Pamukkale Univ Muh Bilim Derg, 24(4), 705-712, 2018

Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi

Pamukkale University Journal of Engineering Sciences

A comparison of sales forecasting methods for a feed company: A case study

Bir yem işletmesi için satış tahmin yöntemlerinin karşılaştırılması: Bir vaka çalışması

Leyla DEMİR[1*], *Selahattin AKKAŞ*[2]

[1]Department of Industrial Engineering, Engineering Faculty, Pamukkale University, Denizli, Turkey
ldemir@pau.edu.tr

[2]Department of Computer Engineering, Engineering Faculty, Pamukkale University, Denizli, Turkey.
sakkas@pau.edu.tr

Received/Geliş Tarihi: 16.11.2017, Accepted/Kabul Tarihi: 17.01.2018 doi: 10.5505/pajes.2018.58235
* Corresponding author/Yazışılan Yazar
Research Article/Araştırma Makalesi

Abstract

Due to global warming in recent years, using natural resources in an effective way has become more and more important to our world. Decreasing natural resources are pushing agriculture and food chains to adopt more efficient management strategies. The first condition for a successful management is to make plans based on accurate and reliable forecasts. In this study, using real-world data, forecasting models are compared for the products of a feed company which is the first chain of agriculture and food chain systems. The traditional statistical time series methods are compared to two popular and effective computational intelligence techniques, i.e. artificial neural networks and support vector regression. The accuracy of the forecasts is calculated by three different error measures, i.e., the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the mean squared error (MSE). The results show that support vector machines produces significantly better results comparing to both time series methods and artificial neural networks.

Keywords: Sales forecasting, Time series methods, Artificial neural networks, Support vector regression, Agriculture and food chains, Feed industry

Öz

Son yıllarda küresel ısınmadan dolayı doğal kaynakların etkin ve verimli kullanımı dünyamız için daha da önemli bir hale gelmiştir. Azalan doğal kaynaklar tarım ve gıda zincirlerini daha etkin yönetim stratejileri benimsemeye zorlamaktadır. Başarılı bir yönetimin ilk şartı doğru ve güvenilir tahminlere dayalı planlar yapmaktır. Bu çalışmada tarım ve gıda zincirlerinin ilk halkası olan yem endüstrisinde yer alan bir işletme için gerçek verilere dayalı olarak oluşturulan tahmin modelleri karşılaştırılmıştır. Geleneksel istatistiksel zaman serisi yöntemleri popüler ve hesaplamalı olarak etkin iki yapay zekâ tekniği olan yapay sinir ağları ve destek vektör regresyonu yöntemleri ile karşılaştırılmıştır. Tahminlerin doğruluğu ortalama mutlak hata (MAE), ortalama mutlak yüzde hata (MAPE) ve ortalama hata kare (MSE) gibi üç farklı hata ölçütü kullanılarak hesaplanmıştır. Sonuçlar destek vektör regresyonu yönteminin zaman serisi ve yapay sinir ağları yöntemlerine göre daha iyi sonuçlar ürettiğini göstermiştir.

Anahtar kelimeler: Satış tahmini, Zaman serisi yöntemleri, Yapay sinir ağları, Destek vektör regresyonu, Tarım ve gıda zincirleri, Yem endüstrisi

1 Introduction

In today's increasingly competitive conditions, companies are able to sustain their assets depends on effective production planning. Particularly, with in case of global warming and diminishing resources in the world, companies need to develop efficient management strategies considering both the use of natural resources and also environmental factors. For that reason, sales forecasting is of great importance in terms of planning the future activities of the companies especially in agriculture and food chains.

There is a vast literature on sales forecasting in supply chains. In 2006, McCarthy et al. [1] present a literature review study in sales forecasting management area. More recently, Syntetos et al. [2] present a detailed analysis on forcasting in supply chain. They provide a compherensive review and identify the research gaps in this area both from researcher and also practitioner perspectives.

Caniato et al. [3] deal with the problem of forecasting irregular demand including structural, manegarial and random variabilities for fresh food industry. The authors devolop a forecasting model based on clustering approach. Another application in fresh food industry is presented in [4]. In this study, the impact of heterogeneity of customer requests on demand forecasting approaches for not only fresh food industry but also spare parts and retails industries is considered. An appropriate methodology for each industry is designed based on the analysis of customer behavior and demand patterns.

The comparison of top-down and bottom-up approaches for forecasting demand is presented by Widiarta et al. [5]. The performances of these strategies are found to be nearly identical even if there is correlation between the demands of items belonging to same product family. Likewise, Sbrana and Silvestrini [6] deal with top-down and bottom-up forecasting strategies for aggregate demand. In this study, the results achieved by Widiarta et al. [7] are extended and generalized. Moreover, the theoretical results are confirmed by simulation.

Kerkkänen et al. [8] show the effects of sales forecast errors in supply chain management by a case study. They present the role of sales forecasts in production management and analyze the characteristics of sales forecasting errors for a large process industry. Martins and Werner [9] identify the differences in the accuracy of forecasts in industrial series obtained by different forecast combinations. Acar and Everette [10] develop forecasting method selection for a global manufacturer of lubricants and fuel additives.

Apart from using the traditional statistical techniques for time series forecasting, in recent years, computational intelligence methods such as artificial neural networks (ANNs) and support vector regression (SVR) are extensively used in many application areas. Ansuj et al. [11] compare ARIMA models with ANN. Alon et al. [12] compare the performances of traditional methods and ANN for forecasting retail sales. Thomassey and Happiette [13] propose an ANN for textile distribution. Kuo and Xue [14], [15] and Kuo et al. [16] are interested in the performance of ANN and fuzzy neural networks for sales forecasting. Efendigil et al. [17] also compare ANN and neuro-fuzzy models for demand forecasting in durable consumer goods industry. Wong et al. [18] propose adaptive neural networks for time series prediction. Lu et al. [19] compare the multivariate adaptive regression splines and ANNs for sales forecasting in computer wholesaler industry. Štěpnička et al. [20] propose a hybrid methodology, including evolutionary neural networks, support vector machines and genuine linguistic fuzzy rules, for time series forecasting. For more information on ANN applications, the interested readers may refer to review study of Tkáč and Verner [21] In this study, the ANN applications in business are classified from 1994 to 2015.

In recent years, there have been many successful applications of SVR for time series forecasting. Garćia et al. [22] propose an intelligent system based on support vector machines for demand forecasting in supply chains. Du et al. [23] present a new algorithm based on the support vector machines for forecasting demand for perishable farm products. Wang and Du [24] employ SVR with differential evolution for forecasting of motherboard shipments manufacturers in Taiwan. Vahdani et al. [25] present a case study in cosmetic industry using a meta-heuristic for training SVR.

To the best of authors' knowledge the only study considered forecasting in feed industry is the study of Reboiro-Jato et al. [26]. However, their study focus on feed production for poultry farms whereas our study considers an industrial firm producing feeds for both poultry and bovine animals. Moreover, our study compares the performances of not only time series methods and artificial neural networks but also support vector regression. The rest of this paper is organized as follows. The next section explains forecasting methods used in this study. In section 3, the details of the case study are presented and the results are discussed. Finally, the findings of the study are summarized and some future research directions are given in Section 4.

2 Methodology

The process of predicting future events is called as forecasting. In this process, evaluating the results of past events or incidents had occurred in the results of potential future events that are tried to be predictable. The aim of sales forecasting is trying to determine the amount of demand for the products/product groups in the future based on the actual sales.

There is an abundance of forecasting method in literature. These methods can be classified as qualitative and quantitative methods. When there is little or no data available qualitative methods are commonly preferred and these methods are mostly based on personal views. On the other hand, quantitative methods are based on mathematical approaches and they use historical data to predict the future.

In this study, using real-world data forecasting models based on time series methods, artificial neural networks and support vector regression are developed for a feed company. In the following subsections, the basic concepts of all methods used in this study are briefly reviewed.

2.1 Time series methods

Times series methods is the most used forecasting methodology in the literature. This methodology uses past sales data to predict the future demand. In most cases, past sales data of a product can be divided six components: average amount of sales for that period, a trend, seasonal element, cyclical elements, random variations and autocorrelation. Trend shows the tendency of a time series having a stable pattern of growth or decline. A seasonal element refers sales data pattern which is repeated at fixed intervals such as month, year etc. If the length and the magnitude of the cycle vary it is called as cyclical element. Random variations are occurred by chance events. Lastly, the persistence of occurrence is denoted by autocorrelation. The interested readers may refer to [27] and [28] for more information on time series methods.

Before proceeding we define some basic notations used in this study. Assume that $D_1, D_2, \dots, D_t \dots$ as the observed values of sales during the periods $1, 2, \dots, t$. Moreover, the forecasted value for period t is defined as F_t. Using this notation the methods employed in this study are briefly explained in the following subsections.

2.1.1 Moving averages (MA)

Moving averages (MA) is the simplest but very popular forecasting method. When the past sales values of a product have not seasonal behavior and the related data is neither growing nor declining rapidly moving averages can be useful for predicting future sales. A moving average is defined as the arithmetic average of the most recent N observations as follows [29]:

$$F_t = (D_{t-1} + D_{t-2} + \cdots + D_{t-N})/N \quad (1)$$

2.1.2 Exponential smoothing (ES)

The major drawback of the moving averages method is the need to continually carry a large amount of historical data. In many applications, the most recent occurrences are more indicative of the future than those in the more distant past. In these cases, exponential smoothing (ES) is the most logical and easiest method to use and the formulation of simple exponential smoothing is given as follows [29]:

$$F_t = \alpha D_{t-1} + (1-\alpha)F_{t-1} \quad (2)$$

where $0 < \alpha \leq 1$ is the smoothing constant.

2.1.3 Holt's linear method (HLM)

Holt [30] extended the simple exponential smoothing method to allow forecasting of data with trends. This method uses following three equations with two smoothing constants α and β ($0 < \beta \leq 1$) [29]:

$$S_t = \alpha D_t + (1-\alpha)(S_{t-1} + G_{t-1}) \quad (3)$$

$$G_t = \beta(S_t - S_{t-1}) + (1-\beta)G_{t-1} \quad (4)$$

$$F_t = S_t + G_t \quad (5)$$

where S_t denotes the level and G_t denotes the slope of the series at time t, respectively.

2.1.4 Winters's method (WM)

Winters's method is also exponential smoothing method having three smoothing constants and it considers seasonality in the series. The advantage of this method is that it can be easily updated as new data become available. Three exponential smoothing equations are used each period to update estimates of deseasonalized series, the seasonal factors, and the trend. These equations have different smoothing constants labeled as α, β, and γ $(0 < \gamma \leq 1)$.

The current level of the deseasonalized series, S_t, is given by [29]:

$$S_t = \alpha(D_t/c_{t-N}) + (1-\alpha)(S_{t-1} + G_{t-1}) \quad (6)$$

The trend is updated in a similar manner as in Holt's linear method:

$$G_t = \beta[S_t - S_{t-1}] + (1-\beta)G_{t-1} \quad (7)$$

Finally the seasonal factors are computed as follows:

$$c_t = \gamma(D_t/S_t) + (1-\gamma)c_{t-N} \quad (8)$$

Here c_t is the multiplicative seasonal component in period t, where $\sum c_t = N$. Using these equations the forecast in period t can be made by the following equation:

$$F_{t+1} = (S_t + G_t)c_{t+1-N} \quad (9)$$

2.2 Artificial neural networks (ANNs)

Artificial neural networks (ANNs) are computational structures that simulate the cognitive learning process of human brain. In contrast to traditional computational techniques, ANNs have capabilities to solve a variety of problems including nonlinearity such as prediction, optimization, recognition and control problems. The basics of ANN are founded on the pioneering work of McCulloch and Pitts [31]. Later, ANNs have received considerable interest and a lot of works have been done in this area.

Neuron is the basic element of ANN that processes information. The structure of a neuron can be explained by a nonlinear model as shown in Figure 1.

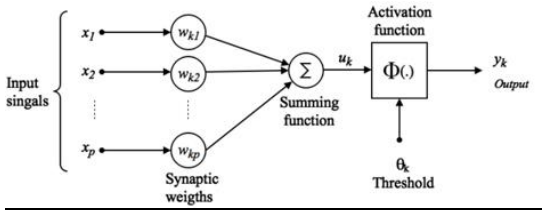


Figure 1: A nonlinear model of a neuron [31].

Three basic elements are identified from this figure. As it can be seen from this figure, a signal x_j at the input of synapse j is multiplied by synaptic weight w_{kj} and it is connected with the neuron k. Then summing function adds the input signals and the respective synapses of the neuron via a linear combiner. Lastly, the amplitude the output of a neuron is limited by an activation function. There are several activation functions used in literature such as threshold, piecewise linear, sigmoid or Gaussian functions.

In mathematical terms, a neuron k is described by the following equations [32]:

$$u_k = \sum_{j=1}^{m} w_{kj}\, x_j \quad (10)$$

$$y_k = \Phi(u_k + b_k) \quad (11)$$

Where u_k is the linear combiner output; $w_{k1}, w_{k2}, \ldots, w_{km}$ are the synaptic weights of neuron k; $x_1, x_2, \ldots, x_m$ are the input signals; y_k is the output signal of the neuron, $\Phi(.)$ is the activation function, and b_k is the bias.

If artificial neurons are depicted by nodes and neuron inputs and outputs are connected with directed edges ANNs become weighted directed graphs. Based on the structure of the graph, ANNs can be classified in two categories. The first one is feed-forward networks having no loops, and the other is recurrent networks having loops because of feedback connections. Figure 2 shows the feed-forward ANN network structure which is also employed in this study. This structure contains two layers with sigmoid function for hidden layer and linear activation function for output layer as seen in Figure 2.

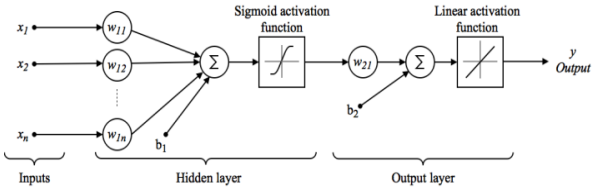


Figure 2: The ANN structure.

Different network structures need appropriate learning algorithms, such as supervised, unsupervised, and hybrid learning. Supervised learning algorithm analyses the data and it tries to produce output values as close as possible to the known input data by determining the appropriate weights. In contrast, in unsupervised learning algorithm, while training data set, there is no need to a correct answer associated with each input pattern. Lastly, hybrid learning algorithms combine supervised and unsupervised learning.

In ANN, when using the supervised learning algorithms the prediction error tries to be minimized. There are different optimization methods for minimization, such as Levenberg-Marquardt, Quasi-Newton, gradient descent, and conjugate gradient methods. Among them Levenberg-Marquardt method is the most widely used in ANN literature. In our study, Levenberg-Marquardt backpropagation algorithm is employed as training algorithm.

For more information on ANN can be found in [32].

2.3 Support vector regression (SVR)

The support vector machines (SVM) method is proposed by Vapnik [33] and it is based on statistical learning theory. Originally this method is designed to solve the classification and regression problems and later it is extended by Drucker et al. [34] to solve the estimation and prediction problems. This extended form of SVM is named as support vector regression (SVR).

The SVR method tries to minimize the prediction error and its aim is to find a function that approximates the training data set in this way. In this process, the risk of over-fitting is reduced by simultaneously trying to maximize the flatness of the function [35].

To explain the SVR, let consider a data set of training points, $\{(x_1, y_1), \dots, (x_l, y_l)\}$, where $x_i \in R^n$ is a future vector and $y_i \in R$ is the target output. In theory, the nonlinear relationship between input and output data should be formulated by a linear function. A SVR function is defined as [33]:

$$f(x) = w^T \Phi(x) + b \tag{12}$$

Where $f(x)$ denotes the forecasted values, $\Phi(\cdot)$ is a nonlinear mapping function, and w ($w \in R^n$) and b ($b \in R$) are the adjustable coefficients. The standard form of SVR, under the given parameters $C > 0$ and $\varepsilon > 0$, is defined as [33]:

$$\text{Min}_{w,b,\xi,\xi^*}(1/2)w^T w + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \tag{13}$$

Subject to

$$\begin{cases} w^T\Phi(x_i) + b - y_i \le \varepsilon + \xi_i \\ y_i - w^T\Phi(x_i) - b \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0, i = 1,2, \dots, l \end{cases} \tag{14}$$

where ξ_i^* denotes training errors above ε and ξ_i denotes training errors below $-\varepsilon$.

After solving the above quadratic optimization problem with inequality constraints, the parameter vector w in Eq. (12) is calculated as follows [33]:

$$w = \sum_{i=1}^{l}(\lambda_i^* - \lambda_i)\Phi(x_i) \tag{15}$$

where λ_i^*, and λ_i are Lagrange multipliers. Lastly, the SVR function is calculated as follows [33]:

$$f(x) = \sum_{i=1}^{l}(\lambda_i^* - \lambda_i)K(x_i, x_j) + b \tag{16}$$

where $K(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right)$ is the Gaussian radial basis kernel function.

2.4 Forecast errors

Forecast error (e_t) for one period defined as the difference between the actual and forecasted values as follows [29]:

$$e_t = F_t - D_t \tag{17}$$

If $e_1, e_2, \dots, e_n$ is defined as the forecast errors observed during n periods, three common measures of forecast error, i.e. the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the mean squared error (MSE), are calculated by the following equations [29]:

$$\text{MAE} = (1/n)\sum_{i=1}^{n}|e_i| \tag{18}$$

$$\text{MAPE} = \left[(1/n)\sum_{i=1}^{n}|e_i/D_i|\right] \times 100 \tag{19}$$

$$\text{MSE} = (1/n)\sum_{i=1}^{n}e_i^2 \tag{20}$$

MAE is the average of the absolute difference between the estimated and actual values. If the absolute deviation is expressed as a percentage of the actual value, it is called as MAPE. In the MSE method, the averages of the squares of the errors are taken so that the positive (+) or negative (-) errors do not affect each other.

3 Case study

In this paper, a sales forecasting study is carried out for five predefined products of a feed company operating in Denizli, Turkey. The firm is the first private feed company in Turkey and it is established in 1969. The company produces all types of feeds for poultry, sheep and bovine animals. It also produces feeds for fish farms. The aim of this study is to manage the production process more efficiently and make plans based on accurate and reliable forecasts. For this purpose, continuously produced and the most sold five products which had also large profit values were chosen after discussing with the managers of the company.

After deciding on which products would have been forecasted, totally 23 months data set covering May 2014 to March 2015 for each product was gathered. The descriptive statistics of the product sales are given in Table 1.

Table 1: Descriptive statistics of the product sales.

	N	Minimum	Maximum	Std. Deviation
Product 1	23	3245	13165	2504.14
Product 2	23	428	1824	365.05
Product 3	23	99	2055	499.00
Product 4	23	8	361	116.36
Product 5	23	24	110	17.65

The first product (Product 1) is produced for milk cows during the lactation period and it is used to increase the milk productivity. This feed is one of the most sold products of the company. As it can be seen from Table 1 the sales amounts are the highest one among the five products. It should be noted that the sales amounts are given as ton for all products. The second feed (Product 2) is produced for calves for their growing in early periods. The third product (Product 3) is a special feed that provides high live weight gain in a short time and because of that it is highly preferred. The fourth type of feed (Product 4) is consumed between the 45th day and the 90th day after birth and it is extensively consumed in that period. Finally, Product 5 is a feed that newborn calves consume from the first week until the end of the third month.

After determining the products, the forecasting methods explained in the previous section were employed to these products and the results were compared by three different measures defined in Section 2.4.

MS Excel is used to perform time series analysis. While employing time series methods the parameter values are chosen so as to minimize the errors. For instance, in MA, three, four and five months moving averages are examined and it is decided to use four months moving averages for products 1, 2, and 5; three months moving averages for product 3 and two months moving averages for product 4. In a similar manner, smoothing parameters of the methods for exponential smoothing, Holt's linear, and Winters's methods are chosen so as to minimize the forecast errors. Excel Solver tool is used for this purpose. The optimum values of these parameters are given in Table 2.

As mentioned in Section 2.2, a two layer feed-forward network having sigmoid and linear activation functions is designed for

ANN using MATLAB nftool. Levenberg-Marquardt backpropagation algorithm is used as training algorithm. The data set is divided in three groups as follows: 70% of data is used for training, %15 is used for validation and %15 is used for testing. Moreover, to minimize the forecast errors different neuron numbers, i.e. 5, 10, and 15 are examined and the values produces minimum forecast errors are chosen for each product. The optimal neuron numbers are determined as 10, 5, 10, 10, and 5 for products 1, 2 ,3, 4 and 5, respectively.

Support vector regression model is also coded using MATLAB. The data are normalized before employing SVR and the Gaussian Kernel function is used to develop the model. Different values are tried for SVR parameters and it is found that taking SVR parameters as $C = 100$, $\varepsilon = 0.01$, $\sigma = 0.1 - 0.2$ is produced best results.

Table 3 shows the forecast errors for five products and Figure 3 shows the graphs for real data and forecasted values for each product, respectively. As it can be seen in Table 3, the SVR method produces significantly better results than other methods for almost all products, except product 4. For this product ANN is the best in respect to both MAE and MSE values but when considering MAPE value SVR produces the best result. A significant advantage of SVR is that whilst ANNs can suffer from multiple local minima, the solution to a SVR is global and unique. Because of this reason SVR is more successful than ANN while predicting product sales.

Table 2: Optimum values of the smoothing parameters.

	Method			
	MA	ES	HLM	WM
Product 1	4	$\alpha = 0.40$	$\alpha = 0.40, \beta = 0.05$	$\alpha = 0.20, \beta = 0.01, \gamma = 0.01$
Product 2	4	$\alpha = 0.40$	$\alpha = 0.35, \beta = 0.05$	$\alpha = 0.20, \beta = 0.01, \gamma = 0.01$
Product 3	3	$\alpha = 0.85$	$\alpha = 0.85, \beta = 0.01$	$\alpha = 0.85, \beta = 0.01, \gamma = 0.85$
Product 4	2	$\alpha = 0.95$	$\alpha = 0.95, \beta = 0.01$	$\alpha = 0.95, \beta = 0.10, \gamma = 0.95$
Product 5	4	$\alpha = 0.95$	$\alpha = 0.95, \beta = 0.01$	$\alpha = 0.95, \beta = 0.01, \gamma = 0.95$

Table 3: Forecast errors for each product.

METHOD	Product 1			Product 2			Product 3			Product 4			Product 5		
	MAE	MAPE (%)	MSE	MAE	MAPE (%)	MSE	MAE	MAPE (%)	MSE	MAE	MAPE (%)	MSE	MAE	MAPE (%)	MSE
MA	2024	28	7178014	278	30	138808	376	70	264347	80	92	9109	13	25	333
ES	1979	28	7141512	268	29	138949	377	71	262198	60	74	5489	18	35	585
HLM	1992	28	7571097	276	30	145006	377	70	266739	57	72	4954	16	32	444
WM	1951	27	6594458	266	29	131118	381	70	271935	57	71	5203	16	32	448
ANN	1088	14	2016629	116	10	29352	145	13	40143	19	33	685	9	13	258
SVR	360	3	304701	67	5	11776	76	5	17569	25	25	2609	5	6	83

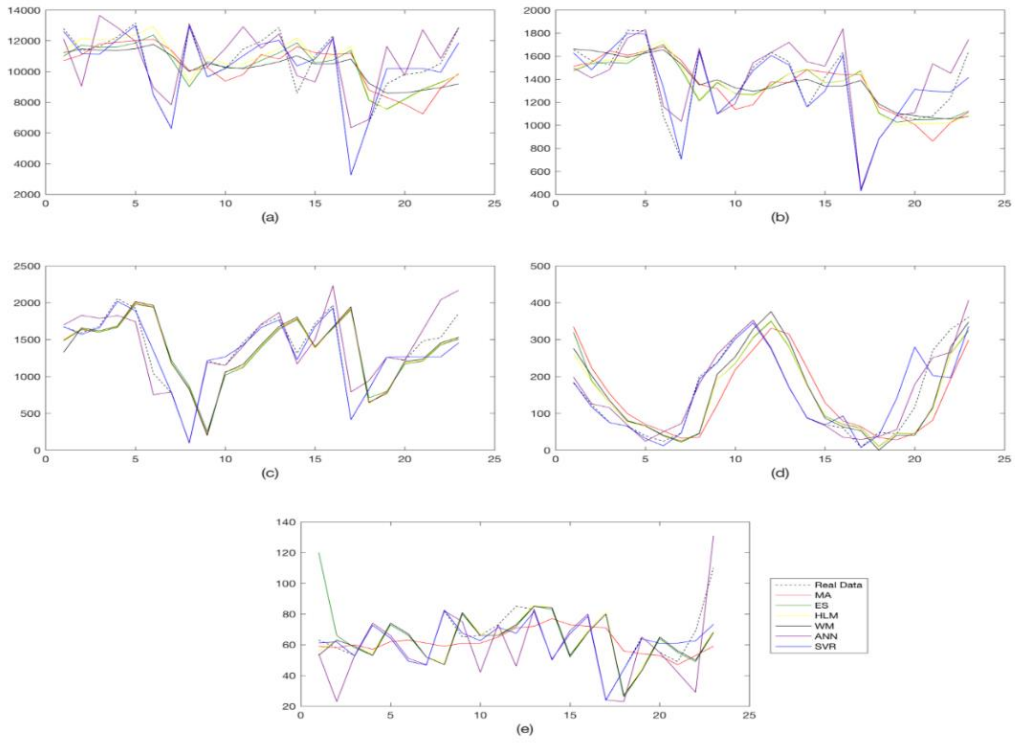


Figure 3: Forecasted values. (a): Product 1, (b): Product 2, (c): Product 3, (d): Product 4, (e): Product 5.

The superiority of SVR method can be seen in also Figure 4 and 5. These figures are the fitted plots and they show whether forecasted values are close or not to real data. SVR method produces best forecasted values for all products except product 4 again. Since R-value for product 4 (0.89) in SVR is not too low and the SVR method has also superiority on MAPE for same product, it can be concluded that SVR method can be used to predict future sales of all these products.

On the other hand, when time series methods compares to each other Winters's method is the best for products 1 and 2 while the superiority of other methods is not clear for other products. If the company tries to forecast using traditional time series methods which are simple and can be known by managers, the decision will be problematic to choose which method should be used. These results show us why company should use the SVR method for forecasting instead of using traditional forecasting methodologies.

The results obtained for considered five products can be extended for other products of the company which has a similar behavior. So, it can be concluded if the SVR method is integrated with the company's database system and also planning programs such as MRP and ERP softwares, the company can make reliable forecasts. In this way, company makes accurate production and purchasing plans which increases the efficiency of the production process and this results more profitable company. This also provides the efficient management of natural resources.

4 Conclusions

Global warming and diminishing resources in the world pushes the companies to develop efficient management strategies considering both the use of natural resources and also environmental factors. Because of that sales forecasting is of great importance in terms of planning the future activities of the companies especially in agriculture and food chains. In this research, a forecasting study is carried out for feed industry which is one of the most important components of the agriculture and food chains.

The sales forecasting is made by employing time series methods, artificial neural networks and support vector regression for predefined five products of a feed company, located in Denizli, Turkey. Using real-world data, the MAE, MAPE, and MSE values of each forecasted method are calculated and compared to determine the most appropriate forecasting method. The results show that the SVR method produces the best results compared to both time series methods and ANN. Based on these results it can be said that if the proposed SVR model is integrated with the company's database system the company can make more reliable forecasts. The production and purchasing plans based on reliable forecasts helps to manage the company more efficiently while using natural resources in an effective way.

As a future study, a decision support tool can be developed to integrate the SVR method into company's database system. Moreover, the efficiency of the SVR method can be compared to other forecasting methodologies such as grey prediction model.

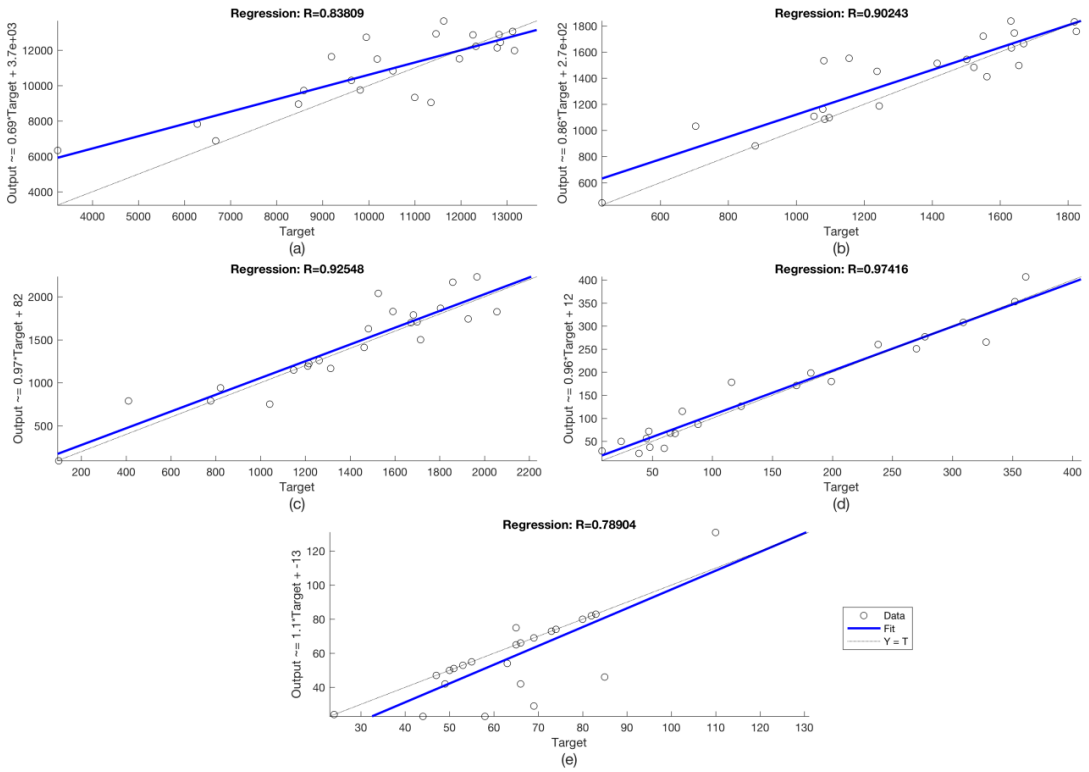


Figure 4: ANN fitted plots: (a): Product 1, (b): Product 2, (c): Product 3, (d): Product 4, (e): Product 5.

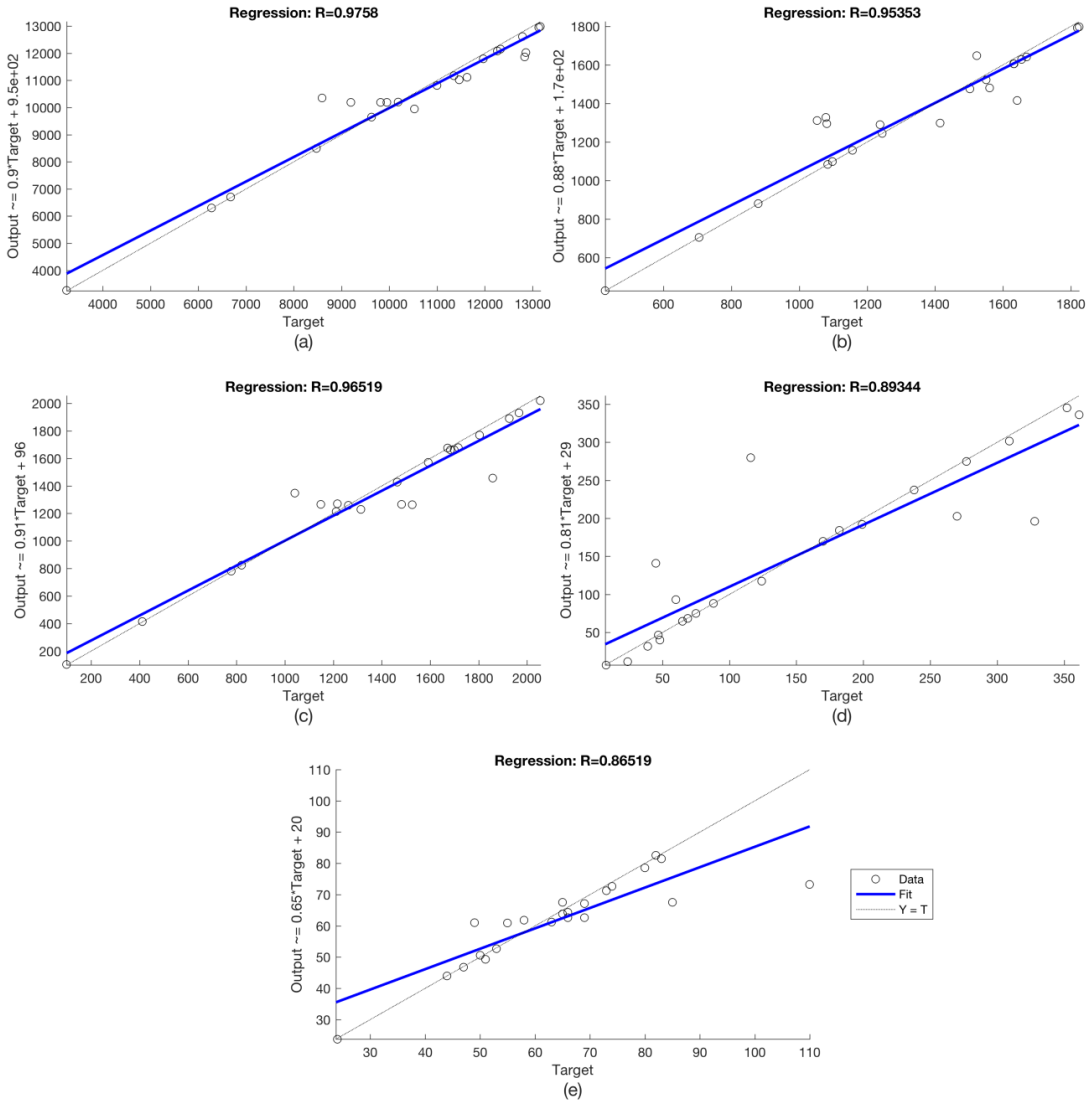


Figure 5: SVR fitted plots. (a): Product 1, (b): Product 2, (c): Product 3, (d): Product 4, (e): Product 5.

5 References

[1] McCarthy TM, Davis DF, Golicic SL, Mentzer JT. "The evolution of sales forecasting management: a 20-year longitudinal study of forecasting practices". *Journal of Forecasting*, 25(5), 303-324, 2006.

[2] Syntetos AA, Babai Z, Boylan JE, Kolassa S, Nikolopoulos, K. "Supply chain forecasting: theory, practice, their gap and the future". *European Journal of Operational Research*, 252(1), 1-26, 2016.

[3] Caniato F, Kalchschmidt M, Ronchi S, Verganti R, Zotteri G. "Clustering customers to forecast demand". *Production Planning and Control*, 16(1), 32-43, 2005.

[4] Kalchschmidt M, Verganti R, Zotteri G. "Forecasting demand from heterogeneous customers". *International Journal of Operations & Production Management*, 26(6), 619-638, 2006.

[5] Widiarta H, Viswanathan S, Piplani R. "Forecasting item-level demands: an analytical evaluation of top–down versus bottom–up forecasting in a production- planning framework". *IMA Journal of Management Mathematics*, 19(2), 207-218, 2008.

[6] Sbrana G, Silvestrini, S. "Forecasting aggregate demand: analytical comparison of top-down and bottom-up approaches in a multivariate exponential smoothing framework". *International Journal of Production Economics*, 146(1), 185-198, 2013.

[7] Widiarta H, Viswanathan S, Piplani R. "Forecasting aggregate demand: an analytical evaluation of top–down versus bottom–up forecasting in a production-planning framework". *International Journal of Production Economics*, 118(1), 87-94, 2009.*International Journal of Production Economics*, 118(1), 87-94, 2009.

[8] Kerkkänen A, Korpela J, Huiskonen J. "Demand forecasting errors in industrial context: measurement and impacts". *International Journal of Production Economics*, 118(1), 43-48, 2009.
[9] Martins VLM, Werner L. "Forecast combination in industrial series: a comparison between individual forecasts and its combinations with and without correlated errors". *Experts Systems with Applications*, 39(13), 11479-11486, 2012.
[10] Acar Y, Everette SGJ. "Forecasting method selection in a global supply chain". *International Journal of Forecasting*, 28(4), 842-848, 2012.
[11] Ansuj AP, Camargo ME, Radharamanan R, Petry DG. "Sales forecasting using time series and neural networks". *Computers & Industrial Engineering*, 31(1-2), 421-424, 1996.
[12] Alon I, Qi M, Sadowski, RJ. "Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods". *Journal of Retailing and Consumer Services*, 8(3), 147-156, 2001.
[13] Thomassey S, Happiette M. "A neural clustering and classification system for sales forecasting of new apparel items". *Applied Soft Computing*, 7(4), 1177-1187, 2007.
[14] Kuo RJ, Xue KC. "A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights". *Decision Support Systems*, 24(2), 105-126, 1998a.
[15] Kuo RJ, Xue KC. "An intelligent sales forecasting system through integration of artificial neural network and fuzzy neural network". *Computers in Industry*, 37(1), 1-15, 1998b.
[16] Kuo RJ, Wu P, Wang CP. "An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination". *Neural Networks*, 15(7), 909-925, 2002.
[17] Efendigil T, Önüt S, Kahraman C. "A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis". *Expert Systems with Applications*, 36(3), 6697-6707, 2009.
[18] Wong WK, Xia M, Chu WC. "Adaptive neural network model for time-series forecasting". *European Journal of Operational Research*, 207(2), 807-8016, 2010.
[19] Lu C-J, Lee T-S, Lian C-M. "Sales forecasting for computer wholesalers: A comparison of multivariate adaptive regression splines and artificial neural networks". *Decision Support Systems*, 54(1), 584-596, 2012.
[20] Štěpnička M, Cortez P, Donate JP, Štěpničková L. "Forecasting seasonal time series with computational intelligence: On recent methods and the potential of their combinations". *Expert Systems with Applications*, 40(6), 1981-1992, 2013.
[21] Tkáč M, Verner R. "Artificial neural networks in business: Two decades of research". *Applied Soft Computing*, 38, 788-804, 2016.
[22] Garćia FT, Villalba LJG, Portela J. "Intelligent system for time series classification using support vector machines applied to supply chain". *Expert System with Applications*, 39(12), 10590-10599, 2012.
[23] Du XF, Leung SCH, Zhang JL, Lai KK. "Demand forecasting of perishable farm products using support vector machine". *International Journal of Systems Science*, 44(3), 556-567, 2013.
[24] Wang F-K, Du T. "Implementing support vector regression with differential evolution to forecast motherboard shipments". *Expert System with Applications*, 41(8), 3850-3855, 2014.
[25] Vahdani B, Razavi F, Mousavi SM. "A high performing meta-heuristic for training support vector regression in performance forecasting of supply chain". *Neural Computing and Applications*, 27(8), 2441-2451, 2016.
[26] Reboiro-Jato M, Glez-Dopazo J, Glez D, Laza R, Gálvez JF, Pavón R, Glez-Peña D, Fdez-Riverola F. "Using inductive learning to assess compound feed production in cooperative poultry farms". *Expert Systems with Applications*, 38(11), 14169-14177, 2011.
[27] Makridakis S, Wheelwright SC, Hyndman JR. *Forecasting: Methods and Applications*. 3rd ed. New York, USA, John Wiley&Sons, 1998.
[28] Montgomery DC, Jennings CL, Kulahci M. *Introduction to Time Series Analysis and Forecasting*. 2nd ed. New Jersey, USA, John Wiley & Sons, 2016.
[29] Nahmias S. *Production&Operation Analysis*. 5th ed. New York, USA, McGraw-Hill, 2005.
[30] Holt CC. *Forecasting seasonal and trends by exponentially weighted moving averages*. Pittsburgh, USA, Carnegie Institute of Technology, Graduate School of Industrial Administration, 1957.
[31] McCulloch WS, Pitts W. "A logical calculus of ideas immanent in nervous activity". *Bulletin of Mathematical Biophysics*, 5(4), 115-133, 1943.
[32] Haykin S. *Neural Networks: A Comprehensive Foundation*. 2nd ed., New Jersey, USA, Pearson Prentice Hall, 1999.
[33] Vapnik VN. *The Nature of Statistical Learning Theory.* New York, USA, Springer-Verlag, 1995.
[34] Drucker H, Burges CJC, Kaufman L, Smola A, Vapnik V. *Support vector regression machines*. Editors: Mozer, MC, Jordan, MI, Petsche, T. Advances in Neural Information Processing Systems 9, 155–161, Cambridge, MA, USA, MIT Press, 1997.
[35] Huang C-F. "A hybrid stock selection model using genetic algorithms and support vector regression", *Applied Soft Computing*, 12(2), 807–818, 2012.