DEMAND FORECASTING OF OVER-PROMOTED FMCG PRODUCTS IN A MANUFACTURING COMPANY

PATAK Michal[1], PECINOVA Zuzana[1], BRANSKA Lenka[1]

[1]*University of Pardubice, Pardubice, Czech Republic, EU*

Abstract

Overuse of trade sales promotion to keep fast-moving consumer goods (FMCG) in the range of retail stores results in a lot of negative impacts on all members of the supply chain network. One of the consequences is also an extreme increase in demand variability for FMCG manufacturers. However, such demand becomes unpredictable if only common forecasting methods are applied. This paper aims to find ways of forecasting the demand that is affected by frequent implementation of promotional events. Based on the case study conducted with a large Czech manufacturer of FMCG products, the paper first discusses the possibilities and barriers of the current theoretical approaches to demand forecasting of promoted products, which subsequently results in a proposal of a statistical forecasting method for over-promoted products. The proposed approach to demand forecasting combines a multiple linear regression (MLR) model with an autoregressive integrated moving average (ARIMA) model. By its application in the company involved in the research, they were able to decrease the simple statistical forecast error by 24%.

Keywords: Autoregressive integrated moving average, demand forecasting, fast-moving consumer goods, multiple linear regression, promotion

1. INTRODUCTION

A high degree of consumer price sensitivity in the Czech FMCG markets and a strong negotiation position of retail chains force manufacturers to make use of promotional events (particularly price discounts) on a frequent basis. As a result of this, most FMCG products are nowadays sold in the consumer markets during a promotional event. The resulting decrease in the sales per unit of goods is compensated by manufacturers and retailers for by an increase in the product sales price in the current period, as a result of which Czech consumers are only more and more encouraged to seek promotional events purposely in their effort to buy as many goods in sale as possible. In the current conditions, it is hard to get out of such a vicious circle, and so FMCG manufacturers have no choice but to adapt to these changes.

Excessive implementation of promotional events has a significant impact in the form of extreme growth of product demand variability. Over-promoted demand is hard to forecast using the common methods, which results in ancillary expenses and losses relating to underestimation (stock-outs) or overestimation (short shelf-life) of expected future sales. This paper aims to draw attention to the problems connected with application of the current theoretical approaches to demand forecasting during promotions and to propose a demand forecasting method for over-promoted products in manufacturing companies.

The paper is organized as follows. In Section 2, we review the literature dealing with forecasting models for demand during promotions. In Section 3, we discuss the possibilities and barriers of the theoretical approaches to promotional forecasting in the case study conducted with a large Czech manufacturer of dairy products. Subsequently, we propose and evaluate a forecasting method for over-promoted products. In Section 4, we present the conclusions and discuss several options for future research.

2. DEMAND FORECASTING IN THE PRESENCE OF PROMOTIONS

Demand forecasting at Stock Keeping Unit (SKU) level is usually based on the combination of a univariate statistical forecasting method and managerial judgment made by forecasters in the organization. The former provides an automatic baseline statistical forecast, while the latter is responsible for manual adjustment of the baseline in order to include market forces, promotions and other influences [1]. Apart from these traditional forecasting approaches, in the recent literature there is an effort propose such statistical methods that would be able to make reliable demand forecasts even during promotions without requiring adjustment by forecasters. Such solutions can holistically model trends, seasonality, sales promotions, prices, and other related factors that influence the demand using predictive analytics [2].

2.1. Traditional approach to promotional demand forecasting

The traditional approach to forecasting, in the practice known as the base-times-lift approach [3], is a two-step procedure which initially generates a baseline forecast with simple time series models and then makes adjustments for any incoming promotional events as follows (1):

$$final\ forecast_t = \begin{cases} M_t, & if\ no\ promotion \\ M_t + adjustment, & otherwise \end{cases} \tag{1}$$

where M_t is the baseline forecast for time *t* (usually generated by a simple exponential smoothing model). The adjustment is estimated based on the lift effect of the most recent promotion and/or the judgments made by forecasters [4]. Promotional and advertising activities are the main reason for judgmental adjustment of statistical forecasts in practice. Unfortunately, such adjustments are costly and they represent a significant source of systematic errors [5].

An alternative approach to making adjustments is causal forecasting [6]. It involves building multivariate statistical models with inputs that correspond to the promotional features - price discounts, type of advertising, etc. [1]. In a simpler case, it could be a linear regression model with promotional variables [7, 8]. Sophisticated causal forecasting could involve particular hypotheses about customer response to promotional activities, and the data is used to estimate its parameters [9]. Machine learning techniques (such as regression trees or artificial neural networks), on the other hand, do not assume a particular relationship between the variables and involve a search through the functional form space as well as parameter estimation [10]. Moreover, Poh et al. [11] found that a neural network is able to capture nonlinear relationships in a causal model without being explicit about the structure, aiding in the short-term forecasting of a variable interest.

As a result, the related sales promotional data needed to be stripped (cleansed) of the demand history, as well as the demand history adjusted for shortages and outliers. This manual process is time consuming and impractical. According to Chase [12], demand planners still spend over 80% of their time cleansing data and information across the organization, rather than using the data and information to improve forecast accuracy.

2.2. Holistic approach to promotional demand forecasting

Statistical forecasting methods can be divided into time series models and causal models. Time series models are based on univariate analysis that analyzes historical sales in order to extract a demand pattern that is projected into the future. Causal models rely on cause-effect relationships between sales and demand influencing factors. Holistic approaches to promotional demand forecasting combine time series and causal models. Therefore, they are often called hybrid models in the literature [13].

One of the possibilities is application of artificial neural networks (ANN) or a multiple linear regression (MLR) with inputs including not only external variables (promotional features), but also time variables. An example of successful application of MLR models is forecasting daily sales in retail stores in the Chicago area [3, 14]. The main advantage of MLR models is a transparent and simple regression style model structure (2) according to the literature [15]:

$$\ln(Y_t) = intercept + \sum_{j=1}^{L} \alpha_j \ln(Y_{t-j}) + \sum_{i=1}^{N} \sum_{j=0}^{L} \beta_{i,j} \ln(X_{i,t-j}) + \varepsilon_t \qquad (2)$$

where Y_t is the demand in time t, L is the order of lags, N is the number of external variables $X_1, \dots, X_N$.

A hybrid model that is most frequently mentioned in the literature is an autoregressive integrated moving average with external variables (ARIMAX) or its seasonal modification (SARIMAX). For example, Bratina and Fanagel [16] used an ARIMAX model to forecast daily beer demand in the Slovenian market. Lee and Hamzah [17] developed a SARIMAX model to forecast monthly sales of clothes in Indonesia. Their model had better forecast results than decomposition method, SARIMA or ANN models. A (S)ARIMAX model can be written in the general form (3) according to the literature [18]:

$$Y_t = b_0 + b_1 X_{1,t} + b_2 X_{2,t} + \dots + b_k X_{k,t} + N_t \qquad (3)$$

where Y_t is the demand in time t modelled as a function of k external variables $X_{1,t}, \dots, X_{k,t}$, and N_t is the residual (autocorrelated) series that is modelled as a (S)ARIMA process.

In the case of promotional forecasting, the effect of a change in an explanatory variable does not show up in the forecasted demand instantaneously, but it is distributed across several time periods. And so the model should include not only explanatory variable X_t, but also previous (lagged) values of the explanatory variable (X_{t-1}, $X_{t-2}, \dots, X_{t-k}$). Such a model has made demand forecasting significantly more accurate, e.g. in a manufacturing company specializing in household products [1].

In addition to hybrid (S)ARIMAX models, some of the past studies also proposed hybrid (S)ARIMA-ANN models. One of the earliest hybrid ARIMA-ANN models was applied to forecast monthly sales of consumer goods for a Danish company [19]. Aburto and Weber [20] proposed another hybrid SARIMA-ANN model to forecast daily demand at a Chilean supermarket.

Finally, Arunraj and Ahrens [13] compared seasonal naïve, SARIMA, SARIMA-ANN, and SARIMAX models used to forecast daily sales of bananas at a retail store in Germany. They found that SARIMAX models have better forecast accuracy than the other models due to their integration of seasonal adjustment using SARIMA and consideration of external variables using MLR. Although a SARIMA-ANN model considers the same factors, its forecast accuracy was lower than that of SARIMAX models due to some of important disadvantages of ANN, such as over-fitting and a complex design of a suitable neural network structure.

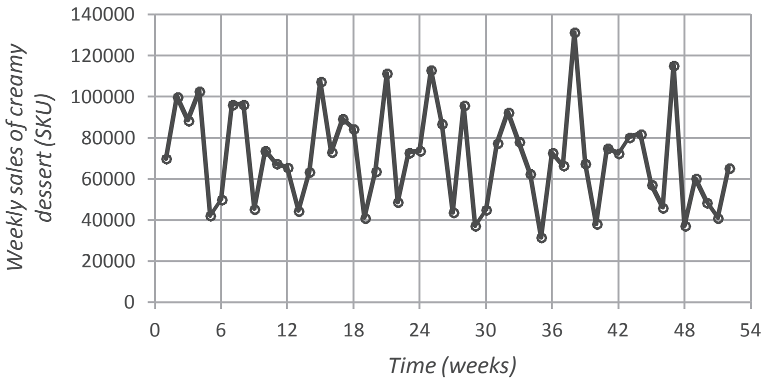
3. CASE STUDY: DEMAND FORECASTING OF OVER-PROMOTED DAIRY PRODUCTS

Problems related to application of theoretical approaches to demand forecasting within excessive use of promotions will be demonstrated through a case study conducted at a large Czech manufacturer of dairy products.

3.1. Description of data

This study uses weekly sales data of dairy product (creamy dessert) measured in SKU from January 2013 to December 2014. The creamy dessert was selected because of its short shelf-life (14 days) and its importance for the company (one of the key products, measured by annual turnover). The chosen temporal aggregation of sales (weekly sales) is directly directed to utilization of the given forecasts as the basic information input for short-term supply chain planning. **Figure 1** illustrates the weekly sales data for 2014 in a time series plot and other statistics pointing to frequent use of promotional events.

This time series is highly variable due to the presence of promotions. Almost two thirds of sales were made during promotions (price discounts). During a promotional event, which lasts a few days (usually a week), a customer buys a significantly larger quantity of product than usual. Due to the short shelf-life, which makes long-term accumulation of the products at the customers' warehouses impossible, promotions only result in short-term deviations of demand (cannibalization effects of own and competing substitute products). A strong negotiation position of large retail chains forces manufacturers to provide irregularly repeating promotional events, which must not overlap in time with events at competing chains. As a result of this, weekly sales are grossly distorted by promotions practically in each observation of the time series.



Creamy dessert 2014 sales statistics:
Sum of sales = 3,694,195 SKUs
Common sales = 1,286,889 SKUs (34.8%)
Sales during promotions = 2,407,306 SKUs (65.2%)
Number of customers = 1,869
Number of orders = 30,529
Mean of weekly sales = 71,042 SKUs
St. deviation of weekly sales = 23,591 SKUs

Figure 1 Weekly sales of the creamy dessert in 2014

3.2. Proposal of a forecasting method

The basic problem of the studied time series is the impossibility of stripping it of the influence of promotional events on the SKU level. One of the possible solutions consists in cleansing time series on a lower level of aggregation, e.g. Stock Keeping Unit by Customer (SKUC) with unambiguous identification of sale of the given item to the given customer. However, the mentioned method would be disproportionately laborious (1869 types of SKUCs) and, moreover, sales time series already show the character of intermittent demand on such a low level of aggregation. In conclusion, the character of the time series makes it impossible to apply traditional approaches to promotional demand forecasting.

A solution to demand forecasting in our case can be found in holistic approaches. The core of the problem is choice of suitable explanatory variables, which would be directly related to the volume of weekly sales on the SKU level. Suitable variables, which are at the same time available in the corporate database, include the above discounts provided within a promotion, the number of implemented events, or the number of customers who were provided with a discount in the given week. However, there is a question whether there is a real dependence between weekly sales and considered variables. Unfortunately, further time series analysis proved that within each time series observation, there were several concurrent events, where different customers were provided with different discounts. Similarly, there were situations where different customers reacted to a comparable discount differently. For this reason, even holistic approaches to sales forecasting on the SKU level would obviously show very low accuracy.

Our solution consists in segmentation of customers into 24 segments and subsequent disaggregation of the sales time series on the SKU level to 24 partial sales time series on the Stock Keeping Unit by Segment (SKUS) level. The segmentation was based on a qualitative judgement of the sale manager aiming to group customers with similar reaction to promotions in the way to make the total volumes of sales in individual segments comparable. Therefore, some segments comprise one customer only (large retail chains), while other segments combine customers using common distribution channels or serving common end consumers. On the SKUS level, it is then much more senseful to forecast demand, as the chosen explanatory variables are already directly related to the volume of sales in the segment. Moreover, the relatively low number of created segments eliminated the problem with intermittent demand, which is typical for lower levels of sales aggregation.

Time series forecast on the SKUS level was made using the hybrid ARIMA-MLR model, which included the influence of two external variables (N=2) and their k-lagged values: factor X_1 (percentage of customers in the segment who were provided with a discount) and factor X_2 (average percentage discount in the segment). The residual series N_t was modelled as an ARIMA process. The forecasting model can be expressed as follows (4):

$$Y_t = \beta_0 + \sum_{i=1}^{N} \sum_{j=0}^{k} \beta_{i,j} X_{i,t-j} + N_t \tag{4}$$

The final forecast for the sales time series on the SKU level can then be obtained through a simple sum of forecasts in *M*-segments using the following equation (5):

$$Y_t(SKU) = \sum_{i=1}^{M} Y_t(SKUS_i) \tag{5}$$

The parameters of the forecasting models were optimized on 2013 weekly sales in the environment of the software of SPSS Statistics on the basis of minimization of the Bayesian Information Criterion. Three types of forecasting models were tested to assess the importance of external variables: Model A (only factor X_1 was considered), Model B (only factor X_2 was considered), and Model C (factors X_1 and X_2 were considered simultaneously).

3.3. Forecast value added analysis

Cross-validation is primarily a way of measuring the forecasting performance of a model. Our way to cross-validate a model is to test it on a set of data not used in fitting (sales during 2014). This study uses the root mean square error (RMSE) as the performance measure. We also use the forecast value added (FVA) analysis for evaluation of forecasting models. The FVA principle is based on using a benchmark model (e.g. a naïve model or an average model) to scale the forecasting performance objective. The FVA analysis outcomes are shown in **Table 1**.

Table 1 FVA analysis

Forecasting model	RMSE (SKUs)	FVA over Average method (SKUs)	Comparative improvement (% points)
Average method	23,497	-	-
Model A (X_1)	17,896	5,601	23.8%
Model B (X_2)	21,628	1,869	8.0%
Model C (X_1, X_2)	17,867	5,631	24.0%

The FVA analysis implies that all the proposed methods can decrease the error of the Average method. As factor X_1 (% of customers) explains the demand variability much better than factor X_2 (discount), Model A and Model C show a comparable accuracy (RMSE error decreased by 24%).

The relatively high residual variability, which we were not able to explain by the proposed models, can be caused by several facts. Some customers were provided by a discount at the turn of two weeks. The increase in the sales during the promotion then showed in two observations of the time series, but in different ratios, which could not be detected by any logical rules. The customer segmentation also resulted in generation of the segment of “Others”, which included diverse customers with random sales, which themselves showed high variability without relation to a company’s promotional activity. Finally, a forecasting model also cannot cover demand fluctuation caused by supply chain effects (bullwhip effect), as well as the situation where the customer does not always react to the same discount in the same way.

4. CONCLUSION

Current approaches to demand forecasting during promotions have a number of barriers for utilization in manufacturing companies, particularly in the cases of excessive implementation of promotional events. Based on the case study conducted with a large Czech manufacturer of dairy products, a forecasting method for over-promoted products was proposed. It is a three-step procedure which initially segments customers according to their reaction to promotional events. Subsequently, a forecast is made separately for each segment using the hybrid ARIMA-MLR model. In the end, the final forecast on the SKU level is calculated as the sum of partial forecasts in individual segments.

Although the proposed method helped to increase the accuracy a benchmark model significantly (RMSE error decreased by as much as 24%), some factors that could additionally increase the model accuracy were not considered in our study. Therefore, future research in this area should be particularly focussed on seeking next (potentially better) explanatory variables. At the same time, it would be interesting to verify how the level of sales aggregation (or the way of customer segmentation) and the type of the used hybrid model affect accuracy of final forecasts.

REFERENCES

[1] TRAPERO, J. R., PEDREGAL, D. J., FILDES, R., KOURENTZES, N. Analysis of judgmental adjustments in the presence of promotions. International Journal of Forecasting, 2013, vol. 29, no. 2, pp. 234-243.

[2] CHASE, C. W. Does demand history really need to be cleaned? Journal of Business Forecasting, 2016, vol. 35, no. 3, pp. 17-32.

[3] HUANG, T., FILDES, R., SOOPRAMANIEN, D. The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. European Journal of Operational Research, 2014, vol. 237, no. 2, pp. 738-748.

[4] FILDES, R., GOODWIN, P., LAWRENCE, M., NIKOLOPOULOS, K. Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. International Journal of Forecasting, 2009, vol. 25, no. 1, pp. 3-23.

[5] FRANSES, P. H., LEGERSTEE, R. Do experts’ adjustments on model-based SKU-level forecasts improve forecast quality? Journal of Forecasting, 2010, vol. 29, no. 3, pp. 331-340.

[6] FILDES, R., NIKOPOULOS, K., CRONE, S., SYNTETOS, A. A. Forecasting and operational research: a review. Journal of the Operational Research Society, 2008, vol. 59, no. 9, pp. 1150-1172.

[7] COOPER, L. G., BARON, P., LEVY, W., SWISHER, M., GOGOS, P. PromocastTM: a new forecasting method for promotion planning. Marketing Science, 1999, vol. 18, no. 3, pp. 301-316.

[8] VAN DONSELAAR, K. H., PETERS, J., DE JONG, A., BROEKMEULEN, R. A. C. M. Analysis and forecasting of demand during promotion for perishable items. International Journal of Production Economics, 2016, vol. 172, no. 1, pp. 65-75.

[9] FADER, P. S., HARDIE, B. G. S. Modeling consumer choice among SKUs. Journal of Marketing Research, 1996, vol. 33, no. 4, pp. 442-452.

[10] ALI, Ö. G., SAYIN, S., VAN WOENSEL, T., FRANSOO, J. SKU Demand forecasting in the presence of promotions, Expert Systems with Applications, 2009, vol. 36, no. 10, pp. 12340-12348.

[11] POH, H-L., YAO, J., JAŠIC, T. Neural networks for the analysis and forecasting of advertising and promotion impact. International Journal of Intelligent Systems in Accounting, Finance and Management, 1998, vol. 7, no. 4, pp. 253-268.

[12] CHASE, C. W. Cleanse your historical shipment data? Why? Journal of Business Forecasting, 2015, vol. 34, no. 2, pp. 29-33.

[13] ARUNRAJ, N. S., AHRENS, D. A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting. International Journal of Production Economics, 2015, vol. 170, no. 1, pp. 321-335.

[14] MA, S., FILDES, R., HUANG, T. Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra- and inter- category promotional information. European Journal of Operational Research, 2016, vol. 249, no. 1, pp. 245-257.

[15] FADER, P. S., HARDIE, B. G. S. The value of simple models in new product forecasting and customer-base analysis. Applied stochastic models in business and industry, 2005, vol. 21, no. 4-5, pp. 461-473.

[16] BRATINA, D., FANAGEL, A. Forecasting the primary demand for a beer brand using time series analysis. Organizacija, 2008, vol. 41, no. 3, pp. 116-124.

[17] LEE, M., HAMZAH, N. Calendar variation model based on ARIMAX for forecasting sales data with Ramadhan effect. In *Proceedings of the Regional Conference on Statistical Sciences 2010.* Kota Bharu, Malaysia, 2010. pp. 349-361.

[18] MAKRIDAKIS, S., WHEELWRIGHT, S. C., HYNDMAN, R. J. *Forecasting: methods and applications. 3rd ed.* New York: John Wiley & Sons, 1998. 642 p.

[19] LUXHOJ, J. T., RIIS, J. O., STENSBALLE, B. A hybrid econometric-neural network modeling approach for sales forecasting. International Journal of Production Economics, 1996, vol. 43, no. 2-3, pp. 175-192.

[20] ABURTO, L., WEBER, R. Improved supply chain management based on hybrid demand forecasts. Applied Soft Computing, 2007, vol. 7, no. 1, pp. 136-144.