Introducing Fuzzy Cognitive Maps for decision making in precision agriculture

Ath. Markinos[1], El. Papageorgiou[2], Chr. Stylios[3] and Th. Gemtos[1]
[1]*Laboratory of Farm Mechanization, Department of Agriculture, Crop Production and Rural Environment, University of Thessaly, Greece*
[2]*Laboratory for Automation and Robotics, Department of Electrical and Computer Engineering, University of Patras, Greece*
[3]*Department of Telecommunications, Informatics and Management, TEI of Epirus, Artas Epirus, Greece*
markinos@agr.uth.gr

Abstract

A Fuzzy Cognitive Maps (FCMs) is a modelling methodology based on exploiting knowledge and experience. It comprises the main advantages of fuzzy logic and neural networks, representing a graphical model that consists of nodes-concepts (describing elements of the system) which are connected with weighted edges (representing the cause and effect relationships among the concepts). FCMs have proved to be a promising modeling methodology with many successful applications in different areas especially for simulating system design, modeling and control. In this work, FCMs are introduced to model a decision support system for precision agriculture (PA). The FCM model developed consists of nodes which describe soil properties and cotton yield and of the weighted relationships between these nodes. The nodes of the FCM model represent the main factors influencing cotton crop production i.e. essential soil properties such as texture, pH, OM, K, and P. The proposed FCM model addresses the problem of crop development and spatial variability of cotton yield, taking into consideration the spatial distribution of all the important factors affecting yield. The first results of the study are very promising; our model achieves a 70% average success rate on yield class prediction between two possible categories (low and high) for three different years. This model will be further investigated to achieve better results by introducing learning algorithms into FCMs.

Keywords: Fuzzy Cognitive Maps, modeling, fuzzy sets, decision making, cotton crop

Introduction

Fuzzy Cognitive Maps (FCMs) constitute an attractive modeling technique for complex systems. FCMs were proposed by Kosko (1986) to represent the causal relationship between concepts and analyze inference patterns. FCM is a soft computing technique that follows an approach similar to the human reasoning and decision-making process. An FCM consists of nodes which illustrate the different aspects of the system behavior. These nodes (concepts) interact with each other, illustrating the dynamics of the model. Human experts who supervise a system and know its behavior under different circumstances develop a FCM model of the system in such a way that their accumulated experience and knowledge are integrated into the causal relationships between factors/characteristics of the FCM model (Stylios and Groumpos, 1999).

FCMs have been used in many disciplines for easy comprehension of complex social systems and for decision-making (Peláez and Bowles, 1996; Miao and Liu, 2000; Papageorgiou *et al.*, 2003). Here, a first study on implementing FCM for decision making in PA has been investigated using seed cotton production as an example. This approach creates a novel simulation model to describe the seed cotton yield spatial distribution based on the spatial distribution of soil properties in a field.

Previous studies in the field of precision agriculture mainly employed linear algorithms, direct processes and statistical methods. Some studies have been made using artificial neural networks (ANNs) and machine-learning algorithms for setting target yields which is one of the problems in PA (Liu *et al.*, 2001; Miao *et al.*, 2006). In the case of knowledge-based systems using fuzzy logic techniques, only a few studies have been undertaken so far (Ambuel *et al.*, 1994; Khan and Khor, 2004), but their first results on predicted yields were preliminary and further work is needed using real measurements.

This work proposes an alternative methodology to yield prediction in precision farming, which is based on the formalization of specialized knowledge and experience from experts (soil scientists, experienced farmers). Here, a first study on implementing FCM for predicting final yield as a part of the decision-making process in precision agriculture has been investigated. FCMs have been applied to model spatial variable seed cotton production.

Materials and methods

In 2001, the Laboratory of Farm Mechanization of the University of Thessaly established an experimental 5ha field at Myrina, Karditsa prefecture, Central Greece. Over the last 6 years, the field was cultivated and managed using spatially uniform applications and a series of measurements were made each year.

Yield mapping was undertaken for the years 2001-06 using a commercial yield monitor system from Farmscan™ installed on a two row John Deere™ cotton picker (Gemtos *et al.*, 2004). After harvesting of a field was complete, a calibration procedure was performed to improve the yield estimation (Markinos *et al.*, 2004).

In May 2006, a VERIS machine was used to measure the apparent soil electrical conductivity (EC) and produce maps at two depths 0-0.30 and 0-0.90 m. The machine consists of a sensor cart with four vertical disks mounted on it (Lund *et al.*, 1999). The machine was pulled through the field at a speed of approximately 7 km/h, in tracks at a spacing of 4 m. Data were recorded every 1 s.

In February 2002, a 16x26 m grid was formed in the north part of the field (4.3 ha). Overall, 114 soil samples were taken at the grid points at 0-30cm depth. The samples were analyzed for texture, N, P, K, pH, Mg, Ca, Na and organic matter.

The SSToolbox™ 3.61 software was used to store, represent, filter and analyze the acquired field data (SSToolbox, 2004). All the collected data were interpolated in order to produce a map (4.3 ha) on a 10x10 m grid size that corresponds to a reliable field management unit (cell). The interpolation method of inverse distance was used for yield and EC due to dense data sampling, while kriging was used for the soil properties maps derived from a sparse spatial sampling-grid (SSToolbox, 2004). Data from 20m strips around the field near the edges were filtered and removed to avoid machinery compacted soil with lower yields. The data from every cell (10x10 m) of the filtered maps represent the data to be used as inputs in the FCM model simulations with the yield from each year as output. Every cell of each input map linked to a scalar value in a GIS database. Each particular cell defines a vector of scalar values for the same spatial point for each corresponding map layer (for every measured parameter) complete with the specified measured yield at the same point. Every vector constitutes a record in the database matrix extracted from GIS and imported into MATLAB code for the specific FCM model.

Fuzzy Cognitive Maps representation

FCMs represent knowledge in a symbolic manner and relate states, processes, events, values and inputs. The knowledge which accumulated for years on the operation and behavior of a system can be adequately explained using FCMs. Figure 1 illustrates a graphical representation of a FCM consisting of five concepts (C1 to C5) and ten weights w_{ji} (cause-effect relationships among the concepts).

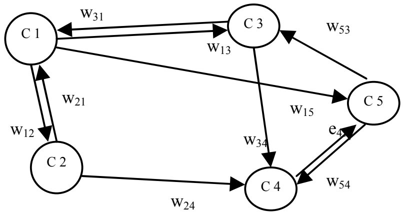

Figure 1. A simple Fuzzy Cognitive Map.

The cause and effect interconnection between two concepts C_j and C_i is described with the weight w_{ji}, taking a value in the range –1 to 1. Three possible types of causal relationships exist: $w_{ji} > 0$ which indicates positive causality between concepts C_j and C_i, $w_{ji} < 0$ which indicates negative causality between concepts C_j and C_i and $w_{ji} = 0$ which indicates no relationship between C_j and C_i.
The value A_i of the concept C_i expresses the degree of its corresponding physical value. At each simulation step, the value A_i of a concept C_i is calculated by computing the influence of other concepts C_j's on the specific concept C_i following the calculation rule:

$$A_i^{(k+1)} = f(A_i^{(k)} + \sum_{j \neq i}^{N} A_j^{(k)} \cdot w_{ji}) \tag{1}$$

where $A_i^{(k+1)}$ is the value of concept C_i at simulation step $k+1$, $A_j^{(k)}$ is the value of concept C_j at simulation step k, w_{ji} is the weight of the interconnection from concept C_j to concept C_i and f is a sigmoid threshold function:

$$f = \frac{1}{1 + e^{-\lambda x}} \tag{2}$$

where $\lambda > 0$ is a parameter that determines its steepness. In our approach, the value $\lambda=1$ has been used. This function is selected since the values A_i lie within [0,1].
The procedure for constructing fuzzy cognitive maps is as follows: experts define the main concepts that represent the model of the system; they describe the structure and the interconnections of the network using fuzzy conditional statements. The fuzzy IF-THEN rule that experts use to describe the relationship among concepts assumes the following form, where A and B are linguistic variables:

*IF value of concept C_i is **A** THEN value of concept C_j is **B**.*

The linguistic variable, describing the causal relationship between the value of concept C_i and concept C_j, is inferred from the fuzzy rule.
More specifically, the linguistic variables describing the causal inter-relationships among concepts are declared using the variable *Influence* which takes values in the universe U=[-1,1]. Its term set T(influence) is suggested to comprise seven variables. Using seven linguistic variables, an expert can describe in detail the influence of one concept on another and can discern between different degrees of influence. The seven variables that are used frequently according to the problem characteristics are: T(influence)={very very low, very low, low, medium, high, very high, and very very high}. The corresponding memberships functions, that describe each linguistic variable, for these terms are shown in Figure 2 and they are: μ_{vvl}, μ_{vl}, μ_l, μ_m, μ_h, μ_{vh} and μ_{vvh}.
Then, all the proposed linguistic variables suggested by experts, are aggregated using the SUM method and an overall linguistic weight is produced which, with the defuzzification method of Center Of Gravity (Jang *et al.*, 1997), is transformed to a numerical weight w_{ji}, within the interval [-1, 1]. A

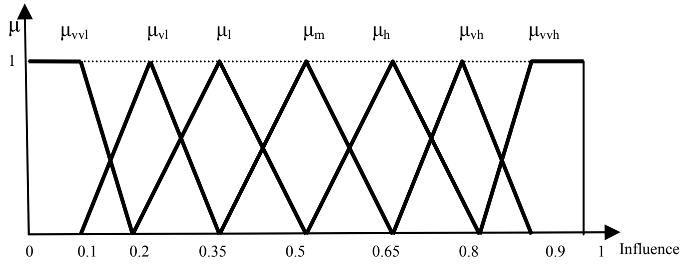

Figure 2. The 7 membership functions corresponding to the 7 linguistic variables.

detailed description of the development of FCM model is given in (Stylios and Groumpos, 2004). The flexibility of FCMs in system design, model and control, as well as their learning properties, make their choice attractive for a variety of modeling and decision support tasks (Papageorgiou *et al.* 2004; Papageorgiou and Groumpos 2005).

Fuzzy Cognitive Map model for precision farming

To design the FCM model for precision farming, one experienced cotton farmer and two experienced soil scientists played the role of experts and they designed the FCM model following the developing methodology described in Stylios and Groumpos (1999). The three experts stated that there are eleven main factors-concepts (which represent soil properties) that determine the cotton yield; (see Table 1). The output concept C_{12} represents the first cotton yield picking measured with the yield monitor. The set of linguistic variables that every concept can take are described in Table 1 and the corresponding membership functions for the four selected soil parameters (e.g. sand, clay, OM and shallowEC) are illustrated in Figure 3.

Then, the experts were asked to describe the degree of influence from one concept to another using IF-THEN rules among factor concepts and yield. An example of this process is given, selecting the relation between concepts C_8 to C_{12} for the calculation of linguistic variable. The following rules were proposed by each expert:

1st Expert: *IF value of concept C_8 is* ***low*** *THEN value of concept C_{12} is* ***low***
IF value of concept C_8 is ***med*** *THEN value of concept C_{12} is* ***med***
2nd Expert: *IF value of concept C_8 is* ***high*** *THEN value of concept C_{12} is* ***med***
3rd Expert: *IF value of concept C_8 is* ***high*** *THEN value of concept C_{12} is* ***high***

These fuzzy rules for each causal relationship are aggregated using the approach described in (Stylios and Groumpos, 2004) and so an overall linguistic rule is produced from which, using the SUM fuzzy inference method, a numerical weight for w_{ij} is calculated. Using this, the weights of the FCM model are inferred and the FCM shown in Figure 4 is developed.

The PA procedure is based on the determination of the value of output concept "Yield" that estimates the cotton yield measured with the yield monitor.

Results

Some of the resulting yield and soil property maps are showed in Figure 5. The proposed FCM approach developed a simulation model for precision farming which can be implemented for decision-making to examine different scenarios and determine the category of the cotton yield. Three different cases have been examined using the proposed FCM model and the simulation results for

Table 1. Concepts of the FCM: Type of values.

C1: ShallowEC (mS/m)	C2: Mg (ppm)	C3: Ca (ppm)	C4: Na (ppm)
Five Fuzzy	*Five Fuzzy*	*Five Fuzzy*	*Five Fuzzy*
0 – 10 Very Low	< 60 Very Low	< 400 Very Low	< 25 Very Low
10 – 20 Low	60 – 180 Low	400 – 1000 Low	25 – 70 Low
20 – 30 Medium	181 – 360 Medium	1001 – 2000 Medium	71 - 160 Medium
30 – 40 High	361 - 950 High	2001 – 4000 High	161 – 460 High
> 40 Very High	> 950 Very High	> 4000 Very High	> 460 Very High
C5: K (ppm)	**C6: P (ppm)**	**C7: N (ppm)**	**C8: OM (ppm)**
Five Fuzzy	*Five Fuzzy*	*Five Fuzzy*	*Three Fuzzy*
< 40 Very Low	< 5 Very Low	< 3 Very Low	< 1.0 Low
40 – 120 Low	5 – 15 Low	3 – 10 Low	1.0 – 2.0 Medium
121 – 240 Medium	16 – 25 Medium	11 – 20 Medium	> 2.0 High
241 – 470 High	26 – 45 High	21 – 40 High	
> 470 Very High	> 45 Very High	> 40 Very High	
C9: Ph	**C10: Sand %**	**C11: Clay %**	**C12: Yield (tons/ha)**
Seven Fuzzy	Four Fuzzy	Three Fuzzy	Three Fuzzy
<4.5 Very Low	< 20 Low	< 15 Low	< 2.5 Low
4.6 – 5.5 Low	20 – 70 Medium	15 – 37 Medium	2.5 - 3.5 Medium
5.6 – 6.5 Slightly Low	71 – 80 High	Texture	>3.5 High
6.6 – 7.5 Neutral	> 80 Very High	> 37 High	
7.6 – 8.5 Slightly High			
8.6 - 9.5 High			
> 9.5 Very High			

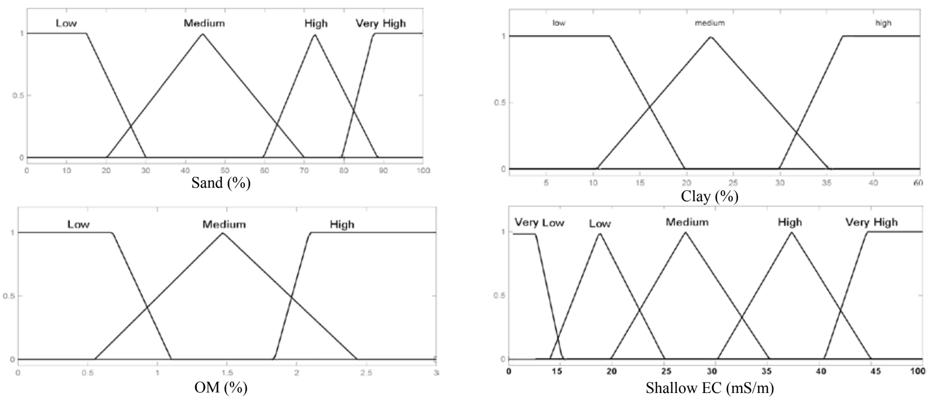

Figure 3. Membership functions for Shallow EC, OM, Sand and Clay.

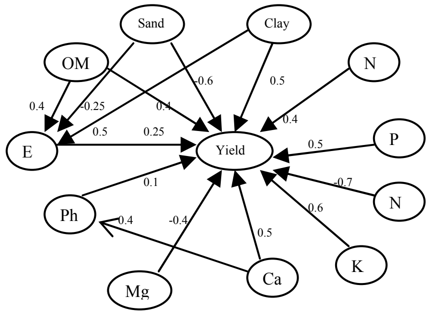

Figure 4. The FCM model for describing the final cotton yield.

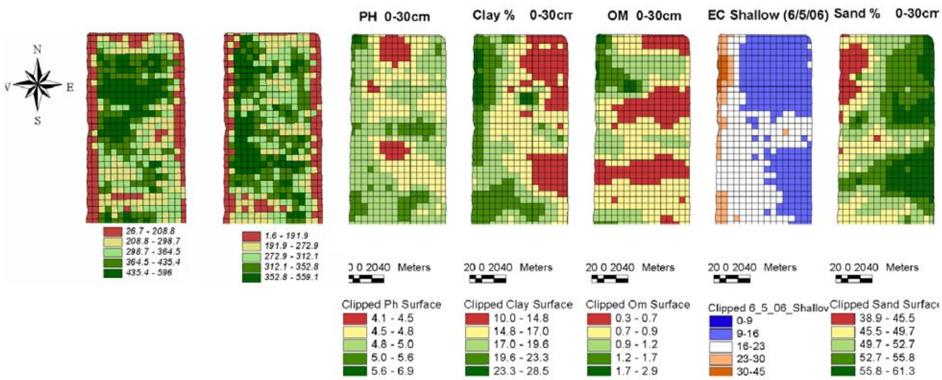

Figure 5. Two of the yield (years 2001 and 2003) and some of the soil properties maps.

the crop were comparable to the real measurements. The initial values of concepts for each case are derived from the real measurements for corresponding concepts of the yield assessment.

First case: In this case, the initial fuzzy values of the concepts (as they have been measured and converted to corresponding fuzzy sets), are the following:

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
Very low	high	med	low	med	med	low	high	low	med	high

The initial vector for the first case of yield production is:

$A^1 = [0.1\ 0.75\ 0.7\ 0.4\ 0.5\ 0.5\ 0.2\ 0\ 0.3\ 0.5\ 0.7\ 0]$,

representing the real data of the physical process (after thresholding), and the initial value of yield production was put equal to zero. These values are used in Equation (1) to calculate the equilibrium region of the process. After 11 iteration steps, the FCM reaches an equilibrium point where the values do not change any more from their previous ones, that is:

$A^{fin\_1} = [0.7201\ 0.75\ 0.7548\ 0.4\ 0.5\ 0.5\ 0.2\ 0\ 0.739\ 0.5\ 0.7\ 0.8226]$

Figure 6 depicts the subsequent values of calculated concepts for every simulation step. It is observed that the final value of concept C_{12} is 0.8226, which means that, in this region, the yield is less than the approximate value 0.85 that has been chosen as the threshold value to achieve desired results. Actually, for this case, the measured yield production was low so that the derived result of the FCM model is the expected one according to the initial measurements.
The concept "yield" takes three values, either *low* when its numerical value is less than the threshold of 0.83, *medium* when its numerical value is between 0.83 and 0.87 and *high* when the calculated numerical weight is greater than 0.87.

Second case: Here the initial measured fuzzy values of the concepts are the following:

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
verylow	low	med	low	med	med	Low	high	low	med	high

Thus, the initial vector for this case is:

A^2 = [0.1 0.25 0.2 0.4 0.5 0.5 0.2 0 0.3 0.5 0.7 0],

representing the real measured data (after defuzzification). These values and the initial weights are used in Equation (1) to calculate the equilibrium region of the process. After 12 iteration steps, the equilibrium region is reached:

A^{fin_2} = [0.7201 0.25 0.7548 0.4 0.5 0.5 0.2 0 0.739 0.5 0.7 0.8539]

In this case, it is observed that the value of concept C_{12} ("yield") in its final state is 0.8539, which means that, in this region, the yield is approximately 0.85 and equal to the *medium* yield. Actually for this case, the measured yield production was medium, thus the derived result is the expected one.

Third case: For this case, the initial fuzzy values of the concepts have been selected from real measurements. Actually, for this case, the measured yield production was high:

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
low	high	med	low	med	med	low	high	med	med	high

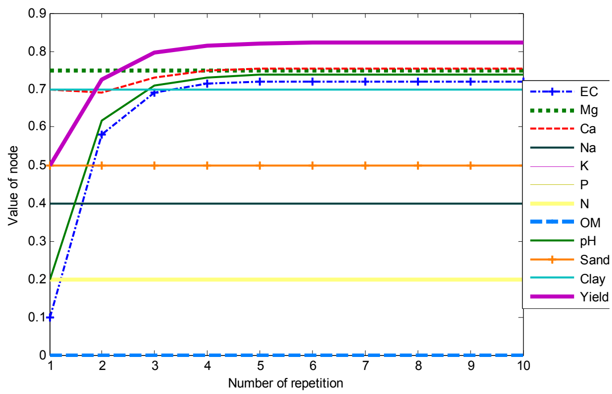

Figure 6. Subsequent values of concepts for first case till convergence

The initial vector with the concept values is:

A^3 = [0.4 0.75 0.7 0.4 0.5 0.5 0.2 1 0.5 0.5 0.7 0.5],

representing the real data of the physical process. Then, using Equation (1), the FCM simulates and, after 12 iteration steps, the equilibrium region is reached in vector:

A^{fin_3} = [0.8073 0.75 0.7548 0.4 0.5 0.5 0.2 0 0.739 0.5 0.7 0.8824]

It is observed that the value of concept C_{12} in its final state is 0.8824, which means the yield is higher than 0.87 so that it is considered *high*. The derived result is the expected according to the real measurements.

The FCM simulation model was tested for all the available 360 cases using the data for 2001in order to calculate the average accuracy of the yield production. For these experiments, two categories of *low* and *high* yield respectively were considered. For decision making reasons, a threshold value has been selected equal to 0.85 to discriminate the two yield categories- *low* and *high*. This means that if the calculated output values of yield are lower than 0.85 then the produced yield is *low* and *vice versa*. The average accuracy for 2001 is almost 74% which is efficient for this first trial using FCMs. For the 182 cases of low yield, 135 were characterized as low yield and the rest as high yield, and for the 178 cases of high yield, 131 were characterized as high yield and the others as low yield.

The same FCM model to estimate the yield output for the years 2003 and 2006 was used with the same threshold value. The results for the three years are gathered in Table 2.

It is important to address here the limitations of the proposed model that have been considered: the pH values at every point of the specified field were well below medium in the acid region and only the specific fuzzy sets were used and the weights of the FCM have not been trained using any training algorithm. Another limitation is that this model is a general one that estimates the intrinsic yield potential for each part of the field according to the soil property values and using the experts' experience and knowledge. As has been mentioned in a previous study (Gemtos *et al.*, 2004), the difference in yield spatial distribution could be attributed to the weather conditions of each year. The high degree of complexity in this problem requires the input of factors related to weather (rainfall, temperature, growing degree days, etc.).

The results of the FCM-model are very promising; our model achieved prediction of the cotton yield production of about 70% average success for the three years. This FCM-based processing approach will be further investigated in order to achieve better results, by using learning algorithms to fine-tune the causal relationships of the FCM model.

Conclusions

In this work, a new modelling and simulation approach based on Fuzzy Cognitive Maps was proposed for the first time to address the issue of crop yield prediction. The main goal of this work was not to propose a new classification technique for soil data analysis to improve accuracy, but

Table 2. Average accuracy for three years (2001, 2003 and 2006).

Accuracy/year	2001	2003	2006
Low yield	(135/182):74.18%	(126/185):67.57%	(123/174):70.69%
High yield	(131/178):73.6%	(117/175):66.86%	(130/186):69.89%
Average accuracy	73.8%	67.2%	70.3%

to propose a new modelling approach for the complex process of precision farming, using the FCM tool. The proposed soft computing technique is an advanced knowledge representation and processing method that can handle the main characteristics and site-specific management behaviour of the cotton crop providing an interpretable and transparent model.
In future work, we are going to further extend the FCM model to work with different soil properties (for example alkaline soils). That could lead to an advanced model which could estimate the yield production of every field. For this reason, it is required to choose specific interfaces for filtering the initial values and adapt dynamically the weights of each factor interaction.

References

Ambuel, J.R., Colvin, T.S. and Karlen, D.L. 1994. A Fuzzy logic yield simulator for prescription farming. Transactions of the ASAE 37(6) 1999-2009.

Gemtos, A.T., Markinos, Ath., Toulios, L., Pateras, D. and Zerva, G. 2004. Precision Farming applications in Cotton Fields of Greece. 2004 CIGR international conference, Beijing, China, 10-14 October 2004, CD-ROM.

Jang, J. S. R., Sun, C. T. and Mizutani, E. 1997. Neuro-Fuzzy & Soft Computing, N.J. Prentice-Hall, Upper Saddle River, USA.

Khan, M.S. and Khor, S.W. 2004. A framework for fuzzy rule-based cognitive maps, PRICAI 2004, LNAI 3157, Springer Verlag, Auckland New Zealand. pp. 454-463.

Kosko B. 1986. Fuzzy Cognitive Maps, International Journal of Man-Machine Studies 24 65-75.

Liu, J., Goering, C.E. and Tian, L., 2001. A neural network for setting target corn yields. Transactions of the ASAE 44 (3) 705-713.

Lund, E.D., Christy, C.D. and Drummond, P.E. 1999. Practical applications of soil electrical conductivity mapping. In: Proceedings of the 2nd European Conference on Precision Agriculture, ed. J.V. Stafford, Sheffield Academic Press, Sheffield, UK. pp. 771-779.

Markinos, A.T., Gemtos, T.A., Pateras, D., Toulios, L., Zerva, G. and Papaeconomou, M. 2004. The influence of cotton variety in the calibration factor of a cotton yield monitor. In: Proceedings of 2nd Hellenic Association of Information and Communication Technology in Agriculture (HAICTA) conference, Thessaloniki, Greece, Vol. 2; pp. 65-74.

Miao, Y. and Liu Z. 2000. On Causal Inference in Fuzzy Cognitive Maps. IEEE Transactions onFuzzy Systems 8 107-119.

Miao, Y., Mulla, D.J. and Robert, P.C. 2006. Identifying important factors influencing corn yield and grain quality variability using artificial neural networks, Precision Agriculture 7 117–135

Papageorgiou, E.I. and Groumpos, P.P. 2005. A weight adaptation method for fine-tuning Fuzzy Cognitive Map causal links. Soft Computing 9 846-857.

Papageorgiou, E., Stylios, C. and Groumpos, P. 2003. An Integrated Two-Level Hierarchical Decision Making System based on Fuzzy Cognitive Maps (FCMs), IEEE Transactions Biomedical Engineering 50 (12) 1326-1339.

Papageorgiou, E.I., Stylios, C.D. and Groumpos, P.P. 2004. Active Hebbian Learning to Train Fuzzy Cognitive Maps. International Journal of Approximate Reasoning 37 219-249.

Peláez, C.E. and Bowles, J.B., 1996. Using fuzzy cognitive maps as a system model for failure modes and effects analysis. Information Sciences 88 177-199.

SSToolbox for agriculture. 2004. User Guide, N. Country Club Rd. Stillwater, OK, USA.

Stylios C. and Groumpos P. 1999. A Soft Computing Approach for Modelling the Supervisor of Manufacturing Systems, Journal of Intelligent and Robotic Systems 26 389–403.

Stylios, C. and Groumpos, P. 2004. Modeling Complex Systems using Fuzzy Cognitive Maps. IEEE Transactions on Systems, Man & Cyberernetics 34 159-165.