BIOSYSTEMS ENGINEERING 102 (2009) 202–218

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper: SW—Soil and Water

Improved irrigation water demand forecasting using a soft-computing hybrid model

Inmaculada Pulido-Calvo, Juan Carlos Gutiérrez-Estrada*

Dpto. Ciencias Agroforestales, EPS, Campus Universitario de La Rábida, Universidad de Huelva, 21819 Palos de la Frontera (Huelva), Spain

ARTICLE INFO

Article history:
Received 5 December 2007
Received in revised form
24 September 2008
Accepted 29 September 2008
Available online 30 November 2008

Recently, Computational Neural Networks (CNNs) and fuzzy inference systems have been successfully applied to time series forecasting. In this study the performance of a hybrid methodology combining feed forward CNN, fuzzy logic and genetic algorithm to forecast one-day ahead daily water demands at irrigation districts considering that only flows in previous days are available for the calibration of the models were analysed. Individual forecasting models were developed using historical time series data from the Fuente Palmera irrigation district located in Andalucía, southern Spain. These models included univariate autoregressive CNNs trained with the Levenberg–Marquardt algorithm (LM). The individual models forecasting were then corrected *via* a fuzzy logic approach whose parameters were adjusted using a genetic algorithm in order to improve the forecasting accuracy. For the purpose of comparison, this hybrid methodology was also applied with univariate autoregressive CNN models trained with the Extended-Delta-Bar-Delta algorithm (EDBD) and calibrated in a previous study in the same irrigation district. A multicriteria evaluation with several statistics and absolute error measures showed that the hybrid model performed significantly better than univariate and multivariate autoregressive CNNs.

1. Introduction. General scope of the work

Information regarding water demand in irrigated areas is basic information for the development and implementation of successful water resource management tools given that irrigated agriculture is the largest user of water throughout the world, accounting for 87% of consumptive uses (ONU, 1997; Sumpsi *et al.*, 1998). Also, forecasting of water demand is one of the main problems in the design, management and modernisation of water supply and distribution systems.

Actually, most pressurised irrigation systems operating on-demand deliver water with the flow rate and pressure required by farm irrigation systems, sprinkling or micro-irrigation, and respecting the time, duration and frequency decided by the farmers. Therefore, they allow farmers to operate their irrigation systems with a large freedom with respect to other types of delivery schedules. Usually the Clément formula (Clément, 1966; Clément and Galand, 1979) is used to design collective irrigation systems operating on-demand. This approach does not permit to take into consideration the variety of flow regimes occurring in a collective irrigation system. So, a risk threshold is accepted, i.e. during the operation of the system, flow rates higher than those assumed at design may occur with low probability due to the seasonal and daily variation in water demand. Consequently, a large spatial and temporal variability of pressure and flow rates available in the hydrants may occur and affect network performance and even crop yield (Pereira, 1999; Lamaddalena

* *Corresponding author.*
E-mail addresses: ipulido@uhu.es (I. Pulido-Calvo), juanc@uhu.es (J.C. Gutiérrez-Estrada).
1537-5110/$ – see front matter
doi:10.1016/j.biosystemseng.2008.09.032

Nomenclature

C	fuzzy curve
CNN	Computational Neural Network
d	total number of observations of the validation set
E_2	efficiency coefficient
EDBD	Extended-Delta-Bar-Delta algorithm
FAM	Fuzzy Associative Memory
HI	fuzzy partition, category High
LM	Levenberg–Marquardt algorithm
LO	fuzzy partition, category Low
NFI	Naïve Forecast I model
NO	fuzzy partition, category Normal
PI	Persistence Index
Q_t	observed water demand at time t, $m^3\ day^{-1}$
$\widehat{Q}_t$	estimated water demand at time t, $m^3\ day^{-1}$
$\overline{Q}$	average of the observed water demand, $m^3\ day^{-1}$
R^2	determination coefficient
RMS	square root of the mean square error, $m^3\ day^{-1}$
SEP	percent standard error of prediction, %
VH	fuzzy partition, category Very High
VL	fuzzy partition, category Very Low
Y	defuzzed value. Correction of the CNN output value, $m^3\ day^{-1}$
$y^{Centre(f)}$	geometrical centre of the output fuzzy values
μ_i	fuzzy membership function

et al., 2007; Calejo *et al.*, 2008). That is why the operating plan of water distribution systems is typically prepared for a period of 24 h in advance in order to programme the pumps and valves control settings (Alvisi *et al.*, 2007).

Alternative methods to the Clément formula are being studied to generate flow rates through the simulation or estimation of the irrigation water demand taking into consideration the farmers' behaviour, i.e. that farmers' irrigation decisions vary relative to the assumptions made at design and planning phase (Pulido-Calvo *et al.*, 2003b; Khadra and Lamaddalena, 2006; Moreno *et al.*, 2007; Calejo *et al.*, 2008). Therefore, the water demand is one of the main parameters to evaluate because its forecasting is fundamental to the real-time operational control of an on-demand water distribution system (Pulido-Calvo *et al.*, 2003a).

For many years, methodologies for consumer demand modelling and prediction in a real-time environment for on-demand water distribution systems have been studied (Zhou *et al.*, 2002; Alvisi *et al.*, 2007). Two of the major approaches presented in the literature to assess the water demand involve conceptual (physical) modelling and system theoretical modelling (sometimes referred as black box approach).

A conceptual model generally aims to formulate the daily water requirements for crop irrigation by the rates of percolation and evapotranspiration that have been predicted at the stage of irrigation planning. Many models have been used to simulate these water requirements, from empirical or functional (Doorenbos and Pruitt, 1977; Doorenbos and Kassam, 1979; Allen *et al.*, 1998) to mechanistic (Van Aelst *et al.*, 1988). However, water requirements calculated for irrigation planning do not always meet the actual use (that is, consumer demand) due to changes in the field environment such as weather conditions and farmers' behaviour, which can influence the actual amounts of water used (Pulido-Calvo *et al.*, 2003b; Khadra and Lamaddalena, 2006; Moreno *et al.*, 2007; Pulido-Calvo *et al.*, 2007). Actual water management in some irrigation districts is carried out depending only on the experience and knowledge of the administrator although there is always a need to forecast daily water demand.

In the system theoretical approach a model is applied to identify a direct mapping between inputs and outputs without detailed consideration about the internal structure of the physical processes. Generally, compared with a conceptual model, the theoretical approach has much less data requirements. Many practical situations may not justify the time and the effort required to develop, to validate and to implement a conceptual model as their main goal is not to deeply access the irrigation water demand physical process, but to get from a given system answers that, despite not being fully understood, prove to be accurate enough. One of those practical situations involves water demand short-term forecasting.

It should be pointed out that the data availability often determines the model choice. In fact, continuous measurements of climatic data (precipitation, temperature, relative humidity, wind speed, etc.) can be easily and cost effectively obtained when compared with continuous measurements of soil characteristics, initial soil moisture, infiltration, etc. Therefore, a black box approach that operates based on accessible and commonly available data is much more suitable for operational forecasting purposes than a conceptual model.

The Artificial or Computational Neural Networks (ANNs or CNNs) can be classified as black box type models. A CNN is a non-linear mathematical structure capable of representing the complex non-linear processes that relate the inputs to the outputs of a system. CNNs models are increasingly being applied in many fields of science and engineering and usually provide highly satisfactory results. Some specific applications of CNN to water resource management and planning include the modelling of monthly, daily and hourly rainfall–runoff process (Hsu *et al.*, 1995; Lorrai and Sechi, 1995; Mason *et al.*, 1996; Abrahart *et al.*, 1999; Tokar and Johnson, 1999; Thirumalaiah and Deo, 2000; Tokar and Markus, 2000; Chiang *et al.*, 2004; Moradkhani *et al.*, 2004; Anctil and Rat, 2005; Agarwal *et al.*, 2006), real-time river and lake stage forecasting (Thirumalaiah and Deo, 1998, 2000; Abrahart and See, 2000, 2002; See and Openshaw, 2000; Cameron *et al.*, 2002; Nayebi *et al.*, 2006; Ondimu and Murase, 2007), rainfall forecasting (French *et al.*, 1992; Zhang *et al.*, 1997; Kuligowski and Barros, 1998), groundwater modelling (Roger and Dowla, 1994; Yang *et al.*, 1997), assessment of stream's hydrologic and ecological response to climate change (Poff *et al.*, 1996), drought analysis (Shin and Salas, 2000), etc.

Neural approaches for water demand prediction in urban (Griñó, 1992; Jain *et al.*, 2001; Bougadis *et al.*, 2005; Firat *et al.*, 2008) and irrigation delivery systems (Pulido-Calvo *et al.*, 2002, 2003a, 2007) have been reported in the literature. Essentially two basic techniques were developed. The first technique consists of establishing models based on the relation between the demand data and other influential factors. In the case of urban delivery systems, the models are based on the relation between the demand data and demographic and environmental factors. In the case of irrigation water-delivery systems, the models are based on the relation between the demand data, climatic and crop data. The second technique calculates the relationship between present and past demand data (stochastic analysis of time series).

In this paper, CNN approach was applied to the short-term forecasting of daily irrigation water demand. In the applications carried out only water demand data (water demand time series) were considered as input data. In the time series forecasting issues, past observations of one or more variables are collected and introduced as input data in a model that describes the underlying relationships among those variables and allows estimating future realisations of one of the same (Zhang, 2003). Recently, CNNs have been extensively applied to time series forecasting (Al-Saba and El-Amin, 1999; Abrahart and See, 2000; Zhang *et al*., 2001; Zhang, 2003; Gutiérrez-Estrada *et al*., 2004; Pulido-Calvo and Portela, 2007).

Some authors (Park, 1998; Abrahart and See, 2000; Pulido-Calvo *et al*., 2002, 2003a; Gutiérrez-Estrada *et al*., 2005; Pulido-Calvo *et al*., 2007; Pulido-Calvo and Portela, 2007) in works applied to forecasting of different kinds of time variables using multiple regression, Autoregressive integrated moving average (ARIMA) and/or neural network models, indicate that the forecasts provided by the models in each time period were systematically very close to the data observed in the previous time period (naïve behaviour). In order to improve the performance of the daily water demand forecasts, and so to solve this systematic displacement between observed and estimated values, a hybrid methodology combining CNN, fuzzy logic and genetic algorithms was applied to take advantage of the strength of these models.

The neural networks, fuzzy logic and genetic algorithms are intelligent computational methods or soft-computing technologies that can be extremely effective when used on their own. However, when combined together, the individual strengths of each approach can be exploited in a synergistic manner for the construction of powerful, hybrid and intelligent systems (See and Openshaw, 2000). Recent works have demonstrated that combining different models to create a single forecast produces better performance than the use of the best individual model alone (Pelikan *et al*., 1992; Ginzburg and Horn, 1994; Luxhoj *et al*., 1996; Wedding and Cios, 1996; Zhang, 2003). Some specific applications of hybrid models to water resource management have been reported by Shamseldin and O'Connor (1999), See and Openshaw (1999, 2000), See and Abrahart (2001), Chang *et al*. (2005), Guan and Aral (2005), Keskin *et al*. (2006), Kişi and Öztürk (2007), Pulido-Calvo and Portela (2007) and Firat and Güngör (2008). The basic idea is that a real-world problem is often complex in nature and any single model may not be able to capture different patterns equally well. Therefore, combining different models can increase the chance to capture different patterns in the data and improve forecasting performance (Zhang, 2003).

The purpose of this paper is to assess the potential improvements in performance that can be achieved by using soft-computing technologies for short-term irrigation water demand forecasting. The methodology for combining each of the techniques into a single forecasting solution is outlined. Water demand historical data from an irrigation distribution system in southern Spain were used to test the methodology.

2. Material and methods

2.1. *Study area and data source*

The proposed methodology was applied to the demand pressurised system of the irrigation district of Fuente Palmera, located in the Guadalquivir valley (Córdoba province, southern Spain) (Fig. 1). The mean water consumption in the zone is 16.5 ± 5.9 h m^3 annually and must be drawn from the Guadalquivir River. The average irrigated area is approximately 5000 ha and is irrigated by sprinkling and micro-irrigation on demand.

The pressurised irrigation system has two pumping stations in series. The first station carries water from the Guadalquivir River to a 5000 m^3 tank, which is the aspiration chamber for the second station, which discharges directly into the distribution line (Fig. 1). Given that the storage capacity of this tank does not allow the two pumping stations to operate independently, it is used to provide pressure to the branched pipeline system.

The main water supply system is branched and carries water from the booster station (second pumping station) to 78 different groups of farmers, each one whom has only one outlet. The minimum, maximum and average areas of the group of farmers are 21.6, 218.3, and 67.4 ha, respectively. From the main network outlets, the water is distributed to the plots through a secondary pipe network that is underground and fixed. This network includes all of the underground pipe networks for each group of farmers; these pipes branch off from the outlet to each hydrant of the farming units. The portable water supply system is made up of a number of pipes and portable elements belonging to each group of farmers and includes anything from hydrants to sprinklers or drips on the plots, whose average area is 6.25 ha. The total lengths of the main and secondary pipes networks are 38,000 and 141,000 m, respectively.

The more representative crops, according to areas occupied in a period of 14 consecutive irrigation seasons (from 1984–1985 to 1997–1998), were $43.31 \pm 18.68\%$ cotton, $23.78 \pm 11.60\%$ sunflower, $14.30 \pm 8.30\%$ wheat, $3.37 \pm 2.79\%$ sugar beet, $2.81 \pm 3.47\%$ olive, $2.61 \pm 3.39\%$ corn, $1.41 \pm 2.52\%$ sorghum, $1.23 \pm 1.07\%$ citric fruits and $1.12 \pm 0.99\%$ melon/watermelon.

Measured values of daily consumer demand data from 1988, 1989, 1990 and 1991 years were available from continuous records of booster station outputs. To determine irrigation water requirements (Fig. 2) and to compare with the consumer water demands, crop and climatic data from these three irrigation seasons were used. The crop data (crop coefficients and duration of the development stages) were obtained from agricultural experimental studies near the area and from other studies. The climatic data were collected at the Córdoba airport meteorological station. Mean annual rainfall in the area was 608 mm, with a standard deviation of 229 mm. Because of the

Fig. 1 – Scheme of main water supply system and localisation of Fuente Palmera irrigation district. The number of control nodes of pressure and flow are showed.

area's Mediterranean climate, there is a wet season in winter and a very dry season in summer. Consequently, monthly rainfall distribution is very irregular. The mean air temperature ranges from 10 °C in winter to over 27 °C in summer. The daily reference evapotranspiration ET_0 was estimated using the Hargreaves method (Pulido-Calvo *et al.*, 2003a,b).

The Fuente Palmera district can be considered representative of irrigation areas located in the interior of the Andalucía region (south Spain) (Rodríguez-Díaz *et al.*, 2004a). A fundamental characteristic of this type of irrigation district is that the mean irrigation water requirements are normally higher than measured values of consumer water demands (Pulido-Calvo *et al.*, 2003a,b; Lorite *et al.*, 2004; Pulido-Calvo *et al.*, 2007; Rodríguez-Díaz *et al.*, 2008). So, the average values for the index Annual Relative Water Supply (ARWS) (Malano *et al.*, 2004; Rodríguez-Díaz *et al.*, 2004b), which relates the total volume of water applied (irrigation plus rainfall) to the volume of water required by the crop (computed as gross irrigation requirements plus rainfall), are normally lower than 1. This indicates a deficit-irrigation situation. This indicator in the Fuente Palmera irrigation district was 0.6 ± 1.8.

The relatively high water productivity found in the Fuente Palmera district is an other characteristic of this type of irrigation district. This is due to a combination of deficit irrigation and the widespread use of extensive crops as cereals, cotton, sunflower and olive trees which efficiently use a substantial proportion of the annual rainfall in Mediterranean climates, thus lowering their irrigation requirements (Lorite *et al.*, 2004; Rodríguez-Díaz *et al.*, 2008).

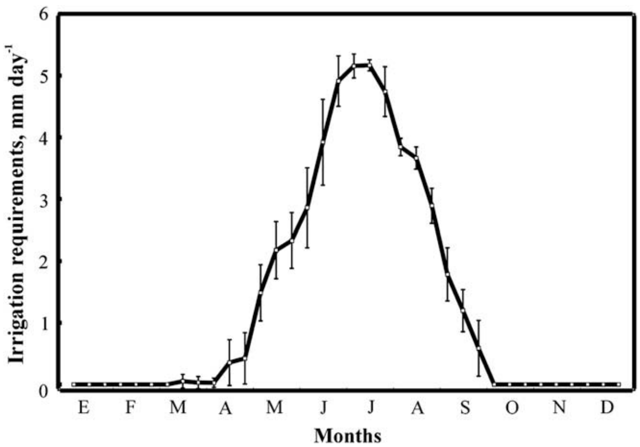


Fig. 2 – Water requirements of cropping pattern. Mean requirements in each 10-day period with their standard deviations are presented.

Actually, the operational control of the distribution system is based on averaged demand profiles for each 10-day period obtained from the experience and knowledge of the administrator. When there is a significant difference between the demand profile assumed and that which materialises as the day progresses, it is necessary to re-run the pump-scheduling software with the revised data. This may imply that there are pressure and flow non-acceptable values during certain time period of the system operation and effect network performance and crop yield. Consequently, the alternative is to develop an adaptive demand-forecasting process which can be updated daily (as the proposed approach in this paper).

The irrigation district selected and the water demand data used for hybrid model calibration and validation were the same that those utilised in the work of Pulido-Calvo *et al.* (2003a). In this study, the performance of CNN models was compared with linear multiple regression analysis and univariate time series models (exponential smoothing and ARIMA models) in predicting consumer demands for irrigation water. In CNN and multiple regression, the relationship between present and past water demand data and climatic and crop data (for previous days) was examined, that is to say that univariate and multivariate models were compared.

2.2. *Computational Neural Network models*

CNNs are mathematical models inspired by the neural architecture of the human brain. CNNs can recognise patterns and learn from their interactions with the environment. The most widely studied and used structures are multilayer feed forward networks (Rumelhart *et al.*, 1986). A typical four-layer feed forward CNN has g, n, m and s nodes or neurons in the input, first hidden, second hidden and output layers, respectively (the notation of the neural network is (g,n,m,s)). The parameters associated with each of the connections between nodes are called weights. All connections are feed forward, that is, they allow information transfer only from an earlier layer to the next consecutive layers.

To determine the set of weights, a corrective–repetitive process called learning or training of the CNN is performed. This training helps to define the interconnections between neurons (weights), and it is accomplished by using known inputs and outputs (training sets or training patterns), and presenting these to the CNN in some ordered manner, adjusting the interconnection weights until the desired outputs are reached. The strength of these interconnections is adjusted using an error convergence technique so that a desired output will be produced for a given input. There are many training methods. In this work, the same neural architectures trained with the Extended-Delta-Bar-Delta algorithm (EDBD) by Pulido-Calvo et al. (2003a) were re-calibrated with a variation of back-propagation algorithm (Rumelhart *et al.*, 1986), known as the Levenberg–Marquardt algorithm (LM) (Shepherd, 1997). This is a second-order non-linear optimisation algorithm with very fast convergence and has been recommended by several authors (Tan and Van Cauwenberghe, 1999; Anctil and Rat, 2005).

Let epoch denote the time period that encompasses all the iterations performed after all the patterns are displayed. In the study presented in this paper, the learning process was controlled by the method of internal validation (20% of calibration data to test the error at the end of each epoch) (Tsoukalas and Uhrig, 1997; Gutiérrez-Estrada *et al.*, 2004). The weights are updated at the end of each epoch. The number of epochs with the smallest error of the internal validation indicates the weights to select.

The optimal numbers of hidden layers and nodes in the hidden layers were determined by trial and error. CNNs with 1–2 hidden layers and 2–14 hidden nodes were successively trained based on the calibration data set. Each CNN architecture was trained with a pool of 10 repetitions due to the random initial values of weights. The CNN having the best performance when applied to the validation data set was selected. CNN models were implemented using STATISTICA 6.0 (Statsoft, Inc., 1984–2002).

The daily water demand data from 1988, 1989 and 1990 years were used for the model calibration (training in the CNN). To check model validation (generalisation capacity in the CNN), data from the year 1991 were used.

2.3. *CNN + fuzzy logic hybrid model*

Fuzzy logic is based on the mathematics of fuzzy set theory where the classical notion of binary set membership has been modified to include partial membership ranging between 0 and 1 (Zadeh, 1965). Fuzzy sets, in contrast to their crisp counterparts, have ambiguous boundaries and therefore gradual transitions between defined sets, allowing for the uncertainty associated with these concepts to be modelled directly. The construction of a fuzzy model implies, in a first step, to define for each model variable a series of overlapping fuzzy sets (or geometrical partitions) and the mapping of inputs to outputs expressed as a set of IF–THEN rules (Kosko, 1997). Subsequently, it is necessary to define the method that transforms the fuzzed inputs in defuzzed outputs or quantitative outputs.

The fuzzy sets and rules are referred as the fuzzy model knowledge base. Crisp inputs to the model are first fuzzified *via* this knowledge base (called Fuzzy Associative Memory or FAM), and a fuzzy inference engine is then used to process the rules in parallel *via* a fuzzy inference procedure such as the max–min or max–product operations (Jang *et al.*, 1997). The explicit relationship between the partitions of the input fuzzy sets and the output fuzzy set is stored in a FAM, which initially is created by a human expert on the basis of his experience.

In this paper, the fuzzy logic rule-based model has two input fuzzy sets (1: forecasts obtained from the best CNN, associated to five triangular partitions labelled as Very Low [VL], Low [LO], Normal [NO], High [HI] and Very High [VH], and 2: representative categories of monthly water demand patterns, associated to five singleton partitions labelled as Very Low [VL], Low [LO], Normal [NO], High [HI] and Very High [VH]) and one output fuzzy set (correction of forecasts obtained from the best CNN, associated to nine triangular partitions labelled as Very High Negative [VH−], High Negative [HI−], Normal Negative [NO−], Low Negative [LO−], Very Low [VL], Low Positive [LO+], Normal Positive [NO+], High Positive [HI+] and Very High Positive [VH+]).

Usually, there is often a large random component superimposed to the monthly water demand patterns, therefore, the data should be pre-processed in order to reduce this non-informative contribution and obtain representative prototypes to be used in the fuzzy reasoning algorithm. The selected method relies on the concept of fuzzy curve. The use of fuzzy curves was introduced by Lin and Cunningham (1995) and Lin et al. (1996) as a way to approximate a set of data with a minimum number of fuzzy rules. In this context, for the relationship between N possible inputs (water demand of N previous days) $Q_{t-N}, \ldots, Q_{t-1}$, and one output (water demand of day t) Q_t, the fuzzy curves were created by the following steps:

(a) Data at M points ($Q_{t-1,k}, \ldots, Q_{t-N,k}$; $Q_{t,k}$), $k = 1, 2, \ldots, M$, were collected to represent the possible relationship between the inputs ($Q_{t-1,k}, \ldots, Q_{t-N,k}$) and the output ($Q_{t,k}$). $N = 5$ was selected based on autocorrelation and partial autocorrelation functions of the demand series.

(b) For each previous day $(t-1, \ldots, t-N)$, a fuzzy membership function $\mu_{i,k}$ was created with i_j points and $i = 1$ for the first previous day $(t-1)$,…, $i = N$ for the Nth previous day $(t-N)$:

$$\mu_{i,k} = \exp\left[-\left(\frac{i_j - i}{b}\right)^2\right], \tag{1}$$

where b is a parameter controlling the spread of the membership function. In this case $b = 1$ (20% of the number of previous days) was selected (Lin and Cunningham, 1995; Marsili-Libelli, 2004).

(c) These fuzzy membership functions were defuzzified to produce a global fuzzy curve C for each month and each category (representative prototype of monthly water demand) for all input variables Q_i (water demand of N previous days) using

$$C = \frac{\sum_{k=1}^{M} \sum_{i=1}^{N} [Q_i \times \mu_{i,k}]}{\sum_{k=1}^{M} \sum_{i=1}^{N} \mu_{i,k}}. \tag{2}$$

As example, Fig. 3 shows how to determine the fuzzy curve of one day (k = constant). The water demand calibration data set (years 1988, 1989 and 1990) was ranged in five categories {Very High (VH): 219,457–274,320 $m^3\,day^{-1}$; High (HI): 164,593–219,456 $m^3\,day^{-1}$; Normal (NO): 109,729–164,592 $m^3\,day^{-1}$; Low (LO): 54,865–109,728 $m^3\,day^{-1}$; Very Low (VL): 0–54,864 $m^3\,day^{-1}$}. Fuzzy curves (or representative prototypes of monthly water demand) were obtained for each month and each category from daily water demand calibration data. Thus, for example, there are three fuzzy curves for April (NO; LO; VL) and four fuzzy curves for June (HI; NO; LO; VL) (Fig. 4).

For each daily water demand data from 1991 year (validation data), the fuzzy curve of one day k for the month m is compared by means the square root of the mean square error (RMS) error value with all representative prototypes of water demand for the month m. This comparison provides the representative category of monthly water demand pattern for day k of validation phase (Fig. 5).

After the inputs (forecasts obtained from the best CNN and representative categories of monthly water demand) have been obtained, the fuzzy solution resulting from the execution of the rule-base is defuzzed to produce the system output (correction of forecasts obtained from the best CNN). In this work, the method used to obtain the system output was the minimum rule. Also, a function that transforms the fuzzy output into a crisp value is necessary. In this case the defuzzication technique was the centre of area:

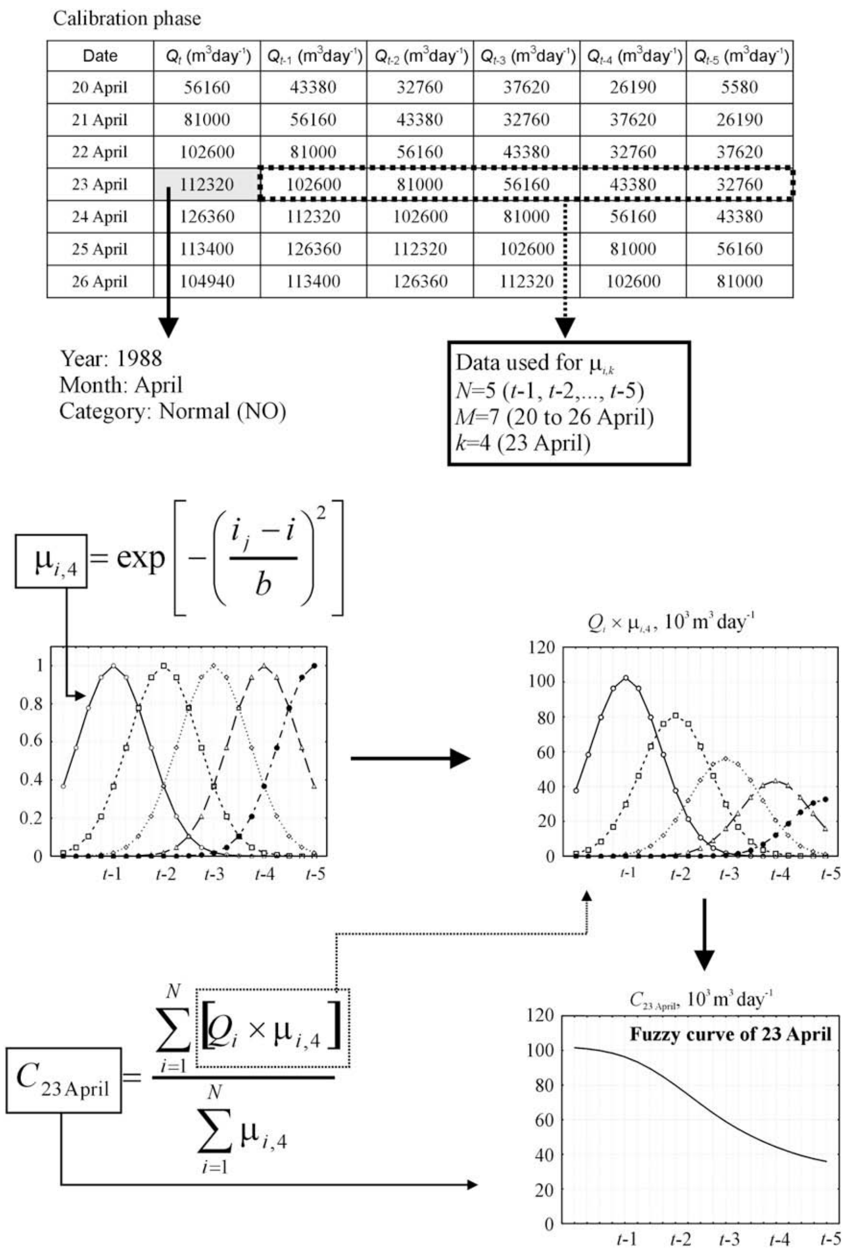
$$Y = \frac{\sum_{f=1}^{S} y^{\mathrm{Centre}(f)} \times \mu_{\mathrm{out}}(y^f)}{\sum_{f=1}^{S} \mu_{\mathrm{out}}(y^f)}, \tag{3}$$

where the crisp value (Y) is the geometrical centre $(y^{\mathrm{Centre}(f)})$ of the output fuzzy values $\mu_{\mathrm{out}}(y^f)$ with $f = 1, \ldots, S$, where S is formed by all contributions of rules whose degree of fulfilment is higher than zero. Fig. 6 shows a scheme in which the correction for one day of forecast obtained from the best CNN is determined.

2.4. *Optimising the fuzzy logic model with genetic algorithms*

In Section 2.3, how to construct a fuzzy logic model that can estimate the corrections of forecasts obtained from a CNN has been shown. The accuracy of this estimation depends on the parameters of the fuzzy logic model: (1) the shape of fuzzy sets or geometrical partitions; (2) the overlapping level of fuzzy sets; and (3) the definition of IF–THEN rules. A genetic algorithm has been used to find the optimal values of these parameters for the fuzzy logic model. Genetic algorithms are non-linear search and optimisation methods inspired by the biological processes of natural selection and survival of the fittest (Holland, 1975; Goldberg, 1989). This algorithm differs from traditional search methods because it considers many points in the search space simultaneously and therefore has a low probability of converging to local optimum.

In a genetic algorithm the basic unit is the gene. Various genes contain the information needed to define a chromosome whose decoding is interpreted as an individual. In this case, the parameters of the fuzzy logic model (boundaries and maximum value of fuzzy triangular membership functions, the overlapping level of fuzzy sets and the definition of IF–THEN rules) were coded as genes in the chromosome. A direct coding method was carried out in this paper (Fig. 7) (Jacob, 2001).



Calibration phase

Date	Q_t (m^{3}day^{-1})	Q_{t-1} (m^{3}day^{-1})	Q_{t-2} (m^{3}day^{-1})	Q_{t-3} (m^{3}day^{-1})	Q_{t-4} (m^{3}day^{-1})	Q_{t-5} (m^{3}day^{-1})
20 April	56160	43380	32760	37620	26190	5580
21 April	81000	56160	43380	32760	37620	26190
22 April	102600	81000	56160	43380	32760	37620
23 April	112320	102600	81000	56160	43380	32760
24 April	126360	112320	102600	81000	56160	43380
25 April	113400	126360	112320	102600	81000	56160
26 April	104940	113400	126360	112320	102600	81000

Fig. 3 – Example showing the determination of fuzzy curve of one day.

Once the initial information has been coded, three types of operators (reproduction, crossover and mutation) were used in order to evolve towards an optimal fuzzy configuration. Reproduction is a process where chromosomes with high fitness values in the generation t yield a high number of sons (copies) in the next generation. Crossover is an operator that mixes two chromosomes through a random process to take advantage the best qualities of each chromosome. The mutation operator changes the values of bits associated to a gene with a very low probability, which can produce unsuitable fuzzy configurations (abortions) or pre-adaptive fuzzy configurations that generate best solutions. In this paper the maximum probability of reproduction for a good individual was 0.7. The probabilities given to the crossover and mutation processes were 0.3 and 0.1, respectively. The mutation radius of GEN1 (CNN output) and GEN3 (Fuzzy output) were selected as 10% of the maximum value considered for each fuzzy set. The mutation radius of GEN4 (FAM) was 2.

A fitness function (a term used in genetic algorithms which is an objective function) is required to apply the genetic algorithm (Chen *et al.*, 2000). In this paper, the RMS error between observed and estimated water demand was used. A leave-one-out cross-validation procedure is used to optimise the unknown parameters of the fuzzy logic model. Cross-validation techniques estimate accuracy of a forecast model from a series of independent data sets over all the available data (year 1991). For a data sample of size K days, the

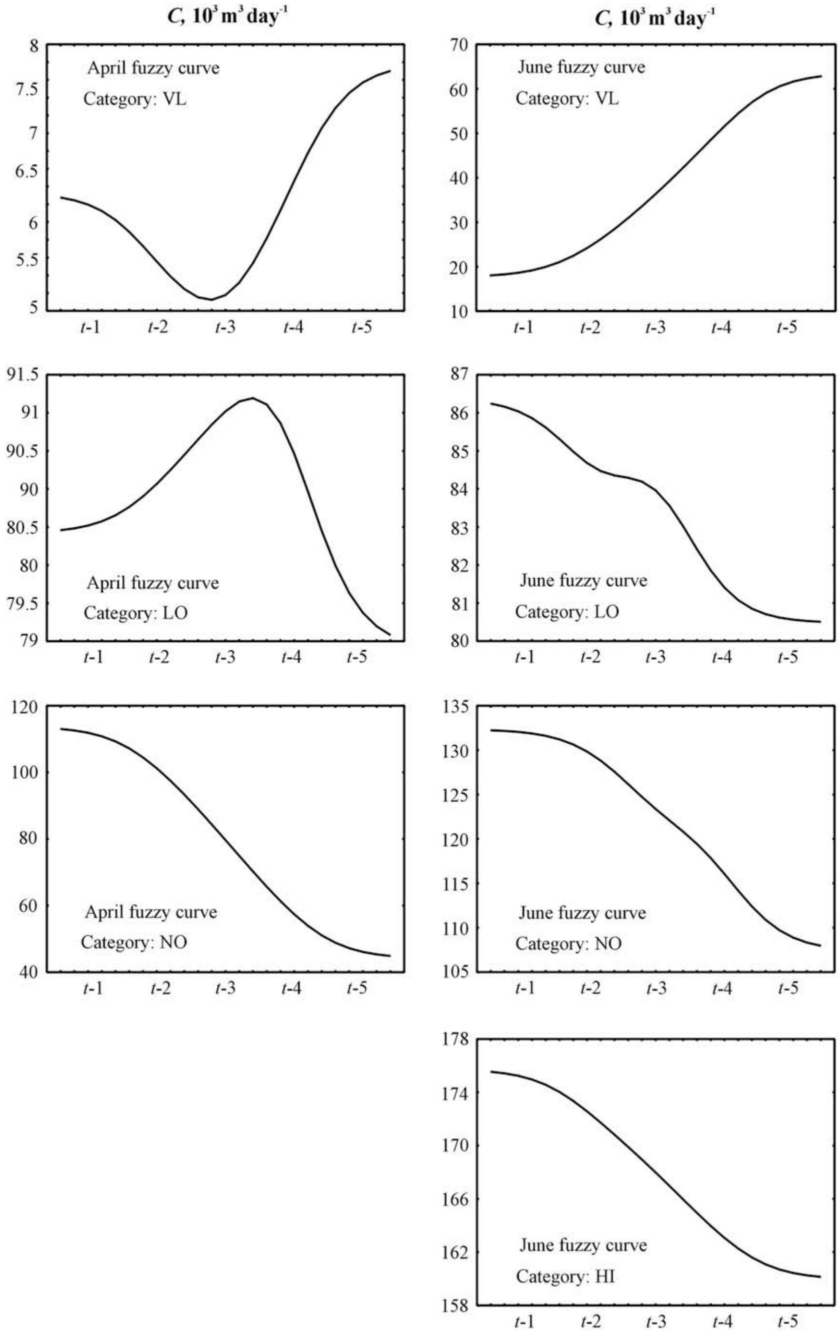


Fig. 4 – Representative prototypes of water demand of April and June months.

leave-one-out cross-validation method involves training the model using the $K-1$ days of data and then the forecasting error value for the day that have been left out is obtained. This is repeated K times until all days have been left out once. Finally, the model error (RMS error) was computed averaging the K forecasting error values (Goutte, 1997; Moradkhani *et al.*, 2004; Ruiz *et al.*, 2006).

2.5. *Measures of accuracy*

To assess the performance of the models to forecast daily irrigation water demands in delivery systems (CNN; CNN + fuzzy logic) during the validation phase, several measures of accuracy were applied (Yapo *et al.*, 1996; Legates and McCabe, 1999; Abrahart and See, 2000). The correlation between observed and predicted water demand was expressed by means of the correlation coefficient R. The coefficient of determination (R^2) describes the proportion of the total variance in the observed data that can be explained by the model. Other measures of variances applied were the percent standard error of prediction (SEP) (Ventura *et al.*, 1995) and the coefficient of efficiency (E_2) (Nash and Sutcliffe, 1970; Kitanidis and Bras, 1980).

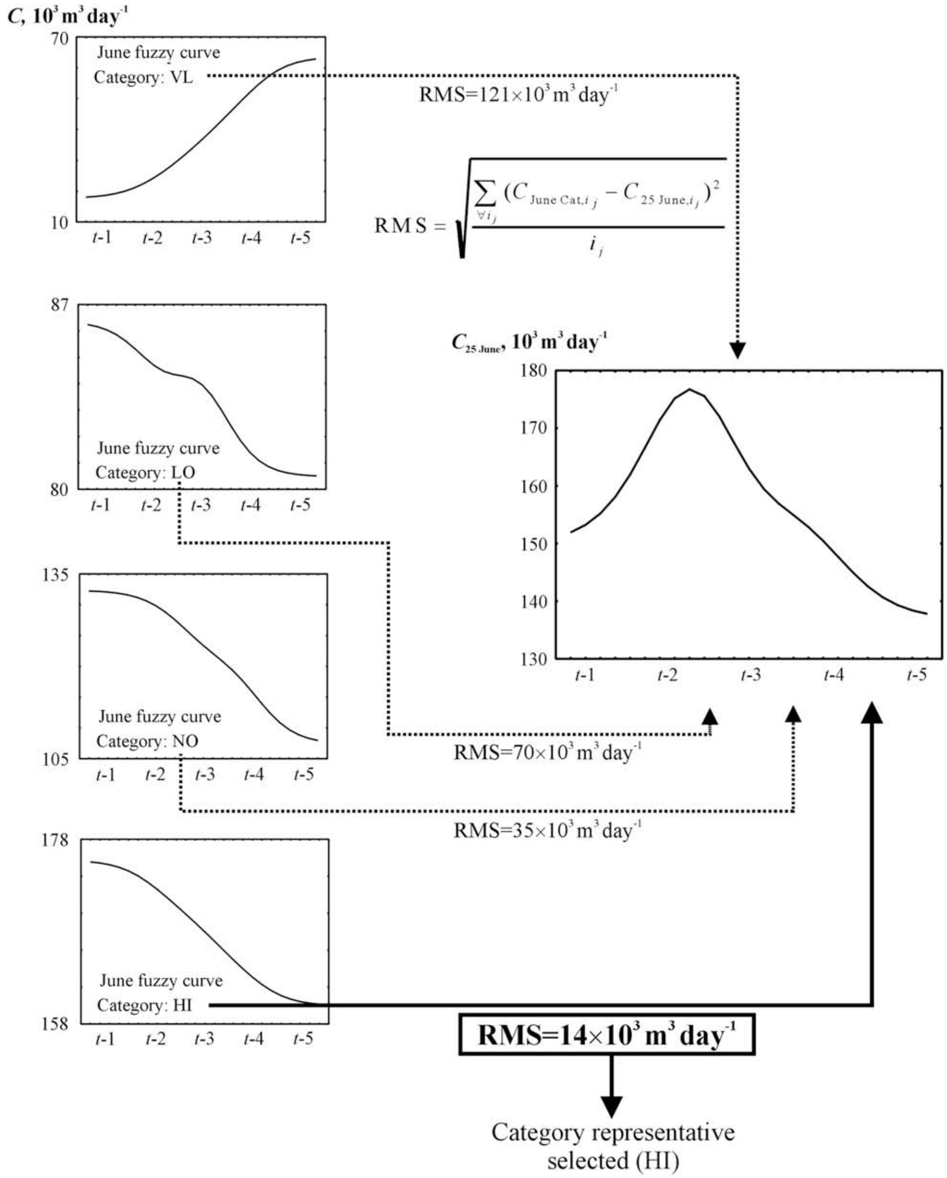


Fig. 5 – Example showing the comparison between the fuzzy curve of June 25 and all prototypes of water demand of June.

In addition, it is advisable to quantify the error in the same units of the variables. These measures, or absolute error measures, include the RMS given by

$$\mathrm{RMS} = \sqrt{\frac{\sum_{t=1}^{d}(Q_t - \widehat{Q}_t)^2}{d}}, \tag{4}$$

where Q_t is the observed water demand at the time step t; $\widehat{Q}_t$ is the estimated water demand at the same time step t; and d is the total number of observations of the validation set.

The SEP is defined by

$$\mathrm{SEP} = \frac{100}{\overline{Q}}\mathrm{RMS}, \tag{5}$$

where $\overline{Q}$ is the average of the observed water demand of the validation set. The coefficient of efficiency E_2 is used to see how the model explains the total variance of the data and represents the proportion of the variation of the observed data considered by the model. E_2 is given by

$$E_2 = 1.0 - \frac{\sum_{t=1}^{d}(Q_t - \widehat{Q}_t)^2}{\sum_{t=1}^{d}(Q_t - \overline{Q})^2}. \tag{6}$$

A value of zero for E_2 indicates that the observed average $\overline{Q}$ is as good predictor as the model, while negative values indicate that the observed average is a better predictor than the model (Legates and McCabe, 1999). For a perfect match, the values of R^2 and E_2 should be close to one and those of SEP close to zero.

Also the Persistence Index, PI, was used for the model performance evaluation (Kitanidis and Bras, 1980):

$$\mathrm{PI} = 1 - \frac{\sum_{t=1}^{d}(Q_t - \widehat{Q}_t)^2}{\sum_{t=1}^{d}(Q_t - Q_{t-L})^2}, \tag{7}$$

where Q_{t-L} is the observed water demand at the time step $t - L$ and L is the lead-time. In the applications carried out L was equal to one, since only one-day ahead forecasts were

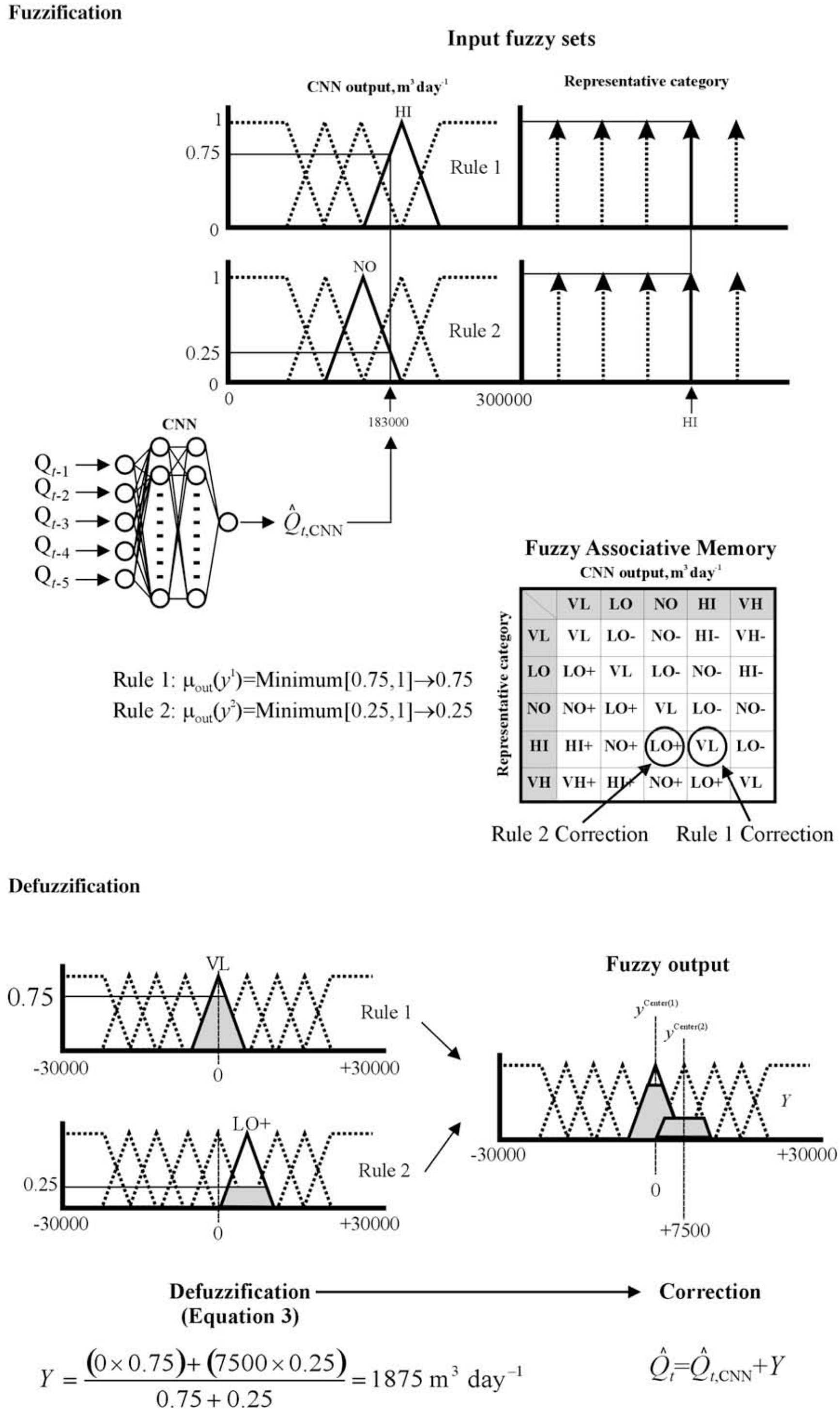


Fig. 6 – Performance scheme of CNN + fuzzy logic hybrid model.

performed. A PI value of one reflects a perfect adjustment between predicted and observed values, and a value of zero is equivalent to say that the model is no better than a naïve model, which always gives as prediction the data observed in the previous time period. A negative PI value would mean that the model is degrading the original information, thus denoting a performance worse than the one of the naïve model (Anctil and Rat, 2005).

For each measure of accuracy the benchmark of the worst permissible error was calculated. McLaughlin (1983) suggests that a naïve model determines the forecasting accuracy benchmark of any model. The basic naïve model, known as Naïve Forecast I (NFI), is defined as the next period's level will be the same as that of the preceding period. This way, if the forecasting model cannot do better than NFI, it should be disqualified.

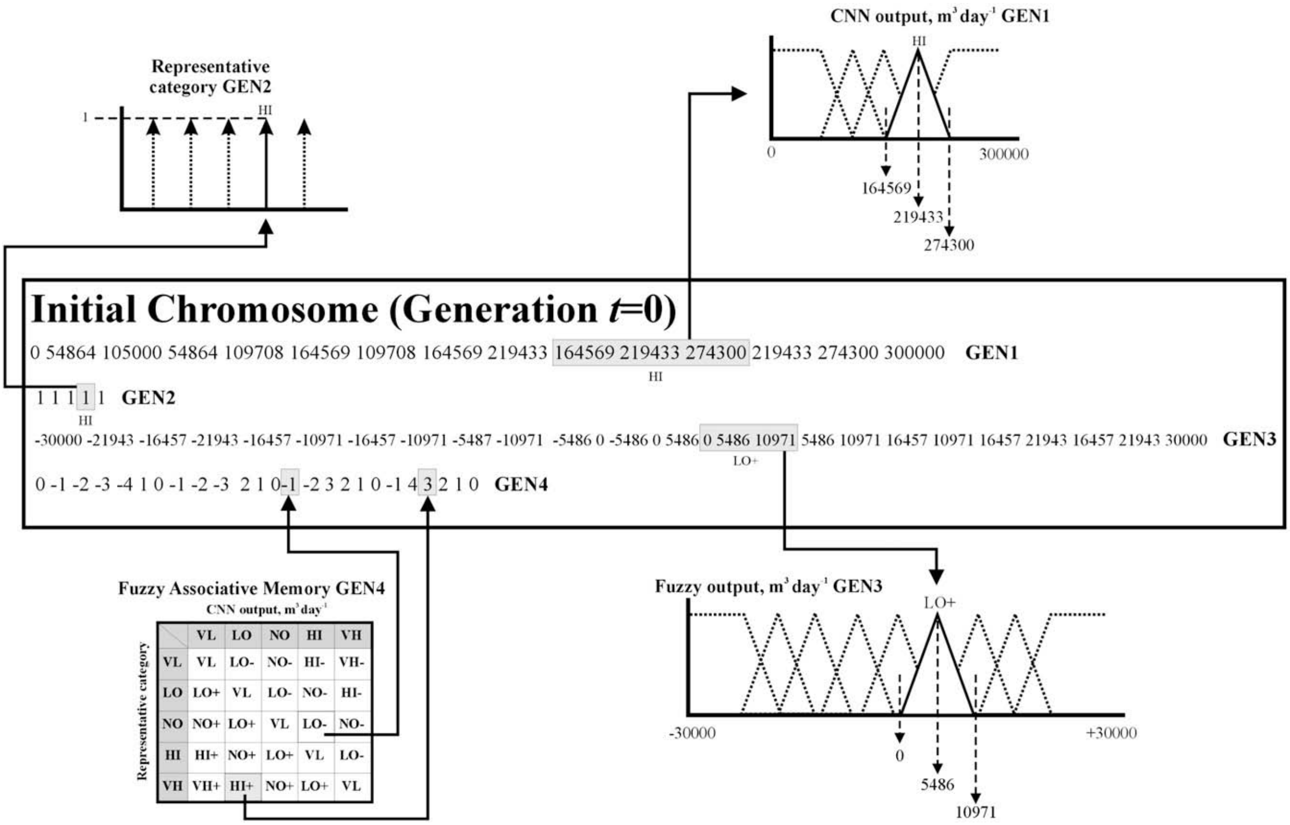


Fig. 7 – Scheme of the initial chromosome configuration.

3. Results

The potential of ARIMA and CNN models for short-term control of daily volumes to pump in irrigation water distribution systems was presented by Pulido-Calvo *et al.* (2003a). In this work, the best estimates were obtained with a multivariate CNN model trained with EDBD algorithm, considering the demands and the maximum temperatures of the two previous days as inputs and 12 nodes in each hidden layer [CNN(4,12,12,1) EDBD]. The best ARIMA model [ARIMA(1,1,2)] produced estimates with higher correlation and determination coefficients. However, this increase of the explained variance level was not associated with an improvement of the error magnitudes (RMS, SEP, E_2 and PI). With respect to univariate autoregressive CNN models with water demand of five previous days as input data, the best result was reached with 10 nodes in each hidden layer [CNN(5,10,10,1) EDBD]. This univariate neural approach provided worse accuracy measures than those obtained with ARIMA(1,1,2) except in RMS, SEP and PI values (Table 1).

The best estimates of the univariate CNN model re-calibrated with LM algorithm were also obtained with an architecture of 10 nodes in the first and second hidden layers [CNN(5,10,10,1) LM]. In this case, improvements in relation to the models mentioned above were not reached. Additionally, the general behaviour of this model was even worse than the NFI model (Table 1).

For all previous cases, the main error source was due to the occurrence of systematic displacement between estimated and observed water demands, that is to say that these models led to predictions in day t very close to the observed water demands in day $t-1$, as is shown in the exemplification of one-step-ahead prediction of water demand for year 1991 for the univariate CNN models trained with EDBD and LM algorithms [CNN(5,10,10,1) EDBD; CNN(5,10,10,1) LM] (Fig. 8). This way, the PI values were not higher than 0.15 for univariate models and were only slightly higher (PI = 0.21) for the best multivariate CNN model [CNN(4,12,12,1) EDBD] (Table 1).

When the developed hybrid methodology combining CNN, fuzzy logic and genetic algorithms was applied to the best univariate CNN model [CNN(5,10,10,1) EDBD], significant improvements of forecasts were achieved. With this hybrid model [CNN(5,10,10,1) EDBD + Fuzzy], the explained variance increased by 8% reaching a level of 89%. Also, best error terms were obtained in the validation phase. The RMS value decreased 7465 $m^3\,day^{-1}$, which supposed an SEP value slightly greater than 20% and an efficiency coefficient very close to 0.9. However, the most important improvement was observed in the PI which means that the water demand in each day resulting from hybrid model did not necessarily approach the water demand in the previous day, as happened in the other cases mentioned above. That is to say, the hybrid model proved to be able to eliminate in many of the validation phase estimates the systematic displacement between observed and forecasted water demands (Table 1; Fig. 8).

Table 1 – Accuracy measures calculated in the validation phase

Model	R	R^2	RMS, 10^3 m^3 day^{-1}	SEP, %	E_2	PI
NFI	0.88	0.78	34.9	28.99	0.77	0
ARIMA(1,1,2)[a]	0.93	0.86	32.2	26.81	0.82	0.13
CNN(5,10,10,1) EDBD[b]	0.90	0.81	31.8	26.48	0.81	0.15
CNN(4,12,12,1) EDBD[c]	0.91	0.82	30.7	25.50	0.82	0.21
CNN(5,10,10,1) LM	0.88	0.77	35.7	29.67	0.76	−0.06
CNN (5,10,10,1) EDBD[b] + Fuzzy	0.94[d]	0.89[d]	24.4[d]	20.27[d]	0.89[d]	0.51[d]
CNN(5,10,10,1) LM + Fuzzy	0.93	0.86	27.2	22.65	0.86	0.38

NFI row indicates the accuracy measures calculated for Naïve Forecast I model. a, b and c results were obtained from Pulido-Calvo et al. (2003a).
a Best ARIMA model.
b Best univariate autoregressive CNN model with water demands of five previous days as input data, trained with EDBD algorithm.
c Best multivariate autoregressive CNN model with water demands and maximum temperatures of two previous days as input data, trained with EDBD algorithm.
d Best results.

Additionally, the hybrid methodology proposed in this paper was very useful when the performance of the non-hybridised CNN model was not statistically acceptable. This was the case of the univariate CNN model trained with LM algorithm [CNN(5,10,10,1) LM]. Its hybridisation with the fuzzy model [CNN(5,10,10,1) LM + Fuzzy] allowed to obtain values for all accuracy measures significantly better than benchmark values of NFI model and even higher that those obtained with the best multivariate CNN model [CNN(4,12,12,1) EDBD] (Table 1; Fig. 8).

The schematic representation of the forecasted water demands as a function of the observed water demands (scatterplots between observed and forecasted water demands) showed that the univariate CNN models [CNN(5,10,10,1) EDBD; CNN(5,10,10,1) LM] presented the highest dispersion along the line 1:1 (that would correspond to the perfect adjustment between the observed and estimated water demands), while the hybrid models [CNN(5,10,10,1) EDBD + Fuzzy; CNN(5,10,10,1) LM + Fuzzy] denoted a higher approximation of observed and estimated values. Scatterplots for univariate CNN models and hybrid models are presented in Fig. 8 that also includes the diagrams of the observed and forecasted water demands from year 1991 (validation period). This figure shows that the hybrid model [CNN(5,10,10,1) EDBD + Fuzzy] presented the closest match between forecasted and observed water demands over the entire water demand range.

The final optimised fuzzy sets and the corresponding rule-base of the best hybrid model [CNN(5,10,10,1) EDBD + Fuzzy] appear in Fig. 9. In the case of the partitions of GEN1 (CNN output), it is possible to observe that uncertainty level increases as the water demand is higher. So the overlapping of the fuzzy sets Very Low [VL] and Low [LO] is minimum compared with the overlapping between High [HI] and Very High [VH]. Nevertheless, the highest overlapping (between High [HI] and Very High [VH]) was lower than initially established. Globally, this implies that the uncertainty associated with the CNN output was smaller than originally proposed. Moreover the genetic algorithm bred different boundaries and maximum values for each membership function.

An evaluation of the Fuzzy Associate Memory (FAM) which was modified by the genetic algorithm showed a reasonable behaviour (Fig. 9). So in a general form, above principal diagonal of FAM matrix, the categories associated with the corrections of forecasts obtained from the neural models had negative values. Just the opposite was found below principal diagonal. On the other hand, the uncertainty level associated to GEN3 (Fuzzy output) was significantly higher than the one obtained in GEN1 and than the one originally anticipated. This reasoning was more marked when the correction of CNN forecast was between −36,000 and 5000 m^3 day^{-1}. In this range, the genetic algorithm removed the Low Negative [LO−] category, which is *quasi* contained among Normal Negative [NO−] and Very Low [VL] categories. This way, any fuzzy rule was associated to this category in the FAM matrix.

4. Discussion

The improvement of water management in an irrigation district requires the analysis of water demand in order to determine ways in which it may be modified and rationalised with a view to making water management policies more efficient and also can provide reference data for the design, modernisation and exploitation of water-delivery systems. For this reason, in the work of Pulido-Calvo *et al.* (2003a) approaches were examined based on linear multiple regressions, univariate time series models (exponential smoothing and ARIMA models) and univariate–multivariate autoregressive CNNs for consumer demand modelling and prediction in a short-term environment for an on-demand irrigation water distribution system. The goal of this paper was to assess possible improvements by using a hybrid methodology combining CNNs, fuzzy logic and genetic algorithms in one-step-ahead daily water demand forecasting.

The results revealed that the hybrid methodology was characterised by a higher accuracy in terms of all standard and relative statistical measures and unbiased forecasts. Additionally, the best hybrid model explained 89% of the variance, produced forecasts that had SEP values less than 21% and with the capacity to predict the amplitude, start and end of the irrigation season. These results were better than those obtained by Pulido-Calvo *et al.* (2003a) in the same irrigation district using the same calibration and validation data and than those obtained by Griñó (1992) in the forecasting of

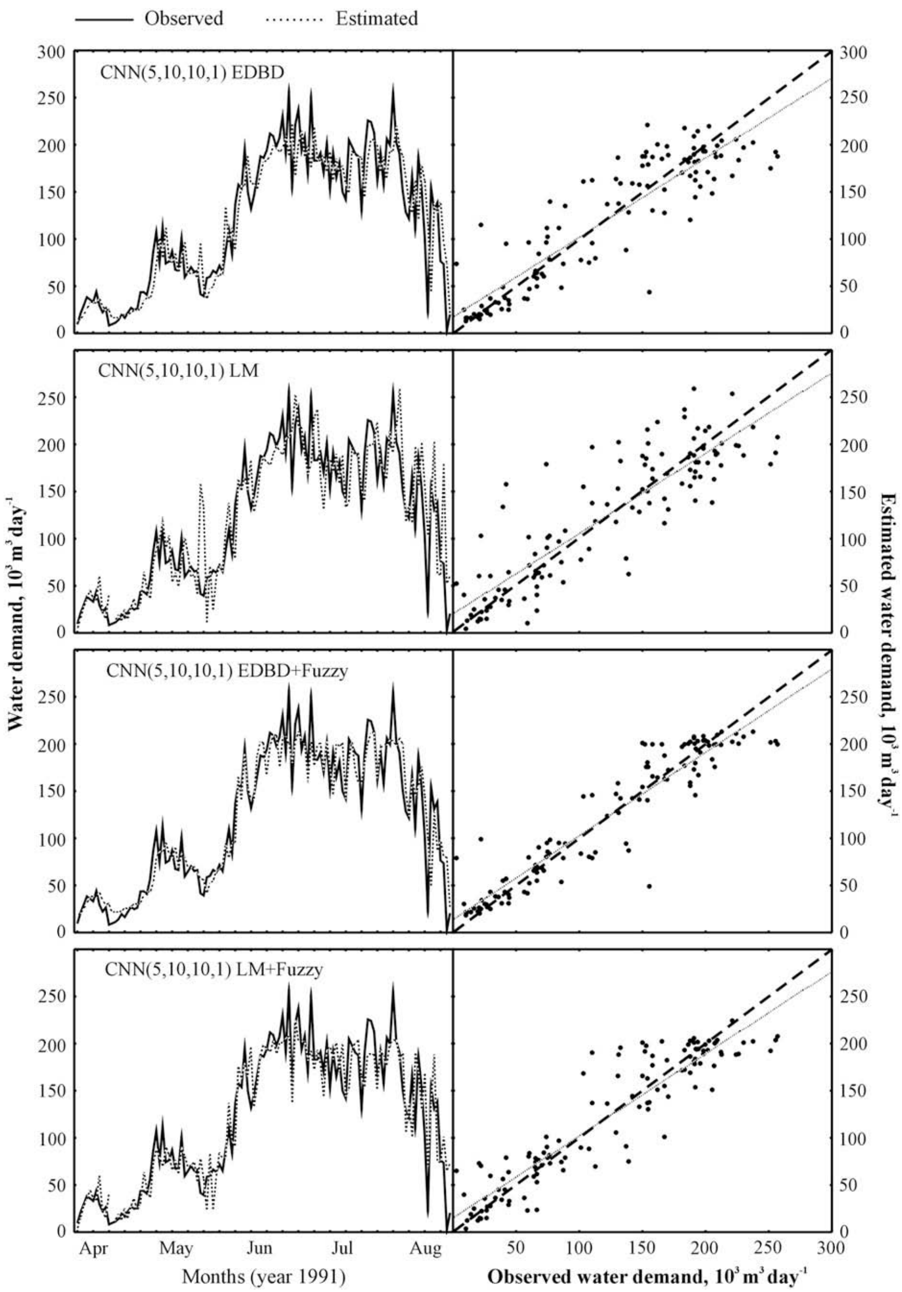
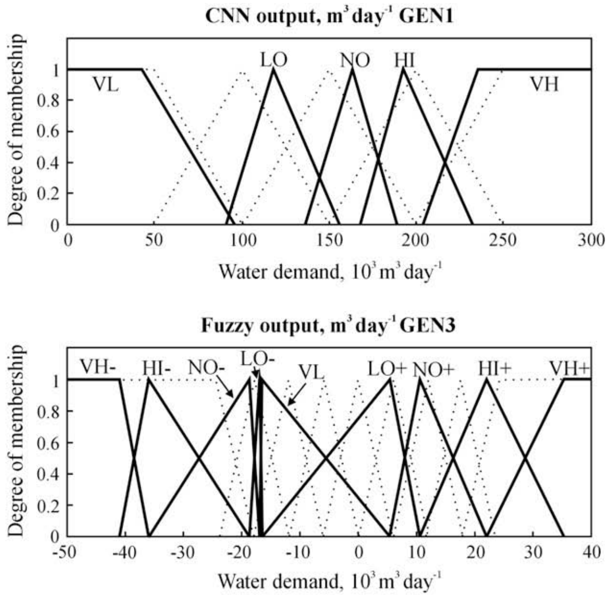


Fig. 8 – Exemplification of one-step-ahead prediction of water demand for year 1991, and scatterplot between observed and estimated water demand (validation period).

daily demand time series in an urban water system. In these works, additional inputs to water demand data as maximum temperatures of previous days or seasonal factors were added to improve the learning capacity of CNNs. These comparatives provide evidences that the hybrid methodology developed in this paper for water demand forecasting can offer a higher degree of reliability and accuracy considering only as input data water uses in previous time steps.

Moreover, the improvements reached in goodness-of-fit with the hybrid model developed in this paper regarding to traditional neural approaches were slightly higher than the results obtained by Zhang (2003) and Pulido-Calvo and Portela (2007) in different time series forecasting combining linear and non-linear models (ARIMA and CNN), and similar to those obtained by Chang *et al.* (2005) for flood forecasting. In this last case, back-propagation neural networks were compared with a hybrid version of EACH learning algorithm and a fuzzy inference system. The higher forecast capacity of hybrid models as the one proposed in this paper can be related with the presence of non-stables or changing patterns typically included in time series data which add a high level of uncertainty that cannot be extracted by inference techniques such as ARIMA models or autoregressive CNNs independently.



Fuzzy Associative Memory GEN4
CNN output, $m^3 day^{-1}$

Representative category	VL	LO	NO	HI	VH
VL	LO+	LO+	NO-	HI-	HI-
LO	LO+	VH-	VH-	VL	VH-
NO	NO+	VH+	VH+	HI-	VH-
HI	HI+	VL	VH+	NO+	VH-
VH	HI+	VH+	NO+	VH-	VH-

Fig. 9 – Final fuzzy sets and rule-base (FAM) configuration.

The higher forecast capacity of hybrid models such as the one proposed in this paper is related to intrinsic properties associated with the types of models hybridised. On one hand, the artificial neural networks have the capacity to combine non-linear relationships between variables and their topologies allow a parallel inter-relationship of input variables. This provides to CNNs a great flexibility to represent several non-linear relationships at the same time, which is very problematic for classic estimation techniques. On second hand, the systems based on fuzzy logic rules present a high error tolerance, behaving as conservative systems. This is related with the capacity of these types of systems to deal with continuous changes and generate small or large control, which is determined by the simultaneous execution of several rules whose influence degrees depend on the input parameter strength in the fuzzy sets (Lee *et al.*, 2000).

An appropriate memory must be added to a multilayer feed forward network so it be used in the time series forecasting as consequence of its static architecture. The simplest form of memory consists of a buffer that contains multiple copies of the input data at various time delays (in this case, the water demand in previous days) (De Vries and Principe, 1991). In the problem analysed, despite the memory added to the CNNs by means of a buffer containing recent inputs (the water demands in the five previous days), the forecasts provided in each day were systematically very close to the water demands observed in the previous day. This circumstance is shown in the PI values that were close to 0 and even negative for the best univariate autoregressive CNNs proved. This is probably due to the fact that the correlation between observed values in any two consecutive time periods was most of the time very high, and so, each time occurrence is highly responsible for the next time realisation.

Authors as Park (1998), Abrahart and See (2000), Pulido-Calvo *et al.* (2003a), Gutiérrez-Estrada *et al.* (2005), Pulido-Calvo and Portela (2007), Gutiérrez-Estrada *et al.* (2007) and Pulido-Calvo *et al.* (2007) also reported this lag-one difference between actual values and values resulting from different models applied to forecasting of different kinds of time variables. One way to decrease this systematic displacement is providing additional or external variables to univariate autoregressive models as was shown with the multivariate CNN model with water demands and maximum temperatures of two previous days as input data. With this multivariate CNN model improvements were obtained in all the accuracy measures in comparison with the univariate CNN model. These improvements were significant for the PI value with an increase of 30% in comparison with the rest of accuracy measures (increases of 1% for R^2 and E_2 values and of 4% for RMS and SEP values). However, with the hybrid methodology developed in this paper higher improvements were obtained (increases of 8% for R^2, of 9% for E_2, of 23% for RMS and SEP values and of 71% for PI value).

Actually, the operational control of most irrigation water distribution systems is based on averaged demand profiles for certain time periods (10-day or weekly periods, normally), which vary according to the development stage of the cropping pattern and the climatic conditions, based on the experience and knowledge of the administrator. When there is a significant difference between the demand profile assumed and that which materialises as day progresses, it is necessary to re-run the pump-scheduling software with the revised data. This may imply that there are non-acceptable values of pressure and flow during certain time periods of the system operation which may affect network performance and even crop yield. Consequently, the alternative is to develop an approach similar to the one developed in this paper that implies an adaptive demand-forecasting process which can be updated daily. Also, this approach for the short-term operational control of the water distribution system is more appropriate since the pump and valve settings need to be re-optimised at short, regular intervals, in response to the highly variable water demands.

5. Conclusions

The objective of this paper was to forecast consumer demands of an irrigation area using on-farm water-use information and a hybrid methodology combining CNNs, fuzzy logic and genetic algorithms. In this model, a fuzzy inference system was used to estimate the corrections of forecasts obtained from a univariate autoregressive neural network and a genetic algorithm was used to find the optimal values of the parameters of the fuzzy system.

This hybrid model was applied to predict water demand one-day ahead in the Fuente Palmera irrigation district, southern Spain. The results indicated that the hybrid model had much better statistics and error measures than those of multivariate and univariate autoregressive neural networks models.

In summary, this hybrid model proved to be a powerful tool that, with not very large data requirements, can be very suitable for the development of policies on irrigation water consumption since information regarding water demand is key to schedule pumping efforts and minimise operation costs of water distribution systems as well as to evaluate the marginal value of irrigation water and the response level to different irrigation water rates.

REFERENCES

Abrahart R J; See L; Kneale P E (1999). Using pruning algorithms and genetic algorithms to optimize network architectures and forecasting inputs in a neural network rainfall–runoff model. Journal of Hydroinformatics, **1**(2), 103–114.

Abrahart R J; See L (2000). Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. Hydrological Processes, **14**, 2157–2172.

Abrahart R J; See L (2002). Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments. Hydrology and Earth System Sciences, **6**(4), 655–670.

Agarwal A; Mishra S K; Ram S; Singh J K (2006). Simulation of runoff and sediment yield using artificial neural networks. Biosystems Engineering, **94**(4), 597–613.

Allen R G; Pereira L S; Raes D; Smith M (1998). Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage Paper 56. FAO, Rome, Italy.

Al-Saba T; El-Amin I (1999). Artificial neural networks as applied to long-term demand forecasting. Artificial Intelligence in Engineering, **13**(2), 189–197.

Alvisi S; Franchini M; Marinelli A (2007). A short-term, pattern-based model for water-demand forecasting. Journal of Hydroinformatics, **9**(1), 39–50.

Anctil F; Rat A (2005). Evaluation of neural network streamflow forecasting on 47 watersheds. Journal of Hydrologic Engineering, **10**(1), 85–88.

Bougadis J; Adamowski K; Diduch R (2005). Short-term municipal water demand forecasting. Hydrological Processes, **19**, 137–148.

Calejo M J; Lamaddalena N; Teixeira J L; Pereira L S (2008). Performance analysis of pressurized irrigation systems operating on-demand using flow-driven simulation models. Agricultural Water Management, **95**, 154–162.

Cameron D; Kneale P; See L (2002). An evaluation of a traditional and a neural net modelling approach to flood forecasting for an upland catchment. Hydrological Processes, **16**, 1033–1046.

Chang L C; Chang F J; Tsai Y H (2005). Fuzzy exemplar-based inference system for flood forecasting. Water Resources Research, **41**, W02005. doi:10.1029/2004WR003037.

Chen D G; Hargreaves N B; Ware D M; Liu Y (2000). A fuzzy logic model with genetic algorithm for analyzing fish stock–recruitment relationship. Canadian Journal of Fisheries and Aquatic Sciences, **57**, 1878–1887.

Chiang Y M; Chang L C; Chang F J (2004). Comparison of static-feedforward and dynamic-feedbackward neural networks for rainfall–runoff modeling. Journal of Hydrology, **290**, 297–311.

Clément R (1966). Calcul des débits dans les réseaux d'irrigation fonctionnant à la demande. Houille Blanche, **20**(5), 553–575.

Clément R; Galand A (1979). Irrigation par Aspersion et Réseaux Collectifs de Distribution Sous Pression. Editions Eyrolles, París.

De Vries B; Principe J C (1991). A theory for neural networks with time delays. In, **vol 3**. Morgan Kaufmann Publishers, California.

Doorenbos J; Kassam A H (1979). Yield Response to Water. FAO Irrigation and Drainage Paper 33. FAO, Rome, Italy.

Doorenbos J; Pruitt W O (1977). Guidelines for Predicting Crop Water Requirements. FAO Irrigation and Drainage Paper 24. FAO, Rome, Italy.

Firat M; Güngör M (2008). Hydrological time-series modelling using an adaptive neuro-fuzzy inference system. Hydrological Processes, **22**, 2122–2132.

Firat M; Yurdusev M A; Turan M E (2008). Evaluation of artificial neural network techniques for municipal water consumption modeling. Water Resources Management. doi:10.1007/s11269-008-9291-3.

French M N; Krajewski W F; Cuykendall R R (1992). Rainfall forecasting in space and time using a neural network. Journal of Hydrology, **137**, 1–31.

Ginzburg I; Horn D (1994). Combined neural networks for time series analysis. Advances in Neural Information Processing Systems, **6**, 224–231.

Goldberg D (1989). Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Reading, Massachusetts, USA.

Goutte C (1997). Note on fee lunches and cross-validation. Neural Computation, **9**, 1211–1215.

Griñó R (1992). Neural networks for univariate time series forecasting and their application to water demand prediction. Neural Network World, **2**(5), 437–450.

Guan J; Aral M M (2005). Remediation system design with multiple uncertain parameters using fuzzy sets and genetic algorithm. Journal of Hydrologic Engineering, **10**(5), 386–394.

Gutiérrez-Estrada J C; De Pedro-Sanz E; López-Luque R; Pulido-Calvo I (2004). Comparison between traditional methods and artificial neural networks for ammonia concentration forecasting in an eel (*Anguilla anguilla* L.) intensive rearing system. Aquacultural Engineering, **31**, 183–203.

Gutiérrez-Estrada J C; De Pedro-Sanz E; López-Luque R; Pulido-Calvo I (2005). Estimación a corto plazo de la temperatura del agua. Aplicación en sistemas de producción en medio acuático. [Water temperature short-term forecasting. Application in aquaculture systems]. Ingeniería del Agua, **12**(1), 77–92.

Gutiérrez-Estrada J C; Silva C; Yáñez E; Rodríguez N; Pulido-Calvo I (2007). Monthly catch forecasting of anchovy *Engraulis ringens* in the north area of Chile: non-linear univariate approach. Fisheries Research, **86**, 188–200.

Holland J (1975). Adaptation in Natural and Artificial Systems. MIT Press, Cambridge, Massachusetts, USA.

Hsu K; Gupta H V; Sorooshian S (1995). Artificial neural network modeling of the rainfall–runoff process. Water Resources Research, **31**(10), 2517–2530.

Jacob C (2001). Illustrating Evolutionary Computation with Mathematica. Morgan Kaufmann Publishers, San Francisco, California, USA.

Jain A; Varshney A K; Joshi U C (2001). Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks. Water Resources Management, **15**(5), 299–321.

Jang J S R; Sun C T; Mizutani E (1997). Neuro-Fuzzy and Soft Computing. Prentice Hall, New Jersey, USA.

Keskin M E; Taylan D; Terzi Ö (2006). Adaptive neural-based fuzzy inference system (ANFIS) approach for modelling hydrological time series. Hydrological Sciences Journal, **51**(4), 588–598.

Khadra R; Lamaddalena N (2006). A simulation models to generate the demand hydrographs in large-scale irrigation systems. Biosystems Engineering, **93**(3), 335–346.

Kişi Ö; Öztürk Ö (2007). Adaptive neurofuzzy computing technique for evapotranspiration estimation. Journal of Irrigation and Drainage Engineering, **133**(4), 368–379.

Kitanidis P K; Bras R L (1980). Real time forecasting with a conceptual hydrological model. 2: Applications and results. Water Resources Research, **16**(6), 1034–1044.

Kosko B (1997). Fuzzy Engineering. Prentice Hall, London, UK.

Kuligowski R J; Barros A P (1998). Experiments in short-term precipitation forecasting using artificial neural networks. Monthly Weather Review, **126**(2), 470–482.

Lamaddalena N; Fratino U; Daccache A (2007). On-farm sprinkler irrigation performance as affected by the distribution system. Biosystems Engineering, **96**(1), 99–109.

Lee P G; Lea R N; Dohmann E; Prebilsky W; Turk P E; Ying H; Whitson J L (2000). Denitrification in aquaculture systems: an example of a fuzzy logic control problem. Aquacultural Engineering, **23**, 37–59.

Legates D R; McCabe Jr. G J (1999). Evaluating the use of ''goodness-of-fit'' measures in hydrologic and hydroclimatic model validation. Water Resources Research, **35**(1), 233–241.

Lin Y; Cunningham G A (1995). A new approach to fuzzy-neural system modeling. IEEE Transactions on Fuzzy Systems, **3**(2), 190–198.

Lin Y; Cunningham G A; Coggeshall S V (1996). Input variable identification-fuzzy curves and fuzzy surfaces. Fuzzy Sets and Systems, **82**, 65–71.

Lorite I J; Mateos L; Fereres E (2004). Evaluating irrigation performance in a Mediterranean environment. I: Model and general assessment of an irrigation scheme. Irrigation Science, **23**, 77–84.

Lorrai M; Sechi G M (1995). Neural nets for modelling rainfall–runoff transformations. Water Resources Management, **9**, 299–313.

Luxhoj J T; Riis J O; Stensballe B (1996). A hybrid econometric-neural network modeling approach for sales forecasting. International Journal of Production Economics, **43**, 175–192.

Malano H; Burton M; Makin I (2004). Benchmarking performance in the irrigation and drainage sector: a tool for change. Irrigation and Drainage, **53**, 119–133.

Marsili-Libelli S (2004). Fuzzy prediction of the algal blooms in the Orbetello lagoon. Environmental Modelling and Software, **19**, 799–808.

Mason J C; Tem'me A; Price R K (1996). A neural network model of rainfall–runoff using radial basis functions. Journal of Hydraulic Research, **34**(4), 537–548.

McLaughlin R L (1983). Forecasting models: sophisticated or naïve? Journal of Forecasting, **2**(3), 274–276.

Moradkhani H; Hsu K; Gupta H V; Sorooshian S (2004). Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. Journal of Hydrology, **295**, 246–262.

Moreno M A; Planells P; Ortega J F; Tarjuelo J (2007). New methodology to evaluate flow rates in on-demand irrigation networks. Journal of Irrigation and Drainage Engineering, **133**(4), 298–306.

Nash J E; Sutcliffe J V (1970). River flow forecasting through conceptual models. I: A discussion of principles. Journal of Hydrology, **10**, 282–290.

Nayebi M; Khalili D; Amin S; Zand-Parsa Sh (2006). Daily stream flow prediction capability of artificial neural networks as influenced by minimum air temperature data. Biosystems Engineering, **95**(4), 557–567.

Ondimu S; Murase H (2007). Reservoir level forecasting using neural networks: lake Naivasha. Biosystems Engineering, **96**(1), 135–138.

ONU (1997). Comprehensive Assessment of the Freshwater Resources of the World. United Nations Department for Policy Coordination and Sustainable Development (DPCSD), Commission on Sustainable Development, New York, USA.

Park H H (1998). Analysis and prediction of walleye pollock (*Theragra chalcogramma*) landings in Korea by time series analysis. Fisheries Research, **38**, 1–7.

Pelikan E; De Groot C; Wurtz D (1992). Power consumption in West-Bohemia: improved forecasts with decorrelating connectionist networks. Neural Network World, **2**, 701–712.

Pereira L S (1999). Higher performance through combined improvements in irrigation methods and scheduling: a discussion. Agricultural Water Management, **40**, 153–169.

Poff L N; Tokar A S; Johnson P A (1996). Stream hydrological and ecological responses to climate change assessed with an artificial neural network. Limnology and Oceanography, **41**(5), 857–863.

Pulido-Calvo I; Roldán J; López-Luque R; Gutiérrez-Estrada J C (2002). Técnicas de predicción a corto plazo de la demanda de agua. Aplicación al uso agrícola. [Short-term forecasting techniques of water demand. Application to agricultural use]. Ingeniería del Agua, **9**(3), 319–331.

Pulido-Calvo I; Roldán J; López-Luque R; Gutiérrez-Estrada J C (2003a). Demand forecasting for irrigation water distribution system. Journal of Irrigation and Drainage Engineering, **129**(6), 422–431.

Pulido-Calvo I; Roldán J; López-Luque R; Gutiérrez-Estrada J C (2003b). Water delivery system planning considering irrigation simultaneity. Journal of Irrigation and Drainage Engineering, **129**(4), 247–255.

Pulido-Calvo I; Montesinos P; Roldán J; Ruiz-Navarro F (2007). Linear regressions and neural approaches to water demand forecasting in irrigation districts with telemetry systems. Biosystems Engineering, **97**, 283–293.

Pulido-Calvo I; Portela M M (2007). Application of neural approaches to one-step daily flow forecasting in Portuguese watersheds. Journal of Hydrology, **332**, 1–15.

Rodríguez-Díaz J A; Camacho-Poyato E; López-Luque R (2004a). Application of data envelopment analysis to studies of irrigation efficiency in Andalusia. Journal of Irrigation and Drainage Engineering, **130**(3), 175–183.

Rodríguez-Díaz J A; Camacho-Poyato E; López-Luque R (2004b). Application benchmarking and data envelopment analysis (DEA) techniques to irrigation districts in Spain. Irrigation and Drainage, **53**, 135–143.

Rodríguez-Díaz J A; Camacho-Poyato E; López-Luque R; Pérez-Urrestarazu L (2008). Benchmarking and multivariate data analysis techniques for improving the efficiency of irrigation districts: an application in Spain. Agricultural Systems, **96**, 250–259.

Roger L L; Dowla F U (1994). Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. Water Resources Research, **30**(2), 457–481.

Ruiz J E; Cordery I; Sharma A (2006). Impact of mid-Pacific Ocean thermocline on the prediction of Australian rainfall. Journal of Hydrology, **317**, 104–122.

Rumelhart D E; Hinton G E; Williams R J (1986). ''Learning'' representations by backpropagation errors. Nature, **323**(9), 533–536.

See L; Abrahart R J (2001). Multi-model data fusion for hydrological forecasting. Computers and Geosciences, **27**, 987–994.

See L; Openshaw S (1999). Applying soft computing approaches to river level forecasting. Hydrological Sciences Journal, **44**(5), 763–778.

See L; Openshaw S (2000). A hybrid multi-model approach to river level forecasting. Hydrological Sciences Journal, **45**(4), 523–536.

Shamseldin A Y; O'Connor K M (1999). A real-time combination method for the outputs of different rainfall–runoff models. Hydrological Sciences Journal, **44**(6), 895–912.

Shepherd A J (1997). Second-order Methods for Neural Networks. Springer, New York.

Shin H S; Salas J D (2000). Regional drought analysis based on neural networks. Journal of Hydrologic Engineering, **5**(2), 145–155.

Sumpsi J M; Garrido A; Blanco M; Varela C; Iglesias E (1998). Economía y Política de Gestión del Agua en la Agricultura. Ministerio de Agricultura, Pesca y Alimentación, Mundi-Prensa, Madrid, Spain.

Tan Y; Van Cauwenberghe A (1999). Neural-network-based d-step-ahead predictors for nonlinear systems with time delay. Engineering Applications of Artificial Intelligence, **12**(1), 21–25.

Thirumalaiah K; Deo M C (1998). River stage forecasting using artificial neural networks. Journal of Hydrologic Engineering, **3**(1), 26–32.

Thirumalaiah K; Deo M C (2000). Hydrological forecasting using neural networks. Journal of Hydrologic Engineering, **5**(2), 180–189.

Tokar A S; Johnson P A (1999). Rainfall–runoff modeling using artificial neural networks. Journal of Hydrologic Engineering, **4**(3), 232–239.

Tokar A S; Markus M (2000). Precipitation–runoff modeling using artificial neural networks and conceptual models. Journal of Hydrologic Engineering, **5**(2), 156–161.

Tsoukalas L H; Uhrig R E (1997). Fuzzy and Neural Approaches in Engineering. Wiley Interscience, New York, USA.

Van Aelst P V; Ragab R A; Feyen J; Raes D (1988). Improving irrigation management by modelling the irrigation schedule. Agricultural Water Management, **13**, 113–125.

Ventura S; Silva M; Pérez-Bendito D; Hervás C (1995). Artificial neural networks for estimation of kinetic analytical parameters. Analytical Chemistry, **67**(9), 1521–1525.

Wedding II D K; Cios K J (1996). Time series forecasting by combining RBF networks, certainty factors, and the Box–Jenkins model. Neurocomputing, **10**, 149–168.

Yang C C; Prasher S O; Lacroix R; Sreekanth S; Patni N K; Masse L (1997). Artificial neural network model for subsurface-drained farmland. Journal of Irrigation and Drainage Engineering, **123**(4), 285–292.

Yapo P O; Gupta H V; Sorooshian S (1996). Automatic calibration of conceptual rainfall–runoff models: sensitivity to calibration data. Journal of Hydrology, **181**, 23–48.

Zadeh L (1965). Fuzzy sets. Information and Control, **8**, 338–353.

Zhang G P (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, **50**, 159–175.

Zhang M; Fulcher J; Scofield R A (1997). Rainfall estimation using artificial neural network group. Neurocomputing, **16**, 97–115.

Zhang G P; Patuwo B E; Hu M Y (2001). A simulation study of artificial neural networks for nonlinear time-series forecasting. Computers and Operations Research, **28**(4), 381–396.

Zhou S L; McMahon T A; Walton A; Lewis J (2002). Forecasting operational demand for an urban water supply zone. Journal of Hydrology, **259**, 189–202.