Using Fuzzy Cognitive Maps to predict the economic sustainability of Jordan Social Security

Ahmad Zyad Alghzawi[1,*], George Sammour[2], Koen Vanhoof[1]
[1] Department of Business Informatics, Hasselt University, Belgium
[2] Department of Management Information system, Princess Sumaya University for Technology, Jordan
(*) ahmad.alghzawi@uhasselt.be

Abstract. Fuzzy Cognitive Maps are emerging as an important new tool in economic modeling. This study investigates the use of fuzzy cognitive maps with their learning algorithms, based on genetic algorithms, for the purposes of economic prediction. The case study data are extracted from the Jordanian social sociality revenues and expanse for the last 120 months; The Real-Code genetic algorithm and structure optimization algorithm were chosen for their ability to select the most significant relationships between the concepts and to predict future development of the Jordanian social sociality revenues and expenses. Furthermore, fuzzy cognitive maps are able to calculate prediction errors accurately. The study shows that fuzzy cognitive maps models clearly predict the future of a complex financial system with incoming and outgoing flows. Consequently, this research confirms the benefits of fuzzy cognitive maps applications as a tool for scholarly researchers, economists and policy makers.

Index Terms — Fuzzy Cognitive Maps, prediction problems, modeling, learning algorithms, Jordanian social sociality.

1 Introduction

Fuzzy Cognitive Maps (FCMs) was developed by B. Kosko as a tool to model and analyze complex systems using causal relations. In social, economic and political systems [1] where soft knowledge is predominant they are particularly applicable since they allow configuration through the combination between fuzzy logic and neural networks [2] into causal maps. This combination also allows to describe and model the behavior of knowledge-based systems with interpretable. Features . This application has recently become popular among scholars. In this last application FCMs are information networks where graph nodes represent objects, states, concepts or entities [3].

Recently FCM became a tool to describe and model the behavior of knowledge-based systems with interpretable features. Accordingly, FCM’s are information networks where graph nodes represent objects, states, concepts or entities. These translate investigated systems and include a more delicate meaning to the problem domain [3].

The use FCMs as pattern classifiers is presented in [4]. An innovative method for forecasting artificial emotions and designing an effective decision system on the basis of FCMs is proposed in [5]. Time series is a sequence of observations, which related to regular intervals. The basic representation of time series is an unprocessed sequence of values quantifying give phenomena. FCMS is represented with weights matrix, this model matrix can be learned from available training data, which time series can model this data was described in [6, 7].

In practical applications to solve the financial sustainability of the Jordanian Social Security System (JSS) using FCMS with time series as a predictive method, finding the most significant concepts and connections plays an important role. It can be based on expert knowledge at all stages of analysis: designing the structure of the FCM model, determining the weights of the relationships and selecting input data. Supervised and population-based algorithms allow the automatic construction of FCMS on the basis of data selected by the experts or all available input data. Also some simplifications strategies by posteriori removing nodes and weights are presented [7].

The objective of this paper are :

- Investigate learning algorithms of FCM for Real-Coded Genetic Algorithm (RCGA) and Structure Optimization Genetic Algorithm (SOGA).
- Predict the future for revenue and expenses in JSS.

The remainder of this paper is organized as follow. Section 2 introduces the fuzzy cognitive maps. Section 3 introduces the proposed method for FCMs learning. Section 4 Simulation ISEMK Software Results. Section 5 predictive value between expenses and revenue in future. The last Section contains conclusions and discusses future research direction.

2 Fuzzy Cognitive Maps (FCMs)

Mathematically FCMs are used to demonstrate and to model the knowledge on the physical system under investigation in terms of concepts. FCMs can be defined by a 4-tuple (*C*,*W*,*A*,*f*) where $C = \{c_1, c_2, \dots, c_N\}$ represents a set of N concepts, furthermore the relation between pair of neurons Ci and Cj, is defined by a function $W: (c_i, c_j) \rightarrow W_{ij}$ this function normally take a value in the rang $W_{ij} \in [-1,1]$ this value allow to describing the interface between two map neurons by using causal relation. There are three type of causal relationships that reflect the weight between two concepts. According to [8], the three causal relation are:

A) If $W_{ij} > 0$ indicates that an increase (decrement) in the cause neuron c_i will produce an increment (decrement) of the effect neuron c_j with intensity $|W_{ij}|$.

B) If $W_{ij} < 0$ indicates that an increase (decrement) in the cause neuron c_i will produce a decrement (increment) of the effect neuron c_j with intensity $|W_{ij}|$.

If $W_{ij} = 0$ denotes the absence of relation between $c_i$and c_j.

Ideally, an FCM computes the activation value of map neuros c_i at each step by using the function $A:(c_i) \rightarrow A_i^{(t)}$. The degree of the concept plays a significant role through the FCM clarification, the higher value of concept, the stronger its impact over the system [9]. Then, a transfer function $f:\mathcal{R} \rightarrow I$ is used to keep the activation value of concepts in the interval $I = [0,1]$ or $I = [-1,1]$. Three widely used transfer function are described in [10] .

Concepts obtain value in the range between [0,1] which can be used in time series prediction. The value of concepts can be calculated according to the form :

$$X_i(t+1) = f\left(X_i(t) + \sum_{j \neq i} W_{j,i} . X_j(t)\right), \qquad (1)$$

where t the discrete time steps, $t = 0, 1, 2, \ldots, T, T$ is the maximal number of iterations, $X_i(t)$ is the value of the ith concept, $i = 1, 2, \ldots, n, n$ is the number of concepts, $f(x)$ is a transformation function. In many examples, the sigmoid function, Eq. (2), is chosen for transformation:

$$f(x) = \frac{1}{1 + e^{-cx}}, \qquad (2)$$

where $c > 0$ is a parameter.

FCMs can be automatically constructed with the use of supervised and population-based learning algorithms. In the next section, selected methods of FCMs learning are described.

3 Fuzzy Cognitive Maps Learning

The aim of the FCM learning process is to estimate the weights matrix W. In Real-Coded Genetic Algorithm and the Structure Optimization Genetic Algorithm for fuzzy cognitive maps learning is analyzed. And the performance of the developed approach is compared. Description of these methods is presented below.

The objective of this section as summarized as follow:

1. To explore learning algorithms of FCM for multivariate time series modeling and prediction.
2. To execute experimental evaluation of the proposed methods.

3.1 Real-Coded Genetic Algorithm

The Real-Coded Genetic Algorithm defines each chromosome as a floating-point vector. This is mathematically expressed as follows [11]:

$$W' = \left[W_{1,2}, \dots, W_{1,n}, W_{2,1}, W_{2,3}, \dots, W_{2,n}, \dots, W_{n,n-1}\right]^T, \quad (3)$$

Where $W_{j,i}$ is the weight of the connection between the $j-th$ and the $i-th$ concept.

Each chromosome in the population is decoded into a candidate FCM and its quality is evaluated on the basis of a fitness function according to the objective. [12].

The fitness function can be described as follows(4):

$$fitness_p = \frac{1}{a.j(l)+1}, \quad (4)$$

where a is a parameter, $a > 0$, p is the number of chromosome, $p = 1, \dots, P$, P is the population size, l is the number of population, $l = 1, \dots, L$, L is the maximum number of populations, $J(l)$ is the learning error function, described as follows:

$$J(l) = \frac{1}{(T-1)n}\sum_{t=1}^{T-1}\sum_{i=1}^{n}(Z_i(t) - X_i(t))^2, \quad (5)$$

Where t is discrete time of learning, T is the number of the learning records, $Z(t) = [Z_1(t), \dots, Z_n(t)]^T$ is the desired FCM response for the initial vector $Z(t-1)$, $X(t) = [X_1(t), \dots, X_n(t)]^T$ is the FCM response for the initial vector $Z(t-1)$, n is the number of output concepts, $X_i(t)$ is the value of the ith output concept, $Z_i(t)$ is the reference value of the ith output concept.

A probability of reproduction is assigned to each population. According to this assigned probability function, parents are selected and new population of chromosomes are generated. Chromosomes with above average fitness tend to reproduce more than those with below average fitness [13]. The basic operator of the selection process is a roulette wheel method. For each chromosome in the population the probability of occurrence of a copy of this chromosome in the next population can be calculated according to the formula [14]:

$$P(p) = \frac{fitness_p(J_l)}{\sum_{i=1}^{P} fitness_i(J_l)}, \quad (6)$$

Where p is the number of the chromosome, P is the population size.The mutation operator modifies elements of a selected chromosome with a probability defined by the mutation probability P_m. The use of mutation prevents the premature convergence of the genetic algorithm to suboptimal solutions [13, 14]. In the analysis Mühlenbein’s and random mutation were used. To ensure the survival of the best chromosome in the population, an elite strategy was applied. It retains the best chromosome in the population [13].

The learning process stops when the maximum number of populations L is reached or the condition (7) is met.

$$fitness_{best}(J_l) > fitness_{max} \ , \tag{7}$$

where $fitness_{best}(J_l)$ is the fitness function value for the best chromosome, $fitness_{max}$ is a parameter.

3.2 Structure Optimization Genetic Algorithm

In this research, Structure Optimization Genetic Algorithm is used as a learning method. It allows to selecting the most important concepts and connections between them for the objective of prediction tasks. SOGA defines each chromosome as a floating-point vector type (2) and a binary vector expressed as follows:

$$C' = [C_1, C_2, \dots . C_n]^T, \tag{8}$$

where C_i is the information about including the i-th concept to the candidate FCM model, whereas $C_i = 1$ means that the candidate FCM model contains the $i-th$ concept, $C_i = 0$ means that the candidate FCM model does not contain the $i-th$ concept.

The quality of each population is calculated based on an original fitness function, described as follows [13]:

$$fitness_p(J_l') - \frac{1}{a.J'(l) + 1}, \tag{8}$$

where a is a parameter, $a > 0$, p is the number of the chromosome, l is the number of population, $l = 1, \dots, L, L$ is the maximum number of populations, $J'(l)$ is the new learning error function with an additional penalty for the high complexity of FCM. It is understood as a large number of concepts and non-zero connections between them [13]:

$$J'(l) = J(l) + b_1.\frac{n_r}{n^2}.J(l) + b_2.\frac{n_c}{n}.J(l), \tag{10}$$

where t is discrete time of learning, T is the number of the learning records, b_1, b_2 are the parameters, $b_1 > 0, b_2 > 0$, n_r is the number of the non-zero weights of connections, n_c is the number of the concepts in the candidate FCM model, n is the number of all the possible concepts, $J(l)$ is the learning error function (10).

4 Simulation ISEMK Software Results

ISEMK is a computer software tool for modeling decision support systems based on FCMs [13].

The aim of using the tool was to acquire a better understanding of the (JSS) by using the available monthly data and to investigate the future evolution based on the current value. Both the expenses and the revenues of the system were investigated.

This study uses both the Structure Optimization Genetic Algorithm and the Real-Coded Genetic Algorithm, based on an analysis of historical data taken from the Jordanian social security System. The dataset contained the revenues and expenses basis of the system over the last 120 months and contained the following fields:

- **Revenues**
  - X1:Aging subscription
  - X2:Work related injuries
  - X3:Maternity insurance
  - X4:Years earlier service
  - X5:Optional Subscriptions
  - X6:Different revenue
  - X7:Stamps
- **Expenses**
  - X1:Pensions
  - X2:One time
  - X3:Work injuries
  - X4:Maternity insurance

The data were related to a ten-year historical log from 2006 until 2015 on. The available data were normalized to the [0,1] range in order to use the FCM model in the same way as we explain before when depicting the normalization process [1].

The revenue and expenses dataset were both divided into two subsets: learning (114 records) and testing (6 records) and also the same to expenses dataset.

The quality of the fitted prediction model are evaluated based on four different quality measures, as depicted in Eq. (11)-(14). lowest Mean Square Error (MSE) (Eq. 11).

$$MSE = \frac{1}{n(T-1)} \sum_{t=1}^{T-1} \sum_{i=1}^{n} \left(Z_i(t) - X_i(t)\right)^2, \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n(T-1)} \sum_{t=1}^{T-1} \sum_{i=1}^{n} \left(Z_i(t) - X_i(t)\right)^2}, \tag{12}$$

$$MAPE = \frac{1}{n(T-1)} \sum_{t=1}^{T-1} \sum_{i=1}^{n} \left| \frac{Z_i(t) - X_i(t)}{Z_i(t)} \right|, \tag{13}$$

$$MAE = \frac{1}{n(T-1)} \sum_{t=1}^{T-1} \sum_{i=1}^{n} |Z_i(t) - X_i(t)|, \tag{14}$$

where t is time of testing, $t = 0, 1, \ldots, T-1$, T is the number of the test records, $Z(t) = [Z_1(t), \ldots, Z_n(t)]^T$ is the desired FCM response for the initial vector $Z(t-1)$, $X(t) = [X_1(t), \ldots, X_n(t)]^T$ is the FCM response for the initial vector $Z(t-1)$, $X_i(t)$ is the value of the ith output concept, $Z_i(t)$ is the reference value of the ith output concept, n is the number of the output concepts..

Table 2 shows the weights matrix for the map in Fig1 which presents the structure of the FCM learned by using the **SOGA**. Table 3 shows the weights matrix for the map in Fig2 which presents the structure of the FCM learned by using the **RCGA** .

The model validation is assessed by means of the MSE measure. The data set were split into training and test sets. The training set is used to generate the models while the test set is fed into the model for prediction. This allows for measuring the performance of the model using an unseen set of data (test set). [7, 12].

Table 1: Revenues dataset analysis results of the RCGA, SOGA.

Method	Learning parametrs	MSE	RMSE	MAE	MAPE
RCGA	P=500, L = 200, a = 10, b_1=0.1, b_2=0.01, $fitness_{max}$ = 0.999, c = 13, Mühlenbein's mutation, P_c = 0.5, P_m = 0.2	0.058	0.241	0.207	0.31
RCGA	P=500, L = 200, a = 10, b_1=0.1, b_2=0.01, $fitness_{max}$ = 0.999, c = 13, Mühlenbein's mutation, P_c = 0.5, P_m = 0.2	0.051	0.227	0.19	0.285
SOGA	P=500, L = 200, a = 10, b_1=0.1, b_2=0.01, $fitness_{max}$ = 0.999, c = 13, Mühlenbein's mutation, P_c = 0.5, P_m = 0.2	**0.018**	**0.135**	**0.112**	**0.168**
SOGA	P=500, L = 200, a = 10, b_1=0.1, b_2=0.01, $fitness_{max}$ = 0.999, c = 13, Mühlenbein's mutation, P_c = 0.5, P_m = 0.2	0.069	0.263	0.234	0.335

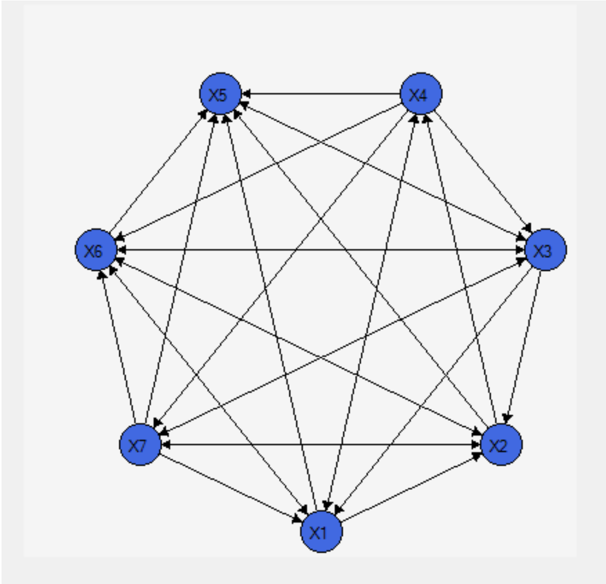

Fig1 The structure of revenue FCM learned with use of SOGA.

Table 2: Revenues weights matrix for learning map with using SOGA.

	X1	X2	X3	X4	X5	X6	X7
X1	0	0.41	0	0.16	0.59	-0.34	0
X2	0	0	0	-0.72	-0.62	0.06	-0.58
X3	0.95	0.66	0	0	0.65	-0.35	0.13
X4	-0.52	0	-0.87	0	-0.6	-0.2	-0.64
X5	0	0	0.93	0	0	0	0
X6	-0.16	-0.16	-0.49	0	-0.455	0	0
X7	-0.79	-0.9	-0.43	0	-0.56	-0.4	0

Table 3: Expenses dataset analysis results of the RCGA, SOGA.

Method	Learning parametrs	MSE	RMSE	MAE	MAPE
RCGA	P=500, L = 200, a = 10, b_1=0.1, b_2=0.01, $fitness_{max}$ = 0.999, c = 13, Mühlenbein's mutation, P_c = 0.5, P_m = 0.2	0.109	0.331	0.277	0.499
RCGA	P=500, L = 200, a = 10, b_1=0.1, b_2=0.01, $fitness_{max}$ = 0.999, c = 13, Mühlenbein's mutation, P_c = 0.5, P_m = 0.2	**0.012**	**0.11**	**0.083**	**0.129**
SOGA	P=500, L = 200, a = 10, b_1=0.1, b_2=0.01, $fitness_{max}$ = 0.999, c = 13, Mühlenbein's mutation, P_c =	0.087	0.295	0.247	0.445

	0.5, P_m = 0.2				
SOGA	P=500, L = 200, a = 10, b_1=0.1, b_2=0.01, $fitness_{max}$ = 0.999, c = 13, Mühlenbein's mutation, P_c = 0.5, P_m = 0.2	0.104	0.322	0.268	0.462

In the revenue case, the lowest error measures for revenue data computed by SOGA are MSE=0.018, $RMSE$=0.135, MAE=0.112 and MAPE=0.168, whereas for the error measures for expenses data computed by RCGA are MSE=0.012, $RMSE$=0.11, MAE=0.083 and MAPE=0.129.

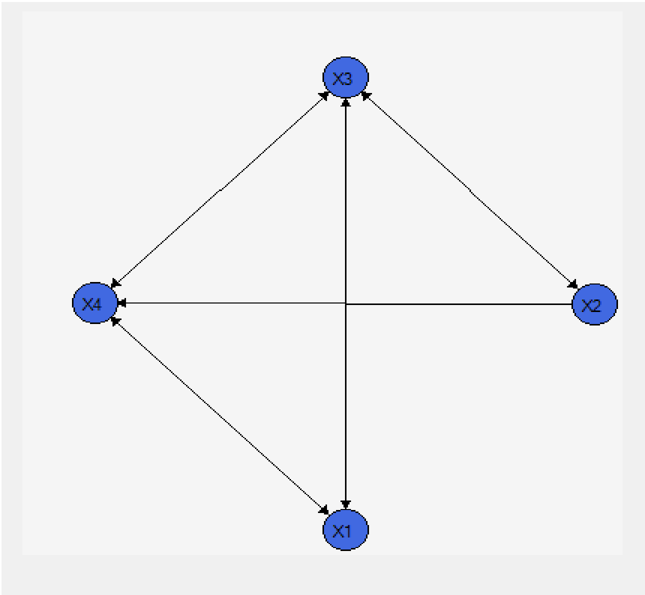

Fig.2 The structure of expenses FCM learned with use of SOGA.

Table 4: Expenses weights matrix for learning map with using SOGA.

	X1	X2	X3	X4
X1	0	0	-0.6338	0.26
X2	0	0	-0.91	-0.77
X3	-0.96	-0.88	0	-0.8
X4	0.16	0	0.68	0

5 Experimental Results

In this section, we introduce a study case concerning to the forecasting social security revenues and expenses in Jordan. The aim of this study is focused on predicting the revenue and expenses values and understanding the underlying interrelation between concepts; the latter is the main motivation to use cognitive mapping models. The dataset contains the last 120 months (from 2006 until 2015) data of revenues and expenses which we received from the Jordanian social security organisation, and comprises the seven fields in revenue and four fields in expenses after that we comprises the total revenue with total expenses (i.e., map concepts) that are listed below, we have allowed the algorithm to indicate a relationship value between the neurons to fluctuate between -1 and 1. We also presented a section on the most important task concepts and connections between them allowing to learn from the use of fuzzy cognitive maps; They suggest that the expenses will not be equal to the revenues in the near future and always revenues will be higher than the expenses. The difference will however not be large.

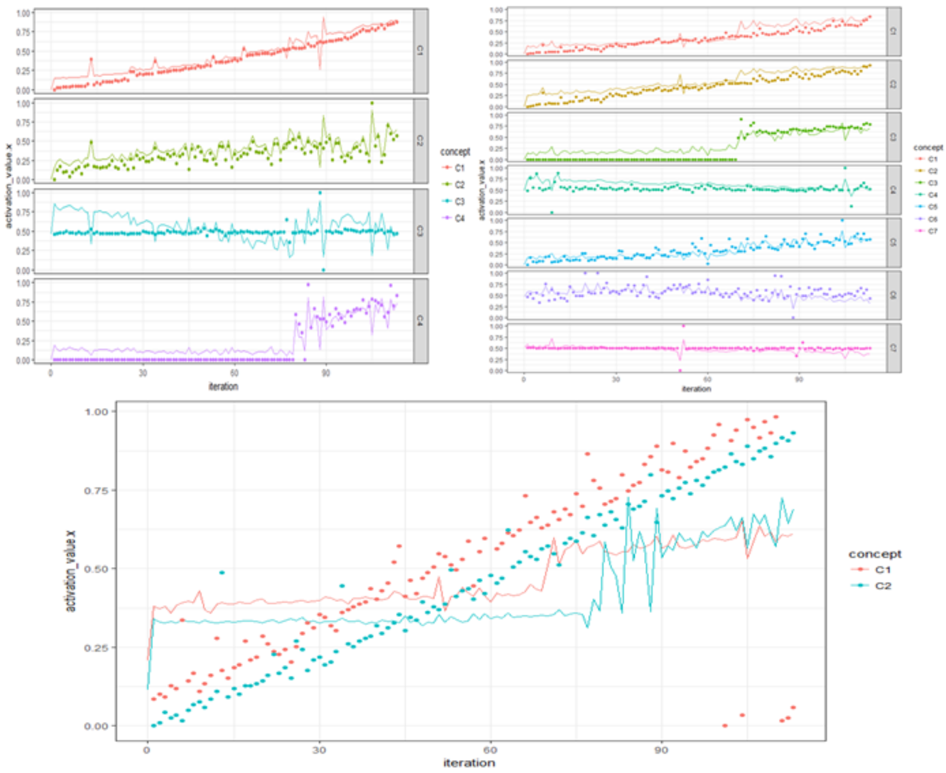

Fig.3 Revenue and expenses prediction

6 conclusions

In this paper, we proposed an FCM-based model to investigate the factors affecting revenues and expenses of the social security system in Jordan. The proposed Real-Code Genetic Algorithm and Structure Optimization Genetic Algorithm are described together with well-known methods to learn the map structure and the lessons we can draw for learning to use FCM's in a better way.

We decided to base our predictions on the use of FCM's because of the multiple advantages they present. FSMs are capable of modeling complex systems in a very transparent way due to the graphical representation in neural networks. Moreover, FSMs allow for the introduction of expert knowledge in the prediction models of the behavior of such complex systems. Finally, FSMs predict fairly accurately. In our case, the predictions using FCM's in Figure 3 allow us to conclude that the expenses of the Jordanian Social Security System will not be equal revenues in the close future. Decision making on the basis of these results will also become more accurate as the use of these models can also map the consequences of eventual policy changes by comparing it to our results. Thus FCM's are a useful policy making tool in complex decision making situations.

The RCGA and SOCG algorithms seems to be promising and effective methods for modeling complex decision support systems based on fuzzy cognitive maps. The analysis of the use of the revenue and expense data did predict the moment at which the expenses will be equal to the revenues, thus indicating the critical points in the actual JSS. Our future research will further focus on interpreting the JSS data in terms of sustainability of the Jordanian Social Security System (JSS) in terms of expenses and revenues.

Acknowledgments. We would like to thank our colleague Frank Vanhoenshoven for his valuable support and constructive comments.

References

1. Kosko, B., *Fuzzy cognitive maps.* International Journal of Man-Machine Studies, 1986. **24**(1): p. 65-75.
2. Groumpos, P.P. and C.D. Stylios, *Modelling supervisory control systems using fuzzy cognitive maps.* Chaos, Solitons & Fractals, 2000. **11**(1–3): p. 329-336.
3. Nápoles, G., et al., *On the convergence of sigmoid Fuzzy Cognitive Maps.* Information Sciences, 2016. **349–350**: p. 154-171.
4. Papakostas, G.A., et al., *Towards Hebbian learning of Fuzzy Cognitive Maps in pattern classification problems.* Expert Systems with Applications, 2012. **39**(12): p. 10620-10629.

5. Salmeron, J.L., *Fuzzy cognitive maps for artificial emotions forecasting.* Applied Soft Computing, 2012. **12**(12): p. 3704-3710.
6. Lu, W., et al., *The modeling of time series based on fuzzy information granules.* Expert Systems with Applications, 2014. **41**(8): p. 3799-3808.
7. Homenda, W., A. Jastrzebska, and W. Pedrycz, *Time Series Modeling with Fuzzy Cognitive Maps: Simplification Strategies*, in *Computer Information Systems and Industrial Management: 13th IFIP TC8 International Conference, CISIM 2014, Ho Chi Minh City, Vietnam, November 5-7, 2014. Proceedings*, K. Saeed and V. Snášel, Editors. 2014, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 409-420.
8. Ketipi, M.K., et al., *A flexible nonlinear approach to represent cause–effect relationships in FCMs.* Applied Soft Computing, 2012. **12**(12): p. 3757-3770.
9. Nápoles, G., et al., *Learning and Convergence of Fuzzy Cognitive Maps Used in Pattern Recognition.* Neural Processing Letters, 2016: p. 1-14.
10. Bueno, S. and J.L. Salmeron, *Benchmarking main activation functions in fuzzy cognitive maps.* Expert Systems with Applications, 2009. **36**(3, Part 1): p. 5221-5229.
11. Stach, W., et al., *Genetic learning of fuzzy cognitive maps.* Fuzzy Sets and Systems, 2005. **153**(3): p. 371-401.
12. Homenda, W., A. Jastrzebska, and W. Pedrycz. *Modeling time series with fuzzy cognitive maps*. in *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. 2014.
13. Pocz, K., et al. *Learning fuzzy cognitive maps using Structure Optimization Genetic Algorithm*. in *Computer Science and Information Systems (FedCSIS), 2015 Federated Conference on*. 2015.
14. Herrera, F., M. Lozano, and J.L. Verdegay, *Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis.* Artificial Intelligence Review, 1998. **12**(4): p. 265-319.