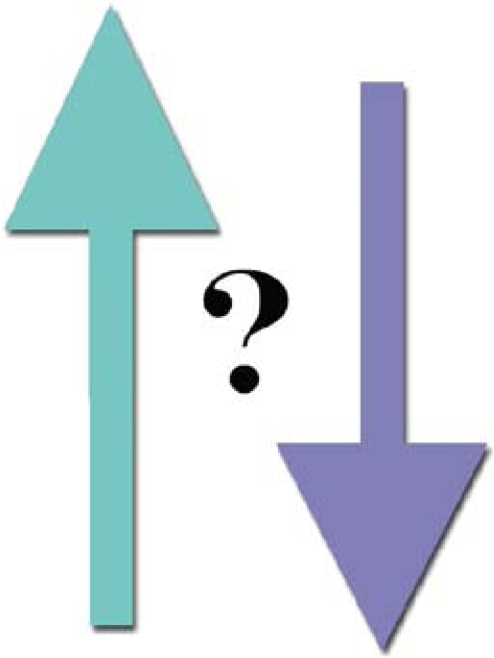
FORESIGHT

THE INTERNATIONAL JOURNAL OF APPLIED FORECASTING

ARTICLES *by* FORECASTERS *for* FORECASTERS

Join the *Foresight* readership by becoming a member of the International Institute of Forecasters

forecasters.org/foresight/



Choosing Levels of Aggregation for Supply Chain Forecasts

John Boylan

PREVIEW *With this issue,* ***Foresight*** *debuts a new column on supply-chain forecasting. John Boylan, our column editor, will offer his insights on the practical challenges of forecasting and reconciling forecasts across the hierarchical levels of a supply chain.*

INTRODUCTION

Supply chains are characterized by multiple hierarchies across products, markets, and geographical sectors. There is a further hierarchy, one often taken for granted and by no means unique to supply chains: a hierarchy of time. For example, quarterly data may be considered as an aggregation of monthly data.

Forecasts are often needed at different levels of aggregation in supply chains and for different purposes. This makes hierarchical forecasting an important topic, but it is only recently that the subject has been receiving the attention it deserves. In this column, I will review some recent developments in hierarchical forecasting and provide some recommendations for forecasting practice.

HIGH-LEVEL FORECASTS

In all areas of management, high-level (strategic) decisions should precede and set the context for lower-level (operational) decisions. Supply chain management is no exception to this general rule.

In recent years, an important advance has been the development of systematic approaches to high-level planning, particularly in a manufacturing environment. Experienced supply-chain consultants have recognised this need. George Palmatier (2007) makes the point that Sales and Operations Planning (S&OP) should be an executive responsibility, led from the top, and not simply delegated to a cross-functional, middle-management team.

The Aberdeen Group (2007) distinguished between Executive and Operational S&OP. *Executive S&OP* is concerned with volume issues, such as capacity, in contrast to *Operational S&OP*, which focuses on detailed sequencing (also known as the "mix"). Bob Stahl's columns in *Foresight* make a strong case for Executive S&OP. In their case study of Jarden Branded Consumables, Stahl and Brad McCollum (2009) point out that, although problems often emerge in the mix, the cause may relate to insufficient manufacturing capacity. In such cases, it is pointless to refine operational planning if strategic planning at the executive level is the real problem.

In a nonmanufacturing environment, a distributor also needs to conduct aggregate planning to support decisions about warehousing and transportation capacity. This means that, for both manufacturers and distributors, aggregate planning must

Key Points

- Approaches to forecasting a product hierarchy include bottom-up, top-down, and hybrid modeling. Each has its merits and drawbacks, but sole reliance on bottom-up forecasting is not a good strategy.
- In top-down, a higher-level forecast must be disaggregated, and the allocation is often based on *historical proportions*. A potentially superior alternative to historical proportions is that of *forecasted proportions*, so it is wise to compare these two techniques.
- If using exponential smoothing, consider a hybrid approach in which we estimate seasonal factors from a product family and other components (level and trend) from the individual series.
- For intermittent-demand data, try approaches based on aggregation over time, particularly if it is possible to match the historical time units to the forecast horizon.

precede operational planning to ensure an ongoing balance between supply and demand. High-level forecasts, outside standard lead times, are therefore crucial for the success of the supply chain.

BOTTOM-UP FORECASTS

Let's suppose that statistical forecasts are required at two distinct levels – an upper level, such as countrywide or product family, and a lower level, such as individual region or item within a product family.

One approach is to use statistical methods to generate all the lower-level (regional) forecasts and then to add these forecasts to give a higher-level forecast for the whole country. This is the *bottom-up* approach. This is often the only approach which is tried, either because alternatives have not been considered or because the forecasting software does not facilitate them.

The bottom-up approach has certain advantages. It is able to model the distinct characteristics of each of the individual regions without having to make assumptions about common properties across regions, such as that they are all growing at the same rate. A study based on real (M-competition) data by Byron Dangerfield and John Morris (1992) supported the bottom-up approach, showing it was more accurate in forecasting the lower levels than top-down methods; however, the analysis was restricted to small families of products.

The bottom-up approach has problems if the lower-level data histories are short, as is often the case in supply-chain forecasting. With a short history, and especially if the series is noisy, it is difficult to identify the best forecast method to use – whether, for example, the forecast should be based on a straight trend or a damped trend. Even if the correct method is identified, it can be difficult to estimate the components of the model, such as the seasonal indices.

The bottom-up approach also restricts the choice of forecasting models. At the lower levels, it is very difficult to fit econometric models that link demand to leading indicators, such as the Consumer Confidence Index; at the higher level, though, this may be feasible. For example, Stahl and McCollum (2009) show how a company was able to define market-facing product families and use regression modeling for each family (higher-level) based on leading economic indicators.

Finally, adding the lower-level forecasts to obtain the higher-level forecast may work well over short forecast horizons, but it is more questionable over longer horizons.

TOP-DOWN FORECASTS AND DISAGGREGATION

Top-down forecasts begin with an aggregate forecast at the higher level, which is the forecast disaggregated to each of the lower levels.

There are instances in supply-chain forecasting when the top-down approach is clearly more appropriate than bottom-up methods. If there is a change in policy, for example, with regard to pack sizes, then aggregate data (in suitable units of measurement) is a better guide to the future than previous sales of a particular pack size. In this case, the aggregate forecast may require disaggregation using judgmental estimates of the effect of introducing new pack sizes.

The first systematic study of disaggregation mechanisms was undertaken by Charles Gross and Jeffrey Sohl (1990). These authors found that two disaggregation methods were particularly effective:

- For each period, calculate the proportion that an individual series contributes toward the aggregate series. Then average the proportions over all periods.
- Calculate the total, over all periods, for the individual series, and then do the same for the aggregate series. Then calculate the ratio of the former to the latter.

George Athanasopoulos and colleagues (2009) refer to the first method as the *average historical proportions* and the second as the *proportions of the historical averages*. For the first approach, they point out that we need not restrict our methods to historical proportions. Instead, we can use *forecasted proportions*. The researchers found that the use of forecasted proportions worked well in predicting Australian domestic tourism and was more accurate than the two standard approaches recommended by Gross and Sohl. Although not definitively tested on supply-chain data, the use of forecasted rather than historical proportions appears to be promising and is a feature typically offered in forecasting software.

The strengths and weaknesses of top-down approaches are the opposite of bottom-up forecasts. Longer data histories make model identification and component estimation more reliable at the aggregate level. On the other hand, when we disaggregate, we are making strong assumptions that the lower-level series are following the same trend and seasonal pattern as the higher-level data, and these assumptions may be incorrect.

In multi-level supply chains, we can also use a mixture of approaches. For example, in a three-level supply chain, a "middle-out" approach can be adopted. This means that the top-level forecasts are calculated as the sum of the middle-level forecasts (bottom-up) but the bottom-level forecasts are disaggregated from the middle-level approaches (top-down). Middle-out forecasting is an issue that is ripe for research.

There have been many studies comparing the merits of top-down and bottom-up forecasts, with some studies favoring one approach and some the other. This lack of conclusive evidence has led to the search for alternative approaches.

HYBRID FORECASTS

Suppose we believe that the level and trend are unique to a particular item or product, but the seasonal pattern is shared by all the items in a product family. Mark Dekker and colleagues (2004) analysed such products from two large wholesalers, one selling food and the other electro-technical products. They suggested the following forecasting procedure, but with changes in level of sales and possibly in the trends from one period to the next:

- Aggregate the sales data for all items belonging to the same product family.
- Use Winters' exponential smoothing to model aggregate sales, obtaining estimates of the seasonal indices (as well as level and trend).
- Use these seasonal indices to deseasonalize the sales data for the individual series.

- Apply nonseasonal exponential-smoothing models to the individual sales series to obtain estimates of the current level and possible trend in these data.
- Reseasonalize the forecasts for the individual series using the aggregate seasonal indices.

This procedure, which is available in some software packages, is a hybrid method: data on the individual series are used to find the demand level (and trend), but the seasonal indices are based on aggregate data only. The originators of this procedure found that it produced more stable and more accurate forecasts than Winters' method applied to individual series, and it was easier to find the seasonal indices.

George Athanasopoulos and colleagues (2009) discuss an alternative hybrid method called the "optimal combination approach." First you independently forecast all series at all levels of a hierarchy. Then you use a statistical formula (based on regression analysis) to combine and reconcile the forecasts. They found this optimal combination method to be a strong competitor to the top-down method using forecasted, rather than historical, proportions (FP).

Whilst the top-down vs. bottom-up debate may not yet be resolved, we can see that the product-family variant of the Winters' method and the forecasted-proportions version of the top-down approach both offer viable alternatives. The optimal combination approach may also be considered, and it would be interesting to test it on supply-chain data.

FORECASTS BASED ON AGGREGATION OVER TIME

In most operational supply chain applications, we need a forecast over lead time (or lead time plus review time). This requirement defines a natural time unit that may not equate to the time buckets in which data is collected. For example, we may need a forecast of total demand over the next three months, but our historical data is available for each month. In this situation, we would typically make forecasts for one, two, and three months ahead, and then sum these forecasts to give the lead time forecast.

Is this the best way to forecast? An alternative approach, known as *temporal aggregation* or *aggregation over time*, is to choose different time buckets for the demand history, and to forecast accordingly. The historical data may be re-presented in a quarterly format, and then a one-step-ahead forecast would immediately give the desired result – a quarterly forecast. Alternatively, longer time buckets could be employed, and then the final forecast would have to be adjusted accordingly.

Kostas Nikolopoulos and colleagues (2010) argue that the temporal-aggregation approach is particularly useful for intermittent-demand forecasting. Intermittent demand is characterized by frequent occurrences of zero demand. Temporal aggregation reduces the frequency of zeroes, making the series easier to forecast. The researchers analyzed the demand of 5,000 Stock-Keeping Units from the Royal Air Force and found that temporal aggregation acted as a self-improvement mechanism for all the forecasting methods investigated, including Croston's method and the Syntetos-Boylan Approximation (see Boylan, 2005 for further details). Setting the aggregation level to one lead time plus one review period showed particularly promising results.

RECOMMENDATIONS

My first recommendation is that you take a long view, based on aggregate data, and use the precepts of Executive S&OP. Doing so will help ensure that the right strategic decisions are made.

At an operational level, there are benefits to basing lower-level forecasts on higher-level forecasts, such as in top-down and hybrid approaches. While there is real-world evidence in favor of these approaches, I advise experimentation on your own data before adopting any particular method.

For intermittent-demand data, the temporal-aggregation methods approach is simple, and there is evidence in its favor. Again, some experimentation on real data would allow for the best time buckets to be identified and for the method's performance to be judged against more conventional approaches.

REFERENCES

Aberdeen Group (2007). *Executive Sales and Operations Planning: Process and Technology*, Aberdeen Group, Inc.

Athanasopoulos, G., Ahmed, R.A. & Hyndman, R.J. (2009). Hierarchical forecasts for Australian domestic tourism, *International Journal of Forecasting*, 25, 146-166.

Boylan J.E. (2005). Intermittent and lumpy demand: a forecasting challenge, *Foresight: the International Journal of Applied Forecasting*, Issue 1 (June 2005), 36-42.

Dangerfield, B.J. & Morris, J.S. (1992). Top-down or bottom-up: Aggregate versus disaggregate extrapolations, *International Journal of Forecasting*, 8, 233-241.

Dekker, M., van Donselaar, K. & Ouwehand, P. (2004) How to use aggregation and combined forecasting to improve seasonal demand forecasts, *International Journal of Production Economics*, 90, 151-167.

Gross, C.W. & Sohl, J.E. (1990). Disaggregation mechanisms to expedite product line forecasting, *International Journal of Forecasting*, 9, 233-254.

Nikolopoulos, K., Syntetos, A.A., Boylan, J.E, Petropoulos, F. & Assimakopoulos, V. (2010). An aggregate-disaggregate intermittent demand approach: an empirical proposition and analysis, *Journal of the Operational Research Society* (doi:10.1057/jors.2010.32).

Palmatier, G.E. (2007). The need to lead, *Business Excellence*, March 2007, 5.

Stahl, R.A. (2009). Sales & Operations Planning: Simpler, Better and Needed More than Ever, *Foresight: The International Journal of Applied Forecasting*, Issue 14 (Summer 2009), 12-16.

Stahl, R.A. & McCollum, B. (2009). How Jarden Branded Consumables made forecasting simpler and better through executive S&OP, *Foresight: The International Journal of Applied Forecasting*, Issue 15 (Fall 2009), 25-31.

John Boylan is Professor of Management Science at Lancaster University. He specializes in supply-chain forecasting, and has researched intermittent-demand forecasting and the interface between forecasting and inventory management. John is the supply-chain forecasting editor for *Foresight*.

j.boylan@lancaster.ac.uk