Minimizing Risk in Ocean Shipping Contracts

PURDUE UNIVERSITY | Krannert School of Management

Jackson Bronkema, Diego Victor Carlos Chavez, Asad Husain, Yuxuan Li, Kumar Rahul, Kshitij Virdi, Yang Wang

Purdue University, Krannert School of Management

jbronke@purdue.edu; carlos2@purdue.edu; husain10@purdue.edu; li4196@purdue.edu; krahul@purdue.edu; kvirdi@purdue.edu yangwang@purdue.edu

ABSTRACT

Working with our client, an industry leader in the high-end furniture business, we forecast container volume by shipping lane. Before this project, our client did not have a robust reference point for their expected lane-wise shipping volume. We applied several time series models for each lane. We effectively modeled 65% of the total shipping volume our client has imported in the past eight fiscal years.

Our suggestions will help them build long-term contracts with the carriers. This will help mitigate our client's financial risk when negotiating contracts with ocean freight carriers.

INTRODUCTION

The ocean-freight industry operates on fixed lane contracts. When demand overwhelms the supply, carriers reject the shipper's load, forcing them to go to the spot market. The spot market refers to the auction mechanism where shippers issue loads and carriers offer bids to fulfill the loads on a real-time basis (Sinha, Thykandi 2019). The spot market is volatile to price fluctuations, leading shippers to pay up to three times more than contract prices. It is possible to experience a high number of load rejections from carriers as a shipper, resulting in higher costs. The company's goal is to reduce the total logistics cost by avoiding the spot market.

A potential approach to this problem would be to forecast the volume shipped in a given period using a time-series model. An insight into the expected number of containers through shipping volume would help shippers commit to the carriers in a more accurate binding contract that would benefit both parties in the long run. This would also strengthen the shipper-carrier relationship.

Our methodology is to use recent ocean shipping data and leverage it in creating many candidate models for each lane using R software. Then, we choose the most accurate model based on its MAPE score and use it to provide our client with a 12-month lane-wise volume aggregate.

Fig 1: Typical value chain for the shipper community

OBJECTIVES

- Is there any trend or seasonality in the historical data that can generate business insights?
- How has COVID-19 changed how shipping volume is modeled and forecasted?
- What is the best way to reduce the logistic costs for the client by leveraging the historical data?

LITERATURE REVIEW

Our methodology focuses on the ARIMA and ETS models. The ETS model is most effective when the data behavior has a strong trend and seasonality. In our study, it is particularly suitable for producing short-term forecasts for time-series data that represents shipping demand. The ARIMA model requires the past data of a time series to generalize the estimates. ARIMA models also perform well on short-term forecasts.

METHODOLOGY

We created functions to repeatedly deal with missing values in the individual lanes and then choose the best model to forecast it. The shipping data is sourced as a 6-month dump from their database. By merging all those datasets together, we get the combined shipping history, which we then use to define lane-wise time series. Our second function then creates another time series which is mean imputed, if needed. After modeling both these time series, we select the best ARIMA or ETS model which gets us the least MAPE score. The end deliverable of our code is the next years' aggregated volumes, that the client can use to form the contract with the carrier.

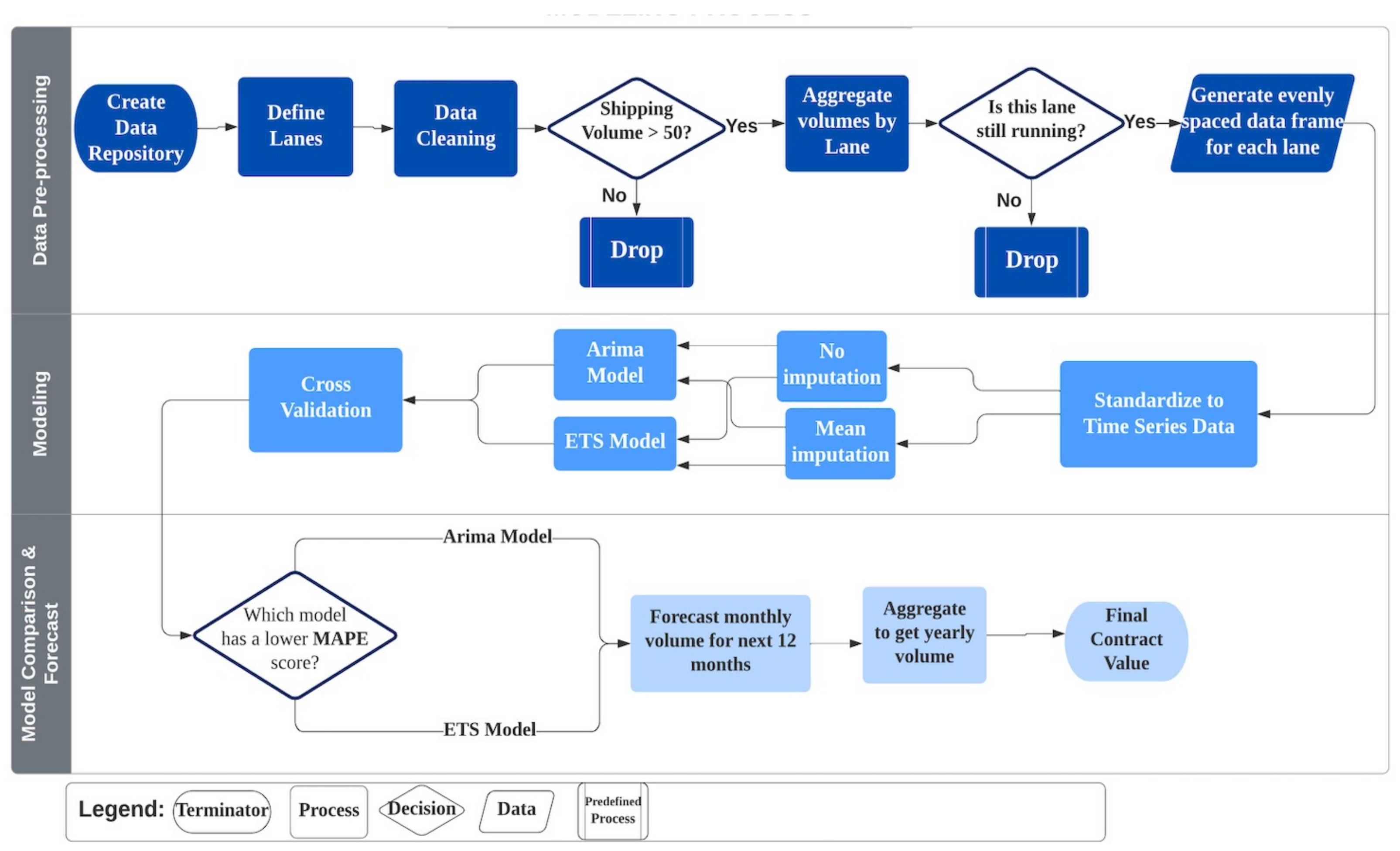

Fig 2: Modeling Process

STATISTICAL RESULTS

Below are the forecasting results of various shipping lanes.
The title represents the shipping lane starting from origin to the destination city. The actual values are black, and our model's fitted values are green. The blue line represents a twelve-month forecast of the shipping volume.
We used cross-validation instead of the train-test type of evaluation to make more efficient use of the data present.
We can see below the top 6 lanes where we could most accurately forecast the volumes.

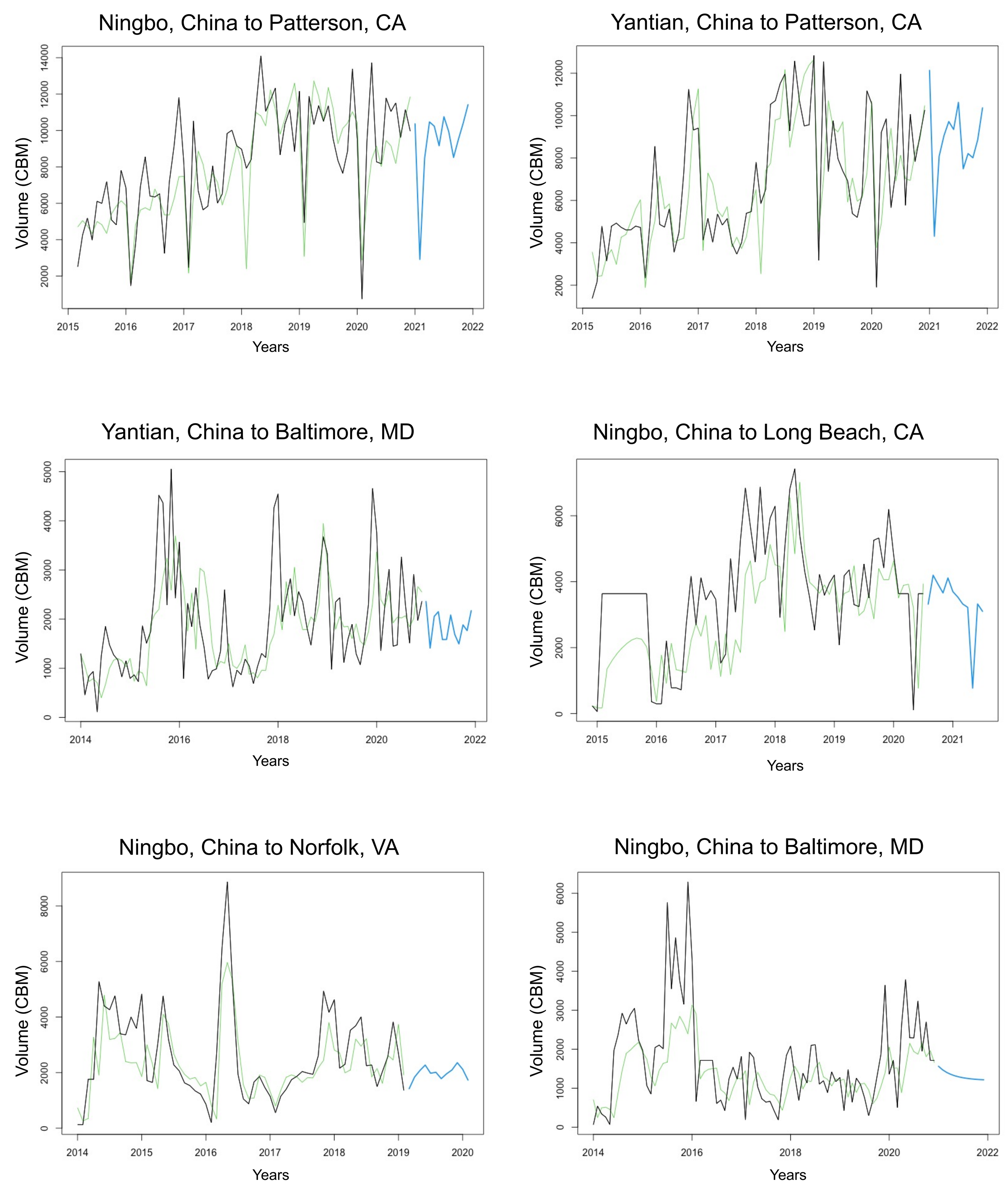

Fig 3: Modelled lanes

EXPECTED IMPACT

The client requirement included delivering a relative evaluation metric enabling them to compare the accuracy of various shipping lane forecasts. To do this, we used MAPE, or mean absolute percentage error. MAPE is an accuracy measure based on the relative errors of each prediction in our projections. We chose the champion model for each lane by choosing the candidate model with the lowest overall MAPE post cross-validation. Using this statistic, we can predict up to a certain percentage of how accurate the forecast for the monthly shipping volume will be.
These predictions will help our client save money on shipping costs. Over twelve months, our forecasts are 32% more accurate on average. Assuming that our sponsor is forced to go to the spot market for all the volume they inaccurately estimate, we can say that the sponsor will save approximately $2.5 million using our time series forecasts over the past twelve months.
Another benefit of more accurate lane-wise forecasting is that it enables our client to build trust with its partner carriers. Building this trust increases the probability that the carriers will fulfill the sponsors' shipping demands, thus fewer contract defaults for our client and better shipping rates.

CONCLUSIONS

We attempted to train over ten candidate models on each of the largest shipping lanes for our client through the modeling process. Each lane differed in volume and the number of missing values. This was a problem that we needed to combat. One of the main drivers of our success in this project was that we tried a wide variety of imputation methods on missing data for each of the lanes. We then trained each of these candidate models on the differing data. We then selected the candidate model with the lowest MAPE. Please refer to the table below for the results.

Lane	Cross-validated MAPE	Champion Model
Ningbo, China to Patterson, CA	37.63	ETS (A, N, A)
Yantian, China to Patterson, CA	34.61	ETS (A, Ad, N)
Yantian, China to Baltimore, MD	27.17	ARIMA (0,1,1) (0,0,2)
Ningbo, China to Long Beach, CA	44.72	ARIMA (1,1,1) (0,0,1)
Ningbo, China to Norfolk, VA	45.77	ARIMA (2,0,0) (1,0,0)
Ningbo, China to Baltimore, MD	27.56	ETS (A, Ad, N)

Table 1: Modeling Results

We come away from this project feeling confident that we could provide our client with powerful time series forecasting models that outperform the previous system that they were using to stipulate lane-wise volume to the carriers they work with. We project that our client will save millions of dollars in the next contract season by implementing our models.

ACKNOWLEDGEMENTS

We want to thank our industry partner for all the guidance that they provided to us throughout the project. They provided us with many helpful insights enabling us to produce accurate forecasts. We would also like to thank Professor Dr. Yang Wang, who oversaw our project and helped us overcome many obstacles throughout the project.