Future LPG Shipments Forecasting Based on Empty LPG Vessels Data

Jou-Tzu Kao, Rong Liao, Hongxia Shi, Joseph Tsai, Shenyang Yang, Matthew A. Lanham
Purdue University, Department of Management, 403 W. State Street, West Lafayette, IN 47907
kao21@purdue.edu; rliao@purdue.edu; shi395@purdue.edu; tsai103@purdue.edu;
yang1469@purdue.edu; lanhamm@purdue.edu

ABSTRACT

This study assesses the feasibility of using information of empty liquefied petroleum gas (LPG) carrier vessels that are moving in the ocean to predict the how much LPG will be shipped in the future. As prices of multiple commodities fluctuate with the supply and demand of LPG, it is crucial to identify effective indicators of LPG commodity flow to the foresee future trends of the market. In order to conduct this analysis, we acquired access to the shipping schedule data of empty and full LPG vessels, and ran multiple types of regressions to understand the correlation between these two factors. Our analysis shows that we were able to build a valid and usable predictive model using ridge regression to predict the amount of LPG being shipped in the future.

Keywords: Predictive Analytics, Liquefied Petroleum Gas, Supply Forecast

INTRODUCTION

The main question that our study will answer is whether the flow of empty shipping movements can serve as an indicator of future flow commodity LPG shipments. The privilege of tracking and predicting commodity movements will cause phenomenal impact in the business world. In general, prices of commodities are decided by both supply and demand. Having insight into the flow of goods can allow businesses to better predict price changes and take advantage on both ends. For example, if oil producers know less oil will be ordered to ship in the coming future, they can communicate and control the amount of production before the products turn into excess inventory and are forced to encounter price wars with other companies. On the other hand, if oil traders know there will be less oil being provided in the future, they can formulate strategies to acquire sufficient amount in advance to the incremental changes in price. What we focus on in this study is the need of predicting future LPG shipments, and the information required on vessel movements.

1. The Need in Predicting Future Shipments

Whether companies build, buy or sell commodities, they normally rely on multiple sources of information with great effort to try to make the correct decisions. They utilize pieces of available information such as sales history and news. However, people are always in search of better indicators that can more quickly and precisely predict the future. This is where the value of data science and our study comes in. We are exploring the potentials of a new source of information which is vessel shipping flows. When companies need to charter, service or trade in commodities, better and quicker information can help them earn or save millions of dollars.

Utilization of shipping forecasts has been the key to success in multiple industries. Shipping industry professionals and commodity traders use the forecasts to optimize their investment, planning and market positioning decisions (IHS Markit 2016). Forecast in shipping is especially important in the liquefied petroleum gas (LPG) market, because there is a lot of variation in demand and price of LPG. Figure 1 provides an example of the variation in price and volume from 2014 thru 2017.

Figure 1. LPG Price fluctuations in the US from 2014 to 2017

According to an article published by Joeri van der Sman (2017), currently the market is in a downward state because transport vessels are supplying beyond demand. What is interesting is it is projected to have an impressive growth in demand in the near future for seaborne LPG trades.

Demand forecasts are also very important in the trading and investment markets. One major differentiator among good and bad investment or commodity traders are how well they can find reliable information and turn it into actionable knowledge. Being able to forecast the future trend of a commodity, such as LPG, they will be able to anticipate and take advantage of the demand and future price changes of their assets.

Misleading forecasts can result in poor decision-support and lead to severe crises in all industries. China, which produces almost half of the world's steel, says it is pursuing painful reforms, but that the glut is at least partly the result of weaker global demand (Paul Page 2017). In fashion markets, inventory management also plays a huge role. After years of struggling to manage their inventory, vendors are using increasingly sophisticated tools to track apparel and accessories through the supply chain, aiming to avoid the typical post-Christmas fire sales (Stephanie Wong 2017).

2. The Sufficient Amount of Precise Information on Vessel Movements

It is important today for commodity traders to dig deeper into shipping flows as sufficient amount of shipping data becomes available. Maritime transportation is known for having rich information in terms of volume. Most of the transportation information is utilized in supporting the global commodity supply chain. Maritime surveillance data are increasingly used to achieve a high-level picture of situational awareness. Cooperative self- reporting vessel location systems, including the Automatic Identification System (AIS) and Long-Range Identification and Tracking system (LRIT), provide a great amount information about the vessels at port and at sea.

This study utilizes the power of data analytics to take full advantage of the shipment data and explore the possibility to using this information to predict future LPG demand, and ultimately predict the price of the commodity.

The remainder of this paper is organized as follows: A review on the literature on various criteria and methods used for vessel information detecting is presented in the next section. In this section, several practices that people use to forecast future shipment and commodity price is discussed and evaluated. In section 2, the data used in this analysis will be introduced. The proposed methodology is presented in section 4 and the assumptions and difficulties is discussed. In Section 5 various models are formulated, tested and evaluated. Section 6 outlines the performance of the models and summaries the prediction results. Section 7 concludes the paper with a discussion of the business insights draw from this study, future research directions, and concluding remarks.

LITERATURE REVIEW

In the pursuit to predict the future amount of shipments of LPG, three logical questions need to be addressed: Whether empty ships are a good indicator of future demand? If so how can the information be used to predict how much LPG will be shipped in the future? Lastly, how do the future shipments define the demand?

1. Vessel Shipment Detection

Studies on vessel surveillance, tracking, and prediction have been for various perspectives based on the data collected. Some have researched ship tracking using visual/image data. For example, how to track ships via visualizations (Robert-Inacio, Raybaud et al. 2007), visual classification, visual ship counting (Chen, Chen et al. 2008), and simultaneous detection and tracking (Hu, Yang et al. 2011). Chen, Chen et al. (2008) found that when trying to count ships using image data, ships in regions where wave ripples were prevalent resulted in inaccurate results. Moreover, wave ripples lead to inaccurate ship detection, leading to mistakes in ship classification. Fei, Qing et al. (2014) investigated video surveillance, specifically Closed-Circuit Television (CCT) video sequences to track a ship inland using a Kalman filter methodology. They found their approach was better than previous known approaches, that suffered from cluttered backgrounds and occlusion.

The information used as predictors in this study is the number of empty vessels and the tonnage capacity that the vessel can hold. When the shipment is at sea, the tonnage is different depending on the load of the shipment. Based on this information, the empty vessels can be detected. One problem associated with this is that it cannot be known if the vessel is empty or not, and it is challenging to detect the volume of the load. Using the volume as an indicator would be more accurate than the number of shipments in prediction because different vessels have different loading capacity. Future research is needed to address this problem.

2. Predictive Modelling

Studies found that are believed to be most related to this project entail trying to predict or understand future behavior or movement between successive AIS points (Hammond 2014). One of the well-known challenges in trying to predict future patterns of ships is the massive amount of information available. There are thousands and ships and millions of tracking points. This is overwhelming information for a human to process and synthesize in an effective manner. Riveiro (2011) points this out when he tries to address this problem by developing a visual analytics approach. Pallotta, Vespe et al. (2013) developed a methodology called Traffic Route Extraction and Anomaly Detection (TREAD) that uses an unsupervised and incremental learning approach on AIS data to detect path anomalies, and project current and future trajectories. As noted in their paper, some have tried to subdivide the area(s) of interest into spatial grids, whose cells might be "*characterized by motion properties of the crossing vessels*" (Vespe, Sciotti et al. 2008). The potential problem with grid-based approaches for small area surveillance is that they are hard if not impossible to scale. Another important consideration for path prediction is understanding of what is referred to as Course Over Ground (COG) turning points. Essentially these areas would be like intermediate nodes in a supply chain network, while entry and exit points are other nodes. Using an Ornstein-Uhlenbeck stochastic process, (Pallotta, Horn et al. 2014) improve their vessel predictions, by using historical AIS data to estimate parameters that are essential characteristics of recurrent routes. In other words, they exploit prior knowledge to predict the position of vessels with greater confidence.

Most of the research conducted to date, try to detect the future trajectories or spatial pattern of shipments.

Much less work has been performed to predict the future number of shipments by analyzing historical data. Spatial projections are important in maritime transportation control, but traders are more interested in how much of the commodity of interest will be transported in the future. Therefore, the focus in our study is to predict shipment pattern in terms of different time grids, i.e. conduct time-series analysis on the shipment data. By grouping the shipments by different time grids, it is possible to find a best time frame for future predictions. The accuracy of the prediction can be improved by putting extra information into the time-series model such as exponential smoothing and seasonality.

3. Demand Forecasting

Jan Tore Klovland (2002) points out there is a close timing relationship between cycles in economic activity, shipping rates and commodity prices. But the relation is very complex. To understand the theory behind this, people needs to have profound economic knowledge and empirical experience.

Instead of relying on theory, the method proposed in this paper is all based on facts available to the public. When a vessel receives a mission of transporting commodities, it may start its voyage to a loading port. Then the vessel is expected to carry commodities to the destination port. The number of empty shipments at sea at one point of time, may indicate the supply of the commodities at a future time point when the shipment arrives at the destination port. Further the supply, or availability of the commodity affects the price of the commodity at that specific market. Based on these series of assumptions, this study investigates the relationship between the drives by interpreting the shipment data.

As these handful of studies demonstrate, shipment time pattern recognition and future vessel behavior is a challenging problem. The key objective of this study is to manipulate the available vessel data, find relationships and patterns of the vessels and shipments. Also, develop a predictive model that can provide more accurate results and thoughtful insights of what will happen in the future. The key performance measure is how reliable the conclusion is in helping traders to uncover maritime market opportunities and make investment decisions.

DATA

For this study, we collaborated with an industry partner that mines this publicly available data. One table called Voyage_lpg provided information about the full ships, including vessel names, date and position of departure and port arrival. The Voyage_lpg_ballast table has similar information, but it is about empty ships as shown in Table 1. We created a more meaningful feature from this data called the average total empty vessels' capacity. The average total empty vessels' capacity is obtained by calculating the empty vessels' capacity from 27 days ago to 21 days ago of specific date as shown in Table 2. The justification of this time window came from trial and error and extensive exploratory data analysis.

file	Variable	Type	Description
voyage_lpg.csv	vessel	str	Name of vessel
	date_depart	date	Date of departing
	date_arrive	date	Date of arriving
	bbls	numberic	Bbls * Probability
voyage_lpg_ballast.csv	vessel	str	Name of vessel
	date_depart	date	Date of departing
	region_sink	str	Region of the sink
	date_arrive	date	Date of arriving
	date_load	date	Date of loading

Table 1: Variables from raw data

Variable	Type	Description
day	date	specific date
total_bbls	numberic	demand of specific date
total_empty_tonnage_21_days_ago	numberic	empty vessel 21 days ago* capacity
total_empty_tonnage_22_days_ago	numberic	empty vessel 22 days ago* capacity
total_empty_tonnage_23_days_ago	numberic	empty vessel 23 days ago* capacity
total_empty_tonnage_24_days_ago	numberic	empty vessel 24 days ago* capacity
total_empty_tonnage_25_days_ago	numberic	empty vessel 25 days ago* capacity
total_empty_tonnage_26_days_ago	numberic	empty vessel 26 days ago* capacity
total_empty_tonnage_27_days_ago	numberic	empty vessel 27 days ago* capacity
average_total_empty_tonnage	numberic	average total empty vessel capacity

Table 2: Data used for prediction models

METHODOLOGY

Predictive modeling-type project

Our target is to determine if there is any relationship between the demand of oil and information of empty ships three weeks prior. Figure 2 illustrates the overall methodology flow chart of our study. To follow, we describe how to create features, train models, evaluate models, and generate decision-support.

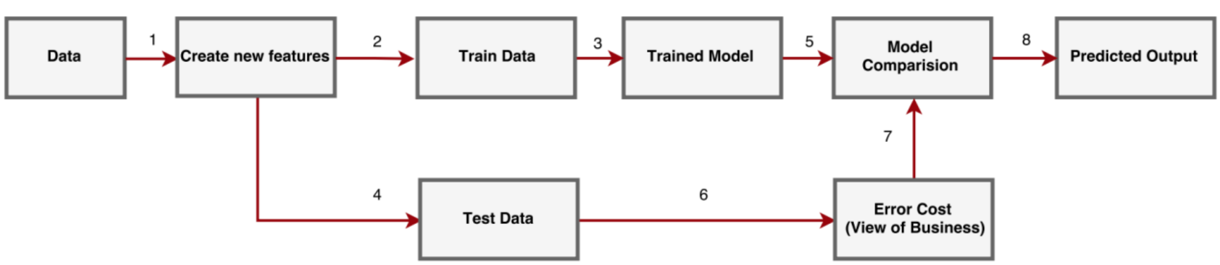

Figure 2. Flow chart of the project

Identifying significant features

We used correlation analysis to identify the most relevant features and reduce redundancy. The most significant feature is the average_total_empty_tonnage with a correlation coefficient of 0.88. As shown in Figure 3, there is a significant relationship between the total tonnages shipping today (y-axis) and the average_total_empty_tonnage (3 weeks prior). We then performed regression analysis to model the relationship between the two variables (i.e., total number of demands and the average tonnages number of empty ships before 21 to 27 days).

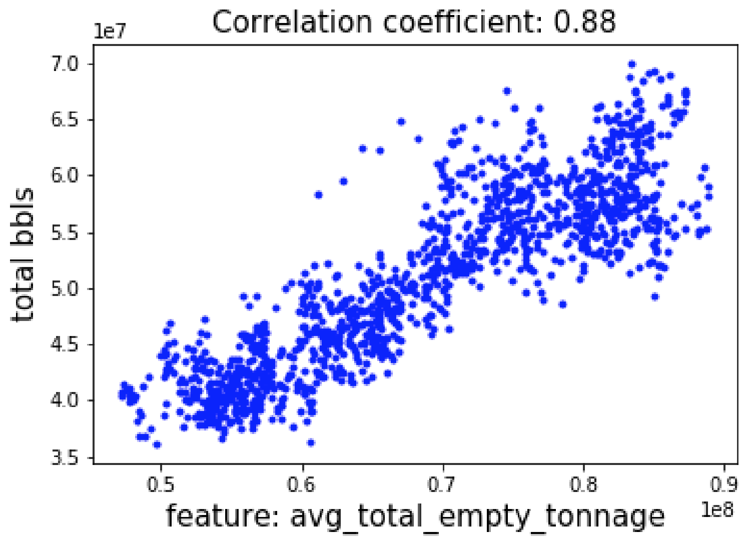

Figure 3. Scatter plot of the average total empty tonnage 21~28 days ago with the total tonnage shipped today. Pearson's correlation is 0.88 between the two set of variables, indicating a high correlation.

Our predictive modeling experiments followed a 5-fold cross-validation scheme, where the data is partitioned into 80% training and 20% test sets 5 times in total. In the end, the average performance over each split was obtained. The justification for this design is to obtain a more reliable estimate of error than using a validation set (or one holdout set) approach, as well as obtain a more robust model. R-squared and absolute error ratio were the statistical performance measures we used, which are common for regression-type problems. We discovered that there was a strong linear relationship between LPG demand and empty vessel capacity, so that we incorporated that as a predictor into our regression model. We were not sure exactly the reason that this relationship existed or if it was even causal. We expect that domain experts would have much more insight in this area that we do. Our goal was to just build a predict model as accurately as possible using the available information.

With the prediction model, we could check what it would bring to the company and choose the optimal model to do prediction. In detail, we will compare the total cost with prediction and the total cost without prediction. In this way, we can see if this model can benefit a trader.

Regression models

In this project, we selected and compared the performance of three regression models: linear regression model, ridge regression model, and support vector regression model.

The linear regression model was selected as a baseline for our study, since it is the simplest and most-widely used regression model. This model can be described with the following equation:

$$y_i = \beta_0 1 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \varepsilon_i$$

where x_i represents the input variables, y_i represents the target variable, β represents the coefficients for each input variable, and finally ϵ_i is the residual error. In this model, the training process consists of estimating the coefficient β given a set of input-output pairs (x_i, y_i) using ordinary least squares (OLS). There are no hyper-parameters to tune in the linear regression model. The advantage is that it is very simple to use and usually will not lead to overfitting (because it is too simple). The disadvantage is that it cannot model non-linear relationships between input and output and thus will suffer from higher bias.

The ridge regression model was selected since it introduces a regularization term to the linear regression model. One potential problem with linear regression is that it does not provide any mechanism to prevent overfitting issues. Ridge regression's regularization term penalizes overly complex models and helps obtain a better bias-variance tradeoff. With such capability, the model can prevent overfitting and can be better generalized to innovative/unseen data points. Additionally, we can apply the kernel trick to ridge regression so that non-linear relationships can be modelled. The tuned parameters in ridge regression include the regularization term, as well as the optimized kernel type (linear, poly, or Gaussian). The formula for ridge regression is shown below:

$$\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij}\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^2 = \mathrm{RSS} + \lambda\sum_{j=1}^{p}\beta_j^2,$$

λ: is the tuning parameter that determines the effect of the impact of the penalty term $\lambda\sum_{j=1}^{p}\beta_j^2$. As λ increases, the impact of the penalty term grows; when λ is 0, the ridge regression parameter coefficients are the same as those generated using ordinary least squares in the linear regression model.

The Support Vector Regression (SVR) is a more advanced regression model. It uses support vectors to keep a max distance between decision margin and the closest data points. Such mechanism would make the model more robust to noise, thus increasing its generalization capability. The advantage is that it can model complicated relationship in the data, while the disadvantage is that it is overly complicated, therefore the training is going to take longer, and it requires more data for the training to converge. The SVR model has been widely used in many domains, including supply chain prediction (Guanghui, Wang et al), flight control prediction (Shin, Jongho, et al), and tourism demand prediction (Chen, Kuan-Yu et al). The SVR model has been found to achieve the best performance in diverse domains, therefore, we consider it in our study. However, since our dataset is quite simple, we did not expect SVR to perform the best due to its complexities.

Grid-search for hyper-parameter tuning

Most of the regression models have certain hyper-parameters that we can tune to achieve the best performance. For example, the SVR model has several kernels that can be chosen from (i.e. linear, polynomial, Gaussian etc). The selection of the most appropriate hyper-parameter would cause the model to achieve the best performance. At the same time, a poor selection of hyper-parameter would lead to bad performance. In our case, we implemented a k-fold grid-search technique to find the best hyper-parameter for each model. The entire dataset was separated into 80% training data and 20% testing data. The 80% training data was further split into five folds (16% in each fold, denoted as a1~a5). The model was trained on a1~a4 and tested on a5, and this process was repeated 5 times among a1~a5. The hyper-parameters generating the best average performance on the held-out validation split was chosen in the end as the optimal ones. And the model is trained with the best found hyper-parameter. This trained model was then tested on the 20% testing data to achieve the final performance.

Optimization modeling-type project

With the prediction model for demand, we can know the future demand of LPG. Predicting the future demand, we can assist company to manage their tonnages of delivering LPG. Neither would they oversupply nor undersupply LPG to market. Therefore, they would able to avoid loss on supplying LPG.

RESULTS

To measure the performance of a regression model, we used the relative Mean Absolute Error (rMAE) on the testing data as the most important metric. The equation of rMAE is listed as follows where y is the actual value and pred_y is the predicted value:

$$rMAE = (|y\text{-}pred\_y|) / y$$

The smaller the rMAE, the better the regression model is with predicting future tonnages. We found that the ridge regression performs the best with an average rMAE value of 5.20%. The linear regression and SVR performed slightly worse with an average rMAE of 5.83%. Also, the ridge regression reached the highest R-squared value on the testing dataset. The p-value for the estimated parameter coefficient was also less than 0.001 indicating strong evidence that the feature had an association with the response. This indicates that the ridge regression can best describe the variations and trends of our dataset. Therefore, it should be used in the real system to provide LPG tonnage predictions for decision making guidance. Table 3 displays the performance of each model on the train and test dataset.

Model	R-squared (train)	Mean of R-squared	R-squared (test)	Mean of R-squared	Error ratio (test)	Mean of Error ratio
Linear	0.778	0.7742	0.751	0.7652	5.93%	5.83%
	0.773		0.771		5.85%	
	0.774		0.764		5.83%	
	0.773		0.769		5.76%	
	0.773		0.771		5.76%	
ridge	0.824	0.8222	0.798	0.8064	5.16%	5.20%
	0.819		0.818		5.13%	
	0.824		0.8		5.29%	
	0.822		0.807		5.20%	
	0.822		0.809		5.20%	

SVR	0.777	0.7726	0.748	0.7636	5.91%	5.79%
	0.771		0.77		5.80%	
	0.773		0.763		5.78%	
	0.771		0.768		5.71%	
	0.771		0.769		5.73%	

Table 3: Model selection summary

In Figure 4, the actual total tonnage value and the predicted tonnage value is shown. As we can see, there is an average training R-squared (R2) score of 0.822, and an average rMAE of 5.20% on the test split. The regression model can accurately predict the total tonnage using the methods proposed.

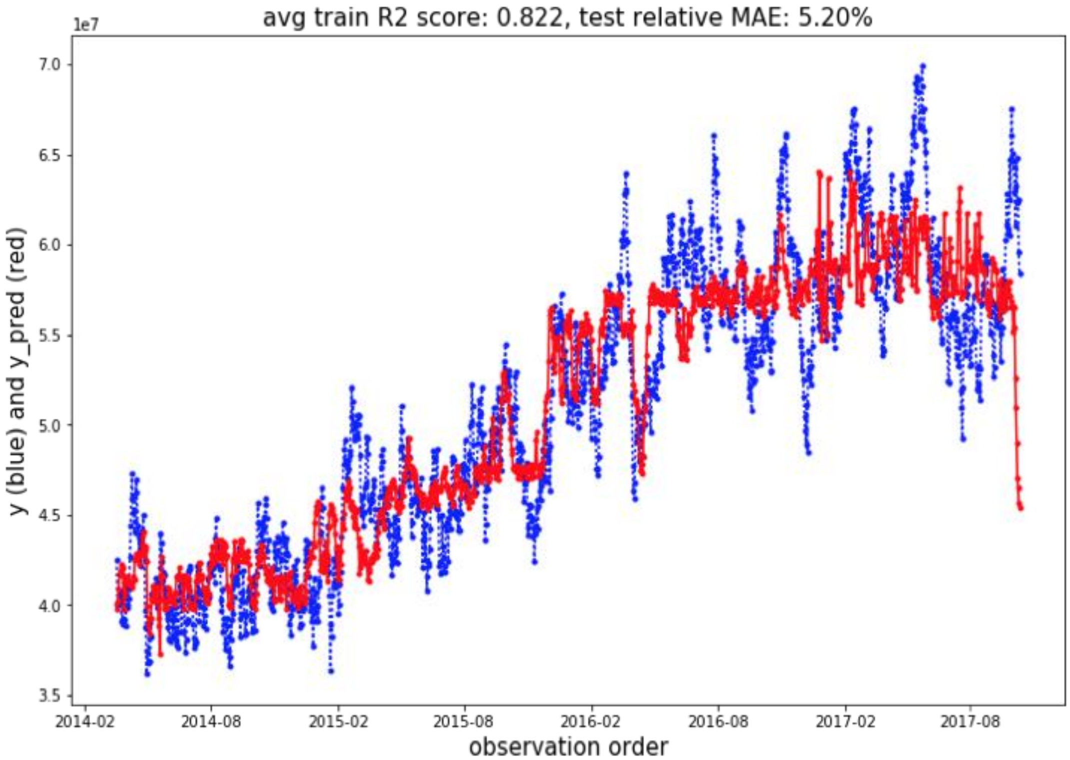

Figure 4. Plot of the actual total tonnage shipped (blue lines) and the predicted tonnage shipped using number of tonnages 3 weeks ago (red lines). As shown, the regression model can accurately model the variances in the tonnage changes.

The four-in-one plot for the regression is displayed in Figure 5 for regression diagnosis purposes. As shown, all the assumptions of a linear model have been satisfied, therefore justifying the correct usage of the regression model.

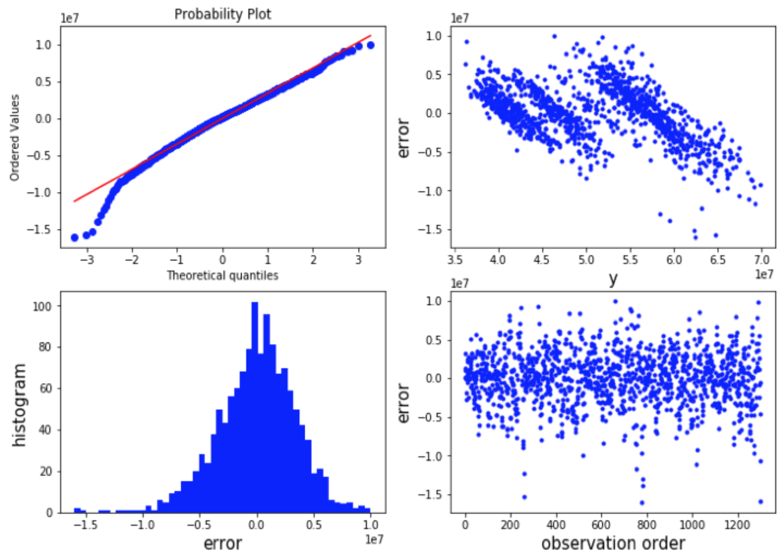

Figure 5. Four-in-one plot for the ridge regression model. The regression assumptions have been met.

Specifically, the errors are normally distributed as shown in the histogram of residuals and probability plots. There does not appear to be a major issue of heteroskedasticity either, thus the constant variance assumption is not violated.

FUTURE WORK

At current stage, we only use the past three weeks' average empty vessels' capacity as the independent variable to build our model. After further exploring data relationships and patterns, we may consider including extra variables in the model in the future. For example, segregating vessels based on different features (e.g. annual total tonnage of LPG, route, vessel size), might allow for improved predicted performance of future LPG amounts.

Route and Schedule Per Vessel

By identifying time sequence of each vessel (by merging on the date_load from empty vessel and the date_depart from full vessel), we are able to produce a full operating schedule of each vessel. For example, the schedule of no.50 in year 2014 is as follow:

	empty_date_depart	empty_to_full_date_arrive	date_load_depart	full_date_arrive
0	NaT	NaT	2014-01-23 21:24:07.490	2014-02-10 00:00:27.243
1	2014-02-11 04:01:40.503	2014-03-09 20:00:48.760	2014-03-11 06:02:22.860	2014-04-01 09:26:38.047
2	2014-02-11 04:01:40.503	2014-03-09 20:00:48.760	2014-03-11 06:02:22.860	2014-04-07 09:00:22.000
3	2014-04-09 04:01:13.483	2014-04-24 08:00:48.927	2014-04-27 09:00:49.137	2014-05-12 04:03:38.130
4	2014-05-12 22:06:42.000	2014-05-14 01:08:40.000	2014-05-15 02:07:49.000	2014-05-16 07:04:34.923
5	2014-05-19 00:23:34.403	2014-06-08 21:28:34.497	2014-06-09 00:00:11.247	2014-06-27 22:12:03.000
6	2014-06-29 04:07:22.637	2014-07-17 09:04:10.567	2014-07-18 00:22:11.963	2014-08-05 21:10:41.000
7	2014-06-29 04:07:22.637	2014-07-17 09:04:10.567	2014-07-18 00:22:11.963	2014-08-09 04:24:24.000
8	2014-08-10 23:03:21.483	2014-09-02 10:22:14.413	2014-09-03 09:01:11.880	2014-09-21 00:00:20.000
9	2014-08-10 23:03:21.483	2014-09-02 10:22:14.413	2014-09-03 09:01:11.880	2014-09-22 01:30:14.000
10	2014-09-22 23:17:36.000	2014-10-14 07:00:13.440	2014-10-15 19:00:06.567	2014-11-02 06:26:13.000
11	2014-09-22 23:17:36.000	2014-10-14 07:00:13.440	2014-10-15 19:00:06.567	2014-11-04 00:04:29.000
12	2014-11-05 05:11:22.000	2014-11-24 23:00:50.757	2014-11-26 05:03:06.043	2014-12-08 15:37:47.000

Table 4: Shipment schedule of vessel no.50 in year 2014

Given the above schedule and related location code, we can develop strategies to optimize operations based on business need.

CONCLUSIONS

The ability to predict future commodity flows of LPG is very beneficial to companies at both the supply side and the demand side. With better predictions, companies can optimize their operations to either save costs or earn more profit. The model we developed can help predict the future amount of LPG being shipped three weeks after the detection of empty moving vessels. This has potential to be useful for a commodities trader whom seek to glean as much information on supply and demand as they can to improve their investment strategies. In the future we can incorporate other features such as seasonality, region, and consumer behavior to help make this model more robust against errors. To sum up, this exploratory step we took in utilizing shipping information has returned valid results; in the future, better refinements will continue to make the predictive model stronger and flexible to industry requirements.

REFERENCES

IHS Markit (2016). “Maritime & Trade Fleet Capacity Forecast.” https://www.ihs.com/products/fleet-capacity-forecast.html

Paul Page (2017). “Today’s Top Supply Chain and Logistics News From WSJ.” The Wall Street Journal https://www.wsj.com/articles/todays-top-supply-chain-and-logistics-news-from-wsj-1512128205

Joeri van der Sman (2017). “Major Turnaround In LPG Shipping By 2018?” Seeking Alpha https://seekingalpha.com/article/4079722-major-turnaround-lpg-shipping-2018

Stephanie Wong (2017). “Christmas Becomes Inventory War Game for US Fashion Brands.” Bloomberg https://www.businessoffashion.com/articles/news-analysis/christmas-becomes-inventory-war-game-for-us-fashion-brands

Akshay, M. et al. (2004) STATE-OF-THE-PRACTICE IN FREIGHT DATA: A REVIEW OF AVAILABLE FREIGHT DATA IN THE U.S. Center for Transportation Research The University of Texas at Austin

Yan T.,(2002)Business cycles, commodity prices and shipping freight rates - some evidence from the pre-WWI period. SNF project no1312,gloablizatoin, economic growth and the new economy.

Chen, T., et al. (2008). Ship-flow analysis and counting system based on image processing. Proceedings of the Fourth International Conference on Intelligent Information Hiding and Multimedia Signal Processing.

Fei, T., et al. (2014). Robust Inland Waterway Ship Tracking via Hybrid TLD and Kalman Filter. Advanced Materials Research, Trans Tech Publ.

Hammond, T. (2014). Applications of Probabilistic Interpolation to Ship Tracking. Proceedings of 2014 Joint Statistical Meetings-American Statistical Association.

Hu, W.-C., et al. (2011). "Robust real-time ship detection and tracking for visual surveillance of cage aquaculture." Journal of Visual Communication and Image Representation 22(6): 543-556.

Pallotta, G., et al. (2014). Context-enhanced vessel prediction based on Ornstein-Uhlenbeck processes using historical AIS traffic patterns: Real-world experimental results. Information Fusion (FUSION), 2014 17th International Conference on, IEEE.

Pallotta, G., et al. (2013). Traffic knowledge discovery from ais data. Information Fusion (FUSION), 2013 16th International Conference on, IEEE.

Pallotta, G., et al. (2013). "Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction." Entropy 15(6): 2218-2245.

Riveiro, M. J. (2011). Visual analytics for maritime anomaly detection, Örebro universitet.

Robert-Inacio, F., et al. (2007). "Multispectral target detection and tracking for seaport video surveillance." Proceedings of the IVS Image and Vision Computing New Zealand: 169-174.

Vespe, M., et al. (2008). Maritime multi-sensor data association based on geographic and navigational knowledge. Radar Conference, 2008. RADAR'08. IEEE, IEEE.

Guanghui, W. A. N. G. (2012). Demand forecasting of supply chain based on support vector regression method. Procedia Engineering, 29, 280-284.

Shin, J., Kim, H. J., Park, S., & Kim, Y. (2010). Model predictive flight control using adaptive support vector regression. Neurocomputing, 73(4), 1031-1037.

Chen, K. Y., & Wang, C. H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. Tourism Management, 28(1), 215-226.