Modelling wholesale distribution operations: an artificial intelligence framework

Modelling Wholesale Distribution Operations: An Artificial Intelligence Framework

Abstract

Purpose. This paper proposes an artificial intelligence-based framework to support decision making in wholesale distribution, with the aim to limit wholesaler out-of-stocks by jointly formulating price policies and forecasting retailer's demand.

Design/methodology/approach. The framework is based on the cascade implementation of two artificial neural networks (ANNs) connected in series. The first ANN is used to derive the selling price of the products offered by the wholesaler. This represents one of the inputs of the second ANN that is used to anticipate the retailer's demand. Both the ANNs make use of several other input parameters and are trained and tested on a real wholesale supply chain.

Findings. The application of the ANN framework to a real wholesale supply chain shows that the proposed methodology has the potential to decrease economic loss due to out-of-stock occurrence by more than 56%.

Originality. The combined use of ANNs is a novelty in supply chain operation management. Moreover, this approach provides wholesalers with an effective tool to issue purchase orders according to more dependable demand forecasts.

Keywords: *Artificial Neural Networks (ANNs); Demand forecasting; Multiple Neural Networks (MNNs); Price determination; Supply chain management; Wholesale distribution.*

1 Introduction

In the wholesale distribution of consumer goods, the role of a wholesaler is to provide its customers with a broad assortment of products from different suppliers, which in most cases are large companies managing business relationships exclusively with their key distributors (Figure 1). In such a system, the wholesaler renders the service of making the desired products available for purchase at a specific time, which creates value for the customer (Ehrenthal et al. 2014). The primary focus of the wholesaler is to maximize the difference between what retailers will pay and what vendors will accept as payment for the products they offer. The distributors typically have a strong bargaining power, which enables them to obtain good contractual conditions with industries, with profit margins up to 6% (Holweg et al. 2016; Wyman 2014).

Defining the amount of product to be kept in stock is a crucial issue for wholesalers, as it leads to inventory management and purchasing decisions, which, if not managed properly, can involve dramatic consequences (Fisher 2009). Inventory management inaccuracies have a significant negative impact on the performance of all the actors in the supply chain (Sari 2008). In some cases, buyers order more than the actual demand due to poor order forecasting, thus issues related to unsaleable products arise (Lee et al. 2006; Holweg et al. 2016). In other cases, incorrect purchasing decisions can lead to out-of-stock (OOS) occurrence, which is a serious issue in supply chains (Roland Berger Consultants 2003; Supermarket Guru Consumer Panel 2011). Indeed, although OOS is mainly observed at the retail stores, supplier related issues contribute to approximately 10-30% of OOS situations (Aastrup & Kotzab 2009). From an economic perspective, OOS cause lost sales as an immediate consequence, but in the long-term dissatisfy shoppers, diminish store loyalty and jeopardize marketing efforts. Therefore, reducing OOS provides retailers with the opportunity of increasing sales and reducing cost (Fernie et al. 2010).

Demand forecasting issues were found to cause approximately 47% of OOS situations in retail (Gruen et al. 2002). An accurate forecasting of retailers' demand largely affects the successful inventory management of the wholesaler, while the inability to match supply with demand is one of the biggest obstacles to supply chain excellence (Fliedner 2001; Esper et al. 2010; Stank et al. 2011). The demand forecast is often grounded on incomplete information and thus tends to underestimate the real demand generating OOSs. In addition, the demand and sales forecasts invariably differ due to sales variances caused by OOS occurrence; this means that OOSs, in turn, contribute to an inaccurate demand forecasting process (Gruen & Corsten 2007). Conversely, improved demand forecasting accuracy results in monetary savings, greater competitiveness, enhanced channel relationships, and customer satisfaction (Moon et al. 2003).

Wholesalers manage two crucial activities that determine the flow of purchased products from the supplier(s) to the wholesaler and of sold products from the wholesaler to the retailer(s), and are, therefore, closely related to the demand forecasting process and to the occurrence of OOS situations. These activities are:

i. the formulation of the price list of products to be sold to the retailers;

ii. the issuance of purchase orders to its suppliers.

With respect to point (i), when planning their product range, wholesalers should decide which products to offer in each store and at each time, as well as products to be promoted weekly, including price discounts or special promotions (e.g., a 3 for 2 promotion) (Kumar et al. 2016; Holweg et al. 2016; Willart 2015). The definition of the selling prices, as well as the choice of the products to be sold through sales promotions, typically depends on several factors, including the available inventory, purchasing price, and product's turnover rate. As the selling price of a product is expected to influence the resulting demand to a relevant extent (Willart 2015), it has the potential to indirectly affect the occurrence of OOS situations. In retail, this is exacerbated by sales promotions, which are among the factors that affect consumers' behavior to the greatest extent (Laroche et al. 2005). However, the impact of the selling price on customer's demand is not linear and, in any case, can vary depending on the considered product (Gruen & Corsten 2007). With respect to point (ii), for each item inventory managers are to decide when to purchase, how much to purchase and from whom to purchase, to meet the future customer's demand (Bala 2012). However, the future demand of an item can depend on a large number of factors, which makes it challenging for the retailers to issue a purchase order that will reflect that demand (Bala 2012). Besides the presence of sales promotions, sometimes the wholesaler can be forced to buy products to ensure in-stock availability of the entire assortment offered to its customers (Kök & Fisher 2007). Other factors, such as the purchasing price, lead time and current inventory level, could affect the wholesaler decision on the purchased product quantity, ultimately affecting OOS occurrence.

The ideal case of a perfect match, in time and quantity, between the flows of purchased products and sold products could be somehow achieved by sharing information between wholesaler and retailer. However, although sharing information is acknowledged as a best practice among the top supply chains, this practice is still not widely diffused among retail channel members (Banerjee & Mishra 2017). More frequently, to synchronize these flows, the retailer's demand should be anticipated, by defining in advance the number of products to buy and avoid OOS occurrence. Hence, a good demand forecasting turns out to be essential.

In line with these premises, this paper proposes and implements an approach based on a particular type of artificial intelligence (AI) tool, i.e. artificial neural networks (ANNs), to predict the retailer's demand in wholesale supply chains. Although the use of ANNs is not new in demand forecasting, this paper contributes to the existing literature in a twofold manner.

First, before being used for demand forecasting, the proposed ANN tool supports the process of formulating product selling prices. This is a relevant aspect for reliably predicting the final customer's demand in the wholesaling context, where products are subject to a certain price elasticity (Sugandi et al. 2016; Syntetos et al. 2016). Second, the proposed approach makes use of two ANNs connected in series, in line with the-need for modeling two subsequent processes. The output of the first network is used as one of the inputs of the second network; this is a further new point compared to the existing literature. The forecast precision achievable using two ANNs is evaluated exploiting real data related to a wholesale distributor.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature, focusing on the studies involving the combined use of ANNs in a cascade and on the use of ANNs in the demand forecasting domain. Section 3 describes the wholesaling context and the relevant variables influencing the estimation of the final customer's demand. The proposed ANN framework is described in section 4, while section 5 evaluates its performance in the case of a real wholesaler. Section 6 summarizes the key findings and outlines future research directions.

Insert Figure 1 Here

2 Literature review

In supply chain management, a typical approach to demand forecasting involves the use of statistical tools, in the form of exponential smoothing, time series regression and autoregressive and integrated moving average (ARIMA) models (Fildes et al. 2009; Silva Filho et al. 2013; Babai et al. 2013; Khosroshahi et al. 2016; Huber et al. 2017). These models all assume linear relationship between independent and dependent variables, which is not always the case for demand forecasting in retail distribution (Gruen & Corsten 2007). ANNs fall within the class of non-linear methods and are, therefore, extensively used in the demand forecasting process (Bala 2012; Cho & Kim 1995; Kittler et al. 1998). ANNs are mathematical tools inspired by biological neural networks that provide, for a given problem, solutions similar to those formulated by the human mind. Multiple neural networks (MNNs), i.e. the connection of more ANNs, were found to be particularly suitable for improving the accuracy of the model prediction (Nguyen & Chan 2001). Connecting more ANNs means considering the input data of the central net to be the output of the preceding networks.

A review of the relevant literature (Table 1) shows that the use of ANNs for demand forecasting problems is not new and has been proposed in different contexts. Among recent works, Mahbub et al. (2013) developed an ANN model to forecast the optimum demand as a function of several factors, including the time of year, festival period, promotional programs, holidays, number of advertisements, cost of advertisements, number of workers and availability. Ferreira et al. (2016) developed a methodology to forecast demand in an order treatment center using ANNs. The methodological approach makes use of an ANN multilayer perceptron, which is trained using an error back-propagation algorithm. Aizenberg et al. (2016) developed a multilayer neural network with multi-valued neurons for long-term time series forecasting in the context of oil production.

Some studies have proposed integrated models where ANNs is combined with other tools to make the demand prediction more effective. Among these studies, Kabir & Hasin (2014) proposed a forecasting model that exploits the adaptive neuro-fuzzy inference system (ANFIS) techniques to manage the fuzzy demand with incomplete information. ANFIS is utilized to combine the power of fuzzy logic with ANNs and specifically exploits the mathematical properties of ANNs in tuning rule-based fuzzy systems that approximate the way humans' process information. Jaipuria & Mahapatra (2014) proposed an approach that integrates discrete wavelet transforms analysis and an ANN to improve demand forecasting and to reduce the bullwhip effect in a supply chain. Singh & Challa (2016) proposed a forecasting methodology using nonlinear customer demand in a multilevel supply chain. The approach consisted of a combination of discrete wavelet theory, ANNs, and an adaptive-network-based fuzzy inference system.

Further studies offered a comparison of ANN with more traditional tools for demand forecasting; to be more precise, Kandananond (2012) and Shahrabi et al. (2009) compared ANNs with ARIMA models, moving average, exponential smoothing and exponential smoothing with trend. Benkachcha et al. (2014) and Lu et al. (2012) have shown that ANN models provide more accurate forecasting compared to multiple linear regression and other AI tools. Azadeh et al. (2011) examined the effect of the number of inputs, hidden nodes, hidden layers as well as the size of the training sample on the in-sample and out-of-sample performance of an ANN used for demand forecasting. They also considered a new forecasting approach based on regression as a benchmark for comparison with ANN.

Besides ANNs, other techniques used for demand forecasting include agent-based models (Liang & Huang 2006), hybrid modes (Aburto & Weber 2007), non-linear machine learning tools (Carbonneau et al. 2008), wavelet transforms (Ferbar et al. 2009), Bayesian networks (Wang et al. 2014), martingale models (Boulaksil, 2016) nearest neighbor approaches (Nikolopoulos et al. 2016) and genetic algorithms (Sakhuja et al. 2016).

The review above demonstrates that although all the studies reviewed focus on demand forecasting, none of them explicitly target the issue of formulating the product selling price, which, nonetheless, do affect the resulting demand. In addition, from a technical perspective to the best of our knowledge, no studies explored the potential of using MNNs or ANNs in cascade for demand forecasting. These gaps will be addressed in this paper. Table 1 highlights the differences between our study and the earlier researches in the literature.

Insert Table 1 Here

3 Problem formulation

3.1 The context

This study targets the wholesale distribution. The company analyzed is a leading distributor of fast moving consumer goods in central and southern Italy and specializes in health and beauty care products. The company is part of a consortium, which, in turn, is a member of a holding corporation. As a part of a large distribution group, the company enjoys high bargaining power against its suppliers. The company issues purchase orders to these suppliers and receives direct deliveries from them. The company's annual turnover is approximately €50 million, with approximately 8,000 different handled products.

3.2 Determination of the selling price

The relevant factors affecting the final price of the products vary depending on the type of product considered (Burnett 2003; Bala 2012); in line with this consideration, some managers of the wholesaling company under examination were asked to provide a list of the key factors affecting the selling price of their products. Suggestions from the literature, when available, were also used to confirm the identification of these key factors. The factors listed and described below emerged as the most relevant for determining the selling price of the products sold.

1. Purchasing cost. It has an obvious direct impact on the selling price: if the wholesaler can buy a specific product from its suppliers at a very favorable cost, it is able to propose the same product to its customers at a competitive price;

2. Mark-up rate. To determine the selling price, the wholesaler typically applies a mark-up rate to the purchasing cost (Burnett 2003). In the considered context, the mark-up oscillates from 6% to 18% depending on the product or on possible variations in the purchasing cost;
3. Stock Turnover Rate (STR). The STR provides useful indications about the sales trend of a product in a given time period. High STR can either indicate that the product is kept in stock for a relatively short time or that the available inventory will decrease shortly and needs to the restored. Conversely, low STR can either suggest that the product has a low demand or that a too high price prevents the possibility of selling the product. To correctly distinguish these situations, it is useful to couple the data about the STR with the inventory level of the product;
4. Average selling price. It is included as the input of the ANN to ensure that the estimated selling price is consistent with the previous ones. In other words, such an input should prevent forecasting an 'unfeasible' price (e.g., too low or too high compared to the market trend);
5. The economic value of the stock. It is computed as the product of the inventory level and the purchasing cost. Determining the selling price of a product requires knowing the economic value of the stock: the high stock of a product could suggest decreasing the selling price to encourage purchase from the final customers;
6. Purchased quantity. It can influence the selling price both directly and indirectly. Purchasing a larger volume of a given product typically involves quantity discounts, meaning that more favorable purchasing costs are likely to be obtained; this variation in the purchasing cost indirectly affects the selling price (Burnett 2003). Conversely, buying more products generates a relevant stock level; the wholesaler could decide to sell out these products by offering them at a more competitive selling price to recover capital.

3.3 Demand forecasting

The key factors affecting the retailer's demand were again identified by integrating the suggestions from the managers of the wholesaling company with the findings available in the literature. The elements listed and described below emerged.

1. Selling price. In the targeted industry, the selling price is expected to be the factor that affects the final product demand to the greatest extent. Indeed, for consumer goods for which shelf life is not an issue, customers are very price sensitive and tend to buy the product in a speculative way;

2. Previous demand value. It reflects the auto-regressive component of the customer's demand. The weekly demand value is used;
3. Average demand. It is computed over a time span of one year using the weekly demand data and is included as input to prevent the network from generating a forecast that is too far from the average value observed during the last year;
4. Promotion. In the presence of a sales promotion, the customer's demand tends to grow as it is triggered by lower prices (Inman et al. 1990; McClure & West 1969). Sales promotions also encourage consumers to purchase non-promotional products (Mulhern & Padgett 1995) and accelerate the number of shopping trips to the store (Walters & Rinne 1986);
5. Backordered quantity. This input takes into account the situations in which the demand forecast could be biased by a previous 'failure to deliver'.

4 ANN modelling and software implementation

The effect of the factors listed above on the selling price and retailers' demand is non-linear and hard to describe using mathematical relationships (Wei & Li 2015). Moreover, besides these factors, the retailer's demand is also characterized by an intrinsic variability (İşlek & Öğüdücü 2015). These considerations motivate the adoption of ANNs.

A graphical representation of the proposed framework is shown in Figure 2. ANN#1 in Figure 2 is used to predict the selling price of a specific product, while ANN#2 aims at forecasting the retailer's demand for the product. The output of ANN#2, reflecting the amount of product to purchase, is finally used by the wholesaler to issue orders to its suppliers (dashed line in Figure 2). To prove the effectiveness of the proposed approach, two scenarios are considered for implementation, namely:

1. The situation in which the two ANNs are used separately (Figure 2-a). In this situation (referred to as *scenario 1*), ANN#1 predicts the selling price and ANN#2 forecasts the retailer's demand;
2. The cascade implementation of the two ANNs (Figure 2-b). In this situation (*scenario 2*) the output of ANN#1 is one of the inputs of ANN#2.

The two scenarios differ in the selling price set as an input in ANN#2. In scenario 1, it reflects the real selling price applied by the wholesaler, while in scenario 2, the selling price is determined by ANN#1 as a function of its input parameters. As it makes use of the real input and output data, scenario 1 is primarily used to tune the parameters of the ANNs and to identify the configuration of both networks that returns the best forecasting performance. Conversely, scenario 2 has been introduced to investigate whether an accurate definition of the selling price (as returned by ANN#1) has potential to help the wholesaler forecast the retailer's demand and, thus, decrease the occurrence of OOS situations. Scenario 2 is also useful to assess whether the cascade implementation of ANNs is more effective in predicting the retailer's demand than the use of separate ANNs. The capability to decrease the occurrence of OOS situations at the wholesaler will be used as a performance parameter to evaluate the framework effectiveness in both scenarios.

Insert Figure 2 Here

The ANNs were modelled in MatLab R2015b, according to the pathway shown in Figure 3, and run on a personal computer with an Intel® Core i5 processor and 4GB RAM. As this pathway is the same for both ANNs, for the sake of brevity its implementation is detailed only for ANN#1, while the corresponding description for ANN#2 is more succinct.

Insert Figure 3 Here

4.1 ANN#1: determination of the selling price

As shown in Figure 3, a preliminary step in the development of the ANN consists in analyzing the problem to be faced and in identifying the inputs of the ANN (*Problem formulation*). For ANN#1, this means identifying the factors that affect the selling price of the products offered by the wholesaler. The corresponding analysis, already proposed, showed that the key inputs of ANN#1 are purchasing cost, mark-up rate, STR, average selling price, economic value of the stock and quantity purchased.

Step 2 (*Choice of the inputs and outputs of the ANN*) consists of selecting the relevant inputs for ANN#1 among those identified. As it is not immediate to detect the parameters that are more relevant to the definition of the selling price, four different configurations of ANN#1 were analyzed, with a different combination of the input parameters.

The necessary data related to the input and output previously selected for ANN#1 needs to be collected (*Data collection*). To be more precise, the first network should be trained with historical data to learn the buyer's logic to determine the products' selling price in order to predict this price for the next period with good precision. Indeed, the selling price is the final decision of the wholesaler, which, in addition to considering the input factors mentioned, could also take into account personal assessments.

The next step (*Network setting*) consists in selecting the algorithm for network training, the transfer function and the termination criterion. A feed-forward back-propagation algorithm for training and, specifically, a fitting neural network (FNN) was selected, as it is suitable for implementation when estimating the value of a given output on the basis of a set of inputs and target values (Mathworks 2008). More precisely, the Levenberg-Marquardt training function (Levenberg 1944; Marquardt 1963) was used; this approach initializes the weights of the network according to the Levenberg-Marquardt optimization and corrects them during training using a back-propagation procedure. A "tansig" function was used as the transfer function from the inputs to the hidden layer(s), while a "purelin" function was used as the transfer function from the hidden layer(s) to the output (Efendigil et al. 2009; Ali et al. 2011). We set the following termination criterion for the training phase: the network training stopped when the validation error grew for six consecutive iterations (Mathworks 2008).

A typical measure for the forecasting accuracy is the mean absolute percentage error (MAPE) (Mentzer and Bienstock 1998) and is also widely used by companies (McHugh and Sparkes 1983; Dalrymple 1987; West 1994). The MAPE reflects the differences between the target values (y_t, $t = 1, \dots n$) and values predicted by the ANN ($\hat{y}_t$) divided by the actual value, i.e.:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad [\%] \qquad (1)$$

where n denotes the number of outputs. A low MAPE demonstrates the ability of the ANN networks to produce accurate estimates.

The last step (*Choice of the hidden nodes and layers*) consists in setting the number of layers and hidden neurons. There are no general methods that can provide the optimal setting of an ANN architecture (Zhang et al. 1998); therefore, we elected to set a single hidden layer and to vary the number of neurons following a "trial and error" procedure. The number of output neurons is always set at 1. For networks with one single hidden layer, it is advisable to try the following numbers of neurons: $2n + 1$, $2n$, $n/2$ and $n + m$, where n and m represent the number of inputs and outputs respectively (Lawrence & Fredrickson 1998; Ali et al. 2011). Following these suggestions, we tested the ANN#1 with 2, 6, 8, 10 and 14 hidden neurons. The analysis was repeated for the four network configurations mentioned previously, obtaining 5×4=20 combinations in total.

The most appropriate setting in terms of inputs and neurons was determined from the analysis of the network capability to return an effective forecasting performance (*Parameter setting*). To this extent, some simulations were launched using Matlab™ to derive the selling price of the products and to select the network configuration that returned the lowest MAPE. To compute the MAPE, the vector of the outputs provided by the network during the testing phase was compared with the corresponding target values derived from a subset of the previously selected dataset. Such a subset was excluded from the training phase. On the basis of the resulting MAPE, some corrections were made regarding the network configuration (i.e., to the number of hidden layers - *correction #1*) or the inputs considered (*correction #2*).

The minimum percentage error and maximum percentage error, reflecting the minimum and maximum deviation between network prediction and the corresponding target value, were also used as complementary parameters to assess the network performance.

4.2 ANN#2: demand forecasting

For ANN#2, the preliminary step (*Problem formulation*) consists in identifying the factors affecting the retailer's demand for the final products offered by the wholesaler. For the case under investigation, these factors were listed and described in section 3.3. Five different network configurations were identified for ANN#2, which differ in the number and combination of the factors chosen as the input *(Choice of the inputs and outputs of the ANN*).

In *Data collection* step, we collected historical data related to all the ANN's input factors. These data were used differently in the two implementation scenarios:

- In scenario 1, ANN#2 is trained using historical data (in particular, the selling prices retrieved from the real wholesaler's data) and predicts the output based on these data;
- In scenario 2, ANN#2 is trained and tested by using as input the selling price resulting from ANN#1.

The *Network setting* of ANN#2 does not vary significantly compared to ANN#1, as ANN#2 was also trained using a feed-forward back-propagation algorithm. The transfer function and termination criterion are the same as those of ANN#1. The architecture of the network is similar too, as ANN#2 consists of a single hidden layer, one output neuron and a variable number of hidden neurons (*Choice of the hidden nodes and layers*). This latter was set at 2, 6, 8, 10 and 14 and the best configuration was defined using a “trial and error” procedure. Once the network was configured, some simulations were launched to derive the retailer’s demand. The results were used to assess the capability of the specific network architecture to return an effective forecast (*Parameter setting*), as well as to identify and select the best configuration in terms of MAPE, minimum percentage error and maximum percentage error.

4.3 Economic evaluation

For ANN#1, MAPE was used as the main performance parameter to assess the network effectiveness, as the formulation of the selling price was not forecasted in its literal sense but rather as an estimate of a value that was finally determined by the wholesaler. Conversely, the ANN#2 is expected to return a true demand forecast, whose ultimate aim is to help the wholesaler limit the occurrence of OOS situations. Therefore, its performance was evaluated also from a quantitative perspective; specifically, a cost/benefit analysis was carried out to assess the results that are achievable when using the ANN framework for demand forecasting. It is easy to realize that the prediction error of ANN#2 can be in excess or in the defect, reflecting an overestimate or an underestimate of the retailer’s demand. In the latter case, the amount of product purchased by the wholesaler will not be sufficient to satisfy the real demand, resulting in an OOS situation and lost sales. To evaluate the benefits resulting from the use of the proposed approach, the actual economic losses experienced by the targeted wholesaler because of OOS occurrence was compared to the losses resulting if the ANN framework was used to predict the customer’s demand in the same period. The underestimate of the retailer’s demand, although causing an extra cost of holding stocks, is by far less critical than the occurrence of OOS situations and has neglected in the chosen context, because of the following reasons:

- The goods are not perishable and are fast moving items. Hence, a higher stock will be easily sold out in a relatively short time;
- The wholesaler under consideration has an extra storage capacity, which allows storing the number of overstock products. Therefore, the cost of holding a higher stock is marginal.

Despite the fact that the ANN performance is evaluated only based on the cost of OOS, it should also be remarked that the model would never predict the demand to be infinite, thanks to the presence of the "previous demand value" and "average demand" among the inputs of the ANN. The cost/benefit evaluation was carried out separately in scenarios 1 and 2.

5 Application of the ANN framework to the case study

5.1 Data collection

The wholesaler's information system was queried to retrieve the data needed to train and test the two ANNs. The data collected refer to the whole set of inputs and outputs of both ANNs plus the occurrence of OOS situations and the corresponding extent of the OOS. As mentioned, the wholesaler handles approximately 8,000 different products. For the application of the proposed approach, the "health and beauty care" category (970 products), representing the wholesaler specialization, was selected as the pilot class of products. A preliminary Pareto analysis was carried out on this class to identify the most critical products in terms of turnover and OOS occurrence. A subset of 208 products (21.44% of the total amount, 73.50% of turnover) was identified as particularly critical for the targeted wholesaler. These products vary significantly in terms of price, demand and average stock level; therefore, the chose subset is appropriate to test the effectiveness of the ANN framework. The data related to these products were retrieved over a one-year time span, i.e. from mid-2015 to mid-2016. Overall, the dataset consisted of 208 (products) × 12 (months) = 2,496 records for each input and output of the ANNs.

With respect to the inputs of ANN#2, the data retrieved from the wholesaler's information system did not explicitly spot sales promotions or special price policies applied in 2015-2016. A preliminary elaboration on the available data was therefore made to identify the presence of sales promotions, by analyzing the deviation of the selling price from its average value and attempting to relate this deviation to the demand trend.

As far as the training, validation, and testing steps are concerned, the allocation of the dataset to these steps is quite arbitrary (Konar 2006); in the case under examination, the following approximate percentages were applied: training subset = 80%; validation subset = 10%; and testing subset = 10%.

5.2 Results for ANN#1

Table 2 summarizes the four different configurations of the input parameters analyzed for ANN#1. The test of each configuration with a different number of neurons in the hidden layer is shown in Table 3, in terms of the MAPE, minimum percentage error, and maximum percentage error.

Insert Tables 2 & 3 Here

From Table 2 it can be seen that regardless of the configuration and number of neurons in the hidden layer, the minimum percentage error scores zero almost always, with the only exception being network configuration 4 with 2 hidden neurons for which the minimum percentage error was 0.07. Therefore, more attention has been paid to the analysis of the maximum percentage error, i.e. the maximum deviation between the network prediction and the corresponding target value. The comparison of the different network architectures in terms of the minimum and maximum percentage error shows that network configuration 3 with 6 input neurons and 14 hidden neurons is the most effective in terms of forecasting performance. This configuration is able to determine, in the testing phase, the selling price with a MAPE of 2.06%, a minimum prediction error of 0.00% and a maximum prediction error of 4.44%.

We recall that the wholesaler determines the selling price of products by applying a mark-up rate to the purchasing cost, ranging from 6% to 18% as a function of the specific product. As the range of the variability of the mark-up is quite large, a maximum error of 4.44% is tolerable from a practical perspective. This conclusion is also supported by the literature: an in-depth study on the price forecast of tomatoes showed that the application of a neural network with back-propagation returned a price forecast with an error ranging from 3.90% to 7.50% (Yu & Ou 2009).

5.3 Results for ANN#2

Table 4 summarizes the five different configurations of ANN#2, in terms of input parameters and neurons. Each of these configurations was tested using 2, 4, 8, 10 and 14 neurons in the hidden layer and the related performance is reported in Table 5.

Insert Tables 4 & 5 Here

The results show that the minimum percentage error of ANN#2 ranges from 0.05% to 2.10%, regardless of the configuration and the number of neurons in the hidden layer. Conversely, the maximum MAPE is quite relevant and varies significantly among the configurations. To be more precise, it always exceeds 35% in the last two configurations, with a peak of 52.48% in network configuration 5. Therefore, the overall forecasting performance of these network configurations is not satisfactory. Looking at the first three configurations, the lowest maximum percentage error is obtained using 6 (configuration 2) or 8 (configurations 1 and 3) neurons in the hidden layers. Overall, network configuration 3 with 8 neurons in the hidden layer returns the best predictive performance, with minimum and maximum percentage errors of 0.02% and 14.27% respectively.

Evaluating the different network architectures allows for the derivation of some key findings. Network configurations 4 and 5, whose forecasting performance is not satisfactory, make use of the "backorder quantity" as an input. Hence, it could be argued that this input parameter probably disturbs the network and causes an inaccurate forecast. This is also confirmed by the fact that network configuration 3 with 8 hidden neurons, that yields the best forecasting performance, does not make use of the "backorder quantity" as an input. Conversely, the "average demand" is a useful input for the ANN, as deduced by comparing the performance of network configurations 2 and 3. As the ANN takes into account 208 different products simultaneously, it is reasonable to expect that network "recognizes" a specific product based on its average demand, which is almost the same in every period, as well as on the previous demand, thus justifying the impact of these factors on the network performance.

5.4 Economic evaluation

To evaluate the performance of the ANN framework in economic terms, the loss/saving resulting when the ANN is used to predict the demand for the 208 products tested was computed. To this end, the cost of the OOS that occurred in the real scenario was compared to that resulting from the ANN prediction. This cost was computed as follows:

$$OOS_{cost} = \sum_i p_i * OOS_{q_i} \qquad (2)$$

Where

- i $(= 1 \dots .208)$ is the product tested;
- p_i is the selling price of a unit[1] of product i [€/unit]. Depending on the scenario considered, this is can be either the real selling price of the product (i.e. the historical value) or the value forecasted by ANN#1;

[1] We use the generic term "unit" to denote the set of items that can be resold. Depending on the product considered, this can refer to an item, a lot of 250 pieces or a full pallet.

- OOS_{q_i} is the amount of OOS of product i [units].

With respect to OOS_{q_i}, in the real scenario, the OOS occurrence and related amount are known, as both data were retrieved from the wholesaler's information system. Conversely, when the ANN is used to predict the retailer's demand, OOS_{q_i} should be computed as the difference between the real demand (d_{real_i}) and demand predicted by the ANN (d_{pred_i}), with d_{real_i} being available as data collected from the wholesaler's information system. Eq.3 is used for the computation:

$$OOS_{q_i} = d_{real_i} - d_{pred_i} \begin{cases} > 0 \; if \; d_{real_i} > d_{pred_i} \\ \leq 0 \; if \; d_{real_i} \leq d_{pred_i} \end{cases} \tag{3}$$

Obviously, OOS situations occur only if the prediction is less than the real demand (i.e., $d_{real_i} > d_{pred_i}$). Accordingly, OOS_{cost} will be computed when $OOS_{q_i} > 0$. The computation of OOS_{cost} has been carried out for all forecasting periods (12 months); for the sake of brevity, we limit the presentation of the outcomes to the worst situation, i.e. the month in which the proposed framework provides the lowest forecasting performance.

5.4.1 Scenario 1

The main economic outcomes in the case of a separate use of the two ANNs are reported in Table 6. In this table, zero denotes the case in which the OOS situations did not occur. As Table 6 shows, the economic loss due to an OOS occurrence in the real scenario is greater than the loss resulting when ANN#2 is exploited to forecast the retailers' demand. From the numerical outcomes (48,031 €/month vs. 28,358 €/month), it can be estimated that the use of the ANN reduced economic loss due to OOS by approximately 40.96% in a sample month; the economic savings could be greater if considering the whole year.

Insert Table 6 here

5.4.2 Scenario 2

To assess the effectiveness of the cascade implementation of the ANNs, we first compare the forecasting performance of the ANN framework under scenarios 1 and 2. The outcomes, shown again in Table 6, indicate that regardless of the scenario considered, the use of the proposed framework slightly underestimates the real demand, although the forecasting error is lower in scenario 2 than in scenario 1 (-1.095% vs. -0.479%).

Despite the underestimate of the real customer's demand, the proposed framework does not increase the occurrence of OOS situations: indeed, the economic analysis shows that, overall, the sale loss due to OOS occurrence when using the cascade ANN framework is significantly lower than that observed in the real case. The aggregated sales loss resulting from the cascade ANN implementation accounts for 21,008.89 €/month for the whole set of 208 products (vs. 48,031.21 €/month in the real scenario). This suggests that the ANN framework tends to provide a more precise forecast for high-value products, while for low-value ones, the forecast is probably less accurate.

Overall, the aggregated outcomes show that the use of the two connected ANNs allows the economic loss to be reduced by more than 56%, which demonstrates a further improvement compared to the scenario where the ANNs are used separately.

6 Conclusions

This study has proposed an ANN framework for predicting retailer's demand in a wholesale supply chain, with the ultimate aim to help the wholesaler decrease the occurrence of OOS situations. The proposed approach makes use of two ANNs. ANN#1 takes 6 factors as inputs and returns the selling price of the products sold by the wholesaler as output, while ANN#2 takes 5 factors as inputs and returns an estimate of the retailer's demand as output. This output can be used by the wholesaler to issue purchase orders to its suppliers.

The framework has been trained and tested on a real wholesale supply chain using the data related to a sample of 208 different products handled by the wholesaler, and two scenarios were considered for ANN implementation. For each scenario, both the forecasting performance of the ANNs in terms of the MAPE and the economic impact of OOS reduction resulting from the use of the ANNs for demand forecasting were evaluated.

From a practical perspective, it is self-evident that significant economic benefits can be achieved if the various supply chain echelons are able to predict the final customer's demand with good accuracy (Moon et al. 2003; Fernie et al. 2010). The proposed approach has proved to be effective in improving the forecasting performance of the wholesaler and thus has potentials to bring relevant savings for the whole supply chain. The numerical outcomes suggest that the use of the ANNs allows the economic loss due to OOS to decrease by approximately 40.9% in scenario 1 and by more than 56% in scenario 2. Looking expressively at the wholesaler, who is the focus of the proposed approach, a better forecasting accuracy represents a great opportunity first of all to reduce OOS, but also to improve business relationships with the retailers and enhance customer's satisfaction (Moon et al. 2003). While this study has looked at the economic aspect of OOS occurrence and at the savings achievable using the proposed model, evaluating different business aspects was not done in this paper, as at the time of writing, the company has still not implemented the ANN framework systematically (this will be the next step of the research).

From a scientific point of view, the proposed approach is effective in predicting both the selling price of products and the retailer's demand. Additionally, the cascade implementation of the ANNs is more effective than the separate use of ANNs in terms of OOS reduction, which provides further justification for the proposed framework. Moreover, although the use of ANNs *per se* is not new in the demand forecasting context, the newness of the proposed approach grounds in proposing the use of cascade ANNs, for predicting the selling price of products and forecasting the relating demand. In this respect, the proposed approach fills two main gaps in the published literature. First, the proposed approach addresses the product selling price formulation, before predicting the demand; this is an important point because in wholesale distribution the selling price is expected to have a relevant impact on the demand of a product. Therefore, a structured approach to determine the selling price of products could be useful to the wholesaler. Second, from a more technical perspective, there are no examples of the application of MNNs or cascade ANNs in the demand forecasting domain. Therefore, the proposed approach is expected to give a relevant contribution to the literature.

Although the structure of the model is of general validity, it should be mentioned that the factors used as input in the ANNs are relevant for the context of wholesale distribution (where the application was contextualized), but could not be equally relevant in a different context. Probably, if changing the application context, additional factors could be identified and included in the analysis, while other ones could be removed. This is also the reason why only one application (mainly for testing purpose) was carried out in this study. Nonetheless, it is reasonable to expect that more applications would contribute to proving the validity of the proposed model. Hence, future research activities can be directed to extending the application of the model to the remaining classes of products handled by the targeted company, to verify whether the forecasting performance of the ANN framework is somehow affected by the amount of input data treated. Similarly, contexts other from the wholesale distribution could be considered for the implementation of the proposed approach, to verify its suitability in case of products with different characteristics.

References

Aastrup, J., Kotzab, H. (2009). Analyzing out-of-stock in independent grocery stores –an empirical study. International Journal of Retail & Distribution Management, 37(9), 765-789.

Aburto, L., & Weber, R. (2007). Improved supply chain management based on hybrid demand forecasts. Applied Soft Computing, 7(1), 136-144.

Aizenberg, I., Sheremetov, L., Villa-Vargas, L., Martinez, J. (2016). Multilayer Neural Network with Multi-Valued Neurons in time series forecasting of oil production. Neurocomputing, 175, 980–989

Ali, S., Sanjoy, P., Kamrul, A., Abdullahil, A. (2011). Forecasting of optimum raw material inventory level using Artificial Neural Network. International Journal of Operations and Quantitative Management, 17(4), 333-348.

Azadeh, A., Beheshtipour, B., Moghaddam, M. (2011). An artificial neural network approach for improved demand estimation of a cool-disk manufacturer. International Journal of Industrial and Systems Engineering, 7(3), 357-380.

Babai, M. Z., Ali, M. M., Boylan, J. E., & Syntetos, A. A. (2013). Forecasting and inventory performance in a two-stage supply chain with ARIMA (0, 1, 1) demand: Theory and empirical analysis. International Journal of Production Economics, 143(2), 463-471.

Bala, P.K., (2012). Improving inventory performance with clustering based demand forecasts. Journal of Modelling in Management, 7(1), 23-37.

Banerjee, M., Mishra, M., (2017). Retail supply chain management practices in India: A business intelligence perspective. Journal of Retailing and Consumer Services, 34, 248–259

Benkachcha, S., Benhra, J., El Hassani, H. (2014). Demand forecasting in supply chain: Comparing multiple linear regression and artificial neural networks approaches. International Review on Modelling and Simulations, 7(2), 279-286.

Burnett, J. (2003). Core concepts of marketing. Hoboken, New Jersey: John Wiley Sons (chapter 9).

Boulaksil, Y. (2016). Safety stock placement in supply chains with demand forecast updates. Operations Research Perspectives, 3, 27-31.

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. European Journal of Operational Research, 184(3), 1140-1154.

Cho, S.B., Kim, J.H. (1995). Combining multiple neural network by fuzzy integral for robust classification. IEEE Transactions on Systems Man and Cybernetics, 25(2), 380-384.

Dalrymple, D.J. (1987). Sales forecasting practices: results of a United States survey. International Journal of Forecasting, 3, 379-392

Efendigil, T., Önüt, S., Kahraman, C. (2009). 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.

Ehrenthal, J.C.F., Gruen, T.W., Hofstetter, J.S. (2014). Value attenuation and retail out-of-stocks: A service-dominant logic perspective. International Journal of Physical Distribution & Logistics Management, 44(1/2), 39-57

Esper, T.L., Stank, T.P., Ellinger, A.E., Flint, D. and Moon, M. (2010). Demand and supply integration: a value and knowledge-based conceptual framework for competitive advantage. Journal of the Academy of Marketing Science, 38(1), 5-18.

Ferbar, L., Čreslovnik, D., Mojškerc, B., & Rajgelj, M. (2009). Demand forecasting methods in a supply chain: Smoothing and denoising. International Journal of Production Economics, 118(1), 49-54.

Fernie. J., Sparks, L., McKinnon, A.C., (2010). Retail logistics in the UK: past, present and future. International Journal of Retail & Distribution Management, 38(11/12), 894-914

Ferreira, R.P., Martiniano, A., Ferreira, A., Ferreira, A., Sassi, R.J. (2016). Study on daily demand forecasting orders using Artificial Neural Network. IEEE Latin America Transactions, 14(3), 1519-1525.

Fildes, R., Goodwin, P., Lawrence, M., Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. International Journal of Forecasting, 25(1), 3-23.

Fliedner, G. (2001). Hierarchical forecasting: issues and use guidelines. Industrial Management & Data Systems, 101(1), 5-12.

Fisher, M., (2009). Rocket science retailing: the 2006 Philip McCord Morse lecture. Operations Research, 57(3), 527-540.

Gruen, T., Corsten, D. (2007). A comprehensive guide to retail out-of-stock reduction in the fast-moving consumer goods industry. http://www.nacds.org/pdfs/membership/out_of_stock.pdf. Accessed 20.06.2016

Gruen, T., Corsten, D., Bharadwaj, S. (2002). Retail out-of-stocks: a worldwide examination of extent causes and consumer responses. http://itsoutofstock.com/wp-content/uploads/2013/04/GMA_2002_-Worldwide_OOS_Study.pdf. Accessed 04.07.2016.

Holweg, C., Teller, C., Kotzab, H., (2016). Unsaleable grocery products, their residual value and instore logistics. International Journal of Physical Distribution & Logistics Management, 46(6/7), 634-658.

Huber, J., Gossmann, A., & Stuckenschmidt, H. (2017). Cluster-based hierarchical demand forecasting for perishable goods. Expert systems with applications, 76, 140-151.

Inman, J.J., McAlister, L., Hoyer, W.D. (1990). Promotion signal: proxy for a price cut? Journal of Consumer Research, 17, 74-81.

İşlek, İ., Öğüdücü, Ş.G. (2015). A retail demand forecasting model based on data mining techniques. Proceedings of the IEEE 24th International Symposium on Industrial Electronics (ISIE), 55-60. DOI:10.1109/ISIE.2015.7281443.

Jaipuria, S., Mahapatra, S.S. (2014). An improved demand forecasting method to reduce bullwhip effect in supply chains. Expert Systems with Applications, 41(5), 2395-2408.

Kabir, G., Hasin, M.A.A. (2014). Application of adaptive neuro fuzzy inference system in demand forecasting for power engineering company. International Journal of Industrial and Systems Engineering, 2, 237-255.

Kandananond, K. (2012). A comparison of various forecasting methods for autocorrelated time series. International Journal of Engineering Business Management, 4(1), 1-6.

Khosroshahi, H., Husseini, S. M., & Marjani, M. R. (2016). The bullwhip effect in a 3-stage supply chain considering multiple retailers using a moving average method for demand forecasting. Applied Mathematical Modelling, 40(21), 8934-8951.

Kittler, J., Hatef, M., Duin, R.P.W., Matas, J. (1998). On combining classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(3), 226-239.

Kök, G., Fisher, M.L. (2007). Demand estimation and assortment optimization under substitution: methodology and application. Operations Research, 55(6), 1001–1021.

Konar, A. (2006). Computational intelligence: principles, techniques and applications. Berlin: Springer Science & Business Media.

Kumar, A., Adlakha, A., & Mukherjee, K. (2016). Modeling of product sales promotion and price discounting strategy using fuzzy logic in a retail organization. Industrial Management & Data Systems, 116(8), 1418-1444.

Laroche, M., Kalamas, M., Huang, Q. (2005). Effects of coupons on brand categorization and choice of fast foods in China. Journal of Business Research, 58, 674–686.

Lawrence, J., Fredrickson, J. (1998). BrainMaker user's guide and reference manual (7th ed.). Nevada City: California Scientific Software.

Lee, Y.H., Jung, J.W., Eum, S.C., Park, S.M., Nam, H.K. (2006). Production quantity allocation for order fulfilment in the supply chain: A neural network based approach. Production Planning and Control, 17(4), 378-389.

Levenberg, K. (1944). A method for the solution of certain problems in least squares. Quarterly of Applied Mathematics, 5, 164-168.

Liang, W. Y., & Huang, C. C. (2006). Agent-based demand forecast in multi-echelon supply chain. Decision support systems, 42(1), 390-407.

Lu, C.-J., Lee, T.-S., Lian, C.-M. (2012). Sales forecasting for computer wholesalers: A comparison of multivariate adaptive regression splines and artificial neural networks. Decision Support Systems, 54(1), 584-596.

Mahbub, N., Kumar Paul, S., Azeem, A. (2013). A neural approach to product demand forecasting. International Journal of Industrial and Systems Engineering, 15(1), 1-18.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics, 11(2), 431-441.

Mathworks (2008). MatlabTM 7 starting guide. http://www.mathworks.com/academia/student_version/learnmatlab.pdf. Accessed 20.06.2016.

McClure, P.J., West, E.J. (1969). Sales effects of a new counter display. Journal of Advertising Research, 9(1), 29-34.

McHugh, A.K., Sparkes, J.R. (1983). The forecasting dilemma. Management Accounting, 61(3), 30-34

Mentzer, J.T., Bienstock, C. (1998). Sales Forecasting Management. Sage Publications, London

Moon, M.A., Mentzer, J.T. Smith, C.D. (2003). Conducting a sales forecasting audit. International Journal of Forecasting, 19(1), 5-25.

Mulhern, F.J., Padgett, D.T. (1995). The relationship between retail price promotions and regular price purchases. Journal of Marketing, 59(4), 83-90.

Nguyen, H.H. Chan, C.W., (2004). Multiple neural networks for a long term time series forecast. Neural Computing and Applications, 13, 90–98.

Nikolopoulos, K. I., Babai, M. Z., & Bozos, K. (2016). Forecasting supply chain sporadic demand with nearest neighbor approaches. International Journal of Production Economics, 177, 139-148.

Roland Berger Consultants, 2003. ECR – Optimal shelf availability increasing shopper satisfaction at the moment of truth. Munich: ECR Europe.

Sari, K. (2008). Inventory inaccuracy and performance of collaborative supply chain practices. Industrial Management & Data Systems, 108(4), 495-509.

Sakhuja, S., Jain, V., Kumar, S., Chandra, C., Ghildayal, S.K. (2016). Genetic algorithm based fuzzy time series tourism demand forecast model. Industrial Management and Data Systems, 116(3), 483-507.

Shahrabi, J., Mousavi, S.S., Heydar, M. (2009). Supply chain demand forecasting: A comparison of machine learning techniques and traditional methods. Journal of Applied Sciences, 9(3), 521-527.

Singh, L.P., Challa, R.T. (2016). Integrated forecasting using the discrete wavelet theory and artificial intelligence techniques to reduce the bullwhip effect in a supply chain. Global Journal of Flexible Systems Management, 17(2), 157-169.

Silva Filho, O. S., Carvalho, M. A., Cezarino, W., Silva, R., & Salviano, G. (2013). Demand forecasting for blood components distribution of a blood supply chain. IFAC Proceedings Volumes, 46(24), 565-571.

Stank, T.P., Dittmann, J.P., Autry, C.W., (2011). The new supply chain agenda: a synopsis and directions for future research. International Journal of Physical Distribution & Logistics Management, 41(10), 940-955.

Sugandi, G., Sule, E.T., Hasan, M. (2016). Corporate reputation and customer relationship management impact on business customers purchase decision. International Journal of Management Sciences and Business Research, 5(1), 68-84.

Supermarket Guru Consumer Panel (2011). Shoppers cite store annoyances. The Woodlands: Supermarket News.

Syntetos, A.A., Babai, Z., Boylan, J.E., Kolassa, S., Nikolopoulos, K. (2016). Supply chain forecasting: Theory, practice, their gap and the future. European Journal of Operational Research, 252, 1–26.

Walters, R.G., Rinne, H.J. (1986). An empirical investigation into the impact of price promotions on retail store performance. Journal of Retailing, 62, 237-266.

Wang, W. J., & Xu, Q. (2014). A Bayesian Combination Forecasting Model for Retail Supply Chain Coordination. Journal of applied research and technology, 12(2), 315-324.

Wei, Y., Li, F. (2015). Impact of heterogeneous consumers on pricing decisions under dual-channel competition. Mathematical Problems in Engineering, Article ID 584084, 9 pages. DOI: 10.1155/2015/584084.

West, D.C. (1994). Number of sales forecast methods and marketing management. Journal of Forecasting, 13(4), 395-407

Willart, S.P.C., (2015). Price competition in retailing: The importance of the price density function. Journal of Retailing and Consumer Services, 26, 125–132

Wyman, O. (2014). Purchasing cost reduction: what distributors can learn from leading retailers and manufacturers.http://www.oliverwyman.com/content/dam/oliverwyman/global/en/files/insights/distributionwholesale/2014/Jan/PurchasingCostReductions/OW_Purchasing%20Cost%20Reduction_2014.pdf. Accessed 20.06.2016.

Yu, S., Ou, J. (2009). Forecasting model of agricultural products prices in wholesale markets based on combined BP neural network-time series model. Proceedings of the International Conference on Information Management, Innovation Management and Industrial Engineering, 1, 558-561. DOI: 10.1109/ICIII.2009.140.

Zhang, G., Patuwo, B.E., Hu, M.Y. (1998). Forecasting with artificial neural network: The state of the art. International Journal of Forecasting, 14, 35–62.

FIGURES

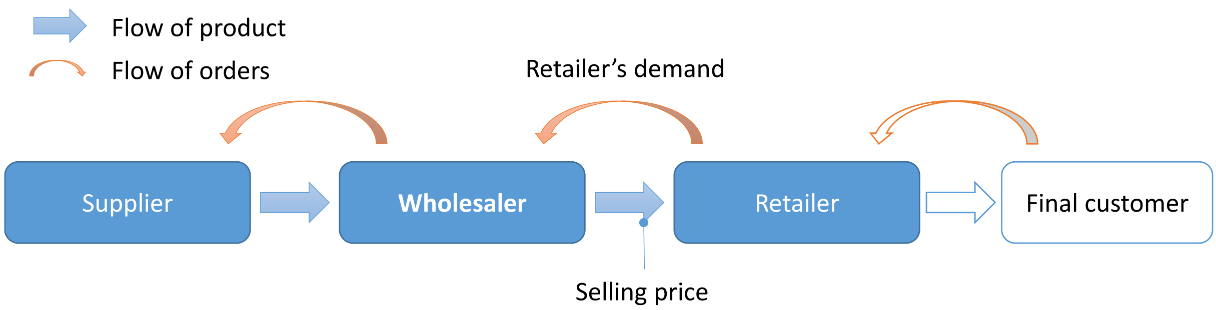

Figure 1. Scheme of the wholesale supply chain

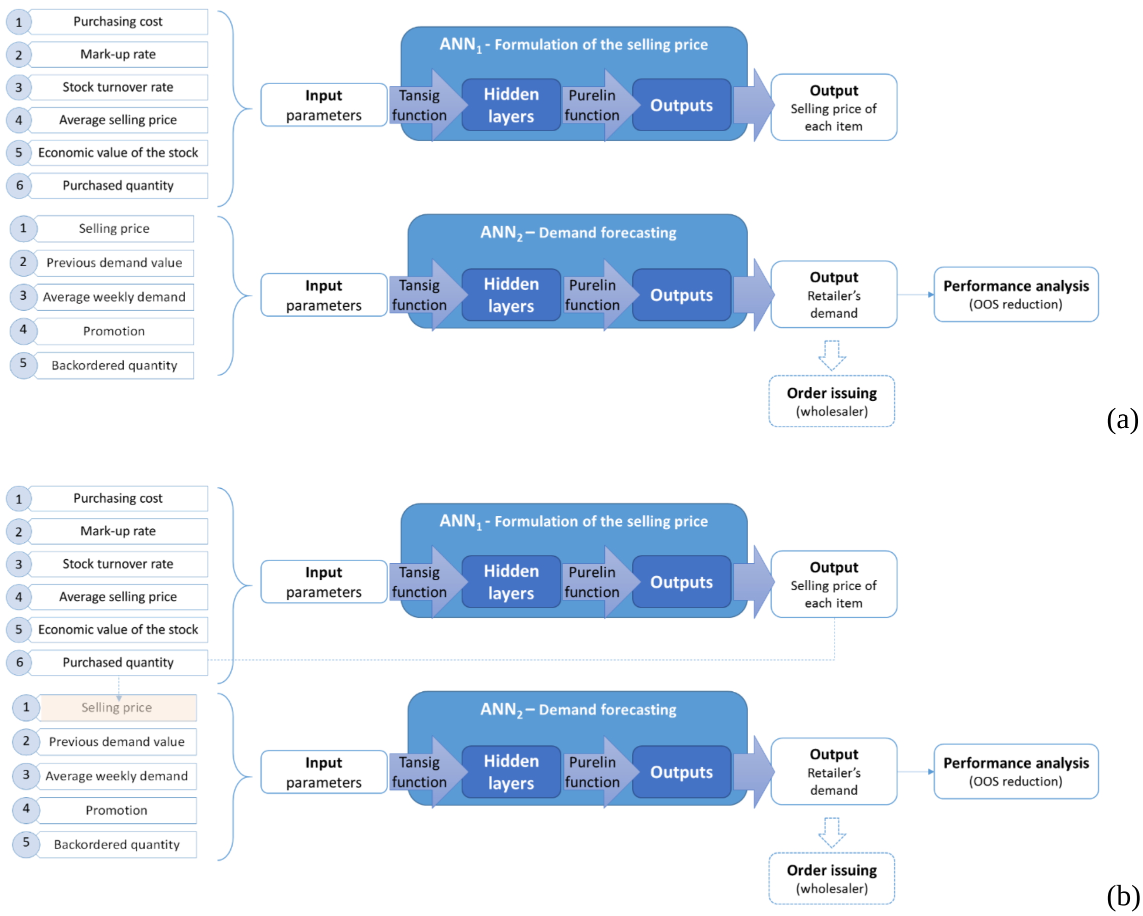

Figure 2. Proposed approach – separate ANN implementation (a) and the cascade ANN implementation (b)

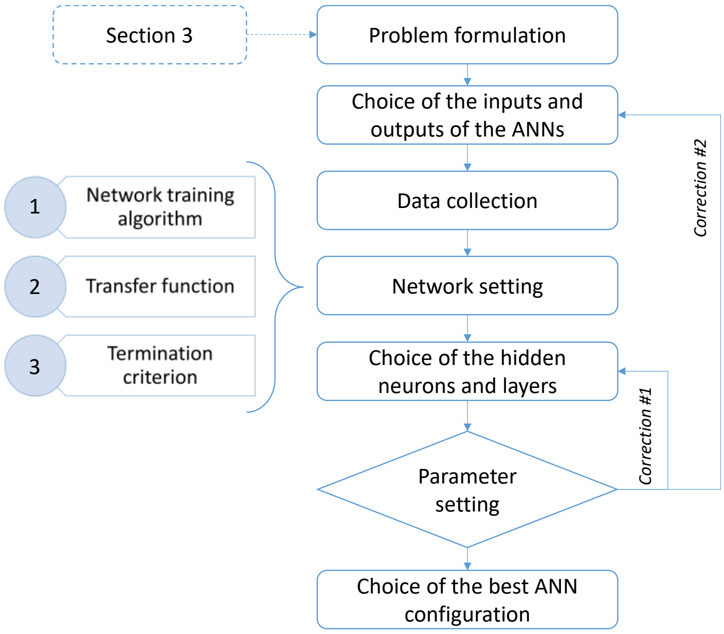

Figure 3. Framework for the development of the ANN architecture

TABLES

Table 1. Comparison of the present study vs. the previous studies

Study	**Area of demand forecasting**	**Applied Approach(es)**
Aburto & Weber	Retailing	Hybrid Autoregressive Integrated Moving Average (ARIMA) model and neural network
Aizenberg et al.	Oil production	Multilayer neural network with multi-valued neurons
Azadeh et al.	Cool-disk manufacturing	ANN vs. regression
Babai et al.	Two-stage supply chain	ARIMA(0,1,1) model
Benkachcha et al.	Not specified	ANN vs. multiple linear regression
Boulaksil	Two-stage supply chain	Martingale model of forecast evolution
Carbonneau et al.	Upstream end of a supply chain	Non-linear machine learning techniques
Ferbar et al.	Multi-echelon supply chain	Wavelet denoising analysis model
Ferreira et al.	Delivery parcel service	ANN multilayer perceptron
Huber et al.	Retail supply chain of perishable goods	Multivariate ARIMA models coupled with cluster analysis
Jaipuria & Mahapatra	Manufacturers in supply chains	Integrated discrete wavelet transforms (DWT) analysis and ANN
Kabir & Hasin	Power engineering	Adaptive neuro-fuzzy inference system
Kandananond	Consumer products	ANNs vs. support vector machines and ARIMA model
Khosroshahi et al.	Three-stage supply chain with multiple retailers	Moving average method
Liang & Huang	Multi-echelon supply chain	Genetic algorithm solution approach incorporates the rough set theory
Lu et al.	Computer wholesaling	ANN vs. multivariate adaptive regression splines and other AI tools
Mahbub et al.	Furniture	ANN
Nikolopoulos et al.	Multi-echelon supply chain with sporadic demand	Nearest neighbor approach
Sakhuja et al.	Tourist arrivals in Taiwan	Fuzzy time series combined with an evolutionary genetic algorithm
Shahrabi et al.	Car components	Series forecasting methods vs. ANNs and support vector machines
Silva Filho et al.	Four-stage blood supply chain	Univariate multiplicative seasonal Box-Jenkins ARIMA model
Singh & Challa	Cement dealer	Combination of wavelet theory, ANNs, and an adaptive-network-based fuzzy inference system
The current study	*Wholesale supply chain*	*Cascade implementation of two artificial neural networks (ANNs)*
Wang & Xu	Retail Supply Chain	Bayesian combination forecasting model

Table 2. Summary of the configurations tested for the determination of the purchasing cost (ANN#1)

Input neurons	Network configuration			
	1	**2**	**3**	**4**
Purchasing cost	x	x	x	x
Mark-up rate	x	x	x	x
Stock turnover rate	-	-	x	x
Average selling price	x	x	x	x
Economic value of the stock	x	x	x	x
Purchased quantity	-	x	-	x
Number of input neurons	*4*	*5*	*5*	*5*

Table 3. Network performance for the determination of the purchasing cost (ANN#1)

Network configuration	Number of hidden neurons	MAPE (%)	Minimum percentage error (%)	Maximum percentage error (%)
1	2	2.82	0	12.74
	6	2.67	0	10.71
	8	2.30	0	8.04
	10	2.01	0	5.19
	14	2.44	0	6.89
2	2	2.90	0	12.26
	6	2.77	0	10.93
	8	2.75	0	12.26
	10	2.66	0	8.37
	14	3.14	0	11.25
3	2	2.33	0	11.88
	6	2.66	0	9.84
	8	2.37	0	10.11
	10	2.38	0	6.80
	14	2.06	0	4.44
4	2	3.41	0.07	17.35
	6	2.53	0	12.26
	8	3.76	0	17.24
	10	2.66	0	11.57
	14	2.50	0	8.10

Table 4. Summary of configurations tested for the demand forecast (ANN#2)

Input neurons	**Network configuration**				
	1	**2**	**3**	**4**	**5**
Selling price	x	x	x	x	x
Previous demand value	-	x	x	-	x
Average demand	x	-	x	x	x
Promotion	x	x	x	x	x
Backordered quantity	-	-	-	x	x
Number of input neurons	*3*	*3*	*4*	*4*	*5*

Table 5. Network performance for the determination of the demand forecast (ANN#2)

Network configuration	Number of hidden neurons	MAPE (%)	Minimum percentage error (%)	Maximum percentage error (%)
1	2	14.76	0.24	32.29
	6	11.37	1.16	22.11
	8	11.37	0.05	21.63
	10	9.20	0.92	22.11
	14	11.96	0.42	25.67
2	2	10.66	1.84	30.79
	6	7.95	0.79	17.96
	8	13.10	0.86	28.30
	10	16.32	0.68	35.18
	14	11.92	0.97	28.30
3	2	7.24	0.43	18.26
	6	9.03	1.83	22.53
	8	5.51	0.02	14.27
	10	8.73	0.81	19.89
	14	8.80	1.38	22.68
4	2	12.91	0.99	44.73
	6	16.35	2.35	40.67
	8	17.81	1.06	42.80
	10	16.40	0.65	43.14
	14	17.34	0.46	45.75
5	2	17.32	0.14	48.26
	6	19.65	0.05	52.48
	8	19.90	2.10	42.55
	10	14.82	1.43	42.20
	14	13.75	0.37	36.23

Table 6. Economic evaluation of the demand forecasting process – scenarios 1 and 2.

		Real scenario		ANN adoption - scenario 1			ANN adoption - scenario 2		
Product code[1]	**Price [€/unit]**	**Real demand [units/month]**	**OOS_{cost} [€/month]**	**Forecasted demand – [units/month]**	**Forecasting error**	**OOS_{cost} [€/month]**	**Forecasted demand [units/month]**	**Forecasting error**	**OOS_{cost} [€/month]**
COD0001	10.75	731	0.00	779	6.57%	0.00	715	-2.19%	172.00
COD0002	0.99	3536	0.00	3234	-8.54%	298.98	3204	-9.39%	328.68
COD0003	1.43	1012	657.80	894	-11.66%	168.74	1040	2.77%	0
COD0004	2.2	557	68.20	611	9.69%	0.00	546	-1.97%	24.20
COD0005	0.46	948	0.00	909	-4.11%	17.94	973	2.64%	0
COD0006	1.63	1030	319.48	1210	17.48%	0.00	1088	5.63%	0
COD0007	7.28	174	0.00	152	-12.64%	160.16	187	7.47%	0
COD0008	2.77	2194	537.38	2266	3.28%	0.00	2338	6.56%	0
COD0009	1.02	326	0.00	283	-13.19%	43.86	358	9.82%	0
COD0010	0.71	599	81.65	571	-4.67%	19.88	549	-8.35%	35.50
COD0011	2.18	470	0.00	532	13.19%	0.00	531	12.98%	0
COD0012	2.49	6556	0.00	6235	-4.90%	799.29	6120	-6.65%	1085.64
COD0013	0.77	9781	19.25	9224	-5.69%	428.89	10094	3.20%	0
COD0014	0.87	619	653.37	528	-14.70%	79.17	596	-3.72%	20.01
COD0015	1.3	4255	279.50	4476	5.19%	0.00	4536	6.60%	0
…	…	…	…	…	…	…	…	…	…
COD0196	0.79	588	0.00	598	1.70%	0.00	672	14.29%	0
COD0197	2.85	5230	28.50	5124	-2.03%	302.10	5231	0.02%	0
COD0198	1.19	822	197.54	865	5.23%	0.00	894	8.76%	0
COD0199	1.85	5577	373.70	5214	-6.51%	671.55	5378	-3.57%	368.15
COD0200	0.85	3300	0.00	3082	-6.61%	185.30	3216	-2.55%	71.40
COD0201	2.24	437	208.32	452	3.43%	0.00	446	2.06%	0
COD0202	2.68	50	0.00	49	-2.00%	2.68	51	2.00%	0
COD0203	1.25	3104	80.00	3046	-1.87%	72.50	3083	-0.68%	26.25
COD0204	1.03	1849	186.43	2097	13.41%	0.00	2032	9.90%	0
COD0205	1.2	572	28.80	622	8.74%	0.00	584	2.10%	0
COD0206	2.91	9315	2661.75	10535	13.10%	0.00	9172	-1.54%	416.13
COD0207	2.03	493	50.75	546	10.75%	0.00	543	10.14%	0
COD0208	0.68	11109	33.32	10276	-7.50%	566.44	10697	-3.71%	280.16

TOTAL	**-**	**558,881**	**€ 48,031.21**	**556,207**	**-1.095%**	**€ 28,358.42**	**555,122**	**-0.479%**	**€ 21,008.89**

(1) The product code is simply a progressive number (from 1 to 201) as the real information about the products tested was kept confidential. Also, the table is limited to a subset of the product, for visualization purpose.