Improving forecast quality in practice

Robert Fildes and Fotios Petropoulos

Foresight readers will know there’s more to improving forecasting in an organisation than just developing an improved algorithm. Recent issues have examined the Sales and Operational Planning (S&OP) process and the use of different error metrics to evaluate forecasting accuracy: but these are just elements in a bigger picture: how does an organisation go about improving the quality of its forecasting activities? To gain insight into the question, first, we conducted a literature survey focussing on forecast quality improvement, a topic that remains a barren field. One of the authors, RF, had carried out some early work in examining how organizations can improve their forecasting process, conducting a detailed case study of the forecasters in ten divisions of a large multinational corporation (Fildes & Hastings, 1994): the results we’ll return to in a moment. Also, Moon, Mentzer and colleagues have conducted a research program into forecasting practice, with the results provocatively summarised in Moon (2006) and Moon, Mentzer, and Smith (2003). The authors suggest how organisations should audit their forecasting function, a theme taken up also by Finney and Joseph (2011). Both studies point to the importance of first understanding and evaluating the current practice and then examining the gap with best practice. The second element in our study was to undertake a survey of practicing forecasters to ask them as to the dimensions of forecasting quality they thought were particularly important in bridging the gap.

Auditing the Forecasting Function

Before an organization can improve the quality of its forecasting, it needs to understand where it currently stands, how its clients perceive its performance and also to get some notion of how it might change to get closer to the best practice. The first and most fundamental question for any forecaster is who needs their forecasts, for what purpose and how far into the future to look. Above all, what is the value to improved forecasting? With no value to improvements there is really nothing to worry. This is never a straightforward question: most forecasts, even if we focus on demand forecasts alone, fulfil a variety of organizational functions from operations to finance to sales and marketing. Time horizons vary too – from short-term operations to mid-term budgeting and long-term strategic planning. Focussing on improving the one can lead to a disjunction with the others. And if we are to improve accuracy, then we need to evaluate the current forecasting performance as it relates to the value over the decision-relevant time horizon.

The Foresight Journal has contained many discussions on the forecasting accuracy and we won’t revisit them (but see: Boylan & Syntetos, 2006; Hyndman, 2006). Two key conclusions come from the discussions are:

1. The accuracy measures used need to be carefully thought out to link to the decisions faced within a specific time horizon.
2. Standard statistical measures won’t work with intermittent demand and the corresponding inventory decisions.

Accuracy also needs to be measured taking explicit account of the time horizon facing the forecast users. For call centres planning half hourly forecasts are used, for retail operations it could be within day, while for manufacturing with frozen lead times forecast horizons are measured in months. For
the macroeconomist horizons range from a nowcast (predicting the current state) to the next few years ahead, while for the climatologist there is no use in mixing up a weather forecast with decadal forecasts of global warming.

Above all, accuracy needs to be measured. While most organizations claim to measure their forecasting accuracy, many of them fail, particularly when it goes to strategic forecasts, but even when the target is operational. Key mistakes are:

- The failure to focus on the important forecast horizons. As a result, forecasts are averaged, giving a rosy picture of the “achieved” accuracy or unbiasedness.
- Averaging over ‘apples’ and ‘oranges’. For example, errors of an important product are averaged with errors of unimportant products.

To evaluate the current forecasting practice within an organisation, the forecasting accuracy needs to be compared to a suitable benchmark. But what should that benchmark be? Many organisations tend to evaluate and rank their performance based on surveys of industry practice. However, this is not a recommended strategy. Every company faces its own problems. Such surveys are never representative, the sample size is usually small, the answers are not verified, one organisations customers are not like another’s (Kolassa, 2008). Most importantly, the philosophy under which each company operates and acts is unique, and this reflects also on the forecasting function. Instead what has to be done is an off-line comparison using the organisation’s own data and top-quality validated and state-of-the-art software. For examples, see the success stories of the Lancaster Centre for Forecasting. The errors need to be measured relatively to these best practice benchmarks (see also the article by Steve Morlidge in this issue).

So far, in the audit we’ve highlighted:

- Purpose and value of forecast improvement, the users of the forecast, and the forecast horizon.
- The evaluation of current accuracy/performance and the measurement of forecast error.
- Establishing a suitable benchmark.

But that leaves two further questions: how are the forecast produced and with what resources? In particular, what is the information available for use in producing the forecasts? Is the information reliable? Is the utilisation of soft info through judgment valid? Are the assumptions underlying the forecasts (such as the promotional plans or future interest rates) open to challenge?

Finally, an audit must consider the resources that can be mustered: the people, the software and, not least, the raw data. What techniques are used, do they have known flaws and could they potentially be improved on? And does relevant information flow smoothly across the organization and from the outside?

**Perspectives on improving forecasting quality**

From our discussion of the principles behind a forecast audit, the potential problem areas that get in the way of improving quality fall under five headings.

1. Organisational constraints and the flow of information
2. The forecasting software
3. Limited resources
4. The techniques that have been implemented
5. The monitoring and evaluation of the accuracy and value of the forecasting activity
So, how can we bridge the performance gap identified through the organisational audit? In RF’s early study of a multinational company and its forecasting functions (Fildes & Hastings, 1994), managers were asked as to what aspects of their forecasting job would lead to the greatest importance. They identified the priorities presented in Table 1.

**Table 1.** Priorities in improving the forecasting accuracy (Fildes & Hastings, 1994)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Respondents scoring the activity as important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developing consistent data</td>
<td>83%</td>
</tr>
<tr>
<td>Increased software support</td>
<td>70%</td>
</tr>
<tr>
<td>Improved techniques</td>
<td>66%</td>
</tr>
<tr>
<td>Improved data bases</td>
<td>61%</td>
</tr>
<tr>
<td>Improved communication with users</td>
<td>35%</td>
</tr>
</tbody>
</table>

This early study and Moon and colleagues (2003) research provided the evidence on which we based to design a questionnaire for demand planners and forecasters. The primary aim of this survey was to explore perceptions of where quality improvements could best be found.

First we needed to establish how forecasts are typically produced. Previous research had established the importance of judgment. And sure enough, the earlier findings of Fildes & Goodwin (2007) were confirmed.

**Table 2.** How forecasts are typically produced.

<table>
<thead>
<tr>
<th>Fildes&amp;Goodwin 2007</th>
<th>This survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>i) Judgment alone</td>
<td>25%</td>
</tr>
<tr>
<td>ii) Statistical methods exclusively</td>
<td>25%</td>
</tr>
<tr>
<td>iii) An average of a statistical forecast and management judgmental forecast(s)</td>
<td>17%</td>
</tr>
<tr>
<td>iv) A statistical forecast judgmentally adjusted by the company forecaster(s)</td>
<td>34%</td>
</tr>
<tr>
<td>Sample Size</td>
<td>42</td>
</tr>
</tbody>
</table>

Other researchers have delved into the type of statistical methods used with McCarthy et al. (2006) offering the most comprehensive study. The key finding was that judgment is extensively used followed by simple methods such as exponential smoothing. There is no evidence on the use of regression models that can capture causal features of a forecasting problem but are used much less
as Weller and Crone (Weller & Crone, 2012) demonstrate in a recent survey of 200 manufacturing companies. We draw two conclusions:

1. Judgment is a key element of any forecasting audit and improvement program.
2. There is the potential to improve accuracy through the use of better techniques.

But what are the principal targets of forecasters? In our survey, 94% identified accuracy as the most important objective. McCarthy et al. (2006) had a longer list with credibility ranked as equally important. Undoubtedly, the focus of any forecasting activity must be the forecasting accuracy and the credibility that goes with a strong tracking record.

With accuracy in mind, the forecaster’s job is a mixture of activities from data management, statistical analysis, judgement call (in order to take into account special factors) and collaboration with colleagues. We use the problem areas identified earlier to analyse the survey evidence on how forecasters see how quality can be improved.

1. **Organisational constraints and the flow of information**

So, how do the forecasters collect the information they need to make the judgment calls and the respective adjustments? Is this facilitated by the organizational processes they experience? Their primary sources of information used by respondents were internal to the company from sales and, to a lesser extent, marketing. Production/operations and finance were of lesser importance. Internal sources were more used than external ones with about 20% of respondents not using external information sources at all, a finding collaborated by Weller and Crone (2012). In our survey, the availability of internal data still figures at the top of the list of areas for improvement, the same leader as 35 years previously in the initial survey of Fildes & Hastings (1994). The limitations of external data were seen as less important with email being the most common medium of communication. Collaborative software, either internal or external still has limited diffusion. The good news, compared with earlier surveys, is that motivational issues and senior management support were seen as less important constraints on quality improvement.

2. **The forecasting software**

Data base related problems still remain a primary concern, even after all these years of hype as to data warehousing and the more recent hype of ‘big data’. Total reliance on Excel-based solutions was still overly prevalent, a problem identified also in an earlier survey by Sanders & Manrodt (2003). One of the organisations in our study was fully relying on Excel with data being manually copied from sheet to sheet with little validity testing. In another case, forecasters were confronting size constraints imposed by Microsoft Excel. As a result, little of the process had been rendered automatic. The quality of other software products remained an issue for respondents. Commercial software packages have, in most cases, too many limitations, particularly their failure to focus on user-requirements and the uniqueness of each company.

3. **Limited resources**

The lack of a training team was perceived as the key resource problem. The day-to-day workload also tended to interfere with any improvement programme. This suggests that data management and forecasting has not been successfully routinized and that too much time is spend making judgment calls.

4. **The techniques that have been implemented**
Surveys of forecasting software suggest that the statistical techniques embodied in forecasting support systems have serious limitations (Robert Fildes & Goodwin, 2012). Our survey suggests that organisations are all too well aware of these limitations and is of no surprise that judgment is a key element of their forecasting process. However, the process by which judgments are made needs improvement, along with the quality of information incorporated, both internal and external. Forecasters are also well aware that they intervene too often basing their adjustments on inadequate information. Previous studies suggest that there are ways to improve the quality of such interventions (Fildes & Goodwin, 2007).

5. The monitoring and evaluation of the accuracy and value of the forecasting activity. As we have conjectured, measuring accuracy still remains an issue. It is all too rarely done well, if done at all.

The survey has provided a long list where respondents believed improvements could be found. They were not pushed to prioritize. But the top of the list remains the availability of internal data, much like the responses 35 years earlier.

Conclusions

Improving forecast quality has many dimensions. Surveys and case studies provide the evidence that we needed an holistic approach to reach this target, which means, tackling the important areas as an integrated whole. Respondents in this most recent survey tended to see the whole list of 17 potential problems as of some importance. So where do we, the authors, see the major improvements arising based on the evidence?

1. Reappraise the focus of the forecasting activity, in particular, lead times and level of aggregation suitable for the forecast users.
2. Improved and expanded information and an integrated data base.
3. Benchmark existing techniques against ‘best’ practice:
   - developing the associated software
   - avoiding reliance on basic Excel
4. Developing forecasting support systems to encourage effective inclusion of judgment.
5. Effective organisational links so that key pieces of information are shared in a timely fashion.
6. Trained, motivated and better resourced managers:
   - forecasters with too much to do produce worse forecasts!
   - without the training, forecasters may fall back on the routines they are most familiar with.

In short, better data, better software, a wider range of reliable information processed by validated statistical methods, and applied by trained forecasters – now there’s a vision for all to work to!

References


