ATLANTIS PRESS
Atlantis Studies in Uncertainty Modelling, volume 3
Joint Proceedings of the 19th World Congress of the International Fuzzy Systems Association (IFSA), the 12th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), and the 11th International Summer School on Aggregation Operators (AGOP)

Linear Regression Approach to Fuzzy Cognitive Maps with History Data

Vesa A. Niskanen
Dept. of Economics and Management, University of Helsinki, PO Box 27, 00014 Helsinki, Finland,
vesa.a.niskanen@helsinki.fi

Abstract

Methods of linear regression analysis are applied to fuzzy cognitive map construction according to history data from the standpoint of quantitative human sciences. When the linearized version of history data is also used, this construction may be reduced to ordinary linear regression analysis. This linearization applies the inverted transformation functions of the fuzzy cognitive maps. Our approach will avoid subjective reasoning and interpretation on these model outcomes by relying on the objective and well-justified statistical theories instead.

Keywords: Fuzzy cognitive maps, History data, Linear regression analysis, Quantitative human sciences.

1 Introduction

Fuzzy cognitive maps (FCM) seem applicable to complex system modelling. They comprise a set of concepts and their interrelationships and we may examine how these concepts will vary in a given time interval. These trends may be collected into the data matrix known as the history data [2, 3, 10, 18-22].

Conversely, if only the history data is available, our aim is to construct such FCM that will yield this data as well as possible. This approach will thus apply optimization methods when the appropriate FCM parameters are specified. Below we will examine this problem-setting and focus on the numeric FCMs from the standpoint of the quantitative human sciences [11, 12].

The prevailing FCM construction methods often seem to apply techniques that are analogous to those of the neural networks, and this approach will arouse two principal problems. First, these FCMs are not necessarily stable systems with respect to their parameters because when repeating the parameter optimization many times, we may obtain distinct parameter values in each run [7, 8, 17].

Second, we will lack a sound theoretical basis when we make the interpretations on our parameter values and thus we often perform subjective reasoning instead.

The Author has studied FCMs from the statistical standpoint for providing a less subjective and still a consistent basis on the FCM interpretation, and this approach is central in the quantitative human sciences [5, 6, 14-16]. Hence, below we will suggest resolutions to the foregoing problems by applying certain statistical methods, in particular the linear regression analysis. In this manner, we should obtain stable parameter values and provide a sound theoretical basis for our FCM construction and interpretation.

Section 2 presents the basic ideas for constructing an FCM. Section 3 introduces our resolutions at a general level. Section 4 provides a concrete example. Section 5 concludes our examination.

2 Fuzzy Cognitive Maps

The prevailing FCM constructions stem from the neural networks [3, 7, 10]. Consider a set of n concepts,

$$C_i,\ 0 \leq C_i \leq 1,\ i = 1,2, \ldots, n,$$

and their $n \times n$ connection matrix, M, with the weights

$$W_{ij},\ -1 \leq W_{ij} \leq 1,\ i, j = 1,2, \ldots, n.$$

These weights will denote the intensities of the interconnections or interrelationships of the concepts C_i. The driver and target concepts are the row and column concepts in M, respectively (Table 2.1.). Hence, in fact, we operate with directed graphs.

In the FCM simulation we will first assign the initial values to the concepts C_i in a state vector $V_0=(C_1,\ldots,C_n)$. Then we may calculate their updated values in a given time interval, t=1, 2, …, m-1, by using the weights in the matrix M,

$V_{t+1} = F(V_t*M)$, t = 0,1, 2, … ,m-1 (2.1.)

in which V_t*M denotes the matrix product of V_t and M and *F* is a transformation function [10].

Since the values in vector $U_t=V_t*M$ should usually range from 0 to 1, in each iteration we should use such nonlinear transformation function as

$F(U_t) = 1 / (1 + exp(-lambda \cdot U_t))$, lambda $\geq$ 1, t = 0, 1, …,m-1 (2.2.)

in which *exp* is the exponential function should also be applied (Fig. 2.1.) [3, 10, 13, 18-22].

If these values may range from -1 to 1, we may define

$F(U_t)=\tanh(lambda \cdot U_t)= (exp(lambda \cdot U_t) - exp(-lambda \cdot U_t)) / (exp(lambda \cdot U_t) + exp(-lambda \cdot U_t))$, lambda > 0. (2.3.)

Below we will apply (2.2.). In statistics, this approach is analogous to the logistic, multinomial logistic and Cox's regression models [11, 12, 14, 15].

Hence, by iterating the foregoing procedure *m-1* times to each target concept, we will obtain a time series of concept values known as the $m \times n$ history data matrix in which the first row contains the initial concept values. We may thus examine the trends of our concepts in a given time interval.

If, on the contrary, only history data are available and we should construct the corresponding FCM, our task is to create the appropriate connection matrix and lambda value. This will be our principal aim below.

	C1	**C2**	**…**	**Cn**
C1	W_{11}	W_{12}	…	W_{1n}
C2	W_{21}	W_{22}	…	W_{2n}
…	…	…	…	…
Cn	W_{n1}	W_{n2}	…	W_{nn}

Table 2.1: Connection Matrix Used in Fuzzy Cognitive Map (drivers in rows, targets in columns).

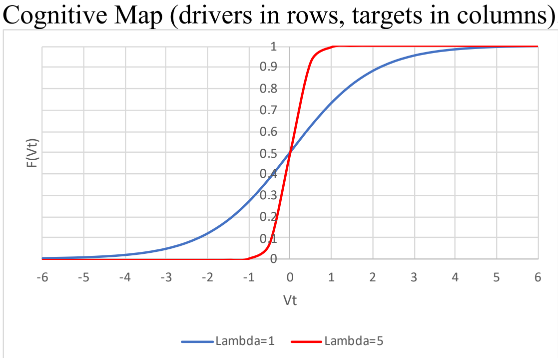

Figure 2.1: Examples on Transformation Functions for FCM with (2.2.).

The prevailing methods usually optimize both the weights of the connection matrix and the lambda value according to the history data, and the obtained FCM should yield the trends similar to this data. This method is analogous to those applied to the neural networks [3, 7, 8, 10, 17].

However, this approach seems problematic from the standpoint of model stability because we will have too much dispersion in our optimized parameters. Hence, when repeating this optimization many times with our history data, in each calculation we may obtain distinct parameter values for the connection matrix and lambda. In the models of the human sciences we nevertheless aim at unique parameter values.

Within the prevailing approaches, the interpretation and significance estimation of the connection weights are also problematic because they often base on subjective reasoning. This problem leads to ad hoc or even arbitrary conclusions in FCM examination and thus to the undesirable outcomes within the quantitative human sciences.

We will aim at resolving these problems. Our alternative method will principally apply linear regression analysis when assigning the weights to the connection matrix, and this approach will be considered below.

3 The Inverse Function Approach

When constructing an FCM according to the history data, we will aim at creating its connection matrix in such a way that this matrix will reveal the interconnections between our concepts in a unique, objective and theoretically well-justified manner.

Hence, below we will suggest a method that reduces the connection matrix examination to the ordinary linear regression analysis, LRA. We may thus utilize the theoretical outcomes of this analysis. Bearing in mind that in the prevailing FCM simulations we will first apply the linear functions with (2.1.) and then the nonlinear transformation function (2.2.), in our case, in turn, the linear modelling of (2.1.) is essentially applied.

As a preliminary work, in the human-scientific approach we may regard our FCM concepts as being the normally distributed random variables and thus we may estimate their central tendencies and measures of dispersion, among others. Our connection weights, in turn, may usually be treated as uniformly distributed variables. When applying (2.1.), their weighted sums will approach normal distributions according to the central limit theorem [11, 12, 16].

Hence, if our original history data are not ranging from 0 to 1, we may usually apply an appropriate transformation for them. The prevailing FCM methods seem to apply the formula

$$0 \leq |x - \min| / |\max - \min| \leq 1,$$

when x is the original observation. However, the sample minima and maxima are usually quite unstable values.

In the human sciences a widely-used method is to calculate first the standard scores of the data, *zs*, and then the scores |*zs*|>3 are excluded as the outliers [11, 12]. Finally, we may apply such modification as (*zs*+3)/6 for having our data ranging from 0 to 1. This method seems more plausible because the sample means and standard deviations are expected to be more stable values [11, 12].

Our FCM approach, in turn, will construct an LRA model without a constant term for each target concept according to the given $m \times n$ history data matrix, H, with the initial concept values and their *m-1* consecutive values. Unlike in the ordinary FCM construction, in which case we are using directly this data in the parameter optimization, we, in turn, will also use the corresponding linearized data. Hence, we will also formulate the inverse images of our data for obtaining the corresponding linear data.

Hence, we will apply the inverse functions of (2.2.) or (2.3.) for obtaining the corresponding linear values of our history data into matrix HL,

$$F^{-1}(y) = \ln(y / (1-y)) / \text{lambda},\ 0 < y < 1,\ \text{lambda} \geq 1, \quad (3.1.)$$

in which y = H(i,j), i=2, 3, …,m, j=1, 2, …,n and *ln* is the natural logarithm. Since in (3.1.)

$$0 < y < 1 \text{ and } -n < F^{-1}(y) < n,$$

we will establish that

$$F^{-1}(0) = -n \text{ and } F^{-1}(1) = n$$

when n denotes the number of drivers. Then, we may also operate with the original history data values 0 and 1 (Fig. 3.1).

Alternatively, for (2.3.) we may apply

$$F^{-1}(y) = \tanh^{-1}(y) / \text{lambda} = 0.5 \cdot \ln((1+y) / (1-y)) / \text{lambda},\ -1 < y < 1. \quad (3.2.)$$

We will only focus on (3.1.) below and this is not applied to our initial values, H(1,•) in our original history data. Hence, HL(1,•)=H(1,•). Since the lambda value is now unknown to us, it should be optimized or estimated when constructing an appropriate FCM.

Thanks for (3.1.), we will also obtain the corresponding linearized $m \times n$ history data matrix, HL, and that is used for our target concepts. Bearing also in mind, that the updated concept values will base on the concept values in the preceding iteration step, we should use a reorganized history data for our LRA.

Hence, if the concept C_k is our response variable (target concept), its data will only include the rows 2 to m in the column k in HL. For its predictors (drivers), C_1 to C_n, we will use the rows 1 to m-1 in the columns 1 to n in the original history data matrix, H. If the connection matrix has zeros in its diagonal, i.e., no self-loops are used, the predictor concept C_k in the column k in H may be excluded (Table 3.1.).

When using our reorganized history data matrix with an appropriate lambda value for the LRA model construction without the constant term, its estimated regression coefficients will be the weights in the FCM connection matrix, M, and we may also utilize such LRA outputs as their confidence intervals, standardized beta values, statistical significances with the t-tests and possible multicollinearity measures. This simple linear method will also yield unique weights, and thus this type of stability of our FCM weights is attained. Naturally, we may also examine such general LRA goodness criteria as the Rsquares and F-tests.

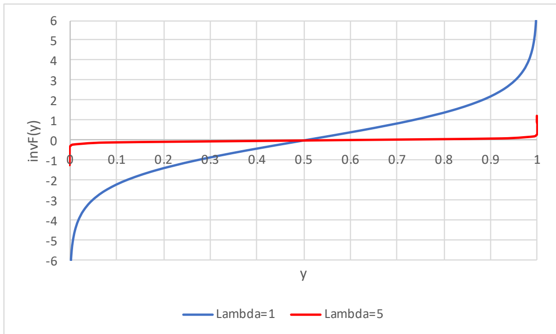

Figure 3.1: Examples on Inverse Transformation Functions of (2.2.).

	C1	C2	…	Ck	…	Cn
1				Not used		
2						
3						
…						
m-1						
m	Not used	Not used	Not used		Not used	Not used

Table 3.1: The History Data Used for an LRA Model when C_k Is the Response Variable and the Other Concepts Are Predictors (white cells: from original history data matrix for the predictors, grey cells: from linearized history data matrix for the response).

This method is applied one by one to each target concept. Finally, we may create our FCM simulation outputs with our new connection matrix by applying (2.1.) and (2.2), and these data are expected to be identical to those of the original history data, H. Below we will provide a concrete example.

4 Application Example

Our concrete application example is based on the studies on Parkinson's disease in [1]. Their medical FCM model included these eight concepts,

C1: Body bradykinesia (slowness of movement)

C2: Rigidity (stiffness of muscles)

C3: Posture

C4: Movement of upper limbs

C5: Gait

C6: Tremor

C7: Self-care

C8: Stage of Parkinson's disease (their principal target)

Their interconnections, that also included self-loops, are presented in Fig. 4.1. and Table 4.1. In this FCM concept C8 is the principal target concept.

If we apply their FCM with lambda=1 and initial concept values with nine iterations, we will obtain their history data presented in Table 4.2. and Fig. 4.2. [1]. These data will be used in our analysis at the outset, and we should be able to construct their FCM.

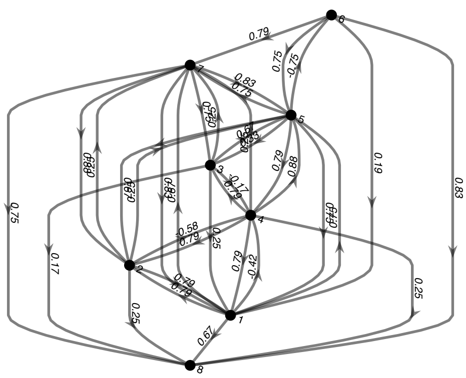

Figure 4.1: The Graph Depicting Parkinson's Disease Model.

	C1	C2	C3	C4	C5	C6	C7	C8
C1	1	0.8	0	-0.4	0.8	0	0.7	0.7
C2	0.8	1	0	-0.6	0.3	0	0.3	0.3
C3	0.3	0	1	-0.2	0.7	0	0.3	0.2
C4	0.8	0.8	0.8	1	0.9	0	0.3	0.3
C5	0.8	0.7	0.3	0.8	1	-0.8	0.8	0
C6	0.2	0	0	0	0.8	1	0.8	0.8
C7	0.8	0.9	0.8	0	0.8	0	1	0.8
C8	0	0	0	0	0	0	0	1

Table 4.1: The Connection Matrix Used in Parkinson's Disease FCM.

For the sake of comparison, the prevailing method for specifying the connection matrix and lambda with the data in Table 4.2. was first carried out by using Matlab's^{TM} Genetic Algorithm Toolbox and then fine-tuning these outputs with Levenberg-Marquardt optimization (Table 4.3., lambda=1.79). However, when repeating this procedure, alternative weights will be obtained and thus the instability problem is encountered. In addition, the interpretations of these weights are essentially based on subjective reasoning.

Our inverse method, in turn, is also using the corresponding linearized history data matrix for our target concepts by applying (3.1.) (Table 4.4.). Hence, for example, if we will construct the LRA model with Parkinson's disease (concept C8) as the response variable, the concepts C1-C8 will be its tentative predictors. Here C8 is also used as the predictor because the foregoing FCM contains self-loops.

	C1	C2	C3	C4	C5	C6	C7	C8
1	0.75	0.75	0.50	0.50	0.75	1.00	0.50	1.00
2	0.95	0.94	0.82	0.56	0.97	0.61	0.94	0.96
3	0.98	0.97	0.91	0.56	0.99	0.47	0.97	0.97
4	0.98	0.98	0.92	0.55	0.99	0.43	0.97	0.96
5	0.98	0.98	0.92	0.55	0.99	0.42	0.97	0.96
6	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96
7	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96
8	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96
9	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96
10	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96

Table 4.2: The Original History Data Based on Connection Matrix in Table 4.1. when Lambda=1 (the initial values in the first row).

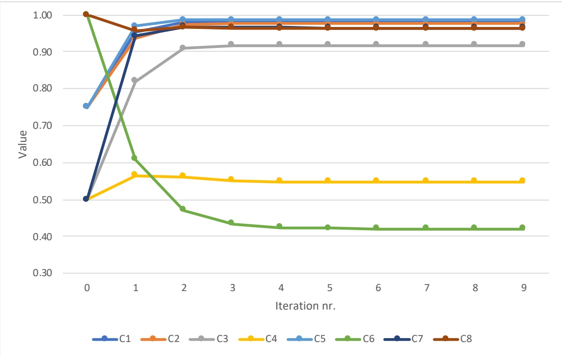

Figure 4.2.: Trends of Concepts in Parkinson's Disease FCM.

	C1	C2	C3	C4	C5	C6	C7	C8
C1	0.74	-0.21	0.97	-0.77	-0.51	0.24	-0.6	-0.5
C2	0.71	0.21	-0.14	-0.01	0.63	0.25	0.5	0.4
C3	-0.05	-0.14	0.31	0.47	0.55	-0.34	0.78	0.47
C4	-0.43	0.76	-0.75	-0.53	-0.5	-0.72	-0.46	0.28
C5	-0.19	0.91	0.7	-0.01	0.41	-0.64	-0.21	-0.71
C6	0.12	-0.06	-0.26	0.03	0.32	0.33	0.71	0.86
C7	0.81	0.75	-0.22	0.08	0.66	0.01	0.98	0.88
C8	0.47	0.2	0.3	0.69	0.85	0.56	0.46	0.83

Table 4.3: An Example of Connection Matrix when the Prevailing Optimization Approach Is Adopted.

	C1	C2	C3	C4	C5	C6	C7	C8
1	0.75	0.75	0.5	0.5	0.75	1	0.5	1
2	3.03	2.68	1.52	0.26	3.44	0.44	2.79	3.11
3	3.97	3.61	2.29	0.25	4.2	-0.12	3.37	3.32
4	4.05	3.7	2.4	0.21	4.22	-0.27	3.35	3.28
5	4.04	3.7	2.4	0.2	4.19	-0.31	3.32	3.24
6	4.04	3.69	2.4	0.19	4.18	-0.31	3.31	3.23
7	4.04	3.69	2.4	0.19	4.18	-0.32	3.31	3.23
8	4.03	3.69	2.4	0.19	4.18	-0.32	3.3	3.23
9	4.03	3.69	2.4	0.19	4.18	-0.32	3.3	3.23
10	4.03	3.69	2.4	0.19	4.18	-0.32	3.3	3.23

Table 4.4: The Inverted (Linearized) History Data of Table 4.2. for Response Variables when (3.1.) and Lambda=1 Are Used (then the initial values in the first row are not inverted and are also irrelevant).

When creating our reorganized history data matrix for the LRA model by combining Table 4.2. with Table 4.4., we will apply the method in (2.1.) to this new history data, i.e., given the two consecutive rows in the history data matrix, the concept values in the former row will yield their updated values in the latter row. After this procedure, we may apply LRA.

In our example, we will thus use the history data matrix rows 1 to 9 in Table 4.2 for the predictors C1-C8 in our example LRA model. For the response concept data of C8, in turn, we will use the rows 2 to 10 in the linearized (inverted) history data matrix in Table 4.4. Hence, we are not completely using the original history data (Table 4.5).

When applying LRA with Matlab without the constant term, i.e., when in practice resolving X in the matrix equation A*X=B in which A includes the predictor and B the response values in Table 4.5., we will obtain (as expected) such estimated regression coefficients, X, for concept C8 that are virtually identical to the weights of C8 in Table 4.1. However, contrary to the subjective reasoning, in principle, our method also enables us to examine the statistical significances of its predictors and the confidence intervals of our weights in an objective and a well-justified manner.

However, since our predictor variables caused much multicollinearity due to their high intercorrelations, their t-tests, for example, are now more or less unreliable. The original history data also converged to the fixed-point values, and that will restrict the use of representative data [4, 9]. Still another problem was that only a small history data set was available.

The similar procedure was also carried out with the other response concepts C1 to C7 and then we also obtained the coefficients identical to those weights in Table 4.1. Hence, the total of eight LRA models were constructed. Thanks for our approach, these regression coefficients will remain unaltered even though the LRA will be repeated.

Since our history data concepts may have distinct ranges in their LRA models, we may also construct alternative LRA models by replacing our data matrices with their standard scores. Hence, their corresponding standardized regression coefficients, the Beta values, will provide even better comparable grounds for evaluating the effects of the predictors on their response variables. This method is widely used in LRA. However, in this case, our coefficients are not necessarily ranging from -1 to 1 [11, 12].

When applying this procedure to our example, i.e., when using the corresponding standard scores of Table 4.5., the concepts C1, C6 and C7 seem the most important predictors for C8 due to their high absolute Beta values. This type of interpretation may also be applied to the other target concepts (Table 4.6.).

	C1	C2	C3	C4	C5	C6	C7	C8	C8r
1	0.75	0.75	0.5	0.5	0.75	1	0.5	1	**3.11**
2	0.95	0.94	0.82	0.56	0.97	0.61	0.94	0.96	**3.32**
3	0.98	0.97	0.91	0.56	0.99	0.47	0.97	0.97	**3.28**
4	0.98	0.98	0.92	0.55	0.99	0.43	0.97	0.96	**3.24**
5	0.98	0.98	0.92	0.55	0.99	0.42	0.97	0.96	**3.23**
6	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96	**3.23**
7	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96	**3.23**
8	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96	**3.23**
9	0.98	0.98	0.92	0.55	0.98	0.42	0.96	0.96	**3.23**

Table 4.5: The Data Used in Our LRA Model Example (predictors use original but response linearized values, the last column is response variable).

	C1	C2	C3	C4	C5	C6	C7	C8
C1	0.23	0.18	0	-1.24	0.23	0	0.29	0.9
C2	0.18	0.22	0	-1.66	0.07	0	0.11	0.33
C3	0.1	0	0.47	-0.9	0.37	0	0.2	0.41
C4	0.04	0.04	0.05	0.71	0.06	0	0.03	0.08
C5	0.17	0.16	0.09	2.36	0.31	-0.23	0.33	0
C6	0.11	0	0	0	0.58	0.77	0.86	2.78
C7	0.38	0.4	0.4	0	0.51	0	0.87	2.02
C8	0	0	0	0	0	0	0	0.22

Table 4.6: Standardized LRA Coefficients (Beta coefficients) as Connection Weights for Parkinson's Disease FCM.

If we aim at simplifying our FCM, we may apply stepwise LRA, among others, in which case only the statistically significant predictors are included in our models. In this manner, our connection matrix will contain less nonzero weights [11, 12].

In our example we knew the "true" connection matrix in advance, and this approach was adopted deliberatively for comparing better our outcomes to those of presented in [1]. However, as generally in the FCM construction, if the history data is only available and the lambda value is unknown, our task may be more challenging. More studies should thus be carried out within this problem area.

Hence, thanks for the general LRA theory, our approach to FCM analysis and interpretation will not rely on inordinate subjective reasoning but rather is more objective and theoretically well-justified by nature. Lack of space precludes additional examples.

5 Conclusions

FCM construction according to the history data was considered from the standpoint of the quantitative human sciences. In this context, the prevailing approaches seem to apply the methods of the neural networks when optimizing the FCM parameter values. Hence, two principal problems may arise.

First, we may obtain distinct parameter values if this optimization is repeated, this state of affairs leading to the instability problems in the FCM constructions. Second, they seem to lack an objective and a well-justified theoretical basis concerning the interpretations on their outcomes.

Our approach reduced the FCM construction to linear regression analysis by also using the linearized transformations of the original history data. Thanks for this approach, we will always obtain unique parameter values and we may also utilize the corresponding statistical theories in our conclusions. Hence, we may avoid better unstable models and ad hoc or subjective reasoning.

The FCM construction was considered first in the light of the prevailing methods. Then, our method was introduced. A concrete example was also provided with the Parkinson's disease model. As was presupposed, we noticed that our LRA outcomes were similar to those in the original Parkinson's model.

If, unlike above, we do not know the correct outcomes in advance, our task may be more challenging. Hence, our approach seems promising in this type of model construction, but still further studies are expected.

Acknowledgements

In memory of the centenary of the birth of the great Prof. Lotfi Zadeh, my inspiring friend and mentor. I also express my thanks to the anonymous reviewers for their comments.

References

[1] A. Anninou and P. Groumpos, "A Theoretical Mathematical Modeling of Parkinson's Disease Using Fuzzy Cognitive Maps", IEEE 12th Int. Conference on Bioinformatics & Engineering, Larnaca, 2012.

[2] R. Axelrod, Structure of Decision, The Cognitive Maps of Political Elites. Princeton University Press, Princeton, 1976.

[3] M. Glykas, (Ed.), "Fuzzy Cognitive Maps", Springer, 2010.

[4] I. Harmati and L. T. Koczy, "Improvements on the Convergence and Stability of Fuzzy Grey Cognitive Maps", in "Information Processing and Management of Uncertainty in Knowledge-Based Systems", pp. 509-523, Springer, 2020.

[5] M. Hatwágner, G. Vastag, V. A. Niskanen and L. Kóczy "Banking Applications of FCM Models" in M. Cornejo & al., (Eds.), "Trends in Mathematics and Computational Intelligence", Studies in Computational Intelligence, Vol. 796, pp. 61-72, Springer, 2018.

[6] M. Hatwagner, G. Vastag, V. A. Niskanen, and L. Koczy, "Improved Behavioral Analysis of Fuzzy Cognitive Map Models", in: L. Rutkowski, R. Scherer, M. Korytkowski, W. Pedrycz, R. Tadeusiewicz and J. M. Zurada (Eds.), "Artificial Intelligence and Soft Computing", Vol. 2, LNAI 10842, pp. 630-641, Springer, 2018.

[7] D. O. Hebb, "The organization of behavior: A neuropsychological theory", Psychology Press, 1949.

[8] A. Kannappan, A. Tamilarasi, E. Papageorgiou, "Analyzing the performance of fuzzy cognitive maps with non-linear Hebbian learning algorithm in predicting autistic disorder", Expert Systems with Applications 38(3), pp. 1282-1292, 2011.

[9] C. J. Knight, D. J. Lloyd, A. S. Penn, "Linear and sigmoidal fuzzy cognitive maps: analysis of fixed points", Applied Soft Computing 15, pp. 193-202, 2014.

[10] B. Kosko, "Fuzzy Engineering", Prentice Hall, Upper Saddle River, New Jersey, 1997.

[11] J. Metsämuuronen, "Essentials in Research Methods in Human Sciences, Multivariate Analysis", Sage, London, 2017.

[12] J. Metsämuuronen, "Essentials in Research Methods in Human Sciences, Advanced Analysis", Sage, London, 2017.

[13] A. Nair, D. Reckien and M. Maarseveen, "Generalised fuzzy cognitive maps: Considering the time dynamics between a cause and an effect", Applied Soft Computing, Vol. 92, pp. 1-11, 2020.

[14] V. A. Niskanen, "Application of Logistic Regression Analysis to Fuzzy Cognitive Maps", in L. Zadeh and R. Aliev (Eds.): "Fuzzy Logic Theory and Applications", pp. 507-530, World Scientific, 2019.

[15] V. A. Niskanen, "Statistical Approach to Fuzzy Cognitive Maps", in J. Kacprzyk, S. Shahbazova and M. Sugeno (Eds.): "Recent Developments in Fuzzy Logic and Fuzzy Sets", Studies in Fuzziness and Soft Computing, Vol. 391, pp. 33-59, Springer, 2020.

[16] V. A. Niskanen, "A Statistical Random Variable Approach to Fuzzy Cognitive Map Modeling", IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, 2019.

[17] E. Papageorgiou, C. Stylios, and P. Groumpos, "Fuzzy Cognitive Map Learning Based on Nonlinear Hebbian Rule", in T.D. Gedeon and L.C.C. Fung (Eds.): AI 2003, LNAI 2903, pp. 256-268, Springer, 2003.

[18] W. Pedrycz, A. Jastrzebska, and W. Homenda, "Design of Fuzzy Cognitive Maps for Modeling Time Series", IEEE Transactions on Fuzzy Systems, Vol. 24, No. 1, pp. 120-130, 2016.

[19] W. Stach, L. Kurgan, W. Pedrycz, "A survey of fuzzy cognitive map learning methods", Issues in soft computing: theory and applications", pp. 71-84, 2005.

[20] W. Stach, L. Kurgan, W. Pedrycz, M. Reformat, "Genetic learning of fuzzy cognitive maps", Fuzzy sets and systems 153(3), pp. 371-401, 2005.

[21] W. Stach, L. Kurgan, W. Pedrycz, "Numerical and linguistic prediction of time series with the use of fuzzy cognitive maps", IEEE Transactions on Fuzzy Systems 16(1), pp. 61-72, 2008.

[22] C. Stylios, P. Groumpos, "Modeling complex systems using fuzzy cognitive maps", IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 34(1), pp. 155-162, 2004.