Predicting Part Lead Times: Big Data Techniques

27 April 2016

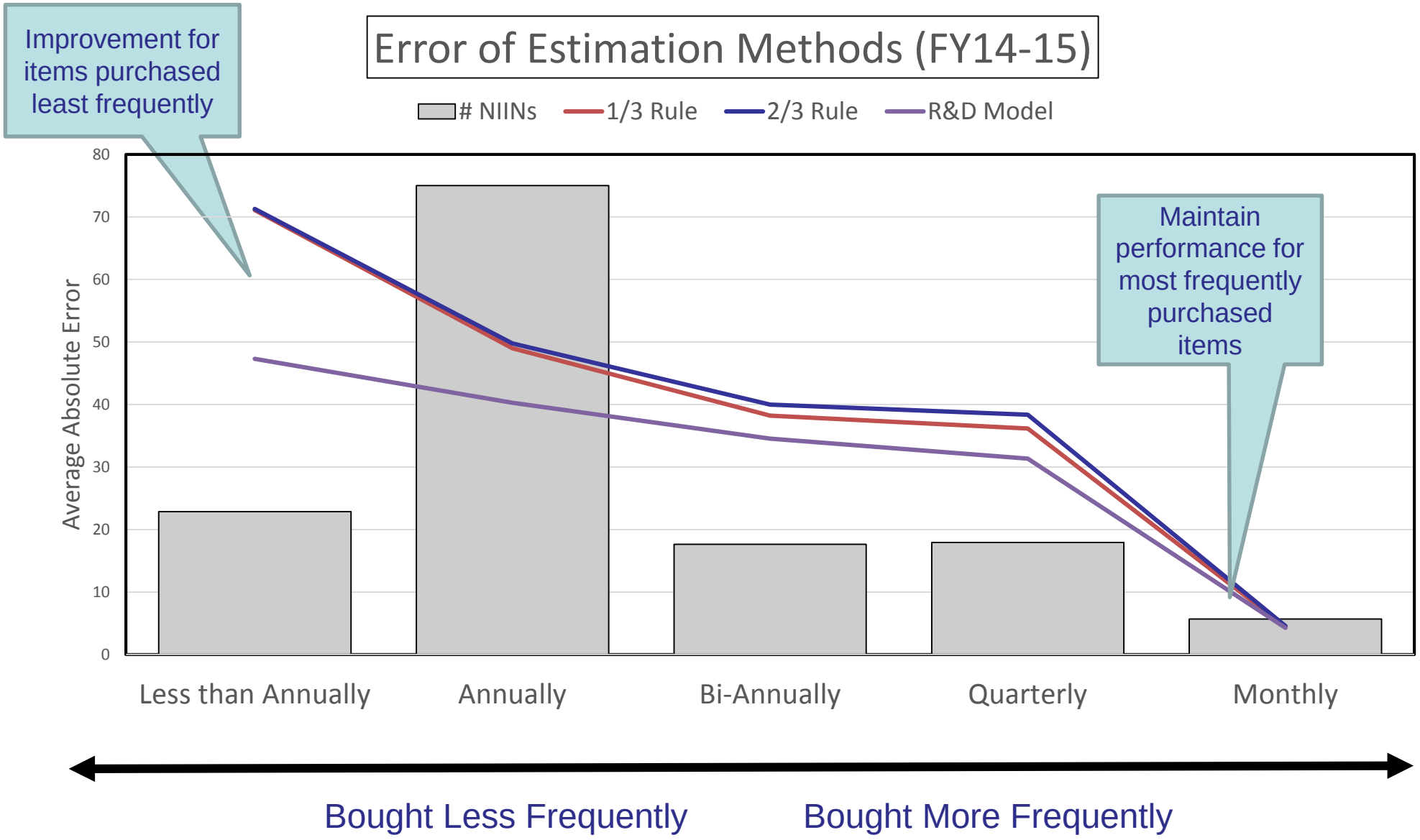
LMI

Table of Contents

Background

- Big data is the next big buzzword, but concrete application examples are sparse
- LMI conducted a research and development project for DLA WSSP, improving estimates of production lead time(PLT) using big data techniques
  - PLT is used by DLA to schedule orders, determine order quantities, and assign safety stock levels
  - Current estimation methods use weighted averages of past PLTs to estimate what PLT would be for an order made today
  - Inaccurate estimates result in delivery delays and increased inventory costs

Performance (order frequency)



The Problem

- Spare parts catalogs are characterized by a small number of frequently ordered parts, and a large number of infrequently ordered parts
- The simplest estimation approach assesses each part individually
  - Parts ordered infrequently have limited history on which to base estimation
- Other standard approaches would require consistent order history
- Ideally want a modeling approach that accommodates highly variable data

Big Data and Machine Learning

- Big data: a collection of tools, platforms, and techniques to squeeze out useful information from extreme quantities of data
- Machine learning: technique that uses algorithms (computer processes) to build a model rather than analysts pre-defining relationships
- Doesn’t presuppose any explicit relationships, but instead dynamically identifies relationships from the data
  - Allows the construction of more complex and nuanced models with less effort

Big Data Applied to PLT Estimation

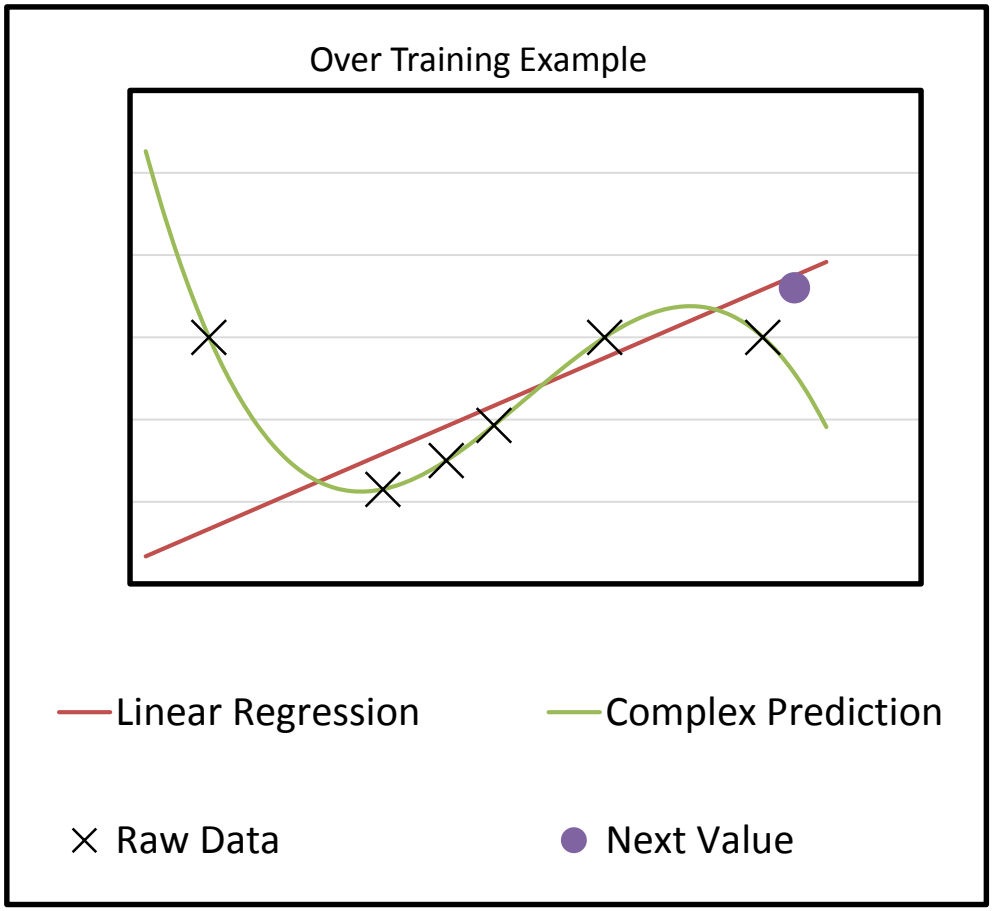
- Our approach was to use gradient boosted regression to:
  - Pool the entire catalog to identify similar items and take advantage of correlations in behavior
  - Account for systemic trends across DLA's system
- Supplement the amount of data available to predict any one part's PLT
  - More valuable for items with limited history
  - Includes items never purchased before

Implementation Requirements

- Enterprise Business Intelligence System capable of delivering quantities of valid data on the problem
- Statistical and/or big data packages
  - SAS, SPSS, HADOOP, etc.
- Analysts who understand the business process to be modeled and the data connections within your Enterprise Business Intelligence System
  - Essential that the analysts understand your business enough to identify false or misleading correlations

Potential Pitfalls

- Big data does not identify cause-effect relationships
  - Care must be taken in using results to make process improvements
- Training set and test set
  - Complex analysis methods can quickly “fit” to a set of data, without having predictive power
  - Consider the example, if we predict out in time on the green line we will predict a dramatic reduction in value
  - Define a test set that can be used to assess the model’s performance



Other Potential Uses of Big Data

- Buyer Assistance Systems
  - Identify similar products procured recently
  - Predict Fair and Reasonable Pricing
  - Identify potential bad vendor behaviors (e.x. Cage Hopping)
  - Identify potential “no quote” solicitations early
  - Identify vendors who may have the capability to make a part but have not bid
- Demand Forecasting
  - Link Maintenance data and Maintenance plans to drive procurement decisions
  - Predict buy/build arounds
  - Predict effects of OPTEMPO changes and deployment cycles