# Reservoir Operation by Fuzzy Reasoning


A. Jolma[a], E. Turunen[b], M. Kummu[a] and T. Dubrovin[c]

[a] *Helsinki University of Technology, P.O.Box 5200, 02015 HUT, Finland (Ari.Jolma@hut.fi)*

[b] *Tampere University of Technology, Tampere, Finland*

[c] *Finnish Environment Institute, Helsinki, Finland*



**Abstract:** The operation of lakes and reservoirs is currently under scrutiny in Finland due to the aging of the old operation permits, climate change, and changing environmental values. Simulation is a viable tool for studying reservoir operation but difficult to use when the operation does not have clear explicit rules. Fuzzy logic has been shown to be useful in modeling simple decision making which requires intelligence. Two fuzzy logic based methods, namely case-based and rule-based reasoning were examined and applied for modeling the operation of five lakes of a river basin. Different mathematical structures and interpretations were tested including the use of fuzzy similarity based approaches. The best results were obtained using a rule based approach with a simple mathematical structure and interpretation of the fuzzy *and* as a product. The rule base in the model was allowed to be inconsistent and each rule had an associated weight or strength. The case-based approach performed rather well in comparison with the rule-based approach and in some cases better in some respect. Both models can mimic the human operator reasonably well in easy circumstances but have problems especially in the downstream lakes with large discharge to storage ratios. The fuzzy logic based models, case-based reasoning or rule-based reasoning require further work to be applicable in studying practical problems like the adjusting the operation to changes in environmental values and climate.



*Keywords:* Reservoir Operation; Fuzzy Logic; Total Fuzzy Similarity; Simulation


## 1. INTRODUCTION

The actual operation of many Finnish lakes is often a compromise between the goals of the operator and the rules set out in an operation permit. Often neither the operator's own rules nor the operation permit state explicitly what the release should be in a given situation. The operation is thus very much based on the experience and skill of the managers. This forms an obstacle for modeling and simulation of the operation of the lake.

Currently many operation permits are being revised. A strong motivation for the revision are the changing environmental values and changing uses of water resources. A thorough analysis of the operation and its impact requires simulation under existing and predicted conditions.

This paper describes an attempt to model the operation of a system of five lakes using a fuzzy logic based approach. The attempt is based on general knowledge and on water level and release data. The goal of the modeling at this stage is to mimic the human operator. We have tested several techniques and mathematical formulations including the total fuzzy similarity of Turunen [1999].

Several approaches to apply fuzzy set theory to reservoir operation have been described in the literature. These include fuzzy optimization techniques, fuzzy rule base systems, and combinations of fuzzy approach with other techniques. Applications can be found in the work of Fontane et al. [1997], Huang [1996], and Saad et al. [1996]. Fuzzy rule base control systems for reservoir operation are presented by Russell and Campbell [1996] and Shrestra et al. [1996]. The fuzzy rule base can be constructed on the basis of expert knowledge or observed data. Methods for deriving a rule base from observations have been presented by Bárdossy and Duckstein [1995] and Kosko [1992].

## 2. FUZZY REASONING OF RESERVOIR RELEASE

### 2.1 Introduction

In reservoir operation the operator determines a suitable release based on observation data and other conditions. To describe the values of various variables a reservoir operator typically employs linguistic values. The basic assumption in fuzzy modeling is that all linguistic values can be modeled as fuzzy sets. Fuzzy sets are defined by membership functions, which map crisp values to degree of membership (DOM) (to the fuzzy set) values. The linguistic values should cover the whole range of possible values a variable may have. For example the variable "water level" may have three linguistic values: "low", "normal", and "high". Each crisp water level value can then be expressed as a vector of three DOM values in this case.

There is not one universal interpretation for logical operators (*and*, *or*, *not*, *implication*, and *equivalence*) as mathematical operations between DOM values in fuzzy logic. Each interpretation defines a different formula for the t-norm ($\otimes$), which is an interpretation of the logical connective *and* and from which interpretations for other operators can be derived. Several interpretations are given in Bárdossy and Duckstein [1995] and in Turunen [1999].

### 2.2 Case-based Reasoning of Release

A direct method for making a decision is to look at the historical data for similar cases and make a similar decision that was made in those cases. Fuzzy similarity is defined by the equivalence relation between two DOM values. To derive the total fuzzy similarity between two observation vectors having one or more variables one must combine component similarities. The combination has to be done in two steps, first by each linguistic value and then by variables. Interestingly, the combination of similarity is mathematically defined only for Lucasiewicz (bounded difference-sum) valued similarity. The total fuzzy similarity is an average of the component similarity measures [Turunen, 1999].

The similarity between the observation vectors is then assumed to equal the similarity between the decisions. Thus the decision derived by case-based reasoning is a weighted average of releases in observation vectors in the historical data similar to the observation vector of the time of the decisionmaking. The weights are the similarities between the observation vectors.

### 2.3 Learning Fuzzy Rules from the Data

A fuzzy rule is an implication $a \rightarrow b$, "if $a$ then $b$" where $a$ and $b$ are sets of fuzzy sets combined using logical operators. Each historical case is a piece of evidence in favor of one or more rules, i.e., it strengthens the weights of those rules. Before any learning has occured every rule has zero weight. By using fuzzy rules we are separating the reasoning into two phases: learning and application of the rule set. The information, or knowledge, available in the historical data is stored into a more condensed format into the rule set.

Using the generalized modus ponens we can calculate the truth value for any rule given a historical case. The generalized modus ponens is a rule for fuzzy reasoning and it states that, if we know the truth values of $a$ and $a \rightarrow b$, then we can calculate the truth value of $b$ as $v(a) \otimes v(a \rightarrow b)$, $v(a)$ is the truth value of the proposition $a$. It is easy to solve from this $v(a \rightarrow b)$ once we have selected the mathematical interpretation for the logical operators.

It is also possible to use the similarity approach in the determination of the support a case gives to a rule. The similarity is then calculated between the observation matrix of DOM values and a (0,1)-valued DOM matrix which defines each if-part of a rule. For example the if-part "water level is low" is written as a DOM valued vector (1,0,0).

It should be noted that each case in the historical data may support several rules. Also there will be rules, which are inconsistent with each other in the sense that one if-part may imply two or more different decisions. This is a consequence of having a strength associated with each rule.

There will inevitably be a situation where one case gives a strength $m_1$ to a rule and another gives a strength $m_2$ to the same rule. Interestingly, this problem is tackled in the Dempster-Shafer theory of evidence as a reduction of two parallel rules using the formula $1 - (1 - m_1)(1 - m_2)$, which happens to be the Gaines-valued (algebraic product-sum) interpretation of fuzzy *or*. *Or*'ing subsequent strengths has the, unwanted in this case, property of locking into value one if at least one strength is one. Other possibilities for combining the evidences is to use the average or logical *and*.

### 2.4 Reasoning a Release Decision Based on a Fuzzy Rule Set

The truth value of each if-part in the rule set can be calculated from the observation vector. Alternatively, the similarity of the observation vector with each if-part can be calculated as described above. This together with the strengths of the rules can be directly fed into the generalized modus ponens to derive DOM values for the release decision. Multiple values associated with the same linguistic value are combined using the logical *or*.

In the last step the thus obtained DOM vector has to be mapped to a crisp decision, i.e., de-fuzzified. We have used here two de-fuzzification methods. In the first the largest element in the vector is selected and the membership function of the respective linguistic value is used. If this membership function is triangular, i.e, there are two crisp values for each DOM value, the values of the elements in the vector to the right and to the left of the largest element are summed, and depending on those sums either the smaller or the larger crisp value is selected. The other technique is to maximize the total fuzzy similarity between a crisp decision and the DOM value vector. The maximization requires knowledge about the form of the membership functions.

## 3. MODELED LAKE SYSTEM

The Kokemäenjoki river basin (figure 1.) is situated in South-Western Finland. There are thousands of lakes in the basin but most of them small. The lakes are connected with rivers and rapids, many of the larger ones have dams or hydro power plants and the watercourse is thus mostly regulated. The basin is the fourth largest in Finland having an area of 27 046 km2, out of which 11 % is lakes. There are two large cities, Tampere (195 000 inh.) and Hämeenlinna (46 000 inh.) and several smaller communities on the basin. The total population on the basin is upwards of 1/2 million.

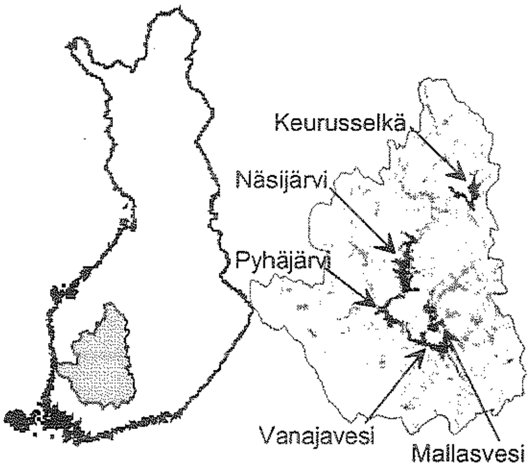


**Figure 1.** Location of the river basin in Finland and the lakes in the basin.

For this study a system of five large and operated lakes, with the exception of Mallasvesi, which is currently not operated, were selected. The lakes represent most of the regulated part of the basin and are Keurusselkä (118 km$^2$), Näsijärvi (262 km$^2$) Mallasvesi (52 km$^2$), Vanajavesi (151 km$^2$), and Pyhäjärvi (124 km$^2$). A schematic of the system is shown in figure 2. The sizes of the triangles in the figure are proportional to the areas of the lakes. The numbers on the flow arrows are the average discharges from the lakes (data is from years 1960–1990, unit is m$^3$/s).

## 4. FUZZY OPERATION MODEL FOR THE LAKES

The development and testing of the models was very much an interactive and a trial and error exercise. Data from years 1970–1989 was used as historical cases and for training and data from years 1990–1999 was used for validation.

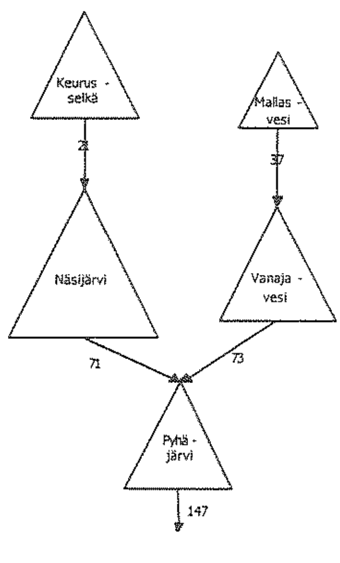


**Figure 2.** A schematic diagram of the five lakes in the modeled system.

The time step of the simulation was set to two weeks. This is long enough for the effect of the travel time of water to be negligible and short enough to take into account the dynamics of the operation. The fuzzy reasoning was used to generate a release proposal, which was then checked by the simulator and changed if minimum or maximum water levels would have been violated. The release was not allowed to be negative.

Two statistics and visual inspection of the operation of the model were employed in the modeling. The two statistics were the average squared difference between the observed and the model output and the number of times the simulator had to change the release proposed by the fuzzy model.

The season, observed storage, observed inflow during the previous time step, inflow forecast, and observed release during the previous time step and some derivative variables, e.g., change in inflow, were tested as input variables. It was deemed valid to use the inflow forecast because in reality one is available (although not a perfect one) and release decisions can be made more often than every two weeks. The effect of the number of linguistic values was tested. The parameters of the membership functions were calculated from the training data.

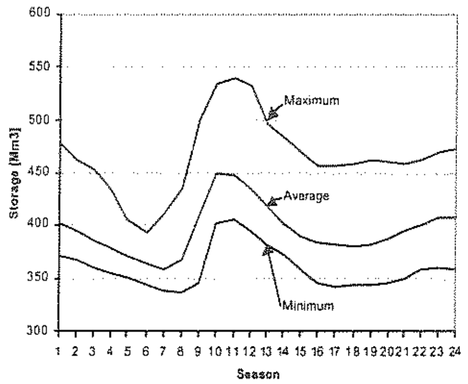


**Figure 3.** The parameters of the membership functions for storage of Keurusselkä.

The whole structure of the model was subject to change and the theoretical considerations discussed previously were used only as suggestions. Different t-norms were tested.

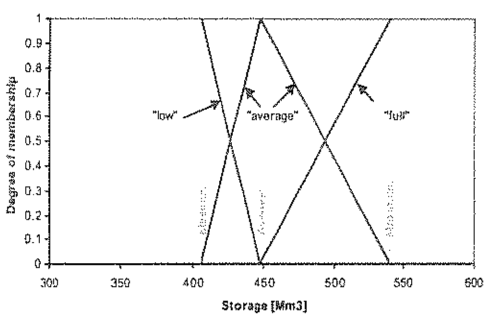


**Figure 4.** An example of fuzzy set membership functions for three linguistic values for storage of Keurusselkä (season is 11, compare to fig. 3).

Figures 3 and 4 depict an example of seasonal variation of the parameters of the membership functions and the triangular shape of the functions. The parameters were obtained from the training data.

## 5. RESULTS

### 5.1 General

Figures 5, 6 and 7 show visually the quality of the operation of the models. All figures visualize the independent operation of the models in the validation period, i.e., the actual data from the validation period, although shown in the figures, was not used in any way by the models. All figures also show season-dependent minimum and maximum storage values as calculated from the data from the training period.

Generally the results were satisfactory for this phase of the modeling. The rule-based approach worked slightly better except for the case of the two upstream lakes, where the case-based approach worked better if judged by the squared error statistic alone. The case-based approach suffered from jerkiness of the operation, which was propagated downstream and amplified. The release suggestion of the best rule-based model had to be disapproved by the simulator 41 times out of 100 in the case of the most downstream lake Pyhäjärvi, and 4–9 times out of 100 in the case of the other lakes in general in the validation period.

### 5.2 Membership Functions

Three linguistic values to represent the whole range of possible values of all variables was found to be enough. Manual smoothing of the season-dependent parameters of the membership functions did not seem to have much effect in the operation of the models and the values obtained directly from the data were used.

### 5.2 Input Variables

The best combination of input variables was found to be storage and inflow forecast. Combinations with more input values produced better results in the training period but failed to reproduce the good performance in the validation period, a sign of over-learning. Also the combination of storage with previous inflow peformed better in the training period but not as well in the validation period.

### 5.4 Effect of the Structure of the Model

The selection of the t-norm had almost as much importance as the selection of the input variables. In general the often used Gaines interpretation (algebraic product-sum) seemed to work the best. The theoretically correct structure did not in general work very well, most notably it was found out that the interpretation of the rule as an implication and subsequent evaluation of its truth value using modus ponens was not useful. Simply stating that $v(a \rightarrow b) = v(a) \otimes v(b)$ and using the product interpretation for $\otimes$ worked the best. The Lukasiewicz interpretation did not produce good results but a more theoretically correct approach produced better results when it was used as opposed to other interpretations. The Lukasiewicz interpretation caused the model to fail to produce a release proposal in many occasions, this is due to the fact that the sum of two truth values has to be more than one in order for the Lukasiwicz t-norm to be greater than zero.

### 5.5 Similarity Approach

The similarity approach did not improve the results in the de-fuzzification but produced almost as good results as the method based on the largest DOM value. The case was similar when the similarity of the input and the if-parts of the rules was used instead of truth values. In the case-based reasoning it was found out that better results were obtained if only cases from the same season were used. The Gödel interpretation (maximum) produced the best results.

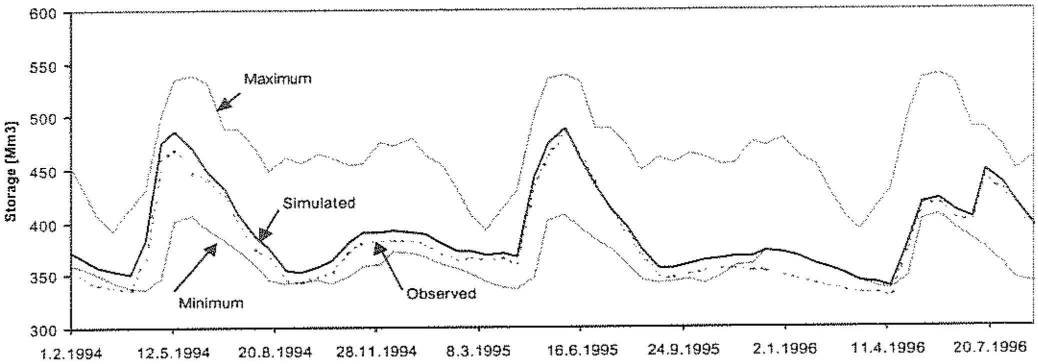


**Figure 5.** Storage trajectory of Keurusselkä in a part of the validation period by the fuzzy rule-based controller and observed data.

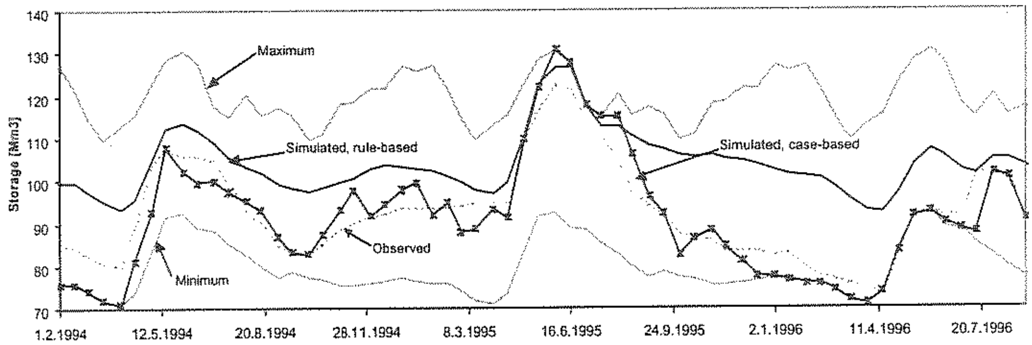


**Figure 6.** Storage trajectory of Mallasvesi in a part of the validation period by the fuzzy rule-based controller, case-based reasoning, and observed data.

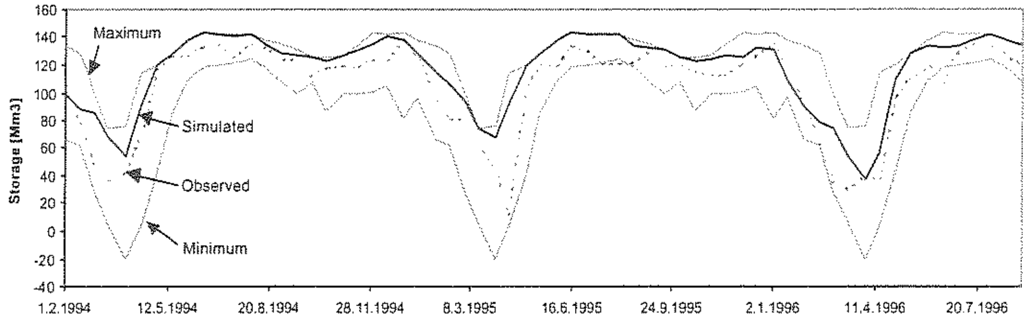


**Figure 7.** Storage trajectory of Pyhäjärvi in a part of the validation period by the fuzzy rule-based controller and observed data.

## 6. DISCUSSION

The fuzzy model succeeds generally rather well in operating the lakes and mimicing the human operator, except in the case of Pyhäjärvi. The problem with Pyhäjärvi is its relatively large throughflow compared to the active storage. The difference in the simulated and observed storage of Mallasvesi in the case of the rule-based model is difficult to explain. Mallasvesi is not operated and thus its outflow is a function of its water level. However, the outflow during a time period of two weeks cannot be calculated from the initial storage and the total inflow during the period.

During the winter 1995–1996 the inflows were exceptionally low and this caused the controller to fail in the operation of Näsijärvi during that period. Similar failure is not visible in other lakes. The model seems to keep the water level of Keurusselkä higher during low flow periods. This could indicate a change in the operating policy between the learning period and validation period.

The membership functions were derived from the data and thus represent only indirectly the objectives of the operators. In reality the membership functions would be the central issue in the discussion about the regulation of the lake. The average water level can be interpreted as the target value and the minimum and maximum as the physical/legal constraints to the regulation. It should be studied if the membership functions could be used in adjusting the model for different conditions.

One of the benefits of the fuzzy model is that the model attempts to keep the lake water level at the average, or target, value only implicitly and not "from the fear of being punished". This can be viewed as more natural than how optimization models behave. The optimization models are often built to try to keep the water level or discharge on the target by imposing a cost for deviating from it.

## 7. CONCLUSION

From the results it can be concluded that the fuzzy approach is in general suitable for simulating reservoir operation. The developed fuzzy model was, however, in some cases rather sensitive for changes in conditions and susceptible to failure when working outside the teaching data. The applicability of the model for practical studies of operation of these lakes will need further work.

## 8. REFERENCES


Bárdossy, A., and L. Duckstein, *Fuzzy Rule-Based Modeling With Applications To Geophysical, Biological And Engineering Systems*. CRC Press, Boca Raton, 1995.

Fontane, D. G., T. K.Gates, and E. Moncada, Planning reservoir operations with imprecise objectives. *J Water Resour. Plng Mgmt*, 123(3), 154-162, 1997.

Huang, W.-C. Decision support system for reservoir operation. *Water Resour. Bull.*, 32(6), 1221-1232, 1996.

Kosko, B. *Neural Networks and Fuzzy Systems*. Prentice Hall, Englewood cliffs, N.J, 1992.

Russell, S. O., and P. F. Campbell, Reservoir operating rules with fuzzy programming. *J Water Resour. Plng Mgmt*, 122(3), 165-170, 1996.

Saad, M., P. Bigras, A. Turgeon, and R. Duquette, Fuzzy learning decomposition for the scheduling of hydroelectric power systems. *Water Resour. Res.*, 32(1), 179-186, 1996.

Shrestra, B. P., L. Duckstein, and E. Z. Stakhiv, Fuzzy rule-based modeling of reservoir operation. *J Water Resour. Plng Mgmt*, 122(4), 262-269, 1996.

Turunen, E. *Mathematics Behind Fuzzy Logic*, Advances in soft computing. Physica-Verlag, Heidelberg, 1999.