HYDROLOGICAL PROCESSES
Hydrol. Process. **17**, 3749–3762 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hyp.1359

Application of fuzzy logic to forecast seasonal runoff

C. Mahabir, F. E. Hicks* and A. Robinson Fayek
Department of Civil & Environmental Engineering, University of Alberta, Edmonton, Alberta T6G 2G7, Canada

Abstract:

Each spring in Alberta, Canada, the potential snowmelt runoff is forecast for several basins to assess the water supply situation. Water managers need this forecast to plan water allocations for the following summer season. The Lodge Creek and Middle Creek basins, located in southeastern Alberta, are two basins that require this type of late winter forecast of potential spring runoff. Historically, the forecast has been based upon a combination of regression equations. These results are then interpreted by a forecaster and are modified based on the forecaster's heuristic knowledge of the basin. Unfortunately, this approach has had limited success in the past, in terms of the accuracy of these forecasts, and consequently an alternative methodology is needed.

In this study, the applicability of fuzzy logic modelling techniques for forecasting water supply was investigated. Fuzzy logic has been applied successfully in several fields where the relationship between cause and effect (variable and results) are vague. Fuzzy variables were used to organize knowledge that is expressed 'linguistically' into a formal analysis. For example, 'high snowpack', 'average snowpack' and 'low snowpack' became variables. By applying fuzzy logic, a water supply forecast was created that classified potential runoff into three forecast zones: 'low', 'average' and 'high'. Spring runoff forecasts from the fuzzy expert systems were found to be considerably more reliable than the regression models in forecasting the appropriate runoff zone, especially in terms of identifying low or average runoff years. Based on the modelling results in these two basins, it is concluded that fuzzy logic has a promising potential for providing reliable water supply forecasts. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS fuzzy logic; fuzzy expert system; water supply; forecast

INTRODUCTION

Owing to the high demand placed on water resources, water managers are frequently required to ensure a continuous water supply to meet such demands as consumption, agriculture and the environment. Water supply forecasting is an important resource management tool in regions where the annual streamflow is dominated by spring and early summer peak flows, since this volume of water often represents the majority of the natural annual runoff for the basin. To assist water managers with both water quantity and water quality issues, water supply forecasts are typically prepared at the end of the winter, approximately when the maximum snow accumulation has been achieved (Milhous, 1982).

Frequently, a water supply forecast volume is provided as a range of possible values to incorporate uncertainties, such as weather forecasts. Roos (1988) explores many of the uncertainties associated with water supply forecasting and their impact on the forecast. Because the reliability of a water supply forecast increases as the season progresses, an effective water supply forecast must balance the need for accuracy of the forecast water volume with the forecast lead-time. For this reason, water supply forecasts are tailored to the needs of the water managers. Some may forecast a runoff volume a year in advance, whereas others sacrifice lead time for higher accuracy.

Typically, the water supply forecast is developed from the available historical hydrological and meteorological data, selecting only parameters that will be available in real-time operations. Variables are selected

* Correspondence to: F. E. Hicks, Department of Civil & Environmental Engineering, University of Alberta, Edmonton, Alberta T6G 2G7, Canada. E-mail: faye.hicks@ualberta.ca

Received 28 June 2002
Accepted 18 December 2002

based on their correlation to seasonal runoff, and empirical regression analysis techniques are often used to determine the forecast equations (Huber, 1984). Regression equations perform best with large sets of data where both the low and high extremes are represented. This characteristic make them less than ideal for water supply forecasting, where the data are often very limited and the need for reliable forecasting of low volumes is critical. In addition, regression equations are site specific and must be explicitly developed for each water supply forecast site.

The purpose of this investigation was to explore the potential viability of an alternate approach to this water supply forecasting based on fuzzy logic. Based on fuzzy set theory (FST), fuzzy logic is being applied in many areas where empirical relationships are not well defined or impractical to model. The foundations of FST, to deal specifically with non-statistical uncertainties, were first developed by Zadeh (1965). Since that time, other researchers have explored the applicability of fuzzy logic to a variety of problems, including engineering applications (e.g. Siskos, 1982; Seo and Sakawa, 1985; Dubois and Prade, 1989) and a number of reference texts are available (e.g. Dubois, 1980; Sakawa, 1993). Despite the subjectivity of establishing the descriptive variables, fuzzy logic model applications have been widely successful in the field of civil engineering, particularly in situations where there are many uncertainties in the relationships between the input variables and the output results. For example, Fontane *et al*. (1997) explored the application of fuzzy logic for regulating reservoir levels. Preliminary results from See and Openshaw (2000) indicate that fuzzy logic can be used with a combination of soft computing techniques to create sophisticated river-level monitoring and forecasting systems. Hundecha *et al*. (2001) demonstrated that a fuzzy logic approach could be used to simulate actual component hydrologic processes (e.g. snowmelt, evaporation, runoff and basin response) in areas where sufficient data were available to model these processes physically.

In this study, the potential applicability of fuzzy logic to water supply forecasting is investigated for two basins in the Cypress Hills, Alberta, an area within an extensive provincial water supply forecasting region that was identified as a problem area for forecasters. With limited historical data, traditional non-linear multi-regression analysis produced less than satisfactory results for this problem, and modelling the physical processes within the basin was not feasible due to a lack of real-time data.

In this investigation, a fuzzy logic expert system for water supply forecasting is first developed for a basin in the Cypress Hills by evaluating the data with linguistic terms and applying fuzzy logic. The model results are then evaluated relative to the forecasts from the regression modelling techniques currently being used. Finally, the viability of transposing the rules governing runoff to an adjacent basin is investigated.

STUDY SITE

Cypress Hills is a geographic area located in the southeast corner of Alberta, Canada (Figure 1). Lodge Creek (effective drainage area of 908 km^2) and Middle Creek (effective drainage area of 305 km^2) both originate in the Cypress Hills. Water from these basins drains southeastward, through the province of Saskatchewan (Canada) into the Milk River (USA), which in turn flows into the Mississippi River and ultimately into the Gulf of Mexico. The Cypress Hills rise 350 m above the surrounding plains. For this reason, the basin characteristics in the headwaters vary significantly from the characteristics of the arid plains in the lower half of these basins. Several small reservoirs have been created on Lodge and Middle Creeks to support local agriculture. Water management in the region is vital to meet interprovincial water apportionment agreements between the provinces of Alberta and Saskatchewan, and international water apportionment agreements between Canada and the USA.

To assist regional water managers with water allocation, a water supply forecast report is provided annually in March. The water supply forecasts are produced at point locations on each creek as defined by the respective Water Survey of Canada hydrometric stations. Lodge Creek at the Interprovincial Boundary (station 11AB082) is located east of the Alberta–Saskatchewan boundary; Middle Creek at the Interprovincial Boundary (station 11AB009) is located in Alberta, west of the boundary (Figure 1). The water supply forecast is produced

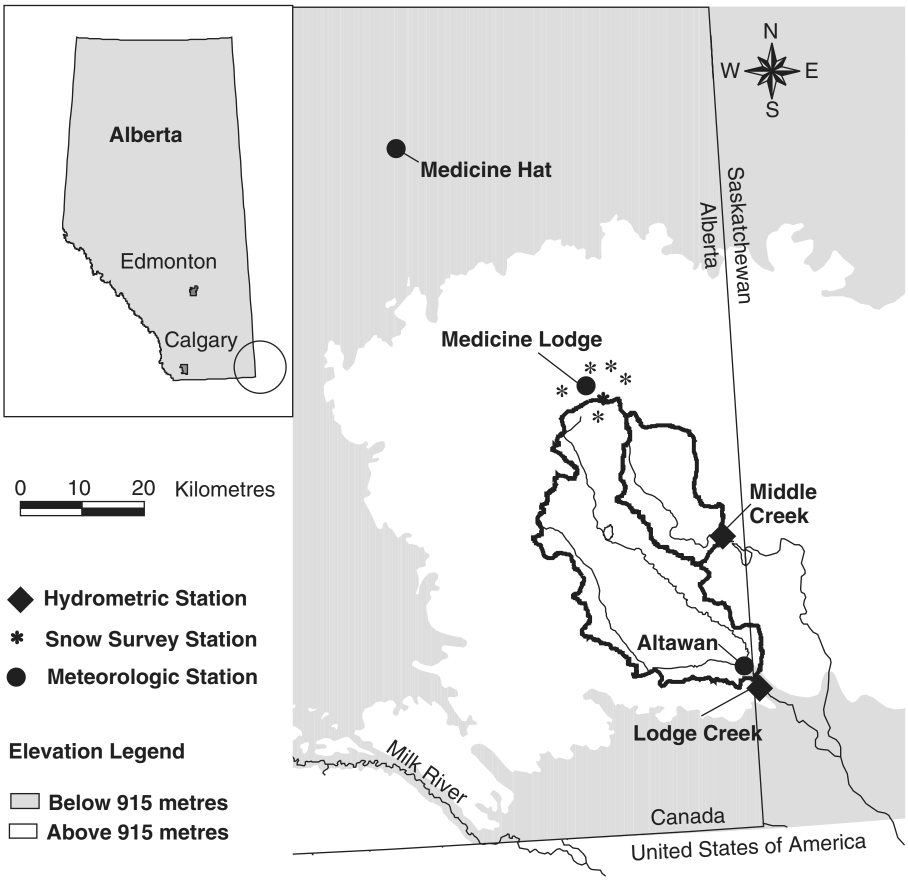

Figure 1. Map of southeast Alberta

for these sites, as they are key monitoring sites in the apportionment agreements. Beginning in 1980, a standard procedure was developed for calculating the natural flow from the Lodge Creek and Middle Creek basins. Details of the calculation are available in the *Handbook for administration of the apportionment agreement*, (Transboundary Waters Unit, 1995). For the purpose of this study, naturalized flow data from the interprovincial apportionment calculations were used to determine seasonal runoff volumes. Since a standard method of calculating the natural flow was applied in 1980, 21 years of flow runoff volumes are available.

Historically, two types of data have been used for water supply forecasts in the Cypress Hills: specifically, late winter snowpack and seasonal precipitation. The snowpack represents the potential runoff stored in the basin. Fall precipitation provides an indication of the antecedent soil moisture conditions leading to the winter period, and the spring precipitation is considered a primary source for generating runoff potential.

The geographic locations of all data collection stations are indicated on Figure 1. Six snow courses are located at varying elevations on the central area of the Cypress Hills. In late February each year, Alberta Environment staff perform standard snow surveys at each of these six sites, with snow depth reported as snow water equivalent (SWE). The arithmetic average of the SWE from these six sites produces a single estimate of SWE for the Lodge Creek and Middle Creek basins.

Precipitation is measured at three sites in the Cypress Hills area. Meteorological Services Canada operates an automated meteorological station in Medicine Hat, approximately 50 km northwest of the Cypress Hills, as well as a manual weather station, Altawan, located on the eastern edge of the Cypress Hills. Alberta

Environment also has an automated meteorological station, Medicine Lodge, located along the plateau crest of the Cypress Hills, which is the newest of these stations (installed in the fall of 1985).

CURRENT REGRESSION FORECASTING TECHNIQUES

Currently, forecasters model the potential basin runoff through regression analysis techniques, relating the natural runoff to various combinations of data from these sites. Several regression models are developed for each site and, operationally, the forecasts from these various models are compared and a potential range of runoff is selected as the forecast. Water management is planned based on the forecasted range of values and adjusted as the year progresses. Therefore, absolute numerical prediction of the regression models is not as important as correctly forecasting a potential range of runoff volume.

For Lodge Creek, the results of three regression models are considered in developing the water supply forecast:

1. Lodge mixed regression, based on a weighted value of fall precipitation from the Medicine Lodge and Medicine Hat stations, and a weighted moisture volume from the spring precipitation at Medicine Lodge and the SWE of the snowpack.
2. Medicine Lodge regression, based on fall and spring precipitation at Medicine Lodge station only, and the SWE of the snowpack.
3. Medicine Hat regression, based on fall and spring precipitation at Medicine Hat station only, and the SWE of the snowpack.

Two regression models have been developed for Middle Creek:

1. Middle mixed regression, based on weighted spring and fall precipitation from the Altawan and Medicine Lodge precipitation data, combined with the SWE snowpack data.
2. Medicine Lodge–Middle Basin, is based on spring and fall precipitation data from Medicine Lodge station only, and the SWE snowpack data.

Four of the five regression models rely on precipitation data from Medicine Lodge, which has only 15 years of data. For the 15 years, these regression models have only been able to forecast the high runoff years qualitatively with any reliability (although it is the low runoff years that are of primary concern). Quantitatively, the variation between the forecasts provided by the various regression models within each basin can be in excess of 40%. This creates low confidence in the models, particularly during low runoff.

FUZZY LOGIC EXPERT SYSTEM

Fuzzy logic has an advantage over many statistical methods in that the performance of a fuzzy expert system is not dependent on the volume of historical data available. Since these expert systems produce a result based on logical linguistic rules, extreme data points in a small data set do not unduly influence these models. Because of these characteristics, fuzzy logic may be a more suitable method for water supply forecasting than current regression modelling techniques.

Creating a fuzzy expert system consists of four basic steps:

1. For each variable, whether an input variable or a result variable, a set of membership functions must be defined. A membership function defines the degree to which the value of a variable belongs to the group and is usually a linguistic term, such as high or low.

2. Statements, or rules, are defined that relate the membership functions of each variable to the result, normally through a series of IF–THEN statements. For example, one rule would be: IF the snowpack (condition) is low (linguistic term represented by a membership function) THEN the runoff (conclusion) is low (linguistic term represented by a membership function).
3. The rules are mathematically evaluated and the results are combined. Each rule is evaluated through a process called implication, and the results of all of the rules are combined in a process called aggregation.
4. The resulting function is evaluated as a crisp number through a process called defuzzification.

Subjective decisions are frequently required in fuzzy logic modelling, particularly in defining the membership functions for variables. In cases such as this, where large data sets are not available to define every potential occurrence scenario for the model, expert opinion is used to create logic.

Membership functions

Membership functions are the most subjective part of fuzzy logic modelling. Each variable must have membership functions, usually represented by linguistic terms, defined for the entire range of possible values. The linguistic terms normally describe a concept related to the value of the variable, such as low, average and high. These linguistic membership functions define the degree μ to which a particular numerical value of a variable fits the concept expressed by the linguistic term. The value of μ ranges from zero (not part of the set) to one (perfectly represents the linguistic concept).

Membership functions can take several shapes, depending on the philosophy behind the concept of the linguistic term. For example, 'average' could be an absolute value, i.e. a single point, as in Figure 2a, or it could be a range of values as in Figure 2b, if factors such as the precision of the measurement were considered to influence the value. With fuzzy logic, membership functions typically overlap, making it possible for a value of a variable to have a membership in more than one linguistic term. For example, 30 mm of precipitation may be considered 'low' to a degree (e.g. $\mu = 0{\cdot}2$) and 'average' to a greater degree (e.g. $\mu = 0{\cdot}5$). The number, shape and range of the membership functions are key factors in the final numerical predictions provided by the fuzzy logic model.

The transition from no membership to 100% membership may take any concave shape. The simplest representation is a straight line. If supported by data or specific knowledge of the subject, then the transition of the membership function can be related by complex equations.

Rule definition

The fuzzy expert system consists of linguistic rules relating the membership functions of the input variables to the membership function of the output variable. A series of IF–THEN statements relates the input to the output. Operators such as AND can be used to relate the input variables to each other to define the result as a

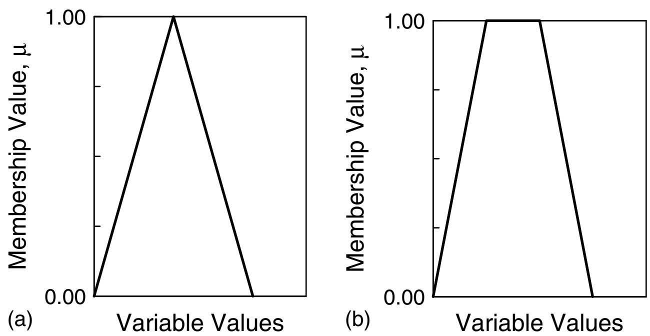

Figure 2. (a) Average as a point value; (b) average as a range of values

combination of the input variables. The AND operator is mathematically applied as an intersection operator by either the 'minimum' or 'product' function. Minimum is commonly used when the input data are independent of each other, and product is often applied if input variables are interdependent (Sun and Robinson Fayek, 2001).

In a rule-based model, the relationship between input variables and the results is easily understood by simply reading the rule. For example: IF snowpack (input variable) is low (membership functions for snowpack) AND fall precipitation (input variable) is low (membership function for fall precipitation) AND spring precipitation (input variable) is low (membership function for spring precipitation) THEN spring runoff (output variable) is low (membership function for spring runoff).

Rules are influential in selecting the number of variables and membership functions to be modelled with fuzzy logic because the model becomes exponentially more complex as the number of variables or membership functions increase. This is because a rule must be available for each possible combination of input variable membership functions and potential outcome membership functions. If a model consists of two input variables m defined by three membership functions n, then nine rules n^m are required to define all situations. For three input variables with three membership functions each, the number of rules increases to 27 (3^3). The addition of unnecessary variables and membership functions would create a more complex model than necessarily required, and may lead to problems defining rules.

Although relationships are easy to define for both the high and low extreme conditions, the intermediate rules are determined by examining historical records. Therefore, although large data sets are not essential, it is desirable to have sufficient data points to define these rules adequately. For cases where there is insufficient data to dictate a rule, rules governing adjacent membership functions can be considered along with knowledge about the relationships. Some rules can be weighted to have more influence on the result than other rules, if there is physical evidence in the data to support a heavier weighting.

Implication and aggregation

'Implication' is a process that evaluates the portion of the membership function that is active for a particular rule. The minimum function defines the implication as follows: if a value of a variable belongs to a set to the degree μ, then the active set area can be considered to be all of the values in the membership function that belong to the membership function to an equal or lesser degree. In contrast, the product method of implication scales the entire membership function by the degree to which the variable belongs. Implication results in one set of values for each rule evaluated.

The sets from implication are combined into a single set in a process called 'aggregation'. If the sets from implication are summed together, then the method of aggregation is called summation. If aggregation of the sets occurs by combining the maximum values obtained for each output membership function after implication, then the maximum method has been used. No firm guidelines have been developed for applying various methods of implication and aggregation. Typically, a sensitivity analysis is performed to determine which methods perform best for a particular fuzzy logic model. More information on these processes is given by Klir *et al*. (1997).

Defuzzification

Defuzzification is the process by which a solution set is converted into a single crisp value. The fuzzy logic solution set is in the form of a function, relating the value of the result to the degree of membership. The entire range of possible solutions may be contained in the fuzzy solution set. Defuzzification is a process to extract an easily comprehendible answer from the set. The centre of area (or centre of gravity) method is one of the most common of the defuzzification methods and consists of selecting the value corresponding to the centre of gravity for the solution set. The bisector method produces a value that will split the area of the solution set in half.

Three other defuzzification methods focus on the maximum membership value attained by the solution set. Frequently the maximum value of the solution set is a range of values rather than a point value. 'Smallest of maxima' selects the lowest value at which the highest membership value is attained. Similarly, 'middle of maxima' and 'largest of maxima' respectively select the middle value and the largest value at which the largest membership value occurs.

The method of defuzzification normally is the most sensitive of the calculation parameters (Robinson Fayek and Sun, 2001). For example, consider the case where a resultant set is composed of 90% membership to 'low' and 10% membership to average. 'Smallest of maxima' will produce the lowest value where 90% membership in 'low' occurs. By contrast, the bisector method will find the value that splits the resultant set in half by area and produce that value as the numerical result. The range of potential resultant values produced by a model can be limited by the method of defuzzification. In the case where the resultant set is shaped as a triangle, the lowest numerical value that can be forecast using the centre of gravity method of defuzzification is the value of the centre of gravity for that particular triangle. However, any of the maxima methods of defuzzification will produce the value where the highest membership occurred, which would be the highest point of the triangle. The objective of the model will influence the selection of defuzzification methods.

FUZZY EXPERT SYSTEM DESIGN

The methods described above were applied to the study basins to investigate the viability of using fuzzy expert systems for water supply forecasting. In order to compare the quality of these forecasts with those obtained using traditional regression methods, the fuzzy expert system was limited to the variables that were used in the regression equations. These included the average SWE, and the fall and spring precipitation data from Medicine Lodge, Medicine Hat, and Altawan.

As discussed above, model complexity increases, in terms of the number of rules to be defined, as each new input variable is added. Therefore, possible cross-correlations between the data from the three precipitation stations were investigated to see if a single station would be sufficiently representative. Medicine Lodge was selected for this purpose, owing to its geographic location in relation to the headwaters of the Middle Creek and Lodge Creek basins. Figure 3 presents the fall and spring precipitation data recorded at the Medicine Hat and Altawan stations versus the corresponding precipitation at the Medicine Lodge station. As the trend lines in Figure 3a illustrate, there was evidence of a reasonable correlation between these stations for the fall precipitation, particularly between Medicine Hat and Medicine Lodge. However, less correlation was apparent in the spring precipitation data between Medicine Hat and Medicine Lodge, and no correlation was found between the Altawan and Medicine Lodge data (Figure 3b).

Based on these results, it is evident that while including the fall precipitation data from Medicine Hat and Altawan may be useful in refining a forecast, Medicine Lodge precipitation is sufficiently representative. Although the spring precipitation is more variable, it will be an unknown quantity in an actual forecasting situation, and the ability to predict precipitation is not refined to the point of differentiating between these two sites. For these reasons, precipitation from a single site, Medicine Lodge, was selected to model spring precipitation.

Membership functions

Three linguistic terms were chosen to describe the input variables and the results: specifically, low, average, and high. Further refinement of the model could be achieved by adding 'extremely low' and 'extremely high' to the groups, but, practically speaking, this was not necessary for this application. In this study, 'average' was considered to be the median value of the data set and, therefore, was a single point value. A triangular membership function was therefore used. Average extends as low as the 25% quartile and as high as the 75% quartile of the data set, as illustrated in for the SWE data set in Figure 4a. Any data value from zero

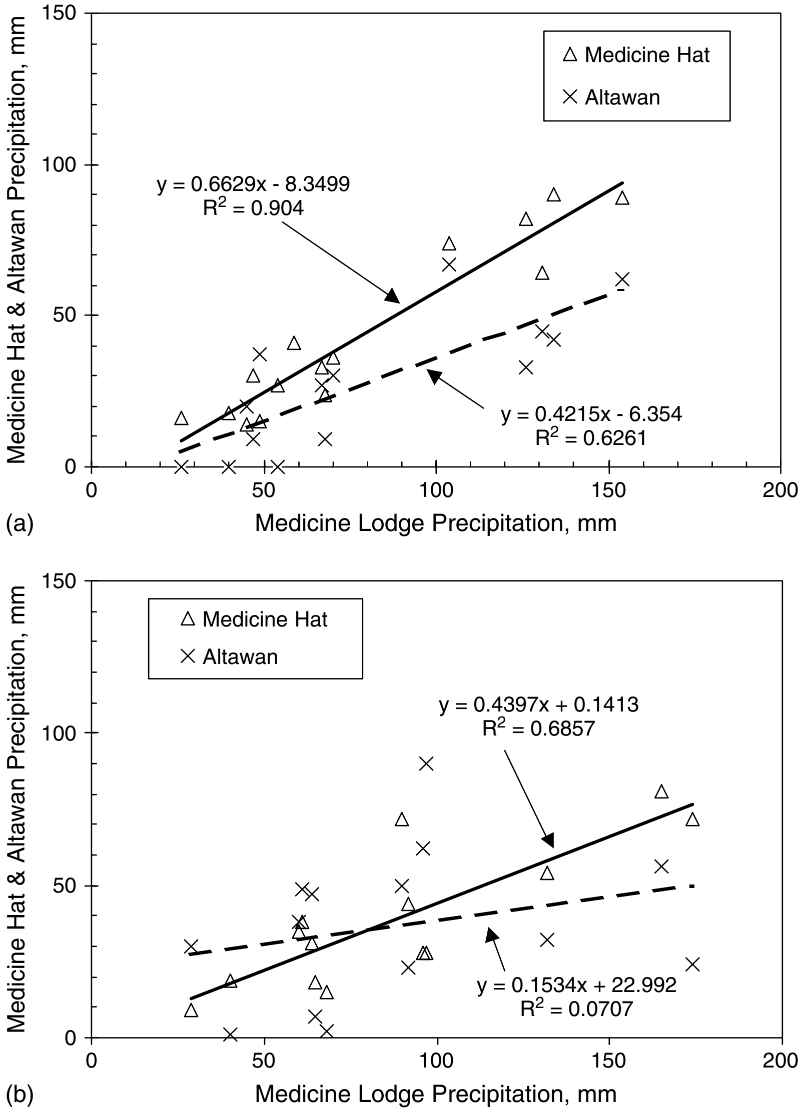

Figure 3. (a) Fall precipitation correlation of stations located on the plains (Medicine Hat, Altawan) with hilltop station (Medicine Lodge). (b) Spring precipitation correlation of stations located on the plains (Medicine Hat, Altawan) with hilltop station (Medicine Lodge)

to below the 25% quartile is definitely 'low', whereas data values between 'low' and the median are 'low' to a lesser degree. Similarly, data points above the 75% quartile to the maximum recorded value are 'high'. Trapezoidal membership functions were used for the low and high terms. Although this defines the range of each function, the transition from 100% membership to 0% membership has not been defined. Since there is not enough data to define the transition functions strongly, straight lines have been used, recognizing that this could be a refinement feature if necessary.

Based on this concept of the data classification, membership functions were determined for fall precipitation (Figure 4b) and spring precipitation (Figure 4c). Applying a similar method of data classification, membership functions were determined for the natural runoff from each basin (Figure 4d and e) from duration curves.

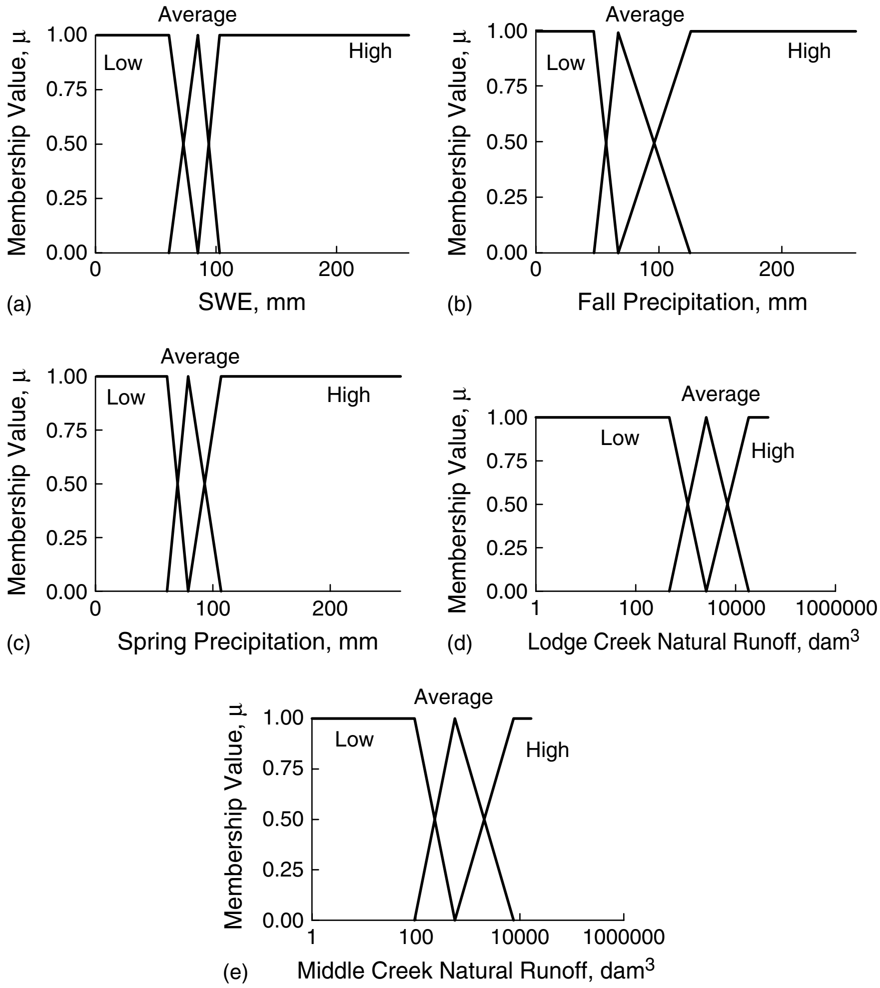

Figure 4. Membership Functions for: (a) SWE; (b) fall precipitation; (c) spring precipitation; (d) natural runoff volume from Lodge Creek; (e) natural runoff volume from Middle Creek

Rule definition

Three years of historical data and expert knowledge were used to create a rule base for the fuzzy logic model. Rules were easy to define for both the high and low extreme conditions, regardless of actual occurrences, because of the physical nature of the relationships. For the intermediate membership functions, expert opinion was combined with the data from three arbitrary years: 1988, 1989 and 1990. Since there are three membership functions for each of the three input variables, the minimum rule base of 27 rules was created. For each data point, all rules were evaluated.

Model construction

The platform selected for the fuzzy logic expert system was MatLab (Version 6.1.0.450) and MatLab's Fuzzy Logic Toolbox (Version 2.1.1). The variables were combined into rules using the concept of 'AND'. The fuzzy operator 'minimum' was applied as the 'AND' function to combine the variables. No weightings were applied, which means no rule was emphasized as more important in respect to estimating the spring runoff. Implication was performed with the minimum function, and aggregation was performed with the maximum function. The centroid, or centre of gravity, method was applied as a means of defuzzification of the output membership functions to determine a crisp set. Based on this structure, a baseline model fuzzy logic expert system for water supply forecasting was constructed for the Lodge Creek basin. Alternate functions for the expert system were investigated through sensitivity analysis.

A fuzzy logic expert system for forecasting water supply for the Middle Creek basin was developed from the Lodge expert system. The same input variables were used, since the same data sets define runoff for the Middle Creek Basin. The rule base remained the same, as it was expected that the relationship between the input variables and the runoff for an adjacent basin should be very similar given the geographic proximity of the two areas. New membership functions for the runoff were developed to represent the runoff conditions for Middle Creek. The operators chosen within the model (implication, aggregation and defuzzification) remained the same.

SENSITIVITY ANALYSIS

A sensitivity analysis was performed for the fuzzy logic operator AND, and for the methods of implication, aggregation and defuzzification. The results of changing a single operator or method while the rest of the model was held constant were compared with the results from the baseline model. The results were evaluated on the basis of correct linguistic matches.

Based on this sensitivity analysis, the AND operator 'minimum' and the implication method 'minimum' were found to perform better than the product method. This supports the use of these operators for independent data input (SWE, fall precipitation and spring precipitation). The model was found to be relatively less sensitive to the method of aggregation, in that both operators, maximum and summation, produced the same number of linguistic matches. Based on this, a prototype model configuration was developed: using minimum for the AND operator; product for the implication; and maximum for the aggregation method. The model results were most sensitive to the method of defuzzification. The centroid and bisector methods produced better results than the smallest, median, and largest of maxima methods.

For the comparison with the regression models, the fuzzy logic expert system consisted of the baseline model configuration modified to use one of the two optimum defuzzification methods, denoting the model name. Specifically, the fuzzy logic expert systems created for the Lodge Basin are the centroid fuzzy logic expert system and the bisector fuzzy logic expert system.

COMPARISON OF FORECAST MODELS

The fuzzy logic and regression models were evaluated based on their ability to forecast in the proper runoff ranges of 'low', 'average' or 'high'. The forecasted runoff values from all models were mapped to the linguistic membership function by evaluating the degree of membership. For example, if a runoff forecast value belonged to 'average' runoff to a degree of 80% and also belonged to 'high' runoff to a degree of 20%, then that particular runoff forecast would be classified as 'average'. By definition, the membership functions for the runoff volume forecasts will cross midway between the median value (100% membership value in 'average') and the 75-percentile value. Runoff values above this location will belong more to 'high' than to 'average'. By applying this definition, a crisp dividing value between membership functions was established. This technique provided a basis for a comparison between the fuzzy model and the regression model in terms of linguistic forecasts.

Tables I and II show the ability of the forecast models to match the actual runoff zone. From these tables, it is clear that High runoff years were adequately modelled both by the regression models and by the fuzzy expert systems. However, for the low to average runoff years, which are of critical importance to water supply forecasting, the fuzzy expert systems were clearly superior. The regression models were only able to forecast an average spring runoff about half of the time, and did very poorly at forecasting low spring runoff years.

Table I. Model results for Lodge Creek spring runoff

	Linguistic matches			
	Low	Average	High	Total correct
Actual runoff	5	4	6	15
Centroid fuzzy logic expert system	5	3	6	14
Bisector fuzzy logic expert system	5	3	6	14
Lodge mixed regression model	0	2	6	8
Medicine Lodge regression model	0	2	6	8
Medicine Hat regression model	2	2	5	9

Table II. Model results for Middle Creek spring runoff

	Linguistic matches			
	Low	Average	High	Total correct
Actual runoff	5	5	5	15
Centroid fuzzy expert system	5	4	5	14
Bisector fuzzy expert system	5	5	5	15
Middle mixed regression model	0	2	5	7
Middle Medicine Lodge regression model	1	3	5	9

In contrast, the fuzzy expert systems were able to forecast all of the high and low runoff years correctly and, at most, missed only one of the average years.

A quantitative comparison between the bisector fuzzy expert system and mixed regression forecast models for Lodge Creek is provided in Figure 5. These results are representative of all fuzzy expert system versus regression comparisons. On the figure, boxed zones have been labelled to assist in visually classifying the boundaries for an accurate linguistic forecast (e.g. low, average, high). It was found that both fuzzy expert systems forecasted the low flow years well, whereas all regression models incorrectly classified the low-flow years into all ranges, including impossible negative values (which were not displayed on the graph). In the Middle Creek basin, similar performances were observed for the fuzzy expert systems compared with the regression models (Figure 6).

CONCLUSIONS

Water supply forecasting is a vital part of water management in many areas. Generally, a range of values for the potential runoff is forecast because of the uncertainty of future meteorological conditions, the lack of available data, and the general lack of knowledge of the empirical interactions between the variables. Regression equations have been used to model potential runoff, but tend to be less reliable for low runoff years when a water supply is the most valuable to water managers. Furthermore, since regression equations are site specific, a model must be created independently for each forecast location. The purpose of this study was to explore the potential to forecast water supply conditions with fuzzy logic by creating a fuzzy logic expert system for Lodge Creek. The model configuration for Lodge Creek was then transposed to forecast runoff from an adjacent basin, Middle Creek.

For the limited data set within this study (15 years), the fuzzy logic expert systems were found to be considerably superior to the regression equations currently used for forecasting the water supply. Although all of the forecast models did equally well forecasting high runoff, the advantage of the fuzzy expert systems

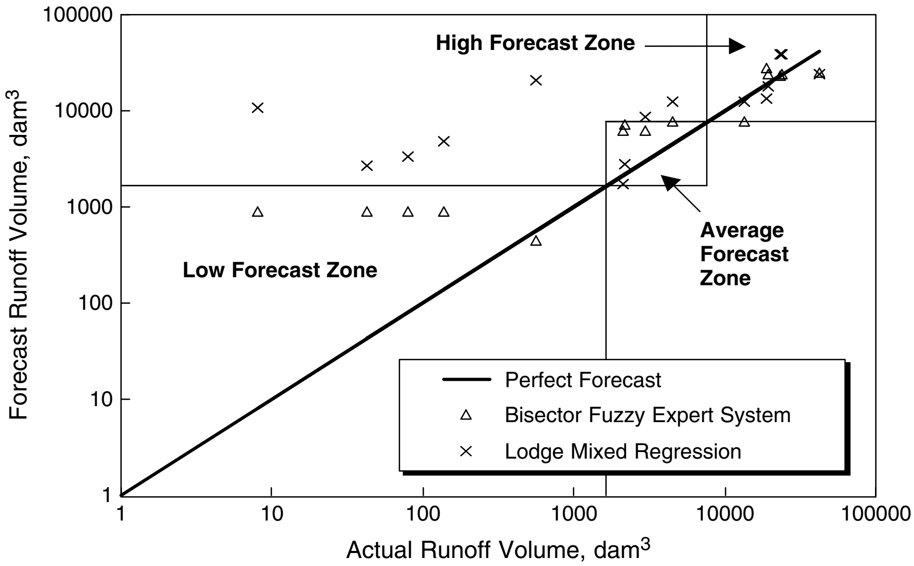

Figure 5. Lodge Creek runoff forecasts accuracy

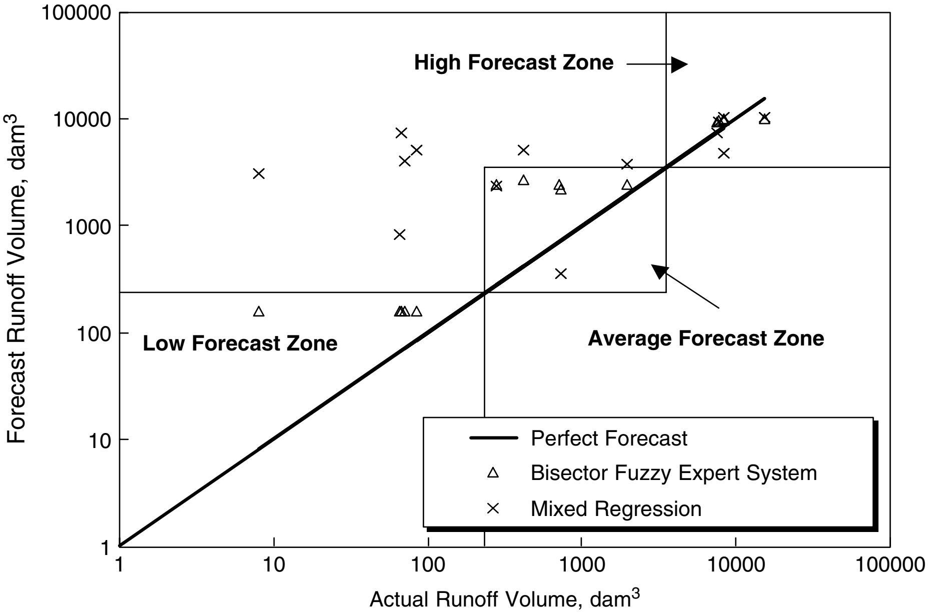

Figure 6. Middle Creek runoff forecasts accuracy

is that they were better able to forecast low runoff years, which are the most critical events in water supply forecasting.

The theory that the general rules relating variables should apply to an adjacent watershed with similar characteristics was tested by transposing the Lodge Creek basin model to the Middle Creek basin. The Middle Creek fuzzy logic expert system, developed from the Lodge fuzzy logic expert system, performed better than the regression models developed specifically for the Middle Creek basin. The rules developed in

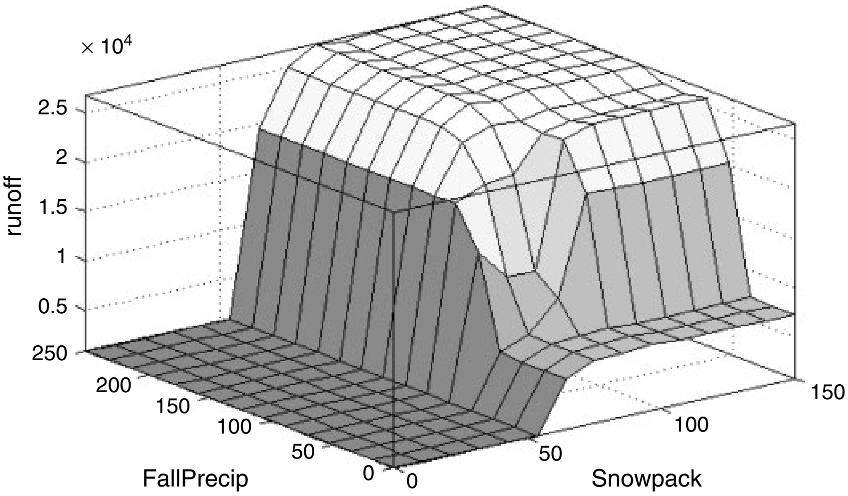

Figure 7. Rule surface for fall precipitation and SWE in relation to runoff

the Lodge fuzzy logic expert system for forecasting runoff were thus found to be applicable to the Middle Creek basin, as the performance of the Middle Creek fuzzy logic expert system was equally as good as the Lodge Creek fuzzy logic expert system.

Although further research in this area is required, this study indicates that fuzzy logic modelling may be very practical for forecasting water supply conditions for small basins with limited data.

It may be possible to obtain a more accurate quantitative forecast with fuzzy logic through the use of more complex membership functions, particularly in terms of the transition between membership groups for both the input and output variables. Results from both the Lodge and the Middle fuzzy expert systems suggest that the transitions in a membership function may be some function other than the assumed linear relationship. By plotting the relationship of the fall precipitation and snowpack (SWE) with respect to runoff (Figure 7), it can be seen that there is a ridge immediately before the maximum plateau is achieved. There is no physical reason why a ridge should occur at this location. Possibilities include that the average runoff membership function should decrease faster than the assumed linear rate or that the high runoff membership function should increase faster than the assumed linear rate. Consequently, had the purpose of this study been to produce a quantitative forecast, the membership functions would need to be investigated further.

It may be worthwhile adding an additional membership function below the 'low' membership to improve the numerical accuracy of the forecast in the low runoff years. Although the fuzzy expert systems adequately forecast the correct runoff zone, it can be seen in both Figures 5 and 6 that the method of defuzzification has limited the lowest value forecast.

To assess the general ability of fuzzy expert systems to forecast spring runoff conditions accurately, more basins in different geographical locations should be analysed. Within Alberta, both mountain and plains runoff watersheds could be investigated to explore the viability of fuzzy logic modelling for water supply forecasting.

ACKNOWLEDGEMENTS

We would like to acknowledge the assistance of Alberta Environment staff, Steve Shio for the preparation of Figure 1 and Neil Assmus for the data collection, and the assistance of the provincial government library staff Mica Wickramasekara, Judy Thomas and Bonnie Kendall.

REFERENCES

Dubois D. 1980. *Fuzzy Sets and Systems, Theory and Applications*. Academic Press: New York.

Dubois D, Prade H. 1989. Fuzzy knowledge in an artificial intelligence system for job-shop scheduling. In *Applications of Fuzzy Set Methodologies in Industrial Engineering*, Evans G, *et al*. (eds). Elsevier: 73–79.

Fontane DG, Gates TK, Moncada E. 1997. Planning reservoir operations with imprecise objectives. *Journal of Water Resources Planning and Management* **123**(3): 154–162.

Huber AL. 1984. Water supply forecasting in real-time. *Water Resources Bulletin*, **20**(2): 167–171.

Hundecha Y, Bardossy A, Theisen HW. 2001. Development of a fuzzy logic-based rainfall-runoff model. *Hydrological Sciences* **46**(3): 363–376.

Klir GJ, St Clair UH, Yuan B. 1997. *Fuzzy Set Theory Foundations and Applications*. Prentice Hall: New Jersey; 245.

Milhous RT. 1982. The importance of streamflow forecasting to water management with instream flow considerations. In *International Symposium on Hydrometeorology*. American Water R esources Association; 207–211.

Robinson Fayek A, Sun Z. 2001. A fuzzy expert system for design performance prediction and evaluation. *Canadian Journal of Civil Engineering* **28**: 1–25.

Roos M. 1988. Water supply forecasting, its uncertainties. *Journal of the American Water Resources Association* 805–814.

Sakawa M. 1993. *Fuzzy Sets and Interactive Multiobjective Optimization*. Plenum Press: New York.

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

Seo F, Sakawa M. 1985. Fuzzy multiattribute utility analysis for collective choice. *IEEE Transactions on Systems, Man, Cybernetics* **15**(1): 37–51.

Siskos J. 1982. A way to deal with fuzzy preferences in multi-criteria decision problems. *European Journal of Operations Research* **10**: 314–324.

Transboundary Waters Unit. 1995. *Handbook for administration of the apportionment agreement*. Report no. 84, Environment Canada.

Zadeh LA. 1965. Fuzzy sets. *Information and Control* **8**: 338–353.

