Outline for IJF paper on Whether Forecast Accuracy Matter Approximately 15 pages

Title:

Does Forecast Accuracy Matter: The Impact of Accuracy Improvement on Supply Chain Outcomes

Article Keywords: Forecast Accuracy, Supply Chain Outcomes, Business Application, Accuracy Relationship, Demand Planning

Abstract: Consulting and commercial organizations have frequently referenced the business and supply chain benefits of increasing forecast accuracy. The question of the relationship between forecast accuracy and business benefits is important for forecasting practitioners and researchers. Deciding whether to attempt to implement new forecasting models or modify parameters of existing models involves significant effort. By identifying the impact of improving forecast accuracy on supply chain outcomes, forecasters can apply a benefits lens to prioritize their efforts and determine the expected returns from their efforts. In this paper, we propose an approach for identifying the business benefits by utilizing a supply chain modeling method and perturbating forecast predications with directed error levels to identify the relationship between forecast accuracy and key supply chain output measures.

Literature review – included on each page of the submission at the bottom of each page.

Subdivision - numbered heads Use major and minor headings to aid readability. Headings should be short. Double space everything. Use wide margins.

# Introduction

State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.

Forecasting demand, particularly to support supply chain decisions, is one of the most frequently used implementations of forecasting in business. While there have been numerous papers, both commercial and academic, that compare forecast accuracy among specific forecasting methods, there are very few published research papers that explain how improving forecast accuracy, without regard to forecasting method, affects supply chain outcomes. Those papers that have been published normally compare specific methods (e.g., Naïve forecast versus exponential smoothing or machine learning methods). The range of differences in forecasting accuracy between methods is normally relatively small. Small differences in accuracy in these papers have then been utilized to measure the change in supply chain outcome. These comparisons in outcomes are developed using simulations of the decisions that are implemented in specific supply chain algorithms. Because the differences in accuracy are small, it is difficult to generalize the results of the supply chain outcomes across a broader range of forecast accuracy levels. One of the key research questions in this paper is, “As you improve forecast accuracy, are the impacts on supply chain outcomes consistent, or are there points where increased accuracy fails to yield the same level of supply chain improvement?”

Answering this question has significant implications for the process of trying to improve your forecast accuracy. In large firms making thousands, or millions of demand forecasts per day or month, determining which forecasts should be worked on for improvement is an important process question. Firms do not have unlimited resources to apply to this effort (human, software, algorithms, or hardware). This research, by applying consistent levels of accuracy (error), to the actual demand value, permits the supply chain outcome assessment for each Stock Keeping Unit (SKU) to be measured without regard to a forecast method. Doing so permits supply chain managers and forecasters to estimate the impact of improving the forecast accuracy for each SKU on key supply chain outcomes. This information can then be used to prioritize forecast improvement efforts and estimate the effect that a given level of forecast accuracy improvement will have on key supply chain outcomes.

Material and methods

Authors will normally be expected to submit a complete set of any data used in electronic form,

or provide instructions for how to obtain them. Exceptions to this requirement may be made at

the discretion of the handling editor. The author must describe methods and data sufficiently so

the research can be replicated. The provision of code as well as data is encouraged, but not

required. Upon acceptance, the data sets of published papers will be posted on the IJF website:

<http://www.forecasters.org/ijf/>.

In order to make this research replicable, the dataset used to represent demand values to be forecast was the daily demand of 30,490 SKUs provided by Walmart to support the MOFC, M5 Forecasting competition (<https://mofc.unic.ac.cy/m5-competition/>). The dataset of daily SKU demand can be retrieved from the Kaggle.com competition website (<https://www.kaggle.com/competitions/m5-forecasting-accuracy>). The file is sales\_train\_validation.csv. This file contains information about each SKU followed by 1913 daily observations of demand for each SKU from 2011-01-29 through 2016-04-04.

Forecasts with known levels of accuracy were developed at these levels of error – 50%, 40%, 30%, 20%, 10%, 5%, 2%, 1%, and 0% (no forecast error). These forecast errors were introduced by utilizing the actual demand for each SKU for each period of time and calculating the error (either plus or minus) and adding (or subtracting) the error from the actual value of demand at each of the different error percentages. This created nine different forecast values per day for each SKU.

In addition to the forecasts at the different known levels of accuracy (error), it is necessary to adopt a supply chain simulation method to capture the impacts of the different levels of accuracy. One of the most popular implementations of supply chain software in the world can be found in software provided by SAP Supply Chain Management (SCM). SAP’s software has been implemented by over 8,229 companies globally (enlyft, 2023). All of the calculations and input variables necessary for the supply chain outcome calculations were derived from the documentation for that software (SAP, 2021).

To estimate the supply chain outcomes for each SKU at each level of forecast accuracy (error), a simulation was run. The code is designed to permit any number of SKUs to be estimated at each level. The simulation code was developed in the data science program R, and utilized RStudio as the IDE. The code is available at: <https://github.com/Hoover-code/Forecasts_matter>. Supply chain outcome results are output in Microsoft Excel format which can then be used to assess results across SKUs.

Theory/calculation

A Theory section should extend, not repeat, the background to the article already dealt with in the Introduction and lay the foundation for further work. In contrast, a Calculation section represents a practical development from a theoretical basis.

# Calculation Methods

Kolassa (2022) and Robette (2022) proposed utilizing a simulation framework to assess the impact of improved forecast accuracy on supply chain outcomes. Neither proposed utilizing multiple known levels of forecast accuracy (error) within the simulation to assess the impact of improving the forecast on supply chain performance. The simulation approach for the assessment followed the steps below:

1. Decide on the key supply chain performance outcome measures. In this case, we utilized the following measures at each level of forecast accuracy:
   1. average inventory on hand
   2. backorder quantity per period
   3. percentage of total periods where backorders occurred
   4. backorder quantity / demand (i.e., fill rate)
   5. inventory holding cost
2. Collect time series and other data for the simulation. I this case, I utilized the SKU demand data from the M5 competition as the actual demand per period. Other fields are required to perform the supply chain simulation. This are not found in M5 competition data, and were assumed for simulation purposes. Some key fields within the simulation that were assumed include:
   1. Reorder Point, Fixed Order Quantity (R,Q) inventory replenishment policy
   2. Replenishment lead time (in days) – assumed fixed
   3. Forecast period (in days) – assumed to match replenishment lead time (in days)
   4. Ordering cost per order
   5. Holding cost per item
   6. Service level
3. Simulate multiple forecasts for each time period. In this step, forecasts were calculated using known intervals of forecast error. Those errors were either subtracted or added to the actual value that day to achieve a forecast. Below is an example calculation of the error and subsequent forecast when the actual value was 52.  
     
   A screenshot of a computer error

   Description automatically generated

Figure - Example calculation of forecast at different levels of forecast error.

Results

Results should be clear and concise.

Discussion

This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

Conclusions

The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section.

Appendices

If mathematical derivations are needed, put them in an Appendix.