Journal of Operations Management 44 (2016) 62–68

Contents lists available at ScienceDirect

Journal of Operations Management

journal homepage: www.elsevier.com/locate/jom

Short communication

Determining an optimal hierarchical forecasting model based on the characteristics of the data set: Technical note

Zlatana D. Nenova*, Jerrold H. May

Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, PA 15260, USA

ARTICLE INFO

Article history:
Received 20 September 2015
Received in revised form
7 April 2016
Accepted 18 April 2016
Available online 28 April 2016
Accepted by: Daniel R. Guide

Keywords:
Forecasting
Hierarchical forecast aggregation
Supply chain management

ABSTRACT

The efficient flow of goods and services involves addressing multilevel forecast questions, and careful consideration when aggregating or disaggregating hierarchical estimates. Assessing all possible aggregation alternatives helps to determine the statistically most accurate way of consolidating multilevel forecasts. However, doing so in a multilevel and multiproduct supply chain may prove to be a very computationally intensive and time-consuming task. In this paper, we present a new, two-level oblique linear discriminant tree model, which identifies the optimal hierarchical forecast technique for a given hierarchical database in a very time-efficient manner. We induced our model from a real-world dataset, and it separates all historical time series into the four aggregation mechanisms considered. The separation process is a function of both the positive and negative correlation groups' variances at the lowest level of the hierarchical datasets. Our primary contributions are: (1) establishing a clear-cut relationship between the correlation metrics at the lowest level of the hierarchy and the optimal aggregation mechanism for a product/service hierarchy, and (2) developing an analytical model for personalized forecast aggregation decisions, based on characteristics of a hierarchical dataset.

1. Introduction

Accurate products' demand forecasts facilitate the smooth movement of goods through a supply network (Sanders and Manrodt, 2003). Not surprisingly, there exists a nontrivial literature on the development of guidelines to improve firms' sales forecasting methodologies (Moon et al., 2003). Such recommendations are particularly useful when creating forecasts for hierarchically organized data, for which models are designed to address managerial requests associated with a single or multiple levels of the firm's hierarchies. Providing data to support decisions involving multiple levels of a hierarchy typically requires consolidating hierarchical estimates, such as products' demand forecasts.

To ensure the reliability of their consolidated multilevel forecasts, statisticians assess the accuracy of various aggregation alternatives, because selecting an optimal consolidation method could significantly improve the accuracy of the overall forecast. The concept of aggregation bias was introduced by Theil (1954). Since then, researchers have examined the performance of various aggregation mechanisms. No consensus appears to have been reached as to which one is optimal. Some authors argue in favor of the top-down method (Fogarty et al., 1990; Grunfeld and Griliches, 1960; Narasimhan et al., 1995), while others support the use of the bottom-up approach (Dangerfield and Morris, 1992; Schwarzkopf et al., 1988). The inconsistent findings lead to the conjecture that there exists some population of data sets for which top-down aggregation is optimal, and a second population of data sets for which bottom-up aggregation is optimal. Ideally, we would like to find a way to describe and separate those two populations of data sets, and to provide a theoretical justification as to why the data in each population is best modeled using the particular aggregation technique. Within the academic community, the search for such descriptors and separators of the two populations has led to the examination of the association between bottom-level series correlations and optimal forecast aggregations.

Schwarzkopf et al. (1988), limiting their discussion to two products, note that estimation precision is a function of product correlation. Estimates based on aggregated data may have a smaller variance than those based on individual forecasts when the items are independent, but they may have greater variance when the items have a strong positive correlation, as well as a much smaller variance if the items have a strong negative correlation. They

* Corresponding author.
E-mail address: zdn3@pitt.edu (Z.D. Nenova).

http://dx.doi.org/10.1016/j.jom.2016.04.001
0272-6963/

suggest that negative correlation benefits top-down approaches; Duncan et al. (2001) claim that positive correlation benefits top-down approaches. Chen and Boylan (2009) resolve that contradiction by noting that negative correlation only reduces variability if the data series involved follow the same model.

Correlation, though, is a bivariate measure. When a group of $n \geq 3$ products is to be forecasted, and the correlation matrix for those products contains both positive and negative entries, it is more challenging to quantify the impact of those correlations on the optimal aggregation technique. When there are three or more products, Chen and Boylan consider the positive and the negative correlations separately, and use the ratio of the sums of the positive and negative correlation coefficients as a predictor.

In a simulation study, Widiarta et al. (2007) show that, when the lag-1 autocorrelation of the demand for at least one of the items in the bottom level is greater than one third, the bottom-up method is consistently better, regardless of the item's proportion in the family, and regardless of the coefficient of correlation between the error terms of the two item demand processes. They also state that there is no significant difference between the aggregation strategies if the lag-1 autocorrelation of the two item demands have a coefficient of serial correlation less than or equal to one third. In a more recent article, Widiarta and Viswanathan (2008) observe that "regardless of the coefficient of correlation between item demands, the items' proportion in the family and the coefficient of serial correlation term" there is no significant difference between the bottom-up and top-down forecast strategies when the examined time series follow an MA(1) process.

Kremer et al. (2015) argue that, from a judgmental forecasting perspective, the correlation structure of the lower hierarchical level determines, to a great extent, the forecast aggregation mechanism for a hierarchically organized time series. The authors conclude that bottom-up forecasting performs better for hierarchical datasets with lower level items with similar responses to long- and short-term shocks. They also state that top-down forecasting is preferable when a data set's lower level products are substitutes for each other either in the long or in the short run.

Our contention is that correlations can be used to create a rule-based model to select the most appropriate aggregation approach. Using correlations for such a model is attractive, because they do not make any assumptions about the structure of the data, and determine a recommendation based only on the characteristics of the data themselves. To the best of our knowledge, no model of that type has appeared in the literature to date.

We propose a time-efficient alternative to the exhaustive aggregation selection mechanism. Our objective is to develop a reliable and easy-to-implement prediction model, which provides quick and accurate results when consolidating multilevel forecasts.

Similar to Loh and Vanichsetakul (1988) and Lopez-Chau et al. (2013), our prediction model has an oblique decision tree structure (Murthy et al., 1994), and uses Fisher's linear discriminant to find optimal splits at each node of the tree. We consider, in our model, four forecast aggregation methods, one bottom-up and three that are variations on top-down. Because of that, our model has two levels. At the first level, we use a discriminant function to classify a dataset as bottom-up or top-down. At the second level, a different discriminant function separates data sets that are predicted to be top-down into the three top-down aggregation groups considered.

We contribute to the forecasting literature in the following ways. First, we establish a direct relationship between the optimal aggregation mechanism and the correlation metrics at the lowest hierarchical level, based on the analysis of a large, real-world database. Second, using those correlation metrics, we develop an analytical model to choose an expected optimal forecast consolidation strategy for each dataset. We show, empirically, the gain from using our adaptive procedure, as opposed to using the same strategy for all datasets.

The remainder of the paper is organized as follows. We discuss the procedure and data used to obtain the model building dataset in §2. We describe the optimal aggregation prediction model and its performance in §3 and 4. Concluding remarks and future research are addressed in §5.

2. Modeling methodology

In an effort to avoid restrictive assumptions, which may hinder our model's generalizability, we use an empirical, rather than a theoretical approach (Schwarzkopf et al., 1988; Widiarta et al., 2007; Widiarta and Viswanathan, 2008). Our model is induced from a real-world data, rather than from a simulated database (Chen and Boylan, 2009), because we believe that a sufficiently complex real dataset is more appropriate than would be randomly generated samples from a population that may not closely resemble real-life data. Our training and test data sets consist of a range of different products from two item groups. They were provided by a major U.S.-based food production and distribution corporation.

Our first steps were to (1) select the candidate aggregation techniques, (2) select the forecasting methods to be used with those aggregation techniques, (3) define the forecast horizon to be used, (4) decide how to measure forecast accuracy, and (5) propose a way to determine the better approach for each of the datasets available. We then (6) examine the results of our choices on the number of datasets used for modeling.

2.1. Selecting the candidate aggregation techniques

We considered four aggregation approaches: bottom-up (**bu**), picked due to its ability to capture time series dynamics at the lowest database level, and three top-down methods, which do not face the **bu** challenge of modeling noisy bottom level data (Grunfeld and Griliches, 1960). Two of the top-down approaches, top-down using average historical proportions (**tdgsa**) and top-down using the proportions of historical averages (**tdgsf**), are among the most common top-down methods, which allocate forecasts based on the historical proportions of the model building data. We chose them because they are among the most promising disaggregation approaches out of the twenty-one schemas in Gross and Sohl (1990). The third approach, top-down using forecasted proportions (**tdfp**), was chosen because of its noteworthy performance in Athanasopoulos et al. (2009).

A different top-down approach, which we do not include in our empirical analysis, is the *optimal reconciliation method* of Hyndman et al. (2011). The authors first generate independent forecasts for all nodes. Those forecasts may not sum to the desired quantity. A reconciliation step is used to adjust the independent forecasts, in order to produce values that are consistent with the hierarchical structure. The revised forecast at each node is a weighted average of the forecasts from all nodes. They obtain the weights for reconciliation from a linear regression of the independent forecasts from all nodes against a set of indicator variables that identify which of the bottom-level series contribute to each node. The weights are "optimal" in the sense that the mean squared reconciliation error, computed using the differences between the reconciled and independent forecasts, is as small as possible. The weights in the *optimal reconciliation method* depend only on the hierarchical structure and not on the values of the observed data. Thus, the reconciliation weights only need to be calculated once for each hierarchy (Hyndman and Athanasopoulos, 2014).