Prognosis of urban water consumption using hybrid fuzzy algorithms

Gergely Bárdossy, Gábor Halász and János Winter

ABSTRACT

For the optimal operation of waterworks it is necessary to predict the expected water consumption of the following days as accurately as possible. However, there are no conventional methods to predict the water demand. In this paper a prediction model based on hybrid fuzzy algorithms is introduced. The software automatically creates a fuzzy rule system out of a training database using the so-called VISIT (Variable Input Spread Inference Training) algorithm. A fuzzy neural network (FNN) system is created. Rules are trained with back propagation (BP) and least squares estimate (LSE) methods. The parameters of the algorithm are optimized with a simple genetic algorithm. As a result, one gets a rule system that delivers higher accuracy than a common statistically based model. Calculations and results are presented in this paper.

Key words | fuzzy logic, hybrid algorithm, learning, prediction, water consumption

Gergely Bárdossy (corresponding author)
Gábor Halász
Department of Hydrodynamic Systems,
Budapest University of Technology and Economics,
Mûegyetem rkp. 3, Budapest 1111,
Hungary
Tel.: +36 1 463 3097
Fax.: +36 1 463 3091
E-mail: *bardossy@hds.bme.hu*

János Winter
Department of Hydraulic and Water Resources Engineering,
Budapest University of Technology and Economics,
Mûegyetem rkp. 3, Budapest 1111,
Hungary

INTRODUCTION

Forecasting urban water consumption is a highly complex and difficult task. The consumption is influenced by numerous parameters. In addition, a part of these parameters is stochastic (for example, the effect of a burst pipe). However, observations prove that air temperature and rainfall have a significant effect on water consumption. Our aim is to predict the water demand with the help of the known input parameters (daily rainfall data, daily average and maximum temperature). There are multiple ways to construct prediction models. Sastri & Valdes (1989) investigated the effect of rainfall interventions. An empirical model was developed to estimate the transient drops in water consumption. Usually statistical methods are applied for water demand estimation. According to the review of Altunkaynak *et al.* (2005) regression analysis and time series models are the most frequently used statistical techniques in the estimation models. Graeser (1958) showed the relation between the maximum daily water demands in Dallas, Texas and the days with a maximum air temperature over an allotted temperature value. Weeks & McMahon (1973) found that the weekly pan evaporation and the average maximum daily temperature were more significant variables than the rainfall data for a linear regression model. Davis *et al.* (1987) developed the so-called IWR-MAIN Water Use Forecasting system software. The software is highly disaggregated; water requirements for the residental, commercial, industrial and public sectors can be estimated. However, a review of Thompson *et al.* (1993) states that, for many cities, all input data needed for accurate forecasts are simply not available. In recent years various soft computing techniques were applied for water demand estimation instead of statistical methods. Jain *et al.* (2004) introduced an algorithm using artificial neural networks for short-term water demand prognostication. The weekly water demand for Kanpur, India was prognosticated with this method. The algorithm outperformed the regression and time series models. There are also many applications of fuzzy logic in the research area of water resources. Faye *et al.* (2000) used fuzzy methods for the long-term management of water resource systems. Kindler (1992) developed a method based on fuzzy logic for optimal water allocation, while Bárdossy & Disse (1993) applied fuzzy methods to

doi: 10.2166/aqua.2009.092

model infiltration and water movement in unsaturated zones. Altunkaynak *et al.* (2005) developed a Takagi–Sugeno fuzzy method to predict the monthly water consumption of Istanbul, Turkey. As for the input variables of the fuzzy rule system three antecedent monthly water demands were used. The parameters of the model are tuned with the ANFIS (Adaptive-Network-Based Fuzzy Inference System) method.

In this paper we introduce a hybrid fuzzy method to prognosticate the daily water consumption. The algorithm constructs a Takagi–Sugeno type fuzzy rule system out of a training database. The VISIT algorithm is used for this process. The rule system is tuned with the ANFIS method. For the optimization of the algorithm parameters a simple genetic algorithm (GA) is used. The goal of the optimization is to get higher accuracy with a lower number of fuzzy rules.

The prediction model was tested on the data of the DMRV (Duna Menti Regionális Vízmû: in English, Danube Inshore Regional Waterworks) waterworks. Water demand forecasts have been carried out in three different regions in Hungary: Gödöllõ (population: 32,000; area: 63 km^2), Csömör (population: 8,000 area: 23 km^2) and Szada (population: 3,400 area: 17 km^2). The predicted values were compared with forecasted data gained from a statistically based model. In this paper the results for the region of Csömör are presented.

FUZZY METHODS

Introduction to fuzzy logic

An extended introduction to fuzzy logic and fuzzy sets can be found in Tanaka (1991) and Cox (2005). The basics of fuzzy logic were invented by Lotfi Zadeh, who defined fuzzy sets for the first time in the journal *Information and Control* in 1965 (Zadeh 1965). In contrast to binary logic, fuzzy logic is able to handle vagueness mathematically. As conventional ("crisp") sets are separated by strict margins, they cannot characterize the transitional areas. In contrast to this, fuzzy sets are able to describe such transitions. Thus fuzzy set theory can be treated as an extension of the crisp set theory. An expressive example regarding the height of people is shown in Figure 1.

A fuzzy set **A** (e.g. "Small") on the universe **X** (e.g. "Height") is defined by its membership function μ_A:

$$\mu_A : X \rightarrow [0, 1] \tag{1}$$

The membership function is an extension of the characteristic function which describes the crisp sets. The membership function can be any arbitrary real value between 0 and 1. In contrast, the characteristic function can be either 0 or 1. Thus the membership function defines the degree, whereby the given element belongs to the fuzzy set. One can execute the same operations with fuzzy sets as with crisp sets, like union, intersection, complement, etc. A detailed description of these operations can be found in Tanaka (1991).

Construction of fuzzy rules

With the use of fuzzy sets one is able to construct so-called fuzzy rules. Fuzzy rules establish the connection between the input and the output parameters. The rules can be given in "IF–THEN" format:

$$\text{IF } x \in A \quad \text{and} \quad y \in B \quad \text{THEN} \quad z \in C$$

with **A**, **B** and **C** fuzzy sets. In the context above, x and y are the so-called premise variables and z is the consequence

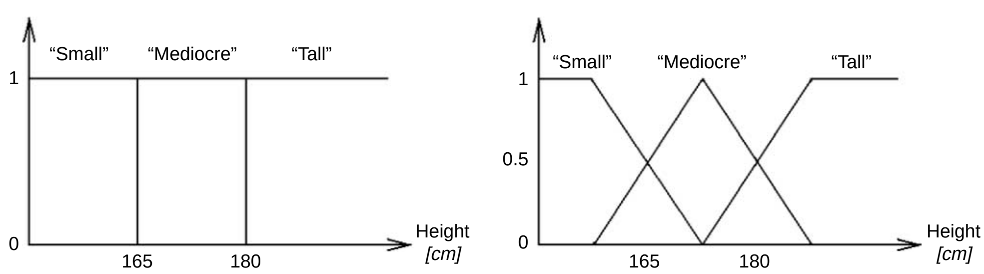


Figure 1 | Classification of a person's height with crisp sets and fuzzy sets.

variable. One has to calculate the membership values for x and y. Out of these values we can determine the adaptability of the given rule. In this paper we use the minimum operator, thus the min value of the two membership values gives the degree of adaptability of the rule. The degree of adaptability is also called the *firing strength*. A fuzzy rule system consists of several fuzzy rules. The resultant output is given by some combination of the output fuzzy sets of each rule. As for generating fuzzy rules, one can use direct or indirect methods. In this paper we use the Takagi & Sugeno direct reasoning method.

THE VISIT ALGORITHM

Our goal is to construct a fuzzy rule system that is able to estimate the water consumption with high accuracy. Therefore one should analyse temperature, rainfall and water consumption data in the preceding period. Out of these data one can create fuzzy rules and build up a whole rule system. However, this procedure is rather time-consuming. Furthermore, there might exist more accurate rule systems than the one obtained. After a while the rule system loses its actuality and should be updated. For instance, water demand has a seasonal trend. In addition, waterworks possess numerous supply zones. In order to get high accuracy, one has to create a rule system for each supply zone. It is obvious that creating suitable fuzzy rules needs an enormous effort. Therefore our aim is to create an algorithm for automated fuzzy rule generation. The algorithm should create the rules out of a training database containing preceding data. This is the so-called learning process. In our research we use the VISIT algorithm of Xiaoguang & Lilly (2004). The VISIT algorithm is in fact an improvement of the MLFE (Modified Learning From Example) fuzzy learning method. It applies asymmetric Gaussian membership functions:

$$\mu_A(x) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma L}\right)^2}, \text{if } x < m \\ e^{-\frac{1}{2}\left(\frac{x-m}{\sigma R}\right)^2}, \text{if } x > m \end{cases}$$

where $\mu_A(x)$ denotes the membership function of the fuzzy set $\mathbf{A}$, m – the so-called "center" – is the location of the peak of the Gaussian membership function $\mu_A(x)$, σ_L and σ_R are the left and right spreads (standard deviations) of the fuzzy set. The VISIT algorithm can be described with the following main steps, after Xiaoguang & Lilly (2004):

1. Read the first data row from the training database: $x_1^{k=1}, x_2^{k=1}, \cdots, x_n^{k=1}, y^{k=1}$ where $x_i^{k=1}, i = 1, \cdots, n$ are the input variables (daily temperature and rainfall) and $y^{k=1}$ is the output variable (water consumption) of the first data row. The index $k = 1, \ldots, m$ denotes the number of the data rows of the training database. Now create the membership functions for the $x_i^{k=1}$ input variables. Center: $m_i^{j=1} = x_i^{k=1}$, spreads: $\sigma_{Ri}^{j=1} = \sigma_{Li}^{j=1} = \sigma_0$. The index $j = 1, \ldots, R$ denotes the rule number of the fuzzy rule system. Here σ_0 initial spread is a given value. As for the output: $b^{j=1} = y^{k=1}$.
2. Define a fuzzy rule using the previously created fuzzy sets. For *Rule 1* one gets

$$\text{IF } x_1 \in A_1^{j=1} \text{ AND } x_2 \in A_2^{j=1} \text{ AND } \cdots \text{ AND } x_n \in A_n^{j=1} \text{ THEN } y = b^{j=1}$$

3. If there is no additional data in the training database then quit; else read the next data row. Apply the rule system on the input parameters of these data. Compare the output of the rule system with the real consumption. If the estimation error is below the error limit ϵ then jump to the next data row. Else continue with the fourth step.
4. Investigate each x_i^k input parameter, what degree they belong to each A_i^j fuzzy set – namely, calculate the membership values for x_i^k in A_i^j. If the membership values are below a prescribed a_i value (α-cut), then a new fuzzy set has to be created. The center of this fuzzy set is $m_i^{j+1} = x_i^k$. As for the output set $b^{j+1} = y^k$.
5. Now the left- and right-hand side spreads have to be calculated. Furthermore the spreads of the adjacent fuzzy sets have to be updated. The right-hand side spread of the new fuzzy set and the left-hand side spread of the adjacent fuzzy set can be calculated as follows:

$$\sigma_{Ri}^{j+1} = \frac{1}{\omega_i}\left|m_i^{j+1} - m_i^{nR}\right| \tag{2}$$

with m_i^{j+1} the center of the new fuzzy set, m_i^{nR} the center of the right-hand side adjacent fuzzy set and ω_i the weight factor. Similarly one can obtain the left-hand side spread of the new fuzzy set and the right-hand side spread of the adjacent fuzzy set:

$$\sigma_{Li}^{j+1} = \frac{1}{\omega_i}\left|m_i^{j+1} - m_i^{nL}\right| \tag{3}$$

6. With the created fuzzy sets and the new output, a new fuzzy rule is obtained. Now it has to be investigated if this rule conflicts with the preceding rules. If one of the preceding rules has nearly the same premises but a differing consequent, then the new rule is inconsistent and thus it has to be neglected. If there are no conflicts, the rule can be added to the existing rule system. Afterwards jump to the third step.

To sum up, the algorithm creates the required fuzzy sets and creates a complete fuzzy rule system. This rule system is capable of predicting the water consumption.

ADAPTIVE-NETWORK-BASED FUZZY INFERENCE SYSTEM (ANFIS)

The goodness of a fuzzy rule system is highly influenced by the shape and position of each fuzzy set. Therefore, by changing these parameters in each fuzzy set of the rule system gained by the VISIT algorithm, the accuracy of the model can be improved. In this paper the ANFIS method is applied for this attempt. This method is precisely described in the work of Jang (1993). Here the parameters of the fuzzy system, such as the input fuzzy sets, firing strengths, consequent parameters and the overall output, are treated as nodes of a neural network. This analogy is shown in Figure 2. The output of the fuzzy rule system is calculated as follows:

$$f = \frac{w_1 f_1 + w_2 f_3}{w_1 + w_2} = \bar{w}_1 f_1 + \bar{w}_2 f_2 \tag{4}$$

Just like the VISIT algorithm, ANFIS needs the training database to optimize the parameters in each network node. The most widely used method is the Back Propagation (BP) algorithm. The basic idea of this method is to distribute the prognostication error among the nodes. The fuzzy rules with high firing strengths w^j get the greater amount of this error. Proportional to this error the parameters of the fuzzy sets constituing this rule are modified. In the current work the algorithm of Kuo *et al.* (2004) is applied. The degree of the modification rate can be specified with a learning factor η.

1. Correction of the output:

$$b^j(n+1) = b^j(n) + \eta(y_k - o_k)w^j \tag{5}$$

2. Correction of the membership function centers:

$$m_i^j(n+1) = m_i^j(n) + \eta(y_k - o_k)b^j(n) \, w^j \frac{1}{\sigma_i^2}\left(x_i - m_i^j(n)\right) \tag{6}$$

3. Correction of the left- and right-hand side spreads:

$$\sigma_i^j(n+1) = \sigma_i^j(n) + \eta(y_k - o_k)b^j(n) \, w^j \frac{1}{\sigma_i^3}\left(x_i - m_i^j(n)\right)^2 \tag{7}$$

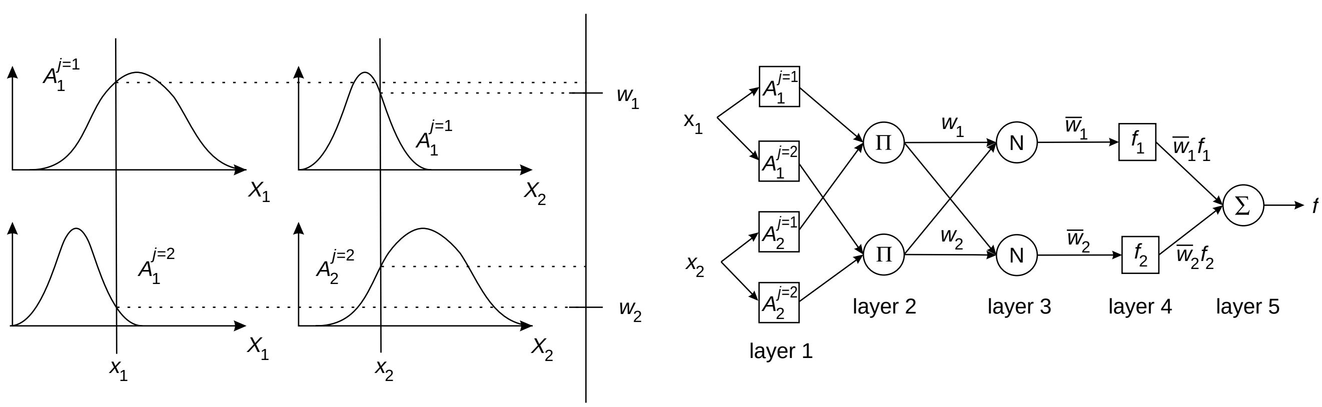


Figure 2 | Fuzzy system mapped on an adaptive neural network.

with $j = 1, \ldots, R$ number of rules, $n = 1, \ldots, N$ iteration counter, $k = 1, \ldots, m$ data rows of the training database, o_k output of the rule system (prognosticated value), y_k real output for the kth data row and w^j firing strength. The learning factor can be changed during the iteration process regarding the variation of the errors. The number of the iterations can be specified arbitrarily. Here the size of the training database does not delimit the calculation: the same data rows can be used many times. The other method in ANFIS is the Least Squares Estimate (LSE). This algorithm modifies the consequence parameters of the fuzzy rule system. As for the prognostication algorithm we use the sequential method proposed in Jang (1993). During the calculations we apply the LSE algorithm only once and the BP method is iterated several times.

OPTIMIZATION OF THE ALGORITHM PARAMETERS

Necessity of optimization

Now our fuzzy algorithm is capable of setting up a fuzzy rule system out of a training database. The parameters of the rule system are fine-tuned using neural networks. However, to start the calculation one needs to determine the initial values for the algorithm. The main initial values are: σ_0 initial spread, a_i α-cuts for each fuzzy set, σ_i weight factors, ε error limit. These initial values have a significant effect on the whole fuzzy rule system and therefore on the accuracy of the prognostication. Incorrect initial values can result in a bad prognostication. To determine the optimal initial parameters an optimization process has to be carried out.

Optimization with genetic algorithms (GA)

To optimize the initial parameters, a genetic algorithm proposed in Xiaoguang & Lilly (2004) is applied. The application of genetic algorithms is very useful if the space, where we search for the optimal solution, is poorly understood and irregular. A good description of genetic algorithms can be found in the book of Cox (2005). The mechanism of genetic algorithms is based on Darwin's evolutionary theory. In the evolution of the species, there are always stronger and weaker individuals. Stronger individuals have a larger chance to survive than the weaker ones. If stronger ones are mated, the produced offspring are also likely to be strong-performing. In contrast, weak individuals will produce weakly-performing offspring. These will be separated from the population with time. To perform an optimization with genetic algorithms one has to define the individuals. These individuals have to contain all the parameters that are to be optimized. In this paper the ith individual is defined as follows:

$$P_t(i) = \left[\omega_1^t(i), \cdots, \omega_n^t(i), a_1^t(i), \cdots, a_n^t(i), \sigma_1^t(i), \cdots, \sigma_n^t(i), \varepsilon\right] \quad (8)$$

with t generation number (tth generation) and i the number of individuals within a generation. Thus the individual $P_t(i)$ contains all the parameters that have to be optimized. In the first step, we randomly generate N individuals. Using the crossover operator and the mutation operator described in Xiaoguang & Lilly (2004), new individuals can be created. Overall we get $2N$ individuals. With each individual a fuzzy rule system is created. Regarding the accuracy and the number of rules in the rule system, one can classify the individuals. Note that the “goodness” (fitness) of a fuzzy rule system is not only influenced by the accuracy, but also by the number of fuzzy rules. Now the N best individuals are selected. These individuals form the new generation. Again, crossover and mutation can be performed. From generation to generation, the fitting of the individuals will increase. After several generations, the optimal individual will be determined.

RESULTS

Input parameters for the algorithm

In order to test the fuzzy learning algorithm, two databases are required. On the one hand, the training database has to be created out of preceding data. The training database is used for setting up and optimizing the fuzzy rule system. On the other hand, a database with data of the period to be prognosticated is necessary. The algorithm will gain the input parameters out of this database. In this case we assume the consumptions to be unknown. The algorithm prognosticates the consumption using the input parameters. The forecasted water consumption is compared with the

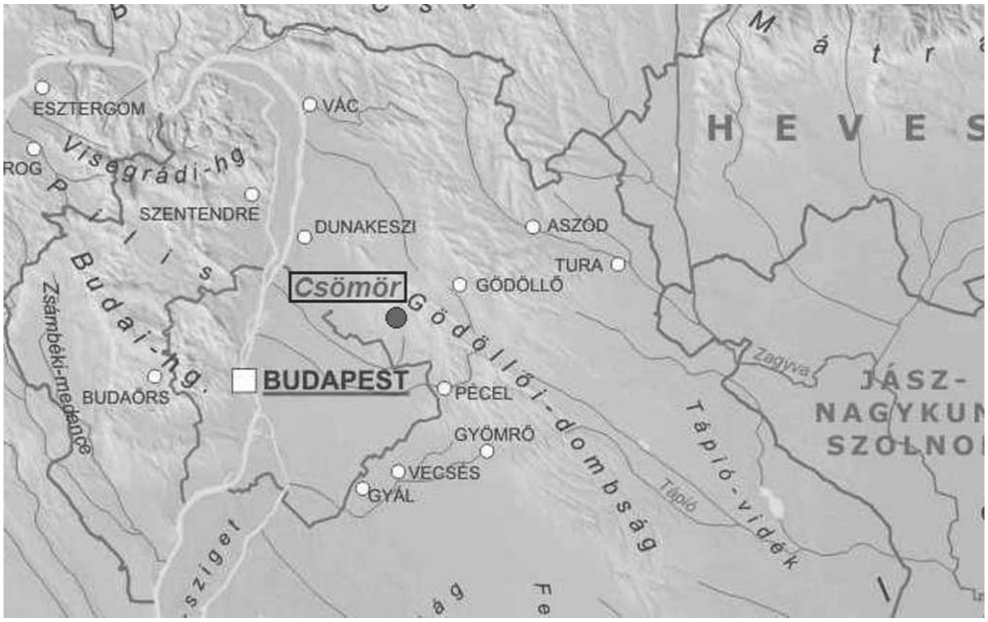


Figure 3 | Location map of Csömör, Hungary.

real consumption, thus the error can be estimated. In the prognostication model two input parameters were defined: exponentially smoothed daily rainfall rate and exponentially smoothed daily maximum temperatures. Both variables show a significant correlation with the actual water consumption. The exponentially smoothed daily rainfall rate is calculated as follows:

$$R_S^{n+1} = \alpha R_S^n + (1-\alpha)R^{n+1} \quad (9)$$

with α smoothing coefficient of rainfall data, R_S^{n+1} smoothed rainfall value for day $n+1$ and R^{n+1} rainfall value on day $n+1$. Similarly, the exponentially smoothed daily maximum temperature:

$$T_S^{n+1} = \beta T_S^n = (1-\beta)T^{n+1} \quad (10)$$

with β smoothing coefficient of temperature data, T_S^{n+1} smoothed temperature value for day $n+1$ and T^{n+1} temperature value on day $n+1$. The testing simulations have shown that the optimal value of the smoothing coefficient of rainfall data is approximately 0.8, while the smoothing coefficient of temperature data is 0.5. The process of water consumption is the following. The algorithm creates a fuzzy rule system out of the training database. With the input parameters of the next day, the water consumption will be estimated. As the real water consumption is known, the training database will be updated with the actual data. At the same time, the oldest data row of the database is erased. Thus the training database is kept up to date. This also means that there is no need for corrections to treat trends in the water consumption.

Prognosis for the region Csömör

During the testing period water prognostication runs were carried out, among others for the region of Csömör, Hungary. Csömör has a population of about 8,000 peole and is situated north-west of the capital Budapest. A location map is given in Figure 3. The water supply of Csömör is ensured by the waterworks DMRV.

The task was to prognosticate the daily water demand during the period from 1 May 2006 to 31 August 2006. The training database contained data for the period from 1 May 2005 to 31 August 2005. These data were automatically updated during the prognostication process. The data consisted of daily rainfall and daily maximum temperature values. The exponentially smoothed input variables of the fuzzy sets were calculated from these data. The fuzzy set automatically generated by the estimation software used

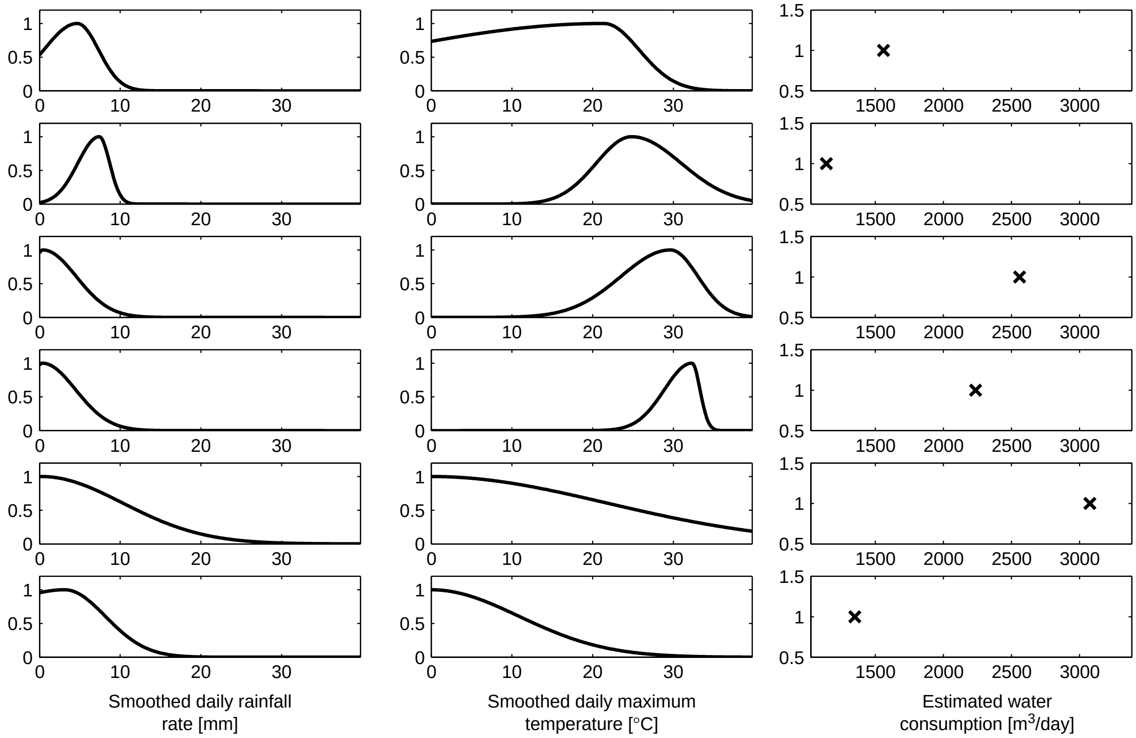
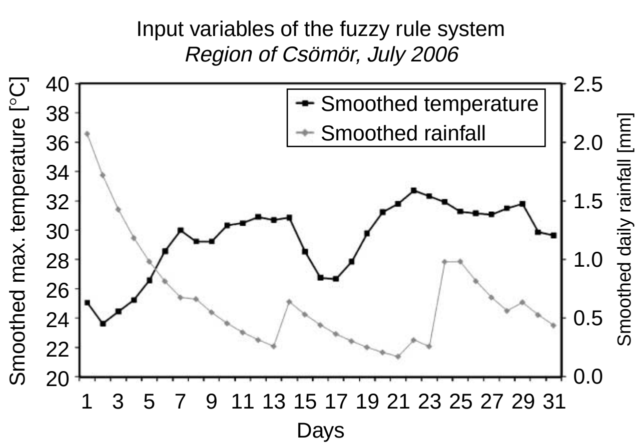


Figure 4 | Sample fuzzy rule set and membership functions.

these input parameters to prognosticate the water consumption. In Figures 4 and 5 the input and output variables of the fuzzy systems are shown as a function of time. A sample fuzzy rule system generated by the estimation software is shown in Figure 4.

The water demand prognostication was also carried out with statistical methods. Here, linear and nonlinear regression was applied. The independent variables were the average daily maximum temperature and the average daily rainfall rate of the actual and the preceding day. The evaluated daily water consumptions were compared with the results of the hybrid fuzzy system. Figure 6 shows the real and the estimated water consumptions for the region Csömör in July 2006. In Figure 7 the deviation of the estimates from the real water consumption is shown.



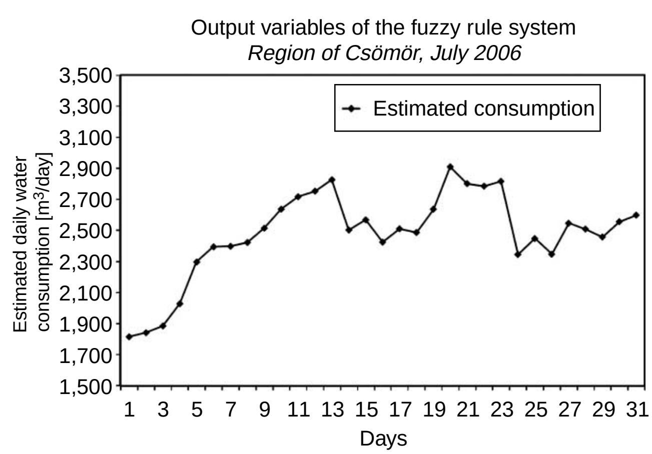


Figure 5 | Input and output variables of the fuzzy rule system.

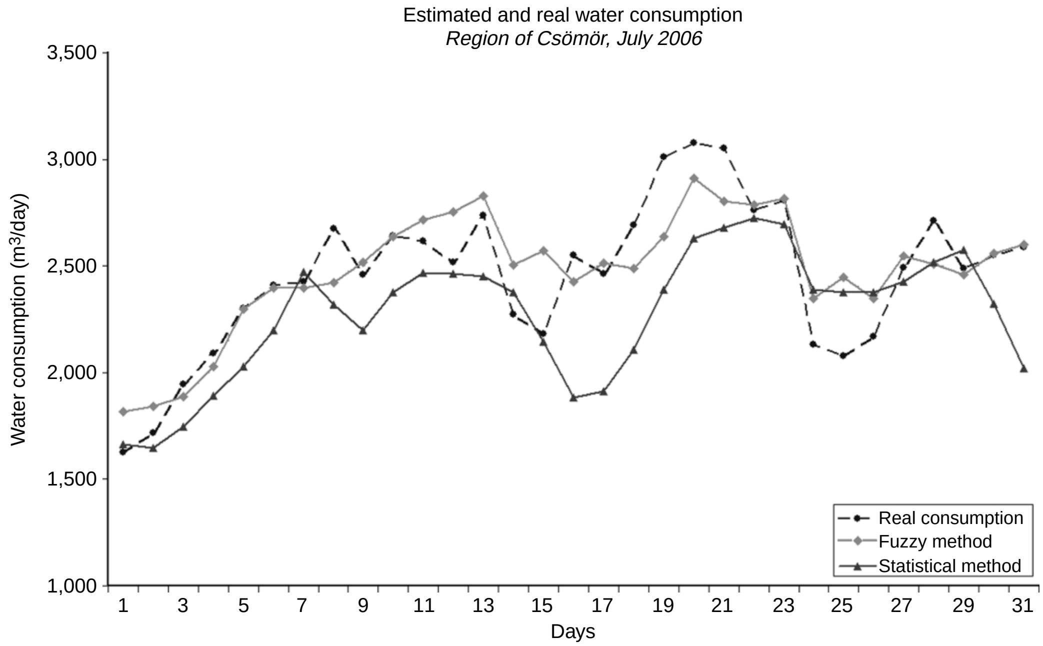


Figure 6 | Comparison of the fuzzy method and the statistical methods.

We can denote that the deviations of the fuzzy prognosis are usually smaller than the deviations of the estimates determined with statistical methods. In Table 1 we compared the monthly average deviation, the correlation coefficient and the Nash–Sutcliffe sufficiency score (NSSS) for the two estimation methods. The hybrid fuzzy method produces a stable deviation between 5–6%. In contrast to this, the results of the statistical method vary between 5.5–10%. One can note that, for July 2006, the statistical prognostication delivers a deviation of 10%. In the next month, August 2006, the average deviation decreases to 5.5%. Thus the regression models deliver a less stable

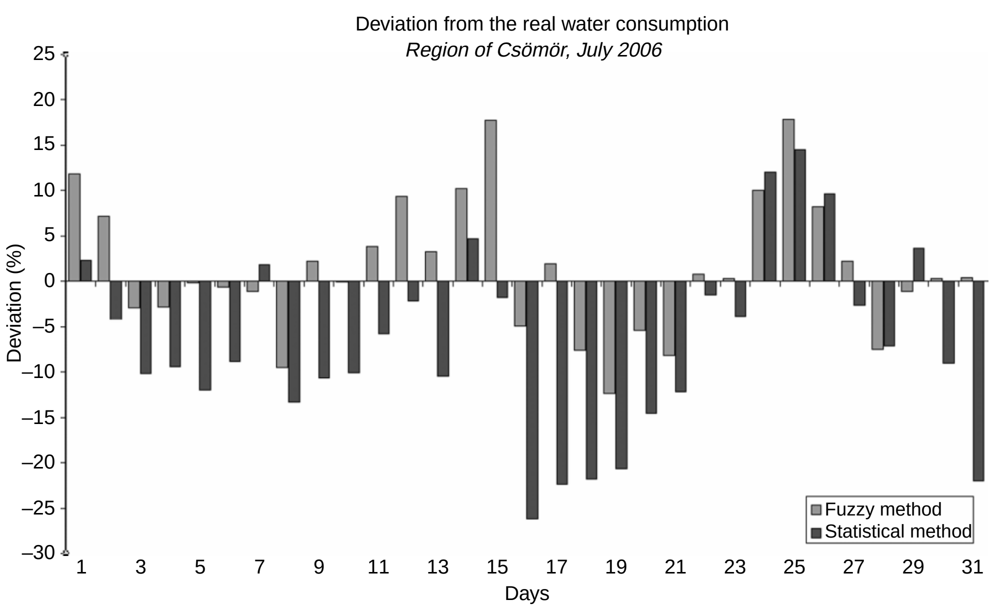


Figure 7 | Comparison of the deviations.

Table 1 | Comparison of the fuzzy and statistical estimation methods

	Average deviation		Correlation coefficient		NSSS	
	Fuzzy	Statistical	Fuzzy	Statistical	Fuzzy	Statistical
May	5.86	7.41	0.71	0.60	0.50	0.15
June	5.74	9.27	0.93	0.87	0.83	0.70
July	5.55	10.05	0.87	0.70	0.74	0.18
August	5.01	5.50	0.83	0.87	0.62	0.65

estimation than the hybrid fuzzy method. The correlation coefficient and the NSSS are also higher for the fuzzy method in the period May–July.

CONCLUSIONS

In this paper we developed a hybrid fuzzy algorithm to prognosticate urban water consumption. The algorithm is able to learn from a training dataset containing preceding rainfall and temperature data. The fuzzy rule system is generated using the VISIT algorithm, after Xiaoguang & Lilly (2004). The automatically generated rule system is fine-tuned with a fuzzy neural network (FNN). Here the ANFIS method of Jang (1993) is used. The initial values of the hybrid fuzzy algorithm are optimized with genetic algorithms.

As for testing the hybrid fuzzy method we performed a water consumption estimation for the region of Csömör. The consumption was also prognosticated using linear and nonlinear regression models. The results of the hybrid fuzzy method and the statistical method were compared with the real consumptions. The fuzzy method proved to be more accurate. The variation of the estimation deviations for the fuzzy model is lower than for the statistical methods.

The big advantage of the hybrid fuzzy model is the easy adaptability to each supply zone of the waterworks. Due to the automatic rule generation and parameter optimization there is no need to manually tune the parameters. Thus the fuzzy estimation model can be used easily in practice.

The hybrid fuzzy algorithm is highly sensitive to the accuracy of the training dataset. Some faulty data in the training dataset can spoil the accuracy of the water consumption estimation.

In order to get more information on the hybrid fuzzy algorithm, more tests should be carried out. In the future, the running time of the algorithm should be optimized. On the other hand a method should be developed to estimate the water consumption distribution during the day.

REFERENCES

Altunkaynak, A., Özger, M. & Çakmakci, M. 2005 Water consumption prediction of Istanbul City by using fuzzy logic approach. *Wat. Res. Mngmnt.* **19**(5), 641–654.

Bárdossy, A. & Dise, M. 1993 Fuzzy rule based models for infiltration. *Wat. Res. Res.* **29**, 373–382.

Cox, E. 2005 *Fuzzy Modeling and Genetic Algorithms for Data Mining and Exploration*. Elsevier, Amsterdam.

Davis, W. Y., Rodrigo, D. M., Opitz, E. M., Dziegielewski, B. & Baumann, D. D. 1987 *IWR-MAIN Water Use Forecasting System. Version 5.1. User's Manual and System Description*. IWR Report 8-R6, US Army Corps of Engineers Institute for Water Resources, Fort Belvoir.

Faye, R. M., Sawadogo, S., Mora-Camino, F. & Achaibou, A. K. 2000 Fuzzy modeling approach for the long term management of water resource systems. In: *IEEE International Conference on Fuzzy Systems, San Antonio, TX, USA*. pp 499–504.

Graeser, H. J. Jr. 1958 Meter records in systems planning. *J. AWWA* **50**, 1395–1402.

Jain, A., Varshney, A. K. & Joshi, U. C. 2004 Short-term water demand forecast modeling at IIT Kanpur using artificial neural networks. *Wat. Res. Mngmnt.* **15**(5), 299–321.

Jang, J. S. R. 1993 Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst.* **23**(3), 665–685.

Kindler, J. 1992 Rationalizing water requirements with aid of fuzzy allocation model. *Wat. Res. Plann. Mngmnt.* **118**, 308–323.

Kuo, H. C., *et al.* 2004 Symbiotic evolution-based design of fuzzy-neural transformer diagnostic system. *Electric Power Syst. Res.* **72**, 235–244.

Sastri, T. & Valdes, J. B. 1989 Rainfall intervention analysis for on-line applications. *J. Wat. Res. Plann. Mngmnt.* **115**, 397–415.

Tanaka, K. 1991 *An Introduction to Fuzzy Logic for Practical Applications*. Springer-Verlag, Berlin.

Thompson, J. G., Parker, M., Templin, W. & Reynolds, R. R. Jr. 1993 A review of application issues of the Metropolitan Water District-Main Water Forecasting System. *J. AWRA* **29**(3), 425–433.

Weeks, C. R. & McMahon, T. A. 1973 A comparison of water utilities in Australia and the U.S. *J. AWWA* **65**, 232–241.

Xiaoguang, Ch. & Lilly, J. 2004 Evolutionary design of a fuzzy classifier from data. *IEEE Cybern.* **34**(4), 1894–1906.

Zadeh, L. A. 1965 Fuzzy sets. *Inf. Control* **8**, 338–353.

First received 28 August 2007; accepted in revised form 7 November 2008. Available online February 2009