Forecasting and Revenue Management: what’s there to learn?

Professor Robert Fildes
Director, Lancaster University Centre for Forecasting
Lancaster Centre for Forecasting

• Objectives:
  – Conduct applied research & consultancy projects with companies;
  – Facilitate knowledge-transfer between academia and business;
  – Establish best practices in forecasting.

• Services:
  – Consultancy;
  – Forecasting courses;
  – Master’s student projects;
  – Doctoral research projects.

• For research into forecasting to support revenue management including hierarchical forecasting, intermittence and OD choice methods contact:
  r.fildes@lancaster.ac.uk

Directors Robert Fildes and John Boylan
and
Drs Sven Crone, Nikos Kourentzes, Nicos Pavlidis and
Ivan Svetunkov
together with 8 doctoral students
Agenda

• Forecasting challenges
  – New method
  – Gaining from using hierarchies
  – Model choice

• Forecast Value Added

• Current issues in ‘forecasting for revenue management’
  – Your views?

• Retail promotional optimization
  – Choice models in action

Finally

• Improving the quality of forecasting
Criticality of Forecasting to revenue management

**Reservation/sales system**

- Transaction data

**Optimization module**
- Prices
- Overbooking limits
- Markdowns
- Promotions
- Decision rules

**Forecasting module**
- Bookings
- No-shows
- Cancellations
- Price sensitivity
- Competitor prices

**Decision Support System**

- User interventions

**Market data**

**Product data**

**Transaction data**

**Decision Support System**
Forecasting challenges

- Micro-data: flight x class x OD
- Multiple hierarchies
- Double/ triple-seasonality:
  - time of day, day, month, holiday
- Trend?
- Intermittent
- Short histories within season
- Missing observations
- Elasticity estimation
- N<p: many variables
Current Issues in Forecasting

• New methods
  – Computer intensive (eg. Neural nets, fuzzy methods etc), econometrics, forecasting quantiles

• Aggregation and pooling
  – Hierarchy
  – Cross-sectional & Over time

• Model choice
  – Measuring accuracy
  – Aggregate vs individual selection
    • one ‘best method’ or matching a method to a data history

• Judgmental interventions in model based forecasts

• Extending the econometric choice models
  – Between flights, operators: Large p > no. observations
New methods

• Computer intensive methods
  – Neural nets, singular value decomposition, k-means

• Aggregation
  – Product hierarchy
  – Cross sectional & Across time

• Econometric (choice) modelling
Challenges of Forecasting

Reality is highly nonlinear, especially when considering multiple variables. Neural networks (and other advanced nonlinear causal models) are a way forward.

- Nonlinear effects of Promotions & Prices and Weather
- ... with nonlinear interactions
Challenges of Forecasting

Neural Network as ARX(p) setup
\[ \hat{y}_{t+h} = f(x_t, x_{t-1}, \ldots, x_{t-n}) + \epsilon_{t+h} \]

Alternatives
- Exponential smoothing developments
- ARIMA
- Multivariate

Competition
- Data sets
- Methodology
- Error measures

New methods
### k-Nearest Neighbour for Time Series Regression

**Matching:**
- Past time series patterns
- Patterns in similar competitive routes

**Ensemble forecast of the k most “similar” days**
**Problem:**
- Many features affecting future demand
- New service, little historical data
- Many past service launches

**Aim:**
- Match new route (and available data) to past analogous services
- Use the realised demand from analogies to predict

**k-Nearest Neighbour for time series prediction**

kNN for Time Series (Regression):
- Measure distance between current and past similar situation

Euclidean Distance between two time series $Q = \{q_1, q_2, ..., q_n\}$ and $S = \{s_1, s_2, ..., s_n\}$

$$D(Q, S) = \sqrt{\sum_{i=1}^{n} (q_i - s_i)^2}$$

Find similar shapes in time series
k-Nearest Neighbour for time series prediction

New methods – analogy

**kNN for Time Series (Regression):**

- Measure distance between current and past similar situation

<table>
<thead>
<tr>
<th>Database</th>
<th>Distance</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.98</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.21</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.43</td>
<td>3</td>
</tr>
</tbody>
</table>
Hierarchical Forecasting

Hierarchies may refer to:
- Product types
- Geographical allocation
- Channels
- ...

Problem: forecasts are different at each aggregation level!

Main approaches for reconciling hierarchical forecasts:
- **Top-down approach**: Forecast at the highest level and disaggregate using historical proportions
- **Bottom-up approach**: Forecast at the lowest level and aggregate the forecasts up to the required level
- **Middle-out approach**
- **Optimal approach**: optimally combines forecasts from each level

Top-down and Bottom Up forecasting for a two level system

| TABLE 12.4 | TOP-DOWN AND BOTTOM-UP FORECASTS FOR A TWO-LEVEL SYSTEM |
| --- |
| **TOP DOWN** ↓  |
| Top | Total \((T_t)\) |
| Level 1 | \(P_{jt} \cdot T_t, j = 1, \ldots, L_1,\) where \(P_{jt}\) is the proportion of total sales in the \(j\)th group |
| Level 2 | \(P_{kjt} \cdot (P_{jt} \cdot T_t), k = 1, \ldots, L_2,\) where \(P_{kjt}\) is the proportion that the \(k\)th SKU contributes to the \(j\)th group’s sales |
| \(T_t = \sum_{j=1}^{L_1} S_{jt}\) | Sales of the \(j\)th group \(S_{jt} = \sum_{k=1}^{L_2} S_{kjt}\) |
| \(S_{kjt} = \) sales of the \(k\)th SKU in the \(j\)th group |
| **BOTTOM UP** ↑  |


And:

- Middle out
- Optimal (consistent across levels)

See: Straightening Out Your Forecasting Systems! Agifors 1999, Chatterjee, Summerbell
Temporal non-overlapping aggregation transforms the original data from the higher observed frequency (e.g. months) to lower frequencies (e.g. quarters or years).

- This strategy **highlights or attenuates different series characteristics** on each level of aggregation:
  
  - At **lower aggregation** (high frequency time series) periodic components, such as **seasonality** will be prominent.
  
  - At **higher aggregation** levels high frequency signals are filtered and more importance is given to the lower frequency components, such as the **level** and **trend** of a time series and common explanatory variables.
Traditionally we model time series at the frequency that we sampled them or take decisions. However, a time series can be viewed in many different ways, adapting the notion of product hierarchies to **temporal hierarchies**:

- Time of day
- Day
- Week
- Month
- Season

![Temporal Aggregation Diagram](image)
What if we do not select a particular time aggregation level? → but use multiple

Issues:
- Different model
- Different length
- Combination
Forecast combination:

• Forecast combination is widely considered as beneficial for forecast accuracy
• Simple combination methods (average, median) considered robust, relatively accurate to more complex methods

Result:

• Combining over different time horizons (usually) produces more accurate forecasts
• Consistent over time horizons
Aggregation and forecasting

- It is not new, but new approaches yet to be widely implemented

- Supply chain applications & revenue management → beneficial to slow and fast moving items forecast accuracy (like everything... not always!):
  - but see Strauss et al. 2016 for a revenue management application to hire cars: 5-10% improvement over baseline constrained demand

- Cross-sectional hierarchical forecasting well established as effective
  - But no clear result as to which method of aggregation to use.

Message:
Compare and test
Model choice

With a large number of time series
  – Flights, hotels, skus
  – Various time series methods
    • E.g. exponential smoothing, NN

Three choices
• Apply ‘best’ method to all series in aggregate
• Select the ‘best’ method for each individual series
• Combine

Research conclusions:
• Select usually, combine sometimes
• ‘Tune’ parameters to problem
Forecast Value Added and the Forecasting Support System

Surveys show:
- Most common type of forecasting

Activity System & Variables

Data

Previous Forecast & Error

Method based forecast

Yes?

Judgemental adjustment

No?

Final Forecast

Judge Interventions
Organisationally based Forecasting combines statistical analysis with managerial judgement

- humans are adaptable and can take into account one-off events, but they are inconsistent and suffer from cognitive biases
- statistical methods are rigid, but consistent, and can take into account large volumes of information

\[ SY_t = \alpha Y_{t-1} + (1 - \alpha) SY_{t-1} + \lambda X_t + \]
Why not rely on the models?

The Problems
- Too complex
- Incomplete data on many drivers
- ‘Unique’ events
- No available statistical expertise
- Management understanding & acceptance
- Belief that managerial expertise is best

In airlines
- New routes, new slots
- Changed competition

In retailing
- Promotions
- Events & holidays

But does the adjustment add value?
- US Air thought so
- In manufacturing and retail mixed results
Current issues in Forecasting for Revenue Management
(Weatherford, J. Revenue and Pricing Management, 2016)

• Unconstraining (retail, transport, hotels)
  – EM and PD: are there better approaches?
    • Worth a lot! >1%?

• Time series (extrapolative) methods
  – seasonality

• Aggregation

• Judgmental adjustment

• Choice models

• Volume of routes

• **Link between forecasting models and revenue optimisation**

• Industry specific issues
  – Retail: promotional intensity, range of competitive variables
Retail demand forecasting and promotion optimisation

Why cross-item information is important?

- One product's promotion can influence the sales of another
- Substitutive in the same product category
- Both complementary and substitutive in cross category
- Much existing research has provided strong empirical evidence as to potential importance
  - (Gupta, 1988; Chiang, 1991; Walters, 1991; Chintagunta, 1993; Bucklin et al., 1998; etc.)

Theoretical findings applicable in a real forecasting system?

- High dimensionality!
- No research has empirically considered the promotional interactive effects in a grocery forecasting system that can work in practice
Empirical models of promotional demand
- Profit consequences

Choice models
• Dynamics
• Within brand
• Across brand
• Across category

Too complex?

Optimisation applied based on retail decision rules. e.g. price points, no. of simultaneous promotions etc.

Results:
• Need to include competitive & category information

• Pruning to simplify ⇒ costly

Intra-category cross SKU promotional effects, worth extra 3 percent of the total profit.
Final thoughts
Improving the Quality of Forecasting

• Specify forecasting problem
  – Hierarchy, Level of Aggregation & Forecast horizon
  – Available information

• Current accuracy
  – Compared to base line method on your data
  – Appropriate measures?
  – Value-added analysis of judgment?

• Software choices
  – Benchmarked statistical methods
  – Tuned software
  – Method selection

• Implementation and Improvement Issues
  – Management of Forecast Function
Conclusions

• Forecasting drives the effectiveness of revenue management optimisation

• New methods available
  – With limited historical data (new routes: new products)
  – Aggregation, pooling, hierarchy

• Too little testing done
  – No clear results from forecasting competition for airlines!
    (some reports in Agifors presentations)

• Judgment seen as beneficial but limited ‘Value added analysis’

• Importance of modelling intra and interbrand effects when optimizing
  – Pruning to simplify problem leads to losses
Take-Aways

• Benchmark your current forecasting performance against best practice
• Tune your software to your current data
• Individual vs aggregate model selection?
• Measure your value added
  – Develop data base to capture adjustments
• Develop choice based econometric models
  – Include inter and intrabrand competition
• Manage a forecasting improvement program
Questions?