Norwegian School of Economics
Bergen, Spring 2019

The Winds of Winter – How Wind Forecasts Predict Next-Day Freight Rates in Dry Bulk Shipping

Aksel Stokseth and Torstein Eika Amsrud

Supervisor: Maximilian Rohrer

Master thesis, Economics and Business Administration

Major: Business Analytics and Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Acknowledgements

We would first and foremost like to thank our supervisor Maximilian Rohrer for valuable input throughout the writing of this thesis. We would also like to thank the Norwegian Coastal Authorities for access to their satellite AIS data, and the Baltic Exchange Ltd for access to FFA rates. Finally, we would like to extend our thanks to the NHH professors we have bothered with questions on shipping markets and time series regressions.

Abstract

This thesis examines the short-run capesize dry bulk shipping market, where shipping theory states that freight rates are driven by supply, as demand is inelastic. Transporting cargo faster is an important way to increase supply if demand increases in the short run. Market participants monitor the speed of the available fleet and form expectations of future supply, and these expectations are key drivers of freight rates in the very short run. Any factor affecting the speed of vessels will therefore have an impact on freight rates. We combine local wind forecasts with positional vessel data, and show how wind forces push vessels forwards and inflate their speed. Based on these findings, we propose a novel positive wind force measure. We then show our positive wind force measure predicts next-day freight rates. For a one standard deviation increase in the positive wind force at average level, we find the next-day spot rate increases by 0.3% and the next-day forward freight agreement (FFA) rate increases by 0.6%. Our measure helps shipowners and charterers make better pricing decisions in the spot market, and hedge their exposure better through the FFA market. Other traders can also profit in the FFA market, as we show changes in our measure can predict changes in the FFA rate greater than transaction costs.

Keywords –

AIS, Capesize, Dry bulk, Forward freight agreements, Freight rates, Wind forecasts

Contents

List of Figures

List of Tables

1 Introduction

In this thesis we study the predictive power of wind affecting vessels' speed on next-day freight rates in the capesize dry bulk shipping market. The dry bulk shipping market is responsible for the transportation of essential industry materials such as iron and coal, as well as grain across the world. Four main vessel classes, categorized by size, operate in this market: handysize, handymax, panamax and capesize, where capesize is the largest vessel class. The capesize market is the most volatile of all shipping markets and is highly sensitive to changes in the key drivers of supply and demand (Stopford, 2009).

In the short-run shipping market, freight rates are determined by the demand for transportation of cargo and the supply of vessels available to transport cargo (Stopford, 2009). According to Stopford (2009), demand is inelastic with respect to freight rates in the short run, as there are no competing ways to transport cargo. It is only possible to increase the supply by either increasing the number of vessels, or by *transporting cargo faster*.

An important part of short-run supply is the speed of the fleet (Stopford, 2009). As demand for freight increases, the vessels increase their speed in order to satisfy the demand by transporting more cargo in the same amount of time. The speed of the vessels cannot be increased indefinitely and will at some point approach a maximum, where even small increases in demand will result in major increases in rates. The market participants know this and monitor the speed, forming expectations of future supply. These market expectations affect freight rates in the very short run (Stopford, 2009).

All capesize vessels continuously report their speed, course and position via the Automatic Identification System (AIS), as they sail around the globe. This vessel information is collected worldwide by coastal authorities and other commercial actors such as MarineTraffic and Clarksons SeaNet using satellite and land-based receivers. The shipping market can therefore monitor the movements of the vessels in real-time using AIS.

The reported speed through AIS does not account for wind pushing vessels forwards or backwards as they sail. Consequently, a vessel's reported speed through AIS is a biased measure of its actual chosen speed through water. Changes in the observed speed can

therefore be a result of changes in wind forces affecting the vessels, and not an active speed decision. This skews the expectations of the market participants when wind is causing observed speed changes and not active speed decisions. When vessels are sailing in increasing tail-wind, market participants relying on the biased AIS speed would erroneously attribute the change in speed to the market heating up, when in reality this should have been attributed to the increase in tail-wind.

We estimate this wind effect through a novel *positive wind force* measure. The measure is calculated by estimating the forward-pushing force of the apparent wind on a vessel, and adjusting for the vessel's size. The resulting measure is an estimate of wind forces pushing vessels forwards per dead weight ton capacity.

We use our measure of positive wind force to estimate a predictive regression on next-day spot freight rates, on the C5 iron ore route from Australia to China. We control for apparent wind speed on vessels, speed of vessels, fuel costs, one-month and forward freight agreement (FFA) rates. We find the C5 spot rate increases after an increase in positive wind force. For a one standard deviation increase in positive wind force on vessels sailing without cargo at average levels, there is an estimated increase in the C5 spot rate of 0.3%.

We also estimate a predictive regression of positive wind force affecting vessels on next-day one-month FFA rates, controlled for apparent wind speed on vessels, speed of vessels and fuel costs. With this model we find that the FFA rate increases after an increase in positive wind force. For a one standard deviation increase in positive wind force on vessels travelling without cargo at average levels, there is an estimated increase in the FFA rate of 0.6%.

We establish that the positive wind force measure is not just a proxy for apparent wind speed on vessels. We validate this claim by performing a regression of positive wind force and apparent wind speed against the speed of vessels. The results reveal that increases in speed can partially be explained by increases in positive wind force and vice versa.

We also establish Granger-causal relationships between positive wind force and both the C5 spot rate and the FFA rate. The results of the causality tests are strengthened by the lack of any other logical confounding time series. The same causal relationship cannot be established for the apparent wind speed.

The research most closely related to ours is conducted by Regli and Nomikos (2019) on the tanker market. The authors find significant explanatory effects of AIS-derived supply measures on next-period freight rates in weekly sampled data on a busy tanker route. Their main finding is a novel availability ratio, but Regli and Nomikos (2019) also measure supply through the speed of vessels. They find that the speed of vessels are significant in predicting next-week spot rates. Based on their findings we use vessel speeds as well as bunker fuel price as controls in our models.

To the best of our knowledge, similar research on the predictive power of positive wind force on freight rates has not been conducted.

2 Motivation

In this section we will provide an overview of the short-run shipping market. From this foundation we will then elaborate on our central economic motivations for investigating the predictive power of positive wind force on future spot and FFA rates.

2.1 Short-run and momentary market equilibrium

Freight is a service that cannot be stored, and the absence of warehouse buffers allows for greater short-term volatility in freight rates, as there are no reserves to be tapped. This also frees it from arbitrage conditions that bind spot and forward rates in other commodity markets together (Batchelor et al., 2007; Alizadeh and Nomikos, 2009). Importantly, the nonstorable nature of freight means that short-run supply is entirely determined by the available fleet at any given time. The available fleet can adjust its operational activity level by laying up vessels if rates are low, and increasing its speed if rates are high. The activity level in turn affects the amount of cargo that can be transported. Freight demand is relatively inelastic with respect to freight rates, shown in the steep demand curves seen in Figure 2.1 (Stopford, 2009). The demand for freight in the short-run is inelastic because there are no viable competing ways of transporting freight. Freight by sea is the cheapest alternative for transporting large volumes even when rates are high (Stopford, 2009).

The supply-demand relationship results in freight rates highly dependent on current fleet utilization levels. If the world fleet is operating at near its maximum short-run capacity, a small outward shift in the demand curve will result in a major jump in freight rate. Conversely, a large shift in the demand curve will have a small effect on freight rates if supply is far from its capacity.

In the very short-run, such as in day-to-day operations, momentary market equilibrium is established via auctions where market sentiment and the number of promptly available vessels play an important role. The supply curve is highly inflexible in this time frame, as seen in Figure 2.2. For instance, new vessels cannot be built nor laid up over night. Should the local demand for freight be greater than the number of local vessels available for orders, the rates will increase greatly (Stopford, 2009). Shipowners and charterers

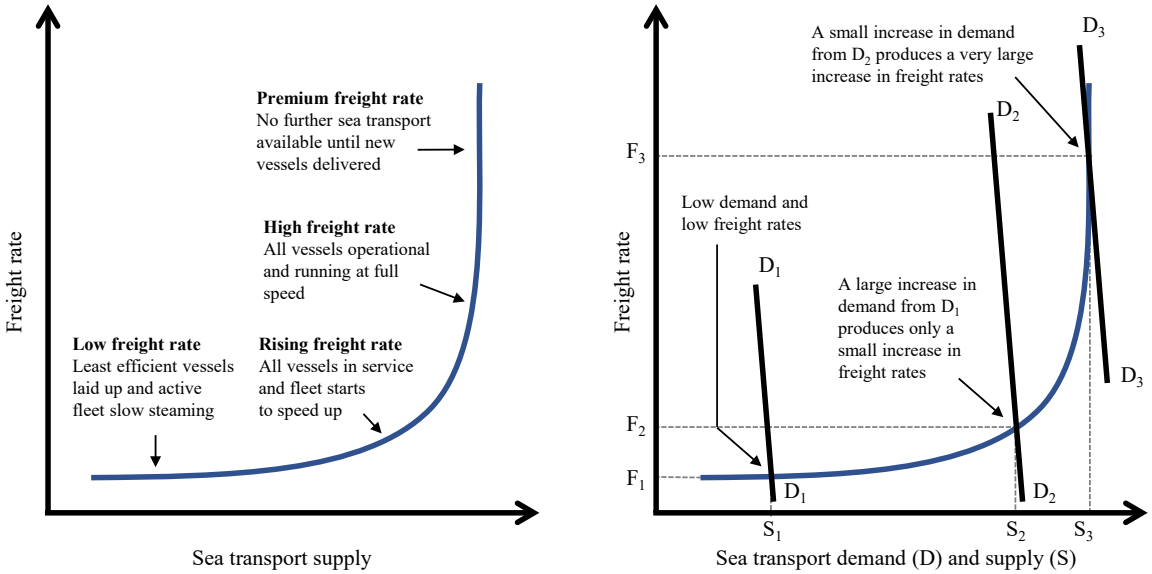


Figure 2.1: Supply and demand mechanics in the freight market

This figure shows the short-run freight market equilibrium according to Stopford (2009). Low demand and low supply result in low freight rates, F_1. A large increase in demand from D_1 to D_2 produces only a small increase in freight rate, from F_1 to F_2. A small increase in demand from D_2 to D_3 results in a large jump in rate from F_2 to F_3, because supply is nearing its maximum capacity. The maximum available supply is limited by the size and speed of the world fleet, while newbuilding of vessels will alleviate this in the longer run.

continually have to decide whether to lock in current rates, or hold out for a better deal. If a charterer believes the rates will rise, it seems rational to lock in and vice versa. As a result, market sentiment is one of the key drivers of rates in the momentary market (Stopford, 2009).

Market sentiment is affected by vessels' speed, as an increase in speed is associated with an increase in freight rates as explained in Stopford (2009). When market participants observe an increase in the fleet speed, they know this is likely related to changes in demand, as the supply curve is inflexible in the very short run. The market observes this increase in speed and expects there to be fewer available vessels for hire. The effect of this expectation can be seen in Figure 2.2. The *expectations curve* limits the supply and freight rates are pushed upward. In sum, we see that an increase in speed, which is associated with increasing prices in the short-run market, forms an expectation in the momentary market for limited supply. Even though an increase in speed does not majorly influence rates in the momentary market, it does so through the expectations curve.

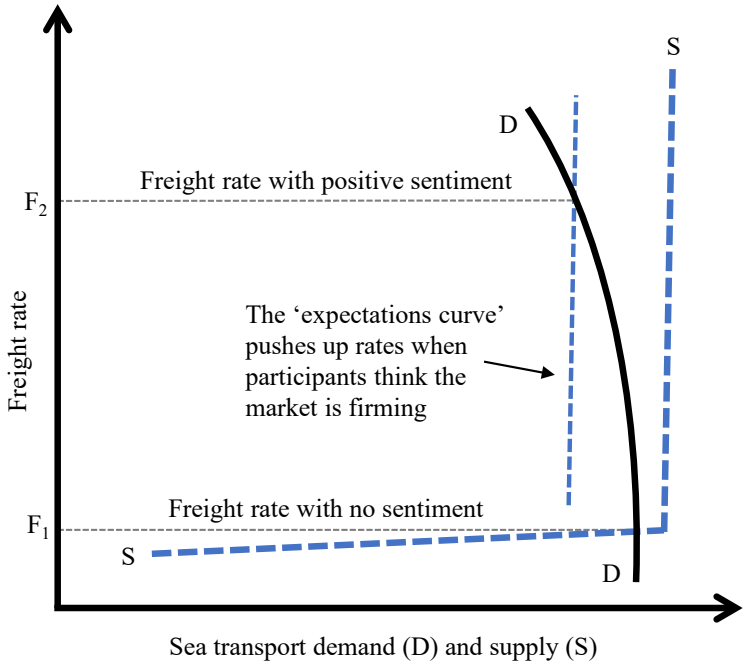


Figure 2.2: The effect of expectations in the momentary market

This figure shows the effect of expectations in the momentary market, adapted from Stopford (2009). When the speed of the fleet is increasing, the expectations curve pushes up rates from F_1 to F_2 as participants think the market is heating up. The sharp bend in the supply curve is due to no more available vessels for hire in the momentary market.

2.2 Optimal speed decisions

Vessels choose their speeds in accordance with the short run market situation and their financial objectives as determined by their *loading condition*.[1] The loading condition of a vessel refers to whether it is in ballast, travelling without cargo, or laden, travelling with cargo (Ronen, 1982).

Important variables considered when deciding on a vessel's optimal speed for a given freight rate are fuel costs, opportunity costs and the mentioned loading condition (Ronen, 1982). Fuel costs, or *bunker costs* in shipping jargon, are the single largest voyage cost, accounting for approximately 66% of all voyage costs (Stopford, 2009). There are no other relevant costs to consider as manning, insurance, maintenance and capital costs are fixed in the short run. If freight rates are high, vessels may speed up to capitalize on the

[1]Ronen (1982) uses *leg type* as a determinant for a vessel's financial objective. This categorization differs from loading condition in that the leg type of a vessel is not necessarily the same as the loading condition. In Ronen (1982), the leg type is laden if the vessel is generating daily profit.

heated market (Stopford, 2009), but in a depressed market, operators may instead choose to keep speeds at an optimal minimum to preserve fuel (Ronen, 1982; Stopford, 2009). Opportunity costs in the speed decision refers to contract-determined fees (bonuses) of late (early) arrival, and lost potential earnings on the next contract.

According to Ronen (1982), vessels in ballast have a different set of criteria for their optimal speed than laden vessels. Generally, vessels in ballast will aim to minimize the costs associated with sailing without cargo. Decreasing the speed will allow vessels to save on fuel, but this is offset by increasing opportunity costs, as arriving earlier can result in winning profitable contracts sooner. Whether a chosen speed is cost-minimizing depends on vessels' fuel consumption, bunker cost, travelling distance, other daily operational cost, and the mentioned opportunity cost. We can see that the only decision variable available to vessels is their speed, as the other variables are fixed.

Laden vessels, in contrast, aim to maximize daily profit. Similarly, this is a trade-off between saving on fuel and spending fewer days at sea. Because laden vessels need to displace more water, they also have a greater fuel consumption than vessels in ballast (Aßmann et al., 2015), making it more expensive to speed up. In the case that the opportunity cost is greater than the maximized daily profit, Ronen (1982) suggests using the cost-minimizing approach of vessels in ballast to arrive at the optimal speed instead. Like in the case of vessels in ballast, speed is the only available decision variable.

The behavioral difference between laden vessels and vessels in ballast serves as the motivation for separating the vessels by their loading conditions for our analysis. A vessel travelling in ballast has different motivations than a laden vessel, and it is relatively cheaper for a vessel in ballast to increase its speed. These differing motivations also mean that vessels respond differently to changes in the market depending on whether they are laden.

2.3 Positive wind force

When monitoring the reported speed of vessels, there is a theoretical difference between their reported speed and their chosen speed through water. This is because the reported speed is measured by satellite navigation systems, not physical movement. Wind, sea

current, swell and other environmental factors impact the actual speed of vessels (Prpić-Oršić et al., 2014). This is also supported by vessel manufacturers often reporting a vessel's fuel consumption to have a 10-15% weather margin (Watson, 2002). The vessels act as huge sails catching the wind, which affects their reported speed. This wind either pushes the vessel forwards or backwards depending on the apparent wind's direction. We define this effect as the *positive wind force*.

The positive wind force exerted on a vessel can be estimated using a few known parameters. It is possible to estimate how the wind pushes a vessel forwards, by combining the methods of Kitamura et al. (2017) and Blendermann (1994). To the best of our knowledge, there is no way to accurately estimate the effect of current and swell on vessels without acquiring wind tunnel data of all vessels. Even if this data was readily available, the drag properties of a vessel's hull deteriorates as a result of marine growth and aging. The marine growth is removed through regular maintenance and information about when this is performed for each vessel is not readily available either. The only parameters needed for estimating the positive wind force on a vessel is the apparent wind velocity, the loading condition and the length and breadth of the vessel. The effect of wind on a vessel is easily estimated, but the effect of swell and current is not.

Vessels in ballast are more likely to have their reported speed affected by positive wind force, because vessels in ballast have larger surface areas for the wind to catch. This surface area is 70% larger for vessels in ballast than laden vessels.[2] The other reason is that laden vessels displace far more water than vessels in ballast, meaning the same amount of wind force exerted on a laden vessel will have a lower impact than on a vessel in ballast. These two reasons suggest that the positive wind force effect on laden vessels is lower than that for vessels in ballast.

The positive wind force affects freight rates through the market sentiment, following the same argument as for the fleet speed. The market responds to increases in speed through the expectations curve, which in turn changes rates. As an example, on a given day, vessels travelling in ballast are reporting an increase in speed. However, the vessels are experiencing a sudden net increase in tail-wind. By taking this tail-wind into account,

[2]We compare the combined frontal and side-view projected surface areas of a laden vessel against a vessel in ballast using the estimation model from Kitamura et al. (2017) as described in Section 4.3. The surface area is calculated for the 180,000 dwt Baltic Exchange defined capesize vessel of 290 meters length and 45 meters breadth (Baltic Exchange Ltd, 2014).

we can assume the reported speed increase can at least partially be explained by the vessel benefiting from increased positive wind force. The reported increase in speed shifts the expectations curve inwards, since the market is unaware of the increase in positive wind force. When we assume the market participants are rational, the expectations curve should only shift inwards when the market is heating up. When the speed increase can be attributed to an increase in positive wind force, the vessels have not actually changed their speed decision and the expectations curve should remain unchanged.

2.4 Expected results

We expect freight rates to increase when positive wind force increases, because changes in observed fleet speed are biased when not accounting for positive wind force.

Automatic Identification System (AIS) data allows for the market to observe vessels in real time, and thus the supply side of the market. We expect any predictive effects on rates from AIS speed data to be promptly traded away, since market participants monitor fleet speed. However, we do not expect that the market is monitoring the effects of positive wind forces on vessels, as the estimation method provided by Kitamura et al. (2017) was published late in our sample period.

Certain frictions exist in the spot freight market which prevents the spot rate from responding as quickly as the FFA rate to new information. These frictions include the physical execution of the trade, costly termination fees as well as default risks (Adland and Alizadeh, 2018). These frictions do not apply to the cash-settled FFA contracts. Consequently, we expect the relatively more liquid FFA market to immediately incorporate any new information from AIS data into its prices. Approximately 70% of FFA volume was speculative in 2011, according to market sources in Nomikos and Doctor (2013).

3 Data collection

In this section we will briefly describe the freight rate data and show how the AIS and wind data was collected and structured. The data sample spans from January 2014 to December 2018.

3.1 Spot rates and forward freight agreements

The spot rate for the C5 route is from Clarksons Research (2019a).[3] The C5 spot rate is priced in dollars per ton based on a 180,000 ton shipment of iron ore from West Australia to Qingdao, China, and a round-trip on this route takes a little less than a month.[4] Australia accounted for 50% of the world's exports of iron ore and China accounted for 65% of all iron ore imports in 2018 (International Trade Centre, 2019).

Forward freight agreements (FFAs) are freight rate derivatives traded by shipowners, charterers, and speculators, in recent years mainly through clearing houses (Alizadeh and Nomikos, 2009). The contracts are popular, as they allow shipowners and charterers to hedge against the movements of the volatile capesize market, and other traders to speculate on market developments. FFAs are cash-settled difference contracts, and do not involve the actual chartering of a vessel.

We obtain the FFA rates from the Baltic Exchange Ltd (2019), where the rate is the average price for traded contracts each day. The one-month contracts used in this thesis are traded by lots, where one lot is equal to one day of time chartering a capesize vessel. The FFAs are settled on the last business day of the month against the same month's average time charter rate. For example, a one-month FFA purchased any trading day in January is settled against the average time charter price in February. When differencing these rates' time series, the rate differences are roll-adjusted to the contract of the same maturity on the last trading day of the month. We use the one-month FFA rate in our

[3]The rates for Monday February 22, 2016 is erroneously reported as Saturday February 20 in the data from Clarksons. This date has been corrected in our dataset.

[4]The traveling distance is 3,583 nautical miles between Port Hedland, Australia and Qingdao, China. The round-trip excluding time for loading and unloading is 25 days, using the mean speeds of both loading conditions in our data sample.

models, as this is the closest maturity FFA contract available, which never enters the averaging settlement period.

Only the FFAs settled on the 5 time charter (5 T/C) average are liquid in the capesize market.[5] Thus, the FFA rates we use in this thesis contain information pertaining to the 5 T/C average, and not the C5 route spot rate we predict. On May 6, 2014, the time charter route composition of the underlying freight rate of FFAs were changed to its current composition. From January 1 to May 2, 2014, we therefore impute the now deprecated 4 T/C average FFA rates to preserve a continuous time series.

3.2 Automatic Identification System

Automatic Identification System (AIS) is a maritime communication system where vessels are equipped with transponders and transmit vessel-specific data to other vessels and coastal authorities. It has been mandatory for large vessels to have an operational AIS system that transmits at all times since December 31, 2004 (IMO, 2019).

AIS data consists of current navigational information such as position in coordinates, speed over ground, course over ground, draught and a unique vessel identifier. The data can be grouped into two different categories: Dynamic data, which is automatically read and transmitted by a vessel's systems, and static data, which is manually entered by crew members. Position, speed and course are dynamic data, and draught is static data. The data are transmitted through 27 different message types in total, and Table 3.1 shows the message types used for the construction of our variables.[6] Figure 3.1 shows an example of an AIS message type 1 before it is decoded. The meaningful data is stored in the highlighted section.

```
!AIVDM,1,1,,B,177KQJ5000G?tO`K>RA1wUbN0TKH,0*5C
```

Figure 3.1: Example AIS transmission.

This figure shows an example of a raw AIS message from Raymond (2016). The binary data payload containing the message fields is highlighted.

We use five AIS message fields for the construction of our variables. We use the Maritime

[5] We were informed by the Baltic Exchange that single route FFAs are illiquid, even though single route FFAs exist in the capesize market.

[6] Readers are referred to Raymond (2016) for a complete specification of all message types.

Table 3.1: AIS message types

Message Type	Information	Frequency
1	Position Report Class A	2 seconds - 3 minutes
2	Position Report Class A (scheduled)	2 seconds - 3 minutes
3	Position Report Class A (interrogation response)	Response
5	Static and Voyage Related Data	6 minutes
27	Position Report For Long-Range Applications	3 minutes

This table shows the AIS message types based on Raymond (2016) used in the creation of our variables. Message types 1, 2 and 3 transmit dynamic position reports, and one of the three are always transmitted every 2 seconds to 3 minutes, depending on whether the vessel is moving or stationary. Message type 5 contains static information such as draught and destination. Message type 27 is a condensed version of message types 1-3 optimized for long-range transmissions including satellites.

Mobile Service Identity (MMSI) to identify unique vessels; speed over ground, to monitor supply; coordinates and course over ground, to match vessels with wind data; and draught, to separate vessels by loading condition. Speed over ground, course over ground, and coordinates are dynamic data from message types 1, 2, 3, and 27, and draught is static data from message type 5 only. All message types are linked to a single vessel through the MMSI.

We obtain raw AIS data from the Norwegian Coastal Authorities (NCA). The raw data consists of all AIS messages received daily, exclusively by satellites. The NCA has collected worldwide AIS data by satellite since July 2010 and in July 2014, a second NCA satellite became operational. Two more satellites with improved detection equipment in busy waters were launched in July 2017. Both these events improved the NCA's global coverage, and in turn the quality of the AIS data.

3.3 Structuring the AIS data

We use the library Libais by Schwehr (2018) in a distributed computing environment to decode the raw AIS messages. Libais does not natively read the type 5 messages in the raw AIS data from NCA, but by relaxing the requirements to message type 5's padded bits (Raymond, 2016), they are decoded without problems.

All transmissions from non-capesize vessels as identified by Clarksons Research's World

Fleet Registry are discarded. In the context of our paper, we define a capesize vessel to be in the range of 140,000 to 350,000 dead weight tons (dwt). The basis of the Baltic Exchange's capesize vessel is 180,000 dwt (Baltic Exchange Ltd, 2014), so smaller capesize vessels of less than 140,000 dwt are discarded. Vessels larger than 350,000 dwt are also rarely traded on the open market and are therefore discarded as well.[7] The resulting observable capesize fleet contains 1,537 vessels (Clarksons Research, 2019b).

We only extract one observation per vessel per day. In order for an observation of a vessel to be valid, we require at least one message of type 1, 2, 3 or 27 to get speed, course and positional information and one message of type 5 for the draught information. We only use the first observation after 00:00 UTC for each vessel if multiple messages are received from the same vessel in a single day. Table 3.2 provides the structure of vessel-level, daily observations. Each observation is grouped by MMSI and date. All observed capesize vessels on April 1, 2018, are plotted in Figure 3.2.

Table 3.2: Vessel observation structure

Field	Format / Range	Unit	Message type / Origin
Date	YYYY-MM-DD	Date	Message metadata
MMSI	9 digits	Integer	1, 2, 3, 5 and 27
Latitude	−90 – 90	Degrees	1, 2, 3 and 27
Longitude	−180 – 180	Degrees	
Speed over ground	0 – 102	Knots	
Course over ground	0 – 359.9	Degrees	
Draught	0 – 25.5	Meters	5
Length over all	267 – 340	Metres	World Fleet Register
Breadth	43 – 60	Metres	

This table shows the structure of a vessel observation constructed from AIS message data. A vessel may only be observed once per day and the message types are combined into a single observation by matching the date and the unique MMSI identifier across messages. The date of a vessel observation is from the AIS message metadata. The length over all and breadth of a vessel is from the World Fleet Register (Clarksons Research, 2019b). Valid data ranges for the AIS message fields are from Raymond (2016).

Due to the nature of the data, a vessel observed on a given day might be unobserved the next. As we aggregate all vessel-level data to the global fleet level in Section 4.4, this should not lead to bias in our analysis. This assumption relies on satellites picking up

[7] A majority vessels larger than 350,000 dwt are owned by companies whose primary activity is Chinese state interests according to the World Fleet Register (Clarksons Research, 2019b), making them unlikely to be traded on the open market.

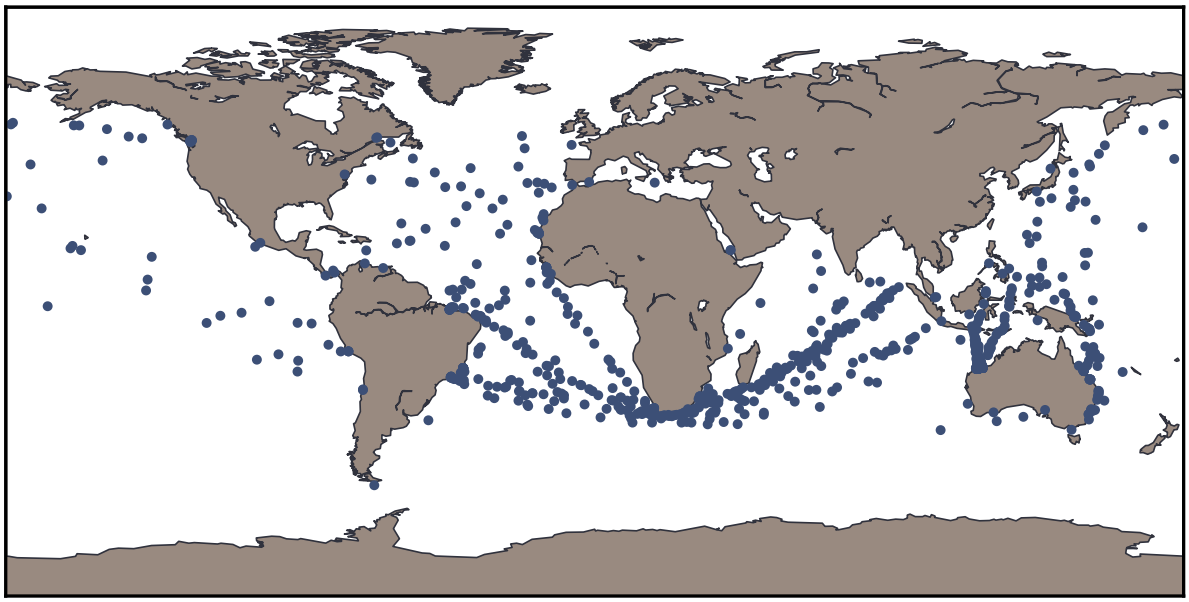


Figure 3.2: Plot of observed vessels
This figure shows all observed capesize vessels on April 1, 2018.

random AIS-messages and not being biased towards certain vessels or regions. Our data exploration does not offer any evidence to the contrary.

3.4 Wind forecasts

We collect the forecast wind data from the European Reanalysis System 5 (ERA5) dataset maintained by the European Centre for Medium-Range Forecasts (ECMWF). This dataset consists of a $0.28125° \times 0.28125°$ latitude/longitude grid of 10-metre u- and v-components of wind forecast at 06:00 UTC each day in the sample period. The u- and v-components are the vector representation of wind used in meteorological applications. We download the data using the Meteorological Archival and Retrieval System (MARS) web-API developed by ECMWF (2019). We read the resulting General Regularly-distributed Information in Binary form (GRIB) files using EcCodes and Pygrib developed by ECMWF (2018a) and Whitaker (2019) respectively. This dataset closely resembles ECMFW's operational model's availability of real-time forecasts at 06:00 UTC (ECMWF, 2018b). We used the ERA5 dataset instead of the operational model's archive due to easier and faster access to the data, as the data is stored on disk instead of tapes (C3S, 2017). A plot of the wind in the Mollucan Sea in Indonesia on April 1, 2018, is shown in Figure 3.3 along with vessels

observed on the same day.

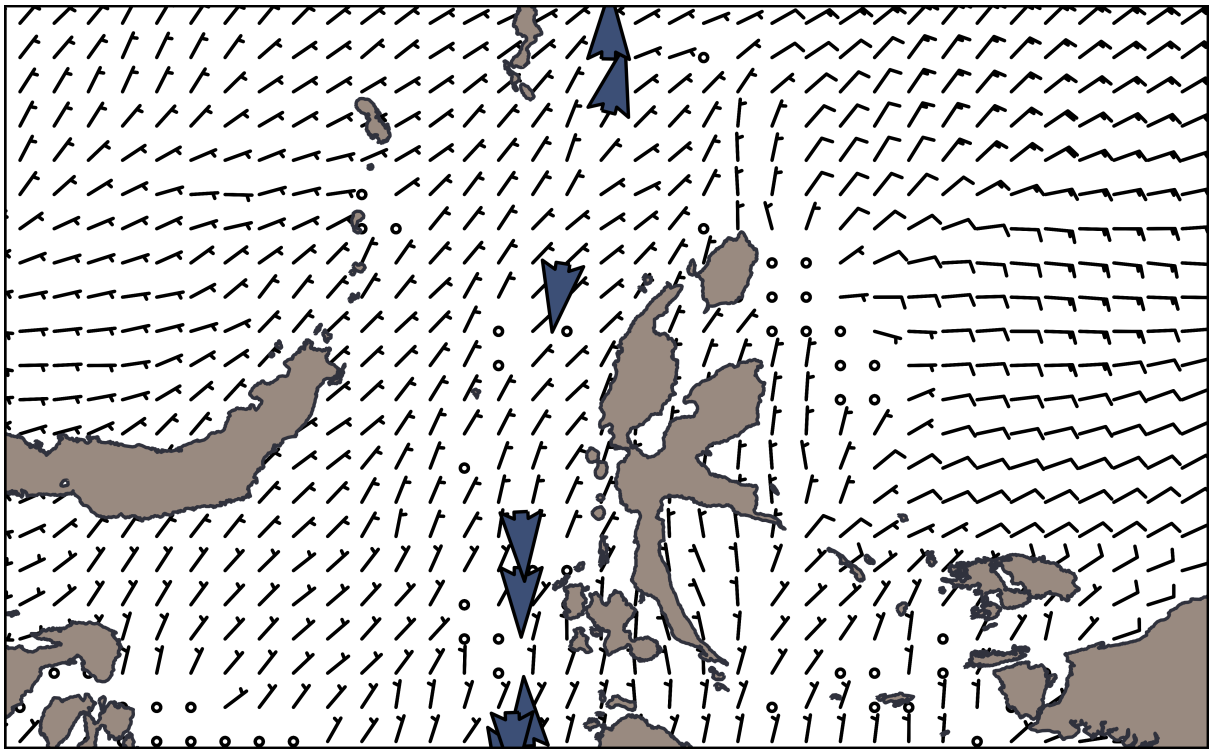


Figure 3.3: Plot of wind data

This figure shows the wind in the northern Mollucan Sea in Indonesia on April 1, 2018. The barbs indicate wind direction and speed. Each full fin on a barb indicates 10 knots of wind. The blue pointers are capesize vessels observed on the same day.

3.5 Erroneous data and missing observations

We discard invalid observations in accordance with the documentation for AIVDM/AIVDO protocol decoding by Raymond (2016). We also remove obvious erroneous values: Vessel observations with unreasonably high speeds of 20 knots and above or invalid coordinates.

We use draught information from AIS message type 5 to determine a vessel's loading condition. This means all vessel observations containing only message type 1, 2, 3 or 27 are discarded. In the AIS data sample, 44.04% of observed vessels across all message types are missing draught information. As argued in Section 2.2, vessels with different loading conditions have different objective functions. We are therefore not able to take advantage of observations missing the draught information. As noted in Section 3.2, message type 5 contains static information input by crew members, and is therefore prone to human error.

We downloaded the World Fleet Registry in January 2019 and the registry only contains active vessels as of that time. This discrepancy means all vessels scrapped between 2014 and 2018 are unobservable, as we need to match each AIS message to a concrete vessel. In total, 247 capesize vessels were demolished in this period (Clarksons Research, 2019a).

There are almost no observations of vessels from August 4 to August 14, 2017, in the raw AIS data due to break-in of the two newly launched satellites. We remove this date range from our sample before running our regressions, as we find no reasonable way to impute values. The missing dates account for 0.6% of the daily data sample from 2014 to 2018. There is a theoretical possibility that this missing dates impact the results of autocorrelation and stationarity tests.

4 Variable construction

In this section we present how we construct and aggregate explanatory variables for our models. We construct six variables from the AIS and wind data, grouped by their loading condition: Positive wind force on vessels, apparent wind speed on vessels and the speed of vessels.

4.1 Separating vessels by loading condition

We use the draught information of a vessel to define a *load ratio*. This ratio is used to determine a vessel's loading condition. We calculate this load ratio because the laden design draught of capesize vessels vary with their size. The load ratio for each vessel observation is then given as $\ell = \frac{\rho}{v}$, where ℓ is the load ratio, ρ is the reported draught in meters and v is the laden design draught in meters. We selected 0.73 as an appropriate cutoff value to determine a vessel's loading condition after consulting Figure 4.1. Vessels with a load ratio below this cutoff point are in ballast, and vessels above are laden. Summary statistics for the loading condition ratios are found in Table 4.1.

Table 4.1: Summary statistics of loading condition ratios

	Count	Mean	Std	Median	Min	Max
Ballast	1,815	47.47	3.18	47.67	35.03	60.00
Laden	1,815	52.53	3.18	52.33	40.00	64.97
Note:					Values in percentages	

This table shows the summary statistics of daily loading condition ratios. All values are in percentages. *Ballast* is vessels travelling without cargo, and *Laden* is vessels travelling with cargo.

4.2 Apparent wind

The estimated apparent wind on a vessel is determined by the forecast wind near the vessel, taking the vessel's speed and course into account. In the AIS data, a vessel's speed is given in knots and its heading is given in degrees relative to north. A heading of 0° is bound north and a heading of 90° is bound east. We calculate the longitudinal x

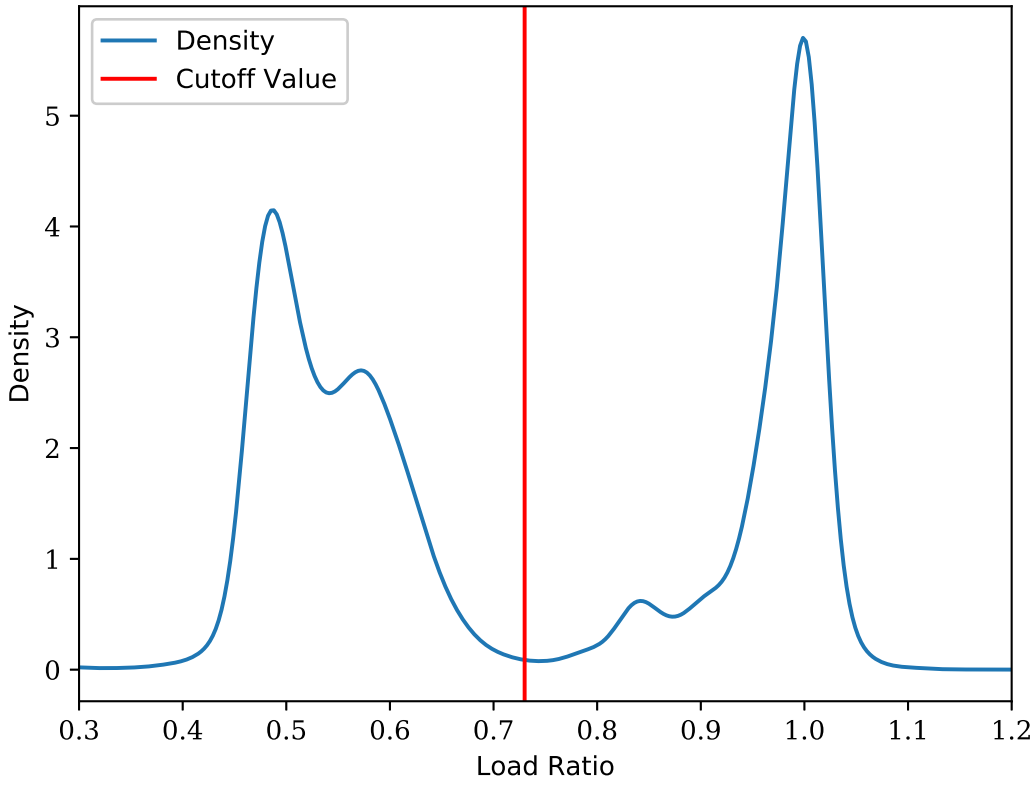


Figure 4.1: Load ratio density plot

The density plot shows if a vessel is under ballast, determined by its load ratio, reported draught over design draught. The vertical line is the cutoff point. Vessels with a load ratio above the cutoff are considered as laden, and below as in ballast.

and latitudinal y components of the vessel's velocity by using basic trigonometry shown in Equation 4.1, where $|\vec{S}|$ is the speed of the vessel and Φ is the heading of the vessel. We find the corresponding wind vector of each vessel by matching the coordinates of a vessel with to the closest data point by absolute degrees in the wind data. We then calculate the apparent wind, $\vec{V}_a$ by subtracting the vessel's velocity from the wind vector, shown in Equation 4.2. This is the same as adding the vessel's induced wind to the wind vector. We finally transform the apparent wind $\vec{V}_a$ back to absolute wind speed, $|\vec{V}_a|$ and relative angle of attack, ϵ as shown in Equation 4.3. An illustration of the apparent wind

calculation is shown in Figure 4.2.

$$\vec{S} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -|\vec{S}| \sin \Phi \\ -|\vec{S}| \cos \Phi \end{bmatrix} \tag{4.1}$$

$$\vec{V}_a = \vec{V} - \vec{S} = \begin{bmatrix} u \\ v \end{bmatrix} - \begin{bmatrix} x \\ y \end{bmatrix} \tag{4.2}$$

$$\begin{aligned} \epsilon &= 180 + \frac{180}{\pi} \arctan 2(u_a, v_a) \\ |\vec{V}_a| &= \sqrt{u_a^2 + v_a^2} \end{aligned} \tag{4.3}$$

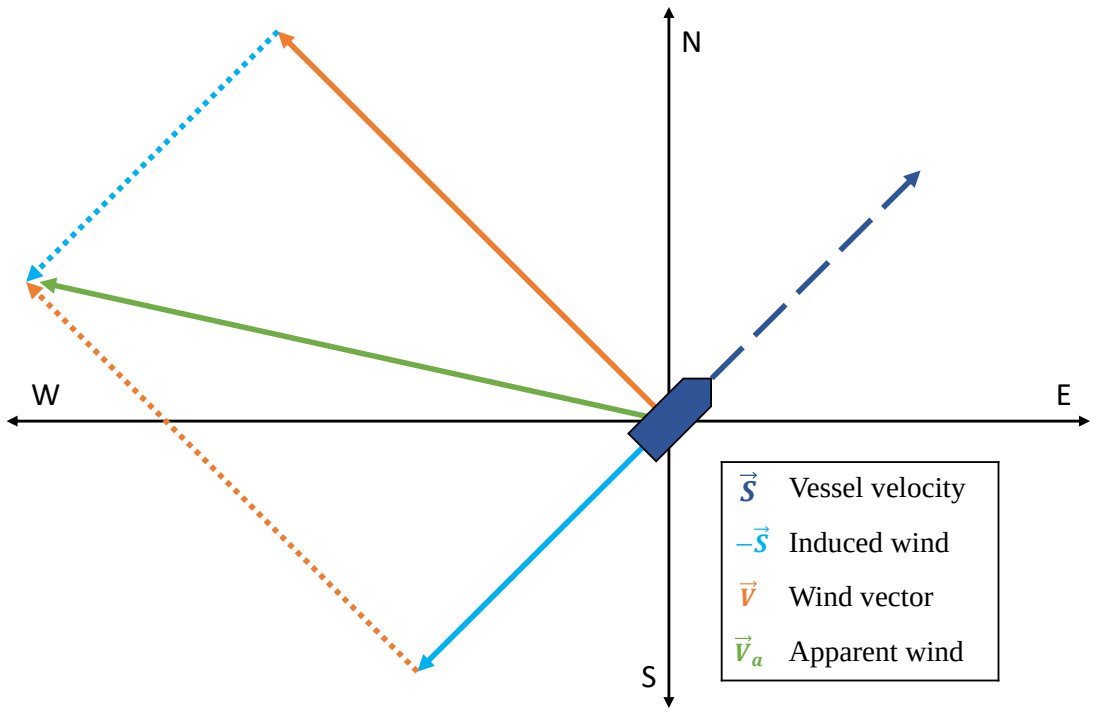


Figure 4.2: Apparent wind on a vessel

This figure shows the apparent wind on a vessel. The vessel is travelling with a course due north-east, producing induced wind from the north-east. The wind in the area is coming from the south-east. The apparent wind is calculated by adding the vessel's induced wind to the wind vector.

4.3 Positive wind force

We propose a novel *positive wind force measure* by combining the estimation method for wind load coefficients by Blendermann (1994) and above-water structural parameters by Kitamura et al. (2017).

We start by estimating the frontal projection of the vessel, A_F, and the side-view projection of the vessel, A_L. We estimate the parameters A_F and A_L by using the linear regression models from Kitamura et al. (2017). The only inputs to these regression models are the vessel's length over all, L_{OA}, breadth, B, and loading condition.

We continue by estimating the wind load, X. We first calculate the wind load coefficient, C_X, using the estimates for A_F and A_L along with the remaining parameters of a cargo vessel as defined by Blendermann (1994): The cross-force, $\delta = 0.40$; the coefficient of lateral resistance, $CD_t = 0.85$; and the coefficient of longitudinal resistance, CD_l. CD_l is given as a linear function of the wind's angle of attack ϵ, shown in Equation 4.5. The density of air, ρ, is approximately equal to 1.225 kg/m^3 at sea level (Cavcar, 2000). The method for estimating the wind load coefficient is shown in Equation 4.4 and 4.5 (Blendermann, 1994). Solving Equation 4.6 for the wind load X, yields the total longitudinal wind load exerted on a single vessel.

We further standardize the data, because larger vessels have larger surface areas resulting in a higher absolute wind load. The resulting longitudinal wind load X is therefore divided by the vessel's dwt capacity. This gives us our positive wind force measure in newton per ton.

$$C_X = -CD_l \frac{A_L}{A_F} \frac{\cos \epsilon}{1 - \frac{\delta}{2}(1 - \frac{CD_l}{CD_t}) \sin^2 2\epsilon} \tag{4.4}$$

$$CD_l = CD_{l_{A_F}} \frac{A_L}{A_F}, \quad CD_{l_{A_F}} = \begin{cases} 0.65 \text{ if } \epsilon = 0^\circ \\ 0.55 \text{ if } \epsilon = 180^\circ \end{cases} \tag{4.5}$$

$$C_X = \frac{X}{0.5\rho|\vec{V}_a|^2 A_F} \tag{4.6}$$

4.4 Data aggregation

Having grouped the observations by their loading condition, we calculate mean positive wind force, $Wind^{Force^+}$, mean apparent wind speed, $Wind^{Speed}$, and mean speed, $Speed$, for each day in our sample. Only vessels with a transmitted speed of 6 knots or greater are included in this aggregated dataset as in Regli and Nomikos (2019), in order to exclude stationary vessels and vessels not travelling on open seas. We use mean aggregation, as we are interested in worldwide short-run supply as the driver for market sentiment. Summary statistics for the vessel observations before data aggregation are found in Table 4.2.

Table 4.2: Vessel-level summary statistics

	Count	Mean	Std	Median	Min	Max
$Wind^{Force^+}_{Ballast}$	333,032	0.20	0.21	0.15	-0.72	1.85
$Wind^{Speed}_{Ballast}$	333,032	12.54	5.71	12.35	0.00	47.64
$Speed_{Ballast}$	333,032	12.07	1.89	12.30	6.00	20.00
$Wind^{Force^+}_{Laden}$	367,069	0.09	0.12	0.05	-0.44	1.25
$Wind^{Speed}_{Laden}$	367,069	13.07	5.68	12.98	0.00	43.86
$Speed_{Laden}$	367,069	10.89	1.45	11.00	6.00	20.00
$Wind^{Force^+}$	700,101	0.14	0.18	0.09	-0.72	1.85
$Wind^{Speed}$	700,101	12.82	5.70	12.68	0.00	47.64
$Speed$	700,101	11.45	1.77	11.40	6.00	20.00

This table shows the summary statistics for the vessel-level observations. The sample data runs from January 2014 to December 2018. $Wind^{Force^+}$ is the mean positive wind force per vessel dwt capacity in N/t. $Wind^{Speed}$ is the mean apparent wind speed on vessels in knots. $Speed$ is the mean speed of moving vessels where $Speed > 6$kn. The $Wind$ and $Speed$ variables are separated by the vessels' loading condition.

5 Summary statistics

In this section, we present summary statistics and unit root tests for our daily data sample running from January 2014 to December 2018. Summary statistics in levels and log-differences are shown in Table 5.2 and time series plots of are shown in Figure 5.1. A correlation matrix between the log-differenced variables is also shown in the appendix in Table A1.1.

In levels, positive wind force on vessels in ballast ranges from 0.10 N/t to 0.34 N/t, with a mean of 0.20 N/t. Apparent wind speed on vessels in ballast varies from 8.83 kn to 17.20 kn, with a mean of 12.57 kn. Positive wind force on laden vessels ranges from 0.04 N/mt to 0.17 N/t, with a mean of 0.09 N/t. Apparent wind speed on laden vessels varies from 9.82 kn to 18.05 kn, with a mean of 13.10 kn. Speed of vessels in ballast is in the range of 11.20 kn to 13.45 kn, with a mean of 12.06 kn. Speed of laden vessels ranges from 10.07 kn to 11.39 kn, and has a mean of 10.88 kn. The bunker fuel price varies from \$155.50/t to \$641.50/t, with a mean value of \$388.61/t. FFA rates vary from \$2,382/day to \$26,680/day and has a mean of \$13,099/day. The C5 spot rate ranges between \$2.84/t and \$11.76/t, with a mean of \$6.36/t. The mean speed of vessels in ballast is higher than for laden vessels, which is in line with our expectations from Section 2.2.

The positive wind force has a positive value, which means vessels more often than not benefit from positive wind force. We can provide two explanatory factors for this positive mean. Blendermann (1994) shows that vessels in general have a higher wind load coefficient when in tail-wind, compared to head-wind. This means wind coming from the aft of the vessel have a higher effect on the vessel's positive wind force. Bialystocki and Konovessis (2016) also state it is rational for vessels to optimize their voyage route by taking the weather into account.

We use multiple tests for unit roots to test the time series for stationarity. We see the unit root tests in Table 5.1 for the time series gives us inconclusive results for $C5$, FFA, and speed of vessels in ballast. The Augmented Dickey–Fuller (1981) (ADF) and Phillips–Perron (1988) (PP) tests suggests the time series are weakly stationary, but the Kwiatkowski et al. (1992) (KPSS) test suggest they have a unit root. Positive wind force, apparent wind speed for vessels in both loading conditions and speed of laden vessels,

are stationary at level. By log-differencing the data, we ensure all the time series are stationary, and avoid the dangers of under-differencing (Plosser and Schwert, 1978).

Table 5.1: Unit root tests

	Levels			Log Differences		
	ADF	PP	KPSS	ADF	PP	KPSS
$Wind^{Force^+}_{Ballast}$	-5.88***	-19.43***	0.2	-15.86***	-78.74***	0.04
$Wind^{Speed}_{Ballast}$	-5.31***	-21.05***	0.15	-13.45***	-84.35***	0.16
$Wind^{Force^+}_{Laden}$	-5.31***	-21.26***	0.45*	-14.29***	-86.68***	0.17
$Wind^{Speed}_{Laden}$	-4.6***	-21.45***	0.24	-15.05***	-84.23***	0.15
$Speed_{Ballast}$	-5.46***	-18.61***	0.52**	-17.64***	-101.26***	0.05
$Speed_{Laden}$	-4.72***	-17.7***	0.45*	-14.08***	-90.36***	0.03
$Bunker$	-1.94	-1.87	1.46***	-25.94***	-26.87***	0.43*
FFA	-2.88**	-3.03**	1.22***	-24.45***	-29.49***	0.43*
$C5$	-3.39**	-3.64***	1.11***	-23.15***	-23.32***	0.11
Note:						* p < 0.1; ** p<0.05; *** p<0.01

This table shows multiple stationarity tests for the sample data. ADF is the Augmented Dickey–Fuller (1981) test with the lag order chosen by minimizing AIC. The ADF regression includes an intercept. PP is the Phillips–Perron (1988) test. The alternative hypothesis of ADF and PP tests state the time series is weakly stationary. The critical values for the ADF and PP tests are −3.44 at 1% level, −2.86 at 5% level, and −2.57 at 10% level. KPSS is the Kwiatkowski et al. (1992) unit root test. The null hypothesis of the KPSS unit root test states the time series is weakly stationary. The critical values for the KPSS test are 0.74 at 1% level, 0.46 at 5% level, and 0.35 at 10% level.

There is a risk that we over-difference the time series and induce negative autocorrelation in the first lag when differencing an already stationary time series (Plosser and Schwert, 1978). We therefore check for negative autocorrelation in the first lag using a autocorrelation function (ACF) plot. We do not see any evidence of over-differencing in the response variables $C5$ and FFA or the explanatory wind variables. The speed of laden vessels and vessels in ballast show weak signs of over-differencing. The ACF-plots are found in the appendix in Figure A2.1.

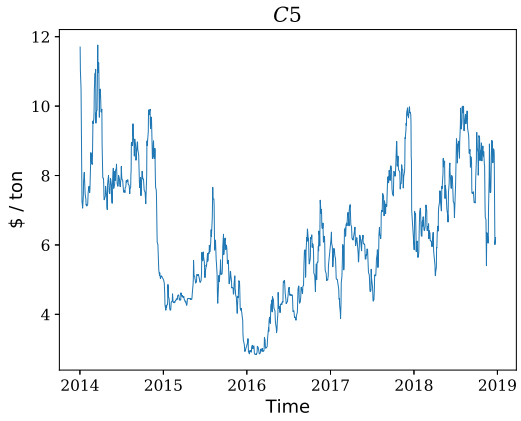
We test the C5 spot rate and FFA for cointegration using the Johansen (1991) cointegration test. Even though Adland and Alizadeh (2018) suggest these rates to be cointegrated,

we fail to find a cointegrating relationship between the rates. A possible reason for this failure might be that the FFA rate is settled against the 5 T/C average and is therefore a slightly different freight commodity than the C5 spot rate.

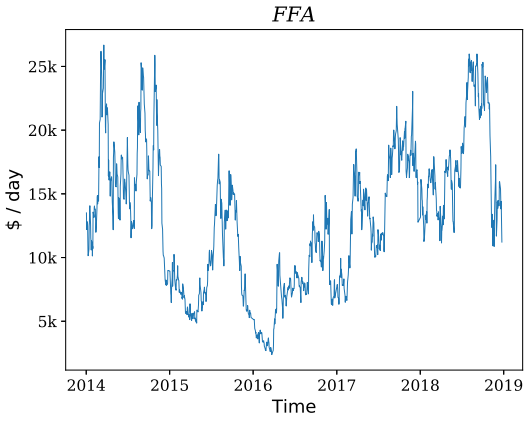
Table 5.2: Summary statistics

Levels								
	Count	Mean	Std	Median	Min	Max	Skewness	Kurtosis
$Wind^{Force^+}_{Ballast}$	1,815	0.20	0.04	0.20	0.10	0.34	0.23	0.18
$Wind^{Speed}_{Ballast}$	1,815	12.57	1.18	12.50	8.83	17.20	0.30	0.41
$Wind^{Force^+}_{Laden}$	1,815	0.09	0.02	0.09	0.04	0.17	0.65	0.66
$Wind^{Speed}_{Laden}$	1,815	13.10	1.08	13.02	9.82	18.05	0.40	0.59
$Speed_{Ballast}$	1,815	12.06	0.29	12.04	11.20	13.45	0.42	0.60
$Speed_{Laden}$	1,815	10.88	0.19	10.87	10.07	11.39	-0.19	0.24
$Bunker$	1,249	388.61	131.00	363.00	155.50	641.50	0.35	-0.85
FFA	1,249	13,099	5,573	13,076	2,382	26,680	0.24	-0.67
$C5$	1,249	6.36	1.88	6.19	2.84	11.76	0.19	-0.73
Differences								
	Count	Mean	Std	Median	Min	Max	Skewness	Kurtosis
$\Delta \ln Wind^{Force^+}_{Ballast}$	1,813	0.00	0.13	-0.00	-0.61	0.50	0.02	0.48
$\Delta \ln Wind^{Speed}_{Ballast}$	1,813	-0.00	0.07	0.00	-0.36	0.27	0.07	0.48
$\Delta \ln Wind^{Force^+}_{Laden}$	1,813	-0.00	0.16	-0.00	-0.65	0.56	0.12	0.51
$\Delta \ln Wind^{Speed}_{Laden}$	1,813	-0.00	0.06	-0.00	-0.26	0.25	0.10	0.32
$\Delta \ln Speed_{Ballast}$	1,813	-0.00	0.02	-0.00	-0.07	0.06	-0.09	0.70
$\Delta \ln Speed_{Laden}$	1,813	-0.00	0.01	0.00	-0.04	0.04	-0.07	0.46
$\Delta \ln Bunker$	1,248	-0.00	0.02	0.00	-0.11	0.09	0.05	4.82
$\Delta \ln FFA$	1,248	-0.00	0.06	-0.01	-0.22	0.24	0.28	1.05
$\Delta \ln C5$	1,248	-0.00	0.04	-0.00	-0.19	0.14	0.07	2.69

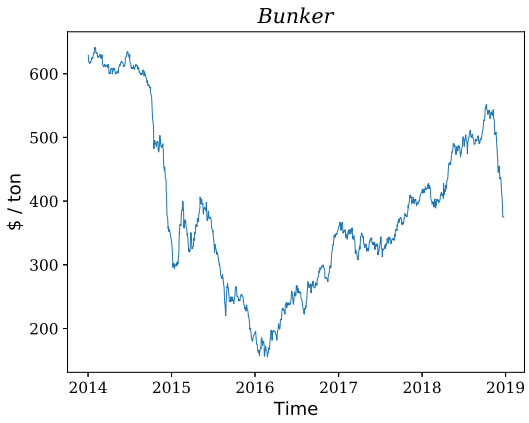
This table shows summary statistics for the sample data in levels and log-differences. The sample data runs from January 2014 to December 2018 and is sampled daily. $Wind^{Force^+}$ is the mean positive wind force per vessel dwt capacity in N/t. $Wind^{Speed}$ is the mean apparent wind speed on vessels in knots. *Speed* is the mean speed of moving vessels in knots where $Speed > 6$kn. *Wind* and *Speed* variables are separated by the vessels' loading condition. *Bunker* is the bunker fuel oil price in \$/t. FFA is the one-month 5 time charter average forward price in \$/day. $C5$ is the C5 route spot price in \$/t. In log-differences, FFA rates are rolled over on the last trading day of the month.



(a) Rate of C5

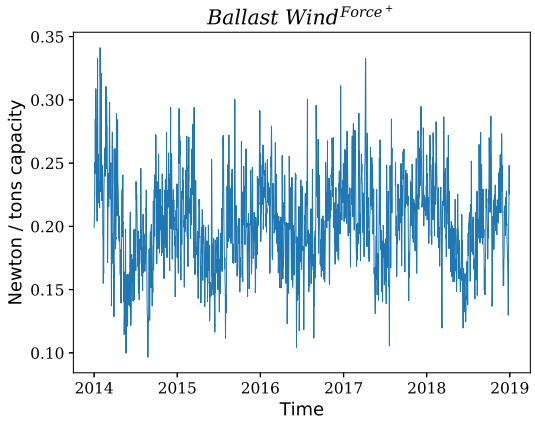


(b) Rate of FFA

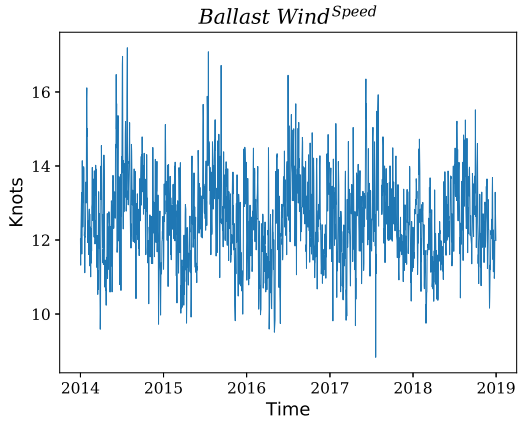


(c) Price of bunker fuel oil

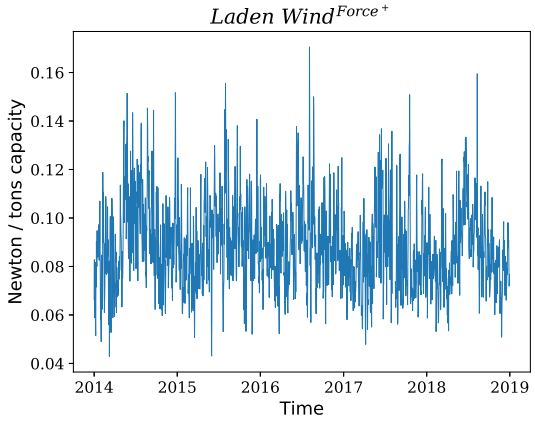
Figure 5.1: Time series plots



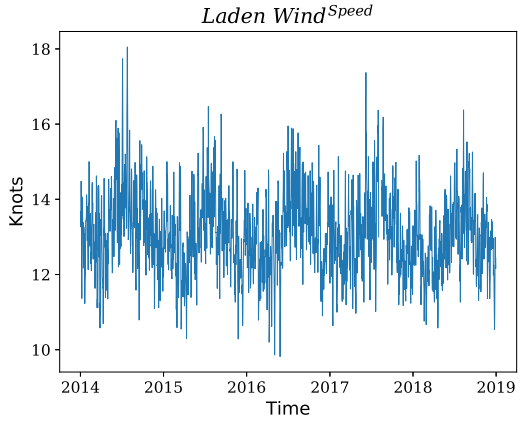
(d) Positive wind force on vessels in ballast



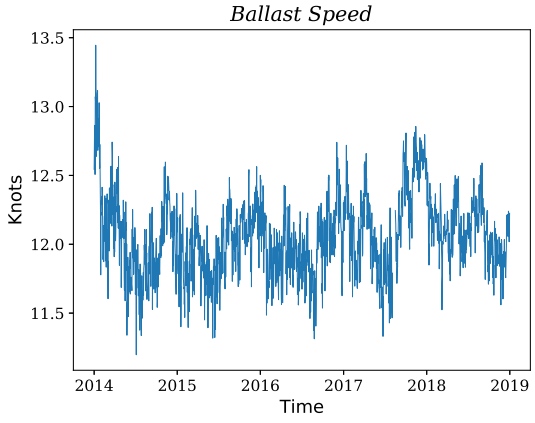
(e) Apparent wind on vessels in ballast



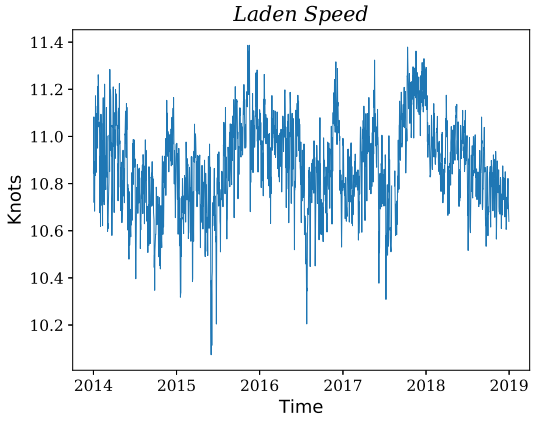
(f) Positive wind force on laden vessels



(g) Apparent wind on laden vessels



(h) Speed of vessels in ballast



(i) Speed of laden vessels

Figure 5.1: Time series plots (cont.)

This figure shows the time series plots for our sample data in levels. The sample data runs from January 2014 to December 2018 and is sampled daily. $C5$ is the C5 route spot rate in \$/t. FFA is the one-month 5 time charter average forward price in \$/day. $Bunker$ is the bunker fuel oil price in \$/t. $Wind^{Force^+}$ is the mean positive wind force per vessel dwt capacity in N/t. $Wind^{Speed}$ is the mean apparent wind speed on vessels in knots. $Speed$ is the mean speed of moving vessels in knots where $Speed > 6$kn. The $Wind$ and $Speed$ variables are separated by the vessels' loading condition.

6 Results

In this section we present our models for predicting the C5 spot rate and FFA rate in daily sampled data using our novel positive wind force measure. The positive wind force measure is made possible by combining the positional information of vessels from AIS data with wind forecasts and vessel dimensions. We show that our measure is able to predict future rates in daily sampled data.

We control our models for several factors. We control our positive wind force measure for the possibility of it only being a proxy for apparent wind, by including a control measure for apparent wind speed on a vessel. The control for apparent wind speed is a proxy for sea conditions. Sea conditions are determined largely by wind, as evident in the commonly used Beaufort-scale, where sea conditions are can be expressed as a function of the wind speed. Worsening sea conditions can for instance increase risk, and an increase in freight rates would therefore be expected. The speed of the fleet is controlled for, in line with our discussion in Section 2.3, since the positive wind force affects the speed of vessels which in turn predicts rates. Our sample period from 2014 to 2018 covers major fluctuations in oil price, and we have therefore included bunker fuel costs as a control variable in our regressions as well. We account for autocorrelation by including a single lagged dependent variable. For the spot rate model, FFA rates are also included, as they are known to be good predictors of future freight rates (Adland and Alizadeh, 2018). Other macroeconomic effects reside in the error term.

We always use the closest temporal data when fitting our models. This means in particular that when fitting values for Mondays in the response variable, we use AIS and wind data from Sunday, as this data is readily available before the market opens on Monday morning.

6.1 Spot freight rate predictive regression

In order to ensure all variables are stationary, we log-difference the variables as discussed in Section 5. We then set up our linear model for predicting the C5 spot rate as shown in Equation 6.1. This model predicts changes in next-day rates by observing changes in the

current supply, as demand is inelastic in the short run.

$$\begin{aligned}\Delta \ln C5_t = {} & \beta_1 \Delta \ln \underset{Ballast}{Wind}_{t-1}^{Force^+} + \beta_2 \Delta \ln \underset{Ballast}{Wind}_{t-1}^{Speed} \\ & + \beta_3 \Delta \ln \underset{Ballast}{Speed}_{t-1} + \beta_4 \Delta \ln \underset{Laden}{Speed}_{t-1} \\ & + \beta_5 \Delta \ln Bunker_{t-1} + \beta_6 \Delta \ln FFA_{t-1} + \beta_7 \Delta \ln C5_{t-1} + u_t\end{aligned} \tag{6.1}$$

We fit the model using ordinary least squares (OLS) regression. The regression results shown in Table 6.1 model 7, show the C5 rate increases after an increase in positive wind force on vessels in ballast. Speed of vessels in ballast is not significant in any of the regressions. This insignificance is in line with our expectations as AIS supply effects should already be captured by changes in the more liquid the FFA rate. The FFA rate is as expected a significant predictor of the future C5 spot rate. The C5 rate increases in models 3 and 6 as the speed of laden vessels decreases. The effect, however, is not significant in our final model specification, model 7, from Equation 6.1.

In economic terms, for an increase in positive wind force on vessels in ballast by one standard deviation from average levels, holding everything else equal, the C5 spot rate is expected to increase by 0.3%.[8] At average levels of the C5 spot rate at \$6.36/t this amounts to an increase of \$0.0197/t. For a 180,000 ton shipment, this leads to an increase in voyage price of \$3,546.

[8] At average levels of 0.20 in positive wind force and a standard deviation of 0.04, the expected change in the C5 spot rate is $0.017 \ln(\frac{0.20+0.04}{0.20}) = 0.00310$ using model 7 from Table 6.1.

Table 6.1: Predictive regression on C5 rates

	Response variable: $\Delta \ln C5_t$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta \ln Wind^{Force^+}_{t-1}$ $_{Ballast}$	0.015*			0.019**	0.013*		0.017**
	(0.008)			(0.008)	(0.008)		(0.008)
$\Delta \ln Wind^{Speed}_{t-1}$ $_{Ballast}$	0.026*			0.021	0.024*		0.02
	(0.013)			(0.014)	(0.013)		(0.014)
$\Delta \ln Speed_{t-1}$ $_{Ballast}$		-0.052		-0.094		-0.045	-0.084
		(0.061)		(0.066)		(0.061)	(0.066)
$\Delta \ln Speed_{t-1}$ $_{Laden}$			-0.183**		-0.134	-0.18**	-0.122
			(0.083)		(0.084)	(0.083)	(0.084)
$\Delta \ln Bunker_{t-1}$	0.079	0.075	0.083	0.08	0.084	0.083	0.085
	(0.059)	(0.059)	(0.059)	(0.059)	(0.059)	(0.059)	(0.059)
$\Delta \ln FFA_{t-1}$	0.233***	0.235***	0.234***	0.233***	0.233***	0.234***	0.233***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
$\Delta \ln C5_{t-1}$	0.184***	0.182***	0.182***	0.184***	0.183***	0.182***	0.184***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Observations	1239	1239	1239	1239	1239	1239	1239
R^2	0.244	0.237	0.239	0.245	0.245	0.239	0.246
Adjusted R^2	0.241	0.234	0.237	0.241	0.242	0.236	0.242
F Statistic	79.496***	95.677***	97.055***	66.645***	66.752***	77.719***	57.472***
BG LM Statistic	31.668*	32.887*	31.871*	32.378*	31.213*	32.22*	31.88*
BP LM Statistic	7.423	7.188	7.715	7.404	7.777	7.654	7.752
KPSS Residuals	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Note:							*p<0.1; **p<0.05; ***p<0.01

This table shows the regression results of the C5 predictive model. Standard errors of coefficients are in parenthesis. $C5$ is the C5 route spot price for in \$/t. All variables are log-differenced. BG LM is the Breusch–Godfrey (1978; 1978) Lagrange multiplier χ^2 test for serial correlation in the residuals. The null hypothesis of BG LM test states that the residuals are not serially correlated. BP LM is the Breusch–Pagan (1979) Lagrange multiplier χ^2 test for contemporaneous heteroskedasticity. The null hypothesis of the BP LM test states that the residuals are contemporaneously homoskedastic. KPSS is the Kwiatkowski et al. (1992) unit root test. The null hypothesis of the KPSS unit root test states the residuals are weakly stationary. Variance inflation factors (VIF) for the explanatory variables are $\text{VIF}(\Delta \ln \underset{ballast}{Wind^{Force^+}_{t-1}}) = 1.43$, $\text{VIF}(\Delta \ln \underset{ballast}{Wind^{Speed}_{t-1}}) = 1.26$, $\text{VIF}(\Delta \ln \underset{ballast}{Speed_{t-1}}) = 1.17$, $\text{VIF}(\Delta \ln \underset{laden}{Speed_{t-1}}) = 1.06$, $\text{VIF}(\Delta \ln Bunker_{t-1}) = 1.03$, $\text{VIF}(\Delta \ln FFA_{t-1}) = 1.44$ and $\text{VIF}(\Delta \ln C5_{t-1}) = 1.30$.

6.2 Forward freight agreement predictive regression

We set up the linear model for the predicting FFA rates similar to the the spot freight rate model. The model expression is shown in Equation 6.2.

$$\begin{aligned}\Delta \ln FFA_t = \beta_1 \Delta \underset{Ballast}{\ln} Wind_{t-1}^{Force^+} + \beta_2 \Delta \underset{Ballast}{\ln} Wind_{t-1}^{Speed} \\ + \beta_3 \Delta \underset{Ballast}{\ln} Speed_{t-1} + \beta_4 \Delta \underset{Laden}{\ln} Speed_{t-1} \\ + \beta_5 \Delta \ln Bunker_{t-1} + \beta_6 \Delta \ln FFA_{t-1} + u_t\end{aligned} \tag{6.2}$$

We again fit the model using OLS regression. The regression results in Table 6.2 show the FFA rate increases after an increase in positive wind force on vessels in ballast. None of the speed variables are significant, which is in line with our expectations from Section 2.2. Positive wind force is the only significant variable apart from the lagged dependent variable.

In economic terms, for an increase in the positive wind force on vessels in ballast by one standard deviation from average levels holding everything else equal, the FFA rate is expected to increase by 0.6%.[9] At the average level of FFA rates of \$13,099/day, this leads to an expected increase of \$78.6/day. For 100 lots traded this amounts to an increase in price of \$7,860.

[9] At average levels of 0.20 in positive wind force and a standard deviation of 0.04, the expected change in the FFA rate is $0.033 \ln(\frac{0.20+0.04}{0.20}) = 0.006$ using model 7 from Table 6.2.

Table 6.2: Predictive regression on FFA rates

	Response variable: $\Delta \ln FFA_t$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta \ln Wind^{Force^+}_{t-1}$ *Ballast*	0.028**			0.035**	0.026*		0.033**
	(0.014)			(0.015)	(0.014)		(0.015)
$\Delta \ln Wind^{Speed}_{t-1}$ *Ballast*	0.003			-0.004	0.002		-0.005
	(0.024)			(0.024)	(0.024)		(0.025)
$\Delta \ln Speed_{t-1}$ *Ballast*		-0.069		-0.163		-0.059	-0.15
		(0.109)		(0.117)		(0.109)	(0.118)
$\Delta \ln Speed_{t-1}$ *Laden*			-0.23		-0.171	-0.225	-0.15
			(0.147)		(0.149)	(0.147)	(0.15)
$\Delta \ln Bunker_{t-1}$	0.057	0.054	0.063	0.058	0.064	0.063	0.064
	(0.105)	(0.105)	(0.105)	(0.105)	(0.105)	(0.105)	(0.105)
$\Delta \ln FFA_{t-1}$	0.171***	0.173***	0.171***	0.171***	0.17***	0.171***	0.17***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Observations	1239	1239	1239	1239	1239	1239	1239
R^2	0.035	0.031	0.033	0.036	0.036	0.033	0.037
Adjusted R^2	0.032	0.029	0.03	0.033	0.032	0.03	0.033
F Statistic	11.165***	13.209***	13.915***	9.325***	9.197***	10.504***	7.937***
BG LM Statistic	30.176	30.364	30.742	29.187	30.218	30.113	29.4
BP LM Statistic	10.927*	10.324	10.604	10.617	10.932*	10.478	10.675*
KPSS Residuals	0.35	0.33	0.34	0.34	0.35	0.33	0.34
Note:							*p<0.1; **p<0.05; ***p<0.01

This table shows the regression results of the FFA predictive model. Standard errors of coefficients are in parenthesis. FFA is the 1-month 5 time charter average forward price in \$/day. All variables are log-differenced and FFA contracts are rolled over on the last trading day of the month. BG LM is the Breusch–Godfrey (1978; 1978) Lagrange multiplier χ^2 test for serial correlation in the residuals. The null hypothesis of BG LM test states that the residuals are not serially correlated. BP LM is the Breusch–Pagan (1979) Lagrange multiplier χ^2 test for contemporaneous heteroskedasticity. The null hypothesis of the BP LM test states that the residuals are contemporaneously homoskedastic. KPSS is the Kwiatkowski et al. (1992) unit root test. The null hypothesis of the KPSS unit root test states the residuals are weakly stationary. Variance inflation factors (VIF) for the explanatory variables are $\text{VIF}(\Delta \ln \underset{ballast}{Wind^{Force^+}_{t-1}}) = 1.43$, $\text{VIF}(\Delta \ln \underset{ballast}{Wind^{Speed}_{t-1}}) = 1.26$, $\text{VIF}(\Delta \ln \underset{ballast}{Speed_{t-1}}) = 1.17$, $\text{VIF}(\Delta \ln \underset{laden}{Speed_{t-1}}) = 1.06$, $\text{VIF}(\Delta \ln Bunker_{t-1}) = 1.02$ and $\text{VIF}(\Delta \ln FFA_{t-1}) = 1.05$.

6.3 Economic implications

There is no economic reason why positive wind force should directly drive freight rates. Market expectations are one of the key drivers of rates in the very short-run market (Stopford, 2009), and expectations are partially formed by participants monitoring the supply through AIS data. Rational market participants should only adjust their expectations if an observed speed increase is a conscious decision. We have proposed a measure which extracts the effect of wind on vessel speeds, which should not impact market expectations in itself.

In the spot market, our measure helps shipowners and charterers make better pricing decisions. Since changes in speed caused by positive wind force is not an active speed decision, the market participants should factor out this effect on speed when negotiating voyage rates.

Our measure helps shipowners and charterers hedge more efficiently with FFA contracts against the highly volatile capesize market. Increased hedging performance is especially relevant in a market where route-specific FFAs are illiquid, and shipowners and charterers have to make due with only the 5 T/C average settled FFA contract. Hedging exposure to specific routes by using a composite FFA contract introduces basis risk, as the instrument used for hedging has a different underlying (Alizadeh and Nomikos, 2009). More accurate pricing models makes hedging cheaper and more effective, so any new effective information should be beneficial to the hedging performance. In particular, improved predictions of future cash flows in both the spot and FFA markets could improve the accuracy of calculated hedge ratios.

For speculative traders in the FFA market, our measure can be incorporated into trading strategies until any potential benefits are traded away. The usual transaction cost of an FFA contract is 0.25% of the contract value (Nomikos and Doctor, 2013), and we have shown that positive wind force can cause increases in rate of beyond this transaction cost. For any change in positive wind force greater than 0.0157 at average levels,[10] the change in rate will be greater than the transaction cost, holding everything else equal.

[10] Solving $0.0025 = 0.033 \ln(\frac{0.20+\Delta}{0.20})$ for Δ gives us the minimum change in positive wind force required for exceeding the transaction cost of 0.25%.

Implementation costs of advanced models can often exceed the potential benefits. This should not be the case here, as market participants should already have access to AIS and wind forecast data. AIS data is available in real time from several commercial actors, and wind forecasts are made available four times a day from ECMWF (2018b). Our positive wind force measure is computationally cheap, and implementing it is straightforward; a modern laptop computer can be used to process the AIS data and wind forecasts.

6.4 Predictive power of speed control variables

Previous research on AIS data in shipping markets has revealed that speed holds significant predictive information on spot freight rates (Regli and Nomikos, 2019), but we are unable to find any conclusive significant effects of speed in our regressions. There is, however, negative correlation between the speed of laden vessels against the future C5 rate in models 3 and 6 shown in Table 6.1. In economic terms, the effect of a one standard deviation decrease in speed of laden vessels, results in an increase of 0.3% in freight rates.[11] In a Ronen (1982) framework, where laden vessels choose their speeds to maximize daily revenue, the negative relationship may be explained by laden vessels expecting C5 rates to increase, as explained in Section 2.2.

We do not see any evidence of speed of vessels in ballast predicting future C5 rates. To the best of our knowledge, no research has been conducted on speed measures from AIS data on daily sampled data in the capesize market, so there might not be observable effects in daily sampled data. Nevertheless, AIS-derived supply measures are not a novel idea, and it is reasonable to believe that market practitioners already monitor them. Any predictive ability they may have on next-day rates should be incorporated in the FFA rate, as the AIS speed information is available to market participants.

[11] At average levels of 10.58 knots in speed of laden vessels and a one standard deviation decrease of 0.19 knots in speed, the expected change in the C5 spot rate is $-0.18 \ln(\frac{10.88-0.19}{10.88}) = 0.003$ using model 6 from Table 6.1.

6.5 Positive wind force measure validation

We validate the positive wind force measure for not being a proxy for apparent wind speed in two ways: We find that increases in positive wind force can partially explain increases in speed. We also check for causal relationships between positive wind force and apparent wind speed, and future freight rates. Causality is only found between positive wind force and future freight rates.

6.5.1 Speed regression

We validate the positive wind force measure's effect on the speed of vessels, to make sure it is a valid proxy for wind pushing the vessels forwards. Increases in speed should be partially explained by increases in positive wind force and vice versa, for our arguments in Section 2.3 to hold. The results of the speed regression show increases in speed can be partially explained by increases in positive wind force.

We define a model for the fleet's speed at a given time t in Equation 6.3. The change in fleet speed is determined by the change in positive wind force, controlled for the apparent wind speed. The error term contains the speed decision of the vessels and all other determinants of speed. The control for apparent wind speed is, as mentioned, a proxy for sea conditions. Vessels will have to involuntarily slow down when apparent wind speed increases, and sea conditions worsen (Prpić-Oršić et al., 2014).

$$\Delta \ln Speed_t = \beta_1 \Delta \ln Wind_t^{Force^+} + \beta_2 \Delta \ln Wind_t^{Speed} + u_t \tag{6.3}$$

The regression results in Table 6.3 show there is a significant effect of positive wind force on the speed of vessels for both loading conditions. Controlling for apparent wind in models 3a and 3b further increases the coefficient of positive wind force. For a one standard deviation increase in positive wind force, the speed of a vessel in ballast increases by 0.84% and the speed of a laden vessel increases by 0.46%.[12] At average levels of 12.06

[12] At average levels of 0.20 in positive wind force for vessels in ballast and a standard deviation of 0.04, the expected change in speed is $0.046 \ln(\frac{0.20+0.04}{0.20}) = 0.0084$. At average levels of 0.09 in positive wind force for laden vessels and a standard deviation of 0.02, the expected change in speed is $0.023 \ln(\frac{0.09+0.02}{0.09}) = 0.0046$.

kn for vessels in ballast and 10.88 kn for laden vessels this amounts to an increase in speed of 0.10 kn for vessels in ballast and 0.05 kn for laden vessels. The regressions also confirm the suspected negative correlation between apparent wind speed and the speed of vessels. These findings support our claim of positive relationship between the positive wind force

Table 6.3: Speed regression

Response variable:	$\Delta \ln Speed_t$ *Ballast*			$\Delta \ln Speed_t$ *Laden*		
	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
$\Delta \ln Wind_t^{Force^+}$	0.036***		0.046***	0.016***		0.023***
	(0.003)		(0.003)	(0.002)		(0.002)
$\Delta \ln Wind_t^{Speed}$		-0.012**	-0.045***		-0.036***	-0.053***
		(0.005)	(0.005)		(0.004)	(0.004)
Observations	1813	1813	1813	1813	1813	1813
R^2	0.101	0.003	0.14	0.057	0.045	0.146
Adjusted R^2	0.1	0.003	0.139	0.056	0.045	0.145
F Statistic	152.157***	5.133**	112.454***	82.603***	76.012***	132.429***
Note:						*p<0.1; **p<0.05; ***p<0.01

This table shows regressions of changes in wind affecting vessels and their speed seperated by their loading condition. Standard errors in parenthesis. The standard errors are heteroskedasticity and autocorrelation robust as in Newey and West (1994) with 8 lags. *Speed* is the mean speed of moving vessels in knots where $Speed > 6$ kn. $Wind^{Force^+}$ is the mean positive wind force per vessel dwt capacity in N/t. $Wind^{Speed}$ is the mean apparent wind speed per vessel in knots.

measure and speed as argued in section 2.3. We find an R^2 of 10% for vessels in ballast and 6% for laden vessels without controlling for apparent wind. When controlling for apparent wind, we find an R^2 of 14% for vessels in ballast, and 15% for laden vessels.

Even though these findings suggest that changes in positive wind force can explain changes in speed for laden vessels as well, we are unable to use this in our predictive freight rate regressions. A possible reason for this failure might be that the explanatory power of the positive wind force on speed of laden vessels is lower, but this is in line with our expectations in Section 2.3. Predictive regression models including these positive wind measures for laden vessels are provided in the appendix in Table A3.1.

6.5.2 Wind causality on spot and FFA rates

We perform the Granger (1969) causality test on our positive wind force measure and apparent wind speed against freight rates. The results of the causality test show there is a causal relationship between our positive wind force measure and freight rates. At the same time, there is no causal relationship between the apparent wind speed and freight rates, which improves the validity of our novel positive wind force measure.

Table 6.4: Wind force Granger causality test results

Causal relationship			χ^2	ML	F	Parameter F
$\Delta \ln Wind^{Force^+}_{t-1,\ Ballast}$	$\Longrightarrow$	$\Delta \ln C5_t$	10.83***	10.78***	10.80***	10.80***
$\Delta \ln Wind^{Speed}_{t-1,\ Ballast}$	$\Longrightarrow$	$\Delta \ln C5_t$	7.30*	7.27*	7.28*	7.28*
$\Delta \ln Wind^{Force^+}_{t-1,\ Ballast}$	$\Longrightarrow$	$\Delta \ln FFA_t$	4.96**	4.95**	4.94**	4.94**
$\Delta \ln Wind^{Speed}_{t-1,\ Ballast}$	$\Longrightarrow$	$\Delta \ln FFA_t$	1.20	1.20	1.20	1.20

Note: * p<0.1; ** p<0.05; *** p<0.01

This table shows the test results of the Granger (1969) causality test with 1 lag of positive wind force and apparent wind speed on vessels in ballast on $C5$, the C5 route spot price and FFA, the one-month FFA rate. The alternative hypothesis of the Granger causality test state that past values of a time series X Granger-causes changes in future values of time series Y. The critical values for the χ^2- and maximum likelihood tests with 1 degree of freedom are 6.63 at 1% level, 3.84 at 5% level and 2.71 at 10% level. The critical values for the F-tests with 1 degree of freedom in the nominator and 1236 degrees of freedom in the denominator are 5.02 at 1% level, 3.84 at 5% level and 2.70 at 10% level.

We check if changes in positive wind force on vessels in ballast actually causes the changes in future freight rates by performing the Granger (1969) causality test. The alternative hypothesis of the Granger causality test states that past values in time series X Granger-causes future values in time series Y. There can however exist a confounding time series Z which causes future changes in both time series X and Y. This makes it difficult to establish an actual causal relationship from the Granger causality test alone.

We see from the results of the tests in Table 6.4, that changes in positive wind forces on vessels in ballast Granger-causes changes in both the C5 spot rate and the FFA rate. We also see that apparent wind speed does not Granger-cause rates. The causal relationship between apparent wind speed and the C5 spot rate is rejected at 5% level with a 7%

p-value for all test statistics.

There is no way for individual vessels to affect their own positive wind force, as long as vessels act rationally and optimize their voyage routes (Bialystocki and Konovessis, 2016) with regards to local wind conditions. It is unlikely that a confounding time series exists which drives both positive wind forces on vessels in ballast and the freight rates, since there is no way for vessels to impact their own positive wind force other than acting irrationally.

6.6 OLS model validation

The freight rate regression models, model 7 from both Table 6.1 and 6.2 are validated against the five assumptions for valid time series OLS estimates as in Wooldridge (2015). All the log-differenced variables are stationary as shown in Table 5.1. The residuals are also stationary according to the Kwiatkowski et al. (1992) (KPSS) unit root test for both of the models. There is no perfect collinearity between any of the explanatory variables, supported by the variance inflation factors. The conditional means of the residuals are zero, as shown in the appendix in Figure A4.1 and A4.2. The residuals are contemporaneously homoskedastic, confirmed by the Breusch–Pagan (1979) Lagrange multiplier χ^2 (BP LM) test. For the FFA rate regression, the BP LM test is rejected at 5% level with a p-value of 7.2%. There is no serial correlation between the residuals, confirmed by the Breusch–Godfrey (1978; 1978) Lagrange multiplier χ^2 (BG LM) test. For the C5 spot rate regression, the BG LM test is rejected at 5% level with a p-value of 7.5%.

7 Conclusion

In this thesis we show how wind forces pushing vessels forwards help predict the C5 spot rate, and the one-month 5 time charter average FFA rate in daily sampled data. These predictions are made possible by our novel positive wind force measure. The measure helps shipowners and charterers make better pricing decisions in the spot market, and hedge their positions better in the FFA market. In the FFA market, the measure also helps other traders improve their trading strategies.

Our estimated models in Section 6.1 and 6.2 have some limitations. We assume the AIS-messages picked up by the satellites are random both on a vessel and geographical level. The wind data is from an 06:00 UTC immediate forecast and the same forecast is used if a vessel is observed right after midnight or just before midnight. Vessels are therefore likely to be somewhat distanced from the local wind forecast in both time and distance.

The NCA has launched multiple new AIS satellites in our sample period, increasing the number of AIS-messages collected, in turn improving AIS data quality over time. Additionally, vessels demolished during our sample period do not appear in the data sample, since the fleet register by Clarksons Research (2019b) only contains active vessels at the time of download. The missing observations from August 4 to August 14, 2017, can also affect the residual validation tests of our regression models.

The success of our positive wind force measure suggests there is a measurable relationship between a vessel's speed over ground, and the chosen speed of a vessel through water. Uncovering the dynamics of this relationship would allow market participants to model supply more accurately, and we suggest this as a topic for further research.

References

Adland, R. and Alizadeh, A. H. (2018). Explaining price differences between physical and derivative freight contracts. *Transportation Research Part E: Logistics and Transportation Review*, 118:20–33.

Alizadeh, A. and Nomikos, N. (2009). *Shipping derivatives and risk management*. Springer.

Aßmann, L. M., Andersson, J., and Eskeland, G. S. (2015). Missing in action? speed optimization and slow steaming in maritime shipping. *NHH Dept. of Business and Management Science Discussion Paper*.

Baltic Exchange Ltd (2014). Baltic exchange to implement new capesize vessel description & routes. https://www.balticexchange.com/news/press-announcements/article/baltic-exchange-to-implement-new-capesize-vessel-description-routes/2874/. Press release. May 1, 2019.

Baltic Exchange Ltd (2019). Forward freight agreement price curves from 2014 to 2018. Unpublished dataset.

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society*, 8(1):27–41.

Batchelor, R., Alizadeh, A., and Visvikis, I. (2007). Forecasting spot and forward prices in the international freight market. *International Journal of Forecasting*, 23(1):101–114.

Bialystocki, N. and Konovessis, D. (2016). On the estimation of ship's fuel consumption and speed curve: A statistical approach. *Journal of Ocean Engineering and Science*, 1(2):157–166.

Blendermann, W. (1994). Parameter identification of wind loads on ships. *Journal of Wind Engineering and Industrial Aerodynamics*, 51(3):339–351.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17(31):334–355.

Breusch, T. S. and Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, pages 1287–1294.

Cavcar, M. (2000). The international standard atmosphere (isa). *Anadolu University, Turkey*, 30:9.

Clarksons Research (2019a). Shipping intelligence network time series. Unpublished dataset.

Clarksons Research (2019b). World fleet register. Unpublished dataset.

Copernicus Climate Change Service (C3S) (2017). Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate. https://cds.climate.copernicus.eu/cdsapp#!/home. Accessed on March 1, 2019.

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, pages 1057–1072.

European Centre for Medium-Range Weather Forecasts (ECMWF) (2018a). eccodes home - eccodes - ecmwf confluence wiki. https://confluence.ecmwf.int/display/ECC/ecCodes+Home. Accessed on February 14, 2019.

European Centre for Medium-Range Weather Forecasts (ECMWF) (2018b). Set i - atmospheric model high resolution 10-day forecast (hres). https://www.ecmwf.int/en/forecasts/datasets/set-i. Accessed on March 25, 2019.

European Centre for Medium-Range Weather Forecasts (ECMWF) (2019). *Access MARS - ECMWF Web API*. Accessed on March 18, 2019.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica: Journal of the Econometric Society*, pages 1293–1301.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438.

International Trade Centre (2019). Trade map - trade statistics for international business development. https://www.trademap.org/Index.aspx. Accessed on May 5, 2019.

International Maritime Organization (IMO) (2019). Automatic identification systems (ais). http://www.imo.org/en/OurWork/safety/navigation/pages/ais.aspx. Accessed on April 28, 2019.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: journal of the Econometric Society*, pages 1551–1580.

Kitamura, F., Ueno, M., Fujiwara, T., and Sogihara, N. (2017). Estimation of above water structural parameters and wind loads on ships. *Ships and Offshore Structures*, 12(8):1100–1108.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and 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? *Journal of econometrics*, 54(1-3):159–178.

Newey, W. K. and West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4):631–653.

Nomikos, N. K. and Doctor, K. (2013). Economic significance of market timing rules in the forward freight agreement markets. *Transportation Research Part E: Logistics and Transportation Review*, 52:77–93.

Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

Plosser, C. I. and Schwert, G. W. (1978). Money, income, and sunspots: measuring economic relationships and the effects of differencing. *Journal of Monetary Economics*, 4(4):637–660.

Prpić-Oršić, J., Vettor, R., Faltinsen, O. M., and Guedes Soares, C. (2014). Influence of ship routes on fuel consumption and co2 emission. In *The 2nd Congress Maritime Technology and Engineering-MARTECH 2014*, pages 857–864.

Raymond, E. S. (2016). Aivdm/aivdo protocol decoding. http://catb.org/gpsd/AIVDM.html. Accessed on February 14, 2019.

Regli, F. and Nomikos, N. K. (2019). The eye in the sky–freight rate effects of tanker supply. *Transportation Research Part E: Logistics and Transportation Review*, 125:402–424.

Ronen, D. (1982). The effect of oil price on the optimal speed of ships. *Journal of the Operational Research Society*, 33(11):1035–1040.

Schwehr, K. (2018). *Libais: C++ decoder for Automatic Identification System for tracking ships and decoding maritime information.* Accessed on February 14, 2019.

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

Watson, D. G. (2002). *Practical ship design*, volume 1. Elsevier.

Whitaker, J. (2019). *Pygrib: Python interface for reading and writing GRIB data.* Accessed on March 18, 2019.

Wooldridge, J. M. (2015). *Introductory Econometrics: A Modern Approach.* Cengage Learning.

Appendix

A1 Variable correlation matrix

Table A1.1: Variable correlation matrix

	$Wind^{Force^+}_{Ballast}$	$Wind^{Speed}_{Ballast}$	$Wind^{Force^+}_{Laden}$	$Wind^{Speed}_{Laden}$	$Speed_{Ballast}$	$Speed_{Laden}$	$Bunker$	FFA
$Wind^{Speed}_{Ballast}$	0.41							
$Wind^{Force^+}_{Laden}$	-0.35	0.12						
$Wind^{Speed}_{Laden}$	0.22	0.61	0.30					
$Speed_{Ballast}$	0.30	-0.07	-0.13	-0.07				
$Speed_{Laden}$	-0.19	-0.13	0.24	-0.24	0.06			
$Bunker$	-0.01	-0.02	0.00	-0.03	0.01	0.06		
FFA	0.03	0.00	-0.05	0.00	0.02	-0.03	0.13	
$C5$	-0.02	0.00	-0.01	0.01	0.00	-0.01	0.11	0.44

Note: All variables in log-differences

This table shows the correlation matrix between all log-differenced variables. $Wind^{Force^+}$ is the mean positive wind force per vessel dwt capacity in N/t. $Wind^{Speed}$ is the mean apparent wind speed on vessels in knots. $Speed$ is the mean speed of moving vessels in knots where $Speed > 6$kn. The $Wind$ and $Speed$ variables are separated by the vessels' loading condition. $Bunker$ is the bunker fuel oil price in \$/t. FFA is the one-month 5 time charter average forward price in \$/day. $C5$ is the C5 route spot price in \$/t.

A2 Autocorrelation function plots

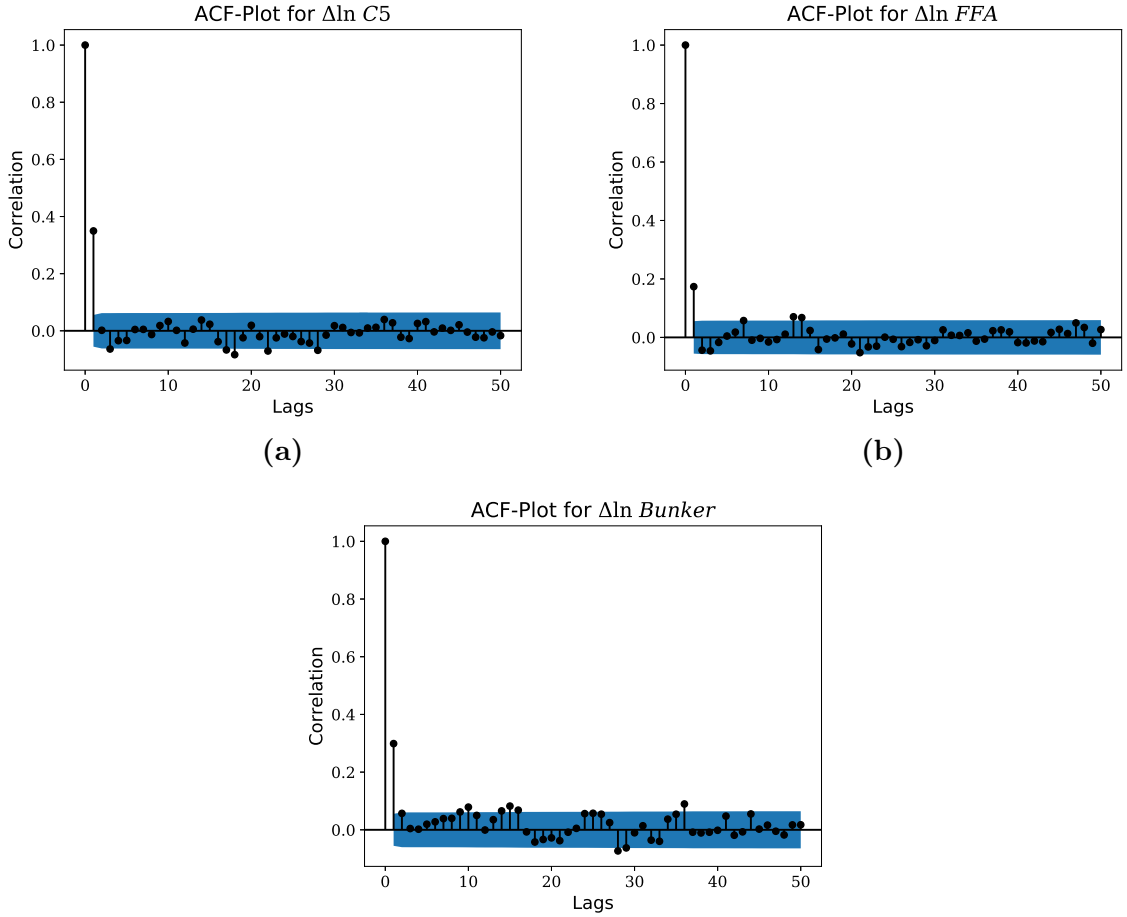


Figure A2.1: ACF-plots for the log-differenced time series

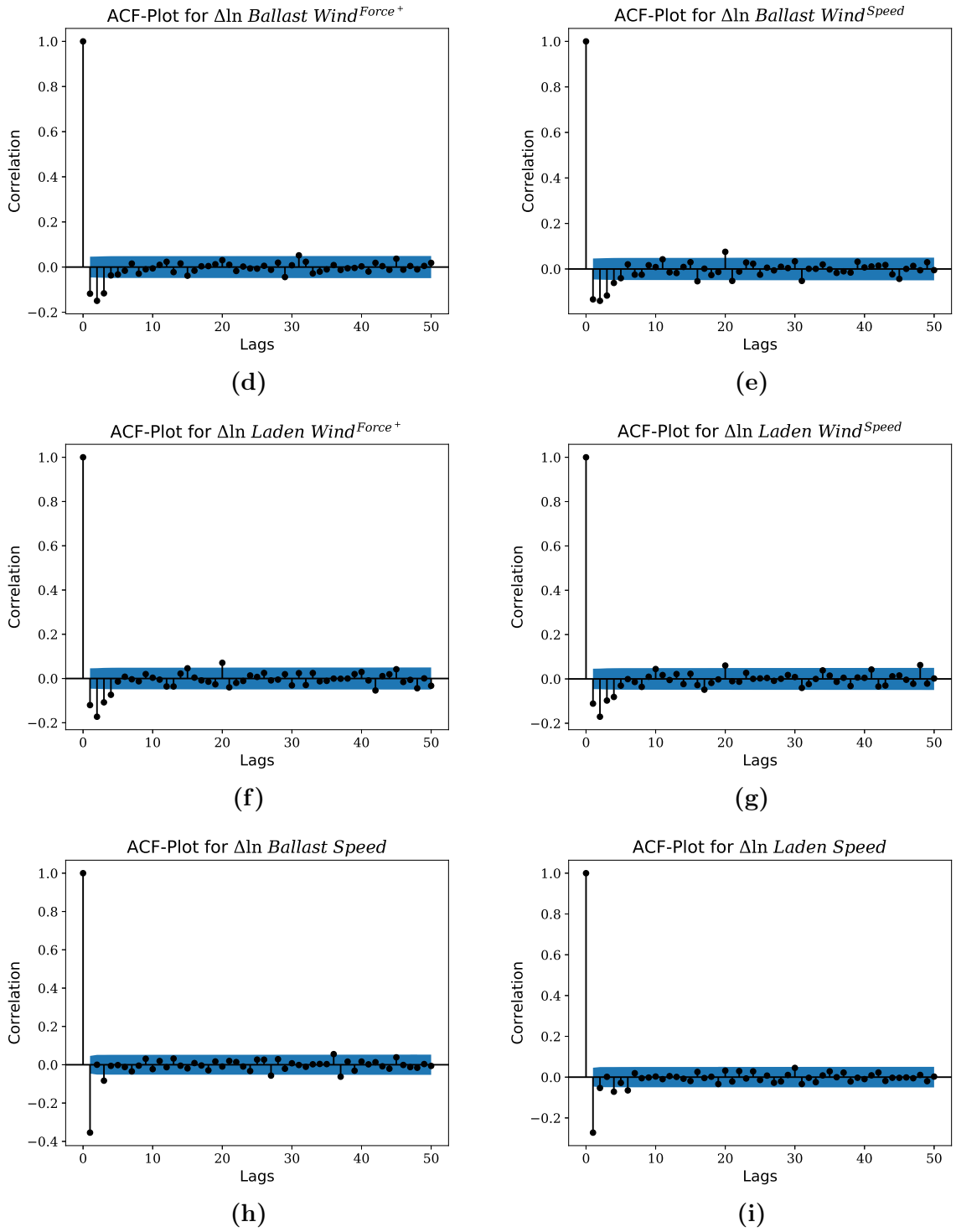


Figure A2.1: ACF-plots for the log-differenced time series (cont.)

This figure shows the autocorrelation function plots for the log-differenced sample data. 95% confidence bands are calculated using Bartlett's (1946) formula. The speed variables (A2.1h and A2.1i) show weak signs of over-differencing. The wind variables (A2.1d, A2.1e, A2.1f and A2.1g) show minimal signs of over-differencing. $C5$ (A2.1a), FFA (A2.1b) and *Bunker* (A2.1c) show no signs of over-differencing.

A3 Additional predictive freight rate regressions

Table A3.1: Additional predictive freight rate regressions

Response variable:	$\Delta \ln C5_t$	$\Delta \ln FFA_t$
$\Delta \ln Wind^{Force^+}_{t-1}$ *Laden*	-0.002	-0.007
	(0.012)	(0.007)
$\Delta \ln Wind^{Speed}_{t-1}$ *Laden*	0.021	0.0
	(0.028)	(0.016)
$\Delta \ln Speed_{t-1}$ *Ballast*	-0.042	-0.069
	(0.11)	(0.062)
$\Delta \ln Speed_{t-1}$ *Laden*	-0.143	-0.2
	(0.161)	(0.09)
$\Delta \ln Bunker_{t-1}$	0.085	0.063
	(0.105)	(0.059)
$\Delta \ln FFA_{t-1}$	0.234***	0.171***
	(0.028)	(0.018)
$\Delta \ln C5_{t-1}$	0.182***	
	(0.028)	
Observations	1239	1239
R^2	0.241	0.033
Adjusted R^2	0.236	0.029
F Statistic	55.774***	7.062***
BG LM Statistic	32.125*	29.922
BP LM Statistic	10.12	11.59*
KPSS Residuals	0.02	0.33
Note:	*p<0.1; **p<0.05; ***p<0.01	

This table shows the predictive freight rate regressions with positive wind force for laden vessels included. Standard errors of coefficients are in parenthesis. $C5$ is the C5 route spot price for in \$/ton. FFA is the one-month 5 time charter average forward price in \$/day. All variables are log-differenced and FFA contracts are rolled over on the last trading day of the month. BG LM is the Breusch–Godfrey (1978; 1978) Lagrange multiplier χ^2 test for serial correlation in the residuals. The null hypothesis of BG LM test states that the residuals are not serially correlated. BP LM is the Breusch–Pagan (1979) Lagrange multiplier χ^2 test for contemporaneous heteroskedasticity. The null hypothesis of the BP LM test states that the residuals are contemporaneously homoskedastic. KPSS is the Kwiatkowski et al. (1992) unit root test. The null hypothesis of the KPSS unit root test states the residuals are weakly stationary.

A4 Regression diagnostic plots

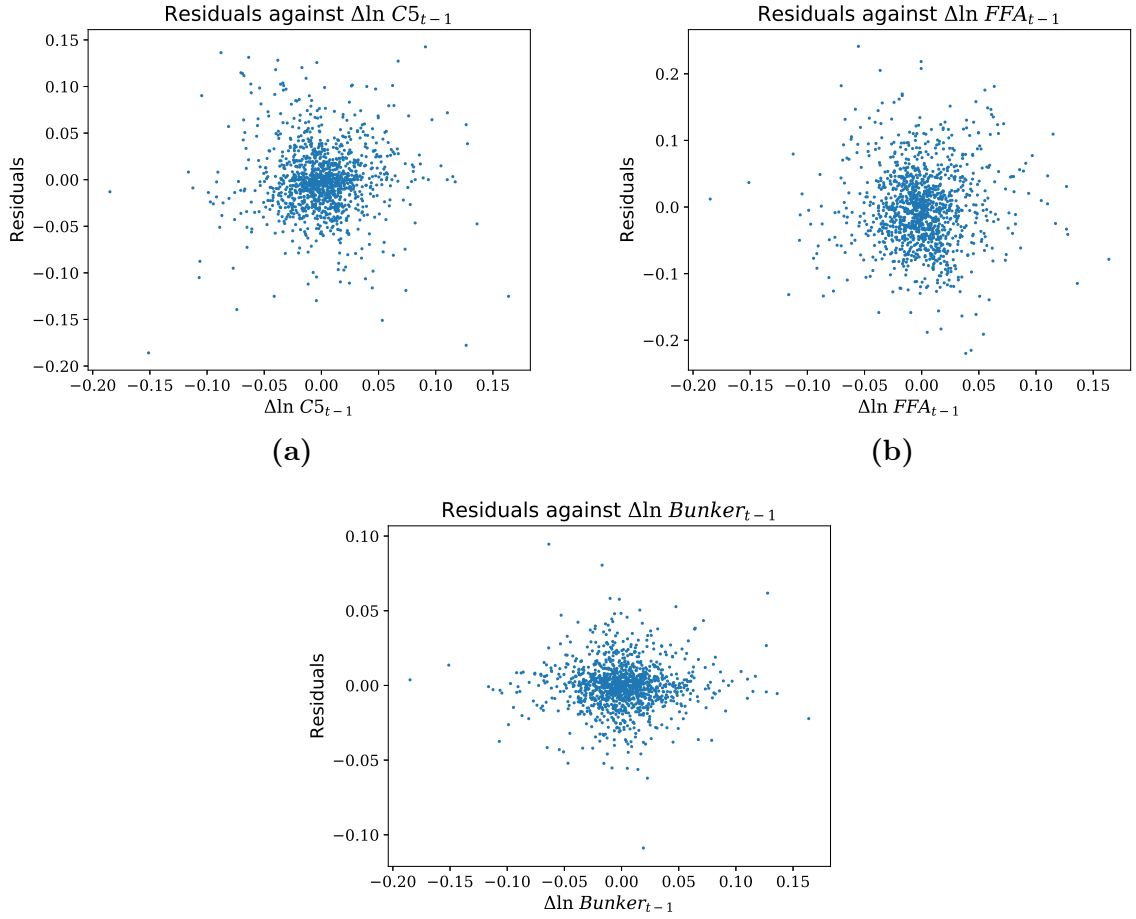


Figure A4.1: Residual plots for the C5 regression model 7 in Table 6.1

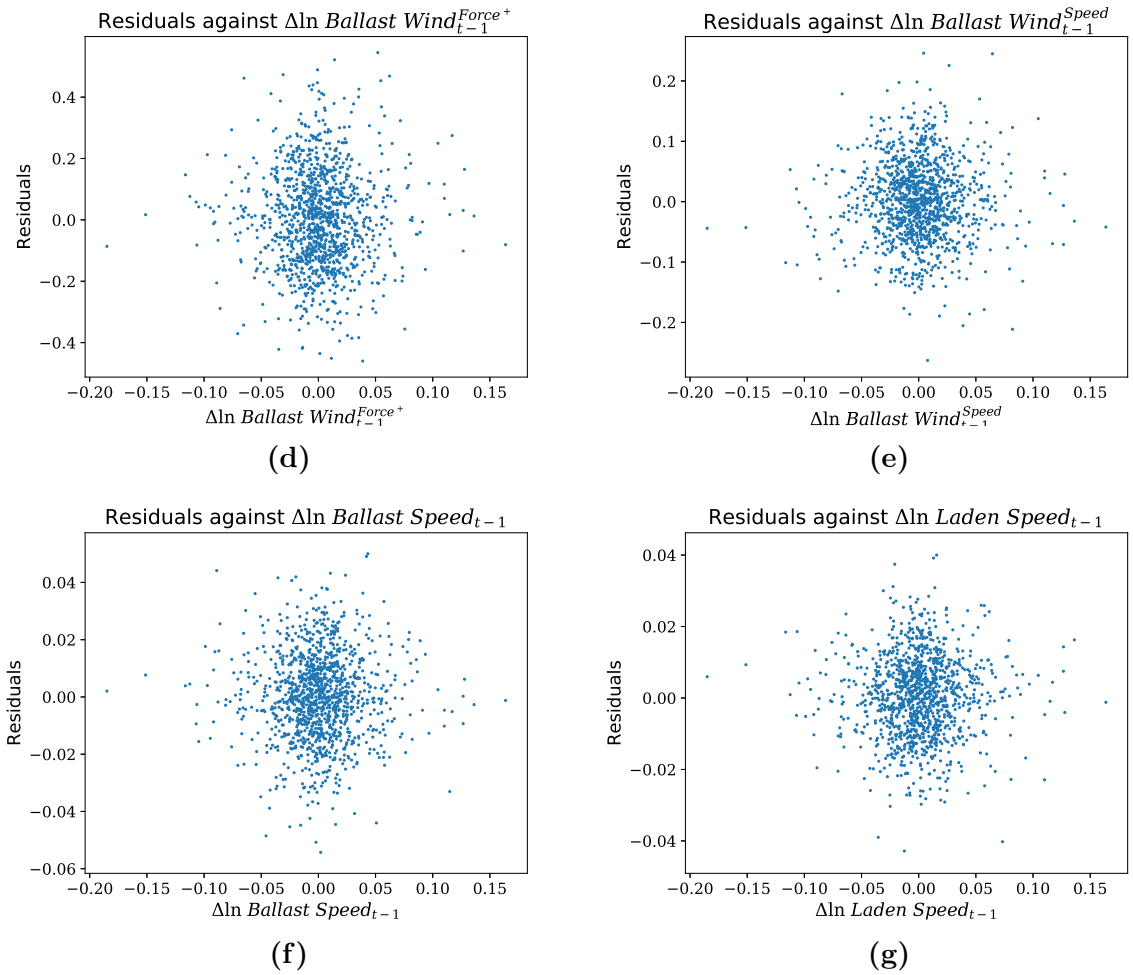


Figure A4.1: Residual plots for the C5 regression model 7 in Table 6.1 (cont.)

This table shows the residuals of model 7 in Table 6.1 plotted against the model's explanatory variables. The plots show no deviance from the OLS assumption of zero conditional mean.

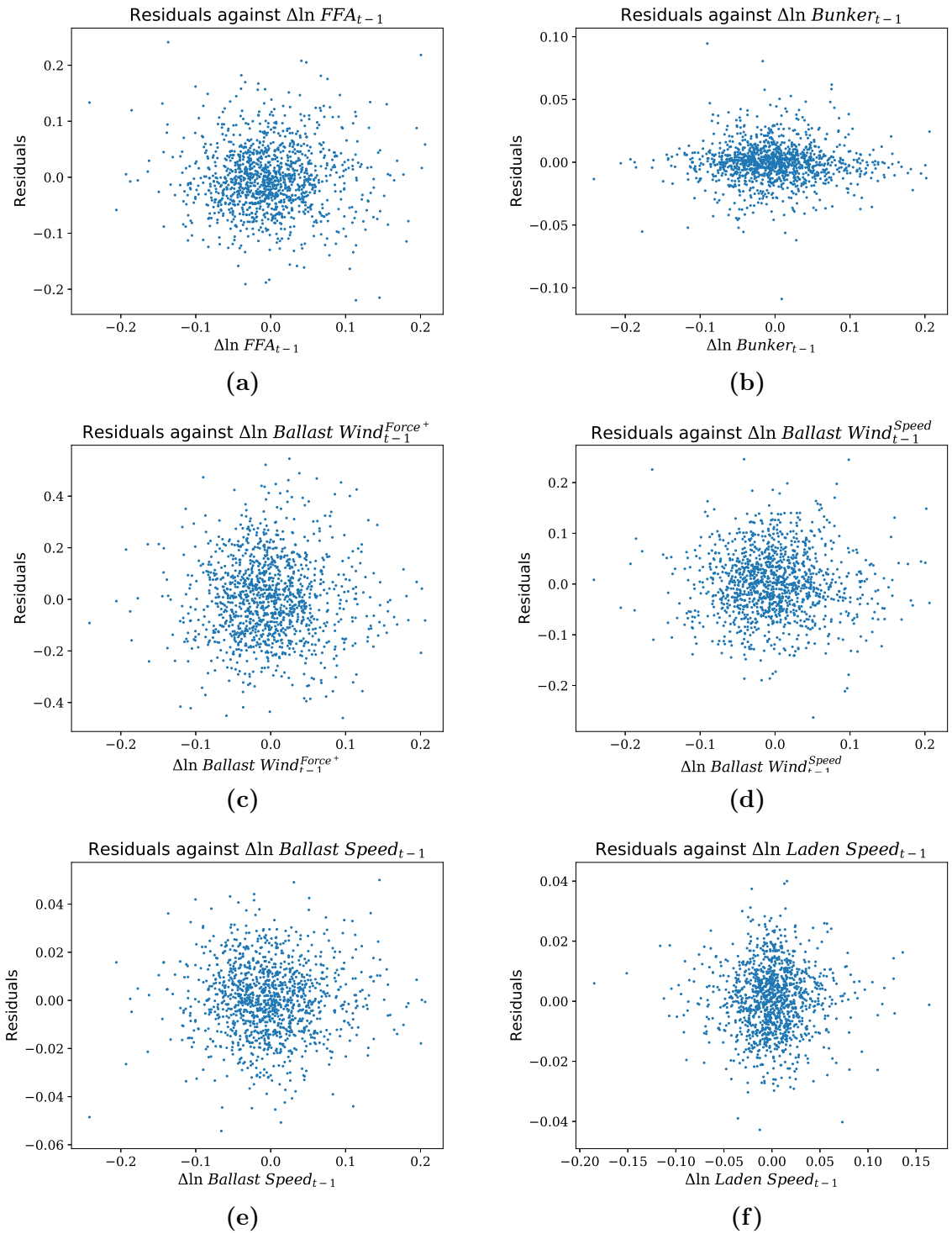


Figure A4.2: Residual plots for the FFA regression model 7 in Table 6.2

This table shows the residuals of model 7 in Table 6.2 plotted against the model's explanatory variables. The plots show no deviance from the OLS assumption of zero conditional mean.