Calculating Insurance Claim Reserves with Hybrid Fuzzy Least Squares Regression Analysis

Furkan BAŞER[1][♠], Ayşen APAYDIN [2]

[1]_Gazi University, Faculty of Commerce and Tourism Education, Department of Computer Applications Education, Ankara, Turkey_
[2]_Ankara University, Faculty of Science, Department of Statistics, Ankara, Turkey_

ABSTRACT

The prediction of an adequate amount of claim reserves is of the greatest importance to face the responsibilities assumed by an insurance company. Although many different deterministic and stochastic methods based on statistical analyses are used for claims analysis, presence of many internal and external factors that increase the uncertainty in insurance environment may lead to considerable loss in reliability of statistical methods. Therefore, in a state of uncertainty that exist in the nature of many actuarial and financial problems, when convenient and reliable data is not held, the use of fuzzy set theory becomes very attractive to get more actual results.

In this paper, a method for calculating insurance claim reserves using hybrid fuzzy least-squares regression analysis is proposed. The results from classical method and this soft computing approach are compared by using original data in automobile liability insurance.

Key Words: _Insurance, Claims reserving, Fuzzy numbers, Fuzzy arithmetic, Fuzzy regression._

1. INTRODUCTION

At the end of the accounting period of an insurance company, some claims have been observed in coverage of policies but, insurance company does not have any information about these claims' presence and cost. Some claims can be reported several years later from the occurrence and this can cause important solvency problems [15]. So, insurance companies must pay great attention to calculation of the claim provisions. Measure of error in provisions will increase the risk of bankruptcy. Although many different statistical methods have been developed until now to estimate outstanding claims, there is no general agreement as to which is the best approach [10].

Among the methods of claims reserving, there exist several different classifications. The highest level of classification is the division method into two main categories, stochastic and non-stochastic methods.

Sometimes it seems useless to apply a stochastic model because there is only inadequate and ambiguous information observed and as a result, the point estimates can not be reliable. So, one will be able to construct intervals for the quantities to be evaluated [9].

Several methods have been outlined in the literature, of which the Chain Ladder (CL) is probably the most widely used. This is mainly due to its practical basis. However, CL method has some problems, that is a purely multiplicative method and the estimate for each origin period is formed only by the most recent value with a development factor [4, 17]. Benjamin and Eagles [3] proposed a slight generalization of the CL method, known as London Chain Ladder (LCL), which is based on the use of ordinary least-squares regression over the accumulated claims.

Even if various methods based on statistical analyses are used to set claim provisions, presence of real world

---

[♠]Corresponding author, e-mail: fbaser@gazi.edu.tr

factors which increase the uncertainty in calculating stage may lead to considerable loss in reliability of statistical methods. Unlike traditional methods in claims analysis, soft computing can tolerate imprecision, uncertainty and partial truth without loss of performance and effectiveness for the end use. The guiding principle of soft computing is: exploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost [19]. It has emerged as an effective tool for dealing with control, modeling and decision problems in complex systems. Thus, as seen in an article that is an overview of fuzzy logic applications in insurance by Shapiro [13]; fuzzy logic, which is the leading constituent of soft computing, has been applied in many insurance areas including risk classification, underwriting, projected liabilities, etc.

Initially, Andrés and Terceño [2] used fuzzy set theory and fuzzy regression method for calculating incurred but not reported reserves. It was a combined approach for claims reserving by integrating fuzzy regression into LCL method. However, the main shortcoming of this fuzzy regression model developed in Tanaka [16] is that the concept of least-squares is not utilized [6].

Randomness and fuzziness are two important sources of uncertainty that exist in actuarial analysis. Randomness models the stochastic variability of all possible outcomes of a situation and describes the inherent variation associated with the environment. On the other hand, fuzziness relates to the unsharp boundaries of the parameters of the model and is more an instrument of a descriptive analysis reflecting the past and its implications [14].

In our present paper, hybrid fuzzy least-squares regression analysis, which is proposed by Chang [7], is applied to predict future claim costs by using the concept of LCL method. Thus, the purpose of the paper is to take advantage of using the hybrid fuzzy regression model by considering both randomness and fuzziness type of uncertainty.

The structure of the paper is as follows. In the next section, hybrid fuzzy least-squares regression analysis is described. In Section 3, after an outline of the insurance claim reserves is given, an approach of hybrid fuzzy least-squares regression to claims reserving is presented. In Section 4, the claims data for a group of ten automobile liability insurance companies is used to demonstrate the proposed method. Finally, based on the results of the numerical application, the most important conclusions of the paper are summarized.

2. HYBRID FUZZY LEAST-SQUARES LINEAR REGRESSION

Regression analysis is a statistical tool to model functional relationship between dependent and independent variables. In ordinary regression analysis, the deviations between observed and estimated values of dependent variable are generally assumed to have a normal distribution, constant variance, and a zero mean [6]. In this sense, the violation of its basic assumption could adversely affect the validity of the regression model. Fuzzy regression, a non parametric method, can be quite useful in estimating the relationships among variables where the available data are very limited and imprecise, and variables are interacting in an uncertain, qualitative, and fuzzy way [12].

Randomness and fuzziness represent two different types of uncertainty. Firstly, consider the points of the sample space as the set of all possible values, which a random variable may take. If the purpose is to predict the probability of event, which can occur, this event is represented in the sample space in a well defined region of that space. For more clear description of situation it must shown that well-defined region has the crisp boundary, because it use a concept of characteristic function. If the well defined region is given by characteristic function, then the sample space is divided into two subspaces: if the point belongs to region, the characteristic function value is one and the event is true, otherwise this point falls outside the region, the value of characteristic function is zero and event is false. That way, the probability value shows the frequency at which the point will occur inside the region.

Case of fuzziness is essentially different from randomness. Here the region is represented by a fuzzy subset. Thus, fuzzy uncertainty is represented by partial membership of a point from the universe of discourse in an imprecisely defined region of space. This membership function gets its values in interval [0, 1], while characteristic function of well-defined region takes values in set {0, 1}, that is, only two values: 0 or 1. Thus, membership function will not divide the universe of discourse into two subspaces. Membership function describes the degree to which the element of universe space corresponds to the property with which fuzzy set is defined [1].

A major difference between fuzzy regression and ordinary regression is in dealing with errors as fuzzy variables in fuzzy regression modeling, and in dealing with errors as random residuals in ordinary regression modeling. To integrate both fuzziness and randomness into a regression model, a concept of hybrid regression analysis is proposed by Chang [7]. This method allows fitting a model to fuzzy data, crisp data, and their mixture by using weighted fuzzy arithmetic and least squares fitting criterion for the purpose of improving the estimations of regression model.

A bivariate regression model for fuzzy data defined by symmetric triangular membership function can be expressed as:

$$\widetilde{Y}_j = \widetilde{a} + \widetilde{b} X_j = (a_C, a_S) + (b_C, b_S) X_j \qquad \ldots\ldots(1)$$

where a_C and b_C are the centers, a_S and b_S are the spreads. Then, each observed value of dependent variable can be expressed as the following symmetric triangular fuzzy number (STFN):

$$\widetilde{Y}_j = (Y_{jC}, Y_{jS}) = (a_C + b_C X_j, a_S + b_S X_j), \qquad j = 1,2,\ldots,n \ \ldots\ldots(2)$$

in which n is the sample size. $\mu_{\widetilde{Y}_{j,L}}$ and $\mu_{\widetilde{Y}_{j,R}}$ are the left bound and the right bound of the observed $\widetilde{Y}_j$ at membership μ level. Similarly, each predicted value $\hat{\widetilde{Y}}_j$ can be expressed as:

$$\hat{\widetilde{Y}}_j = (\hat{Y}_{jC}, \hat{Y}_{jS}) = (\hat{a}_C + \hat{b}_C X_j, \hat{a}_S + \hat{b}_S X_j), \quad j=1,2,...,n \ ....(3)$$

where $\hat{a}_C$, $\hat{b}_C$ are the centers and $\hat{a}_S$, $\hat{b}_S$ are the spreads of estimated regression coefficients, $\hat{\widetilde{a}}$ and $\hat{\widetilde{b}}$. $\mu_{\hat{\widetilde{Y}}_{j,L}}$ and $\mu_{\hat{\widetilde{Y}}_{j,R}}$ are the left bound and right bound of predicted $\hat{\widetilde{Y}}_j$ at membership μ level. These quantities are given by:

$$\mu_{\widetilde{Y}_{j,L}} = Y_{jC} - (1-\mu) Y_{jS} \qquad ....(4a)$$

$$\mu_{\widetilde{Y}_{j,R}} = Y_{jC} + (1-\mu) Y_{jS} \qquad ....(4b)$$

$$\mu_{\hat{\widetilde{Y}}_{j,L}} = [\hat{a}_C - (1-\mu)\hat{a}_S] + [\hat{b}_C - (1-\mu)\hat{b}_S] X_j$$

$$= (\hat{a}_C + \hat{b}_C X_j) - (1-\mu)(\hat{a}_S + \hat{b}_S X_j) \qquad ....(5a)$$

$$\mu_{\hat{\widetilde{Y}}_{j,R}} = [\hat{a}_C + (1-\mu)\hat{a}_S] + [\hat{b}_C + (1-\mu)\hat{b}_S] X_j$$

$$= (\hat{a}_C + \hat{b}_C X_j) + (1-\mu)(\hat{a}_S + \hat{b}_S X_j) \qquad ....(5b)$$

Using the definition of weighted fuzzy arithmetic, the sum of the residual errors between the predicted $\hat{\widetilde{Y}}_j$ and the observed $\widetilde{Y}_j$ is formulated as:

$$\sum (residual\ errors)^2 = \sum_{j=1}^{n} (\widetilde{Y}_j - \hat{\widetilde{Y}}_j)^2$$

$$= \sum_{j=1}^{n} \left[\int_0^1 (\mu_{\widetilde{Y}_{j,L}} - \mu_{\hat{\widetilde{Y}}_{j,L}})^2 \mu \, d\mu + \int_0^1 (\mu_{\widetilde{Y}_{j,R}} - \mu_{\hat{\widetilde{Y}}_{j,R}})^2 \mu \, d\mu \right]. (6)$$

In order to derive formula for the unknown regression coefficients, the objective function for the principle of least squares is to minimize the sum of the squares of residual errors. Consequently, the unknown regression coefficients (a_C, b_C, a_S, b_S) are obtained by following two sets of 2*2 simultaneous equations.

$$n\hat{a}_C + \hat{b}_C \sum X_j = \sum Y_{jC}$$
$$\hat{a}_C \sum X_j + \hat{b}_C \sum X_j^2 = \sum X_j Y_{jC}$$

can solve for $\hat{a}_C$ and $\hat{b}_C$, $\qquad ....(7a)$

$$n\hat{a}_S + \hat{b}_S \sum X_j = \sum Y_{jS}$$
$$\hat{a}_S \sum X_j + \hat{b}_S \sum X_j^2 = \sum X_j Y_{jS}$$

can solve for $\hat{a}_S$ and $\hat{b}_S$ $\qquad ....(7b)$

If asymmetric triangular fuzzy numbers were used, the calculations could be extended to three sets of 2*2 simultaneous equations for centers, right and left spreads (see, Chang [7]).

3. INSURANCE CLAIM RESERVES

Outstanding claim reserves in general insurance are a type of technical reserve or accounting provision in the financial statements of an insurer. They seek to quantify the outstanding liability for insurance claims which have been reported and not yet settled. They may or may not include IBNR (**I**ncurred **B**ut **N**ot **R**eported) reserves. IBNR is a technical reserve of an insurance company, and is established to provide for the future liability for claims which have occurred but which have not yet been reported to the insurance company (for more detailed, see Hossack et al. [10]). Although there are many variations on the way the premiums are paid, the company generally receives the premiums before the claims are actually paid. Therefore, the company must properly account for these claim liabilities by setting up provisions within its balance sheet to reflect as accurately as possible the eventual claims cost [4].

In actuarial literature, many different deterministic and stochastic methods based on statistical analyses are used for estimating outstanding claims (Chain Ladder, London Chain Ladder, London Pivot, Cape-Cod, etc.). Special feature of both of these methods is to group data within a table called run off triangle.

3.1 Run-off Triangle

The claims experience of an insurer in respect of a particular class of business can be summarized in a run-off triangle showed by Table 1. The idea of the run-off triangle as well as a first estimation procedure goes back to an article of Verbeek [18].

Claims run-off data are generated when delay is incurred in settling insurance claims. Typically the format for such data is that of a triangle in which the rows (i) denote origin years and the columns (j) denote delay or development years. "Origin year (period)" is the calendar year (or financial year) in which the incident leading to a claim occurred.

In Table 1, $Z_{i,j}$ is the amount of accumulated incurred losses during development period j ($j = 0,1,...,n$) in respect of claims whose year of origin is i ($i = 0,1,...,n$). Obviously, it is not known, for the ith year of occurrence ($i = 1,2,...,n$), the accumulated losses in the development years $j = n$-i+1, n-i+2, ... , n and therefore, these losses must be predicted.

Table 1. Run-off triangle.

Occurrence / origin period (year)	Development period (year)						
	0	**1**	**⋯**	**j**	**⋯**	**n-1**	**n**
0	$Z_{0,0}$	$Z_{0,1}$	⋯	$Z_{0,j}$	⋯	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1,0}$	$Z_{1,1}$	⋯	$Z_{1,j}$	⋯	$Z_{1,n-1}$	
⋮	⋮	⋮		⋮			
i	$Z_{i,0}$	$Z_{i,1}$	⋯	$Z_{i,j}$			
⋮	⋮	⋮					
n-1	$Z_{n-1,0}$	$Z_{n-1,1}$					
n	$Z_{n,0}$						

3.2. Calculating Insurance Claim Reserves with Hybrid Fuzzy Regression

London Chain Ladder (LCL) method which uses ordinary least-squares regression to estimate claim provisions is proposed by Benjamin and Eagles [3]. This method is extended by using hybrid fuzzy least squares regression analysis thus; it will allow using all the information provided by the run-off triangle more efficiently. On the other hand, estimating fuzzy coefficients in fuzzy regression by using the principle of least-squares will allow us to provide theoretical basis, e.g., the estimates are unbiased, the variances are the smallest, etc.

So, the evolution of the accumulated claims of the accidents happened in the year i from the jth to the j+1th developing years can be adjusted using a fuzzy linear relation $\widetilde{Z}_{i,j+1} = \widetilde{a}_j + \widetilde{b}_j Z_{i,j}$. If $\widetilde{a}_j$ and $\widetilde{b}_j$ are regarded as STFN, these parameters are estimated by hybrid fuzzy regression for $\widetilde{a}_j = (a_{jC}, a_{jS})$ and $\widetilde{b}_j = (b_{jC}, b_{jS})$ in which a_{jC} and b_{jC} are the centers and a_{jS} and b_{jS} are the spreads. Consequently, $\widetilde{Z}_{i,j+1}$ can be denoted as:

$$\begin{aligned}\widetilde{Z}_{i,j+1} &= (Z_{(i,j+1)C}, Z_{(i,j+1)S}) = (a_{jC}, a_{jS}) + (b_{jC}, b_{jS})Z_{i,j} \\ &= (a_{jC} + b_{jC}Z_{i,j}, \ a_{jS} + b_{jS}Z_{i,j})\end{aligned} \quad (8)$$

The estimates of $\widetilde{a}_j$ and $\widetilde{b}_j$ are symbolized as $\hat{\widetilde{a}}_j = (\hat{a}_{jC}, \hat{a}_{jS})$ and $\hat{\widetilde{b}}_j = (\hat{b}_{jC}, \hat{b}_{jS})$, respectively. If hybrid fuzzy regression model (8) is considered, it will be seen that each of the observation for dependent variable takes part in model as a fuzzy number. Additionally, the symmetric triangular membership functions for observations belonging to dependent variable can be determined by definition of confidence interval (see, Buckley [5]). To do that, $(1-\gamma)100\%$ confidence interval for a specific value of dependent variable, Z_0, is denoted as follow:

$$\left[\hat{Z}_0 - t_{n-2,\gamma/2} \ \hat{\sigma}_{Z_0-\hat{Z}_0} \ , \ \hat{Z}_0 + t_{n-2,\gamma/2} \ \hat{\sigma}_{Z_0-\hat{Z}_0}\right] \dots(9)$$

In Eq. (9):

$1-\gamma$: Confidence level

$n-2$: Degrees of freedom. n, sample size

$t_{n-2,\gamma/2}$: Value for $t-$ distribution

$\hat{\sigma}_{Z_0-\hat{Z}_0}$: Standard error.

If simultaneous equations, given in Eqs. (7a) – (7b), are rearranged for STFNs, $\hat{\widetilde{a}}_j$ and $\hat{\widetilde{b}}_j$, the normal equations for centers and spreads are determined as follows:

$$n.\hat{a}_{jC} + \left(\sum Z_{i,j}\right).\hat{b}_{jC} = \sum Z_{(i,j+1)C}$$
$$\left(\sum Z_{i,j}\right).\hat{a}_{jC} + \left(\sum Z_{i,j}^2\right).\hat{b}_{jC} = \sum \left(Z_{i,j}.Z_{(i,j+1)C}\right)$$

can solve for $\hat{a}_{jC}$ and $\hat{b}_{jC}$, ...(10a)

$$n.\hat{a}_{jS} + \left(\sum Z_{i,j}\right).\hat{b}_{jS} = \sum Z_{(i,j+1)S}$$
$$\left(\sum Z_{i,j}\right).\hat{a}_{jS} + \left(\sum Z_{i,j}^2\right).\hat{b}_{jS} = \sum \left(Z_{i,j}.Z_{(i,j+1)S}\right)$$

can solve for $\hat{a}_{jS}$ and $\hat{b}_{jS}$. ...(10b)

Then, the prediction of final cost of the accidents produced in year i, $\hat{\widetilde{Z}}_{i,n}$, is obtained by using fuzzy arithmetic as:

$$\hat{\widetilde{Z}}_{i,n} = \hat{\widetilde{a}}_{n-1} + \hat{\widetilde{b}}_{n-1}\left\{\dots\hat{\widetilde{a}}_{i+2} + \hat{\widetilde{b}}_{i+2}\left[\hat{\widetilde{a}}_{i+1} + \hat{\widetilde{b}}_{i+1}\left(\hat{\widetilde{a}}_i + \hat{\widetilde{b}}_i Z_{i,i}\right)\right]\right\} \quad (11)$$

It is obvious that $\hat{\widetilde{Z}}_{i,n}$ is not STFN but, the results of this nonlinear operations can be converged to a STFN (see, Dubois and Prade [8], Kaufmann and Gupta [11]).

For instance, multiplication of two STFNs, for $\tilde{k} = \tilde{p} \cdot \tilde{q}$, $\tilde{p} = (p_C, p_S)$ and $\tilde{q} = (q_C, q_S)$, is obtained approximately as:

$$\tilde{k} \approx (k_C, k_S) = (p_C \, q_C, \, p_C \, q_S + q_C \, p_S) \qquad ....(12)$$

So, $\hat{\tilde{Z}}_{i,n}$ can be determined as $\hat{\tilde{Z}}_{(i,n)} \approx (\hat{Z}_{(i,n)C}, \hat{Z}_{(i,n)S})$. Therefore, estimated reserve for the ith year of occurrence as fuzzy number $\hat{\tilde{R}}_i = (\hat{R}_{iC}, \hat{R}_{iS})$ is:

$$\hat{\tilde{R}}_i = \hat{\tilde{Z}}_{i,n} - Z_{i,n-i} = (\hat{Z}_{(i,n)C} - Z_{i,n-i}, \, \hat{Z}_{(i,n)S}) \, ...(13)$$

Finally, the whole claims reserve, $\hat{\tilde{R}} = (\hat{R}_C, \hat{R}_S)$, which is the STFN is obtained as:

$$\hat{\tilde{R}} = \sum_{i=1}^{n} \hat{\tilde{R}}_i = (\sum_{i=1}^{n} \hat{R}_{iC}, \sum_{i=1}^{n} \hat{R}_{iS}) \qquad ....(14)$$

4. APPLICATION

In this section, run-off triangle, showed by Table 2, gives the claims of a group of ten Belgian insurance companies. The data which is similar to the one in Goovaerts et al. [9] is used to illustrate the proposed method for calculating insurance claims reserves.

In this cumulative run-off triangle, in respect of claims originating in the first year (origin period, 0), payment totaling 2062 were made in the same year (development period 0), and payment totaling 1629 were made in the following year (development period 1). We need to estimate claim payments after 6th origin period (year) from the past data, in order to deduce the outstanding claims provision required in 31 December of 6th origin year.

After determining hybrid fuzzy linear regression model for each development period ($j = 0,1,...,6$), the observations for dependent variable in each model are fuzzified by using the confidence limits for $\gamma = 0.05$.

The amount of future claims is estimated with hybrid fuzzy regression models by using the amount of accumulated incurred losses which can be settled in previous years. For example, the evolution of the accumulated claims of the accidents happened in the year i from the 0 to the 1st developing years can be adjusted using a fuzzy linear relation $\tilde{Z}_{i,1} = \tilde{a}_0 + \tilde{b}_0 \, Z_{i,0}$. For $j = 0$, fuzzy regression coefficients, $\tilde{a}_0$ and $\tilde{b}_0$, are determined by using simultaneous equations given in Eqs. (10a) – (10b) as:

$$n.\hat{a}_{0C} + \left(\sum Z_{i,0}\right).\hat{b}_{0C} = \sum Z_{(i,1)C}$$
$$\left(\sum Z_{i,0}\right).\hat{a}_{0C} + \left(\sum Z_{i,0}^2\right).\hat{b}_{0C} = \sum \left(Z_{i,0} . Z_{(i,1)C}\right)$$

can solve for $\hat{a}_{0C}$ and $\hat{b}_{0C}$,

$$n.\hat{a}_{0S} + \left(\sum Z_{i,0}\right).\hat{b}_{0S} = \sum Z_{(i,1)S}$$
$$\left(\sum Z_{i,0}\right).\hat{a}_{0S} + \left(\sum Z_{i,0}^2\right).\hat{b}_{0S} = \sum \left(Z_{i,0} . Z_{(i,1)S}\right)$$

can solve for $\hat{a}_{0S}$ and $\hat{b}_{0S}$.

Therefore, a hybrid fuzzy least-squares regression model for $\mu = 0.0$ is given by:

$$\hat{\tilde{Z}}_{i,1} = (273.857, \, 192.449) + (1.694, \, 0.030)Z_{i,0}$$

The spreads of fuzzy regression coefficients, at any membership level, can be calculated according to the symmetric triangular membership function [7]. For $\mu = 0.7$, the spreads are obtained as:

$$(1 - \mu)\hat{a}_{0S} = (1 - 0.7)\,192.449 = 57.735$$

$$(1 - \mu)\hat{b}_{0S} = (1 - 0.7)\,0.030 = 0.009$$

Consequently, the hybrid fuzzy least-squares regression equation for $\mu = 0.7$ is:

$$\hat{\tilde{Z}}_{i,1} = (273.857, \, 57.735) + (1.694, \, 0.009)Z_{i,0}$$

Table 3 shows the main results and hybrid fuzzy regression models for each development period at $\mu = 0.0$ membership level. Table 4 shows fuzzy claim reserves at different membership level for $\mu = 0.0, 0.1, 0.2, ... , 0.9, 1.0$

As seen in Table 3, for $\mu = 0.0$ membership level, the amount of estimated fuzzy claims reserve is obtained as (11914.890, 2173.729) where 11914.890 is the center and 2173.729 is the spreads. In case of the fact that the effects of parameters which give rise to uncertainty decrease relatively, the spreads for fuzzy reserve can be reduced by raising the membership level μ. For example, in Table 4, for $\mu = 0.7$, the amount of estimated fuzzy claims reserve is obtained as (11914.890, 652.119) and for $\mu = 1.0$, reserve is obtained as (11914.890, 0) because hybrid fuzzy regression models are consist of the crisp coefficients.

When all data were included in model as a crisp number, hybrid regression would produce the same results as ordinary regression. So, the value of the center in fuzzy reserves must be the same as the result of ordinary least-squares regression. Therefore, as seen in Table 5, the amount of outstanding is determined as 11914.890 according to LCL method.

Table 2. Cumulative run-off triangle for a group of ten automobile liability insurance companies.

Occurrence / origin period (year)	Development period (year)						
	0	**1**	**2**	**3**	**4**	**5**	**6**
0	2062	3691	4274	4695	5036	5312	5540
1	2031	3737	4380	4828	5163	5470	
2	2164	4051	4718	5172	5541		
3	2320	4180	4851	5314			
4	2462	4371	5107				
5	2651	4809					
6	3084						

Table 3. Reserves with hybrid fuzzy regression at $\mu = 0.0$ membership level.

$\mu = 0.0$	Regression Model		Reserve	
Year of development (j)	$\hat{\tilde{a}}_j$	$\hat{\tilde{b}}_j$	**Origin Period (i)**	$\hat{\tilde{R}}_i$
0	(273.857, 192.449)	(1.694, 0.0300)	**0**	--
1	(-38.702, 91.991)	(1.174, 0.0008)	**1**	(234.782, 0.000)
2	(181.637, 52.254)	(1.058, 0.0008)	**2**	(654.234, 0.000)
3	(21.548, -27.274)	(1.067, 0.0420)	**3**	(1074.643, 254.570)
4	(-953.260, 0)	(1.244, 0)	**4**	(1657.481, 347.216)
5	(0, 0)	(1.043, 0)	**5**	(2690.750, 516.527)
			6	(5602.997, 1055.416)
			$\hat{\tilde{R}}$	**(11914.890, 2173.729)**

Table 4. Fuzzy claims reserves at different membership level.

μ – cuts	$\hat{\tilde{R}}$
0.0	(11914.890, 2173.729)
0.1	(11914.890, 1956.356)
0.2	(11914.890, 1738.983)
0.3	(11914.890, 1521.610)
0.4	(11914.890, 1304.238)
0.5	(11914.890, 1086.865)
0.6	(11914.890, 869.492)
0.7	(11914.890, 652.119)
0.8	(11914.890, 434.746)
0.9	(11914.890, 217.373)
1.0	(11914.890, 0)

Table 5. Claims reserves with LCL method.

Year of development (*j*)	Regression Model		Reserve	
	$\hat{a}_j$	$\hat{b}_j$	Origin Period (*i*)	$\hat{R}_i$
0	273.857	1.694	**0**	--
1	-38.702	1.174	**1**	234.782
2	181.637	1.058	**2**	654.234
3	21.548	1.067	**3**	1074.643
4	-953.260	1.244	**4**	1657.481
5	0	1.043	**5**	2690.750
			6	5602.997
			$\hat{R}$	**11914.890**

5. CONCLUSION

Estimating the fair value of provisions for outstanding claims is one of the most important and complex operations that insurance and reinsurance companies are faced with. Improper reserving, either inadequate or excessive, can give a false picture of the company's liabilities on the balance sheet and leads to fatal consequences. Many different statistical methods have been outlined in actuarial literature. It is important to remember, however, that the traditional methods for claims reserving with scarce data, the lack of a theoretical statistical basis prevents information about the reliability of the resulting provisions from being calculated. Therefore, fuzzy set theory becomes very attractive and suitable instrument in modeling problems which require subjective judgment significantly and when inadequate and ambiguous information observed.

In our claims reserving method, hybrid fuzzy least-squares regression analysis is applied to predict future claim costs by using the concept of LCL method. The main parts of our method are fuzzification and hybrid fuzzy regression modeling. In fuzzification, we use probabilistic confidence limits for designing the triangular fuzzy number. Thus, it will allow us to reflect variability measure contained in data set to the prediction of future claim costs. The purpose of using hybrid fuzzy least-squares regression model is to take advantage of integrating both randomness and fuzziness type of uncertainty into a regression model. On the other hand, estimating fuzzy coefficients in fuzzy regression by using the principle of least-squares will allow us to provide theoretical basis, e.g., the estimates are unbiased, the variances are the smallest, etc.

The special feature of hybrid fuzzy least-squares linear regression analysis is that, if all data were handled as crisp numbers, hybrid fuzzy regression could produce the same results as ordinary regression so; the centers of estimated fuzzy values would be the same as the result of LCL method. By way of addition, in insurance environment, in case of the fact that the effects of parameters which give rise to uncertainty decrease relatively, spreads for fuzzy reserve can be reduced by raising the membership level μ.

Existing ordinary regression programs can be used for the hybrid fuzzy regression analysis. The method demonstrates the practicability of performing fuzzy regression in applications involving fuzzy numbers. Consequently, our extension of LCL method forms a complete methodology for claims reserving. And, calculating the amount of outstanding by utilizing the fuzzy intervals provides insurance company with an advantage of determining liability.

REFERENCES

[1] Aliev, R.A., Fazlollahi, B., Aliev, R.R., "Soft Computing and its Applications in Business and Economics", *Springer-Verlag*, New York, 80 – 81 (2003).

[2] Andrés, J.de, Terceño, A., "Applications of fuzzy regression in actuarial analysis", *Journal of Risk and Insurance*, 70(4): 665–699 (2003).

[3] Benjamin, S., Eagles, L.M., "Reserves in Lloyd's and the London market", *Journal of the Institute of Actuaries*, 113(2): 197–257 (1986).

[4] Boulter,A., Grubbs, D., "Late Claims Reserves in Reinsurance", *Swiss Re Press*, Zurich, 5-28 (2000).

[5] Buckley, J.J., "Fuzzy Probability and Statistics", *Springer-Verlag*, New York, 171-195 (2006).

[6] Chang, Y.-H.O., Ayyub, B.M., "Fuzzy regression methods – a comparative assessment", *Fuzzy Sets and Systems*, 119(2): 187–203 (2001).

[7] Chang, Y.-H.O., "Hybrid fuzzy least-squares regression analysis and its reliability measures", *Fuzzy Sets and Systems*, 119(2): 225–246 (2001).

[8] Dubois, D., Prade, H., "Fuzzy Sets for Intelligent Systems", *Morgan Kaufmann Publishers*, San Francisco, CA, 45-61 (1993).

[9] Goovaerts, M.J., Kaas, R., Heerwaarden, A.E., Bauwelinckx, T., "Effective Actuarial Methods", *North-Holland*, Amsterdam, 243-252 (1990).

[10] Hossack, I.B., Pollard, J.H., Zehnwirth, B., "Introductory Statistics with Applications in General Insurance", ***Cambridge University Press***, Cambridge, 206-242 (1999).

[11] Kaufmann, A., Gupta, M.M., "Introduction to Fuzzy Arithmetic: Theory and Applications", ***Van Nostrand Reinhold***, New York, 9-42 (1991).

[12] Moskowitz, H., Kim, K., "On assessing the H value in fuzzy linear regression", ***Fuzzy Sets and Systems***, 58: 303–327 (1993).

[13] Shapiro, A.F., "Fuzzy logic in insurance", ***Insurance: Mathematics and Economics***, 35(2): 399–424 (2004).

[14] Shapiro, A.F., "Fuzzy random variables". ***Insurance: Mathematics and Economics***, 44(2): 307–314 (2009).

[15] Straub, E., "Non-life Insurance Mathematics", ***Springer-Verlag***, Berlin, 102-115 (1997).

[16] Tanaka, H., "Fuzzy data analysis by possibilistic linear models", ***Fuzzy Sets and Systems***, 24: 363–375 (1987).

[17] Taylor, G., "Separation of inflation and other effects from the distribution of non-life insurance claim delays", ***Astin Bulletin***, 10(1): 217–230 (1977).

[18] Verbeek, H.G., "An approach to the analysis of claims experience in motor liability excess of loss reassurance", ***Astin Bulletin***, 6: 195–202 (1972).

[19] Zadeh, L.A., "The roles of fuzzy logic and soft computing in the conception, design and deployment of intelligent systems", ***BT Technology Journal***, 14(4): 32-36 (1996).