

A Statistical Random Variable Approach to Fuzzy Cognitive Map Modeling

Vesa A. Niskanen
Dept. of Econ. & Management
University of Helsinki
Helsinki, Finland
vesa.a.niskanen@helsinki.fi

Abstract—Fuzzy cognitive maps are studied from the quantitative human-scientific standpoint. The concepts and weights of these maps are thus examined as statistical random variables based on probability distributions in order to make more feasible interpretations on these entities. First, the available fuzzy cognitive maps are considered. Second, our statistical approach is introduced which assumes that the concepts and weights are uniformly or normally distributed random variables. Third, the dynamics and application possibilities of these fuzzy cognitive maps are studied.

Keywords—fuzzy cognitive maps, statistical random variables, statistical analysis, human sciences

I. INTRODUCTION

Fuzzy cognitive maps (FCM) seem to be an applicable method in complex system modeling. Given a set of concepts and their interrelationships, we may study how these concepts will vary in a given time interval. Both numeric and linguistic FCMs are available, and below we focus on the former type of models [1,3,7,10,11].

However, the numerical FCMs encounter certain well-known problems due to their mathematical properties [3]. For example, they only allow monotonic relationships between the concepts. Also, they still seem to have quite much methodological commitments to neural networks, and thus the interpretations of their concept values and relationships may be ambiguous or only base on user's ad hoc meanings. The Author has studied FCMs from the statistical standpoint for providing a more feasible and consistent basis on FCM interpretation, and this approach is also central in the quantitative human sciences, in particular in the social and behavioral sciences as well as in economics. For example, the Author has applied linear, logistic and multinomial logistic regression analyses because these methods seem analogous to the FCM modeling [6,15,16].

Below we continue this examination and assume that FCM's entities are not any arbitrary factors but rather statistical random variables, and then we study how this approach may enrich and clarify our model construction. We also emphasize the human-scientific aspect.

We will focus on the initial stage of FCM model construction which is still a quite novel territory, whereas the subsequent stages are already much analyzed in the FCM literature [3,5,6,8,9,17-22].

Section 2 presents basic theory on the FCMs. Section 3 examines the role of the random variables in the context of the FCMs and also provides some application examples. Section 4 concludes our study.

II. FUZZY COGNITIVE MAPS

If the prevailing methods are used in the numeric FCM model construction, the concept (node) values, C , range from 0 to 1 and their weights, or their intensities of interrelationship, W , range from -1 to 1 [3,10]. Hence, the weight of a relationship indicates how a concept affects another concept. These weights are presented in the $N \times N$ connection matrix, M , when N is the number of concepts in an FCM model. These matrices are constructed according to our expertise or the given historic data [3]. In the latter case we may apply various optimization techniques for finding the appropriate weights [3,4,8,17-21]. Hence, in the typical FCM computer simulations in a given time interval, $t=1,2,\dots,n$, we may basically apply the matrix product, *

$$V_{t+1} = f(V_t * M), t = 1, 2, \dots, n \quad (2.1)$$

in which the state vector, $V_t = (C_1, \dots, C_N)$, contains the driver concept values at time t , f is the transformation function and the target vector, V_{t+1} , contains the obtained concept values at time $t+1$ [3]. Thus, at each iteration, $t > 1$, we will update our concept values.

The transformation function, f , is usually the logistic function with the parameter $\lambda > 0$ (lambda) and the exponential function, \exp [3,10],

$$0 < f(x) = 1 / (1 + \exp(-\lambda \cdot x)) < 1, \quad (2.2)$$

or the hyperbolic tangent function

$$-1 < f(x) = (\exp(\lambda \cdot x) - \exp(-\lambda \cdot x)) / (\exp(\lambda \cdot x) + \exp(-\lambda \cdot x)) < 1. \quad (2.3)$$

Hence, this function transforms the initial target concept values at time $t+1$ into the desired interval. Below we only use (2.2), and then lambda is quite often having the values of 1 or 5 (Fig. 2.1).

A simple example is the well-known public city-health model in which we may apply the connection matrix in Table 2.1 and the drivers and targets are its row and column concepts, respectively [12].

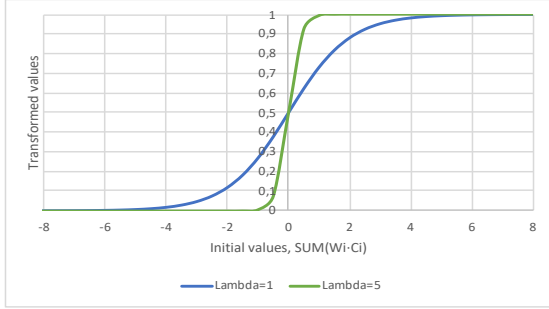


Fig. 2.1. Transformation functions (2.2) with two values of lambda, 1 and 5.

Given now, for example, the driver concept values Migration=0.7 and Nr_diseases=0.2 at time=t, the target concept value for Nr_people at time=t+1 will be

$$\text{Nr_people} = f(0.7 \cdot 0.5 + 0.2 \cdot (-0.3)) = 0.57$$

when lambda=1. The same procedure may also be applied to the other target concepts. In the conventional statistics we have studied this type of problem-setting with correlations and such structural equation models as LisrelTM, AmosTM or MplusTM, but they are more complicated in simulations.

TABLE 2.1. CONNECTION MATRIX FOR PUBLIC CITY-HEALTH MODEL

Concept	Concept						
	Nr_pe ople	Mi-gra-tion	Mod-erni-zation	Gar-bage	Sani-tation	Nr_di-seases	Bac-teria
Nr_pe ople	0	0	0.6	0.9	0	0	0
Mi-gra-tion	0.5	0	0	0	0	0	0
Mod-erni-zation	0	0.6	0	0	0.8	0	0
Gar-bage	0	0	0	0	0	0	0.9
Sani-tation	0	0	0	0	0	-0.8	-0.9
Nr_di-seases	-0.3	0	0	0	0	0	0
Bac-teria	0	0	0	0	0	0.8	0

However, in the quantitative human sciences we expect that we may unambiguously and consistently interpret these weight and concept values and this aspect is not necessarily taken into account in the prevailing FCM models. In particular, in [10] the positive and negative weight values mean causal increase or decrease, respectively, but this quite general definition has led to quite incoherent and ambiguous interpretations in applications.

Another problem is that the goodness of an FCM model in practice is preferred to its thorough interpretation of the weights, and this due to the neural network origin of the FCMs in the

engineering sciences. Hence, in fact, the FCMs may even be black or at least grey boxes for their users [8,10,17,20]. For the sake of comparison, the corresponding statistical correlation matrix would provide us with deeper and more consistent information.

Hence, below we will suggest a statistical interpretation when we study these models within the quantitative human sciences because then we may better make unique, consistent and empirically justified interpretations to our FCM entities.

III. STATISTICAL APPROACH TO FUZZY COGNITIVE MAPS

In the quantitative human sciences, we operate with statistical distributions when we study our variables in model construction. Hence, within the human-scientific FCMs we might also apply statistical theories. Thanks for this approach, we may have a firm and plausible empirical basis for our examinations. We may also make consistent interpretations on our model entities and even study their significances with statistical tests. Below we consider how this approach may provide additional information on FCM construction and we used MatlabTM in our computer simulations. We will focus on the individual target concepts, whereas examples on the entire FCM simulations are already found quite much in the FCM literature, e.g., in [4,6,8,9,17-22].

A. Statistical Random Variables and FCMs

Due to the lack of space, we will only study below two widely-used distributions, the uniform and normal distribution. Consider first, that the FCM concept values, C, are uniformly distributed from 0 to 1, $C \sim \text{Uni}(0,1)$, which seems to be usual in engineering applications.

Consider also, that the connection weights, W, obey the uniform distribution, $W \sim \text{Uni}(-1,1)$, because this seems to be the usual assumption to all models. Thus, we will study their products, $C \cdot W$, in the FCMs (Figs. 3.1-3.3). Since according to (2.1), each initial target concept value is the sum of these products, $\sum C \cdot W$, (Figs. 3.5 and 3.7), the elementary statistical theory indicates that their expected values are $E(C)=0.5$ and $E(W)=0$. Their variances are $\text{var}(C) = 1/12 \cdot 1 = 1/12$ and $\text{var}(W) = 1/12 \cdot 2^2 = 1/3$ [2,13,14]. Hence, in statistics their product distribution will have the expected value

$$E(C \cdot W) = E(C) \cdot E(W) = 0.5 \cdot 0 = 0, \quad (3.1)$$

and the variance

$$\text{var}(C \cdot W) = \text{var}(C) \cdot \text{var}(W) + \text{var}(C) \cdot E(W)^2 + \text{var}(W) \cdot E(C)^2 = 1/36 + 0 + 1/12 = 1/9 = 0.11 \quad (3.2)$$

As regards the corresponding standard deviation, $\text{std}(C \cdot W) = 0.33$, and this statistic is preferred to variance below. Hence, the sum of N product distributions, which is the initial value of a target concept with its N drivers in each FCM iteration,

$$\sum_i C_i \cdot W_i, \quad i = 1, 2, \dots, N$$

will yield the expected value

$$N \cdot E(C \cdot W) = N \cdot 0 = 0 \quad (3.3)$$

and the standard deviation

$$\text{sqrt}(N) \cdot \text{std}(C \cdot W) = \text{sqrt}(N) \cdot 0.33. \quad (3.4)$$

Naturally, these initial target values always range from $-N$ to N .

In the human sciences we are often more interested in the normally distributed FCM concepts because many real-world phenomena seem to obey this distribution. We assumed that each driver concept is normally distributed with parameters $C \sim \mathcal{N}(0.5, 0.17)$, i.e., with the expected value of 0.5 and $\text{std}=1/6=0.17$, because then approximately 99 % of the possible driver concept values belong to the interval $[0.5-3 \cdot \text{std}, 0.5+3 \cdot \text{std}] = [0, 1]$. The rest of the possible concept values were excluded in our random variables which is quite usual procedure in empirical studies (Fig. 3.2) [13,14]. We notice in Fig. 3.2 that, for example, the concept value of 0.5 is more frequent within the normal than the uniform distributions.

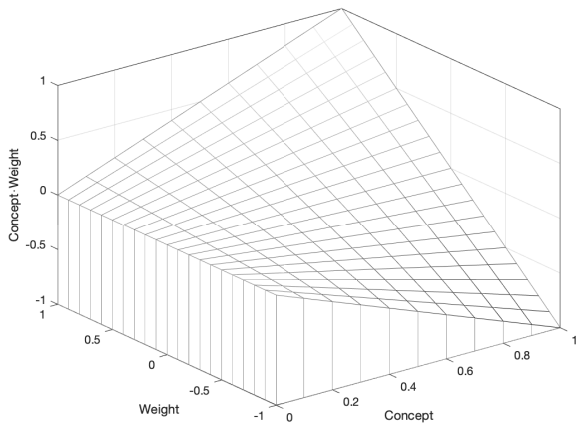


Fig. 3.1. The possible product values, $C \cdot W$, for an FCM concept.

When applying (3.1) and (3.2) to the normally distributed concepts, $E(C \cdot W)=0$ and $\text{std}(C \cdot W)=0.30$, and for N concepts we may again use (3.3) and (3.4) (Figs. 3.4, 3.6 and 3.8).

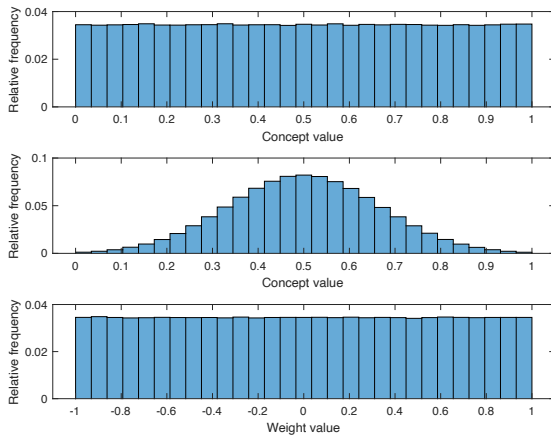


Fig. 3.2. Relative statistical distributions of the concept values when they are uniformly (top) or normally distributed (middle). Relative distribution of uniformly distributed weight values (bottom).

We thus notice in Figs. 3.3-3.8 that the shapes of the uniformly distributed initial target values differ from their normally distributed counterparts, and obviously the type of distribution also affects the corresponding transformed values. Naturally, the lambda values are also essential these calculations.

We also notice, for example, that by only using ten driver concepts for a target concept, the probability of obtaining transformed target concept values around 0.5 is clearly less than 5 % with our random distributions and lambda values. On the other hand, the proportions of the extreme values among the transformed target values will increase when N increases, even exceeding 80 %, and these outcomes seem counterintuitive in applications (Figs. 3.5-3.8).

Hence, thanks for our statistical approach, we may better anticipate the dynamics of the FCMs and also specify plausible statistical distributions according to the concepts' background theories and empirical results. If this information is unavailable, the uniform distribution, for example, seems appropriate. This approach also enables us to make more feasible interpretations on our FCM entities, and this issue will be considered next.

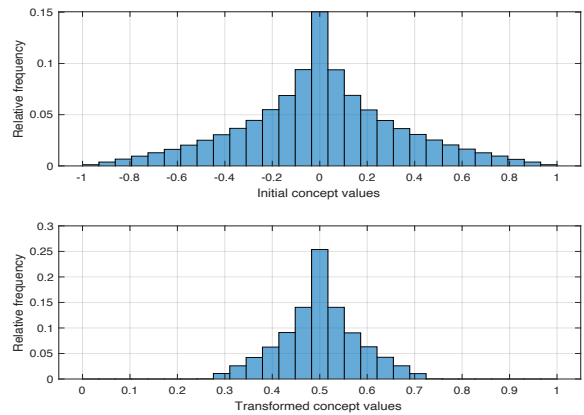


Fig. 3.3. Relative statistical distributions of the initial (top) and transformed target concept values with one million uniformly distributed random values when $\lambda=1$, $C_{\text{target}}=C \cdot W$, $C \sim \text{Uni}(0,1)$ and $W \sim \text{Uni}(-1,1)$.

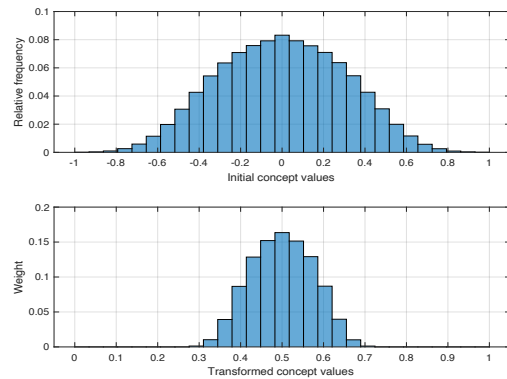


Fig. 3.4. Relative statistical distributions of the initial (top) and transformed target concept values with one million random values when $\lambda=1$, $C_{\text{target}}=C \cdot W$, $C \sim \mathcal{N}(0.5, 0.17)$ and $W \sim \text{Uni}(-1, 1)$.

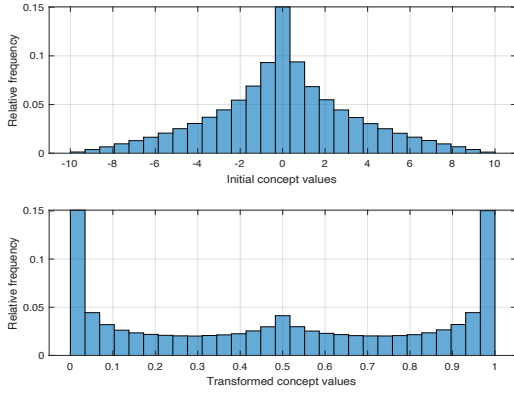


Fig. 3.5. Relative statistical distributions of the initial (top) and transformed target concept values with one million random values when $\lambda=1$, $C_{\text{target}}=\sum_i C_i \cdot W_i$, $i=1, 2, \dots, 10$, $C \sim \text{Uni}(0, 1)$ and $W \sim \text{Uni}(-1, 1)$.

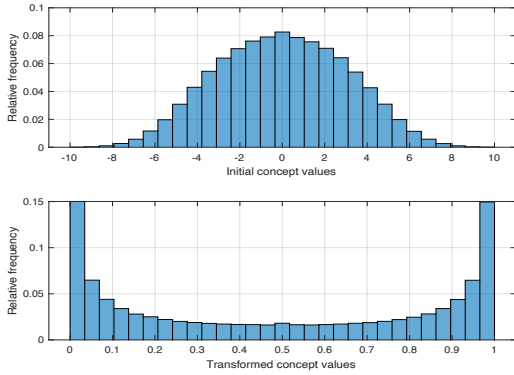


Fig. 3.6. Relative statistical distributions of the initial (top) and transformed target concept values with one million random values when $\lambda=1$, $C_{\text{target}}=\sum_i C_i \cdot W_i$, $i=1, 2, \dots, 10$, $C \sim \mathcal{M}(5, 0.17)$ and $W \sim \text{Uni}(-1, 1)$.

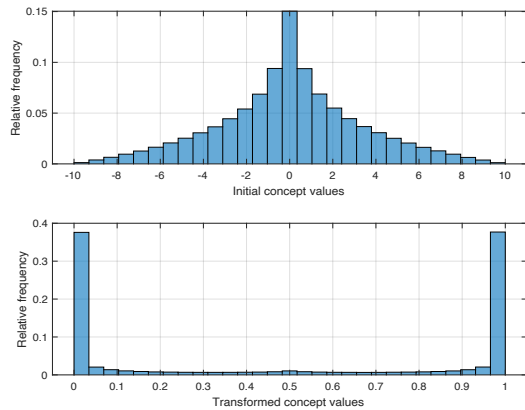


Fig. 3.7. Relative statistical distributions of the initial (top) and transformed target concept values with one million random values when $\lambda=5$, $C_{\text{target}}=\sum_i C_i \cdot W_i$, $i=1, 2, \dots, 10$, $C \sim \text{Uni}(0, 1)$ and $W \sim \text{Uni}(-1, 1)$.

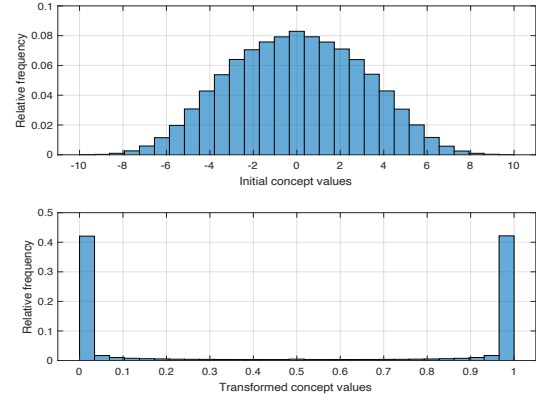


Fig. 3.8. Relative statistical distributions of the initial (top) and transformed target concept values with one million random values when $\lambda=5$, $C_{\text{target}}=\sum_i C_i \cdot W_i$, $i=1, 2, \dots, 10$, $C \sim \mathcal{N}(0.5, 0.17)$ and $W \sim \text{Uni}(-1, 1)$.

B. Problems with Target Concept Values

As stated above, one crucial problem with the numeric FCMs is that the more we have driver concepts for a target concept, the larger is the standard deviation of their sum, $std(\sum_i C_i \cdot W_i)$, because this initial value is always $N \cdot std(C \cdot W)$. This, in turn, means that in transformations (2.2) large N values will yield much extreme values, and these outcomes often seem counterintuitive in practice (Figs. 3.9-3.10). As Fig. 3.9 shows, if $N > 8$ and $\lambda = 1$, the proportion of the extreme transformed values will clearly increase. This situation is even worse with the higher λ values. We should also notice that, in fact, we will never obtain the exact values of 0 or 1 in the foregoing transformations.

For resolving this problem, we should apply alternative transformations. One simple resolution is to use distinct λ values for each target concept [4-6,8]. Another method is to apply such alternative aggregation operators which are used in fuzzy decision making.

In the human sciences we may often use the standard scores, Z , of the variables in this context,

$$Z = (X - \text{mean}(X)) / \text{std}(X) \quad (3.5)$$

in which X is the original variable. The means and standard deviations of the standard scores are 0 and 1, respectively. Then, we may operate with similar scales of measurement [13,14]. Their values virtually range from -3 to 3 when their original variables are normally distributed. This transformation is applied to regression analysis, cluster analysis, principal component analysis and factor analysis, among others [2,13,14]. Within the FCMs, (3.5) seems useful with such additional modification as

$$(Z+3) / 6 \in [0, 1] \quad (3.6)$$

when we transform our original concepts with various scales of measurement into the appropriate FCM concept values [2,13,14]. Due to the lack of space, this examination is precluded here.

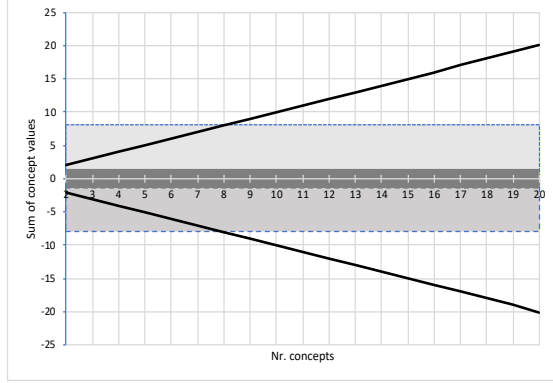


Fig. 3.9. The extreme values of the initial target concept values, $Y = \sum C_i \cdot W_i$, with various number of driver concepts, N , (lines), and the image intervals, $0 < Y < 1$, of the transformed values when $\lambda = 1$ (grey) and $\lambda = 5$ (dark grey).

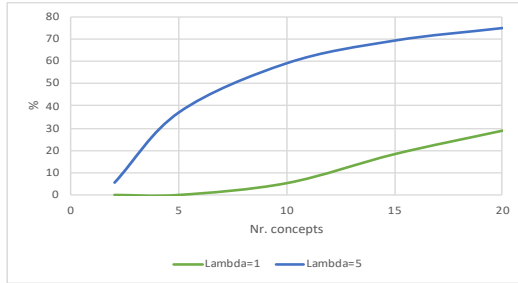


Fig. 3.10. The proportions of the target concept's extreme transformed values, approximately 0 or 1, according to the number of its driver concepts and two lambda values.

We thus notice that in addition to the FCM concept distributions, the transformations also affect significantly their dynamics when the simulations are carried out. These factors will in turn affect the subsequent simulations. The simulation aspect is briefly considered in the next Section.

C. Prospects for Applications with Random Variables

When we study our FCM concepts as statistical random variables, we may utilize various statistical tests and analyses, in particular when historic data is involved. Below we sketch some application ideas.

First, if we apply the idea of the level of significance in statistics, we may estimate how usual our concept values are [2,13,14]. For example, if we have ten uniformly distributed driver concepts and $\lambda = 1$ for a given target concept, then 3.1 % of the obtained initial values and 35.8 % of their transformed values are at least 7 or 0.7, respectively. Hence, at the 5 % level of significance, the former case is quite rare for the target value from the statistical standpoint. We may also apply Kolmogorov-Smirnov and Shapiro-Wilk tests when we analyze the normality of our distributions.

Second, we may draw similarity comparisons between our concept values [2,13,14]. For example, we may study the Pearson correlations between the uniformly and normally distributed concepts (Fig. 3.11).

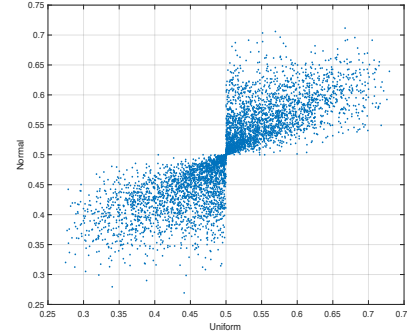


Fig. 3.11. Scatterplot of transformed concept values $C_1 \cdot W$ and $C_2 \cdot W$ with $\lambda = 1$ when C_1 and W are uniformly and C_2 normally distributed.

Third, since the calculation of the initial target values is analogous to linear regression analysis, we may utilize these techniques and thus gain certain added outcomes to FCM construction. Hence, the drivers will be the independent variables and their target is the dependent variable. In particular, we may apply this idea when we examine how our connection matrix yields various target values [15,16]. The regression coefficients are now identical to our weight values (the constant is excluded) and their corresponding standardized beta coefficients and t-test statistics reveal to us the significance of each driver, and thus they may also help us to remove insignificant drivers. The collinearity diagnostics also helps us for reducing the number of drivers and thus yielding simpler FCM models.

Fourth, we may even apply logistic and multinomial logistic regression analyses which enable us to calculate the odds ratios and the probabilities of obtaining desired target concept values (Fig. 3.12) [15,16]. In fact, this logistic modeling, which applies iterative maximum likelihood method, is analogous to the calculation in (2.1) when $\lambda = 1$ [13-16].

The foregoing methods will enable us to make well-justified, consistent and less ambiguous interpretations on our FCM entities. They also reveal to us well the real nature and dynamics of our FCMs, and thus we may construct simpler models and perform plausible predictions on their values more fluently and accurately.

IV. CONCLUSIONS

Fuzzy cognitive maps have been considered from the statistical standpoint. Their concepts and weights were examined as statistical random variables because this approach is usual in particular in the quantitative human sciences. Only uniformly and normally distributed variables were examined but other random distributions may also be used. We noticed that, by applying statistical theories on random distributions to the driver con-

cepts, the distinct random variables yielded distinct FCM outputs. In addition, if the target concepts were the sums of many driver concepts, their transformed outputs tend to approach the extreme values, and this feature seems problematic in practical applications. We suggested one resolution to this problem with the standard scores.

We also sketched certain ideas for applying statistical analyses for the iterated FCMs even though this problem area is already more examined in the literature.

In general, thanks for our statistical approach, we may better make unambiguous and consistent interpretations to our concepts and weights by utilizing the prevailing statistical tools. We may also better examine, anticipate and predict the dynamics or behavior of our FCMs, and statistical theories will support us in this task.

Our study focused on the first stage in the FCM simulation and only on two usual random variables because this issue is still more or less new territory and thus more examinations are still expected in this problem area.

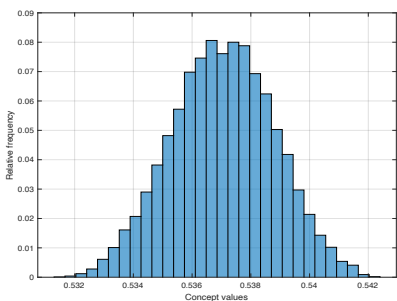
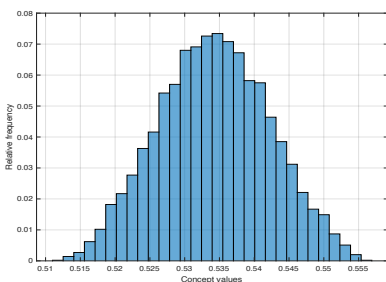
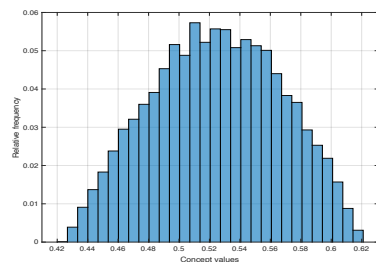


Fig. 3.12. Relative histograms of target values for Nr_people in Public City-Health model after the first (top), second (middle) and third iteration (bottom) when $\lambda=1$ and 10 000 uniformly distributed random initial driver vectors are used.

V. REFERENCES

- [1] R. Axelrod, *Structure of Decision, The Cognitive Maps of Political Elites*. Princeton University Press, Princeton, 1976.
- [2] D. Freedman, "Statistical models: Theory and practice", Cambridge University Press, Cambridge, 2005.
- [3] M. Glykas, (Ed.), "Fuzzy Cognitive Maps", Springer, 2010.
- [4] M. Hatwagner, "Parameterization and concept optimization of FCM models", IEEE International Conference on Fuzzy Systems, 2015.
- [5] M. Hatwagner, 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. Kóczy, "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] V. Kreinovich, C. Stylios, "Why fuzzy cognitive maps are efficient", *International Journal of Computer Communications & Control* 10(5), pp. 825-833, 2015.
- [12] K.C. Lee, W.J. Lee, O.B. Kwon, J.H. Han and P.I. Yu, "Strategic planning simulation based on fuzzy cognitive map knowledge and differential game", *Simulation*, Vol. 71 (5), pp. 316-327, 1998.
- [13] J. Metsämuuronen, "Essentials in Research Methods in Human Sciences, Multivariate Analysis", Sage, London, 2017.
- [14] J. Metsämuuronen, "Essentials in Research Methods in Human Sciences, Advanced Analysis", Sage, London, 2017.
- [15] V. A. Niskanen, "Application of Logistic Regression Analysis to Fuzzy Cognitive Maps", in L. Zadeh and R. Aliev (Eds.): "Fuzzy Logic Theory and Applications", World Scientific, forthcoming.
- [16] V. A. Niskanen, "Statistical Approach to Fuzzy Cognitive Maps", in J. Kacprzyk and S. Shahbazova (Eds.): "Recent Developments in Fuzzy Logic and Fuzzy Sets", Studies in Fuzziness and Soft Computing, Springer, forthcoming.
- [17] E. Papageorgiou, C. Stylios, and P. Groumos, "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.