The Cost of Weather: Modeling Weather Delay in Bulk Shipping

BY Vegard Enerstvedt

DISCUSSION PAPER

Institutt for foretaksøkonomi
Department of Business and Management Science

FOR 4/2025

ISSN: 2387-3000
February 2025

The Cost of Weather: Modeling Weather Delay in Bulk Shipping

Vegard Enerstvedt[1*]

[1]Department of Business and Management Science, Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway.

[*]Corresponding author: Vegard.Enerstvedt@nhh.no

February 3, 2025

Abstract

Weather is an ever-present factor influencing shipping operations at every stage, including port operations. This paper examines the determinants of weather-induced delays in port operations, the probability and duration of such delays, the predictive capability of various statistical models, and the potential for improving upon standard industry methods for estimating port margins. A wide range of models are investigated, including Generalized Linear (GLM) models, Cox Propotional Hazard models, and Autoregressive Conditional Duration (ACD) models. The findings reveal that a GLM with gamma distributed dependent variables provides the best fit for data on delay duration, while a linear multiple regression offers the highest predictive accuracy for delay duration. Similarly, probit and logit models are found to perform comparably well for both predicting delay probabilities and data fit. Moreover, the analysis demonstrates that there are significant potential cost savings when using

a linear regression model with a probit model to predict delays compared to a common industry rule-of-thumb of half a day delay. These results underscore the potential for improving operational efficiency and accuracy in port margin estimation through statistical modeling techniques.

Keywords: Weather, Shipping, Risk, Delay, Ports, ACD, GLM

JEL Classification: C40, C41, C53, R40, R41

1 Introduction

Weather is a critical factor in the shipping industry, influencing both maritime and port operations. If not adequately accounted for, adverse weather conditions can lead to significant financial costs. Rainstorms, for example, may compromise the integrity of cargo, while high winds can render the operation of cranes and other port infrastructure hazardous. In such instances, operations must be suspended until conditions improve, resulting in delays. In the context of a low-margin industry such as shipping, where operational efficiency is paramount, these weather-related risks must be carefully considered when fixing and scheduling vessel movements.

Precipitation presents a major challenge when transporting weather-sensitive cargo, such as cotton, which can deteriorate in quality if exposed to moisture. Similarly, paper-based products, such as cardboard, are highly susceptible to rain, often losing their structural integrity upon contact with moisture. Such degradation not only affects the recipient, who receives damaged goods, but also the sender, who may be held liable for the compromised quality, and the vessel operator, who may also face liability for damages incurred during transport. To mitigate these risks, loading and unloading operations are frequently halted during periods of precipitation, particularly when the cargo is vulnerable. While such operational suspensions help to preserve cargo integrity, they nonetheless

impose financial costs in the form of delays.

Strong winds also pose operational risks. When high winds occur during cargo loading or unloading, unsecured items may shift or fall, creating safety hazards for both the crew and port workers. In these circumstances, operations must be paused, and the cargo must be secured, leading to additional delays caused by adverse weather conditions.

Weather-related delays carry substantial financial implications. During such delays, vessels are unable to generate revenue, even though fixed costs—such as crew wages, insurance, fuel, and port fees—continue to accrue. As weather conditions persist, so do the financial outflows, while income remains halted. It is these financial risks associated with weather disruptions that form the central focus of this paper.

We addresses two key dimensions of weather delays: duration and probability. Although these factors are distinct, they are interrelated and jointly influence decision-making by vessel operators. For example, a port characterized by infrequent but prolonged rain delays may represent an acceptable risk for non-urgent cargo. Conversely, a port that experiences frequent, shorter delays may present a higher risk for time-sensitive shipments, despite the relatively brief interruptions. To analyze this, we apply various statistical techniques, including linear regression, survival analysis, and generalized linear models (GLMs), to model the duration and probability of delays. Furthermore, we compare the performance of these models against the prevalent industry assumption, which typically accounts for a half-day delay during port calls.

This paper makes several contributions to the existing body of literature by addressing five key research questions: (1) What factors influence the duration of weather delays? (2) What factors determine the probability of a delay? (3) Which model performs best in predicting delay duration? (4) Which model

is most effective in predicting the likelihood of a delay? (5) How do these models compare to the industry-standard rule of thumb, which assumes a half-day delay? To the best of our knowledge, this is the first comprehensive analysis of the implications of weather-related delays for vessel operators at ports in terms of time and money. Moreover, access to a novel dataset provided by a vessel operator enables the development of a uniquely realistic model, yielding insights previously unavailable in the literature.

Our analysis demonstrates that by employing a combination of generalized linear models (GLMs), linear regression, and statistical techniques for cross-validation and variable selection, we achieve significantly more accurate predictions of delay durations compared to the industry's standard rule-of-thumb. We find that a linear regression model incorporating climate, seasonal, and congestion-related variables is the model with the best prediction performance among the models tested in this paper. In contrast, the probability of a delay is best predicted using a GLM with the cumulative distribution function of the standard normal distribution as the link function, commonly referred to as the probit model in econometrics.

The factors influencing both the probability and duration of delays vary across different ports, with variables such as seasonality, climate, congestion, and anticipated time in port playing key roles. When applying a GLM with gamma-distributed dependent variables, we find that the relevance of specific factors differs depending on the port in question. This suggests that a universal model for identifying influential factors is not feasible. For instance, some ports may handle weather very poorly, thus making weather a critical factor for predicting delays. While in others, the estimated time required for loading or discharging takes precedence due to a propensity for precipitation to only happen for a short period.

The structure of the paper is as follows: Section 2 reviews the relevant literature, while Section 3 presents the dataset and preliminary analysis. Section 4 details the methodology employed in the study, and Section 5 discusses the results and their implications. Section 7 concludes with a summary of key findings and recommendations.

2 Literature review

The study of freight markets has a long-established history, which can be categorized into three main streams of research. The first stream dates back to foundational works such as (Tinbergen, 1934) and (Koopmans, 1939), which employed economic theory and econometric modeling to analyze various shipping markets. This approach has been a focal point in maritime economics for decades, as seen in the contributions of (Zannetos, 1964),(Beenstock, 1985), (Beenstock & Vergottis, 1989),(Norman & Wergeland, 1981)(Strandenes, 1986), (Tvedt, 2003) and (Tvedt, 2011). These studies have consistently aimed to model the freight markets using rigorous theoretical and empirical frameworks.

The second stream of literature applies time-series methods, including Auto Regressive (AR) models and their extensions like Auto Regressive Integrated Moving Average (ARIMA) models (Munim & Schramm, 2017), (Chen et al., 2012), as well as Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models (Kavussanos, 1996), (Gavriilidis et al., 2018), (Drobetz et al., 2012) (Alizadeh & Nomikos, 2011). Later models also allow for the modeling and forecasting of freight rate volatility over time.

The third stream focuses on continuous-time stochastic differential equations (SDEs) to model shipping freight rates. Notable contributions in this area include (Bjerksund & Ekern, 1995), (Adland & Cullinane, 2006), (Adland et al., 2008), (Poblacion, 2015), (Benth et al., 2015) and (Población, 2017). For

a comprehensive review of the literature on spot freight rates, see Alexandridis (Alexandridis et al., 2018) or (Ke et al., 2022).

Much of the aforementioned literature primarily examines the temporal evolution of freight markets. A separate body of research investigates how freight rates vary across different vessels. (Adland et al., 2017), for example, utilizes hedonic regression to explore how offshore market freight rates evolve over time, establishing a hedonic price index to track price movements. (Tamvakis & Thanopoulou, 1995), (Tamvakis & Thanopoulou, 2000), (Köhn & Thanopoulou, 2011), and others have examined how vessel characteristics, such as age, equipment, and technology, influence freight rates, often exploring whether quality differences create a two-tiered market. However, findings on these relationships have been inconsistent, with some studies suggesting that factors like vessel characteristics and contract terms (e.g., laycan periods) significantly affect fixture prices(Alizadeh & Nomikos, 2011).

One of the persistent risk factors in the maritime industry is weather, which has profound impacts on both port and vessel operations. Extreme weather events can lead to catastrophic disruptions, as highlighted in studies by (Cao & Lam, 2019) and (Doll et al., 2014). Even under normal conditions, weather significantly influences port operations (Athanasatos et al., 2014) (Garcia-Alonso et al., 2020), and vessel operations are similarly affected, as evidenced by research on weather routing (Zis et al., 2020). However, the literature on port disruptions due to weather largely focuses on low-frequency, high-impact events like natural disasters. (Lam & Su, 2015) examines port disruptions in East Asia from 1900, noting that natural disasters account for nearly half of all disruptions, a finding consistent with (Adam et al., 2016) study of UK ports between 1950 and 2014. Other works, such as (Berle et al., 2011), (Zhang et al., 2020) and (Stecke & Kumar, 2009), also emphasize high-severity, low-frequency events.

Weather is not the sole cause of port delays. Congestion, which occurs when a port's capacity cannot meet the demand, is another major factor. A variety of models have been proposed to study port congestion. For instance, queuing theory has been applied by (Oyatoye et al., 2011) and (Leachman & Jula, 2011) to analyze vessel operations in ports, while deep learning and AIS data have been used by (Peng et al., 2023) to predict port congestion. (Wu, 2024) extended this approach by employing neural networks to forecast short-term traffic flows and ARMA models for long-term trends. Markov chains have also been utilized in congestion modeling (Pruyn et al., 2020), with the issue gaining more visibility during the COVID-19 pandemic, prompting innovative approaches such as (Gui et al., 2022) fuzzy Bayesian model for congestion risk and (Bai et al., 2022) application of system dynamics to study the economics of port congestion in LPG transport. Common themes in this literature include optimizing port operations and quantifying congestion.

The fields of meteorology and climatology, with their deep historical roots, offer valuable insights into maritime operations. Reliable and abundant weather observations have facilitated the use of stochastic modeling, with early works dating back to the 18th century (Sheynin, 1984). Modeling wind, for instance, is critical in several industries, and research has focused on fitting distributions to wind direction and speed data (van Donk et al., n.d.) (Morgan et al., 2011) and (Efthimiou et al., n.d.). Although simple distributions like Beta and Weibull often provide good fits, more complex models, such as the bimodal Weibull and Wakeby distributions, have been developed for enhanced accuracy (Morgan et al., 2011).

Flooding is another extensively studied weather phenomenon, with recent advances in satellite technology enabling more sophisticated forecasting (Mason et al., 2010). Much of the flood literature emphasizes the spatial dimension,

utilizing machine learning techniques such as neural networks (Kia et al., 2012) and logistic regression (Pradhan & Lee, 2010) to model spatial dependencies like land use and topology. (Tehrany et al., 2014) suggests that combining multiple models can improve flood forecasting, while (Thompson et al., 2019) proposes a model for predicting days when sea levels exceed certain thresholds in Honolulu, the capital of the US state of Hawaii. The relationship between extreme rainfall and flooding has also been studied, with (Zheng et al., 2014) exploring various methods, including point process models, to capture the dependence between rainfall and flooding.

Rainfall itself has been a focal point in stochastic weather modeling, with (Srikanthan & McMahon, 2001) identifying four main types of models for daily precipitation: two-part models, transition probability models, resampling models, and time series models such as ARIMA. Two-part models, in which the duration and amount of rainfall are modeled separately, are exemplified by studies such as (Acreman, 1990) and (Srikanthan & McMahon, 1985). These models have been applied in various regions, with (Chapman, 1997) testing multiple two-part stochastic models on data from 22 Western Pacific islands. (Jimoh & Webster, 1999) investigated inter-annual weather patterns in Kenya using a Markov chain, while (Srikanthan & McMahon, 2001) concluded that first-order Markov chains are typically sufficient for rainfall modeling. For a comprehensive review of stochastic weather modeling, see (Wilks & Wilby, 1999).

Survival analysis, another relevant field, is widely used in actuarial science and reliability analysis to model the time until an event occurs. Seminal contributions include the Kaplan-Meier estimator (Kaplan & Meier, 1958) and the Nelson-Aalen estimator (Nelson, 1969) (Nelson, 1972), both of which are non-parametric methods for estimating survival curves. The log-rank test, introduced in (Mantel, 1966) and generalized in (Peto & Peto, 1972), is commonly

used to compare survival curves, while hazard models, notably the Cox proportional hazards model (Cox, 1972), are extensively treated in the literature. (Andersen et al., 1993) and (Fleming & Harrington, 1991) provide a theoretical treatment. (Hosmer & Lemeshow, 1999) and (Rao et al., 1998) gives a more applied treatment.

Generalized linear models (GLMs), introduced in (Nelder & Wedderburn, 1972), represent an important extension of linear regression, allowing for more flexible distributional assumptions. This approach unified previously disparate models such as probit regression (Finney, 1952) and Poisson regression (Nelder, 1974). GLMs were further generalized into Generalized Additive Models (GAMs) by (Hastie & Tibshirani, 1990).

Model selection and goodness-of-fit measures are critical areas in econometrics and statistics. The Akaike Information Criterion (AIC), introduced by (Akaike, 1974), is one of the most widely used model selection criteria, balancing fit and model complexity (Hastie et al., 2009). The Bayesian Information Criterion (BIC), introduced in (Schwarz, 1978), offers a similar approach with a different penalty structure.

Despite the wealth of research on maritime operations, a gap remains in the literature regarding delays from the perspective of vessel operators during normal conditions—situations where ports are not affected by extreme, low-frequency weather events such as natural disasters. Delays are normal occurrences and constitute an economic risk for vessel operators. Suppose the current spot rate is in the tens of thousands per day, a few hours of delay can thus have thousands of dollars in opportunity costs alone. Therefore, operators need a good understanding of the delay risk and such that the risk can be priced when fixing voyages. This study addresses this gap by modeling the factors influencing delay and proposes a prediction model to predict the consequences of

weather-related port delays under typical operating conditions.

3 Data

This study utilizes a proprietary dataset on port calls and deduction events, generously provided by Western Bulk. The dataset comprises 1,114 port calls and 3,331 individual deduction events for the six most frequently visited ports by Western Bulk vessels during the period from October 2015 to March 2024. Deduction events refer to reported delays, including those not related to weather. Non-weather-related deduction events were excluded from the analysis. It is important to note that multiple deduction events may occur per port call. To isolate weather-related delays, each port call was matched with its associated deduction events, if any. The total delay time for each port call was calculated as the sum of the weather-related deduction events. This dataset also includes the economic cost per day for the vessel when at port.

The ports included in the analysis are Richard's Bay (South Africa), Al-Jubail (Saudi Arabia), Koh Si Chang (Thailand), Owendo (Gabon), San Lorenzo-San Martín (Argentina, hereafter referred to as San Lorenzo), and Vila do Conde (Brazil). These ports are geographically distributed, primarily in the southern hemisphere, introducing an additional layer of complexity to the analysis. The spatial distribution suggests that weather patterns across the ports may be influenced by different climatic factors. For instance, seasonality varies between South Africa and Saudi Arabia, and the El Niño/La Niña cycle, also known as the El Niño-Southern Oscillation (ENSO), impacts weather in South America differently than in Southeast Asia. The geographical locations of these ports are depicted in Figure 1.

The dataset includes both delayed and non-delayed port calls, leading to an overdispersion of zero-delay observations. When modeling the duration of

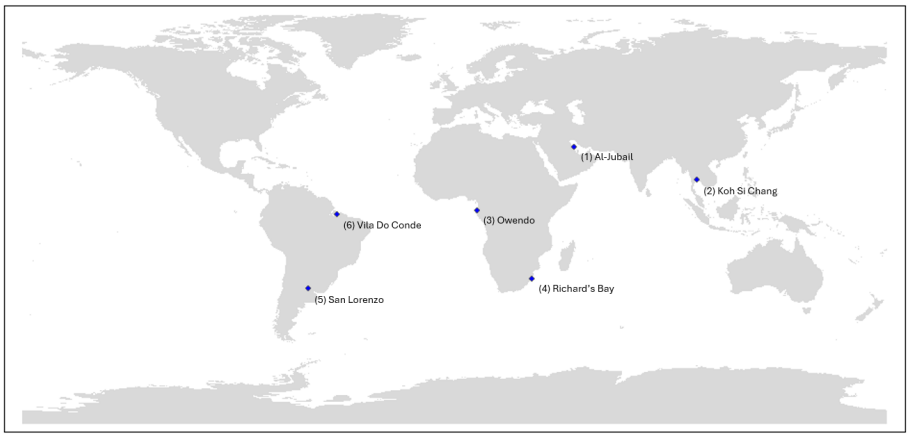

Figure 1: Locations of the six global ports studied: Koh Si Chang, Al-Jubail, Richard's Bay, Owendo, Vila do Conde, and San Lorenzo.

delays, only delayed port calls were included, whereas all port calls were considered when modeling the probability of delay. This leads to more flexibility in modeling since duration and probability can be modeled individually.

Several climate-related variables were included to account for their impact on delays. One of the most prominent climate phenomena affecting precipitation is the El Niño-Southern Oscillation (ENSO), which significantly influences weather in the Americas and Southeast Asia. Differences in sea temperature and wind patterns cause a movement in warm air. This movement in warm air affects precipitation in both the Americas and Asia ("El Niño & La Niña (El Niño-Southern Oscillation) | NOAA Climate.gov", n.d.). The Madden-Julian Oscillation (MJO), an eastward-moving weather pattern, primarily affects weather over the Indian and western Pacific Oceans (Gottschalck & Higgins, n.d.). The North Atlantic Oscillation (NAO), which involves changing air pressure over the Atlantic Ocean, can affect weather as far south as South America ("North Atlantic Oscillation (NAO) | National Centers for Environmental Information (NCEI)", n.d.). Weather-related data were sourced from the National Oceanic

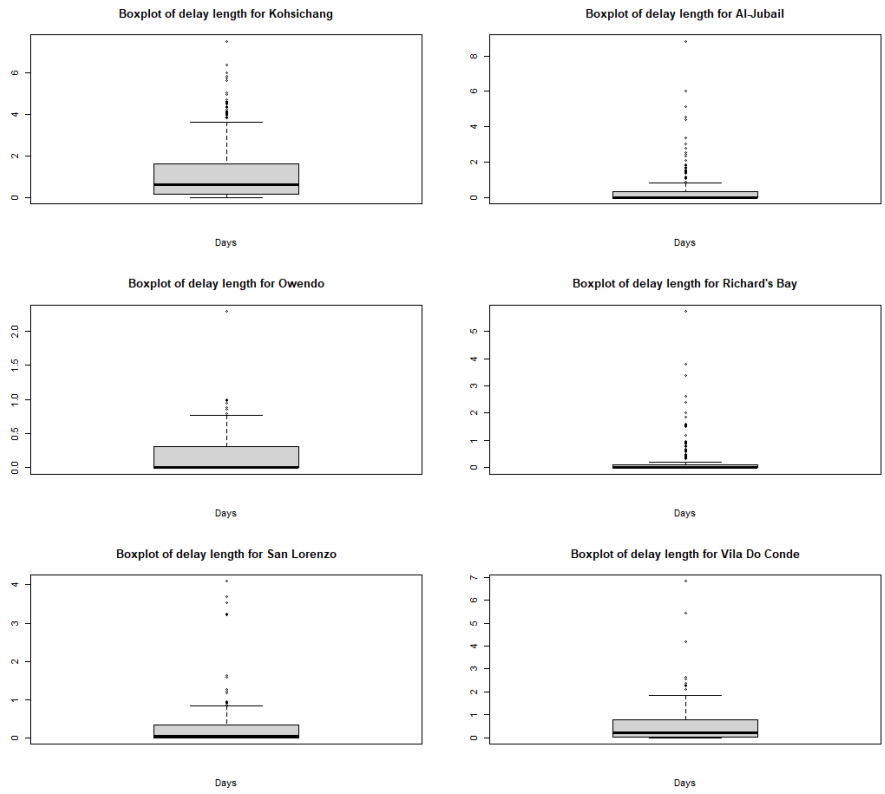

Figure 2: Boxplots of delay length, including port calls with no delay

and Atmospheric Administration (NOAA).

Additional variables, such as cargo size and loading/discharging rates, were included to estimate time spent in port. Congestion was also considered, with data on the number of vessels waiting, tonnage waiting, and both mean and median waiting times. The number of vessels waiting will be used as a proxy for congestion.

3.1 Seasonality

Shipping markets often exhibit seasonality, typically driven by fluctuations in trade demand. For example, demand for LNG tankers increases in the summer

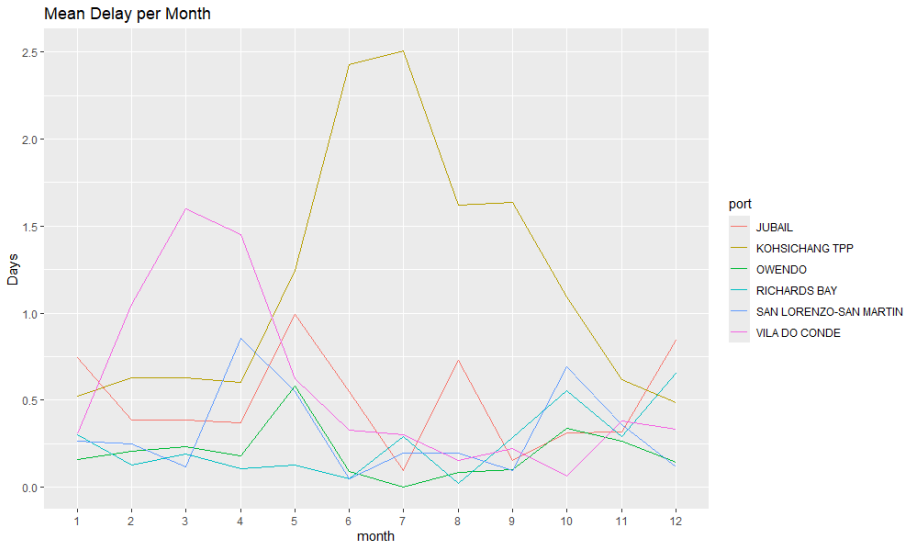

Figure 3: Average delay per month, including port calls with no delay

as European countries stockpile LNG before winter. Although seasonal trade patterns are less relevant in this context, weather-related seasonality may still impact delays. To explore this, the average delay due to weather were calculated for each port and month. The findings are presented in Figure 3.

Consider Figure 3. The two most pronounced curves are the mean delays in Vila Do Conde and Koh Si Chang. Vila Do Conde is a port in northern Brazil that is close to Belém. According to the World Meteorological Organization's (WMO) weather normal, i.e., the average weather from 1991 to 2020, the climate in this part of Brazil is wet during the winter and spring and dry from the summer and fall ("WMO Climatological Normals | World Meteorological Organization", n.d.). We can see a similar pattern in the delay data. The average number of days delayed increases at the start of the year and then peaks during March and April before falling in May and remaining at a stable low number of days for the rest of the year. Koh Si Chang is the other pronounced curve. This port in Thailand also has a wet season. The rainy season starts

in May and continues until October ("WMO Climatological Normals | World Meteorological Organization", n.d.). This mirrors the pattern in Figure 3. The average number of days delayed is stable at around half a day until May, when it increases drastically. The mean delay continues to increase until it reaches a peak in July. There is an interesting divergence here between normal weather and mean delays as September and October are the wettest month during the year. Still, June and July are the month with the most protracted average delays. The remaining ports show less pronounced seasonal fluctuations, if any. An interesting feature is Al-Jubail. Al-Jubail is in Saudi Arabia and is situated a little north of Bahrain. On average, there are 12.6 days of rain in this area, and on average, there are less than three days of rain every month ("WMO Climatological Normals | World Meteorological Organization", n.d.). However, our data shows that the mean number of days of delay varies between about half a day and almost one day. This could be due to the port being very ill-equipped to handle rain, or it could be spurious considering the small number of observations. Seasonality will be investigated further in the next section when we model the delays further.

Figure 3 reveals pronounced seasonal patterns at Vila do Conde and Koh Si Chang. Vila do Conde, located in northern Brazil near the city of Belém, experiences a rainy season during the winter and spring and a dry season during the summer and fall, according to the World Meteorological Organization (WMO) Weather Normals ("WMO Climatological Normals | World Meteorological Organization", n.d.). This pattern is reflected in the delay data, where average delays increase early in the year, peaking in March and April, before dropping in May and remaining low throughout the year. Similarly, Koh Si Chang in Thailand experiences a rainy season from May to October ("WMO Climatological Normals | World Meteorological Organization", n.d.). Correspondingly,

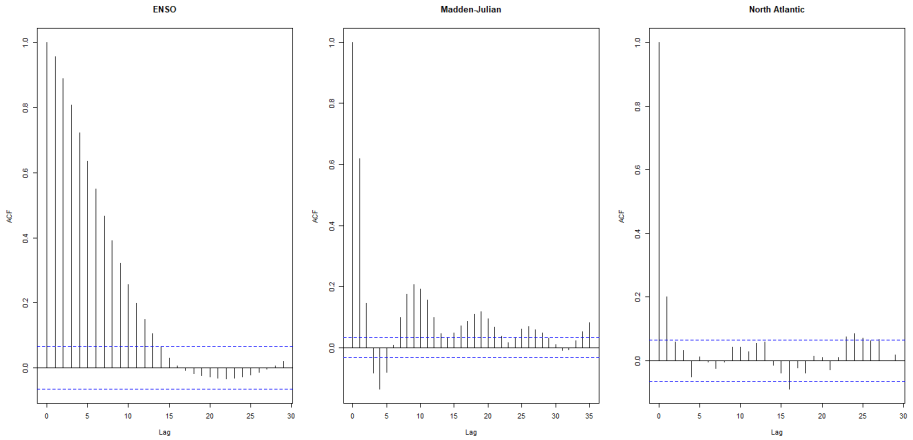

Figure 4: Auto-correlation plots for climate variables: ENSO, Madden-Julian Oscillation, and North Atlantic Oscillation

average delays rise significantly in May and peak in July, despite September and October being the wettest months, suggesting that the most prolonged delays occur earlier in the rainy season.

The other ports show less pronounced seasonality. An interesting exception is Al-Jubail, located in Saudi Arabia. Despite receiving an average of only 12.6 rainy days per year, with fewer than three rainy days per month ("WMO Climatological Normals | World Meteorological Organization", n.d.) , delays fluctuate between half a day and one day. This inconsistency may reflect the port's inadequate infrastructure to handle rain or weather delays due to other weather disruptions, such as strong winds. Further investigation into seasonality will be conducted in the subsequent section through more detailed modeling.

3.2 Auto-Correlation

Figure 4 displays the auto-correlation of the climate indices, indicating clear signs of auto-correlation for ENSO, but less so for the Madden-Julian and North Atlantic Oscillations. The Augmented Dickey-Fuller test rejects the null hypoth-

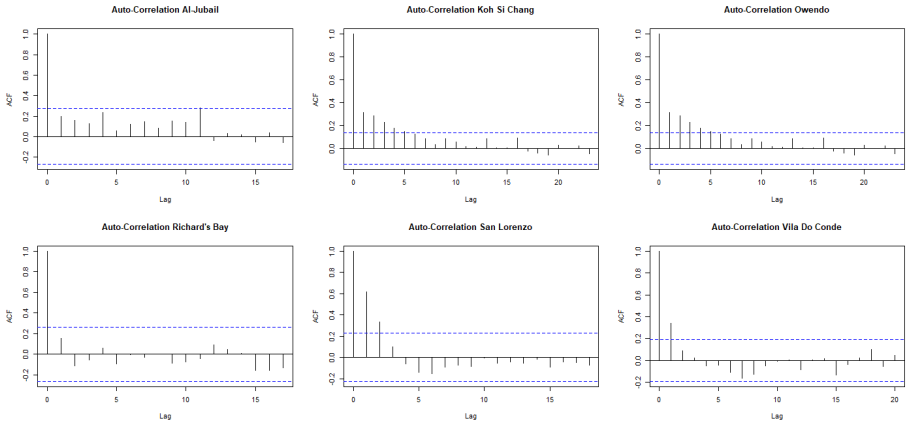

Figure 5: Auto-correlation plots for delay length by port

esis of a unit root at a 1% level, suggesting mean-reverting behavior for ENSO. The MJO and NAO patterns point to potential cyclical, reflected in the wavy nature of their auto-correlation plots. Meteorological evidence suggests that ENSO, MJO, and NAO are cyclical at irregular intervals ("Climate Variability", 2009) ("What is the MJO, and why do we care?", 2014)("El Niño & La Niña (El Niño-Southern Oscillation) | NOAA Climate.gov", n.d.).

Figure 5 shows auto-correlation functions for delays by port. The data suggests potential auto-correlation in Koh Si Chang, Owendo, and San Lorenzo, though this may also reflect seasonality. An added complexity arises from varying interarrival times between delayed port calls.

Robust estimators will be employed in the subsequent analysis to ensure the validity of the results in light of possible auto-correlation.

4 Methodology

In this section, we outline the methodologies employed to analyze and model delays and their economic costs. The selected methods encompass a diverse range of analytical approaches, including some that are uncommon in the field of

shipping economics, such as survival analysis, alongside more traditional econometric techniques.

Given the novelty of this problem in the literature, six different models are explored to examine the duration of delays. These models are: the Cox proportional hazards model, multivariate linear regression, multivariate log-linear regression, a Generalized Linear Model (GLM) with a gamma-distributed dependent variable, a GLM with a negative binomial-distributed dependent variable and an Autoregressive Conditional Duration (ACD) model. The probability of delay is also examined in this paper. Three models are explored to model the probability of delay: multiple linear regression, probit, and logit models. Lastly, the same models are also evaluated with respect to their predictive ability.

4.1 Survival Curves

Survival curves offer a visual and intuitive, non-parametric method for representing the probability of an event occurring as a function of time. In the context of delays, a survival curve represents the probability that a delay lasts longer than a specific time t. Formally, let T be a stochastic variable denoting the length of a delay, defined on a complete probability space $(\Omega, \mathcal{F}_t, P)$, where Ω represents the set of all possible outcomes, $\mathcal{F}_t$ is a filtration, and P is a probability measure. The survival function F is then given by (Aalen, Borgan, & Gjessing, 2008):

$$F(t) = P(T \geq t), \tag{1}$$

Censored data, where the exact value of an observation is only partially known, is common in many fields, such as medicine, where patients may be observed for a limited time, and thus their final outcome (e.g., mortality) is unknown by the end of the study (Aalen, Borgan, & Gjessing, 2008). The Keplan-

Meier estimator is a standard estimator used to estimate survival curves in fields such fields. In cases where no censoring occurs, the Kaplan-Meier estimator is equal to the complement the empirical cumulative distribution function, as $P(T > t) = 1P(Tt)P(T > t) = 1P(Tt)$ (Bland & Altman, 1998). In this study, data is not censored, and therefore, the complement of the empirical cumulative distribution function is used.

4.2 Log-Rank Test

The log-rank test is a non-parametric statistical test designed to compare two or more survival curves based on censored data. Originally introduced by (Mantel, 1966) and later extended by (Peto & Peto, 1972). The same assumptions that underlie the Kaplan-Meier estimator also apply to the log-rank test (Bland & Altman, 2004). These assumptions are, first, censoring and survival probabilities are independent. Second, the survival curve remains constant over time, meaning survival probabilities are the same for port calls that occur earlier and later in the observation period. Thirdly, events occur at the recorded time without measurement error.

This test provides a method to determine whether survival curves differ significantly across groups, making it a valuable tool for comparing delay distributions across different ports.

4.3 Generalized Linear Models

Generalized Linear Models (GLMs) is a flexible class of models that allows for general assumptions on the distributional properties of variables and the functional form of independent variables (Hastie et al., 2009). (Dobson, 2002) describes GLMs as having three properties. Firstly, the dependent variables $Y = Y_1, Y_2, \ldots, Y_n$ are assumed to be identically distributed where the distri-

bution is from the exponential family. The exponential family of distributions encompasses a wide range of distributions, including the normal, Poisson, negative binomial, and gamma distributions (Dobson, 2002). Secondly, have a set of parameters β and independent variables $\mathbf{X} \in R^{n \times M}$. That is, $\mathbf{X}$ is a matrix with n observations of M variables. Thirdly, there is a link function g such that,

$$g(E(X)) = \beta \mathbf{X}. \tag{2}$$

GLMs extend linear regression models in two important ways: the error term need only be independent and need not follow a normal distribution, and the dependent variable Y must be i.i.d. from an exponential family distribution, with a transformed Y (through the link function) linearly related to the independent variables (Dobson, 2002).

The most common methods for estimating GLM parameters include Maximum Likelihood Estimation (MLE) and Iteratively Reweighted Least Squares (IRLS) algorithms (Dobson, 2002).

Several distributions from the exponential family are particularly relevant for this study, including the normal, negative binomial, and gamma distributions. The negative binomial distribution is a discrete probability distribution, whereas the gamma and normal are continuous distributions (Devore & Berk, 2012).

Multiple linear regression is a special case of GLMs where the link function is the identity function, and errors are normally distributed. Multiple linear regression is a model that is widely used in economics, and it's strength lies in it's versatility, ease of interprate and simple estimation using Ordenary Least Square. A further introduction can be found in most introductory econometric textbooks like (Stock, 2015). In this paper, multiple linear regression will be

used to model the length of delay and the probability of delay.

The negative binomial distribution models the number of failures in a series of independent Bernoulli trials before achieving a specified number of successes r. The probability mass function is given by:

$$p(X = x; r, p) = \binom{k + r - 1}{k} (1 - p)^k p^r. \tag{3}$$

(Devore & Berk, 2012). Unlike the linear regression model case, an explicit expression for the coefficients β cannot be explicitly calculated and must be estimated using IRLS, as implemented in the glm.nb function in R ("glm.nb function - RDocumentation", n.d.). We use the logarithm as the link function.

The gamma distribution is particularly suited for modeling strictly positive data. Moreover, the literature on modeling the duration of precipitation suggests that the duration of rainstorms can be modeled using a gamma distribution (Srikanthan & McMahon, 2001). The gamma distribution's probability density function is given by:

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha - 1} e^{-\beta x}, \tag{4}$$

where Γ is the gamma function, defined as:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \tag{5}$$

(Devore & Berk, 2012). This model is also estimated using IRLS, with the inverse of the dependent variable as the link function.

GLMs are popular starting points for modeling probabilities with models such as the Probit and Logit models. GLMs like probit and logit can usually

be fitted using the same methods as other GLMs, either IWLS or Maximum likelihood (Stock, 2015).

The Probit model is a GLM where the linking function is the inverse of the cumulative distribution function of the standard normal distribution. The standard normal distribution is a normal distribution where the mean is 0, and the standard deviation is 1. The Probit model can be expressed as follows,

$$P(Y = 1|\mathbf{X}) = \Phi(\beta\mathbf{X}). \tag{6}$$

In this case, the linking function is $g = \Phi^{-1}$. Moreover, note that the probability of an event, like $Y = 1$, is the expectation of the indicator function for that event.

The Logit function is similar to the Probit model, but the linking function is the inverse of the cumulative standard logistic distribution function (Stock, 2015). This means that the logit model can be expressed as follows,

$$P(Y = 1|\mathbf{X}) = \frac{1}{1 + \exp(-\beta\mathbf{X})} \tag{7}$$

4.4 Cox Proportional Hazard Model

Hazard models are widely employed in survival analysis to model the hazard rate in renewal processes. A renewal process is a stochastic counting process where the times between consecutive events are independent and identically distributed (i.i.d.) (Ross, 2014). In such models, the hazard rate determines the distribution of inter-arrival times between events.

The Cox proportional hazard model, introduced in (Cox, 1972), is a widely used parametric relative risk model. The hazard rate for an individual at time t is given by:

$$\mu(t|x_i) = \mu_0(t) \exp(\beta x(t)) \tag{8}$$

where $\mu_0(t)$ is the baseline hazard, which may vary over time, $x(t)$ is a vector of independent covariates, and β is a vector of coefficients to be estimated (Aalen, Borgan, & Gjessing, 2008).

The Cox model is called "proportional" because a unit increase in an independent variable x leads to a multiplicative increase in hazard rate by a constant $\exp(\beta)$. An underlying assumption is that the hazard rate of the process we study does this.

Estimating the coefficients β using traditional likelihood methods is not feasible for the Cox model (Aalen, Borgan, & Gjessing, 2008). Instead, (Cox, 1975) developed a partial likelihood method, which is commonly used for parameter estimation in the Cox model. This method is further elaborated in (Aalen, Borgan, & Gjessing, 2008).

4.5 Autoregressive Conditional Duration models

Autoregressive Conditional Duration (ACD) models is a class of time series models specifically developed to analyze interarrival times. These models were initially introduced in (Engle & Russell, 1998) to study microstructures in financial markets, and is similar to GARCH models. Let τ_i be the i'th interarrival time then ACD(m,q) models are on the form,

$$x_i = \psi_i \epsilon_i \tag{9}$$

$$\psi_i = \omega + \sum_{j=0}^{m} \alpha_j x_{i-j} + \sum_{j=0}^{q} \beta_j \psi_{i-j} \tag{10}$$

where ϵ_i is an i.i.d. random variable with $E(\epsilon_i) = 1$. Model estimation is typically performed using maximum likelihood estimation (MLE).

4.6 Goodness-of-fit and Model Evaluation

Evaluating model goodness-of-fit is a complex task that requires balancing the accuracy of the model's fit to the data with the risk of overfitting. Generally, models that fit the observed data well are desirable. However, a model that fits a particular data sample too closely may perform poorly when applied to other samples from the same underlying process, a phenomenon known as overfitting. To address this issue, this paper employs goodness-of-fit metrics that balance model fit and complexity, where complexity is measured by the number of parameters in the model. Models with more parameters are more likely to be overfitted, as their increased complexity can result in poorer performance on new data.

The Akaike Information Criterion (AIC) is a common goodness-of-fit metric, often used in model selection, that incorporates both fit and complexity (Hastie et al., 2009). The AIC assesses how well a model fits the data by calculating the likelihood of the model given the data sample and imposes a penalty for model complexity by multiplying the number of parameters by 2 (Devore & Berk, 2012). A model with a lower AIC value is considered preferable to one with a higher AIC, as it suggests a better balance between goodness-of-fit and complexity.

AIC is defined as follows,

$$AIC = 2k - 2\log(\hat{L}), \tag{11}$$

where k is the number of parameters in the model, and $\hat{L}$ is the maximized likelihood of the model.

The Bayesian Information Criterion (BIC) is closely related to the AIC, but the two criteria differ in their theoretical underpinnings and their treatment of model complexity. The BIC is based on Bayesian statistics and penalizes model

complexity logarithmically rather than linearly as AIC does. A model with a lower BIC is also preferable to one with a higher BIC. The BIC is defined as:

$$BIC = 2\ln(k) - 2\log(\hat{L}). \tag{12}$$

where k is the number of parameters and $\hat{L}$ is the maximized likelihood. Since the penalty term in BIC grows logarithmically with the number of parameters, BIC generally penalizes complexity less severely than AIC. An advantage of the BIC is that, as the sample size approaches infinity, it tends to select the true model (Hastie et al., 2009).

In addition to AIC and BIC, non-parametric alternatives such as Mean Square Error (MSE) and Mean Absolute Error (MAE) are frequently used to evaluate model performance. These metrics are particularly versatile in both statistical and machine-learning contexts due to their numerical simplicity and non-parametric nature. The primary difference between MSE and MAE is their treatment of outliers. MSE penalizes large errors more heavily by squaring the error terms. MAE is also sometimes used to evaluate classification models. MAE rewards models that not only predict the correct outcome but also have higher "confidence" in their predictions, i.e., predictions close to zero and one. The formulas for MSE and MAE are:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2 \tag{13}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{Y}_i - Y_i| \tag{14}$$

where $\hat{Y}_i$ is the i'th predicted Y from a model, and Y_i is the i'th true model.

The "0-1 Error" will also be used to evaluate classification models with binary

outcomes. The 0-1 Error measures the proportion of incorrect predictions made by the model, where the predicted outcome is either 0 or 1. The formula for the 0-1 Error is:

$$\text{0-1 Error} = \sum_{i=1}^{n} |\hat{Y}_i - Y_i|. \tag{15}$$

In this context, if the predicted outcome matches the true outcome ($\hat{Y}_i = Y_i$), the error term is 0. If the prediction is incorrect ($\hat{Y}_i \neq Y_i$), the error term is 1.

This paper distinguishes between "in-sample error" and "out-of-sample error." In-sample error refers to the prediction error when predictions are made on the data used to train the model, often referred to as residuals in other contexts (Hastie et al., 2009). Out-of-sample error, on the other hand, measures how well the model performs on data not part of the training set, providing a better indication of the model's ability to generalize to new data.

Cross-validation is a critical method to estimate in-sample and out-of-sample errors, and it provides a more robust evaluation by reducing the impact of arbitrary data separation into training and testing sets. In cross-validation, the dataset is divided into n equal parts, or "folds." The model is trained on n-1 folds and evaluated on the remaining fold. A process that is repeated for each fold. This approach ensures that the model's performance is less dependent on a particular data partitioning (Hastie et al., 2009). Cross-validation is particularly important for models used in forecasting, as it helps ensure that the model generalizes well to new, unseen data.

Sensitivity and specificity are performance measures of a classification model. Sensitivity, also referred to as the true positive rate, is defined as the number of true positive predictions compared to the total number of actual positive observations. Specificity, or the true negative rate, measures the proportion of

true negatives correctly identified relative to the total number of actual negatives (Hastie et al., 2009). These two metrics typically exhibit an inverse relationship. Hence, there is an inherent trade-off between minimizing false positives and minimizing false negatives.

Receiver Operating Characteristic (ROC) curves provides a visual representation of the sensitivity-specificity trade-off. In an ROC curve, the true positive rate (sensitivity) is plotted against the false positive rate (1-specificity) across different decision thresholds. The resulting curve characterizes the model's ability to balance sensitivity and specificity. A random classifier would produce a diagonal line, indicating no discrimination ability, whereas a model with strong discriminatory power generates a curve that deviates substantially from the diagonal, appearing more concave or "hooked." The degree of curvature reflects the model's overall effectiveness in managing the trade-off between sensitivity and specificity.

A confusion matrix is a structured representation of classification outcomes from a classification model. Summarizing the counts of the true positive, false positive, true negative, and false negative in a matrix format. Conventions differ, but in this paper, the row correspond to the true value of an observation, and the column correspond to model prediction. Consequently, the top-left element is the number of true positive cases, the top-right is the number of false negative cases, the bottom-left element is the number of false positive cases, and the bottom-right is the number of true negative cases. This arrangement facilitates the evaluation of model performance by providing a comprehensive overview of classification results.

5 Results

In this section, the results are presented and analyzed in relation to the research questions. A detailed discussion is also provided to interpret the findings.

5.1 Survival Curves

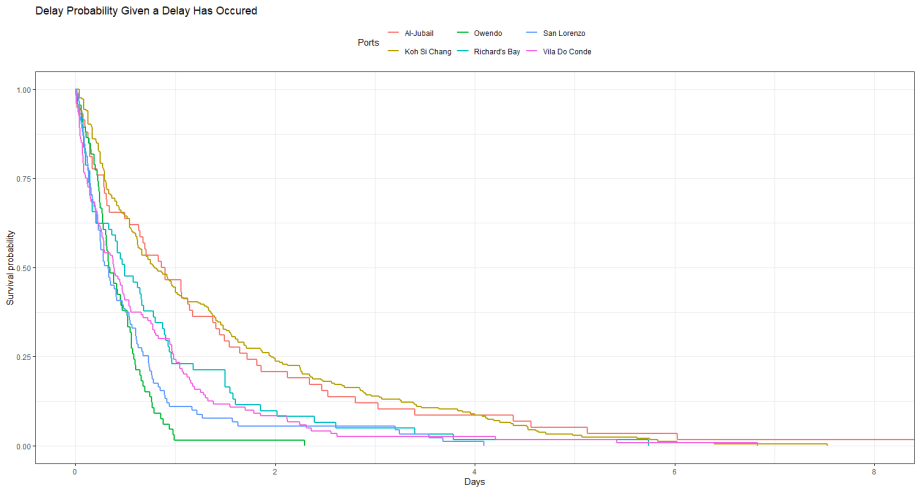

Figure 6: Survival curves for delays per port given that a delay has occurred.

First, note that the assumptions outlined in the methodology section are satisfied. Censoring is independent of survival probability since no data is censored. The second assumption, that survival probability remains constant over time, implies that the underlying climatic phenomena such as El Niño or seasonal effects do not fluctuate significantly over the short term. While long-term changes in climate could lead to variations in these effects, the short-term impact is, therefore, most likely stable and negligible. The third assumption is trivially satisfied in this context. Delays are immediately observable and not latent event observable in retrospect, as in the case of certain diseases.

Table 1 presents the p-value for the pairwise log-rank test and indicates which survival curves are significantly different from each other. Eleven pairs

Table 1: P-Values for pairwise log-rank test

	Al-Jubail	Koh Sic Chang	Ownedo	Richard's Bay	San Lorenzo
Koh Sic Chang	0.947				
Ownedo	0.00000	0			
Richard's Bay	0.040	0.002	0.004		
San Lorenzo	0.0001	0	0.334	0.097	
Vila Do Conde	0.003	0	0.024	0.502	0.254

exhibit significant differences at the 10% significance level; at the 5% level, ten pairs are significantly different; and at the 1% level, eight pairs show significant differences.

The analysis reveals that geographic proximity does not necessarily ensure similar survival curves for port delays. For instance, although Vila Do Conde and San Lorenzo are geographically distinct, their survival curves are statistically different. On the other hand, ports such as Owendo, Richard's Bay, and Al-Jubail exhibit no significant differences in their survival curves. This suggests that the delays experienced in ports within the same region or continent may not follow a uniform pattern, indicating the potential challenges in generalizing a single model across multiple ports.

5.2 What drives delays?

5.2.1 Delay duration

Tables 2 and 3 show that the Generalized Linear Model (GLM) with Gamma-distributed dependent variables offers the best fit for the data while also accounting for model complexity. The second best model is the Autoregressive Conditional Duration (ACD) model, which fits the data better for Al-Jubail and Richard's Bay. However, the performance is only slightly better for these two ports and worse for the remaining four ports. This could indicate a dif-

Table 2: AIC scores of five different models for each port. Model (1) is Al-Jubail, Model (2) corresponds to Koh Si Chang, Model (3) corresponds to Owendo, Model (4) corresponds to Richard's Bay, Model (5) corresponds to San Lorenzo and Model (6) corresponds to Vila Do Conde.

	(1)	(2)	(3)	(4)	(5)	(6)
Log-linear	231.901	749.643	200.510	230.456	284.248	406.843
Linear	240.996	826.784	54.385	186.514	216.081	327.149
Cox	374.046	2,173.190	428.016	393.037	630.317	877.763
GAM Gamma	146.860	599.831	28.165	109.732	59.870	133.326
GAM Negative Binomial	198.585	724.031	122.762	158.159	182.978	256.013
ACD	134.007	618.608	34.932	93.629	618.608	617.868

Table 3: BIC scores of five different models for each port. Model (1) is Al-Jubail, Model (2) corresponds to Koh Si Chang, Model (3) corresponds to Owendo, Model (4) corresponds to Richard's Bay, Model (5) corresponds to San Lorenzo and Model (6) corresponds to Vila Do Conde.

	(1)	(2)	(3)	(4)	(5)	(6)
Log-linear	266.928	809.165	235.545	266.341	326.933	454.230
Linear	276.024	886.306	89.419	222.398	258.766	374.536
Cox	404.953	2,225.708	458.671	424.700	667.980	919.575
GLM-Gamma	186.008	666.355	67.578	149.838	107.576	186.288
GLM-Negative Binomial	233.613	783.552	157.797	194.044	225.663	303.400
ACD	171.095	681.631	72.156	131.625	681.631	680.890

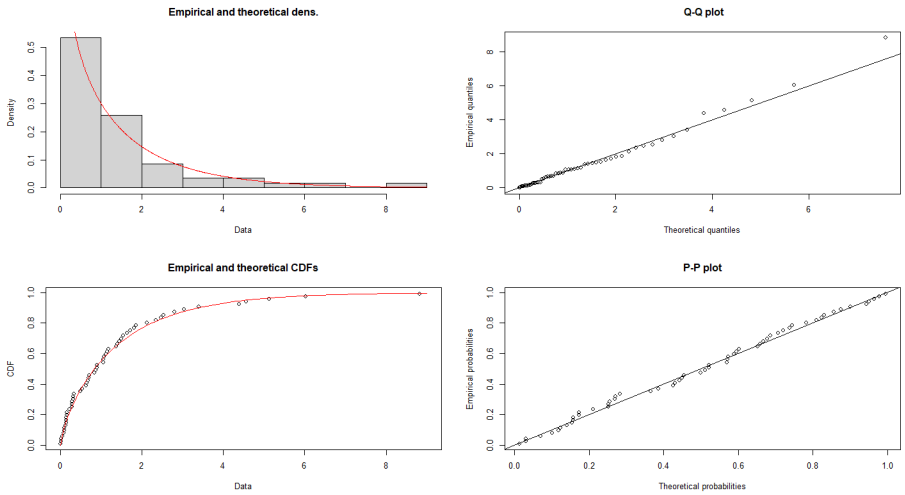

Figure 7: Different figures are comparing the length of delays for port calls to Al-Jubail with the Gamma distribution. The figures are theoretical and empirical density, Q-Q plot, theoretical and empirical cumulative distribution function and P-P plot

ference in autocorrelation structure between ports. This conclusion is further supported by Figure 7. Figure 7 visually illustrates how well the duration of delays at the Al-Jubail port aligns with the Gamma distribution. Notably, only five data points deviate from the expected quantiles in the Q-Q plot, with just one showing a substantial deviation. Although the significance of these plots is limited, given that they focus solely on delays at a single port, they do suggest that assuming the length of delays follows a Gamma distribution is reasonable. Thus, the results in Tables 2 and 3 are well-founded. Consequently, the GLM with Gamma-distributed dependent variables will be employed for inferential analysis.

Consider the coefficients for the Generalized Additive Model (GAM) with gamma-distributed dependent variables and an inverse link function presented from Table 4. Unlike linear regression models, these coefficients cannot be directly interpreted; the link function must be considered. The expected delay

Table 4: Estimates of coefficients for the General Additive Model (GAM) with gamma-distributed dependent variables. Each model corresponds to a different port.

	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
ENSO	0.110	0.181**	−0.209	−0.054	−0.609	0.057
	(0.232)	(0.089)	(0.297)	(0.240)	(0.372)	(0.148)
NAO	0.105	−0.012	0.116	−0.253	0.273*	−0.068
	(0.157)	(0.043)	(0.237)	(0.260)	(0.165)	(0.133)
MJO	−0.030	0.038	0.459**	−0.188	−0.282*	−0.167*
	(0.091)	(0.038)	(0.208)	(0.185)	(0.157)	(0.098)
Cargo volume	0.00001	0.00001***	0.00001	−0.00000	0.00002	−0.00002
	(0.00001)	(0.00000)	(0.00004)	(0.00003)	(0.00003)	(0.00002)
Load/Discharge Rate	0.0001**	0.00003***	0.00001	−0.00000	0.0001	0.00004
	(0.00005)	(0.00001)	(0.0001)	(0.00003)	(0.0002)	(0.0001)
Vessels Waiting	0.106**	−0.010***	−0.272	−0.010	−0.151*	0.001
	(0.049)	(0.003)	(0.175)	(0.015)	(0.078)	(0.020)
February	1.404***	0.157	−0.942	0.817	0.430	−1.067
	(0.410)	(0.483)	(1.084)	(1.667)	(0.828)	(0.683)
March	1.694**	0.090	−0.944	−0.740	1.282	−1.612**
	(0.781)	(0.497)	(1.489)	(1.519)	(1.309)	(0.676)
April	1.266	0.114	−1.774*	−1.312	−1.381**	−1.653**
	(0.819)	(0.485)	(1.004)	(1.698)	(0.628)	(0.727)
May	0.865	−0.283	−1.492	−0.176	−0.932	−0.682
	(0.564)	(0.428)	(1.074)	(1.682)	(0.642)	(0.786)
June	0.956	−0.496	−0.062	−0.981	4.284***	0.789
	(0.714)	(0.418)	(1.185)	(1.879)	(1.487)	(1.061)
July	2.298	−0.562		−0.862	1.285	0.178
	(1.488)	(0.416)		(1.170)	(1.007)	(1.168)
August	1.013	−0.567	−1.088	6.012*	0.011	2.388
	(0.622)	(0.422)	(1.348)	(3.244)	(1.046)	(1.655)
September	1.038***	−0.427	−0.624	−1.117	2.418**	0.640
	(0.364)	(0.424)	(1.421)	(1.562)	(0.967)	(1.521)
October	1.165	−0.176	−2.168**	−1.383	−0.896	4.933**
	(0.748)	(0.434)	(1.019)	(1.390)	(0.734)	(2.091)
November	1.156*	0.137	−0.662	−0.524	−0.258	−0.414
	(0.649)	(0.490)	(1.019)	(1.444)	(0.625)	(0.949)
December	0.840	0.236	−1.106	−1.099	3.031	0.322
	(0.736)	(0.616)	(1.434)	(1.074)	(1.971)	(1.024)
Constant	−1.882**	0.781*	4.121**	2.752**	1.251	2.410***
	(0.740)	(0.444)	(1.629)	(1.380)	(1.354)	(0.842)

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

length is given by the equation:

$$\frac{1}{E(T_i|X)} = \beta X. \tag{16}$$

where a one-unit increase in a covariate x_1 will increase the delay's expected length by $\frac{1}{\beta_i}$. Thus, a positive coefficient indicates that an increase in x_i leads to a decrease in the expected delay length, while a negative coefficient suggests the opposite—an increase in x_i results in an increase in the expected delay.

In Model (1), which estimates the expected delay length for the port of Al-Jubail in Saudi Arabia, the climate variables ENSO, NAO, and MJO are statistically insignificant, which aligns with expectations given the port's location in the arid Persian Gulf, where such weather phenomena have minimal impact. Only the monthly dummy variables for February, March, September, and November are significant, which aligns with the winter season. The winter is typically wet, and a significant part of the rain falls between November and March ("WMO Climatological Normals | World Meteorological Organization", n.d.). Rainfall in the region has generally short duration but high intensity (Metz, 1994). Interestingly, the expected delay duration decreases during this wet season. Given that the region experiences fewer than three days of rainfall per month on average, other weather-related factors, such as strong winds or high temperatures, may play a role. For example, lower temperatures during winter can make working conditions easier and port operations more time-efficient. Additionally, the number of vessels waiting is significant at the 10% level, suggesting that higher congestion corresponds to shorter delays, potentially due to favorable seasonal trade patterns. The load and discharge rate are also significant factors.

Model (2), which models delays at Koh Si Chang in Thailand, shows that the El Niño-Southern Oscillation (ENSO) is significant at the 5% level. As the

ENSO index increases, indicating an El Niño phase, the expected delay length decreases due to drier conditions in Southeast Asia. Conversely, a lower index indicates the wetter La Niña phase, leading to longer delays ("El Niño & La Niña (El Niño-Southern Oscillation) | NOAA Climate.gov", n.d.). None of the seasonal dummies are statistically significant, implying that ENSO primarily drives seasonal variations in delay length. The significance of the cargo volume and load/discharge rate further suggests that vessels with faster operations experience shorter delays as they are less susceptible to weather-related disruptions. However, increased vessel congestion leads to longer delays, likely due to port congestion.

In Model (3), which represents the port of Owendo in Africa, the only significant climate variable is the Madden-Julian Oscillation (MJO). The MJO is known to influence rainfall patterns in Africa (Maybee et al., 2023), and its significance aligns with the onset of the rainy season in October ("WMO Climatological Normals | World Meteorological Organization", n.d.), which is significant at the 5% level. April is also significant at the 10% level, though the low number of observations for this port, with no data for July, limits our ability to detect additional seasonal effects. Neither time spent in port nor congestion shows significant effects.

Model (4), which analyzes delays at Richard's Bay in South Africa, reveals no significant climate, congestion, or cargo-related variables. Only August shows significance at the 10% level, and since it is a dry month, shorter delays are expected. The low number of observations (61 delayed port calls) may obscure more meaningful patterns.

Model (5) examines delays at the port of San Lorenzo in Argentina, located inland along the Paraná River. This port is subject to unique weather-related effects due to its riverine location. The North Atlantic Oscillation (NAO) and

MJO are both significant at the 10% level, but ENSO is surprisingly insignificant. The MJO, which affects precipitation in the Americas, is associated with longer delays as its index increases. Seasonal effects are apparent, with longer delays in April (a wet month) and shorter delays in June and August (drier months) ("WMO Climatological Normals | World Meteorological Organization", n.d.). In general, winters appear to be wetter than summers ("WMO Climatological Normals | World Meteorological Organization", n.d.). April is a wet month, and therefore, longer delays are expected. On the other hand, June and August are dryer and, therefore, should have shorter expected delays. The results might be affected by run-offs or other hydrological effects on the river. Vessel congestion is significant at the 10% level, likely reflecting port congestion effects.

Model (6) models delays at the port of Vila Do Conde in Brazil, situated along the Atlantic coast. The only significant climate variable at the 10% level is the MJO, Like previously stated this phenomenon is known to affect precipitation in the Americas (Gottschalck & Higgins, n.d.). Seasonal effects are significant, with longer delays in March and April, consistent with these being particularly rainy months ("WMO Climatological Normals | World Meteorological Organization", n.d.). No other climate variables are statistically significant for this port.

A notable observation across all models is the relatively weak evidence of clear seasonal effects. Few ports exhibit statistically significant monthly dummy variables. Several factors could explain this. First, the limited number of observations for each port makes it difficult to detect seasonal patterns. Even if seasonal effects exist, insufficient data may prevent their statistical detection. Second, variations in cargo types and port susceptibility to different weather conditions could contribute to the lack of clear seasonal effects. Some cargo

types may be more resilient to adverse weather, such as rain, while other ports may be more affected by wind. Finally, there is significant variability in the number of rainy days across ports. For example, Al-Jubail experiences only two to three rainy days per month, even during wet periods, while Vila Do Conde has over 20 rainy days on average. Such differences may dilute the observed seasonal effects, particularly when comparing ports with markedly different weather conditions.

5.2.2 Delay probability

Table 5: AIC scores of three different models for each port.

	(1)	(2)	(3)	(4)	(5)	(6)
Linear	218.031	186.448	208.650	260.887	266.067	165.576
Logit	207.579	205.805	195.236	247.336	252.910	159.489
Probit	207.918	205.228	195.170	246.522	253.107	159.851

Table 6: BIC scores of three different models for each port.

	(1)	(2)	(3)	(4)	(5)	(6)
Linear	272.696	252.193	263.198	319.801	322.930	220.124
Logit	259.207	267.897	246.753	302.977	306.614	211.006
Probit	259.547	267.321	246.688	302.163	306.811	211.369

Tables 5 and 6 display the goodness-of-fit metrics for each of the three models estimated across six ports. The results indicate that both the probit and logit models outperform the linear model on both criteria for most ports, with the exception of Koh Si Chang. Between the probit and logit models, performance is comparable, with each model slightly outperforming the other for half of the ports. For subsequent analysis, we select the probit model due to its more straightforward interpretation. In the probit model, a positive coefficient

increases the probability of the event occurring—in this case, the likelihood of a vessel being delayed.

Table 7 presents the estimated coefficients for the Probit-model where the probability of delay is presented.

Model (1) in Table 7 indicates that the expected load and discharge time is the only significant predictor of a vessel's probability of being delayed in Al-Jubail, with the variable reaching statistical significance at the 1% level. Neither seasonal nor climate-related variables demonstrate any significant effect.

Model (2) reveals that the probability of delay in Koh Si Chang is influenced by seasonal patterns, climatic conditions, and port congestion. The El Niño-Southern Oscillation (ENSO) is the only significant climate-related variable, with the El Niño phase resulting in drier weather and thus reducing the likelihood of delays in Thailand ("El Niño & La Niña (El Niño-Southern Oscillation) | NOAA Climate.gov", n.d.). The probability of delays increases during the rainy season, which spans from May to November ("WMO Climatological Normals | World Meteorological Organization", n.d.). Additionally, congestion plays a significant role in increasing the probability of delays.

Model (3) demonstrates that the likelihood of delays in Owendo is primarily driven by the expected load and discharge time, alongside a seasonal effect in July. July is a particularly dry month, leading to a significant decrease in the probability of delays during this period ("WMO Climatological Normals | World Meteorological Organization", n.d.).

Delays in Richard's Bay are predominantly influenced by the expected duration of port calls and seasonal variations. Model (4) shows that expected loading and discharging time is significant at the 1% level. The probability of delays decreases during the spring and summer months, corresponding with reduced precipitation in Eastern South Africa from March to August ("WMO Climato-

Table 7: Estimates for the coefficients of the Probit model.

	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
ENSO	−0.195	−0.569***	−0.183	0.0004	0.271*	−0.075
	(0.183)	(0.174)	(0.200)	(0.155)	(0.164)	(0.163)
NAO	−0.076	0.204	0.035	−0.022	0.066	0.009
	(0.111)	(0.128)	(0.139)	(0.095)	(0.109)	(0.161)
MJO	0.075	−0.082	−0.093	−0.160	−0.120	−0.150
	(0.094)	(0.108)	(0.128)	(0.105)	(0.105)	(0.115)
Expected Load/Discharg time	0.149***	−0.013	0.312***	0.138***	0.236**	0.072
	(0.048)	(0.048)	(0.110)	(0.051)	(0.118)	(0.100)
Vessels Waiting	0.037	0.039***	−0.084	0.024*	0.075	0.052*
	(0.068)	(0.012)	(0.063)	(0.013)	(0.076)	(0.029)
February	0.511	0.550	−0.029	−0.885	0.292	1.108*
	(0.585)	(0.444)	(0.732)	(0.571)	(0.488)	(0.662)
March	0.588	0.442	−0.115	−1.136**	0.002	0.591
	(0.573)	(0.428)	(0.762)	(0.551)	(0.559)	(0.624)
April	0.705	0.789	−0.927	−1.321**	0.463	5.437***
	(0.536)	(0.519)	(0.704)	(0.561)	(0.467)	(0.477)
May	−0.245	1.207**	0.609	−1.067*	0.150	1.121
	(0.514)	(0.565)	(0.794)	(0.624)	(0.499)	(0.754)
June	−0.305	1.707**	−0.870	−2.125***	−0.398	5.372***
	(0.553)	(0.766)	(0.691)	(0.686)	(0.479)	(0.460)
July	−0.437	1.842***	−6.157***	−0.710	0.647	−0.137
	(0.663)	(0.706)	(0.635)	(0.556)	(0.562)	(0.736)
August	−0.132	6.370***	−1.184	−1.311**	−0.049	0.031
	(0.527)	(0.475)	(0.740)	(0.555)	(0.518)	(0.572)
September	−0.324	6.264***	−1.004	−0.537	0.253	−0.062
	(0.681)	(0.590)	(0.701)	(0.533)	(0.497)	(0.716)
October	−0.295	1.629***	−0.253	−0.366	−0.143	−0.386
	(0.624)	(0.575)	(0.707)	(0.567)	(0.543)	(0.609)
November	−0.185	1.003**	0.312	−0.681	0.421	0.004
	(0.594)	(0.476)	(0.705)	(0.608)	(0.458)	(0.541)
December	0.391	0.693	−0.669	−0.905	0.263	0.723
	(0.517)	(0.517)	(0.731)	(0.572)	(0.540)	(0.613)
Constant	−1.167**	−1.065*	−0.376	−0.587	−1.449**	−0.434
	(0.480)	(0.580)	(0.696)	(0.513)	(0.703)	(0.650)

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

logical Normals | World Meteorological Organization", n.d.), which lowers the likelihood of weather-related delays.

In San Lorenzo, Model (5) identifies time spent in port and ENSO as the only significant factors affecting the probability of delays. El Niño-driven precipitation likely contributes directly to delays, while runoff from increased rainfall may have an indirect impact ("El Niño & La Niña (El Niño-Southern Oscillation) | NOAA Climate.gov", n.d.). Additionally, the model shows that longer expected loading and discharging times increase the probability of delays.

Model (6) for Vila Do Conde identifies congestion and seasonality as significant predictors of delay probability. Congestion increases the likelihood of delays, though the coefficient is only significant at the 10% level. Delays are more probable in February, April, and June. Given that the summer months are typically dry in Northern Brazil, the significant increase in delay probability during this period is unexpected, possibly due to other weather-related factors.

5.3 What is the best prediction model?

In this subsection, we focus on prediction and will evaluate six models to determine which are most effective at predicting delay duration, and three models for probability of delay.

5.3.1 30-Day Prediction Model

Six models for predicting delay duration were estimated, with their performance summarized in Tables 8 and 9. Variable selection was conducted using AIC forward selection and cross-validation. This approach begins with a simple model containing only a constant and incrementally adds variables that enhance model performance until no further improvements are achieved. The use of cross-validation ensures more robust results, reducing the risk of overfitting to

Table 8: 30-Day predicting out-of-sample mean square error

	(1)	(2)	(3)	(4)	(5)	(6)
Log-linear	0.513	0.089	0.234	0.391	0.241	0.200
Linear	0.322	0.032	0.013	0.173	0.050	0.044
Cox	0.735	0.182	0.330	0.178	0.166	0.154
GLM-Gamma	0.529	0.061	0.271	0.276	0.315	0.355
GLM-Negative Binomial	0.399	0.065	0.133	0.243	0.153	0.137
ADE	1.046	1.203	0.092	0.372	0.271	0.726

Table 9: 30-Day prediction in-sample mean square error

	(1)	(2)	(3)	(4)	(5)	(6)
Log-linear	0.492	0.086	0.209	0.344	0.201	0.186
Linear	0.153	0.025	0.005	0.055	0.023	0.027
Cox	0.897	0.162	0.325	0.213	0.178	0.173
GLM-Gamma	0.559	0.063	0.273	0.241	0.291	0.377
GLM-Negative Binomial	0.373	0.061	0.131	0.186	0.125	0.114
ADE	1.412	1.552	0.091	0.472	1.655	0.762

specific training and test set splits.

The linear regression model demonstrated the lowest mean squared error across all six ports, both in terms of in-sample error (evaluated on the training data) and out-of-sample error (evaluated on independent test data). The Generalized Linear Models (GLMs) also performed well, with the second-best model being the GLM with a Negative Binomial distribution for the dependent variable. This finding contrasts with the results from the last section, where the best-performing model was the GLM with Gamma-distributed dependent variables.

Model performance for the probability of delay is presented in Tables 10, 11, 12, and 13, using two error metrics: mean absolute error and 0-1 error. These metrics provide different insights into model performance. Notably, the mean absolute error rewards models that not only predict the correct outcome but

Table 10: 30-Day probability of delay: out of sample mean absolute error for

	(1)	(2)	(3)	(4)	(5)	(6)
Linear	0.020	0.017	0.011	0.011	0.014	0.022
Probit	0.033	0.028	0.038	0.026	0.013	0.044
Logit	0.045	0.071	0.087	0.044	0.017	0.111

Table 11: 30-Day probability of delay: in-sample mean absolute error for

	(1)	(2)	(3)	(4)	(5)	(6)
Linear	0.020	0.017	0.011	0.010	0.014	0.021
Probit	0.033	0.026	0.038	0.025	0.012	0.048
Logit	0.045	0.067	0.088	0.041	0.015	0.131

Table 12: Out of sample 0-1 error for 30-day prediction of the probability of delay

	(1)	(2)	(3)	(4)	(5)	(6)
Linear	0.017	0.018	0.012	0.010	0.016	0.025
Probit	0.018	0.017	0.008	0.007	0.010	0.022
Logit	0.018	0.017	0.009	0.008	0.012	0.024

Table 13: In sample 0-1 error for the 30-day prediction of the probability of delay

	(1)	(2)	(3)	(4)	(5)	(6)
Linear	0.018	0.017	0.015	0.011	0.019	0.026
Probit	0.017	0.017	0.011	0.008	0.011	0.024
Logit	0.017	0.017	0.012	0.009	0.013	0.025

also have higher "confidence" in their predictions.

Overall, linear regression performs well considering the mean absolute error, whereas Probit performed well when considering 0-1 error. Linear regression has support on the set of real numbers also outside the interval $[0, 1]$, i.e., linear regression can predict probabilities greater than 1 and less than 0. Therefore, we move forward with the Probit model.

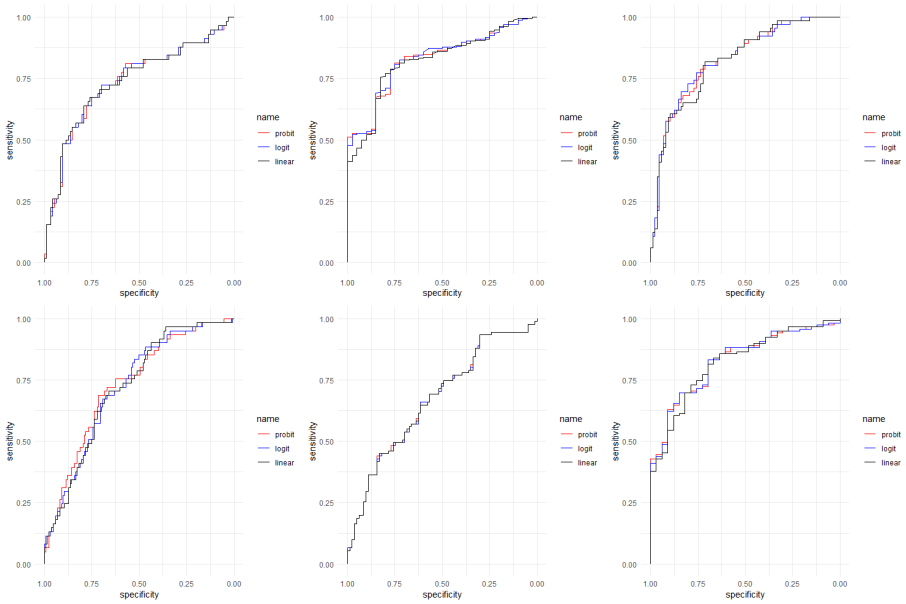

Figure 8: ROC plots for all three models and all ports. Top row: Al-Jubail, Koh Si Chang and Owando. Bottom row: Richard's Bay, San Lorenzo, Vila Do Conde

Figure 8 presents the ROC curves for all three models, evaluated on data from all six ports. Each curve exhibits a concave, "hocked" shape, which signifies that the models demonstrate predictive performance exceeding that of a random classifier. The degree of concavity in the ROC curves reflects the models' ability to balance sensitivity and specificity, with a more pronounced curve indicating superior discriminative power. However, none of the models consistently outperforms the others across all scenarios.

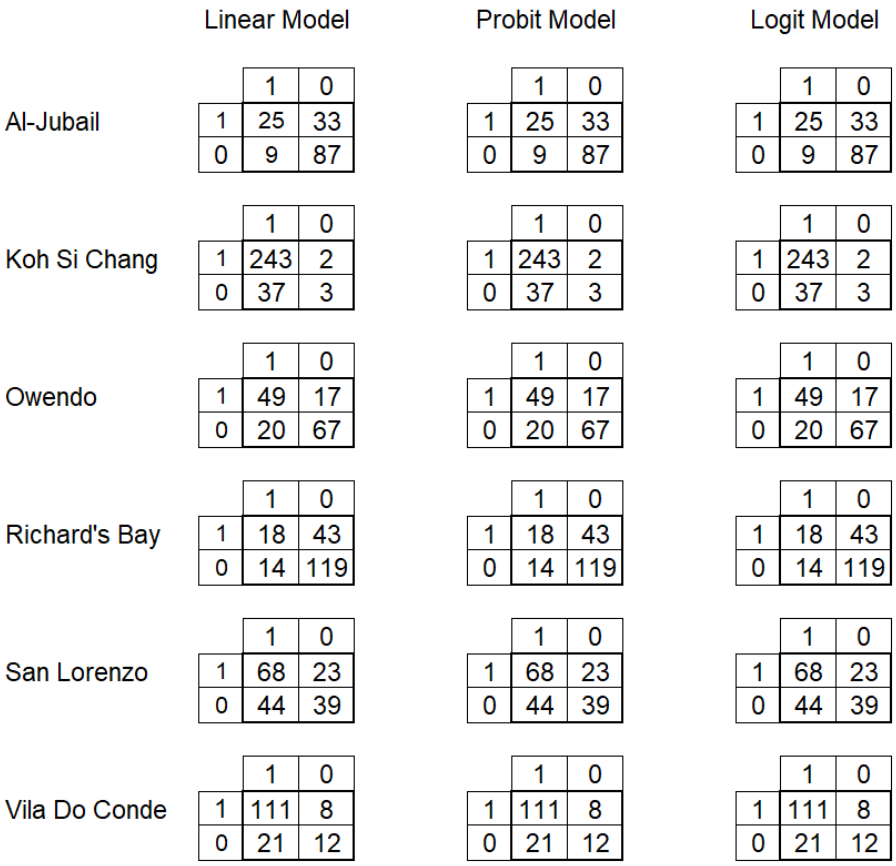

Figure 9: Confusion matrixes for all model and all ports. Top left row are true positives, top right is false negatives, bottom left is false positive and bottom right is true negatives.

Figure 9 presents the confusion matrix for the three models, evaluated on data from all six ports. In this matrix, the column represents the model prediction, and the columns represent are the true values. The predictions made by the models are consistent across all ports. This result aligns with results presented in Tables 13 and 12 where the model results were shown to be very similar. The key distinction between the confusion table 9 and Tables 13 and 12 lies in the data used for evaluation, Table 9 incorporates the entire data sample without dividing into a training and a test set. Consequently, the results are

not contradictory but rather reflect differences in the underlying data samples used for analysis.

The results indicate that all models demonstrate more true positive and true negative predictions than false positives and false negatives, respectively. Consequently, the sensitivity and specificity of all models exceed 0.5. However, sensitivity and specificity vary greatly between ports. This supports the findings in Figure 8. To illustrate how the relationship between sensitivity and specificity varies by port consider Koh Si Chang and Richard's Bay. For Koh Si Chang, the three models exhibit high sensitivity, as a substantial proportion of positive predictions correspond to true positives. However, specificity is lower, as the number of true negatives is only slightly greater than the number of false negatives. On the other hand, the model predictions of delays at Richard's Bay reveals a low sensitivity, with almost as many false positives as true positives, but greater specificity, a greater ratio of true negatives to false negatives. These findings underscore the port-specific nature of the sensitivity-specificity trade-off across the data.

5.4 How do the models compare to the half-day rule?

Table 14: Error for the prediction model and the "half-day"-rule in days. The errors presented is the In-Sample (IS) Error for the model and the "half day"-rule and the respective Out-Of-Sample (OOS) Errors.

	Model In-Sample Error	Model Out-of-Sample Error	Half-Day Error
(1)	0.230	0.250	0.728
(2)	0.818	0.883	0.994
(3)	0.200	0.250	0.395
(4)	0.133	0.194	0.545
(5)	0.257	0.235	0.479
(6)	0.618	0.595	0.570

The fifth research question examines how the prediction models developed

Table 15: Error for the prediction model and the "half-day"-rule in USD. The errors presented are the In-Sample (IS) Error for the model and the "half-day"-rule, and the respective Out-Of-Sample (OOS) Errors.

	Model In-Sample Error	Model Out-of-Sample Error	Half-Day Error
(1)	6,633.024	7,096.505	10,030.750
(2)	9,399.101	9,034.089	12,638.760
(3)	2,289.908	2,499.168	5,780.307
(4)	3,255.476	4,129.169	7,037.433
(5)	3,575.131	4,810.556	6,126.375
(6)	6,155.934	10,200.140	9,011.022

in this study compare to the industry-standard rule, which assumes a delay of half a day for each port call. To address this, the best-performing prediction models for both the duration and probability of delay were estimated, and the prediction errors were calculated in terms of both time and monetary impact. The results are presented in Tables 14 and 15.

Overall, the best-performing prediction model demonstrates greater accuracy for most ports, both in terms of in-sample and out-of-sample errors. A notable exception is Vila Do Conde, where the out-of-sample error is higher for the prediction model when considering time. However, when evaluating the error in monetary terms, the prediction model outperforms the rule of thumb. This suggests that errors in Vila Do Conde tend to occur when the rate is low or when smaller vessels are involved.

The small difference between in-sample and out-of-sample errors indicates that the prediction model generalizes well across data sets, minimizing the risk of overfitting.

While results vary by port, the prediction model generally outperforms the industry rule, suggesting that substantial cost savings can be realized for some ports. For example, prediction for Owendo and Richard's Bay are nearly twice as accurate as the rule of thumb. In contrast, for Vila Do Conde, the rule of

thumb performs slightly better when considering out-of-sample errors.

6 Conclusions

In this paper, we examined the risks associated with weather-related delays during the loading and discharging of cargo. These delays pose a significant financial risk to vessel operators, a topic that, to our knowledge, has not been comprehensively addressed in the existing literature.

Our findings indicate that the expected duration of delays, and consequently the associated costs, vary considerably across ports and, to a lesser extent, across different seasons. This suggests that a uniform approach—wherein a standard delay is assumed across all ports—can result in substantial inaccuracies in delay prediction. Such inaccuracies could have serious implications for the pricing of fixtures, influencing decisions about which contracts to accept and at what price.

Using Kaplan-Meier estimates, we constructed survival curves for delays at different ports, which revealed significant variations in the distribution of delays between locations. These findings were further corroborated by the log-rank test, which supported the notion that delay distributions differ significantly across ports. While geographic proximity appeared to influence delay patterns to some degree, we found that even ports located on the same continent could exhibit significantly different survival curves. Moreover, in some cases, survival curves for short delays were similar across ports, while others exhibited greater similarity in the tail of the delay distribution.

To model the length of delays, we evaluated various potential models and found that a generalized linear model (GLM) with gamma-distributed dependent variables offered the best performance in terms of the Akaike Information Criterion (AIC) across all ports. Other models tested included linear regression (with normally distributed dependent variables), log-linear regression, Cox

proportional hazard models, and GLMs with negative binomial distributions. Our results also provide evidence suggesting that delays may follow a gamma distribution.

As expected, the regression analysis indicated some seasonal effects on delays, with weather phenomena such as El Niño potentially influencing outcomes. However, the overall number of statistically significant effects was limited. This could reflect either the weak influence of these factors or the limitations of our data.

Finally, we proposed a model to estimate the expected costs associated with delays. While the calculations were theoretical, and many of the inputs hypothetical, our results highlight the importance of caution when negotiating vessel fixtures. Although our model is relatively simple, it is realistic and provides a useful tool for determining the margin a vessel operator should account for when fixing a vessel.

References

Tinbergen, J. (1934). Annual Survey of Significant Developments in General Economic Theory. *Econometrica*, *2*(1), 13. https://doi.org/10.2307/1907948

Koopmans, T. C. (1939). Tanker Freight Rates and Tankship Building. *The Netherlands Economic Institute nr. 27.*

Finney, D. J. (1952). Probit Analysis. *Journal of the Institute of Actuaries*, *78*(3), 388–390. https://doi.org/10.1017/S0020268100052938

Kaplan, E. L., & Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, *53*(282), 457–481. https://doi.org/10.1080/01621459.1958.10501452

Zannetos, Z. S. (1964). *The theory of oil tankship rates: An economic analysis of tankship operations.* [Publisher: [Cambridge, Mass., MIT]]. Retrieved November 2, 2023, from https://dspace.mit.edu/bitstream/handle/1721.1/49200/theoryofoiltanks00zann.pdf?sequence=1

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*, *50*(3), 163–170.

Nelson, W. (1969). Hazard Plotting for Incomplete Failure Data [Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00224065.1969.11980344]. *Journal of Quality Technology*, *1*(1), 27–52. https://doi.org/10.1080/00224065.1969.11980344

Cox, D. R. (1972). Regression Models and Life-Tables [_eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1 6161.1972.tb00899.x]. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), 370. https://doi.org/10.2307/2344614

Nelson, W. (1972). Theory and Applications of Hazard Plotting for Censored Failure Data [Publisher: ASA Website _eprint: https://www.tandfonline.com/doi/pdf/10.1080/004017 *Technometrics*, *14*(4), 945–966. https://doi.org/10.1080/00401706.1972.10488991

Peto, R., & Peto, J. (1972). Asymptotically Efficient Rank Invariant Test Procedures. *Journal of the Royal Statistical Society. Series A (General)*, *135*(2), 185. https://doi.org/10.2307/2344317

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Nelder, J. A. (1974). Log Linear Models for Contingency Tables: A Generalization of Classical Least Squares. *Applied Statistics*, *23*(3), 323. https://doi.org/10.2307/2347125

Cox, D. R. (1975). Partial likelihood. *Biometrika*, *62*(2), 269–276. https://doi.org/10.1093/biomet/62.2.269

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2). https://doi.org/10.1214/aos/1176344136

Norman, V. D., & Wergeland, T. (1981). *Nortank: A simulation model of the freight market for large tankers* [Accession Number: 998420341884702202 ISSN: 0332-8333 Publication Title: Norbok Source: NO-TrBIB Type: book]. Retrieved June 13, 2024, from https://urn.nb.no/URN:NBN:no-nb_digibok_2016071909120

Sheynin, O. B. (1984). On the history of the statistical method in meteorology. *Archive for History of Exact Sciences*, *31*(1), 53–95. https://doi.org/10.1007/BF00330243

Beenstock, M. (1985). A theory of ship prices [Publisher: Routledge _eprint: https://doi.org/10.1080/03088838500000028]. *Maritime Policy & Management*, *12*(3), 215–225. https://doi.org/10.1080/03088838500000028

Srikanthan, R., & McMahon, T. A. (1985). Recurrence interval of drought events through stochastic analysis of rainfall and streamflow data. *Hydrological Sciences Journal*, *30*(2), 197–206. https://doi.org/10.1080/02626668509490984

Strandenes, S. P. (1986). *NORSHIP: A simulation model for bulk shipping markets*. Norwegian School of Economics; Business Administration.

Beenstock, M., & Vergottis, A. (1989). An Econometric Model of the World Tanker Market [Publisher: [London School of Economics, University of Bath, London School of Economics and University of Bath, London School of Economics and Political Science]]. *Journal of Transport Economics and Policy*, *23*(3), 263–280. Retrieved January 17, 2024, from https://www.jstor.org/stable/20052891

Acreman, M. C. (1990). A simple stochastic model of hourly rainfall for Farnborough, England. *Hydrological Sciences Journal*, *35*(2), 119–148. https://doi.org/10.1080/02626669009492414

Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*. Chapman & Hall/CRC.

Fleming, T. R., & Harrington, D. P. (1991). *Counting processes and survival analysis*. Wiley.

Andersen, P. K., Borgan, Ø., Gill, R. D., & Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer US. https://doi.org/10.1007/978-1-4612-4348-9

Metz, H. C. (1994). *Persian Gulf States: Country studies* (3rd ed. compl). Government printing office.

Bjerksund, P., & Ekern, S. (1995). Contingent claims evaluation of mean-reverting cash flows in shipping. In *Real options in capital investment: Models, strategies, and applications* (pp. 207–219). Preager Westport, CT.

Tamvakis, M. N., & Thanopoulou, H. A. (1995). An investigation into the existence of a two-tier spot freight market for crude oil carriers. *Maritime Policy & Management*, *22*(1), 81–90. https://doi.org/10.1080/03088839500000034

Kavussanos, M. G. (1996). Comparisons of Volatility in the Dry-Cargo Ship Sector: Spot versus Time Charters, and Smaller versus Larger Vessels

[Publisher: [London School of Economics, University of Bath, London School of Economics and University of Bath, London School of Economics and Political Science]]. *Journal of Transport Economics and Policy*, *30*(1), 67–82. Retrieved January 12, 2024, from https://www.jstor.org/stable/20053097

Chapman, T. G. (1997). Stochastic models for daily rainfall in the Western Pacific. *Mathematics and Computers in Simulation*, *43*(3-6), 351–358. https://doi.org/10.1016/S0378-4754(97)00019-0

Bland, J. M., & Altman, D. G. (1998). Statistics Notes: Survival probabilities (the Kaplan-Meier method). *BMJ*, *317*(7172), 1572–1580. https://doi.org/10.1136/bmj.317.7172.1572

Engle, R. F., & Russell, J. R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data [Publisher: [Wiley, Econometric Society]]. *Econometrica*, *66*(5), 1127–1162. https://doi.org/10.2307/2999632

Rao, M. B., Klein, J. P., & Moeschberger, M. L. (1998). Survival Analysis Techniques for Censored and Truncated Data. *Technometrics*, *40*(2), 159. https://doi.org/10.2307/1270658

Hosmer, D. W., & Lemeshow, S. (1999). *Applied survival analysis: Regression modeling of time to event data*. Wiley.

Jimoh, O. D., & Webster, P. (1999). Stochastic modelling of daily rainfall in Nigeria: Intra-annual variation of model parameters. *Journal of Hydrology*, *222*(1), 1–17. https://doi.org/10.1016/S0022-1694(99)00088-8

Wilks, D. S., & Wilby, R. L. (1999). The weather generation game: A review of stochastic weather models [Publisher: SAGE Publications Ltd]. *Progress in Physical Geography: Earth and Environment*, *23*(3), 329–357. https://doi.org/10.1177/030913339902300302

Tamvakis, M. N., & Thanopoulou, H. A. (2000). Does quality pay? The case of the dry bulk market. *Transportation Research Part E: Logistics and Transportation Review*, *36*(4), 297–307. https://doi.org/10.1016/S1366-5545(00)00005-3

Srikanthan, R., & McMahon, T. A. (2001). Stochastic generation of annual, monthly and daily climate data: A review. *Hydrology and Earth System Sciences*, *5*(4), 653–670. https://doi.org/10.5194/hess-5-653-2001

Dobson, A. J. (2002). *An introduction to generalized linear models* (2. ed). Chapman & Hall / CRC.

Tvedt, J. (2003). A new perspective on price dynamics of the dry bulk market. *Maritime Policy & Management*, *30*(3), 221–230. https://doi.org/10.1080/0308883032000133413

Bland, J. M., & Altman, D. G. (2004). The logrank test. *BMJ : British Medical Journal*, *328*(7447), 1073. Retrieved August 23, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC403858/

Adland, R., & Cullinane, K. (2006). The non-linear dynamics of spot freight rates in tanker markets. *Transportation Research Part E: Logistics and Transportation Review*, *42*(3), 211–224. https://doi.org/10.1016/j.tre.2004.12.001

Aalen, O. O., Borgan, Ø., & Gjessing, H. K. (2008). *Survival and Event History Analysis* (M. Gail, K. Krickeberg, J. Samet, A. Tsiatis, & W. Wong, Eds.). Springer New York. https://doi.org/10.1007/978-0-387-68560-1

Aalen, O. O., Borgan, Ø., & Gjessing, S. (2008). *Survival and event history analysis: A process point of view* [OCLC: ocn213855657]. Springer.

Adland, R., Jia, H., & Lu, J. (2008). Price dynamics in the market for Liquid Petroleum Gas transport. *Energy Economics*, *30*(3), 818–828. https://doi.org/10.1016/j.eneco.2007.02.008

Climate Variability: North Atlantic Oscillation | NOAA Climate.gov. (2009, August). Retrieved January 10, 2025, from http://www.climate.gov/news-features/understanding-climate/climate-variability-north-atlantic-oscillation

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed). Springer.

Stecke, K. E., & Kumar, S. (2009). Sources of Supply Chain Disruptions, Factors That Breed Vulnerability, and Mitigating Strategies. *Journal of Marketing Channels*, *16*(3), 193–226. https://doi.org/10.1080/10466690902932551

Mason, D. C., Speck, R., Devereux, B., Schumann, G. J.-P., Neal, J. C., & Bates, P. D. (2010). Flood Detection in Urban Areas Using TerraSAR-X [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(2), 882–894. https://doi.org/10.1109/TGRS.2009.2029236

Pradhan, B., & Lee, S. (2010). Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environ Earth Sci*.

Alizadeh, A. H., & Nomikos, N. K. (2011). Dynamics of the Term Structure and Volatility of Shipping Freight Rates. *Journal of Transport Economics and Policy (JTEP)*, *45*(1), 105–128.

Berle, Ø., Asbjørnslett, B. E., & Rice, J. B. (2011). Formal Vulnerability Assessment of a maritime transportation system. *Reliability Engineering & System Safety*, *96*(6), 696–705. https://doi.org/10.1016/j.ress.2010.12.011

Köhn, S., & Thanopoulou, H. (2011). A gam assessment of quality premia in the dry bulk time–charter market. *Transportation Research Part E:*

Logistics and Transportation Review, *47*(5), 709–721. https://doi.org/10.1016/j.tre.2011.01.003

Leachman, R. C., & Jula, P. (2011). Congestion analysis of waterborne, containerized imports from Asia to the United States. *Transportation Research Part E: Logistics and Transportation Review*, *47*(6), 992–1004. https://doi.org/10.1016/j.tre.2011.05.010

Morgan, E. C., Lackner, M., Vogel, R. M., & Baise, L. G. (2011). Probability distributions for offshore wind speeds. *Energy Conversion and Management*, *52*(1), 15–26. https://doi.org/10.1016/j.enconman.2010.06.015

Oyatoye, E., Adebiyi, S. O., Okoye, J. C., & Amole, B. B. (2011). Application of Queueing theory to port congestion problem in Nigeria. *European Journal of Business and Management*, *3*(8), 24–36.

Tvedt, J. (2011). Short-run freight rate formation in the VLCC market: A theoretical framework. *Maritime Economics & Logistics*, *13*(4), 442–455. https://doi.org/10.1057/mel.2011.23

Chen, S., Meersman, H., & Voorde, E. v. d. (2012). Forecasting spot rates at main routes in the dry bulk market. *Maritime Economics & Logistics*, *14*(4), 498–537. https://doi.org/10.1057/mel.2012.18

Devore, J. L., & Berk, K. N. (2012). *Modern mathematical statistics with applications* (2. ed). Springer.

Drobetz, W., Richter, T., & Wambach, M. (2012). Dynamics of time-varying volatility in the dry bulk and tanker freight markets [Publisher: Routledge _eprint: https://doi.org/10.1080/09603107.2012.657349]. *Applied Financial Economics*, *22*(16), 1367–1384. https://doi.org/10.1080/09603107.2012.657349

Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2012). An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environ Earth Sci*.

Athanasatos, S., Michaelides, S., & Papadakis, M. (2014). Identification of weather trends for use as a component of risk management for port operations. *Natural Hazards*, *72*(1), 41–61. https://doi.org/10.1007/s11069-012-0491-z

Doll, C., Papanikolaou, A., & Maurer, H. (2014). The vulnerability of transport logistics to extreme weather events. *International Journal of Shipping and Transport Logistics*, *6*(3), 293. https://doi.org/10.1504/IJSTL.2014.060787

Ross, S. M. (2014, January). *Introduction to Probability Models*. Academic Press.

Tehrany, M. S., Lee, M.-J., Pradhan, B., Jebur, M. N., & Lee, S. (2014). Flood susceptibility mapping using integrated bivariate and multivariate statistical models. *Environ Earth Sci*.

What is the MJO, and why do we care? | NOAA Climate.gov. (2014, December). Retrieved January 10, 2025, from http://www.climate.gov/news-features/blogs/enso/what-mjo-and-why-do-we-care

Zheng, F., Westra, S., Leonard, M., & Sisson, S. A. (2014). Modeling dependence between extreme rainfall and storm surge to estimate coastal flooding risk [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013WR014616]. *Water Resources Research*, *50*(3), 2050–2071. https://doi.org/10.1002/2013WR014616

Benth, F. E., Koekebakker, S., & Taib, C. M. I. C. (2015). Stochastic dynamical modelling of spot freight rates. *IMA Journal of Management Mathematics*, *26*(3), 273–297. https://doi.org/10.1093/imaman/dpu001

Lam, J. S. L., & Su, S. (2015). Disruption risks and mitigation strategies: An analysis of Asian ports. *Maritime Policy & Management*, *42*(5), 415–435. https://doi.org/10.1080/03088839.2015.1016560

Poblacion, J. (2015). The stochastic seasonal behavior of freight rate dynamics. *Maritime Economics & Logistics*, *17*(2), 142–162. https://doi.org/10.1057/mel.2014.37

Stock, J. H. (2015). *Introduction to econometrics* (Updated Third, Global edition.). Pearson.

Adam, E. F., Brown, S., Nicholls, R. J., & Tsimplis, M. (2016). A systematic assessment of maritime disruptions affecting UK ports, coastal areas and surrounding seas from 1950 to 2014. *Natural Hazards*, *83*(1), 691–713. https://doi.org/10.1007/s11069-016-2347-4

Adland, R., Cariou, P., & Wolff, F.-C. (2017). What makes a freight market index? An empirical analysis of vessel fixtures in the offshore market. *Transportation Research Part E: Logistics and Transportation Review*, *104*, 150–164. https://doi.org/10.1016/j.tre.2017.06.006

Munim, Z. H., & Schramm, H.-J. (2017). Forecasting container shipping freight rates for the Far East – Northern Europe trade lane. *Maritime Economics & Logistics*, *19*(1), 106–125. https://doi.org/10.1057/s41278-016-0051-7

Población, J. (2017). Are recent tanker freight rates stationary? *Maritime Economics & Logistics*, *19*(4), 650–666. https://doi.org/10.1057/mel.2016.7

Alexandridis, G., Kavussanos, M. G., Kim, C. Y., Tsouknidis, D. A., & Visvikis, I. D. (2018). A survey of shipping finance research: Setting the future research agenda. *Transportation Research Part E: Logistics and Transportation Review*, *115*, 164–212. https://doi.org/10.1016/j.tre.2018.04.001

Gavriilidis, K., Kambouroudis, D. S., Tsakou, K., & Tsouknidis, D. A. (2018). Volatility forecasting across tanker freight rates: The role of oil price shocks. *Transportation Research Part E: Logistics and Transportation Review*, *118*, 376–391. https://doi.org/10.1016/j.tre.2018.08.012

Cao, X., & Lam, J. S. L. (2019). Simulation-based severe weather-induced container terminal economic loss estimation. *Maritime Policy & Management*, *46*(1), 92–116. https://doi.org/10.1080/03088839.2018.1516049

Thompson, P. R., Widlansky, M. J., Merrifield, M. A., Becker, J. M., & Marra, J. J. (2019). A Statistical Model for Frequency of Coastal Flooding in Honolulu, Hawaii, During the 21st Century [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/1 *Journal of Geophysical Research: Oceans*, *124*(4), 2787–2802. https://doi.org/10.1029/2018JC014741

Garcia-Alonso, L., Moura, T. G. Z., & Roibas, D. (2020). The effect of weather conditions on port technical efficiency. *Marine Policy*, *113*, 103816. https://doi.org/10.1016/j.marpol.2020.103816

Pruyn, J., Kana, A., & Groeneveld, W. (2020). Analysis of port waiting time due to congestion by applying Markov chain analysis. In *Maritime Supply Chains* (pp. 69–94). Elsevier. https://doi.org/10.1016/B978-0-12-818421-9.00005-7

Zhang, Y., Wei, K., Shen, Z., Bai, X., Lu, X., & Soares, C. G. (2020). Economic impact of typhoon-induced wind disasters on port operations: A case study of ports in China. *International Journal of Disaster Risk Reduction*, *50*, 101719. https://doi.org/10.1016/j.ijdrr.2020.101719

Zis, T. P. V., Psaraftis, H. N., & Ding, L. (2020). Ship weather routing: A taxonomy and survey. *Ocean Engineering*, *213*, 107697. https://doi.org/10.1016/j.oceaneng.2020.107697

Bai, X., Jia, H., & Xu, M. (2022). Port congestion and the economics of LPG seaborne transportation. *Maritime Policy & Management*, *49*(7), 913–929. https://doi.org/10.1080/03088839.2021.1940334

Gui, D., Wang, H., & Yu, M. (2022). Risk Assessment of Port Congestion Risk during the COVID-19 Pandemic. *Journal of Marine Science and Engineering*, *10*(2), 150. https://doi.org/10.3390/jmse10020150

Ke, L., Liu, Q., Ng, A. K., & Shi, W. (2022). Quantitative modelling of shipping freight rates: Developments in the past 20 years. *Maritime Policy & Management*, 1–19. https://doi.org/10.1080/03088839.2022.2138595

Maybee, B., Ward, N., Hirons, L. C., & Marsham, J. H. (2023). Importance of Madden–Julian oscillation phase to the interannual variability of East African rainfall. *Atmospheric Science Letters*, *24*(5), e1148. https://doi.org/10.1002/asl.1148

Peng, W., Bai, X., Yang, D., Yuen, K. F., & Wu, J. (2023). A deep learning approach for port congestion estimation and prediction. *Maritime Policy & Management*, *50*(7), 835–860. https://doi.org/10.1080/03088839.2022.2057608

Wu, Y. (2024). Predictive Analysis of Maritime Congestion Using Dynamic Big Data and Multiscale Feature Analysis (P.-F. Pai, Ed.). *Journal of Electrical and Computer Engineering*, *2024*(1), 5225558. https://doi.org/10.1155/2024/5225558

Efthimiou, G. C., Hertwig, D., Andronopoulos, S., Bartzis, J., & Coceal, O. (n.d.). A Statistical Model for the Prediction of Wind-Speed Probabilities in the Atmospheric Surface Layer.

El Niño & La Niña (El Niño-Southern Oscillation) | NOAA Climate.gov. (n.d.). Retrieved January 10, 2025, from http://www.climate.gov/enso

Glm.nb function - RDocumentation. (n.d.). Retrieved July 30, 2024, from https://www.rdocumentation.org/packages/MASS/versions/7.3-61/topics/glm.nb

Gottschalck, J., & Higgins, W. (n.d.). Madden Julian Oscillation Impacts.

North Atlantic Oscillation (NAO) | National Centers for Environmental Information (NCEI). (n.d.). Retrieved June 20, 2024, from https://www.ncei.noaa.gov/access/monitoring/nao/

van Donk, S., Wagner, L. E., Skidmore, E. L., & Tatarko, J. (n.d.). Comparison Of The Weibull Model With Measured Wind Speed Distributions For Stochastic Wind Generation. *TRANSACTIONS OF THE ASAE*, *48*.

WMO Climatological Normals | World Meteorological Organization. (n.d.). Retrieved July 11, 2024, from https://community.wmo.int/en/wmo-climatological-normals

NORGES HANDELSHØYSKOLE
Norwegian School of Economics

Helleveien 30
NO-5045 Bergen
Norway

T +47 55 95 90 00
E nhh.postmottak@nhh.no
W www.nhh.no