HOW TO TRACK FORECAST ACCURACY TO GUIDE FORECAST PROCESS IMPROVEMENT

by Jim Hoover

PREVIEW
While considerable attention has been paid to the measurement of forecast accuracy for individual items at particular points in time, issues around an aggregated forecast-accuracy metric and its tracking over time still present opportunities for discussion. Jim Hoover talks about why organizations have neglected the task of tracking forecast accuracy and offers a step-by-step guide for getting back on the track.

INTRODUCTION
One of the more important tasks in supply-chain management is improving forecast accuracy. Because your investment in inventory is tied to it, forecast accuracy is critical to the bottom line. If you can improve accuracy across your range of SKUs, you can reduce the safety-stock levels needed to reach target fill rates.

The desire to improve accuracy is a principal factor behind Enterprise Resource Planning (ERP) systems. Phil Robinson, an ERP implementation consultant (2008), found that the typical reported reduction in inventory levels achieved for companies utilizing ERP systems is about 30%, with some organizations reaching 75% reductions.

I have seen a great deal of information in the forecasting literature on measuring forecasting accuracy for individual items at a point in time but see very little attention paid to the issues of tracking changes in forecasting accuracy over time, especially for the aggregate of items being forecast. Foresight has begun to address this topic with a case study from Robert Rieg (2008).

In practice, the portion of firms tracking aggregated accuracy is surprisingly small. Teresa McCarthy and colleagues (2006) reported that only 55% of the companies they surveyed believed that forecasting performance was being formally evaluated. When I asked the same question at a recent conference of forecasting practitioners, I found that approximately half of the participants indicated that their company tracked forecast accuracy as a key performance indicator; less than half reported that financial incentives were tied to forecast-accuracy measurement.
OBSTACLES TO TRACKING ACCURACY

Why aren’t organizations formally tracking forecast accuracy? One reason is that forecasts are not always stored over time. Many supply-chain systems with roots in the 1960s and 1970s did not save prior-period forecasts because of the high cost of storage in that era. Technology advances have reduced storage costs and, while the underlying forecast applications have been re-hosted on new systems, they have not been updated to retain prior forecasts, thus forfeiting the possibility of tracking performance over time.

A second reason is that saving the history in a useful manner sometimes requires retention of the original customer-level demand data. These are the data that can later be rebuilt into different levels of distribution-center activity, when DCs are added or removed. This additional requirement creates a much larger storage challenge than saving just the aggregated forecasts.

Third, there are companies that haven’t settled on a forecast-accuracy metric. While this may seem to be a simple task, the choice of metric depends on the nature of the demand data. For intermittent demands, popular metrics such as the Mean Absolute Percentage Error (MAPE) are inappropriate, as pointed out in Hoover (2006).

Finally, some companies don’t have processes in place that factor forecast-accuracy metrics into business decisions. So they lack the impetus to track accuracy.

MULTISTEP TRACKING PROCESS

A process for effective tracking of forecasting accuracy has a number of key steps, as shown in Figure 1.

**Step 1. Decide on the Forecast-Accuracy Metric**

For many forecasters, the MAPE is the primary forecast-accuracy metric. Because the MAPE is scale-independent (since it is a percentage error, it is unit free), it can be used to assess and compare accuracy across a range of items. Kolassa and Schutz (2007) point out, however, that this virtue is somewhat mitigated when combining low- and high-volume items.

The MAPE is also a very problematic metric in certain situations, such as intermittent demands. This point was made in a feature section in *Foresight* entitled “Forecast-Accuracy Metrics for Inventory Control and Intermittent Demands” (Issue 4, June 2006). Proposed alternatives included the MAD/Mean ratio, a metric...
which overcomes many problems with low-demand SKUs and provides consistent measures across SKUs. Another metric is the Mean Absolute Scaled Error, or MASE, which compares the error from a forecast model with the error resulting from a naïve method. Slightly more complex is GMASE, proposed by Valentin (2007), which is a weighted geometric mean of the individual MASEs calculated at the SKU level. Still other metrics are available, including those based on medians rather than means and using the percentage of forecasts that exceed an established error threshold.

In choosing an appropriate metric, there are two major considerations. The metric should be scale-independent so that it makes sense when applied to an aggregate across SKUs. Secondly, the metric should be intuitively understandable to management. The popularity of the MAPE is largely attributable to its intuitive interpretation as an average percentage error. The MAD-to-Mean is nearly as intuitive, measuring the average error as a percent of the average volume. Less intuitive are the MASE and GMASE.

I would recommend the more intuitive metrics, specifically MAD-to-Mean, because they are understandable to both management and forecasters. Using something as complicated as MASE or GMASE can leave some managers confused and frustrated, potentially leading to a lack of buy-in or commitment to the tracking metric.

**Step 2. Determine the Level of Aggregation**

The appropriate level of aggregation is the one where major business decisions on resource allocation, revenue generation, and inventory investment are made. This ensures that your forecast-accuracy tracking process is linked to the decisions that rely on the forecasts.

If you have SKUs stored both in retail sites and in a distribution center (DC), you will have the option to track forecast error at the individual retail site, at the DC, or at the overall aggregate level. If key business decisions (such as inventory investment and service level) are based on the aggregate-level SKU forecasts and you allocate that quantity down your supply chain, then you should assess forecast accuracy at the aggregate level. If you forecast by retail site and then aggregate the individual forecasts up to the DC or at the overall SKU aggregate, then you should be mea-
suring forecasting accuracy at the individual site level. Again, the point is to track accuracy at the level where you make the important business decisions.

Additionally, you should consider tracking accuracy across like items. If you use one service-level calculation for fast-moving, continuous-demand items, and a second standard for slower- and intermittent-demand items, you should calculate separate error measures for the distinct groups.

Table 1 illustrates how the aggregation of the forecasts could be accomplished to calculate an average aggregate percent error for an individual time period.

**Step 3. Decide Which Attributes of the Forecasting Process to Store**

There are many options here, including:
- the actual demands
- the unadjusted statistical forecasts (before override or modifications)
- when manual overrides were made to the statistical forecast, and by whom
- when outliers were removed
- the method used to create the statistical forecast and the parameters of that method
- the forecaster responsible for that SKU
- when promotions or other special events occurred
- whether there was collaboration with customers or suppliers
- the weights applied when allocating forecasts down the supply chain

Choosing the right attributes facilitates a “forecasting autopsy,” which seeks explanations for failing to meet forecast-accuracy targets. For example, it can be useful to know if forecast errors were being driven by judgmental overrides to the statistical forecasts. To find this out requires that we store more than just the actual demands and final forecasts.

Figure 2 presents a flowchart illustrating the sequence of actions in storing key attributes. Please note that the best time to add these fields is when initially designing your accuracy-tracking system. It is more difficult and less useful to add them later, it will cost more money, and you will have to baseline your forecast autopsy results from the periods following any change in attributes. It is easier at the outset to store more data elements than you think you need, rather than adding them later.

<table>
<thead>
<tr>
<th>SKU</th>
<th>History Current Period</th>
<th>Forecast for Current Period</th>
<th>Error</th>
<th>Absolute Error</th>
<th>Absolute Percent Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKU 1</td>
<td>20</td>
<td>18</td>
<td>2</td>
<td>2</td>
<td>10.0%</td>
</tr>
<tr>
<td>SKU 2</td>
<td>10</td>
<td>15</td>
<td>-5</td>
<td>5</td>
<td>50.0%</td>
</tr>
<tr>
<td>SKU 3</td>
<td>50</td>
<td>65</td>
<td>-15</td>
<td>15</td>
<td>30.0%</td>
</tr>
<tr>
<td>SKU 4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>60.0%</td>
</tr>
<tr>
<td>SKU 5</td>
<td>3</td>
<td>8</td>
<td>-5</td>
<td>5</td>
<td>166.7%</td>
</tr>
<tr>
<td>SKU 6</td>
<td>220</td>
<td>180</td>
<td>40</td>
<td>40</td>
<td>18.2%</td>
</tr>
</tbody>
</table>

**Table 1. Calculation of an Aggregate Percent Error (MAPE)**
Figure 2. Flowchart for Storing Attributes of a Forecasting Process

Step 4. Apply Relevant Business Weights to the Accuracy Metric

George Orwell might have put it this way: “All forecasts are equal, but some are more equal than others.” The simple truth: you want better accuracy when forecasting those items that, for whatever reason, are more important than other items.

The forecast-accuracy metric can reflect the item’s importance through assignment of weights. Table 2 provides an illustration, using inventory holding costs to assign weights.

As shown in this example, SKUs 3 and 6 have the larger weights and move the Weighted APE metric down from the average of 55.8% (seen in Table 1) to 21.4%.

Use the weighting factor that makes the most sense from a business perspective to calculate your aggregated periodic forecast-accuracy metric. Here are some weighting factors to consider:

- inventory holding costs
- return on invested assets
- expected sales levels
- contribution margin of the item to business bottom line
- customer-relationship metrics
- expected service level
- “never out” requirements (readiness-based)
- inventory

Weighting permits the forecaster to prioritize efforts at forecast-accuracy improvement as shown in Step 6 in Table 2, below.

Step 5. Track the Aggregated Forecast-Accuracy Metric over Time

An aggregate forecast-accuracy metric is needed by top management for process review and financial reporting. This metric can serve as the basis for tracking process improvement over time. Similar to statistical process-control metrics, the forecast-accuracy metric will assess forecast improvement efforts and signal
major shifts in the forecast environment and forecast-process effectiveness, both of which require positive forecast-management action.

Figure 3 illustrates the tracking of a forecast-error metric over time. An improvement process instituted in period 5 resulted in reduced errors in period 6.

**Figure 3. Illustration of a Tracking Signal**

You can see that SKU 6 has the largest impact on the weighted APE tracking metric. Even though SKU 4 has the second-highest error rate of all of the SKUs, it has very little effect on the aggregated metric.

**Step 7. Apply Best Forecasting Practices**

Once you have identified those items where forecast improvement should be concentrated, you have numerous factors to guide you. Did you:

- apply the Principles of Forecasting (Armstrong, 2000)?
- try automatic forecasting methods and settings?
- analyze the gains or losses from manual overrides?
- identify product life-cycle patterns?
- determine adjustments that should have been made (e.g., promotions)?
- evaluate individual forecaster performance?
- assess environmental changes (recession)?

As Robert Reig reported in his case study of forecast accuracy over time (2008), significant changes in the environment may radically affect forecast accuracy. Events like the current economic recession, the entry of new competition into the market space of a SKU, government intervention (e.g., the recent tomato salmonella scare), or transportation interruptions can all dramatically change the accuracy of your forecasts. While the change might not be the forecaster's "fault," tracking accuracy enables a rapid response to deteriorating performance.

**Step 8. Repeat Steps 4 through 7 Each Period**

All of the factors in Step 7 form a deliberative, continuous responsibility for the forecasting team. With

**Table 3. Targets for Forecast Improvement**

<table>
<thead>
<tr>
<th>SKU</th>
<th>History Forecast</th>
<th>Error</th>
<th>Absolute Error</th>
<th>Absolute Percent Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKU 1</td>
<td>20</td>
<td>10</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>SKU 2</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>SKU 3</td>
<td>50</td>
<td>65</td>
<td>-15</td>
<td>15</td>
</tr>
<tr>
<td>SKU 4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>SKU 5</td>
<td>3</td>
<td>8</td>
<td>-5</td>
<td>5</td>
</tr>
<tr>
<td>SKU 6</td>
<td>220</td>
<td>100</td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th>Cost of Item</th>
<th>Inventory Holding Cost</th>
<th>Percentage of Total Cost</th>
<th>Weighted APE Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKU 1</td>
<td>$50.00</td>
<td>$0.00</td>
<td>0.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>SKU 2</td>
<td>$50.00</td>
<td>$750.00</td>
<td>4.4%</td>
<td>2.2%</td>
</tr>
<tr>
<td>SKU 3</td>
<td>$25.00</td>
<td>$1,625.00</td>
<td>9.6%</td>
<td>2.0%</td>
</tr>
<tr>
<td>SKU 4</td>
<td>$5.00</td>
<td>$10.00</td>
<td>0.1%</td>
<td>0.0%</td>
</tr>
<tr>
<td>SKU 5</td>
<td>$15.00</td>
<td>$120.00</td>
<td>0.8%</td>
<td>0.2%</td>
</tr>
<tr>
<td>SKU 6</td>
<td>$75.00</td>
<td>$13,500.00</td>
<td>79.9%</td>
<td>14.5%</td>
</tr>
</tbody>
</table>
the proper metrics in place, forecasters can be held accountable for the items under their purview. Steps 4-7 should be repeated each period, so that the aggregated forecast-accuracy metric is continually updated for management and new targets for improvement emerge.

CONCLUSIONS AND RECOMMENDATIONS
Forecast accuracy has a major impact on business costs and profits. The forecasting process must be evaluated by individual and aggregated forecast-accuracy metrics. Tracking these metrics over time is critical to driving process improvement.

See if your company has included forecast accuracy as a key performance indicator for management. If it has not, create a plan to begin recording accuracy at the aggregated level, and sell the idea to management. Build a tracking database that saves the key attributes of the forecasting process. Doing so will permit forecasting autopsies, which drive improvement efforts and prioritization of forecaster workload. See if you have weighted the forecasts to include the relative business impact, and make sure you have a structured approach to improving the individual and aggregated forecast accuracy over time. The data gathered in a good tracking process should lead to any number of improved business outcomes.

REFERENCES


Bio: Jim Hoover retired from the U.S. Navy in mid June. He had served for 25 years in logistical and supply chain functions, most recently as Chief of Staff of the Naval Supply Systems Headquarters. Jim will continue to contribute to Foresight as Software Editor and commentator on issues of forecast accuracy.
foresight: n.

1: an act or the power of foreseeing: prescience
2: provident care: prudence
   <had the foresight to invest his money wisely>
3: an act of looking forward; also: a view forward

- Merriam-Webster Dictionary

Foresight: The International Journal of Applied Forecasting is a practical guide that relates to business forecasting like no other professional journal can. Four times each year, Foresight’s pages are packed with articles, reviews, and opinions that showcase the best thinking and writing in the field of forecasting.

Every issue features earned expertise from practitioners around the globe that will challenge and inspire you, and that you’ll incorporate in your day-to-day work, whatever types of forecasting you do.

Visit forecasters.org/foresight to learn more and subscribe or renew today.

Foresight.
Do you have it?
Do you want it?
Get it here.

www.forecasters.org/foresight/subscribe
Putting collaboration into practice.
Join your peers, leading S&OP and supply chain practitioners, and noted researchers for 1.5 days of practical professional development. You’ll be immersed in real-world experiences and lessons learned about how you can access the benefits of internal collaboration that S&OP affords, as well as expanding collaborative endeavors with supply chain partners.

What can you expect from the Foresight Practitioner Conference?
Information and experiences that have been tested and proven in the real world. You’ll learn from successful practitioners whose insights come from firsthand experience. You’ll benefit from presentations of practical research from top business school professors, and you’ll find yourself thinking outside the box in an interactive, business school-style simulation.

Practical information you can put to use right away. You’ll return to your organization with a deeper understanding of the methods and practices that will equip you to enable better collaboration internally and across the supply chain. You won’t find vendor presentations or sales pitches in the program.

Opportunities to interact and network. Registration is limited to 100 to enable you to interact with our expert speakers and to ensure that you have maximum opportunities to network with your peers from over 50 companies working through challenges similar to yours.

EARLY REGISTRATION SAVINGS: REGISTER BY JUNE 30TH AND SAVE $200
FORESIGHT SUBSCRIBERS AND IIF MEMBERS SAVE AN ADDITIONAL $100

REGISTRATION AND EVENT DETAILS: FORECASTERS.ORG/FORESIGHT/FPC2014