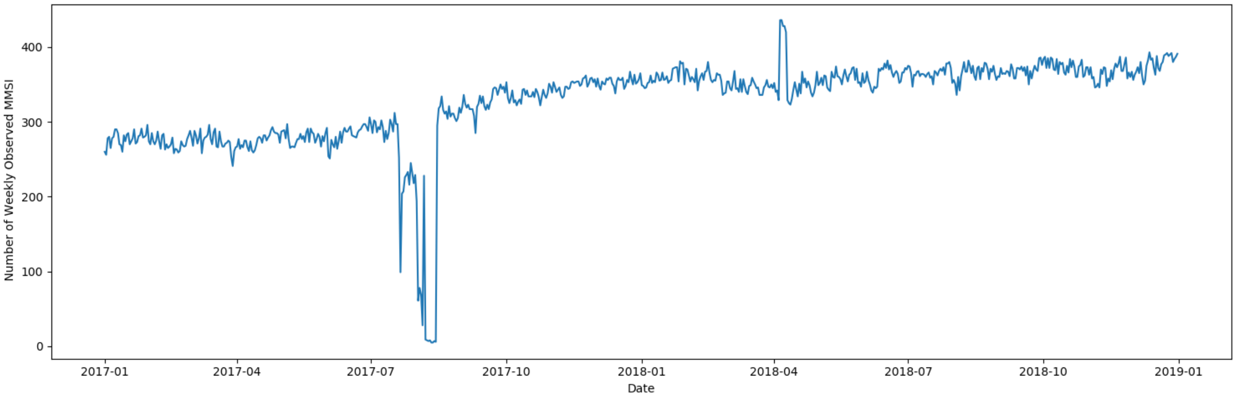


Figure 6.3: Weekly MMSIs observed

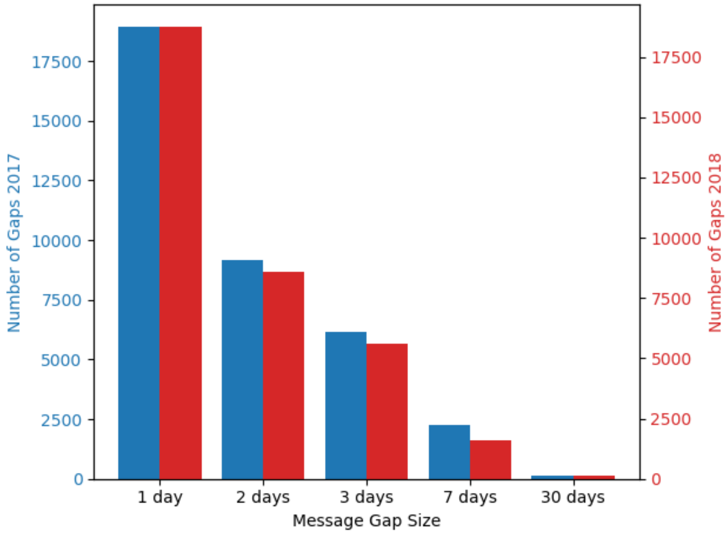


Figure 6.4: Message gaps

6.2 LNG Trade Flow Estimation

In this section we apply the methodology described in section 5.2 in order to analyse LNG trade in or out of a single geo-fence at a time. The method will be conducted on eight LNG exporters consisting of totally ten export terminals, and six LNG importers consisting of fourteen import terminals. Three different geo-fence sizes will be tested for each terminal, which we refer to as "Diameter 1", "Diameter 2", and "Diameter 3". These will have a diameter of 5 km, 15 km and 50 km respectively. A fourth case will use the "Diameter 1" geo-fence, but add two additional conditions: Vessel speed less than two knots and navigational status anchored or moored. This case will be referred to as "Diameter 1 LSMA" [1]. Figure 6.5 visualise the four test cases. All cases will be tested for both 2017 and 2018 and benchmarked against IGU (2019).

Both the importers and exporters have been picked to get various geographical position, different trade volumes and different complexity in terms of nearby shipping activity. The countries and the associated terminals can be seen in table 6.2. Here, the coordinates of the terminal and the import/export benchmark volume is given. The terminals has also been classified by how remote or central they are, as well as by their nearby LNG shipping activity. Three degrees are set for both categories: Central, Semi Central and Remote, and High, Medium and Low nearby LNG shipping traffic. Each node is classified based on visual inspection of activity inside and nearby the geo-fences. Figure 6.6 exemplifies this, while all the case study relevant terminals are plotted in figure 6.7. The results are presented and discussed in chapter 7.

---

[1] Low Speed Anchored or Moored

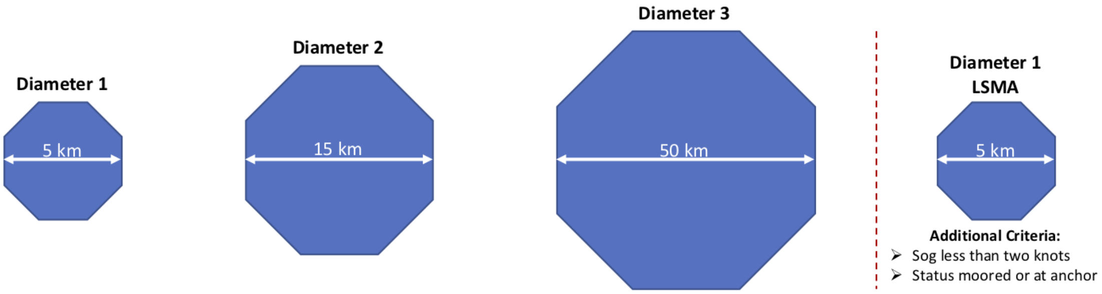


Figure 6.5: Geo-fence size, four test cases

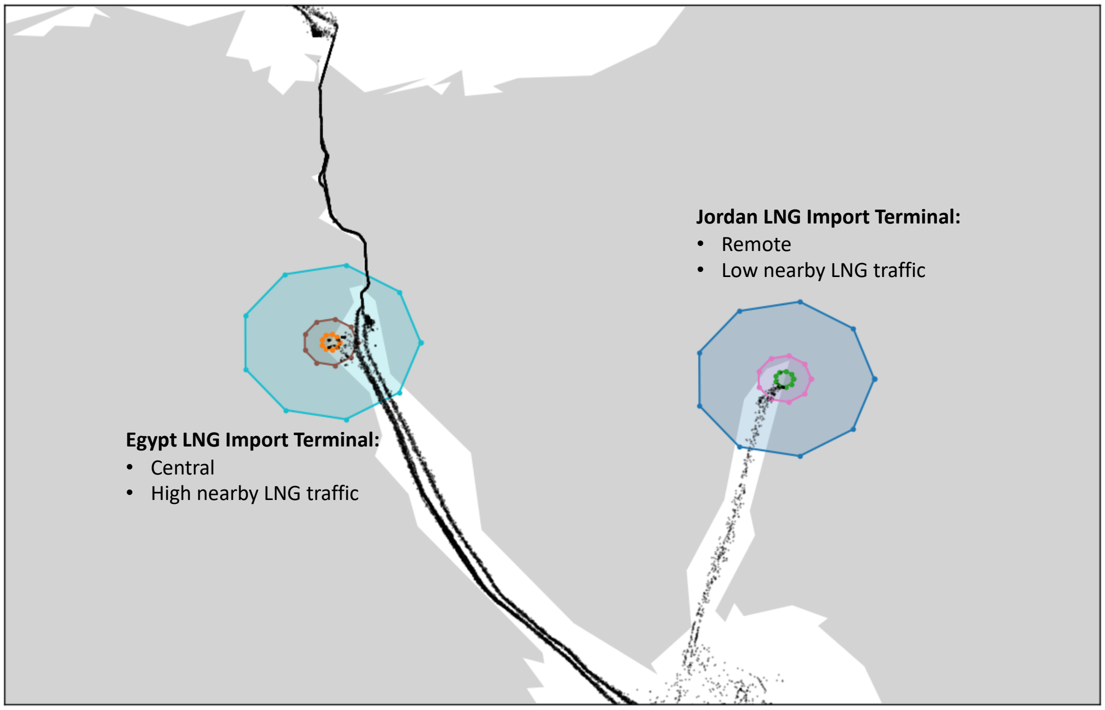


Figure 6.6: Central and remote terminals exemplified with Egypt and Jordan

Table 6.2: LNG import and export terminals investigated in single port case study

Node	**Country**	**Terminal**	**Latitude**	**Longitude**	**Type**	**Trade Flow 2017* [MTPA]**	**Trade Flow 2018* [MTPA]**	**Location Type**	**LNG Traffic**
1	Algeria	Skikda	36.9	6.941	Export	12.17	10.3	Central	High
2	Algeria	Arzew	35.83	-0.24	Export			Central	High
3	Angola	Angola LNG	-6.119	12.335	Export	3.67	4.14	Remote	Low
4	Cameroon	Kribi	3.017	9.836	Export		0.61	Semi Central	Medium
5	Norway	Snøhvit LNG	70.652	23.52	Export	4.04	4.63	Remote	Low
6	Qatar	RasGas I	25.93	51.5891	Export	76.71	78.69	Central	High
7	Russia	Sakhalin	46.615	142.911	Export	11.11	18.93	Remote	Low
8	Russia	Yamal LNG	71.273	72.072	Export			Remote	Low
9	Trinidad	Atlantic LNG	10.2	-61.7	Export	10.76	12.23	Remote	Low
10	United Arab Emirates	ADGAS	25.16	52.88	Export	5.2	5.48	Central	Medium
11	Egypt	Sumed BW	29.631	32.37	Import	5.97	2.26	Central	High
12	France	Dunkirk	51.0336	2.19653	Import	7.58	8.43	Central	Medium
13	France	Fos Cavaou	43.405	4.911	Import			Remote	Low
14	France	Montoir-de-Bretagne	47.31	-2.142	Import			Remote	Low
15	India	Hazira	21.098	72.624	Import	19.3	23.26	Remote	Low
16	India	Ratnagiri	17.54	73.16	Import			Semi Central	Medium
17	India	Jaigarh	17.304	73.206	Import			Semi Central	Medium
18	India	Ennore LNG	13.278	80.33	Import			Remote	Low
19	India	Kochi	9.978	76.226	Import			Central	high
20	India	Mundra	22.774	69.683	Import			Remote	Low
21	India	Dahej	21.674	72.536	Import			Remote	Low
22	Jordan	Aqaba	29.42	34.97	Import	3.36	2.55	Remote	Low
23	Kuwait	Mina Al-Ahmadi	29.083	48.16	Import	3.55	3.48	Remote	Low
24	Pakistan	PGPC Port Qasim	24.766	67.313	Import	4.74	7.15	Remote	Low

*Total import/export for given country, not single terminal, data from IGU (2019)

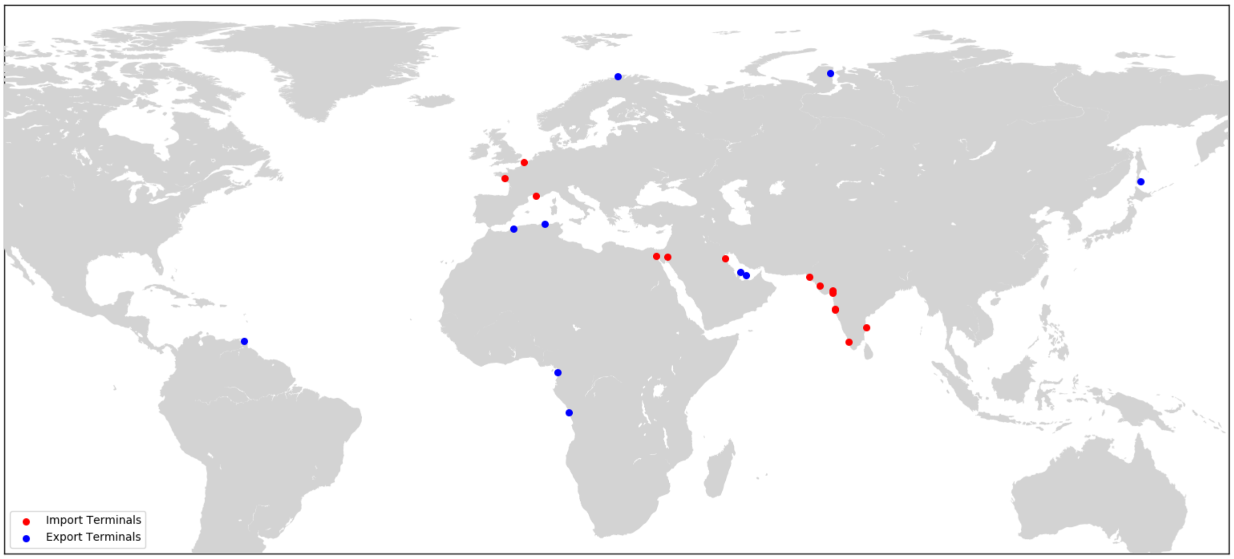


Figure 6.7: Terminals investigated in the LNG trade flow case study

Chapter 7

Results

This chapter presents the results obtained from the LNG trade estimation case study from chapter 6, section 6.2. The results are presented with an associated discussion country by country starting with the exporting countries. Next, a general comparison of the different methods are presented, and the chapter ends with a summary of the findings.

As a reminder, the four different geo-fences and criteria where as follows:

- Diameter 1: Geo-fence with a diameter of 5 km
- Diameter 2: Geo-fence with a diameter of 15 km
- Diameter 3: Geo-fence with a diameter of 50 km
- Diameter 1 LSMA: Geo-fence with a diameter of 5 km and low speed moored/anchored criteria

All the results are benchmarked against import and export values obtained from the International Gas Union's (IGU) annual LNG report, IGU (2019). There is no guarantee that the benchmark values are 100% correct, but in the presentation of the results with corresponding discussion this is assumed. Lastly, three additional assumptions are important to keep in mind when reading the results:

1. The results are presented as "visited gas capacity", which means that all vessels are assumed to be fully loaded. Fully loaded means 97%-99% of the tank volume, but the vessel database source (IHS (2019)) has already subtracted the 1%-3% in their presentation of gas capacities. Hence, their numbers are used directly.
2. A LNG density of 0.41 $tonnes/m^3$ has been used when converting LNG from m^3 to MTPA, although LNG density do vary based on composition.

3. The results for the diameter 1 LSMA are completely off target, and will not be discussed in the country by country presentation of results.

7.1 Exporting Countries

7.1.1 Algeria

The results for Algeria give an estimated trade flow between 40% and 138% of the benchmark value, where the measured trade flow clearly increase with increased diameter. The trade flow estimation also increase from 2017 to 2018 as a result of the much higher data amount in 2018. Both Algerian terminals are classified as central with high amount of near-going traffic. The large overshoot in the trade flow estimation for diameter 3 is a direct result of this, as the largest geo-fence interfere with a shipping lane. This is visualised in figure 7.2. Diameter 1 and diameter 2 on the other hand should not be be affected by passing by traffic. Therefore diameter 2 for 2018 with 112% of benchmark should be a good estimate with large enough diameter and enough data to detect close to all visits without including pass through vessels. The small overshoot is then either caused by part of the LNG carriers not filling all cargo tanks and/or varying LNG density.

Table 7.1: Algeria, estimations

Algeria	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSMA**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	12.17	100 %	4.92	40 %	9.26	76 %	13.24	109 %	0.86	7 %
2018	10.3	100 %	6.99	68 %	11.53	112 %	14.26	138 %	1.20	12 %

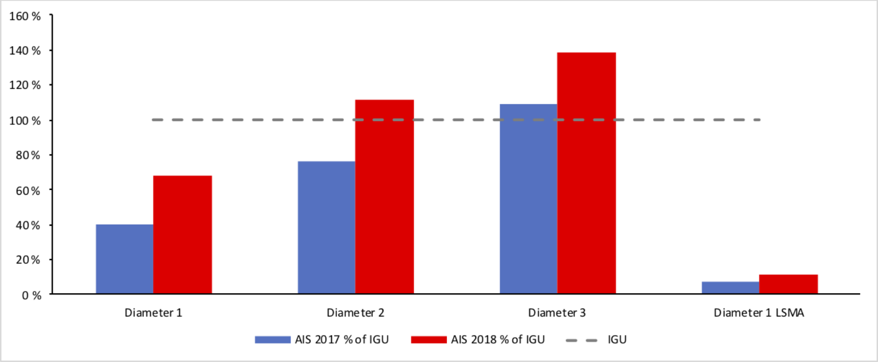


Figure 7.1: Algeria LNG export based on AIS data

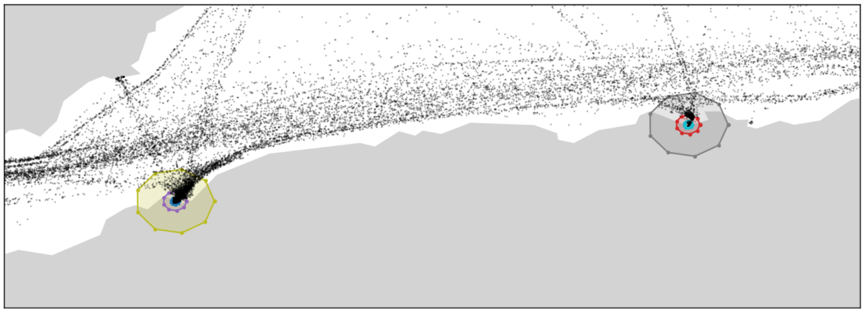


Figure 7.2: Algeria LNG export terminals

7.1.2 Angola

The results for Angola is between 71% and 86% of benchmark when looking at 2017, and between 93% and 96% when looking at 2018, which imply increased accuracy with increased data amount. The terminal is classified as remote with low nearby traffic, and figure 7.4 show how all shipping lanes inside the geo-fences go in to the actual terminal spot, excluding the possibility of carriers passing by affecting the results.

Table 7.2: Angola, estimations

Angola	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSMA**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	3.67	100 %	2.59	71 %	2.79	76 %	3.17	86 %	0.00	0 %
2018	4.14	100 %	3.85	93 %	3.85	93 %	3.98	96 %	0.00	0 %

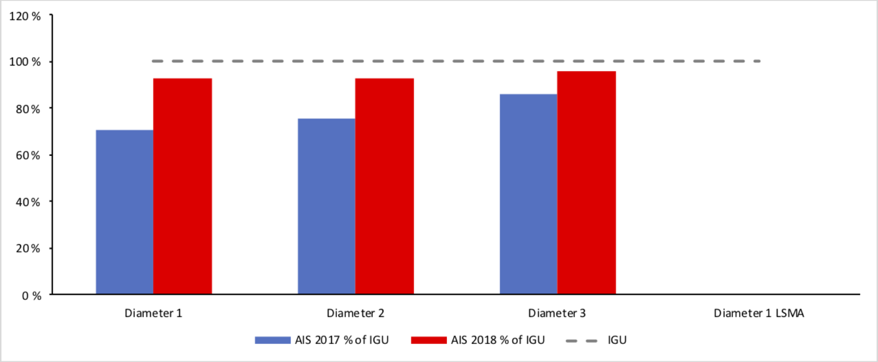


Figure 7.3: Angola LNG export based on AIS data

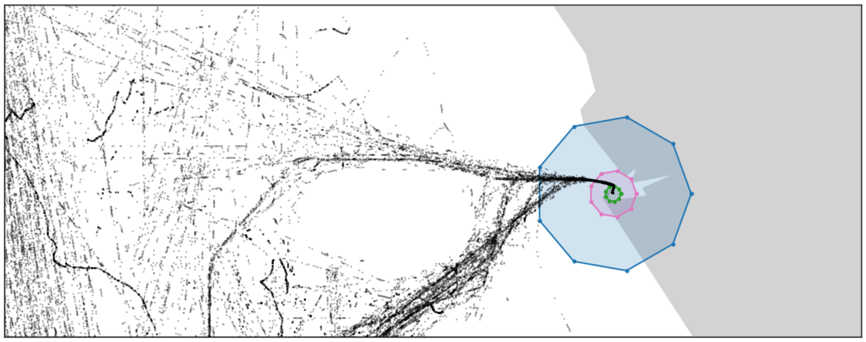


Figure 7.4: Angola LNG export terminal

7.1.3 Cameroon

Cameroon started with LNG export in 2018 after the startup of the floating liquefaction vessel Golar Hilli. The results are all overshooting from 195% to 252%. The terminal is classified as semi central with medium nearby traffic amount, due to the high LNG activity in neighbouring country, Nigeria. The high overshoot is unexpected, and especially that the smallest diameter gives the highest overshoot. From figure 7.6 we can observe that there is activity passing by in the largest diameter, while the activity in diameter 1 is high, and there could be vessels circling around the area giving many vessel detections. This could also be related to the startup of the facility, and we should therefore be careful about putting to much into these results. Lastly, it is also important to keep in mind that this facility has by far the lowest export volume. This means that one wrongly counted vessel will give a much higher impact than for a terminal with higher export volume, since we are looking at percentages. An export of 0.61 MTPA is equivalent to only approximately 9 New Panamax carriers.

Table 7.3: Cameroon, estimations

Cameroon	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSMA**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	0	100 %	0.06	NA	0.06	NA	0.06	NA	0.06	NA
2018	0.61	100 %	1.54	252 %	1.41	231 %	1.19	195 %	0.53	87 %

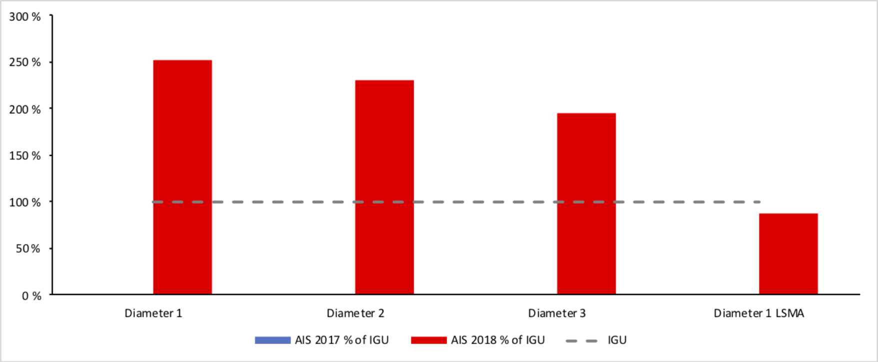
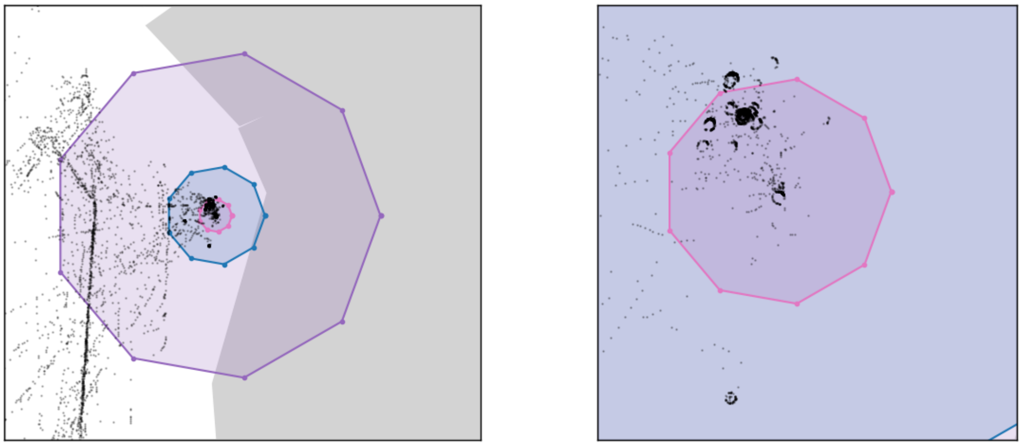


Figure 7.5: Cameroon LNG export based on AIS data



(a) All geo-fences (b) Diameter 1 geo-fence

Figure 7.6: Cameroon LNG export terminal

7.1.4 Norway

The results for Norway is between 98% and 105% of benchmark for 2017 and between 95% and 100% of benchmark for 2018. The remote position of the terminal (Melkeøya) in combination with low nearby LNG traffic makes the estimation for all three diameters come very close to the benchmark. An interesting observation is that the smallest diameter give the highest estimation for this terminal. Figure 7.8 show high activity at the outer parameter of diameter 1. This can potentially make the algorithm double count vessels, if the entering is more than 24 hours later than last entry.

Table 7.4: Norway, estimations

Norway	IGU		Diameter 1		Diameter 2		Diameter 3		Diameter 1 LSMA	
	MTPA	% of IGU	MTPA	% of IGU	MTPA	% of IGU	MTPA	% of IGU	MTPA	% of IGU
2017	4.04	100 %	4.22	105 %	3.77	93 %	3.95	98 %	1.46	36 %
2018	4.63	100 %	4.39	95 %	4.32	93 %	4.64	100 %	0.72	16 %

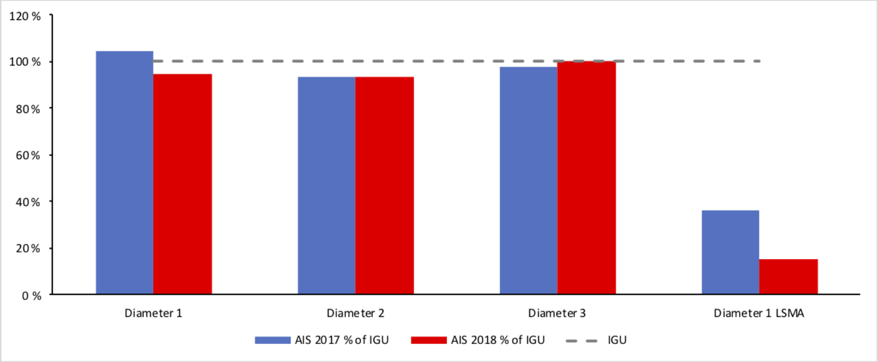


Figure 7.7: Norway LNG export based on AIS data

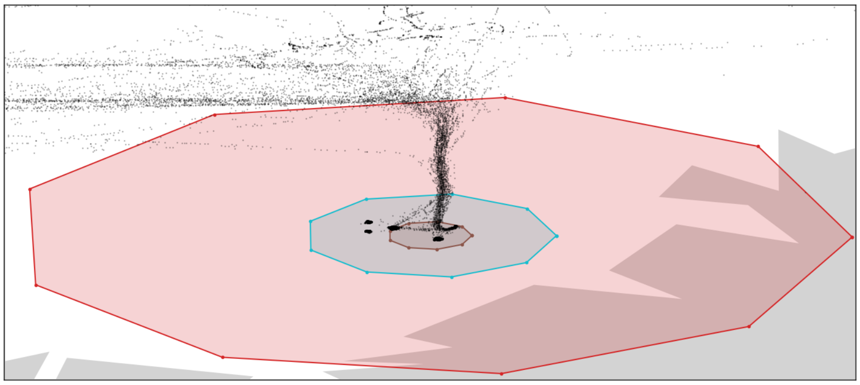


Figure 7.8: Norway LNG export terminal

7.1.5 Qatar

The estimation results for Qatar goes from 19% to 82% of benchmark value where 2018 value in combination with the largest diameter is by far the best estimate. In other words, estimations for this terminal perform significantly better with a larger amount of AIS data collected. Figure 7.10 show that much activity is captured with diameter 3 that is not captured with diameter 1 and diameter 2. This is because qatar have multiple export terminals placed close to each other, while the analysis have only accounted for one terminal. As a result, the diameter 3 is the only geo-fence covering all the terminals.

Table 7.5: Qatar, estimations

Qatar	IGU		Diameter 1		Diameter 2		Diameter 3		Diameter 1 LSM	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	76.71	100 %	14.56	19 %	22.80	30 %	43.10	56 %	0.11	0 %
2018	78.69	100 %	24.19	31 %	36.10	46 %	63.42	81 %	0.17	0 %

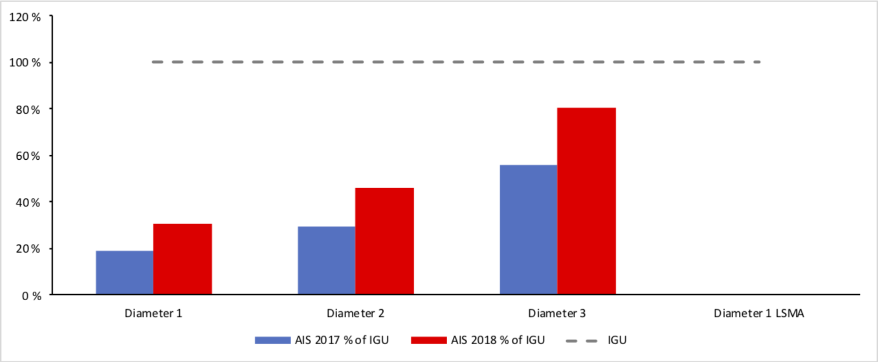


Figure 7.9: Qatar LNG export based on AIS data

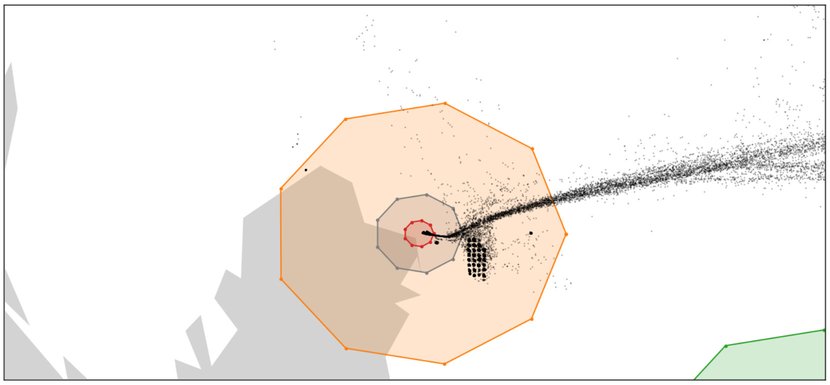


Figure 7.10: Qatar LNG export terminals

7.1.6 Russia

The results for Russia are promising with estimations in between 47% and 104% of the benchmark value. Both terminals are remotely placed and close to all the nearby traffic goes into the terminals, as shown in figure 7.12. The estimations are best for the large diameters, but the small diameters perform much better for 2018 due to the higher data resolution.

Table 7.6: Russia, estimations

Russia	IGU		Diameter 1		Diameter 2		Diameter 3		Diameter 1 LSM	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	11.11	100 %	5.21	47 %	6.87	62 %	11.19	101 %	0.07	1 %
2018	18.93	100 %	16.63	88 %	17.92	95 %	19.68	104 %	0.00	0 %

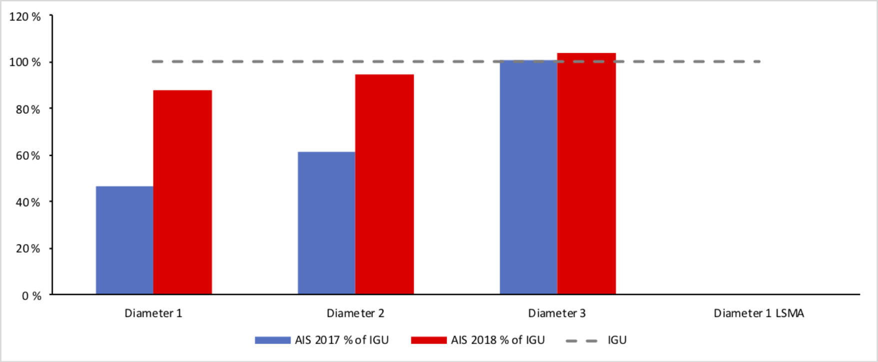
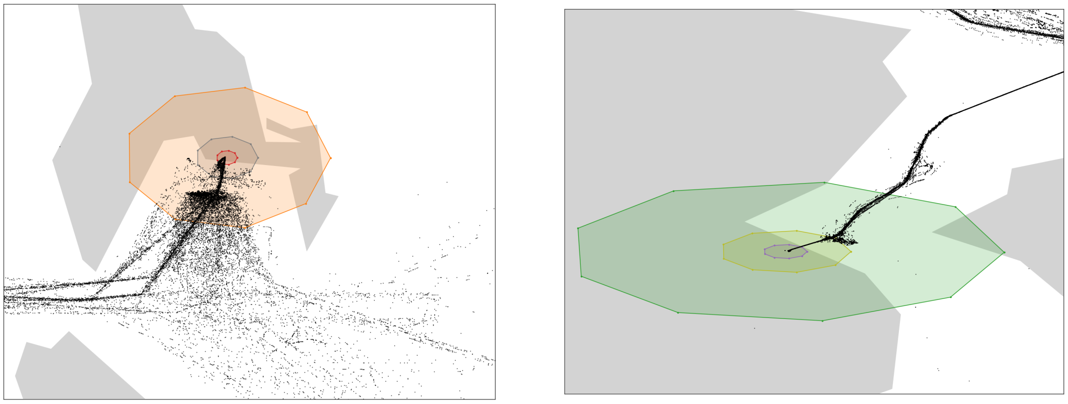


Figure 7.11: Russia LNG export based on AIS data



(a) Sakhalin LNG export terminal

(b) Yamal LNG export terminal

Figure 7.12: Russia LNG export terminals

7.1.7 Trinidad

As for Russia, Trinidad have promising results with estimations in between 73% and 115%. With its remote position in a bay, the terminal have very little probability of false detections. The results for 2018 lay higher due to data resolution, and the overshoot of 15% indicate that not all LNG carriers that entered Trinidad where fully loaded.

Table 7.7: Trinidad, estimations

Trinidad	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSM**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	10.76	100 %	7.82	73 %	9.81	91 %	11.77	109 %	0.17	2 %
2018	12.23	100 %	12.32	101 %	12.62	103 %	14.12	115 %	0.06	1 %

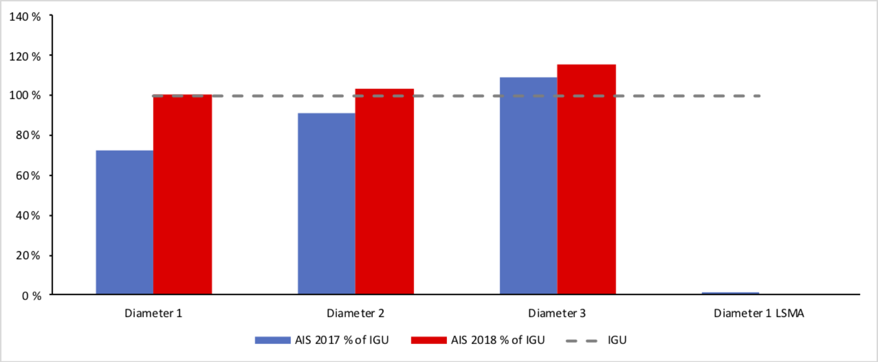


Figure 7.13: Trinidad LNG export based on AIS data

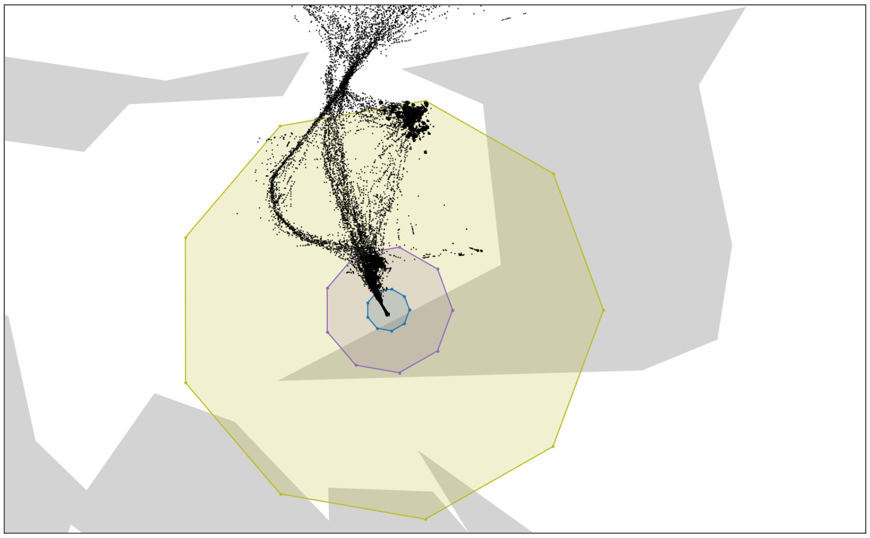


Figure 7.14: Trinidad LNG export terminal

7.1.8 United Arab Emirates

Estimates for the United Arab Emirates are poor with estimates between 16% to 41%. As with Cameroon, the production at this terminal is very low, making the impact one undetected ship significant. The nearby LNG traffic is categorised as medium, due to Qatar being nearest neighbour.

Table 7.8: United Arab Emirates, estimations

UAE	IGU		Diameter 1		Diameter 2		Diameter 3		Diameter 1 LSM	
	MTPA	% of IGU	MTPA	% of IGU	MTPA	% of IGU	MTPA	% of IGU	MTPA	% of IGU
2017	5.2	100 %	0.81	16 %	1.42	27 %	1.92	37 %	0.00	0 %
2018	5.48	100 %	1.17	21 %	1.84	34 %	2.23	41 %	0.00	0 %

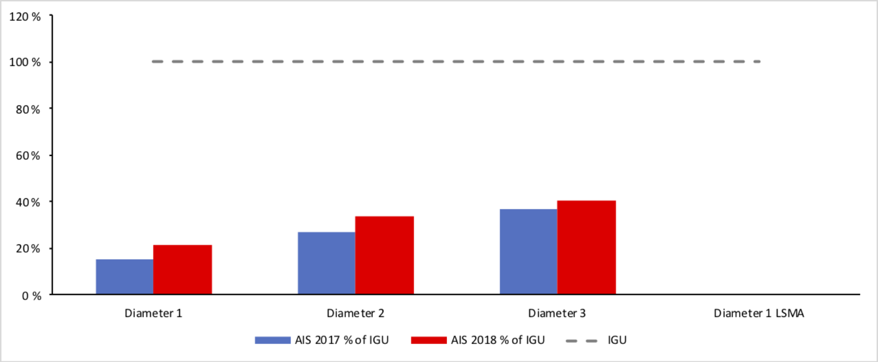


Figure 7.15: United Arab Emirates LNG export based on AIS data

7.2 Importing Countries

7.2.1 Egypt

Egypt with its close position to the Suez Canal is categorised as central and with high nearby traffic. The estimations perform from 24% to 1482% of benchmark value. The extraordinary high overshoot is easily explained by looking at figure 7.17 where we can observe that diameter 3 potentially counts all vessels passing by the Suez Canal, while diameter 2 also partly covers the Suez Canal. Diameter 1 have disappointingly low results with 24-25% of benchmark.

Table 7.9: Egypt, estimations

Egypt	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSMA**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	5.97	100 %	1.45	24 %	2.76	46 %	17.46	292 %	0.00	0 %
2018	2.26	100 %	0.56	25 %	3.09	137 %	33.49	1482 %	0.00	0 %

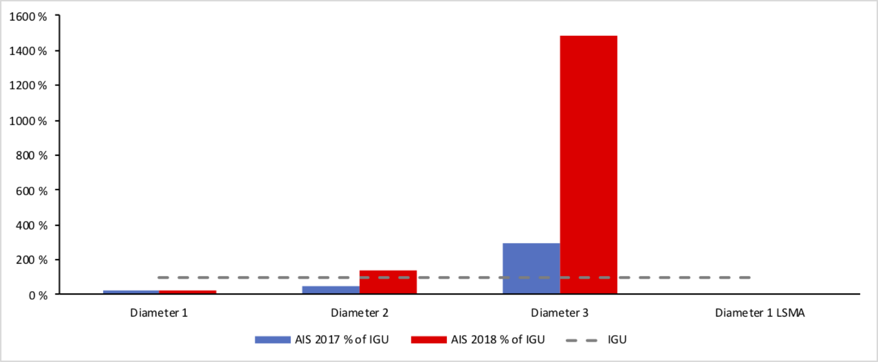


Figure 7.16: Egypt LNG import based on AIS data

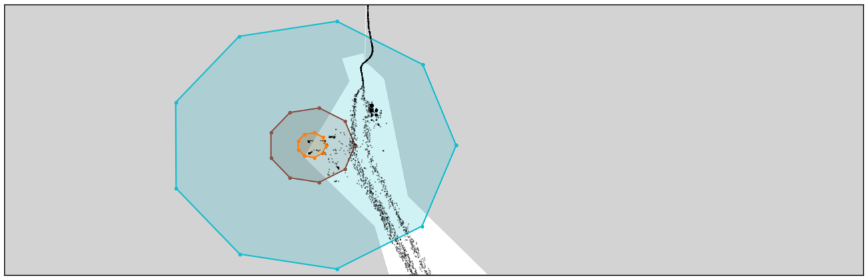


Figure 7.17: Egypt LNG import terminal

7.2.2 France

Estimations on France LNG imports spike from 8% to 69%. France have three LNG import terminals, where 2/3 are classified as remote with low nearby traffic, while the third is classified as central with medium nearby traffic due to a close by shipping lane, seen in figure 7.19. The estimations for France are weak, and a main reason is the low import volume spread between three terminals, which means that an undetected vessel gives a high impact on the results. In addition the satellite coverage of these areas are not optimal according to Eriksen et al. (2018).

Table 7.10: France, estimations

France	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSMA**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	7.58	100 %	0.59	8 %	1.04	14 %	1.66	22 %	0.00	0 %
2018	8.43	100 %	2.66	32 %	3.21	38 %	5.81	69 %	0.06	1 %

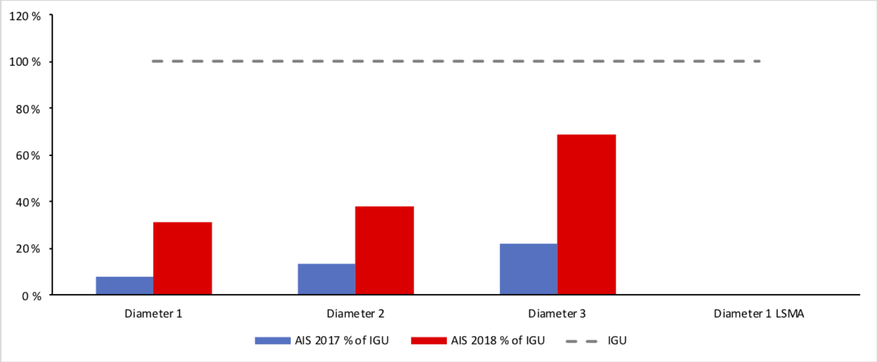


Figure 7.18: France LNG import based on AIS data

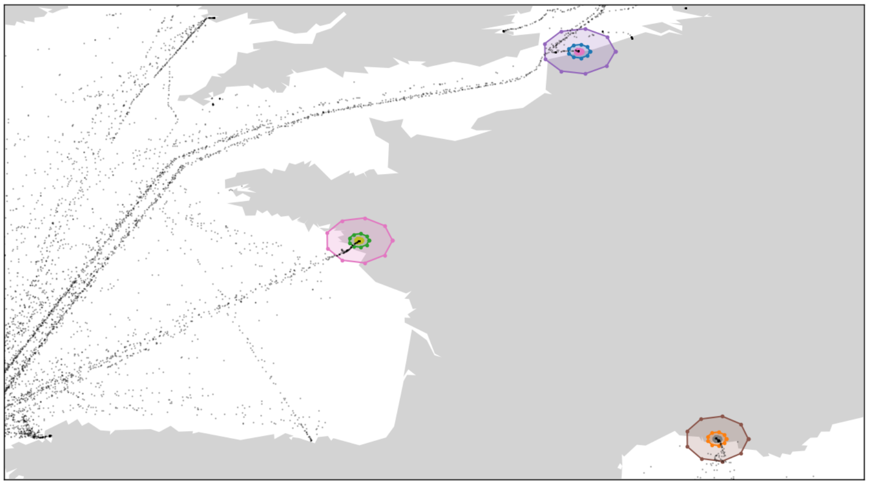


Figure 7.19: France LNG import terminals

7.2.3 India

The estimated LNG imports to India spans from 34% to 143 % of benchmark value. The categorisation of the Indian terminals spans from remote with low traffic to central with high traffic. 2018 perform better than 2017, and we get a large overshoot with diameter 3. This is both due to nearby shipping lane and two terminals being so close to each other that vessels potentially gets double counted. This could have been solved by making one large geo-fence covering both nodes. Lastly, one terminal is completely unvisited in the period, according to the AIS data. Figure 7.21 show the terminals referred to.

Table 7.11: India, estimations

India	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSMA**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	19.3	100 %	6.65	34 %	9.07	47 %	22.42	116 %	0.00	0 %
2018	23.26	100 %	10.74	46 %	14.53	62 %	33.24	143 %	0.00	0 %

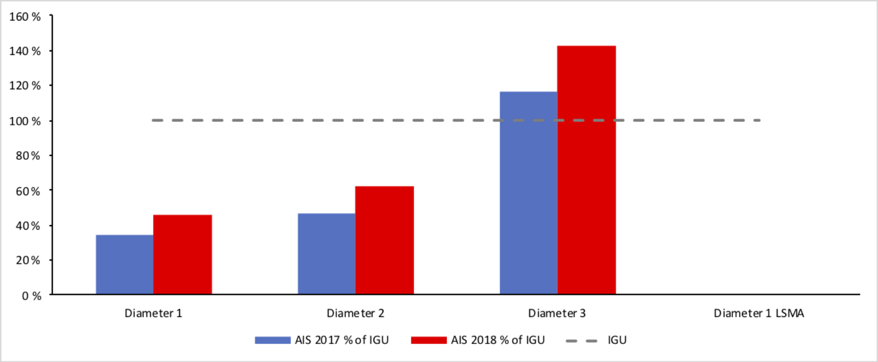


Figure 7.20: India LNG import based on AIS data

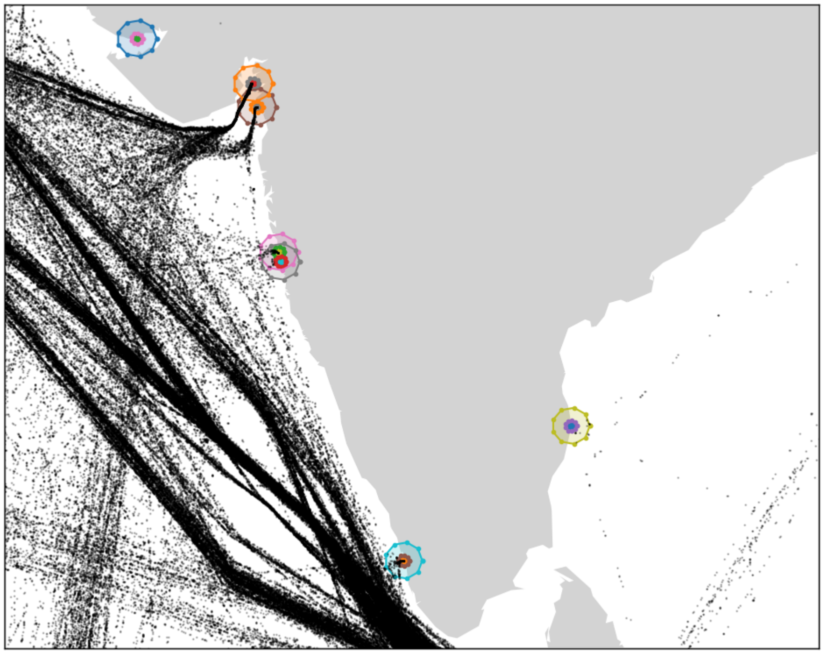


Figure 7.21: India LNG import terminals

7.2.4 Jordan

The estimated imports to Jordan spans from 3% to 63% where data amount and geo-fence size is important for this terminals results, as diameter 3 for 2018 performs best. The terminal is remote with low traffic nearby, as the terminal lay within a bay, shown in figure 7.23. The chance of counting vessels only passing by is therefore zero, and the diameter could have been doubled without risk to increase the chance of detecting visits. Jordans' import amount is in the same range as Norways export amount, and we have classified both terminals as remote and with low nearby traffic. The main difference is that although Jordan is remote in a bay with low traffic, it is still close to the Suez Canal and the Red Sea, which has extremely high traffic compared to northern Norway.

Table 7.12: Jordan, estimations

Jordan	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSM**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	3.36	100 %	0.12	4 %	1.53	46 %	1.79	53 %	0.00	0 %
2018	2.55	100 %	0.07	3 %	1.10	43 %	1.61	63 %	0.00	0 %

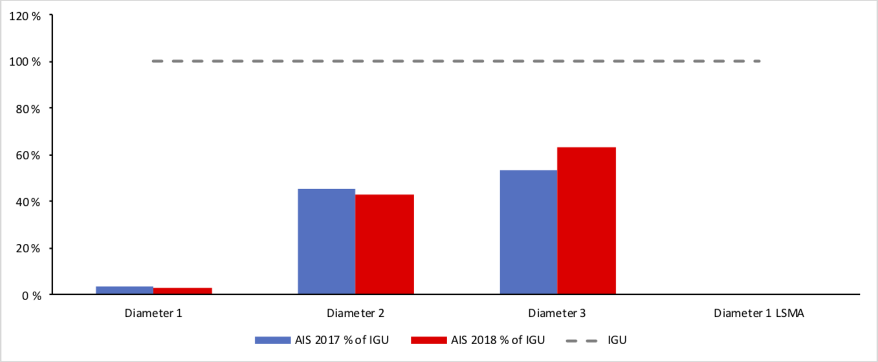


Figure 7.22: Jordan LNG import based on AIS data

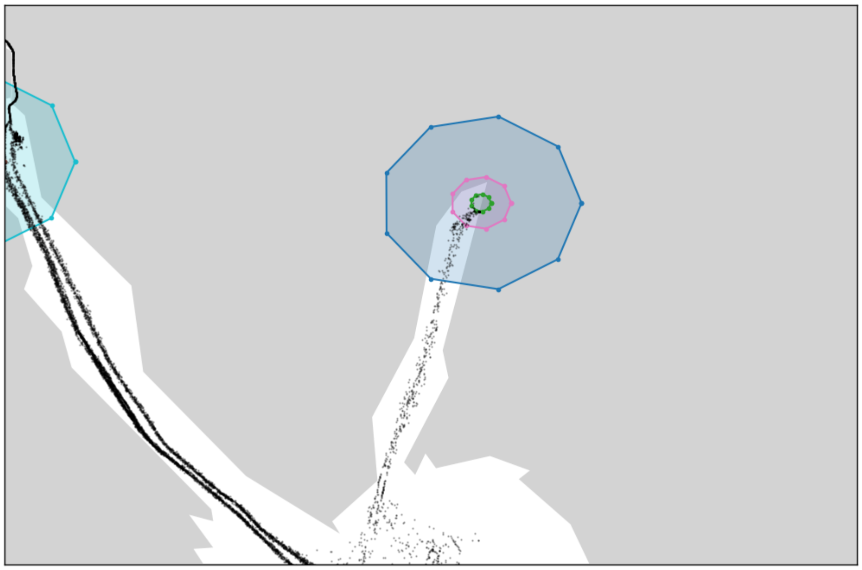


Figure 7.23: Jordan LNG import terminal

7.2.5 Kuwait

The performance of the estimations on Kuwaits LNG import is in the same area as for Jordan, with span between 28% to 71%. The terminal location is categorised as remote with low nearby traffic (same as Jordan), and the import amount is also similar.

Table 7.13: Kuwait, estimations

Kuwait	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSMA**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	3.55	100 %	1.01	28 %	1.29	36 %	2.15	61 %	0.00	0 %
2018	3.48	100 %	0.98	28 %	1.81	52 %	2.47	71 %	0.00	0 %

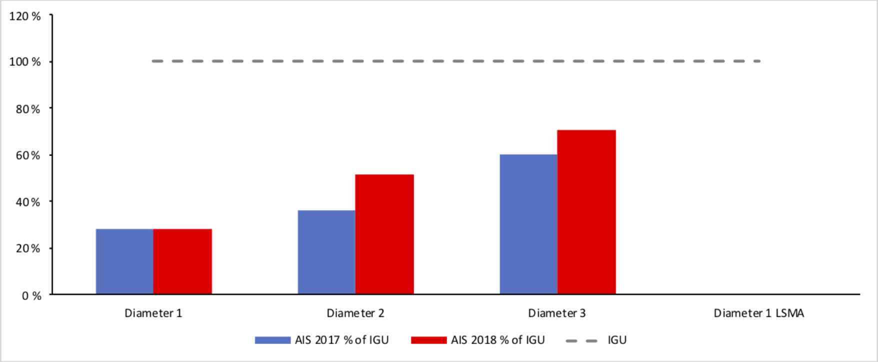


Figure 7.24: Kuwait LNG import based on AIS data

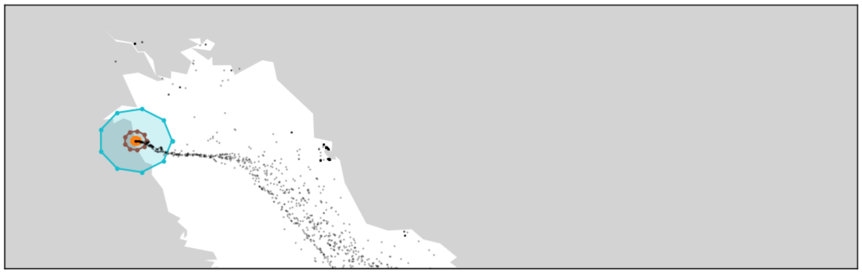


Figure 7.25: Kuwait LNG import terminal

7.2.6 Pakistan

The estimation results for Pakistan is very promising with a span between 51% and 102%, where the large data amount in 2018 combined with large radius gives by far the best results. As seen in figure 7.27 the terminal is remote with no passing by shipping lane and all nearby LNG traffic going into the terminal. Hence, we have no reason to believe that the very close estimate of 102% for diameter 3 in 2018 is just a coincidence.

Table 7.14: Pakistan, estimations

Pakistan	**IGU**		**Diameter 1**		**Diameter 2**		**Diameter 3**		**Diameter 1 LSM**	
	MTPA	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**	**MTPA**	**% of IGU**
2017	4.74	100 %	2.40	51 %	2.93	62 %	4.26	90 %	0.00	0 %
2018	7.15	100 %	3.88	54 %	4.79	67 %	7.29	102 %	0.00	0 %

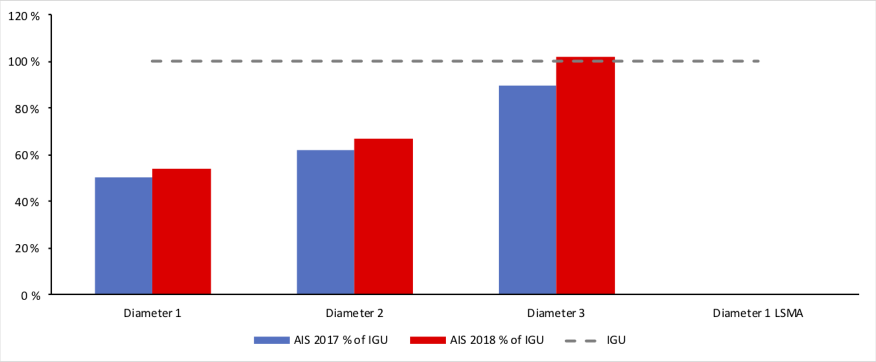


Figure 7.26: Pakistan LNG import based on AIS data

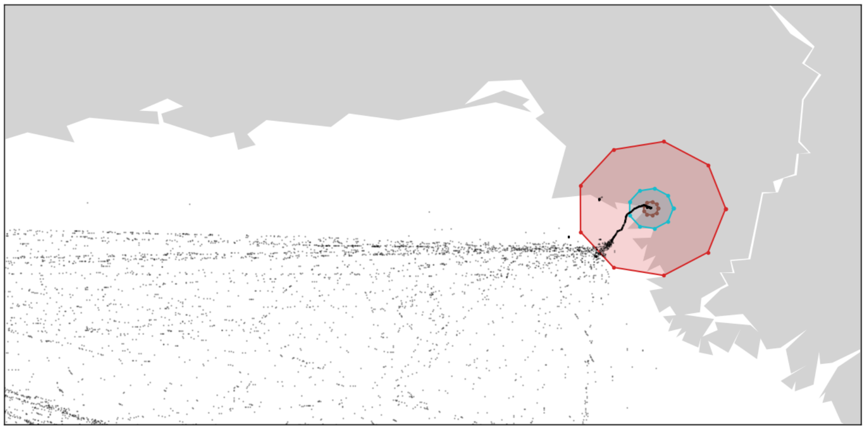


Figure 7.27: Pakistan LNG import terminal

7.3 Result Comparison and Discussion

In the previous section, the results for one and one country where presented and shortly discussed individually. In this section the various countries and diameters are presented and discussed more side by side in order to detect the main drivers to accurate and inaccurate estimations. Additional figures supporting this section can be found in appendix B.

Table 7.15 show the percentage errors for all diameters for all countries, both 2017 and 2018. The table is sorted by the average minimum error of 2017 and 2018, in order to rank the performance. The most easily spotted significantly difference between countries are the results from exporters versus importers. Both weighted average absolute percentage error, and average absolute percentage error is significantly lower for export countries than import countries in general. This can be seen in the figures, and trough the fact that five of the six lowest average minimum errors are exporters. These six countries are all within an error of 10%, while the next on the list have an error of 27%. It is however not the fact that they are exporters that make the estimations better, rather the location and nearby shipping traffic. All top six countries are classified as remote or semi central in the case study, and none of them have interfering shipping lanes through their geo-fences. This is however also the case for several of the terminals with less accurate estimations. In Eriksen et al. (2010), Eriksen et al. (2018) and Skauen and Øystein Olsen (2016), the satellite coverage and AIS transmission to the Norwegian satellites are discussed. Two main factors are presented as the main reasons to incomplete AIS data. One is the fact that the satellites orbit around and can only cover a certain range of the earth at a time, which is changed as the satellites orbit. Different areas of the world have smaller time periods of coverage than other. Since the Norwegian satellites have polar orbits, they cover equator 2-3 times a day, while the north pole is covered as much as 15 times a day. The second factor is interference also referenced to as message collision, Eriksen et al. (2018). Satellites cover a much greater area than land based receivers, and the AIS system was not originally intended for satellites. This means that satellites will receive a much higher amount of messages than land based receivers. This leads to interference, and error in the data collection. Naturally, this problem is greater where the vessel density is high. Wu et al. (2017) develop a method to find vessel density from AIS data. Their results in combination with observations from "Marine Traffic"[1] suggest that the vessel density around the top performing countries Norway, Russia and Trinidad are very low compared to the countries close to Suez and in the Arabian Gulf. Skauen and Øystein Olsen (2016) presents a method to determine vessel detection probability for different geographical areas, and the their results state that the detection probability in the Arabian gulf is low, compared to e.g Angola and Northern Norway. Figure 7.28 is a scatter plot of all messages from the 50 vessels that have the highest number of message gaps over 72 hours. The plot show that these vessels operate in the exact same areas as the countries with the biggest errors and lowest detection probability (obtained in Skauen and Øystein Olsen (2016)), except Cameroon which is a special case due to the startup and low volume. These areas are also as mentioned the areas with the highest ship density in the world.

The vessel detection probability can also be directly linked to the large geo-fences performing better than the small, especially in the areas with low detection probability. A large geo-fence implies that the vessel will be longer within the geo-fence, leading to more messages being sent from the vessel, which increase the probability that a message is captured by a satellite, detect-

---

[1] `https://www.marinetraffic.com/`

Table 7.15: Estimation errors for all countries sorted by average minimum estimation error for 2017 and 2018

		2017					2018					
Country	**Type**	**IGU [MTPA]**	**PE D1**	**PE D2**	**PE D3**	**Min Error 2017**	**IGU [MTPA]**	**PE D1**	**PE D2**	**PE D3**	**Min Error 2018**	**AVG Min Error 2017, 2018**
Norway	Export	4.04	5 %	-7 %	-2 %	2 %	4.63	-5 %	-7 %	0 %	0 %	1 %
Russia	Export	11.11	-53 %	-38 %	1 %	1 %	18.93	-12 %	-5 %	4 %	4 %	2 %
Trinidad	Export	10.76	-27 %	-9 %	9 %	9 %	12.23	1 %	3 %	15 %	1 %	5 %
Pakistan	Import	4.74	-49 %	-38 %	-10 %	10 %	7.15	-46 %	-33 %	2 %	2 %	6 %
Angola	Export	3.67	-29 %	-24 %	-14 %	14 %	4.14	-7 %	-7 %	-4 %	4 %	9 %
Algeria	Export	12.17	-60 %	-24 %	9 %	9 %	10.3	-32 %	12 %	38 %	12 %	10 %
India	Import	19.3	-66 %	-53 %	16 %	16 %	23.26	-54 %	-38 %	43 %	38 %	27 %
Qatar	Export	76.71	-81 %	-70 %	-44 %	44 %	78.69	-69 %	-54 %	-19 %	19 %	32 %
Kuwait	Import	3.55	-72 %	-64 %	-39 %	39 %	3.48	-72 %	-48 %	-29 %	29 %	34 %
Jordan	Import	3.36	-96 %	-54 %	-47 %	47 %	2.55	-97 %	-57 %	-37 %	37 %	42 %
Egypt	Import	5.97	-76 %	-54 %	192 %	54 %	2.26	-75 %	37 %	1382 %	37 %	45 %
France	Import	7.58	-92 %	-86 %	-78 %	78 %	8.43	-68 %	-62 %	-31 %	31 %	55 %
United Arab Emirates	Export	5.2	-84 %	-73 %	-63 %	63 %	5.48	-79 %	-66 %	-59 %	59 %	61 %
Cameroon	Export						0.61	152 %	131 %	95 %	95 %	95 %

PE = Percentage Error

ing the vessel. The same goes for the comparison of 2017 and 2018 errors. With two additional satellites in 2018, the number of collected messages is three times higher, increasing the probability of vessel detection, giving less error.

Although the trade flow estimations are very promising for some of the countries, it is important to state that analysis over a larger time frame is needed to draw a solid conclusion.

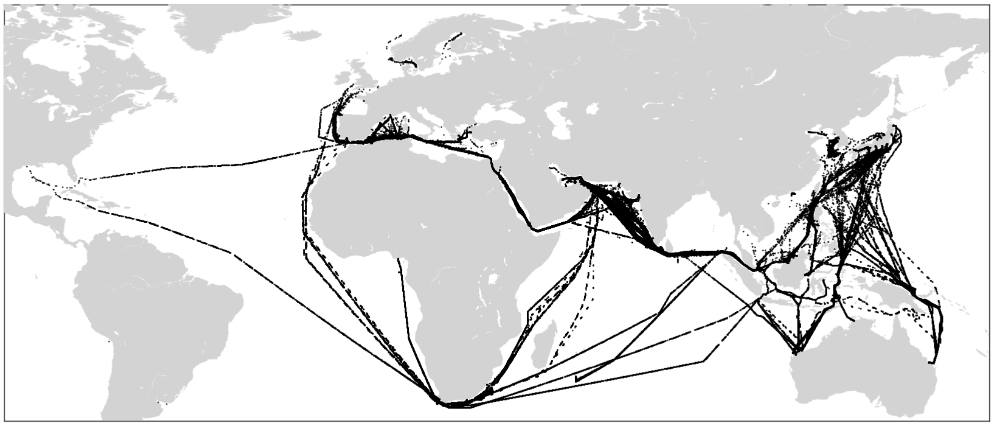


Figure 7.28: LNG carriers (50 carriers) with the highest amount of message gaps over 72 hours

7.4 Summary of Results

The ability to estimate LNG trade flow with the use of AIS data varies from 1% error to 95% error when looking at average minimum error from 2017 and 2018, compared to the benchmark value, IGU (2019). Norway, Trinidad, Pakistan, Angola and Algeria stands out as the best results with percentage errors under 10%. Cameroon has the worst performance with 95% error after a massive overshoot, but we choose to consider it as an outlier, as the facility had startup in 2018 and the production level is low. The remaining countries have an error between 27% and 61%. The inclusion of low speed and moored/anchored criteria drastically increase the error and is concluded useless.

We have found that the main reason to increased error is closely linked to vessel detection probability. This is mainly a function of vessel density and satellite coverage. All countries in the Red Sea and the Arabian gulf have large errors due to both low satellite coverage and high vessel density, while Norway, Angola, Russia and Trinidad are either far north with good satellite coverage or in areas with low vessel density, or a combination of the two. Skauen and Øystein Olsen (2016) show that the vessel detection probability has large local variations, which is a probable reason to why Algeria perform as well as it does.

It has further been found that a large geo-fence increase the probability of detection, hence the error for diameter 3 is generally lower, especially for terminals in low detection probability areas. Lastly, the results for 2018 is generally better than for 2017 due to three times more AIS messages (which is a result of two new satellites) increasing the vessel detection probability. Again, the improvement is more significant in the areas with low vessel detection probability, such as the Arabian gulf.

It has also been found that no countries gets large overshoots in estimations unless the geo-fence is clearly interfering with a nearby shipping lane. This indicate that the majority of LNG tankers sail with all cargo tanks fully loaded. However, this needs further investigation to conclude.

To conclude, the results indicate that LNG trade flows can be estimated with a low error given that the terminal is within an area with high vessel detection probability. Nevertheless, it is important to state that analysis over a larger time frame is needed to draw a solid conclusion.

Chapter 8

Discussion

8.1 Methodology

8.1.1 Vessel Search

In this study we used a methodology including an additional data source to obtain the worlds LNG fleet, in contrast to the approach used by Smestad (2015), who obtain vessel type with the use of heuristics. The reason for using an external datasource is linked to the weak results in obtaining LNG vessels with the heuristics, and the fact that the trade flow methodology requires a highly correct database. There is, however a large trade off by using an external database. Complete and reliable ship databases are expensive to acquire.

The methodology used in this thesis obtained 714 IMO numbers from Seaweb based on all LNG carriers ever built. As we use IMO numbers instead of MMSI numbers we secure getting only getting the vessels we want, as IMO numbers follow the vessels through its life, while MMSI numbers can change. The output of the method is the MMSI numbers, and we obtained 521 unique MMSI, which is in line with reported fleet size by IGU (2019). There is however no guarantee that our obtained list of vessels is without error.

8.1.2 Trade Flow Estimation Method

An advantage with the trade flow estimation methodology introduced in this thesis is the simplicity. A geo-fence is set around a node, and the accumulated gas capacity entering the geofence is measured with an additional time delta criteria. By only using data within and close to the geo-fence, all other data can easily be eliminated, making the algorithm fast.

As discussed in the result chapter 7, the method is dependent of a reliable AIS database with high vessel detection probability in and around the geo-fences. It is also dependent of literally all LNG carriers delivering AIS messages stable to the satellites without many large gaps, in order to obtain reliable estimates. This makes the method vulnerable to for example ships with faulty or partly wrongly mounted AIS systems. The results however show that the additional amount of data that comes with new satellites increase the methods reliability drastically in the low vessel detection probability areas. The results also showed that large geo-fences decreased the error estimate percentage, which means that areas that can use large geo-fences without interfering with nearby shipping lanes are more suitable for the method.

The basis for the method is as mentioned to find the sum of LNG imported or exported. This sum is a sum of LNG volume conducted in m^3, as the LNG carriers gas capacity is given as a volume. Unfortunately the benchmark source for LNG import and export give their values in tonnes of LNG, thus a conversion is needed. We have used a conversion factor of 410 kg/m^3, while LNG density can vary between approximately 400-480 kg/m^3 depending on the LNG composition. In a worst case scenario this alone could give an error of 17% alone, however the majority of LNG has a density closer to 410 kg/m^3 making the potential error much less.

Another weakness or potential source of error with the method is the chance of LNG carriers not being fully loaded, as assumed. By looking at change in draught for the vessels from AIS message type 5 in the methodology, partly filled vessels could have been detected. However, results suggest that LNG carriers are normally filled completely, but further investigation is needed to conclude on the matter.

In the case study conducted in this study, we used approximate circles with different diameters as geo-fences. A key improvement with the method would have been to customise the geo-fences to each node (terminal), as a circle rarely is the optimal geo-fence. Figure 8.1 show how an optimal geo-fence for Jordan could have looked, marked as a red rectangle. Here the vessel detection area has been maximised, while the probability of false detection is still held to a minimum.

Lastly, LNG can be re-exported after being received at an import terminal if the terminal has the capability. For know, re-exports are only a small fraction and only done at certain terminals making the impact on the results small. However, as we introduced in the introduction to LNG chapter 4, the LNG market is liberalising more and more making the trading more dynamic with an increasing amount of vessels trading in the spot market. This again could lead to higher re-export volumes. The method does not take re-exports into consideration, making the model "as is" useless if the scenario happens. Including draught as a variable in the method could solve

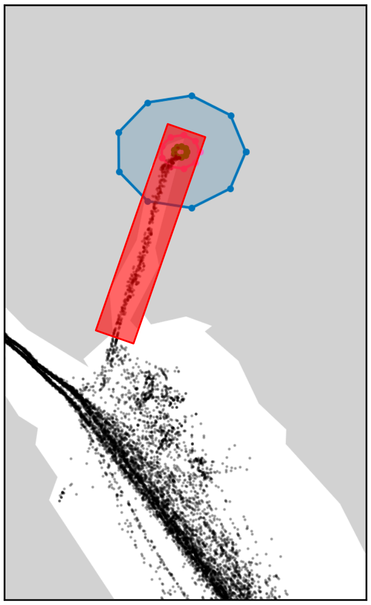


Figure 8.1: Optimal geo-fence size exemplified by the red square covering the whole bay going into Jordan

this.

While the simplicity of the method is a strength, it also puts limitations to the usage by not using any information in between geo-fence visits. By utilising this information, a global trade network could have been built, where the LNG could have been traced continuously from liquefaction to regasification terminal. In addition this could potentially opened for predictive analysis and not only estimations. The methodology for building such a network where constructed, but as the database where delivered to the author to late to be able to conduct a case study on the network creation. The methodology can be seen in appendix A.

8.2 Data Foundation and Benchmark LNG Trade Flow

As we discovered in the data preparation, the AIS data had as much as 10% obvious error data, which where removed. This alone is an action that could lead to less accurate results. There is also always a chance that there are more error data not discovered decreasing the reliability of the results.

In the case study of this thesis, two additional data sources has been used. Seaweb from IHS (2019) to obtain the LNG fleet and the gas capacity for each vessel and IGU (2019) as a benchmark for LNG trade, which has been assumed to be the actual import or export value. It is important to question the reliability of these datasources, and keep in mind that these sources can include flaws. For instance their can be incorrect gas capacities on vessels in the Seaweb database, or the trade flow into a country could be inaccurately reported by IGU (2019). However, to the authors knowledge there is no official data on LNG import or fleet data including vessel specific information. In light of this, IGU (2019) and IHS (2019) are sources with a good reputation.

Chapter 9

Conclusion

9.1 Concluding Remarks

The main objective of this thesis has been to investigate whether AIS data can be used as a reliable tool to obtain LNG trade flows in or out of countries. Circular geo-fences where put up around import and export terminals to detect unique visits by LNG carriers with known gas capacity, and accumulate the visited gas capacity. Three different geo-fence sizes where tested with diameters of 5 km, 15 km and 50 km respectively. In addition, a fourth geo-fence with a diameter of 5 km with additional low speed criteria (sog < 2 knots) and navigational status set to "moored" or "anchored" where tested. Trade flows where analysed for both 2017 and 2018, and it is important to point out that the AIS data amount for 2018 is three times higher than for 2017 due to new satellites launched in 2017. All methods where compared to IGU (2019) which was used as a benchmark for the import and export values.

Overall, the results from our analysis establish evidence that LNG trade flows can be estimated with a low error by using vessel position data from AIS. Six countries got estimation errors of less than 10% compared to benchmark, and it was found that geographical position, the amount of close by ship traffic, size of geo-fence, and data resolution are all crucial factors for how small the error of the estimation can get. The results are best for 2018 due to higher data resolution, and for terminals placed far north and/or in low vessel density areas. In addition, a large geo-fence decrease the error as the time frame given to detect the vessel increases. The inclusion of low speed and moored/anchored criteria drastically increase the error and is concluded useless.

There are some important takeaways from the methodology used. First of all, using a separate database with LNG carrier identification and gas capacity added value to the AIS data and is a necessary step to conduct the trade flow analysis. Secondly, using circular geo-fences is not an optimal solution as different terminals have different surroundings. Customising each geo-fence to the respective terminal will decrease the error rate, as large geo-fences increase the vessel detection probability. Thirdly, the methodology lacks the ability to detect re-export of LNG from import terminals, which is not a big issue at the moment, but a potential issue for the future.

9.2 Recommendation for Further Work

Further research within the use of AIS data to obtain trade flows and shipping networks is recommended. Based on the findings in this study, there is substantial potential of using AIS data to obtain LNG trade flows. However, it is recommended that the methodology used gets tested over a larger time frame in order to check the reliability over time. It is also recommended to test the methodology on data from another data provider with higher data resolution as well as using customised geo-fences for all terminals in the study.

Implementing draught analysis in the methodology would allow both flow inn and flow out of geo-fences, which both could make the trade flow estimations more reliable and increase the usage area to for example making a complete LNG trading network or utilise the methodology on other seaborne commodity flows. The author recommends to both investigate the potential of constructing a complete LNG trade flow network with the use of AIS data, as well as testing the methodology on other seaborne commodities.

Bibliography

Adland, R. and Jia, H. (2016). Dynamic speed choice in bulk shipping. *Maritime Economics & Logistics*.

Adland, R., Jia, H., and Strandenes, S. P. (2017). Are ais-based trade volume estimates reliable? the case of crude oil exports. *Maritime Policy & Management*, 44(5):657–665.

Arguedas, V. F., Pallotta, G., and Vespe, M. (2014). Automatic generation of geographical networks for maritime traffic surveillance. In *17th International Conference on Information Fusion (FUSION)*, pages 1–8.

Arksey, H. and O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8(1):19–32.

Axelsen, J. (2018). A study of the operational patterns of lng carriers from ais data.

Babri, S., Jørnsten, K., and Viertel, M. Application of gravity models with a fixed component in the international trade flows of coal, iron ore and crude oil. *Marit Econ Logist*, 19(2):334–351.

Bosma, P. and Nagelvoort, R. K. (2009). Liquefaction technology; developments through history. In Alfadala, H. E., Reklaitis, G. R., and El-Halwagi, M. M., editors, *Proceedings of the 1st Annual Gas Processing Symposium*, volume 1 of *Advances in Gas Processing*, pages 19 – 31. Elsevier, Amsterdam.

BP (2019). Bp energy outlook 2019 edition.

Eriksen, T., Greidanus, H., and Delaney, C. (2018). Metrics and provider-based results for completeness and temporal resolution of satellite-based ais services. *Marine Policy*, 93:80 – 92.

Eriksen, T., Skauen, A. N., Narheim, B., Helleren, O., Olsen, O., and Olsen, R. B. (2010). Tracking ship traffic with space-based ais: Experience gained in first months of operations. pages 1–8.

Geng, J.-B., Ji, Q., and Fan, Y. (2014). A dynamic analysis on global natural gas trade network. *Applied Energy*, 132:23 – 33.

IEA (2018). Offshore energy outlook. *International Energy Agency*.

IGU (2018). 2018 world lng report. *https://www.igu.org/*, (accessed 2019-01-15).

IGU (2019). 2019 world lng report. *https://www.igu.org/*, (accessed 2019-04-20).

IHS (2019). Seaweb ship database ihs markit. *https://ihsmarkit.com/*, (accessed 2019-04-01).

IMO (2002). Guidelines for the onboard operational use of shipborne automatic identifcation systems (ais),resolution a.917(22). *Guidelines for the onboard operational use of shipborne automatic identifcation systems (ais),resolution a.917(22)*.

ITU (2014). Technical characteristics for an automatic identification system using time division multiple access in the vhf maritime mobile frequency band m series mobile, radiodetermination, amateur and related satellite services.

Jia, H., Smith, T., and Prakash, V. (2019). Estimating vessel payloads in bulk shipping using ais data. *International Journal of Shipping and Transport Logistics*, 11:25.

Kaluza, P., Koelzsch, A., Gastner, M., and Blasius, B. (2010). The complex network of global cargo ship movement. *Journal of the Royal Society, Interface / the Royal Society*, 7:1093–103.

Kontopoulos, I., Spiliopoulos, G., Zissis, D., Chatzikokolakis, K., and Artikis, A. (2018). Countering real-time stream poisoning: An architecture for detecting vessel spoofing in streams of ais data. In *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, pages 981–986.

Leonhardsen, J. H. (2017). Estimation of fuel savings from rapidly reconfigurable bulbous bows - exemplifying the value of agility in marine systems design.

Makholm, D. J. D. and Olive, D. L. T. (2016). A petroleum tanker of a different color: Obstacles to an lng-based global gas spot market.

Maxwell, D. and Zhu, Z. (2011). Natural gas prices, lng transport costs, and the dynamics of lng imports. *Energy Economics*, 33(2):217–226.

Millefiori, L. M., Zissis, D., Cazzanti, L., and Arcieri, G. (2016). A distributed approach to estimating sea port operational regions from lots of ais data. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 1627–1632.

Nersesian, R. L. (2007). *Energy for the 21st century : a comprehensive guide to conventional and alternative sources / Roy L. Nersesian*. M.E. Sharpe, Inc Armonk, N.Y.

Nikhalat-Jahromi, H., Angeloudis, P., Bell, M. G. H., and Cochrane, R. A. (2017). Global lng trade: A comprehensive up to date analysis. *Maritime Economics & Logistics*, 19(1):160–181.

Næss, P. (2018). Investigation of multivariate freight rate prediction using machine learning and ais data.

Pallotta, G., Vespe, M., and Bryan, K. (2013). Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction. *Entropy*, 15:2218–2245.

Skauen, A. N., Helleren, Ø., Olsen, Ø., and Olsen, R. (2013). Operator and user perspective of fractionated ais satellite systems. In *Proceedings of the AIAA/USU Conference on Small Satellites, Around the Corner, SSC13-XI-5*.

Skauen, A. N. and Øystein Olsen (2016). Signal environment mapping of the automatic identification system frequencies from space. *Advances in Space Research*, 57(3):725 – 741.

Smestad, B. B. (2015). A study of satellite ais data and the global ship traffic through the singapore strait.

Spiliopoulos, G., Zissis, D., and Chatzikokolakis, K. (2018). A big data driven approach to extracting global trade patterns. In Doulkeridis, C., Vouros, G. A., Qu, Q., and Wang, S., editors, *Mobility Analytics for Spatio-Temporal and Social Data*, pages 109–121, Cham. Springer International Publishing.

Stopford, M. (2009). *Maritime Economics, 3rd Edition*. Routledge, 3 edition.

Tu, E., Zhang, G., Rachmawati, L., Rajabally, E., and Huang, G. (2018). Exploiting ais data for intelligent maritime navigation: A comprehensive survey from data to methodology. *IEEE Transactions on Intelligent Transportation Systems*, 19(5):1559–1582.

Wang, Y. and Wu, L. (2017). A functional model of ais data fusion. pages 191–199.

Wu, L., Xu, Y.-J., Wang, Q., Wang, F., and Xu, Z. (2017). Mapping global shipping density from ais data. *Journal of Navigation*, 70:67–81.

Appendix A

Trade Network Methodology

As the AIS database was delivered to the author on the 30th of May, there was not sufficient time to implement the complete trade network methodology. The following section will however present the methodology made.

A.1 LNG Trade Flow Analysis - Worldwide Trade Network

In section 5.2 we described a method to obtain LNG trade flow in our out of a polygon (port). This section on the other hand will describe a method made to obtain a worldwide LNG trade network. In specific this means that the methods goal is to track the LNG trade flux between all countries or areas in the world. While the described "Single port analysis" method for example obtain the export amount out of Trinidad, the worldwide trade network method has the goal of tracking all cargoes going out of Trinidad (and all other countries), including their destination. This adds much complexity and sources of error.

A method containing four main steps has been derived in order to obtain a worldwide LNG trade network:

- Step 1: Node selection
- Step 2: Tagging messages to nodes
- Step 3: Creating LNG shipping Network
- Step 4: Deriving LNG trade flow from the shipping network

A.1.1 Node Selection

The first main step is the selection of the nodes in the trade network. A flow chart of the methodology used to obtain the wanted nodes can be seen in figure A.1. The nodes are import and/or export areas defined as geo-fences (polygons). All areas where LNG is loaded/unloaded should be included inside a geo-fence in order to keep track of all LNG flows. IGU (2019) include a complete list of all liquefaction and regasification terminals. This list is used as the initial node selection in this method. Several of the terminals in the list are very close or even placed in the exact same geographical position. The next step is therefore to merge nodes placed very closely to one node. A script to merge all nodes closer than a certain distance using haversine distance has been developed for this purpose. In addition other nodes with greater distance are merged manually based on geographical location, nearby shipping lanes and the value of the added information by keeping the nodes unmerged. After the reduced set of nodes are selected, each node needs to be analysed in order to obtain an expedient polygon size, as discussed in section 5.2.1. The AIS data set should also be analysed with the purpose of detecting loading/unloading outside the defined polygons. This can be done by plotting all messages at low speeds, including checks against navigational status. The output from the node selection should be a list of all nodes including latitude and longitude position of the centre, associated area category, and information about import and export capabilities. Nodes with import only are defined "1", nodes that are export only are defined "2", and nodes with both import and export capabilities are defined "3". An example of a node list is presented in table A.1.

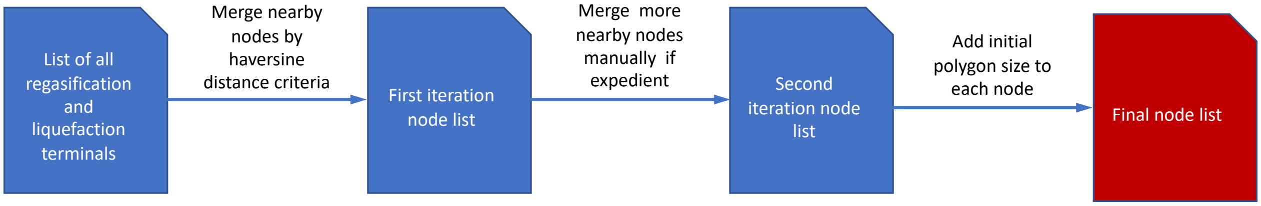


Figure A.1: Flow chart of node creation methodology

Table A.1: Node list example

node	country	terminal	latitude	longitude	type	radius[km]
...						
14	Brazil	Pecém	-3.546	-38.831	1	20
15	Brunei	Brunei LNG	4.669	114.467	2	10
16	Canada	Canaport	45.211	-65.978	1	20
...						

A.1.2 Tagging AIS Messages to Nodes

The second main step is to connect all AIS messages of type 1 to a node. All messages inside a geo-fence (polygon) gets tagged with the nodes index from the node list. The ray casting algorithm, which is described in section 5.2.2, is used in order to do this. A large part of the AIS messages are not inside any geo-fence (polygon). These messages are tagged "-1". The method is described by the following algorithm:

Algorithm 2 Tagging AIS Messages To a Polygon

```
for all AIS Messages of type 1 do
    for all Nodes do
        if RayCasting(AIS Message,Node) == True then
            Tagg message by node index
        end if
    end for
    Tagg message as outside all nodes
end for
```

A.1.3 Creating LNG Shipping Network

In the same manner as we did in the single port trade flow analysis, we want to derive all the unique visits by all LNG carriers inside all the geo-fences. In section A.1.2 we tagged all the AIS messages to the various geo-fences, and the trade network can now be derived from these tags.

The network is derived in the same manner as we derived the unique visits in the single port trade flow analysis giving a port log with only one AIS message per visit in a geo-fence. We start by selecting all AIS messages of type 1 for one LNG carrier at a time. Then we look at the added node tag column, and remove all messages where the previous message has the same tag. This gives us only one message tagged to a node per unique visit. Then all messages with tag "-1" are removed, and the time criteria saying that unique visits in the same node must be at least 24 hours apart each other is added. The output is a port log with all visits into the geo-fences areas added, including information on previous visit. With a perfect complete dataset, this would let us keep track of all LNG carrier movements. Table A.2 show an example port log.

A.1.4 Deriving LNG Trade Flow From the Shipping Network

The objective of this step is to derive the LNG trade volume between all the specified nodes by using the port log obtained in section A.1.3. There are two main difficulties related to this. The first is to keep track of whether the LNG carriers are loaded or empty at all times in order to know

Table A.2: Port log example from world wide trade network analysis

	DateTime	**MMSI**	**node**	**previous node**	**Gas Capacity**
0	2015-03-25 23:50:00	CENSORED	71	No prev	151900
1	2015-04-17 00:29:06	CENSORED	5	71	151900
2	2015-05-22 01:02:10	CENSORED	5	5	151900
...					

if the carrier is loading or unloading in a node. The second main difficulty is related to the data quality and the fact that the Satellites that collect AIS data have major blind zones which can result in node visits being missed. Missing a node visit not only results in a missed trade flow, but it also makes it harder to keep track of the carriers loading condition. It is therefore crucial to distinguish between information we assume to be certain, and less certain information. Therefore we sum the trade from "previous node" to "node" when this is possible, otherwise the trade is market either "from unknown" or "to unknown" based on what information we are missing.

Appendix B

Extra Figures Supporting the Results

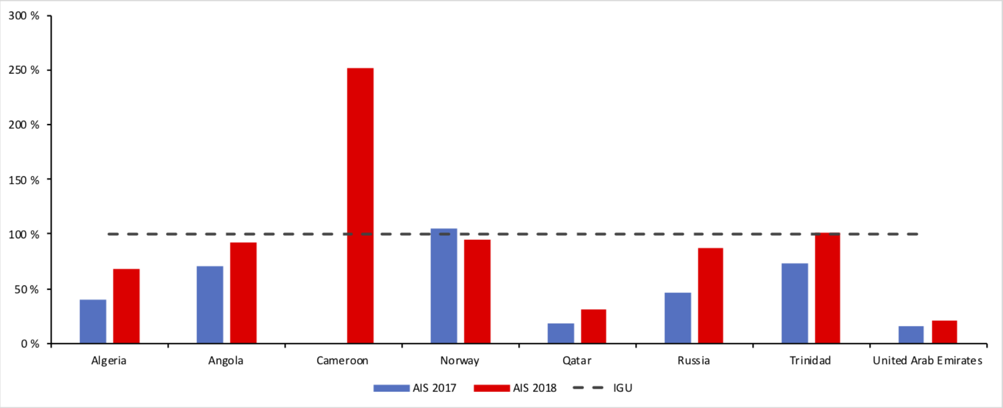


Figure B.1: LNG trade estimation with diameter 1, exporters

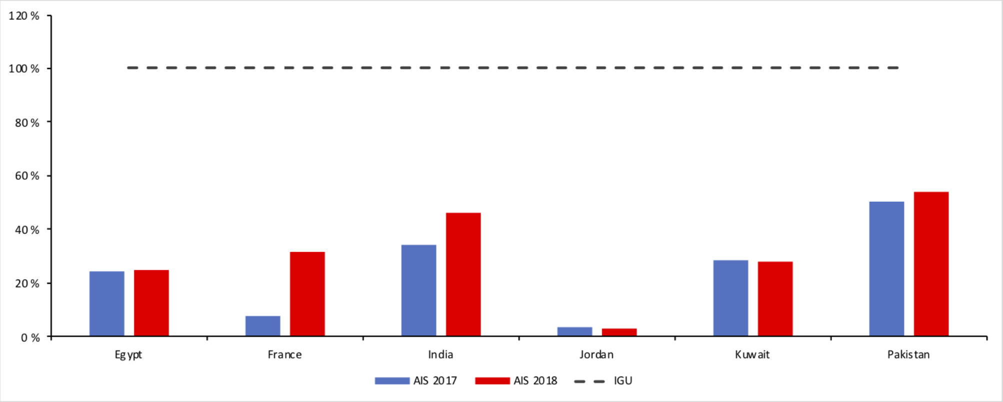


Figure B.2: LNG trade estimation with diameter 1, importers

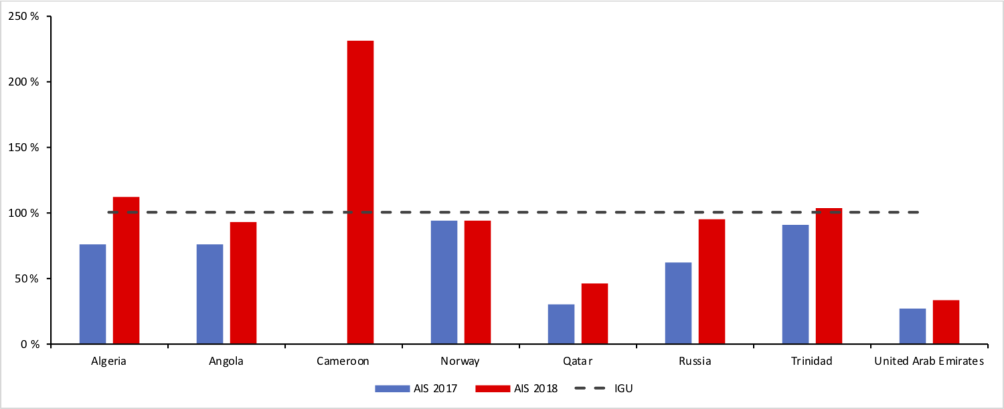


Figure B.3: LNG trade estimation with diameter 2, exporters

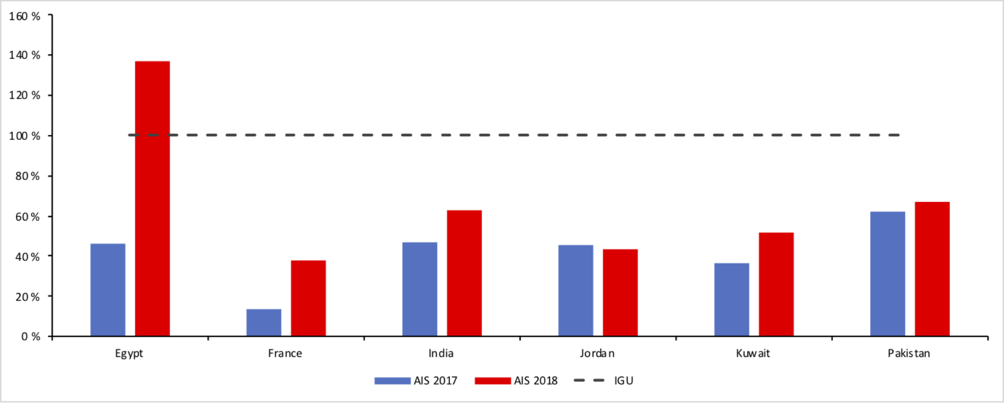


Figure B.4: LNG trade estimation with diameter 2, importers

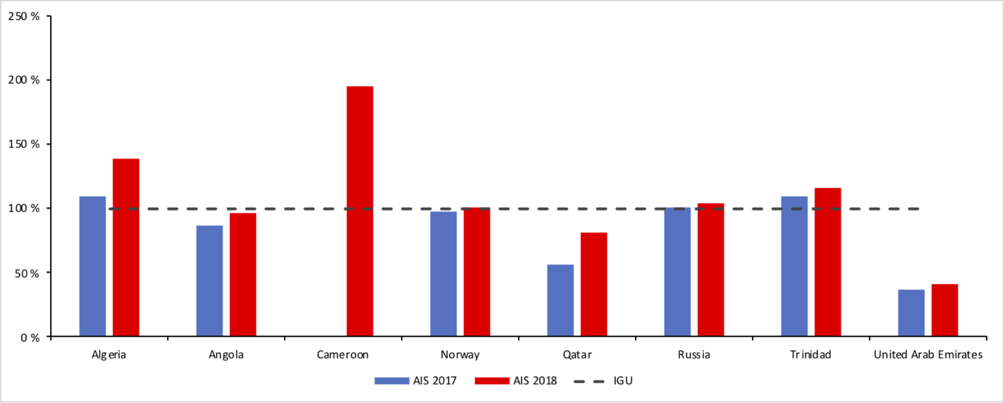


Figure B.5: LNG trade estimation with diameter 3, exporters

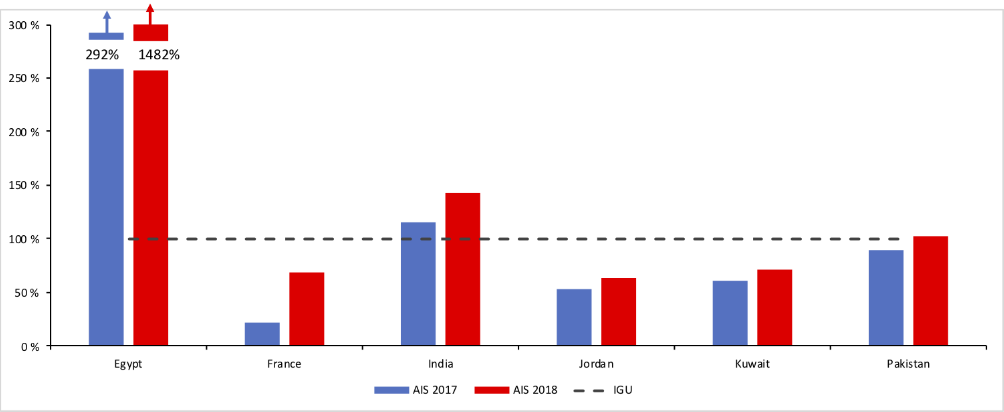


Figure B.6: LNG trade estimation with diameter 3, importers

B.1 Errors

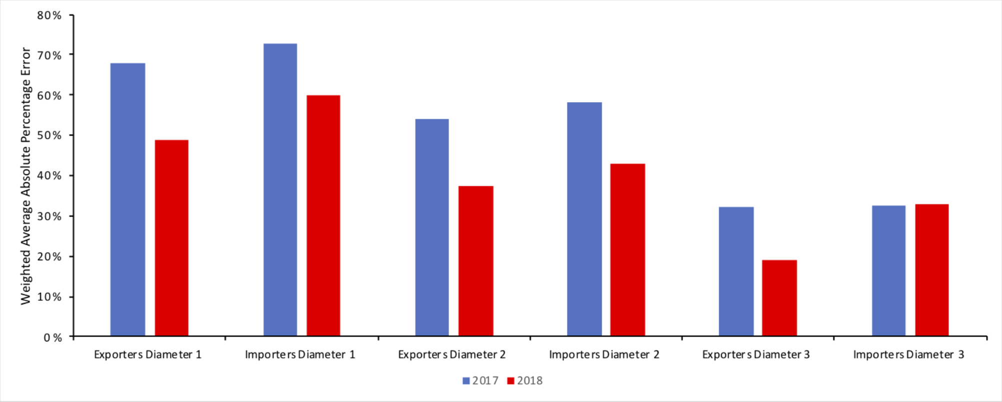


Figure B.7: Weighted average absolute percentage error

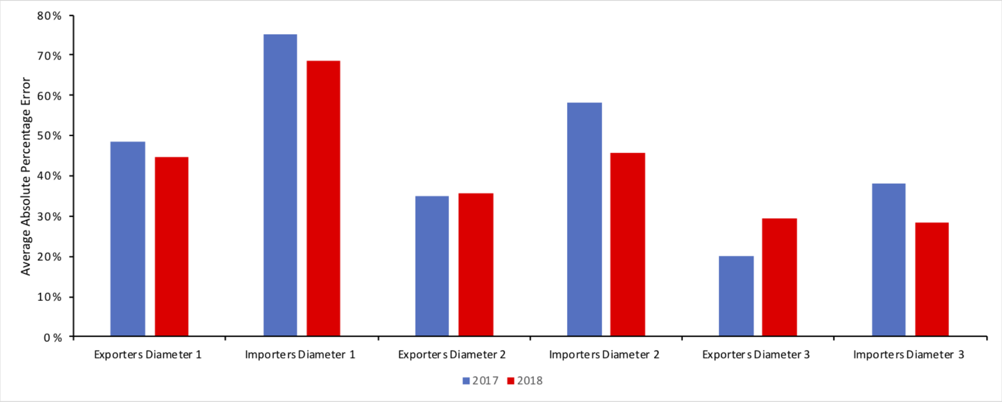


Figure B.8: Average absolute percentage error

Appendix C

Python Code

CONTACT THE AUTHOR

Torjus Halden

Estimation of LNG Trade Flows With the Use of AIS Data

NTNU
Norwegian University of
Science and Technology