Neuro-fuzzy Models for Geomagnetic Storms Prediction

Using the Auroral Electrojet Index

Mahboobeh Parsapoor[1,2], Urban Bilstrup[1], Bertil Svensson[1]
[1]{Mahboobeh.Parsapoor,Urban.Bilstrup, Bertil.Svensson }@hh.se
[2]mahboobeh.parsapoor@cs.man.ac.uk
[1]School of Information Science, Computer and Electrical Engineering (IDE),Halmstad University, Halmstad, Sweden
[2]School of Computer Science, Faculty of Engineering & Physical Science,
The University of Manchester, Manchester, UK

Abstract— **This study presents comparative results obtained from employing four different neuro-fuzzy models to predict geomagnetic storms. Two of thes neuro-fuzzy models can be classified as Brain Emotional Learning Inspired Models (BELIMs) These two models are BELFIS (Brain Emotional Learning Based Fuzzy Inference System) and BELRFS (Brain Emotional Learning Recurrent Fuzzy System). The two other models are Adaptive Neuro-Fuzzy Inference System (ANFIS) and Locally Linear Model Tree (LoLiMoT) learning algorithm, two powerful neuro-fuzzy models to accurately predict a nonlinear system. These models are compared for their ability to predict geomagnetic storms using the AE index.**

Keywords—*Adaptive Neuro-fuzzy Inference System; Auroral Electrojet; Brain Emotional Learning-inspired Model; Locally linear model tree learning algorithm.*

I. Introduction

The auroral electrojet (AE) (Table I lists the abbreviations of this paper) index, which measures auroral zone magnetic activity [1], has been proposed for use as a global quantitative index to characterize the magnetosphere's geomagnetic activities and geomagnetic sub storms. In other words, geomagnetic sub storms can occur in the peak points of the AE index. Empirical studies [2]-[4] aiming to model AE time series, that is a nonlinear time series of the AE index, started in 1971. So far, different models such as neural networks, linear filters and nonlinear filters such as the nonlinear moving average (MA) filter and the nonlinear autoregressive moving average (NARMA) have been proposed to model AE time series (recorded by the World Data Center for Geomagnetism and Space Magnetism at Kyoto University) [2]-[8]. For example, in [6], a real-time learning model was proposed to model the AE index and predict geomagnetic sub storms. A neuro-fuzzy method using a locally linear neuro-fuzzy model tree algorithm has been used to predict the AE index in [7].

An emotion-based machine-learning approach called BEL was utilized to predict the AE index in [9]. The above mentioned algorithms aimsat providing accurate results when predicting geomagnetic storms, using the AE index.

TABLE I. ABBREVIATIONS

Phrase	Definition
AE	Auroral Electrojet
LoLiMoT	Locally Linear Model Tree
BEL	Brain Emotional Learning
BELFIS	Brain Emotional Learning Based Fuzzy Inference System
BELRFS	Brain Emotional Learning Based Recurrent Fuzzy System
BELIMs	Brain Emotional Learning Inspired Models
ANFIS	Adaptive Neuro-Fuzzy Inference System
MA	Moving Average (MA)
NARMA	Nonlinear AutoRegressive Moving Average (NARMA)
TH	THalamus
AMYG	AMYGdala
ORBI	ORBItofrontal cortex
CX	Sensory CorteX,

This paper compares the performance of four neuro-fuzzy models to predict AE time series. The goal is to find a powerful neuro-fuzzy model in order to develop an alert system of geomagnetic sub storms.

The rest of the paper is organized as follows: Section II describes the general structure of BELFIS and BELRFS. Section III redefines the algorithms of two well-known neuro-fuzzy models: ANFIS and LoLiMoT. In Section IV, the results of different methods to model AE time series are compared. Finally, conclusions about this research work are presented and the future work of this study is discussed in Section V.

II. BRAIN EMOTIONAL LEARNING INSPIRED MODELS

Brain Emotional Learning-Inspired Models (BELIMs) [10] can be classified as computational intelligence models. The general structure of BELIMs is an extension of the amygdala-orbitofrontal system [11] by adding new internal parts that have been inspired by the neural structure of the brain emotional system, in particular the circuit of fear conditioning, which emphasizes the role of the amygdala and its components.

The function of each part of a BELIM can be implemented by assigning an adaptive network to that part. The general structure of BELIMs is described in Fig. 1. The components of this structure and their interconnections are imitated by those parts of the brain that are responsible for the emotional learning process. The structure of the model consists of four main parts: TH, CX, AMYG and ORBI which refer to the THalamus, sensory CorteX, AMYGdala, and ORBItofrontal cortex, respectively. These parts have important roles in emotional learning. Certainly, the emotional system's regions are very complex, and this structure has of course not mimicked all their connections in detail. The suggested structure has been the basis of Brain Emotional Learning-Inspired Models (BELIMs) [10], [12] such as the Brain Emotional Learning-based Fuzzy Inference System (BELFIS) [13],[14], the Brain Emotional Learning-based Recurrent Fuzzy System (BELRFS) [15],[16], and the Emotional Learning Inspired Ensemble Classifier (ELiEC) [17].

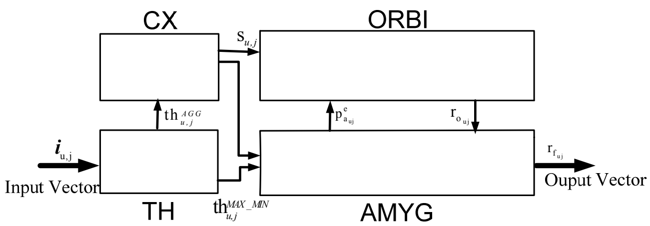


Fig.1. The structure of BELIM.

The following steps describe the input and output of each part of BELPM; when it receives an input as $i_{u,j}$ from the training data set, $\mathbf{I}_u = \{\mathbf{i}_{u,j}\}_{j=1}^{N_u}$.

1. First, $\mathbf{i}_{u,j}$ the j^{th} input vector from $\mathbf{I}_u = \{\mathbf{i}_{u,j}\}_{j=1}^{N_u}$ (taking the assumption that the number of training samples is equal to N_u ; the subscript u has been used to determine the input data is chosen from the training data set) enters the TH, which provides two outputs, $\mathbf{th}_{u,j}^{Max\_Min}$ and $\mathbf{th}_{u,j}^{AGG}$ which are sent to the AMYG and the CX, respectively.

2. The CX provides $\mathbf{s}_{u,j}$ and sends it to both the AMYG and the ORBI.

3. The AMYG receives two inputs: $\mathbf{th}_{u,j}^{Max\_Min}$ and $\mathbf{s}_{u,j}$. It provides the primary output ($r_{a_{u,j}}$) and expected punishment ($p^e_{a_{u,j}}$), that is sent to the ORBI (the subscript a has been used to show the outputs of AMYG).

4. The ORBI receives $\mathbf{s}_{u,j}$ and $p^e_{a_{u,j}}$. It provides the secondary output, $r_{o_{u,j}}$, and sends it to the AMYG.

5. The AMYG receives $r_{o_{u,j}}$ and provides the final output, $r_{f_{u,j}}$ (the subscript f has been used to show the final outputs).

A. *Functional Aspect of BELPM*

The function of BELPM is implemented by assigning adaptive networks to different parts. Figure 2 describes how the adaptive networks can be assigned to each part to implement the functionality of that part. The adaptive network consists of a number of nodes that are connected by directional links. The nodes of the adaptive network can be classified into circle and square nodes. A circle node has a function without adjustable parameters; in contrast, the square nodes have been defined by a function with adjustable parameters. The learning parameters of an adaptive network are a combination of linear and nonlinear parameters and can be adjusted by using a learning algorithm.

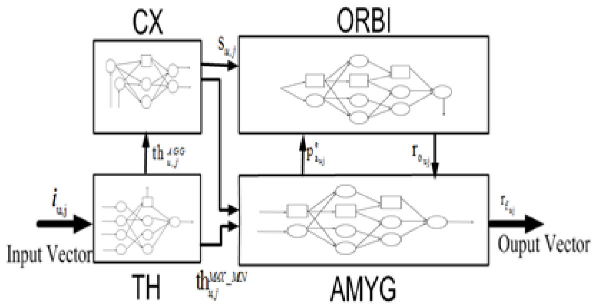


Fig. 2. Assigning different adaptive networks to the parts of BELPM.

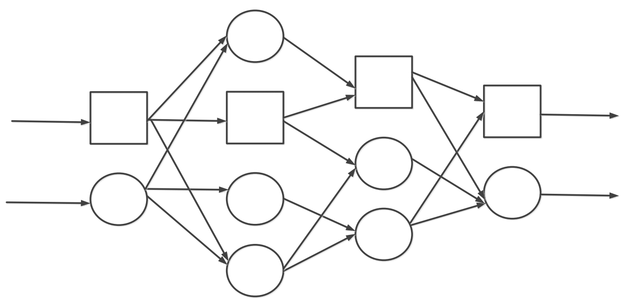


Fig. 3. A structure of an adaptive network.

III. TWO WELL-KNOWN ADAPTIVE NEURO-FUZZY METHODS

A. *Adaptive Neuro-fuzzy Inference System (ANFIS)*

A simple adaptive neuro-fuzzy inference system (ANFIS) is depicted in Fig. 4; this network receives a two-dimensional input vector ($\mathbf{i}_{u,j} = \{i_{u,j1}, i_{u,j2}\}_{j=1}^{N_u}$). An adaptive network can be adapted to the Sugeno fuzzy inference system or the Mamdani fuzzy inference system [18].

The following steps explain the function of each layer of the adaptive network depicted in Fig. 4 with a two-dimensional input vector, $\mathbf{i}_{u,j}=\{i_{u,j1},i_{u,j2}\}_{j=1}^{N_u}$. Note that $\mathbf{i}_{u,j}$ is a vector in two dimensions.

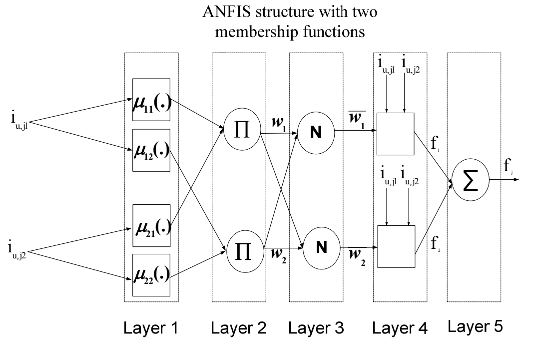


Fig. 4. The third type of ANFIS with two rules.

Layer 1: This layer consists of four square nodes, which are known as adaptive nodes. Each adaptive node can be assigned a Gaussian or a bell-shaped function; the function is denoted by $\mu(.)$. In general, the first layer of an adaptive network has an R-dimensional input vector such as $\mathbf{i}_{u,j}=\{i_{u,j1},...,i_{u,jR}\}$, with m square nodes for each dimension of the input vector; thus, this layer could consist of $m\times R$ square nodes. Equations 1 and 2 calculate a Gaussian or a bell-shaped function for the k^{th} node of the l^{th} dimension. The parameters c_{kl},σ_{kl} are the parameters of the Gaussian function, while a_{kl},c_{kl},b_{kl} are the parameters of the bell-shaped function.

$$\mu_{kl}(i_{u,jl})=\exp\left(-\frac{1}{2}\frac{(i_{u,jl}-c_{kl})^2}{\sigma_{kl}^2}\right) \tag{1}$$

$$\mu_{kl}(i_{u,jl})=\frac{1}{1+\left|\frac{i_{u,jl}-c_{kl}}{a_{kl}}\right|^{2b_{kl}}} \tag{2}$$

Layer 2: This layer has two circular nodes that are labelled with $\prod$. The outputs of the first layers are the inputs of the nodes of this layer, and the outputs of these nodes are the multiplication of their inputs. In a general case, this layer could have $k_a=m^R$ circular nodes, assuming that the $k_2{}^{th}$ node receives the outputs of the second square node of each dimension. The output of this node is calculated as equation (3), where R is the dimension of an input vector.

$$w_{k_2l}(\mathbf{i}_{u,j})=\prod_{l=1}^{R}\mu_{2l}(i_{u,jl}) \tag{3}$$

Layer 3: This layer has two circle nodes with the normalization functions; each node is labeled N. In the general case, this layer could have $k_a=m^R$ circle nodes. Assuming that the $k_2{}^{th}$ node receives all w_{k_2l} s from the previous layer, the output of this node is calculated as equation (4), where R is the dimension of the input vector.

$$\bar{w}_{k_2l}(\mathbf{i}_{u,j})=\frac{\prod_{l=1}^{R}\mu_{2l}(i_{u,jl})}{\sum_{k_2=1}^{k_a}w_{k_2l}(\mathbf{i}_{u,j})} \tag{4}$$

Layer 4: This layer has two square nodes. In the general case, the fourth layer could have $k_a=m^R$ square nodes; the function of $k_2{}^{th}$ node is calculated as f_{k_2}, equation (5). The parameters of the nodes of this layer are linear; each node receives the set of linear parameters as $\left\{\{q_{k_2l}\}_{l=1}^{R},q_{k_2R+1}\right\}$.

$$f_{k_2}(\mathbf{i}_{u,j})=\sum_{l=1}^{R}q_{k_2l}i_{u,jl}+q_{k_2(R+1)} \tag{5}$$

Layer 5: The fifth layer has a single node (circle) that calculates the summation of its input vector $\{f_{k_2}\}_{k_2=1}^{k_a}$. The output of the $k_2{}^{th}$ node is f_k, and has an important role in producing the final output F_j, equation (6).

$$F_j(\mathbf{i}_{u,j})=\sum_{k_2=1}^{k_a}\bar{w}_{k_2}f_{k_2} \tag{6}$$

B. *Locally Linear Model Tree (LoLiMoT)*

Locally linear neuro-fuzzy models or Takagi-Sugeno fuzzy models are another class of neuro-fuzzy models that are based on a divide-and-conquer strategy [19]. The model has a structure with two layers, the nodes of the first layer have two functions: a linear function that is referred to as a 'locally linear model' (LLM) and a nonlinear function that is referred to as a 'validity function'. Receiving an input vector as $\mathbf{i}_{u,j}=\{i_{u,j1},...,i_{u,jR}\}$, the output of the k^{th} node of the first layer is $y_{j_k}\varphi_{j_k}(\mathbf{i}_{u,j})$. Here y_{j_k} is the output of the LLM function that is given as equation (7).

$$\hat{y}_{j_k}=\omega_{k_0}+\omega_{k_1}i_{u,j_1}+\ldots+\omega_{k_R}i_{u,j_R} \tag{7}$$

The parameters $\left[\omega_{k_0},\omega_{k_1},...,\omega_{k_R}\right]$ are the linear parameters of the LLM of the k^{th} node. And $\varphi_k(.)$ is the validity function of the k th node and is calculated by equation (8). Where $\mu_k(\mathbf{i}_{u,j})$ is given as equation (9).

$$\varphi_k(\mathbf{i}_{u,j}) = \frac{\mu_k(\mathbf{i}_{u,j})}{\sum_{k=1}^{M}\mu_k(\mathbf{i}_{u,j})} \tag{8}$$

$$\mu_k(\mathbf{i}_{u,j}) = \prod_{l=1}^{R}\exp\left(-\frac{1}{2}\frac{(i_{u,jl} - c_{kl})^2}{\sigma_{kl}^2}\right) \tag{9}$$

The parameters c_{kl}, σ_{kl} are the parameters of the Gaussian function. The output of the model is calculated as the weighted summation of locally linear models [15] according to equation (10).

$$\hat{y}_j = \sum_{k=1}^{M}\hat{y}_{j_k}\varphi_{j_k}\left(\mathbf{i}_{u,j}\right) \tag{10}$$

The performance of these models depend on the division strategy; one good strategy is locally linear model tree algorithm (LoLiMoT) that was proposed by Nelles [19].

The local linear model tree algorithm (LoLiMoT) is an incremental heuristic algorithm used to optimize learning parameters, linear and nonlinear parameters which correspond to validity function and linear function, respectively. The algorithm consists of two loops: the first loop updates the nonlinear parameter, and the nested loop optimizes the linear parameters. The nested loop is based on the least square method to optimize the $M(R+1)$ linear learning parameters, where M represents the number of nodes of the first layer, and R is the dimension of the input vector. The following steps explain LoLiMoT [19]:

1. An initial structure with one neuron in the hidden layer (M=1).
2. The output vector of the model, $\hat{\mathbf{y}}$, is calculated based on the least square method [19].
3. The local cost function is calculated for each node. The worst locally linear model (LLM) that has maximum value for the lost function is selected to be divided into two nodes. The number of locally linear models is incremented by one.
4. The algorithm is stopped if the termination condition is satisfied; otherwise, it goes to Step 2 and continues.

The main preference of LoLiMoT is its low time complexity because of its linear growth with the number of fuzzy neurons. However, the curse of dimensionality [19] is a significant issue for this algorithm.

IV. Prediction of Geomagnetic Storms using the AE index

This paper utilizes normalized mean square error (NMSE) and the correlation coefficient $\rho_{y,\hat{y}}$, the average percentage error (AVE) given as equations (11) and (12) and (13) to assess the performance of the prediction models and provide comparable results with other studies. The parameters $\hat{\mathbf{y}}$ and $\mathbf{y}$ refer to the predicted values and desired targets, respectively. The parameter $\bar{\mathbf{y}}$ is the average of the desired targets.

$$\mathrm{NMSE} = \frac{\sum_{j=1}^{N_c}(y_j - \hat{y}_j)^2}{\sum_{j=1}^{N_c}(y_j - \bar{y}_j)^2} \tag{11}$$

$$\rho_{y,\hat{y}} = \frac{\mathrm{Cov}(y,\hat{y})}{\sigma_y\sigma_{\hat{y}}} \tag{12}$$

$$\mathrm{AVE} = \frac{1}{N_c}\sum_{j=1}^{N_c}\frac{y_j - \hat{y}_j}{y_j}\times 100 \tag{13}$$

In the first case, the AE index with one minute resolution of the 9th of March 1992 is predicted by BELFIS and BELRFS and the obtained results are compared with the results of ANFIS and LoLiMoT. For this purpose, the obtained AE index of the 7th March 1992 is utilized as training data to predict the AE index of the 9th of March 1992. Figure 5. shows the values of the AE index of 7th March 1992 that have been used as the training samples.

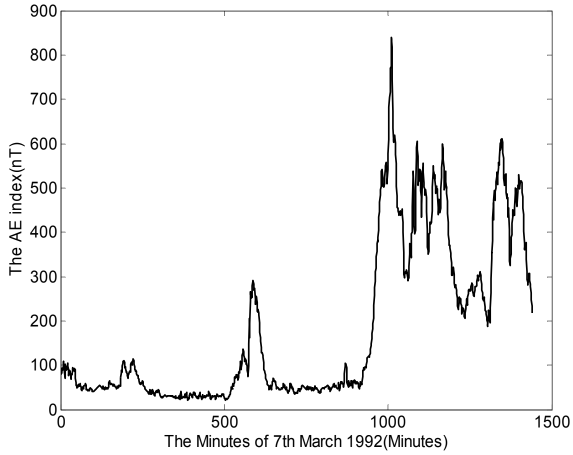
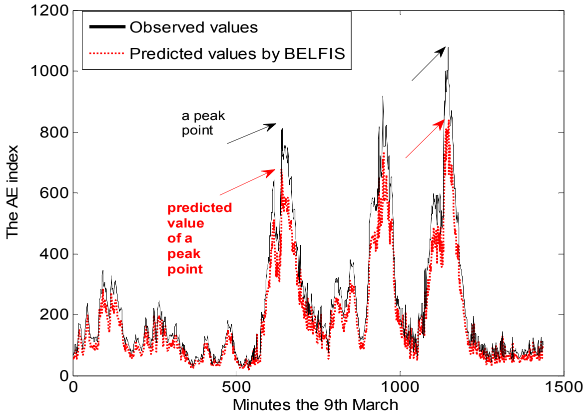


Fig. 5. The AE index of the 7th March 1992.

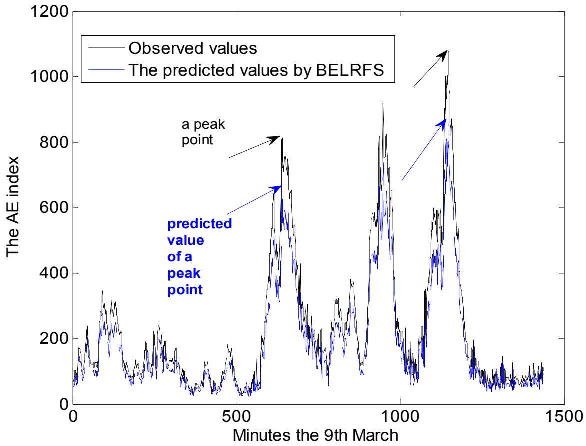
The graphs of Figure 6 describe the predicted values of the 9th March 1992 versus the observed values. As can be observed both BELRFS and BELFIS can predict the peak points of the AE index. Table II compares the obtained NMSE indices, correlation coefficient and the average percentage of error in this case. It can be seen that for the short-term prediction of the AE index, the NMSE indices of BELFIS and BELRFS are lower than the obtained NMSE indices of LoLiMoT and ANFIS.

TABLE II. COMPARISON BETWEEN METHODS FOR PREDICTING THE 9TH MARCH OF 1992 USING THE AE INDEX OF 7TH MARCH.

Method	**Different Methods**		
	NMSE	***Correlation***	***Average of error***
BELRFS	0.1169	0.9931	21.6
BELFIS	0.1161	0.9934	21.5
LoLiMoT	0.1229	0.9885	22.15
ANFIS	0.1188	0.9934	22.3



(a)



(b)

Fig. 6.The predicted values of the AE index of the 9th March 1992 (a).using the BELFIS; (b).using the BELRFS.

In the second case, the neuro fuzzy models have been trained by the AE index of seven days of March 1992 to predict the 9th of March 1992. Figure 7 shows the values of the AE index that have been used as the training samples.

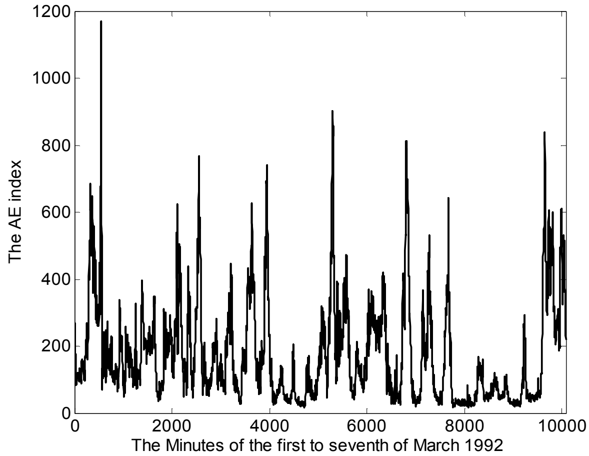
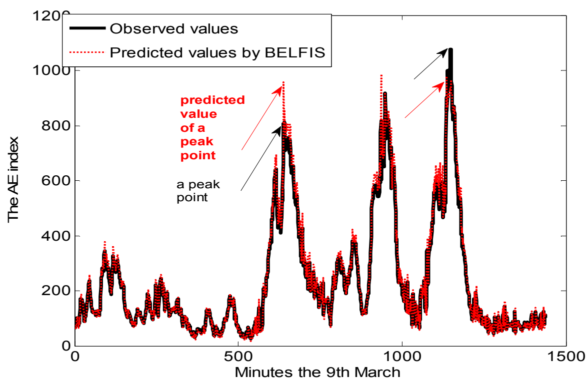


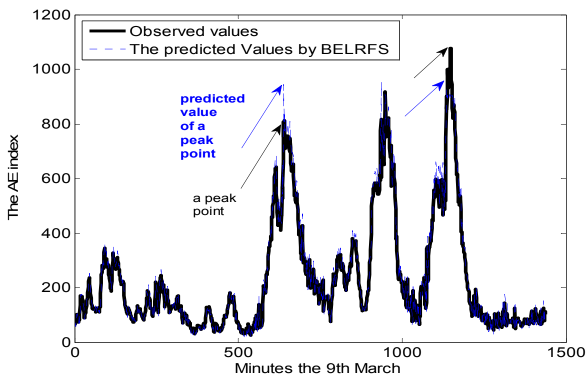
Fig. 7. The AE index of the first to seventh March 1992.

TABLE III. COMPARISON BETWEEN METHODS FOR PREDICTING THE 9TH MARCH OF 1992 USING THE AE INDEX OF FIRST TO SEVENTH DAY OF MARCH 1992.

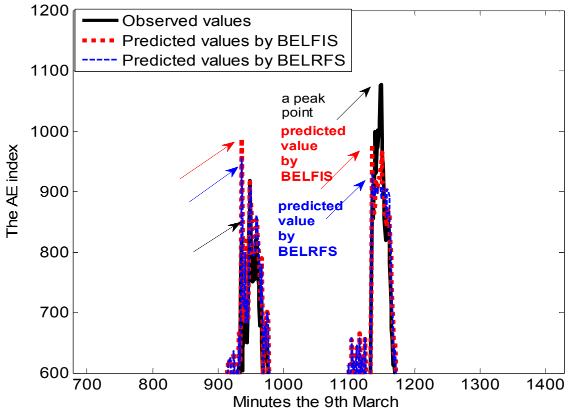
Method	**Different Methods**		
	NMSE	***Correlation***	***Average of error***
BELFIS	0.0125	0.9922	10.2
BELRFS	0.0224	0.9911	10.5
ANFIS	0.0215	0.9926	10.66
LoLiMoT	0.0273	0.9917	22.9



(a)



(b)



(c)

Fig. 8. The predicted values of the AE index of the 9th March 1992 (a).using the BELFIS; (b).using the BELRFS (c).the predicted values of the peak points using BELFIS and BELRFS.

The graphs of Figure 8 depict the predicted values of the 9th March 1992 versus the observed values. As can be seen, using a large number of training samples, the BELRFS and the BELFIS can accurately predict the peak points of the AE index. Table III lists the obtained NMSE indices, correlation coefficient and the average percentage of error in this case. It can be observed taht in this case the NMSE index of BELFIS is equal to 0.0125 which is less than the corresponding value in the case with a small number of training samples. As Table III indicates the obtained results of BELFIS and BELRFS are similar to the obtained results of ANFIS; however, they are slightly better than the obtained results of LoLiMoT.

V. Conclusion

In this paper, four neuro-fuzzy models were tested to predict the geomagnetic storms using the AE index. Two neuro fuzzy models, the BELRFS and the BELFIS, have been developed by taking inspiration from brain emotional learning and these are compared with ANFIS and LoLiMoT, two powerful neuro-fuzzy methods.

The obtained results indicate that the BELRFS and the BELFIS could predict the AE index with the same accurate results as the ANFIS and LoLiMoT models. Thus, the BELFIS and BELRFS can be utilized equally well as ANFIS and LoLiMoT to develop an alert system for geomagnetic storms. The obtained results are motivations for utilizing the BELFIS and BELRFS for other applications of neuro-fuzzy models such as stock market prediction.

In the future, we will combine Singular Spectrum Analysis (SSA) with the proposed models to increase the prediction accuracy for long-term prediction.

The next prediction applications would be other indices of geomagnetic storms such as the Disturbance storm time (Dst) index and the global geomagnetic storm index (Kp index). Moreover, we intend to examine BELFIS and BELRFS as a nonlinear identification method.

References

[1] http://www.ngdc.noaa.gov/stp/geomag/ae.html

[2] L. Xinlin Li, B.Luo and M.Temerin, Prediction of the Dst, AL, AU Indices Using SolarWind Parameters, *Geophysical Research Abstracts*,Vol. 15, EGU2013-3645, 2013.

[3] S. Kugblenu, S.Taguchi and T.Okuzawa, : Prediction of the geomagnetic storm associated Dst index using an artificial neural network algorithm", *Earth Planets Space*, 51, 307–313, 1999.

[4] Z. Voros, D. Jankovicova, "Neural network prediction of geomagnetic activity: a method using local Holder exponents," *Nonlinear Processes in Geophysics*, no. 9, pp. 425 - 433, 2002.

[5] J. Takalo, and J. Timonen, Neural network prediction of AE data, *Geophys. Res. Lett*., 24, 2403, 1997.

[6] W.K. Tobiska, D. Knipp, W. J. Burke, D. Bouwer, J. Bailey, D. Odstrcil, M. P. Hagan, J. Gannon, and B. R. Bowman (2013), The Anemomilos prediction methodology for Dst, *Space Weather*, 11, 490–508, doi:10.1002/swe.20094.

[7] A. Golipour, B. N.Araabi., C. Lucas, "Predicting Chaotic Time Series Using Neural and Neurofuzzy Models A Comparative Study,'' J. *Neural. Process Lett*., vol. 24, no. 3, pp. 217-239, 2006.

[8] H. Gleisner, H. Lundstedt, and P. Wintoft (1996), Predicting geomagnetic storms from solar wind data using time delay neural networks, *Ann. Geophys*., 14, 679.

[9] T. Babaie, R. Karimizandi, C. Lucas, "Learning based brain emotional intelligence as a new aspect for development of an alarm system,'' J. *Soft Computing*., vol. 9, issue 9, pp.857-873, 2008.

[10] M. Parsapoor, *Brain Emotional Learning-Inspired Models. Licentiate dissertation*, Halmstad: Halmstad University Press, 2014.

[11] J.Moren, C.Balkenius,"A computational model of emotional learning in the amygdala,'' *in From Animals to Animats*, MIT, Cambridge, 2000.

[12] M. Parsapoor, C. Lucas, S. Setayeshi, "Reinforcement Recurrent Fuzzy Rule Based System Based on Brain Emotional Learning Structure to Predict the Complexity Dynamic System", in Proc, *Digital Information Management* Third *IEEE International Joint Conf.* , pp. 22-32, 2008.

[13] M. Parsapoor, U. Bilstrup, "Brain Emotional Learning Based Fuzzy Inference System (BELFIS) for Solar Activity Forecasting," in Proc. *IEEE Int. Conf.* ICTAI 2012, 2012.

[14] M. Parsapoor, U. Bilstrup, "Brain Emotional Learning Based Fuzzy Inference System (Modified using Radial Basis Function)," *8th IEEE International Joint Conference for Digital InformationManagement*, 2013. DOI: 10.1109/ICDIM.2013.6693994.

[15] M. Parsapoor and U. Bilstrup, "Chaotic Time Series Prediction Using Brain Emotional Learning Based Recurrent Fuzzy System (BELRFS)," *International Journal of Reasoning-based Intelligent Systems*, *Vol. 5, No. 2, 2013.*

[16] M. Parsapoor, M, U. Bilstrup, "Neuro-fuzzy models, BELRFS and LoLiMoT, for prediction of chaotic time series," in Proc. *IEEE Int. Conf. INISTA*., pp.1-5, 2012.

[17] M. Parsapoor, U. Bilstrup, "An Emotional Learning-inspired Ensemble Classifier (ELiEC)," in Proc. *8th International Symposium Advances in Artificial Intelligence and Applications (AAIA'13),* At Kraków, Poland, 2013.

[18] R. Jang, C. Sun and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A computational approach to Learning and Machine Intelligence.*Upper Saddle River, NJ: Prentice Hall, 1997.

[19] O. Nelles, *Nonlinear System Identification*:*From classicical Approches to Neural Networks and Fuzzy Models.* Berlin, Germany: Springer-Verlag, 2001.