

Demand Forecast in a Supermarket using a Hybrid Intelligent System

Luis Aburto¹, Richard Weber²

1) DADS S. A., Santiago, Chile; luis.aburto@dads.cl

2) RCAST, University of Tokyo, Japan, rweber@mpeg.rcast.u-tokyo.ac.jp
on leave from: Department of Industrial Engineering, University of Chile;
rweber@dii.uchile.cl

Abstract. Demand forecasts play a crucial role in advanced systems for supply chain management. Determining the future demand for a certain product is the basis the respective systems. Several forecasting techniques have been developed, each one with its particular advantages and disadvantages compared to other approaches. This motivates the development of hybrid systems combining different techniques and their respective advantages. In this paper we propose a hybrid forecasting system combining ARIMA models and neural networks. We show improvements in forecasting accuracy and develop a replenishment system based on the respective forecasts for a Chilean supermarket chain.

Keywords: Neural networks, hybrid systems, demand forecasting, supply chain management

1 Demand forecast in supermarkets

The Chilean supermarket chain Economax, as well as any retailing company, offers a broad range of products (about 5,000 different products) purchased from a large number of manufacturers and distributors. In order to successfully provide such a variety of products to its customers at competitive prices, the supermarket and its providers have to manage efficiently the respective supply chain. Based on the data flow generated by the consumers the supermarket has to decide what, how much, and how often to buy.

In order to solve this problem satisfactorily, a reliable prediction of future demand is necessary. This task, however, presents difficulties, since sales depend on many factors, such as: Past sales, Prices, Advertising campaigns, Seasonality, Holidays, Weather, Sales of similar products, Competitors' promotions, among others.

Chapter 2 of this paper describes related work and the traditional way Economax predicted its sales. Chapter 3 provides a comparison of two techniques for time series prediction (ARIMA and neural networks) and analyzes the respective advantages and weaknesses. Based on this analysis we present a hybrid forecasting system in chapter 4 and show the results of different models for demand forecasts of a certain product. The impact of sophisticated forecasts for inventory and supply chain management is presented in chapter 5. Chapter 6 concludes this work and points at future developments.

2 Related Works

Neural networks are mathematical models that “learn” patterns from data. These networks have proved to be very effective in order to solve classification and regression problems by handling non-linearity between input and output variables, being able to approximate any function under certain conditions [7].

Thanks to the above capacities, these models have been used to solve problems in different areas, such as e.g. time series prediction. It has been shown that these neural models work well in the forecast of stock exchange indexes [9] and corporative bonds [10]. There are also successful applications that have been developed in relation to operations management and have led to huge inventory cost savings [1] [11].

The supermarket we worked with had used so far exponential smoothing and naive prediction in order to estimate future demand, but these approaches did not provide satisfactory results.

3 Descriptions and Comparison of Forecasting techniques

We describe ARIMA models and neural networks for time series prediction and provide a comparative analysis of these two techniques.

3.1 ARIMA models

The problem to predict time series has been solved mainly by applying the ARIMA model family (Autoregressive Integrated Moving Average) proposed by Box and Jenkins [2]. It is defined as:

- X_t is the observation of a time series at time t and has a probability distribution $f(X_t)$
- A is a time series of n white noise observations with average zero and variance σ_A^2
- B is the delay operator. e.g. $BX_t = X_{t-1}$ and $BA_t = A_{t-1}$
- $\nabla = 1 - B$ is the differentiating operator. e.g. $\nabla X_t = (1 - B)X_t = X_t - X_{t-1}$

The ARIMA process (p,d,q) is based on a series that has been differentiated d times, with p autoregressive terms and q mobile average process terms. The respective equation is:

$$\phi_p(B)(\nabla^d X_t - \mu) = \theta_q(B)A_t \quad (1)$$

The result of these models is the continuous μ and the parameter vectors θ_q (moving average) and ϕ_p (autoregressive) that best fit the data.

The process can be generalized even more when incorporating seasonal elements. First, the seasonal differentiating operator is defined: $\nabla_s = 1 - B^s$, where s is the seasonal factor. Besides, the X_t time series can be explained by external variables or predictors (also called regressors). In this way, the most general model is defined by SARIMAX $(p,d,q)(sp,sd,sq)$ Y , where Y are the external variables of the process. Finally, the general equation of the model is:

$$\phi_p(B)\Phi_{sp}(B)\left[\nabla^d\nabla_s^{sd}(X_t - \sum_{i=1}^r c_i Y_i) - \mu\right] = \theta_q(B)\Theta_{sq}(B)A_t \quad (2)$$

where $\Phi_{sp}(B)$ is the sp seasonal autoregressive polynomial, $\Theta_{sq}(B)$ is the sq seasonal mobile average polynomial and c_i are the regressors' coefficients.

3.2 Neural Networks

A neural network is a net of many units, linked by connections [14]. Each unit receives and gives numerical data through the connections. Neural networks generally have some kind of training rule, in which the connection weights are adjusted in accordance with the data that the network receives. In other words, these neural networks “learn” the data and, under certain circumstances, they can generalize beyond the data seen during training, i.e. they can give approximate results for new cases that were not found in their training.

One of the most popular models among neural networks is the Multi Layer Perceptron (MLP) [5], which often is trained with the backpropagation learning rule. This rule minimizes the errors by adjusting the weights in the network. The problem produced in the modeling, though, is related to the fact that the learning process can lead to an overfitting from model to data; i.e. the model learns the received data by heart, losing this way the ability to generalize.

The models mentioned above have been used to solve different problems such as classification, optimization, clustering, and prediction. The present work will focus on prediction using a MLP type neural network. An architecture typically used for this kind of problem is shown in the following figure.

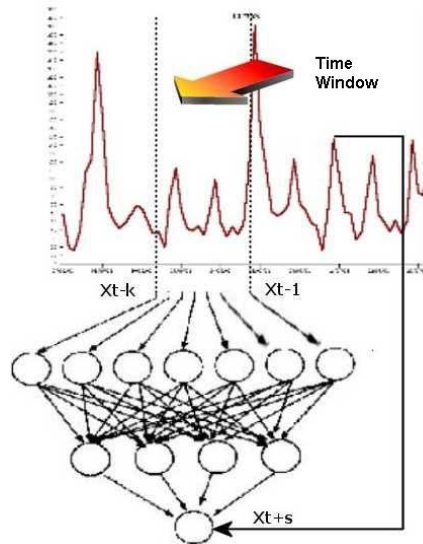


Fig. 1. MLP network for time series forecasting

The architecture used for time series forecast requires two parameters. The first is K , which indicates the length of the time window, which will be used as entrance pattern to predict the time series in the future. The second parameter is s , which represents the number of intervals in which the time series is to be predicted in the future.

3.3 Comparison between ARIMA and neural network models

There are many publications comparing ARIMA and neural models, both theoretically [4][15] and empirically [5][8]. According to Dorffner [4], the main limitation that the ARIMA models have is that they assume a linear relationship between the independent and dependent variables. Besides, as Dorffner [4] and Wan [15] state, MLP networks allow us to model NARX processes, i.e. they are able to model autoregressive non-linear processes with exogenous variables. According to Dorffner, MLP possesses the following advantages:

- By modeling non-linear processes, MLP can represent more complex time series.
- MLP does not assume that the time series to be modeled must be stationary.

However, there are also many advantages the ARIMA models have over neural networks. One of them is the information contained in the model. The ARIMA models allow us to analyze the regressors' coefficients, in this way being able to determine the degree of influence that each one has, in relation to the dependent variable. This is very helpful because it allows us to generate knowledge on which variables are more important in the short-term product sale explanation.

Another disadvantage neural network models have is the high degree of freedom in their architecture. This implies several problems, such as:

- In order to obtain good, reliable results, it is necessary to have a large amount of training examples.
- Having many weights can easily lead to overfitting of the model or provide local minima as result.

As a summary, we present the following table resuming the most important aspects.

Table 1: Comparison between ARIMA and MLP models

ARIMA	Neural Networks (MLP)
Linear Model: assumes <i>a priori</i> behavior of the time series.	Nonlinear model: more degrees of freedom for the model.
Modeling requires the series to be stationary.	Any time series can be analyzed.
Requires interaction with the user.	Requires fewer interactions with the user.
The model provides insight and information through its parameters.	Difficult to interpret the model (black box).
No overfitting.	Overfitting is possible.

4 Development of a Hybrid Forecasting Model

Motivated by the comparison between the two forecasting approaches, it becomes interesting to combine the advantages of both models. We applied a hybrid forecasting system in order to predict demand in the Economax supermarket.

4.1 Additive Hybrid Model

An approach that can be used to solve the problem is to consider the time series as a composition of various series. We represent the original time series with an ARIMA process and the error associated to the forecast as another time series, which shall be modeled by a neural network. The hybrid forecast $\hat{X}(t)$ from the original series is expressed consequently as an addition of an ARIMA process and a neural network model as shown in the following equation.

$$\hat{X}(t) = \hat{Y}(t) + \hat{e}(t)$$

Where $\hat{Y}(t)$ is the forecast from the original series using a SARIMAX (1,0,0) (2,0,0) process (see 3.1 above) that has shown to perform best among ARIMA techniques.

The error of this SARIMAX process has been analyzed as a separate time series and modeled by a neural network. The respective output $\hat{e}(t)$ is the MLP forecast for the errors in the SARIMAX process. The neural network we used has the following architecture:

Input variables: (Past Sales with k lags; dummies variables; prices variables)

Hidden units: 15

Output unit: Present Sale of the product.

Learning rate: 0.3

Momentum rate: 0.1

Stop Conditions: Minimize the RMS error in the Test subset.

4.2 Application of the hybrid model

We applied traditional forecasting techniques, a SARIMAX process, several neural networks and the proposed hybrid system in order to predict demand of the 50 best-selling products in the Economax supermarket. The performance of each technique will be determined by the following two error functions:

Mean Absolute Percentage Error
$$MAPE = \frac{1}{n} \sum_{k=1}^n \left| \frac{(X_k - \hat{X}_k)}{X_k} \right|$$

Normalized Mean Square Error
$$NMSE = \frac{\sum_k (X_k - \hat{X}_k)^2}{\sum_k (X_k - \bar{X})^2} = \frac{1}{\sigma^2 n} \sum_k (X_k - \hat{X}_k)^2$$

Below, we analyze demand forecasts for product 100595 (vegetal oil, 1 liter) in more detail. The variables we used both in the ARIMA models and in the neural networks are:

- Time window with past sales of the product (twk: time window with k days)
- Prices of the product in the past, in the store and from relevant competitors
- Data related to daily characteristics: holidays, end of month, fortnights, among others.

The ARIMA models were constructed using SPSS 8.0 whereas the MLP models were developed with DataEngine 4.0. All proposed models have been compared with the techniques the supermarket currently applies (naive forecast, seasonal naive, and unconditional average). Their performance is evaluated through NMSE (normalized error) and MAPE (percentage error), over both sets (training and test). The results are shown in the following table.

Table 2: Results from different forecasting approaches for product 100595

	100595	Training set		Test set	
		Percentage Error	Normalized Error	Percentage Error	Normalized Error
M1	Naive	44.28%	0.6972	56.83%	1.2481
M2	Seasonal Naive	64.67%	1.2212	45.75%	1.9217
M3	Unconditional average	59.98%	0.7759	48.54%	0.9689
M4	SARIMAX(1,0,0)(2,0,0)	36.21%	0.3301	40.49%	0.6090
M5	MLP-tw21	32.93%	0.4633	31.85%	0.4973
M6	MLP-tw14	31.15%	0.3115	34.64%	0.5703
M7	MLP-tw3	29.61%	0.3002	34.36%	0.5281
M8	MLP-tw1	30.00%	0.3405	35.31%	0.5340
M9	MLP-tw21 with SARIMAX	26.12%	0.2760	28.80%	0.3544

As can be seen SARIMAX (model M4) and neural networks (M5, M6, M7, M8) outperform traditional techniques (models M1, M2, M3). The additive hybrid model (M9) gives best results among all approaches employed.

5 An inventory control system based on demand forecasts

Based on the proposed hybrid forecasting model we suggested a system for inventory replenishment in the supermarket chain. Replenishment from the suppliers is done for most products every P days and the purchase order has to be sent at least L days before the delivery date.

The desired inventory level (T) has to be fixed every period. This is calculated by the equation $T = m' + Z\sigma$. Where m' is the average demand during $P+L$ days and $Z\sigma$ is the security stock, which depends on the desired service level (Z) and on σ , standard deviation of the demand during $P+L$ days.

The benefit of the short-term forecast based on the proposed hybrid system is the dynamic estimation of the average demand during the period between orders (m'). Using sales data from one year we simulated the inventory level for product 100595 applying the replenishment model mentioned before. The following figure shows for this product (vegetal oil, 1 liter) the real inventory level and the desired inventory level (upper line).

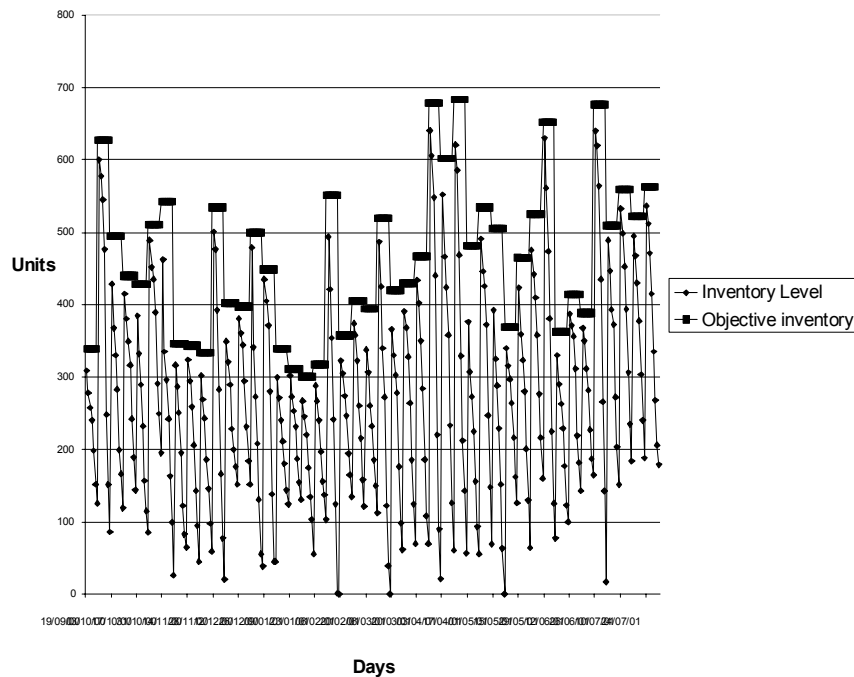


Figure 2: Product 100595 Daily Inventory Level Graph

Comparing the current situation (average) to the results obtained by the proposed model, there are improvements both in the customer service level (measured by sales failure) and in inventory level (measured as inventory/sales average). The results are summarized in the following table.

Table 3: Performance Comparison between the current system and the system proposed for replenishment.

Inventory Management Control indicator	Current System	Proposed System
Reaching days (Inventory /sale average)	30 days	5 days
Sales failures (% of days without products)	6%	0.9%

6 Conclusions

The developed forecasting models leave a considerable amount of valuable information for the business. Using these models it is possible to quantify the effect of every event (holidays, end of month, etc.) in the behavior of every product purchase. Regarding forecast accuracy, neural network outperformed ARIMA models and the proposed additive hybrid model gave best results.

Better short-term forecasts allow the Economax supermarket chain to reduce inventory costs and improve their operation margins, in this way achieving a competitive advantage in the supermarket business.

Acknowledgment: The project described in this paper has been funded by the Millenium Nucleus on Complex Engineering Systems.

References

- [1] K. Bansal, S. Vadhavkar and A. Gupta. Neural Networks Based Forecasting Techniques for Inventory Control Applications. *Data Mining and Knowledge Discovery* 2, pp. 97-102. 1998.
- [2] G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco. 1976.
- [3] C. Brax. *Recurrent Neural Networks for Time Series Prediction*. Department of Computer Science, University of Skövde. Skövde, Suecia. 2000.
- [4] G. Dorffner. Neural Networks for Time Series Processing. *Neural Network World*. 6 (4) pp. 447-468. 1996.
- [5] J. Faraway and C. Chatfield. Time series forecasting with neural networks: a comparative study using the airline data. *Applied Statistics* 47 (2), pp. 231-250. 1998.
- [6] J. Han and M. Kamber. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers. San Francisco. 2001.
- [7] T. Hill, M. O'Connor and W. Remus. Neural Networks for Time Series Forecasts. *Management Science*, 42 (7), pp. 1082-1092. 1996.
- [8] C. Kuo and A. Reitsch. Neural networks vs. conventional methods of forecasting. *The Journal of Business Forecasting Methods and Systems*, 14(4), pp. 17-22. 1995.
- [9] P. G. McCuskey. *Feedforward and Recurrent Neural Networks and Genetic Programs for Stock Market and Time Series Forecasting*. Technical Report CS-93-36. Brown University, USA. 1993.
- [10] J. Moody. *Prediction Risk and Architecture Selection for Neural Networks*. From Statistics to Neural Networks: Theory and Pattern Recognition Applications. 1994.
- [11] C. C. Reyes-Aldasoro, A. Ganguly, G. Lemus and A. Gupta. A Hybrid Model Based on Dynamic Programming, Neural Networks and Surrogate Value for Inventory Optimisation Applications. *Journal of Operational Research Society*, Vol.50, No.1, pp. 85-94. 1999.
- [12] R. G. Schroeder. *Administración de Operaciones*, University of Minnesota. Mc Graw Hill. 1992.
- [13] K. Smith and J. Gupta. Neural networks in business: techniques and applications for the operations researcher. *Computers and Operations Research*, volume 27, number 11-12, pp. 1023-1044. 2000.

- [14] F. M. Thiesing and O. Vornberber. Sales Forecasting using Neural Networks. Proceedings International Conference in Neural Networks 97, Houston. Vol. 4 pp. 2125-2128. 1997.
- [15] E. A. Wan. Finite impulse response neural networks for autoregressive time series prediction. In A. Weigend and N. Gershenfeld, editors, Predicting the Future and Understanding the Past, volume XVII. Addison-Wesley, MA. 1993.