Inventory Control in Sales Periods

Tamás Szántai, Edith Kovács, Attila Egri

Department of Differential Equations, Budapest University of Technology and
Economics, Műegyetem rkp. 3-9, 1111 Budapest, Hungary
e-mail: szantai@math.bme.hu, kovacsea@math.bme.hu, egri@math.bme.hu

Abstract: Sales promotion aims to capture the market and increase sales volume. Therefore, an important task is the forecasting of the demand during the sales period. We present two dynamic methodologies for calculating the quantity which has to be placed on the shelves at the beginning of each day such that we keep some constraints expressing lower and upper bounds on the quantities. Both methodologies are new to this field and are useful because of some specific properties of the problem. Our new methods use historical data of the demands in previous promotions and the consumptions registered in the previous days. Since the promotion period is relatively short, other methods such as time series analysis can hardly be used.

Keywords: inventory control; dynamic forecasting; information driven forecasting

1   Introduction

Many businesses use sales promotions to increase the demand of a product or service. Promotions and sales are important strategies of a successful business. Their effects include growth within the market segment involved, the discovery of new products. Promotions attract new and old customers and can keep the company relevant when competitors appear. Price reductions can substantially boost the sales of the given product, but also cause brand switching.

Effective sales promotions lead to inventory reductions, because customers buy more products. Therefore companies use these actions at the end of a buying season. For example, when Christmas Eve is past, very often, retailers offer discounts to make room on the shelves for other products.

Paper [12] highlights how promotions affect the buying habits of costumers as a consequence of a changed price conditions.

Some interesting statistics on demand in sales period can be found in [1]: "Demand during many promotions is often dramatically greater than median daily demand: demand in 54% of promotions is > 15 standard deviations greater than

daily median demand, and demand in 3% of promotions is > 100 standard deviations greater than normal daily demand. However, promotion demand represents a relatively small percentage of total yearly demand for most products. For 90% of products, promotion demand is <15% of total yearly demand and promotion demand is <20% of total demand among 90% of the products with one or more promotions."

The time series method forecasts the new demand values, on the basis of historical demand data. In [11] time series forecasting models with extending an exponential smoothing approach were proposed. However, exponential smoothing methods have been criticized for their inability to capture the effects of special events such as promotions, announcements. When demand for an item is being driven by such factors as trends and seasonal patterns, time series methods tend to work quite well [6]. However, business data often contain responses to actions, such as promotions, that cannot be captured as part of the level, trend and seasonal components. When a significant amount of demand is being driven by these types of events, time series methods will not work very well.

Fildes and Goodwin ([4]) indicated promotional and advertising activity as one of the main drivers behind adjustments of the statistical forecasts by managerial judgments.

An alternative approach to the problem of forecasting promotional sales is to use regression models, which use past promotional information to formulate causal models, Fildes et al. [5].

Although the information of human judgment cannot be captured by simple promotional models, yet Trapero et al. in [12] showed that a simple model could beat judgmental forecasting. Therefore, there is a need for developing more sophisticated promotional models.

In a recent paper [3], different models of forecasting the demand during a promotion are developed and tested, including a moving average forecast and several regression models. In the paper it is investigated how different factors such as price variation, advertising influence the demand.

Another recent paper on the topic of forecasting demand in sales period is [7]. The presented method consists of the identification of potentially influential categories, and then of the selection of the explanatory variables by using multistage LASSO regression and of the use of a rolling scheme to generate forecasts. The success of the method is also based on dealing with high dimensionality which brings improvements in forecasting accuracy compared to other methods which used also a reduced variable space.

In [13] Trapero et al. proposed a Principal Component Analysis based promotional model that overcomes the limitations caused by multicollinearity and high dimensionality.

In this paper we address the problem of stochastic inventory control during a retail or promotion time. The pricing of the products in sales period is also an important optimization problem, but in this paper we suppose the promotion price was already fixed. This paper is dealing with the problem of daily updating the quantity of a given product on the shelves during a promotion sales period. A specific characteristic of this problem is that, products are sold at a lower price during a relatively short period only. We present two methodologies for making decisions on the quantity of product which has to be placed on the shelves. These are based on historical data of similar promotions that have occurred in the past. To our best knowledge we are the first who introduce the following models to the problem of inventory management during promotional sales.

The dynamic of our models is as follows. At the beginning of each day, a quantity of a given product is placed on the shelves. The demand on each day is observed and based on this cumulative set of information one has to decide the quantity to be placed on the shelves at the beginning of the next day. This way the decision is made day by day and uses beyond the information accumulated from the previous days also historical data collected from previous sales periods. In addition the experts may put some constraints on the quantity of products being on the shelves.

2 Dynamic Inventory Control in Sales Periods by Adapting the Lake Balaton Water Level Regulation Model

2.1 Preliminaries

In paper [8] the following dynamic control model was developed for regulation of the water level of Lake Balaton.

Let us introduce the following notations:

V_0 initial water content of the lake,

ξ_k random water input in month k,

z_k water quantity to be released through the channel Sio in month k,

$\varsigma_k = V_0 + \sum_{i=1}^{k} \xi_i$ initial water content plus the cumulated monthly random water inputs at the end of month k,

a_k lower bound for the water quantity being in the lake at the end of month k,

b_k upper bound for the water quantity being in the lake at the end of month k.

For determining the optimal decisions z_1^*, K, z_N^* according to the first N months one should solve the following stochastic programming problem:

$$\max P\left\{ a_k \leq \varsigma_k - \sum_{i=1}^{k} z_i \leq b_k, k = 1, K, N \right\}$$

supposing that   (1)

$$0 \leq z_k \leq K, \quad k = 1, K, N,$$

where K is the monthly capacity of the channel Sio.

The authors of paper [8] proposed to accept the optimal value $z_1 = z_1^*$ of the first decision variable only, apply it as water release in the actual month and then formulate the next stochastic programming problem of type (1) and so on.

If one observed the realized values x_1, K, x_n of the random water inputs ξ_1, K, ξ_n and the realized water releases were z_1^*, K, z_n^* in the first n months, then the knowledge of these values can also be utilized in the following way. Let us modify the initial water content of the lake for the water content at the end of the nth month, i.e. let be $V_n = V_0 + \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} z_i^*$ and $\varsigma_k = V_n + \sum_{i=n+1}^{k} \xi_i$. Then instead of stochastic programming problem (1) one can regard the problem

$$\max P\left\{ a_k \leq \varsigma_k - \sum_{i=n+1}^{k} z_i \leq b_k, k = n+1, K, n+N \Big| \xi_1 = x_1, K, \xi_n = x_n \right\}$$

supposing that   (2)

$$0 \leq z_k \leq K, \quad k = n+1, K, n+N.$$

Now one may accept the optimal value of the first decision variable $z_{n+1} = z_{n+1}^*$ only, apply it as water release in the actual month and then formulate the next stochastic programming problem of type (2) and so on.

If the random process ξ_1, ξ_2, K is Gaussian, these stochastic programming problems can be solved as it was shown in [8]. In this paper the special case of $n = N = 2$ was taken and the authors successfully applied this method for the monthly dynamic control of the water level of the Lake Balaton for a fifty years long time horizon.

2.2 Adaptation of the Lake Balaton Water Level Regulation Model

Let us now regard a line of goods in a department store which is on sale for a fourteen days' time period. The main difference is that while the water level of the Lake Balaton can be controlled only by decreasing its value, in this case the amount of the line on the shelves can be controlled only by increasing its value. In the same time while the water level of the Lake Balaton increases (changes) randomly and it can be decreased deterministically, in this case the amount of the line on the shelves decreases randomly and it can be increased deterministically.

For describing the stochastic programming models let us now introduce the following notations:

V_0 — the starting amount of the line on the shelves at the beginning of the sale,

ξ_k — the random consumptions of the line on the kth day of the sale,

z_k — decision variable belonging to the kth day of the sale, this is the quantity of the line to be placed on the shelves when opening the kth day of the sale,

$\varsigma_k = \sum_{i=1}^{k} \xi_i$ — the cumulated daily random consumption at the end of the kth day of the sale,

K — the capacity of the shelves over the sale,

a_k — lower bound for the line amounts to be placed on the shelves at the end of the kth day of the sale,

b_k — upper bound for the line amounts to be placed on the shelves at the end of the kth day of the sale.

The notations above are introduced for all days $k = 1,2,K,14$ of the sale.

If we suppose that the line of goods is put on the shelves each day morning then the following inequalities must be fulfilled:

$$V_0 + z_1 \le b_1 \quad \text{quantity at first day opening time,}$$

$$V_0 + z_1 + z_2 - \xi_1 \le b_2 \quad \text{quantity at second day opening time,}$$

$$\mathbb{M} \qquad \mathbb{M} \qquad \mathbb{M}$$

$$V_0 + z_1 + \Lambda + z_N - \xi_1 - \Lambda - \xi_{N-1} \le b_N \quad \text{quantity at } N^{th} \text{ day opening time.}$$

Taking into account the daily random consumption values at the end of the day the following inequalities must be fulfilled:

$$V_0 + z_1 - \xi_1 \qquad\qquad\qquad \geq a_1 \qquad \text{quantity at first day closing time,}$$

$$V_0 + z_1 + z_2 - \xi_1 - \xi_2 \qquad\qquad \geq a_2 \qquad \text{quantity at second day closing time,}$$

$$\text{N} \qquad\qquad\qquad\qquad\qquad \text{N} \qquad \text{N}$$

$$V_0 + z_1 + \Lambda + z_N - \xi_1 - \Lambda - \xi_N \qquad \geq a_N \qquad \text{quantity at } N^{\,th} \text{ day closing time.}$$

For determining the optimal decisions z_1^*, K, z_N^* according to the first $N(\leq 14)$ days one should solve the following stochastic programming problem:

The problem which accords with/corresponds to problem (1) is formally the following:

$$\max P \left\{ \begin{array}{c} V_0 + \sum_{i=2}^{k+1} z_i - b_{k+1} \leq \varsigma_k \leq V_0 + \sum_{i=1}^{k} z_i - a_k, k = 1, \mathrm{K}, N-1 \\[2mm] \varsigma_N \leq V_0 + \sum_{i=1}^{N} z_i - a_N \end{array} \right\}$$

supposing that $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (3)

$$z_1 \leq b_1 - V_0, \; z_1 \geq 0, z_2 \geq 0, \mathrm{K}, z_N \geq 0.$$

If one observed the realized values x_1, K, x_n of the random consumptions ξ_1, K, ξ_n and the quantities z_1^*, K, z_n^* of the line placed on the shelves on the first n days of the sale, then the knowledge of these values can also be utilized in the following way. Let us modify the starting amount of the line on the shelves in the morning of the $n+1^{\,th}$ day, i.e. let be $V_n = V_0 + \sum_{i=1}^{n} z_i^* - \sum_{i=1}^{n} x_i$. Then instead of stochastic programming problem (3) one can regard the problem

$$\max P \left\{ \begin{array}{c} V_n + \sum_{i=n+2}^{n+k+1} z_i - b_{n+k+1} \leq \varsigma_{n+k} \leq V_n + \sum_{i=n+1}^{n+k} z_i - a_{n+k}, k = 1, \mathrm{K}, N-1 \\[2mm] \varsigma_{n+N} \leq V_n + \sum_{i=n+1}^{n+N} z_i - a_{n+N} \end{array} \right| \left. \begin{array}{c} \xi_1 = x_1 \\ \mathrm{M} \\ \xi_n = x_n \end{array} \right\}$$

supposing that $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (4)

$$z_{n+1} \leq b_{n+1} - V_n, \; z_{n+1} \geq 0, z_{n+2} \geq 0, \mathrm{K}, z_{n+N} \geq 0.$$

Now one may accept the optimal value of the first decision variable $z_{n+1} = z_{n+1}^*$ only, apply it as quantity of the line to be placed on the shelves in the actual day and then formulate the next stochastic programming problem of type (4) and so on.

If the random process ξ_1, ξ_2, K is Gaussian, these stochastic programming problems can be solved. For detailed calculation procedure, see paper [8]. Relatively small values of n and N (say, $n = N = 2$) may be enough for achieving good control in this case, too.

2.3   Application of the Algorithm

As it can be seen in Table 1, the random consumptions had relatively large standard deviations according to their mean values, so the modified Lake Balaton inventory control model was not applicable for these data. This model could be applied when the standard deviation of the random consumptions is not larger than one third of the mean value, otherwise one should be able to interpret negative valued consumptions.

3   Dynamic Inventory Control in Sales Periods by using Information-driven Forecasting

3.1   Preliminaries

Sales periods are relatively short, one or two weeks typically, therefore, the popular time series forecasting methods cannot be applied for the goods in promotion sales.

We distinguish the following two kinds of promotional sales. The first one is the case of a product which already exists on the market, the second one is the sale promotion applied to a new product which has to be introduced into the market.

For the second case we have no proper historical data. To overcome this drawback, we can search for products which are similar to the new one, and use their historical data. Having these we may apply our methodology.

We consider now the case of forecasting the demand of an existing product on the market, for which we have earlier data registered, during the sale periods of the same length. We associate a random variable X_i to the daily consumption registered at the end of each day. We can define a random vector $\mathbf{X} = (X_1, K, X_d)$, where d is the length of the sales period expressed in days.

At the end of the ith day, $i = 1, K, d-1$ we have to decide on the quantity of product to place on the shelves. For this we have to forecast the consumption of day $i+1$ based on the consumption of the first i days. Based on the forecasted

consumption we make sequential decisions on the quantity which have to be displayed.

We regard now the problem of forecasting the consumption of the $i+1^{th}$ day.

We consider the random vector: $\mathbf{X} = (X_1, \text{K}, X_{i+1})$ which is a margin of the random vector $\mathbf{X} = (X_1, \text{K}, X_d)$.

The main idea behind our method is that we use 1, 2, 3 (rarely 4) out of the previous days to forecast the consumption of day $i+1$. We emphasize here that we do not use necessarily the days i, i-1, i-2. Instead we will choose those days from all previous days which minimize the uncertainty of the day i+1.

For this task, we use the following informational theoretical concepts.

The uncertainty amount of a random vector can be quantified by its entropy. The entropy does not depend on the values of the random vector; it depends only on the probabilities with which the different values are taken on.

The concept of entropy has its roots back in 1854 in a memoir of Rudolph Clausius. However, in this paper we will use the expression given by Claude Shannon published in his famous paper [10]. More general definitions for entropy were also given by Rényi [9].

We introduce the reader into some information theoretical concepts, which have to be reminded for the understanding of our method. The interested reader can find more details about these concepts in [2].

In the present work, we use the following formula for entropy, which is related to a random vector with m realizations. If χ_i represents the range of the variable X_i then the range of $\mathbf{X}$ is a subset of $\bigotimes_{i=1}^{d} \chi_i$

$$H(\mathbf{X}) = -\sum_{k=1}^{m} p_k \ln p_k$$

where m indicates the number of all distinct realizations of the random vector $\mathbf{X}$ and p_k denotes the probability of the kth realization of the random vector $\mathbf{X}$ (the ordering of the realization has no importance, but is fixed).

In order to quantify how much the uncertainty of a given random variable is reduced by knowing the values taken on by the other random variables we use the concept of conditional entropy denoted by $H(X_i \mid \mathbf{X}_{V-i})$, where we use the notation $\mathbf{X}_{V-i}$ for the random vector of all random variables with indices in V except X_i.

For a better understanding of the concept of conditional entropy we first define the following random variables. Let us fix $i \in V$ and an arbitrary realization $\mathbf{x}_{V-i}^k$ of $\mathbf{X}_{V-i}$. The conditional random variable denoted by $X_i \mid \mathbf{x}_{V-i}^k$ takes on the values x_{i_j} by probabilities $p_{x_{i_j} \mid \mathbf{x}_{V-i}^k} = P\left(X_i = x_{i_j} \mid \mathbf{X}_{-i} = \mathbf{x}_{V-i}^k\right), \quad j = 1, \mathrm{K}, i_s$:

$$X_i \mid \mathbf{x}_{V-i}^k : \begin{pmatrix} x_{i_1} & \mathrm{K} & x_{i_j} & \mathrm{K} & x_{i_s} \\ p_{x_{i_1} \mid \mathbf{x}_{V-i}^k} & \mathrm{K} & p_{x_{i_j} \mid \mathbf{x}_{V-i}^k} & \Lambda & p_{x_{i_s} \mid \mathbf{x}_{V-i}^k} \end{pmatrix}.$$

Here $\mathbf{x}_{V-i}^k$ stands for the kth realization of the random vector $\mathbf{X}_{V-i}$. This way a conditional random variable $X_i \mid \mathbf{x}_{V-i}^k$ is assigned to each realization $\mathbf{x}_{V-i}^k$, $k = 1, \mathrm{K}, m_{V-i}$. We take now their entropies denoted by $h\left(X_i \mid \mathbf{x}_{V-i}^k\right)$, $k = 1, \mathrm{K}, m_{V-i}$ and define a new random variable as follows.

$$\tilde{h}\left(X_i \mid \mathbf{X}_{V-i}\right) : \begin{pmatrix} h\left(X_i \mid \mathbf{x}_{V-i}^1\right) & \mathrm{K} & h\left(X_i \mid \mathbf{x}_{V-i}^k\right) & \mathrm{K} & h\left(X_i \mid \mathbf{x}_{V-i}^{m_{V-1}}\right) \\ p_{\mathbf{x}_{V-i}^1} & \Lambda & p_{\mathbf{x}_{V-i}^k} & \Lambda & p_{\mathbf{x}_{V-i}^{m_{V-1}}} \end{pmatrix},$$

where $p_{\mathbf{x}_{V-i}^k} = P\left(\mathbf{X}_{-i} = \mathbf{x}_{V-i}^k\right), \quad k = 1, \mathrm{K}, m_{V-i}$.

Finally, we arrive to the definition of conditional entropy $H\left(X_i \mid \mathbf{X}_{V-i}\right)$ that is defined as the expected value of $\tilde{h}\left(X_i \mid \mathbf{X}_{V-i}\right)$.

From these it can be seen, that the conditional entropy $H\left(X_i \mid \mathbf{X}_{V-i}\right)$ quantifies the amount of the uncertainty of X_i when there are given the realizations of X_{V-i}. As the smaller the conditional entropy is the better we can reduce the uncertainty of X_i by knowing realizations of $\mathbf{X}_{V-i}$.

This leads to the introduction of the concept of mutual information $I\left(X_i, \mathbf{X}_{V-i}\right)$, which is defined as the following difference:

$$I\left(X_i, \mathbf{X}_{V-i}\right) = H\left(X_i\right) - H\left(X_i \mid \mathbf{X}_{V-i}\right).$$

3.2   The Central Idea of our Method

The main idea of our method is the way we decide on the quantity which has to be placed on shelves next day. The decision is based on the consumptions registered in few previous days. These days are chosen in such a way that these minimize the conditional entropy of the next day's consumption. This is equivalent with maximizing the information gain.

Since the sales period is relatively short, one or two weeks long, we use only one, two or three previous days in forecasting.

From a theoretical point of view taking more than three days leads to over fitting and poor generalization of the model.

3.3   The Consumption Forecasting Algorithm

Based on the historical data we have the joint empirical probability distribution of the random vector $\mathbf{X} = (X_1, \mathrm{K}, X_d)$, where the random variables $X_i, i = 1, \mathrm{K}, d$ represent the daily consumptions.

We introduce the following notations.

The realized consumption of the ith day in the actual sales period is denoted by C_i. Forecasting the consumption of the $i+1^{th}$ day means choosing one of the possible realizations of the random variable X_{i+1} in a certain way. Let $Q_{i+1}^k \quad k = 1, \mathrm{K}, s_{i+1}$ denote the possible values of X_{i+1}. The forecasted consumption of the $i+1^{th}$ day is denoted by Q_{i+1}.

We introduce the following notations:

$$\Lambda_k = \{X_1, \mathrm{K}, X_k\},$$

$$\Lambda_k^2 = \{(X_l, X_m) \mid X_l, X_m \in \Lambda_k, l < m\}$$

$$\Lambda_k^3 = \{(X_l, X_m, X_n) \mid X_l, X_m, X_n \in \Lambda_k, l < m < n\}$$

- In the first day we usually consider the mean value of X_1 as forecasted consumption. This will be the quantity Q_1.

- In the second day we can use the registered consumption C_1 of the first day and forecast the second day consumption by the following maximization:

$$Q_2 = \underset{k=1,\mathrm{K},s_2}{\arg\max} \, P(X_2 = Q_2^k \mid X_1 = C_1)$$

- From $i=3$ to n

  Step 1: Choose a number f^* from 1 to 3, this is the number of the previous days used in the forecast – this decision can be made by interacting with the user.

  Step2. The selection of the informative variables:

if $f^* = 1$, $X_l^* = \underset{X_l \in \Lambda_i}{\arg\min} \, H(X_{i+1} \mid X_l)$

if $f^* = 2$, $(X_l^*, X_m^*) = \underset{(X_l, X_m) \in \Lambda_i^2}{\arg\min} \, H(X_{i+1} \mid X_l, X_m)$

if $f^* = 3$, $(X_l^*, X_m^*, X_n^*) = \underset{(X_l, X_m, X_n) \in \Lambda_i^3}{\arg\min} \, H(X_{i+1} \mid X_l, X_m, X_n)$

Step 3. The forecast of the consumption of day $i+1$:

$$Q_{i+1} = \underset{k=1,\mathrm{K},s_{i+1}}{\arg\max} \, P\!\left(X_{i+1} = Q_{i+1}^k \mid \mathbf{X}^* = \mathbf{C}^*\right).$$

Here the notation $\mathbf{X}^* = \mathbf{C}^*$ stands for

$X_l^* = C_l$, if $f^* = 1$,

$X_l^* = C_l, X_m^* = C_m$, if $f^* = 2$,

$X_l^* = C_l, X_m^* = C_m, X_n^* = C_n$, if $f^* = 2$.

The algorithm was implemented in such a way that the user can make interactive decisions when the code is running. In Step 1 the user can specify, based on the value of conditional entropies, how many previous days should be taken into account in the forecast.

In Step 3 two, three or four dimensional marginal probability distributions are used, depending on conditioning one, two or three earlier days consumptions. We may face to the following problems:

a) X_{i+1} takes on more values with the same probability. In this case, we take their mean value as forecast.

b) In the marginal probability distribution of the historical data never occurs the realization $\mathbf{X}^* = \mathbf{C}^*$. We overcome this problem by using lower marginals.

For example, it may happen that the probability

$$P\!\left(X_{i+1} = Q_{i+1}^k \mid X_l^* = C_l, X_m^* = C_m, X_n^* = C_n\right)$$

cannot be calculated since the conditioning realization did not occur in the historical data. For these cases we have to apply a forecasting scheme which is based on lower marginal probability distributions.

Let us denote by r the dimension of the largest marginal probability distribution of $\mathbf{X}^*$ with the property that the corresponding conditioning set occurs with positive probability.

If the conditioning set contains 3 variables, r can be 2 or 1.

For illustration let us suppose that $r=2$. In this case at least one of the following cases occurs. There exists at least one k such that

$$P\left(X_{i+1} = Q_{i+1}^k \mid X_l^* = C_l, X_m^* = C_m\right) > 0,$$ (5)

or

$$P\left(X_{i+1} = Q_{i+1}^k \mid X_l^* = C_l, X_n^* = C_n\right) > 0,$$ (6)

or

$$P\left(X_{i+1} = Q_{i+1}^k \mid X_m^* = C_m, X_n^* = C_n\right) > 0.$$ (7)

Let us denote by $P_{X_{i+1}}^r$ the sum of the above nonzero probabilities for all $k = 1,\ldots,s_{i+1}$.

For each k we calculate a probability $p_{Q_{i+1}^k}$ as the sum of nonzero probabilities of (5)-(7) for which $X_{i+1} = Q_{i+1}^k$ is taken on, divided by $P_{X_{i+1}}^r$. Using these we define the following random variable.

$$\gamma_r : \begin{pmatrix} Q_{i+1}^1 & \mathrm{K} & Q_{i+1}^k & \mathrm{K} & Q_{i+1}^{s_{i+1}} \\ p_{Q_{i+1}^1} & \mathrm{K} & p_{Q_{i+1}^k} & \Lambda & p_{Q_{i+1}^{s_{i+1}}} \end{pmatrix}$$

The forecasted consumption of the $i+1^{th}$ day is $Q_{i+1} = \underset{k=1,\mathrm{K},s_{i+1}}{\arg\max}\, p_{Q_{i+1}^k}$.

3.4   Decision Making Procedure

Let us suppose that we decided on the amounts of goods to be placed on shelves in the first i days. We have to decide on the amount of goods to be placed on the shelves in the morning of the $i+1^{th}$ day, based on the forecasted consumption. We want the end-of-day amount on display to be equal to the arithmetic mean of the prescribed lower and upper levels. If z_{i+1} denotes the quantity of the line to be placed on the shelves at the beginning of the $i+1^{th}$ day its value have to fulfill the following equality:

$$V_0 + z_1^* + \Lambda + z_i^* - C_1 - \Lambda - C_i + z_{i+1} - Q_{i+1} = \frac{a_{i+1} + b_{i+1}}{2},$$ (8)

where V_0 is the starting amount of line on the shelves at the beginning of the sale; z_1^*,K,z_i^* are the decisions applied in the first i days; C_1,K,C_i are the realized random consumptions in the first i days; Q_{i+1} is the forecasted consumption in the $i+1^{th}$ day, and a_{i+1}, b_{i+1} are the prescribed lower resp. upper bounds on the amount of goods placed on display at the end of the $i+1^{th}$ day.

Solving the equation (8) we get for the optimal decision:

$$z_{i+1}^* = \frac{a_{i+1} + b_{i+1}}{2} - \left(V_0 + z_1^* + \Lambda + z_i^* - C_1 - \Lambda - C_i\right) + Q_{i+1}.$$

3.5   Data Preprocessing

In the practical application of our model the following problem may appear which have to be solved before running the algorithm.

The problem is caused by the relatively small learning data set, and the relatively large range of values which are taken on by each random variable. Therefore we decided to group the values into intervals. Based on the historical data the range of the consumption for each day was divided into 4 intervals as follows.

For each day there was calculated the minimum consumption, maximum consumption, mean value and standard deviation. These divide the range into four intervals. The intervals were delimited by the minimum consumption, the mean value minus the standard deviation, the mean value, the mean value plus the standard deviation and the maximum consumption. Each interval was characterized by the mean value calculated from the historical data.

First we forecasted an interval then on the basis of this we accepted the mean value assigned to this interval as forecast for the consumption.

3.6   Application of the Algorithm

Our dynamic decision making algorithm has been applied to the real-life data set of a 14 day sales period. We got observed data for 46 sales periods. Data of randomly selected 40 sales periods was used as learning data set and data of the remaining 6 sales periods was used as testing data set. In Table 1 there are given the mean values and the standard deviations of the daily consumptions and the prescribed lower resp. upper bounds for the amount of goods to be placed on the shelves. The runs of our dynamic control for the testing data sets can be seen on Figures 1-6.

Table 1

The mean values and standard deviations of the daily consumptions and the prescribed lower and upper bounds for the amount of goods to be placed on the shelves

		Exp. val.	Std. dev.	Lower bound	Upper bound
1	Friday	6.9500	3.5515	6	20
2	Saturday	5.9000	3.7403	6	20
3	Sunday	3.6000	3.3344	6	20
4	Monday	3.2250	2.8328	6	18

5	Tuesday	3.9250	3.1816	5	18
6	Wednesday	3.8250	3.2415	5	18
7	Thursday	4.4500	3.9481	5	16
8	Friday	3.5500	3.2734	5	16
9	Saturday	3.1750	2.5709	5	16
10	Sunday	2.3750	1.8904	5	14
11	Monday	2.1000	1.7802	4	14
12	Tuesday	3.2500	2.4469	4	14
13	Wednesday	2.3250	1.9133	4	12
14	Thursday	2.3250	1.9792	4	12
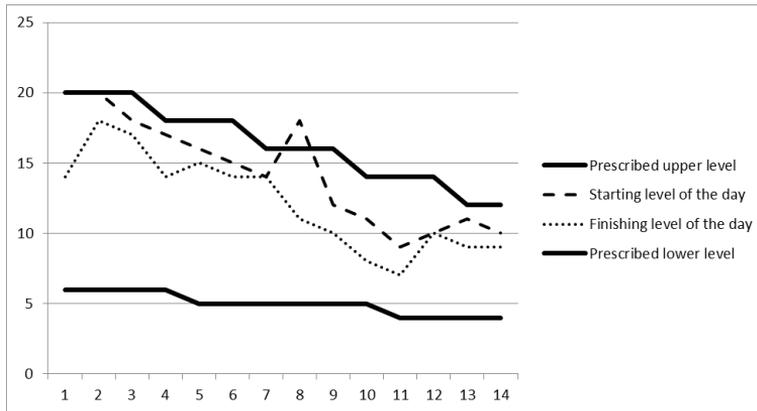
Figure 1

Run of the dynamic control for the first testing data set
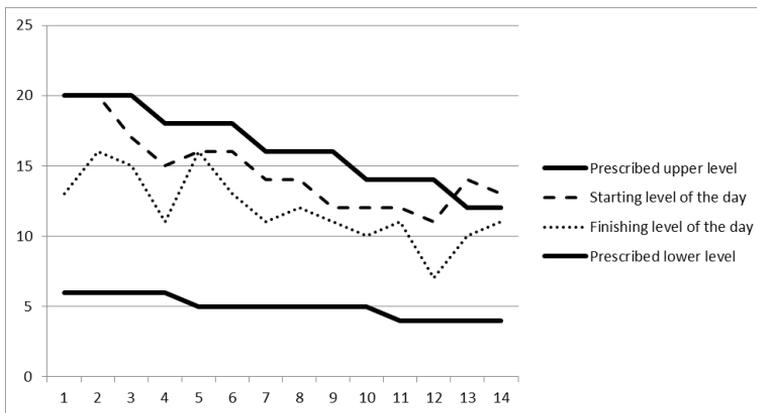
Figure 2

Run of the dynamic control for the second testing data set
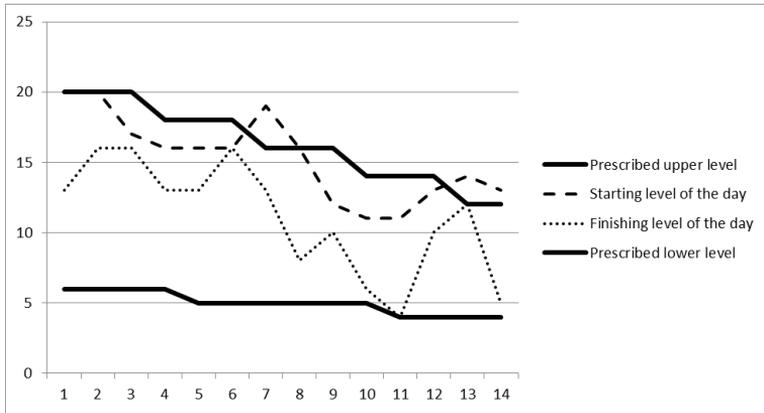
Figure 3
Run of the dynamic control for the third testing data set
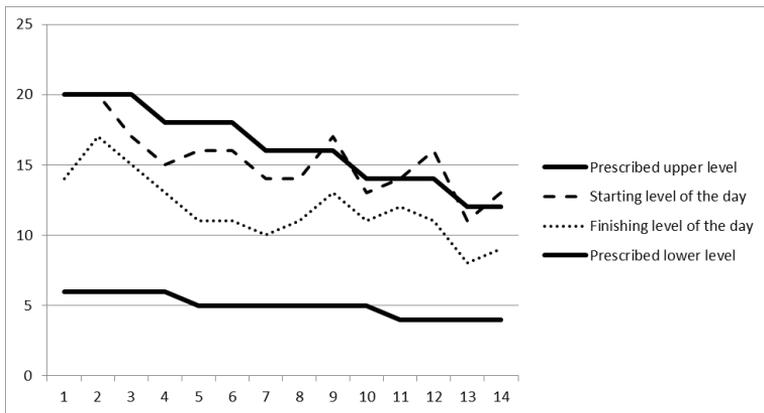
Figure 4
Run of the dynamic control for the fourth testing data set
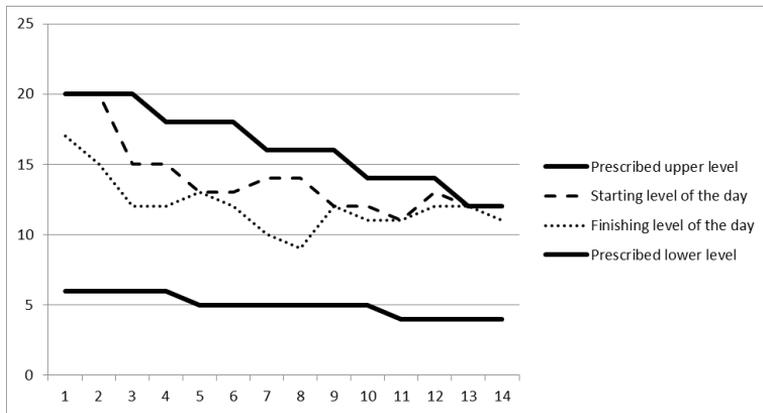
Figure 5
Run of the dynamic control for the fifth testing data set
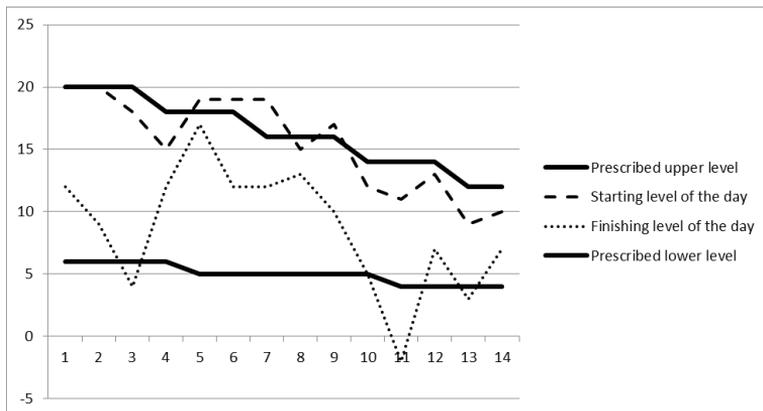
Figure 6
Run of the dynamic control for the sixth testing data set

4   Some Methods to Improve the Quality of the Registered Data and Ideas for Future Work

One of the problems, which occurred, is that the registered consumption was messy, due the fact that there were days when at the end of the day the shelves were empty. In such cases the registered consumptions were considered as the quantities displayed at the beginning of these days. We recommend that in such cases the registered quantity should somehow indicate this fact, for example by

C_i^+, and this fact should be taken into account in the procedure of forecasting the future consumptions.

Very often, the promotional sale for a given line, has effects on the demand of other goods. It would be important to investigate these effects and include them in the calculation.

Conclusions

We presented two new methods of dynamic forecasting for the consumptions within sales periods and a decision procedure based on the forecasted consumption and the prescribed levels. The first method can be applied in cases when the probability distribution of the random consumptions can be supposed to be normal, i.e. the standard deviation of the data is relatively small according to the mean value.

The advantage of the second method is that it needs no hypothesis on the theoretical probability distribution, but for accurate forecasting, it needs a larger historical dataset. The dataset could be enlarged by other observed sales periods for similar items.

Both methods presume that the sales periods were observed under equal market and advertising conditions.

References

[1]   G. Cachon and M. Fisher, Campbell Soup's continuous replenishment program: Evaluation and enhanced inventory decision rules, *Production and Operations Management*, **6**, 266-276, 1997

[2]   T. M. Cover and J. A. Thomas, *Elements of Information Theory,* John Wiley & Sons, New York, 2006

[3]   K. H. Van Donselaar, J. Peters, A. De Jong and R. A. C. M. Broekmeulen, Analysis and forecasting of demand during promotions for perishable items, *International Journal of Production Economics*, **172**, 65-75, 2016

[4]   R. Fildes and P. Goodwin, Against your better judgment? How organizations can improve their use of management judgment in forecasting, *Interfaces*, **37**, 570-576, 2007

[5]   R. Fildes, K. Nikolopoulos, S. F. Crone and A. A. Syntetos, Forecasting and operational research: a review, *Journal of the Operational Research Society*, **59,** 1150-1172, 2008

[6]   M. Leonard, Promotional analysis and forecasting for demand planning: a practical time series approach, https://support.sas.com/rnd/app/ets/papers/PromotionalAnalysis.pdf, 2001

[7]   S. Ma, R. Fildes and T. Huang, 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*, **249**, 245-257, 2016

[8]   A. Prékopa and T. Szántai, On Optimal Regulation of a Storage Level with Application to the Water Level Regulation of a Lake, *European Journal of Operational Research*, **3**, 175-189, 1979

[9]   A. Rényi, On measures of information and entropy, in: Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability, 547-561, 1961

[10]   C. E. Shannon, A mathematical theory of communication, *The Bell System Technical Journal*, **27**, 379-423, July 1948

[11]   J. W. Taylor, Multi-item sales forecasting with total and split exponential smoothing, *Journal of the Operational Research Society*, 62, 555-563, 2011

[12]   J. R. Trapero, D. J. Pedregal, R. Fildes and N. Kourentzes, Analysis of judgmental adjustments in the presence of promotions, *International Journal of Forecasting*, **29** 234-243, 2013

[13]   J. R. Trapero, N. Kourentzes and R. Fildes, On the identification of sales forecasting models in the presence of promotions, *Journal of the Operational Research Society*, **66**, 299-307, 2015