An Innovative Prototype to Create Real-Time Commodity Product Market Intelligence and Move the Market

Abhijit Kumar, Fan Lu, Shengye Guo, Charul Bagla, Matthew A. Lanham

Purdue University Krannert School of Management

kumar354@purdue.edu; lu714@purdue.edu; guo506@purdue.edu; cbagla@purdue.edu; lanhamm@purdue.edu

Abstract

This project aims to build and evaluate a **predictive model** to estimate the **aggregate weekly export of corn, soy beans, and wheat** from major export regions in the US that are listed on the weekly USDA grain inspection report. For our business partner, having earlier market transparency of weekly numbers would provide traders and marketers in the grain industry vital insights into constantly evolving patterns and export flows. We developed an **efficient crawler** to obtain relevant data from USDA and built **time series models** to predict the agricultural commodities inspection export numbers.

Introduction

Agricultural exports are one the top 5 exports from the US. Corn, soybeans and wheat account for **25% of the total agricultural revenue**. Having a predictive tool that could precisely forecast the export quantity information for the given commodities would be advantageous, especially for traders whom require timely market decision-support. This information would also provide market transparency in a more efficient way.

This project first explores the trend and seasonality of agricultural commodities (corns, soybeans, wheats) to generate export quantity predictions for the upcoming week.

We also aim to assess if incorporating information about vessel movements from major ports will help improve the quality of predictions by providing an upper limit on the shipping capacity of each export port.

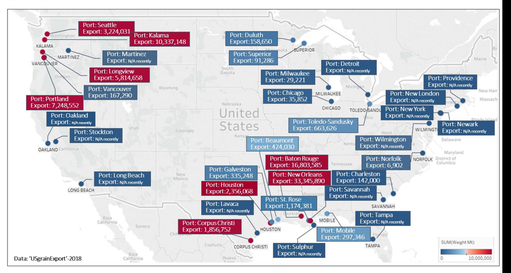


Figure 1. Major Ports Distribution in US

Research Goals:

- Predict future grain exports from major US ports through time-series modeling (ARIMA, decomposition methods, and LSTM neural networks).
- Detect seasonality and trend to accommodate for the grain exports.
- Come up with a rolling-based automatic tool that will accurately predict for the future week that our partner and sell to try and move the market.

Literature Review

The existing literature in this domain covers the importance of using the ARIMA models for forecasting aggregate export quantities for highly seasonal commodities and methods for stabilizing a non-stationary time series to generate accurate predictions. Not a lot of significant research has been done on the latter half of our project covering the effects of incorporating the vessel movement information on improving prediction accuracy for exports. However, research has been conducted on the effect of exchange rates on the export trends for commodities which could be a possible addition to our model.

Methodology

Process Flow

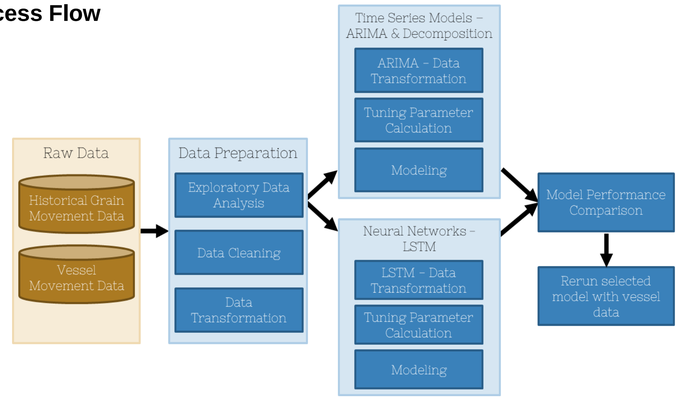


Figure 2. Study Design

Data: EDA, Data Cleaning & Pre-Processing

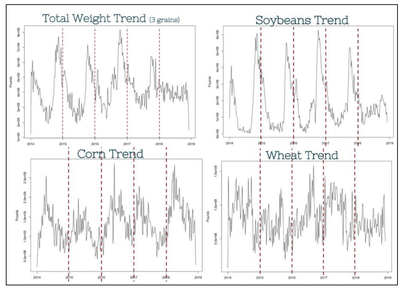


Figure 3. Seasonality in Corn, Wheat and Soybean Exports

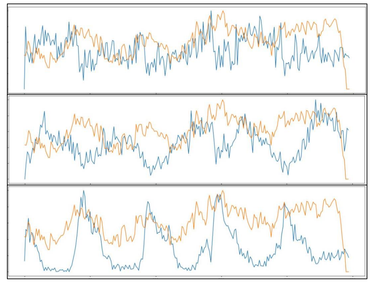


Figure 4. Vessel movement trends with wheat, corn and soybean export trends

- Open source datasets containing yearly export total for each interested commodity from USDA Federal Grain Inspection Services
- Empty vessel information from business partner

Model Design:

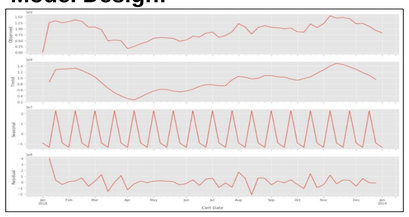


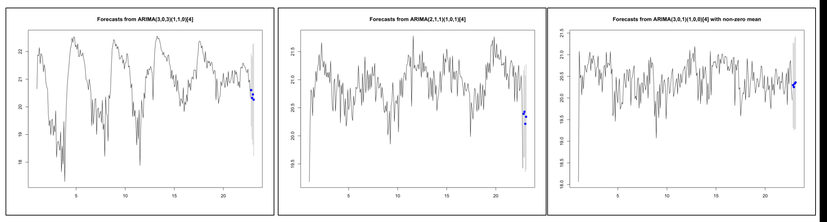
Figure 5. Series Differencing

The model setup for decomposition & ARIMA model requires testing the series for stationarity. ACF and PACF plots are used to perform this analysis.

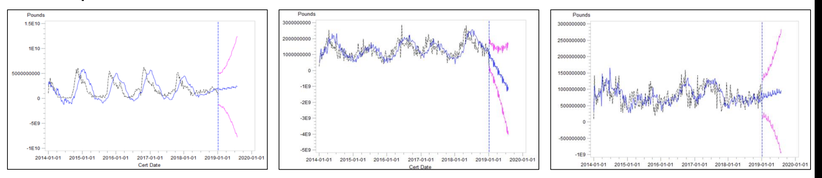
For LSTM neural networks, a lookback and cross validation was setup. K-week lookback uses last K weeks history to forecast the exports in the upcoming week. A lookback based forecasting models require a recursive prediction model where predictions for only 1 week are made at a time and the same is fed into the lookback matrix to get the forecast for consecutive weeks.

Results

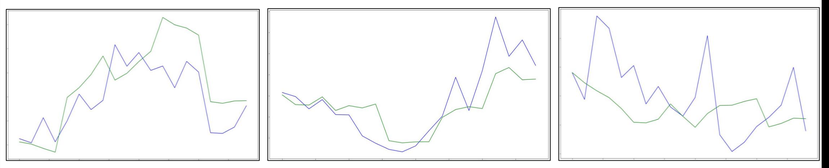
ARIMA Model:



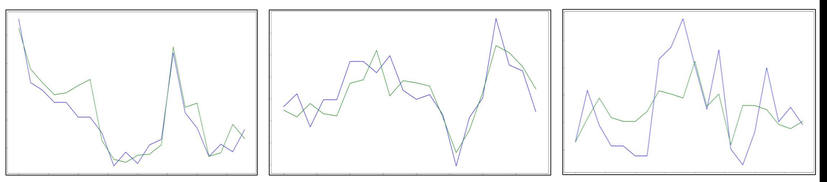
Decomposition Model:



LSTM Neural Network:



LSTM Neural Network with Vessel Data:



Conclusions

The predictive models to estimate the aggregate weekly corn, wheat and soybean exports were compared on the MAPE. The LSTM Neural Network was the best model for Corn, Wheat and Soybeans providing 83%, 78% and 60% accuracy respectively. Incorporation of vessel information helped improve accuracy of soybean prediction by 8% with no major impact on wheat and corn.

Acknowledgements

We thank Professor Matthew Lanham for constant guidance on this project.