Improving Forecast Quality

Steve Morlidge
Steve Morlidge

Unilever (1978–2006) roles include:

- Controller, Unilever Foods UK ($1 billion turnover)
- Leader, dynamic performance management change project (part of Unilever’s Finance Academy), 2002–2006

Outside Unilever

- Chairman of the BBRT, 2001–2006
- BBRT Associate, 2007 to present
- Founder/director, Satori Partners Ltd., 2006
- Ph.D., Hull University (Management Cybernetics), 2005
- Visiting Fellow, Cranfield University, 2007
- Editorial Board, *Foresight* magazine, 2010
- Founder, CatchBull (forecasting performance management software), 2011
Six Key Design Principles

- Mastering purpose
- Mastering time
- Mastering models
- Mastering measurement
- Mastering risk
- Mastering process
# Measuring forecast ‘quality’

<table>
<thead>
<tr>
<th>Period Date</th>
<th>Unit A FA %</th>
<th>Unit B FA %</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 Aug 2010</td>
<td>72.8</td>
<td>79.1</td>
</tr>
<tr>
<td>01 Sep 2010</td>
<td>67.7</td>
<td>78.3</td>
</tr>
<tr>
<td>01 Oct 2010</td>
<td>69.0</td>
<td>75.5</td>
</tr>
<tr>
<td>01 Nov 2010</td>
<td>68.4</td>
<td>77.6</td>
</tr>
<tr>
<td>01 Dec 2010</td>
<td>70.5</td>
<td>72.6</td>
</tr>
<tr>
<td>01 Jan 2011</td>
<td>57.0</td>
<td>70.4</td>
</tr>
<tr>
<td>01 Feb 2011</td>
<td>67.1</td>
<td>77.1</td>
</tr>
<tr>
<td>01 Mar 2011</td>
<td>70.6</td>
<td>76.0</td>
</tr>
<tr>
<td>01 Apr 2011</td>
<td>60.5</td>
<td>75.5</td>
</tr>
<tr>
<td>01 May 2011</td>
<td>62.0</td>
<td>76.4</td>
</tr>
<tr>
<td>01 Jun 2011</td>
<td>60.4</td>
<td>71.2</td>
</tr>
<tr>
<td>01 Jul 2011</td>
<td>61.2</td>
<td>66.4</td>
</tr>
<tr>
<td>01 Aug 2011</td>
<td>64.9</td>
<td>71.3</td>
</tr>
<tr>
<td>01 Sep 2011</td>
<td>63.4</td>
<td>68.4</td>
</tr>
</tbody>
</table>

**Unanswered questions**

- Is this good or bad performance?
- How much of this error is avoidable?
- Is Unit B better at forecasting or is it easier to forecast?
- Is performance declining because it's getting more difficult to forecast?
- What is driving this performance?
- Is your forecasting methodology adding value or destroying it?
- Is the application of judgement improving or degrading performance?
- How much are forecast ‘failures’ costing?
- What are the implications for stock and customer service?
- Is this better or worse than the norm?
- What is the scope for improvement?
- What do I do now?
Forecast Quality

• How to measure it
• What we find in practice
Quality: a practical definition

- Better than ‘not forecasting at all’ (higher bound of forecast error)
- As close to minimum avoidable error (lower bound of error) as possible
- At the decision making level (e.g. supply chain = low level stock replenishment point)
- At affordable cost
Why forecast? 101 (for the Supply Chain)

Replenishment based on consumption

Replenishment based on forecast

Since replenishment based on consumption is the same as using the prior periods actual as a forecast, the upper bound of forecast error should be the naive forecast error…

…which also allows for forecastability because items with volatile demand are usually more difficult to forecast

…so to avoid stock out we need safety stock based on the standard deviation of demand

.. leading to less safety stock as it is based on the (lower) standard deviation of error
The lower bound of forecast error is also related to the naïve forecast...expressed as Relative Absolute Error (RAE)

**Range of Naïve Forecast Errors** (including noise)

- Changing Signal: Min RAE = < 0.71
- Flat Signal: Min RAE = $\sqrt{0.5} = 0.71$

**Noise = totally unforecastable**

**Goodwin: Foresight Summer 2013**

**THE ASSUMPTIONS:**
- We have the perfect forecasting algorithm.
- The remaining errors are pure noise in the statistical sense that they are "stationary and independently and identically distributed with a mean of zero".
- The change in the signal from period to period is unaffected by the previous period's noise.

**THE UNAVOIDABILITY RATIO**

Under these assumptions, the ratio of the variance of pure error (that is, error from a perfect forecasting algorithm) to that of the errors from a naïve forecast model will be:

$$
\frac{\text{Variance of pure error}}{\text{Variance of period-to-period changes in signal}} = \text{Variance in signal}^2
$$

If there are no systematic changes in the signal (e.g., no trend or cyclical pattern), the second and third terms in the denominator become zero, leaving us with:

$$
\frac{\text{Variance of noise}}{2 \times \text{Variance of noise}} = 0.5
$$

For the best possible performance, and thus the definition of what constitutes unavoidable error.
New thinking: new measures

Total Error

Forecast Error

Increasing volatility = increasing difficulty of forecasting

Zero Error

How low can you go?

Unavoidable Error

Relative Absolute Error (RAE)

- 1.0 (b/e)  
- 0.7 (good)  
- 0.5 (best)  
- 0.0

Value is added or destroyed at stock holding level (item/location)

Simple replenishment = Naïve Forecast

Theoretical limit of Forecastability (no trend)

Practical limit of Forecastability (any trend)

Destroying Value

Adding Value
## The evidence: research 2013

9 samples from 8 businesses – 330,000 data points

<table>
<thead>
<tr>
<th></th>
<th>Median RAE</th>
<th>Wtd Av RAE</th>
<th>Median MAPE</th>
<th>Forecast Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>0.94</td>
<td>0.89</td>
<td>56%</td>
<td>49%</td>
</tr>
<tr>
<td>Sample 2</td>
<td>1.15</td>
<td>1.04</td>
<td>34%</td>
<td>77%</td>
</tr>
<tr>
<td>Sample 3</td>
<td>0.97</td>
<td>0.81</td>
<td>89%</td>
<td>34%</td>
</tr>
<tr>
<td>Sample 4</td>
<td>1.00</td>
<td>1.53</td>
<td>56%</td>
<td>35%</td>
</tr>
<tr>
<td>Sample 5</td>
<td>0.99</td>
<td>1.14</td>
<td>56%</td>
<td>45%</td>
</tr>
<tr>
<td>Sample 7</td>
<td>1.06</td>
<td>1.89</td>
<td>42%</td>
<td>8%</td>
</tr>
<tr>
<td>Sample 7</td>
<td>0.94</td>
<td>0.99</td>
<td>10%</td>
<td>35%</td>
</tr>
<tr>
<td>Sample 8</td>
<td>1.05</td>
<td>0.87</td>
<td>105%</td>
<td>53%</td>
</tr>
<tr>
<td>Sample 9</td>
<td>1.10</td>
<td>0.99</td>
<td>110%</td>
<td>51%</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>1.02</strong></td>
<td><strong>1.13</strong></td>
<td><strong>62%</strong></td>
<td><strong>43%</strong></td>
</tr>
<tr>
<td><strong>Excl Outliers</strong></td>
<td><strong>0.96</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Traditional measures unhelpful

Very little value added
**Research 2013**

9 samples from 8 businesses – 330,000 data points

<table>
<thead>
<tr>
<th>Sample</th>
<th>Median RAE</th>
<th>Wtd Mean RAE</th>
<th>Excl Outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>0.94</td>
<td>0.89</td>
<td>0%</td>
</tr>
<tr>
<td>Sample 2</td>
<td>1.15</td>
<td>1.04</td>
<td>34%</td>
</tr>
<tr>
<td>Sample 3</td>
<td>0.97</td>
<td>0.81</td>
<td>89%</td>
</tr>
<tr>
<td>Sample 4</td>
<td>1.00</td>
<td>1.53</td>
<td>56%</td>
</tr>
<tr>
<td>Sample 5</td>
<td>0.99</td>
<td>1.14</td>
<td>56%</td>
</tr>
<tr>
<td>Sample 7</td>
<td>1.06</td>
<td>1.89</td>
<td>7%</td>
</tr>
<tr>
<td>Sample 7</td>
<td>0.94</td>
<td>0.99</td>
<td>4%</td>
</tr>
<tr>
<td>Sample 8</td>
<td>1.05</td>
<td>0.87</td>
<td>6%</td>
</tr>
<tr>
<td>Sample 9</td>
<td>1.10</td>
<td>0.99</td>
<td>2%</td>
</tr>
<tr>
<td>Mean</td>
<td>1.02</td>
<td>1.13</td>
<td>5%</td>
</tr>
</tbody>
</table>

**Distribution of RAE**

<table>
<thead>
<tr>
<th></th>
<th>&lt;0.5</th>
<th>0.5-0.7</th>
<th>0.7-1.0</th>
<th>&gt;1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>0%</td>
<td>6%</td>
<td>52%</td>
<td>42%</td>
</tr>
<tr>
<td>Sample 2</td>
<td>4%</td>
<td>5%</td>
<td>33%</td>
<td>62%</td>
</tr>
<tr>
<td>Sample 3</td>
<td>8%</td>
<td>12%</td>
<td>33%</td>
<td>47%</td>
</tr>
<tr>
<td>Sample 4</td>
<td>13%</td>
<td>11%</td>
<td>27%</td>
<td>49%</td>
</tr>
<tr>
<td>Sample 5</td>
<td>1%</td>
<td>9%</td>
<td>42%</td>
<td>48%</td>
</tr>
<tr>
<td>Sample 7</td>
<td>1%</td>
<td>10%</td>
<td>27%</td>
<td>56%</td>
</tr>
<tr>
<td>Sample 7</td>
<td>7%</td>
<td>11%</td>
<td>44%</td>
<td>40%</td>
</tr>
<tr>
<td>Sample 8</td>
<td>4%</td>
<td>2%</td>
<td>35%</td>
<td>57%</td>
</tr>
<tr>
<td>Sample 9</td>
<td>2%</td>
<td>3%</td>
<td>31%</td>
<td>64%</td>
</tr>
<tr>
<td>Mean</td>
<td>5%</td>
<td>8%</td>
<td>36%</td>
<td>52%</td>
</tr>
</tbody>
</table>

What “good” looks like:

Most forecasts are destroying value.

Few forecasts can beat RAE of 0.5…natural limit?

Distribution of RAE:

- Few forecasts can beat RAE of 0.5…natural limit?
- What “good” looks like:
- Most forecasts are destroying value.

Median MAPE Forecast Accuracy:

- 56%
- 49%
- 34%
- 77%
- 89%
- 34%
- 56%
- 35%
- 56%
- 45%
- 42%
- 8%
- 10%
- 35%
- 110%
- 53%
- 110%
- 51%
- 62%
- 43%
- 58%
- 36%

Research 2013
Forecast Quality

...Putting the research to use
Key Concepts

- Identify and cost **avoidable** error:
  ....to provide an objective business assessment of quality and its value

- Separate **two types** of error:
  ....target bias and variation to improve forecast quality

- Translate into **Forecast Value Added**:
  ....the one metric/benchmark for **all** users

- Continuously **track performance** at all levels:
  ....to stimulate speedy corrective action
Key Concepts: Forecast Value Added

Value Added Score (VAS)

-100 to 0 = ‘Unacceptable’

simple replenishment = Naïve Forecast

VAS 0 - 30 = ‘Acceptable’

VAS 30 - 60 = ‘Good’

VAS 60 - 100 = ‘Excellent’

Relative Absolute Error (RAE)

1.0

0.85

0.7

0.5

0.0

100% Error

Zero Error

Forecast Error

Forecast

Error

Bias/Variation

DRILL

Unavoidable Error

Portfolio Analysis

Cost of Avoidable Error

Limit of Forecastability

Destroying Value

Adding Value

Stock Holding Level

1.0

0.85

0.7

0.5

0.0

Stock Holding Level
Improving Forecast Quality

In practice
I. Issue management: eliminating bias

Traditional Metrics focus on high level performance...

...but cost and customer service are driven by the quality of low level forecasts

Continuous Improvement

Guided by system alarms, drill down to identify the source of forecasting issues

Tackle issues at the lowest level to improve high level metrics

<table>
<thead>
<tr>
<th>Level</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0</td>
<td>All Products (1)</td>
</tr>
<tr>
<td>Level 1</td>
<td>Brands (70)</td>
</tr>
<tr>
<td>Level 2</td>
<td>Minor Brands (150)</td>
</tr>
<tr>
<td>Level 3</td>
<td>Products (1500)</td>
</tr>
</tbody>
</table>
Issue Management Example

FMCG: <£1b revenue

High level bias (red): c0%, but low level (grey): c 15%...

...generating over and under forecasting bias alarms

Bias Impact Overview

Impact of Alarm Classes: 73,890
Impact of Other Alarms: 0
Other Impact: 85,753
Total Impact: 159,643
2. Improving Methods: where to use judgement

- **High volume, high variability.**
  - Forecasting involves judgement. Difficult but possible to excel. Aim for blue.

- **Low volume, high variability.**
  - Use simplest/cheapest methods – aim for amber or low red.

- **High volume, low variability.**
  - Optimise forecast algorithm and restrict judgemental input.
  - Should be easy to beat naive. Aim for green...

---

**ABC/XYZ Analysis**

- **Horses**
- **Mad Bulls**
- **Mules**
- **Jack Rabbits**

**Notes:**

- VAS<0
- VAS 0-30
- VAS 30-60
- VAS >60

---

**CatchBull**

MAKE FORECASTING PAY
2. Improvement Example

FMCG: >£1b revenue

Value Added Score mostly red (value destruction)

51% of SKU’s have negative VAS

High Bias

High Variation
## What is this worth?

<table>
<thead>
<tr>
<th>Ready Reckoner</th>
<th>Cost of Sales</th>
<th>Per €1b revenue*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost of Error</td>
<td>4%-8%</td>
<td>€20-40m</td>
</tr>
<tr>
<td>Forecast Value Added</td>
<td>0%-2%</td>
<td>€0-10m</td>
</tr>
<tr>
<td>Avoidable Error</td>
<td>2%-4%</td>
<td>€10-20m</td>
</tr>
</tbody>
</table>

*Assuming 50% Gross Margin*
Key points

1. Measurement is key
2. Need to account for forecastability
3. Measures should be actionable
4. Improvement is from a) tuning and b) choosing models
5. Forecasts add value by beating the naïve forecast
6. The first step is to stop destroying value: ‘easy’
7. Differentiate between bias and variation: the impact of interventions and of model choice
8. Drill from high to low level to tune forecasts
9. Differentiate to help choose modelling approach
In summary
Contact details

Thank you

steve.morlidge@catchbull.com
www.catchbull.com