Forecasting Mortality Rate Using a Neural Network with Fuzzy Inference System

George Atsalakis[1], Dimitrios Nezis[1], George Matalliotakis[1], Camelia Ioana Ucenic[2], and Christos H. Skiadas[1]

[1] Data Analysis and Forecasting Laboratory,
Technical University of Crete,
73100 Chania, Crete, Greece
(e-mail: atsalak@otenet.gr, dnezis84@yahoo.gr, matal@ermes.tuc.gr, skiadas@ermes.tuc.gr)

[2] Department of Management and Industrial Systems
Technical University Cluj Napoca, Romania
(e-mail: cameliaucenic@yahoo.com)

Abstract. Various methods have been developed to improve mortality forecasts. The authors proposed a neuro-fuzzy model to forecast the mortality. The forecasting of mortality is curried out by an ANFIS model which uses a first order Sugeno-type FIS. The model predicts the yearly mortality in a one step ahead prediction scheme. The method of trial and error was used in order to decide the type of membership function that describe better the model and provides the minimum error. The output of the models is the next year's mortality. The results were presented and compared based on three different kinds of errors: RMSE, MAE, and MAPE. The ANFIS model gives good results for the case of two gbell membership functions and 500 epochs. Finally, the ANFIS model gives better results than the AR and ARMA model.
Keywords. ANFIS, Forecasting, Mortality, Modeling.

1 Introduction

Mortality forecast is very important for the governments and the private insurance companies. Government agencies use the results of these forecasts when planning and developing health policy. Insurance companies, also use these predictions to plan their policies and their strategy for retirement systems.

The issue of mortality risk and, in particular, of longevity risk has been largely caused the interest of scientific community in recent years, when facing the main problem of pricing the insurance products. This risk, is known in the scientific literature as Longevity Risk, and is the risk derived from a future mortality rate. (Brouhns, et all 2002). The price of any insurance product on the duration of life is based on two types of assumptions - demographical and financial. The assumptions are the main reason for which is necessary to include a forecast of the future trends of mortality.

Thus, it is very important to provide improved mortality forecasts. Moreover, it is a well-known fact that the human mortality globally declined. Improvements in standards of living, sanitary conditions and medicine led to rapidly decreasing mortality rates in the early part of the century. People from all over the world, live more and the costs of the public or private insurance become bigger.

Various methods have been developed to improve mortality forecasts. A very popular method of mortality forecasting was introduced by Lee and Carter (1992) to model and forecast U.S. mortality. The main statistical tool of Lee and Carter is the least-squares estimation via singular value decomposition of the matrix of the log age-specific observed forces of mortality together with Box–Jenkins modelling for time series. The method was designed for long-term forecasting based on a lengthy time series of historic data. However, significant structural changes have occurred in mortality patterns over our century, reducing the validity of experience in the more distant past for present forecasts.

The future trend of mortality is modelling as a stochastic process. Furthermore, there is always a systematic danger, which is very difficult to eliminate and it is independent from the plenty of policies that an insurance organisation can apply (Denuit and Frostig, 2005). Specifically, Lee-Carter model describes the logarithmically transformed age-specific central rate of death as a sum of:

- the age-specific component that is independent of time
- the product of a time-varying parameter, also known as the mortality index, that summarizes the general level of mortality and an additional age-specific component that represents how rapidly or slowly mortality at each age varies when the mortality index changes.

As it is easily understandable, the model is very simple to be applied because it has a limited number of parameters. The Lee Charter model became so popular that it was called "leading statistical model of mortality forecasting in the demographic literature" (Deaton and Paxson, 2004). However, there were some criticisms for the model. For example, Bell (1997) noted that the model did not fit the jump off data very well. In addition, it was appeared a number of variants, which tried to optimize the initial model of Lee and Charter. Some of them are presented to the second part of the paper.

The authors proposed a neuro-fuzzy model to forecast the mortality. Fuzzy logic has gained recognition and was intensively applied in mathematics and computer sciences. In the field of data modelling, regression models based on fuzzy data were first developed by Tanaka et al. (1982) and Diamond (1988).

There are many papers related with the issue of mortality forecasting. Each of them, proposes a different variant of Lee Charter model or a new practical method, which is based on basic statistical tools. See C. Pedroza

(2006), Booth H. et al. (2006), Hyndman R.J. and Shahid U. (2005), Skiadas C. (2007) and Skiadas et al. (2007).

2 Model Presentation

The fuzzy logic theory was first formulated by Zadeh (1965) as a new way of characterizing non-probabilistic uncertainties. In contrast to the Boolean 1-0 logic, fuzzy logic also permits in-between values for any judged statement, i.e., it applies a continuous, multi-valued logic between 0 and 1. A fuzzy inference system (FIS) is a computing framework that combines the concepts of fuzzy logic, fuzzy decision rules, and fuzzy reasoning Jang(1993). The fuzzy decision rules are the way an FIS relates an input variable x to an output variable y. In case than more than one variable are involved in the premise side, the structure of the rules take the form:

[1]if *x1* is *A* and *x2* is *B*, then *y* is *Z*

where *x1* and *x2* are the input variables

A, B and *Z* are linguistic values (small or big etc.) defined as membership function (MF) in the input and output spaces.

The steps for creating a fuzzy inference model are:

1. Fuzzification: the input variables are compared with the MFs on the premise part of the fuzzy rules to obtain the probability of each linguistic label.
2. Combine (through logic operators) the probability on the premise part to get the weight (fire strength) of each rule.
3. Application of firing strength to the premise mf of each rule to generate the qualified consequent of each rule depending on their weight.
4. Defuzzification: Aggregate the qualified consequents to produce a crisp output.

A neuro-fuzzy system is defined as a combination of Neural Networks and Fuzzy Inference System. Jang (1993) introduced an Adaptive Neuro-Fuzzy Inference System (ANFIS) where the MF parameters are fitted to a dataset through a hybrid-learning algorithm. The basis of the ANFIS model is the theory of artificial neural networks (ANN). The following figure presents the fuzzy reasoning process.

The example consists on a first order Sugeno type FIS, with two inputs variables (x and y), one output (z), and two if-and-then rules. Each input space has been characterized by two intuitively labeled gauss MF, drawn separately for clarity and to give graphical representation of each rule.

In general, when ANFIS model is used, the shape of the membership functions depends on parameters that can be adjusted to change the shape of the membership function. In addition, the parameters can be automatically adjusted depending on the data that we try to model. Using a given input data set, the toolbox function anfis constructs a fuzzy inference system (FIS)

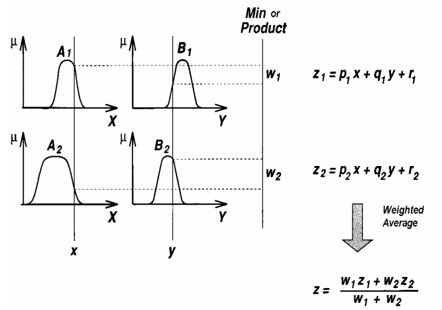

Fig. 1. Fuzzy reasoning process

whose membership function parameters are tuned (adjusted) using either a backpropagation algorithm alone, or in combination with a least squares type of method. This allows fuzzy systems to learn from the data that are modelled

Because of the linear dependence of each rule on the input variables of a system, the Sugeno method is ideal for acting as an interpolating supervisor of multiple linear controllers that are to be applied, respectively, to different operating conditions of a dynamic nonlinear system. A Sugeno fuzzy inference system is extremely well suited to the task of smoothly interpolating the linear gains that would be applied across the input space; it is a natural and efficient gain scheduler.

Advantages of the Sugeno Method are:

1. the computationally efficiency, that it works well with linear techniques, with optimization and adaptive techniques and it has guaranteed continuity of the output surface
2. it is suited for modelling nonlinear systems by interpolating between multiple linear models.

3 Model Description

The forecasting of mortality is curried out by an ANFIS which uses a first order Sugeno-type FIS. The model predicts the yearly mortality in a one step ahead prediction scheme. The method of trial and error was used in order to decide the type of membership function that describe better the model and provides the minimum error.

The gbell membership function made available better results than gauss2mf, trapezodial, triangular and gauss membership functions. Finally two-membership functions of gbell shape were chosen.

The least-squares method and the backpropagation gradient descent method are used for training the Fuzzy Inference System (FIS) membership function parameters. Also it is used a checking data set, for checking the model over fitting. ANFIS model was created using different input variables for each model. The input of each model is lagged values (k-1, or k-2) of the independent variables. In fact they represent the mortality one year before and two years before. The output of the models is the next year's mortality.

4 Results

The data are real word data and has been taken from the site www.mortality.org which provides data of mortality for many countries. The data is the mortality of America. The samples are the yearly deaths of the ages of 50, 55, 60, 65 and 70 years. Data concern the period from 1933 until 2004. The first 70% of data was used for training and the 30% for testing.

Different types of ANFIS models were trained and tested. The disposition of the values is determined by the phases of training and testing data, which have been defined to the algorithm.

The analysis of model quality was done related to three main types of errors: RMSE, MAE and MAPE. A comparison of the results of ANFIS for the case of two gbell membership functions and the models AR and ARMA is presented in the following tables.

Table 1. Forecasting results age (50)

Errors	**ARMA**	**AR**	**ANFIS**
RMSE	0.3071	0.3066	1.1764
MAE	216.3503	218.2318	308.9341
MAPE	2.8042	2.8308	4.0867

Table 2. Forecasting results age (55)

Errors	**ARMA**	**AR**	**ANFIS**
RMSE	408.6207	408.6731	231.4310
MAE	275.3421	282.0063	117.9749
MAPE	2.6748	2.7312	1.1292

Table 3. Forecasting results age (60)

Errors	**ARMA**	**AR**	**ANFIS**
RMSE	548.8466	574.1797	414.4632
MAE	399.9773	406.1126	233.1209
MAPE	2.9573	2.9949	1.7718

Table 4. Forecasting results age (65)

Errors	**ARMA**	**AR**	**ANFIS**
RMSE	590.8704	616.9826	657.9868
MAE	446.0150	458.2732	373.0914
MAPE	2.5928	2.6373	2.2006

Table 5. Forecasting results age (70)

Errors	ARMA	AR	ANFIS
RMSE	748.8269	777.3899	619.8892
MAE	575.4017	584.3968	446.8840
MAPE	2.7717	2.8022	2.2528

The model was tested using different number of epochs. Finally good results were obtained at 500 epochs. As it seems from the above types of errors, the model with the best results, is the

ANFIS model. The comparison between actual values and ANFIS forecasted values is illustrated in figure 2 for the age of 55 years.

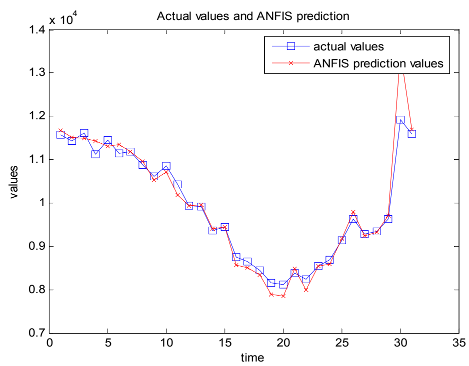

Fig. 2. Comparing actual values – forecasted values (age 55)

5 Conclusion

The paper presented an ANFIS forecasting model that depends on previous load values. The results were presented and compared based on three different kinds of errors: RMSE, MAE, and MAPE. The ANFIS model gives good results for the case of two gbell membership functions and 500 epochs. Finally, the ANFIS model gives better results than the AR and ARMA model. The research will be continued in the future works.

References

Bell, W.R. (1996): “Comparing and Assessing Time Series Methods for Forecasting Age Specific Demographic Rates” *Journal of Official Statistics,* vol.13, pp.279-303.

Booth, H., Hyndman, J.R. (2006): Tickle, L. and P. Jong, “Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions”, *Demographic research,* vol. 15, pp. 289–310.

Brouhns, N., Denuit M. and J.K. Vermunt. (2002): “Measuring the Longevity risk in mortality projections”, *Bulletin of the Swiss Association of Actuaries*, pp.105-130.

Deaton, A., and C. Paxson. (2004): “Mortality Income, and Income Inequality Over Time in the Britain and the United States” *Technical Report 8534 National Bureau of Economic Research Cambridge, MA,* http://www.nber.org/papers/w8534.

Denuit, M., and E. Frostig. (2005): “Association and heterogeneity of insured lifetimes in the Lee–Carter framework”, *Working Paper*, Institut des Sciences Actuarielles, Universit Catholique de Louvain, Louvain-la-Neuve, Belgium.

Diamond, P. (1988): “Fuzzy least-squares”, *Information Sciences*, vol. 3, pp. 141–157.

Harvey, A.C. (1989): “Forecasting, structural time series models and the Kalman filter.” *working paper*, Cambridge University Press.

Jang JS. (1993): ANFIS: Adaptive –Network –based Fuzzy Inference System. *IEEE Trans systems*, *working paper*, Man and cybernetics, vol.23, (**??**),pp. 665–684.

Jang JS and C.Sun. (1993): Predicting chaotic time series with fuzzy if-then rules, *Proceedings of the IEEE international conference on fuzzy systems*, pp.1079–1084.

Khalaf, A.M., Haberman, S. and R.J. Verrall. (2006): “Measuring the effect of mortality improvements on the cost of annuities”, *Insurance: Mathematics and Economics, Elsevier,* vol. 39, pp. 231–249.

Lee, R. D. and L.R. Carter. (1992): ”Modelling and forecasting U.S. mortality”, *Journal of the American Statistical Association,* vol.87, pp. 659–671.

Murray, C. J. and A.D. Lopez. (1996): “The Global Burden of Disease”, *working paper,* Cambridge, *Harvard University Press.*

Pedroza, C. (2006): “ A Bayesian forecasting model: predicting U.S. male mortality”, *Biostatistics*, vol. 4, pp. 530–550.

Rob, J., Hyndman J.R. and M.S. Ullah. (2005): “Robust forecasting of mortality and fertility rates: a functional data approach”, *Working Paper*, Department of Econometrics and Business Statistics, Monash university, Australia.

Shiro, H. and J.R. Wilmoth. (1995): “The aging of mortality decline.” Presented at the Annual Meeting of the Population Association of America, San Francisco, *working paper.*

Sithole, T.Z., Haberman, S. and R..J. Verrall. (2000): “An investigation into parametric models for mortality projections, with applications to immediate annuitants and life office pensioners’ data”, *Insurance: Mathematics and Economics, Elsevier,* vol. 27, pp. 285–312.

Skiadas C. and C.H. Skiadas. (2007): “A modelling approach to life table data sets”, in *Recent advances in stochastic modelling and data analysis*, C. H. Skiadas, Ed., Word Scientific, Singapore, pp. 350–359.

Skiadas C., Matalliotakis G. and Skiadas C.H., “ An extened quadratic health state function and the related density function for life table data.”, in *Recent advances in stochastic modelling and data analysis*, C. H. Skiadas, Ed., Word Scientific, Singapore, pp. 360–369.

Tanaka, H., Uejima, S. and K. Asai. (1982) “Linear regression analysis with fuzzy model”. *IEEE Transactions on Systems*, *Man and Cybernetics,* vol. 6, pp. 903–907.

Zadeh, L. A. (1965): *Fuzzy sets,* Inform. Contr., vol. 8, pp. 338–353.