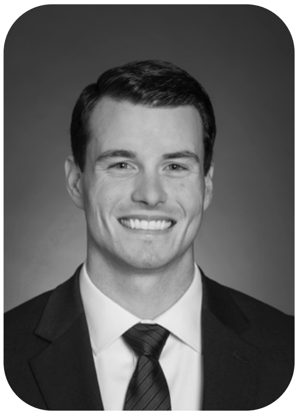
Forecasting International Movements of Returnable Transport Items

By: Patrick Jacobs and Rajdeep Walia
Thesis Advisor: Dr. Eva Ponce-Cueto

Topic Areas: Forecasting, Supply Chain Planning

Summary: This research was focused on developing a one-month-ahead forecasting model to predict international movements of Returnable Transport Items between the United States of America and Canada. The project's original premise was to utilize two countries' macro-economic variables, such as the foreign exchange rates and GDP figures, to predict the number of international returnable transport item movements. While this method was evaluated, more traditional and validated forecasting methods were also utilized. Ultimately, 36 unique forecasting models were developed and compared using various metrics.



Before coming to MIT, Patrick Jacobs graduated with his Bachelor's in Operations Management from Georgia Southern University. Following his undergraduate education, he worked in manufacturing at Rolls-Royce. Upon graduation, Patrick will join Apple as an Operations Program Manager.

Before coming to MIT, Rajdeep Singh Walia graduated with an MBA from Indian Institute of Foreign Trade, and then worked for Royal Dutch Shell for five years. Prior to that, he graduated with a Bachelor's degree in Computer Science Engineering from Indraprastha University. Upon graduation from MIT, Rajdeep will join Amazon as a Finance Business Partner for their Retail Operations function.

Key Insights

1. Utilizing macro-economic variables to forecast business supply conditions does not outperform standard industry forecasting practices.

2. The time-series decomposition framework provides the most methodical approach to identifying appropriate forecasting models.

3. Changing the aggregation level of the forecast greatly impacts model performance and selection.

Introduction

Modern logistic systems rely on Returnable Transport Items to effectively move materials both locally and globally. Returnable Transport Items (RTI) are defined by ISO as objects used for the purpose of "transportation, storage, handling, and product protection in the supply chain, which are returned for further usage". These items often include shipping platforms such as TEUs, crates, and pallets. While some companies choose to manage these components internally, many utilize a logistics services provider to manage their RTI supply chain operations.

Our partner firm, hereafter referred to as Company X, is one of the largest players in the logistics solutions industry, with RTI leasing management generating a significant portion of the company's revenues. Since 2014, Company X has experienced unprecedented levels of international shipments, specifically between the United States and Canada. The sharp increase in international movements has amplified the volatility of forecasting the net international surplus of RTI in Canada. Net international surplus is calculated by subtracting the Canada to USA flow of RTI from the corresponding USA to Canada movements.

Concurrent to the increasing movements between Canada and the USA, the value of the United States Dollar appreciated significantly to that of the Canadian dollar. This seemingly simultaneous shift led Company X to hypothesize that foreign exchange rates were a potential driver of the increased movements from Canada to the USA. The RTI customers were believed to be increasing their imports from Canada in an effort to take advantage of the relatively cheap Canadian dollar.

In order to more quickly respond to the changing market conditions, Company X sought to improve its current abilities to forecast international movements by incorporating macro-economic variables. In addition to improving forecasting accuracy, Company X imposed two additional constraints. The first required the forecasting model to take minimal time to update. The final constraint specified that the model would need to operate on a relatively common software platform, such as Microsoft Excel.

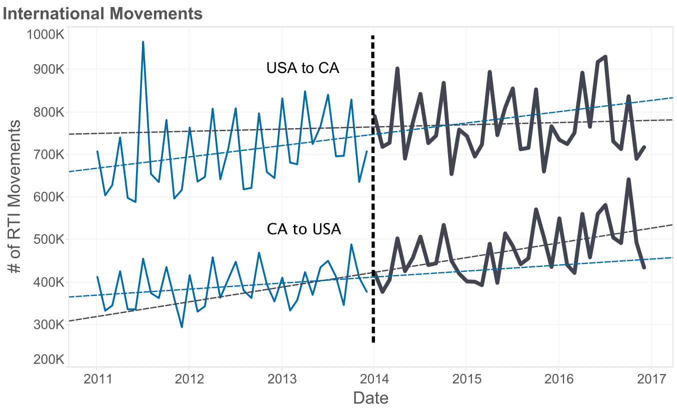


Chart 1: USA to CA and CA to USA flows from January 2011 to December 2016. The hashed line marks January 2014 when the exchange rate the United States Dolllar began to appreciate significantly in comparison to the Canadian Dollar.

Utilizing Macro Economic Variables

In addition to the foreign exchange rate of CAD to USD, the influence of other macro-economic variables were also evaluated. These variables included USA GDP, USA diesel #2 prices, gold prices, number of exports from Canada, and number of Canadian imports. Furthermore, since incorporating macro-variables into forecasting models requires the use of historical values, the variables were lagged 1 to 12 months prior to conducting a correlation analysis. Six years of Company X's historical time series data was utilized in conjunction with the exogenous data. Prior to performing the correlation analysis, we used scatterplots to visually determine if any non-linear relationships existed. No non-linear relationships were identified. Correlation values were only deemed to be significant if the values were less than -.6 or greater than .6.

Movement	Variable	Lag (months)	Correlation
CA to USA	US Quarterly GDP	12	0.7197
	Canada Quarterly GDP	12	0.6868
	Average CA to USD FEx	5	-0.6995
USA to CA	--		
Net International	--		

Table 1: Macro-economic variable correlation analysis. Only values less than -.6 or greater than .6 are shown.

As can be seen in Table 1, only the CA to USA flow had high correlation values with the lagged macro-economic variables. The lack of significant variables for the USA to CA flow required other methods be used rather than purely macro-economic based forecasting models.

The Method

In order to determine the optimal forecasting methods that should be considered, we utilized the highly-researched decomposition method, which reduces time series data sets into level, trend and seasonality components. After performing this analysis, both the USA to CA and CA to USA time series demonstrated strong seasonality components, level values, and slight linear trends. Further research suggested the best models given our decomposition findings would be the exponential smoothing state space methods: seasonal exponential and Holt-Winter. Additional methods included least squares regression and Seasonal Autoregressive Integrated Moving Average (SARIMA).

Prior to model development, we split the six years of historical time series data provided by Company X into a training set (4.5 years) and validation set (1.5 years). Model parameters were then developed and optimized to the training sets.

Models were developed using two primary methods. The first used the same forecasting methodology for the both the USA to Canada and the Canada to USA models. The second approach identified the top 5 best performing USA to Canada models and paired them with the top 5 performing Canada to USA models. A hybrid forecasting model, which utilized two separate methods, was then created in attempt to utilize the optimal models for the two separate international movements.

Furthermore, the Naïve method was used as a benchmark forecast in order to provide a comparison for the developed forecasts. In order to evaluate the models' performances three different metrics were used: Mean Absolute Percent Error (MAPE), Mean Absolute Scaled Error (MASE), and Mean Absolute Deviation (MAD). These three metrics were chosen as they complimented each individual's deficiency. For example, MAPE allowed for an easy to comprehend metric, but failed to provide a direct comparison to the bench mark, naïve method. This led to the inclusion of MASE, which directly compares the forecasting model's error term to the naïve method's error term, thus providing an easy to compare indices.

Model Selection

Ultimately, 36 models (not including the naïve benchmark) were developed using the aforementioned methods. The models were then ranked according to their relative performances on the validation sets in regards to the MAPE, MASE, and MAD metrics. The value of 1 was assigned to the best performing model in each respective metric, while the worst performing model received a value of 36. A composite score was then created using both mean and multiplicative formulations. The mean method simply took the average rank across the 3 metrics, while the multiplicative multiplied the rank values. The average composite score favored models that performed consistently across all metrics, while multiplicative favored models that performed well on at least 2 out of 3 metrics. While using two composite scores did create some redundancy, the separate measures were used to ensure a model was not disregarded simply because of penalizing arithmetic.

After evaluating the composite scores, only 7 out of the original 36 models outperformed the naïve benchmark in both composite scores. These models are shown in Table 2 below. However, the two different composite scores yielded different model selections. Seasonal Exponential | Seasonal Exponential was the best performing model when the mean composite score was utilized, while the multiplicative score suggested SARIMA | SARIMA was the better performing method. The Seasonal Exponential model performed 5% better on the MASE metric, but 8% worse across the other two metrics when compared to the SARIMA model. To decide between the two models, a qualitative comparison was done to evaluate 1) the ease of updating the models and 2) the ability to operate in commonly held software programs. Both models required only minimal time to update. However, the ability to operate in commonly used software programs served as the differentiating factor. Seasonal Exponential is relatively simple to create in commonly used software programs such as Microsoft Excel. Conversely, implementing SARIMA requires more advanced software programs. Ultimately, Company X decided that the 8% improvement to MAPE and MAD that could be achieved using SARIMA would not offset the required cost of the licensure. Thus, Seasonal Exponential was selected not only for its quantitative performance, but also its performance across qualitative metrics.

Model	MAPE Rank	MASE Rank	MAD Rank	Mean Score	Mult. Score
Seasonal Exponential | Seasonal Exponential	7	10	5	7.33	350
Holt-Winter | Simple Regression	6	21	2	9.67	252
SARIMA | SARIMA	1	28	1	10.00	28
Seasonal Exponential | Simple Regression	2	26	4	10.67	208
Simple Regression | Simple Regression	3	23	6	10.67	414
Seasonal Exponential | Endogenous Regression	4	33	3	13.33	396
Seasonal Exponential | Simple Regression M2Y	5	29	7	13.67	1015

Table 2: Models that outperformed the naïve benchmark's multiplicative and mean composite scores. The model listed to the left of the "|" was used for the USA to CA movement and the model on the right was used to forecast the CA to USA flow.

Additional Findings

Additional analysis was done to compare the robustness of the recommendation to utilize the seasonal exponential approach rather than SARIMA when forecasting for extended time horizons. A forecasting time horizon of one-year was used to compare the two models on a different aggregate level. The parameters were then re-optimized. Given the one year forecasting time horizon, the SARIMA model outperformed Seasonal Exponential by 40% across MAPE, MASE and MAD. The significant difference found by comparing monthly forecasting performance to yearly, reiterated the criticality of creating a new forecasting selection study rather than extrapolating model selection across multiple time horizons.