Computers and Electronics in Agriculture 91 (2013) 19–29

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Yield prediction in apples using Fuzzy Cognitive Map learning approach

E.I. Papageorgiou [a,*], K.D. Aggelopoulou [b], T.A. Gemtos [b], G.D. Nanos [b]

[a] Department of Informatics and Computer Technology, Technological Educational Institute (TEI) of Lamia, 3rd Km Old National Road Lamia-Athens, 35100 Lamia, Greece
[b] University of Thessaly, Department of Agriculture, Crop Production & Rural Environment, Fytoko Street, N. Ionia, Magnesia 38446, Greece

ARTICLE INFO

Article history:
Received 15 May 2012
Received in revised form 3 November 2012
Accepted 13 November 2012

Keywords:
Fuzzy Cognitive Maps
Modeling
Knowledge representation
Fuzzy sets
Decision making
*Malus*domestica*

ABSTRACT

This work investigates the yield modeling and prediction process in apples (cv. Red Chief) using the dynamic influence graph of Fuzzy Cognitive Maps (FCMs). FCMs are ideal causal cognition tools for modeling and simulating dynamic systems. They gained momentum due to their simplicity, flexibility to model design, adaptability to different situations, and easiness of use. In general, they model the behavior of a complex system, have inference capabilities and can be used to predict new knowledge. In this work, a data driven non-linear FCM learning approach was chosen to categorize yield in apples, where very few decision making techniques were investigated. Through the proposed methodology, FCMs were designed and developed to represent experts' knowledge for yield prediction and crop management. The developed FCM model consists of nodes linked by directed edges, where the nodes represent the main soil factors affecting yield, [such as soil texture (clay and sand content), soil electrical conductivity (EC), potassium (K), phosphorus (P), organic matter (OM), calcium (Ca) and zinc (Zn) contents], and the directed edges show the cause-effect (weighted) relationships between the soil properties and yield. The main purpose of this study was to classify apple yield using an efficient FCM learning algorithm, the non-linear Hebbian learning, and to compare it with the conventional FCM tool and benchmark machine learning algorithms. All algorithms have been implemented in the same data set of 56 cases measured in 2005 in an apple orchard located in central Greece. The analysis showed the superiority of the FCM learning approach in yield prediction.

1. Introduction

Soil variability within a farm exists in most soils and regions. This variability interacts with weather, inputs (which practically cannot be applied homogeneously) and the variability of genetic material to produce crop and yield (quantitative and qualitative) variability. These different spatial or temporal variabilities have to be properly managed by the farmers to achieve the best profit with the lowest inputs, thus reducing the adverse environmental effects. Precision agriculture aims at better managing the variability. Several data are collected during each growing period over the years and stored in databases to assist in this management. The problem is that historical data cannot always predict yield variability and final yield. It seems that yield variability is canceling out after 3 years. This was the case in the data reported by Blackmore et al. (2003) for cereals and Fountas et al. (2004) for cotton. It is quite possible to have more stable zones of high and low yielding in perennial crops like orchards. But, in most farms and regions, alternate bearing is encountered in apple orchards and, to a lesser extend, many other fruit crops (Childers et al., 1995). This unpredictable yield and yield variation makes the successful application of variable rate of inputs difficult. It also causes a failure of the development of decision support systems (DSS) for precision agriculture, which is considered one of the reasons of the relatively low adoption.

Precision agriculture applications in fruits and vegetables are rather limited in the literature. Extensive work was reported in citrus (Shumann et al., 2006; Sakai et al., 2007; Maja and Ehsani, 2010; Mann et al., 2011), and vineyards (Bramley et al., 2003), but it is very limited in other crops. In particular, Aggelopoulou et al. (2010, 2011a, 2011b) have worked on precision agriculture applications in apples, Fountas et al. (2011) in olives, and Konopatzki et al. (2009) in pears. Such data would be useful for site-specific yield prediction calculation with proper methods.

Yield prediction in apples, and fruit trees in general, is very important, because it could be used to improve crop management and plan fruit marketing. Furthermore, when yield is predicted site-specifically, the inputs (water, fertilizers, pesticides) and field operations (e.g. fruit thinning) could be applied with variable rates depending on the real needs of the trees in the different areas of the field. As most crop models used did not successfully predict yield and yield spatial variation, different authors have used other parameters to predict them during the growing season. The earlier

* Corresponding author.
E-mail address: epapageorgiou@teilam.gr (E.I. Papageorgiou).

0168-1699/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.compag.2012.11.008

the prediction, the better it can be used to improve orchard management. In apples, analysis of multispectral images of the fruits (Alchanatis et al., 2007) was used for yield prediction. Digital images of the trees at full bloom (Aggelopoulou et al., 2011b) were used to predict yield variability.

Besides, a large number of approaches like crop models, algorithms and statistical tools have been proposed and used for yield prediction in precision agriculture using historical data. Correlation and multiple linear regression (MLR) have been commonly used to predict yield and identify important factors influencing yield (Kravchenko and Bullock, 2000; Park et al., 2005; Gutiérrez et al., 2008; Huang et al., 2010), but the results are not so encouraging due to the existence of polynomial and interaction terms, which were not considered (Kitchen et al., 2003). In the case of MLR analysis, the description of linear relationships between crop parameters and site variables is limited, and the results may be misleading when these relationships are not linear (Kitchen et al., 2003; Schultz et al., 2000; Miao et al., 2006). Another common approach is combining multivariate techniques, like principal component analysis (PCA) and factor analysis (FA), with multiple regressions (Jiang and Thelen, 2004; Huang et al., 2010; Fortin et al., 2010). These methods were used to minimize the problem caused by interacting variables, facilitate the interpretation of complex relationships, reduce the dimensionality of the dataset or select a subset of appropriate variables from a large data set (Huang et al., 2010). Subsequent attempts have been made by applying artificial intelligence principles and soft computing techniques in precision agriculture for spatial analysis and crop management (Drummond et al., 2003; Savin et al., 2007; Jutras et al., 2009; Huang et al., 2010).

Artificial neural networks (ANNs), as no-linear statistical techniques, have also been applied to investigate yield response to soil variables (Jutras et al., 2009; Fortin et al., 2010; Park et al., 2005; Effendi et al., 2010). Specifically, ANN analysis has been applied in precision agriculture for spatial analysis and crop management (Kitchen et al., 2003; Drummond et al., 2003; Liu et al., 2001; Irmak et al., 2006). In the case of ANNs, the observed dataset of the selected variables is fitted to describe the problem by adjusting the weights of linkages connecting input and output variables and can be regarded as multivariate non-linear analytical tools. The ANNs can be combined with other artificial intelligence techniques or other statistical methods to benefit from the advantages of ANN modeling, and to also avoid some of their limitations such as the need for large amounts of data for training. Huang et al. (2010) summarized the soft computing techniques and their applications in agricultural and biological engineering.

An alternative soft computing technique, the fuzzy cognitive mapping (FCM), could be used for yield prediction and is used in our case study for apple yield prediction. FCM is a method for analyzing and depicting human perception of a given system with the development of a conceptual model, which is not limited by exact values and measurements (Kosko, 1986). The advantageous modeling features of FCMs, such as simplicity, adaptability and capability of approximating abstractive structures, encourage us to use them for complex problems. They gained momentum due to their dynamic characteristics and learning capabilities (Salmeron, 2009). The learning approaches for FCMs are concentrated on learning an adjacency matrix, based either on expert intervention and/or on the available historical data. According to the available type of knowledge, the learning techniques can be categorized into three groups: Hebbian-based, population-based and hybrid combining the main aspects of Hebbian-based and evolution-based type learning algorithms. The most used learning approaches in the literature involve the non-linear Hebbian learning (NHL) algorithm and the genetic algorithm learning (Papageorgiou, 2012), which are the most efficient in FCM training.

FCMs have been widely used in many different scientific fields such as engineering, business and management, environment, medicine and telecommunications (Papageorgiou, 2012). In agriculture, the FCM methodology with its learning capabilities was applied in cotton yield prediction (Papageorgiou et al., 2009, 2011), producing also a modeling tool for helping farmers make decisions in precision agriculture (Papageorgiou et al., 2011). The aim of the present study was to construct the FCM model for classifying yield in apples based on experts' knowledge and then to use efficient learning approach to train this model with field data and therefore exploit yield predictions. The NHL method was used for yield classification and its inference capabilities were compared with the FCM tool without learning and with the most used and known machine learning algorithms. It was shown that the NHL method for FCMs gives better prediction accuracies than those obtained with the use of conventional FCMs and machine learning techniques.

2. Material and methods

2.1. The data

The present study was carried out in a commercial apple orchard located in Agia area, central Greece (22°35′33″E, 39°40′28″N) in a 5 ha field in 2005. The main cultivar was Red Chief grafted on MM106 with Golden Delicious as pollinator. The trees were planted at 3.5 m × 2 m, trained as free palmette and intensively cultivated including regular irrigation and fertilization, winter and summer pruning and precise hand thinning 2 weeks after petal fall. This work is a part of an experiment that lasted 3 years (2005–2007) including soil, yield, and quality mapping. Yield was consistent over the two of the 3 years of the study (2005 and 2006). In the third year (2007) yield was lower than the two previous years because a number of trees died out (Aggelopoulou et al., in press). There were no sights of alternate bearing in the bearing trees due to proper weather conditions during fruit set, the intensive cultivation practices applied and the low vegetative vigor of apple cv. Red Chief. In this paper the data of year 2005 are presented.

For yield mapping, apples were collected manually in September 2005 and placed in plastic bins along the tree rows at commercial harvest time (Fig. 1). Yield per 10 trees was weighted and the geographical position in the center of the 10 trees was recorded using a hand-held computer with GPS (Trimble pathfinder).

In December 2005, 20 soil samples were taken before winter crop fertilization to a sampling depth of 0–30 cm. The sampling positions were geo-referenced using a hand-held computer with GPS. The samples were air-dried, passed through a 2 mm sieve and analyzed for the following properties: soil texture (% sand, %

Fig. 1. Orchard under study with the harvesting bins placed along the rows. Apples from groups of 10 adjacent trees were weighed to create yield maps.

silt and % clay), phosphorus (P), exchangeable calcium (Ca), exchangeable potassium (K), available zinc (Zn), and organic matter (OM) concentration.

In December 2005, a Veris machine was used to measure the apparent soil electrical conductivity (ECa) at depths 0–30 cm and 0–90 cm and produce relevant maps.

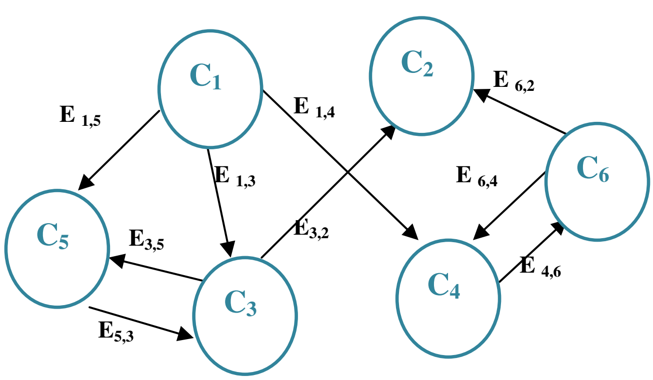
All data were interpolated on a 30×30 m grid, which corresponds to a reliable field management unit, in order to create the maps. Yield and ECa were interpolated using kriging and the rest of the data using nearest neighbor interpolation.

2.2. The model Fuzzy Cognitive Map

2.2.1. Fuzzy Cognitive Map main aspects

Fuzzy Cognitive Map can be conveniently represented either by a graph, which is easily understood by a human, or by a connection matrix, which is useful for computational purposes when running simulations. In the graph representation, concepts are represented as nodes and relationships are depicted by directed edges between the nodes. Each edge is also associated with a number (weight) that quantifies the strength of the corresponding relationship. In the matrix representation, concepts are represented by successive rows/columns and cells store all relationships' values. Fig. 2 shows a generic representation of the FCM model.

For the purposes of this paper, we define FCM as an order pair $\langle C, E \rangle$, where C is the set of labels and E is the connection matrix. Every label $A_i \in C$ is mapped to its activation value $A_i \in [0, 1]$, where 0 means no activation, and 1 means full activation. The labels from C can be interpreted as linguistic terms (Kosko, 1986; Stylios and Groumpos, 2004) that point to fuzzy sets. In such case, the activation value A_i is interpreted as the value of fuzzy membership function that measures the degree in which an observed value belongs to the fuzzy set pointed by the related term. The other, simplified interpretation of C can be such that the labels C_i denote the real valued variables, the domains of these variables are assumed as normalized into the [0, 1] interval. Note that the latter, simplified



	C_1	C_2	C_3	C_4	C_5	C_6
C_1	0	0	$E_{1,3}$	$E_{1,4}$	$E_{1,5}$	0
C_2	0	0	0	0	0	0
C_3	0	$E_{3,2}$	0	0	$E_{3,5}$	0
C_4	0	0	0	0	0	$E_{4,6}$
C_5	0	0	$E_{5,3}$	0	0	0
C_6	0	$E_{6,2}$	0	$E_{6,4}$	0	0

Fig. 2. Example of FCM model: (a) FCM graph, and (b) *Adjacency* connection matrix.

interpretation of C, in most cases, does not influence the computational methods that stand behind the reasoning process based on FCM (Sharif and Irani, 2006).

The matrix W does not change in time and stores the weights assigned to the pairs of concepts. The weights represent the generalized (over a given period of time) causal dependency between the concepts. The weights assume the values $E_{ij} \in [-1, 1]$, where the value of weight $E_{ij} = 1$ expresses the full positive and $E_{ij} = -1$ the full negative impact of ith causal-concept on jth effect-concept. The intermediate values of weight refer to partial causality.

The development and design of the appropriate FCM for the description of a system requires the contribution of human knowledge. Usually, knowledgeable experts, who are familiar with the FCM formalism, are required to develop FCM using an interactive procedure of presenting their knowledge on the operation and behavior of the system.

Experts are asked to determine the concepts that best describe the model of the system, since they know which factors are the key principles and functions of the system operation and behavior, and they introduce a concept for each one. Experts have observed the operation and behavior of the system during its operation, since they are the operators and supervisors of the system, who control it using their experience and knowledge. They have stored in their mind the correlation among different characteristics, states, variables and events of the system and in this way they have encoded the dynamics of the system using fuzzy if-then rules. Each fuzzy rule infers a fuzzy weight, which in procedure is translated to a numerical one used in the FCM reasoning process. The procedure described here is an approach based in previous works for constructing FCMs (Papageorgiou et al., 2003; , 2006, 2011).

In most cases, FCMs are constructed manually, and, thus, they cannot be applied when dealing with large number of variables. In such cases, their development could be significantly affected by the limited knowledge and skills of the knowledge engineer. Therefore, it is essential to use data-driven learning algorithms to accomplish this task. Diverse adaptive and evolutionary-based learning methods (Papageorgiou et al., 2003, , 2004; Papageorgiou and Groumpos 2005a,b; Stach et al., 2005), have been explored by various authors and these types of learning algorithms have been proved in the literature as efficient ones for training FCMs (Papageorgiou et al., 2011, 2005; Papageorgiou and Groumpos, 2005a,b; Stach et al., 2005; Papageorgiou, 2011a). A recent review study on learning algorithms for FCMs and their advanced applications was published by Papageorgiou (2011b).

2.2.2. Fuzzy Cognitive Map inference process

Once the FCM is constructed, it can receive data from its input concepts, perform reasoning and infer decisions as values of its output concepts (Papageorgiou et al., 2008, 2011). During reasoning the FCM iteratively calculates its state until convergence. The state is represented by a *state vector* A^k, which consists of real node values $A_i^{(k)} \in [0, 1]$, $i = 1,2,\ldots,N$ at an iteration k. The value of each node is calculated by the following equation:

$$A_i(k+1) = f((2A_i(k) - 1) + \sum_{\substack{j \neq i \\ j=1}}^{N} (2A_j(k) - 1) \cdot E_{ji}) \quad (1)$$

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 step k, e_{ji} is the weight of the interconnection between concept C_j and concept C_i and f is a sigmoid threshold (activation) function that transform the values of concepts in the range [0, 1] (Bueno and Salmeron, 2009).

This rescaled simulation process removes the spurious influence of inactive concepts (with $C_i = 0$) on other concepts, and avoids the conflicts emerging in cases where the initial values of concepts are 0 or 0.5. Also, the insufficient knowledge and/or missing information for each node can be handled with less deviation from reality. The iteration stops when a limit vector is reached, i.e., when $\boldsymbol{A}^k = \boldsymbol{A}^{k-1}$ or when $\boldsymbol{A}^k - \boldsymbol{A}^{k-1} \leq e$; where e is a residual, whose value depends on the application type (and, in most applications, is equal to 0.001). Thus, a final vector ***A_f*** is obtained.

2.2.3. Non-linear Hebbian learning algorithm

The Hebbian paradigm is perhaps the best-known unsupervised learning theory in connectionism (Papageorgiou et al., 2004, Papageorgiou et al., 2006). Hebbian-based learning algorithms, such as Active Hebbian Learning (AHL) and NHL were proposed for learning the weight matrix of FCMs based on experts' intervention (Papageorgiou et al., 2006; Papageorgiou and Groumpos, 2005a; Stach et al., 2008). Since FCMs have non-linear structure of their concepts, the non-linear Hebbian learning has already been used to train FCMs for classification of the autistic disorder (Kannappan et al., 2011). In this algorithm, the learning rule for FCMs integrates a learning rate parameter η_k, weight decay parameter γ, and the decision output concepts. When the NHL algorithm is applied, only the initial non-zero weights suggested by the experts are updated for each iteration step. All the other weights of weight matrix e_{ji} remain zero, which is their initial value. The steps of the NHL algorithm for FCM training task to predict the class of apple yield are described as follows:

Algorithm **FCM–NHL()**
Input: Concept values in vector ***A*** and weight matrix ***E*** (*initial weight matrix assigned by experts*)

Process:
Step 1: Initialize the input values of eight concepts A_i, the weights E_{ji} for each interconnection given by the experts, determine learning rate parameter $\eta_k = 0.001$, weight decay parameter $\gamma = 0.98$ and the three target values related with each one of the three categories, $T_{i\min} \leqslant T_i \leqslant T_{i\max}$, where $i = 1$–3. For this work, one decision output concept (DOC) is assigned to express the output category, which can be categorized in the classes [i.e. low yield (T1), medium yield (T2) and high yield (T3)] and thus produce the output of the system
Step 2: For each iteration step *k go to next step*
Step 3: Update the weights according to Eq. (2). Only the non-zero weights are updated

$$e_{ji}^{(k)} = \gamma \cdot e_{ji}^{(k-1)} + \eta A_i^{(k-1)}(A_j^{(k-1)} - sgn(e_{ji}^{(k-1)})e_{ji}^{(k-1)}A_i^{(k-1)}) \tag{2}$$

Only the non-zero weights are updated
Step 4: Assign $\mathbf{A}(k+1)$ using the formulation presented in Eq. (1). Calculate values $A_i^{(k+1)}$ using Eq. (3) for the candidate weight set $e_{ji}^{(k)}$, until a fixed state is reached
Step 5: Evaluate stopping conditions, using $A_{24}^{(k+1)}$ and $A_{24}^{(k)}$ from Step 4, and candidate $\boldsymbol{E}(k)$
Step 6: When one of the two termination conditions is met, go to step 2
Step 7: Return the produced $\mathbf{E}(k)$, which is the final one. (In this step, we also check the termination conditions for weights, where the weights do not change any more their values.)
Check the value $A_{24}^{(k)}$ for the DOC_i and categorize it in one of the three related classes (low, medium and high)
Stopping conditions: Either the conditions 1 or 2 has to be satisfied to terminate the iterative process
Condition 1: Calculate $F1 = \sqrt{\sum_{i=1}^{m}(DOC_i - T_i)^2}$
where m is the number of DOCs and T_i defines the target value which can be calculated as $T_i = (T_{i\min} + T_{i\max})/2$ where $i = 1$–3
Condition 2: Calculate $F2 = |DOC_i^{(k+1)} - DOC_i^{(k)}| < e$
where e takes a value of 0.001 (Papageorgiou and Groumpos, 2005a,b)

The termination condition that is usually accomplished in all cases is the fixed-point state and in this fixed-state we check if the calculated value of DOCi (for Concept A_{24}) meets one of the three restrictions (T-targets) and thus we finally classify this pattern in one of the three classes. If the produced class is the same with the initial one, then this pattern is classified correctly.

2.2.4. Development of FCM model for yield variability prediction in apples

According to the FCM development process, the number and kind of concepts were determined by a group of experts. Three experts were used: one soil scientist from Technological Education Institute of Larisa, a pomologist from the University of Thessaly and a pomologist from the Zagora Cooperative of apple growers. Zagora is well-known high-elevation area producing high quality apples in Central Greece. The experts' answers were pooled together to determine the main factors of the model, which represented concepts as well as the interconnections among concepts. Each interconnection was described by the experts using if-then rules that infer a fuzzy linguistic variable from a determined set *T*{*influence*}, which associates the relationship between the two concepts and determines the grade of causality between the two concepts. The knowledge of experts can be easily encoded to a number of if-then rules, which is later used for the determination of adjacency weight matrix.

The experts help in designing the FCM model following the step by step approach described in what follows.

2.2.4.1. Step I: Determination of concepts and fuzzy membership functions for each one concept. The problem of yield prediction in apples was described using a number of factors that mainly affect apple fruit production. At first, the experts stated that there are eight main factors – variables that determine apple fruit yield (Table 1). Each one of the variables represents soil properties and has three or five fuzzy values. Then, using a questionnaire which was created for this process that was completed by the team of experts, the fuzzy linguistic weights among these concepts (which describe the strengths of connections among concepts) were assigned.

The quantitative values of concepts measured in the orchard under study were classified into qualitative values, using the FuzMe toolbox. This toolbox is used to delineate management zones in the field. Management zones are defined as zones with similar properties which can be managed homogeneously. This toolbox clusters the input data into classes for each one factor according to the real measurements and then these classes are assigned as fuzzy (see examples in Fig. 4). The depicted fuzzy values for each one concept, using the defuzzification method of Center of Area, were transformed into the range [0, 1], thus were able to be used in FCM simulation process (Papageorgiou et al., 2011). Fig. 3 illustrates the corresponding membership functions for the three soil parameters, shallow EC, Clay and K content, and the apple yield.

Table 1
Concepts of the FCM model.

Concepts	Description: soil properties measured over 0–300 mm soil depth	Range of measured values	Type and number of scaled values
C_1: Shallow EC	Shallow soil electrical conductivity (mS m^{-1})	*5.1–59.5*	*Five fuzzy*
C_2: Ca	The exchangeable calcium in the soil (mg kg^{-1})	*125–492*	*Three fuzzy*
C_3: K	The exchangeable potassium in the soil (mg kg^{-1})	*60–355*	*Five fuzzy*
C_4: OM	The organic matter content in the soil (%)	*1.1–3.2*	*Three fuzzy*
C_5: P	The measured phosphorus in the soil (mg kg^{-1})	*0.5–9.2*	*Five fuzzy*
C_6: Zn	The available zinc in the soil (mg kg^{-1})	*0.8–2.7*	*Five fuzzy*
C_7: Clay	The clay content in the soil (%)	*15.1–31.3*	*Three fuzzy*
C_8: Sand	The sand content in the soil (%)	*51.7–65.6*	*Three fuzzy*
C_9: Yield	Apple yield (Mg ha^{-1})	*0–91.2*	*Three fuzzy*

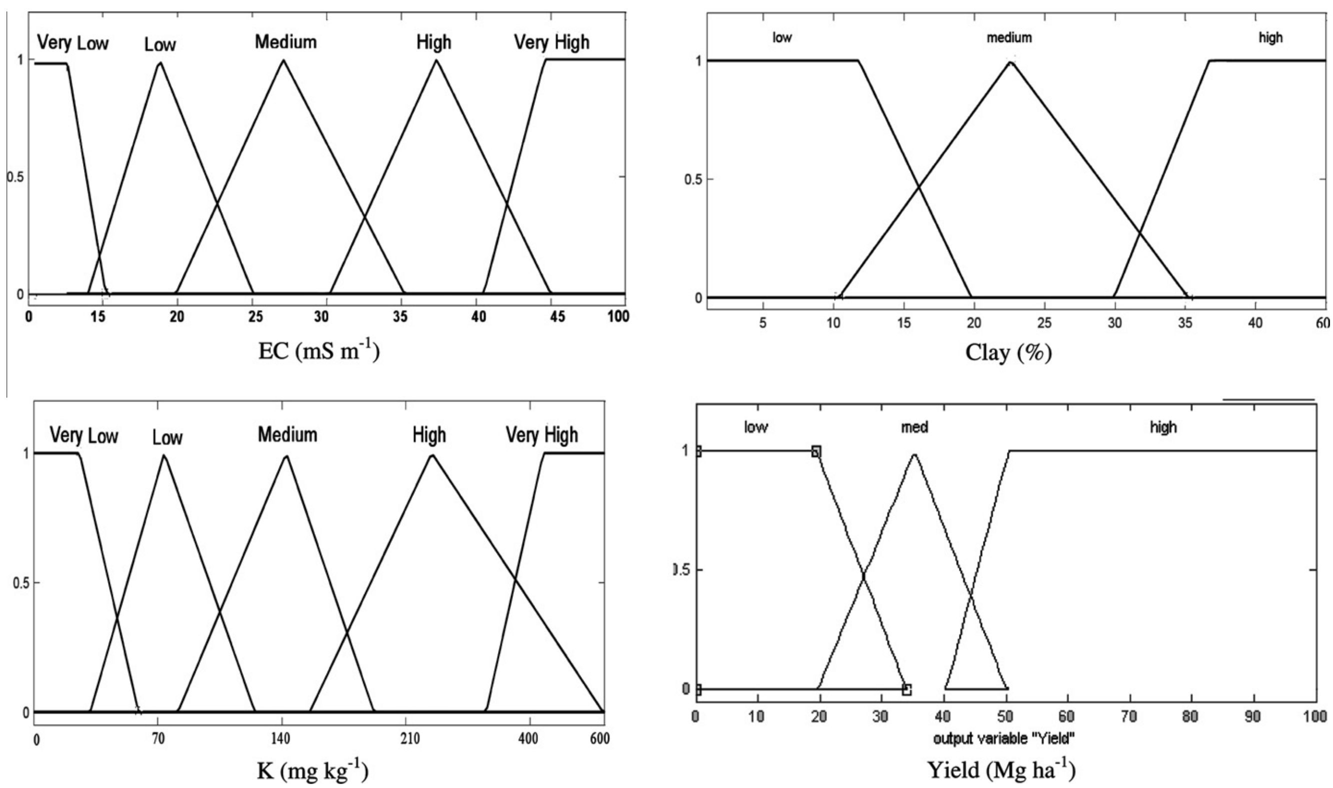


Fig. 3. Membership functions for some soil properties (ShallowEC, K, Clay) and Yield.

2.2.4.2. Step II: Fuzzy rules and linguistic terms that describe the influences among concepts. The FCM model for describing spatial variation in apple yield is constructed by experts using their knowledge to describe the relationships between concepts. The three experts assigned relationships for the FCM model, and determined the relationships among all the nine factors. They were asked to describe the degree of influence from one concept to other using if-then rules among factor concepts and yield from the term set *T{influence}*.

In the case of fuzzy rules for the specific problem of yield description in apples, some examples, as derived from experts, are given:

IF a **small** change occurs in the value of concept C_1 (Shallow EC), **THEN** a **small** change in the value of concept C_9 (yield) is caused.

This means that: the influence from concepts C_1 to C_9 is **low**.

IF a **high** change occurs in the value of concept C_6 (Potassium), **THEN** a **high** change in the value of concept C_9 (yield) is caused.

This means that: the influence from concepts C_6 to C_9 is **high**.

IF a **medium** change occurs in the value of concept C_7 (Clay), **THEN** a **negative medium** change in the value of concept C_9 (yield) is caused.

This means that: the influence from concepts C_7 to C_9 is **negatively medium**.

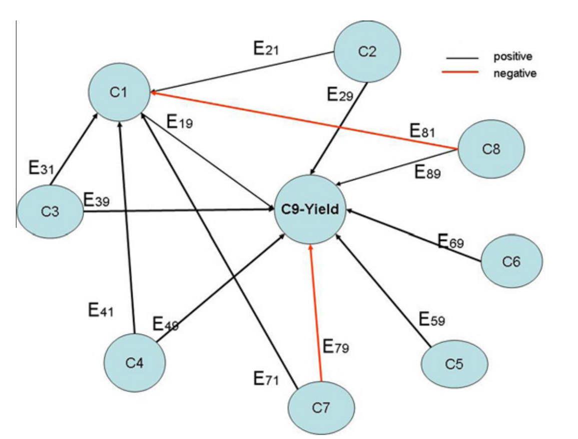


Fig. 4. The FCM tool illustrating the problem of apple fruit yield prediction.

2.2.4.3. Step III: Determination of numerical weights. The inferred fuzzy linguistic weights (strengths of influences) are combined using the SUM method as described by Zadeh (1976). Then, using the defuzzification method of centroid (Zadeh, 1976), the linguistic weights y with membership functions $\mu(y)$ are transformed into a numerical weight e_{ji} which lies in the range $[-1,1]$:

$$e_{ji} = \frac{\int_y y \cdot \mu(y)dy}{\int_y \mu(y)dy}$$

Table 2 gathers the produced numerical weights of the FCM model.

The weights are gathered into a weight matrix ($\boldsymbol{E} = (E_{ji})$) 9×9, where 9 is the number of concepts. The produced FCM model for modeling and predicting yield in apples including the initial calculated values of weights is shown in Fig. 4.

Thus, the FCM is an abstract conceptual model depending on the three experts who decide about the input and output decision concepts as well as for the causal relationships among them. All the eight concepts are considered as factor concepts by the experts as they are listed in Table 1 to design the FCM model and each concept represents three, or five fuzzy values. In this problem, the concept C_9 has been considered from the experts as Decision Output Concept (DOC) and could be categorized as low yield (“low”), medium yield (“med”) and high yield (“high”).

The measured data were stored in a two-dimensional matrix that represents the spatial distribution of every factor in the field. Each cell of the matrix corresponds to an area of 30×30 m which is the spatial resolution of the yield data model. The data of every cell of filtered maps represent the data that will be used as input variables in the proposed FCM model. Each one vector of the data record represents the initial concept values of the proposed FCM model that interact through the FCM simulation process until an equilibrium point or decision is reached.

The measured data used were categorized by the FCM tool using the NHL algorithm into three yield categories. For comparison purposes with FCM learning tool, the most often used machine learning techniques applied in a large number of scientific fields are used. WEKA toolbox is an easy to use tool containing the most known classification techniques, such as decision trees, conjunction rules, *k*-nn, Bayesian networks and artificial neural networks (ANNs). In this paper, we assume readers’ familiarity with machine learning methods and it is not worthwhile for the paper’s readability to present them analytically. A brief presentation of each one of the techniques is given in the literature (Witten and Frank, 1999).

3. Results

After the construction of FCM tool for predicting yield in apples, 56 cases were considered for training FCM tool using NHL algorithm for the evaluation of the model. Thus, the decision making capabilities of the proposed FCM tool were presented by simulating these cases and finding the predicted outcomes according to the available data. An important part of fuzzy inference system construction is to check that the obtained outputs match the real measured yield. Therefore, according to this approach, the rate of “correct” decisions, i.e. the decisions complying with the yield category, is considered as the classification accuracy of the compared models.

In the next subsection, the simulation results of two representative case studies for apple fruit yield in Central Greece apple farms are presented. These case studies were derived from the existing dataset. In each of the test cases, we have an initial vector **A**, representing the data at a given time of the process, and a final vector **A_f**, representing the last state that we can arrive at.

The initial values of concepts are transformed into the range [0,1], with quantification based on the fuzzy sets theory in order to be used for the simulation of FCM (the related fuzzy sets have been assigned by the FuzMe toolbox). The FCM simulates at first through the inference algorithm described in Section 2.2.2, and, then, using the FCM–NHL learning algorithm described in Section 2.2.3. For this FCM model, the values of concepts are changed at the same time units and referred to as an iteration step. New values for concepts are calculated until the FCM/FCM–NHL tools predicting yield in apples reached an equilibrium point (steady state region for the specific set of concepts), where the values of concepts do not change any more from their previous ones. Eq. (1) can be applied to find the equilibrium final state after the creation of FCM. This procedure is based on the determination of the value A_9 of output concept C_9 “yield” which works as decision concept (DOC) that represents the apple production measured in the corresponding pixel. The DOC was initially categorized as low, medium and high, taking real measured values of yield for each category such as $0 \leqslant$ “low” $\leqslant 20$ Mg ha^{-1}, 20 Mg ha$^{-1} \leqslant$ “medium” $<$ 50 Mg ha^{-1} and “high” $\geqslant 50$ Mg ha^{-1}. After training, the ranges of the “yield” concept values for the three classes were changed in [0,1] in order to calculate the classification accuracy for each category, thus $0 \leqslant$ “low” yield $\leqslant 0.25$, $0.251 \leqslant$ “medium” yield < 0.50 and $0.501 \leqslant$ “high” yield $\leqslant 1$ (Papageorgiou et al., 2008).

Using the FCM–NHL algorithm for the developed FCM model and for the 56 data records of the specific field, a new FCM model is produced establishing new weighted relationships among concepts. Due to the learning algorithm characteristics, only the initial non-zero values of weights update their values after training.

3.1. Simulation results of two cases examined by FCM and FCM–NHL tool

Two different cases of “med” and “high” yield categories were examined to evaluate the proposed methodology based on FCM and NHL learning FCM tools for determining category of yield.

In the first case, a “med” yield pixel 30×30 from the production area was selected, with the following real measured values for each factor; EC = 32.989 mS m^{-1} (C_1), Ca = 125 mg kg^{-1} (C_2), K = 151 mg kg^{-1} (C_3), OM = 1.69% (C_4), P = 5.5 mg kg^{-1} (C_5), Zn = 0.98 mg kg^{-1} (C_6), Clay = 15.1% (C_7) and Sand = 63.1% (C_8). The real value of yield was 31.714 Mg ha^{-1} at this pixel. These numerical values of measured soil parameters are transformed to corresponding fuzzy sets, normalized and then, using defuzzification, the initial activated values of concepts are produced. The value of EC corresponds to the fuzzy set “high” which is transferred to normalized numerical initial value $A_1 = 0.66$. The measured value of Ca corresponds to the fuzzy set “very low” and transferred to numerical value $A_2 = 0.1$. The same happens with all other concepts. K has a “med” value which is transferred to $A_4 = 0.34$, OM has a “medium” value which is transferred to $A_6 = 0.43$, P has a “low” value which is transferred to $A_3 = 0.16$, Zn has a “very low”

Table 2
Initial weight matrix assigned by experts. The zero values mean that we do not consider interactions.

Concepts Ci/Cj	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9-yield
C_1	0	0	0	0	0	0	0	0	0.25
C_2	0.8	0	0	0	0	0	0	0	0.6
C_3	0.2	0	0	0	0	0	0	0	0.2
C_4	0.2	0	0	0	0	0	0	0	0.52
C_5	0	0	0	0	0	0	0	0	0.52
C_6	0	0	0	0	0	0	0	0	0.43
C_7	0.65	0	0	0	0	0	0	0	−0.6
C_8	−0.3	0	0	0	0	0	0	0	0.3
C_9	0	0	0	0	0	0	0	0	0

value which is transferred to $A_5 = 0.12$, Clay has a "med" value which is transferred to $A_7 = 0.32$ and sand has a "med" value which is transferred to $A_8 = 0.59$. The initial value of yield production is equal to 0.345 and it describes a "med" yield category. The initial fuzzy qualitative values of the concepts which correspond to "med" yield production are represented at the following vector of concept values:

$$\mathbf{A}^{\text{first}} = [0.66\ 0.1\ 0.34\ 0.43\ 0.16\ 0.12\ 0.32\ 0.59\ 0.345]$$

This vector represents the numerical values of measured soil parameters of the physical process, after transformation, normalization to (0,1) and correspondence to numerical values. Then, using, the FCM inference process as described in Section 2.2.3, the equilibrium region is reached after 10 iteration steps and the calculated values of concepts are depicted at the following vector:

$$\mathbf{A\_FCM}^{\text{first}} = [0.1525\ 0.1000\ 0.3400\ 0.4300\ 0.1600\ 0.1200\ 0.3200\ 0.5900\ 0.1243]$$

The output of this FCM approach reveals that this case is classified as "low" yield, and does not concur with the initial category/state.

For the same input vector, using the FCM–NHL algorithm, the resulting output vector is produced.

$$\mathbf{A\_NHL\_FCM}^{\text{first}} = [0.672\ 0.111\ 0.340\ 0.432\ 0.161\ 0.119\ 0.320\ .591\ 0.4996]$$

From the calculated DOC value (0.4996), which prescribes the output "yield" of the FCM–NHL approach, it is assigned that the yield is categorized as "med" and it is the same with its initial category. The produced weight matrix is illustrated in Table 3. The algorithm performance keeps the initial zero values of weights due to non-consideration of influences.

It is concluded, from the produced weight matrix after FCM–NHL algorithm, that almost all the weights have changed their values. Some of them show smaller and others higher changes from their initial ones. Generally, all the weights were modified to lower values. Examples of some weights (E_{ij}), where we observe minor changes, without any significant difference to the strength of influence among concepts, are the E_{19} [where a small modification to the initial weight value occurred after training, the value 0.25 (Table 2) was modified to 0.2059 (Table 3)], E_{29}, (where also a small modification to the initial weight value occurred after training, the value 0.6 was modified to 0.4744), E_{31}, E_{41}, E_{39}, E_{49}, E_{59}, E_{71}, E_{81}, E_{79} and E_{89}. We observe that approximately two causal influences-weights have changed their values at an important level. These influences (weights) are: E_{21}, which describes that an initial high influence among C_2 and C_1 (weight = 0.8) has changed to a medium influence between these concepts ($E_{21} = 0.62$), and causal influence E_{69}, which means that an initial medium influence among C_6 and C_9 (weight = 0.45) has changed to a low influence between these concepts ($E_{69} = 0.34$). It is observed that only in second technique of DDNHL-based FCM, the case of "med" yield was classified correctly as "med" yield.

In the second case, a high yield pixel from the production area was selected, with the following real measured values for each factor; EC = 8.034 mS m^{-1} (C_1), Ca = 276 mg kg^{-1} (C_2), K = 354.8 mg kg^{-1} (C_3), OM = 1.8% (C_4), P = 1.98 mg kg^{-1} (C_5), Zn = 1.15 mg kg^{-1} (C_6), Clay = 17.3% (C_7) and Sand = 64% (C_8) and yield = 59.46 Mg ha^{-1} which is categorized as "high" yield. These numerical values of measured soil parameters are transformed to corresponding fuzzy sets, then defuzzified and produce the following vector $\mathbf{A}^{\text{second}}$ of numerical values. The initial value of yield production is considered equal to 0.75. The initial fuzzy qualitative values of the concepts, which correspond to "high" yield production, are represented at the following vector of concept values:

$$\mathbf{A}^{\text{second}} = [0.16\ 0.25\ 0.77\ 0.45\ 0.057\ 0.144\ 0.37\ 0.708\ 0.657]$$

Then, using Eqs. (1) and (2), the FCM tool simulates, the equilibrium region is reached after 10 iteration steps and the calculated values of concepts are depicted in the following vector:

$$\mathbf{A\_FCM}^{\text{second}} = [0.2479\ 0.2500\ 0.7700\ 0.4500\ 0.0570\ 0.1440\ 0.3700\ 0.7080\ 0.1843].$$

The calculated value of DOC (yield) is 0.1843 which assigns a "low" yield value for this specific case. This case was not classified correctly using the FCM tool.

For the same input vector, using the FCM–NHL algorithm, the resulted output vector is produced, which is:

$$\mathbf{A\_NHL\_FCM}^{\text{second}} = [1.00\ 1.00\ 0.659\ 1.00\ 0.659\ 1.00\ 0.659\ 1.00\ 0.659].$$

In the case of FCM–NHL algorithm, the yield is categorized as "high" and it concurs with the real value. Thus, for this case, the NHL-based FCM learning approach gave correct results. The produced weight matrix is illustrated in Table 4.

3.2. Prediction results

The FCM tool and the FCM–NHL algorithm for training the FCM tool were evaluated for 56 cases using the measured data of soil parameters of year 2005 to calculate the average classification accuracy of the crop yield production. For these experiments, three categories of "*low*", "*med*" and "*high*" yield were considered for the prediction process.

All the 56 cases were used for the training of FCM model, using the proposed algorithm FCM-NHL. The real measurement data gathered for the specific field for year 2005 were given in Aggelopoulou et al. (2011a). The 10-fold cross validation was used only for the FCM collective model to calculate the confusion matrix after the yield prediction. Table 5 presents the confusion matrices of the FCM tool without learning, whereas Table 6 gathers the results of confusion matrices and overall accuracies from FCM–NHL

Table 3
Produced weight matrix calculated by FCM–NHL algorithm for first case of medium yield.

Concepts Ci/Cj	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9-yield
C_1	0	0	0	0	0	0	0	0	0.2059
C_2	0.6268	0	0	0	0	0	0	0	0.4744
C_3	0.1633	0	0	0	0	0	0	0	0.1637
C_4	0.1639	0	0	0	0	0	0	0	0.4158
C_5	0	0	0	0	0	0	0	0	0.4124
C_6	0	0	0	0	0	0	0	0	0.3433
C_7	0.5133	0	0	0	0	0	0	0	−0.4670
C_8	−0.2312	0	0	0	0	0	0	0	0.2395
C_9	0	0	0	0	0	0	0	0	0

Table 4
Produced weight matrix calculated by yield FCM–NHL algorithm for second case of high yield.

Concepts Ci/Cj	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9-yield
C_1	0	0	0	0	0	0	0	0	0.1983
C_2	0.61837	0	0	0	0	0	0	0	0.4659
C_3	0.16004	0	0	0	0	0	0	0	0.1600
C_4	0.16201	0	0	0	0	0	0	0	0.4069
C_5	0	0	0	0	0	0	0	0	0.4046
C_6	0	0	0	0	0	0	0	0	0.3371
C_7	0.5072	0	0	0	0	0	0	0	−0.4534
C_8	−0.2243	0	0	0	0	0	0	0	0.2372
C_9	0	0	0	0	0	0	0	0	0

Table 5
Confusion matrix for FCM tool without learning.

	Low	Med	High
Low	3	4	2
Med	1	29	1
High	2	11	3
Correctly classified	35/56		
Accuracy	62.50%		

algorithm and for the different examined classification techniques we ran for comparison of the results.

The results of the FCM model after learning with NHL algorithm were compared with some of the well-known machine learning techniques such as decision tree learner C4.5 (Release 8) (Quinlan, 1990), the multi layer perceptron (MLP) neural network back propagation training program (Duda et al., 2001), the Naïve Bayes classifier, the recurrent backpropagation (RBF) network, the conjunctive rules, the association rule Zero-R, the multiclass classifier, and the *k*-means algorithm. The WEKA *(Waikato Environment for Knowledge Analysis)* toolbox (WEKA, 2003) was used to test these eight machine learning algorithms for the same data set of 56 cases.

Considering the decision trees, the J48 algorithm, an implementation of C4.5 release 8 (Quinlan, 1990), was used to test data sets and to categorize yield into three categories. For the Bayesian networks, the Naïve Bayes classifier was implemented to categorize the yield data to three output nodes (''*low*'' yield, ''*med*'' yield and ''*high*'' yield), (Berthold and Hand, 2003). The back propagation algorithm for a multi-layer perceptron to categorize the yield data to three output nodes (''low'' = 1, ''med'' = 2 and ''high'' = 3) was used. The multilayer perceptron used in our case study has nine input nodes, one hidden layer and three output nodes (describe each decision).

According to the 10-fold cross-validation procedure, the dataset of the 56 entries was randomly divided into 10 subsets. Then, each subset was used to test the performance of the classifiers trained on the union of the remaining nine subsets. This procedure was repeated 10 times, choosing a different part for testing each time. The results of the 10-fold cross validation of each machine learning algorithm are illustrated in Table 7 (Witten and Frank, 1999).

4. Discussion of results

The aim of the FCM model was to predict the apple yield in three categories (low, medium and high yield) based on each of the eight factors using simulated data from real measurements. Through the literature, more studies have used statistical analysis techniques and only a few have used computational intelligent algorithms to classify and predict yield from large datasets. Most of these studies were in arable crops (Liu et al., 2001; Miao et al., 2006; Shearer et al., 1999; Canteri et al., 2002), but no work has been done in apples. For comparison purposes, we employed eight classification techniques such as decision trees, MLP NNs with back propagation algorithm, Naïve Bayes (Bayesian networks), the RBF network, the conjunctive rules, the association rule Zero-R, the multiclass classifier, and the *k*-means algorithm Ibk, on the same data set of year 2005 to estimate the classification accuracy into three yield classes (Witten and Frank, 1999; Quinlan, 1990; Jang et al., 1997). All methods are machine learning techniques that efficiently accommodate numerical data as well as categorical or symbolic data (Jang et al., 1997).

For this approach, the classification accuracy for all machine learning algorithms was measured by 10-fold cross-validation using a simple discriminant method on the values of output concept ''yield'' that represents the category of yield in apples. The NHL algorithm was run with standard configuration (Stach et al., 2008).

The results from the application of FCM-based techniques (FCM and FCM–NHL approaches) and those obtained by the other eight classification techniques (Table 6) indicated that the FCM-based approaches outperform the machine learning algorithms. More specifically, the proposed learning tool based on FCM–NHL, produced the best results for year 2005 with overall prediction rate of 75%. The rest machine learning techniques gave much lower overall prediction rates than FCM-based methodologies.

It is concluded that, based on the overall accuracy of each method, the FCM–NHL algorithm for training FCMs outperformed when compared with the FCM tool without learning and the benchmark machine learning approaches, showing its superiority in classification task. Thus, the proposed FCM learning simulation model is able to estimate yield with reasonably high overall accuracy, sufficient for this specific application area.

Next, the classification performance of each algorithm was tested using mean absolute error (MAE) index, root mean square error (RMSE) index, classification rate (CR), which is defined as the percentage of the correctly classified patterns over the entire testing set, and True Positive (TP) rate, which is defined as the percentage of all true positive rates of the three categories.

The objective function that is being minimized in all cases is the Mean Absolute Error (MAE) between the actual and the desired output concepts of the prediction tools. The MAE performance index is defined as follows:

Table 6
Classification results from each examined approach for yield data.

Results/technique		FCM–NHL			DTs (J48)			BNs (Naïve)		
		Low	Med	High	Low	Med	High	Low	Med	High
Confusion Matrix	Low	4	5	–	1	7	1	1	6	2
	Med	1	28	2	2	28	1	8	15	8
	High	–	6	10	3	12	1	3	12	1
Accuracy		75% (42/56)			53.5714% (30/56)			30.36% (17/56)		
		RBF network			MLP NNs			Zero R conjunction		
Confusion Matrix	Low	–	9	–	4	3	2	–	9	–
	Med	–	31	–	3	17	11	–	31	–
	High	–	16	–	3	9	4	–	16	–
Accuracy		55.36% (31/56)			44.64% (25/56)			55.36% (31/56)		
		Conjunctive Rules			Multiclass Classifier			Ibk *k*-means		
Confusion Matrix	Low	–	9	–	–	9	–	2	4	3
	Med	–	31	–	–	31	–	3	18	10
	High	–	16	–	–	16	–	2	9	5
Accuracy		55.36% (31/56)			55.36% (31/56)			44.64% (25/56)		

Table 7
Classification performance analysis of the studied prediction tools.

	FCM–NHL	FCM	DTs	RBF	BNs	Conjunctive rules	ZeroR classifier
MAE	0.2545	0.4821	0.3803	0.3932	0.4624	0.389	0.3945
RMSE	0.2886	0.4564	0.4746	0.4456	0.5512	0.4451	0.4431
TP rate (All)	0.65756	0.48544	0.35900	0.33333	0.21933	0.33333	0.33333
Classification accuracy	75.00%	62.50%	53.37%	55.37%	30.35%	55.37%	55.37%

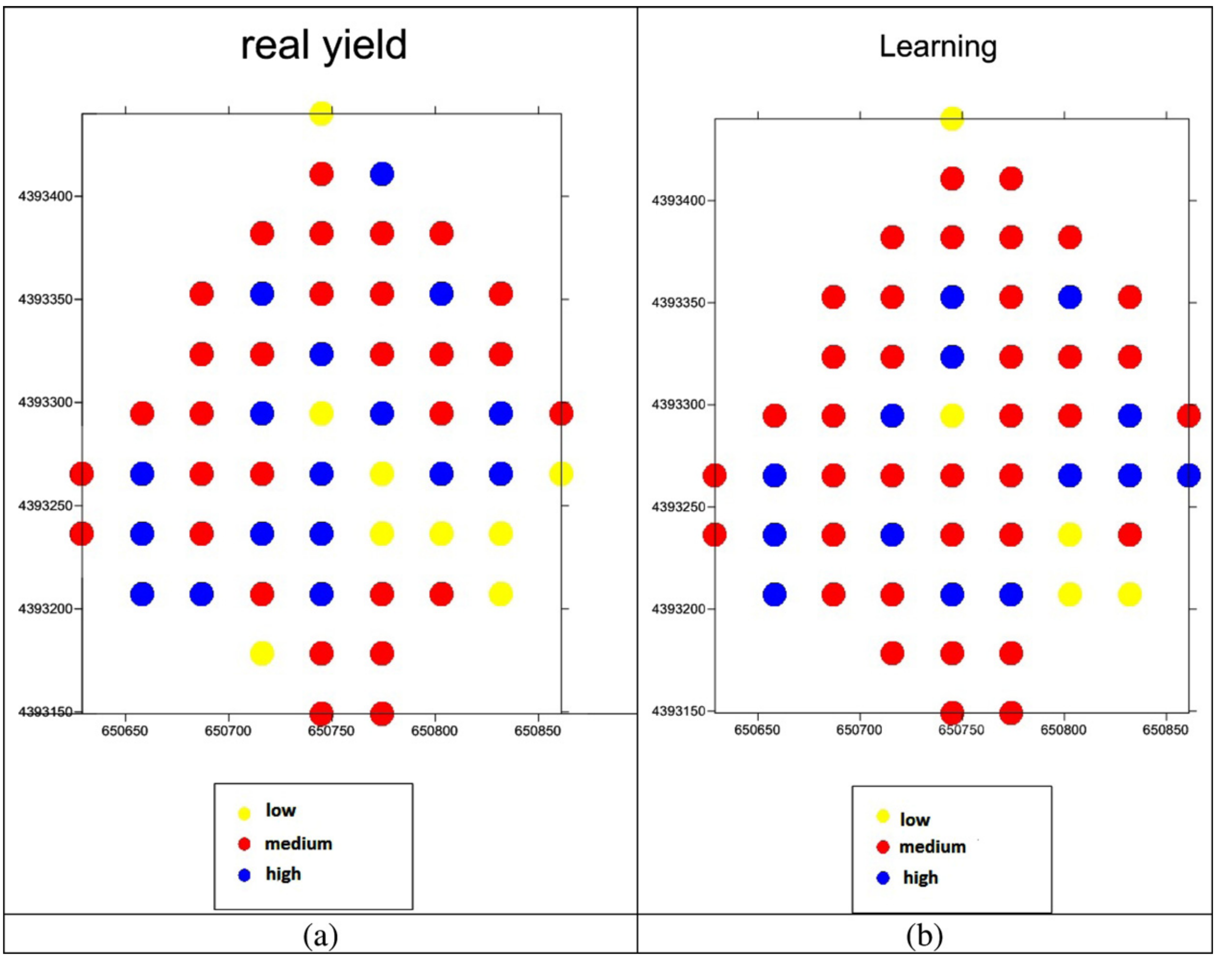


Fig. 5. Maps of real yield (a) and predicted yield with FCM–NHL (b).

$$\text{MAE} = \frac{1}{M}\left(\sum_{k=1}^{M}\sum_{l=1}^{N}|DOC_{lk}^{\text{Real}} - DOC_{lk}^{\text{Predicted}}|\right)$$

where M is the number of training sets ($N = 56$), N the number of system output concepts ($N = 3$) and $(DOC_k^{\text{Real}} - DOC_k^{\text{Predicted}})$ is the difference between the lth decision output concept (DOC) and its corresponding real value (target), when the kth set of input concepts appears to the tool's input.

The RMSE performance index is defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{MN}\left(\sum_{k=1}^{M}\sum_{l=1}^{N}(DOC_{lk}^{\text{Real}} - DOC_{lk}^{\text{Predicted}})^2\right)}$$

where M and N are the number of training sets and the system outputs, respectively.

The classification analysis was performed and the results depicted in Table 6 show that the DDNHL algorithm for FCM tool (second column in bold face) outperforms the other prediction tools, giving better classification rate of 75%, with smaller training error.

In order to test the performance of the two different approaches, FCM and NHL algorithm for FCM, a statistical analysis was explored. We implemented the well known t-test paired two samples of mean among FCM–NHL tool and FCM without learning to show if the results are statistically significant or not. This study is built on the hypothesis that there is a significant difference between the FCM–NHL approach for yield prediction system, and the FCM of collective experts' experience. In the case of FCM/FCM–NHL, both methods produced different mean and standard deviations. Calculated t-value was 0.024 for one tail and 0.048 for two-tail, which was not significant at both one-tail critical (1.67) and two-tail critical (2.004), respectively. It is observed from the statistical analysis, that the FCM–NHL outperforms the FCM without learning and the results are statistically significant.

The strongest point of the FCM methodology is the insight it can provide on the role of key feedbacks in the system. These feedbacks can remain hidden and can be uncovered by applying a tool such FCM. Also, FCM represents a system in a form that closely corresponds to the way humans perceive it. Therefore, the model is easily understandable, even by a non-technical audience and each parameter has a perceivable meaning. The resulting fuzzy model was used to analyze, simulate, test the influence of parameters and predict the behavior of the system. Further experiments on more cases are necessary to validate our observations. In Fig. 5 the maps of real yield and predicted yield using FCM–NHL are presented.

The prediction of the FCM–NHL model of apple yield proved correct for 42 of the 56 cases. This prediction is based on eight easily measured parameters proposed by the experts. Having this information available early in the growing season, two management options are applicable. To adjust the cultivation practices to the predicted yield thus saving inputs or to manipulate inputs to achieve the values the experts believe that would result in improved yield. Obviously this is not the case for ECa or soil texture or even for organic matter that cannot change in the short run, although the latter can be improved in the long run. But other parameters can be changed (P, K, Ca) during the growing period and achieve the proper level to possibly improve yield.

5. Conclusions

The approach used in this work focuses on the soft computing technique of Fuzzy Cognitive Maps enhanced by the NHL training method for the estimation of yield category in apples with respect to precision agriculture aspects. The presented solution has been raised by some of the requirements imposed by the targeted application: the causal association of soil parameters on yield prediction in apples that seem to be crucial for the right yield classification. We have also sketched in this paper the exemplary problem of prediction and its simulation using the proposed solution. The main objective of this work was to present a method based on FCM learning technique to develop a computational intelligent tool for categorizing apple fruit yield. The results have shown that the FCM learning approach predicts properly the phenomenon, gives a front-end decision about the class of apple fruit yield, and provides similar results to those obtained from horticulturist experts. Modeling with this knowledge-based tool resembles closely to the way horticulturist experts perceive it. Therefore, the suggested model is easily understandable and each parameter has a perceivable meaning.

References

Aggelopoulou, K., Wulfsohn, D., Fountas, S., Gemtos, T.A., Nanos, G.D., Blackmore, S., 2010. Spatial variability of yield and quality in a small apple orchard. Precision Agriculture 11 (5), 538–556.
Aggelopoulou, K.D., Pateras, D., Fountas, S., Gemtos, T.A., Nanos, G.D., 2011a. Soil spatial variability and site-specific fertilization maps in an apple orchard. Precision Agriculture 12, 118–129.
Aggelopoulou, A.D., Bochtis, D., Fountas, S., Swain, K.C., Gemtos, T.A., Nanos, G.D., 2011b. Yield prediction in apple orchards based on image processing. Precision Agriculture 12, 448–456.
Aggelopoulou, K.D., Castrignanò, A., Gemtos, T.A., De Benedetto, D., 2012. Delineation of management zones in an apple orchard in Greece using a multivariate approach. Computers and Electronics in Agriculture 90, 119–130.
Alchanatis, V., Safren, O., Levi, O., Ostrovsky, V., 2007. Apple yield imaging using hyperspectral machine vision. In: 6th European Conf. on Precision Agriculture, Skiathos, Greece, pp. 555–562.
Berthold, M., Hand, D.J., 2003. Intelligent Data Analysis, second ed. Springer-Verlag, Berlin.
Blackmore, S., Godwin, R., Fountas, S., 2003. The analysis of spatial and temporal trends in yield map data over six years. Biosystems Engineering 84, 455–466.
Bramley, R.G.V., Pearse, B., Chamberlain, P., 2003. Being Profitable Precisely – A Case Study of Precision Viticulture from Margaret River. Australian & New Zealand Grape grower & Winemaker – Annual Technical Issue, 473a, pp. 84–87.
Bueno, S. and Salmeron, J.L. (2009). Benchmarking main activation functions in Fuzzy Cognitive Maps, Expert Systems with Applications, 36 (3 part. 1), 5221–5229.
Canteri, M., Avila, B.C., Dos Santos, E.L., Sanches, M.K., Kovaleschyn, D., Molin, J.P., and Gimenez, L.M., 2002. Application of data mining in automatic description of yield behavior in agricultural areas. In: Proc. of the World Congress of Computers in Agriculture and Natural Resources, Iguacu Falls, Brazil, pp. 183–189.
Childers, N.F., Morris, J.R., Sibbett, G.S., 1995. Modern Fruit Science, 10th ed. Horticultural Publications, Gainsville, FL.
Drummond, S.T., Sudduth, K.A., Joshi, A., Birrell, S.J., Kitchen, N.R., 2003. Statistical and neural methods for site-specific yield prediction. Transactions of the ASAE 46 (1), 5–14.
Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification, second ed. J. Wiley & Sons Inc., New York.
Effendi, Z., Ramli, R., Ghani, J.A., 2010. A back propagation neural networks for grading *Jatropha curcas* fruits maturity. American Journal of Applied Sciences 7 (3), 390–394.
Fortin, J.G., Anctil, F., Parent, L.É., Bolinder, M.A., 2010. A neural network experiment on the site-specific simulation of potato tuber growth in Eastern Canada. Computers and Electronics in Agriculture 73 (2), 126–132.
Fountas, S., Aggelopoulou, K., Bouloulis, C., Nanos, G.D., Wulfsohn, D., Gemtos, T.A., 2011. Site-specific management in an olive tree plantation. Precision Agriculture 12, 179–195.
Fountas, S., Blackmore, S., Gemtos, T.A., Markinos, T., 2004. Trend yield maps in Greece and the UK. In: Proc. of the 2nd HAICTA Conference, Thessaloniki, Greece, 18–20 March, 2, 65–74.
Gutiérrez, P.A., López-Granados, F., Peña-Barragán, J.M., Jurado-Expósito, M., Hervás-Martínez, C., 2008. Logistic regression product-unit neural networks for mapping *Ridolfia segetum* infestations in sunflower crop using multitemporal remote sensed data. Computers and Electronics in Agriculture 64 (2), 293–306.
Huang, Y., Lan, Y., Thomson, S.J., Fang, A., Hoffmann, W.C., Lacey, R.E., 2010. Development of soft computing and applications in agricultural and biological engineering. Computers and Electronics in Agriculture 71, 107–127.
Irmak, A., Jones, J.W., Batchelor, W.D., Irmak, S., Boote, K.J., Paz, J.O., 2006. Artificial neural network model as a data analysis tool in precision farming. Transactions of the ASABE 49 (6), 2027–2037.
Jang, J.S.R., Sun, C., Mizutani, E., 1997. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Upper Saddle River. Prentice-Hall International, Inc., NJ.
Jiang, P., Thelen, K.D., 2004. Effect of soil and topographic properties on crop yield in a north-central corn-soybean cropping system. Agronomy Journal 96, 252–258.
Jutras, P., Prasher, S.O., Mehuys, G.R., 2009. Prediction of street tree morphological parameters using artificial neural networks. Computers and Electronics in Agriculture 67, 9–17.
Kannappan, A., Tamilarasi, A., Papageorgiou, E.I., 2011. Analyzing the performance of Fuzzy Cognitive Maps with non linear hebbian learning algorithm in predicting autistic disorder. Expert Systems with Applications 38, 1282–1292.
Kitchen, N.R., Drummond, S.T., Lund, E.D., Sudduth, K.A., Buchleiter, G.W., 2003. Soil electrical conductivity and topography related to yield for three contrasting soil-crop systems. Agronomy Journal 95, 483–495.
Konopatzki, M.R.S., Souza, E.G., Nobrega, L.H.P., Uribe-Opazo, M.A., Suszek, G., Rodrigues, S., et al., 2009. Pear tree yield mapping. Acta Horticulturae 824, 303–312.
Kosko, B., 1986. Fuzzy cognitive maps. International Journal of Man-Machine Studies 24 (1), 65–75.
Kravchenko, A.N., Bullock, D.G., 2000. Correlation of corn and soybean grain yield with topography and soil properties. Agronomy Journal 92 (1), 75–83.
Liu, J., Goering, C.E., Tian, L., 2001. A neural network for setting target corn yields. Transactions of the ASAE 44 (3), 705–713.
Mann, K.K., Schumann, A.W., Obreza, T.A., 2011. Delineating productivity zones in a citrus grove using citrus production, tree growth and temporally stable soil data. Precision Agriculture 12 (4), 457–472.
Maja, J.M., Ehsani, R., 2010. Development of a yield monitoring system for citrus mechanical harvesting machines. Precision Agriculture 11, 475–487.
Miao, Y., Mulla, D.J., Robert, P.C., 2006. Identifying important factors influencing corn yield and grain quality variability using artificial neural networks. Precision Agriculture 7 (2), 117–135.
Papageorgiou, E.I., Stylios, C., Groumpos, P., 2003. An integrated two-level hierarchical decision making system based on fuzzy cognitive maps (FCMs). IEEE Transactions on Biomedical Engineering 50 (12), 1326–1339.
Papageorgiou, E., Stylios, C.D., Groumpos, P.P., 2004. Active hebbian learning algorithm to train fuzzy cognitive maps. International Journal of Approximate Reasoning 37 (3), 219–249.
Papageorgiou, E.I., Groumpos, P.P., 2005a. A new hybrid learning algorithm for Fuzzy Cognitive Maps learning. Applied Soft Computing 5, 409–431.
Papageorgiou, E.I., Groumpos, P.P., 2005b. A weight adaptation method for fine-tuning Fuzzy Cognitive Map causal links. Soft Computing Journal 9, 846–857.
Papageorgiou, E.I., Stylios, C.D., Groumpos, P.P., 2006. Unsupervised learning techniques for fine-tuning Fuzzy Cognitive Map causal links. International Journal of Human-Computer Studies 64, 727–743.
Papageorgiou, E.I., Spyridonos, P., Glotsos, D., Stylios, C.D., Groumpos, P.P., Nikiforidis, G., 2008. Brain tumor characterization using the soft computing technique of Fuzzy Cognitive Maps. Applied Soft Computing 8, 820–828.

Papageorgiou, E.I., Markinos, A., Gemtos, T.A., 2009. Application of fuzzy cognitive maps for cotton yield management in precision farming. Expert Systems with Applications 36 (10), 12399–12413.
Papageorgiou, E.I., Markinos, A., Gemtos, T.A., 2011. Soft computing technique of Fuzzy Cognitive Maps to connect yield defining parameters with yield in cotton crop production in central Greece as a basis for a decision support system for precision agriculture application. In: Fuzzy Cognitive Maps: Advances in Theory, Methodologies, Tools, Applications, edited by M. Glykas, Springer Verlag, Berlin.
Papageorgiou, E.I., 2011a. A new methodology for decisions in medical informatics using Fuzzy Cognitive Maps based on Fuzzy Rule-Extraction techniques. Applied Soft Computing 11, 500–513.
Papageorgiou, E.I., 2011b. A review study of FCMs applications during the last decade. In: Proc. of IEEE International Conference of Fuzzy Systems (FUZZ IEEE), Taipei, Taiwan, pp. 828–835.
Papageorgiou, E.I., 2012. Learning algorithms for Fuzzy Cognitive Maps – a review study. IEEE Transactions on Systems, Man, and Cybernetics (SMC) – Part C 42 (2), 150–163, <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5766766>.
Park, S.J., Hwang, C.S., Vlek, P.L.G., 2005. Comparison of adaptive techniques to predict crop yield response under varying soil and land management conditions. Agricultural Systems 85, 59–81.
Quinlan, J.R., 1990. Decision trees and decision making. IEEE Transactions on System, Man, and Cybernetics 20 (2), 339–346.
Salmeron, J.L., 2009. Supporting decision makers with Fuzzy Cognitive Maps. Research-Technology Management 52 (3), 53–59.
Schultz, A., Wieland, R., Lutze, G., 2000. Neural networks in agroecological modeling-stylish application or helpful tool? Computers and Electronics in Agriculture 29, 73–97.
Sharif, A.M., Irani, Z., 2006. Exploring Fuzzy Cognitive Mapping for IS evaluation. European Journal of Operational Research 173, 1175–1187.
Savin, Y., Stathakis, D., Negre, T., Isaev, V.A., 2007. Prediction of crop yields with the use of neural networks. Russian Agricultural Sciences 33, 361–363.
Shearer, S.A., Thomasson, J.A., Mueller, T.G., Fulton, J.P., Higgins, S.F., Samson, S., 1999. Yield Prediction using a Neural Network Classifier Trained using Soil Landscape Features and Soil Fertility Data, ASAE, Paper No. 993042, St. Joseph, MI, USA.
Shumann, A.W., Miller, W.M., Zaman, Q.U., Hostler, K.H., Buchanon, S., Cugati, S., 2006. Variable rate granular fertilization of citrus groves: spreader performance with single-tree prescription zones. Applied Engineering in Agriculture 22 (1), 19–24.
Stach, W., Kurgan, L.A., Pedrycz, W., Reformat, M., 2005. Genetic learning of Fuzzy Cognitive Maps. Fuzzy Sets and Systems 153 (3), 371–401.
Stach, W., Kurgan, L., Pedrycz, W., 2008. Data driven nonlinear Hebbian learning method for Fuzzy Cognitive Maps. In: Proc. IEEE World Congress on Computational Intelligence, Hong Kong, pp. 1975–1981.
Stylios, C.D., Groumpos, P.P., 2004. Modeling complex systems using Fuzzy Cognitive Maps. IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans 34, 155–162.
WEKA, Toolbox, 2003. Available in: http://www.cs.waikato.ac.nz/~ml/weka (last accessed 30/9/12).
Witten, I., Frank, E., 1999. Data Mining: Practical Machine Learning Tools and Techniques, second ed. Morgan Kaufmann Publishers, San Mateo, CA, USA.
Zadeh, L.A., 1976. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3–28.