A fuzzy logic model with genetic algorithm for analyzing fish stock–recruitment relationships

D.G. Chen, N.B. Hargreaves, D.M. Ware, and Y. Liu

Abstract: A new fuzzy logic model with a genetic algorithm is developed that overcomes some of the inherent uncertainties in the fish stock–recruitment process. This model is applied to stock–recruitment relationships for the Southeast Alaska pink salmon (*Oncorhynchus gorbuscha*) and the West Coast Vancouver Island Pacific herring (*Clupea pallasi*) stocks. In both examples, the annual mean sea surface temperature is used as an environmental intervention in the model. The fuzzy logic model provides the functional relationship between the number of fish spawners and the sea surface temperature that is used to reconstruct the historical fish recruitment time series and also to predict the number of fish that will recruit in the future. Globally optimized genetic learning algorithms are used to find the optimal values of the parameters for the fuzzy logic model. The results from this fuzzy logic model are compared with results from both a traditional Ricker stock–recruitment model and a recent artificial neural network model. These comparisons demonstrate the superior capability of the fuzzy logic model for addressing problems of uncertainty and vagueness in both the data and the stock–recruitment relationship. The fuzzy logic model approach is recommended as a useful addition to the analytical tools currently available for fish stock assessment and management.

Résumé : Nous avons élaboré un nouveau modèle à logique floue avec algorithme génétique qui contourne certaines des incertitudes inhérentes au processus de recrutement des stocks de poissons. Nous appliquons ce modèle aux relations stock–recrutement pour les stocks de saumon rose du sud-est de l'Alaska et de hareng de la côte ouest de l'île de Vancouver. Dans les deux exemples, la moyenne annuelle de la température de la surface de la mer représente l'intervention de l'environnement dans le modèle. Le modèle à logique floue fournit la relation fonctionnelle entre le nombre de géniteurs et la température de la surface de la mer qui sert à reconstituer la série chronologique historique du recrutement des poissons, et aussi à prédire le nombre de poissons qui vont recruter dans l'avenir. Nous avons recours à des démarches d'apprentissage génétique globalement optimisées pour trouver les valeurs optimales des paramètres pour le modèle à logique floue. Les résultats obtenus par ce modèle sont comparés à ceux d'un modèle traditionnel stock–recrutement de Ricker et à ceux d'un modèle récent de réseau neuronal artificiel. Ces comparaisons font ressortir la capacité supérieure du modèle à logique floue pour régler les problèmes d'incertitude et d'imprécision tant dans les données que dans la relation stock–recrutement. L'approche du modèle à logique floue nous semble un ajout utile à la panoplie des outils d'analyse qui existent actuellement pour l'évaluation et la gestion des stocks.

[Traduit par la Rédaction]

Introduction

The analysis of stock–recruitment relationships is frequently an important step in developing or evaluating fishery policies, such as establishing optimal escapement goals for salmon or optimal size of spawning stocks for fish species that reproduce in the ocean. Traditional stock–recruitment analyses are based on the assumption that a functional relationship exists between the size of the spawning stock and the number of fish that recruit in the future. If they are included at all, the effects of changing environmental conditions and human interventions (e.g., harvest by fisheries) on the stock–recruitment parameters are typically estimated using least-squares or maximum likelihood methods (Ricker 1975; Hilborn and Walters 1992; Quinn and Deriso 1999).

Received March 20, 2000. Accepted June 30, 2000.
J15670

D.G. Chen,[1] N.B. Hargreaves, and D.M. Ware. Fisheries and Oceans Canada, Pacific Biological Station, Nanaimo, BC V9R 5K6, Canada.
Y. Liu. 16973 60A Avenue, Surrey, BC V3S 5X8, Canada.

[1]Author to whom all correspondence should be sent at the following address: International Pacific Halibut Commission, University of Washington, 250 Oceanography Teaching Building, Seattle, WA 98195, U.S.A.
e-mail: din@iphc.washington.edu

Three major sources of error are embedded in these traditional stock–recruitment analyses. The first source is measurement error that arises from the collection of the data in the field or from experiments. This type of error is commonly referred to as the "errors-in-variables" and has been closely examined (Ludwig and Walters 1981; Walters and Ludwig 1981; Hilborn and Walters 1992).

The second source of error arises from the assumptions made about processes in the analytical model. The Ricker (1975) model has traditionally been widely used to characterize the stock–recruitment relationship for fish in both marine and freshwater populations. Alternative stock–recruitment models have been described by Beverton and Holt (1957), Schnute and Kronlund (1996), and others. Application of these various models to the same stock–recruitment data frequently yields substantially different results.

The third source of error results from the "fit" of the analytical model to the data. Even if the correct model is identified and the data are collected without error, the methods that the model uses to fit the data may produce errors in the parameter estimates. This situation is most severe when there are many local minima or maxima in the stock–recruitment parameter space.

The numerous analytical approaches developed to attempt to resolve these difficulties (e.g., nonlinear regression, likelihood method, Bayesian analysis, bootstrapping, and nonparametric methods) have been reviewed by Hilborn and Walters (1992) and Quinn and Deriso (1999).

Recently, there has been increasing interest in applying machine learning algorithms such as artificial intelligence (AI) neural network models to stock–recruitment analyses (e.g., Lek et al. 1995; Aoki and Komatsu 1997; Guegan et al. 1998). Saila (1996) provided a comprehensive guide for the application of AI models, including neural network and fuzzy logic, to research and management of fisheries and oceans. This approach appears to offer substantial advantages and improvements to the more traditional methods used to analyze stock–recruitment relationships. A neural network model frequently provides a better fit to the data and improved recruitment forecasts because there are fewer constraints on the assumed form of the functional relationships between recruitment, spawning stock, and other environmental or human intervention variables (e.g., Jarre-Teichmann et al. 1995; Chen and Ware 1999; Mackinson et al. 1999). Two of the most attractive features of the genetic algorithm approach are that it appears to be widely applicable and the search procedure for the model parameters is very efficient.

Similar to the AI neural network, fuzzy logic models are designed to mimic how the human brain tends to classify some information or data imprecisely, such as concluding that water temperatures are "cool," "warm," or "hot." In fuzzy logic models, information is processed in terms of fuzzy sets. The fuzzy set is made precise through the definition of an associated membership function. The specific inference is then processed by the fuzzy set combined with some fuzzy rules (e.g., "if the number of fish spawners is high, and the sea surface temperature is cold, then the fish recruitment is high"). The fuzzy logic model combines one or more input signals, which are defined by the fuzzy sets, with a collection of fuzzy rules to produce an output that can be compared with the actual values observed in the real world.

Similar to the AI neural networks, fuzzy logic systems can be used to model nonlinear relationships. Unlike neural networks, however, fuzzy logic systems can provide insight into their own operation because the fuzzy rules provide an easily understood and common sense description of the action of the fuzzy logic system. In this sense, the fuzzy logic system complements the neural network approach for dealing with complex nonlinear problems. A fuzzy logic system is a well-defined function that maps real-value inputs to real-value outputs. Recent results have also shown that fuzzy logic systems are universal approximators (Kosko 1993) for general nonlinear functional relationships, to any desired degree of accuracy. This makes fuzzy logic modelling a powerful tool for exploring complex, nonlinear biological problems like stock–recruitment analyses and forecasting.

In this paper, a fuzzy logic model is developed to analyze the stock–recruitment relationship, including the effects of an environmental variable. The parameters in this model are estimated using a genetic algorithm that minimizes the root mean square error (RMSE) between the observed and modelled recruitment. This new fuzzy logic model with genetic algorithm (hereafter called the Fuzzy–GA model) (*i*) characterizes the uncertainties in the data (the first source of error) by fuzzification of each type of stock–recruitment data into a fuzzy set with associated degrees of membership data (the second source of error), (*ii*) provides a functional or data-driven approximation to the underlying theoretical model, (*iii*) provides estimates of the model parameters that are guarantee to be optimized (the third source of error), and (*iv*) provides a functional relationship that is fitted from the historical stock–recruitment data and can be used to forecast future recruits even if the input data are uncertain or vague (e.g., the size of the spawning stock can be characterized as either "small" or "large" but cannot be quantified more precisely). The performance of the Fuzzy–GA model is compared with the results from a traditional Ricker stock–recruitment model and also with the neural network model proposed by Chen and Ware (1999). Two real data sets, one for pink salmon (*Oncorhynchus gorbuscha*) from Southeast Alaska (SEAK), U.S.A., and the second for Pacific herring (*Clupea pallasi*) from West Coast Vancouver Island (WCVI), British Columbia, Canada, are evaluated.

Methods

Fuzzy logic model

A fuzzy logic model is also known as a fuzzy inference system or a fuzzy rule based system. Basically, any fuzzy logical model constitutes three parts: the fuzzy membership function, fuzzy decision rules, and fuzzy reasoning. In this paper, a fuzzy logic model similar to that described by Takagi and Sugeno (1983) and Liu and Kojima (1993) was used to develop a fuzzy logic model of fish stock–recruitment. There are two input variables in this fuzzy logic model: x_1 is the biomass of spawners (i.e., parents) in the brood year in which the recruits were born and x_2 is the annual mean sea surface temperature (SST) in the same brood year. The output of fuzzy logic model is y, which is the biomass of fish recruits that results from the corresponding spawning in brood year t.

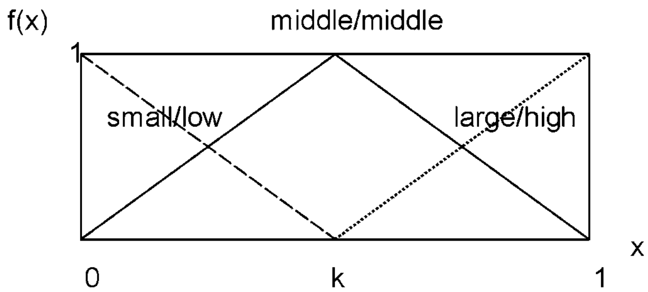
Membership function

The membership function for the input variables x_1 (spawner biomass) and x_2 (SST) is shown in Fig. 1. The values of k in the figure for both input variables (k_1 and k_2) are equal to the mean values of each time series. Using these mean (k) values, the data for the two input variables were classified into three linguistic categories: "small," "middle," or "large" for spawner biomass and "low," "middle," or "high" for SST. The real data for both the SST and spawner biomass are actually much more quantitative than is reflected by these three categories. However, the advantage of using these categories is that we think this more accurately reflects the actual level of uncertainty that is inherent in these data, especially for spawner biomass. For example, estimates of fish spawning biomass tend to be highly variable, and even knowledgeable experts might debate and fail to reach agreement about the "true" value of the spawning biomass for any given year. However, they would likely easily agree on whether the spawning biomass was small, middle (average), or large relative to the long-term average.

Fuzzy rules

There are nine fuzzy rules that are totally determined by the choice of the fuzzy membership function defined in Fig. 1. In general form, each fuzzy rule is written as

Fig. 1. Example of the membership function for the input variables (spawner biomass and SST). The parameter k in each case is chosen to be the mean value of the input variable to classify the input into three fuzzy sets ("small," "middle," and "large" for spawner input and "low," "middle," and "high" for SST input). In the figure, the dashed line is the membership function for the fuzzy "small/low" set, the dotted line on the right is the membership function for the fuzzy "large/high" set and the middle solid lines are the membership function for the fuzzy "middle/middle" set.



$$\text{Rule: if } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } A_2, \text{ then } y = ax_1 + bx_2 + c$$

where A_1 and A_2 are the fuzzy sets that describe the nature of the inputs, such as "small," "medium," or "large," and a, b, and c are the parameters that are to be estimated using the genetic algorithms. Therefore, the "consequent" parts of the nine fuzzy rules are defined by the nonfuzzy equations of the two input variables (x_1 and x_2) as follows:

Rule 1: if x_1 is small and x_2 is low, then
$$y_1 = a_1x_1 + b_1x_2 + c_1$$

Rule 2: if x_1 is small and x_2 is middle, then
$$y_2 = a_2x_1 + b_2x_2 + c_2$$

Rule 3: if x_1 is small and x_2 is high, then
$$y_3 = a_3x_1 + b_3x_2 + c_3$$

Rule 4: if x_1 is middle and x_2 is low, then
$$y_4 = a_4x_1 + b_4x_2 + c_4$$

Rule 5: if x_1 is middle and x_2 is middle, then
$$y_5 = a_5x_1 + b_5x_2 + c_5$$

Rule 6: if x_1 is middle and x_2 is high, then
then $y_6 = a_6x_1 + b_6x_2 + c_6$

Rule 7: if x_1 is large and x_2 is low,
$$y_7 = a_7x_1 + b_7x_2 + c_7$$

Rule 8: if x_1 is large and x_2 is middle, then
$$y_8 = a_8x_1 + b_8x_2 + c_8$$

Rule 9: if x_1 is large and x_2 is high, then
$$y_9 = a_9x_1 + b_9x_2 + c_9$$

where a_i, b_i, and c_i (i = 1 to 9) are the fuzzy parameters to be estimated from the genetic algorithms.

Fuzzy reasoning

Suppose that we have the above "implications" Rule i (i = 1 to 9) and we are also given that ($x_1 = X_1$, $x_2 = X_2$), where X_1 and X_2 are the observed values of x_1 (the fish spawner biomass) and x_2 (the observed annual SST). The model value of y is then inferred from the following steps.

Step 1: For each implication Rule i, y_i is calculated by the function g_i in the consequence (Takagi and Sugeno 1983)

$$y_i = g_i(X_1, X_2) = a_iX_1 + b_iX_2 + c_i.$$

Step 2: The truth value of the "proposition" $y = y_i$ is calculated by the equation

$$|y = y_i| = |X_1 \text{ is } A_{i1}, X_2 \text{ is } A_{i2}| \wedge |R_i| = (A_{i1}(X_1) \wedge A_{i2}(X_2)) \wedge |R_i|$$

where $|*|$ means the truth value of proposition "*" and "^" stands for min operation and $|X \text{ is } A| = A(X)$, i.e., the grade of the membership of X. For simplicity, we assume that $|R_i| = 1$, so the truth value of the consequence obtained is

$$|y = y_i| = A_{i1}(X_1) \wedge A_{i2}(X_2).$$

Step 3: The final output y that is inferred from the nine implications is given as the average of all y_i with the weights $|y = y_i|$:

$$y = \frac{\sum |y = y_i| \times y_i}{\sum |y = y_i|}$$

Table 1 shows the reasoning process for each implication when we are given $x_1 = X_1$ and $x_2 = X_2$. The column "Premise" in Table 1 shows the membership functions of the fuzzy sets "small," "middle," and "large" for the spawner biomass variable and "low," "middle," and "high" for the SST in the premises. The column "Consequence" shows the value of y_i calculated by the function g_i of each consequence for the inputs and corresponding parameters, and the column "Weight" is obtained from the fuzzy "and" operation for the memberships of the inputs. Finally, the fuzzy logic model output value inferred by the implications is obtained as follows:

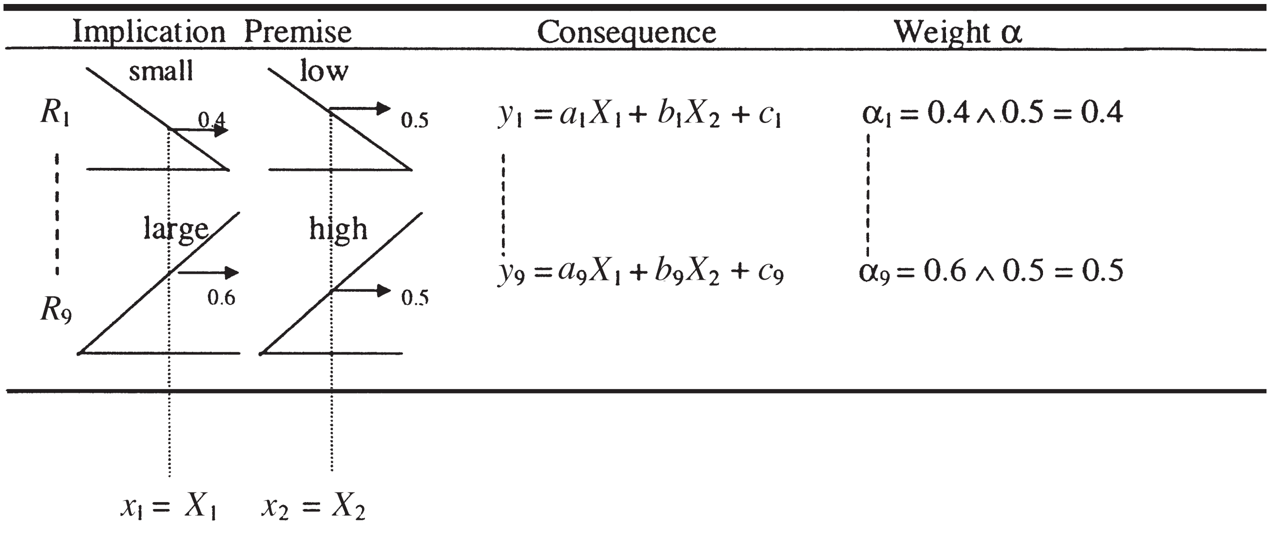
$$(1) \qquad y = \frac{\sum_{i=1}^{9} y_i \alpha_i}{\sum_{i=1}^{9} \alpha_i}$$

A numerical example of the fuzzy reasoning process using data from the WCVI Pacific herring stock is given in the Appendix.

Genetic learning optimization algorithm

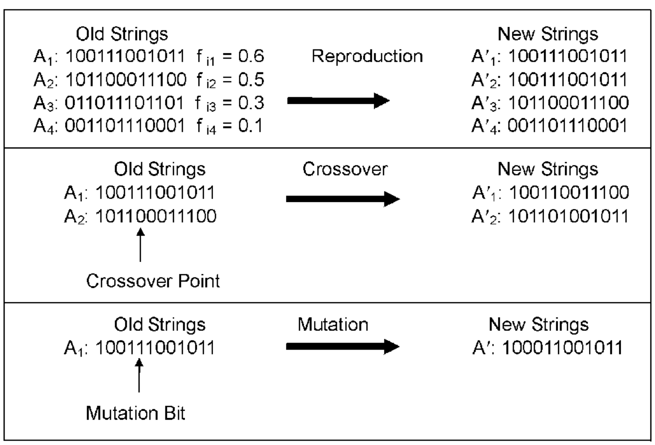
In the above section, we discussed how to construct a fuzzy logic model that can be used to estimate recruitment from various input data. The consequent-part parameters (a_i, b_i, c_i) in the fuzzy logic model depend on the input–output data and usually can be identified by any mathematical minimization procedure. However, it is difficult to find the optimal set of consequent-part parameters when there are a large number that have to be estimated (27 in this case). In such instances, conventional mathematical search algorithms are likely to converge on some local minimum instead of the global minimum. In this paper, we use genetic algorithms to find the optimal consequent-part parameters. Beasley et al. (1993) discussed the advantages of using genetic algorithms in multimodal function optimization. In his programming guide in C/C++, Welstead (1994) compared the genetic algorithm with the most commonly used optimization algorithms, such as the steepest descent and the conjugate gradient method, and concluded that the

Table 1. Fuzzy reasoning process.



Note: "Implication premise" describes the fuzzy membership function for the two inputs spawner abundance (or biomass) and SST under nine fuzzy rules, "Consequence" is the value calculated from each consequence for the inputs and corresponding parameters, and "Weight α" is calculated from the fuzzy "and" operation for the memberships from the two inputs.

Fig. 2. Example of three genetic algorithm operators: reproduction, crossover, and mutation. The string in this case represents one possible combination of all the model parameters.



genetic algorithm is a global optimization algorithm. Detailed discussions of genetic algorithms can be found in Karr (1991) and Liu and Kojima (1993).

Genetic algorithms

A genetic algorithm is a searching algorithm based on the mechanics of natural selection and genetics. A genetic algorithm combines the "survival of the fittest" concept among string structures with a structured yet randomized information exchange to form a search algorithm with some of the innovative flair of human search. In every generation, a new set of strings is created using bits and pieces of the "fittest" strings of the previous generations. While randomized, a genetic algorithm does not perform a random walk through the search space. Instead, it uses random choice efficiently in its exploitation of prior knowledge to locate optimal solutions rapidly. A genetic algorithm differs from more conventional search techniques because it considers many points in the search space simultaneously and therefore has a reduced chance of converging to some local optima.

A genetic algorithm utilizes three fundamental "genetic" operators: reproduction, crossover, and mutation (Fig. 2). Reproduction is a process where strings with large fitness value (i.e., good solutions to the problem at hand) receive correspondingly large numbers of copies in the new population. Crossover provides a mechanism for strings to mix and match their desirable qualities through a random process. The crossover process occurs in three steps. First, two newly reproduced strings are selected. Second, a position along the two strings is selected at random. The third step is to exchange all characters after the selected crossover position. Finally, mutation is the occasional alteration of a value at a particular string position. It is an insurance policy against the permanent loss of any simple bit and enhances the ability of the genetic algorithm to find the optimal solution. By analogy with a biological genetic system where mutations are rare, a higher probability (0.6) is given to the crossover process than to the mutation process (0.005).

Learning algorithm

A common "concatenated mapped, unsigned binary coding" is used to code the consequent parameters of fuzzy logic model. For example, a parameter p, which is the consequent parameter of the fuzzy logic model, is discretized by mapping from a minimum value $p_{\min}$ to a maximum value $p_{\max}$ using an m-bit, unsigned binary integer according to the equation

$$p = p_{\min} + (p_{\max} - p_{\min}) \times \frac{b}{2^m - 1}$$

where b is the integer value represented by an m-bit string (for this application, m = 8). All the parameters are coded using the same method. Accordingly, all of the parameters needed to construct a fuzzy logic model can be represented by one long string with length 216 bits (total 27 parameters with m = 8 bits for each parameter).

A fitness function (a term used in the genetic algorithm, which acts like an objective function in any optimization problem) is required to apply the genetic algorithm. We use the RMSE between the observed recruit numbers (or biomass) and the output from the fuzzy logic model as the fitness function. Figure 3 illustrates the steps and data flow in the genetic learning algorithm. As an example, Fig. 4 shows the change in the fitness function (RMSE) in sub-

Fig. 3. Genetic learning algorithm program and data flow.

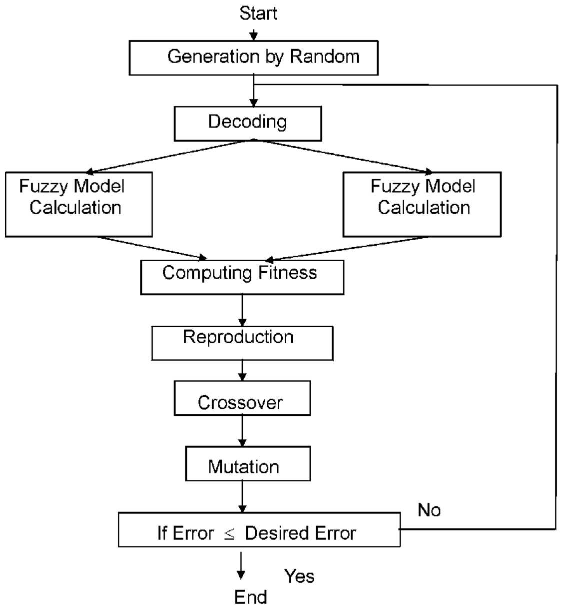
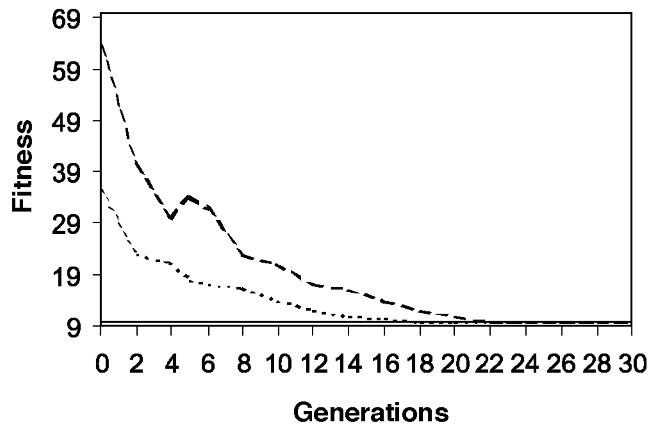


Fig. 4. Changes of fitness versus generation (or iteration) in the genetic algorithm learning process for fitting the WCVI Pacific herring data. A similar figure can be produced for the SEAK pink salmon data. The short-dashed, long-dashed, and solid lines from the top to the bottom represent the worst, median, and best fitness values in the population in each generation.



sequent generations (the term "generation" used in genetic algorithm terminology is equivalent to an "iteration" in other optimization methods). In this case, the uppermost line shows the worst fitness values in the population, the lowest line shows the best fitness values, and the middle line shows the median fitness value in the population. In this example, it takes about 20 generations or learning cycles to find the optimal solution. The genetic learning algorithm is written in Turbo C++. The two input variables and one output variable are standardized to lie in the interval [0,1].

Data sources and results

Two stock–recruitment data sets were analyzed. The first example, taken from Quinn and Deriso (1999), is for SEAK pink salmon from 1960 to 1989. The second example is for the WCVI Pacific herring stock (Chen and Ware 1999). To illustrate the capabilities of the Fuzzy–GA model to reconstruct historical stock–recruitment time series and to forecast future recruits, the analysis is divided into two parts. The first part uses the complete time series to test the ability of the Fuzzy–GA model to reconstruct the historical pink salmon and Pacific herring recruitment time series. In the second part, we reserved 5 years of data (with largest range of variation in recruitment during this period) from the fitting process to independently test the ability of the Fuzzy–GA model to forecast future recruits. To facilitate model comparison, the data in both parts were analyzed using the simple Ricker climatic stock–recruitment model, a neural network model, and the Fuzzy–GA model.

SEAK pink salmon

Data description and preliminary analysis

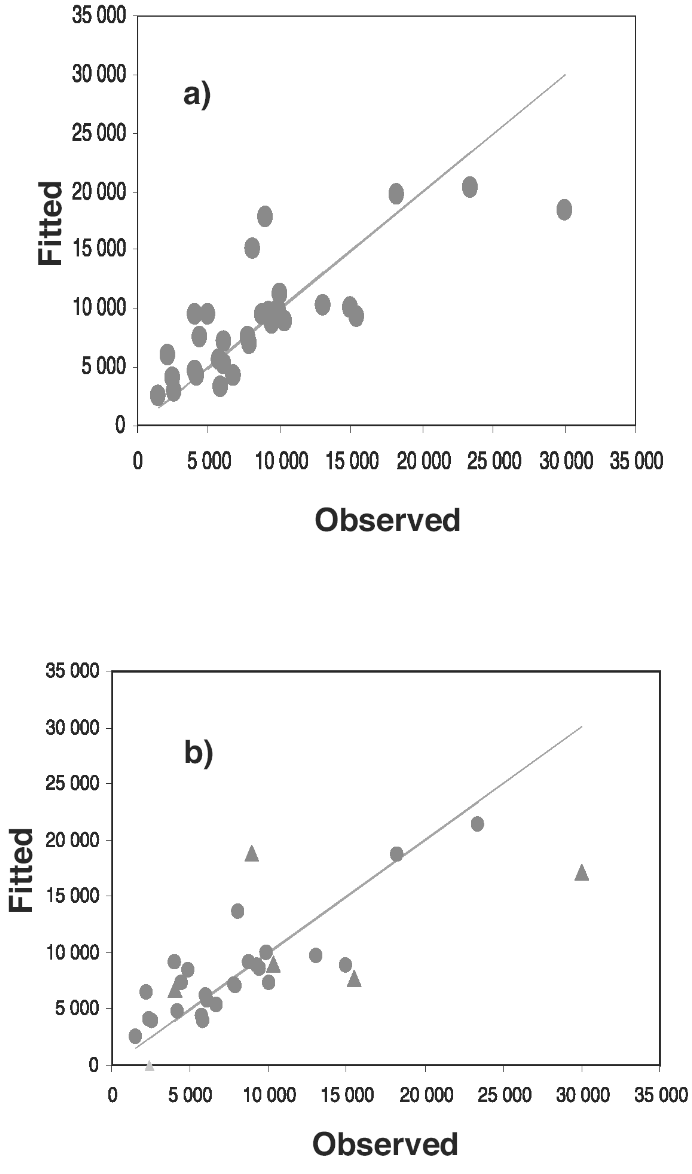
Detailed data descriptions, analyses, and the stock–recruitment time series for SEAK pink salmon from 1960 to 1989 were taken from Quinn and Deriso (1999). The right-hand panels in Fig. 5 reproduce the recruitment time series. To determine the stock–recruitment relationship, Quinn and Deriso (1999) explored eight models, which included the constant model, density-independent model, Bevertion–Holt model, Ricker model, Cushing model, Deriso–Schnute model, Shepherd model, and gamma model. Table 3.2 in Quinn and Deriso (1999) summarized the fitted results for these nine models. In brief, the density-independent model fit the data better than the constant model and had a statistically significant coefficient, which implies that recruitment tends to increase with spawner biomass. For the six density-dependent stock–recruitment models, the estimated density-dependent parameter was not significantly different from zero, and the coefficient of determination (R^2) was only 20–25%.

To try to account for some of the unexplained variation in recruitment, Quinn and Deriso (1999) introduced an environmental factor, average annual SST off Sitka, Alaska, into the analysis. They found that a Ricker stock–recruitment model, which included SST, produced a statistically significant fit to the recruitment time series: the RMSE was 4729.5 and the correlation coefficient (R) was 0.67.

Fuzzy–GA model analysis

To perform the Fuzzy–GA model analysis, the data were rescaled from 0 to 1. For the fuzzy membership function in Fig. 1, k_1 and k_2 for the spawner and SST input variables were chosen to be the mean values (0.25 and 0.5 for the rescaled data) to reflect the fuzzy membership function for fuzzy set "middle." As shown in Fig. 1, the spawner and SST inputs are classified into three categories ("small," "middle," and "high" for spawner input and "low," "middle," and "high" for SST input). This classification leads to 27 fuzzy parameters: a_i, b_i, and c_i (i = 1 to 9) in the nine fuzzy rules. These parameters are estimated from genetic algorithm operations to minimize the RMSE (Table 2). The resulting RMSE of the observed and fitted pink salmon recruits is 3838.5 (units in thousands of pink salmon) and R is 0.79. The top panel plots in Fig. 5 illustrate the Fuzzy–GA

Fig. 5. SEAK pink salmon recruitment data. (*a* and *c*) The results using all of the data. (*a*) The relationship between the observed and the Fuzzy–GA modelled recruits with the 1:1 line. (*c*) The recruitment time series, where the line indicates the modelled recruitment estimates. (*b* and *d*) The results for the forecasting analysis. The points indicated by triangles were not included in the fitting process for forecasting. (*b*) The relationship between the observed and the Fuzzy–GA modelled recruits with the 1:1 line. (*d*) The reconstructed recruitment time series. Circles indicate the observed recruitment data points and the five triangles represent the Fuzzy–GA model forecasts.



model reconstruction of the historical SEAK pink salmon recruitment time series.

A 2–2–1 neural network model was also applied to these data (Fig. 6). A detailed description of the model can be found in Chen and Ware (1999). This model produced an RMSE of 4351.3 and R of 0.72. Model fits using one and three hidden neuron models (i.e., a 2–1–1 neural network and a 2–3–1 neural network) produced similar results.

The ability of the Ricker climatic, neural network, and Fuzzy–GA models to fit the pink salmon data is summarized in Table 3. The Fuzzy–GA model produced the best fit based on the RMSE and the value of R between the observed and modelled recruitment. The neural network model performed better than the Ricker climatic model.

Pink salmon recruitment forecasting

Data from 1980 to 1984 were reserved for forecasting because there was a large range in recruitment during this period. The Fuzzy–GA, neural network, and the Ricker climatic models were fit to the rest of the data to train the model. The performance of these models in fitting the data is summarized in Table 3. For the training process, both the Fuzzy–GA and neural network models outperformed the Ricker climatic model.

The resulting models were then used to forecast recruitment for the 1980–1984 period. In this case, the RMSE was used as the performance criterion because the correlation is not a reliable criterion based on only 5 years of data. Table 3 indicates that the Fuzzy–GA model produced a better fore-

Table 2. Optimal fuzzy logic model parameters obtained from the genetic learning algorithm.

	SEAK pink salmon			WCVI Pacific herring		
Rule	*a*	*b*	*c*	*a*	*b*	*c*
1	0.73	1.18	–0.08	–1.02	1.37	–1.15
2	0.26	–0.22	0.27	–0.03	–0.83	0.75
3	–1.63	0.7	–0.04	0.15	0.76	–0.24
4	0.41	0.34	0.03	–0.13	0.07	0.8
5	0.07	0.82	0.04	–0.3	–0.99	0.77
6	0.31	0.3	–0.02	0.04	0.93	–0.52
7	0.03	–0.76	1.24	0.76	1.44	0.82
8	0.5	–0.16	0.02	–1.07	0.8	0.5
9	1.57	1.69	1.85	–1	0.03	0.45

cast than the neural network model, which in turn was better than the Ricker climatic model.

WCVI Pacific herring

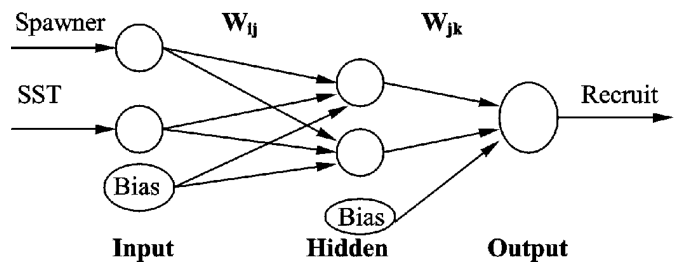
Data description

A long-term research program on the WCVI Pacific herring stock indicates that the key variables that affect the biomass of 3-year-old Pacific herring recruits (in year *t*) are the biomass of spawners (i.e., parents) in the year in which the recruits were born (t – 3 years), the annual mean SST in year t – 3, and the biomass of Pacific hake (*Merluccius productus*) (an important predator) in year t – 3 (Ware 1991; Ware and McFarlane 1995; Chen and Ware 1999). Temperature is believed to be a proxy "signal" that reflects interannual variability in the relative biomass of larval and juvenile Pacific herring predators and possibly some important components of the Pacific herring food supply. In general, cooler (warmer) temperatures tend to produce larger (smaller) recruitments 3 years later. This negative correlation between the annual SST and the biomass of WCVI Pacific herring recruits has been a consistent feature of the recruitment time series since Tester (1948) discovered it about 50 years ago. Recent work indicates that the relationship between parent spawners and recruits is masked by the temperature effect. The underlying dome-shaped spawner–recruit relationship becomes apparent if the recruitment data are sorted into two groups: year-classes born in years of above-average temperature and year-classes born in years of below-average temperature. For each group, there is a significant Ricker-like relationship between spawner and recruit biomass (Ware 1996). To be consistent with the SEAK pink salmon example, we examined the importance of the Pacific herring spawning biomass and the annual SST for recruitment using the Fuzzy–GA model.

Fuzzy–GA model analysis

To implement the model, the input data were rescaled from 0 to 1. Values for k_1 (for spawner biomass) and k_2 (for SST) in the fuzzy membership function illustrated in Fig. 1 were chosen to be the mean values (0.22 for spawner biomass and 0.5 for SST) corresponding to the rescaled data to reflect the fuzzy membership function for fuzzy set "middle." Twenty-seven fuzzy parameters, a_i, b_i, and c_i based on nine fuzzy rules, were estimated by the genetic algorithm and are summarized in Table 3. Figure 4 illustrates the performance of the genetic algorithm for different generations. The final result produced an RMSE between the observed and fitted WCVI Pacific herring recruits of 9.98 and a linear R of 0.72. For comparison, a Ricker climatic model was also fit to the data, with a resulting RMSE of 12.77 and an R of 0.53. We also examined the fit of the data to a 2–1–1, 2–2–1 (Fig. 6), and 2–3–1 neural network model (see Chen and Ware 1999 for details). Similar RMSE and R values were obtained. For example, the RMSE from the 2–2–1 neural network model was 10.22 and R was 0.70.

Fig. 6. Topology of a two-input, two-hidden, and one-output (2–2–1 in neural network terminology) neural network model. The two inputs are spawner biomass and SST. The output is the resulting biomass of recruits. W_{ij} and W_{jk} are weights, which represent the link between the inputs and the output. The weights contain all the information about the network. Therefore, the objective is to train the network to find a series of weights that yield an output signal that has a small error relative to the observed output.



A comparison of the results produced by the three different models is summarized in Table 3. The Fuzzy–GA model produced the best fit to the original recruitment time series based on the estimated RMSE and the R values (Figs. 7*a* and 7*c*). The neural network model performed better than the Ricker climatic model, which is consistent with the results obtained by Chen and Ware (1999).

Pacific herring recruitment forecasting

To compare the forecasting abilities of the three models, we created a new time series where the data for the years 1972–1976 were omitted. This 5-year period was selected because it contained the largest contrast in recruitment (Fig. 7). The Fuzzy–GA, neural network, and Ricker climatic models were fit to the remaining data. The RMSE and the R between the observed and estimated recruits are shown in Table 3. In terms of performance, we obtained the same results with the shorter time series as we did with the complete time series: the Fuzzy–GA model produced the best results followed by the neural network model. The Ricker climatic model produced the poorest fit.

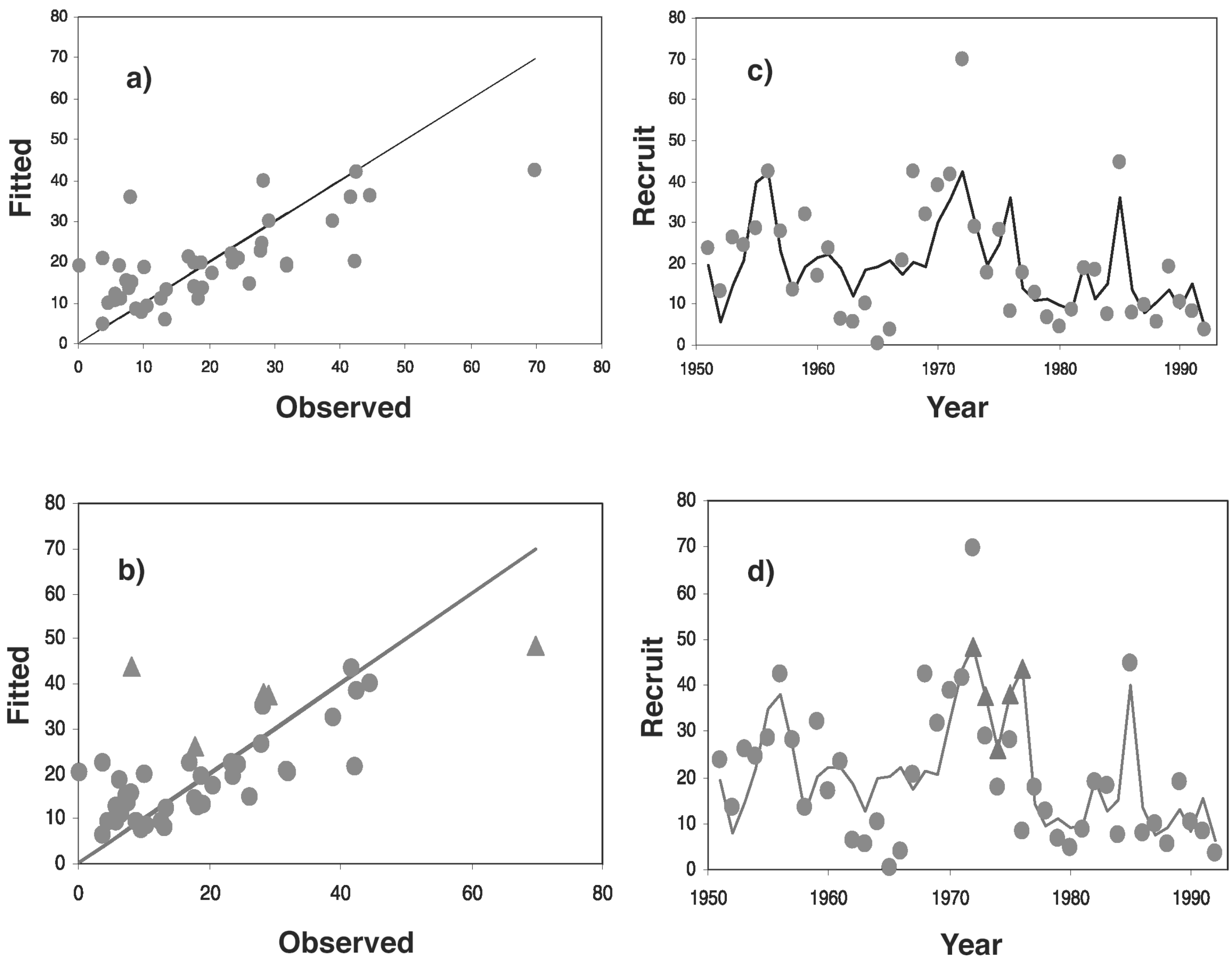
The three models were then used to forecast the abundance of recruits in the 1972–1976 year-classes. In this case, the RMSE was used as the performance criterion because with only 5 years of data, the correlation is not a reliable criterion. The forecasting results for the three models are summarized in Table 3. Once again, the Fuzzy–GA model yielded the best results.

In summary, the descriptive and forecasting capabilities of a Takagi–Sugeno type fuzzy logic model with a genetic learning algorithm are evaluated for two independent sets of fish stock–recruitment. The model is capable of incorporat-

Table 3. Fitting and forecasting summary information for the Ricker climate stock–recruitment, neural network, and Fuzzy–GA models.

		SEAK pink salmon		WCVI Pacific herring	
	Model	*R*	RMSE	*R*	RMSE
All data	Ricker	0.6729	4729.46	0.5291	12.77
	Neural network	0.7191	4351.25	0.7042	10.22
	Fuzzy–GA	0.7899	3838.54	0.7201	9.98
Five years of data reserved					
Fitting	Ricker	0.7640	3357.16	0.4580	11.85
	Neural network	0.7882	3113.09	0.6139	9.88
	Fuzzy–GA	0.8556	2613.82	0.7657	8.06
Forecast	Ricker	0.3722	9111.13	0.5833	26.15
	Neural network	0.4794	8368.87	0.3474	21.69
	Fuzzy–GA	0.4641	8167.48	0.5431	19.82

Fig. 7. WCVI Pacific herring recruitment data. (*a* and *c*) The results using all of the data. (*a*) The relationship between the observed and the Fuzzy–GA modelled recruits with the 1:1 line. (*c*) The reconstructed recruitment time series, where the line indicates the modelled recruitment estimates. (*b* and *d*) The results for the forecasting analysis. The points indicated by the triangles were not included in the fitting process for forecasting. (*b*) The relationship between the observed and the Fuzzy–GA modelled recruits with the 1:1 line. (*d*) The reconstructed recruitment time series. Circles indicate the observed recruitment data points and the five triangles represent the Fuzzy–GA model forecasts.



ing the effects of environmental interventions on the stock–recruitment process. Fuzzy logic operations are used to categorize the input–output information on spawner biomass, recruitment, and environmental interventions into fuzzy sets with an associated degree of membership function. The inherent uncertainties in the data were taken into account by the fuzzification process. The Fuzzy–GA model is capable of empirically approximating the underlying nonlinear stock–recruitment relationship and can also provide a crisp and simple functional relationship among the inputs and outputs according to the fuzzy rules (nine in this application). An important feature of the Fuzzy–GA model is that the functional relationships described by the fuzzy rules can be chosen to more realistically describe the biological processes that affect recruitment.

The application of the genetic algorithm surpasses any conventional mathematical search algorithm because the algorithm used here utilized 240 simultaneous starting points in the searching space. Consequently, the genetic algorithm has a much lower chance of converging to some local optima, especially when several parameters are involved in the optimization. Accordingly, we believe that a fuzzy logic stock–recruitment model with a genetic optimization algorithm can be a useful tool for forecasting recruitment to fish population.

Acknowledgements

We sincerely thank Dr. Ken Newman for his constructive suggestions and comments on this paper. We also thank two anonymous reviewers and the Editors for their helpful comments.

References

Aoki, I., and Komatsu, T. 1997. Analysis and prediction of the fluctuation of sardine abundance using a neural network. Oceanol. Acta, **20**: 81–88.

Beasley, D., Bull, D.R., and Martin, R.R. 1993. A sequential niche technique for multimodal function optimization. Evol. Comput. **1**: 101–125.

Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. 2. Vol. 19. U.K. Ministry of Agriculture and Fisheries, London, U.K.

Chen, D.G., and Ware, D.W. 1999. A neural network model for forecasting fish stock recruitment. Can. J. Fish. Aquat. Sci. **56**: 2385–2396.

Guegan, J., Lek, S., and Oberdorff, T. 1998. Energy availability and habitat heterogeneity predict global riverine fish diversity. Nature (Lond.), **391**: 382–384.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, London, U.K.

Jarre-Teichmann, A., Brey, T., and Haltof, H. 1995. Exploring the use of neural networks for biomass forecasts in the Peruvian upwelling ecosystem. Naga. ICLARM Q. Rep. **18**: 38–40.

Karr, C. 1991. Genetic algorithms for fuzzy controllers. AI Expert, February 1991: 26–33.

Kosko, B. 1993. Fuzzy systems as universal approximators. Proceedings of the 1992 IEEE Conference on Fuzzy Systems (FUZZ-92), March 1992, San Diego, Calif. IEEE Trans. Comput. pp. 1153–1162.

Lek, S., Belaud, A., Dimopoulos, I., Lauga, J., and Moreau, J. 1995. Improved estimation, using neural networks, of the food consumption of fish populations. Mar. Freshwater Res. **46**: 1229–1236.

Liu, Y., and Kojima, H. 1993. One design method of a fuzzy controller using genetic algorithms. *In* Proceedings of the 1993 JSME International Conference on Advanced Mechatronics, August 1993, Tokyo, Japan. The Japan Society of Mechanical Engineers, Shinjuku-ku, Tokyo. pp. 333–338.

Ludwig, D., and Walters, C.J. 1981. Measurement errors and uncertainty in parameters for stock and recruitment. Can. J. Fish. Aquat. Sci. **38**: 711–720.

Mackinson, S., Vasconcellos, M., and Newlands, N. 1999. A new approach to the analysis of stock–recruitment relationships: "model-free estimation" using fuzzy logic. Can. J. Fish. Aquat. Sci. **56**: 686–699.

Quinn, T.J., II, and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. No. 191.

Saila, S.B. 1996. Guide to some computerised artificial intelligence methods. *In* Computers in fisheries research. *Edited by* B.A. Megrey and E. Moksness. Chapman and Hall, London. pp. 8–40.

Schnute, J.R., and Kronlund, A.R. 1996. A management oriented approach to stock recruitment analysis. Can. J. Fish. Aquat. Sci. **53**: 1281–1293.

Takagi, T., and Sugeno, M. 1983. Derivation of fuzzy control rules from human operator's control actions. *In* Proceedings of the IFAC Symposium on Fuzzy Information, Knowledge Representation and Decision Analysis, July 1983, Laxenburg, Austria. pp. 55–60.

Tester, A.L. 1948. The efficacy of catch limitations in regulating the British Columbia herring fishery. Trans. R. Soc. Can. **62**: 135–163.

Walters, C.J., and Ludwig, D. 1981. Effects of measurement errors on the assessment of stock–recruitment relationships. Can. J. Fish. Aquat. Sci. **38**: 704–710.

Ware, D.M. 1991. Climate, predator and prey: behavior of a linked oscillating system. *In* Long-term variability of pelagic fish populations and their environment. *Edited by* T.Kawasaki et al. Pergamon Press, Tokyo, Japan. pp. 279–291.

Ware, D.M. 1996. Herring carrying capacity and sustainable harvest rates in different climate regimes. Pac. Stock Assess. Rev. Comm. Working Pap. H96-3.

Ware, D.M., and McFarlane, G.A. 1995. Climate-induced changes in Pacific hake (*Merluccius productus*) abundance and pelagic community interactions in the Vancouver Island upwelling system. *In* Climate change and northern fish populations. *Edited by* R.J. Beamish. Can. Spec. Publ. Fish. Aquat. Sci. No. 121. pp. 509–521.

Welstead, S.T. 1994. Neural network and fuzzy logic applications in C/C++. John Wiley & Sons, Inc., New York.

Appendix. An example of the fuzzy reasoning process

To illustrate the fuzzy reasoning process, we take the data from year 1951 WCVI Pacific herring stock. The corresponding values for Pacific herring recruitment, spawner biomass, and SST after standardization are 0.34, 0.17, and 0.35, respectively. Using the spawner biomass and SST input values, the modelled values y_i from Rules 1–9 can be calculated from Step 1 with the parameters from Table 2, which are (y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9) = (–0.84, 0.45, 0.05, 0.81, 0.37, –0.18, 1.45, 0.60, 0.30). From the fuzzy membership function defined in Fig. 1, the weight vector α corresponding to Rules 1–9 can be inferred from the min operation (^) in Step 2: $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8, \alpha_9) = (0.23, 0.23, 0, 0.3, 0.7, 0, 0, 0, 0)$. Therefore, the output value from the fuzzy logic model can be obtained from eq. 1 as $y = 0.28$. The model error for the year is then 0.34 – 0.28 = 0.06. The fitness function (i.e., the objective function) is therefore the sum of squares of the errors.