New Method of Learning and Knowledge Management in Type-I Fuzzy Neural Networks

Tahseen A. Jilani and Syed Muhammad Aqil Burney

Department of Computer Science, University of Karachi,
Karachi-74270, Pakistan
tahseenjilani@ieee.org, burney@computer.org

Abstract. A new method for modeling and knowledge extraction at each neuron of a neural network using type-I fuzzy sets is presented. This approach of neuron modeling provides a new technique to adjust the fuzzy neural network (FNN) structure for feasible number of hidden neurons and efficient reduction in computation complexity. Through repeated simulations of a crisp neural network, we propose the idea that for each neuron in the network, we can obtain reduced model with high efficiency using wavelet based multiresolution analysis (MRA) to form wavelet based fuzzy weight sets (WBFWS). Triangular and Gaussian membership functions (MFs) are imposed on wavelet based crisp weight sets to form Wavelet Based Quasi Fuzzy Weight Sets (WBQFWS) and Wavelet Based Gaussian Fuzzy Weight Sets (WBGFWS). Such type of WBFWS provides good initial solution for training in type-I FNNs. Thus the possibility space for each synoptic connection is reduced significantly, resulting in fast and confident learning of FNNs. It is shown that propsed modeling approach hold low computational complexity as compared to existing type-I fuzzy neural network models.

Keywords: extraction of fuzzy rules, fuzzy neural networks, neuro-fuzzy modeling Wavelet based multiresolution analysis.

1 Introduction

Fission of artificial neural networks [10] and fuzzy sets have attracted the growing interest of researchers in various scientific and engineering areas due to the growing need of adaptive intelligent systems to solve the real world problems. A crisp or fuzzified neural network can be viewed as a mathematical model for brain-like systems. The learning process increases the sum of knowledge of the neural network by improving the configuration of weight factors. An overview of different FNN architectures is discussed by [5] and [9]. It is much more difficult to develop the learning algorithms for the FNN than for the crisp neural networks. This is because the inputs, connection weights and bias terms related to a FNN are fuzzy sets; see [19] and [22].

The new technique in mathematical sciences called wavelets can be introduced to reduce the problem complexity as well as the dimensions so that a FNN may provide a fast track for optimization. Wavelet based MRA provides better analysis of complex signals than Fourier based MRA, see [7].

O. Castillo et al. (Eds.): Theor. Adv. and Appl. of Fuzzy Logic, ASC 42, pp. 730–740, 2007.
springerlink.com

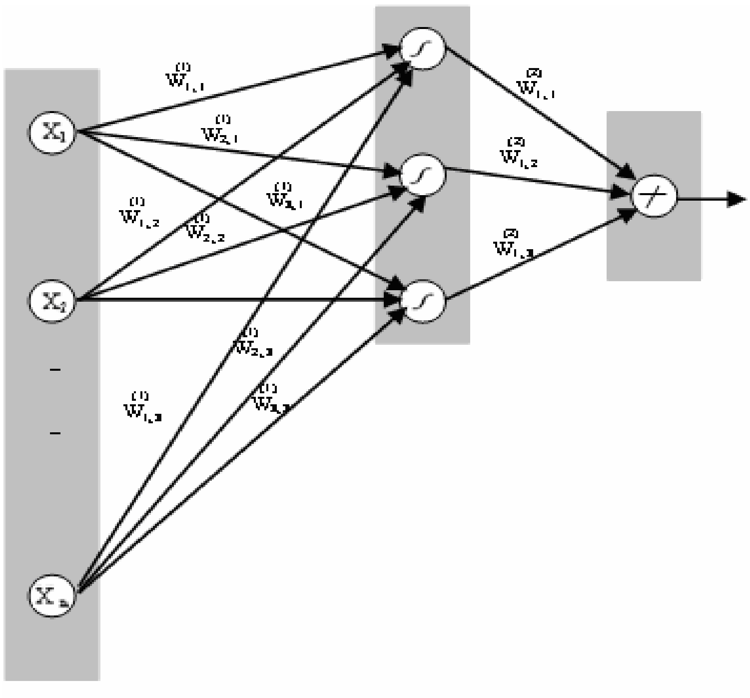

Fig. 1. Structure of a crisp neural network

Based on this approach of knowledge discovery for each synaptic connection, we can convert the probability concept of network connections into possibility perception. These sets provide the initial design for type-I neuro-fuzzy networks as discussed in [1], [3] and [11]. When jumbled with [20], this new approach assures lower computational complexity due to improved selection of seed values of the network. To our knowledge, the concept of obtaining WBFWS through crisp neural networks has not been investigated in the literature. The work is organized as follows. In section 2 we have briefly discussed wavelet based multiresolution analysis technique. In section 3, we have given new approach for fuzzy modeling of network connections. In section 4, simulation experiments are presented. To determine accuracy of WBFWS, a comparison is also made between proposed WBFWS and Gaussian confidence intervals for each hidden synaptic connection of the neural network. Finally, discussions and concussions are given in section 5.

2 Wavelet Based Multiresolution Analysis

In recent years, researchers have developed powerful wavelet techniques for the multi scale representation and analysis of signals see [6], [7] and [17]. These new methods differ from the traditional Fourier technique. Wavelets localize the information in the time-frequency space which makes them especially suitable for the analysis of non-stationary signals [15]. One important area of applications where wavelets have been found to be relevant is fuzzy neural systems is discussed in [9] and [14]. This whole area of research is still relatively new but is evolving very rapidly. We examine the very important property of wavelet transformation i.e. maximization of signal energy using data compression for FNNs. There are essentially two types of wavelet decompositions, Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT), see [21]. Continuous wavelets are usually preferred for signal analysis, feature extraction and detection tasks where as the second type is obviously more adequate whenever it is desirable to perform some kind of data reduction or when the orthogonality of the representation is an important factor [7]. However, the choice between them is optional depending upon the computational considerations. We will use the decomposition in terms of DWT using Mallat's pyramid algorithm which is faster than a CWT and obtained very satisfactory results. Let $f(t)$ be a signal defined in $L^2(R)$ space, which denotes a vector space for finite energy signals, where R is a real continuous number system. The wavelet transformation of $f(t)$ in terms of continuous wavelets is then defined as

$$CWT_{\psi} f(a,b) = \langle f, \psi_{a,b} \rangle = \int_{\infty}^{\infty} f(t)\psi_{a,b}(t)\,dt \tag{1}$$

$$\text{where} \quad \psi_{a,b}(t) = |a|^{\frac{1}{2}} \psi\left(\frac{t-b}{a}\right)$$

$\psi(t)$ is the base function or the mother wavelet with $a,b \in R$ are the scale and translation parameters respectively. Instead of continuous dilation and translation, the mother wavelet may be dilated and translated discretely by selecting $a = a_0^{\ m}$ and $b = nb_0 a_0^{\ m}$, where a_0 and b_0 are fixed values with $a_0 > 1$, $b_0 > 0$, $m,n \in Z$ and Z is the set of positive integers. Then the discretized mother wavelet becomes

$$\psi_{m,n}(t) = a^{-\frac{m}{2}} \psi\left(\frac{t - nb_0 a_0^{\ m}}{a_0^{\ m}}\right) \tag{2}$$

and the corresponding discrete wavelet transform is given by

$$DWT_{\psi} f(m,n) = \langle f, \psi_{m,n} \rangle = \int_{-\infty}^{\infty} f(t)\psi_{m,n}(t)\,dt \tag{3}$$

DWT provides a decomposition and reconstruction structure of a signal using MRA through filter bank. The role of mother scaling and mother wavelet functions $\phi(t)$ and $\psi(t)$ are represented through a low pass filter L and a high pass filter H. Consequently, it is possible to obtain a signal f through analysis and synthesis by using wavelet based MRA [21].

$$f(t) = \sum_{n \in Z} c_{p,n} \phi_{p,n}(t) + \sum_{0 \le m \ge p} \sum_{n \in Z} d_{m,n} \psi_{m,n}(t) \quad (4)$$

where the sum with coefficients $c_{p,n}$ represents scaling or approximation coefficients and sums with coefficients $d_{m,n}$ represent wavelet or detail coefficients on all the scales between 0 and p. Data compression and energy storage in wavelets can be achieved by simply discarding certain coefficients that are insignificant. We combine this property of wavelets with neural networks and found a special class of mother wavelets db4, the most appropriate based on our data. We studied the effect of crisp weights on different neurons by reducing them using wavelets according to their energy preservation.

3 New Technique for Fuzzy Neural Network Synaptic Weights

Methods of fuzzy logic are commonly used to model a complex system by a set of rules provided by the experts [1]. To form WBQFWS from crisp neural network, we can obtain an appropriate unique MF by applying fuzzy aggregation operations. Such aggregation operations allow modeling each free parameter by combining multiple possible alternatives, see [16] and [23]. As in our proposed method, each synaptic fuzzy weight is obtained by ordered selection from crisp simulations, thus we can also impose ordered weighted aggregation operation (OWA) on re-simulated crisp neural networks. The quasi fuzzy weight sets follows all axioms of fuzzy aggregation functions like boundary condition, monotonicity, continuity, idempotency and strict buoyancy.

Let $h:[0,1]^n \rightarrow [0,1]$ denote the fuzzy aggregation from n dimensional space to a unique value given as $A(x) = h(A_1(x), A_2(x), ..., A_n(x))$ for each $x \in X$, where $A_i(x)$ are fuzzy sets defined over X. Thus a quasi-fuzzy set **A** is defined as

$$\min(a_1, a_2, ..., a_n) \le h(a_1, a_2, ..., a_n) \le \max(a_1, a_2, ..., a_n) \quad (5)$$

for all n-tuples $(a_1, a_2, ..., a_n) \in [0,1]^n$. We call $a_l = \min(a_1, a_2, ..., a_n)$ and $a_r = \max(a_1, a_2, ..., a_n)$. For middle parameter, we define simple averaging of quasi fuzzy sets as

$$h_\alpha(a_l, a_r) = \left(\frac{a_l^\alpha + a_r^\alpha}{2} \right)^{\frac{1}{\alpha}} \quad (6)$$

For simplicity, we have assumed $\alpha = 1$. For FNN with WBQFWS, we can also introduce non-symmetric fuzzy MFs by varying the parameter α.

4 Experiment

A crisp neural network with three hidden and one output neuron is trained and repeated the simulations for 700 times with an average rate of 7 simulations before a successful simulation. Through wavelet decomposition, we reduced dimensions by preserving 95% of the energy of original signal. The decomposed signal at level 5 using db4 wavelets for one of the input weight vectors is shown in Fig. 2(a), and its compressed version along with original signal in Fig. 2(b). A threshold of 5% is used for the compression of signals, thus reducing the data dimensionality up to 50%. So that in place of high data requirements of QFWS in [4], we obtained better performance using WBFWS, see [2]. From Fig. 1, the parameters of triangular MF based WBQFWS are given in Eq. (1) and Eq. (2). Although other MF can also be used depending upon the problem addressed.

As most of the actuarial problems are full of fuzziness [24], thus skewed MFs like Gamma and Beta can also be imposed. In Fig. 5 and Fig. 6, we have shown Gaussian MF based modeling for synaptic connections of a neural network.

In Table 1, results of WBQFWS provide superior mapping of weight space obtained using repeated simulations of crisp neural network than 95% Gaussian confidence interval. Interesting to note that the mean and standard errors are similar up to five decimal places in WBQFWS showing consistency of proposed interval sets but not in the case of Gaussian bounds. When each of the 100 simulated values of weights are validated for significance then we observe considerable differences in the prediction capacity of two types of interval sets. Similarly, the 95% WBGFWS also perform better then 95% Gaussian intervals. In validation process, the proposed WBFWS provide more accurate bound as compared to 95% Gaussian confidence intervals (less than 93% close sets) as show in Table 3 for WBQFWS and in Table 6 for WBGFWS.

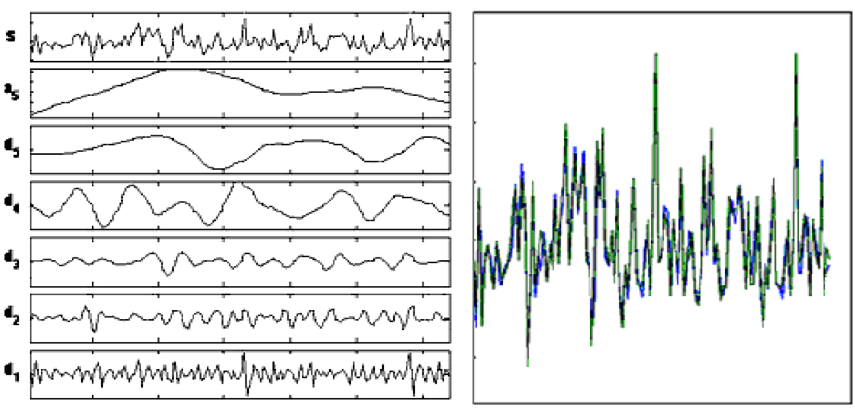

Fig. 2. (a) Signal decomposition (b) original and compressed signals

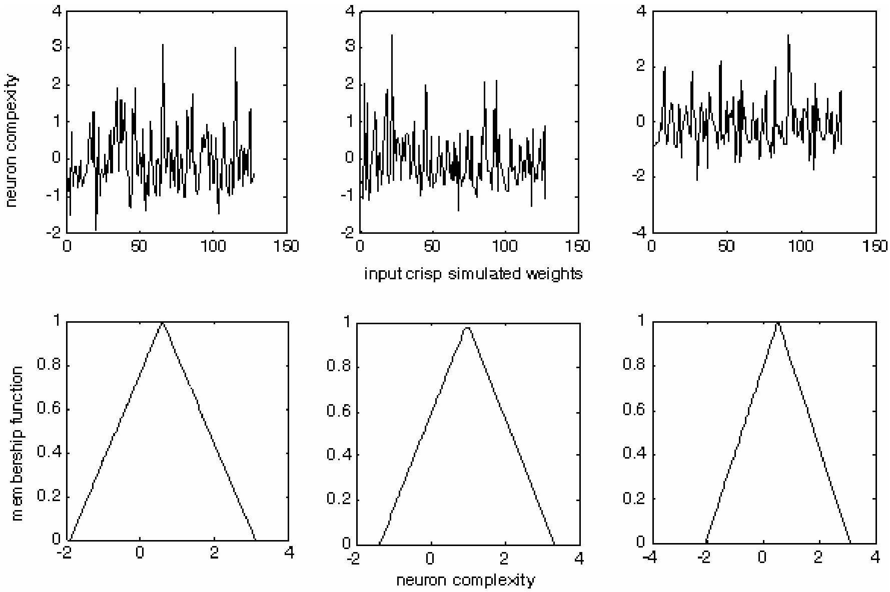

Fig. 3. Wavelet based weights and corresponding triangular MF

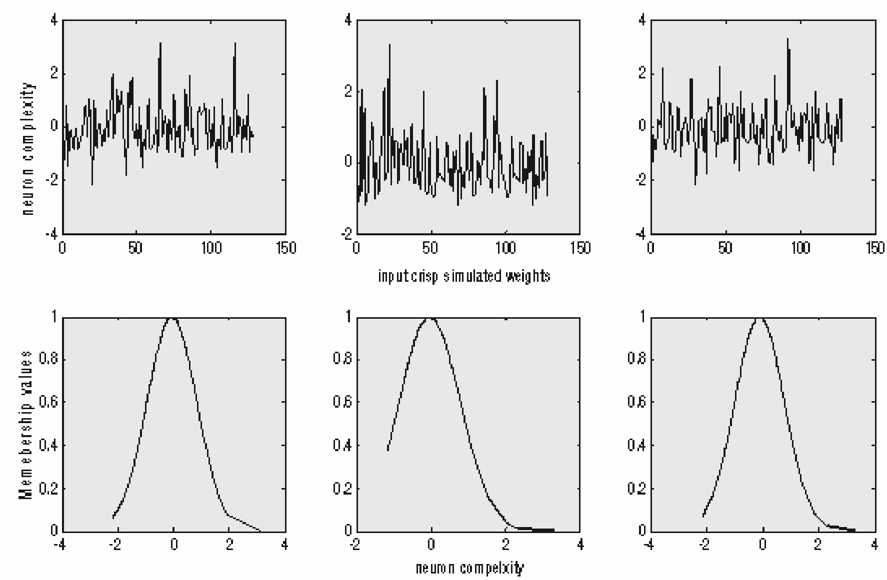

Fig. 4. (a) Input weight matrix for first input vector, (b) Corresponding Gaussian MF

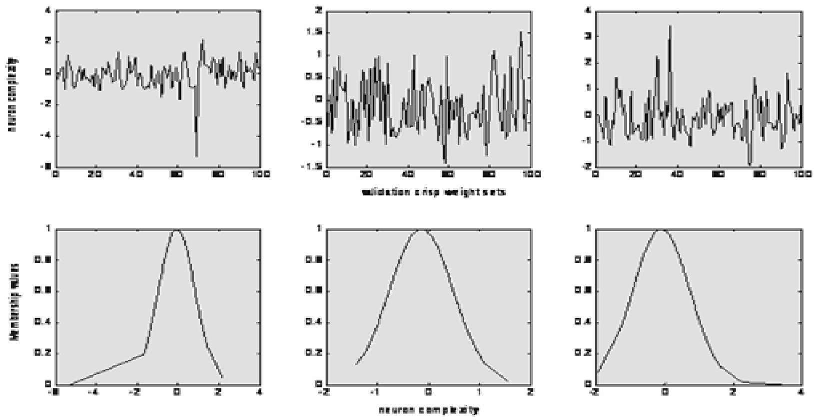

Fig. 5. (a) Validation weight matrix for first input vector, (b) Corresponding Gaussian MF

Table 1. A Comparison of WBQFWS and 95% Gaussian Confidence Intervals

Weight position (from Fig. 1)	(a) WBQFWS				(b) 95% Gaussian C. I.			
	Min	Max	Mean	S. E.	Confidence Bound		Mean	S. E.
$W_1(1,1)$	-1.92	3.09	-0.06	0.86	-1.52	1.40	-0.06	0.89
$W_1(2,1)$	-1.37	3.36	-0.04	0.80	-1.39	1.30	-0.04	0.82
$W_1(3,1)$	-2.09	3.13	-0.11	0.85	-1.56	1.33	-0.11	0.88
$W_2(1,1)$	-2.90	3.22	0.26	1.20	-1.82	2.33	0.25	1.26
$W_2(1,2)$	-1.98	3.13	0.19	1.09	-1.65	2.04	0.19	1.12
$W_2(1,3)$	-2.26	2.92	0.31	1.03	-1.45	2.07	0.31	1.07

Table 2. Validation of WBQFWS and 95% Gaussian Confidence Intervals

Weight position (from Fig. 1)	(a) WBQFWS				(b) 95% Gaussian C. I.			
	Min	Max	Mean	S. E.	Confidence Bound		Mean	S. E.
$W_1(1,1)$	-1.92	3.09	-0.06	0.89	-1.52	1.40	-0.06	0.89
$W_1(2,1)$	-1.37	3.36	-0.13	0.62	-1.16	0.88	-0.13	0.62
$W_1(3,1)$	-2.09	3.13	-0.09	0.83	-1.46	1.27	-0.09	0.83
$W_2(1,1)$	-2.90	3.22	-0.01	0.92	-1.52	1.52	-0.01	0.92
$W_2(1,2)$	-1.98	3.13	0.09	1.13	-1.77	1.97	0.09	1.13
$W_2(1,3)$	-2.26	2.92	0.39	1.12	-1.45	2.23	0.39	1.12

Table 3. Number of Insignificant WBQFWS and Gaussian Confidence Interval

Weight position (from Fig. 1)	WBQFWS	95% Gaussian C.I
$W_1(1,1)$	1	5
$W_1(2,1)$	1	2
$W_1(3,1)$	1	7
$W_2(1,1)$	0	7
$W_2(1,2)$	0	13
$W_2(1,3)$	0	13
Deficiency	0.43%	7.80%
Accuracy through validation	99.57%	92.20%

Table 4. A Comparison of WBGFWS with 95% Gaussian Confidence Intervals

Weight position (from Fig. 1)	(a) WBGFWS				(b) 95% Gaussian C. I.			
	Min	Max	Mean	S. E.	Confidence Bound		Mean	S. E.
$W_1(1,1)$	-1.92	3.09	-0.06	0.86	-1.52	1.40	-0.06	0.89
$W_1(2,1)$	-1.37	3.36	-0.04	0.80	-1.39	1.31	-0.04	0.82
$W_1(3,1)$	-2.09	3.13	-0.11	0.85	-1.56	1.3	-0.11	0.88
$W_2(1,1)$	-2.90	3.22	0.26	1.21	-1.82	2.36	0.25	1.26
$W_2(1,2)$	-1.98	3.13	0.19	1.09	-1.65	2.04	0.19	1.12
$W_2(1,3)$	-2.26	2.92	0.31	1.04	-1.45	2.07	0.31	1.07

Table 5. Validation of 95% WBGFWS and 95% Gaussian Confidence Intervals

Weight position (from Fig. 1)	(a) 95% WBGFWS				(b) 95% Gaussian C. I.			
	Confidence Bound		Mean	S. E.	Confidence Bound		Mean	S. E.
$W_1(1,1)$	-1.81	1.68	-0.06	0.89	-1.53	1.40	-0.06	0.89
$W_1(2,1)$	-1.36	1.08	-0.14	0.62	-1.16	0.89	-0.15	0.63
$W_1(3,1)$	-1.73	1.54	-0.09	0.83	-1.47	1.28	-0.09	0.84
$W_2(1,1)$	-1.81	1.82	-0.01	0.92	-1.53	1.52	-0.00	0.93
$W_2(1,2)$	-2.13	2.33	0.09	1.14	-1.77	1.97	0.09	1.14
$W_2(1,3)$	-1.80	2.59	0.39	1.12	-1.45	2.24	0.39	1.12

Table 6. Number of Insignificant WBGFWS and Gaussian Confidence Interval

Weight position (from Fig. 1)	95 % WBGFWS	95% Gaussian C.I
$W_1(1,1)$	2	5
W1(2,1)	1	2
$W_1(3,1)$	2	7
$W_2(1,1)$	1	7
$W_2(1,2)$	9	13
$W_2(1,3)$	8	13
Deficiency	3.80%	7.80%
Accuracy through validation	96.20%	92.20%

5 Conclusions

Learning with compressed wavelet neural networks using fuzzy weights is efficient and demonstrates much higher level of generalization and shorter computing time as compared to FNNs. We described the architecture of WBFWS based FNN that provide better initial search for synaptic weights. For WBQFWS, results showed that less than 1% chance of bound independent values is possible, thus providing above 99% accurate mapping, in comparison with Gaussian bounds that is below 93%.

In the following, we have presented some aspects of WBFWS regarding the computational complexity of learning in FNN,

1. In FNNs, one of the methods of learning is based on level sets where each fuzzy synaptic connection is divided into υ intervals that satisfy fuzzy arithmetic operations. When assuming υ level sets, we get 2υ parameters of the form

$$\left\{\left(w_L^{h_1}, w_R^{h_1}\right), \left(w_L^{h_2}, w_R^{h_2}\right), \ldots, \left(w_L^{h_\upsilon}, w_R^{h_\upsilon}\right)\right\}$$

Using boundary conditions, monotonicity and weak buoyancy conditions of fuzzy weights, we can write

$$w_L^{h_1} \le w_L^{h_2} \le \ldots. \le w_L^{h_\upsilon} \le w_R^{h_1} \le w_R^{h_2} \le \ldots \le w_R^{h_\upsilon}$$

Thus in FNN with crisp inputs and fuzzy weights, we can define fuzzy MF for each synaptic connection using level sets using WBFWS. If a certain MF consists of k parameters, then total free parameters for a single fuzzy synaptic connection are $2k\upsilon$. For example in a FNN structure $10\times3\times1$ with triangular MF and 100 level sets contains $(10\times3\times3\times2\times100)+(3\times3\times2\times100)=19800$ free parameters to be tuned in each iteration. The computational complexity of even a very small FNN makes them nearly impossible to apply on small scale problems because due to very large number of free parameters of the network, see [8], [12] and [13]. We need huge amount of input data to minimize the performance function (usually Mean Square Error) and to control the degrees of freedom of the network, see [20]. Thus our proposed weight sets with limited possibility space gives faster and more reliable convergence of learning using level sets based FNNs.

2. With the help of our proposed WBFWS, we can decide the architecture of a FNN. If for any connection $w_L^h = w_R^h = 0$ then that connection can be considered as dead connection resulting in a reduced model and thus raising the speed of learning in level sets based FNNs.

6 Future Work

Further improved identification of suitable MF is possible by determining the underlying probability structure of synaptic connections of a crisp neural network using non-parametric statistics like Kernel estimation and learning. Thus based on this idea, we can form fuzzy inference systems with varying rules based on neuron complexity. This may provide new research directions to compare different WBFWS based FNNs. For most of the actuarial problems with non-negative limits, we can propose SANFIS (Skewed Adaptive Neuro-Fuzzy Inference System) to give new impression for actuaries towards this comparatively new field of modeling, forecasting and control. A comparison for most suitable wavelet and optimization algorithms with varying learning parameters is also possible. As future work we will extend this concept on multivariate kernel estimation techniques, and type-II fuzz logic systems as worked by [18].

References

1. Buckley, J.J., Hayashi, Y.: Neural Networks for Fuzzy Systems, Fuzzy Sets and Systems. 71(3) (1995) 265-276.
2. Burney, S.M.A., Jilani, T.A., Saleemi, A.: Approximate Knowledge Extraction Using MRA for Type-I Fuzzy Neural Networks. Proc. of IEEE-ICEIS, (2006) 464-469. Islamabad Pakistan.
3. Burney, S.M.A., Jilani, T.A., Saleemi, A.: Optimizing Neuronal Complexity Using Wavelet Based Multiresolution Analysis for Type-I Fuzzy Neural Networks. Proc. of 4th WSEAS-CIMMACS, (2005) 210-216. Miami, Florida, USA.
4. Burney, S.M.A., Jilani, T.A., Ardil, C.: Approximate Bounded Knowledge Extraction Using Type-I Fuzzy Logic. Proc. of the 5th World Enformatika Conference, WEC'05 (2005). Prague, Czech Republic.
5. Castellano, G., Fanelli, A.M.: Fuzzy Inference and Rule Extraction Using Neural Network. Neural Network World Journal. 3 (2000) 361-371.
6. Cohen, A.: Numerical Analysis of Wavelet Methods. Studies in Mathematics and Its Applications. Vol. 32. Elsevier publishers (2003)
7. Daubechies, I.: Ten Lectures on Wavelets. CBMS-NSF Regional Conference Series. SIAM, Philadelphia (1992)
8. Duch, W.: Uncertainty of Data, Fuzzy Membership Functions, and Multilayer Perceptron. IEEE Trans. on Neural Networks, 16(1) 10-23.
9. Dunyak, J., Wunsch, D.: Training Fuzzy Numbers Neural Networks With Alpha-cut Refinement. Proc. of IEEE Int. Conf. on System, Man, Cybernetics. 1, (1997) 189-194.
10. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice Hall USA (1999)
11. Hee, Y.Y., Chong, F.L., Zhong, B.L.: A Hierarchical Evolutionary Algorithm for Constructing and Training Wavelet Networks. Neural Computing & Applications. 10 (2002) 357-366
12. Ishibuchi, H., Nii, M.: Numerical Analysis of the Learning of Fuzzified Neural Networks from If-Then Rules. Fuzzy Sets and System. 120(2) (2001) 281-307.
13. Ishibuchi, H., Fujioka, R., Tanaka, H.: Neural Networks That Learn from Fuzzy If-Then Rules. IEEE Trans. on Fuzzy System. 1(2) (1993) 85-97
14. Jang, J.-S.R, Sun, C.-T, Mizutani, E.: Neuro-Fuzzy Logic and Soft Computing, a Computational Approach to Learning and Machine Intelligence. Practice-Hall, USA (2003)
15. Kaiser, G.: A Friendly Guide to Wavelets. Birkhauser Boston MA USA (1994)
16. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic, Theory and Applications. Prentice Hall, India (2005)
17. Mallat, S.G.: A Theory of Multiresolution Signal Decomposition: the Wavelet Representation. IEEE Trans. on Pattern Analysis and Machine Intelligence. 11(7) (1989) 674-693
18. Mendel, J.M.: Uncertainly Rule-Based Fuzzy Logic Systems, Introduction and new Directions. Prentice Hall, PTR NJ USA (2001)
19. Mitra, S., Hayashi, Y.: Neuro-Fuzzy Rule Generation: Survey in Soft Computing Framework. IEEE Trans. on Neural Networks. 11(3) (2000) 748–768
20. Park, S., Han, T.: Iterative Inversion of Fuzzified Neural Networks. IEEE Trans. on Fuzzy Systems. 8(3) (2000) 266-280.

21. Percival, D.B., Walden, A.T.: Wavelet Methods for Time Series Analysis. Cambridge University Press (2002)
22. Puyin, L. Xing, H.: Efficient Learning Algorithms for Three-Layer Regular Feedforward Fuzzy Neural Networks. IEEE Trans. on Fuzzy Systems. 15(3) (2004) 545-558
23. Yager, R.R., Filev, D.P.: Essentials of Fuzzy Modeling and Control. John Wiley and Sons (2002)
24. Zimmerman, H.J.: Fuzzy Set Theory and Its Applications. Kluwer Publishers. Boston MA (2001)