Norwegian School of Economics

Bergen, Spring 2017

Investigating the Predictive Ability of AIS-data

The case of Arabian Gulf tanker rates

Mathias Olsen & Truls Rønne Kopke da Fonseca

Supervisor: Roar Os Adland

Master in Finance

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

This thesis investigates whether information derived from AIS-data incorporates superior information about future freight rates. Specifically, we assess if such data improve the ability to predict TD3 spot rates between August 2015 and mid-February 2016. The ability to anticipate short-term fluctuations in freight rates is a key component to long-term profitability for both shipowners and charterers, making the purpose of this thesis an important objective.

The AIS-data contain information about 81,728 individual shipments of crude oil between 2013 and mid-February 2016, and are reduced to 53,116 observations after proper cleansing. For analysis purposes, information deemed relevant are converted into weekly time series, ending up with 162 observations in total. Data-driven selection tools are then used to identify the most powerful predictors of future TD3 rates, and a multivariate VAR is specified in line with these results. In order to investigate the relative performance of information derived from AIS-data, a one-step-ahead forecast is conducted, and evaluated against an univariate ARMA and a multivariate VAR solely based on publicly available data.

Our results suggest that multivariate models perform relatively better than univariate models to predict future freight rates. Further, comparing error measures from the two multivariate VAR models specified, we find weak evidence in favour of using information from AIS-derived data for predictive purposes.

Preface

This master thesis is written as a concluding part of our Master of Science in Economics and Business Administration at The Norwegian School of Economics (NHH). The thesis is written within the field of our major in Finance and corresponds to one semester of full time studies.

First and foremost, we would like to thank our supervisor, Roar Os Adland, for prolific discussions and mentoring skills throughout the process. Guidance and constructive criticism have been essential for our progress. We also want to show gratitude to PhD student Vit Prochazka for helpful counseling on how to properly structure our dataset for time series analysis. Finally, we would like to thank The Norwegian Shipowners Associations Fund at NHH for their grants; hopefully our work will be of interest.

Bergen, June 2017

Mathias Olsen   Truls R. K. da F.

Mathias Olsen   Truls Rønne Kopke da Fonseca

Contents

List of Figures

List of Tables

1 Introduction

In late 2004, the International Maritime Organization (IMO) adopted new regulations regarding carriage requirements for all sea-borne traffic. These regulations required ships of 300 gross tonnage and upwards, and passenger ships irrespective of size, to carry an Automatic Identification System (AIS). The AIS primary objective was to provide information on vessel whereabouts to coastal authorities and other sea-borne traffic, and by doing so reduce the danger of collision (IMO, 2017). However, the advent of the global AIS, in combination with other data sources on the cargo onboard, has also enabled maritime researchers to access disaggregated data on the sea-borne transportation of commodities. Such data are more timely than traditional custom-based trade flows,[1] and can potentially provide valuable information on the driving factors of freight rate formation.

As for all other free markets, the price for sea-borne transportation is determined by the interplay between supply and demand. The general notion among maritime researchers is that the supply of freight not only depends on the overall cargo capacity, but also on the operational efficiency of the fleet. However, the latter is often assumed constant in the context of shipping market analysis, as reliable data on these dynamics are not generally available as historical time series. Moving to the demand dynamics of the market, both the volume of cargo being transported and the distance over which the cargo are being moved affect the overall demand for freight. Hence, the true demand is measured on a tonne-mile basis, multiplying the number of metric tonne shipped by the distance you ship them. However, due to the lack of disaggeregated data on the sea-borne transportation of commodities, maritime researchers often fail to properly account for the distance effect on demand. The AIS-data are disaggregated to individual shipments and provide reliable information on sailing distance and cargo size, which allows us to calculate tonne-mile with great accuracy. Further, information on vessel specifications and time spent at sea enable us to derive measures of operational efficiency on a per shipment basis. Hence, by utilizing information derived from AIS-data in our analysis, we aim to circumvent both of these shortcomings in the exiting literature.

[1] Customs-based trade flows is only provided at the monthly frequency, and usually with a time-lag of several weeks, which makes such data less relevant for the purpose of this thesis.

This thesis investigates whether information derived from AIS-data incorporates superior information about future freight rates. Specifically, we assess whether such data improve the ability to predict TD3[2] spot rates between August 2015 and mid-February 2016. We have examined this hypothesis in the following way: First, we use data-driven selection tools to identify most powerful predictors of future TD3 rates among a pool of potentially relevant variables. Second, we apply standard time series techniques to our in-sample data to investigate whether any causal relationships with respect to the TD3 rate are captured. Finally, we conduct a one-step-ahead forecast from August 2015 to mid-February 2016, and we use standard evaluation measures to compare the out-of-sample performance against a univariate ARMA and a multivariate VAR solely based on publicly available data.

Freight rates are considered one of the most volatile commodities that exist, with weekly changes all the way up to 47% between January 2013 and mid-February 2016. The volatile nature of freight rates makes it difficult to forecast, but also highly rewarding if a good forecasting model were found. Shipowners cash flow, as well as charterer's costs, are to a large extent determined by fluctuations in freight rates, which in turn makes reasonable anticipations of future freight rates a key component to long-term profitability. Today, ships are typically fixed on a lump sum basis, paying a fixed amount of money for a given period of time. This makes short-term forecasts of freight rates primarily benefit shipowners in the decision of when to fix, or whether to fix a ship on a short or long voyage, keeping in mind the alternative cost of not fixing.

The remainder of this thesis is structured as follows: Section 2 reviews relevant literature within the field of freight rate formation and forecasting. In Section 3, we introduce the methodological framework used to examine the predictive ability of information derived from AIS-data. Data preparation and descriptive statistics are described in Section 4, while Section 5 contain results and discussion of our analyses. Finally, a conclusion to sum up our findings, criticism to our findings and suggestions to further research, are presented in Section 6.

[2]The TD3 rate is the benchmark price for VLCC laden trips between the Arabian Gulf and Japan, in which VLCC refers to crude oil tankers ranging from 180-320 deadweight tonnage.

2 Literature review

The price for sea-borne transportation is determined by the interplay between supply and demand, where the freight rates realized in the market are mutually acceptable by both shipowners and charterers. Cheng et al. (2010) states that the main determinants of supply are the overall cargo capacity and the operational efficiency of the fleet, in which Wergeland and Wijnolst (1996) find load factor, sailing speed and waiting days per trip to constitute the most important elements of the latter. The driving factors for demand are more ambiguous and directly derive from the market of the cargo transported. For tanker services, the demand depends on the volume of crude transported, which in turn is linked to the world economic activity and the overall consumption of oil products (Stopford, 2009). However, the demand also depends on the distance over which the cargo is transported: A tonne of crude oil shipped from the Arabian Gulf to the US generates a higher demand for tanker freight than the same tonne going to Europe. Hence, the true demand for tanker freight is measured on a tonne-mile basis, multiplying the number of metric tonnes shipped by the distance you ship them (Aadland et al., 2017).

As evident in Beenstock and Vergottis' (1993) review of the early efforts to model the market for shipping, there have been few attempts to model the demand side of the market. Tinbergen (1934) investigates the sensitivity of freight rates to changes in supply and demand. He considers demand to be perfectly inelastic with respect to freight rates, while supply to respond to freight rates with shifts in the overall fleet capacity and the price of fuel. An increase in the overall cargo capacity leads to a higher supply of freight, which in turn puts negative pressure on freight rates. Conversely, an increase in the price of fuel forces more vessels into economical slow steaming and lay-up, leading to lower supply and higher freight rates. Koopmans (1939) further extends Tinbergen's work by investigating the behaviour of tanker freight rates under different market conditions. He finds supply to be rather inelastic in periods with low freight rates, while in periods with high freight rates the supply is almost perfectly elastic. However, he also problematizes the practical difficulties in modelling the demand side of the market: ”a quantitative analysis of the demand factors of the tanker freight market is a task which far surpasses the limits imposed by practical considerations...”.

Moving to more recent literature on the determinants of Arabian Gulf tanker rates, Tham (2008) utilizes Bayesian selection methods to identify the leading indicators of TD3 rates. He finds refining margin in Asia, crude oil production in the Middle East, vessel utilization rate and Brent/Dubai spread the most significant drivers of TD3 front month swaps. Assmann et al. (2016) argues on similar grounds, using data-driven selection tools to select the most powerful predictors of spot TD3 rates. In addition to refining margins in Asia and crude oil production in The Arabian Gulf (AG), they find gasoline price in Singapore, West/East differences in gasoline price, VLCC available spot in AG, and VLCC future deliveries useful to predict TD3 spot rates. However, they emphasize there are several drawbacks to the research, with one being to not properly account for variation in operational efficiency over time. Aadland et al. (2017) investigate the reliability of AIS-based trade volumes, in which they find reasonable alignment with official customs-based export numbers. As a concluding part of the article, they also argue on the implications of such data in modelling shipping markets. First, they note that tonne-mile for individual shipments are observed within the data set, which provide valuable insight into the demand dynamics in the market. Second, information on sailing speed and load factor allows maritime researchers to incorporate variation in operational efficiency.

An academic review by Kavussanos and Visvikis (2006) reveals that the majority of the existing literature on freight rate forecasting is concerned with the relationship between tradable forward contracts and future spot rates. Such studies find theoretical support in the unbiassedness hypothesis, in which forward rates are expected to be an unbiased predictor of future spot rates in markets for non-storable commodities.[3] Kavussanos and Nomikos (1999) use a Vector Error Correction Model (VECM) to investigate the lead-lag relationship between spot rates and the former BIFFEX futures contracts. They find VECM to perform relatively better than univariate ARIMA and random walk in forecasting spot freight rates, which indicate that such contracts fulfill their unbiased role. Batchelor et al. (2007) argue on similar grounds, investigating whether Forward Freight Agreements (FFA) contracts contain information about future spot rates. In line with previous research, they also find evidence in favour of the unbiassedness hypothesis. However, in forecasting the price for tradeable forward contracts, neither of the two studies have found a model that are able to beat

[3] The unbiassedness hypothesis only holds if the market is speculatively efficient.

the random walk. As become evident, only limited attention has been paid to examine the forecasting performance of publicly available data on market fundamentals. One such study is covered in Assmann et al. (2016), where the aim is to evaluate if variables chosen based on fundamental reasoning add value to produce good forecasts. However, they find univariate models perform relatively better than multivariate models based on publicly available data. Hence, they have not been able to identify any group of past or currently available data that incorporate superior information about the future. As a concluding part of the article, they emphasize that further further research could include a longer data set or possibly the more recently available AIS-data.

To our knowledge, there has been no attempt in using information derived from the AIS-data for an analysis of freight rate formation. In addition, the above review highlights considerable gaps in the existing literature that such data could potentially fill. Accurate estimates of tonne-mile provide better insight into the demand dynamics of the market, while data on sailing speed and load factor allow maritime researchers to incorporate variation in operational efficiency. Thus, the contribution of our thesis to the existing literature is twofold: First, we conduct a more detailed analysis within the field of freight rate formation, by properly accounting for the distance effect on demand and the variation in operational efficiency over time. Second, we assess the applicability of such data in the context of freight rate prediction, keeping in mind that such data have not been used before.

3 Methodology

3.1 Freight rate determinants

Table 1 presents all the variables we have considered in our analysis of the predictive ability of information derived from AIS-data. The table also includes the variables' expected impact,[4] the unit of measurement,[5] and, for comparison, whether the variables are found useful to predict TD3 rates in either Tham (2008) or Assmann et al. (2016). Descriptions of how the variables are constructed, and the rationale behind each variable, are further elaborated on in Section 3.1.1, Section 3.1.2, and Section 3.1.3.

Table 1: List of variables

Variables	Exp.sign	Unit	Assmann et al. (2016)	Tham (2008)
Dependent variable				
TD3_rate		\$/tonne		
Tonne-mile demand				
AG_VLCC	+	tm MMM		
AG_other	+	tm MMM		
global_VLCC	+	tm MMM		
global_other	+	tm MMM		
Operational efficiency				
fleet_speed	−	knots		
voyage_speed	?	knots		
load_factor	−	%		
Publicly available data				
mideast_production	+	\$/bbl MM	x	x
VLCC_fleetchange	−	\$/bbl MM	x	
spot_VLCC	−	dwt M	x	
gasoil_sing	?	\$/bbl	x	
diff_ussing	+	\$/bbl	x	
refinery_margin	+	\$/bbl	x	x
Brent/Dubai_spread	+	\$/bbl		x

[4] The impact of *(gasoil_sing)* and *(voyage_speed)* on TD3 rates are not obvious.
[5] Each M indicates 1,000.

3.1.1 Publicly available data

Variables selected from publicly available data are largely based on findings in Tham (2008) and Assmann et al. (2016), and aim to capture dynamics discussed in existing literature on freight rate formation. The supply of tanker freight depends on the overall cargo capacity and the operational efficiency of the fleet, in which data on operational efficiency are not generally available as historical time series. However, historical data on VLCC fleet development are accessible from Clarkson Research (2017) and will in our analysis serve as a proxy for the global supply of tanker freight (*VLCC_fleetchange)*. Local supply dynamics more specifically related to the TD3 route are approximated by VLCC tonnage available for chartering in AG (*spot_VLCC)* and, similarly to global supply, will be expected to *ceteris paribus* have a negative impact on the TD3 rate.

The demand for tanker services is derived by the volume of crude oil transported, which in turn is linked to world economic activity and the overall consumption of oil products. With refineries as the only conventional buyer of crude oil, the volume of crude oil transported closely relates to refinery profitability: In periods with high refinery margins, we expect refineries to be in demand of more crude oil, which in turn increases the demand on tanker transportation and put upwards pressure on freight rates. In order to account for these dynamics, we have approximated refinery margins in Asia as the difference between the retail gasoline price in Singapore and the price for Dubai crude oil (*refinery_margin)*,[6] in which both data series are obtained from Bloomberg Professional Services (2017). Despite crude oil not being a homogeneous commodity, there might be considerable cross demand between the different crude oil types. Hence, we expect the demand for tanker transportation in AG to increase when the price for Dubai crude oil gets relatively cheaper. To capture this effect, we have constructed a variable on the price difference between Brent crude oil and Dubai crude oil (*brentdubai_spread)*, in which we expect a relatively lower price for Dubai crude oil to push the TD3 rate upwards.

Local demand for tanker transportation will naturally also be highly affected by the amount of crude oil available for trading in a particular region. Hence, we have gathered data on the average

[6]Refinery margins are the difference between the price of crude oil and the petroleum products extracted from it, and they will typically vary across different crude oil types and refinery configurations.

production volume of crude oil in the Middle-East *(mideast_production)*[7] to capture demand dynamics specifically related to AG. Initially, we would also expect retail gasoline price in Singapore *(gasoil_sing)* to capture demand dynamics related to the TD3 rate, as the price of such product would reflect the overall demand for crude oil in South-east Asia. However, in the early literature, gasoline price has also been used as a proxy for supply, contemplating that a rise in the gasoline price would increase the operational costs of shipowners, which in turn forces more vessels into slow steaming or lay-up (Tinbergen, 1934). Accodingly, the relationship between the price of retail gasoline in Singapore and the TD3 rate is not obvious.

3.1.2 Tonne-mile demand

In order to properly account for the distance effect on demand, we calculate tanker demand on a tonne-mile basis. As underlined in Aadland et al. (2017), the main advantage of using AIS-data in the context of shipping market analysis is that such data allows for accurate estimates of tonne-mile on a per shipment basis. However, for an analysis of the formation of freight rates, tonne-mile is only relevant on a more aggregate level. With no prior guidance on how to aggregate tonne-mile that capture demand dynamics specifically related to the TD3 rate, we have to consider the following: First, whether to separate tonne-mile based on the origin of the shipment and target distinctive regions, countries or routes, contemplating that there could be significant differences in local and global demand dynamics. Second, how to properly account for potential cross demand between different tanker types, keeping in mind that subsegments within the tanker market are not completely isolated from each other (Stopford, 2009).

As the TD3 route refers to VLCC laden trips between AG and Japan, we would expect the TD3 rate to be more sensitive to tonne-mile arising in AG than in other parts of the world. Hence, we approximate the local demand for VLCC vessels as the aggregated sum of tonne-mile originating in AG (*AG_VLCC)*. However, Babri et al. (2017) suggest that only a fraction of international trade flows need to be subject to choice, as a large number of shipments are based long-term contracts or trading habits. Hence, in periods when vessels on long-term contracts in AG fail to meet the

[7]The data are gathered from Clarkson Research (2016) and measured in barrels per day (bpd).

local demand for tanker transportation, an increase in global tonne-mile would make it harder to get capacity elsewhere. This will essentially increase the demand for tanker freight in AG, and put upwards pressure on the TD3 rate. We have for this reason also included a variable on the aggregated sum of tonne-mile originating in other parts of the world (*global_VLCC)*.

The theory of economies-of-scale suggest that the marginal cost for transporting one barrel of crude oil always should be lower for a VLCC compared to a Suezmax, given that both vessels are fully loaded. Thus, charterers with sufficient amount of barrels to fill a VLCC should always prefer to fix one VLCC instead of the similar capacity of two Suezmax vessels. Keeping in mind that it takes several years from ordering a new VLCC until it is available in the market, we would expect substantial cross demand between different subsegments in periods with undercapacity.[8] Hence, we reckon an increase in the price for VLCC freight with positive shifts in the demand for vessel types other than VLCC. In order to account for this, we include variables on the aggregated sum of tonne-mile from vessel types other than VLCC, both originating in AG (*AG_other)* and in other parts of the world (*global_other)*.

3.1.3 Operational efficiency

In line with Wergeland and Wijnolst's (1996) notion of the most important elements of operational efficiency, we have taken sailing speed and load factor into consideration. Average sailing speed for the overall VLCC fleet (*fleet_speed*) is extracted from a more complex dataset and provided by Vit Prochazka, while load factor on a per shipment basis (*load_factor*) can be directly derived from the AIS-data. Ideally, we would also have a measure for waiting time per trip, but this is neither observable in the AIS-data nor is it possible to derive an adequate proxy from the information provided. We only consider measures of operational efficiency in relation to VLCC shipments, as this is what we *a priori* would expect to be most relevant to the TD3 rate. Neither would it make sense to consider efficiency measures across vessel segments, as distinct technical specifications and trade routes makes them differ by nature.

[8]Except from the differences in size, the different tanker types essentially offer the same product, and it will for this reason be competition among the subsegments in the tanker market.

Sailing speed is generally decided upon through optimization, unless otherwise specified in the contract. In theory, there exists a positive relationship between vessel speed and freight rates: When freight rates are low, charterers reduce sailing speed to trim operating expenses related to fuel consumption. Conversely, when freight rates are high, the alternative cost of time exceeds the gains from fuel optimization, which in turn promotes a higher sailing speed (Ronen, 1982). Following this argument, freight rates seems to lead changes in speed, not the other way around. However, when all available tonnage is in use, the short-term supply can only be altered by higher speed and more efficient operation (Cheng et al., 2010). Hence, we would expect an increase in average sailing speed for the overall VLCC fleet (*fleet_speed*) to give rise to more supply, leading to higher competition among VLCC vessels and lower TD3 rates.

Even though sailing speed is not directly observable in the AIS-derived data, we utilize information on time spent at sea and the distance between port pairs to derive the implied speed for each shipment (*voyage_speed*). Due to waiting time in port and time spent on loading or discharge, hours laden[9] will be excessively high relative to the actual time spent in open waters, which essentially will lower the implied speed calculation. Thus, we might argue the implied speed calculation (see *Equation 1*) to simultaneously pick up variation in sailing speed and waiting time without being able to separate the two effects. While an increase in sailing speed would increase the supply of tanker freight, the effect of more waiting time per shipment would be the opposite, leading to conflicting expectations on the possible impact of this variable on TD3 rates.

$$Voyage\ speed = \left(\frac{Distance}{Hours\ laden}\right) \tag{1}$$

Equation 2 shows how we have derived the load factor *(load_factor)* for individual shipments. The maximum cargo capacity for a VLCC vessel is approximately 2.1 million barrels, while cargo intake can be directly observed in the AIS-derived data. It is worth mentioning that each observation in the AIS-data pertains to a shipment from a single loading terminal to a single discharge terminal,

[9] $Hours\ laden\ = (Discharge\ date\ - Load\ date) \times 24$

leading to a cargo size that is significantly below maximum cargo capacity for a tanker visiting more than one loading/discharge terminal before reaching its final destination. In order to account for this nuance, we have sorted our dataset such that multiple shipments are merged to a single observation, with the cargo size being the aggregated sum of the observations involved.[10]

$$Load\ factor = \left(\frac{Cargo\ intake}{Cargo\ capacity}\right) \tag{2}$$

The rationale for including load factor is closely tied to the theory of economies-of-scale, in which we expect an inverse relationship between load factor and freight rates: When the tonnage of cargo transported increases, the marginal cost of transporting one extra unit of crude oil is expected to decrease, lowering the freight rate on a USD/tonne basis (Adland et al., 2016). On a more aggregate level, it can also be argued that that a higher average load factor for the overall fleet will increase the available supply of transportation, which in turn puts negative pressure on TD3 rates.

[10]This is further elaborated on in Section 6.2 on data preparation.

3.2 Subset selection method

3.2.1 Variable selection

Having identified a considerable number of variables that we believe to capture dynamics related to the formation of TD3 rates (see Table 1), we apply a subset[11] selection method to narrow down a reasonable set of predictors. As we want to investigate the ability to predict future TD3 rates, we are not interested in variables that exclusively respond to current events. Hence, all variables are included in lagged form, which allows us to identify the combination of variables that best incorporate feedback about future TD3 rates. The subset selection method we apply to our data is based on the leaps-and-bounds algorithm by Furnival and Wilson (1974), which returns the best subsets based on linear regression in accordance with some prespecified selection criteria. For the purpose of this thesis, we have chosen to base the variable selection on R^2 adjusted, Akaike's Corrected Information Criterion (AIC_C) and Mallow's C_p, in which the optimal subset would have the following characteristics: The highest value of R^2 adjusted, the smallest value of AIC_C, and a value of C_p that is close to the number of predictors +1, or the smallest Mallow's C_p relative to other variable combination. The three different selection criterion considered are defined by:

$$R^2adj = 1 - \frac{n-1}{n-k-1}\frac{RSS}{SST} \tag{3}$$

$$AIC_C = AIC + \frac{2k(k+1)}{n-k-1} \tag{4}$$

$$C_p = (n-m-1)\frac{RSS}{RSS_{FULL}} - (n-2p) \tag{5}$$

where *RSS* stands for *Residual Sum of Squares* and the RSS_{FULL} refer to the residual sum of squares in a model containing all predictors considered. For each predictor size *k*, the best model under each selection criteria is the model that minimizes the *RSS*. For the same predictor size all other terms are constant, which allows us to identify the combination of variables that obtain the smallest *RSS*.

[11] A subset will in this context be understood to as a specific combination of variables.

In order to investigate the relative performance of AIS-derived data, we apply the subset selection method on two different pools of variables: The first pool of variables contains the full amount of variables considered, while the the second is restricted to only include variables from publicly available data. Because the leaps-and-bounds method is based on a linear regression, we find it useful to run an OLS on the best subset from each pool. This also allow us to investigate whether the impact of each variable considered is in line with our *a priori* expectations. The OLS results are estimated on a linear model of the form:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 \ldots \beta_n X_n \tag{6}$$

where Y is the response variable, which in this case is the TD3 freight rate. The coefficient (β_n) can be interpreted as the change in freight rate from a one unit increase in X_n, holding all other independent variables constant. The alpha (α) corresponds to the intercept, indicating the value of the response variable when all $X_n = 0$.

3.2.2 Unit root test

As the variable selection method is based on a linear regression, we have to test for possible unit root behaviour in the data. Running the Furnival and Wilson (1974) algorithm on non-stationary variables could result in spurious subset selection results, in which the algorithm interpret relationships among the variables as causal, even though they only exists at random. Investigating results from the Augemented Dickey-Fuller test reported in *Appendix A*, we find the following three variables to be non-stationary: *mideast_production*, *refinery_margin* and *gasoil_sing*. Hence, the above-mentioned variables is only considered in first-difference, as test statistics suggests all three variables to be first-difference stationary.

3.3 Forecasting models

3.3.1 VAR model

In order to predict future TD3 rates, we have applied variables from each subset to a Vector Autoregressive (VAR) model. The VAR fits a multivariate time series regression of each dependent variable on lags of itself and lags of all the other variables (Zivot and Wang, 2006). It often provides superior forecasting performance to those from univariate models and has proven useful for describing the dynamic behaviour among time series. After running the VAR model, we perform diagnostic tests to determine whether our estimates are reliable. We conduct a stability test for the model, look for autocorrelation among the residuals and perform a Jarque-Bera test. This allows us to control for stability and normally distributed residuals in the model. The basic p-lag vector autoregressive (VAR(p)) has the form

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + ... + A_p Y_{p-t} + \upsilon_t \tag{7}$$

where A_i are $(K * K)$ coefficient matrices for $i = 1, .., p$ and υ_t is an $(K * 1)$ unobservable zero mean white noise process, also called the error term.

Furthermore, we will apply a Granger causality test to investigate whether causal relationships with respect to the TD3 rate are captured by any of the variables included. A variable x is said to cause y if past values of x help predict the current level of y, given all other appropriate information (Granger, 1969). If we find that y in fact causes x, then it is unlikely that information on x would help predict the freight rate, given the past history of y. In a classical philosophical sense, Granger causality is not identical to causation, but the likelihood for such an event to occur is demonstrated in this test. The test requires estimating the following two equations, where m is the number of lags, while ε and η are considered to be two uncorrelated white noise series (error terms):

$$y_t = \sum_{j=1}^{m} a_j X_{t-j} + \sum_{j=1}^{m} b_j Y_{t-j} + \varepsilon_t \tag{8}$$

$$x_t = \sum_{j=1}^{m} c_j X_{t-j} + \sum_{j=1}^{m} d_j Y_{t-j} + \eta_t \tag{9}$$

3.3.2 ARIMA model

When the variable we wish to forecast is the only variable, i.e. when Y_t is scalar, the VAR model is reduced to *AR(p)*. Including the integrated order (I(d)) and the moving average term (*MA(q)*), we build an ARIMA model of order p, d, q. The *AR*-term(p) accounts for the autoregressive order in the time series, d is considered if the time-series exhibits unit root behaviour and q is meant to capture the effect of autocorrelation in the variable Y_t. Given a dependent time series ($Y_t : 1 \leq t \leq n$), the ARIMA model is expressed as

$$Y_t = c + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \theta_1 e_{t-1} + \ldots + \theta_p e_{t-p} + e_t \tag{10}$$

where Y_t is the TD3 rate explained in time t and c is the constant/intercept. The coefficients for each autoregressive parameter (p) and the number of errors (q) caused by lagged predicting, are explained by ϕ and θ, respectively.

The process of constructing an ARIMA starts by model identification, i.e. select the order of p, d and q that best fit the time series at hand. We find our univariate model to be autoregressive of order one, while the moving average term has a order of five. A stationary TD3 rate causes an integrated order of zero, and consequently the general ARIMA collapses into an ARMA model.[12] Further, we run diagnostic tests to determine whether the model is stable or not, and postestimation results show that the ARMA model both satisfy the stability condition and the invertibility condition.

[12]The procedure is similar to the Box-Jenkins approach (Box and Jenkins, 1968).

3.3.3 Evaluation of models

For evaluation purposes, we split our data into a training set (in-sample) and an evaluation set (out-of-sample). The multivariate VAR models are based on the in-sample data, whereas the evaluation is done by comparing the errors of the forecasts against the out-of-sample TD3 rate. Using the in-sample data, we estimate a one-step-ahead forecast throughout the evaluation period, stretching from August 2015 to mid-February 2016. The difference between the one-step-ahead forecast and the first evaluation set observation is seen as the forecasting error. For the next step ahead, the training set is extended with the first observation in the evaluation set and a new forecast is made. A new forecasting error is displayed, and the training set is once again augmented with the second observation in the evaluation set.

In order to examine the relative performance of the multivariate VAR models, forecast errors for a univariate ARMA are included as a benchmark. Performances are compared using the following error measures: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Even though there is no consensus on the most appropriate error measures, these three are widely used to assess the performance of forecasting techniques (Zhang and Hu, 1998). Their mathematical expressions are as follows:

$$MAPE = \left(\frac{1}{n} \sum_{t=1}^{n} \frac{Actual - Forecast}{|Actual|} \right) * 100 \tag{11}$$

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |Actual - Forecast| \tag{12}$$

$$RMSE = \sqrt{\frac{1}{n} \sum (Actual - Forecast)^2} \tag{13}$$

4 Data

4.1 Data preparation

The AIS-data are provided by Clipper Data Ltd. (2016) and contain information about 81,728 individual shipments of crude oil between January 2013 and mid-February 2016, and results from a combination of AIS-tracking of individual vessels and information on cargo size from lineup reports. After proper cleansing and necessary adjustments to the dataset, we end up with 53,116 observations in total. Information within the AIS-data deemed relevant is further converted into 162 weekly time series, with variables on operational efficiency as the weekly average of commencing shipments, and tonne-mile as the weekly sum that arises during the week. All other time series considered are either collected from Clarksons Research (2017) or Bloomberg Professional Services (2017). In order to maintain consistency throughout the paper, all time series considered are stated in a weekly number. If we were to use monthly, quarterly or semi-annual time series in our analysis, the number of observations would become very low due to the relatively short time span of the AIS-data. Hence, we have deemed an analysis of weekly time series most appropriate for the data at hand.

Each observation in the AIS-derived data pertains to a shipment from a single loading terminal to a single discharge terminal. Hence, if one of the tankers in question visits more than one loading/discharge terminal before reaching its final destination, this will be accounted for as multiple shipments in the AIS-data. This nuance is consequential for the purpose of this thesis in two ways: First, the average shipment size will appear to be substantially below the average capacity of the tanker, which in turn would bias our estimates of load factor. Second, estimates of tonne-mile will be artificially high as the distance effect on demand are factored in more than one time.

Accounting for this nuance, we have sorted our data in the following way: If a specific vessel visits more than one loading terminal on the same day, the cargo size from each loading terminal are aggregated to one observation, with the longest distance remaining. Similarly, if a specific vessel visits more than one discharge terminal on the same day, the cargo size from each discharge terminal is aggregated to one observation, with the longest distance remaining.[13] We are aware that this might not adequately solve the problem, but as we can see from Table 2 the variable on load factor averages to 86% with a standard deviation of 3.55%, which is similar to what is found in Thomassen and Oestensen (2016). In addition, Wergeland and Wijnolst (1996) suggests that the overall load factor hardly ever will be higher than 95%, in which the highest observed load factor in our data is 95.5%.

To our knowledge, there exists no weekly data on VLCC fleet development (*VLCC_fleetchange*) and Middle East crude oil production (*mideast_production*). Hence, we have used a cubic spline interpolation to create weekly estimates between the known monthly data points (Anacleto et al., 2013). We are aware that the cubic spline interpolated variables will introduce a systematic source of autocorrelation in their respective time series. However, with the overall cargo capacity being the most important determinant of freight supply, and the fact that 82% of all crude oil imported to Japan is shipped from the Middle East, we find it necessary to include both variables as part of our analysis. To further substantiate our decision to include these variables, both Assmann et al. (2016) and Tham (2008) find VLCC fleet development and Middle East crude oil production useful to predict TD3 rates.

[13]The first sorting reduce the number of observations with 20,935, while the second sorting reduce the number observations with 7,677, leading to 53,116 observations in total.

4.2 Data description

Table 2 summarizes the descriptive statistics[14] for all variables considered, categorized by the dependent variable, tonne-mile demand, operational efficiency and publicly available data. In addition, we have included a correlation matrix between the variables, presented in *Appendix B*.[15]

Table 2: Descriptive statistics

Variables	Unit	obs	mean	Std. Dev.	Min	Max
Dependent variable						
TD3_rate	$/tonne	162	14.05	3.82	8.01	24.26
Tonne-mile demand						
AG_VLCC	tm MMM	162	71.95	9.11	47.33	101.23
AG_other	tm MMM	162	9.24	2.53	3.31	17.04
global_VLCC	tm MMM	162	30.77	7.12	12.68	49.83
global_ other	tm MMM	162	40.40	4.11	30.47	50.13
Operational efficiency						
fleet_speed	knot	162	12.43	0.26	11.86	14.03
voyage_speed	knot	162	9.40	0.37	7.57	10.31
load_factor	%*100	162	86.21	3.55	71.12	95.55
Publicly available data						
mideast_production	bpd MM	162	30.79	0.82	29.46	32.70
VLCC_fleetchange	dwt MM	162	0.08	0.15	-0.24	0.56
spot_VLCC	dwt M	162	1,234.46	1,021.33	0	4,922
gasoil_sing	$/bbl	162	94.05	26.48	37.17	132.31
diff_ussing	$/bbl	162	43.33	5.12	28.99	56.53
refinery_margin	$/bbl	162	12.03	4.05	1.86	21.55
Brent/Dubai_spread	$/bbl	162	2.33	1.86	-3.73	8.71

Source: Authors' calculations.

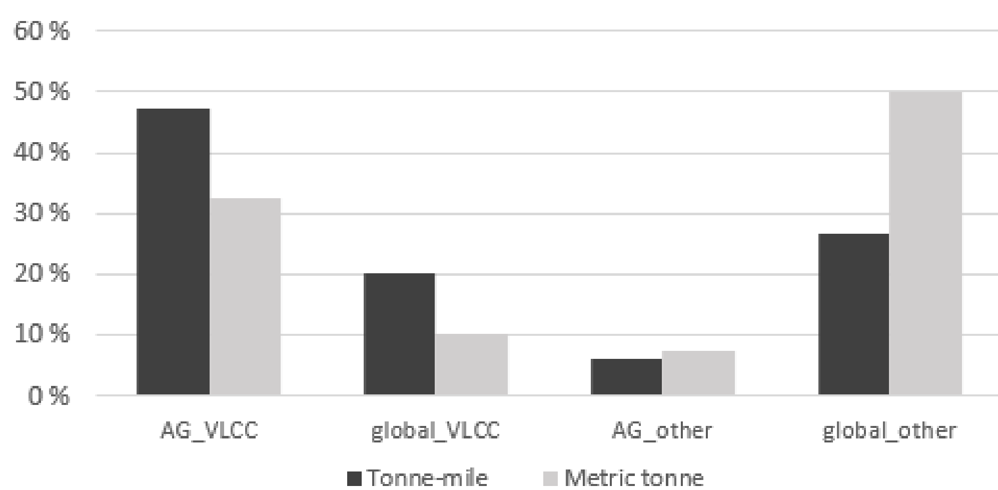
[14]Each M indicates 1,000.

[15]All calculations are performed using the statistical software STATA.

The dependent variable *(TD3_rate)* indicates the cost of shipping one metric tonne and is found by using the following formula: $USD/mt = (\frac{WS}{100}) * FlatRate$, where the annual flat rates are provided by the Worldscale Association (Worldscale, 2017) and fixtures reported in worldscale (*WS*) gathered from Clarksons Research (2017). The freight rate incurs high variation in the period between year-end 2012 and mid-February 2016, fluctuating between a minimum of 8.01 USD/tonne and a maximum of 24.26 USD/tonne.

VLCCs are handling two-thirds of the tonne-mile demand within the period, with a standard deviation of 12.41 billion tonne-miles. VLCCs within AG has a mean of 71.95, occupying 47.22% of the total demand, but only 32.43% total metric tonne is exported (see Figure 1). Long haul routes are usually more present from AG than most other parts of the world, hence the different fractions of metric tonnes and tonne-miles in regards to the total amount. From the correlation matrix in *Appendix B*, both *AG_VLCC* and *global_VLCC* have a negative correlation with their respective *other* variables. Even though correlation give no indication of causality, the negative relationship could indicate cross demand relationships, where some would prefer chartering two suezmaxes rather than one VLCC and vice versa.

Figure 1: Total tonne-mile and metric tonne



Source: Authors' calculations

We notice from the variables regarding operational efficiency, a significant difference in speed, where voyage speed averages 9.4 knots and fleet speed centres around 12.43 knots. In Section 3.1.3 we argue that implied speed can pick up some of the waiting days exposed to a shipment, which in this case is illustrated in the difference of 3 knots. Both variables have an opposite correlation in regards to the freight rate, and we cannot know whether an increase in potential waiting days causes the rate to decline or if an increase in the freight rate force the ships to be more efficient. Further, load factor seems stable during the period, averaging 86% utilization of the ships' capacity. Having a negative correlated load factor indicates the marginal cost of transporting òne extra unit of crude is expected to decrease when the cost of freight increases.

Statistics from publicly available data show a positive mean in *VLCC_fleetchange*, weekly increasing with an average of 80,000 DWT during the period. From VLCC spot in AG there seems to be an average of 5 VLCCs available for chartering, assuming a VLCC of 250,000 DWT. Moreover, the cost of gasoline in Singapore is highly negatively correlated with the dependent variable, and because lower oil prices have a positive impact on global trades, it consequently impacts transportation.

5 Empirical results

5.1 Subset selection results

Furnival and Wilson's (1993) leaps-and-bonds algorithm suggests the following two linear regression models of lagged variables to best incorporate feedback about future TD3 spot rates.[16] As shown in Table 3, model (2a) is based on the full amount of variables considered, while model (2b) is restricted to only include variables from publicly available data.

Table 3: Linear regression results

	(2a) TD3_rate		(2b) TD3_rate
L1spot_VLCC	-0.00140*** (-6.55)	L1spot_VLCC	-0.0015*** (-5.20)
L2global_other	0.1620*** (3.89)	L3Δmideast_production	7.039** (3.44)
L2global_VLCC	0.0878** (3.21)	L3VLCC_fleetchange	-11.193* (-1.83)
L3global_VLCC	0.0889** (2.91)	L3spot_VLCC	-0.0004* (-1.91)
L1load_factor	-0.1624** (-3.27	L1diff_ussing	-0.1127 (-1.66)
L1AG_VLCC	0.0330 (1.60)	L2Δgasoil_sing	-0.4005*** (-3.61)
L1global_other	0.0896* (1.96)	L2BD_spread	-0.6153*** (-3.91)
L3spot_VLCC	-0.0005** (-2.37)	L2VLCC_fleetchange	5.8165 (1.13)
L1global_VLCC	0.0833** (3.00)	L2Δrefinery_margin	0.303** (2.19)
L3voyage_speed	-0.9889 (-1.65)	_cons	22.414*** (7.53)
L3load_factor	-0.1395** (-2.68)		
L1fleet_speed	-2.0816** (-2.45)		
L3VLCC_fleetchange	-3.4237** (-2.54)		
_cons	56.7377*** (4.00)		
R^2-*adj.*	0.62	R^2-*adj.*	0.42
AIC_C	586.39	AIC_C	636.55
Mallow's C_p	3.67	Mallow's C_p	7.56

*** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations

[16]Both linear regression models are estimated with heteroskedasticity-consistent standard errors.

Comparing the realized values of AIC_C, Mallow's C_p and R^2*-adjusted* for the two models in question, we find model (2a) to perform relatively better than model (2b) in our in-sample period. With higher explanatory power of future TD3 rates and better in-sample fit, the model based on the full amount of variables considered performs better along all relevant criteria. Under the assumption that a good representation of the past leads to reasonable anticipations of the future,[17] a combination of variables on tonne-mile demand, operational efficiency and publicly available data exhibits most promise in an out-of-sample forecast. Investigating more extensive subset selection results reported in *Appendix C*, we observe that the relative improvement for each quantity of predictors is reduced when more variables are included. This suggests that the most powerful predictors is selected first and the more inferior later. To emphasize the relative importance of each predictor, the variables in Table 3 is presented in the same order as they were selected by the subset selection method, in which both models suggests *L1spot_VLCC* to be the most powerful predictor. Because many of the models are similar in fit, we do not lose or gain much explanatory power choosing one or another. For this reason, we make an arbitrary choice among them, and choose the most desirable value of AIC_C and Mallow's C_p, keeping in mind that the optimal R^2-adjusted includes slightly more predictors.

Investigating the variables assigned to each of the two models, we only find VLCC available spot in AG and the weekly change in overall VLCC capacity present in both models. Our estimates of one week lagged spot VLCC are significant on a 99% level, and suggest that one additional VLCC available for chartering in AG reduces prevailing TD3 rates with approximately 0.35 USD/tonne.[18] Three week lagged change in the overall VLCC capacity is significant on a 95% level, which suggests that sudden changes to the global supply of VLCC vessels first are fully reflected in the TD3 rate 3 weeks into the future. However, the impact of changes to the fleet is significantly different across the two models, indicating a reduction of 0.86 USD/tonne and 2.79 USD/tonne respectively, for each additional vessel added to the fleet.

[17] A good in-sample fit does not necessarily lead to a good forecast, and vice-versa: An overfit model will typically have very small in-sample errors, but be terrible at forecasting.

[18] $-0.00140 * 250 = 0.35$ USD/tonne, assuming a VLCC of 250,000 DWT.

Except for tonne-mile originating in AG from vessels types other than VLCC, all tonne-mile variables are included in model (2a). Estimates of VLCC tonne-mile originating from other regions than AG are significant on all three lags, having an impact of similar magnitude on prevailing TD3 rates. Further, we notice that VLCC tonne-mile originating in AG fail to significantly incorporate feedback about about future TD3 rates, which is in contrast to our *a priori* expectations. This might suggest that the large amount of vessels stationed in AG on long-term contracts makes the TD3 rate more sensitive to VLCC tonne-mile originating in other parts of the world. With estimates for the variable on tonne-mile from vessel types other than VLCC being significant on 99% level, there also seems to be substantial cross demand between subsegements in the tanker market. All variables on operational efficiency are also represented in model (2a), with one week lagged load factor and one week lagged fleet speed being significantly able to explain future TD3 rates. In line with *a priori* expectations, an increase in average sailing speed by one knot reduce TD3 rates with 2.1 USD/tonne, while increasing the average load factor with one percentage point reduce the freight rate in question with 0.162 USD/tonne.

In model (2b), both Brent-Dubai spread and weekly changes in Singapore gasoline prices are significant on a 99% level. Estimates of Brent-Dubai spread suggests a reduction in the TD3 rate when the price for AG crude oil gets relatively cheaper. This is not in line with our *a priori* expectations and might indicate that the cross demand between the two crude oil types is not strong enough to have a visual impact on the freight rate. Estimates of weekly changes in Singapore gasoline price suggests that such dynamics are more related to the supply of freight than the demand, which is in line with findings in Tinbergen (1934). Estimates of Middle East crude oil production and weekly changes in refinery margins are both significant on a 95% level, and exhibits a positive relationship with respect to TD3 rates.

5.2 VAR results

Having identified the variables that best incorporate feedback about future TD3 rates, we want to specify our two VAR models based on these variables. Keeping in mind that we only have 132 observations in our in-sample data, we have to carefully consider whether it would be sensible to include all variables suggested: *linear regression (2a)* suggest 8 different variables,[19] with a maximum lag length of 3, to best incorporate feedback about future TD3 rates. If we were to specify a VAR model based on these variables, the model would be required to calculate 252[20] estimates in total, which far exceeds the number of observations in our in-sample data. Similarly, *linear regression (2b)* suggests 7 different variables, with a maximum lag length of 3, which in a VAR model would equal 154 estimates. Hence, including all variables suggested by Furnival and Wilson's (1993) leaps-and-bonds algorithm would in both cases lead to negative degrees of freedom, which in turn would penalize the accuracy of the VAR estimates and impair the predictive ability of the two models. (Baltagi, 2008).

Following this argument, we have specified our two VAR models in the following way: For each of the two VAR models we include the 6 most powerful predictors of TD3 rates (see Table 3), and for the number of estimates to not exceed the number of observations in our in-sample data, the maximum lag order is set to 3. Further, we use lag-order selection statistics to identify the most appropriate lag-length for each of the two models, in which both Akaike's Information Criterion (AIC) and Bayesian information criterion (BIC) suggest 2 lags. The model based on the full amount of variables considered will now be referred to as the *preferred VAR*, while the one solely based on publicly available data will be the *restricted VAR*. Postestimation results (see *Appendix D*) for our preferred VAR model indicate the model estimates satisfy the stability condition, and there are no signs of autocorrelation in the residuals for either of the two lag-orders. However, due to severe kurtosis in the disturbance term, the model fails to satisfy the Jarque-Bera test of normally distributed residuals. Turning to the restricted VAR, neither the Jarque-Bera test are satisfied, nor can we reject the possibility of autocorrelation in the error term.

[19]The total number of variables selected by the subset selection method is 13, but as each variable are included with more than one lag-order some of the variables appear more than once.

[20]Each matrix of coefficients for a given lag length is $8 * 8 = 64$, and the vector of constant has 8 elements, so a total of $64 * 3 + 8 = 252$ estimates are calculated.

In order to determine if our preferred VAR model or the restricted VAR can be used for predictive purposes, we investigate whether causal relationships with respect to the TD3 rate are captured by any of the variables included. Table 4 below presents the variables selected to each of the two VAR models and postestimation results from a Granger causality test.

Table 4: Granger causality test

Preferred VAR	Prob > chi2	Restricted VAR	Prob > chi2
spot_VLCC	0.465	spot_VLCC	0.608
global_other	0.055	B/D_spread	0.044
global_VLCC	0.153	fleet_netchange	0.310
load_factor	0.258	diff_ussing	0.123
AG_VLCC	0.775	Δmideast_production	0.438
voyage_speed	0.023	Δgasoil_sing	0.025
ALL	0.031	ALL	0.307

Source: Authors' calculations

Starting with our preferred VAR model, results from the Granger causality test suggest that the two lags of voyage speed is the only variable that exhibits a causal relationship with respect to TD3 rates. The two lags of tonne-mile originating outside AG also show some promise, being significant on a 90% level. Despite mediocre feedback from individual variables, test statistics for the joint significance of the full set of endogenous variables suggest the model is useful for forecasting purposes. For the restricted VAR model, both weekly changes in Singapore gasoline prices and Brent/Dubai spread are significant on a 95% level. However, unlike our preferred VAR model, the combination of variables included in the restricted VAR show no sign of causality. Hence, we would expect our preferred VAR model to perform best in an out-of-sample prediction.

5.3 Forecast evaluation

The out-of-sample performance of our preferred VAR model and the restricted VAR are summarized Table 5, and are evaluated based on error measures introduced in Section 3.3.3. In order to examine the relative performance of the two multivariate VAR models, forecast errors for a univariate ARMA are included as a benchmark.

Table 5: Forecast error - TD3 spot rate

Error measure	Preferred VAR	Restricted VAR	ARMA
MAE	2.11	2.11	2.38
MAPE	0.124	0.125	0.143
RSME	2.69	2.70	2.91

Source: Authors' calculations

Comparing the realized forecast errors for the univariate ARMA and the restricted VAR, we find the univariate ARMA performs significantly worse along all error measures considered. This indicates that weekly TD3 rates are somewhat predictable between August 2015 to mid-February 2016, and that publicly available data on freight rate determinants incorporate superior information about future spot rates. This result is nothing new to the existing literature on freight rate forecasting, as both Batchelor et al. (2007) and Kavussanos and Nomikos (1999) find evidence in favour of predictable spot rates when utilizing the cointegrated relationship between spot freight rates and tradable forward contracts. However, contrary to our results, Assmann et al. (2016) find univariate models to perform relatively better than multivariate VAR models based on publicly available data. One reason for the conflicting results might be the relative difference in sample size: Assmann et al. (2016) use monthly data between 2006 and 2012, amounting to approximately 72 observations in total, which is significantly below the 132 observations that we have utilized in our VAR models. As the accuracy of VAR estimates are highly sensitive to loss of degrees of freedom when the sample size is small, the relatively low number of observations used in Assmann et al. (2016) might favour the performance of the univariate models.

In order to specifically examine the predictive ability of information derived from AIS-data, we compare the out-of-sample performance of both multivariate models considered. In line with Granger causality results and our *a priori* expectations, we find that the VAR model including AIS-derived variables performs better than the VAR model solely based on publicly available data. With slightly more desirable values of MAPE and RMSE, and more or less equal values of MAE, we have established weak evidence in favour of using AIS-data for forecasting purposes. However, investigating Diebold-Mariano comparison of forecast accuracy (Baum, 2003)[21], test statistics suggest the mean difference in out-of-sample performance is statistically insignificant across all three error measures. Hence, the slightly better performance between August 2015 and mid-February 2016 might not hold for other periods in time.

As both multivariate models show promise in forecasting spot TD3 rates, we also want to investigate whether the same models could be applied to forecast the price of FFA contracts. A reasonable forecast of FFA contracts would not only assist charterers and shipowners in when or whether to hedge their position, but also could be used to earn speculative profits. In an efficient market, the price for tradable assets should reflect all publicly available data about the market already, but with the AIS-derived data only made available for a small number of people, it would not be considered publicly available data *per se*. However, investigating realized forecast errors reported in Table 6, we find no evidence of forecastable +1 month FFA prices, as neither our preferred VAR nor the restricted VAR can outperform the univariate ARMA. Hence, in line with previous literature, we find the market for FFA contracts to be seemingly efficient.

Table 6: Forecast error - TD3 +1 month FFA

Error measure	Preferred VAR	Restricted VAR	ARMA
MAE	1.40	1.37	1.14
MAPE	0.079	0.077	0.067
RSME	1.80	1.77	1.55

Source: Authors' calculations

[21] Allow the null-hypothesis of equal accuracy to be tested.

6 Concluding remarks

The purpose of this thesis has been to investigate whether information derived from AIS-data incorporates superior information about future freight rates. Specifically, we have assessed if such data improve the ability to predict TD3 spot rates between August 2015 and mid-February 2016. The ability to anticipate short-term fluctuations in freight rates is a key component to long-term profitability for both shipowners and charterers, making the purpose of this thesis an important objective.

There are some important takeaways from our empirical results that we want to highlight. First, Furnival and Wilson (1974) leaps-and-bonds algorithm find information derived from AIS-data to significantly improve the ability explain future TD3 rates. With 10 out of 13 variables within the optimal subset being directly derived from AIS-data, the variables seems to capture important dynamics related to the formation of TD3 rates. Second, unlike our *a priori* expectations, tonne-mile originating in AG fail to significantly incorporate feedback about future TD3 rates. Hence, the price for tanker freight between AG and Japan seems to be unaffected by the amount of occupied tonne-mile originating in the same region. One reason for the conflicting result might be that the large amount of vessels stationed in AG on long-term contracts makes the TD3 rate more sensitive to tonne-mile originating in other parts of the world. However, in line with our *a priori* expectations, we find tonne-mile from vessel types other than VLCC to significantly explain future TD3 rates, which suggests considerable cross demand between subsegments in the tanker market. Third, variables on operational efficiency also shows promise in the context of shipping market analysis. Variables on load factor and fleet speed both incorporates feedback about future TD3 rates, while the variable on voyage speed exhibits a causal relationship with respect to TD3 rates. Final, we find weak evidence in favour of using information derived from AIS-data for predictive purposes, as the multivariate VAR model based on the full amount of variables considered outperform the univariate ARMA and the multivariate VAR model solely based on publicly available data. In line with previous literature, we also find the market for FFA contracts seemingly efficient, as neither of our preferred VAR nor the restricted VAR can outperform the univariate ARMA in forecasting +1 month FFA prices.

Even though our results show promise, we acknowledge there are several limitations to our research, in which the most decisive are related to the short time span of the AIS-data. For the purpose of this thesis, neither load factor, sailing speed nor tonne-mile are interesting on a per shipment basis, and each must be converted into one-dimensional time series in order to make sense. However, due to the relatively short time span of the AIS-data, even an analysis of weekly time series leads to a relatively low number of observations, which in turn is consequential for how we can specify our models. First, a great number of time series on shipping market dynamics are only available in a monthly interval, which reduce the number of variables considered in our analysis. Second, the chosen lag-length should ideally reflect a whole shipping cycle, which for all practical considerations is impossible with the data at hand. Other limitations are mainly related to how we have prepared the AIS-derived data for analyses purposes. Due to the complexity of such data, we have not been able to properly account for multiple shipments, which essentially lead to biased estimates of tonne-mile and load factor. Moreover, we have not been able to obtain sailing distance for all 81,728 shipments, which reduces the number of observations made possible to use. Ideally, we would also include a measure for waiting time per trip in our analysis, but this is neither observable in the AIS-data nor is it possible to derive an adequate proxy from the information provided.

Further research using AIS-data is required to exploit the full potential of the information within it. With AIS-data stretching over a longer period of time becoming available, it would be possible to include more variables and a higher lag-order, which in turn could lead to a more desirable model. Moreover, if a more desirable model were found, it would be interesting to extend on the timing decision for shipowners of when to fix, or whether to fix on a short or long voyage. We also acknowledge that the current method used for estimating tonne-mile and measures of operational efficiency may not be optimal, and could possible be resolved using more sophisticated programming techniques. Another potential area for future research using AIS-data would be to investigate whether physical movements of crude oil lead changes in the price of the commodity, keeping in mind that the such data covers all seaborne trade of this commodity.

References

Aadland, R., Haiying, J., and Strandenes, S. P. (2017). Are ais-based trade volume estimates reliable? the case of crude oil exports. *Maritime Policy and Management*, 0(0):1–9.

Adland, R., Haiying, J., and Strandenes, S. P. (2016). The determinants of vessel capacity utilization: the case of brazilian iron ore exports. *Transportation Research Part A: Policy and Practice*, 94.

Anacleto, O., Queen, C., and Albers, C. (2013). Forecasting multivariate road traffic flows using bayesian dynamic graphical models, splines and other traffic variables. *Australian and New Zealand Journal of Statistics*, 55(2):69–86.

Assmann, L., Andersson, J., and Aadland, R. (2016). Using fundamental variables for tanker freight forecasting - a comparison with time series models.

Babri, S., Jørnsten, K., and Viertel, M. (2017). Application of gravity models with a fixed component in the international trade flows of cloal, iron ore and crude oil. *Maritime Economics Logistics*, 19(2):334–351.

Baltagi, B. H. (2008). *Econometrics*. Springer-Verlag Berlin Heidelberg, 5th edition.

Batchelor, R., Alizadeh, A. H., and Visvikis, I. (2007). Price discovery, causality and forecasting in the freight futures market. *International Journal of Forecasting*, 23(1):101–114.

Baum, C. F. (2003). Dmariano: Stata module to calculate diebold-mariano comparison of forecast accuracy. https://ideas.repec.org/c/boc/bocode/s433001.html. Retrieved 1 June 2017.

Beenstock, M. and Vergottis, A. (1993). *Econometric Modelling of World Shipping*. Chapman and Hall, London, UK.

Box, G. E. P. and Jenkins, G. M. (1968). Some recent advances in forecasting and control. *Royal Statistical Society*, 17(2):91–109.

Cheng, T., Luen, Y. H., and Lai, K. H. (2010). *Shipping and Logistics Management*. O'Reilly Media, Inc., Sebastopol, CA.

Furnival, G. M. and Wilson, R. W. (1974). Regressions by leaps and bounds. *Technometrics*, 16(4):499–511.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

IMO (2017). Automatic identification systems (ais). http://www.imo.org/en/OurWork/safety/navigation/pages/ais.aspx. Retrieved 27 March 2016.

Kavussanos, M. G. and Nomikos, N. K. (1999). The forward pricing function of the shipping freight futures market. *Journal of Futures Markets*, 19(3):353–376.

Kavussanos, M. G. and Visvikis, I. D. (2006). Shipping freight derivatives: a survey of recent evidence. *Maritime Policy Management*, 3(33):233–255.

Koopmans, C. T. (1939). *Tanker Freight Rates and Tankship Building: An Analysis of Cyclical Fluctuations*. F. Bohn and P. S. King, Harleem, Netherlands.

Research, C. (2017). Shipping intelligence network. www.clarksons.net. Retrieved 1 February 2017.

Ronen, D. (1982). The effect of oil price on the optimal speed of ships. *The Journal of the Operational Research Society*, 33(11):6.

Stopford, M. (2009). *Maritime economics*. Routledge Taylor Francis Group, London and New York, 3rd edition.

Tham, E. (2008). Leading indicators for arabian gulf oil tanker rates. *OPEC Energy Review*, 32(2):139–149.

Thomassen, K. and Oestensen, E. R. (2016). *Energy Efficiency and Voyage Rates: An assessment of a potential two-tier market. The case of the VLCC spot market*. Master Thesis. Norwegian School of Economics.

Tinbergen, J. (1934). Tonnage and freight. *The nederlandsche Conjuncturr*, 44(2):271–288.

Wergeland, T. and Wijnolst, N. (1996). *Shipping*. Delft University Press, Mekelweg 4, 2628 CD Delft.

Worldscale (2017). Tanker nominal freight rate scale. recieved through email. Retrieved 15 February 2017.

Zhang, G. and Hu, M. Y. (1998). Neural network forecasting of the british pound/us dollar exchange rate. *Omega*, 26(4):495–506.

Zivot, E. and Wang, J. (2006). *Modelling Financial Time Series with S-PLUS*. Springer New York, New York, NY.

Appendicies

Appendix A - Dickey-Fuller test for unit root

Variable	Test Statistics	1 % Critical Value	5 % Critical Value	p-value
mideast_production	-0.132	-3.490	-2.886	0.946
refinery_margin	-1.74	-3.490	-2.886	0.072
gasoil_VLCC	0.717	-3.490	-2.886	0.990

Source: Authors' calculations

Variable	Test Statistics	1 % Critical Value	5 % Critical Value	p-value
Δmideast_production	-3.479	-3.490	-2.886	0.009
Δrefinery_margin	-14.248	-3.490	-2.886	0.000
Δgasoil_VLCC	-10.092	-3.490	-2.886	0.000

Source: Authors' calculations

Appendix B - Correlation matrix

Correlation	TD3_rate	TD3_FFA	AG_VLCC	AG_other	global_VLCC	global_other	fleet_speed	voyage_speed
TD3_rate	1.0000							
TD3_FFA	0.8527	1.0000						
AG_VLCC	0.2045	0.1995	1.0000					
AG_other	-0.2203	-0.1796	-0.0156	1.0000				
global_VLCC	0.3520	0.2865	0.1582	-0.1841	1.0000			
global_other	0.2410	0.3333	0.0881	0.0806	-0.0157	1.0000		
fleet_speed	0.3572	0.2747	0.1628	-0.1825	0.1276	0.1896	1.0000	
voyage_speed	-0.1222	-0.1130	0.1442	0.0025	-0.0112	-0.0335	-0.0119	1.0000
load_factor	-0.2198	-0.2319	0.0889	-0.0495	-0.0139	-0.1487	0.0947	0.0813
gasoil_sing	-0.5769	-0.6684	-0.1793	0.2290	-0.2656	-0.2324	-0.5689	0.2714
diff_ussing	0.0606	0.1644	0.0721	0.0737	0.2056	0.0089	0.1028	-0.0548
BD_spread	-0.1237	-0.2338	-0.0763	0.1047	-0.1930	-0.1887	-0.1244	0.0164
refinery_margin	0.1190	0.0675	0.0363	-0.0249	0.1301	-0.0758	0.1898	-0.0800
ME_production	0.3508	0.4501	0.1192	-0.1370	0.1432	0.0770	0.5089	-0.2474
VLCC_fleetchange	0.0373	-0.0515	-0.0427	0.0803	0.0638	-0.1210	0.1411	-0.1411
spot_VLCC	-0.5753	-0.4089	-0.1014	0.2193	-0.0896	-0.0195	-0.4057	0.0593

Correlation	load_factor	gasoil_sing	diff_ussing	BD_spread	refinery_margin	ME_production	VLCC_fleetchange	spot_VLCC
load_factor	1.0000							
gasoil_sing	0.1424	1.0000						
diff_ussing	-0.0265	-0.5348	1.0000					
BD_spread	0.1792	0.3708	-0.3782	1.0000				
refinery_margin	0.1427	-0.1904	0.0250	0.1381	1.0000			
ME_production	0.1019	-0.7409	0.4500	-0.2221	0.4760	1.0000		
VLCC_fleetchange	0.1758	-0.1659	0.2347	0.0136	0.4693	0.3570	1.0000	
spot_VLCC	-0.0653	0.3072	-0.0252	0.0275	-0.0681	-0.2403	-0.1719	1

Source: Authors' calculations

Appendix C - Subset selection results

	Results from unrestricted pool			Results from restricted pool		
Preds	R2ADJ	C	AICC	R2ADJ	C	AICC
1	0.247735	107.6454	661.9789	0.247735	38.0987	661.9789
2	0.363282	71.76802	640.9069	0.306041	25.97057	652.3563
3	0.439088	48.75638	625.1805	0.338012	19.80082	647.2163
4	0.481657	36.32465	615.8422	0.368448	14.07242	642.1155
5	0.512308	27.73218	608.9216	0.373516	13.91028	642.2302
6	0.541657	19.68101	601.8808	0.394059	10.45886	639.0097
7	0.563806	13.93199	596.5356	0.408205	8.433152	637.1103
8	0.582928	9.190544	591.845	0.4153	7.9379	636.7777
9	0.595949	6.34308	588.9281	0.42197	**7.555257**	**636.5533**
10	0.602329	5.503833	588.1439	0.42304	8.353884	637.6395
11	0.61034	4.221034	586.8013	0.428343	8.287525	637.7755
12	0.614922	3.964872	586.6248	0.429145	9.155042	638.9856
13	0.619721	**3.668912**	**586.3872**	0.430609	9.891635	640.0743
14	0.622739	3.900461	586.7921	**0.431561**	10.73569	641.3164
15	0.622519	5.056417	588.3699	0.431233	11.83688	642.8933
16	0.622799	6.072066	589.808	0.430363	13.0437	644.6333
17	0.622721	7.187792	591.4096	0.426886	14.75583	647.0169
18	0.623193	8.152502	592.8557	0.424199	16.3034	649.2517
19	0.623414	9.187807	594.4304	0.420308	18.07256	651.7999
20	**0.623489**	10.26325	596.0973	0.415419	20.01805	654.6103

Source: Authors' calculations

Appendix D - Postestimation results

Postestimation results		
	Included AIS-derived data	**Excluded AIS-derived data**
Tests	**Prob >chi2**	**Prob >chi2**
Jaque-Bera*	0.004	0.000
Skewness*	0.375	0.000
Kurtosis*	0.01	0.000
Autocorrelation**	Lag 1: 0.1254 Lag 2: 0.2249	Lag 1: 0.000 Lag 2: 0.000
Eigenvalue stability condition	All eigenvalues inside unit circle	All eigenvalues inside unit circle

* H0: the disturbance in VAR is normally distributed

** H0: no autocorrelation at lag order

Source: Authors' calculations

Eigenvalue stability condition
for preferred VAR

Eigenvalue	Modulus
0.887918	0.887918
0.73087	0.73087
-0.11417 + .5737608i	0.585009
-0.11417 - .5737608i	0.585009
0.473354 + .272196i	0.546035
0.473354 - .272196i	0.546035
-0.4287 + .0261793i	0.429503
-0.4287 - .0261793i	0.429503
-0.25502 + .3330084i	0.419439
-0.25502 - .3330084i	0.419439
0.089974 + .3736635i	0.384343
0.089974 - .3736635i	0.384343
0.328985 + .1960071i	0.382949
0.328985 - .1960071i	0.382949

Eigenvalue stability condition
for restricted VAR

Eigenvalue	Modulus
0.825544 + .4898264i	0.959923
0.825544 - .4898264i	0.959923
0.850071 + .4025342i	0.940561
0.850071 - .4025342i	0.940561
0.926446	0.926446
0.857174	0.857174
0.673951 + .1116233i	0.683132
0.673951 - .1116233i	0.683132
0.654571	0.654571
-0.009 + .3612729i	0.361385
-0.009 - .3612729i	0.361385
-0.338344	0.338344
0.150285	0.150285
-0.037494	0.037494

Source: Authors' calculations