Improving Supply Chain Planning with Advanced Analytics:

Analyzing Lead Time as a Case Study

By: Darryl Yau
Thesis Advisors: Dr. Christopher Caplice
Topic Areas: Supply Chain Planning, Advanced Analytics, Forecasting

Summary: While considerable research has been done in formulating accurate and robust demand forecasts, there remains many areas for improvement on the supply planning side. In particular, many planning parameters (e.g., lead time, waste, yield, run rate, capacity, etc.) are often not given the consideration they deserve as part of the planning process. Oftentimes, the values of these parameters are not scientifically derived in the first place, or their actual values may have changed since its original inception and now differ significantly from its planned value. This research showed there is room for improvement in how lead time is managed and considered within the current planning process. Using predictive analytics to predict lead time (predictive lead time) can reduce the deviation between the planned and actual values in the lead time parameter. Predictive lead time can reduce the safety stock cost, the manual labor required in exception management (re-planning), and the manual labor in purchase order management (since the suggested lead time will be more accurate than the current process).

Before coming to MIT, Darryl Yau worked as an Operational Effectiveness Consultant for PwC, and an S&OP and APS Solutions Consultant for flexis AG. Darryl holds a Bachelor of Applied Science degree in Engineering Science, majoring in Manufacturing Systems Engineering from the University of Toronto, Canada, and a Bachelor of Laws from the University of London, UK. Upon graduating from MIT, Darryl will join Apple's Supply Demand Management Team in Cupertino, CA.

KEY INSIGHTS

1. Major disconnect among the lead time variables used along the planning to execution process.
2. Forecast error of lead time can be notably reduced by using historical data to support the planning process, which can lead to financial benefits
3. Lanes (vendor to plant combinations) appear to have different demand patterns that can dictate the use of different forecasting techniques

Introduction

Over the years, supply chain management has continued to change and evolve to become a major component in competitive strategy to enhance organizational productivity and profitability. There is a growing idea that it is supply chains that compete and not and that the success or failure of supply chains is ultimately determined in the marketplace by the end consumer. Matching the right product, at the right time and place to the customer is seen as not only essential to competitive success, but also the key to survival.

While considerable research has been done in creating accurate and robust demand forecasts, there remains many areas for improvement on the supply planning side. In particular, many planning parameters (e.g., lead time, waste, yield, run rate, capacity, etc.) are often not given the consideration they deserve as part of the planning process. Oftentimes, the values of these parameters may not have been objectively and scientifically derived in the first place, or their actual values may have changed since its original inception and now differ significantly from its planned value.

The concept of self-healing supply chain (or self-learning supply chain) improves the planning process by enabling planners to create more accurate and actionable plans. In a self-healing supply chain, a process or system is in place that focuses on surfacing the differences between as-designed (what was planned) and as-demonstrated (what actually happened) planning parameters. The differences are then continuously minimized through the use of predictive analytics.

This thesis focuses on examining lead time as a case study for how management of planning parameters can be improved. In particular, this thesis seeks to answer the following questions:

1) To what extent are the lead time variables found in the Enterprise Resource Planning (ERP) system used in predicting lead time and how accurate is the prediction?
2) Can predictive analytics on historical lead time data be used to improve the forecast accuracy and what are the benefits in doing so?

Data and Methodology

The data used for this research was provided by a global supply chain planning software company (the Consulting Firm) and contains operational data from one of their clients (the Company) in the industrial product industry. The data received describes the incoming orders for a manufacturing plant/warehouse from a vendor. In total, the flat file, for the data received, contained over 4 million entries, with records ranging from 2004 to 2017. Each entry in the record represents a line item of a purchase order (the file contained 500,000+ purchase orders for 80,000+ different SKUs), with corresponding attributes related to that SKU and order.

Lead time is defined as the time between an order being placed to the time the inbound shipment was received. In most ERP systems, including SAP, standard lead time variables were maintained for planning purposes. In the data received for this study, two standard lead time variables were present: 1) Lead time based on SKU and Vendor (LTv) and 2) Lead time based on SKU and Plant (LTp). These two variables appeared to be static, meaning they did not appear to be updated on a regular basis and are collectively referred to as “static lead time variables” in this paper. These two variables were the only lead time planning parameters found in the system and served as reference points in the formulation of the planned lead time found in the purchase orders (Planned Lead Time). Figure 1 illustrates the different lead time variables involved in the planning to execution process.

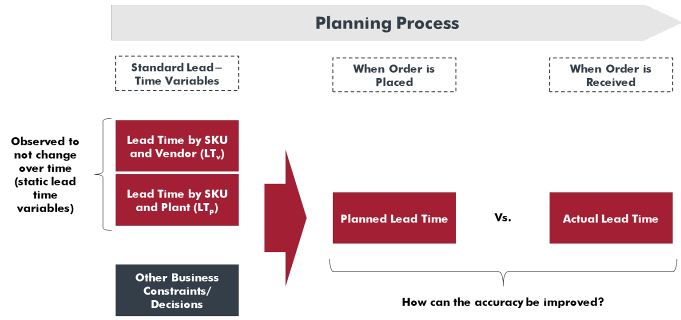

Figure 1. Overview of Various Lead Time Variables

This research first baselined the effectiveness of the current practice of using static lead time variables through:

1) Descriptive statistics of the current data
2) Regression Analysis
3) Hypothesis Test: t-test analysis of the lead time variables

This research then examined the viability of alternative solutions using time series analysis. The analysis was split into two phases:

1) Detailed analyses of three SKU & Lane combinations
2) High-level analysis on the entire dataset

In the time series analysis, the following forecasting methods were used to compare against the baseline lead times: Naïve Method; Simple Mean; Moving Average (n=5 and n=10); Simple Exponential Smoothing; Holt’s Method; and Holt-Winter’s Method. The resulting forecasts were compared against two baselines: planned lead time (lead time from purchase order) and the lead time gathered from regression analysis. The regression analysis baseline was used to provide a more lateral comparison, since the planned lead times were undoubtedly influenced by additional information occurring during execution that is not captured within the models.

The majority of the analyses in this research was performed at the lane level (unique vendor to plant combination) since both attributes appear to be significant factors in the variability of the actual lead time.

Results

The first phase of the research found that there were significant disconnects in the lead time variables used along the planning and execution process. The regression analyses performed on the static lead time variables against the planned lead time and the actual lead time showed very low R^2 values, with the R^2

values for regression on actual lead time being consistently worse than the values for the analysis on planned lead time. This means the static lead time variables were not good explanatory variables for either the planned lead time or the actual lead time and that there was a systematic difference in how the planned lead time is regularly closer to the static lead time variables than the actual lead time. Perhaps, the planners often added a buffer in their planning, so they would not be in understock situations. However, the R^2 values do appear to improve over time, which could indicate a more data-driven planning process was adopted by the Company.

Three sets of hypothesis tests (*t* tests) were performed at the SKU-Lane level. Over 25,000 SKU-Lanes were involved in each set of tests. The three tests examined whether the following set of data were statically the same:

- Static Lead Time based on SKU and Vendor (LTv) vs Planned Lead Time
- Static Lead Time based on SKU and Plant (LTp) vs Planned Lead Time
- Planned Lead Time vs Actual Lead Time

The tests showed that in all three set of tests for all the SKU-Lanes involved, for the majority of the SKU-Lanes, the two set of data are statistically different. This means that from the beginning of the planning process to the execution process, the lead times used were consistently different from each other.

The second phase of the research found that the forecast error can notably be reduced by using historical data to support the planning process. The lead time data was split between training set and testing set, and a rolling forecast was performed on SKU-Lanes with orders on over 100 distinct dates. Over 2,500 SKU-Lanes fit this criteria. Refer to Figure 2 for results.

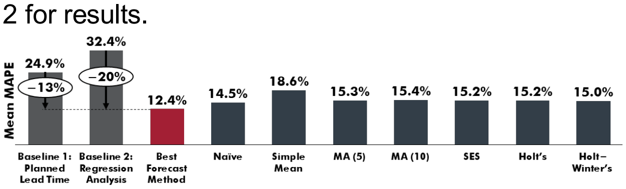

Figure 2. Mean MAPE of Forecast Methods vs. the Baselines

On average, the lowest forecast methods provided a 13% reduction in MAPE against the planned lead time baseline; and a 20% reduction in MAPE against the regression baseline. The analysis also showed that, on average, using any of the forecast methods would perform better than both baselines. Note that the Naïve method is advantaged by the gaps being front-filled during the analyses.

The Holt-Winter's method outperformed the other methods in over 40% of the SKU-Lanes examined (best), while Holt's method only outperformed the others in only 7% of the SKU-Lanes (last). This research then reviewed the different characteristics of the time series (e.g., level, trend, seasonality) to get a better context of the results. This research used Dickey-Fuller test, time series decomposition, and auto-correlation within descriptive analytics format to contextualize the different time series characteristics. It was found that trend did not play a major factor in the dataset, and that Holt-Winter's method outperformed other methods regardless of the 'magnitude' of seasonality present in the data.

This research also found that different lanes appear to have different profiles. Figure 3 illustrates how certain methods appear to be consistently more successful in certain lanes than others.

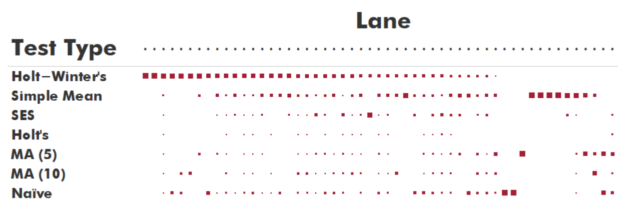

Figure 3. Winning Method by Lane

The more accurate predictions of lead time also has another quantifiable impact, and that was the safety stock required. For the same service level, using predictive lead time can result in a 21.3% and 6.4% reduction in safety stock quantity when compared to the regression and planned lead time baselines, respectively.

Conclusion

This research showed there is room for improvement in how lead time is managed and considered within the current planning process. Using predictive analytics to predict lead time (predictive lead time) can reduce the deviation between the planned and actual values in the lead time parameter. Predictive lead time can reduce the safety stock cost, the manual labor required in exception management (re-planning), and the manual labor in purchase order management (since the suggested lead time will be more accurate than the current process). A company implementing predictive lead time should consider assigning the forecast method by lane, focusing the implementation on the high volume, low cost items, and categorizing the different SKU-Lanes by level, trend and seasonality.