“EVALUATION OF UNIVARIATE AND MULTIVARIATE FORECASTING METHODS FOR THE DRY BULK FREIGHT RATES”

By

Nikolaos Triantafillou

Acknowledgements

First of all, I would like to thank my mother and my brother who supported me in many ways throughout the period of the MEL programme. Also, I would like to thank my professor Mr Albert Veenstra who supported me and guided me during the whole academic year and especially during the time where I did this study as well as the whole academic and office staff from the MEL programme which successfully managed to deliver us the whole MEL programme efficiently and in time during those hard and unsure times of the global pandemic. Last but not least, I would like to thank all of my classmates with whom I had excellent cooperation and wonderful moments which I will never forget.

Abstract

The dry bulk freight rates bring an increased amount of interest in many different parties in the shipping industry and this emerges the need for utilisation of different forecasting techniques. Throughout the years, numerous researchers produced a few forecasting models for the dry bulk shipping industry which included simple and complex forecasting models with mixed results in terms of efficiency. This study concentrates on the evaluation of various univariate and multivariate forecasting models including a Naïve forecasting approach, an ARIMA process, S-GARCH and E-GARCH models and the VAR and VECM models. Also, 3 different models with the implementation of an exogenous variable including the ARIMA-X, VAR-X and VECM-X are produced. The reasoning for the selection of these models can be found in the methodology. Panamax and Capesize time series are examined including both Spot and Time Charter freight rates of 8 years (2010-2018). Forecasting results of 12 months are produced. Also three different forecast measures are used to evaluate the results which include RMSE, MAE and MAPE. Fleet Size of each ship category as well as the BDCI and BDPI are used as endogenous variables for the multivariate models and Global Crude Oil Production is used as an exogenous variable for the models above. The results show that overall, the multivariate models slightly outperform the univariate models in 3 out of 4 time series which are examined and show more accurate result when measured as an average price. The VAR model produced the better results in terms of MAPE with an average price of 10.22% for Spot Rates and 7.92% for Time Charter rates. Also it is concluded that the multivariate models produce slightly more accurate results in general in comparison to univariate forecasting models. Finally, the results show that in terms of MAPE, the standard models produce better results in comparison to the model which include the exogenous variable and those models only show some more accurate results in terms of RMSE and MAPE. Overall, the differences in the results the evaluation indicators of this study do not show any major difference between the simple univariate and more complex multivariate models in general and also show that the complex models with an exogenous variable are unable to produce more accurate results.

List of Tables

List of Figures

Introduction

The need for trading different goods from all over the world began a thousand of years ago. Maritime transport is considered as the cheapest way to carry different kinds of goods between significant distances all over the world. Until the 19th century, ships were using steam coal for bunkering purposes. Apart from the difficulty of finding the appropriate coal fuel in ports, coal for bunkering required a lot of storage room which resulted in the restriction of net deadweight for the usage of cargo transportation and thus carrying cargo was kind of restricted. After the introduction of gasoil engines and the appropriate distribution of storage tanks for cargo, even bigger ships started to be built. Moreover, heavy industrialization born the need for enormous quantities of raw materials. Dry bulk vessels vary enormously in their size, from small coastal vessels up to 400.000 deadweight capacity vessels which serve specific Atlantic-Pacific routes for the transportation of iron ore. As a result, a very dynamic market has been built around this kind of vessels and unlike the tanker market where a big tonnage concentration has been seen among big oil companies, there is a distribution of tonnage among small ship-owning companies and manufacturing companies with an integrated supply chain. Nowadays, dry bulk vessels have served hundreds of trading routes around the world, either serving a specific trading pattern and route or trading in the sport market with different routes, depending on the market conditions. As in every industry, the main goal of investors is a satisfying return of investment (Stopford, 2009). Vessels consist of an investment which requires a lot of capital and shipping industry in general has distinctive characteristics and volatility which makes it a kind of unique in comparison with other industries. As any other industry, business cycles are a part of the shipping industry which have varied in duration and results throughout the past years. (Stopford, 2009). This volatility and uncertainty has emerged the need for a series of tools which could help shipowners to "predict" the future and be able to make the right investments. Throughout the years, different models have been used to predict different elements the shipping industry such as the average haul, demand and supply for the shipping industry, fleet productivity and of course freight rates. Freight is the type of payment to the shipowners for the carriage of goods and raw materials (Institute of Chartered Shipbrokers, 2016). The forecasting of freight rates is a topic which has been researched and experimented with many kinds of forecasting models throughout the years. Various univariate and multivariate models have been utilised for the forecast of the freight rates and have been proved inaccurate throughout the years and the reason behind of that is that there are some aspects of the shipping industry which cannot be predictable. (Stopford, 2009). As the future freight rates depend on many different variables such as the future fleet, the new orders, the supply of vessels, the price of commodities which are carried and the demand of those commodities, there are also some unexpected variables such are economic crisis, wars between countries and even global pandemics such as the COVID-19 which has spread to all over the globe, that can considerably affect the freight rates, even when all the other predictable variables are taken into consideration. There are essential steps for improving forecasting in general with the most important being acquiring the right information at the right time. (Stopford 2009). In this dissertation there are going to be used some different kinds of forecasting models, including univariate models of AR(I)MA and GARCH models, multivariate models such as the Vector Autoregression Model and the Vector Error Correction Model and some alterations of the ARIMA, VAR and VECM model with the inclusion of an exogenous variable after a comparison and close selection of the optimal exogenous variables. Several studies have been conducted for the evaluation of different forecasting method for the dry bulk, the tanker and container shipping industry with mixed results but this dissertation is going to include a few endogenous variables which include the number of the global fleet for each dry bulk vessel category as well as the specific Baltic Dry Index for each vessel category to check whether the multivariate models can produce more accurate results in comparison to the univariate models. Also, the comparison is going to include the same models with the inclusion of an exogenous variables to see if the more complex models can produce more accurate results. Apart from the different vessel categories, different freight rates are going to be included in the study, including time charter and spot rates to spot any differences between the forecasting efficiency of the models for each category. Different forecasting efficiency measures are going to be included into this study such as the Root Mean Square Error and the Mean Absolute Percentage of Error to quantify the efficiency of each forecasting model which will contribute to the evaluation of each model and the comparison between the simple univariate and the more complex multivariate models. As the freight rates of the shipping market in general present an increased amount of variability this can lead to inaccurate results from the standard univariate models and brings up the need for more complex models to bring more accurate forecasting results. The initial hypothesis of the research is that it is expected that the more complex the forecasting model the more accurate the results.

Structure of the Thesis

As far as the structure of the dissertation concerned, the first section includes a literature review concerning the different aspects which affect the demand and supply of the dry bulk shipping industry, the shipping markets of the shipping industry, introduction to time chartering and voyage chartering employment of the vessels, information about the most significant commodities and vessels of the dry bulk shipping industry as well as some information for some forecasting approaches which have been used in the past by researchers as well as the components of time-series forecasting.

The second section includes the theory behind the forecasting methods which are going to be used, as well as the theory behind every evaluating measure and every validation test.

The third section includes the analysis of the results for every time series as well as the comparison between every forecasting model.

The fourth section includes the conclusion of the study along with answers for the research questions of the study.

Research and Sub-Research Questions

Many univariate and multivariate models have been utilised in many studies to produce effective forecasting results for the dry bulk industry and the shipping industry in general. Many of the models have resulted with adequate and inadequate forecast results which can be explained by the variables which have been used, as well as some unexpected variables which cannot taken into consideration by the forecasting models. As this study concentrates in the evaluation of the different forecasting models for each time series, the most important question which is generated is the following:

Does the inclusion of an exogenous variable lead to more accurate results of the existing univariate and multivariate models?

After developing the different forecasting models for each time series and evaluating the results, the complexity of some forecasting models which are included in this study, generates some more sub-research questions which can be presented below.

i) **Which variables affect the dry bulk freight rates?**

ii) ***Does the complexity of the multivariate models lead to more accurate results?***

iii) ***Which forecasting method is the most appropriate for each time series variables?***

Structure of the Thesis

As far as the structure of the dissertation concerned, the first section includes a literature review concerning the different aspects which affect the demand and supply of the dry bulk shipping industry, the shipping markets of the shipping industry, introduction to time chartering and voyage chartering employment of the vessels, information about the most significant commodities and vessels of the dry bulk shipping industry as well as some information for some forecasting approaches which have been used in the past by researchers as well as the components of time-series forecasting.

The second section includes the theory behind the forecasting methods which are going to be used, as well as the theory behind every evaluating measure and every validation test.

The third section includes the analysis of the results for every time series as well as the comparison between every forecasting model.

The fourth section includes the conclusion of the study along with answers for the research questions of the study.

1. Literature Review

1.1 Shipping Markets

1.1.1 **Newbuilding Market**

This market consists of the place of order and purchase of newbuilding vessels. Current shipowners or investors that would like to involve in the shipping industry contact with shipyards and after expressing their exact specifications that they would like for a vessel to contain or choosing a pre-fixed vessel type and model, they negotiate for the price of the vessel, which is correlated with the current market condition, the financing terms which include any shipyard guarantee which may not favour the shipyard and it is usually needed from a lending bank and the method of payment for the vessel as it might be in stages before the delivery of the vessel or upon the delivery of the vessel to the shipowner (Stopford, 2009). It should be stated that the vessel will not be available to the future owner for use after 2 or 3 years upon the order date, depending on the sophistication of the vessel. The buyer side may involve different kinds of reasons such as a casual replacement of fleet of a big shipping company or an investment of a shipowner with "bright" future expectations about the conditions of the market. (Stopford, 2009). The price of newbuilding vessels is determined by a series of factors which may include the freight rates, the price of modern second-hand ships in the market at a particular time, the financial position of the buyer as he might be in a strong liquidity position and might achieve better prices from the shipyards and of course the supply of the existing world fleet. By talking about future expectations, we can see how the shipyard orderbook increases as there are positive future expectations about shipowners.

1.1.2 Sale and Purchase Market (S&P)

The S&P market involves the sale of existing vessels (second-hand) into new buyers with brokers existing as intermediates in order to arrange the transactions. (Institute of Chartered Shipbrokers, 2016). There are many different reasons which affect the supply of existing tonnage for sale and the demand for vessels. From the point of buyers, there might be a particular policy for fleet replacement according to an age limit, an positive expectation about a future trend of fleet demand and freight rates or a spontaneous purchase to fulfil a current demand situation. From the point of sellers there might be need for liquidity, future expectations about the state of the market, change of policy of a shipping company about the type or size of the fleet and other reasons. Second-hand vessels are often sold free of any mortgages or long-term charter commitments from the previous owner but there are some exceptions where a vessel has been sold with an ongoing time charter contract. (Stopford, 2009). This market has a significant price volatility which Is interdepended from the other shipping markets, as well as the demand and supply of vessels in general.

1.1.3 Demolition Market

The demolition market is a market for vessels, typically for those which exceed their operating age which are sold in scrapyards or intermediates which eventually sell them in those scrapyards where they are destroyed for obtaining the metals that a vessel has. (Institute of Chartered Shipbrokers, 2016). Demand for scrapping is mainly driven by the demand and price of steel from the manufacturing countries. Supply is driven mainly by the fleet age of the market. (Stopford, 2009). Although in the recent years, due to market recessions and even depression, there have been many ships which have been sold by shipowners because there was no other choice due to ships being idle for a long time. (Stopford, 2009). Another more unusual factor is that some shipowners after replacing an older ship with a new one, they have decided to sell the ship for recycling rather then selling it in the second-hand market, even at a better price, for the reason that the anticipated that their old vessel would be a competitor in the “hands” of another owner. (Institute of Chartered Shipbrokers, 2016). The countries which have the most scrapyards are Bangladesh, Pakistan, India, China and Turkey.

1.1.4 Freight Market

This market consists of the shipowners, which provide available tonnage for the transportation of cargo, charterers who can either be sole traders who purchase commodities and charter vessels to transport the cargo between regions or producers who sell their commodities to manufacturing companies or even ship operators who charter ships in order to transport cargo and seek profit between the difference of the hire for chartering a vessel and the freight obtained for carrying cargo. (Institute of Chartered Shipbrokers, 2016). In this markets, shipbrokers act as intermediates where they undertake the task of either finding available cargo or available tonnage, depending on their principal’s interest. Today there is an international freight market which consists of different categories of markets according to ship type, size and trade routes. (Stopford, 2009). Although we do have different sub-categories of freight markets, these are interconnected because we either have combined vessels like Oil-Bulk-Ore (OBO) which can switch between the Dry Cargo and Tanker market or we frequently see the situation where liner companies time-charter dry cargo vessels to meet seasonal demand in the liner shipping industry (Institute of Chartered Shipbrokers, 2016). The freight market mainly consists of two types of transaction, the voyage chartering contract where a vessel is chartered for one trip, with different sub-categories, and the time-chartering contract where the shipowner charters the ship for a particular time horizon and receives daily hire payment. (Stopford, 2009). This market is described by intense volatility and many factors drive the price of the freight.

1.1.5 Interdependence of the four Shipping Markets

It can be examined, how those four different shipping markets are closely interdependent with each other. First of all, the same shipowners are trading in all four markets and their activities are closely correlated (Stopford, 2009). Cash flow is the main driver of the interaction of the four different markets, in the way which can be seen below.

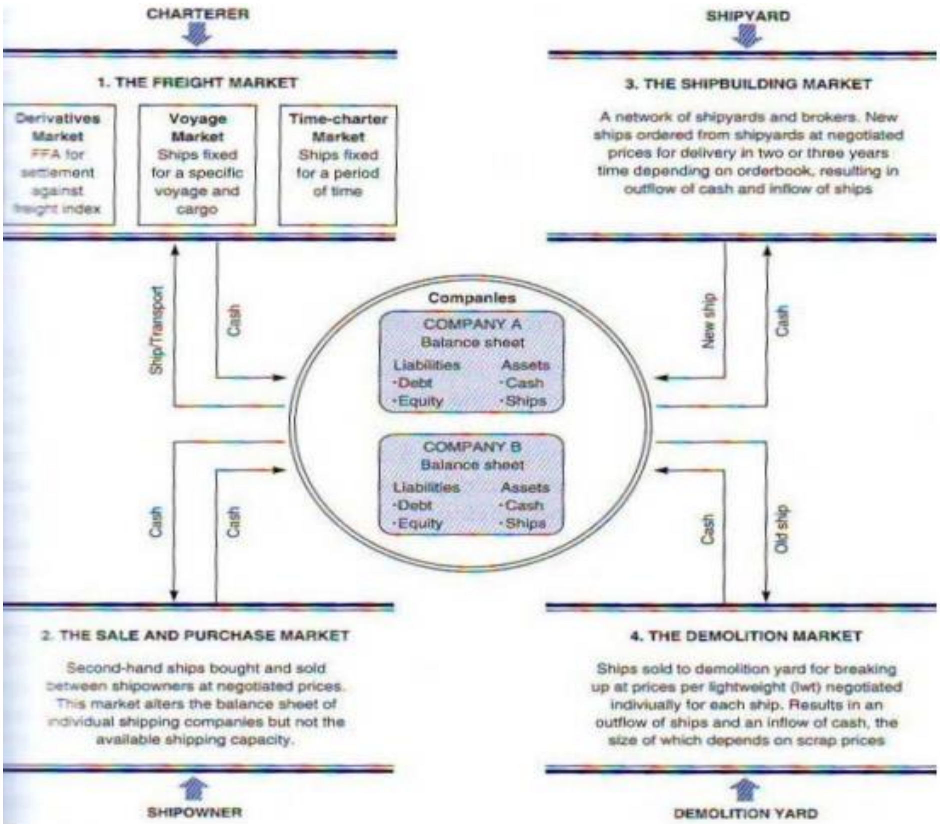


Figure 1: The linkage between the four shipping markets

Source: Stopford, (2009, p.179), Maritime Economics

1.2 The Dry Cargo Market

1.2.1 Introduction

The Dry Cargo market, is the market where the movement of significant commodities are carried either in bulk or in packages and pallets. This market is closely linked with the manufacturing industry where raw material is being carried in manufacturing countries. (Institute of Chartered Shipbrokers, 2016)

1.2.2 Iron Ore

Today, most of the world's metals are produced by processing mineral ores such as iron ore, after which precious metals such as steel or bauxite are produced. It is mainly exported from developing countries and imported from developed and industrialised nations. Iron ore consists of the single largest dry bulk cargo which is traded nowadays (Institute of chartered Shipbrokers, 2016). Also, the procedure of processing the iron ore into precious metals it taken place in the developing countries before reaching their last destination and this is for different reasons, with the main being the cost-savings for the end-user and the gathering of foreign currency for developing countries. (Institute of Chartered Shipbrokers, 2016). China has been the largest importer of iron ore in order to fulfil the industrial needs and the biggest exporters are Australia and Brazil. We have seen very big vessels of over 200.000 deadweight being built for the carriage of Iron Ore between Brazil and China, constructed for the giant mineral ore extracting company "Vale". Iron Ore has a small stowage factor, typically of 0.3 cubic meters per tonne. Capesize, VLBC and Panamax vessels are the main vessels which carry iron ore. Bulk carriers which carry iron ore are considered as "specialists", which have small holds above double-bottom tanks. (Institute of Chartered Shipbrokers, 2016). This raw material is infamous for causing damage in the vessels which carry it during port operations and also at the time of transportation and thus this structural arrangement provides strength while it reduces stress at the ship which can be caused by a large metacentric height. (Institute of Chartered Shipbrokers, 2016). Special care is taken during the loading and discharging operations of iron ore because it is considered as the most infamous dry cargo commodity in terms of maritime claims. (Institute of Chartered Shipbrokers, 2016). Iron ore is typically carried in a humid condition and thus, draught surveys may be needed to determine the bill of lading weight and thus the freight payment based on the weight. Sophisticated terminals are responsible for the loading and discharge of iron ore, usually being loaded by big-sized conveyors along with a ship-to-shore loading gantry and discharged usually by gantry-type grab fitted cranes from the conveyor belts to the port storage stack. (Institute of Chartered Shipbrokers, 2016). As the demand for tonnage volume is dependent on the supply of iron ore, no one can be sure about the future, as enormous amounts of iron ore are yet to remain unexploited. (Institute of Chartered Shipbrokers, 2016)

1.2.3 Coal

Coal is not a simple material and it has many properties. Coal is a source of energy used by many industries. There are two types of coal, steam coal which is used for power generation mainly in electric plants and in the past as an energy source for the propulsion of steam ships and coking coal which is mainly used as a heat source for industrial purposes, particularly for steel production. (Institute of Chartered Shipbrokers, 2016). Coal is found in various quantities in many countries. The patterns of shipment of coal are primarily dictated by the consumption of the major industrial centres of the world. Also, georgraphic and economic factors play a role where those major industrial nations tend to receive the raw material from neighbour countries for cost saving purposes (Institute of Chartered Shipbrokers, 2016). Moreover, the demand for coking coal is closely linked with the "health" of the international steel industry. Similarly with iron ore, coal is loaded by conveyor belts or a chute. Discharging is being held usually by shore cranes with grabs (Institute of Chartered Shipbrokers, 2016). Capesize and Panamax vessels are the main types of vessels which carry coal around the world.

1.2.4 Grain

There are many types of grains which are carried by bulk carriers including soy beans, wheat, oats, corn and barley. They are either destined for human or animal consumption. Countries with a surplus in grain production like Canada, US and Russia, export the commodity into China, most European Countries and Australia. (Institute of Chartered Shipbrokers, 2016). Some types of grains are processed into meals prior to shipment. A special feature of this commodity is the fact that the supply varies year by year due to the destruction of harvest by extreme weather conditions. The heavy grains have a stowage factor of about 1.3 cubic meters per tonne, which means that the cargo compartment can be filled by volume way before the deadweight capacity is reached. (Institute of Chartered Shipbrokers, 2016). Grains usually undergo through a trimming process upon loading according to grain calculations that each bulk carrier has. This happens because there is a danger for the ship to capsize. Handymax, Handysize and Panamax are the types of vessels which usually transport grains. (Institute of Chartered Shipbrokers, 2016). Up until the 20th century, the grain trade patterns remained unchanged until a significant change of former developing countries to alter their diet and demand for more dairy products. Grains typically are stored in grain elevators which are located in all over the country and they are stored until they are carried by trucks or barges in the exporting ports. Usually, ports in the developed countries are sophisticated and use conveyor belts to load the grains into the vessel. In less sophisticated terminals, cranes with grabs are used. Discharge of grain is being executed by grabs. (Institute of Chartered Shipbrokers, 2016). Also, cargo compartments must be cleaned before loading and fumigated to remove insects and rodents. In ports which are less sophisticated, bags are used for the loading of grains and it is usually required by the shipowner in the charter party to carry extra bags to replace the old ones in case of cargo bleeding. (Institute of Chartered Shipbrokers, 2016)

1.2.5 Other Bulk Cargoes

There are other products that are being carried in bulk except grains, coal and iron ore. Some of those include Process material including plywood which are very susceptible to damage and thus are carried in specialist bulk carriers, forest products which include wood-derived product and timber which is carried and loaded in specialist bulk carriers as well, steels where there is a significant variety among them and are carried by Handymax and Panamax vessels. (Institute of Chartered Shipbrokers, 2016). Steel products are notorious for causing damage in the hull of the ship. Also apart from iron ore, there are other minerals which are carried by Dry bulk vessels such as bauxite. Finally there are some significant agricultural products which are carried by bulk carriers with the most important being sugar which is carried in raw state in bulk into processing refineries before turning into its final state or into processed foods into the final destinations. Also, cassava is another agricultural product carried in bulk mainly exported from Thailand and Malaysia. (Institute of Chartered Shipbrokers, 2016)

1.3 Dry Cargo Vessels

The dry cargo vessels are designed to transport large quantities of the cargoes explained above. The capacity of dry cargo vessels can vary from 1000 deadweight, up to 400.000 deadweight. These vessels have usually from one up to nine holds and hatches, depending on their size and they have different types and differences in their characteristics and design according to their size. (E. Plomaritou, 2018)

1.3.1 Handymax/Handysize

These vessels have a varying size of 20000 deadweight and up to 50000 deadweight. Handymax vessels with a bigger capacity tend to imitate the trade routes and preferences of Panamax vessels but can also carry a bigger variety of cargoes because they have ship gear. This is mainly because they trade in ports where there are not developed terminals with sophisticated equipment (Institute of Chartered Shipbrokers, 2016. Handysize vessels tend to trade in short haul routes and carry a variety of minor bulks and also smaller quantities of grain and bauxite. Some of them also have the equipment to load and discharge logs and cement. (E. Plomaritou, 2018). Handymax vessels typically trade in relatively short or medium haul routes.

1.3.2 Panamax

Panamax vessels are usually characterised as the work horses of the shipping industry. Those vessels were considered as these having the maximum size capable of entering the Panama Canal up until 2016, before the canal undertook expansion work (Institute of Chartered Shipbrokers, 2016). These vessels have a size up to 80.000 deadweight and in their majority are gearless with the exception of some older vessels. Many ports around the world are designed with a maximum draft close to the draft which Panamax vessels have. (E. Plomaritou, 2018). They are mainly used for transporting coal, iron ore and grain. The maximum cargo capacity which those vessels could have in order to transit the Panamax Canal when fully loaded was up to 52.500 tonnes (E. Plomaritou, 2018). Those vessels usually trade in medium and long haul routes, including the regions of Atlantic Basin, Pacific and Indian Ocean with the freight rates depending on the trade route (Institute of Chartered Shipbrokers, 2016). After the expansion of the Panama Canal, a new category arised, named as “New Panamax” vessels. Panamax vessels typically have seven holds and seven hatches. (Institute of Chartered Shipbrokers, 2016)

1.3.3 Capesize

Capesize vessels are gearless bulk carriers with a varying size from 100.000 to almost 400.000 deadweight. They are trading almost exclusively in long haul trade routes including Brazil or Australia to China. (E. Plomaritou, 2018) They mainly carry iron ore and coal. Due to their big size they are not able to navigate through the Panama Canal and thus they navigate through the “Cape of Good Hope” from which they took their name from. Those vessels have subdivisions including “Small Capes” which have a size up to 150.000 deadweight, Normal Capes which are up to 180.000 deadweight, Large Capes with over 180.000 deadweight including the “Very Large Bulk Carriers” and “Very Large Ore Carriers” with a deadweight up to 400.000 deadweight, serving the route of Brazil-China. (E.Plomaritou, 2018). Those large Capesize vessels are also called “Chinamax” or “Valemax” getting their name from the trade route that they serve and from the Brazilian iron ore processing company Vale, which serves the Chinese market. Capesize vessels have seven or nine holds. Those vessels are also benefited from the so called “economies of scale” and are able to carry cargo at a lower transportation cost. (Institute of Chartered Shipbrokers, 2016). The larger type of Capesize vessels has been heavily criticized due to the fact that they tend to oversupply the already “overstuffed” dry cargo market. (E. Plomaritou, 2018)

1.4 Demand and Supply Factors in the Dry Bulk Market

1.4.1 The Demand and Supply model

The shipping industry is a very volatile market which is influenced by many supply and demand factors of tonnage and cargo and as well as geopolitics (Institute of Chartered Shipbrokers, 2016). The most significant factors can be simplified and turned into a model which is shown below. (Stopford, 2009) The model has three different components, demand (module A), supply (module B) and the freight market (module C) which is the linking point of demand and supply by regulating the movement of cashflow from module A to module B. (Stopford, 2009) Business cycles and growth percentage of different commodities determine the demand and thus the volume of those transferred by maritime transportation. Continuing to the supply module, in the short term, the world fleet provides a fixed tonnage number for the transport of goods. This number fluctuates and can be lowered or increased according to the demand. (Stopford, 2009). The following demand and supply factors of the model will be analysed further below.

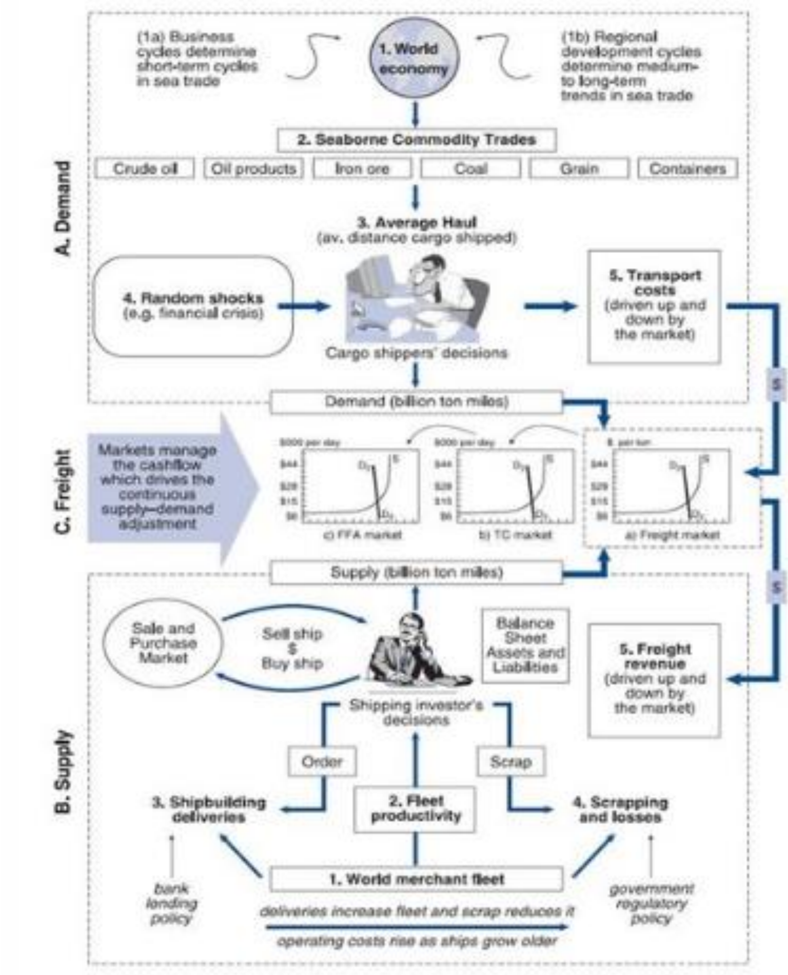


Figure 2.

“The shipping market demand and supply model “

Source: M. Stopford, Maritime Economics, 2009

1.4.2 Demand Factors of the Shipping Industry and the Dry Bulk Market

1.4.2.1 The World Economy and Business Cycles

The world economy is considered as the most significant factor for the determination of demand for the shipping industry as it generates the demand through the manufacturing and exporting of raw material and as well as trade in material and products. (Stopford, 2009) The most significant factor that may provide the demand with changes are the business cycles (Stopford, 2009). The business cycle creates a cyclical pattern of demand for vessels through the fluctuations in the rate of economic growth through into the maritime trade (Stopford, 2009). Those business cycles in the world economy are mirrored into the shipping industry with the so called “Maritime Business Cycles”. Many economist nowadays believe that those cycles arise from a number of factors with the most important of them being:

- **Time Lags**: These are the delays between decision making and the actual time of fulfilment. The perfect example for this can be the order of new vessels from shipyards, where the shipowners make the orders in a time of economic boom and the vessels are ready for operations after two to three years, usually inn times of recession after the oversupply of the market (Stopford. 2009)

- **Stockbuilding**: This procedure produces the exact opposite result from the previous factor, as the shipping market faces sudden bursts of demand which drives the shipping industry into a positive growth. This stage happens into times of economic recovery. (Stopford, 2009)
- **Psychologic Factors:** Individuals tend to act according to personal beliefs of optimism and pessimism rather than actin rational.

Business cycles tend to vary in duration. Statisticians have introduced the so called "leading indicators" which provide advance information and warnings about the turning times of the business cycles. (Stopford, 2009).

1.4.2.2 Average Haul

According to Stopford, the transport demand is determined by the average haul, which is a type of measurement of the tonnage of cargo shipped, multiplied by the average distance over which is transported. (Stopford, 2009) Throughout the years, the alternation of average haul has affected the overall demand on the shipping industry. In most shipping trades, the average haul has been alternated throughout the years. Along the years, many statisticians have taken into account the average haul in order to forecast the shipping market.

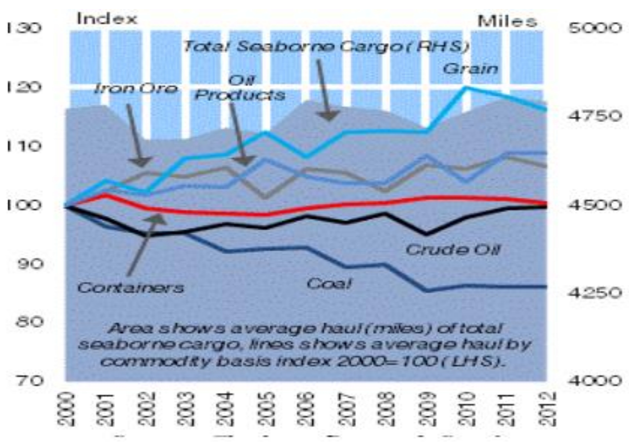


Figure 3.

"Average haul of total seaborne cargo"

Source: Clarkson

1.4.2.3 Random Shocks on the demand in the Shipping Industry

These shocks include many different categories such as Acts of God, commodity price volatility and sudden price changes, any political events such as a war or a revolution or maybe even "Trade Wars" which we have recently seen between China and U.S and various economic shocks which have occurred throughout the years. All these different types of shocks are able to disturb the shipping business cycle . (Stopford, 2009)

1.4.2.4 Transport Cost

The cost of transport the so called "freight" is determined by the balance of demand and supply, as well as the type of cargo being carried, the distance which the vessel has to cover to transport the cargoes and the stowage factor of the cargoes which may determine the freight (Institute of Chartered Shipbrokers, 2016). Maritime Transportation is considered as the "cheapest" means of transportation of large quantities of cargo and according to Stopford, raw materials will be transported through vessels if the cost is acceptable and reasonable and value is added in the supply chain. Over the last years we have encountered some major improvements in the quality of services of vessels as well as the so called "economies of scale" with the introduction of larger vessels and as a result, more attractive costs can be generated for the transportation of cargo. Transport costs rather have a long-term effect in the trade development, than a short-term effect. (Stopford, 2009)

1.4.2.5 Other more specific demand factors for the Dry Bulk Market

- The overall policy of China regarding the various industrial industries, including imports, exports and production of steel and iron ore and demand for coking and steam coal.
- Overall steel production on a global level.
- Seasonality.
- Grain exports from US and Canada and imports in China and Africa and weather conditions which might affect the production. (Plomaritou, 2018)

1.4.3 Supply Factors of the Shipping Industry and the Dry Bulk Market

1.4.3.1 Total merchant fleet

The first factor is the total fleet of vessels in the market (including newbuilding orders). There have been many fluctuations in the number of the merchant fleet throughout the years. A very important element in this topic, is the mechanism by which the world fleet number adjusts when the total demand did not go as expected. (Stopford, 2009) The world fleet is constantly either into a phase of expansion or on a contraction, depending on many different factors. There are also many factors which can contribute to the world fleet number, including the ability of some vessels which were mostly used in the past to change markets (Ore-Bulk-Oil) and also some coastal vessels competing into the medium haul route markets. In the dry bulk market, we have witnessed the expansion of the world fleet at a steady rate throughout the years with a trend of building bigger vessels to benefit from the economies of scale. Those larger rates have caused the depression of the charter rate for smaller vessels in the market and thus there are many examples of dry bulk vessels competing into different trade routes or carrying a wider variety of cargo in order to be able to survive.

1.4.3.2 The fleet productivity

The productivity of the merchant fleet is typically measured in ton miles per deadweight. (Stopford, 2009) This measure has a flair of flexibility in the annual number which is based in different factors. The first one is speed where it shows that a vessel is at sea transport. Speed can vary, and in the last years shipowners tend to choose lower speed and fuel combinations in order to reduce the bunker costs. Also, technology and age are important factors for the speed of a vessel as well as fouling matters of the hull, which can cause more fuel consumption (Stopford, 2009). Port time is another critical factor for the fleet productivity. Dry Bulk vessels seem to be the ships with the most incalculable port-time since most tanker charter parties stipulate a 72-hour laytime and containerships are trading in sophisticated ports which provide a quick turnround time. Also, days where a vessels either travels in ballast or might be laid-up will have a contribution in the fleet productivity. (Stopford, 2009)

1.4.3.3 Shipbuilding Orderbook

The shipbuilding process plays a very significant role in the total annual fleet number. (Stopford, 2009) The output of the shipbuilding market is measured annually. It must be mentioned the fact that the shipbuilding process takes from 2 up to 4 years which depends on the size of the orderbook. New orders are placed either based on a rational decision or after beliefs about the future demand. This procedure has an impact on the cyclicality of the market. Regarding the dry bulk newbuilding market, is has been stable in the past decades with a flair of cyclical investment. (Stopford, 2009)

1.4.3.4 Freight Rates

The supply of the world fleet is heavily influenced by the freight rates. This is considered as a critical regulator for decision-makers, in order to invest in the shipping industry or existing shipowners to make further investments (Stopford, 2009)

1.4.3.5 Scrapping

Along with the shipbuilding market, it determines the total merchant fleet. There are many factors which determine the decision for selling a ship for scrap. The first and most significant is the age of the vessel. Vessels typically do have an operating age of around 25 years, depending on the type. Also technical obsolescence of a type of a vessel may lead to the decision for scrapping a vessel earlier than the end of the operating years (Stopford, 2009). Market recession and even depression along with tempting scrap metal prices may also lead to the decision for scrapping (Stopford, 2009). Finally, a less significant factor that ha snot been witnessed in many occasions but has still occurred, is when a shipowner decides to sell the old vessel for scrap metal instead of selling in into the second-hand market even at a better price after replacing it with a new vessel, for the reason that the old vessel will be faced as a competitor in the market. (Institute of Chartered Shipbrokers, 2016)

1.5 Chartering

1.5.1 Introduction of the Dry Bulk Chartering Market

The dry cargo market is not divided rigidly into different divisions such as the tanker market (crude oil, products, chemicals e.t.c) but there can be identified many different markets around the world based on different characteristics which include ship type, ship size, type of commodities, coastal and intercontinental and deep-sea transport and other characteristics. (Institute of Chartered Shipbrokers, 2016). Nowadays there is not really a centre for the dry cargo market as the communication technology that exists today, makes business to conduct easier even if the parties involved in the trade are residents of different continents. Although, some years ago when communication matters were far more complicated, dry freight markets were developed in various traditional shipping centres all over the globe, some of which remain active even nowadays with the most significant being London, Ney York, Singapore, Shanghai, Piraeus e.t.c.

1.5.2 Categories of Dry Bulk Chartering Methods

1.5.2.1 Voyage Chartering

In this type of employment, a vessel is employed to load cargo from one or more ports and perform a single trip in order to discharge cargo in one or more ports. The reward for this type of chartering business is called freight which is payed to the shipowner in return for the transportation of cargo either in a lump sum or pro rata against the quantity of cargo loaded and less commonly against the net tonnage capacity of the hatches of the vessel. (Institute of Chartered Shipbrokers, 2016). The shipowner is responsible for the Operational and voyage costs of the vessel and calculates the total costs with the method of voyage estimation. It is usual for the operations department of a company to conduct multiple voyage estimations for multiple destinations to choose the most profitable, manage trip costs with respect to the next employment. It would be not profitable at all to choose an offer of cargo transportation which will not help the shipowner to find cargo at the port of discharge or near of it. Laytime is the time allowed for the charterer to load and discharge cargo. It can be seen that time is a very significant factor and plays a big role in the profitability of the employment. Thus, If laytime is exceeded, the charterer is obligated to pay demurrage and it is payable upon calculating the days of extra laytime used in respect of terms agreed in the charter party (f.e SSHEX, NOR extra time etc) and the log book and statement of facts written by the port agent. (Institute of Chartered Shipbrokers, 2016). In contrast with demurrage, if the time used for port operations is less than the allowed laytime, the charterer should demand for despatch. In terms of charter parties, despatch is usually agreed to be the half of demurrage. (Institute of Chartered Shipbrokers, 2016). Provided that the charterer supplied the ship with cargo which is less than the agreed in the charter party, the shipowner is entitled to deadfreight, (Institute of Chartered Shipbrokers, 2016). There are also some sub-categories of Voyage Chartering, called Consecutive Voyages Chartering which involves a vessel to perform a pre-determined number of voyages or as much voyages at a fixed time frame with the difference that the freight rate might be agreed to fluctuate according to market rates and also demurrage and despatch will be different in each voyage. In case where the vessel is unable to perform the trips or lost, the contract is terminated. (Institute of Chartered Shipbrokers, 2016)

1.5.2.2 Time Chartering

In this type of employment, the vessel is hired for a specific period which can be long or short depending on the needs of the charterer. There are many reasons where time charter contracts are being conducted and determine the length of the contract with some being the urgent need for extra tonnage due to maintenance of existing fleet or supplementary demand for transportation which drives many major companies to employ extra tonnage. (Stopford, 2009). Also, due to the nature of this type of employment, it can be seen that it is being preferred against Voyage Chartering in times where freight rates seem unfavourable for the shipowners, demand can't match tonnage supply and thus time charter contracts offer a stable income. (Stopford, 2009). Also, in econometric analysis charts, it can be seen that time charter contracts consists of a smoother line of voyage charter contract with less fluctuation. In other words, many shipowners strive for voyage charter contracts in times of attractive freight market rates and seek safety with time charter contracts in times where low freight rates exist. There are many differences with voyage charter contracts. The most significant is that the shipowner is responsible only for the capital costs and operational costs which include all the costs for the maintenance of the vessel and the charterer is responsible for the voyage costs. (Institute of Chartered Shipbrokers, 2016). In addition to voyage chartering, the shipowner is entitled the payment of hire, which is paid semi monthly in advance and in which several additions and deductions charter party clauses are included in the charter party contract. The term “trip-chartering” is relatively new and is about employing a vessel for a single or round-trip voyage which is similar to the term of voyage chartering but with the responsibilities that a time charter contract has which include the responsibility of voyage costs to the charterer. Some charterers prefer this type of employment in case where they seem to achieve lower bunker prices and voyage costs in addition to paying a lump sum. (Institute of Chartered Shipbrokers, 2016). Also under this type of employment, charterers often negotiate to change the name of the vessel to be more easily connected with the route that it might operate. The charterer who is involved in this type of employment is also called as the “disponent owner” where he/she acts like he/she were the actual owner of the vessel. (Institute of Chartered Shipbrokers, 2016)

1.5.2.3 Bareboat Chartering

This type of employment is considered as a very long-term time charter contract. Shipowners hire-out their vessels for periods of more than 5-8 years to charterers. The difference with the former is that the charterer virtually runs the vessel as he is the actual owner, having almost every responsibility of the vessel, including running costs. In return the shipowner receives a hire payment which would be significantly lower than if he had a normal time charter contract. (Institute of chartered Shipbrokers, 2016). Bareboat Chartering can be considered as a ship-finance tool where investors with no expertise or prior experience in the shipping industry can be involved, leaving the management and employment matters of the vessel to a charterer. Also it can be a ship financing method for a charterer who hasn’t got enough funds to purchase a vessel. Even leasing method of a vessel could be agreed after some years of bareboat charter contract with a purchase at a lower price. (Institute of Chartered Shipbrokers, 2016).

1.5.3 Baltic Dy Index

The BDI is considered as a shipping index which measures the changes in the cost of sea transportation of various commodities. Created by Baltic Exchange, the physical freight market in London in 1985, it started to publish a series of daily route assessments and a group of brokerage companies voted on those routes. (Institute of Chartered Shipbrokers, 2016). Also a formula where those route assessments where weighted by their importance and as a result the Baltic Dry Index was created which covered every type of ships which were used for the carriage of Dry Cargo Vessels but later more route assessments have been added and as a result the BDI is broken into more specialised indices, depending on the type of vessel.

1.6 Forecasting in the Shipping Industry

Forecasting is a very crucial aspect in the shipping industry because anticipation of the future leads to more profitable decisions regarding charter contracts, acquisition of vessels and other entrepreneurial moves of the shipowners and other parties which are involved in the shipping industry such as the Lender Banks and shipyards. (Stopford, 2009).

In 1992, Culinnane developed a model in which the Box Jenkins method is applied to forecast the Baltic International Freight Futures Exchange and compared this type of forecast with other types of forecasts. (Culinnane, 1992). The study concluded to the fact that the ARIMA-Box Jenkins model is optimal when it comes to short term forecasts and short term market movements. Also the study showed that this model needs an optimal level of specification in order to produce efficient results as highly specified models are restricted by their parameters and less specific models, by their accuracy. (Cullinane, 1992). Also this study concluded to the fact that the use of univariate time-series models provide the most cost-effective way to produce forecasts for the BIFFEX. (Cullinane, 1992).

In 1997, Veenstra and Franses developed the VEC model to examine a sample of ocean freight rates, where this model does not include any other variables because the market is considered as efficient. The results showed that an economically meaningful structure exists in the set of ocean dry bulk freight rates and that there are stable long-run relationships between them (Veenstra, 1999). The conclusion of the study is that a significant part of movement of the dry bulk freight rates is stochastic in nature and that despite the fact that long relationships exists, those relationships do not produce better forecasts. (Veenstra 1997) (Franses, 1997)

In 1999 Culinnane again with a revisited study, produced a Box Jenkins model in order to evaluate the changes to the Baltic Dry index after the Handysize trades expunged from the index and compare it to the results of the previous Box Jenkins model where the Handysize trades were included within the index. The results showed that the models are remarkably similar and that there is no alteration in the behaviour of Baltic Dry Index despite the revision. (Culinnane, 1999).

In 2003, Kavussanos and Nomikos investigated the relationship between futures and spot prices in the freight futures market. They produced VECM, VAR, ARIMA and other random walk models and concluded to the fact that VECM model produces more more accurate forecasts of spot prices compared to tother models. (Kavussanos, 2003)

In 2004, Kavussanos and Visvikis investigated the unbiasedness hypothesis of FFA prices in the freight over the counter forward market trades with the use of cointegration techniques. The results indicated the fact that one and two month before maturity FFA's are unbiased predictors of the spot freight rates while the three month before maturity FFA's for Panamax Pacific routes are unbiased predictors of spot prices. (Kavussanos, 2004)

In 2002, Bessler, Jonnala and Fuller developed a GARCH model to estimate the ocean grain rate equation. The results of this study showed that ship size, voyage distance, contract terms flag and season are important to determine the rates and also that efficient port infrastructure and its ability to accommodate the constant increasing size of bulk carriers is very important in maintaining competitiveness in tha global grain markets (Bessler, 2002)

In 2007, Batchelor and Visvikis tested the performance of various forecasting time-series models of spot and forward rate in major trade routes using VECM, VAR and ARIMA models. The study concluded to the fact that when it comes to predicting forward rates, ARIMA and VAR models forecast better even though VECM models seem to give the best in-sample fit. (Batchelor, 2007). Also the study indicates the danger of forecasting with the use of equilibrium correction models when the market structure is evolving (Batchelor, 2007).

In 2013, Papailias and Thomakos, focused on the identification of cyclical patterns in the shipping industry, as well as tried to forecast the BDI using linear and trigonometric regression. Results showed that commodities and trigonometric regression can lead to improved forecasts (Papailias, 2013).

In 2014, Wong used a combination of forecasting models in order to generate short term and long term predictions of the BDI. By using ARIMA, Grey system and fuzzy heuristic modelling, the study came into a conclusion that the ARIMA model produces more accurate forecasts when it comes to longer term results (Wong, 2014).

Also in 2014, Zhang examined the relationship between the spot and time charter rates as well as the relationship between spot rates and forward freight rates and concluded to the fact that a VECM forecasting model seems to be the most appropriate for more accurate forecasts (Zhang, 2014)

In 2018, Taib C. and Mohtar developed a study about forecasting the spot freight rates based on FFA's and T/C contracts. VECM model was used to forecast the spot rates and also VECM with the ordinary least square method, while MAD and RMSE were used to evaluate the forecasting performance of the two models concluding to the fact that FFA was more suitable for managing the volatility of the spot market (Taib 2018)

In 2017, Geomelos and Xideas conducted a study which used forecasting techniques for the spot prices for both bulk shipping for dry cargo and tanker market in a disaggregated level. By using multivariate and univariate forecast models and comparing them they came into a conclusion that better and more accurate results could be driven out ot the combination of forecast models. (Geomelos, 2017)

In 2019, Epameinondas-Nektarios Diakodimitris conducted a study for the evaluation of various forecasting approaches for the Crude Oil Tanker Rates. By comparing ARIMA, ARMA-GARCH, VECM-VAR forecasting approaches with the evaluation of MAE RMSE and MAPE. The conclusions were that the forecasting methods led to significant forecasting methods attributed to the fact that the time series had a significant variability. Also ARIMA models produced better results for spot freight rates, while VAR-VECM models, outperformed every cased of time charter rates that has been examined. (Diakodimitris, 2019)

Time series forecasting

As the name tells by itself, times series is a collection of data through a specific time frame. Time series forecast techniques have been used throughout the year in many sectors with the purpose of trying to forecast the future. (Chatfield, 2001)

There are three types of forecasting methods

- **Judgemental forecasts** which are based on a judgement or any kind of information (Chatfield,2001)
- **Univariate time series forecast** which only depend on present and past values and being augmented by a time function (Chatfield, 2001)
- **Multivariate time series forecast methods** which forecasts depend on values of one or more additional time series variables called explanatory variables. (Chatfield, 2001)

The time series forecast models is a form of extrapolation which involves the fitting of past data in a model and then using that model outside that fitting of data with the expectation of the future data will be similar to the past data. (Chatfield, 2001). This type of extrapolation is regarded with disfavour in other statistical methods but it is considered as unavoidable in the time series analysis. Time series forecasts are also considered as conditional statements meaning that they are primarily based on the assumptions which are built into the model. (Chatfield, 2001)

Objectives of time series analysis

The main objectives of a time-series analysis are identification of the nature of phenomenon through the description of the data with the help of summary statistics and identification of any trends which can be easily obtained by the time plot of the series. Also, after the identification of this nature, the result will lead to the selection of the most appropriate forecasting model for the time series. After the selection of the model, the next objective of the time series is the forecast of the future values of the time series. It is very important to evaluate the results of the forecast and also to control the outcome which is the last objective of time series analysis and has led to the implementation of various forecasting performance measures and tools for the analysis of the outcome. (Chatfield, 2001)

Software used for analysis

R Studio and Eviews 11 (Student Version) software have been utilised for the purpose of forecasts and analysis of this study. These software contributed to the analysis of the data, conduction of various tests, forecasts and graphical analysis of every model which will be used in the study.

Components of time-series analysis

It is necessary to refer the four main components which are presented in the time-series

Seasonal Variation.

This type of variation is annual in period and is being used in many series which may be weekly or monthly and do present similar patterns of behaviour at particular times of the year. (Chatfield, 2001). A significant example to see where the seasonal variation is being implemented, is the annual economic report of almost every business is separated into quarterly seasonal components.

Cyclical Variation.

This type of variation can describe the movement of the series which has been occurred under cyclical patterns. (Chatfield, 2001). These cyclical patterns are a result of the uncertain nature of the economy which leads to certain phases in the economy such as peaks, recessions, toughs and recoveries which lead to business and economy cycles.

Trend.

This type of variation exists when there is observed an upward growth, downward decline or stagnation over several successive time periods. (Chatfield,2001). Thus it can be inferred that the trend is a long term movement in a time series (Adhikari, Agrawal, 2013). Also it can be inferred that the trends can be linear and non-linear with the first category being the most common among trend patterns.

Irregular fluctuations.

This includes every variation after the removal of any trend or seasonality and may be completely random and cannot be forecasted. (Chatfield, 2001) Those fluctuations can be a result of incidents such as wars, physical destructions e.t.c.

All of the above components can be combined to create the values of Y ordered over time. An important purpose of time series analysis is to isolate each of these three components and by this way, it can be concluded how much each component affects the Y value. It is also worth noting that identification of the pattern of a time series, is complicated by the presence of random error (Gerbing, 2016)

2. Methodology

The purpose of this thesis is to evaluate different existing forecasting methods including univariate and multivariate ones and more complex methods and compare them with univariate and multivariate models which have an exogenous variable implemented. The reason for this research is to test whether the most complex models can improve the poor results of the existing forecasting models. The data that will be used for the evaluation of the forecasting models are the spot and time charter rates for a Panamax and Capesize trading route and historical data of fleet size and the Baltic Dry Panamax and Capesize index are going to be incorporated as endogenous variables of the multivariate models. Also, historical data of the monthly global Crude Oil Production are going to be incorporated for the development of the models with exogenous variables.

Univariate and multivariate models are going to be used into this thesis for various reasons. From the point of univariate models, Naïve, ARIMA, ARIMAX and 3 different types of GARCH models are going to be used for the comparison and evaluation of forecasts for the dry bulk freight market. As far as the multivariate models concerned, a Vector Autor Regressive model is going to be implemented into the study. Apart from this method, a Vector Autoregression Model with an exogenous variable is going to be used along with a Vector Error Correction model to check if it provides better forecasts and to examine any cointegration relationship between the endogenous variables before converting it into a Vector Autoregression Model. The priority of this study is to produce valid models which can pass every diagnostic test. After interpreting the results, comparisons are going to be made between the results of each model to see which one is the most accurate as well as comparisons between multivariate and univariate models to check If the complexity of a model leads to more accurate results and also comparisons between standard models and the same models with the implementation of an exogenous variable to check if the implementation of an exogenous variable leads to more accurate results.

2.1 Data Specification

The data of freight rates that are going to be used for this study are retrieved from the Clarksons Shipping Intelligence Network which is provided by the Clarksons, a leading shipbrokerage institution. The main function of this Network is to obtain and store all different kinds of data of the shipping industry which are later used for various studies, researches and reports from different individuals, companies and institutions linked with the shipping industry. Data from January 2010 to December 2018 will be used from two different types of vessels, Panamax and Capesize which are mostly employed for medium and long-distance routes and these trade routes seem to be the most volatile in terms of freight rates throughout the past decades. Also, the data are going to include both voyage charters and time charters because it is important to identify the efficiency of the forecasting techniques on the different types of employment. As far as the supplementary data, Fleet size of the Capesize and Panamax global market and historic data of the Baltic Dry Capesize and Panamax index from January 2010 to December 2018 are going to be used. The implementation of this data to the Vector Autoregression model leads to sufficient forecasting results. Global Crude Oil Production historic data from January 2010 to December 2018 are going to be implemented into this study as an exogenous variable of the shipping markets.

CATEGORY	TIME SERIES
PANAMAX VOYAGE CHARTER (Coal, Baltimore/ARA, 75.000 dwt)	January 2010-December 2018
PANAMAX TIME CHARTER (75.000 dwt)	January 2010-December 2018
CAPESIZE VOYAGE CHARTER (Ore, Tubarao/Qingdao 170.000 dwt)	January 2010-December 2018
CAPESIZE TIME CHARTER (75.000 dwt)	January 2010-December 2018
CAPESIZE MARKET FLEET SIZE (in 000’s)	January 2010-December 2018
PANAMAX MARKET FLEET SIZE (in 000’s)	January 2010-December 2018
BALTIC DRY PANAMAX INDEX	January 2010-December 2018
BALTIC DRY CAPESIZE INDEX	January 2010-December 2018
GLOBAL CRUDE OIL PRODUCTION	January 2010-December 2018

Time series selected for the models (Table 1)

Also, time plots are developed for these different kinds of data in order to observe the plot throughout those years. As it can be seen, 2008 has been a difficult year for the shipping industry and thus it can explain the drop in rates.

Below, the descriptive statistics for each time series will be presented

	SPOT_FREI...	SPOT_FREI...	TIME_CHAR...	TIME_CHAR...	BDCI	BDPI	FLEET_CAP...	FLEET_PAN...
Mean	17.93894	11.29649	16449.00	11917.77	1936.450	1320.600	293.9766	184.1237
Median	18.01250	10.99375	14928.13	11006.25	1797.548	1183.158	308.3180	194.2235
Maximum	31.58750	22.90000	37600.00	29625.00	4626.000	4304.895	361.0310	228.6209
Minimum	5.625000	4.820000	5093.750	4781.250	-243.0500	324.1905	170.1557	120.2100
Std. Dev.	5.795767	3.354425	7095.886	4926.168	1031.735	731.8756	48.13451	28.91301
Skewness	0.093724	0.770360	1.133857	1.648486	0.383490	1.858694	-0.980081	-0.711519
Kurtosis	2.559170	4.079204	4.121328	5.955637	2.651639	7.257197	3.108624	2.485461
Jarque-Bera	1.262073	19.46176	35.19948	107.8320	3.902874	175.6848	21.19720	12.59384
Probability	0.532040	0.000059	0.000000	0.000000	0.142070	0.000000	0.000025	0.001842
Sum	2367.940	1491.137	2171268.	1573146.	255611.4	174319.2	38804.91	24304.33
Sum Sq. Dev.	4400.411	1474.034	6.60E+09	3.18E+09	1.39E+08	70169086	303518.0	109511.1

Descriptive Statistics of time series

From these values, the coefficient of variation for each time series can be calculated to capture the variability of each time series and underline the difficulty of the univariate forecasting methods to produce accurate results.

Values	Cape SR	Cape TC	Pana SR	Pana TC	BDCI	BDPI	FLEET CAPE	FLEET PANA
Coefficient of Variation	32.3%	29.6%	43.13%	41.33%	53.24%	55.41%	16.37%	15.7%

This table shows the increased variability that the time series have and this shows that more complex forecasting methods might show better results.

2.2 Seasonality

Seasonality is the presence of different variations which can be seen in the time series data and occur at specific intervals which last less than a year (typically monthly or quarterly). Those effects are a result of different natural and social effects such as seasons and holidays of people. Seasonality in the time-series data can usually be identified by observing the regular peaks in the autocorrelation function and the partial autocorrelation function (Diakodimitris, 2019). In the ACF, it shows how data points are related to the preceding data points and the PACF shows a summary of the relationship between an observation in a time series with observations of its lags with the relationships of intervening observations removed (Box, Jenkins and Reinsel, 1994). When forecast methods are going to be used in the time series, seasonality has to be removed from the data because seasonality can create some fluctuations in the data which can outrun other significant fluctuations that they might have.

2.3 Stationarity

A stationary time series is the one whose descriptive statistical properties such as the mean and variance do remain constant throughout the time. In other words that the variable distribution is time independent. Most univariate and multivariate statistic methods assume that stationarity of the data exists. (Duke, 2010). Another reason for the need for stationary time series data is to be able to obtain meaningful descriptive statistics data (Duke, 2010). There are three different tests which are used to see if stationarity is achieved in the time series. For the data used, time series are not stationary and as a result the first differences (D1) of the data are going to be used to the methods.

Dickey-Fuller test

This test examines two hypotheses where the null hypothesis of the following equation

$$\Delta Yt = a + \beta_t + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-j} + \cdots + \delta_{p-1} \Delta Y_{t-p+1} + \varepsilon_t$$

Where β is a coefficient of time, α represents a constant and p represents the lag order of the process.

The H0 hypothesis implies that γ=0, while the alternative hypothesis implies that γ<0.

The DF test has some serious weaknesses when it comes to rejection of the null hypothesis of a unit root, especially under the presence of mean reversion which is long compared to the length of the sample (Maddala & Kim, 1998).

This is remedied by performing an alternative test, the KPSS test, which assumes that the variable is stationary and tests the alternative hypothesis of a unit root. The KPSS test is generally more suitable for relatively small samples (Caner & Kilian, 2001; Kuo and Tsong, 2005).

KPSS test

Unlike the Dickey Fuller test which tests the null hypothesis that the time series is integrated of order one, the KPSS test describes the opposite hypothesis. (Kwiatkowski, Phillips 1992).

This test is based on linear regression, breaking up the

series into three parts: a deterministic trend (βt), a random walk (r_t), and a stationary error (ε_t),

with the regression equation:

$$x_t = r_t + \beta t + \varepsilon_t$$

where $r_t = r_{t-1} + u_t$ and $u \sim (0, \sigma^2)$ and are white noise(iid).

The KPSS test is used as a complement to the DF test so as a result, both tests are going to be implemented into the data series for the stationarity. (Palachy, 2019)

Phillips-Perron Test (PP)

This test also tests the stationarity of the data set and is set to be as an alternative of the DF test and is mostly used to series which have structural breaks (Perron, 1989). Also, according to Davidson, the PP test performs no better than the DF test and for those reasons it is not going to be used in this study.

2.4 Forecasting Main Evaluation Measures

After producing a series of forecasting results with the different univariate and multivariate methods, it is important to be able and evaluate each forecast. There is a number of indicators which provide us with a comparison and evaluation of the different models which will be presented below. The forecasting accuracy of the models can be shown by the forecast errors that they present. The forecast error is the difference between the value of the forecast and the actual value. There are many reasons that a forecast error occurs with the most significant being the randomness which shows errors that have nothing to do with the model and the bias which refers to a consistent mistake that is being made (Tsioumas, 2016).

2.4.1 Root Mean Square Error (RMSE)

The MSE shows the measurement of the average of the squared difference between the result of the forecast and the actual value which occurred. The RMSE likewise, measures the deviation of this measurement.

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2}$$

So n shows the number of the forecasted observations, yj are the real values at the period j and the other are the forecasted values. (Diakodimitris, 2019). In opposition to the MAE which is going to be explained below, the MSE seems to be more appropriate for the type of forecasts which do not show a big number of errors. The main disadvantage of MSE is that it is presented to be highly vulnerable to outliers where by squaring them makes their effect even more significant (Tsioumas, 2016). The RMSE is the square root of MSE but this does not affect the performance of the measure of MSE neither positive nor negative. (Tsioumas, 2016)

2.4.2 Mean Absolute Error (MAE)

MAE is the average of absolute errors and it indicates an approximation of the forecasted values to the actual values. (Tsioumas, 2016) It measures the average importance of the errors in their absolute shape.

$$\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N}\left|F_t - A_t\right|$$

Where Ft is the value of the forecast, At is the actual value and N is the size of the sample (Tsioumas, 2016)

2.4.3 Mean Absolute Error Percentage (MAPE)

This indicator shows many similarities to the MAE but unlike this indicator, it measures the percentage of error between the forecasted values and the actual values. As this indicator presents the results in a form of percentage, it makes it more comprehensive. (Diakodimitris, 2019).

$$MAPE = \frac{100\%}{n}\sum_{j=1}^{n}\left|\frac{y_j - \hat{y}_j}{y_j}\right|$$

Where n shows the number of the observations that have been forecasted, yj shows the real values at period j and the other shows the forecasted values

2.4.4 Akaike information criterion (AIC)

This is an indicator which shows the grade of how much suitable a forecasting model is for the initial data. The AIC is used for the comparison between different forecasting models to show the most suitable. According to Sakamoto, the AIC score is useful only if it is used to compare two models. (Sakamoto, 1986). The lower the number of the AIC is, the more optimal the model will be.

2.4.5 Schwarz criterion (SC) / Bayesian Information criterion (BIC)

It is defined as a criterion in econometrics used similarly with AIC described above, to select the best model and it can be defined as:

$-2L_m + m\ln n$

Where n is the size f the sample, m is the number of the parameters of the model and Ln represents the Log Likelihood of the model. (Tsioumas, 2016)

$$Y_t = a + bY_{t-1} + e_t$$

2.5 Forecasting Methods

2.5.1 Naïve method

In this method several assumptions are made that the series are stationary and have no seasonality. An AR(1) process is going to be used. The following equation represents the model.

$$Y_t = a + bY_{t-1} + e_t$$

Where Yt represents the value at period t, b is the coefficient at time t-1 and a represents a constant. (Diakodimitris, 2019)

2.5.2 ARIMA Approach and Box-Jenkins methodology

In the case of the AR(1) process, ψ1 is equal to the rate of the decay that ψi has. Instead of increasing the the order of the AR model as normal, Wold in 1938 combined it with the add of an MA(1) term which will adjust ψ1 while this change will have no effect on the exponential decay of ψι. This combination resulted in the creation of a mixed autoregressive moving average model (ARMA(1,1)). The ARMA (p,q) is given as:

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

$$= \delta + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \qquad 2.5.2.1$$

or

$$\Phi(B)\, y_t = \delta + \Theta(B)\, \varepsilon_t \qquad 2.5.2.2$$

Where εt is a white noise process with zero mean and variance σ^2. B is the backshift operator and ϕ(z) and θ(z) are polynomials of order p and q, assuming that they have n roots for |z|<1 to ensure causality (Brockwell and Davis, 1986).

Stationarity of the ARMA (p,q) process

The stationarity of this model is linked to the AR component of the model and can be examined through the roots of the polynomial.

$$m^p - \phi_1 m^{p-1} - \phi_2 m^{p-2} - \cdots - \phi_p = 0. \qquad (2.5.2.3)$$

Where if all the roots of the above polyonimal are less than 1 (<1) which is an absolute value, then the stationarity of the model is confirmed

Invertibility of ARMA (p,q) process

The invertibility of an ARMA forecast is related solely to the MA component of the polyonym and it can be examined through the following

$$m^q - \theta_1 m^{q-1} - \theta_2 m^{q-2} - \cdots - \theta_q = 0 \qquad (2.5.2.4)$$

Similarly as the stationarity process, if the roots of the above equation are less than one in the absolute value, the ARMA (p,q) forecast is considered as invertible .

Box-Jenkins Methodology

There are 4 different steps which are included in this methodology and are going to be explained below.

The first step includes short and seasonal differentiation to the model to achieve stationarity of the mean and also logarithmic transformation in order to achieve stationarity in the variance. It has to be answered whether the data is stationary or non-stationary by checking the fluctuation of the values with the plot or other means. The plot of ACF can be very useful to check the stationarity of the values. According to Diem Ngo (Diem Ngo, 2013) by looking at the ACF plot, if the ACF cuts down quickly then the time series can be considered as stationary, otherwise they must be considered as non- stationary. In case that the series are seasonal, multiplicative seasonal models combined with long-term differencing might be the most optimal way to achieve stationarity of the mean but on the other side it can be observed that there are never enough data to determine the appropriate seasonality level for the ARIMA model (Diakodimitris 2019)

The second step includes the selection of the most appropriate model. To achieve the perfect fit, partial and non-partial autocorrelation coefficients can be used to understand which are the most optimal p and q values (Makridakis, 1997). Also, many kinds of graphs of the stationary and differentiated data are used to understand which is the most suitable ARIMA model. (Hyndman, 2001). Again, there might be a difficulty that more than one model might be suitable and this will push the researcher to choose on the most suitable forecasting method, without knowing the implications of his or her choice (Diakodimitris, 2016). To solve this problem, the principle of parsimony was introduced which explains in case of many options available, the simpler choice must be chosen.

The third step includes the estimation of the coefficients of the parameters. This can be easily achieved by the method of non-linear optimization, based on the method of the steepest descent which is used to estimate the coefficient of each parameter (Marquardt, 1963). As this procedure is pretty straightforward, there is one difficulty which can be occurred and this is the inability to guarantee a global optimum. (Diakodimitris, 2016). Computational algorithms are used, mainly the Ordinary Least Square (OLS) and Maximum Likelihood (ML).

The fourth step includes the diagnostic check of the model where the residuals of the forecasting procedure are examined on whether they are white noise or not (Makridakis 1997). By conducting the Diagnostic check, if it is resulted that the model is inadequate, the Box-Jenkins method has to be repeated from the second step. (Chistidoulopoulos, 2020). Otherwise, the forecast result can be produced. In other words, the diagnostic test checks the degree of randomness of the residuals.

In this method, several packages from the R Studio are going to be utilised. First of all, stationarity tests are going to be conducted and ACF and PACF plots are going to be drawn to find the optimal lags for the model. Then, the auto.arima function is going to be utilised to find the optimal AR() I() and MA() components for the model. After that, a Ljung Box test is going to be conducted to the residuals and the forecast() function is going to be used and then the evaluation measure values are going to be obtained.

2.5.3 ARIMA-X with the implementation of an exogenous variable

This is an extension of the standard ARIMA model with the incorporation of an exogenous variable. The model that is going to be incorporated are of the form of :

$$
\begin{aligned}
y_t &= \beta x_t + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - \theta_1 z_{t-1} - \cdots - \theta_q z_{t-q} + z_t y_t \\
&= \beta x_t + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - \theta_1 z_{t-1} - \cdots - \theta_q z_{t-q} + z
\end{aligned}
$$

Where xt represents the covariate and β represents the coefficient of the covariate. The existence of the lagged values of the variable show that β can only be interpreted on the value of the previous values of the response variable (Hyndman, 2010). Just like in the ARIMA model, stationarity of the variables must be ensured. After checking the correlation of several exogenous variables, the most optimal variable is going to be obtained

Just as in the ARIMA model. Stationarity of all the variables including the exogenous ones is going to be ensured before proceeding to the model.

In this particular research, the following version of ARIMAX model is the following

$$
\begin{aligned}
y_t &= \beta \chi_{t-1} + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - \theta_1 z_{t-1} - \cdots - \theta_q z_{t-q} + z_t y_t \\
&= \beta x_{t-1} + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - \theta_1 z_{t-1} - \cdots - \theta_q z_{t-q} + z
\end{aligned}
$$

2.5.4 ARCH and GARCH approach including different GARCH models

2.5.4.1 ARCH approach

The autoregressive heteroscedasticity (ARCH) model was firstly developed in 1982 and was used to develop a forecast model to make the model based on the changing variance of a times series (Diakodimitris, 2019). By assuming the fact that a financial asset has a return series of rt is often a serially uncorrelated sequence with a zero mean (Diakodimitris, 2019). The result of this is the fact that its conditional variance is not constant. In case that the model presents a large squared return, this might indicate a volatile period and a small squared return may indicate a low volatility period. (Diakodimitris, 2019). Basically, this model represent a regression model with the conditional volatility as a response variable and the variable being the past lags of the squared return those variables which can be seen below. (Diakodimitris 2019)

where,
$$r_t = \sigma_{t|t-1}\varepsilon_t$$
$$\sigma^2{}_{t|t-1} = \omega + \alpha r_{t-1}^2$$

α and ω in the equation represent the unknown values, εt represent the independently distributed random variables with a mean that equals to zero.

The main use of this forecast model is for the prediction of the future conditional variances which can be seen by the equation example below. For example, If the h-step is asked to be forecasted:

$$\sigma^2{}_{t+h|t} = E(r^2{}_{t+h}|r_t, r_{t-1}, \dots)$$

For h=1 the above equation can be computed as followed:

$$\sigma^2{}_{t+1|t} = \omega + \alpha r_t^2 = (1 - a)\sigma^2 + \alpha r_t^2$$

This equation represents the weighted average of the variance and the current squared return (Diakodimitris, 2019)

2.5.4.2 GARCH approach

In 1986 Taylor and Bollerslev, made another approach with new introductions into this model, including the p lag of the conditional variance. This inclusion of the p lags resulted into the formation of a new model, the autoregressive conditional heteroscedasticity (GARCH (p,q)) model with the following formation

$$\sigma^2{}_{t|t-1} = \omega + \beta_1\sigma^2{}_{t-1|t-2} + \cdots + \beta_p\sigma^2{}_{t-p|t-p-1} + \alpha_1 r_{t-1}^2 + \cdots + \alpha_q r_{t-q}^2$$

Something that must be underlined for this model is the fact that the coefficients of the GARCH model are constrained to be non-negative, because the conditional variances have to be non-negative. (Nelson and Cao, 1992)

2.5.4.3 ARMA-GARCH approach

Referring to this forecasting model, the conditional mean structure is modelled by an ARMA(u,v) model and the white noise term will be modelled by a GARCH(p,q) model. (Diakodimitris, 2019). It can be seen by the following equation that:

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_u Y_{t-u} + \theta_0 + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots \theta_v e_{t-v}$$

where,

$$e_t = \sigma_{t|t-1}\varepsilon_t$$

and

$$\sigma^2{}_{t|t-1} = \omega + \alpha_1 e_{t-1}^2 + \cdots + \alpha_q e_{t-q}^2 + \beta_1\sigma^2{}_{t-1|t-2} + \cdots + \beta_p\sigma^2{}_{t-p|t-p-1}$$

Where Yt are considered to be the freight rate time series.

For this procedure, the Eviews software is going to be utilised. First of all, stationarity of the time series is going to be checked with ADF test in the software and the values are going to be transformed into logarithms. After that, the time series are going to be checked for ARCH effects, before proceeding further. After testing the validity of the time series, the auto.arima function is going to be utilised to find the optimal AR() and MA() components to fit in the GARCH model. After running the model into the eviews and checking the Log Likelihood and AIC that is the most optimal for the model, the E-Garch model is going to be produced as well to find which model is the most optimal based on AIC. After choosing the optimal model, the forecast procedure is selected from the software and the forecast results are obtained, along with the forecast evaluation measures.

2.5.4.4 E-GARCH approach

This can be specified as an exponential Garch model which has been developed by Nelson and Cao and can be seen below by the following equation:

$$\log \sigma_t^2 = \omega + \sum_{k=1}^{q} \beta_k g(Z_{t-k}) + \sum_{k=1}^{p} \alpha_k \log \sigma_{t-k}^2$$

This type of GARCH model is mainly used in the sector of asset pricing. The g(Zt) allows the variable to have different effects on the volatility. (Nelson and Cao, 1992)

2.5.4.5 Arch effects and testing

Lagrange multiplier (LM) Test: The use of this test is for the investigation of whether the variables of the model have ARCH effects which is crucial before developing the GARCH model. It can be defined as

LM = TR2 (Greene, 2003)

Ljung-Box Test: This test is used to check whether autocorrelation exist of order m in the residuals. The equation for this test can be seen below:

$$Q(m) = T(T+2)\sum_{l=1}^{m}\frac{\hat{\rho}_l^2}{T-l}.$$

2.5.5 Vector Multivariate Models

2.5.5.1 Vector Autoregression model (VAR)

The Vector Autoregression model includes endogenous variables in the linear function as a function with the variable's lags and all the other endogenous variables lags in the system. (Tsioumas 2016). It is considered a multivariate econometric model. The equation of the model can be described below

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-n} + \varepsilon_t \tag{2.5.1}$$

Where c is a vector of intercepts, Ai is a time variant matrix of m x m and εt is a vector of error terms which are uncorrelated with zero mean and no serial correlation. (Tsioumas 2016).

Apart from those conditions, another condition that should exist is the following

$$\det(I_K - A_1 z - \cdots - A_p z^p) \neq 0 \; for \; |z| \leq 1 \tag{2.5.2}$$

Where the polynomial will not have unit roots and time series must be stationary. VAR consists an econometric approach which can contribute to the investigation of many relationships between the variables such as Granger's causality apart from the forecasting matters (Tsioumas 2016)

For this procedure, the Eviews software is going to be utilised. After conducting a Unit Root test to check the stationarity, the variables of the model are converted into first differences I(1). Later, the optimal lags of the model are utilised, having in mind that the degrees of freedom are at a rational level and also the model does not have any serial correlantion at the chosen lag level. After this procedure, the VAR model is created. After that, Granger's Causality tests are conducted for the endogenous variables and also, the Impulse response and Variance Error Decomposition Graphs are obtained. After this procedure, the forecast results are obtained for the selected period using the "Static Forecast" option along with the forecast evaluation measures.

2.5.5.2 Vector Autoregression model with Exogenous Variable (VAR-X)

This model is an extension of the standard Vector Autoregression model which includes one or more exogenous variable along with the endogenous ones. (Tsioumas 2016). The following equation represents the standard VAR-X model:

$$Y_t = \lambda_0 + J_1 Y_{t-1} + \cdots + J_n Y_{t-n} + L_1 X_{t-1} + \cdots + L_m X_{t-m} + U_t \tag{2.5.3}$$

Where Y represents the endogenous variables, X represents a vector of the exogenous variables L and J are coefficient matrices and λ0 is a vector of intercepts (Tsioumas 2016)

All of the endogenous and exogenous variables of the model do need to be stationary or in first level differentiated.

There is an important requirement on the VAR-X model which is represented by the following condition (Tsioumas 2016).

$$E[U_t \mid \{Y_{t-i}\}_{i=1}^{\infty}, \{X_{t-i}\}_{j=1}^{\infty}] = 0 \quad (2.5.4)$$

For this procedure, the Eviews software is going to be utilised. After conducting a Unit Root test to check the stationarity of every variable, along with the exogenous one, the variables of the model are converted into first differences I(1). Later, the optimal lags of the model are utilised, having in mind that the degrees of freedom are at a rational level and also the model does not have any serial correlantion at the chosen lag level. After this procedure, the VAR-X model is created. After that the same procedure is being followed as in the VAR model process which is described above.

2.5.5.3 Granger's Causality

This test is used in the Vector Autoregression model and shows whether the particular variable Y Granger causes a variable X . It is very important to conduct this kind of test when using a VAR model in I(1), otherwise the model might not be valid if the criteria from the test will not be met.

2.5.5.4 Impulse Response

This type of analysis is considered as a complementary analysis further to the Granger's and shows graphically the way that each testing variable interacts with the other. (Tsioumas 2016) This analysis finds the way each variable reacts to the other to an impulse into a particular system. It is important to consider that this type of analysis is only helpful if the data of the Vector Autoregression model are either stationary or in first levels (I1).

2.5.5.5 Forecast Error Variance Decomposition (FEVD)

This type of analysis can show the percentage amount of the variance of the forecasted data which can be explained by the randomness of the endogenous values of the Vector Autoregression model (Tsioumas 2016). This analysis can show the amount of information which is being presented by each value of the model and this can be pretty important towards the decision making process of the model. (Tsioumas 2016)

2.5.6 Vector Error Correction Model (VECM)

The Vector Error Correction model actually is a restricted version of a VAR model which has an additional error correction after following a process to check whether the variables included in the model are cointegrated. (Tsioumas 2016) The term of cointegration is actually describing the speed of adjustment to a long-term equilibrium after a change of a variable which is independent (Tsioumas 2016). The co-integration of the variables is a pre-requisite before running the VECM model. If the variables are not co-integrated, a Vector Autoregression model might be the optimal solution. The VECM model is mainly used for variables which are not stationary, but co-integrated. The equation of a VECM model is the following

$$\Delta Y_t = \mu_t + \Pi Y_{t-1} + \sum_{i=1}^{n-1} \Gamma_i \Delta Y_{t-i} + \varepsilon_t \qquad (2.5.5)$$

Where μt is the deterministic term, Yt is a vector which has the endogenous variables , Δ indicates the differentiated data, Γi and Π are coefficient matrices which show the adjustment of long and short term and εt is a vector of the error terms. (Tsioumas 2016)

For this procedure, the Eviews software is going to be utilised as with the VAR and VAR-X model. First of all, the first differences of the endogenous variables are going to be obtained. After that, the optimal lags of the model are going to be determined in respect to the validity of the model (serial-correlation). Trace and maximum Eigenvalue Co-Integration tests are going to be conducted to find the number of co-integration relationship. A VECM model with p-1 lags and n as the number of the co-integration relationships is created and after that the Impulse response and Variance Error Decomposition graphs are obtained. For the last step, the forecast is obtained for the selected period, along with the forecast evaluation values.

2.5.7 Vector Error Correction With Exogenous Variable (VECM-X)

Similarly with the VAR-X model, this model is an extension of the standard Vector Error Correction model with the inclusion of one or more exogenous variables. The following equation represents the model : (Tsioumas 2016)

$$\Delta Y_t = \mu_t + \Pi Y_{t-1} + \sum_{i=1}^{n-1} \Gamma_i \Delta Y_{t-i} + \sum_{j=1}^{m-1} B_j \Delta X_{t-j} + \varepsilon_t \qquad (2.5.6)$$

Where Xi is a vector which includes the exogenous variables, μt is the deterministic trend Yt is a vector which has the endogenous variables, Δ indicated the diffetentiated data, Γi and Π are coefficient matrices and εt is the vector of the error terms (Tsioumas, 2016).

In this model, the same procedure is being followed for the VECM-X model but with the difference that the matrix of the differentiated exogenous variable is being implemented before proceeding to find the co-integration and optimal lags.

3. Results

After utilising the data of all the time series which are examined from 2010 to 2018, forecasts for the period of January 2019 to December 2019 are produced and the evaluation criteria of RMSE, MAE and MAPE, measure the accuracy of these forecasts.

3.1 Naïve Method

Forecasting results produced by the Naïve model (Table 2)

Time Series	RMSE	MAE	MAPE
Capesize Spot	2.9589	2.0811	11.5237%
Capesize Time	2163.756	1623.855	10.0321%
Panamax Spot	1.4972	1.1542	10.0892%
Panamax Time	1211.171	834.3575	7.2115%

3.2 AR(I)MA Method

First of all, stationarity tests will be conducted to see If the variables need to be transformed into levels, as well as normality tests. For these tests, Augmented Dicky Fuller test and Shapiro Wilk normality tests are going to be conducted to the variables. At the table below, the rejected results will be highlighted with the yellow colour.

ADF and Shapiro-Wilk Normality tests before proceeding to the orders of the AR(I)MA model (Table 3)

Time Series	Augmented Dickey Fuller Test	Shapiro Wilk Normality Test
Capesize Spot	0.5142	0.1009
Capesize Time	0.2128	5.2e
Panamax Spot	0.1681	0.0072
Panamax Time	0.056	8.885e

Orders of AR(I)MA model after the utilisation of the Auto.Arima function in Rstudio (Table 4)

Time Series	**AR() I() MA()**
Capesize Spot	(2,1,1)
Capesize Time	(0,1,0)
Panamax Spot	(1,1,2)
Panamax Time	(1,1,1)

After finding the proper components for the model, the Ljung-Box Test will be conducted for the test of the residuals of the model.

Ljung-Box test on residuals (Table 5)

Time Series	**Ljung-Box Test**
Capesize Spot	0.08303
Capesize Time	0.6739

Panamax Spot	0.2018
Panamax Time	0.4135

ACF and PACF tests have been conducted each model as well as further tests on residuals which are graphically represented and can be found at the latter stage of this thesis.

After finding the optimal components for each model, the accuracy of each model can be measured below

Forecasting results produced by the AR(I)MA model (Table 6)

Time Series	RMSE	MAE	MAPE
Capesize Spot	2.8159	2.0844	11.8995%
Capesize Time	2153.718	1609.168	9.9401%
Panamax Spot	1.397694	1.042919	9.3122%
Panamax Time	1151.441	794.7175	6.5830%

3.3 AR(I)MA-X method

Before proceeding into the model, a Pearson-Correlation test has been conducted between the endogenous and the exogenous variable of Global Crude Oil Production that has been selected. Apart from the Global Crude Oil production, some other exogenous variables have been tested and has been decided that the Oil production is the most appropriate variable for this model.

Pearson's Correlation Test on exogenous variables, before determining the variable which will be used (Table 7)

	Libor	Steel_Chi	Oil Production	Inflation
Cape_sr	-0.08	-0.33	-0.49	-0.08
Cape_ts	-0.04	-0.3	-0.44	-0.04
Pana_sr	-0.08	-0.41	-0.56	-0.08
Pana_ts	-0.001	-0.34	-0.49	-0.001

Proceeding into the process, a matrix is being created with the exogenous variable and the AR and MA components can be found, before proceeding into the model evaluation

Orders and Results (Table 8)

Time Series	AR() I() MA()	Log Likelihood	AIC
Capesize Spot	(1,0,0)	-262.61	533.22
Capesize Time	(2,0,0)	-976.63	1963.67
Panamax Spot	(1,1,2)	-188.35	386.7
Panamax Time	(0,1,0)	-910.25	1824.5

After finding the components, the Ljung-Box test will be conducted to test the residuals

Ljung-Box test on residuals (Table 9)

Time Series	Ljung-Box Test
Capesize Spot	0.0648 (22 lags)
Capesize Time	0.6217 (22 lags)
Panamax Spot	0.1602 (22 lags)
Panamax Time	0.1755 (22 lags)

After conducting the residual test, the accuracy of each test can be shown below based on the accuracy measures.

(Forecasting results of the AR(I)MA-X model (Table 10)

Time Series	RMSE	MAE	MAPE
Capesize Spot	2.738809	2.05481	12.05249%
Capesize Time	2022.094	1633.192	10.6760%

Panamax Spot	1.396669	1.044732	9.34177%
Panamax Time	1192.046	873.0215	7.3486%

3.4 GARCH MODELS

For the first step, every variable has to be transformed into a log, because it is not stationary and a Hederoskedasticity Test must be conducted for the testing for ARCH effects, before proceeding into the modelling phase.

ARCH-LM testing for ARCH effects (Table 11)

Time Series	Heteroskedasticity Test
Capesize Spot (log)	0.7520
Capesize Time (log)	0.000
Panamax Spot (log)	0.000
Panamax Time (log)	0.000

Since the Capesize spot rates are transformed into a log and a level 1 difference and they still don't pass the test for ARCH effects, no model will be computed.

After the ARCH test, the components of the AR() and MA() will be conducted before proceeding into a GARCH(1,1) process

AR()MA() and Garch() orders (Table 12)

Time Series	AR() MA() / GARCH()
Capesize Spot	N/A
Capesize Time	(0,1)/ (1,1)
Panamax Spot	(0,2)/ (1,1)
Panamax Time	(1,1) / (1,1)

After finding the AR() , MA() and GARCH() components, 2 different models are being developed to select the most optimal based on AIC value and run the forecast.

1. Capesize Time Series

Model	R-Sqrd	Adj R-Sq	Log Lkl	Durbin-W	Akaike	Schwarz	HannanQ
Garch	0.6832	0.6802	-1025.029	0.40745	19.074	19.198	19.124

2. Panamax Spot Series

Comparison between S-GARCH and E-GARCH model results (Table 14,15)

Model	R-Sqrd	Adj R-Sq	Log Lkl	Durbin-W	Akaike	Schwarz	HannanQ
Garch	0.41918	0.41371	21.7011	0.5031	-0.2868	-0.1378	-0.2264
E-Garch	0.41906	0.41358	21.6949	0.4652	-0.2721	-0.0982	-0.2016

3. Panamax Time Series

Comparison between S-GARCH and E-GARCH model results (Table 16)

Model	R-Sqrd	Adj R-Sq	Log Lkl	Durbin-W	Akaike	Schwarz	HannanQ
Garch	0.9513	0.9504	-879.607	2.1563	16.5534	16.7032	16.6141
E-Garch	0.9503	0.9493	-895.463	2.3058	16.8684	17.0433	16.9393

Based on the testing procedure, we can conclude to the fact that the most optimal method for each time series is :

Selection of the optimal Garch model (Table 16)

Time Series	Model
Capesize Spot	N/A
Capesize Time	GARCH (standard)
Panamax Spot	E-GARCH
Panamax Time	GARCH (standard)

Now, after the decision about the most optimal model for each time series, we proceed with the accuracy evaluation of each model forecast

Forecasting results of the GARCH model (Table 17)

Time Series	RMSE	MAE	MAPE
Capesize Spot	N/A	N/A	N/A
Capesize Time	4311.719	3207.887	20.6852%
Panamax Spot	0.2373	0.1807	8.0482%
Panamax Time	1155.958	780.8209	6.56002%

3.5 VECTOR AUTOREGRESSION MODEL (VAR)

For the first step we determine the endogenous variables of the model which will be the following

Determination of the endogenous variables for the VAR model (Table 18)

Time series	Endogenous Variables
Capesize Spot	CP Fleet number, Baltic Dry Capesize Index
Capesize Time	CP Fleet number, Baltic Dry Capesize Index
Panamax Spot	PN Fleet number, Baltic Dry Panamax Index
Panamax Time	PN Fleet number, Baltic Dry Panamax Index

As none of the variables are found to be stationary, we convert every variable of the model into 1st differences (I1).

The next step is to find the optimal lags of the model with respect to the Serial-Correlation LM Test and also the lags used for the model must be within a reasonable range.

Selection of the Optimal Lags based on AIC and Serial-correlation test on Lag Level (Table 19)

Time Series	Lags	AIC Lag Criteria	LM test for Lag Level
Capesize Spot	4	22.1828	0.0944
Capesize Time	4	36.1742	0.6250
Panamax Spot	4	16.9650	0.2231
Panamax Time	4	13.8456	0.7798

As it has been found that there is no serial correlation at these lag levels, Granger's causality test can be conducted and later, the model can be constructed

Granger's Causality Test (Table 20)

Time Sr	Cape S	Cape T	Pana S	Pana T	Cape Fl	Pana Fl	BDCI	BDPI
Cape s					0.3452		0.4868	
Cape T					0.1679		0.0139	
Pana S						0.1018		0.6386
Pana T						0.0927		0.0011
Cape Fl	0.3692	0.4677					0.1823	0.7018
Pana Fl			0.0321	0.3452				0.0721
BDCI	0.4551	0.22227			0.5885			
BDPI	0.3846			0.0278		0.0659		

Results of the VAR model (Table 21)

Time Series	Determinant adjusted resid covariance	Determinant resid covariance	Log Likelihood	Akaike information criterion	Schwarz criterion	Number of coefficients
Cape Spot	657870.8	394307.8	-1091.324	22.3396	23.5749	48
Cape Time	7.62E+11	5.08E+11	-1826.595	36.2251	37.22277	39
Pana Time	4.16E+09	2.49E+09	-1537.694	31.09204	32.3273	48
Pana Spot	5276.886	3520.413	-859.0182	17.4372	18.4348	39

After examining the components of the VAR model for each variable, the evaluation accuracy procedure can be conducted

Forecasting results of the VAR model (Table 22)

Time Series	RMSE	MAE	MAPE
Capesize Spot	2.5875	2.0015	11.5155%
Capesize Time	1768.369	1413.353	9.7118%
Panamax Spot	1.2389	0.9601	8.9431%
Panamax Time	801.369	642.428	6.0756%

3.6 VECTOR AUTOREGRESSION MODEL WITH EXOGENOUS VARIABLE (VAR-X)

Selection of the endogenous and exogenous Variables for the VAR-X model (Table 23)

Time Series	Endogenous Variables	Exogenous Variable
Capesize Spot	CP Fleet number, Baltic Dry Capesize Index	Global Crude Oil Production
Capesize Time	CP Fleet number, Baltic Dry Capesize Index	Global Crude Oil Production
Panamax Spot	PN Fleet number, Baltic Dry Panamax Index	Global Crude Oil Production
Panamax Time	PN Fleet number, Baltix Dry Panamax Index	Global Crude Oil Production

As conducted in the VAR model, all of the variables are transformed into I(1) differences, since we have non-stationary values.

Proceeding into the Lag structure and criteria,

Selection of appropriate Lags based on AIC criteria and Serial-Correlation tests on Lag Level (Table 24)

Time Series	Lags	**AIC Lag Criteria**	**LM test for Lag Level**
Capesize Spot	5	22.20605	0.1356
Capesize Time	4	36.0195	0.2968
Panamax Spot	4	17.1525	0.0612
Panamax Time	4	31.2278	0.5762

Also, no serial correlation between the variables at the selected lags which means that the model can be valid, Granger's Causality Tests for the VARX model (Table 25)

Time Sr	Cape S	Cape T	Pana S	Pana T	Cape Fl	Pana Fl	BDCI	BDPI	Oil
Cape s					0.4727		0.4881		
Cape T					0.1137		0.0021		
Pana S						0.0800		0.6590	
Pana T						0.0271		0.0003	
Cape Fl	0.4602	0.1056					0.6365		
Pana Fl			0.0155	0.2890				0.0491	
BDCI	0.3833	0.2742			0.5388				
BDPI			0.3903	0.0077		0.0701			
Oil									

Results of the VAR-X model (Table 26)

Time Series	Determinant adjusted resid covariance	Determinant resid covariance	Log Likelihood	Akaike information criterion	Schwarz criterion	Number of coefficients
Cape Spot	595772.7	357088.1	-1086.268	22.2405	23.47583	48
Cape Time	6.61E+11	4.41E+11	-1819.286	36.0832	37.0808	39
Pana Time	3.96E+09	2.38E+09	-1535.228	31.04369	32.27897	48
Pana Spot	5106.821	3294.821	-855.6048	17.4292	18.5035	42

Forecasting Results of the VAR-X model (Table 27)

Time Series	RMSE	MAE	MAPE
Capesize Spot	2.5496	1.9930	11.8845%
Capesize Time	1739.088	1434.912	10.187%
Panamax Spot	1.2371	0.9548	8.8707%
Panamax Time	802.538	647.68	6.2294%

3.7 VECTOR ERROR CORRECTION MODEL (VECM)

As the literature insists, as long as the variables are non-stationary and they are co-integrated, it is more optimal to use a Vector Error Correction Model but for the sake of research, every possible model has been created for comparison purposes. Eviews 11 was utilised for this model.

Trace and Max Eigenvalue co-integration tests (Table 28)

Time Series	Cointegration Number (Trace)	Cointegration Number (Maximum Eigenvalue)
Capesize Spot	2	2
Capesize Time	3	1
Panamax Spot	2	2
Panamax Time2	1	1

These results insist on developing a VECM model for each of the time series with P-1 lags

Results of the VECM model (Table 29)

Time Series	Determinant adjusted resid covariance	Determinant resid covariance	Log Likelihood	Akaike information criterion	Schwarz criterion	Number of coefficients
Cape Spot	661604.7	410540.3	-1093.382	22.4388	23.7513	51
Cape Time	9.03E+11	6.44E+11	-1838.748	36.40288	37.3237	36
Pana Time	4.56E+09	2.93E+09	-1545.903	31.19418	32.35226	45
Pana Spot	5278.408	3640.118	-860.7402	17.5289	18.6032	42

Forecasting Results of the VECM model (Table 30)

Time Series	RMSE	MAE	MAPE
Capesize Spot	2.5998	2.009	11.6344%
Capesize Time	1769.359	1416.094	9.7289%
Panamax Spot	1.2418	0.9556	8.8971%
Panamax Time	862.230	679.788	6.4856%

3.8 VECTOR ERROR CORRECTION MODEL WITH EXOGENOUS VARIABLE (VECM-X)

Similarly with the VECM model but now with an exogenous variable, the Cointegration tests are being conducted.

Trace and Max Eigenvalue co-integration tests (Table 31)

Time Series	Cointegration Number (Trace)	Cointegration Number (Maximum Eigenvalue)
Capesize Spot	2	2
Capesize Time	1	1
Panamax Spot	2	2
Panamax Time2	1	1

Again as every model has one or more cointegrations, we proceed into the development of the VECM-X model.

Model Results (Table 32)

Time Series	Determinant adjusted resid covariance	Determinant resid covariance	Log Likelihood	Akaike information criterion	Schwarz criterion	Number of coefficients
Cape Spot	605878.6	363145.3	-1087.126	22.3750	23.7647	54
Cape Time	7.77E+11	5.36E+11	-1829.308	36.27782	37.27543	39
Pana Time	4.44E+09	2.75E+09	-1542.903	31.19046	32.42574	48
Pana Spot	5112.790	3410.938	-857.3913	17.5221	18.6732	45

Forecasting Results of the VECM-X model (Table 33)

	RMSE	MAE	MAPE
Capesize Spot	2.5456	1.9836	11.6556%
Capesize Time	1832.894	1494.544	10.5881%
Panamax Spot	1.2401	0.9503	8.8212%
Panamax Time	854.8678	677.5604	6.5515%

4. Comparison of Forecasting Results by vessel category

4.1 Capesize Spot Rates

Comparison between the different forecasting models for the Capesize Spot Rates (Table 34)

Forecasting Method	RMSE	MAE	MAPE
Naïve	2.9589	2.0811	11.5237%
AR(I)MA	2.8159	2.0844	11.8995%
AR(I)MA-X	2.7388	2.0548	12.0524%
GARCH	N/A	N/A	N/A
VAR	2.5875	2.0015	11.5155%
VAR-X	2.5496	1.99030	11.8845%
VECM	2.5998	2.009	11.6344%
VECM-X	2.5456	1.9836	11.6556%

In terms of the Capesize Spot rates, it can be observed that every forecasting method produces a similar results in RMSE, where the best result is produced by the VECM-X model and the second best by the VAR-X model, so this results in the fact that the multivariate models with the interpretation of an exogenous variable can produce the best results in terms of RMSE. In terms of MAPE, the results are very close to each other with the simple VAR model producing the best result. It is also observed that although the models with exogenous variables produce a less better result in terms of RMSE in comparison with the models without an exogenous variable. The weakest results are being produced by the ARIMA-X model in terms of RMSE and MAPE which shows the weakness of this forecasting method to produce a competitive forecast result. In terms of Theil's U, the VECM-X produced the best results. As a result, with VECM-X having the best results in 3 out of 4 evaluation measures, it can be concluded that it produced the best results but the difference between the most simple and the most complicated forecasting method is non-signficant.

ii) Comparison of the Standard vs the models which have an exogenous variable included

Comparison of standard and models with exogenous variables (Table 35)

Forecasting method	RMSE	MAE	MAPE
ARIMA	2.8159	2.0844	11.8995%
ARIMA-X	2.7388	2.0548	12.0524%

Forecasting method	RMSE	MAE	MAPE
VAR	2.5875	2.0015	11.5155%
VAR-X	2.5496	1.99030	11.8845%

Forecasting method	RMSE	MAE	MAPE
VECM	2.5998	2.009	11.6344%
VECM-X	2.5456	1.9836	11.6556%

As far as the Capesize Spot Rates concerned, every forecasting model which included an exogenous variable, produced worse results than the standard models without an exogenous variable in Terms of the MAPE measure. Although, in terms of RMSE MAE the models which included an exogenous variable produced better results than the standard models but also with a non significant difference. It is worth noting that the ARIMA-X produced the least satisfying results in terms of MAPE, out of eery forecasting method.

4.2 Capesize Time Charter Rates

Comparison between the different forecasting models for the Capesize TC Rates (Table 36)

Forecasting Method	RMSE	MAE	MAPE
Naïve	2163.756	1623.855	10.0321%
AR(I)MA	2153.718	1609.168	9.9401%
AR(I)MA-X	2020.094	1633.192	10.6760%
GARCH	4311.719	3207.887	20.6852%
VAR	1768.369	1413.353	9.7718%
VAR-X	1739.088	1434.912	10.187%
VECM	1769.359	1416.094	9.7289%
VECM-X	1832.894	1494.544	10.5881%

In terms of Capesize Time charter rates, it can be observed that the VAR, VECM and AR(I)MA models produced results with a MAPE of just under 10% with the best MAPE result being produced by the VECM model. The GARCH model did not perform adequately and there is no comparison between any other result in terms of MAE, MAPE and RMSE. The VAR-X model produced the best results in terms of RMSE and the in terms of MAE evaluation, the VAR model produced the best result. Also , the ARIMA model, although it produced an interesting MAPE result, the RMSE result, could not compete the results of the multivariate models. Also the AR(I)MA model, outperformed the ARIMA-X model.

ii) Comparison of the Standard vs the models which have an exogenous variable included

Forecasting method	RMSE	MAE	MAPE
ARIMA	2153.718	1609.168	9.9401%
ARIMA-X	2020.094	1633.192	10.6760%

Forecasting method	RMSE	MAE	MAPE
VAR	1768.369	1413.353	9.7718%
VAR-X	1739.088	1434.912	10.187%

Forecasting method	RMSE	MAE	MAPE
VECM	1769.359	1416.094	9.7289%
VECM-X	1832.894	1494.544	10.5881%

Comparison of standard and models with exogenous variables (Table 37)

As far as the Capesize time charter rates concerned, the standard models produced better results in terms of MAPE and MAE in comparison to the models with the inclusion of an exogenous variable. In terms of RMSE, the ARIMA-X and VAR-X model produced a better RMSE result than the standard models and the VECM-X model produced a worse result than the standard VECM model. Also, in this situation the standard univariate and multivariate models, outperformed the models with an exogenous variable but with a small difference in the percentage.

4.3 Panamax Spot Rates

Comparison between the different forecasting models for the Panamax Spot Rates (Table 38)

Forecasting Method	RMSE	MAE	MAPE
Naïve	1.4972	1.1542	10.0892%
AR(I)MA	1.3976	1.0429	9.3122%
AR(I)MA-X	1.3966	1.0447	9.3417%
E-GARCH	0.2373	0.1807	8.0482%
VAR	1.2389	0.9601	8.9431%
VAR-X	1.2371	0.9548	8.8707%
VECM	1.2418	0.9556	8.8971%
VECM-X	1.2401	0.9503	8.8212%

In terms of Panamax Spot rates, it can be observed that the E-GARCH model outperformed every other model in every evaluating measure. Particularly, the E-GARCH model had the lowest RMSE number by far, compared with every other forecasting method, as well as a MAE value. In terms of MAPE, the E-GARCH model almost reached the value of 8%. The VECM-X model produced the second best MAPE results. And the VAR-X model produced the second best RMSE results. In this situation apart from the E-GARCH model which produced the best results, it can be observed that the multivariate models with an exogenous variable, produced slightly better results than the simple multivariate models in every evaluating measure.

ii) Comparison of the Standard vs the models which have an exogenous variable included

Comparison of standard and models with exogenous variables (Table 39)

Forecasting method	RMSE	MAE	MAPE
ARIMA	1.3976	1.0429	9.3122%
ARIMA-X	1.3966	1.0447	9.3417%

Forecasting method	RMSE	MAE	MAPE
VAR	1.2389	0.9601	8.9431%
VAR-X	1.2371	0.9548	8.8707%

Forecasting method	RMSE	MAE	MAPE
VECM	1.2418	0.9556	8.8971%
VECM-X	1.2401	0.9503	8.8212%

As far as the Panamax Spot Rates concerned, the ARIMA model produced better results in every evaluating measure in comparison to the ARIMA-X model. The VAR-X model produced better results than the standard VAR model and it can be observed the same thing in terms of the VECM-X model which produces better results in every evaluating measure in comparison to the standard VECM model. Again, the difference in the results in not very significant.

4.4 Panamax Time Charter Rates

Comparison between the different forecasting models for the Panamax TC Rates (Table 40)

Forecasting Method	RMSE	MAE	MAPE
Naïve	1211.171	834.3575	7.2215%
AR(I)MA	1151.441	794.7115	6.5830%
AR(I)MA-X	1192.046	873.0215	7.3486%
E-GARCH	1155.958	780.8209	6.56002%
VAR	801.369	642.428	6.0756%
VAR-X	802.538	647.68	6.2294%
VECM	862.230	679.788	6.4856%
VECM-X	854.867	677.560	6.5515%

In terms of Panamax Time Charter Rates, it can be observed that the VAR model produced the greatest score in every evaluating measure, closely followed by the VAR-X model. The VAR model produced a MAPE value of almost 6% which is the lowest score and a value of almost 800 on RMSE. Every forecasting model produced good results and it can be observed the fact that the AR(I)MA-X model produced the worst results in terms of MAPE. This table also shows that the multivariate models performed slightly better than the univariate models in terms of MAPE and also the difference on the RMSE value between the multivariate and univariate models is sizeable.

ii) Comparison of the Standard vs the models which have an exogenous variable included

Comparison of standard and models with exogenous variables (Table 41)

Forecasting method	RMSE	MAE	MAPE
ARIMA	1151.441	794.7115	6.5830%
ARIMA-X	1192.046	873.0215	7.3486%

Forecasting method	RMSE	MAE	MAPE
VAR	801.369	642.428	6.0756%
VAR-X	802.538	647.68	6.2294%

Forecasting method	RMSE	MAE	MAPE
VECM	862.230	679.788	6.4856%
VECM-X	854.867	677.560	6.5515%

As far as the Panamax Time Charter Rates concerned, the standard ARIMA, VAR and VECM models produced the best MAPE results in comparison to the models with an exogenous variable. In terms of the RMSE, the ARIMA and VAR models slightly outperformed the models with an exogenous variable, while the VECM model produced worse results than the VECM-X model. Exactly the same situation with RMSE goes with MAE for every forecasting method examined.

4.5 Comparison of Forecasting Results based on Average Price of MAPE

i) For Spot Rates (Panamax and Capesize)

Average MAPE results for Spot Rates

Forecast Model	MAPE (AVERAGE)
NAIVE	**10.8%**
ARIMA	**10.6%**
ARIMAX	**10.69%**
GARCH	**N/A**
VAR	**10.22%**
VARX	**10.37%**
VECM	**10.26%**
VECMX	**10.23%**

ii) For Time Charter Rates (Panamax and Capesize)

Average MAPE results for Time Charter Rates

Forecast Model	MAPE (AVERAGE)
NAIVE	**8.62%**
ARIMA	**8.26%**
ARIMAX	**9%**
GARCH	**13.62%**
VAR	**7.92%**
VARX	**8.2%**
VECM	**8.1%**
VECMX	**8.56%**

By calculating the average MAPE number for each forecasting method, the fact that the VAR model slightly outperformed out of every forecasting model can be observed. Also, it is worth mentioning the fact that the models with an exogenous variable implemented have produced slightly worse results than the simple multivariate and univariate models, especially for the ARIMA-X model, which produced to worst result for the time charter rates.

5. Conclusion

5.1 Answers of the research and sub-research questions

In this part of the thesis, answers and conclusions for the main research question as well as the sub-research questions will be provided

For the main research question ***Does the inclusion of an exogenous variable lead to more accurate results of the existing univariate and multivariate models?***

The Results provided some very useful results for the models with the exogenous variables. First of all, the same exogenous variable has been used for every forecasting model.

For the RMSE evaluation measure, the ARIMA-X model provided better results for the Capesize freight rates and the ARIMA for the Panamax freight rates. The VAR-X model provided the best results in 3 out of 4 time series except for the Panamax TC rates. Same thing for the VECM-X as is provided the best results in 3 out of 4 time series except for the Capesize Time Charter rates. It can be concluded that in terms of RMSE, the models with an exogenous variable provided more accurate results in comparison to the standard models.

For the MAPE measure, the ARIMA model outperformed the ARIMA-X model in every time series. Also the VAR model, outperformed the VAR-X model in 3 out of 4 time series except of the Panamax Spot Rates. Finally, the same thing applies to the VECM model which outperformed the VECM-X model in 3 out of 4 time series. In conclusion, the standard models provided more accurate results in comparison to the models with an exogenous variable in terms of MAPE.

For the MAE measure, the ARIMA model outperformed the ARIMA-X model in 3 out of 4 time series. The VAR-X model outperformed the standard VAR model in 3 out of 4 time series. Finally, the VECM-X model, provided more accurate results in 3 out of 4 time series in comparison to the standard VECM model. In conclusion, in terms of MAE, the most accurate results were provided by the models with the implementation of an exogenous variable.

Overall, in terms of RMSE and MAE evaluating measures, the models with the implementation of an exogenous variable, outperformed the standard models almost in every time series but only with a slight difference. Only in terms of the MAPE evaluating measure the standard models came up with more accurate results. In conclusion, the models with an exogenous variable provided better results in comparison to the standard models but only with a very slight difference in terms of RMSE and MAE but the standard models outperform the models with an exogenous variable in terms of MAPE and as a result the fact that the more complex models with an exogenous variable did not lead to more accurate results.

After providing an answer to the main research question, the sub-research questions are going to be answered.

i) Which variables affect the dry bulk freight rates?

The answer for this question can be found in the Literature review. As there is a model created by the demand and supply for the dry bulk shipping industry, any change in the demand or supply can alter the freight rates. Therefore, any variable which could alter the demand and supply of the tonnage and cargo for the market as well as the world economy and the business cycles which go hand-to-hand with the shipping cycles and any random shocks of the economy created by unexpected events which can not be forecasted or quantified, can alter the freight rates of the dry bulk shipping industry.

ii) Does the complexity of the multivariate models lead to more accurate results?

According to the results, the multivariate models produced the best results in 3 out of 4 time series in terms of MAPE, MAE and RMSE except of the Panamax Spot rates, where the GARCH model outperformed the other models in every evaluating measure. To compare the results more accurately, the average of every evaluating measure is going to be calculated for the sum of every univariate model and is going to be compared with the average price of every evaluating measure for the sum of the multivariate models for every time series

Comparison of Average price of univariate and Average price of multivariate evaluating measure variables (Table 42)

Capesize Spot Rates	RMSE	MAE	MAPE
Avg Univariate Price	2.837	2.073	11.82%
Avg Multivariate Price	2.570	1,995	11.67%

Capesize TC Rates	RMSE	MAE	MAPE
Avg Univariate Price	2662.27	2018.52	12.83%
Avg Multivariate Price	1777.24	1439.72	10.06%

Panamax Spot Rates	RMSE	MAE	MAPE
Avg Univariate Price	1.131	0.855	9.19%
Avg Multivariate Price	1.239	0.954	8.88%

Panamax TC Rates	RMSE	MAE	MAPE
Avg Univariate Price	1177.65	820.72	6.92%
Avg Multivariate Price	830.25	661.86	6.33%

As it can be observed from the tables above, the average price of the multivariate models slightly outperforms the average price of the univariate models in almost every evaluating measure except in the case of the Panamax Spot rates where the E-Garch model presented very accurate results and this fact contributed to the average number. By observing these results it can be mentioned that the multivariate results presented more accurate results in comparison to the univariate models in almost every evaluation measure but these results are not much better from the results of the univariate models and thus, the conclusion that that the multivariate models did not adequately outperformed the univariate models can be drawn.

iii) Which forecasting method is the most appropriate for each time series variables?

Different results have been observed for each time series, even with different results in the evaluating measures in some of them.

As far as the Capesize Spot Series concerned, the VECM-X models provided the best results in terms of RMSE, MAE and Theil's U evaluating measures, while the best MAPE result was provided by the simple multivariate VAR model

As far as the Time Charter Rates time series concerned, the results were mixed with the best RMSE and Theil's U results provided by VAR-X, the best MAE by VAR and the best MAPE results by VECM.

Overall, for the Capesize market, the VAR-X and VECM-X models provided the best results as a sum of all the results. Many different results have been observed for the time series with a relatively high score of MAPE which started from 9% up to 12% with the exception of a high 20% which shows the volatility in freight rates and the fact that the models have not been in a position to capture the errors adequately enough. Also this shows that the data had an increased variability and they have a stochastic trend and even other endogenous and exogenous variables cannot achieve a significant increase in the errors.

For the Panamax Spot Rates Time Series, the Exponential Garch model provided the best results in terms of RMSE MAE and MAPE, while the results for Theil's U could not be obtained for this model. This fact shows that the Garch model could capture the volatility of the time series effectively in comparison to the other time series.

For the Panamax Time Charter Rates Time Series, every forecasting method provided interesting MAPE results with 7% or less. The standard VAR model provided the best results in every evaluating measure.

The Panamax market had less variability than the Capesize market and this can be proved by the lower MAPE scores.

Overall, with the exception of the Panamax Spot Rates, the multivariate models provided better results than the univariate models in every evaluating measure and for the Capesize market, the multivariate models with the implementation of an exogenous variable provided the best results in many evaluating measures for the time and spot rates time series.

5.2 Limitations of the study

The first limitation that is being presented for this study is the fact that the results for the evaluating measures might be different for another sample of freight rates of a different date or freight rates of different freight routes based on the different variability of the freight rates.

The second limitation is the fact that a priority has been given to the legitimacy of every forecasting model which means that the models are able to come through any evaluation test. This means that less efficient AIC numbers have been taken into consideration as well as less lags than the optimal ones for the sake of the validity of models. Better results might have been occurred with the use of bigger lags.

The third limitation of the model is that the GARCH forecasting model could not be generated for the Capesize Spot Freight Rates due to the lack of validity of the model.

The last limitation is the fact that no neural network has been generated for further comparison of the existing forecasting models.

5.3 Recommendations

As an extension of this study, Neural Networks and Random Forest forecasting models could be created and compared with the existing models.

Also, further studies could concentrate on the analysis of other possible endogenous variables which could be used to the multivariate models which could lead to more accurate forecasting results. The same thing can be applied for the exogenous variables.

Last but not least, further research could be conducted to the combination of existing forecasting techniques such as using a multivariate forecasting model to the data and combining a univariate model or Random Forest to the residuals of the Time series.

Bibliography

- Institute of Chartered Shipbrokers. (2016) "*Dry Cargo Chartering*".

- Stopford, M. (2009) "*Maritime Economics*".

- Gorton L. ,Ihre R. , Plomaritou E. and Papadopoulos, A. (2018) "*Shipbroking and chartering practice*."

- Diakodimitris, E., (2020). "*Evaluation of Forecasting Approaches for Crude Oil Tanker Freight Rates*." Available at: https://thesis.eur.nl/pub/50898

- Tsioumas, V., (2016) "*Quantitative analysis of the dry bulk freight market, including forecasting and decision making*." Available at: https://dione.lib.unipi.gr/xmlui/bitstream/handle/unipi/9411/Tsioumas_Vangelis.pdf?sequence
- Culinnane, (1992) "*A short-term adaptive forecasting model for Biffex speculation: A Box-Jenkins approach*." Available at: https://www.tandfonline.com/doi/abs/10.1080/03088839200000018
- Culinnane, (1999). "*A Comparison of Models for Forecasting the Baltic Freight Index: Box-Jenkins Revisite*d." Available at: https://link.springer.com/article/10.1057/ijme.1999.10

- Kavussanos, M. G. (1996). "*Price Risk Modelling of Different Size Vessels in the Tanker Industry Using Autoregressive Conditional Heteroskedasticity (Arch) Model*s." Available at: https://trid.trb.org/view/463604
- Clarksons Shipping Intelligence network "http://sin.clarcksons.net"

- Nikolaos D. Geomelos & Evangelos Xideas (2017). "*Econometric estimation of second-hand shipping markets using panel data analysis*" Available at: https://www.econstor.eu/handle/10419/169191
- A.W. Veenstra, P.H. Franses, (1997) "*A co-integration approach to forecasting freight rates in the dry bulk shipping sector*" Available at: https://www.sciencedirect.com/science/article/abs/pii/S0965856497000025
- A.W. Veenstra (1999) "*Quantitative analysis of shipping markets*"

- M. G. Kavussanos, (2003). "*Price Discovery, Causality and Forecasting in the Freight Futures Market*." Available at: https://link.springer.com/article/10.1023/B:REDR.0000004824.99648.73
- M.G. Kavussanos, I.D. Visvikis, (2004), "*Market interactions in returns and volatilities between spot and forward shipping freight markets*" Available at: https://www.sciencedirect.com/science/article/abs/pii/S0378426603002279
- Papailias, F., and Thomakos, D.D. (2013), "*The Baltic Dry Index: Cyclicalities, Forecasting and Hedging Strategies*." Available at: https://link.springer.com/article/10.1007/s00181-016-1081-9
- H.L. Wong, (2014), "*BDI Forecasting Based on Fuzzy Set Theory Grey System and ARIMA*"

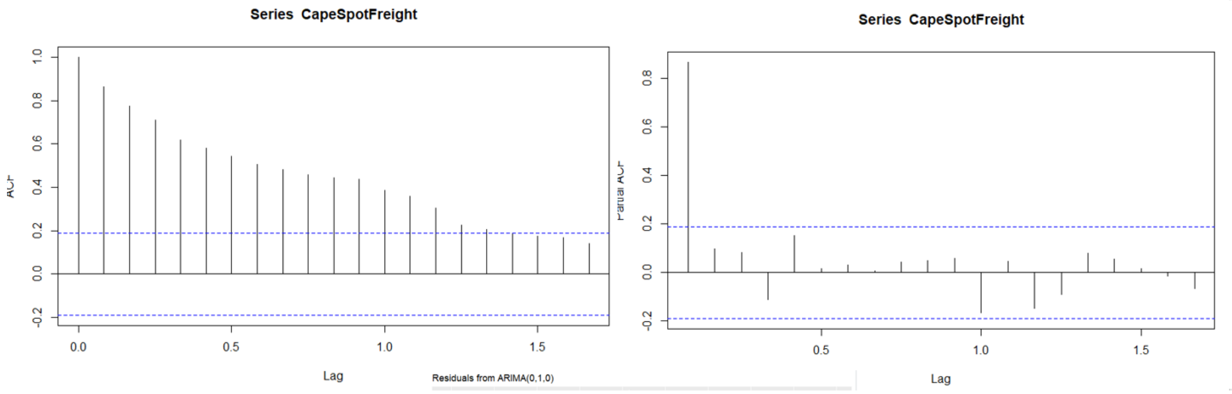
Available at: https://link.springer.com/chapter/10.1007/978-3-319-07467-2_15

- J. Zhang, Q. Zeng, X. Zhao, (2014), "*Forecasting spot freight rates based on forward freight agreement and time charter contract*" Available at: https://www.tandfonline.com/doi/abs/10.1080/00036846.2014.937038?journalCode=raec20
- Taib, C. and Mohtar, Z., (2018). "*Forecasting spot freight rates using vector error correction model in the Dry Bulk Market*." Available at: https://einspem.upm.edu.my/journal/fullpaper/vol12no2/6.%20Che%20mohd%20Taib.pdf
- Bessler, D. and Jonnala, S., (2002) "*A GARCH Approach to Modelling Ocean Grain Freight Rates*." Available at: https://link.springer.com/article/10.1057/palgrave.ijme.9100039
- Chatfield, C. (2001). "*Time-Series Forecasting"*
- Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994) "*Time Series Analysis; Forecasting and Control. 3rd Edition*"
- Duke, (2010). "*Stationarity and Differencing*."
- Maddala and Kim, (1998). "*Unit Root, Cointegration and Structural Change*."
- Caner, M., & Kilian, L. (2001). "*Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate*." Available at: https://www.sciencedirect.com/science/article/abs/pii/S0261560601000110
- Kuo, B. S., & Tsong, C. C. (2005). "*Bootstrap inference for stationarity. Helsinki Center for Economic Research, Discussion Paper*" Available at: https://researchportal.helsinki.fi/en/publications/bootstrap-inference-for-stationary
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). "*Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?"* Available at: https://www.sciencedirect.com/science/article/abs/pii/030440769290104Y
- Palachy, S. (2019). "*Detecting stationarity in time series data*." Available at: https://towardsdatascience.com/detecting-stationarity-in-time-series-data-d29e0a21e638
- Perron, P. (1989). "*Testing for a Unit Root in a Time Series with a Changing Mean*". Available at: https://www.jstor.org/stable/1391977?seq=1
- Y. Sakamoto, (1986). "*Akaike Information Criterion Statistics*"
- Brockwell, Davis. (1991) "*Time Series: Theory and Methods*"
- S Makridakis. (1997). "*Forecasting: Methods and Applications*"
- R. Hyndman. (2001). "*Forecasting with Exponential Smoothing*"
- R. Hyndman. (2010). "*The ARIMAX model muddle*" Available at: https://robjhyndman.com/hyndsight/arimax/
- D. Nelson , C. Cao. (1992). "*Inequality Constraints in the Univariate GARCH Model*" Available at: https://www.jstor.org/stable/1391681?seq=1

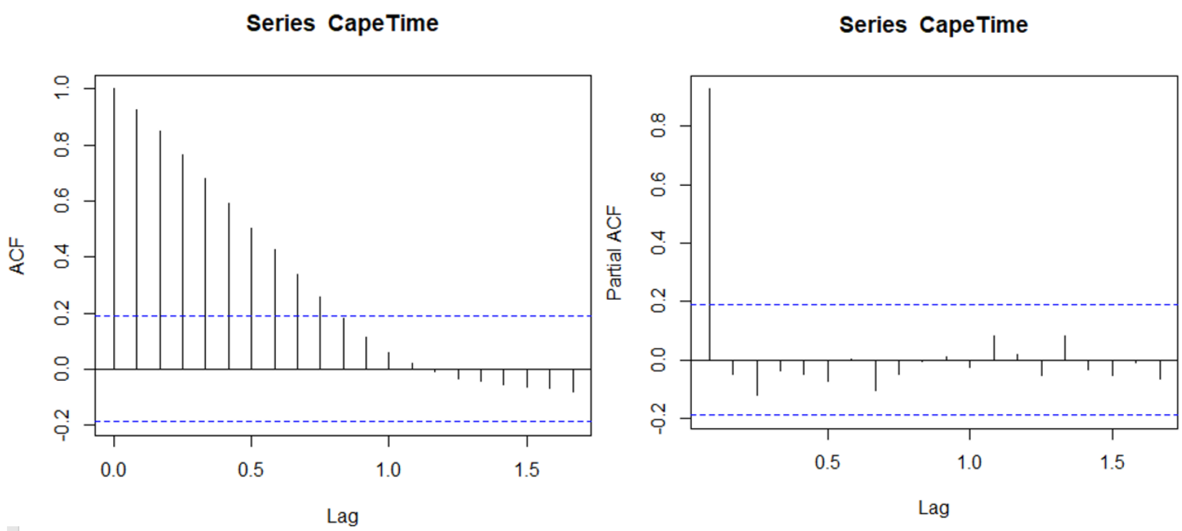
- W. H. Greene. (2003). “*Econometric Analysis 5th Edition*”

- Gerbing. (2016) “*Time Series Components*”

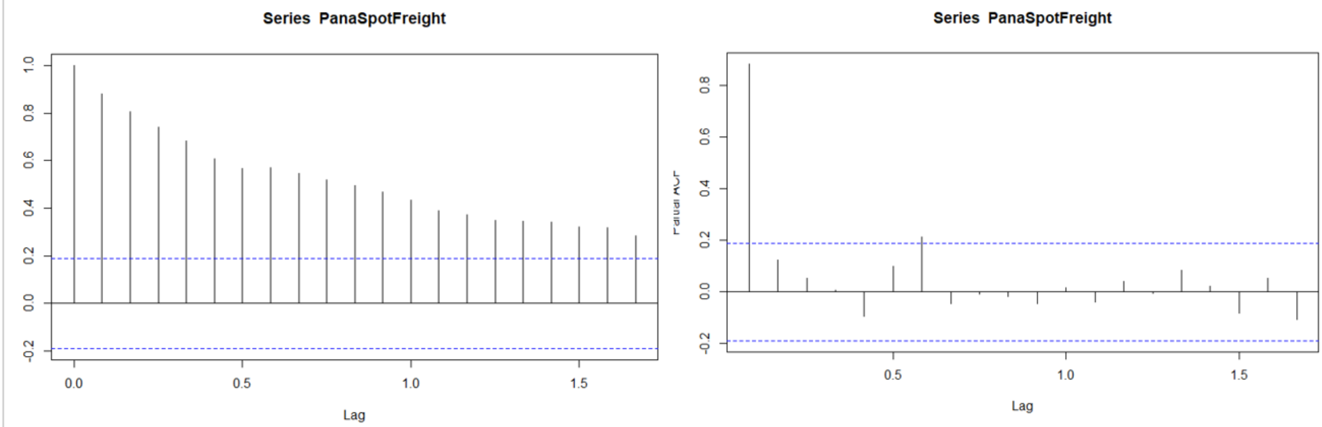
APPENDIX



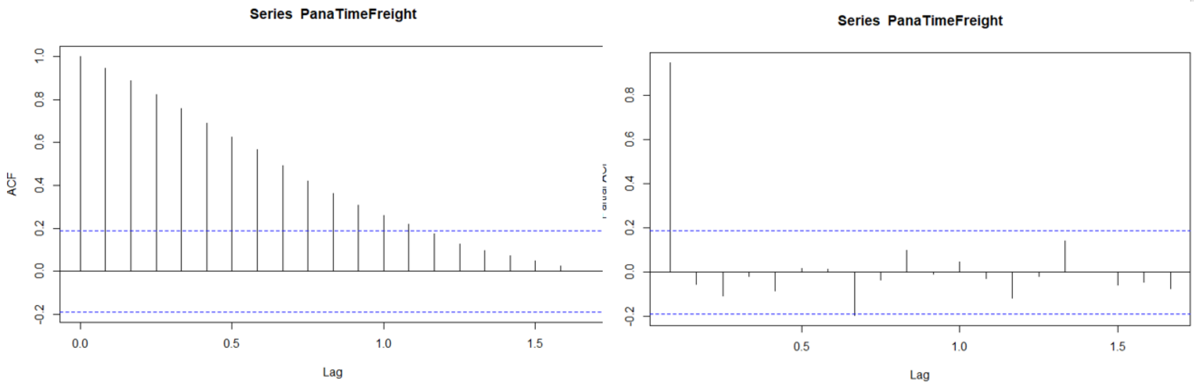
(ACF-PACF for Capesize Spot Rates)



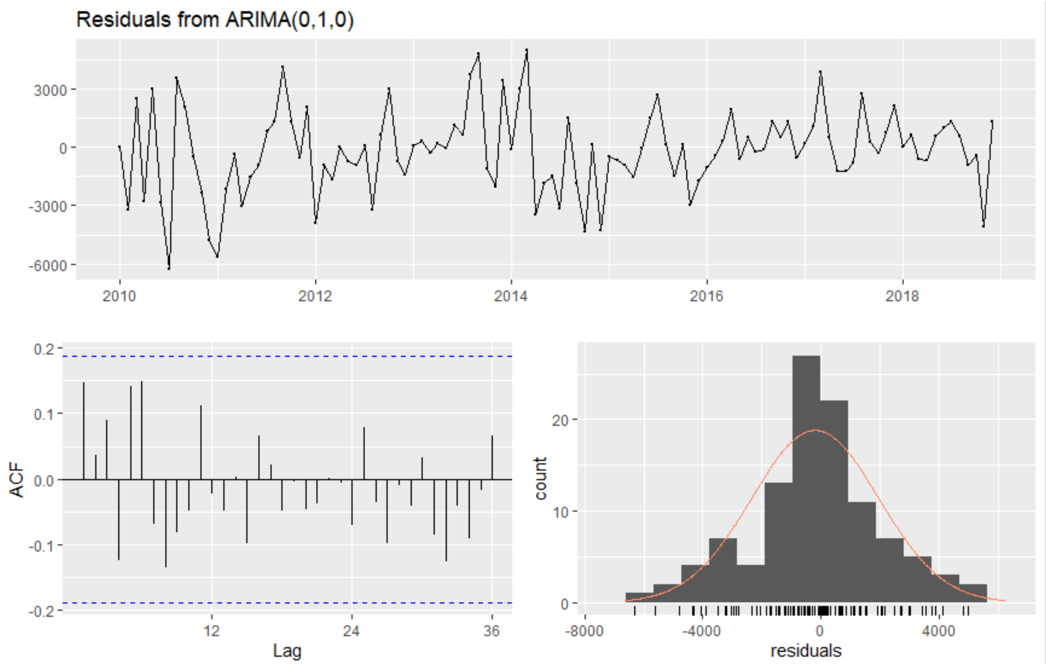
(ACF-PACF for Capesize TC Rates)



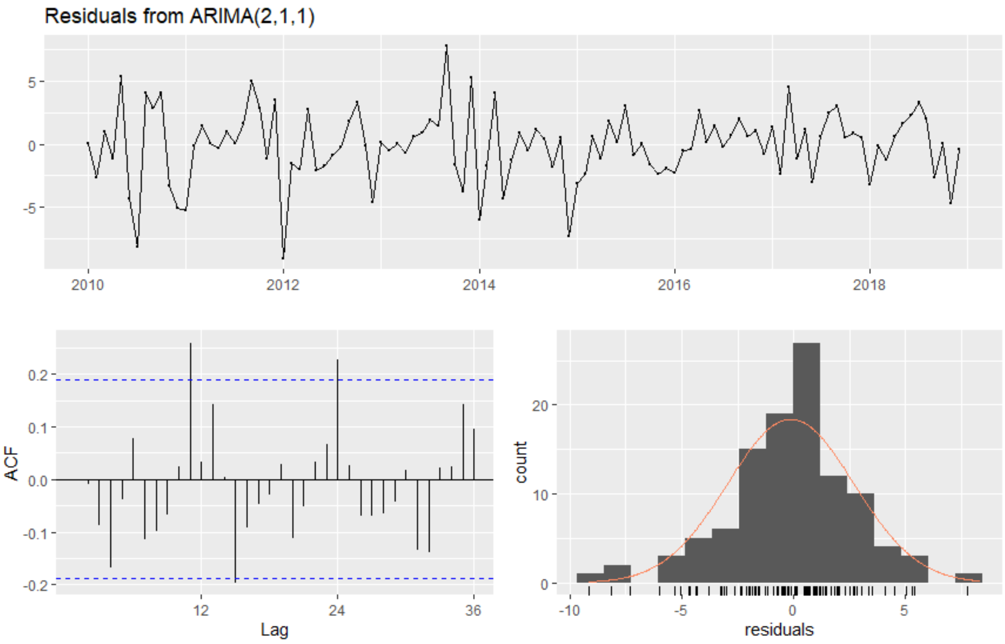
(ACF-PACF for Panamax Spot Rates)



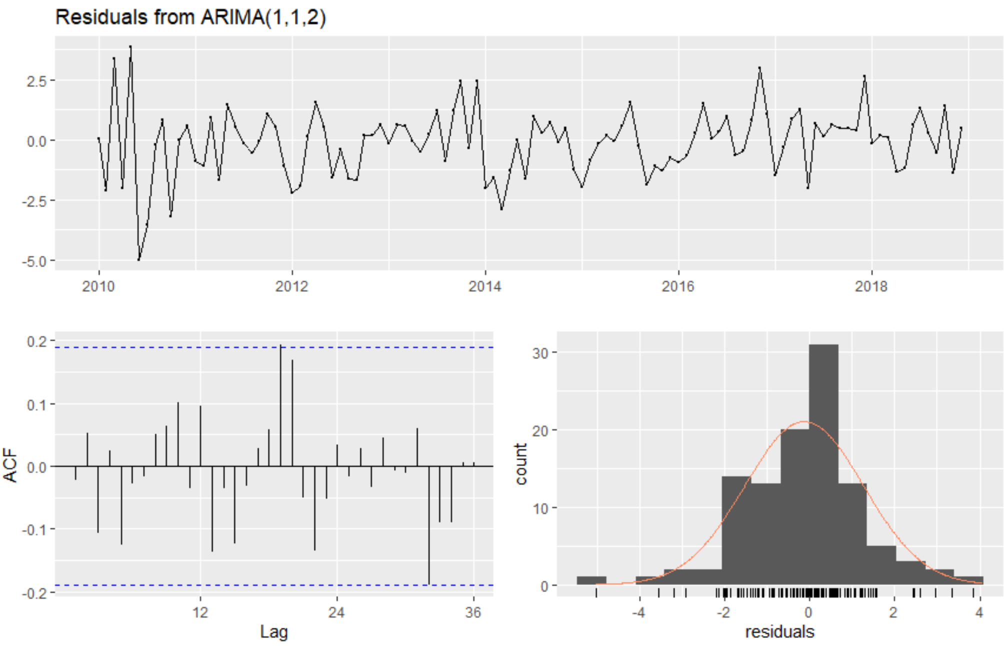
(ACF-PACF for Panamax TC Rates)



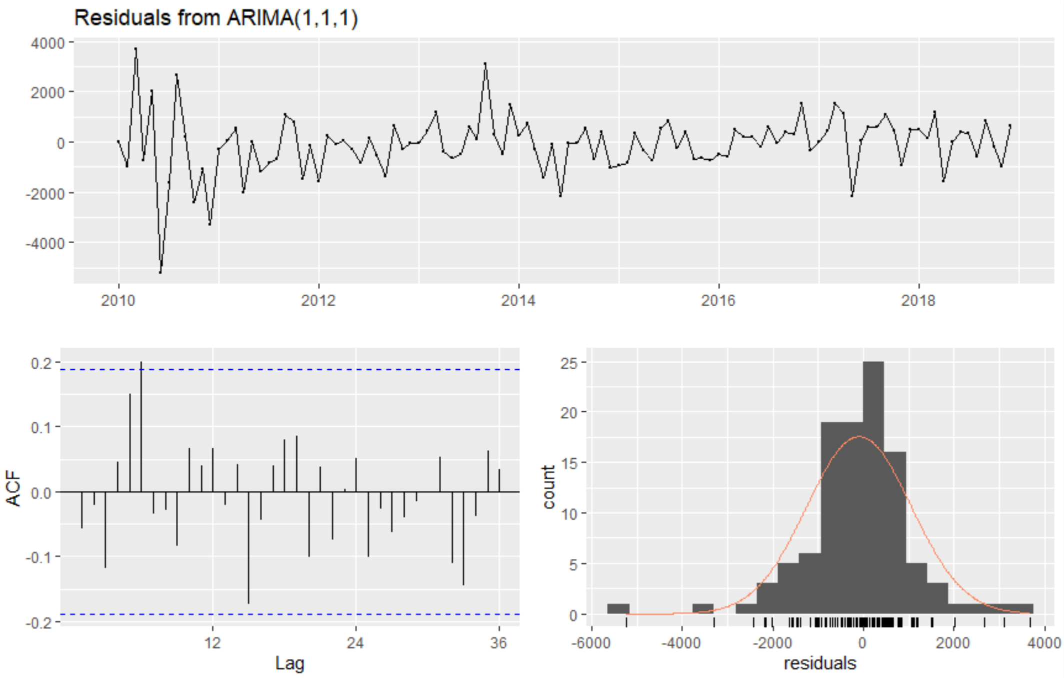
(Residuals of Capesize Spot Rates ARIMA model)



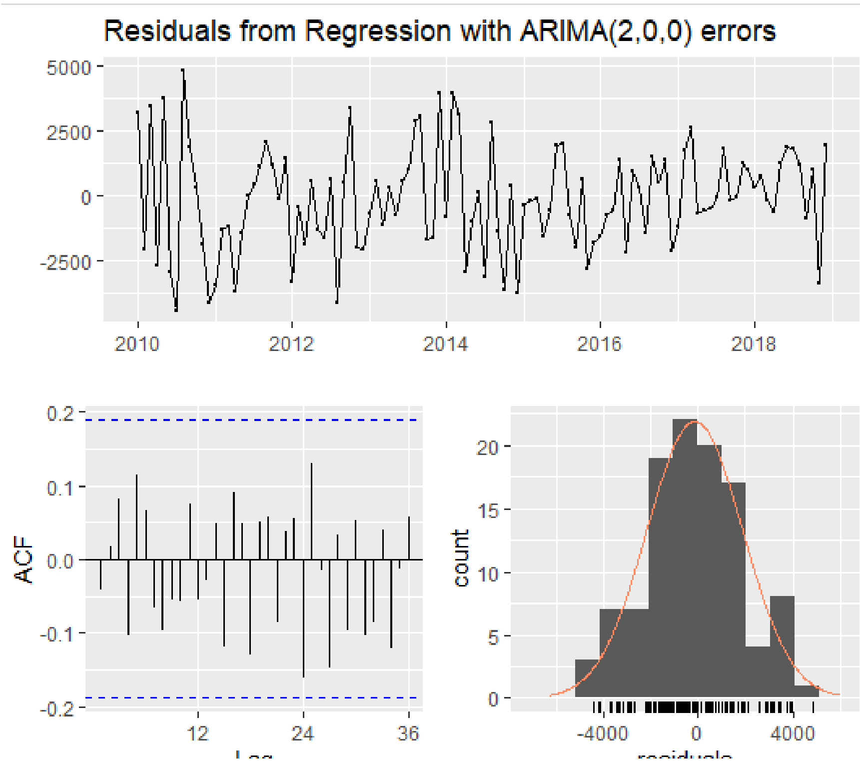
(Residuals of Capesize TC Rates ARIMA model)



(Residuals of Panamax Spot Rates ARIMA model)

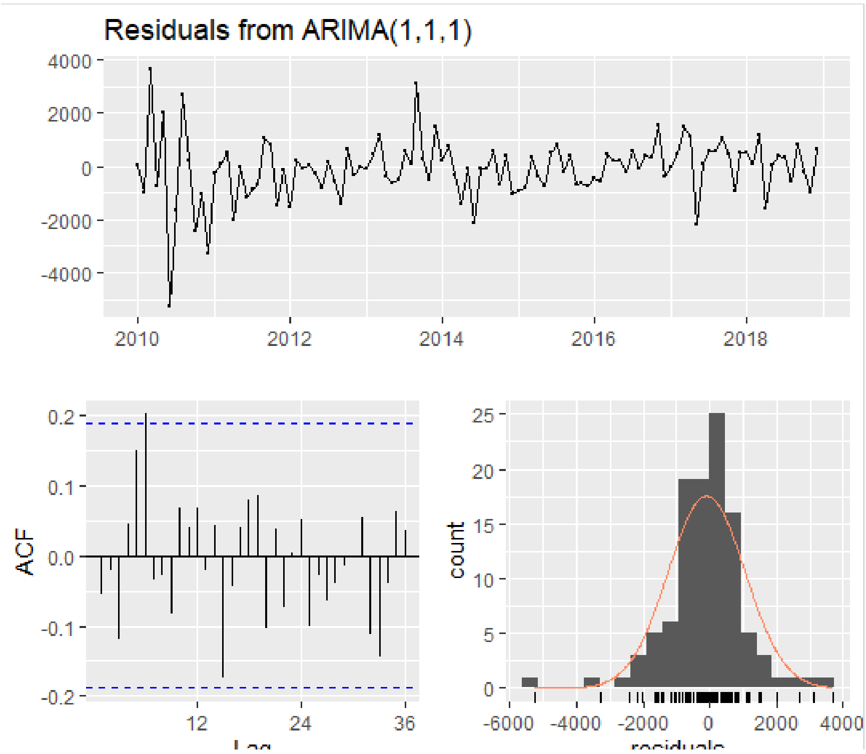


(Residuals of Panamax TC Rates ARIMA model)

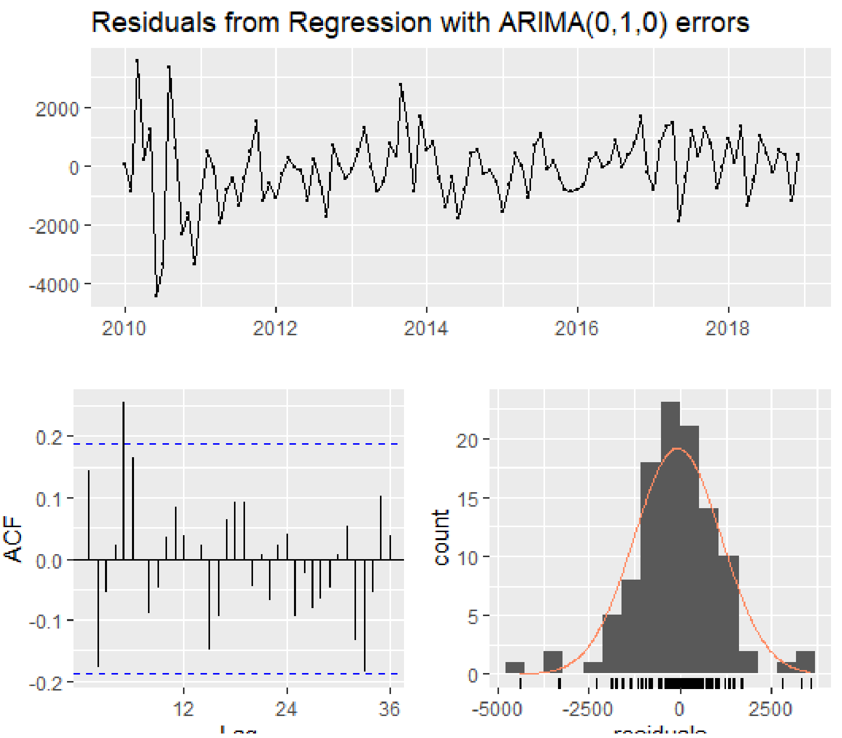


(Residuals of Capesize Spot Rates ARIMA-X model)

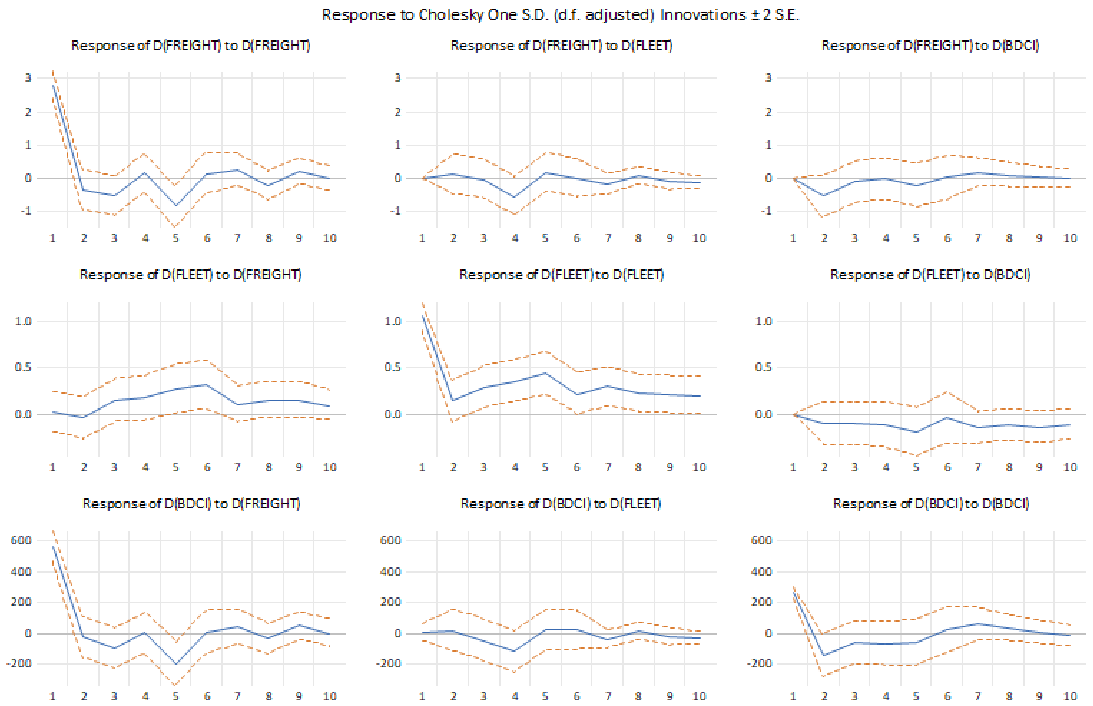
(Residuals of Capesize TC Rates ARIMA-X model)



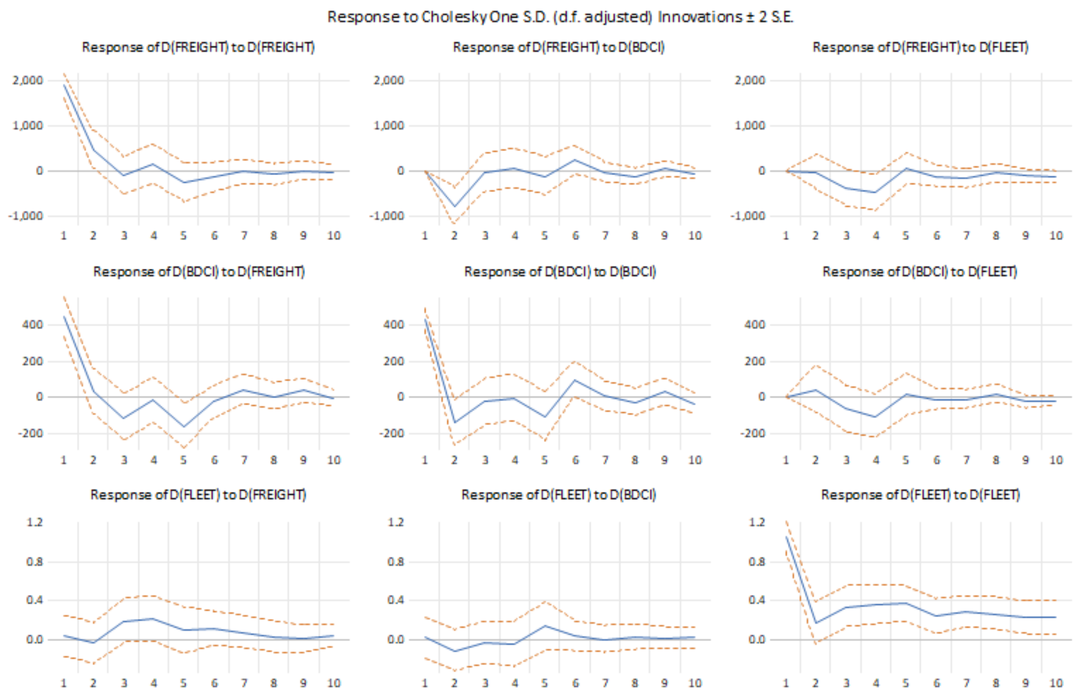
(Residuals of Panamax Spot Rates ARIMA-X model)



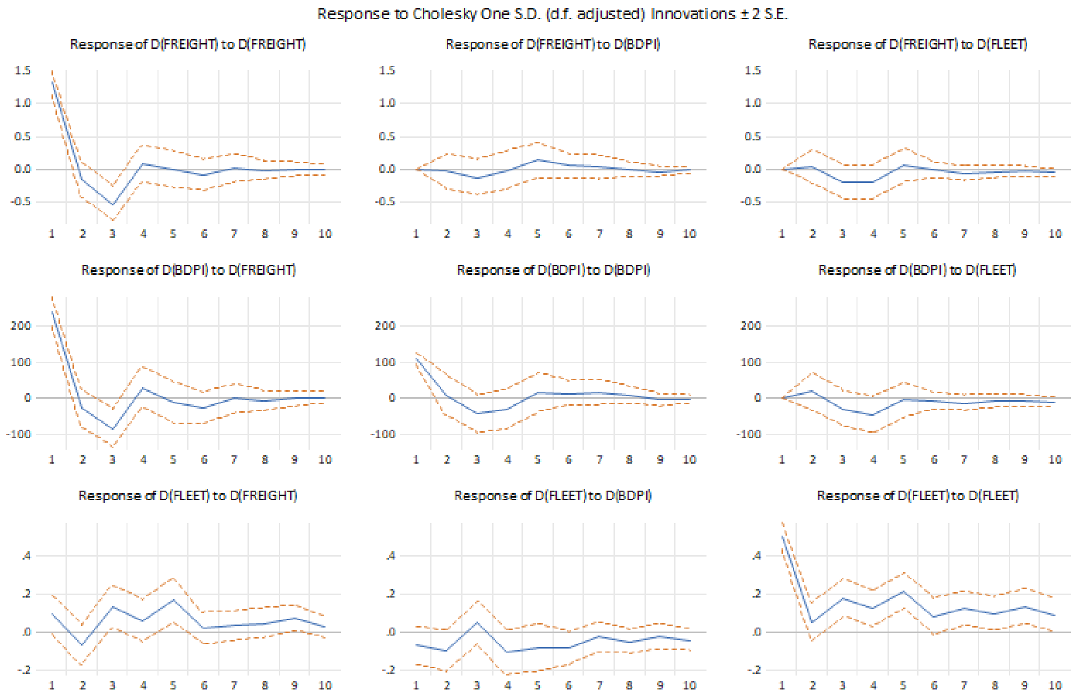
(Residuals of Panamax TC Rates ARIMA-X model)



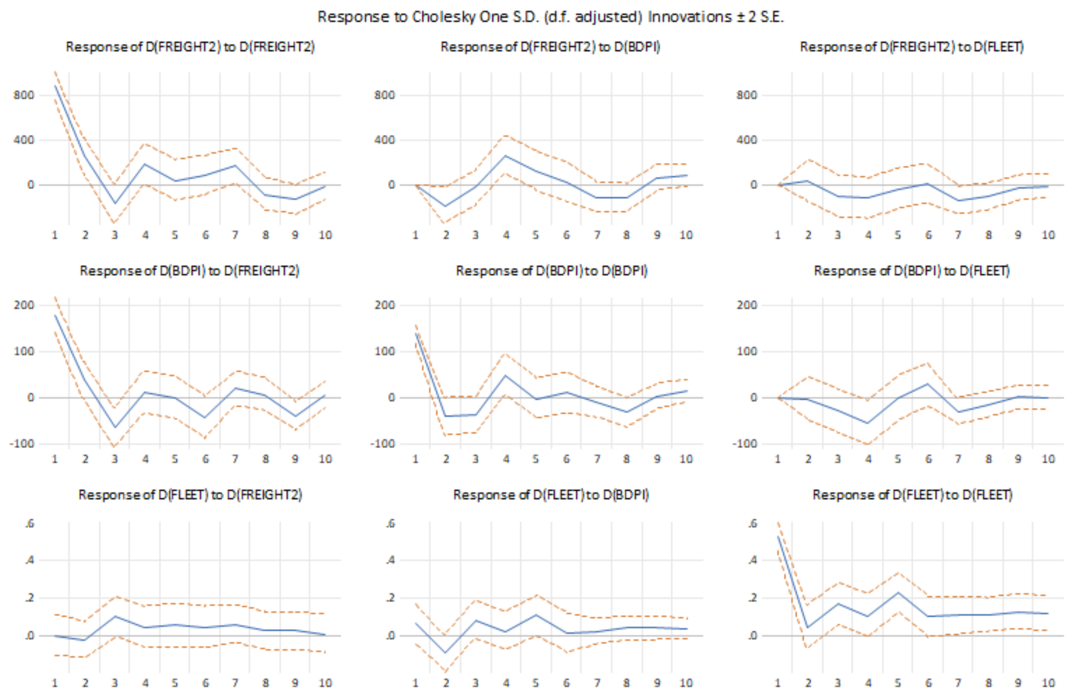
(Impulse Response of Capesize Spot Rates Var Model)



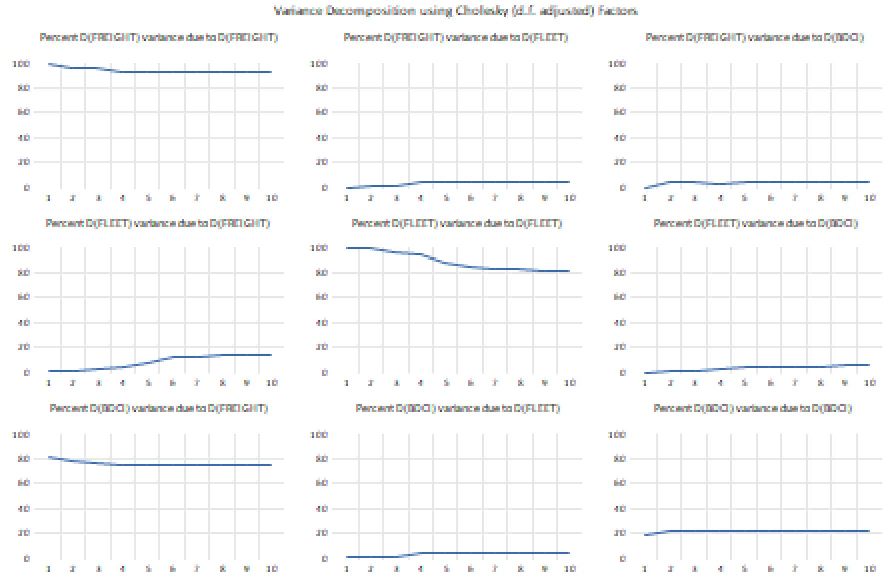
(Impulse Response of Capesize TC Rates Var Model)



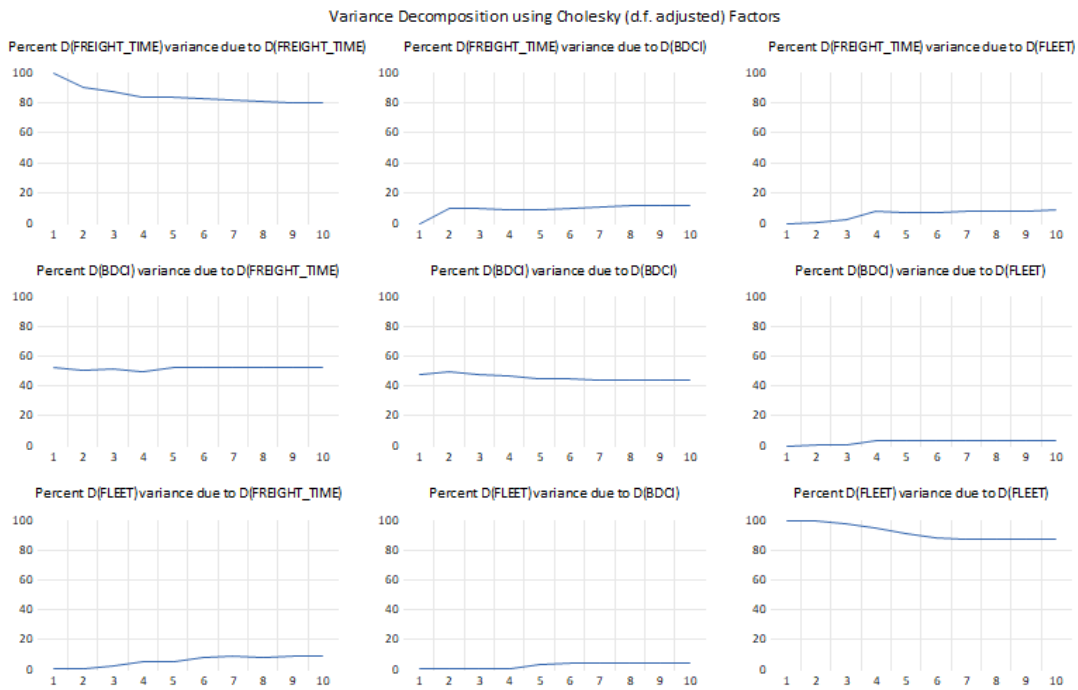
(Impulse Response of Panamax Spot Rates Var Model)



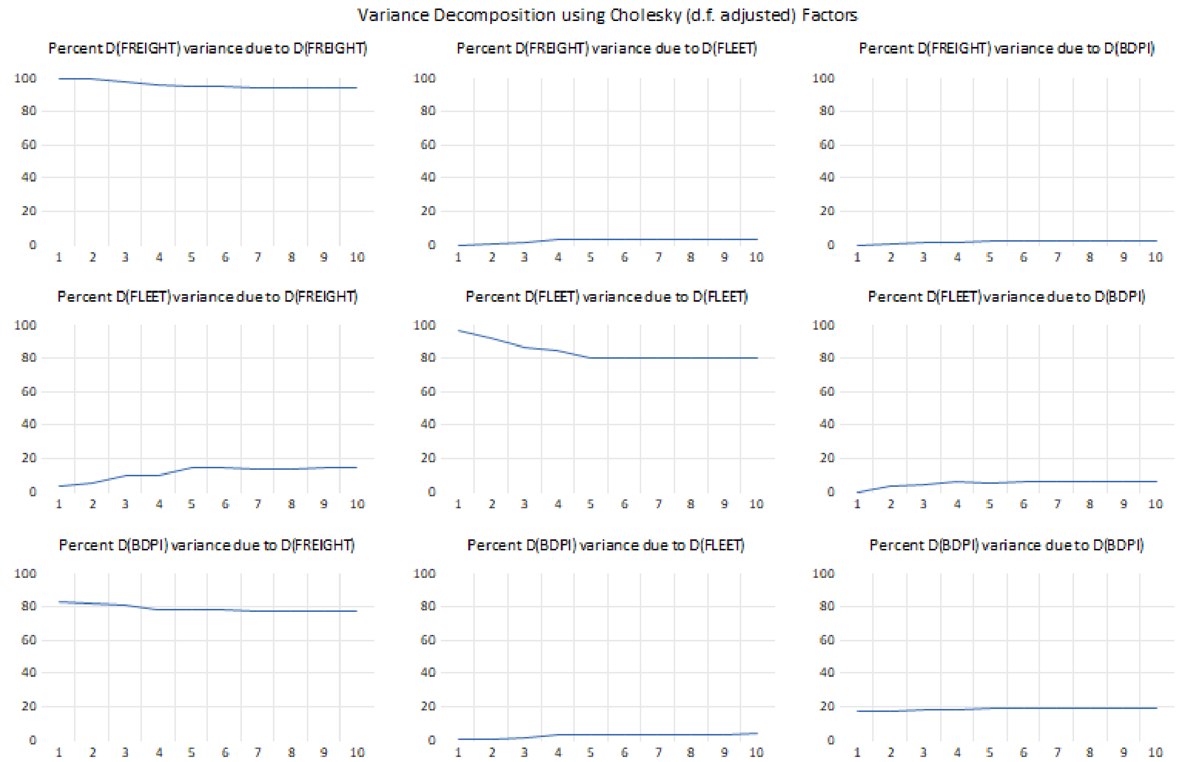
(Impulse Response of Panamax TC Rates Var Model)



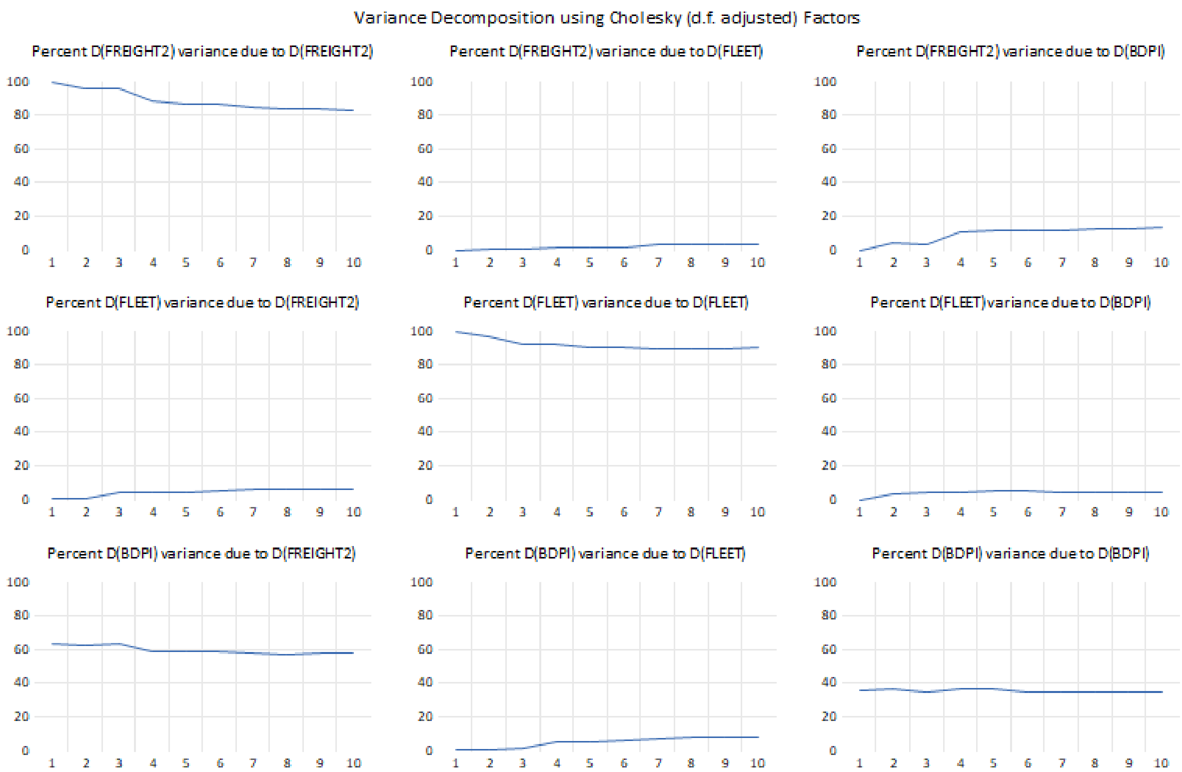
(Variance Decomposition of Capesize Spot Rates VAR model)



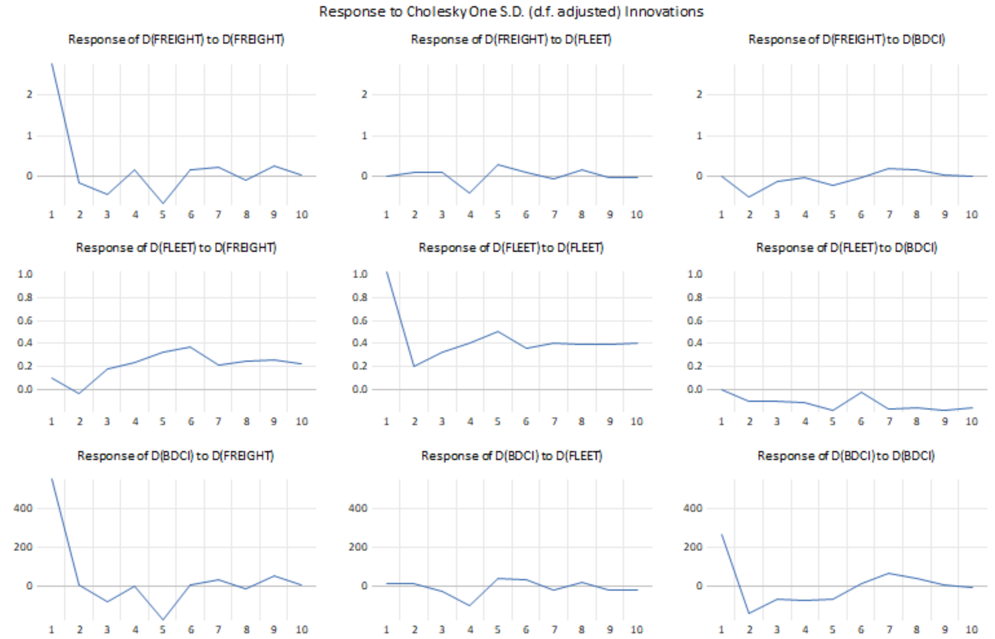
(Variance Decomposition of Capesize TC Rates Var Model)



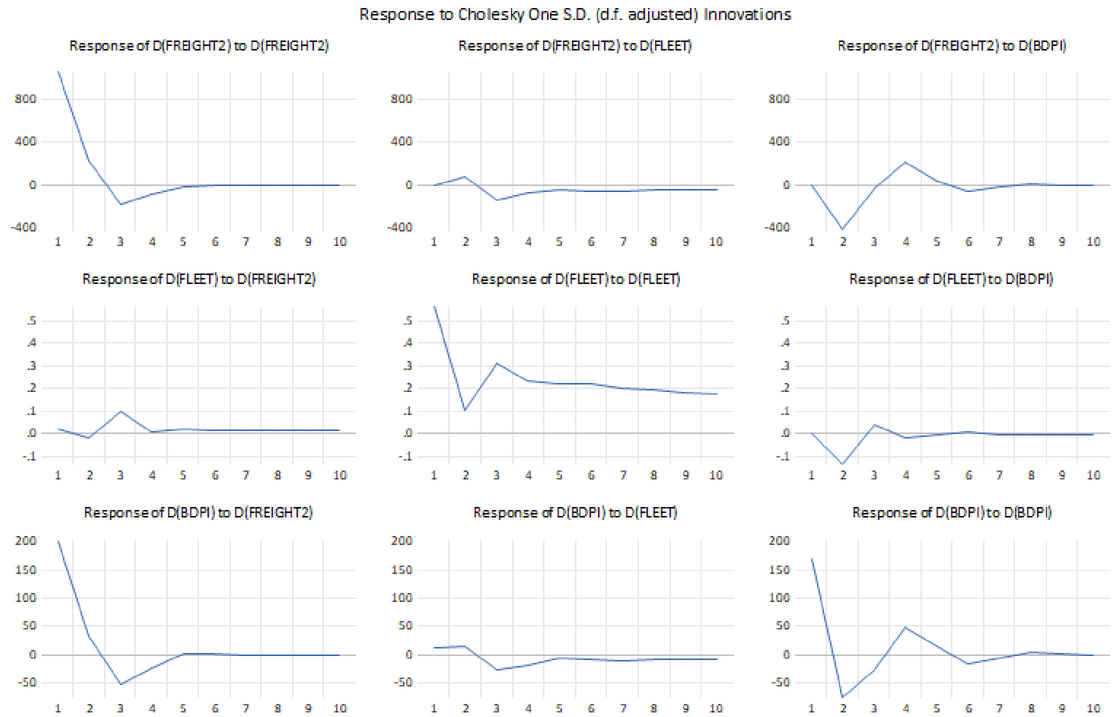
(Variance Decomposition of Panamax Spot Rates Var Model)



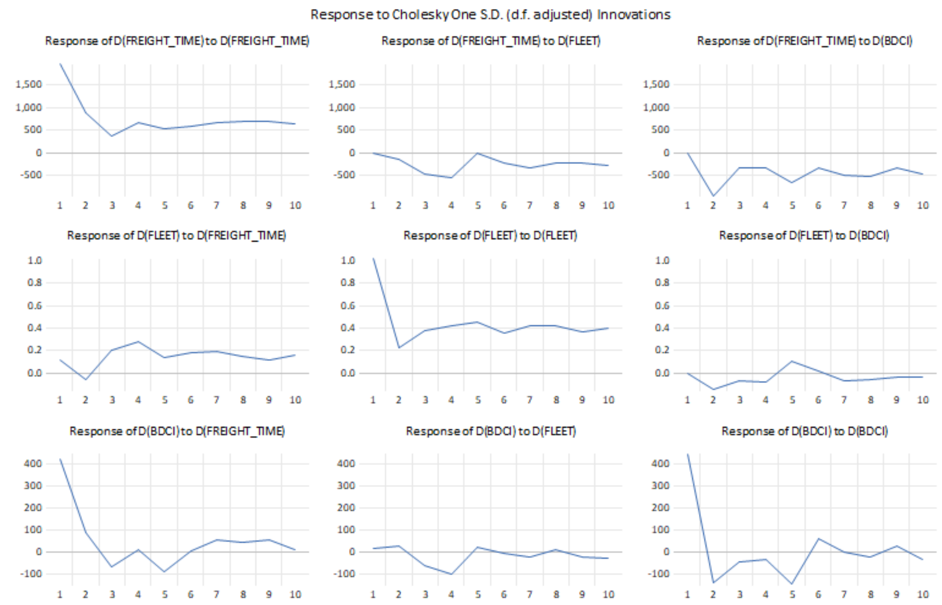
(Variance Decomposition of Panamax TC Rates Var Model)



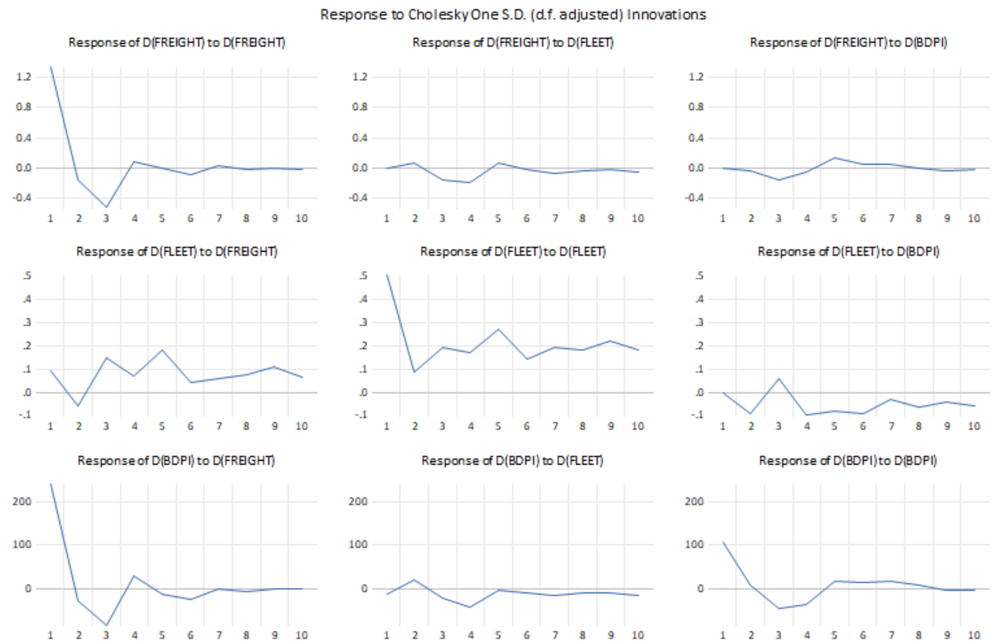
(Impulse Response of Capesize Spot Rates Vecm Model)



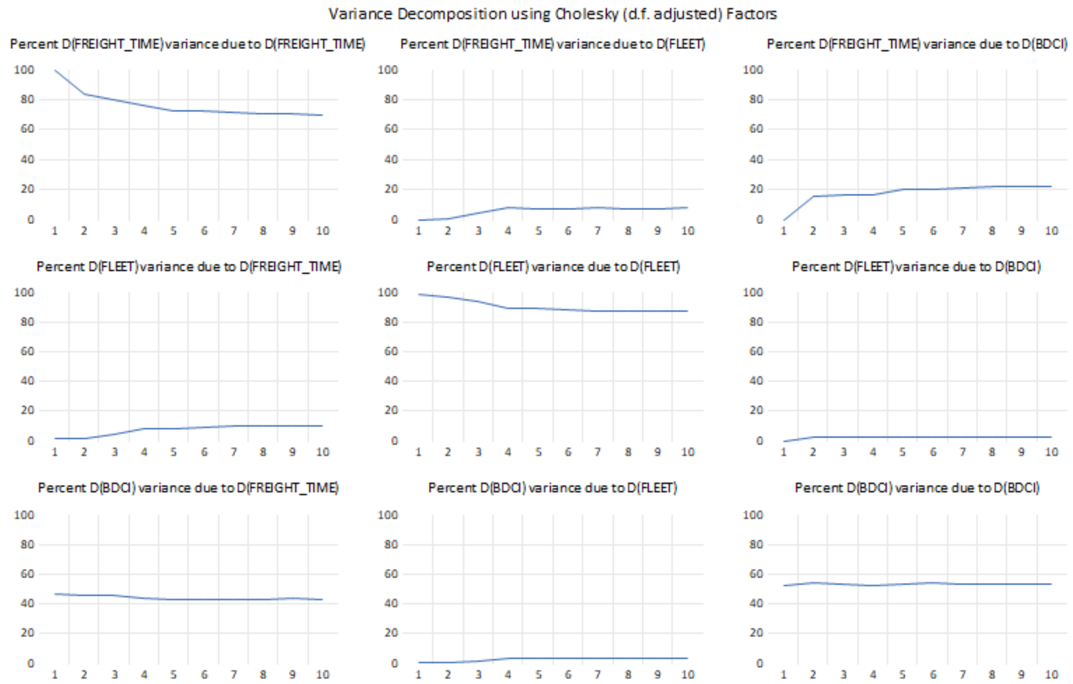
(Impulse Response of Capesize TC Rates Vecm Model)



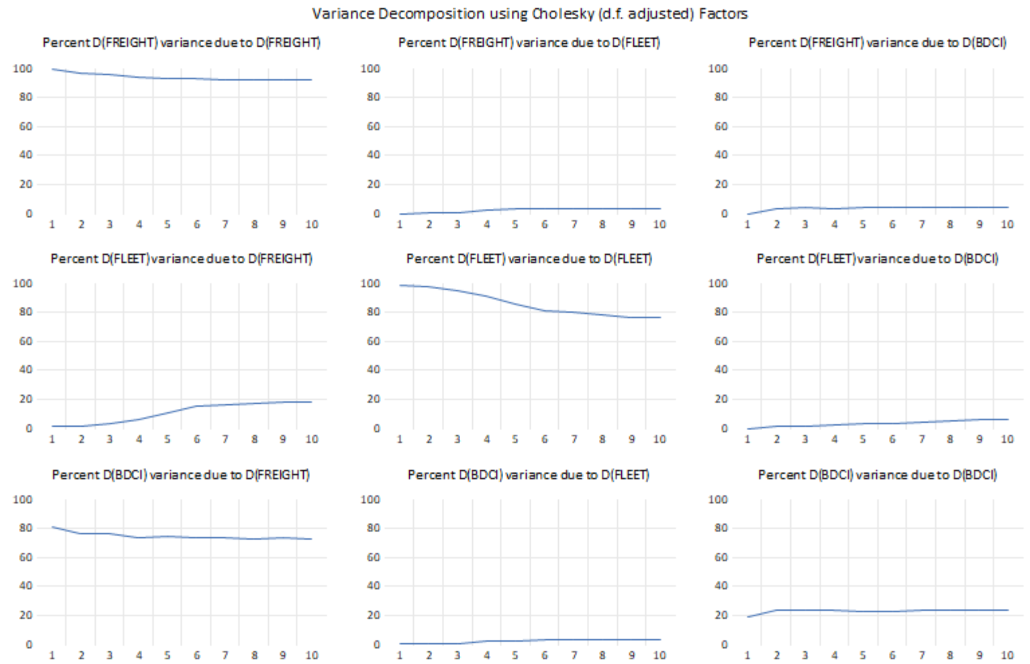
(Impulse Response of Panamax Spot Rates Vemc Model)



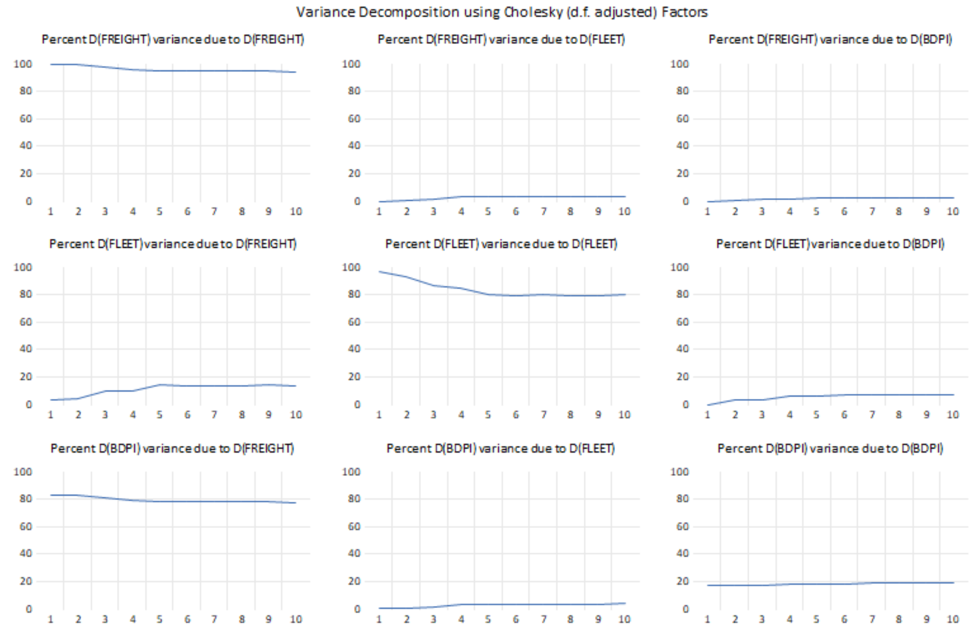
(Impulse Response of Panamax TC Rates Vecm Model)



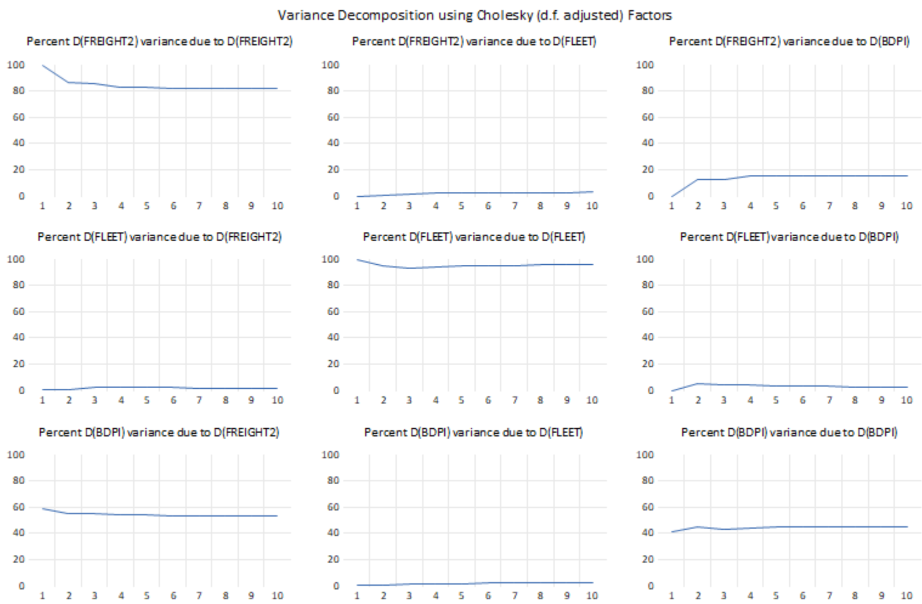
(Variance Decomposition of Capesize Spot Rates Vecm Model)



(Variance Decomposition of Capesize TC Rates Vecm Model)



(Variance Decomposition of Panamax Spot Rates Vecm Model)



(Variance Decomposition of Panamax TC Rates Vecm Model)