

The Bullwhip effect in water demand management: taming it through an artificial neural networks-based system

Borja Ponte, Laura Ruano, Raúl Pino and David de la Fuente

ABSTRACT

The Bullwhip effect (BE) refers to the amplification of the variance of orders and inventories along the supply chain as they move away from the customer. This is considered as the main cause of inefficiencies in the management of a traditional supply chain. However, the BE is not relevant in the classic system of water distribution, based on long-term supply management. Nevertheless, current circumstances have drawn a new context, which has introduced the concept of water demand management, in which efficiency and sustainability are of great importance. Then, the time horizon of management has decreased enormously and the supply time takes on an important role. Therefore, the BE must be considered, as it significantly raises the costs of management. On the one hand, this paper brings evidence that the BE appears in a system of real-time management of water demand. On the other hand, it proposes the application of artificial intelligence techniques for its reduction. More specifically, an advanced forecasting system based on artificial neural networks has been used. The BE is heavily damped.

Key words | artificial neural networks, Bullwhip effect, water demand management

Borja Ponte (corresponding author)
Laura Ruano
Raúl Pino
David de la Fuente
Polytechnic School of Engineering,
University of Oviedo,
Campus de Viesques s/n,
33204, Gijón,
Spain
E-mail: ponteborja@uniovi.es

INTRODUCTION

The concept of Bullwhip effect (BE) emerged in the early 1990s in some large companies, when the new competitive context conceded strategic importance to supply chain management (SCM). Some businesses began to understand SCM as a source of competitive advantages and studied it in detail, trying to optimize its performance. At that time, Procter and Gamble realized that the purchase orders received in one of its flagship products, Pampers diapers, fluctuated significantly, while the product demand in the retailer was almost constant. They also found out that the variability in orders transmitted to their suppliers were much higher. It was called the BE (Lee *et al.* 1997).

The growing importance of logistics in the doubtful environment currently faced by businesses has prompted the development of this concept, which is considered to be the main cause of inefficiencies in SCM (Disney *et al.* 2005). For this reason, various supply chains have focused on reducing

the BE, with the aim of minimizing the derivatives overruns. In contrast, in some particular supply chains, this phenomenon has not been relevant and it has not been widely studied. The water supply system is one of them.

Nevertheless, the perspective of municipal policies about water management has changed significantly over the last two decades, mainly due to the pressures generated by population growth and industrialization. Hence, the concept of water demand management (WDM) has developed significantly. Brooks (2006) proposed a current definition of WDM with five components: (1) reducing the quantity or quality of water required to accomplish a specific task; (2) adjusting the nature of the task so it can be accomplished with less water or lower quality water; (3) reducing losses in movement from source through use to disposal; (4) shifting time of use to off-peak periods; and (5) increasing the ability of the system to operate during droughts.

This burgeoning concern over efficiency and sustainability around WDM (Charlesworth & Adeyeye 2013) has led to a reduction in the time horizon. Some years ago, long-term forecasting was enough for the design of the system and the development of plans (among others, Willsie & Pratt (1974)). However, nowadays, short-term forecasts are required for attaining high efficiency in operation and management (among others, Gato et al. (2007)). Herrera et al. (2010) defend that the ready availability of hourly predictions of water demand is crucial due to three main reasons: (1) it allows to determine the optimal regulation and pumping systems to meet the predicted demand, which promotes energy efficiency (operative point of view); (2) it allows to combine water sources in the most appropriate way to achieve a preset standard in the supply water (quality point of view); and (3) it allows to detect failures and network losses through the comparison of the actual and expected flow (vulnerability point of view). It can be called real-time WDM.

In a long-term WDM system, the BE does not arise. If the time horizon is very long, the supply time becomes trivial and does not determine the performance of the replenishment policy. However, reducing this time horizon introduces in the study the need to consider the supply time, and therefore the menace of BE surges. It must be taken into account in order to avoid the negative consequences that it can have on the supply system. Thereby, one of the main objectives of this paper is to bring evidence via simulation of the appearance of the BE in a real-time WDM.

Furthermore, this work proposes a solution to the identified problem, based on the application of artificial intelligence techniques in forecasting the hourly water demand. More specifically, an advanced forecasting system, whose core is an artificial neural network (ANN), has been developed. This methodology has been widely used in the forecasting of series of a similar nature, as the short-term electricity load (see Hippert et al. (2001) for a review). Herrera et al. (2010) showed that predictive models, among which ANNs are included, provide great performance in forecasting the hourly water consumption. This research has tried to reduce the error even further by developing a double-loop system that chooses at all times the optimal network structure (both input variables and

hidden neurons). Therefore, the second goal of this paper is to demonstrate that these smart tools can cause a large decrease in the BE generated in the water distribution system and, consequently, it can lead to an improvement in the management.

BACKGROUND: THE BE IN SUPPLY CHAINS

Although research on the BE was strengthened two decades ago when large companies looked at the problem, Forrester (1961) long before noted the amplification of demand variability along a generic supply chain through a simulation model. Thereby, many authors express mathematically the BE generated at level n of a linear supply chain (BE^n) as the quotient of the variance of the orders issued to the upper level supply chain ($\sigma_{POE}^2{}^n$) and the orders received from the lower level of the same ($\sigma_{POR}^2{}^n$). As this metric only evaluates the output variance compared with the input variance, it should be supplemented by another one that provides the variation in the level of inventories (i.e. the structure that causes the above variation). Therefore, some authors (e.g. Disney & Towill 2003) propose an alternative metric of the quotient of the variance of the stock ($\sigma_{STOCK}^2{}^n$) and the variance of the demand ($\sigma_{POR}^2{}^n$). It can be named the alternative bullwhip effect (ABE^n) and is expressed by Equation (2).

$$BE^n = \frac{\sigma_{POE}^2{}^n / \mu_{POE}^n}{\sigma_{POR}^2{}^n / \mu_{POR}^n} = \frac{\sigma_{POE}^2{}^n}{\sigma_{POR}^2{}^n} \quad (1)$$

$$ABE^n = \frac{\sigma_{STOCK}^2{}^n}{\sigma_{POR}^2{}^n} \quad (2)$$

The BE involves large economic losses in the supply chain, by increasing missing sales, obsolescence, and labor, transportation and storage costs, so it can be considered a major cause of inefficiencies within SCM (Disney et al. 2005).

Lee et al. (1997) showed that there are five main causes that lead to this phenomenon: (1) errors in demand forecasting; (2) non-zero lead times; (3) order batching; (4) price fluctuations; and (5) supply shortages. The famous 'Beer

Game', proposed by the MIT and analyzed by [Sterman \(1989\)](#), brings evidence that the BE is generated along the supply chain even if the last three causes are not considered. Obviously, if lead time was null, the supply from the factory would instantly respond to customer requirements and the BE would not appear, and if there were no errors in the forecasting, each level would know exactly what it needs, so the BE would not surge either.

SCM is a very complex problem, which is conditioned by the interaction of multiple agents, each one of which has to weight a large number of variables. Thus, modern artificial intelligence tools have been widely used in order to optimize the management and to buffer the BE. Next, a brief literature review on this subject is shown. In the beginning, the Metamorph tool, based on multi-agent methodology and developed by [Maturana *et al.* \(1999\)](#), can be highlighted. In 2010, [Hong *et al.*](#) designed an ANNs-based controller and using radio-frequency identification (RFID) technology ([Hong *et al.* 2010](#)). [Jaipuria & Mahapatra \(2014\)](#) developed an advanced forecasting system (ANNs and wavelet discrete transform) to reduce the BE in a generic supply chain. Also, the recent and relevant works carried out by [Bahroun *et al.* \(2010\)](#), [Saber *et al.* \(2012\)](#) and [Zarandi & Gamasae \(2013\)](#) should be mentioned.

THE BE IN REAL-TIME WATER DEMAND MANAGEMENT

The main hypothesis of this work is that the BE appears in real-time WDM systems, and therefore it must be controlled due to the consequences that it could bring to the system.

Under these conditions, the BE in a water supply network is the increasing variability of the demand transmitted along the same as it moves away from the final points of consumption. This phenomenon directly causes the increase of the variations in the water flow conveyed along the distribution network and also in the increase of the variations in the water stored in the supply tanks. Therefore, it tends to oversize the system (distribution network, supply tanks and treatment equipments), although the infrastructure oversize is more influenced by other reasons – reliability and security against unforeseen, but possible, events. Moreover, the BE generates cost overruns in the

works of water pumping, collection and purification, as the contracted power is greater when the variability of the system requirements over time is large. Hence, taming the BE leads to improvements in management.

Simulation model

To demonstrate the generation of the BE along a real-time WDM system, this research has considered a simple structure of a water supply network, which consists of three main levels interconnected by the distribution piping: (1) natural sources (catchment points), where water is collected; (2) points-of-use (POU), representing the distributed water demand; and (3) supply tanks (storage reservoirs), which receive water from the natural sources and send it to the POU. Then, a discrete simulation model has been developed in MATLAB R2014a of a supply system managed hourly, focused on the supply tanks.

Other assumptions adopted to model the supply system are the following: (1) stochastic POU demand (see Results and discussion section, as the same time series has been used); (2) fixed supply time: 1 hour (on the one hand, from natural sources to supply tanks and, on the other hand, from supply tanks to POU); (3) unconstrained catchment, storage and transportation system; (4) water is pumped to the supply tanks in order for them to store at the beginning of each hour – order-up-to point – the forecast plus a security level, with the aim of protecting against shortage; and (5) non-negative condition of the order quantity (water cannot be returned to the previous level). Obviously, it is a simplified model of the reality, but it considers the main causes that surge the BE in real-time WDM systems.

Next, the mathematical formulation of the model is described. Water pumped at the end of each hour from natural sources to supply tanks (WP_t) can be expressed as the difference between the demand forecast for the next period (\widehat{D}_{t+1}) and the water stored in the tanks at the end of that period (WT_t), also considering security level which must be kept in the tank (SL), by Equation (3). Along the same line, the water stored in the tanks at the end of each period (WT_t) is the water stored in the tanks at the end of the previous period (WT_{t-1}), adjusted by the water pumped in the previous period from natural sources (WP_{t-1}) – as the lead time is 1 hour – and by the demand

(D_t), unless this difference is less than 0, according to Equation (4). In that case, it is not possible to meet all the demand, and a deficit of unmet demand (UMD_t) is generated, by Equation (5). Furthermore, logically, the water sent from the supply tanks to the POU (WS_t) is the demand (D_t), unless the water stored at the end of the previous period (WT_{t-1}) was lower, according to Equation (6).

$$WP_t = \max\{\widehat{D}_{t+1} - WT_t + SL, 0\} \tag{3}$$

$$WT_t = \max\{WT_{t-1} - D_t + WP_{t-1}, 0\} \tag{4}$$

$$UMD_t = \max\{-(WT_{t-1} - D_t + WP_{t-1}), 0\} \tag{5}$$

$$WS_t = \min\{D_t, W_{t-1}\} \tag{6}$$

The operational logic of the simulation system is illustrated in Figure 1. As mentioned above, it is based on the supply tanks, and the BE can be observed when comparing the demand transmitted from POU to supply tanks and from supply tanks to natural sources.

the user through an interface, and it is connected to a database with the aim of storing and analyzing the results. It should be noted that there are two main flows the water flow, from natural sources to POU and constrained by the lead time (supply time), and the order flow, in the opposite direction. The flow chart of the operations in the supply tanks corresponds to the previous equations.

Simulation results

To calculate the forecast for the next time period (\widehat{D}_{t+1}), moving averages (Holt 2004) of 3–6 periods and simple exponential smoothing (Gardner 2006) with coefficients 0.5–0.9 have been used. In addition, three different tests with each forecasting method (FM) have been carried out, as the value of the security level with which the tanks works has also been modified. Table 1 shows the results of the 12 simulations using in all cases the same week (randomly chosen) of the time series.

Table 1 demonstrates the generation of BE in the 12 tests (since the ratio is greater than 1 in all cases), in which different FMs and security levels have been used. In

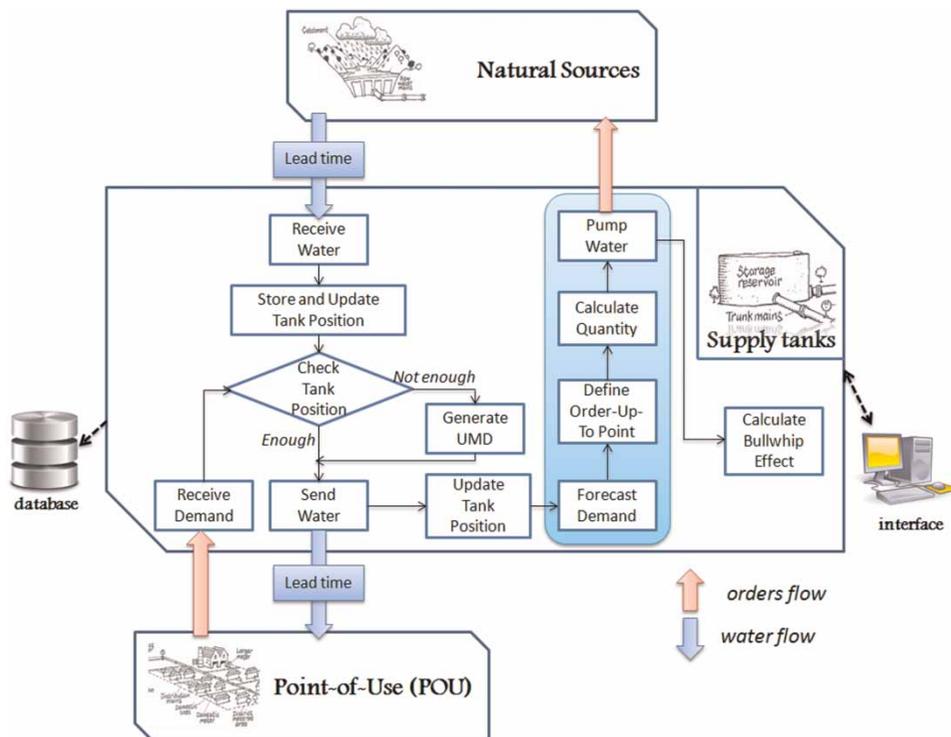


Figure 1 | Outline of the simulation system.

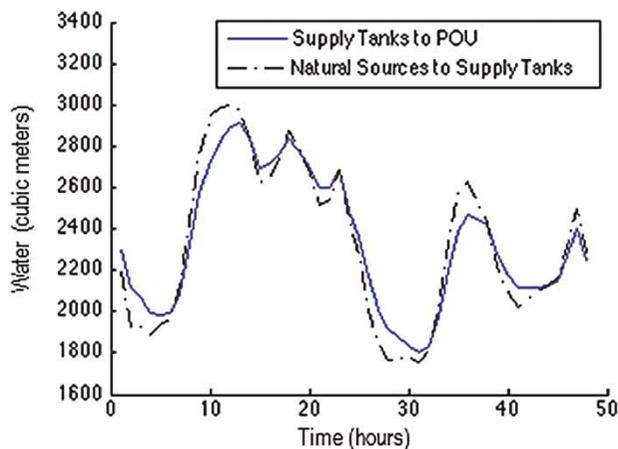
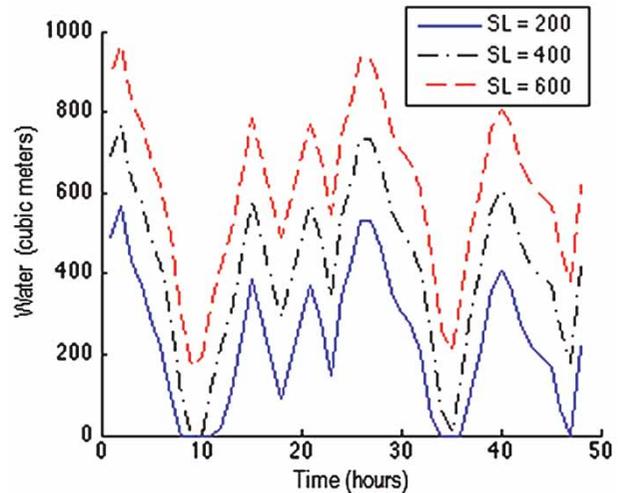
Table 1 | Results of the simulation

Test	FM	SL	BE	ABE	UMD
1	MA3	200	1.19	0.18	5,730
2	MA3	400	1.26	0.27	744
3	MA3	600	1.27	0.29	0
4	MA6	200	1.09	0.29	13,336
5	MA6	400	1.19	0.45	4,656
6	MA6	600	1.28	0.58	515
7	ES0.5	200	1.17	0.16	4,439
8	ES0.5	400	1.23	0.23	269
9	ES0.5	600	1.23	0.24	0
10	ES0.9	200	1.19	0.10	1,229
11	ES0.9	400	1.20	0.11	0
12	ES0.9	600	1.19	0.11	0

The columns refer to the number of the test (Test), the FM, the security level of the tanks in cubic meters (SL), the BE, the ABE, and the UMD in cubic meters.

the best situation (test 4), the amplification of the variance of the demand is 9%. Although not included for the sake of simplicity, tests carried out with changes in the supply time or the pumping policy also evidence the existence of this phenomenon. Thereby, in this real-time WDM system, there is amplification in the variability of the demand.

The results presented in Table 1 show a straightforward (and easy to understand) relationship: the higher the security level, the lower the UMD. Furthermore, it brings evidence that the higher the security level, the higher the variations along the system, which typically results in an increase of the BE.

**Figure 2** | Water received and sent by the supply tanks for 48 hours (corresponding to Friday and Saturday) in test 3.**Figure 3** | Volume of water in the supply tanks for 48 hours (corresponding to Friday and Saturday) in tests 1, 2 and 3.

The BE generation on the water supply network, by way of example, can be seen graphically in Figure 2, which represents the water conveyed between supply tanks and POU and between natural resources and supply tanks for 2 days of test 3. In it, the amplification of the variance is 27%. Figure 3 displays, for the same time period, the volume of water in the supply tanks in tests 1, 2 and 3. These variations produce the magnification of the ABE when the security level increases, although the UMD obviously decreases. Thereby, the consequences of the BE in the WDM system are evidenced.

DESCRIPTION OF THE FORECASTING SYSTEM

The forecasting errors are the main cause of the BE. Hence, a system based on an ANN structure has been developed to forecast the hourly demand with the aim of minimizing the errors. The results will be evaluated by comparing them with the ones provided by statistical methods, which will be detailed afterward.

ANN forecasting system

ANNs are computational models inspired by an animal's central nervous system, which are capable of machine learning, as well as pattern recognition. They are systems of

interconnected neurons, distributed in different layers, which can compute values from inputs. Two characteristics of ANNs that make them particularly useful for forecasting time series are the ability to approximate practically any function (even non-linear ones) and the opportunity for ‘piece-wise’ approximations of the functions. For a more detailed description of ANNs as a FM and its contrast with other traditional tools, see Pino et al. (2008).

In particular, the model used for this study is the non-linear autoregressive network with exogenous inputs (NARX), where the next value of the dependent output signal is forecast ($\hat{y}(t) = \hat{D}_t$) as a regression on previous values of the output signal ($y(t) = D_t$) and previous values of an independent (exogenous) input signal ($x_t = x(t)$). The NARX model is developed, among others, in the work of Piroddi & Spinelli (2003). The software that has been used is MATLAB R2014a.

Figure 4 shows the architecture of the forecasting system – it is called multi-layer perceptron (MLP). From a set of inputs, the system is capable of building a response. In particular, the program takes not only the previous demands, but also the hour (ranged from 00 to 23 h), the week day (from 1, corresponding to Mondays, until 7, corresponding

to Sundays) and an extra variable, related to the main feature of the day, which differences working days (1), Saturdays (2) and Sundays and holidays (3) – due to the nature of this time series from Monday to Friday, consumption of water remains pretty similar, while it decreases on Saturday, and keeps falling on Sundays (holidays can be approximated to Sundays).

MLP are networks that have more than one layer of adaptive weights (Bishop 1995). It has three layers of units taking values in the range 0–1, and each layer is nourished with the previous ones. Any number of weighted connections can be used, but MLPs with two weighted connections are very much capable of approximating just about any functional mapping. The MLP can be mathematically represented by Equation (7), where y_t represents the output (forecast), f_{outer} represents the output layer, f_{inner} represents the input layer transfer function, w_{xy} represents the weights and biases ($i \in [1, (3m + 3)]$ refers to the input neurons and $j \in [1, n]$ refers to the hidden neurons) and (z) represents the z th layer.

$$\hat{D}_t = y_t = f_{outer} \left[\sum_{j=1}^n w_{1j}^{(2)} f_{inner} \left(\sum_{i=1}^{3m+3} w_{ji}^{(1)} \cdot x_i + w_{j0}^{(1)} \right) + w_{10}^{(2)} \right] \tag{7}$$

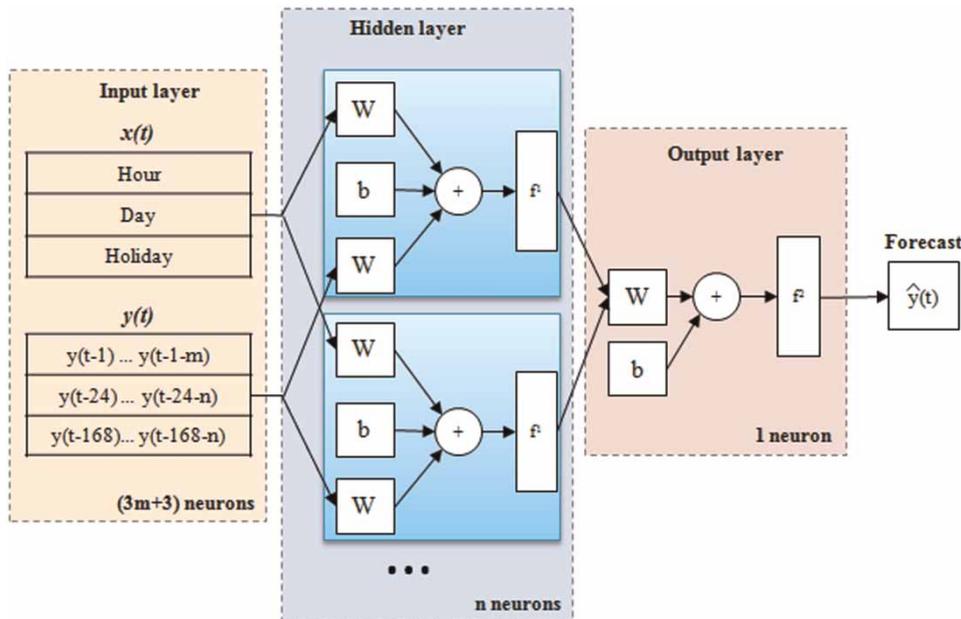


Figure 4 | ANN architecture of the forecasting system.

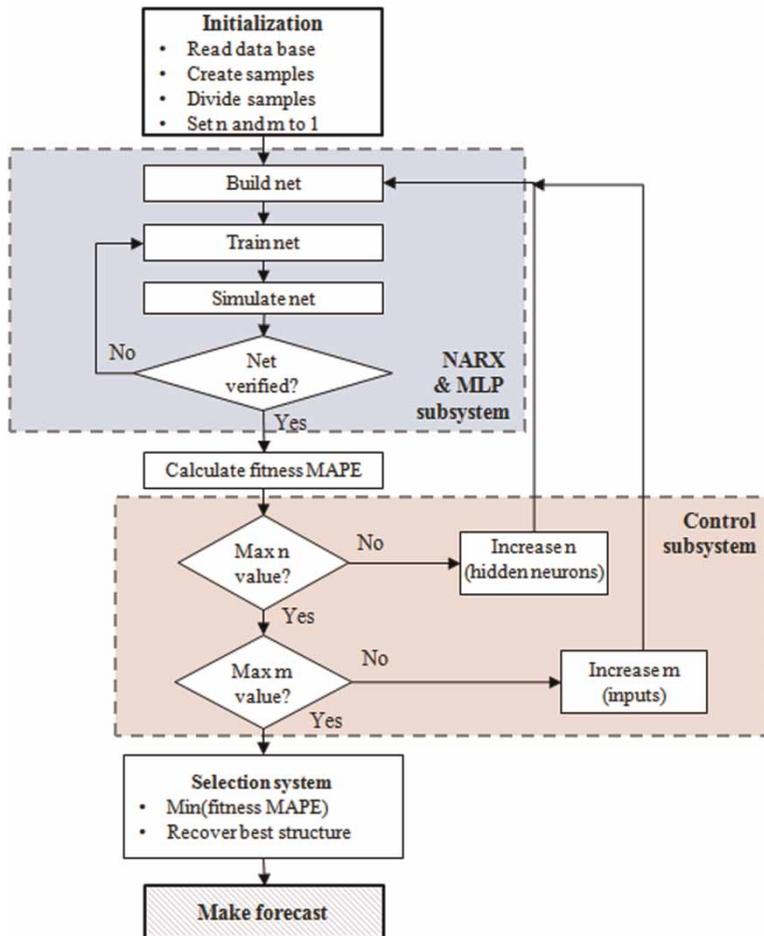


Figure 5 | Flow chart followed by the forecasting system to make the prediction.

Figure 5 provides a brief explanation of the structure and operation of the ANN forecasting system. It makes an hourly forecast, when it receives the last demand from the measurement equipments and the information is stored in the database. Then, it reads the database and selects the last 1,008 samples, which correspond to an entire period of 6 weeks (the hourly demands of 42 days). Samples are randomly divided (except the last 12) into three different groups 70% of them are classified as training data, for adjusting the network according to its error; 15% as validation data, used to measure network performance and to halt training when it stops improving; and the remaining 15% as testing data, which provides an independent measure of network performance during and after training.

The used training function updates weight and bias values according to Levenberg–Marquardt optimization, which uses

this approximation to the Hessian matrix in a Newton-like update (see Moré 1978). To verify the training of the ANNs and to avoid overfitting, the early-stopping method (Sarle 1995) has been used, as the number of training examples is sufficiently large. It presents interesting advantages in terms of speed and ease of application in comparison with cross-validation (Kohavi 1995), which is much more suitable when the number of examples is low. Training stops when any of these conditions occurs: the maximum number of repetitions (100) is reached; the maximum amount of time is exceeded (10 minutes); the performance gradient falls below the value defined (10^{-10}); validation performance has increased more than the times defined (6) since the last time it decreased; or the scalar value exceeds its maximum value (10^{10}).

In the search of the structure that fits best the time series, two things are varied by the control subsystem the

number of neurons in the hidden layer and the number of delays (hence the number of variables that are considered to forecast). Therefore, the system chooses at each time the optimal structure of the network, seeking for a better performance of the tool than if the same structure was always imposed.

About the first loop, it should be kept in mind that the higher number of hidden neurons (n) are chosen, the more complexity the structure will have, and it requires a higher time of carrying out. However, this does not always translate into a better outcome of the system because overfitting problems are more prone to occur if there are too many neurons in the hidden layer. The system builds a network for each one of the different values of the variable from $n = 2$ to $n = 20$, with jumps of two units.

In addition, another loop is made in order to look for the combination of delays that best fits the input data. As the hourly consumption time series shows trend and double periodicity, the best way of defining a new value for the curve of water consumption is by choosing the demands of the previous hours (from $y(t - 1)$ to $y(t - 1 - m)$), the demands of the previous day at the same hour and the previous ones (from $y(t - 24)$ to $y(t - 24 - m)$), and the demands of the previous week at the same day hour and the previous ones (from $y(t - 168)$ to $y(t - 168 - m)$). The system evaluates alternatives from $m = 1$ to $m = 8$.

To evaluate the performance of the forecasting system, the criterion of the mean absolute percentage error (MAPE), introduced by Makridakis (1993), is used. It can be expressed by Equation (8), where p is the time horizon

$$\text{MAPE} = \frac{1}{p} \sum_{t=1}^p \left| \frac{D_t - \hat{D}_t}{D_t} \right| \quad (8)$$

The last 12 demands are saved as testing samples, in order to orientate the double-loop to determine the optimal ANN architecture at every moment. Therefore, after each iteration, the system calculates the MAPE of the last 12 demands as an indicator of network performance in recent hours (fitness MAPE). Once the double-loop process of building networks has ended, the system chooses the structure that has generated a minimum fitness MAPE and new predictions are made with this architecture.

Statistical models

Traditional methods are used to compare their results with the developed system, and to show the improvements on the BE reduction. Three statistical techniques have been used. The system chooses the best of the three at any time using the same criterion (fitness MAPE minimization).

First, an autoregressive model (AR) is used (Akaike 1969). The algorithm for computing the least squares AR model is the forward-backward approach, which minimizes the sum of a least squares criterion for a time-reversed model. The second model corresponds to IVAR, which estimates the AR model using the instrumental variable method (Arellano & Bover 1995). Both algorithms treat noise differently. AR assumes white noise, while the IVAR is not sensitive to noise color. The third one corresponds to the ARMA model (Jones 1980). It includes a moving average component to consider the relation of the series with past values of the errors.

RESULTS AND DISCUSSION

To evaluate the effectiveness of the forecasting system in the BE reduction, a simulated time series with the hourly water demand in 2009 and 2010 in Gijón (a municipality of 300,000 inhabitants in the north of Spain) has been used. Validated by the municipal water company – real data are not available as this company still do not carry out an hourly management – this series was created through the monthly water demand of the city, a distribution model of hourly water demand for a city in south-eastern Spain (Herrera *et al.* 2010), and random parameters. The information obtained from the literature was used to create a consumption modulation curve describing the behavior of the hourly water demand along the different days of the week. To adjust properly the vertical scale (in cubic meters), and hence including the long-term trend of the series, each month's water demand (known for 2009 and 2010) has been applied. This simulation was run for the above-mentioned time horizon, adding random parameters with the aim of slightly modifying the curve at every moment and creating short-term trends in the series. Holidays have also been considered.

This way, the time series replicates a real hourly water demand series, which is a complex series with double seasonality and trend. On the one hand, it has a daily periodicity, as every 24 hours the series shows a similar structure. On the other hand, the consumption significantly varies on Saturdays and Sundays (and on holidays if there are), hence there is a weekly periodicity (168 hours). Moreover, the time series does not remain in a constant range, but it exhibits the above-mentioned trends both in mean and variance.

In this study, different days and hours have been selected randomly with the aim of evaluating the performance of the system in different situations. Table 2 presents the different periods that have been chosen. In its last

Table 2 | Training period (6 weeks) and testing periods (24 hours) of the eight tests that have been performed

Test	Training period		Testing period		Testing day
	From	To	From (hour)	To (hour)	Kind
1	24/01/09	06/03/09	0	23	Saturday
2	28/07/09	08/09/09	5	4	Holiday
3	17/12/09	28/01/10	17	16	Working day
4	30/12/09	10/02/10	12	11	Working day
5	10/01/10	21/02/10	4	3	Holiday
6	22/01/10	05/03/10	21	20	Saturday
7	12/03/10	23/04/10	5	4	Working day
8	28/07/10	08/09/10	14	13	Holiday

column, it differentiates between working days (1), Saturdays (2) and Sundays and holidays (3), according to the classification above mentioned.

The discrete simulation model described in the third section has been used to calculate the BE and the ABE with the ANN forecasting system in the eight tests. The chosen security level in the supply tanks is 500 m³ – this value has been selected because there would not be UMD in none of the eight cases.

Table 3 depicts the final results obtained in this research. They point out, broadly speaking, the huge efficiency of the ANN forecasting system versus the statistical methods in the reduction of the BE. As expected, an improvement in the forecasting MAPE usually implies an improvement in both indicators of the BE.

The ANN forecasting system leads to the achievement of minor errors. By selecting at each time the best architecture of the network, forecasting errors around 1% are obtained in the tests performed, below those achieved by the traditional statistical methods. Thus, the BE – that is evident and a major threat to the WDM system with the statistical models (the amplification varies between the 11% in test 2 and the 53% in test 5) – experiences a great reduction when using the ANN system. In other words, this forecasting system makes the amplification of the variability of the demand along the supply network non-significant. Similarly, variations in the water volume at the supply tanks are largely

Table 3 | Results of the simulation

Test	Artificial neural networks				Statistical methods			Reduction		
	Struct.	MAPE (%)	BE	100 ABE	MAPE (%)	BE	100 ABE	MAPE (%)	BE (%)	ABE
1	12-2-1	0.70	0.98	0.59	2.64	1.33	10.94	73.48	26.32	18.54
2	12-2-1	0.98	1.03	0.86	1.58	1.11	2.38	37.97	7.21	2.77
3	9-8-1	1.65	1.00	6.71	2.09	1.28	7.89	21.05	21.88	1.18
4	9-10-1	0.58	1.02	0.57	1.86	1.29	7.73	68.82	20.93	13.56
5	9-12-1	0.90	1.00	0.99	3.38	1.53	16.37	73.37	34.64	16.54
6	9-12-1	0.76	0.95	0.92	2.65	1.34	12.65	71.32	29.10	13.75
7	15-4-1	1.02	1.00	0.76	2.73	1.12	5.21	62.64	10.71	6.86
8	12-2-1	1.03	1.00	0.83	2.64	1.33	10.94	60.98	24.81	13.18

The columns contain the MAPE of the forecasting in percentage, the BE generated in the distribution system and the ABE multiplied by 100, both when the ANN forecasting system is used and when the best statistical model is used to forecast.

In addition, the comparison between both methodologies is displayed through the percentage reduction of MAPE and BE and through the quotient between the ABE obtained in both cases. It also includes the ANN structure used by the system to forecast the hourly demand in each test (struct.).

reduced. This leads to the conclusion that the negative consequences of the BE in the hourly managed water distribution system are remarkably attenuated with the system that has been implemented.

Regarding the system's architecture, Table 3 brings evidence that there is not a direct relationship between the complexity of the network and the accuracy of their forecasts. For working days, in most cases, the system finds that the best architecture corresponds to the selection of the minimum value of m , so that the number of inputs is usually smaller (tests 3, 4 and 6) than in weekends and holidays. However, if the number of hidden neurons in each test is analyzed, it can be noted that weekends and holidays generally need fewer neurons in the hidden layer (tests 1, 2 and 8).

By way of example, test 1 is a clear example in which the results of the ANN forecasting system significantly decreases the MAPE obtained with the statistical methods, and as a result the BE is minimized. Figure 6 shows the real consumption and the two forecasts. The ANN system (0.70% MAPE) offers better performance (2.64% MAPE of the best statistical model). The graph shows that it captures very accurately the periodicity and the trend of the consumption. Meanwhile, Figure 7 displays the difference between the POU's consumption (from supply tanks to POU) and the transmitted demands (from natural sources to supply tanks), one when ANNs are used and the other with statistical methods. It shows that the distortion introduced to the WDM system is much smaller with the ANNs, so that the tank requirements vary much less. The

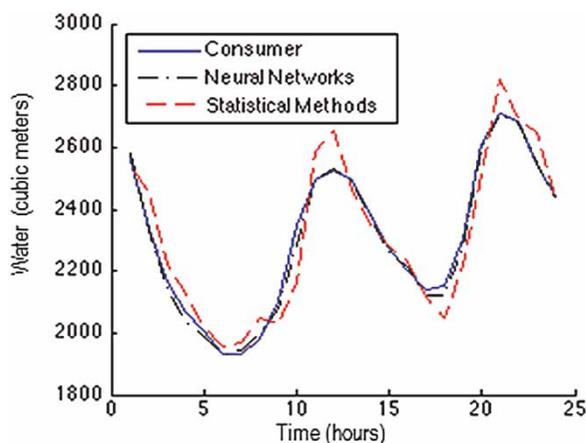


Figure 6 | Differences between the demand and the two forecasts for test 1.

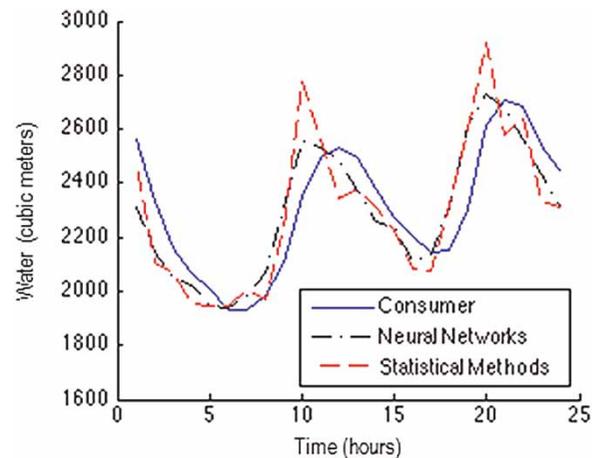


Figure 7 | Variations of the real and the transmitted demand for test 1.

figure shows that the accuracy of the forecasting system causes the water conveyed between natural resources and supply tanks to approximate closely to the POU's consumption, but displaced – the supply time is the time difference between them. Thus, the BE is greatly reduced.

CONCLUSIONS AND FUTURE RESEARCH LINES

In this paper, the BE is studied for the first time in the context of water supply networks. Even though it was not a relevant concept in a traditional long-term WDM system, the BE is emphasized nowadays with new approaches based on hourly management, that look for efficiency optimization. Under these circumstances, demand forecasting is an essential practice and supply time must be taken into account. As a consequence of both, the BE comes out. Through a discrete simulation model, its generation has been shown, as well as the consequences it has on a real-time system system's oversize, risks of shortage, and energy expenditure increase. Therefore, the BE should be considered as a head cause of inefficiencies in WDM.

One way to reduce the BE and to mitigate its damage is the use of advanced forecasting tools. Hence, this research has developed a double-loop forecasting system which chooses at each time the most appropriate architecture of the network (both the inputs to be considered and the neurons in the hidden layer). With this ANNs-based system, very-low errors in forecasting the hourly demands are

achieved in comparison with traditional statistical methods. The tests performed at random moments of time point out that the MAPE reduction leads to a large decrease of the BE. Thereby, the use of the intelligent forecasting system reduces the distortion induced in the water supply network, so that the inefficiencies in WDM are significantly mollified.

There are two main lines of future work that this research group is planning as next steps on this topic. The first is to extend this model to a larger noise conditions scenario, as well as to use a more complex supply structure. Considering these new factors can provide insights to other relevant insights on this issue. More specifically, it is planned to study the BE in WDM from a supply approach, as many real systems are greatly influenced by hydrological uncertainty (could a reverse BE exist?). The second line is to integrate this forecasting system within a larger system aimed at optimizing the management.

ACKNOWLEDGEMENTS

The authors acknowledge the Government of the Principality of Asturias for the financial support of this research, through the Severo Ochoa program (reference BP13011), as well as the help provided by the municipal water company of Gijón. We also would like to thank the referees for their insightful comments on the first version of this paper.

REFERENCES

- Akaike, H. 1969 Fitting autoregressive models for prediction. *Ann. Inst. Statist. Math.* **21** (1), 243–247.
- Arellano, M. & Bover, O. 1995 Another look at the instrumental variable estimation of error-components models. *J. Econom.* **68** (1), 29–51.
- Bahroun, Z., Moalla, M., Baazaoui, G. & Campagna, J. P. 2010 Multi-agent modelling for replenishment policies simulation in supply chains. *Eur. J. Indust. Eng.* **4** (4), 450–470.
- Bishop, C. M. 1995 *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford.
- Brooks, D. B. 2006 An operational definition of water demand management. *Int. J. Water Resour. Dev.* **22** (4), 521–528.
- Charlesworth, S. & Adeyeye, K. 2013 Editorial: water efficiency and management. *J. Water Suppl. Res. Technol. Aqua* **62** (8), 493–495.
- Disney, S. M. & Towill, D. R. 2003 The effect of Vendor Managed Inventory (VMI) dynamics on the Bullwhip effect in supply chains. *Int. J. Prod. Econ.* **85**, 199–215.
- Disney, S. M., Farasyn, I., Lambrecht, M., Towill, D. R. & Van de Velde, W. 2005 Taming the Bullwhip effect whilst watching customer service in a single supply chain echelon. *Eur. J. Oper. Res.* **173** (1), 151–172.
- Forrester, J. W. 1961 *Industrial Dynamics*. MIT Press, Cambridge, MA.
- Gardner, E. S. 2006 Exponential smoothing: the state of the art – Part II. *Int. J. Forecast.* **22** (4), 637–666.
- Gato, S., Jayasuriya, N. & Roberts, P. 2007 Temperature and rainfall thresholds for base use urban water demand modeling. *J. Hydrol.* **337** (1), 364–376.
- Herrera, M., Torgo, L., Izquierdo, J. & Pérez-García, R. 2010 Predictive models for forecasting hourly urban water demand. *J. Hydrol.* **387** (1), 101–140.
- Hippert, H. S., Pedreira, C. E. & Sourza, R. C. 2001 Neural networks for short-term load forecasting: a review and evaluation. *IEEE Trans. Power Syst.* **16** (1), 44–55.
- Holt, C. C. 2004 Forecasting seasonals and trends by exponentially weighted moving averages. *Int. J. Forecast.* **20** (1), 1–5.
- Hong, S. R., Kim, S. T. & Kim, C. O. 2010 Neural network controller with on-line inventory feedback data in RFID-enabled supply chain. *Int. J. Prod. Res.* **48** (9), 2613–2632.
- Jaipuria, S. & Mahapatra, S. S. 2014 An improved demand forecasting method to reduce Bullwhip effect in supply chains. *Expert Syst. Appl.* **41** (5), 2395–2408.
- Jones, R. H. 1980 Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics* **22** (3), 389–395.
- Kohavi, R. 1995 A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Vol. 2, pp. 1137–1143.
- Lee, H. L., Padmanabhan, V. & Whang, S. 1997 The Bullwhip effect in supply chains. *Sloan Manage. Rev.* **38** (3), 93–102.
- Makridakis, S. 1993 Accuracy measures: theoretical and practical concerns. *Int. J. Forecast.* **9** (22), 527–529.
- Maturana, F., Shen, W. & Norrie, D. H. 1999 Metamorph: an adaptive agent-based architecture for intelligent manufacturing. *Int. J. Prod. Res.* **37** (10), 2159–2174.
- Moré, J. J. 1978 The Levenberg–Marquardt algorithm: implementation and theory. *Lect. Notes Math.* **630** (1), 105–116.
- Pino, P., Parreño, J., Gómez, A. & Priore, P. 2008 Forecasting next-day price of electricity in the Spanish energy market using artificial neural networks. *Eng. Appl. Artif. Intell.* **21** (1), 53–62.
- Piroddi, L. & Spinelli, W. 2003 An identification algorithm for polynomial NARX models based on simulation error minimization. *Int. J. Contr.* **77** (16), 1767–1781.
- Saberi, S., Nookabadi, A. S. & Hejazi, S. R. 2012 Applying agent-based system and negotiation mechanism in improvement of inventory management and customer order fulfillment in multi echelon supply chain. *Arabian J. Sci. Eng.* **37** (3), 851–861.

- Sarle, W. S. 1995 Stopped training and other remedies for overfitting. In: *Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics*, pp. 352–360.
- Sterman, J. 1989 [Modelling managerial behaviour: misperceptions of feedback in a dynamic decision making experiment](#). *Manage. Sci.* **35** (3), 321–339.
- Willsie, R. H. & Pratt, L. H. 1974 [Water use relationships and projection corresponding with regional growth Seattle region](#). *Water Resour. Bull.* **10** (2), 360–371.
- Zarandi, M. H. F. & Gamasae, R. 2013 A type-2 fuzzy system model for reducing Bullwhip effect in supply chains and its applications in steel manufacturing. *Sci. Iranica.* **20** (3), 879–899.

First received 1 July 2014; accepted in revised form 4 December 2014. Available online 5 January 2015