Predicting Ocean Freight Rates Using Machine Learning Methods

Article · February 2025

DOI: 10.34218/JAIML_04_01_001

Journal of Artificial Intelligence and Machine Learning (JAIML)
Volume 4, Issue 1, January-June 2025, pp. 1-26, Article ID: JAIML_04_01_001
Available online at https://iaeme.com/Home/issue/JAIML?Volume=4&Issue=1
Journal ID: 5979-A561; DOI: https://doi.org/10.34218/JAIML_04_01_001

PREDICTING OCEAN FREIGHT RATES USING MACHINE LEARNING METHODS

Prithviraj Lakkakula[a], William W. Wilson[b], Shantanu Awasthi[c], Indranil SenGupta[d]

[a]Former Research Assistant Professor in the Department of Agribusiness and Applied Economics, North Dakota State University, Fargo, North Dakota, USA. Currently, Data Scientist at Mann+Hummel USA Inc.

[b]Distinguished Professor in the Department of Agribusiness and Applied Economics at North Dakota State University, Fargo, North Dakota, USA

[c]Visiting Assistant Professor in the Department of Statistics, Wake Forest University, Winston-Salem, North Carolina, USA

[d]Professor in the Department of Mathematics at North Dakota State University, Fargo, North Dakota, USA.

ABSTRACT

Ocean freight rates (hereafter referred to as ocean rates) have seen unprecedented growth (over 150-175% increase) and volatility in recent years due to many factors, including energy prices, global supply chain logistics and transportation challenges. In this paper, we use machine learning methods (regularized regression and support vector machine regression) to predict ocean rates using daily data between January 2015 and May 2022. The models include global supply chain pressure index, Baltic Dry Exchange Index, Brent crude oil prices, time charter rates, total bulker sales, commodity price, and several other global trade indices as features to predict ocean rates. For model selection, evaluation, and improving accuracy, we employed time series cross validation as well as hyperparameter tuning. Predictive accuracy results of ocean rates will help trading firms in their risk management strategies and strategic decisions.

Keywords: Ocean Freight Rates, Machine Learning, Regression, Supply Chain Indices, Baltic Dry Index, Brent Oil Prices, Time Series Validation, Risk Management.

Cite this Article: Prithviraj Lakkakula, William W. Wilson, Shantanu Awasthi, Indranil SenGupta. (2025). Predicting Ocean Freight Rates Using Machine Learning Methods. *Journal of Artificial Intelligence and Machine Learning (JAIML)*, 4(1), 1-26.

https://iaeme.com/MasterAdmin/Journal_uploads/JAIML/VOLUME_4_ISSUE_1/JAIML_04_01_001.pdf

1. Introduction

Ocean rates have escalated in importance in international trade of bulk commodities and have become increasingly volatile. For example, ocean rates account for about 35-53% of the total shipping costs for soybean shipments from the United States and Brazil to China and for about 6-10% of the total landed cost [USDA, 2022]. Due to the numerous events that occurred concurrent with and following the pandemic, volatility of ocean rates escalated drastically, resulting in risk and opportunities for market participants. For these reasons, understanding how critical factors impact ocean rates and predicting them has become more important.

Figure 1 show the change in ocean rates from US Gulf to various destination ports, including China, Indonesia, Japan, and South Korea between January 2016 and May 2022.

Ocean rates have increased over 175% in case of shipments to China and at least over 150% for other origins. Additionally, ocean rates experienced significant volatility in the recent years, especially after the corona virus disease 2019 (COVID-19). Ocean rates impact almost every business/agribusiness due to their dependence on international trade.

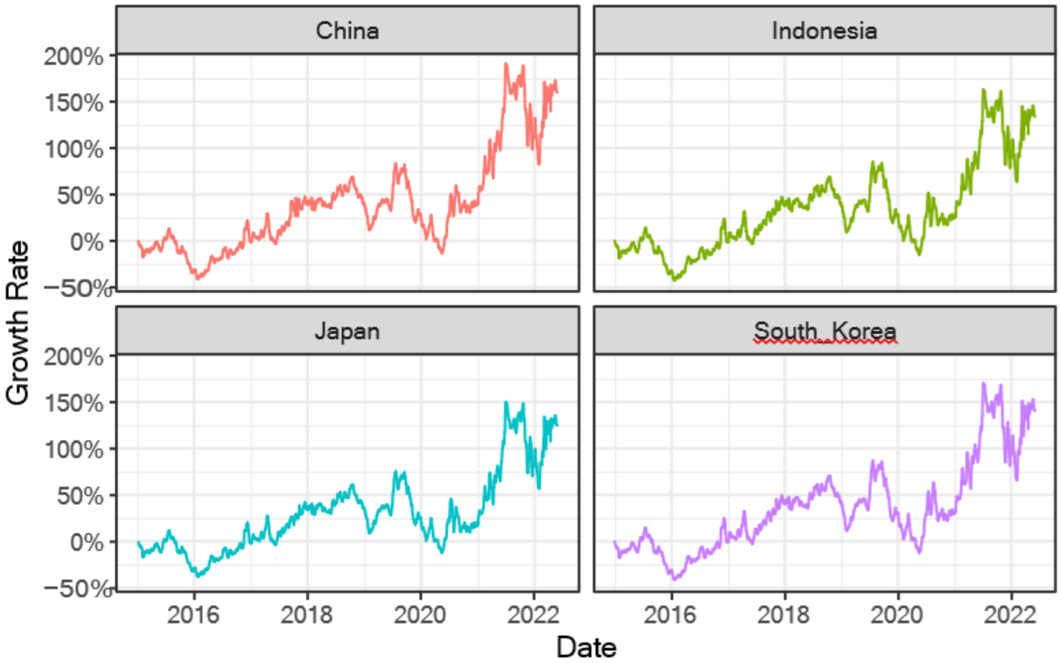


Figure 1: Growth Rate of Ocean Rates from US Gulf, January 2015 to May 2022

Many factors have increased volatility in world agricultural trade. These include the 2021 drought in the United States Northern Plains, the emergence of Renewable Diesel and Sustainable Aviation Fuels, oil prices increasing (from negative values), pressures related to the post-COVID economic expansion, in addition to labor shortages constraining rail, trucks, and other logistical functions, and followed by the Russian invasion of Ukraine "Special Military Operation" [Drewry, 2022]. Additionally, numerous factors impact ocean rates. These have been described in detail in reference to the container shipping sector [Adjemian and Wilson, 2022] but have similar implications for dry-bulk shipping industry.

The purpose of this paper is to develop machine learning (ML) models to predict ocean rates for dry-bulk from the US Gulf and US Pacific Northwest (PNW) origins to selected destinations using two different machine learning specifications. They are regularized regression and support vector machine regression.1 Our primary goal is predictive accuracy of ocean rates and to evaluate and compare predictions of two models. To predict ocean rates, we use features, including time charter rates, Brent crude oil prices, Baltic dry exchange index, total bulker sales, global supply chain pressure index, order book percent fleet, commodity prices, and several other trade indices.

Machine learning accounts for inherent correlations among the features using regularization for improving accuracy of the predictions. Both regularized regression and support vector machine regression models are evaluated using time series cross validation and hyper-parameter tuning for consistency and robust performance of the models, and to improve the overall accuracy of the models, respectively.

Supply chains for most of the products and commodities being traded were impacted by the post-COVID developments. Gamio and Goodman [2021] provide an explanation on how the supply chain crisis developed, including increasing freight rates, port congestion, interior logistical congestion, among others. Drewry [2020] described the increase in shipping costs due to COVID-19. Probably the most important factors impacting the dry-bulk sector in the period post-pandemic includes the volatility in oil prices, in addition to congestion at ports and interior shipping, labor and insurance costs, increased wait and transit times, and changing in global shipping patterns. Additionally, port restrictions in some countries constrained ships from loading and unloading (e.g., China, among others). All these factors caused greater focus on supply chains including risks and management. By late 2022, the global supply chain crisis seems to have abated [Fung et al., 2022].

These elevated risks have important implications for dry-bulk shippers, and for agricul- tural shipping in particular. Exporters confront at least three risks related to changes in ocean freight costs. First is the relationship between short-term and long-term shipping rates which affects strategies related to forward coverage. Second is how global shipping patterns change which in turn affects marketing and asset strategies. A classic example is to switch origins for grain from US Gulf (USG) to the US Pacific Northwest (PNW), and from Ukraine to Brazil for corn, among others [Mano, 2023]. Third is the change in shipping

[1]Alternate machine learning models, including ensemble models such as random forests and gradient boosting models were analyzed but they overfit for our data. Therefore, our analysis was restricted to these two machine learning models described in this paper.

costs between the commodity transaction and shipping date. If these changes are sufficient, traders may negotiate a change of origin, which escalates the value of "switching options" in international commodity trading and agricultural trading in particular [Meersman et al., 2012]. The implications of these changes for traders is the advantage of being able to supply from all origins as suggested in recent trade strategy literature Meersman et al. [2012], and quantified by Johansen and Wilson [2019]. Indeed, this is a virtue of multiple-origins capability. Finally, as a result of these issues, there is a trend in agricultural trade between more geographically nearby origins [Freight-News, 2022a], and to allow greater optionality as reflected in recent tenders [Freight-News, 2022b].

Several studies analyzed the relationships among commodity prices and ocean freight costs, including Bandyopadhyay and Rajib [2021], Tsioumas et al. [2021], Melas and Michail [2021], Barua et al. [2020], Kanamoto et al. [2019], Hathikal et al. [2020] and among others. However, only a few studies have analyzed the ocean rates [Han et al., 2014, Eslami et al., 2017, Yang and Mehmed, 2019] although they account for a critical component of the overall shipping cost.

Han et al. [2014] developed models to predict dry bulk rates using a combined model of wavelet transform and support vector machine. Their model was superior in accuracy to other models. Eslami et al. [2017] sought to predict tanker rates using an artificial neural network and an adaptive genetic algorithm. The results were superior to regression and moving average models. Yang and Mehmed [2019] developed several Artificial Intelligence (AI) models to project ocean rates and compared them to traditional time series.

Others have studied commodity markets using a data-science-based analysis. For in- stance, SenGupta et al. [2019] used a machine/deep learning-based refined stochastic process to analyze optimal hedging strategies in the commodity market. Traditionally, stochastic models assume fixed or nonrandom production or inventory. However, this assumption is not supported by empirical data. This quantity is stochastic, and it is an important feature of risk management strategies. This clearly impacts the hedging decisions. In another paper [Wilson et al., 2019] a data-science-driven analysis is provided for handling the quantity risk in connection to the various stochastic models.

This paper contributes to the literature in several ways. First, ocean rates have been historically high, especially after COVID pandemic, and given the escalation in volatility, predicting rates with improved accuracy considering recent data is important. Second, we perform the best practices of machine learning and data science, including time series cross validation and hyperparameter tuning of both our models unlike other studies in the lit- erature. Finally, we study ocean rates using the unique combination of data generated from Thomson Reuters Eikon and Clarkson, which are rich sources of shipping intelligence information.

2. Background

Organization of the ocean rate markets is well documented [Alizadeh and Nomikos, 2009, Stopford, 2008, Clarkson, 2022]. There are a number of contractual alternatives that are relevant in dry-bulk shipping, including spot transactions, voyage, trip and time charters, all of which are traded and quoted in multiple sources. These are in addition to shippers vertically integrating into ship ownership and operations.

The shipping industry plays an essential role in global trade. For shipping companies, having an accurate assessment of the markets is essential to their marketing and risk management strategies. A number of important determinants impact ocean rates. Most important are fuel costs, in addition to distance, and time-in-transit. Stopford [2008] indicated that the single most important cost category is fuel costs, and more recently Hellenic-Shipping-News [2022] indicated that 47% of the cost of dry-bulk shipping was fuels costs, and it is expected these would be of increasing importance (and which will be exacerbated by decarbonization initiatives). As a result of the escalation in fuel costs in mid-2020, there were radical changes in ocean rates and spreads that impact shipping decisions [Wilson et al., 2022]. Of interest, are the change in rates between the US Gulf and PNW to Japan (one of the most commonly quoted ocean rates spreads), in addition to changes on rate spreads from the Black Sea, among others. Ocean freight costs, of which ocean rates comprise an critical component, were also impacted by the post-COVID-19 recovery and its multitude of impacts on supply chains. While most of the attention has focused on container ships and port congestion, there were several important impacts on bulk shipping. One is the increase in fuel prices as described above. In addition, congestion, primarily at ports, had the impact of increasing ocean rates. In normal times, ships are allowed time for loading and if exceeded, demurrage is charged. Typically, this would compensate the ship owner for the loss related to idled capacity. Under the post-COVID recovery, this relationship was compounded. Frequent situations around the world, resulted in ship-wait times being elongated in part by COVID-related protocols (e.g., China, Brazil), and more recently in the Black Sea due to operations of the Grain Corridor.

The impact of increased wait times was to have a consequential reduction in effective world ship capacity. This resulted in increasing ocean rates. Shih [2022] described how congestion has the effect of reducing capacity and suggested that the rise in freight rates was due to the Chinese lockdown (a two-stage lockdown which disrupted land-based logistics which adversely impacted port congestion) which exacerbated port congestion. Specifically, Sadden [2022] indicated there were “300 vessels waiting outside Shanghai, up five-fold from just two and a half weeks ago, with around 125 dry bulk carriers included among them.”
As a result of these developments, traders became exposed to abnormally large changes in ocean rates either when bidding in export tenders, or post-tender but prior to shipping. Both of these make risk management strategies important for ocean freight [Alizadeh and Nomikos, 2009]. Specifically, traders become exposed to change in ocean rates either when bidding in export tenders, or post-tender but prior to shipping.

3. Model Framework

The objective of this paper is to predict ocean rates at two primary ports of the United States- US Gulf and US Pacific Northwest (PNW) - to selected destinations, including China, Indonesia, Japan, and South Korea. The ocean rate, which is the target variable, is dependent on several feature variables. We discuss this in detail in the next section. We propose a general mathematical model for this analysis. We label US Gulf and US Pacific Northwest by i = 1 and j = 2 respectively, and the destinations (China, Indonesia, Japan, and South Korea) by j = 1, 2, 3, 4. The model is motivated by the previous studies [Awasthi and SenGupta, 2021, Awasthi et al., 2022].

We denote the ocean rate from i to j as S_t^(i,j) . For i = 1,2, and j = 1,2,3,4, we we model:

$$S_t^{i,j} = S_0^{i,j} e^{X_t^{i,j}}, \text{ where } dX_t^{i,j} = b_t^{i,j}\, dt + \sum_{k=1}^{n} \theta_t^{(i,j,k)} \left(\sigma_t\, dW_t^{(i,j,k)} + dJ_t^{(i,j,k)}\right), \qquad (1)$$

where $b_t^{i,j}$ is a deterministic function of t, $W_t^{(i,j,k)}, k = 1, \ldots, n,$ are independent Brownian motions and $J_t^{(i,j,k)}$ is the jump process with intensities $\lambda_k, k = 1, \ldots, n.$ We assume that $W_t^{(i,j,k)}$ and $J_t^{(i,j,k)}$, for $k = 1, \ldots, n,$ are independent. The coefficients $\theta_t^{(i,j,k)}$, at every t satisfy $\sum_{k=1}^{n} \left(\theta_t^{(i,j,k)}\right)^2 = 1$. In addition to that, σ_t is assumed to be stochastic, and its dynamics is governed by

$$d\sigma_t = F(\sigma_t^2, \beta_t^{(1)} H_t^{(1)}, \beta_t^{(2)} H_t^{(2)}, \ldots, \beta_t^{(n)} H_t^{(n)}), \tag{2}$$

for an appropriate function F, where $H_t^{(k)}$, for $k = 1, \ldots, n$ are jump processes with intensities

$\mu_k, k = 1, \ldots, n$. The coefficients $\beta_t^{(k)}$, at every t satisfy $\sum_{k=1}^{n} \left(\beta_t^{(j)}\right)^2 = 1$

We observe, that for a fixed i and j, if $\sum_{k=1}^{n} \left(\theta_t^{(i,j,k)}\right)^2 = 1$ and $\sum_{k=1}^{n} \theta_t^{(i,j,k)}$, $dW_t^{(i,j,k)}$ can be represented by $dB_t^{(i,j)}$, where $B_t^{(i,j)}$ is a Brownian motion. Consequently, (1) can be written as

$$S_t^{i,j} = S_0^{i,j} e^{X_t^{i,j}}, \text{ where } dX_t^{i,j} = b_t^{i,j}\, dt + \sigma_t\, dB_t^{(i,j)} + \sum_{k=1}^{n} \theta_t^{(i,j,k)}\, dJ_t^{(i,j,k)}, \tag{3}$$

The expression (3) provides an alternative explanation for the coefficients $\theta_t^{(i,j,k)}$. These coefficients pick up the fluctuations due to the “unusual” jumps. When we analyze the empirical data, these terms contribute to the detection of unusual or significant feature variable activities.

4. Support Vector Machine Regression

Support Vector Machine (SVM) is one of the most popular machine learning tools used both for classification and regression contexts. We use SVM regression to analyze ocean rates. SVM regression is generally considered as a non-parametric technique given it’s reliance on the kernel functions [MathWorks, 2022]. Finding the optimum fit line is the fundamental tenet of SVM regression.

Given the training data with relevant features, X_ijtand Y_ijt representing the outcome variable, that is, ocean rates for each origin (i)and destination (j), the linear function of SVM regression is shown below [Fan et al., 2005, Chen et al., 2006].

Y =f(X)= X'β + b (4)

The goal of the SVM regression is to formulate the above setup as a convex optimization problem to minimize

$$J(\beta) = 1/2\ \beta'\beta$$

subject to $|Y\_n\ -X'\beta+b| \leq \epsilon$, where $|Y\_n\ -X'\beta+b|$ is the residual of the linear SVM regression function specified in the equation (4) and ϵ is the threshold value [MathWorks, 2022]. The SVM Regression matches the best line within a threshold value, in contrast to other regression models that aim to reduce the error between the actual and predicted value. The distance between the boundary line and the hyperplane is the threshold value. While linear SVM Regression merely takes into account the linear kernel, it offers a faster implementation than non-linear SVM regression. This is because the samples' whose prediction is close to the objective are ignored by the cost function and the resulting model created relies on a portion of the training data.

Some advantages of using SVM regression include 1) it's robustness to anomalies, 2) it's simplicity in updating the decision model, 3) it's strong capability for generalization and predictive accuracy, and 4) it's simple to use empirically. Finally, a few disadvantages using SVM regression include 1) it's use is not appropriate for huge datasets, 2) it does not function well when there is noise in the training data.

5. Regularized Regression

Regularization is a technique to discourage learning complex models and thus help to avoid over-fitting, which is often present in regression models and neural networks. The main idea of regularization is to constrain or shrink the coefficient estimates towards zero. Empirically, if there is noise in the training data, then the estimated coefficients won't generalize well to the new data. In this scenario, regularization plays a key role in shrinking the coefficients towards zero discouraging the learning of noise. There are three main types of regularization, including L1 regularization (LASSO), L2 regularization (ridge regression), and Elastic net (combines both LASSO and Ridge regression) [James et al., 2013].

Given the linear relationship between the outcome (Y) and predictor (X_t) variables, the regression equation is written as follows:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \ldots + \beta_p * X_p \qquad (5)$$

The residual sum of squares (RSS) is the loss function used during the model fitting process. This loss function's minimization is achieved by selecting the coefficients.

$$RSS = \sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}(\beta_j x_{ij}))^2 \qquad (6)$$

Least Absolute Shrinkage and Selection Operator (LASSO) regression consists of adding a shrinkage term to RSS, which is the L1 norm of the coefficient vector. The LASSO regression penalizes the large coefficient more severely and thereby results in the coefficients becoming zero [James et al., 2013].

$$RSS + \|\vec{\beta}\|_1 = RSS + \lambda \sum_{j=1}^{p} \backslash \beta_j \backslash \qquad (7)$$

LASSO regression can be examined in a different perspective. When summation of the squares of the coefficients is less than or equal to s, the ridge regression (described later) can be viewed as the solution to the equation. Additionally, the LASSO can be viewed as an equation where sum of the coefficients' moduli is smaller than or equal to s. In this case, the constant s holds true regardless of the shrinkage factor's value.

Consider there are two parameters in a given problem. Then according to above formulation, the ridge regression is expressed by β_1^2 + β_2^2≤ s. This implies that ridge regression coefficients have the smallest RSS(loss function) for all the points that lie within the circle given by β_1^2 + β_2^2≤ s. Similarly, for LASSO, the equation becomes, |β_1 | + |β_2 | ≤ s. This implies that LASSO coefficients have the smallest RSS (loss function) for all points that lie within the diamond given by |β_1 | + |β_2 | ≤ s.

Ridge regression adds a shrinkage term to the RSS objective function as shown below

$$RSS + \|\vec{\beta}\|_2 = RSS + \lambda \sum_{j=1}^{p} \beta_j^2 \qquad (8)$$

The RSS of the ridge regression is changed by including the shrinkage term. This RSS function is minimized in order to estimate the coefficients. If we want to reduce the size of the above function, then the coefficients of a model that represents an improvement in flexibility must be minimal by limiting the upward trend of coefficients. The estimated associations of each variable with the response have been shrunk, with the exception of the intercept.

The estimates generated by ridge regression are equal to least squares estimates when λ= 0, as the penalty term has no impact at that point. However, as, the shrinkage penalty's effect increases and the estimations of the ridge regression coefficients get closer to zero. Therefore, choosing an appropriate value for is crucial. Combining techniques such as time series cross validation and hyper-parameter tuning is useful to compute the appropriate value of λbased on the chosen accuracy metric. The L2 norm refers to the coefficient estimates generated by this procedure.

As shown in the equation (8), the shrinkage term uses the L2 norm of the coefficient vector. λis the regularization hyper-parameter which gives the flexibility to set the amount of of model complexity to be penalized. If λequal zero, then it is equivalent to the original RSS function. As a requirement it is a general practice to standardize the regressor or predictors before applying ridge regression.

Finally, the Elastic net linearly combines both the L1 and L2 penalties of the LASSO and ridge regression methods as shown in the following equation.

$$RSS + \lambda_1 \|\vec{\beta}\|_1 + \lambda_2 \|\vec{\beta}\|_2 \qquad (9)$$

In the above equation, some of the special cases include the following [James et al., 2013]:

- $\lambda_1 = \lambda, \lambda_2 = 0$: LASSO regression
- $\lambda_1 = 0, \lambda_2 = \lambda$: Ridge regression
- $\lambda_1 = \lambda_2 = 0$: ordinary least squares (OLS)

The similarity and differences between the LASSO and ridge regression [James et al., 2013] are important. A few similarities include 1) both the methods improve generalization by penalizing model complexity, 2) their computational complexity is quite similar, and 3) penalization hyper-parameter λ must be carefully set using hyper-parameter tuning. Differences between the LASSO and ridge regressions include 1) ridge regression shrinks large coefficients but does not perform feature selection, 2) LASSO regression performs both shrinkage and feature selection.

The bias-variance trade-off is critical for model learning and model generalization. Bias is generally related to model learning and caused due to incorrect assumptions in the model. For instance, if you fit a linear model to variables that are associated non-linearly, then the model suffers from high bias. Bias is indicated by the training error. In contrast, variance is related to model generalization. For instance, if the variance of a least squares model is high then we conclude that the model does not generalize well to the new data. Variance is obtained by subtracting training error from test error.

Regularization dramatically lowers the model's variance while maintaining or even in- creasing its bias depending the tuning parameters. The bias-variance trade-off is partly controlled by the tuning parameter, which is employed in the regularization methods. As the value of λ increases, the coefficients' values decrease, lowering the variance. Up to a certain degree, this rise in λ is advantageous because it reduces variance (avoiding over-fitting), without losing any significant data features. However, after a certain value, the model begins to lose crucial characteristics, leading to bias and under-fitting. Therefore, the value of should be carefully selected using time series cross validation and hyperparamter tuning.

For our empirical analysis, we include two hyperparameters related to regularization methods, including 1) penalty (λ) term that was already discussed and 2) mixture term that specifies the type of regularization method from the LASSO (mixture = 1), ridge (mixture = 0), and Elastic net (mixture = 0.5) [James et al., 2013].

6. Data

In this paper, our goal is to predict ocean rates from two US origins (US Gulf and US Pacific Northwest) to four destinations (China, Indonesia, Japan, and South Korea). Many factors contribute to the prediction of ocean rates such as supply and demand of ships or vessels, crude oil prices, commodity prices, and global trade and its volume indices etc. Additionally, supply chain pressure or port congestion indices play a major role as well and has been especially critical during the COVID time. Therefore, our target or dependent variable is the ocean rate ($/Metric Ton) from each of the US origin to each destination.

In order to predict the ocean rates, features included global supply chain pressure index, Baltic exchange dry index, Brent crude oil price, total bulker sales, time charter rates, order book percent fleet, global trade index, global trade volume index, global bulk trade volume index, and commodity (corn) price. All these features are same for US Gulf and US PNW models, except the commodity price. We explored other variables such as port congestion index, idle number of ships at the port, and dry bulk trade index etc but could not include them in our analysis either due to the data being incomplete for our period or due not strongly associated with the ocean rates. In our models, congestion is captured by the global supply chain pressure index.

We use daily data between January 2015 and May 2022. Data are from three sources, including Thomson Reuters Eikon [TR-Eikon, 2022], Clarkson's Research Portal [Clarkson, 2022], Federal Reserve Bank of New York [NY-Fed, 2022]. We collected data on ocean rates, Baltic exchange dry index, Brent crude oil price, US corn prices were from Thomson Reuters Eikon. Other variables such as total bulker sales, time charter rates, order book percent fleet, US corn price, global trade index, global trade volume index, global bulk trade volume index were collected from Clarkson's Shipping Intelligence's Research Portal. Finally, global supply chain pressure index was collected from New York Federal Reserve. Summary statistics of the data variables are presented in Table 1.

Table 1: Summary Statistics

Characteristic	**US_Gulf**, N = 16,242[1]	**US_PNW**, N = 10,828[1]
Ocean Rate	37 (15), 10, 90	23 (7), 10, 47
Global Supply Chain Pressure Index	0.79 (1.37), -0.81, 4.38	0.79 (1.37), -0.81, 4.38
Baltic Exchange Dry Index	1,369 (870), 290, 5,650	1,369 (870), 290, 5,650
Brent Crude Oil Price	60 (16), 19, 128	60 (16), 19, 128
Total Bulker Sales	729 (415), 31, 2,055	729 (415), 31, 2,055
Time Charter Rates	12,443 (6,043), 4,750, 30,400	12,443 (6,043), 4,750, 30,400
Order Book Percent Fleet	11.6 (4.0), 7.0, 22.9	11.6 (4.0), 7.0, 22.9
US Corn Price	78 (32), 26, 205	205 (54), 113, 383
Global Trade Index	4 (9), -14, 27	4 (9), -14, 27
Global Trade Volume Index	122 (14), 92, 157	122 (14), 92, 157
Global Bulk Trade Volume Index	109 (7), 90, 121	109 (7), 90, 121

[1] Mean (SD), Range

Notes: Range is the spread of the data from the lowest (minimum) to the highest (maximum) value in the distribution

We conducted a series of data preprocessing steps as the values range differently for dif- ferent features and to make it ready for use in the machine learning models–regularized regression and support vector machine regression. Specifically, we conducted Z-score nor- malization, which is robust for any potential anomalies/outliers in the time series. We also preprocessed the data for features containing near zero variance, and created a set of dummy variables for the destinations (to account for unobserved factors for the origin-destination combination) and month (for capturing seasonality in ocean rates).

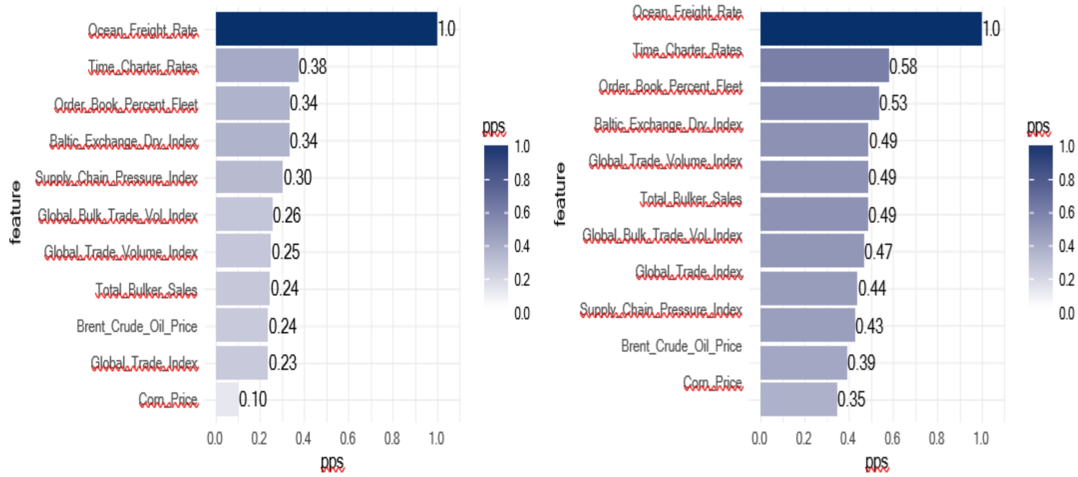
We split the data into training and test sets by taking into account the time dependency. Specifically, the training set includes entire data between January, 2, 2015 and May, 31, 2021 in each of the origin-destination combinations which the test set includes the between June, 1, 2021 and May, 31, 2022 (last 365 days) of data for each origin-destination combinations. Therefore, for both the US Gulf and US PNW, total observations in the training set includes 9368 observations while the test set includes 1460 (365×4 [destinations]) for each origin.

7. Results and Discussion

We first describe the exploratory data analysis, results of both machine learning models, including regularized regression and support vector machine regression. We then discuss the results of the study.

As part of the exploratory data analysis, instead of using correlations between features and the target variable, we used predictive power score of each feature on the target variable [Wetschoreck, 2020, van der Laken, 2021]. The predictive power score is relatively new technique and developed as an alternative to the correlation matrix [Wetschoreck, 2020] and has three main advantages. First, the correlations between any two variables are symmetric while predictive power score is asymmetric. That is, the correlation relationship between two variables are same if you flip the axis (symmetric) while the relationship may not be the same in reality (asymmetric). Second, the predictive power score works for both numeric as well as categorical features unlike correlation matrix. Finally, predictive power score allows non-linear relationships between the feature and the target variable. The predictive power score captures the predictive power of each feature on the target variable on a scale of 0 to 1 with the values close to 1 indicating high predictive power while the values close to zero indicating low predictive power.

Results of predictive power scores are presented in Figure 2 for both the US Gulf (panel a) and US PNW (panel b). There are four main takeaways from the results of predictive power scores. First, the predictive power scores are relatively high for the ocean rates at US PNW compared with the US Gulf. Second, the top three features, including time charter rates, order book percent fleet, Baltic exchange dry index, with high prediction scores are same for the both the origins. Third, the global supply chain pressure index ranked much higher in the case of US Gulf compared with the US PNW. Finally, the commodity price



(a) US Gulf (b) US Pacific Northwest (PNW)

Figure 2: Predictive Power Score (pps)

(Corn Price) ranked the lowest among all the features considered in predicting ocean rate for both the origins. Overall, the results of these predictive power scores indicate that time charter rates, order book percent fleet, Baltic exchange dry index play an important role while commodity price play a minimal role in predicting ocean freight.

We specified two machine learning models: the regularized regression and support vector machine regression. Before finalizing on the two models, we explored several other machine learning models, including ensemble models such as random forest and gradient boosting models, which are based on the decision/regression trees as the base learners. But, these ensemble models suffered from overfitting for these data highlighting the importance that the use of the machine learning model specification depends on the characteristics of the data.

Predictive accuracy (rmse) results for both the machine learning models by each origin and destination are presented in Table 2. These results are based on the best (lowest) root mean square error (rmse) metric. We considered several other accuracy metrics, including mean absolute error (mae), mean absolute percentage error (mape), symmetric mean ab- solute percentage error (smape), mean absolute scaled error (mase), and R-squared (rsq). We chose rmse as our accuracy metric due to its interpretability and is one of the common accuracy metrics used in the regression-based supervised predictive models. However, the results of the other accuracy metrics are similar to the rmse results (and hence not shown in the results).

Table 2: Root Mean Square Error (rmse): Prediction Accuracy Results by Origin and Des- tination

		Model			
		Regularized Regression		Support Vector Machine Reg.	
Origin	Destination	Train	Test	Train	Test
US Gulf	China	2.65	4.27	2.27	4.52
	Indonesia	2.89	4.22	2.52	5.15
	Japan	2.98	4.58	2.62	4.67
	South Korea	2.37	5.38	2.25	7.15
US PNW	China	1.66	3.16	1.62	3.36
	Indonesia	1.58	3.20	1.22	2.35
	Japan	1.76	2.50	1.46	3.46
	South Korea	1.23	1.75	1.07	2.08
US Gulf	Overall	2.73	4.64	2.42	5.48
US PNW	Overall	1.57	2.72	1.36	2.88

The rmse for the regularized regression models for the US Gufl to China indicate that the ocean rate predictions are off by \$2.65/MT on average in the training data while they are off by \$4.27/MT on average in the new (test) data. Similarly, between US Gulf and China, the support vector machine regression model rmse results indicate that ocean rate predictions are off by \$2.27/MT in training data while they are off by \$4.52/MT in the test data. As the variance (\$4.27 - \$2.65 = \$1.62) is lower in regularized regression compared with the variance (\$4.52-\$2.27 = \$2.25) of support vector machine regression, we conclude that the regularized regression has generalized (slightly) well in predicting new ocean rates between US Gulf and China.

The bias-variance trade-off is important to consider when evaluating the performance of the machine learning models. Bias-variance trade-off is essentially the trade-off between model learning and model generalization as it ensures that the models do not overfit. A model that overfits is very sensitive to the small changes in training data and does not generalize to the new data. An ideal model should have less bias and less variance.

Predictive accuracy (rmse measures in Table 2) show that the regularized regression results consistently performed well in the test set when compared with the results of support vector machine regression. The training error approximates the bias while the difference in the test and training error indicates the variance. The bias is less if the training error is small. The model will have lower variance if the difference between test and training error is small. Based on the results, the US PNW has less bias and less variance compared with the results of the US Gulf. The training error shows the approximate estimate of the bias, which signifies how well the model is learning while the variance is the difference between test error and training error. Variance shows whether a model is learning the patterns or the noise. For instance, if the variance is high, then the model overfits, which means that the model is learning noise. In contrast, if the model variance is low, then the model does not overfit, which means that the model is learning patterns.

Figure 3 shows the predictions for the test set of both the models (regularized regression is shown as *1_GLMNET* while SVM regression is shown as *2_LIBLINEAR* in Figure 3) compared with the actual observations in the case of US Gulf. Overall, the predictions for all destinations have higher accuracy except for South Korea, which has greater confidence bands compared with all other destinations. Results are similar in the case of US PNW but with a greater accuracy in the predictions as outlined in the Table 2 as well.

8. Time Series Cross Validation

Test error can vary depending on the data split and the sample at hand. Cross validation, popularly known as K-fold cross validation, is widely used approach for estimating test error. The idea of K-fold cross validation is to randomly divide the data into K (roughly) equal-sized sets leaving one fold aside [James et al., 2013]. The training is performed on K-1 number of folds and tested on the left out fold. This process is repeated K times to compute the estimates of test error K times. Finally, the results are combined to compute the accuracy measures of interest.

Since our data has time stamps, the usual cross validation may not be appropriate be- cause of its random split of data. As there is a need to maintain time dependency of the observations, we instead use time series cross validation. Time series cross validation involves dividing the whole data set into a K-folds or slices to perform the training in the first slice and test the model in the consequent slice [Hyndman and Athanasopoulos, 2018].

Results of time series cross validation are shown in Figure 4, which shows the accuracy

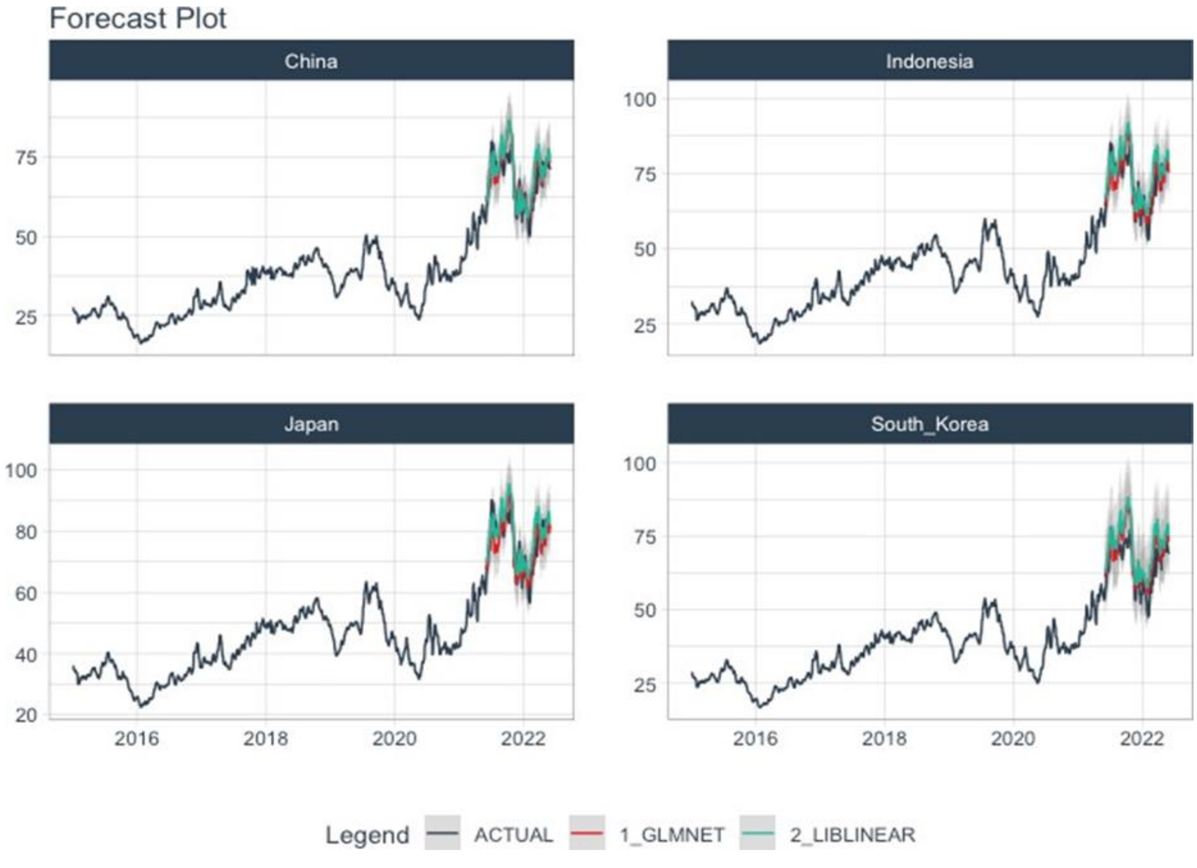


Figure 3: US Gulf: A Comparison of Actual versus Predictions on Test Data by Model and Destination

metrics for all the 10 slices of data samples. The average results of various predictive accuracy measures indicate that the regularized regression (shown as *1_GLMNET* in Figure 4) performs consistently (and slightly) better than the SVM regression model (shown as *2_LI-*

BLINEAR in Figure 4) across all the slices. These results are consistent compared with the results shown in Table 2.

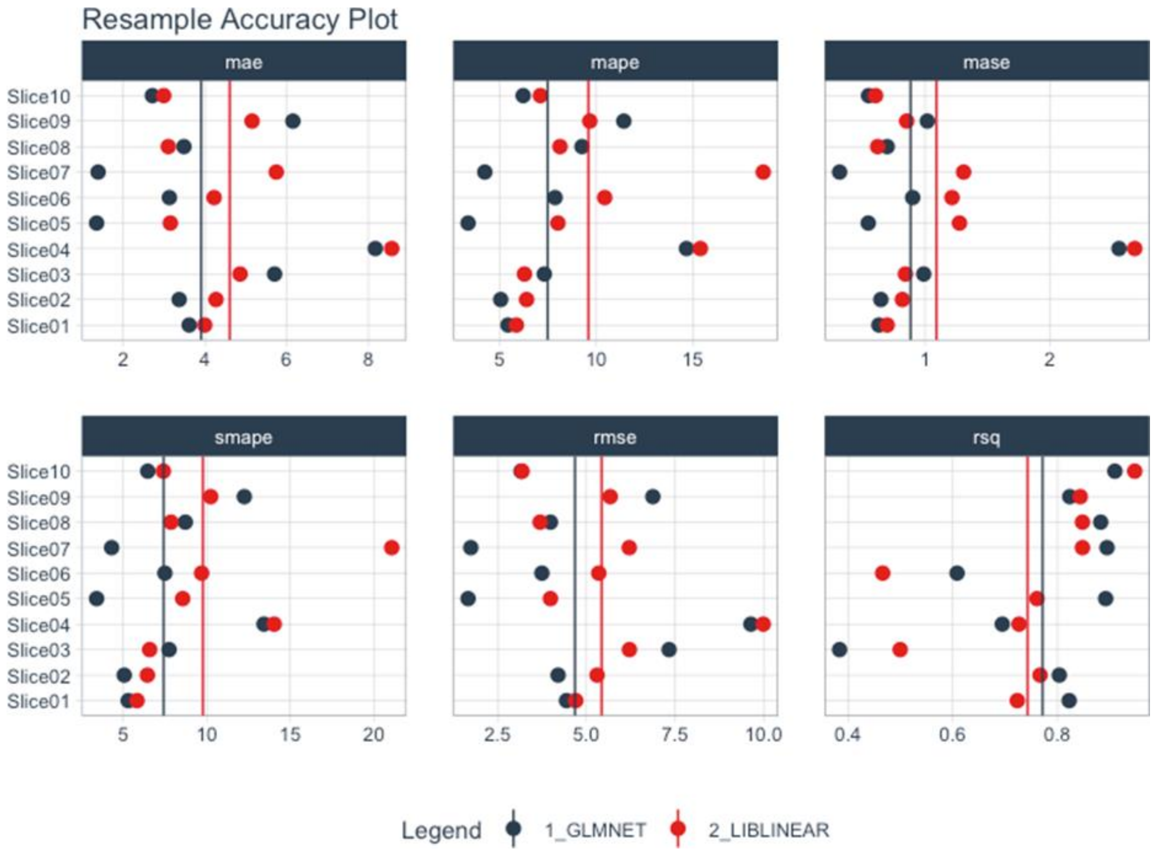


Figure 4: US Gulf: Time Series Cross Validation Resamples: Accuracy Metrics

9. Hyperparameter Tuning

Hyperparameter tuning is the procedure for finding the optimal values for hyperparameters in a machine learning model by evaluating predictive accuracy based on a selected accuracy metric. Hyperparameters are specified before the training of the model as opposed to model parameters which are obtained after the training procedure in the form of weights or the coefficient estimates of the features.Hyperparameter tuning can significantly affect the predictive accuracy of a machine learning model. Instead of evaluating machine learning model using default or a single set of hyperparameters, it is advised to evaluate over a wide range in order to obtain optimal values at which the accuracy of the model is the best.

Each machine learning model has its own hyperparameters for tuning. In the case of regularized regression, there are two hyperparameters, including a *penalty* term for regular- ization and a *mixture* term for the type of model (*mixture* = 0, ridge regression; *mixture*

= 1, LASSO regression; *mixture* = 0.5, elastic net regression) to be specified [James et al., 2013, Kuhn et al., 2013]. In the case of support vector machine regression, there are two hyperparameters for tuning, including the *cost* and *margin* [James et al., 2013, Kuhn et al., 2013].

For each model, a tuning grid is specified through a random search of 1000 values for each hyperparameter of the model [Kuhn et al., 2013]. Results of the hyperparameter tuning are shown in Table 3. The hyperparameter values are different for both the US Gulf and US PNW models, which indicate that the hyperparameter values depend on the dataset and their tuning procedure help to improve the accuracy of the model.

Table 3: Hyperparameter Tuning Parameters Based on the Lowest RMSE

Model	Hyperparameter	US Gulf	US PNW
Regularized Reg. Support Vector Machine Reg.	Penalty Mixture Cost	0.880 0.374 0.126	0.372 0.554 0.015
	Margin	7.1×10^{-5}	0.125

10. Summary and Concluding Remarks

Ocean rates increased recently between 150% and 175% depending on the shipping route. This unprecedented growth in ocean rates has coincided with COVID and post COVID era. Predicting ocean rates is important given it's contribution to the total cost of transportation of the bulk commodities. For example, ocean rates account for about 35-50% of the total shipping costs of soybean shipments from the origins such as the United States and Brazil to destination such as China [USDA, 2022].

The purpose of this paper is to to predict ocean rates for dry-bulk from the US Gulf and US Pacific Northwest (PNW) origins to selected destinations using machine learning speci- fications. We specify regularized regression and support vector machine regression methods to analyze and predict ocean rates using daily data obtained on relevant features between January 2015 and May 2022. In order to predict ocean rates, we use features, including global supply chain pressure index, Baltic exchange dry index, Brent crude oil price, total bulker sales, time charter rates, order book percent fleet, commodity (corn) price, and global trade indices. Specifically, for predicting ocean rates, we analyzed US origins (US Gulf and US Pacific Northwest) to selected destinations, including China, Indonesia, Japan, and South Korea. For improving model selection and its performance, we used time series cross validation and hyperparameter tuning. We used root mean square error as the accuracy metric for selecting the hyperparameters in case of both the models.

Predictive power score results indicate that time charter rates, order book percent fleet, and Baltic exchange dry index have higher predictive power scores compared with the others in predicting ocean rates highlighting their importance. While commodity (corn) price con- tributes the least relatively in predicting the ocean rates. Predictive accuracy results show that both the regularized regression and support vector machine regression models performed well in predicting ocean rates, however, regularized regression method yielded slightly bet- ter predictive accuracy both in terms of bias and variance compared with support vector machine regression. In the case of regularized regression, the mixture type indicate that the models are close to representing Elastic net rather than the LASSO and ridge regressions.

The methodology and results of this study have implications for both commodity analysts and for trading and shipping companies in particular. Specifically, ocean rates are volatile, and their volatility has escalated in recent years. As a result, prediction of ocean rates is integral to understanding commodity flows and competition. The models developed in this paper provide superior predictions, and the regularized regression models in particular performed better than the support vector machines. Second, the increased volatility of ocean rates results in greater risks to trading firms. Ultimately, trading firms have to find ways to mitigate these risks. The risk of changes in ocean rates can be managed either through including predicted rates in formulated bids to importers, and/or using forward contracts for ocean rates. Additionally, in some cases, trading firms may facilitate ‘switching’ options (as discussed in [Meersman et al., 2012] and quantified in [Johansen and Wilson, 2019]) to their import customers in anticipation of predicted changes in ocean shipping rates.

11. Declarations

Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest

Ethical Approval

On behalf of all authors, the corresponding author affirm that all procedures followed in this article subscribe to relevant ethical rules.

Funding

Funding for this study was from the Center of Risk and Trading at [University Name with- held] and a major international grain trading firm**.**

Data Availability

Data are available from the various sources as listed in the Data section of this article. In some cases there may be some limits on sharing to non-subscribers. Upon request, we can and will provide the data formatted as used in this study and the results reported in this submission.

References

[1] Michael K Adjemian and William W Wilson. Implications of rising ocean freight rates for agri-food product markets. usda agricultural marketing service report. Technical report, 2022.

[2] Amir Alizadeh and Nikos Nomikos. Shipping derivatives and risk management. Springer, 2009.

[3] Shantanu Awasthi and Indranil SenGupta. First exit-time analysis for an approximate barndorff-nielsen and shephard model with stationary self-decomposable variance process. Journal of Stochastic Analysis, 2(1):5, 2021.

[4] Shantanu Awasthi, Indranil SenGupta, William Wilson, and Prithviraj Lakkakula. Machine learning and neural network based model predictions of soybean export shares from us gulf to china. Statistical Analysis and Data Mining: The ASA Data Science Journal, 15 (6):707–721, 2022.

[5] Arunava Bandyopadhyay and Prabina Rajib. The asymmetric relationship between baltic dry index and commodity spot prices: evidence from nonparametric causality-in-quantiles test. Mineral Economics, pages 1–21, 2021.

[6] Limon Barua, Bo Zou, and Yan Zhou. Machine learning for international freight transportation management: a comprehensive review. Research in Transportation Business & Management, 34:100453, 2020.

[7] [Pai-Hsuen Chen, Rong-En Fan, and Chih-Jen Lin. A study on smo-type decomposition methods for support vector machines. IEEE transactions on neural networks, 17(4):893– 908, 2006.

[8] Clarkson. Shipping intelligence network. https://sin.clarksons.net, 2022.

[9] Drewry. Covid-19 drives up ship operating costs. https://www.drewry.co.uk/dry-bulk-shipping, 2020. Accessed: 2022-11-30.

[10] Drewry. Russia-ukraine war to crash dry bulk freight rates. https://www.drewry.co.uk/dry- bulk-shipping, 2022. Accessed: 2022-11-30.

[11] Payman Eslami, Kihyo Jung, Daewon Lee, and Amir Tjolleng. Predicting tanker freight rates using parsimonious variables and a hybrid artificial neural network with an adaptive genetic algorithm. Maritime Economics & Logistics, 19(3):538–550, 2017.

[12] Rong-En Fan, Pai-Hsuen Chen, Chih-Jen Lin, and Thorsten Joachims. Working set selection using second order information for training support vector machines. Journal of machine learning research, 6(12), 2005.

[13] Freight-News. Taiwan's mfig buys about 55,000 tonnes corn in tender -traders. , 2022a. Accessed: 2022-08-29.

[14] Freight-News. Algerian wheat buy on wednesday still seen at around 500,000 tonnes - traders., 2022b. Accessed: 2022-08-29.

[15] Esther Fung, Costas Paris, and Sharon Terlep. Supply chains upended by covid are back to normal. https://www.wsj.com/articles/supply-chains-upended-by-covid-are-back-to-normal-11671746729, 2022. Accessed: 2022-12-26.

[16] Lazaro Gamio and Peter. S. Goodman. How the supply chian crisis unfolded. https://www.nytimes.com/interactive/2021/12/05/business/economy/supply-chain.html, 2021. Ac-cessed: 2022-06-26.

[17] Qianqian Han, Bo Yan, Guobao Ning, and B Yu. Forecasting dry bulk freight index with improved svm. Mathematical Problems in Engineering, 2014, 2014.

[18] Saraswathi Hathikal, Sung Hoon Chung, and Martin Karczewski. Prediction of ocean import shipment lead time using machine learning methods. SN Applied Sciences, 2(7):1–20, 2020.

[19] Hellenic-Shipping-News. Shipping fuels, already a major component of shipping's operating costs, will only increase in the future. https://www.hellenicshippingnews.com/shipping- fuels-already-a-major-component-of-shippings-operating-costs-will-only-increase-in-the- future/, 2022. Accessed: 2022-09-26.

[20] Rob J Hyndman and George Athanasopoulos. Forecasting: principles and practice. OTexts, 2018.

[21] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An introduction to statistical learning, volume 112. Springer, 2013.

[22] Stephan Johansen and William Wilson. Valuing switching options in international grain marketing. Agribusiness, 35(3):423–438, 2019.

[23] Kei Kanamoto, Yujiro Wada, and Ryuichi Shibasaki. Predicting a dry bulk freight index by deep learning with global vessel movement data. In Transdisciplinary Engineering for Complex Socio-technical Systems, pages 105–114. IOS Press, 2019.

[24] Max Kuhn, Kjell Johnson, et al. Applied predictive modeling, volume 26. Springer, 2013.

[25] Ana Mano. China becomes biggest Brazil corn buyer in January, revised trade data show. https://www.reuters.com/article/brazil-corn-china-idAFL1N34O2B3, 2023. [On-line;accessed 22-Feb-2023].

[26] MathWorks. Understanding support vector machine regression. https://www.mathworks.com/help/stats/understanding-support-vector-machine-regression.html, 2022.

[27] Steven Meersman, R Reichtsteiner, and Graham Sharp. The dawn of a new order in com- modity trading. Oliver Wyman Risk Journal, 2(1):67–75, 2012.

[28] Konstantinos D Melas and Nektarios A Michail. The relationship between commodity prices and freight rates in the dry bulk shipping segment: A threshold regression approach. Maritime Transport Research, 2:100025, 2021.

[29] NY-Fed. Global supply chain pressure index. federal reserve bank of new york. https://www.newyorkfed.org/research/policy/gscpi, 2022. Accessed: 2022-07-26.

[30] Euan Sadden. Freight rates rise as shanghai lockdown exacerbates port con- gestion. https://www.agricensus.com/Article/Freight-rates-rise-as-Shanghai-lockdown-exacerbates-port-congestion-21418.html, 2022. Accessed: 2022-09-26.

[31] Indranil SenGupta, William Wilson, and William Nganje. Barndorff-nielsen and shephard model: oil hedging with variance swap and option. Mathematics and Financial Economics, 13:209–226, 2019.

[32] Willy Shih. Seven things you should know to understand the supply chain. https://www.wsj.com/articles/what-to-know-supply-chain-11658152794, 2022. Accessed: 2022-09-26. Martin Stopford. Maritime economics 3e. Routledge, 2008.

[33] TR-Eikon. Thomson reutern eikon. , 2022. Accessed: 2022-06-26.

[34] Vangelis Tsioumas, Yiannis Smirlis, and Stratos Papadimitriou. Capturing the impact of economic forces on the dry bulk freight market. Maritime Transport Research, 2:100018, 2021. USDA. Grain transportation report. https://www.ams.usda.gov/sites/default/files/media/ GTR08112022.pdf, 2022. Accessed: 2022-09-26.

[35] Paul van der Laken. ppsr: Predictive Power Score, 2021. URL https://CRAN.R-project.org/package=ppsr. R package version 0.0.2.

[36] Florian Wetschoreck. Rip correlation. introducing the predictive power score. https://towardsdatascience.com/rip-correlation-introducing-the-predictive-power-score- 3d90808b9598, 2020. Accessed: 2022-12-16.

[37] William Wilson, William Nganje, Semere Gebresilasie, and Indranil SenGupta. Barndorff- nielsen and shephard model for hedging energy with quantity risk. High Frequency, 2(3-4): 202–214, 2019.

[38] William W Wilson, Prithviraj Lakkakula, and David W Bullock. Logistical competition for corn shipments from the united states and ukraine to targeted international market. Technical report, 2022.

[39] Zaili Yang and Esin Erol Mehmed. Artificial neural networks in freight rate forecasting. Maritime Economics & Logistics, 21(3):390–414, 2019.

Citation: Prithviraj Lakkakula, William W. Wilson, Shantanu Awasthi, Indranil SenGupta. (2025). Predicting Ocean Freight Rates Using Machine Learning Methods. Journal of Artificial Intelligence and Machine Learning (JAIML), 4(1), 1-26.

Abstract Link: https://iaeme.com/Home/article_id/JAIML_04_01_001

Article Link:
https://iaeme.com/MasterAdmin/Journal_uploads/JAIML/VOLUME_4_ISSUE_1/JAIML_04_01_001.pdf

editor@iaeme.com