Scientia Horticulturae 246 (2019) 797–804

Contents lists available at ScienceDirect

Scientia Horticulturae

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

Selection of the most important features affecting pistachio endocarp lesion problem using artificial intelligence techniques

Somayeh Sadr[a,b], Vahid Mozafari[a,*], Hossein Shirani[c,e], Hossein Alaei[d], Ahmad Tajabadi Pour[a]

[a] Department of Soil Science, College of Agriculture, Vali-e-Asr University of Rafsanjan, Iran
[b] Faculty Member of Payame Noor University of Kerman, Iran
[c] Department of Soil Science, College of Agriculture, Shahrekord University, Iran
[d] Department Plant Protection, College of Agriculture, Vali-e-Asr University of Rafsanjan, Iran
[e] Soil Science, Department of Soil Science, College of Agriculture, Vali-e-Asr University of Rafsanjan and Shahrekord University, Iran

ARTICLE INFO

Keywords:
Rafsanjan region
Fungal contamination
Feature selection
Hybrid genetic algorithm - artificial neural network
Sensitivity analysis

ABSTRACT

Pistachio endocarp lesion (PEL) problem is one of the most important causes of damage to pistachios, and so far, no research has been done on the factors influencing it. The present study is designed to investigate the biotic and abiotic agents influencing the occurrence of this problem. To study the biotic agents, six orchards having pistachio with PEL symptoms, were sampled and parts of these pistachio fruits were cultured on PDA. In studying abiotic agents, 90 orchards with symptoms of the PEL were identified in pistachio orchards of Rafsanjan and Anar. Then, from the leaves of the trees, soils in 0–40 and 40–80 cm depths, damaged branches of the trees and irrigation water were sampled. Damage level of the problem was determined in 4 classes for each tree (target variable). Then, 68 features including soil, plant and water chemical properties were measured in these samples. These features were analyzed using the Hybrid Genetic Algorithm - Artificial Neural Network and sensitivity analysis. Our results showed that the fungal factors had no effect on this problem. Among the soil features, calcium magnesium ratio, soil salinity and fruit magnesium concentration were introduced as the most important factors. Increased air temperature and reduction in rainfall, water shortage and its salinity and much extraction of underwater resources could be great factors in increasing the magnesium concentration and occurrence of the symptoms of false calcium deficiency in pistachio orchards in Rafsanjan.

1. Introduction

Pistachio (Pistacia vera) has been one of the most important exporting products of Iran for a long time and its contribution to the export of non-oil goods is about 14% (The Islamic republic of Iran Customs Administration (IRICA, 2011). There are more than 60–65% of pistachio orchards in Iran (Food and Agriculture Organization of the United Nations (FAO, 2015) and about 200–250 thousand tons of pistachio is produced in this area. Rafsanjan Township is considered as the largest pistachio producing center in Iran and exports about 45–50 thousand tons to the global markets (Ahmadi et al., 2015).

So many factors reduced the quality of this product. Besides factors such as salinity, imbalanced nutrients in the soil and water shortage, the emergence of some complications has also increased the severity of the damage. In recent years, pistachio endocarp lesion (PEL) problem was more expanded and imposed huge losses to Iranian farmers cultivating pistachio each year and the damage has been observed to be more than 50% in some orchards.

For the first time, PEL problem was observed in some parts of California pistachio orchards in 1984 and was then introduced as “stylar end lesions” (Bolkan et al., 1984; Rice et al., 1985). This problem starts from the beginning of the formation of endocarp (under weather conditions of Kerman province, early May) and continues until mid-June. In the early stage of endocarp formation, darkening of the endocarp starting from fruit tip or margin. The seeds with symptoms eventually fall down and the remaining ones often get soft and flexible in the infected section (Fig. 1).

In the early years of the occurrence, some researchers discovered this problem as a result of the attack of an insect named *Leptoglossus clypealis* in addition to the enzymes secreted from the bite of this pest (Bolkan et al., 1984). In recent decades, the symptoms of this problem were compared with those of pests like *Acrostenum* spp.، *Brachynema* spp.، *Lygaeus panderus* and *Apodiphus amygdali*. Thus, the effect of insects on this problem was not shown (Hashemi Rad and Rajabi, 1998;

* Corresponding author at: Department of Soil Science, College of Agriculture, Vali-e-Asr University of Rafsanjan, Postal Code: 7718897111, Iran.
E-mail addresses: s.sadr@stu.vru.ac.ir (S. Sadr), mozafari@vru.ac.ir (V. Mozafari), shirani@vru.ac.ir (H. Shirani), hossein.alaei@vru.ac.ir (H. Alaei), tajabadi@vru.ac.ir (A. Tajabadi Pour).

https://doi.org/10.1016/j.scienta.2018.11.048
Received 1 July 2018; Received in revised form 15 November 2018; Accepted 16 November 2018

Fig. 1. Symptoms of pistachio endocarp lesion on a pistachio cluster (Iran-Rafsanjan, May 2016).

Hashemi Rad, 2005).

Agricultural experts and farmers of the region experimentally agreed that this disease occurs as a result of the development of fungal species in high humidity conditions such as rainfall and irrigation and thus, they use fungicides and stop irrigating the orchards in the event of the disease occurrence. Zhang (2004) noted that the cause of the problem is calcium deficiency and disorientation of calcium magnesium ratio and also mentioned that the pistachios endocarp strength at the time of formation depends on the availability of calcium ion and disorientation of calcium magnesium ratio causes lack of calcium absorption by the plant. Unfortunately, despite the fact that PEL causes a high damage to pistachio production, no detailed study has been done to determine the factors affecting the formation and development of this problem. Due to the complexity and ambiguities that exist in this regard, there is necessity for closer and more complex study on the possible influencing factors of this problem.

It should be noted that, in natural problems, which have very complex and non-linear relationships, classical statistics and linear regression analyses have revealed no accurate and interpretable results. In this regard, using the artificial intelligence techniques such as combined evolutionary algorithms with a powerful modeling method like the artificial neural network would be very efficient in analyzing the complex processes.

Among all possible features influencing PEL such as chemical properties of soil, water and plant, some features have greater influence on this problem. Among the massive amounts of data, it would be necessary to use some mentioned methods in order to reduce features and select only a few of them as superior features and eliminate the rest of them so that a lot of data are not lost (Guyon and Elisseeff, 2003). As a result, the feature selection is very important. Among the goals to follow from feature selection is reduction of the amount of data, speeding up the operation, better understanding of the results and selecting the most effective features (Zhu and Chipman, 2006). Numerous studies have been done in the fields of features selection using modern methods of the artificial intelligence in various areas (Ramadan et al., 2001; Lou and Nakai, 2001; Nguyen and Won, 2016; Chouhan et al., 2016; Shirani et al., 2018). But another approach which has been focused on the feature selection studies is application of the combined models of meta-heuristic algorithms such as neural networks, particle swarm optimization, support vector machine, genetic algorithm, etc. (Tirelli and Pessani, 2011; Yang and You, 2013; Besalatpour et al., 2014; Moncada et al., 2014; Shirani et al., 2015).

Since there is no detailed and serious study on the factors affecting this destructive problem and the annual development of this disease in pistachio orchards is causing great loss to pistachio farmers and Iranian economy, this research was done with the following objectives: (1) To investigate the influence of fungal agents in the occurrence of endocarp lesion problem (2) To select the most important factors affecting the formation and development of pistachio endocarp lesion problem among the properties of the soil, water and plant using the model of the hybrid genetic algorithm - artificial neural network and (3) To determine the importance of the selected features on the formation and development of pistachio endocarp lesion problem using the sensitivity analysis method.

2. Materials and methods

2.1. Study area and sampling

This study was carried out in some parts of the pistachio orchards in Kerman province located in the center of Iran. From the beginning until the end of May 2016, 90 trees with several levels of PEL problem were randomly selected and marked. Since the dominant rootstock of pistachio trees in the study area is Badami Rize Zarand, all the trees in this study had this rootstock. To investigate the effect of the fungal factors on this problem among 90 identified trees, 15 trees that showed severe symptoms were selected and then the infected pistachio fruits of the trees were sampled. In order to maintain the samples from infection of fungus and saprophytic bacteria in the environment, they were sent to the laboratory in containers filled with ice. Geographical location of each tree and their height was recorded by GPS (model 76CSx, Gramin Co, Taiwan).

The stady area has extended within the geographical range of 55° 1’ 36.159” to 56° 6’ 31.691” to the East and 30° 10’ 35.78” to 31° 5’ 26.675” to the North. The mean altitude in the sampled points is 1498 m above sea level.

In order to assess and select the abiotic agents affecting PEL including properties of the soil, water and plant, in the range of the shading area of the tree, soil samples were taken from two depths of 0–40 and 40–80 cm of the soil level. At the same time, the leaf sample (from the central branches which had no fruits) and sample of the pistachio cluster were also taken. Water of the engine pump used to irrigate the trees was also sampled. The infection rate of the selected trees were scored in four levels of (I) = less than 10%, (II) = 11 to 35%, (III) = 36 to 60% and (IV) = more than 60%. These scores were used as a target variable in feature selection using classification method.

2.2. Investigation of the biotic agents effecting PEL problem

2.2.1. Cultivation of fruits with symptoms

For etiology of the PEL problem in terms of fungal factors, three samples were obtained from each pistachio fruit having PEL symptoms. To isolate the causal agents, the samples were cultured on to potato dextrose agar (PDA). In this way, small pieces of the infected tissue were cut and disinfected with 0.5% sodium hypochlorite solution for 1–2 min. the samples were washed using sterilized distilled water, dried and placed on the culture media. Then, the cultures incubated at 25 °C under dark condition. Meanwhile, in order to assess the possibility of the presence of the pathogenic agents limited to the vessel, vessels of the cluster and stem end of the infected pistachio fruit were sampled and cultures. After 72 h, the grown fungus was assessed using the microscope. When fungus hyphae were observed in any of the cultivated areas, isolated renewal and purification of the fungus colonies were performed using Single Sporing method (Booth, 1971). Moreover, pure fungus colonies were stored in refrigerator for next assessments.

2.2.2. Identification of isolated fungi

The initial identification of all fungal isolates was made based on morphological characteristics of conidia and conidiophores as well as the colonies. Microscopic slides of each isolates was examined in lactophenol cotton blue mounts by light microscopy (BH2, Olympus, Japan) Pertinent (Barnett and Hunter, 1972).

2.2.3. Pathogenicity of isolated fungi

Detached clusters method was used to study and prove the

pathogenicity of isolated fungi under in vitro conditions. For this purpose, healthy clusters of Kalleh Quchi cultivar trees with no symptoms were collected from the orchards. Ordinary sterilization was done using 0.5% sodium hypochlorite. Then, they were individually placed in sterile Petri dishes and after that, they were inoculated by spraying using each fungal spores in concentration of 107 spore ml^{-1}. The Petri dishes were placed inside plastic containers and in their floor, sterilized distilled water was poured in order to have saturated moisture. They were stored at a temperature of 25 ± 0.5 °C and then checked every day. After, similarity of the observed symptoms was objectively compared with the signs of the problem.

2.3. Investigation of the abiotic agents affecting PEL problem

2.3.1. Samples preparation and measurement of some properties of the soil, water and plant

Soil samples were air dried and passed through 2 mm sieve. The soil texture was measured using hydrometer method (Bouyoucos, 1951; Gee and Bauder, 1986). ECe (Electrical conductivity in saturated extract) was determined in the extract using an electrical conductivity meter (Model Ohm-644, Metrohm AG Herisus, Switzerland) (Rhoades, 1996). Soil pH was measured in saturated paste using digital pH-meter (Model 691, M0065trohm AG Herisus, Switzerland) (Thomas, 1996). CCE (Calcium Carbonate Equivalent) was measured neutralization method (Allison and Moodie, 1965). SOC (Soil Organic Carbon) content was measured by Walkly-Black method with dichromate extraction and titrimetric quantization (Walkley and Black, 1934). Available phosphorus was measured using colorimetric method with Olsen extraction (Olsen and Sommers, 1982). Exchangeable potassium was extracted with ammonium acetate and was measured using flame photometer (410 Cornig model) (Knudsen et al., 1982). Sodium solution was measured using atomic emission (Chapman and Pratt, 1961). Absorbable calcium and magnesium was measured using titration method with EDTA (Tucker and Kurtz, 1961). DTPA–extractable iron, copper, manganese and zinc was measured using atomic absorption method (Lindsay and Norvell, 1978). Boron was measured using Azomethine-H (Gupta et al., 1985; Wolf, 1974) and Spectrophotometer Methods.

The leaves were washed with 0.1 N HCl solution, rinsed twice in distilled water, and oven-dried for 24 h at 65 °C. Pistachio clusters were separately dried after detachment of the fruits having the problem from the fruits without the problem dried leaf and fruit samples were ground to pass through a 1-mm sieve. One g of the dried and ground leaf sample was ashed at 550 °C for 5 h. The ash was dissolved in 2 ml of 1 N HCl and adjusted to a volume of 100 ml for determination of elements concentration (Bremner, 1996). Phosphorus concentration was measured in the extract by colorimetric method (Xie et al., 2006). Potassium and sodium concentrations were measured by flame photometer (410 Cornig model). Calcium and magnesium concentrations were measured by titration method with EDTA. Iron, copper, zinc and manganese concentrations were measured by atomic absorption and boron concentration was measured by Azometin-H and Spectrophotometer Methods (Keren, 1996).

Electrical conductivity, pH, calcium, magnesium, potassium, sodium and boron concentrations in irrigation water were measured using a similar method and also, water carbonate and bicarbonate concentrations were measured using titration with acid. Chlorine concentrations was measured using titration with silver nitrate solution.

2.3.2. Data analysis

Statistical indicators including average, maximum, minimum and variation coefficient in the measured properties of the soil, plant and water samples were analyzed using SPSS for Windows (ver. 24).

2.4. Feature selection

2.4.1. Artificial neural network

Artificial Neural Network (ANN) is a simple imitation of the human brain and includes networks with strong relationships of simple processing units called neurons. Neurons of the ANN are found in groups called layer. All neurons of every layer are in relationship with each other and these relationships are used to determine the burden of the input parameters. This network like the human brain has the power to learn, generalize and make decisions and it can create a mapping from a multivariate space with received data. The most common type of artificial neural network is a set of basic neurons which constitute input vector, one or several hidden layers and one output layer. ANN implementation consists of two phases: training and network testing. In training phase, the purpose is to improve the amounts of the weights of network for numerous samples so that parts of the desired features are introduced into the network as training data and the network during the learning process modifies the amounts of its burdens based on error or the difference between the network output and target variable carried out using learning algorithms. Then, by repeating this practice several times, burdens are timely arranged in such a way that by observing the new information, the network provides the right answer (Fadere, 2009). In the testing phase, the network is tested for a series of known variables which are not used in the network training and thus the possible defects are modified. ANN method has some advantages over the other methods. That is independent of the statistical distribution of data. In this study, SCGBP (Scaled Conjugate Gradient Backpropagation) algorithm was used for optimizing. This algorithm is specifically used for classification studies.

2.4.2. Genetic algorithm

Genetic algorithm (GA) is an inspiration from genetics and Darwin's theory of evolution and it is based on the superiors' survival or natural selection. GA is the most well-known method among the evolutionary algorithms. In the GA, a group of individuals survives in an environment according to their desirability. Those with superior functionality will have greater chance of reproduction. So, after several generations, there will be children with better efficiency (responses). In the GA, each individual (response) of the population is introduced as a chromosome. Chromosomes get more complete after several generations. Chromosomes are evaluated in each generation and according to their value, they will survive and reproduce. In the discussion of the GA, reproduction is carried out with cross over and mutation operators and then superior parents are selected based on a propriety function. In this study, a tournament selection was used to choose parents and in this mechanism, a subset of the attributes of one society is selected and its members compete with each other and eventually only one attribute of each subset is selected for reproduction (Mitchell, 1996).

2.4.3. Selection of the features influencing the formation of PEL problem using hybrid GA -ANN

Data set often has a number of characteristics in the real world and some of them are not important or unrelated to the target variable. It occurs when the range of knowledge in relation to the characteristics effective on the target variable is unknown or incomplete. In this situation, many features are introduced to show this domain that the presence of unsuitable and inactive features may cause confusion in the distribution of really relevant features for that target and may cause damage to the process of data processing and knowledge derived from it. Feature selection and its reduction techniques may be useful in solving this problem (Fayyad et al., 1996; Besalatpour et al., 2014; Chandrashekar and Sahin, 2014). In this study, to determine the factors affecting the formation of PEL problem, a feature selection method based on the classification of the input data was carried out. The target variable was the different levels of the progress of the PEL problem which was divided into 4 classes and used in the process of feature

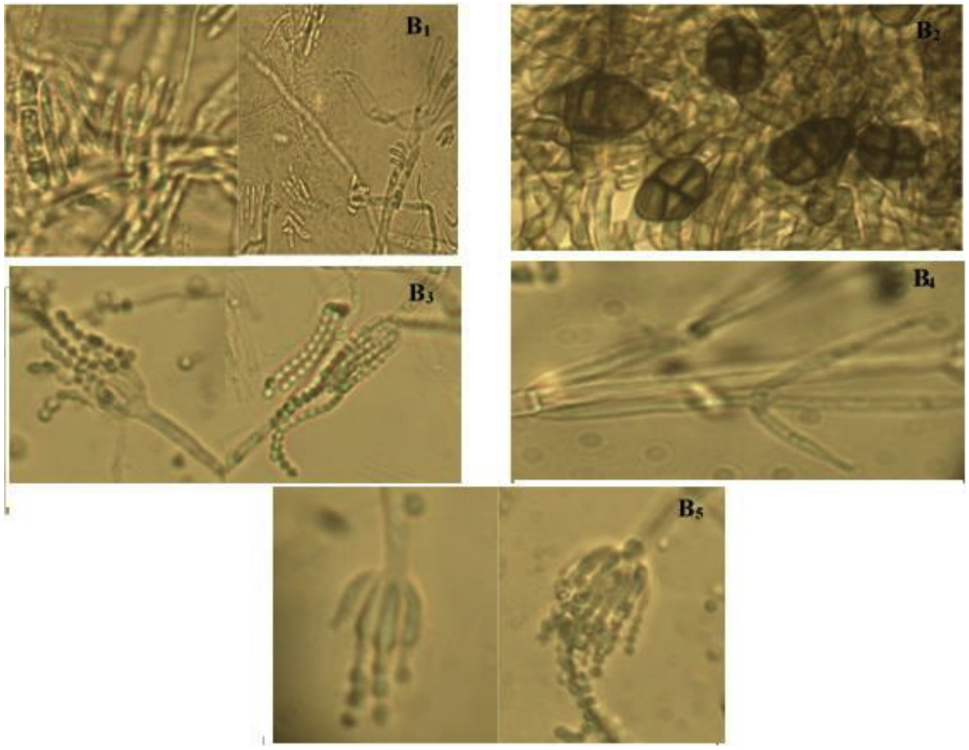

Fig. 2. Information on the extracted species of fungi from pistachios with PEL.
(B_1 = *Fusarium* sp.; B_2 = *Ulocladium* sp.; B_3 = *Penicillium* sp.; B_4 = *Verticillium* sp.; B_5 = Penicillium sp.)

selection of the GA, which was specifically designed for working with ANN (that is, GA-ANN). In this study, the process of the feature selection was performed in few steps as follows:

First step: Different permutations from all the input features are created using GA.

Second step: These permutations are introduced to the ANN and classification of the problem is modeled using the ANN and modeling error is determined.

Third step: The error obtained from the ANN function is again feedbacked to the GA and the GA deals with choosing the next generation among from the most worthy permutations using this error and comparing it with errors of the previous permutations which are introduced to the modeling.

Fourth step: If condition to stop the algorithm is to be provided, in other words, the error rate to become equal with the defined rate for the algorithm or the number of considered repetition for selecting the superior chromosomes is to be provided, then the algorithm stops, otherwise, the algorithm returns to the third step, so, the loop will continue until the stop conditions is provided.

Fifth step: Superior features are selected among all the features and reported with code 1. Other features that are not selected are reported with code 0. MATLAB (Math Works, Inc., Natick, MA, USA) software was used to perform calculations related to the feature selection in this study. Modeling error is equal to the sum of the burden of the training and testing error as follows:

Error = 0.8 test Error + 0.2 train Error

2.4.4. Determination of the most important parameters influencing the creation of PEL problem through sensitivity analysis

The process of sensitivity analysis shows the sensitivity of the model to its input variables. In this research, StatSoft method was used for sensitivity analysis of the model. In this method (StatSoft Inc, 2004), the sensitivity ratio for each input feature is obtained from dividing by the network error in the absence of the desired input feature to the network error in the presence of all the input variables. That is, the model was first created with all the input variables and after achieving the best performance or the least error, the value of the error index (that is, MSE1) was calculated and considered. Then, a certain input variable was deleted and the model was created again with other input features. After obtaining the most proper structure and performance in the model, the error index ratio (MSE2) was determined in this case. The value of the output sensitivity towards the desired input variable was calculated from the ratio of the error index in the second mode (removing one input feature) than the first mode (the presence of all inputs). Sensitivity parameter is expressed as:

$$sensitivity = \frac{MSE_2}{MSE_1}$$

If the ratio of the sensitivity coefficient for one input variable is more than 1 then that variable may have a significant contribution to the performance and output of the model. If the ratio is less than 1, it means that the error in the absence of the input feature is less than the error in its presence. Therefore, this variable not only has a positive effect on the accuracy of the model but also has negative influence. Thus, it must be removed and not used in the model. This should be performed for all the individual input variables and the sensitivity ratio to be calculated (Liu et al., 2001).

3. Results

3.1. Biotic agents influencing PEL problem

A total of 18 symptomatic pistachio samples with different PEL symptoms were selected and cultured onto PDA. Fungal isolates were obtained only from five cultured samples. Based on morphological and cultural characteristics, the isolates were identified as *Ulocladium* sp., *Penicillium* sp., *Verticillium* sp., and *Fusarium* sp. (Fig. 2).Numerous isolates of *Aspergillus* spp., *Penicillium* spp., were always associated with symptomatic pistachio nuts in different areas.

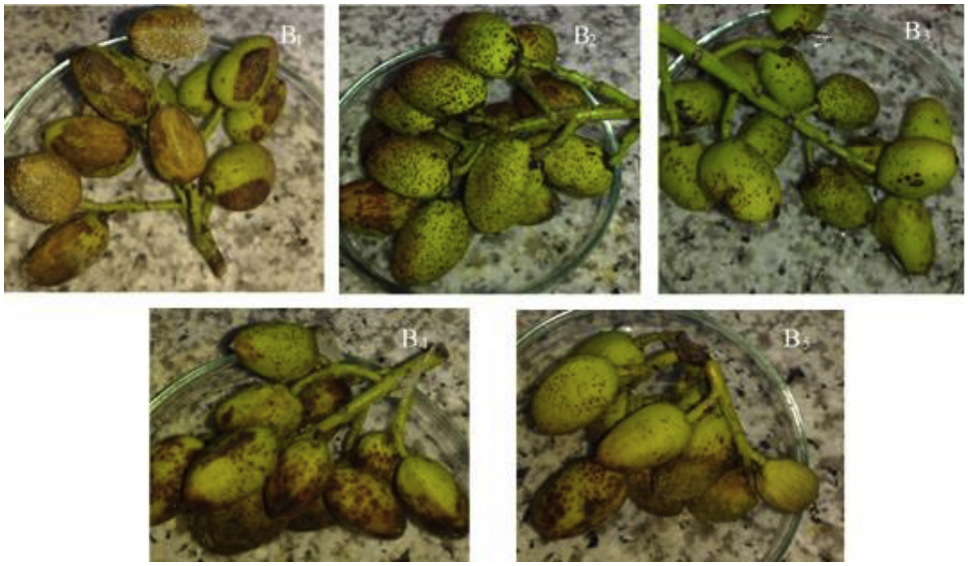

Fig. 3. Evaluation of pathogenicity of isolated fungal strains from different types of PEL contamination on detached clusters in laboratory conditions. (B_1 = *Fusarium* sp.; B_2 = *Ulocladium* sp.; B_3 = *Penicillium* sp.; B_4 = *Verticillium* sp.; B5 = *Penicillium* sp.)

No growth was observed in vessels of the cluster and stem end of the infected pistachio fruits in the cultivated environment.

Results of the pathogenicity tests on detached healthy pistachio clusters showed that all fungal species were produced brown to dark necrotic spots on inoculated pistachio nuts but none of them were similar to the endocarp lesion problem (Fig. 3).

3.2. Abiotic agents influencing PEL problem

3.2.1. Statistical analysis

In this research, chemical properties of water and plant soil were considered but the features like the rootstock was considered the same for all the trees due to very low changes in the area. Statistical details of the probable chemical properties affecting the formation of the endocarp lesion problem including properties of the plant (healthy leaf and fruit), irrigation water and soil, are shown in Table 1–3. Among the studied parameters of the plant, phosphorus of the leaf had the lowest coefficient of variation (CV) (0.16) and sodium had the highest CV (0.91) (Table 1). Among the parameters of water, water acidity had the lowest CV (0.11) and water boron had the highest CV (1.79) (Table 2). Among the soil parameters, soil acidity had the lowest CV (0.03) and soil zinc had the highest CV at the depth of 0–40 (1.4) (Table 3). Estimates of CV for different soil properties are widely available. The CV of the soil properties were less than 15%, 15%–35%, 35%–75% and 75%–150% respectively and were classified into four: low, medium, high and very high (Carter and Gregorich, 2007). Among the features examined in this study, only water and soil acidity were categorized in low range, equivalent calcium carbonate, soil organic materials at a depth of 40–80 cm, sand at a depth of 0–40 cm, sodium, potassium and boron of the fruit and potassium, calcium, magnesium and phosphorus of the leaf were in the medium class and other features were in the high and very high class. Variations of the soil properties in each region were due to changes in soil formation factors and geomorphic processes. The importance of the CV in modelling – based studies is that, as the change in the value of a variable in the study area are higher, modelling has a higher credit.

3.2.2. Selection of the features influencing the formation of PEL problem using Hybrid genetic algorithm and artificial neural network

The process of feature selection was performed using the Hybrid genetic algorithm and artificial neural network (GA-ANN). Out of 68 features investigation in this study, 27 were selected as features influencing in formation of PEL. These features are shown in Table 4. This table shows that most of the selected features are related to the depth of 40–80 cm of soil and the water features in the region have had the least effect on this problem.

3.2.3. Determination of the most important features influencing in PEL problem through sensitivity analysis

Fig. 4 shows the most important features influencing the formation of PEL problem through sensitivity analysis using Statsoft method. As shown in the Fig. 4, from 27 input features in Hybrid (GA-ANN), 19 features had sensitivity coefficient more than 1 which represented the effectiveness of this model in determining the characteristics influencing the formation of PEL problem. Fig. 4 shows the order of the importance of the parameters influencing the formation of PEL problem. Among the chosen features, the ratio of calcium to magnesium and salinity at a depth of 40–80 cm from the soil surface and magnesium of

Table 1
Summary statistics of plant properties used in developing the prediction.

parameter plant Sample			Na (%)	K	Ca	Mg	P	B ($mg\,kg^{-1}$)	Fe	Mn	Zn	Cu
Descriptive statistics	Mean	Leaf	0.07	1.24	2.2	1.15	0.122	175.0	73.00	27.9	13.5	11.72
		fruit	0.03	1.39	0.44	0.40	0.05	38.39	40.22	6.84	10.2	10.0
	Minimum	Leaf	4.20	0.60	1.24	0.3	0.08	47.00	38.60	13.9	5.20	0.80
		fruit	0.01	0.90	0.08	0.05	0.05	21.50	16.60	2.90	3.30	2.30
	Maximum	Leaf	0.56	1.94	3.9	2.37	0.21	497.5	236.8	53.7	90.0	68.0
		fruit	0.09	2.72	1.00	0.98	0.22	80.17	87.90	19.3	86.0	58.0
	CV (%)	Leaf	91	23	23	33	16	49	38	30	79	60
		Fruit	33	24	41	50	72	28	37	42	89	72

Table 2
Summary statistics of water properties used in developing the prediction.

Parameter water Sample		EC (dSm^{-1})	pH	Na ($meq\,L^{-1}$)	K	Ca	Mg	Cl	CO_3^{-2}	HCO_3^{-}	B ($mg\,L^{-1}$)
Descriptive statistics	Mean	4.37	7.9	14.5	28.8	0.07	0.11	18.6	0.91	1.46	5.52
	Minimum	0.58	6.9	4.20	4.90	0.01	0.02	3.44	0.43	0.01	1.14
	Maximum	12.6	8.7	60.5	86.0	0.38	0.70	83.0	2.92	10.4	11.6
	CV (%)	59	11	72	87	86	91	84	41	179	102

the leaf were known as the most important factors. In other words, the formation of PEL problem has the highest sensitivity to the mentioned features.

4. Discussion

4.1. Biotic agents influencing PEL problem

Most experts and local farmers of the study regions agreed that the symptoms of the problem are associated with fungal activity. In our study, fungal species includeing *Ulocladium* sp., *Penicillium* sp., *Verticillium* sp. and *Fusarium* sp. were isolated and identified from pistachio fruits with symptoms of endocarp lesion problem in which had high variability in terms of distribution and species. *Penicillium* with more than 150 known species has the second largest population of saprophyte in nature (Alexopoulos et al., 1996). Also, most species of *Ulocladium* sp. were saprophytes which exist abundantly in soil, air, dust and bodies of dead creatures (Doster and Michailides, 1999). *Fusarium* sp. also exists in all farms containing less moisture and organic matter worldwide. They can also exist in oceans and deserts. Some of the *Fusarium* species are pathogens and some exist as saprophytes around plants root and spend most of their life time in the soil, plants derbies and soil organic matter (Saremi, 2005). Species of Verticillium also exist in soil and in the presence of a suitable host, causing plant wilting and destruction and great damage to the vessels and thus the plant (Pegg and Brady, 2002). Symptoms of this fungus in infected pistachio tree gradually reduce the pistachio trees: leaves lose their color, there is less growth in the branches and trees overall, and branches do not expand so much. The apparent health of the chosen trees in the sampling stage and lack of growth of this fungus in cultivated samples in vessels near the pistachio did not show the influence of this fungus on this problem.

Due to the use of livestock manure and the presence of plant remains in pistachio gardens, the possibility of the presence of different species of these fungi in the soil is very high and tillage activities, soil movement by rain and wind can transfer this species from the soil to the foliage which justifies the existence of these fungi in the fruits of the cultivated pistachio. Also, in the stage of inoculating these fungal species on the detached pistachios clusters, none of the observed symptoms on the clusters were similar to the symptoms of endocarp lesion; thus, the direct impact of such fungus on the endocarp lesion is reject.

Table 4
Selected features (soil, water and plants properties) affect the PEL by (GA-ANN) hybrid algorithm.

Water	Fruit	Leaf	Soil[b]	Soil[a]	
SAR	K	Ca	clay%	Ca	Selected features
	Mg	Mg	sand %	SAR	
	P	P	Ca/Mg	P	
	Zn	Mn	Ec_s	Mn	
	Cu	Zn	P	Cu	
		Cu	Cu	$CaCo_3$	
		Fe	Zn		
			OM%		

[a] Depth of 0–40 soil.
[b] Depth of 40–80 soil.

4.2. Abiotic agents affecting PEL problem

With regards to the results of this study, among all the studied features, soil texture, amount of calcium, magnesium, phosphorus, soil salinity and also, micronutrient elements in the soil and plants were effective in the formation of PEL; thus, the influence of nutrients factors on the formation of PEL can be approved. Hashemi Rad (2005) and Zhang (2005) also emphasized the possibility of the influence of nutrients factors on formation of PEL in their study. Based on the results of

Table 3
Summary statistics of soil properties used in developing the prediction.

Parameter soil Sample		depth cm	clay (%)	silt	sand	CCE	SOM	EC (dSm^{-1})	Na ($meq\,L^{-1}$)	Ca	Mg
Descriptive statistics	Mean	0-40	23.0	22.5	54.5	18.0	1.15	6.40	18.0	20.7	20.8
		40-80	25.3	29.7	45.0	18.4	0.84	8.00	25.2	27.6	23.9
	Minimum	0-40	0.50	6.5	24.0	8.38	0.48	2.10	5.10	1.20	2.80
		40-80	0.50	1.5	9.50	7.57	0.45	2.10	5.80	2.00	0.60
	Maximum	0-40	43.0	59.8	76.0	38.4	3.69	20.9	58.3	72.8	79.2
		40-80	63.0	79.4	84.5	39.5	1.76	18.6	68.9	74.0	97.2
	CV (%)	0-40	43.0	45.0	21.0	31.0	41.0	63.0	56.0	85.0	92.0
		40-80	55.0	54.0	41.0	29.0	29.0	55.0	56.0	73.0	82.0

Parameter soil Sample		depth cm	pH	SAR	B ($mg\,kg^{-1}$)	P	Fe	Mn	Cu	Zn	K
Descriptive statistics	Mean	0-40	7.9	1.20	8.80	20.0	4.60	4.40	0.95	0.50	406
		40-80	7.9	1.40	11.3	13.3	5.00	3.10	0.80	0.30	475
	Minimum	0-40	7.0	0.24	2.71	3.84	0.70	0.83	0.19	0.03	159
		40-80	7.4	0.17	2.61	4.71	1.26	0.59	0.11	0.02	120
	Maximum	0-40	8.9	5.00	17.64	103	19.0	13.2	650	5.70	1511
		40-80	8.5	8.20	30.24	59.0	30.8	8.50	2.10	3.75	1842
	CV (%)	0-40	4.0	58.0	42.0	106	63.0	50.0	53.0	140	45.0
		40-80	3.0	84.0	53.0	92.0	76.0	45.0	50.0	133	55.0

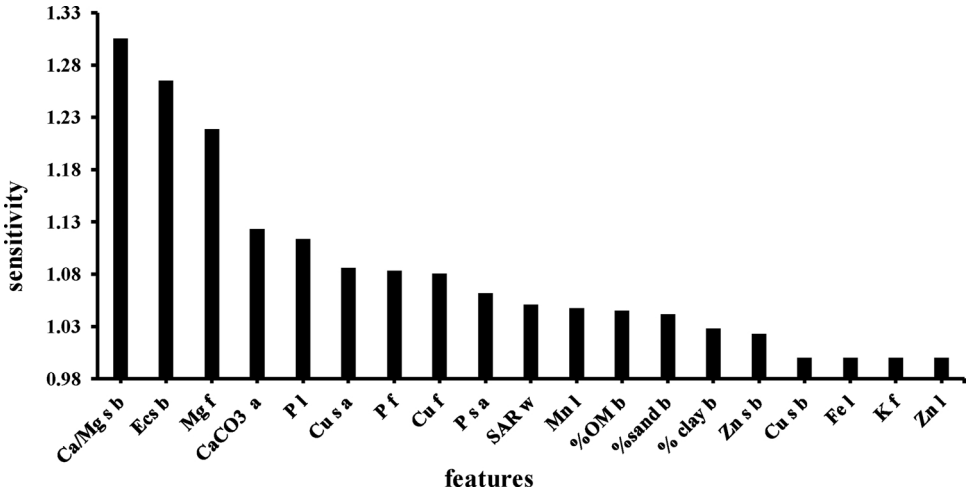

Fig. 4. The ordering of the importance of effective parameters on the formation of PEL problem using sensitivity analysis by Statsoft method (S: soil, L: leaf, F: fruit, W: Water, a: Depth of 0–40 soil, b: Depth of 40–80 soil).

the sensitivity analysis, the ratio of calcium to magnesium and soil salinity around the tree root and also magnesium concentration of the healthy fruit in the trees having PEL problem symptoms were introduced as the most important features affecting PEL problem.

Initiation of the process of endocarp lesion in pistachio coincides with filling of the pistachio kernel. At the time of kernel filling, the fruit needs to be increased with a variety of nutrients and in order to meet this need, the phloem sap is introduced to the fruit. This is why calcium is considered as sedentary element in the phloem sap and its concentration in phloem sap is less than that of other inorganic elements including magnesium. On the other hand, in organs such as fruits where there is less transpiration, introducing more solution substances in the phloem sap reduces or reverses the xylem sap flow and since calcium is the only element whose concentration in the xylem sap is more than that in the phloem sap, less calcium enters into these organs and the plant shows calcium deficiency symptoms. There is a direct relationship between the distribution of calcium and transpiration in the end organs of the plant. Organs such as fruit which have less transpiration show more capacity for calcium deficiency because of low amount that enters the xylem sap. So, any factor that decreases the transpiration thus increases the severity of these symptoms (Marschner, 1995).

Reduction of about 50% of annual rainfall and climate changes during 14 years ago (warming of about 2 °C) and thus warming of the soil caused drought and increased interval in Rafsanjan regions producing pistachio to more than 60 days. Also, reduced relative humidity and soil salinity caused water tension in trees and restricted the movement of calcium from wood vessels to the fruit (Freitas and Mitcham, 2012). Water tensions in plants cause reduced pressure in roots, reduced calcium absorption, increased reactive oxygen in fruit tissue and damage to the plasma membrane. Thus, signs of calcium deficiency occur and develop in the fruit (Ho and White, 2005; Freitas and Mitcham, 2012). This calcium deficiency causes damage to the structure of the pectin in the cell wall and therefore cause disintegration to the cell wall by affecting the enzyme activity of polygalacturonase, especially in the parts of the plant that more calcium is needed like endocarp. Unlike the other macronutrients, lots of calcium exists in the cell wall of the plant textures (Marschner, 1995).

In arid and semi-arid regions such as Iran, probably due to the high amount of magnesium in irrigation water which is caused by an increased level of water use in the pistachio orchards, the balance between calcium and magnesium is changed despite the abundance of calcium in the local matter and soil of these regions, and in fact, an increase in magnesium causes the false calcium deficiency in the fruit.

5. Conclusion

Endocarp lesion problem on pistachio fruits in Rafsanjan Township, contrary to the opinions of many local experts is not caused by the influence of different fungal species. The role of nutrients factors is obvious in forming it. Factors such as the ratio of calcium to magnesium, salinity and the amount of magnesium in fruit are introduced as the most important features affecting this problem. Increased air temperature and reduction in rainfall, water shortage and its salinity and much extraction of underwater resources could be great factors in increasing the magnesium concentration and occurrence of the symptoms of false calcium deficiency in pistachio orchards in Rafsanjan. It is suggested that using the results of this study, the effect of rosettook type on the resistance of pistachio trees to this problem should be investigated.

Acknowledgements

The authors are grateful to Department of Soil Science, College of Agriculture, Vali-e-Asr University of Rafsanjan.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.scienta.2018.11.048.

References

Alexopoulos, C.J., Mims, C.W., Blackwell, A., 1996. Introductory Mycology. John Wiley and Sons, New York.
Allison, L.E., Moodie, C.D., 1965. Carbonate. In: Black, C.A., Evans, D.D., White, L.J., Ensminger, L.E., Clark, F.E. (Eds.), Methods of Soil Analysis. American Society Agronomy, Madison, Wisconsin, pp. 1379–1396.
Ahmadi, K., gholi zade, H.A., et al., 2015. Office of Statistics and Information Technology. Ministry of Jihad-e- Agriculture, Tehran, IRAN (in persian).
Barnett, H.L., Hunter, B.B., 1972. Illustrated General of Imperfect Fungi. Burgess Publications Co., Minneapolis.
Besalatpour, A.A., Ayoubi, S., Hajabbasi, M.A., Yousefian Jazi, A., Gharipour, A., 2014. Feature selection using parallel genetic algorithm for the prediction of geometric mean diameter of soil aggregates by machine learning methods. Arid Land Res. Manag. 28, 383–394.
Bolkan, H.A., Ogawa, J.M., Rice, T.R.E., Bostock, R.M., Crane, J.C., 1984. A leaffooted bug (Hemiptera: Coreidae) and epicarp lesion of pistachio fruit. J. Econ. Entomol. 77, 1163–1165.
Booth, C., 1971. The Genus Fusarium. Common wealth Mycological Institute, Kew, Surrey, England.
Bouyoucos, G.J., 1951. A recalibration of the hydrometer method for making mechanical analyses of soils. Agron. J. 43, 434–438.
Bremner, J.M., et al., 1996. Nitrogen-total. In: Sparks, D.L. (Ed.), Method of Soil Analyses. Part3: Chemical Methods. Soil Science Society of America. American Society

Agronomy, Madison, Wisconsin, pp. 1085–1121.
Carter, M.R., Gregorich, E.G., 2007. Soil Sampling and Methods of Analysis, second edition. CRC Press.
Chandrashekar, G., Sahin, F., 2014. Survey on feature selection methods. Comput. Electr. Eng. 40, 16–28.
Chapman, H.D., Pratt, P.F., 1961. Methods of Analysis for Soils, Plants and Waters. University of California, Riverside, CA.
Chouhan, S., Singh, D., Singh, A., 2016. An improved feature selection and classification using decision tree crop datasets. Int. Comput. Appl. 142, 5–7.
Doster, M.A., Michailides, T.J., 1999. Relationship between shell discoloration of pistachio nuts and incidence of fungal decay and insect infestation. Plant Dis. 83, 259–264.
Fadere, D.A., 2009. Modelling of solar energy potential in Nigeria using an artificial neural network model. Appl. Energy 86, 1410–1422.
Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., 1996. From data mining to knowledge discovery in databases. Artif. Intell. 17, 37–54.
Food and Agriculture Organization of the United Nations (FAO), 2015. Statistical. book. www.fao.org.
Freitas, S.T., Mitcham, E.J., 2012. Factors involved in fruit calcium deficiency disorders. Hortic. Rev. 40, 107–146.
Gee, G.W., Bauder, J.W., 1986. Particle size analysis. In: Klute, A. (Ed.), Methods of Soil Analysis, Part 1. Physical and Mineralogical Methods, Agronomy Monograph No. 9, 2nd edition. American Society of Agronomy. Soil Science Society of America, Madison, WI, pp. 383–411.
Gupta, U.C., James, Y.W., Campbell, C.A., Leyshon, A.J., Nicholaichuk, W., 1985. Boron toxicity and deficiency: a review. Can. J. Soil Sci. 65, 381–409.
Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182.
Hashemi Rad, H., 2005. Identification and Study of the Causative Agents for Pistachio Endocarp Lesion. Report for Pistachio Research Institute.
Hashemi Rad, H., Rajabi, A., 1998. Introduction a new disorder on pistachio fruit in Kerman and the study of Hemipteren relationship to the disorder. 13th Iranian Plant Protection Congress.171. (In Persian with English Summary).
Ho, L.C., White, P.J., 2005. A cellular hypothesis for the induction of blossom-end rot in tomato fruit. Ann. Bot. 95 (4), 571–581.
Keren, R., et al., 1996. Boron. In: Sparks's, D.L. (Ed.), Method of Soil Analyses. Part 3: Chemical Methods. Soil Science Society of America. American Society Agronomy, Madison, Wisconsin, USA, pp. 603–626.
Knudsen, D., Peterson, G.A., Pratt, P.F., 1982. Lithium, sodium and potassium. In: Page, A.L., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis: Part 2: Chemical and Microbiological Properties. Soil Science Society of America, American Society of Agronomy, Madison, Wisconsin, USA, pp. 225–246.
Lindsay, W.L., Norvell, W.A., 1978. Development of a DTPA soil test for zinc, iron, manganese and copper. Soil Sci. Soc. Am. J. 42, 421–428.
Liu, J., Goering, C.E., Tian, L., 2001. A neural network for setting target yields. Am. Soc. Agri. Biol. Eng. 44, 705–713.
Lou, W., Nakai, S., 2001. Application of artificial neural networks for predicting the thermal inactivation of bacteria: a combined effect of temperature, pH and water activity. Food Res. Int. 34 (7), 573–579.
Marschner, H., 1995. Mineral Nutrition of Higher Plants. Academic Press, Londan.
Mitchell, M., 1996. An Introduction to Genetic Algorithms. MIT Press, Cambridge, MA.
Moncada, M.P., Gabriels, D., Cornelis, W.M., 2014. Data-driven analysis of soil quality indicators using limited data. Geoderma 235–236, 271–278.
Nguyen, T.L., Won, Y., 2016. Frequency Features Selection Using Decision Tree for Classification of Sleep Breathing Sound, Lecture Notes in Electrical Engineering; Information Science and Applications (ICISA). Springer Singapore Book Chapter.
Olsen, S.R., Sommers, L.E., 1982. Phosphorus. In: page, A.L., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis: Part 2, Chemical and Microbiological Properties. Soil Science Society of America, American Society of Agronomy, Madison, Wisconsin, USA, pp. 403–429.
Pegg, G.F., Brady, B.L., 2002. Verticillium Wilts. CABI Publishing, Wallingford, UK.
Ramadan, Z., Song, X.H., Hopke, P.K., Johnson, M.J., Scow, K.M., 2001. Variable selection in classification of environmental soil samples for partial least square and neural network models. Anal. Chim. Acta 446 (1–2), 231–242.
Rice, R.E., Uyemoto, J.K., Ogawa, J.M., Pemberton, W.M., 1985. New findings on pistachio problems. Calif. Agric. 39, 1–2.
Rhoades, J.D., 1996. Salinity: electrical conductivity and total dissolved solids. In: page, A.L., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis: Part 2, Chemical and Microbiological Properties. Soil Science Society of America, American Society of Agronomy, Madison, Wisconsin, USA, pp. 417–435.
Saremi, H., 2005. Fusarium, Biology, Ecology and Taxonomy. JD Press Mashhad Branch (In Persian).
Shirani, H., Habibi, M., Besalatpour, A.A., Esfandiarpour, I., 2015. Determining the features influencing physical quality of calcareous soils in a semiarid region of Iran using a hybrid PSO-DT algorithm. Geoderma 259–260, 1–11.
Shirani, H., Hosseinifard, S.J., Hashemipour, H., 2018. Factors affecting cadmium absorbed by pistachio kernel in calcareous soils, southeast of Iran. Sci. Total Environ. 616–617, 881–888.
StatSoft Inc, 2004. Electronic Statistics Textbook (Tulsa, OK). Available at Web site. (Verified 5 March 2015). http://www.statsoft.com/textbook/stathome.html.
Thomas, G.W., 1996. Soil pH and soil acidity. In: page, A.L., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis: Part 2, Chemical and Microbiological Properties. Soil Science Society of America, American Society of Agronomy, Madison, Wisconsin, USA, pp. 475–490.
Tirelli, T., Pessani, D., 2011. Importance of feature selection in decision-tree and artificial-neural-network ecological applications. Alburnus alburnus alborella: a practical example. Ecol. Info. 6, 309–315.
Tucker, B.B., Kurtz, L.T., 1961. Calcium and magnesium determinations by EDTA titrations. Soil Sci. Soc. Am. J. 25 (1), 27–29.
Walkley, A., Black, I.A., 1934. An examination of the DEGTJAREFF method for determining soil organic matter and a proposed modification of the chromic acid titration method. Soil Sci. 37, 29–38.
Yang, X., You, X., 2013. Estimating parameters of van Genuchten model for soil water retention curve by intelligent algorithms. Appl. Math. Inf. Sci. Lett. 7, 1977–1983.
Zhang, J., 2004. Stylar end lesion in pistachio, part 2: elemental analysis. Aust. Nutgrow 18, 28–29.
Zhang, J., 2005. Stylar end lesion in pistachio, part 3: difference or tissue analysis between pistachio rootstocks. Aust. Nutgrow 19, 37–40.
Zhu, M., Chipman, H.A., 2006. Darwinian evolution in parallel universes: a parallel genetic algorithm for variable selection. Technometrics 48, 491–502.
The Islamic republic of Iran Customs Administration (IRICA), 2011. Annual Statistic. http://www.irica.gov.ir/Portal/home/?7068/-Annual-Statistic.
Wolf, B., 1974. Improvements in the azomethine-H method for determination of boron. Commun. Soil Sci. Plant Anal. 5, 39–44.
Xie, Y.H., An, S.Q., Wu, B.F., Wang, W.W., 2006. Density-dependent root morphology and root distribution in the submerged plant Vallisneria nutans. Environ. Exp. Bot. 57, 195–200.