Supply Chain Forecasting for Operations

Report on 19th IIF Workshop

Lancaster University, 28 & 29 June 2016

Venue

This two-day event attracted 30 delegates and was held at Lancaster University Management School.

All presentations were held in a lecture theatre that was adjacent to the breakout space shown in the above picture.

Conference Proceedings

The full programme and the slides from the presenters may be found at:

http://www.lancaster.ac.uk/lums/scfo/proceedings

Presentations on Day 1

The first presentation, by **John Boylan**, from Lancaster University, gave an overview of the current state of the art by adopting three different perspectives. The first was from the viewpoint of Operational Research more generally. While forecasting and simulation have often worked hand-in-hand, there are opportunities for closer linkages between forecasting and optimisation models. The second perspective was both longitudinal (in time and across the supply chain) and hierarchical (by product, location and customers). Opportunities for further research in hierarchical forecasting were highlighted. The third perspective was the dissemination of research into practice. John argued for a closer nexus between academics, software developers and practitioners, in order to speed up the adoption of new approaches which have proved their worth.

In the next presentation concerning the field as a whole, **Thanos Goltsos,** from Cardiff University, highlighted that although the forecasting and inventory literatures are vast, with over 100,000 publications between them, the amount of papers addressing both topics is more modest, with fewer than 1,000 publications. Thanos outlined a classification scheme for the degree of integration between forecasting and inventory control: i) demand is considered deterministic or assumed to be known; ii) need to forecast demand is merely mentioned; iii) forecasting is an input to inventory control; iv) forecasting and inventory control are fully integrated. Classification was performed by looking at forecasting methods and metrics, inventory rules, methodology (analytical/simulation) and data (empirical /synthetic). The papers reviewed were found to pay insufficient attention to empirical data, to be over-reliant on Mean Square Error and to rarely progress to the stage of full integration.

The next session focussed on temporal aggregation. The first presentation was by **Matt Weller**, from Lancaster University, who is investigating the relative merits of forecasting based on weekly and monthly data. Matt noted that there were very few studies on supply chain forecasting using weekly data. His own study examined 1800 SKUs in 20 product categories using item-level Point Of Sale data and promotional variables. The study considered the temporal aggregation and disaggregation of the weekly and monthly forecasts, evaluating at both frequencies. The findings were that more complex multivariate methods (such as stepwise regression and ARIMAX) perform better than simpler univariate methods (such as Exponential Smoothing and ARIMA) when evaluating weekly errors but this result is not replicated for monthly forecasting when simpler methods are better. However, weekly ARIMAX forecasts aggregated to monthly buckets overtakes simpler methods.

The next presentation was from **Bahman Rostami-Tabar**, from Coventry University. The focus of his recent work has been on non-overlapping temporal aggregation when demand follows AR(1) or ARMA(1,1) processes. He has shown that when Single Exponential Smoothing (SES) is used to forecast demand, the forecasting accuracy benefit of using non-overlapping temporal aggregation depends on the autoregressive parameter of the demand process, the aggregation level, and the smoothing constant used for SES. In more recent research, not yet published, Bahman has gone on to consider the effect of non-overlapping temporal aggregation for MA(1) and AR(1) demand processes when an optimal (MMSE) forecasting method is used. His research shows that when the MMSE forecasting method is used, the temporal aggregation method is out-performed by the non-aggregation method.

Our next session moved on to the theme of supply risk mitigation. The first presentation was by **Sridhar Seshadri**, from the Indian School of Business. In this work, market-based forecasts were augmented by the inclusion of an additional variable, namely lagged return on an aggregate financial market index. This model was motivated by the permanent income hypothesis in economics, which says that the amount of consumer spending and the mix of spending on discretionary and necessity items depend on the returns achieved on equity portfolios held by consumers. This new market-based forecast achieved an average 15% reduction in Mean Absolute Percentage Error compared with forecasts by equity analysts. These ideas were then applied to a supply chain network. This resulted in the identification of the following mechanisms that can affect the correlation between sales and the state of the economy: propagation of systematic risk into production decisions, aggregation of orders from multiple customers in a network, and aggregation of orders over time. The latter two factors intensify correlation and result in the amplification of correlation upstream in supply networks.

The next presentation by **Qinyun Li** (Cardiff University) investigated the impact of demand information sharing and production smoothing on order forecast accuracy. In the model developed, each player uses a proportional order-up-to policy with (unbiased) exponential smoothing forecasts to calculate replenishment quantities. The researchers have derived analytical expressions for the frequency response of order forecast errors. When a supply chain only implements a demand information sharing strategy, the accuracy of order forecasts will be improved for high frequency noises, mitigating the bullwhip effect. When a supply chain applies only a production smoothing strategy, production smoothing is found to be more effective than demand information sharing in improving order forecast accuracy. When both strategies are adopted, and production smoothing is present, demand information sharing does not necessarily further improve order forecast accuracy.

Erica Pastore, from the Politecnino di Torino, continued the theme of the bullwhip effect and its mitigation. Her work focussed on an empirical investigation of the bullwhip effect on 46000 SKUs form a three-echelon European automotive spare parts supply chain (dealers, local warehouses, central warehouse). This research was motivated by asking whether the literature results concerning a monotonic demand variability amplification through the echelons is observed in practice. Preliminary analyses confirmed the presence of the bullwhip effect when considering demand aggregated across all SKUs. Demand at the top level was found to be more than twice as variable as demand at the bottom level (using the measure of the Coefficient of Variation). The effect was the strongest from dealers to local warehouses, probably because of the incentives structure of the company. Also, the bullwhip effect was stronger for fast moving items than slow moving items.

The next section focussed on cross-sectional aggregation in supply chain forecasting. **Michele Trovero**, from the software company, SAS, spoke about judgemental overrides to model forecasts (a common practice in supply forecasting and demand planning). Numerous model based forecasts may be used to form an aggregated forecast that is not required to follow a fixed hierarchy. After aggregation, judgmental forecasts can be applied to override the statistical forecasts. These forecasts need to be disaggregated to the lowest level of aggregation. Michele outlined an approach, recently developed at SAS, that allows for large-scale automatic forecast aggregation and bounded constrained optimized judgmental forecast disaggregation. This meets the user needs of allowing the over-rides of multiple series at a time, and not being restricted by pre-determined hierarchies. Work is currently ongoing on the enhancement of forecast tracking and feedback to users.

The next presentation, from **Christina Phillips**, from Bangor University, covered a project to align production with demand in a manufacturing facility experiencing a high degree of uncertainty across the value chain. Christina began by calling for a more 'human-centred analytics' which combines analytics, behavioural operations and action research. The researchers ran a participative simulation experiment to try different family groupings of products with different cycle schedules. From a cross-sectional perspective, aggregate demand is relatively flat but disaggregate demand may be lumpy because of batching and customer rationing, leading to bullwhip problems. Temporal aggregation may increase this effect, in part due to information loss. Current work is seeking to identify the best aggregation level forecast that can be trialled during the participative simulation and verified through use.

The final presentation on the first day of the workshop was by **Rogelio Oliva**, Texas A&M University. This research was motivated by the observation that the specialisation in supply chain planning (eg by marketing, operations) is notorious for generating conflicts. Semantic knowledge boundaries reflect interpretational differences across groups with respect to knowledge and meaning. These differences arise from knowledge that is not explicit but context-specific. Pragmatic knowledge boundaries acknowledge that each functional group has accumulated its knowledge base over time, so the groups have knowledge that is localised, embedded and invested in practice. The researchers found that the adoption of a "business assumption package", based on multiple sources of information, was beneficial in reducing conflict caused by different assumptions. Consensus forecasting, with an independent group for managing the process, helped, and constructive engagement can have a direct positive impact on the process.

Presentations on Day 2

The second day started with a session on forecast evaluation. **Laura Turrini** (Kűhne Logistics University) This research is motivated by the question of how to measure the goodness-of-fit of a statistical distribution of demand, particularly in difficult cases such as intermittent demand. The benchmark method in this study is the Kolmogorov-Smirnov (K-S) test. The major problem with this method, from a supply chain planning perspective, is that it does not put sufficient emphasis on the tail of the distribution. Laura proposed the application of a modified K-S test, where the distance function (between theoretical and actual cumulative distributions) is inflated in the right tail and deflated in the left tail. Examination of empirical datasets showed the modified K-S test to have the best inventory performance for service level up to 94%, with a modified Anderson-Darling test beinbg the best for higher service levels.

The next paper was a joint presentation by **Paul Wang** and **Fotios Petropoulos**, from Cardiff University. They noted that the M and M3 competitions were limited to forecast accuracy and did not consider any other measures of business performance. The researchers focussed on industrial monthly data from the competitions, as most of these series represent demand, sales or consumption of different products. They sought to extend the competition results by undertaking an evaluation based on inventory performance, examining costs, amplification of orders and service levels. They compared the Naïve, Single Exponential Smoothing, Holt's, Damped Holt's and Holt-Winters methods (all used in the original competitions) and also included the ETS package, Auto ARIMA. MAPA and a dynamic optimised Theta model. They simulated an order-up-to inventory system, with service levels of 90%, 95% and 99%, and lead-times up to 12 periods. The results showed Holt-Winters performing poorly, newly proposed methods (eg MAPA) performing well, and simple combinations also doing well.

The session after coffee, on "Advanced Forecasting Methods", started with a presentation by **Sha Zhu**, of Erasmus University, Rotterdam. This research looked at the lumpiness in spare part demand being triggered by lumpiness in component repairs and the uncertainty of individual components generating spare part demand. The researchers investigated a periodic review inventory model with lost sales, minimising total inventory costs. With the information of component arrivals provided by the maintenance plan, spare part demand is binomially distributed and it is possible to obtain the order policy from the inventory model. A case study, based on real data, showed that cost reductions of around 8% may be achieved.

The next presentation was given by **Adriana Martins** from ENSIAT, University of Lille. This research concerned 'fast fashion', a business model that has been adopted by large international retail fashion chains. The shelf-life may be very short (few weeks) and so historical sales data are very limited. The authors have proposed a Clustering Based Sales Forecasting (CSBF) model that takes into account the descriptive characteristics of the products to organise the data into clusters of similarity before computing the forecasts. Data from the past two years have been used as a training set and one year as a test set. Evaluations of the new model are encouraging, showing that it has the potential to be implemented in the real world of fast fashion retailers.

The session on Advance Forecasting Methods concluded with a talk by **Sven Crone** (Lancaster University) on an application of Artificial Neural Networks to supply chain forecasting for the consumer packaged goods manufacturer, Beiersdorf AG. The presentation focussed on a methodology to automatically specify Neural Networks for large industry assortments. The efficacy of the approach was determined on real-life industry time series from multiple countries, in comparison with established statistical methods of Exponential Smoothing and Seasonal Linear Regression. The Neural

Network solution is implemented as an add-in to the software package SAP APO DP, running forecasts for over 50,000 products across 56 countries for Beiersdorf on a monthly basis.

The next session, on Forecast Uncertainty, started with a presentation by **Dennis Prak**, University of Groningen. This paper addressed estimation uncertainty in inventory models. The approach is based on a cost equation that depends on a demand distribution and its unknown parameters. The traditional approach is to substitute their point estimates into the cost equation and minimise the resulting expression by choosing the inventory decision accordingly. An alternative approach is to use not only the point estimates but also a random variable that models the difference between the point estimate and the true parameter. This new approach is demonstrated on a discrete-time, continuous review model with linear holding and shortage costs. Numerical results indicate cases were there can be a cost-benefit of 20% to 30% for mean-stationary demands.

Zied Babai (Kedge Business School, France) presented recent work on variance forecasting by aggregation. The results are based on a first-order autoregressive moving average, ARMA(1,1) demand model and a stochastic lead-time. Analytical expressions for the variance of forecast error over lead-time have been derived for three strategies: 1. Calculating the variance of the aggregated forecast error over lead-time; 2. Aggregating the variances of the per-period forecast error; 3. Considering the forecasts of the demand aggregated time series. The outperformance conditions are insensitive to the type of lead-time distribution. For high positive autocorrelation, both strategies 3 and 2 lead always to better forecasting performance than strategy 1.

The final session, on structural models, began with a presentation by **Juan Trapero,** from the Universidad Castilla-la-Mancha, Spain. He pointed out that demand forecasting error is typically assumed to have zero mean and constant variance. However, this Considering the forecasts of the demand aggregated time series itais not always true in practice, leading to lower service levels than anticipated. The aim of the current research is to explore: i) non-parametric approaches such as kernel density estimation; ii) parametric estimators based on GARCH and exponential smoothing models; iii) optimal combination between uncertainty measures by maximizing the conditional coverage test. Results are showing that robust performance may be achieved by the use of these approaches.

The next paper, given by **Ivan Svetunkov** (Lancaster University) focussed on intermittent state-space models for demand forecasting. A simple intermittent state-space model is proposed, based on a Single Source of Error (SSOE) approach to exponential smoothing. It is shown that the continuous demand SSOE state-space model is a special case of the proposed model, and it underlies a wide variety of processes. It gives a statistical rationale for several popular intermittent demand forecasting methods. The talk concluded with demonstrations of the application of the model and discussion of the evaluation of the methods in terms of appropriate error metrics and inventory performance.

The final talk, by **Giacomo Sbrana**, concerned solving the Ricatti equation for structural time series models. In a structural state-space model, the error covariance matrix for model time-invariant series converge to the Ricatti equation, whose solution greatly simplifies the Kalman Filter. This presentation showed the derived solutions of the Ricatti equation for several structural time series models. The results are valid for models with correlated noise terms, encompassing Single Source and Multiple Source of Error models. They provide a simple, efficient and flexible algorithm for the estimation, prediction and smoothing of these models, avoiding the use of specialised software. The talk closed with an application to retail sales forecasting,

Special Part Issue of the International Journal of Forecasting

Papers were submitted in late 2016 for consideration for a part special issue of the IJF (guest editors John Boylan and Robert Fildes). These papers are currently going through peer review.

Social Events

An informal get together, including a meal and drinks was held at The Borough pub in central Lancaster on the evening of Sunday 27 June for delegates arriving early.

The conference dinner was held on the evening of Monday 28 June at the Midland Hotel, Morecambe.

The dinner was held in the Sun Terrace at the rear of the Midland Hotel. The terrace lived up to its name, with fair weather, glorious views over Morecambe Bay and a thoroughly enjoyable meal, followed by a summer evening's stroll along the pier.

Workshop Evaluation and Feedback

A feedback questionnaire was issued to all delegates. Eighteen questionnaires were returned. Each question had a four point scale (1- highest, 4 –lowest). None of the questions received feedback scores of 3 or 4 from any of the delegates, with all questions receiving an 'excellent' score of 1 from at least two-thirds of those who responded. The detailed results are given below:

Conference organisation

1. Registration management
2. Accuracy and timeliness of information before the Workshop's start

3 Meals
4 Number and duration of breaks
5 Assistance on-site
6 Overall duration of the Workshop
7 Value for money
Conference programme and content
8 My expectations about the Workshop were fulfilled
9 The time allocated for speeches etc was adequate
10 The contributions aroused my interest
11 The contributions were coherent with the topics of the Workshop
12 The contributions aroused were scientifically relevant
13 The feedback I received was very useful for my activity
14 The Workshop helped me to expand my network
15 I'm looking forward to attending similar events

Score	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
1	18	15	12	15	17	15	15	16	17	16	17	17	14	14	17
2	0	3	6	3	0	3	2	2	1	2	1	1	4	4	1
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Summary of Delegates Scores (1= Excellent, 4 = Poor)

The question with the lowest score was Q3, which related to the food during the workshop. One comment received on food specifically related to the lunches. We shall pay closer attention to this aspect of provision when organising future workshops.

The qualitative comments were very positive, with delegates pleased with the time allocated to each talk (35 minutes including 10 minutes for comments from the discussant and for questions from the audience). The respondents thought that discussant approach had worked well. The smooth running of the workshop was also commented upon in several emails after the conference. Those attending were pleased that the nexus between forecasting and inventory management received attention during the workshop, with some delegates asking for even greater emphasis on that theme.

Some delegates asked that this should become an annual or two-yearly event. Our experience of running the IIF Workshop would encourage us to run such an event again, either on Supply chain Forecasting or on another topic, such as Marketing Analytics.

John Boylan
Chair, Organising Committee for 19th IIF Workshop