

# Fuzzy Logic in Insurance: the First 20 Years

Arnold F. Shapiro

Smeal College of Business, Penn State University, University Park, PA 16802, USA  
Phone: 814-865-3961 Fax: 914-865-6284 E-mail: afs1@psu.edu

## Abstract

It has been twenty years since DeWit (1982) first applied fuzzy logic (FL) to insurance. That article sought to quantify the fuzziness in underwriting. Since then, the universe of discourse has expanded considerably and now the FL methodologies include used fuzzy set theory, fuzzy numbers and fuzzy arithmetic, fuzzy inference systems, fuzzy clustering algorithms, fuzzy linear programming, fuzzy regression, and hybrids involving other soft computing technologies. Similarly, the applications have been expanded to include classification, projected liabilities, future and present values, pricing, asset allocations and cash flows, and investments. This article presents an overview of these studies. The specific purposes of the article are twofold: first, to review FL applications in insurance so as to document the unique characteristics of insurance as an application area; and second, to document the extent to which FL technologies have been employed.

*Keywords:* actuarial, fuzzy logic, fuzzy sets, fuzzy arithmetic, fuzzy inference systems, fuzzy clustering, insurance

## Acknowledgments:

This work was supported in part by the Robert G. Schwartz Faculty Fellowship and the Smeal Research Grants Program at the Penn State University. The assistance of Asheesh Choudhary, Bharath Nemali, Laura E. Campbell and Michelle L. Fultz is gratefully acknowledged.

# Fuzzy Logic in Insurance: the First 20 Years

## 1 Introduction

It has been twenty years since the first article on fuzzy logic (FL) in insurance [DeWit (1982)].<sup>1</sup> That article sought to quantify the fuzziness in underwriting. Since then, the universe of discourse has expanded considerably and now includes FL applications involving classification, underwriting, projected liabilities, future and present values, pricing, asset allocations and cash flows, and investments.

This article presents an overview of these FL applications in insurance. The specific purposes of the article are twofold: first, to document the FL technologies have been employed in insurance-related areas; and, second, to review the FL applications so as to document the unique characteristics of insurance as an application area.

Before continuing, the term FL needs to be clarified. In this article, we generally follow the lead of Zadeh, the founder of FL, and use the term FL in its broad sense. According to Zadeh,<sup>2</sup>

Fuzzy logic, FL, has four principal facets. First, the logical facet, FL, [the logic of approximate reasoning], which is fuzzy logic in its narrow sense. Second, the set-theoretic facet, FLs, which is concerned with classes having unsharp boundaries, that is, with fuzzy sets. Third, the relational facet, FLr, which is concerned with linguistic variables, fuzzy if-then rules and fuzzy relations. It is this facet that underlies almost all applications of fuzzy logic in control, decision analysis, industrial systems, and consumer products. And fourth, the epistemic facet, FL<sub>e</sub>, which is concerned with knowledge, meaning, and linguistics.

The methodologies of the studies reviewed in this article are based primarily on the first three of these facets, in that they involve fuzzy sets, fuzzy set theory (FST), fuzzy numbers, fuzzy arithmetic, fuzzy inference systems and fuzzy clustering. The term "fuzzy systems" also is used to denote these concepts, as indicated by some of the titles in the reference section of this paper, and will be used interchangeably with the term FL.

This article is subdivided by fuzzy technique.<sup>3</sup> The topics covered include linguistic variables and fuzzy set theory, fuzzy numbers and fuzzy arithmetic, fuzzy inference systems, fuzzy c-means algorithm, fuzzy linear programming, and soft computing. Each section begins with a description of the technique and is followed by a chronological review of the insurance applica-

---

<sup>1</sup> While DeWit was the first to write an article that gave an explicit example of the use of FL in insurance, FL, as it related to insurance, was a topic of discussion at the time. Joseph (1982), for example, remarked that "... not all expert knowledge is a set of "black and white" logic facts - much expert knowledge is codifiable only as alternatives, possibles, guesses and opinions (i.e., as fuzzy heuristics)."

<sup>2</sup> <http://www.eecs.berkeley.edu/IPRO/Summary/03abstracts/zadeh.13.html>

<sup>3</sup> This article could have been structured by fuzzy technique, as was done by Yakoubov and Haberman (1998) or by actuarial topic, as was done by Derrig and Ostaszewski (1999). Given the focus of this conference, the former structure was adopted.

tions of that technique.<sup>4</sup> However, when an application involves more than one technique, it is only discussed in one section. The article ends with a comment on the future of FL in insurance.

## 2 Linguistic Variables and Fuzzy Set Theory

Linguistic variables are the building blocks of FL. They may be defined (Zadeh, 1975, 1981) as variables whose values are expressed as words or sentences. Risk capacity, for example, may be viewed both as a numerical value ranging over the interval  $[0,100\%]$ , and a linguistic variable that can take on values like high, not very high, and so on. Each of these linguistic values may be interpreted as a label of a fuzzy subset of the universe of discourse  $X = [0,100\%]$ , whose base variable,  $x$ , is the generic numerical value risk capacity. Such a set, an example of which is shown in Figure 1, is characterized by a membership function (MF),  $\mu_{\text{high}}(x)$  here, which assigns to each object a grade of membership ranging between zero and one.

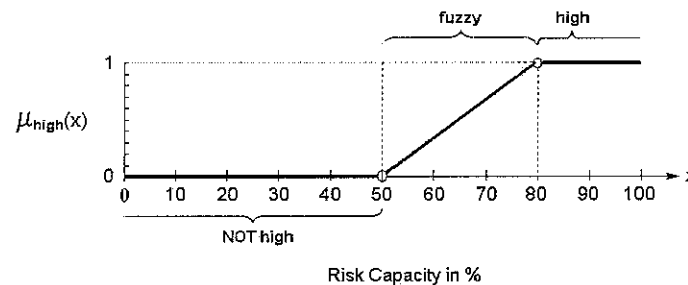


Figure 1: (Fuzzy) Set of Clients with High Risk Capacity

In this case, which represents the set of clients with a high risk capacity, individuals with a risk capacity of 50 percent, or less, are assigned a membership grade of zero and those with a risk capacity of 80 percent, or more, are assigned a grade of one. Between those risk capacities, (50%, 80%), the grade of membership is fuzzy.

If the MF has the shape depicted in Figure 1, it is characterized as S-shaped. Figure 2 shows examples of four other commonly used classes of MFs: triangular, trapezoidal, Gaussian, and generalized bell.

---

<sup>4</sup> Only a cursory review of the FL methodologies is discussed in this paper. Readers who prefer a more extensive introduction to the topic, with an insurance perspective, are referred to Ostaszewski (1993). Those who are interested in a comprehensive introduction to the topic are referred to Zimmermann (1996) and DuBois and Prade (1997). Readers interested in a grand tour of the first 30 years of fuzzy logic are urged to read the collection of Zadeh's papers contained in Yager et al (1987) and Klir and Yuan (1996).

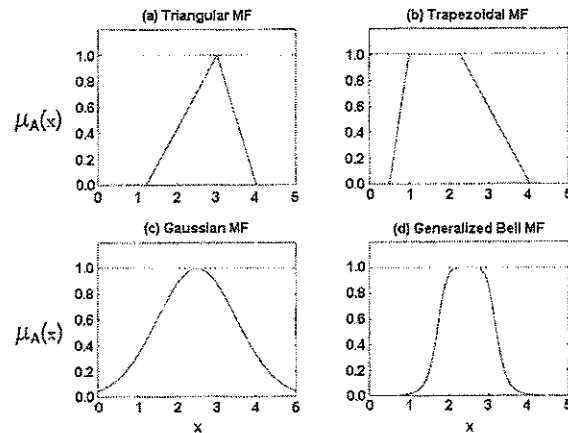


Figure 2: Examples of Classes of MFs

A useful concept insofar as MFs is the  $\alpha$ -cut, an example of which is depicted in Figure 3<sup>5</sup>

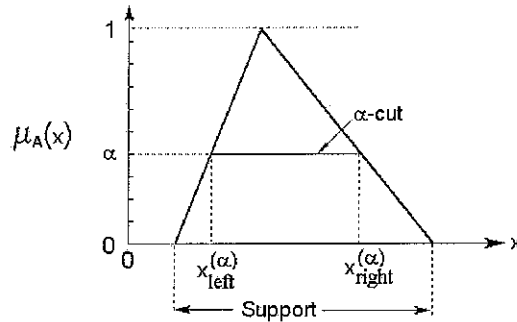


Figure 3: Alpha Cut

As indicated, the essence of the  $\alpha$ -cut is that it limits the domain under consideration to the set of elements with degree of membership of at least alpha. Thus, while the support of fuzzy set A is its entire base, its  $\alpha$ -cut is from  $x_{left}^{(\alpha)}$  to  $x_{right}^{(\alpha)}$ . Values outside that interval will be considered to have a level of membership that is too insignificant to be relevant and should be excluded from consideration, that is, cut out.

Fuzzy sets are implemented by extending many of the basic identities that hold for ordinary sets. Thus, for example, the union of fuzzy sets A and B often is defined as the smallest fuzzy set containing both A and B, and the intersection of A and B often is defined as the largest fuzzy set which is contained in both A and B

## 2.1 Applications

This subsection presents an overview of some insurance applications of linguistic variables and

<sup>5</sup> Adapted from Sinha and Gupta (2000), Figure 7.13.

fuzzy set theory. The topics addressed include: earthquake insurance, optimal excess of loss retention in a reinsurance program, the selection of a "good" forecast, where goodness is defined using multiple criteria that may be vague or fuzzy, resolve statistical problems involving sparse, high dimensional data with categorical responses, the definition and measurement of risk from the perspective of a risk manager, and deriving an overall disability Index.

An early study was by Boissonnade (1984), who used pattern recognition and FL in the evaluation of seismic intensity and damage forecasting, and for the development of models to estimate earthquake insurance premium rates and insurance strategies. The influences on the performance of structures include quantifiable factors, which can be captured by probability models, and non-quantifiable factors, such as construction quality and architectural details, which are best formulated using fuzzy set models. For example, he defined the percentage of a building damaged by an earthquake by fuzzy terms such as medium, severe and total. Two methods of identifying earthquake intensity were presented and compared. The first method was based on the theory of pattern recognition where a discriminative function was developed using Bayes' criterion and the second method applied FL.

Lemaire (1990) envisioned the decision-making procedure in the selection of an optimal excess of loss retention in a reinsurance program as essentially a maximin technique, similar to the selection of an optimum strategy in noncooperative game theory. As an example, he considered four decision variables (two goals and two constraints) and their membership functions: probability of ruin, coefficient of variation, reinsurance premium as a percentage of cedent's premium income (Rel. Reins. Prem.) and deductible (retention) as a percentage of cedent's premium income (Rel. Deductible). His focus was on grades of membership for the decision variables (where the vertical lines cut the MFs) and their degree of applicability (DOA), or rule strength.

In the example he presented, the relative reinsurance premium has the minimum membership value and defines the degree of applicability for this particular excess of loss reinsurance program. The optimal program is the one with the highest degree of applicability.

Cummins and Derrig (1993 p. 434) studied fuzzy trends in property-liability insurance claim costs as a follow-up to their assertion that "the actuarial approach to forecasting is rudimentary." The essence of the study was that they emphasized the selection of a "good" forecast, where goodness was defined using multiple criteria that may be vague or fuzzy, rather than a forecasting model. They began by calculating several possible trends using accepted statistical procedures<sup>6</sup> and for each trend they determined the degree to which the estimate was good by intersecting the fuzzy goals of historical accuracy, unbiasedness and reasonableness.

As an example they investigated the interaction of fuzzy membership values for 72 forecasts according to historical accuracy (goal 1), ordered from best to worst, and unbiasedness (goal 2), before intersection and after intersection. They suggested that one may choose the trend that has

---

<sup>6</sup> Each forecast method was characterized by an estimation period, an estimation technique, and a frequency model. These were combined with severity estimates to obtain pure premium trend factors [Cummins and Derrig (1993: Table 1)]

the highest degree of goodness and proposed that a trend that accounts for all the trends can be calculated by forming a weighted average using the membership degrees as weights. They concluded that FL provides an effective method for combining statistical and judgmental criteria in insurance decision-making.

Another interesting aspect of the Cummins and Derrig (1993) study was their  $\alpha$ -cut for trend factors, which they conceptualized in terms of a multiple of the standard deviation of the trend factors beyond their grand mean. In their analysis, an  $\alpha$ -cut corresponded to only including those trend factors within  $2(1-\alpha)$  standard deviations

A novel classification issue was addressed by Manton et. al. (1994), who used FST to resolve statistical problems involving sparse, high dimensional data with categorical responses. They began with a concept of extreme profile, which, for the health of the elderly, two examples might be "active, age 50" and "frail, age 100." From there, their focus was on  $g_{ik}$ , a grade of membership (GoM) score that represents the degree to which the  $i$ -th individual belongs to the  $k$ -th extreme profile in a fuzzy partition, and they presented statistical procedures that directly reflect fuzzy set principles in the estimation of the parameters. In addition to describing how the parameters estimated from the model may be used to make various types of health forecasts, they discussed how GoM may be used to combine data from multiple sources and they analyzed multiple versions of fuzzy set models under a wide range of empirical conditions.

Jablonowski (1996) investigated the use of FST to represent uncertainty in both the definition and measurement of risk, from the perspective of a risk manager. His conceptualization of exposure analysis was composed of a fuzzy representation of (a) the perceived risk, as a contoured function of frequency and severity, (b) the probability of loss, and (c) the risk profile. The risk profile is the intersection of the first two, using only the min operator. He concluded that FST provides a realistic approach to the formal analysis of risk.

Jablonowski (1997) examined the problems for risk managers associated with knowledge imperfections, under which model parameters and measurements can only be specified as a range of possibilities, and described how FL can be used to deal with such situations. However, unlike Jablonowski (1996), not much detail was provided

The last example of this section is from the life and health area. Chen and He (1997) presented a methodology for deriving an Overall Disability Index (ODI) for measuring an individual's disability. Their approach involved the transformation of the ODI derivation problem into a multiple-criteria decision-making problem. Essentially, they used the analytic hierarchy process, a multicriteria decision making technique that uses pairwise comparisons to estimate the relative importance of each risk factors (Saaty 1980), along with entropy theory and FST, to elicit the weights among the attributes and to aggregate the multiple attributes into a single ODI measurement.

### **3 Fuzzy Numbers and Fuzzy Arithmetic**

Fuzzy numbers are numbers that have fuzzy properties, examples of which are the notions of

“around six percent” and “relatively high”. The general characteristic of a fuzzy number (Zadeh, 1975 and Dubois and Prade, 1980) often is represented as shown in Figure 4, although any of the MF classes depicted in Figure 2 can serve as a fuzzy number, depending on the situation.

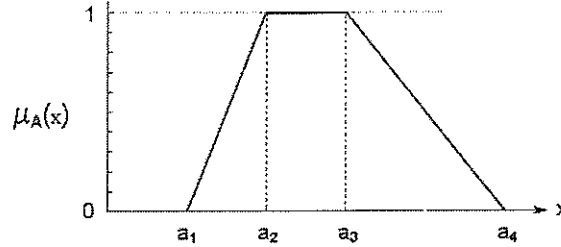


Figure 4: Flat Fuzzy Number

This shape of a fuzzy number is referred to as trapezoidal or “flat” and its MF often is denoted as  $(a_1, a_2, a_3, a_4)$  or  $(a_1/a_2, a_3/a_4)$ ; when  $a_2$  is equal to  $a_3$ , we get the triangular fuzzy number. A fuzzy number is positive if  $a_1 \geq 0$  and negative if  $a_4 \leq 0$ , and, as indicated, it usually is taken to be a convex fuzzy subset of the real line

### 3.1 Fuzzy Arithmetic

As one would anticipate, fuzzy arithmetic can be applied to the fuzzy numbers. Using the extension principle (Zadeh, 1975), the nonfuzzy arithmetic operations can be extended to incorporate fuzzy sets and fuzzy numbers<sup>7</sup>. Briefly, if  $*$  is a binary operation such as addition (+), min ( $\wedge$ ), or max ( $\vee$ ), the fuzzy number  $z$ , defined by  $z = x * y$ , is given as a fuzzy set by

$$\mu_z(w) = \bigvee_{u,v} \mu_x(u) \wedge \mu_y(v), \quad u, v, w \in \mathfrak{R}, \quad (1)$$

subject to the constraint that  $w = u * v$ , where  $\mu_x$ ,  $\mu_y$ , and  $\mu_z$  denote the membership functions of  $x$ ,  $y$ , and  $z$ , respectively, and  $\bigvee_{u,v}$  denotes the supremum over  $u, v$ .<sup>8</sup>

A simple application of the extension principle is the sum of the fuzzy numbers  $A$  and  $B$ , denoted by  $A \oplus B = C$ , which has the membership function:

$$\mu_C(z) = \max \{ \min [\mu_A(x), \mu_B(y)]: x+y=z \} \quad (2)$$

<sup>7</sup>Fuzzy arithmetic is related to interval arithmetic or categorical calculus, where the operations use intervals, consisting of the range of numbers bounded by the interval endpoints, as the basic data objects. The primary difference between the two is that interval arithmetic involves crisp (rather than overlapping) boundaries at the extremes of each interval and it provides no intrinsic measure (like membership functions) of the degree to which a value belongs to a given interval. Babad and Berliner (1995) discussed the use interval arithmetic in an insurance context.

<sup>8</sup>See Zimmermann (1996), Chapter 5, for a discussion of the extension principle.

The general nature of the fuzzy arithmetic operations is depicted in Figure 5 for  $A = (-1,1,3)$  and  $B = (1,3,5)$ <sup>9</sup>.

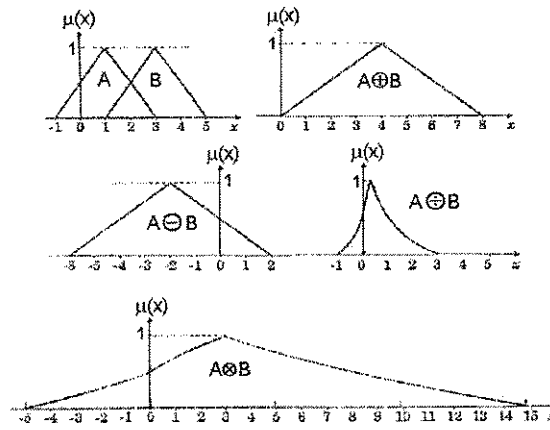


Figure 5: Fuzzy Arithmetic Operations

The first row shows the two membership functions A and B and their sum; the second row shows their difference and their ratio; and the third row shows their product.

### 3.2 Applications

This subsection presents an overview of insurance applications involving fuzzy arithmetic. The topics addressed include: the fuzzy future and present values of fuzzy cash amounts, using fuzzy interest rates, and both crisp and fuzzy periods; the computation of the fuzzy premium for a pure endowment policy; fuzzy interest rate whose fuzziness was a function of duration; net single premium for a term insurance; the effective tax rate and after-tax rate of return on the asset and liability portfolio of a property-liability insurance company; cash-flow matching when the occurrence dates are uncertain; and the financial pricing of property-liability insurance contracts.

Buckley (1987) appears to have been the first author to address the fuzzy time-value-of-money aspects of actuarial pricing, when he investigated the fuzzy future and present values of fuzzy cash amounts, using fuzzy interest rates, and both crisp and fuzzy periods. His approach, generally speaking, was based on the premise that "the arithmetic of fuzzy numbers is easily handled when  $x$  is a function of  $y$ " [Buckley (1987: 258)] For a flat fuzzy number and straight line segments for  $\mu_A(x)$  on  $[a_1, a_2]$  and  $[a_3, a_4]$ , this can be conceptualized as shown in Figure 6

<sup>9</sup>This figure is similar to Musilek and Gupta (2000) Ch 7, Fig 18, p 157, after correcting for an apparent discrepancy in their multiplication and division representations



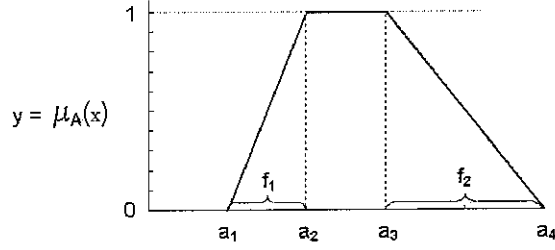


Figure 6: MF and Inverse MF

where  $f_1(y|A) = a_1 + y(a_2 - a_1)$  and  $f_2(y|A) = a_4 - y(a_4 - a_3)$ . The points  $a_j$ ,  $j=1, 2, 3, 4$ , and the functions  $f_j(y|A)$ ,  $j=1,2$ , "A" a fuzzy number, which are inverse functions mapping the membership function onto the real line, characterize the fuzzy number.

If the investment is A and the interest rate per period is i, where both values are fuzzy numbers, he showed that the accumulated value ( $S_n$ ), a fuzzy number, after n periods, a crisp number, is

$$S_n = A \otimes (1 \oplus i)^n \quad (3)$$

because, for positive fuzzy numbers, multiplication distributes over addition and is associative. It follows that the membership function for  $S_n$  takes the form

$$\mu(x|S_n) = (s_{n1}, f_{n1}(y|S_n)/s_{n2}, s_{n3}/f_{n2}(y|S_n), s_{n4}) \quad (4)$$

where, for  $j=1,2$ ,

$$f_{nj}(y|S_n) = f_j(y|A) \cdot (1 + f_j(y|i))^n \quad (5)$$

and can be conceptualized in a manner similar to the above, except that  $a_j$  is replaced with  $S_{nj}$

Then, using the extension principle [Dubois and Prade (1980)], he showed how to extend the analysis to include a fuzzy duration

Buckley then went on to extend the literature to fuzzy discounted values and fuzzy annuities. In the case of positive discounted values, he showed (Buckley 1987 pp. 263-4) that:

If  $S > 0$ , then  $PV_2(S, n)$  exists; otherwise it may not, where:

$$PV_2(S, n) = A \text{ iff } A \text{ is a fuzzy number and } A = S \otimes (1 \oplus i)^{-n} \quad (6)$$

The essence of his argument was that this function does not exist when using it leads to contradictions such as  $a_2 < a_1$  or  $a_4 < a_3$ .

The inverse membership function of  $PV_2(S, n)$  is:

$$f_j(y|A) = f_j(y|S) \cdot (1 + f_{3-j}(y|i))^n, \quad j = 1,2. \quad (7)$$

Both the accumulated value and the present value of fuzzy annuities were discussed<sup>10</sup>

Lemaire (1990), using Buckley (1987) as a model, discussed the computation of the fuzzy premium for a pure endowment policy using fuzzy arithmetic. This involved a MF of the discounted value after ten years at the fuzzy effective interest rate per annum of (.03, .05, .07, .09), and a MF of  $_{10}p_{55}$ . The product of the two resulted in the MF for the present value of the pure endowment.

Ostaszewski (1993: 29-38) extended the pure endowment analysis of Lemaire (1990). First, he incorporated a fuzzy interest rate whose fuzziness was a function of duration. This involved a current crisp rate of 6 percent, a 10-year Treasury Note yield of 8 percent, and a linearly increasing fuzzy rate between the two. Then he investigated the more challenging situation of a net single premium for a term insurance, where the progressive fuzzification of rates plays a major role.

Along the same lines, Terceno et al. (1996) explored the membership functions associated with the net single premium of some basic life insurance products assuming a crisp morality rate and a fuzzy interest rate. Their focus was on  $\alpha$ -cuts, and, starting with a fuzzy interest rate, they gave fuzzy numbers for such products as term insurance and deferred annuities, and used the extension principle to develop the associated membership functions.

Derrig and Ostaszewski (1995b, 1997) illustrated how FL can be used to estimate the effective tax rate and after-tax rate of return on the asset and liability portfolio of a property-liability insurance company. They began with the observation that the effective tax rate and the risk-free rate fully determine the present value of the expected investment tax liability. This leads to differential tax treatment for stocks and bonds, which, together with the tax shield of underwriting losses, determine the overall effective tax rate for the firm. They then argued that the estimation of the effective tax rate is an important tool of asset-liability management and that FL is the appropriate technology for this estimation.

The essence of their paper was contained in an example that showed the membership functions for the fuzzy investment tax rates of a beta one company<sup>11</sup>, with assumed investments, liabilities and underwriting profit, before and after the effect of the liability tax shield. In the assets-only case, the non-fuzzy tax rate is 32.4 percent, but when the expected returns of stocks, bonds, dividends and capital gains are fuzzified, the tax rate becomes the fuzzy number (31%, 32.4%, 32.4%, 33.6%). A similar result occurs when both the assets and liabilities are considered. The authors conclude that, while the outcomes generally follow intuition, the benefit is the quantification, and graphic display, of the uncertainty involved.

Bouet and Dalaud (1996) investigate the use of Zadeh's extension principle for transforming crisp financial concepts into fuzzy ones and the application of the methodology to cash-flow matching. They observe that the extension principle allows them to rigorously define the fuzzy

---

<sup>10</sup> While not pursued here, the use of fuzzy arithmetic in more general finance applications can be found in Calzi (1990) and Simonelli (2001)

<sup>11</sup> A beta one company has completely diversified stock holding, and thus has the same amount of risk ( $\beta = 1$ ) as the entire market

equivalent of financial and economical concepts such as duration and utility, and to interpret them. A primary contribution of their study was the investigation of the matching of cash flows whose occurrence dates are uncertain

The final study of this section is Cummins and Derrig (1997), who used FL to address the financial pricing of property-liability insurance contracts. Observing that much of the information about cash flows, future economic conditions, risk premiums, and other factors affecting the pricing decision is subjective and thus difficult to quantify using conventional methods, they incorporated both probabilistic and nonprobabilistic types of uncertainty in their model. The authors focused primarily on the FL aspects needed to solve the insurance-pricing problem, and in the process "fuzzified" a well-known insurance financial pricing model, provided numerical examples of fuzzy pricing, and proposed fuzzy rules for project decision-making. Their methodology was based on Buckley's inverse membership function and their conceptualization involved a fuzzy loss, the fuzzy present value of that loss, and the fuzzy premium, net of fuzzy taxes. They concluded that FL can lead to significantly different pricing decisions than the conventional approach.

## 4 Fuzzy Inference Systems

The fuzzy inference system (FIS) is a popular methodology for implementing FL. FISs are also known as fuzzy rule based systems, fuzzy expert systems (FES), fuzzy models, fuzzy associative memories (FAM), or fuzzy logic controllers when used as controllers (Jang et al. 1997 p. 73), although not everyone agrees that all these terms are synonymous. Berkan and Trubatch (1997 p. 77), for example, observe that a FIS based on IF-THEN rules is practically an expert system if the rules are developed from expert knowledge, but if the rules are based on common sense reasoning then the term expert system does not apply. The essence of a FIS can be represented as shown in Figure 7.<sup>12</sup>

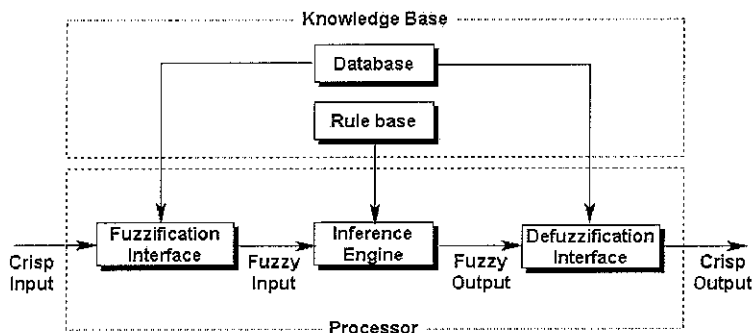


Figure 7: Fuzzy Inference System (FIS)

As indicated in the figure, the FIS can be envisioned as involving a knowledge base and a processing stage. The knowledge base provides MFs and fuzzy rules needed for the process. In the

<sup>12</sup> Adapted from Peña-Reyes and Sipper (1999), Figure 2.

processing stage, numerical crisp variables are the input of the system.<sup>13</sup> These variables are passed through a fuzzification stage where they are transformed to linguistic variables, which become the fuzzy input for the inference engine. This fuzzy input is transformed by the rules of the inference engine to fuzzy output. These linguistic results are then changed by a defuzzification stage into numerical values that become the output of the system.

The operations t-norms (triangular-norms) and t-conorms (its dual) are used in FISs to combine the incoming signals and weights and to aggregate their products. The simplest examples of the t-norm and the t-conorm are the min-operator and max-operator, respectively. Frees and Valdez (1998) show that many copulas can serve as t-norms

The Mamdani FIS, a representation of which is shown in Figure 8, has been the most commonly mentioned FIS in the insurance literature.

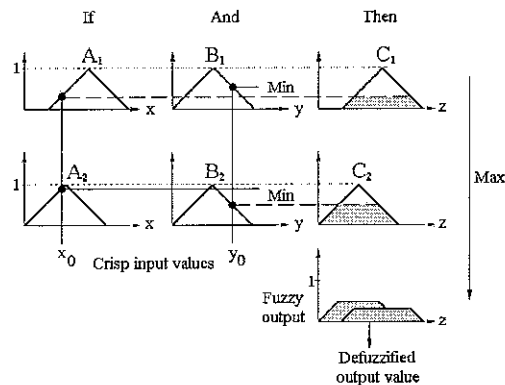


Figure 8: Mamdani FIS

In this case, there are two crisp inputs,  $x_0$  and  $y_0$ , and two sets of membership functions,  $A_j$ ,  $B_j$  and  $C_j$ ,  $j=1,2$ , each set of which represent the rule  $A_j$  and  $B_j \Rightarrow C_j$ , where the conjunction "and" is interpreted to mean the fuzzy intersection. The minimum of the fuzzy inputs in the first two columns gives the levels of the firing (shown by the dashed lines) and their impact on the inference results (shown by the shaded areas in the third column). Taking the union of the shaded areas of the first two rows of column three results in the fuzzy set shown in the third row, which represents the overall conclusion.

Defuzzification converts the fuzzy overall conclusion into a numerical value that is a best estimate in some sense. A common tactic in insurance articles is to use the center of gravity (COG) approach, which defines the numerical value of the output to be the abscissa of the center of gravity of the union. In practice, this is computed as  $\sum_j w_j x_j$ , where the weight  $w_j$  is the relative value of the membership function at  $x_j$ , that is,  $w_j = \mu(x_j) / \sum_j \mu(x_j)$ .

<sup>13</sup> In practice, input and output scaling factors are often used to normalize the crisp inputs and outputs. Also, the numerical input can be crisp or fuzzy. In this latter event, the input does not have to be fuzzified

## 4.1 Applications

This subsection presents an overview of insurance applications of FISs. In most instances, as indicated, an FES was used. The application areas include: life and health underwriting; classification; modeling the selection process in group health insurance; evaluating risks, including occupational injury risk; pricing group health insurance using fuzzy ancillary data; adjusting workers compensation insurance rates; financial forecasting; and budgeting for national health care.

As mentioned above, the first recognition that fuzzy systems could be applied to the problem of individual insurance underwriting was due to DeWit (1982). He recognized that underwriting was subjective and used a basic form of the FES to analyze the underwriting practice of a life insurance company.

Using what is now a common approach, he had underwriters evaluate 30 hypothetical life insurance applications and rank them on the basis of various attributes. He then used this information to create the five membership functions: technical aspects ( $\mu_t$ ), health ( $\mu_h$ ), profession ( $\mu_p$ ), commercial ( $\mu_c$ ), and other ( $\mu_o$ ). Next, by way of example, he combined these membership functions and an array of fuzzy set operations into a fuzzy expert underwriting system, using the formula:

$$W = \left( I(\mu_t) \mu_h \sqrt{\mu_p} \mu_o^2 \sqrt{2 \min(0.5, \mu_c)} \right)^{[1 - \max(0, \mu_c - 0.5)]} \quad (8)$$

where intensification ( $I(\mu_t)$ ) increases the grade of membership for membership functions above some value (often 0.5) and decreases it otherwise, concentration ( $\mu_o^2$ ) reduces the grade of membership, and dilation ( $\sqrt{\mu_p}$ ) increases the grade of membership. He then suggested hypothetical underwriting decision rules related to the values of  $W$ .<sup>14</sup>

Lemaire (1990) used a FES to provide a flexible definition of a preferred policyholder in life insurance. As a part of this effort, he extended the insurance underwriting literature in three ways: he used continuous membership functions; he extended the definition of intersection to include the bounded difference, Hamacher and Yager operators; and he showed how  $\alpha$ -cuts could be implemented to refine the decision rule for the minimum operator, where the  $\alpha$ -cuts is applied to each membership function, and the algebraic product, where the minimum acceptable product is equal to the  $\alpha$ -cut. Whereas DeWit (1982) focused on technical and behavioral features, Lemaire focused on the preferred policyholder underwriting features of cholesterol, blood pressure, weight and smoker status, and their intersection

---

<sup>14</sup> The hypothetical decision rules took the form:

- 0.0  $\leq$  W < 0.1 refuse
- 0.1  $\leq$  W < 0.3 try to improve the condition, if not possible: refuse
- 0.3  $\leq$  W < 0.7 try to improve the condition, if not possible: accept
- 0.7  $\leq$  W < 1.0 accept

An early classification study was Ebanks et. al (1992), which discussed how measures of fuzziness can be used to classify life insurance risks. They envisioned a two-stage process. In the first stage, a risk was assigned a vector, whose cell values represented the degree to which the risk satisfies the preferred risk requirement associated with that cell. In the second stage, the overall degree of membership in the preferred risk category was computed. This could be done using the fuzzy intersection operator of Lemaire (1990) or a fuzzy inference system. Measures of fuzziness were compared and discussed within the context of risk classification, both with respect to a fuzzy preferred risk whose fuzziness is minimized and the evaluation of a fuzzy set of preferred risks.

Following Lemaire's (1990) lead, Hosler (1992) and Young [(previously Hosler) (1993)] used FES to model the selection process in group health insurance. First single-plan underwriting was considered and then the study was extended to multiple-option plans. In the single-plan situation, Young focused on such fuzzy input features as change in the age/sex factor in the previous two years, change in the group size, proportion of employees selecting group coverage, proportion of premium for the employee and the dependent paid by the employer, claims as a proportion of total expected claims, the loss ratio, adjusted for employer size, and turnover rate. She completed the section with a discussion of a matrix of the interaction between the features (criteria) and their interpretation in the context of fuzzy intersection operators.

In the multiple-option case, the additional fuzzy features include single and family age factors, desired participation in each plan, age/sex factors, the difference in the cost of each plan, and the relative richness of each plan. The age factors depended on the possibility of participation, given access cost, the richness of the benefits, employee cost, marital status, and age. The underwriting decision in this case included the single-plan decision as a criterion.

Carreno and Jani (1993) developed a knowledge based system (KBS) that combines fuzzy processing with a rule-based expert system in order to provide an improved decision aid for evaluating life insurance risks. Their system used two types of inputs: the base inputs age, weight and height; and incremental inputs, which deal with particular habits and characteristics of prospective clients. The output of their system was a risk factor used to develop the premium surcharge. One of the advantages that Carreno and Jani identify is the ability of FL to smooth out functions that have jump discontinuities and are broadly defined under traditional methods. By way of example, they investigated risk as a function of age, other characteristics held constant, and replaced a risk function that had jumps at ages 30, 60 and 90, with a FL function where the risk increased smoothly along the entire support.

Another expert opinion-based study was Hellman (1995), which used a FES to identify Finnish municipalities that were of average size and well managed, but whose insurance coverage was inadequate. The study was prompted by a request from the marketing department of his insurance company.

The steps taken included: identify and classify the economic and insurance factors, have an expert subjectively evaluate each factor, preprocessing the conclusions of the expert, and incorporate this knowledge base into an expert system. The economic factors included population size, gross margin rating (based on funds available for capital expenditures), solidity rating, potential

for growth, and whether the municipality was in a crisis situation. The insurance factors were non-life insurance premium written with the company and the claims ratio for three years. Hellman pre-processed the expert's opinion regarding his amount of interest in the non-life insurance premiums written, to construct the associated membership function.

In this instance, two modifications were imposed: first, the piece-wise linear interest function of the expert was replaced with a smooth interest function (equation); and second, a minimum of 20 percent was imposed on the function, in recognition of the advantage gained because the municipality was already a customer. Finally, convex combinations of the interest membership functions became the knowledge base of the FES.

Hellman concluded that the important features of the FESs included that they were easily modified, the smooth functions give continuity to the values, and adding new fuzzy features is easy. Other applications she envisioned included customer evaluation and ratings for bonus-malus tariff premium decisions.

McCauley-Bell and Badiru (1996a, 1996b) discussed a two-phase research project to develop a fuzzy-linguistic expert system for quantifying and predicting the risk of occupational injury of the forearm and hand. The first phase of the research focused on the development and representation of linguistic variables to qualify risk levels. These variables were then quantified using FST. The second phase used analytic hierarchy processing<sup>15</sup> (AHP) to assign relative weights to the identified risk factors. Using the linguistic variables obtained in the first part of the research, a fuzzy rule base was constructed with all of the potential combinations for the given factors.

The study was particularly interesting because, unlike studies such as Derrig and Ostaszewski (1994 and 1995a), which rely on unprocessed expert opinion, McCauley-Bell and Badiru use processed expert opinion. The essential difference is that they use concept mapping to capture a detailed representation of the expert's knowledge relative to the problem space as well as an understanding of the relationships between concepts.

Young (1996) described how FL can be used to make pricing decisions in group health insurance that consistently consider supplementary data, including vague or linguistic objectives of the insurer, which are ancillary to statistical experience data. She conceptualized the building of a fuzzy inference rate-making model as involving: a prescriptive phase based on expert opinion, which verbalizes the linguistic rules, creates the fuzzy sets corresponding to the hypotheses, and determines the output values for the conclusions; and a descriptive phase based on past actions of the company, which involves fine-tuning the fuzzy rules, if applicable. By way of a benchmark, she compared the resulting fuzzy models with linear regressions to judge their performance. Young did not necessarily advocate using such a model without considering experience studies but presented the simplified model to demonstrate more clearly how to represent linguistic rules.

---

<sup>15</sup> Analytic hierarchy processing (AHP) analysis [Saaty (1980)] is a multicriteria decision making technique that uses pairwise comparisons to estimate the relative importance of each risk factors

Young (1996) was extended to include claim experience data in Young (1997). In this later article, Young described step-by-step how an actuary/decision maker could use FL to adjust workers compensation insurance rates. Expanding on her previous article, she walked through the prescriptive and descriptive phases with a focus on targeted adjustments to filed rates and rate departures. In this case, the fuzzy models were fine-tuned by minimizing a weighted sum of squared errors, where the weights reflected the relative amount of business in each state. Young concludes that even though a given FL model may fit only slightly better than a standard linear regression model, the main advantage of FL is that an actuary can begin with verbal rules and create a mathematical model that follows those rules.

A practical application was reported by Horgby et. al. (1997), who applied a FES to the issue of diabetes mellitus in the medical underwriting of life insurance applicants. This was an interesting application because it involved a complex system of mutually interacting factors where neither the prognosticating factors themselves nor their impact on the mortality risk was clear cut. Moreover, it was good example of medical situations where it was easy to reach a consensus among physicians that a disease or a symptom is mild, moderate, or severe, but where a method of quantifying that assessment normally is not available.

In a step-by-step fashion, the authors show how expert knowledge about underwriting diabetes mellitus in life insurance can be processed. Briefly, focusing on the therapy factor, there were three inputs to the system: the blood sugar level, which was represented as very low, low, normal, high, and very high; the blood sugar level over a period of around 90 days, with the same categories; and the insulin injections per week, which had the categories low, medium, high, and very high. The center of gravity (COG) method was used for defussification.

Given the success of the application, the authors concluded that techniques of fuzzy underwriting will become standard tools for underwriters in the future.

Romahi and Shen (2000) developed an evolving rule based expert system for financial forecasting. Their approach was to merge FL and rule induction so as to develop a system with generalization capability and high comprehensibility. In this way the changing market dynamics are continuously taken into account as time progresses and the rulebase does not become outdated. They concluded that their methodology showed promise.

The final review of this section is of a study by Mosmans et. al (2002), which discussed the development of methodological tools for investigating the Belgium health care budget using both global and detailed data. Their model involved four steps: preprocessing the data, segregating the health care channels, validating the channels and data analysis, and calculating the assignment of various categories of the insured to these channels using a multicriteria sorting procedure that was based on t-norms weighted through fuzzy implication operators. The authors concluded that their fuzzy multicriteria sorting procedure could be applied in a more general context and could open new application fields.



## 5 Fuzzy c-Means Algorithm

The foregoing fuzzy system allows us to convert and embed empirical qualitative knowledge into reasoning systems capable of performing approximate pattern matching and interpolation. However, these systems cannot adapt or learn because they are unable to extract knowledge from existing data. One approach for overcoming this limitation is to use a fuzzy clustering method such as the fuzzy c-means algorithm (Bezdek 1981).<sup>16</sup> The essence of the c-means algorithm is that it produces reasonable centers for clusters of data, in the sense that the centers capture the essential feature of the cluster, and then groups data vectors around cluster centers that are reasonably close to them.

A flowchart of the c-means algorithm is depicted in Figure 9:

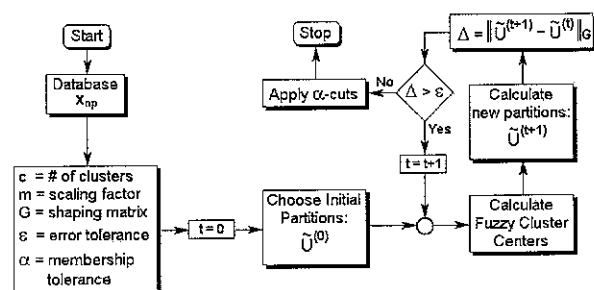


Figure 9: Flowchart of c-Means Algorithm

As indicated, the database consists of the  $n \times p$  matrix,  $\mathbf{x}_{np}$ , where  $n$  indicates the number of patterns and  $p$  denotes the number of features. The algorithm seeks to segregate these  $n$  patterns into  $c$ ,  $2 \leq c \leq n - 1$ , clusters, where the within clusters variances are minimized and the between clusters variances are maximized. To this end, the algorithm is initialized by resetting the counter,  $t$ , to zero, and choosing:  $c$ , the number of clusters;  $m$ , the exponential weight, which acts to reduce the influence of noise in the data because it limits the influence of small values of membership functions;  $G$ , a symmetric, positive-definite (all its principal minors have strictly positive determinants),  $p \times p$  shaping matrix, which represents the relative importance of the elements of the data set and the correlation between them, examples of which are the identity and covariance matrices;  $\varepsilon$ , the tolerance, which controls the stopping rule; and  $\alpha$ , the membership tolerance, which defines the relevant portion of the membership functions.

Given the database and the initialized values, the counter,  $t$ , is set to zero. The next step is to choose the initial partition (membership matrix),  $\tilde{U}^{(0)}$ , which may be based on a best guess or experience. Next, the fuzzy cluster centers are computed, which, in effect, are elements that capture the essential feature of the cluster. Using these fuzzy cluster centers, a new (updated) parti-

<sup>16</sup> The c-means algorithm is discussed here because it is referred to by a number of studies mentioned in this review. However, not all authors advocate the method. Yeo et al (2001), for example, found that the c-means clustering model produced inferior results

tion,  $\tilde{U}^{(t+1)}$ , is calculated. The partitions are compared using the matrix norm  $\|\tilde{U}^{(t+1)} - \tilde{U}^{(t)}\|_G$  and if the difference exceeds  $\varepsilon$ , the counter,  $t$ , is increased and the process continues. If the difference does not exceed  $\varepsilon$ , the process stops. As part of this final step,  $\alpha$ -cuts are applied to clarify the results and make interpretation easier, that is, all membership function values less than  $\alpha$  are set to zero and the function is renormalized.

## 5.1 Applications

This subsection presents an overview of insurance applications of the c-means algorithm. The application areas include: an alternate tool for estimating credibility; risk classification in both life and non-life insurance; and age groupings in general insurance.

Ostaszewski and Karwowski (1992) explored the use of fuzzy clustering methods as an alternate tool for estimating credibility. Given  $\{x_{ij}, i=1, \dots, n, j=1, \dots, p\}$ , a data set representing historical loss experience, and  $\underline{y} = \{y_j, j=1, \dots, p\}$ , a data set representing the recent experience (risk characteristics and loss features), the essential idea is that one can use a clustering algorithm to assign the recent experience to fuzzy clusters in the data. Thus, if  $\mu$  is the maximum membership degree of  $\underline{y}$  in a cluster,  $Z = 1 - \mu$  could be used as the credibility measure of the experience provided by  $\underline{y}$ , while  $\mu$  gives the membership degree for the historical experience indicated by the cluster.

As an example they consider an insurer with historical experience in three large geographical areas extending its business to a fourth large area. The insurer can cluster new data from this fourth area into patterns from the other areas, and thereby derive a credibility rating for its loss experience in the new market. Using the c-means algorithm, the means and standard deviations as features, and two partitions ( $c=2$ ), they arrived at the credibility factor for the data of the fourth area.

Ostaszewski (1993: Chapter 6) observed that lack of actuarially fair classification is economically equivalent to price discrimination in favor of high risk individuals and suggested “... a possible precaution against [discrimination] is to create classification methods with no assumptions, but rather methods which discover patterns used in classification.” To his end, he was among the first to suggest the use of the c-means algorithm for classification in an insurance context.

By way of introducing the topic to actuaries, he discussed an insightful example involving the classification of four prospective insureds, two males and two females, into two clusters, based on the features height, gender, weight, and resting pulse.<sup>17</sup> The two initial clusters were on the basis of gender. In a step-by-step fashion through three iterations, Ostaszewski developed a more efficient classification based on all the features.

---

<sup>17</sup> Age also was a factor, but it was irrelevant to the analysis since the applicants were all the same age. Moreover, the other feature values were intentionally exaggerated for illustrative purposes.

Derrig and Ostaszewski (1994 and 1995a) extended the work of Ostaszewski (1993, Chapter 6) by showing how the c-means clustering algorithm could provide an alternative way to view risk and claims classification. Their focus was on applying fuzzy clustering to the two problems of grouping towns into auto rating territories and the classification of insurance claims according to their suspected level of fraud. Both studies were based on Massachusetts automobile insurance data.

The auto rating territories portion of the study involved 350 towns and the 10 Boston rated subdivisions, with the features bodily injury (BI) liability, personal injury protection (PIP), property damage liability (PDL), collision, comprehensive, and a sixth category comprising the five individual coverages combined. The parameters of the c-means algorithm were five coverage partitions ( $c = 5$ ), which was the number of categories in a previous territory assignment grouping, a scaling factor of 2 ( $m=2$ ), a tolerance of 5 percent ( $\epsilon=0.05$ ), and an  $\alpha$ -cut of 20 percent.

In their study, the initial groups are crisp in the sense that the memberships of the territories were unique. In contrast, as a consequence of applying the c-means algorithm, the optimum classification resulted in some towns belonging to more than one cluster (risk class).

Their second study was based on an interesting use of information derived from experts. Beginning with 387 claims and two independent coders, 62 claims that were deemed fraudulent by either coder were identified. Then, starting with 127 claims (the 62 deemed fraudulent plus 65 from remaining 325), experienced claim managers and experienced investigators were each asked to rank each claim on a scale of 0 to 10. Their responses were grouped into the five initial clusters ( $c=5$ ): none (0), slight (1-3), moderate (4-6), strong (7-9), and certain (10).<sup>18</sup> In this instance, the three features were the adjuster suspicion value, the investigator suspicion value, and a third category labeled the "fraud vote," which was equal to the number of reviewers who designated the claim as fraudulent. The results of their analysis supported the hypothesis that adjuster suspicion levels can serve well to screen suspicious claims. The authors concluded that fuzzy clustering is a valuable addition to the methods of risk and claim classification, but they did not conclude that it was the best way.

The last study of this section is by Verrall and Yakoubov (1999), who showed how the fuzzy c-means algorithm could be used to specify a data-based procedure for investigating age groupings in general insurance. Their database included the total cost of claims associated with more than 50,000 motor policies. Starting with the assumption that distortion effects have already been removed and policyholder age was the only significant factor, they focused on the coverages of automobile material damage and bodily injury.

The heuristic nature of their approach was interesting. They pre-processed the data by grouping the low ages and high ages where data was sparse and categorized it by adjusted frequency, computed as the product of the frequency at each age and the average severity. Then, using an ad hoc approach, they settled on six clusters ( $c=6$ ) and an  $\alpha$ -cut of 20 percent, from which the c-means algorithm results led them to a first approximation of the appropriate age grouping.

---

<sup>18</sup> This data has been used for a number of studies.

They concluded that, while other methods can be used, the flexibility of the fuzzy approach makes it most suitable for grouping policyholder age. They noted also that the algorithm could be applied to other explanatory variables and in other types of insurance, such as the classification of vehicles into vehicle rating groups, the grouping of car engine sizes, and the classification of excess mortality risk in life insurance according to blood pressure.

## 6 Fuzzy Linear Programming

Many of the fuzzy logic studies in insurance involve decision making, and most of these studies rely on the framework established by Bellman and Zadeh (1970). The essential notion is that, given a non-fuzzy space of options,  $X$ , a fuzzy goal,  $G$ , and a fuzzy constraint,  $C$ , then  $G$  and  $C$  combine to form a decision,  $D$ , which is a fuzzy set resulting from the intersection of  $G$  and  $C$ . Assuming the goals and constraints enter into the expression for  $D$  in exactly the same way, a simple representation of the relationship between  $G$ ,  $C$  and  $D$  is given in Figure 10.

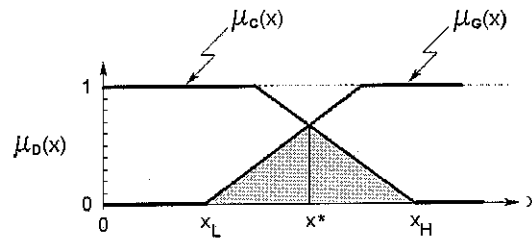


Figure 10: Decision Making

As indicated, the decision involves the fuzzy intersection of the goal and constraint MFs, and the set of possible options in the interval  $x_L$  to  $x_H$ . If the optimal decision is the option with the highest degree of membership in the decision set, the crisp solution to this problem would be

$$x^* = \arg[\max_x \min \{ \mu_G(x), \mu_C(x) \}].$$

In this section, we focus on the role of fuzzy linear programming (LP) in decision making. Like its crisp counterpart, fuzzy LP might involve finding an  $x$  such that (Zimmermann 1996: 289)

$$\begin{aligned} C &= \sum_{ij} c_{ij} x_{ij} \tilde{\leq} C_0 \\ z_i &= \sum_j a_{ij} x_{ij} \tilde{\geq} b_i \\ x_{ij} &\geq 0 \end{aligned}$$

where  $C_0$  is the aspiration level for the objective function, " $\tilde{\sim}$ " over a symbol denotes the fuzzy version of that symbol, and the coefficients  $a_{ij}$ ,  $b_i$ , and  $c_{ij}$  are not necessarily crisp numbers.

This fuzzy LP problem can be resolved by reformulating it as a crisp LP problem. The essence of one approach<sup>19</sup> to doing this is depicted in Figure 11

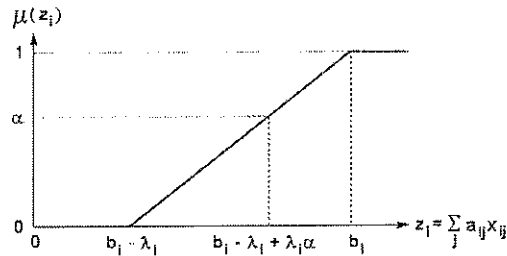


Figure 11: Equivalent Crisp Constraint

As indicated,  $z_i$  is a fuzzy number, whose membership function is zero for  $z_i \leq b_i - \lambda_i$ , one for  $z_i \geq b_i$ , and linearly increasing in the interval  $[b_i - \lambda_i, b_i]$ . Zimmermann refers to  $\lambda$  as a tolerance interval. Using an  $\alpha$ -cut to provide a minimum acceptable satisfaction level, that is,  $\mu(z_i) \geq \alpha$  is an acceptable constraint, we see from the diagram that an equivalent constraint is  $z_i \geq b_i - \lambda_i + \lambda_i \alpha$ . Similarly,  $C \leq C_0 + \lambda - \lambda \alpha$

Thus, given the values of  $\lambda$ , the equivalent crisp programming problem becomes one of maximizing  $\alpha$  subject to the equivalent constraints, that is:

$$\begin{aligned} \text{Maximize:} \quad & \alpha \\ \text{Subject to:} \quad & z_i - \lambda_i \alpha \geq b_i - \lambda_i; \\ & C + \lambda \alpha \leq C_0 + \lambda; \text{ and} \\ & 0 \leq \alpha \leq 1. \end{aligned}$$

## 6.1 Applications

A number of the foregoing articles used decision making, but, since they have already been reviewed, they will not be revisited here. Instead, we focus on three articles that explicitly incorporate linear programming. The topics addressed include optimal asset allocation, insurance pricing, and immunization theory and the matching of assets and liabilities.

Guo and Huang (1996) used a possibilistic linear programming method for optimal asset allocation based on simultaneously maximizing the portfolio return, minimizing the portfolio risk and maximizing the possibility of reaching higher returns. This was analogous to maximizing mean return, minimizing variance and maximizing skewness for a random rate of return.

<sup>19</sup> Adapted from Brockett and Xia (1995), pp 34-38

The authors conceptualize the possibility distribution ( $\pi_{\tilde{r}_i}$ ) of the imprecise rate of return of the  $i$ -th asset of the portfolio as involving  $\tilde{r}_i = (r_i^p, r_i^m, r_i^o)$ , where  $r_i^p$ ,  $r_i^m$ ,  $r_i^o$  are the most pessimistic value, the most possible value, and the most optimistic value for the rate of return, respectively. Then, taking the weighted averages of these values, they defined the imprecise rate of return for the entire portfolio as  $\tilde{r} = (r^p, r^m, r^o)$ , the portfolio risk as  $(r^m - r^p)$ , and the portfolio skewness as  $(r^o - r^m)$ . The authors then showed in a step-by-step fashion how the portfolio could be optimized using Zimmermann's (1978) fuzzy programming method. They concluded that their algorithm provides maximal flexibility for decision makers to effectively balance the portfolio's return and risk.

Carretero and Viejo (2000) investigated the use of fuzzy mathematical programming for insurance pricing decisions with respect to a bonus-malus rating system<sup>20</sup> in automobile insurance. They used the max-min operator and followed Zimmermann's approach (1983, 1985), which led to an optimal solution of the form:

$$\mu_D(x^*) = \max_x \{ \min_x \{ \mu_O(x), \mu_R(x) \} \}, i = 1, \dots, k \quad (9)$$

where  $\mu_D$ ,  $\mu_O$ , and  $\mu_R$  denote the membership function for the fuzzy set "decision D," the fuzzy objective function, and the fuzzy constraints, respectively. Their assumed objective was "attractive income from premiums" while the constraints involved the spread of policies among the risk classes, the weighted sum of the absolute variation of the insured's premium, and the deviation from perfect elasticity of the policyholder's payments with respect to their claim frequency. The system was tested on a large database of third-party personal liability claims of a Spanish insurer and they concluded that their fuzzy linear programming approach avoids unrealistic modeling and may reduce information costs. Carretero (2003) provides further commentary on the approach

Finally, Chang (2003) developed fuzzy mathematical analogues of the classical immunization theory and the matching of assets and liabilities. Essentially, he reformulated concepts about immunization and the matching of assets and liabilities into fuzzy mathematics, and then expressed the objective in terms of a fuzzy linear programming problem. He concluded that his approach offers the advantages of flexibility, improved conformity with situations encountered in practice and the extension of solutions

## 7 Soft Computing

Most of the previously discussed studies focused on FL to the exclusion of other technologies. While their approach has been productive, it may have been sub-optimal, in the sense that stud-

---

<sup>20</sup> A bonus-malus rating system rewards claim-free policyholders by awarding them bonuses or discounts and penalizes policyholders responsible for accidents by assessing them maluses or premium surcharges

ies may have been constrained by the limitations of FL, and opportunities may have been missed to take advantage of potential synergies afforded by other technologies.

This notion was embodied in the concept of soft computing (SC), which was introduced by Zadeh (1992).<sup>21</sup> He envisioned SC as being “concerned with modes of computing in which precision is traded for tractability, robustness and ease of implementation.” For the most part, SC encompasses the technologies of fuzzy logic, genetic algorithms<sup>22</sup> (GAs), and neural networks<sup>23</sup> (NNs), and it has emerged as an effective tool for dealing with control, modeling, and decision problems in complex systems.<sup>24</sup> In this context, FL is used to deal with imprecision and uncertainty, GAs are used for search and optimization, and NNs are used for learning and curve fitting. In spite of these dichotomies, there are natural synergies between these technologies, the technical aspects of which are discussed in Shapiro (2002).

## 7.1 Applications

This section provides a brief overview of a few representative insurance-related articles that have merged FL with either GAs or NNs. The application areas considered are classification and investments. The former involves four representative SC articles in insurance, two on the property-casualty side and two on the life-health side. The latter involves market forecasting.

Our first example of a SC approach is the study of Yoo et al (1994), which proposed it as an auto-insurance claim processing system for Korea. In Korea, given personal and/or property damage in a car accident, the compensation rate depends on comparative negligence, which is assigned using responsibility rates. The authors first describe the expert knowledge structure and the claims processing system. They then explain in general terms how they determined the

---

<sup>21</sup> There are a number of ways that hybrid models could be defined. One approach would be to focus on all adaptive techniques, including such things as chaos theory and fractal analysis. Another approach could be that taken by Yakoubov and Haberman (1998: 75-81), who defined hybrid models as fuzzy techniques in combination with other deterministic and statistical methods. In this article we concentrate on the SC technologies.

<sup>22</sup> GAs are a methodology to perform a randomized global search in a solution space. In this space, a population of candidate solutions, each with an associated fitness value, are evaluated by a fitness function on the basis of their performance. Then, using genetic operations, the best candidates are used to evolve a new population that not only has more of the good solutions but better solutions as well. A working knowledge of GAs can be obtained by reading Shapiro *et. al.* (1999) and Wendt (1995) for genetic algorithms.

<sup>23</sup> NN are computational structures with learning and generalization capabilities. Conceptually, they employ a distributive technique to store knowledge acquired by learning with known samples. Operationally, they use a training set of samples of input-output relationships and a learning algorithm to formulate a supervised learning algorithm that performs local optimization. A working knowledge of NNs can be obtained by reading Francis (2001) and Brockett *et. al.* (1998).

<sup>24</sup> While FL, NNs, and GAs are only a subset of the soft computing technologies, they are regarded as the three principal components. [Shukla (2000): 406]

responsibility rate, and hence the compensation rate, using a fuzzy database, a rule-based system, and a feed-forward NN learning mechanism, and the problems associated with implementing their system.

Cox (1995) reported on a SC-based fraud and abuse detection system for managed healthcare. The essence of the system was that it detected “anomalous” behavior by comparing an individual medical provider to a peer group. The preparation of the system involved three steps: identify the proper peer population, identify behavior patterns, and analyze behavior pattern properties. The peer population was envisioned as a three-dimensional space composed of organization type, geographic region, and organization size.

The behavior patterns were developed using the experience of a fraud-detection department, an unsupervised NN that learnt the relationships inherent in the claim data, and a supervised approach that automatically generate a fuzzy model from a knowledge of the decision variables. Finally, the behavior pattern properties were analyzed using the statistical measures mean, variance, standard deviation, mean absolute deviation, Kolmogorov-Smirnov (KS) test, skewness, and kurtosis.

The discovery properties of the fuzzy model were based on three static and one time varying criteria metrics. The static metrics were the insurer’s exposure to fraudulent behavior, as measured by total claim dollars, the degree of variance from the center of the peer population for each behavior pattern, which was referred to as the population compatibility number, and the number of behaviors that are significantly at variance. The time varying metric was the change in the behavior population dynamics over time. Given the prepared system and the discovery properties, the distribution of data points for the behavior patterns of any individual provider within this population could be computed and compared with all the providers of a similar type, a similar organization size, and within the same geographic area. Thus, the fuzzy system-based fraud and abuse detection system identifies a provider that has significant variance from the peer population.

Cox concluded that the system was capable of detecting anomalous behaviors equal to or better than the best fraud-detection departments.

Peña-Reyes and Sipper (1999) used GA-constructed FISs to automatically produce diagnostic systems for breast cancer diagnosis. The Pittsburgh-style<sup>25</sup> of GAs was used to generate the database and rulebase for the FISs, based on a data furnished by specialists, which contained 444 benign cases and 239 malignant cases, which had been evaluated based on 9 features. They claimed to have obtained the best classification performance to date for breast cancer diagnosis and, because their final systems involve just a few simple rules, high human-interpretability.

Bentley (2000) used an evolutionary-fuzzy approach to investigate suspicious home insurance claims, where genetic programming was employed to evolve FL rules that classified claims into “suspicious” and “non-suspicious” classes. Notable features of his methodology were that it

---

<sup>25</sup> Every individual in the GA is encoded as a string with variable length



used clustering to develop membership functions and committee decisions to identify the best-evolved rules. With respect to the former, the features of the claims were clustered into low, medium, and high groups, and the minimum and maximum value in each cluster was used to define the domains of the membership functions. The committee decisions were based on different versions of the system that were run in parallel on the same data set and weighted for intelligibility, which was defined as inversely proportional to the number of rules, and accuracy. Bentley reported that the results of his model when applied to actual data agreed with the results of previous analysis

By far, the greatest number of SC articles involving fuzzy systems in insurance-related areas is associated with investment models. While most articles are directed at other financial intermediaries, a good deal of the SC research on investment models has implications in the insurance area. The rest of this section provides a brief review of two of these recent studies in order to give a flavor for the types of analysis that have been done<sup>26</sup>. The topic addressed is market forecasting.

Abraham et al. (2001) investigated hybridized SC techniques for automated stock market forecasting and trend analysis. They used principal component analysis to preprocess the input data, a NN for one-day-ahead stock forecasting and a neuro-fuzzy system for analyzing the trend of the predicted stock values. To demonstrate the proposed technique, they analyzed 24 months of stock data for the Nasdaq-100 main index as well as six of the companies listed therein. They concluded that the forecasting and trend prediction results using the proposed hybrid system were promising and warranted further research and analysis.

Finally, Kuo et al. (2001) developed a GA-based fuzzy NN (GFNN) to formulate the knowledge base of fuzzy inference rules, which can measure the qualitative effect (such as the political effect) in the stock market. The effect was further integrated with the technical indexes through the NN. Using buying-selling points and buying-selling performance based on the Taiwan stock market to assess the proposed intelligent system, they conclude that a NN based on both quantitative (technical indexes) and qualitative factors is superior to one based only on quantitative factors.

## 8 Conclusions

The purpose of this article has been to provide an overview of FL applications in insurance. As we have seen, many of the FL techniques have been applied in the insurance area, including fuzzy set theory, fuzzy arithmetic, fuzzy inference systems, fuzzy clustering and fuzzy programming. By the same token, FL has been applied in many insurance areas including classification, underwriting, projected liabilities, fuzzy future and present values, pricing, asset allocation, cash flows, and investments.

The overviews verify that FL has been successfully implemented in insurance. Given this success, and the fact that there are many more insurance problems that could be resolved using

---

<sup>26</sup> See Shapiro (2003) for a review of capital market applications of SC.

fuzzy systems, we are likely to see a number of new applications emerge. There are at least two reasons for this. One reason is that the industry should now have a greater appreciation of potential areas of application, to wit, areas that are characterized by qualitative conditions for which a mathematical model is needed that reflects those conditions. The second reason is that, while fuzzy systems have made inroads into many facets of the business, in most instances the applications did not capitalize on the synergies between the SC technologies and, as a consequence, there are opportunities to extend the studies. These things considered, FL applications in insurance and related areas should be a fruitful area for exploration for the foreseeable future.

## References

- Abraham, A., B. Nath and P. K. Mahanti. (2001) "Hybrid Intelligent Systems for Stock Market Analysis," *Computational Science*, Springer-Verlag Germany, Vassil N. Alexandrov et. al. (Eds.), ISBN 3-540-42233-1, San Francisco, USA, 337-345.
- Babad, Y. M. and B. Berliner. (1995) "Reduction of Uncertainty through Actualization of Intervals of Possibilities and its Applications to Reinsurance," *25th TICA 2*: 23 - 35.
- Bellman, R. and L. A. Zadeh. (1970) "Decision-Making in a Fuzzy Environment," *Management Science*, Vol. 17, 141-164.
- Bentley, P. J. (2000) "Evolutionary, my dear Watson. Investigating Committee-based Evolution of Fuzzy Rules for the Detection of Suspicious Insurance Claims," *Proceedings of the second Genetic and Evolutionary Computation Conference (GECCO)*, 8-12, July.
- Berkan, R. C. and S. L. Trubatch. (1997) *Fuzzy Systems Design Principles : Building Fuzzy If-Then Rule Bases*, Wiley-IEEE Press.
- Bezdek, J.C. (1981) "Pattern Recognition with Fuzzy Objective Function Algorithms," New York, Plenum Press.
- Boissonnade, A. C. (1984) "Earthquake Damage and Insurance Risk," Ph.D. Dissertation, Stanford University.
- Bouet, G. and R. Dalaud. (1996) "Extension de Concepts Financiers en Calcul Flou et Application à l'Adossement de Flux," *6th AFIR*, Vol. 2, 827-843.
- Brockett, P. L., X. Xia. (1995) "Operations Research In Insurance: A Review," *Transactions Of Society Of Actuaries*, Vol. 47, pp. 7-87.
- Brockett, P. L., X. Xia, and R.A. Derrig. (1998) "Using Kohonen's Self-Organizing Feature Map to Uncover Automobile Bodily Injury Claims Fraud," *J Risk and Insurance* 65(2), 245-274.
- Buckley, J. J. (1987) "The Fuzzy Mathematics of Finance," *Fuzzy Sets and Systems*, 21: 257-

- Buehlmann, N. and B. Berliner. (1992) "Fuzzy Zooming of Cash Flows," 24th IICA, 6:437-453.
- Calzi, M.L. (1990) "Towards a General Setting for the Fuzzy Mathematics of Finance," *Fuzzy Sets and Systems*, 35:265-280.
- Carreno, L.A., Y. Jani, (1993) "A Fuzzy Expert System Approach to Insurance Risk Assessment Using FuzzyCLIPS," WESCON Conference Record, 536-541. 28-30 September.
- Carretero, R.C. (2000) "New Approaches for Designing a Bonus-Malus System," 2000 IME Congress., Barcelona, Spain, July.
- Carretero, R.C. (2003) "Fuzzy Logic Techniques in the Non-Life Insurance Industry," in A. F. Shapiro and L. C. Jain (Eds) *Intelligent and Other Computational Techniques in Insurance Theory and Applications*, World Scientific Publishing Company, pp. 229-257
- Carretero, R.C. and A.S. Viejo. (2000) "A Bonus-Malus System in the Fuzzy Set Theory [insurance pricing decisions]," IEEE International Conference on Fuzzy Systems 2, 1033-1036 7-10 May.
- Chang, C. (2003) "A Fuzzy Set Theoretical approach to Asset and Liability Management and Decision Making," in A. F. Shapiro and L. C. Jain (Eds) *Intelligent and Other Computational Techniques in Insurance Theory and Applications*, World Scientific Publishing Company, pp. 301-333.
- Chang, C. and P. Wang. (1995) "The Matching of Assets and Liabilities with Fuzzy Mathematics," *25th International Congress of Actuaries*, 123.
- Chen, J.J.G, and Z He. (1997) "Using Analytic Hierarchy Process and Fuzzy Set Theory to Rate and Rank the Disability," *Fuzzy Sets and Systems*, 88:1-22.
- Chorafas, D. N. (1994) *Chaos Theory In The Financial Markets. Applying Fractals, Fuzzy Logic, Genetic Algorithms* Chicago: Probus Publishing Company.
- Cox, E. (1995) "A Fuzzy System for Detecting Anomalous Behaviors in Healthcare Provider Claims," *Intelligent Systems for Finance and Business*, edited by S. Goonatilake and P. Ireleven, John Wiley & Sons, 111-135.
- Cummins, J. D. and R. A. Derrig. (1993) "Fuzzy Trends in Property-Liability Insurance Claim Costs," *Journal of Risk and Insurance*, Vol. 60, 429-465.
- Cummins, J. D. and R. A. Derrig. (1997) "Fuzzy Financial Pricing of Property-Liability Insurance," *North American Actuarial Journal*, Vol 1, No. 4, 21

- Derrig, R. A. and K.M. Ostaszewski. (1994) "Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification," 4th AFIR International Colloquium, 1:141 - 171.
- Derrig, R. A. and K. Ostaszewski. (1995a) "Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification," *Journal of Risk and Insurance*, Vol. 62, pp. 447-482
- Derrig, R. A. and K.M. Ostaszewski. (1995b) "The Fuzzy Problem of Hedging the Tax Liability of a Property - Liability Insurance Company," 5th AFIR International Colloquium, 1:17 - 42.
- Derrig, R. A. and K. Ostaszewski (1997) "Managing the Tax Liability of a Property-Liability Insurance Company," *Journal of Risk and Insurance*, Vol. 64, 694.
- Derrig, R. A. and K. Ostaszewski. (1999) "Fuzzy Sets Methodologies in Actuarial Science," Chapter 16 in H. J. Zimmerman, *Practical applications of Fuzzy Technologies*, Boston: Kluwer Academic Publishers.
- DeWit, G. W. (1982) "Underwriting and Uncertainty," *Insurance: Mathematics and Economics*, Vol. 1, 277-285
- Dubois, D. and H. Prade (1980) *Fuzzy Sets and Systems: Theory and Applications*. San Diego, CA: Academic Press.
- DuBois, D. and H. Prade. (1997) *Fuzzy Sets and Systems: Theory and Applications*, San Diego, CA: Academic Press
- Ebanks, B., W. Kanvowski, and K. M. Ostaszewski. (1992) "Application of Measures of Fuzziness to Risk Classification in Insurance," *Computing and Information*, IEEE Computer Society Press, Los Alamitos, California, 290-291.
- Erbach, D. W. and E. Seah. (1993) Discussion of "The Application of Fuzzy Sets to Group Health Underwriting" by V. R. Young, *Transactions of the Society of Actuaries*, 45, 585.
- Francis, L. (2001) "Neural Networks Demystified," *Casualty Actuarial Society Forum*, Winter, 253-319.
- Frees, E. W. and E. A. Valdez. (1998) "Understanding Relationships Using Copulas," *NAAJ*, 2 (1), 1-25
- Guo, L. and Z. Huang. (1996) "A Possibilistic Linear Programming Method for Asset Allocation," *Journal of Actuarial Practice*, 2:67-90.
- Hellman, A. (1995) "A Fuzzy Expert System for Evaluation of Municipalities - an Application," 25h TICA, 1:159 - 187.
- Horgby, P., R. Lohse and N. Sittaro. (1997) "Fuzzy Underwriting: An Application of Fuzzy Logic to Medical Underwriting," *Journal of Actuarial Practice*, Vol. 5, No. 1, 79.

- Hosler, V. R. (1992) "The Application of Fuzzy Sets to Group Health Underwriting," *ARCH* 2, 1-63
- Jang, J-S. R., C-I Sun and E, Mizutani. (1997) *Neuro-fuzzy and soft computing. a computational approach to learning and machine intelligence*, Prentice Hall, Upper Saddle River, NJ.
- Jablonowski, M. (1996) "A New Perspective on Risk," *Chartered Property and Casualty Underwriters Journal* 49 (4), 225-236
- Jablonowski, M. (1997) "Modeling Imperfect Knowledge in Risk Management and Insurance," *Risk Management and Insurance Review*, Vol. 1, No. 1, 98
- Joseph, E. C. (1982) "Future Expert Computer Systems For Actuarial Professionals," *ARCH* 1982 1, pp. 1-9
- Kieselbach, R. 1997. "Systematic Failure Analysis Using Fault Tree and Fuzzy Logic," *Technology, Law and Insurance*, No. 2, 13-20.
- Klir, G J., B. Yuan. (1996) "Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh," World Scientific, New Jersey
- Krasteva, E. B., M. G. Singh, G R. Sotirov, J. C. Bennavil, N. C. Mincoff (1994) "Model Building for Pricing Decision Making in an Uncertain Environment," *Systems, Man, and Cybernetics. Humans, Information and Technology*, IEEE International Conference on, Volume: 1, 194 -199. (pdf new)
- Kuo, R. J., C. H. Chen and Y. C. Hwang. (2001) "An Intelligent Stock Trading Decision Support System Through Integration of Genetic Algorithm Based Fuzzy Neural Network and Artificial Neural Network," *Fuzzy Sets and Systems*: 118 (1), 21-45.
- Lemaire, J. (1990) "Fuzzy Insurance," *ASTIN Bulletin*, Vol. 20, No. 1, 33-55.
- Lu, Y, L. Zhang, Z. Guan and H. Shang. (2001) "Fuzzy Mathematics Applied to Insurance Game," *IFSA World Congress and 20th NAFIPS International Conference*, Vol. 2: 941
- Manton, K. G., M. A. Woodbury and H. D. Tolley. (1994) *Statistical Applications Using Fuzzy Sets*, Joh. Wiley & sons, Inc.
- McCauley-Bell, P. R. and A. Badiru. (1996) "Fuzzy Modelling and Analytic Hierarchy Processing to Quantify Risk Levels Associated with Occupational Injuries," Part I and II, *IEEE Transactions on Fuzzy Systems*, 4:124-138.
- McCauley-Bell, P. R., L. L. Crumpton, and H. Wang. (1999) "Measurement of Cumulative Trauma Disorder Risk in Clerical Tasks Using Fuzzy Linear Regression," *IEEE Transactions*

- On Systems, Man, And Cybernetics-Part C: Applications And Reviews, VOL 29, NO 1, pp. 1-14.
- Mosmans, A., J.-C. Praet and C. Dumont. (2002) "A Decision Support System for the Budgeting of the Belgian Health Care System," *European Journal of Operational Research*, Vol. 139, Issue 2, pp. 449-460
- Musilek, P. and M. M. Gupta (2000) "Neural Networks and Fuzzy Systems," in Sinha, N. K. and M. M. Gupta (Eds), *Soft Computing and Intelligent Systems* Academic Press
- Ostaszewski, K. (1993) *Fuzzy Set Methods in Actuarial Science* Schaumburg, IL: Society of Actuaries.
- Ostaszewski, K. and W. Karwowski. (1992) "An Analysis of Possible Applications of Fuzzy Set Theory to the Actuarial Credibility Theory," *Proceeding of the Annual Meeting of the North American Fuzzy Information Processing Society*, Puerto Vallaria, December.
- Peña-Reyes, C. A. and M. Sipper (1999) "A Fuzzy-Genetic Approach to Breast Cancer Diagnosis," *Artificial Intelligence in Medicine* 17, 131–155.
- Romahi, Y. and Q. Shen. (2000) "Dynamic Financial Forecasting with Automatically Induced Fuzzy Associations." *IEEE*, 493-498.
- Saaty, T. (1980) *The Analytic Hierarchy Process* New York: McGraw-Hill
- Siegel, P. H., A. de Korvin and K. Omer, eds. (1995) *Applications of Fuzzy Sets and the Theory of Evidence to Accounting* Greenwich, Conn : JAI Press.
- Shapiro, A. F. (2000) "A Hitchhiker's Guide to the Techniques of Adaptive Nonlinear Models," *Insurance: Mathematics and Economics* 26, 119-132.
- Shapiro, A. F. (2002) "The Merging of Neural Networks, Fuzzy Logic, and Genetic Algorithms," *Insurance: Mathematics and Economics* 31, 115–131.
- Shapiro, A. F. (2003) "Capital Market Applications of Neural Networks, Fuzzy Logic and Genetic Algorithms," *Proceedings of the 13th International AFIR Colloquium*, Vol 1, 493-514
- Shapiro, A. F., J. S. Pflumm and T. A. DeFilippo. (1999) "The Inner Workings of Neural Networks and Genetic Algorithms," *ARCH*, 1999 Vol 1, 415-426.
- Shukla, K. K. (2000) "Soft Computing Paradigms for Artificial Vision," In: Sinha, N. K., Gupta, M.M. (Eds.), *Soft Computing and Intelligent Systems: Theory and Applications*. Academic Press, San Diego, 405–417.
- Simonelli, M. R. (2001) "Fuzziness in Valuing Financial Instruments by Certainty Equivalents," *European Journal of Operational Research* 135, 296-302.

- Sinha, N. K. and M. M. Gupta (Eds). (2000) "Soft Computing and Intelligent Systems : Theory and Applications," San Diego, Cal.: Academic Press.
- Smith, E., J. Eloff. (2000) "Cognitive Fuzzy Modeling for Enhanced Risk Assessment in a Health Care Institution," IEEE Intelligent Systems, Volume: 15 Issue: 2 , March-April, 69 - 75. (pdf new)
- Terceno, A., J. De Andres, C. Belvis and G. Barbera. (1996) "Fuzzy Methods Incorporated to the Study of Personal Insurances," Neuro-Fuzzy Systems. AT'96., International Symposium on, Page(s): 187 -202 (pdf new)
- Verrall, R. J. and Y. H. Yakoubov. (1999) "A Fuzzy Approach to Grouping by Policyholder Age in General Insurance," Journal of Actuarial Practice, Vol 7, pp. 181-203.
- Wendt, R. Q. (1995) "Build Your own GA Efficient Frontier," *Risks and Rewards*, December.
- Yager, R. R., S. Ovchinnikov, R. M. Tong and H. T. Ngugen. 1987. *Fuzzy Sets and Applications: Collected Papers of Lotfi A. Zadeh*, New York: John Wiley & Sons, Inc.
- Yakoubov, Y. H. and S. Haberman (1998) "Review of Actuarial Applications of Fuzzy Set Theory," Actuarial Research Paper No. 105, Department of Actuarial Science and Statistics, City University, London.
- Yager, R. R., S. Ovchinnikov, R. M. Tong, H. T. Ngugen (1987) "Fuzzy Sets and Applications: Collected Papers of Lotfi A. Zadeh," Wiley, New York.
- Yeo, A. C., K. A. Smith, R. J. Willis, and M. Brooks. (2001) "Clustering Technique for Risk Classification and Prediction of Claim Cost in the Automobile Insurance Industry", International Journal of Intelligent Systems in Accounting, Finance and Management, vol. 10, 39-50.
- Yoo, J-H, B-H Kang and J-U Choi (1994) "A Hybrid Approach to Auto-insurance Claim Processing System." IEEE International Conference on Systems, Man, and Cybernetics 1, 537-542 2-5 October.
- Young, V. R. (1993) "The Application of Fuzzy Sets to Group Health Underwriting," *Transactions of the Society of Actuaries*, Vol. 45, 551-590.
- Young, V. R. (1996) "Insurance Rate Changing: a Fuzzy Logic Approach," Journal of Risk and Insurance 63, 461-83.
- Young, V. R. (1997) "Adjusting Indicated Insurance Rates: Fuzzy Rules that Consider Both Experience and Auxiliary Data," *Proceedings of the Casualty Actuarial Society*, 84, 734.
- Zadeh, L. A. (1965) "Fuzzy Sets," *Information and Control*, Vol. 8, 338-353.

- Zadeh, L. A. (1975, 1976) "The Concept of Linguistic Variable and its Application to Approximate Reasoning (Parts 1-3)," *Information Sciences*, Vol. 8, 199-249, 301-357, and Vol 9, 43-80.
- Zadeh, L. A. 1981. "Fuzzy systems theory: A framework for the analysis of humanistic systems." In: Cavallo, R. E., ed., *Recent Developments in Systems Methodology in Social Science Research*, Boston: Kluwer, pp 25-41.
- Zadeh, L. A. (1992) "Foreword of the Proceedings of the Second International Conference on Fuzzy Logic and Neural Networks," xiii-xiv, Iizuka, Japan.
- Zimmermann, H. J. (1983) "Fuzzy mathematical programming," *Computers and Operations Research*, pp 281-298
- Zimmermann, H. J. (1985) "Applications of fuzzy set theory to mathematical programming," *Information Sciences* 36, pp 29-58.
- Zimmermann, H. J. (1987) *Fuzzy Sets, Decision Making and Expert Systems*, Kluwer Academic Publishers, Boston, Massachusetts.
- Zimmermann, H. J. (1996) *Fuzzy Set Theory and its Applications, Third Edition*, Kluwer Academic Publishers, Boston, Massachusetts.