



*Persian J. Acarol.*, 2017, Vol. 6, No. 4, pp. 315328  
<http://dx.doi.org/10.22073/pja.v6i4.30295>  
Journal homepage: <http://www.biotaxa.org/pja>



## Article

### Evaluation of geostatistical method and hybrid Artificial Neural Network with imperialist competitive algorithm for predicting distribution pattern of *Tetranychus urticae* (Acari: Tetranychidae) in cucumber field of Behbahan, Iran

Alireza Shabaninejad<sup>1</sup>, Bahram Tafaghodinia<sup>2\*</sup> and Nooshin Zandi-Sohani<sup>3</sup>

1. Department of Plant Protection, Faculty of Agriculture, Shahrood University, Shahrood, Iran; E-mail: shabanielireza565@gmail.com

2. Department of Plant Production and Sustainable Agriculture, Iranian Research Organization for Science and Technology, Tehran, Iran; E-mail: tafaghodinia@gmail.com

3. Department of Plant Protection, Faculty of Agriculture, Ramin Agriculture and Natural Resources University of Khuzestan, Ahvaz, Iran; E-mail: nzandisohani@yahoo.com

\* Corresponding author

#### ABSTRACT

In this study, the statistical methods and artificial neural network (ANN) were used to estimate the spatial distribution of *Tetranychus urticae* in cucumber field of Behbahan, Iran. Pest density assessments were performed following a 10 × 10 m<sup>2</sup> grid pattern on the field and a total of 100 sampling units on field. In both methods latitude and longitude information were used as input data and output of each methods showed number of pest. In Geostatistics methods ordinary kriging, and ANN with imperialist competitive algorithm were evaluated. Comparison of ANN and geostatistical showed that ANN capability is more than ordinary kriging method so that the ANN predicts distribution of this pest dispersion with 0.98 coefficient of determination and 0.0038 mean squares errors lower than the Geostatistical methods. In general, it can be concluded that the ANN with imperialist competitive algorithm approach with combining latitude and longitude can forecast pest density with sufficient accuracy. Our map showed that patchy pest distribution offers large potential for using site-specific pest control on this field.

**KEY WORDS:** Algorithm; kriging; pest dispersion; statistical methods; variogram.

**PAPER INFO.:** Received: 23 April 2017, Accepted: 24 September 2017, Published: 15 October 2017

## INTRODUCTION

Cucumber (*Cucumis sativus* L.), is a diploid species ( $2n = 14$ ), which belongs to the family of Cucurbitaceae. This plant is native to India and, as the most economical plant in the family of Cucurbitaceae, it is considered one of the valuable vegetable and kitchen garden products in the Middle East (Nario *et al.* 2010). Cultivation area and production rate of this plant in Iran is equal to 80,000 ha and 140,000 tons respectively and it is often planted in field and greenhouse systems (Anonymous 2011). Cucumber has numerous pests and diseases. *Tetranychus urticae* Koch (Acari: Tetranychidae) is one of the major pests that damages cucumbers as well as some summer crops in most parts of the world, especially in warm and temperate areas. Feeding on leaf, *Tetranychus urticae* Koch leads to loss of chlorophyll, reduction of Photosynthetically active areas, and loss of leaves. In case of severe damage (Gorman *et al.* 2001). Plenty of synthetic pesticides are annually used to

control this mite. Having a short life cycle and quick reproduction rate, this pest quickly becomes resistant to synthetic compounds, thus frequent use of acaricide only causes pollution, and increase harmful side effects on non-target organisms and the environment (Isman 2000). Development of integrated pest management programs for this important pest requires knowledge on the changes in the population of the pest in fields over time. However, increased accuracy and precision of interpolation methods and functional mapping of the pests' population are the only factors that can help us achieve these goals. Spatial Interpolation includes estimation of variables such as pest density in the non-sampled points using data obtained from sampled points. In other words, an ideal interpolation method is capable of using the information on the pest population densities obtained from limited number of sampling points (with observation), to properly estimate the pest density in the non-sampled points (Makarjian 2007). The Interpolation methods used in Entomology include kriging and artificial neural network methods.

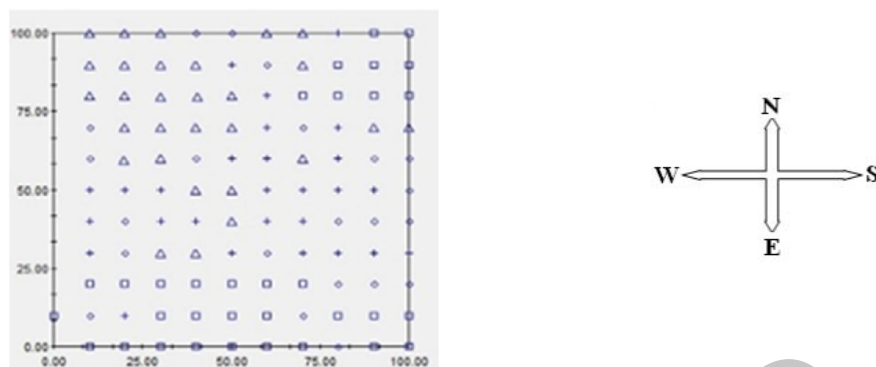
Successful examples of kriging interpolation include prediction of gypsy moth (*Lymantria dispar* L.) population. With the help of this method, we can predict the migration of this pest in forest areas (Liebhold *et al.* 1991). Distribution of *Ostrinia nubilalis* (Hübner) (Lepidoptera: Crambidae) in North America farms was estimated and tracked using geostatistical techniques and their activities was subsequently predicted at the beginning of the season (Wright *et al.* 2002). In recent years that has turned into an international problem, has been conducted with the help of geostatistical methods (Story and Congalton 1994). The geostatistical features that attract *Cydia funebrana* T. (Lepidoptera: Crambidae) are studied by pheromone traps in the plum garden in order to better manage the populations of this insect (Sciarretta *et al.* 2001). Geostatistical characteristics of coffee borer beetle, *Hypothenemus hampei* (Ferrari) (Coleoptera: Curculionidae) and leafminer *Leucoptera coffeella* (Guérin-Méneville) (Lepidoptera: Lyonetiidae) in the coffee plantation, provided the ground for prediction of the insects population and the extent of damage caused by them (De Alves *et al.* 2011). Geostatistics and geographic information system (GIS) were also used in order to implement precision farming system in the management of *Cydia pomonella* L. (Lepidoptera: Tortricidae.) (Ribes-Dasi *et al.* 2005). Geostatistical characteristics and distribution patterns of *Bemisia tabaci* at different growth stages of tobacco were investigated in China (Zhao *et al.* 2011). Similar research has been done in Iran. For example, the distribution and extent of damage incurred by *Batrachedra amydraula* M. (Lepidoptera: Batrachedridae) was specified by geostatistical methods in Khuzestan province (Latifian and Soleymannejadian 2009). Other examples in this field include utilization of geostatistical characteristics to specify the direction and density of *Heliothis virescens* (H.) (Lepidoptera: Nuctoidae) in chickpea fields in Delfan city of Lorestan province (ShafieeNasab *et al.* 2015). In another study (Zhang *et al.* 2008). Learning vector Quantization Neural Network (LVQ) was used to study the spatial distribution of insects in the pasture lands and showed a good performance. According to the literature, no study has ever been conducted to evaluate the performance of geostatistics and ANN in conduction of pest management program in the cucumber farms. Therefore, the present study is an attempt to determine the distribution of the pest in the cucumber field of Behbahan city. It seems that knowledge of the routes through which the insect enter and exit the farm can help us utilize better natural control agents or toxins in order to control pest in the foci of infection.

## MATERIAL AND METHOD

### *Geographical location and development of sampling maps*

A cucumber farm with an area of one hectare around within the geographic area between 47° 41' – 50° 39' east of Greenwich and 29° 58' – 33° 4' north of the equator was chosen for this study. In this farm, the location of each point was considered fixed. To mark the location of sampling points, the north direction of the farm was determined and a point on the southern edge of the farm was

considered the origin of coordinates. The farm was divided into 10 meter segments, and a total of 100 segments were specified on it.



**Figure 1.** Sampling points in the field.

### Sampling method

In all sampling points, a block of  $2 \times 2 \text{ m}^2$  was selected and four plants within each block were randomly selected as the sampling unit. After wards, the number of adult mites behind the plant leaves were counted and recorded.

### Geostatistics

Geostatistical methods are based on the theory of spatial variability. A spatial variable is in fact any environmental or biological feature that is distributed in a two-dimensional or three-dimensional space. Changes in these variables are quite clear and continuous from one point to another. Damage and density, are examples of spatial variables (Katherine 2001). The main difference between this method and classical statistics is that in classical statistics. Samples taken from a population are independent of each other and no sample provides any information about the next sample. But geostatistical methods, deal with spatial correlation between the values of a variable in an area. Spatial dependence between samples can be studied as a mathematical model known as spatial structure (Hassani Pak 2005). In general, geostatistics has two main sections known as variogram and kriging (Gressie 1993).

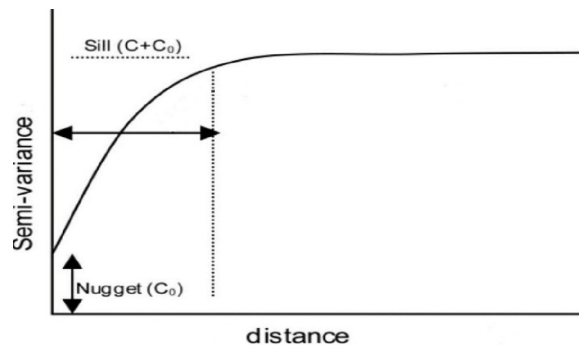
### Variogram

Variogram is in fact the diagram of Variance between samples, which shows the structure of the spatial dependence between samples. For calculation of the experimental variogram, first the square of dissimilarity (in terms of a specific feature) between two points with a distance of  $h$  is calculated, then the average of the squared dissimilarities is calculated and finally an experimental model is fitted on it. The mean squared dissimilarity of the quantity in question is equal to distance  $h$  in all points and is referred to as variogram ( $h$ ) and calculated from equation (1) (Habashi *et al.* 2007).

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(X_i) - z(X_i + h)]^2 \quad (1)$$

( $H$ ) is the amount of variogram for  $N$  pairs of samples at a distance of  $h$  from each other,  $z(x_i)$  and  $z(x + h)$  denote the amount of the  $x$  variable at the  $i$  and  $i + h$  points. With any increase in the distance  $h$ , the variogram value is also gradually increased until it reaches a fixed distance known as threshold. It should be noted that only one group of variogram models have a threshold, for example (exponential, spherical, Gaussian models) and another group such as linear models don't have any threshold. This suggests that with any increase in the distance  $h$ , the variogram value increases as well but it doesn't reach any fixed level (Journel and Huijbregts 1978).

The distance at which the variogram reaches a fixed value is referred to as scope of impact. In other words, the scope of impact is a distance beyond which the samples have no impact on one another and are assumed separated (Goovaets 1997).



**Figure 2.** Generalized semivariogram showing the range of spatial dependence, nugget effect ( $C_0$ ) variability associated with spatial dependence ( $C$ ), and sill ( $C + C_0$ ).

The parameters of the model selected for the variogram can be used to estimate the optimal value of a property based on the collected data. Such an optimized estimation can be conducted by statistical estimators known as kriging (Krige and Magri 1982).

#### *Kriging*

Kriging is basically a generalized name for all statistical methods of spatial variable estimation in which the sum of a weight factor multiplied by the observation points is estimated. This means that the nearer variable point is to the origin of coordinates, the heavier its weight will be and vice versa (Hassani Pak 2005).

#### *Artificial Neural Network*

ANN has a structure similar to the biology of the human brain nervous system (Torrecilla *et al.* 2004). Today neural networks are used in many fields, including classification, pattern recognition, prediction and process modeling in different sciences. The advantage of neural network is direct learning from the data, without any need to estimate their statistical outline (Vakil-Baghmisheh and Pavešić 2003). Regardless of any initial assumptions and prior knowledge of the relationships between parameters, the neural network is capable of finding a relationship between a set of inputs and outputs to predict each output corresponding to a given input (Torrecilla *et al.* 2004; Kaul *et al.* 2005). Error tolerance is another feature of neural network (Azadeh *et al.* 2006). These advantages clearly show why neural network is used to predict pest density.

#### *Data preprocessing*

First, the data is divided into two categories namely training set with 70 members (70 percent of total data) and experiment set with 30 members (30 percent of total data). However, in this case division doesn't provide the desired results, this step can be repeated (Zhang *et al.* 1998). Before using the raw data in the network training, the data should be normalized in an appropriate range because the learning algorithm can't have a good performance along with the raw data. On the other hand, considering the output variation ranges in the sigmoid transfer function used in the intermediate layer, this process seems to be necessary. Otherwise, the network will not converge during the training phase and the desired results will not be achieved (Yuxin *et al.* 2006). When the sigmoid transfer function is used, the best data conversion range stretches from 0.1 to 0.9 (Vakil-Baghmisheh and Pavešić 2003). Linear normalization method in equation 2 is used to convert the data:

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \times (r_{max} - r_{min}) + r_{min} \tag{2}$$

Where, X denotes the input raw data, X<sub>n</sub> shows the normalized data, X<sub>max</sub> and X<sub>min</sub> show maximum and minimum values of input data respectively, r<sub>max</sub> and r<sub>min</sub> show upper and lower bounds of variation range in the converted data.

*Multilayer Perceptron Neural Network (MLP)*

MLP neural networks are composed of one or more intermediate layers. The input signals are normalized by normalizing coefficients and retake their actual values after the calculations. The initial weight values are considered to be random (Kim 2006). In this network, first each neuron in the hidden layer calculates the products of input data multiplied by connection weights and using an activation function, send the product to the neuron in the subsequent layer. The calculated output values are compared to their actual values and the error rate is calculated. In case the error rate deviates from the pre-determined standard error rate, and go back to the previous step and calculate new outputs by changing the connection coefficients and repeating the previous steps. In this network, learning takes place based on back propagation algorithm (Choudhury and Bartarya 2003).

*MLP Neural Network architecture*

The model parameters and optimal designs are selected through the 8-step process introduced by Castera and Boyd (1996) where 5 neurons in the hidden layer show the best performance. The active functions used here are the sigmoid function in the hidden layer and the linear function in the output layer. The number of iterations in the entire process was considered 1000.

*Adjusting the MLP neural network weights using imperialist competition*

This algorithm is a new algorithm in the field of evolutionary computation that is based on social-political development of human. The first stage of this algorithm begins with a number of initial populations known as countries, and then the objective function is obtained for each of the countries. By comparing the objective function values of all countries, some of the best are selected and called imperialist countries and other countries are called colonies. It should be noted that, on equal terms, where multiple values of the same objective function are available for selection of colonial countries, imperialist countries are randomly selected from them. Now, colonial countries are assigned to the imperial states until Empires are formed (every empire consists of country and several colonial countries). It should be noted that the greater the power of an imperialist country (better objective function) a greater number of colonial countries will be assigned to it. An overview of the algorithm is presented in Figure 3 (Atashpaz-Gargari 2009).

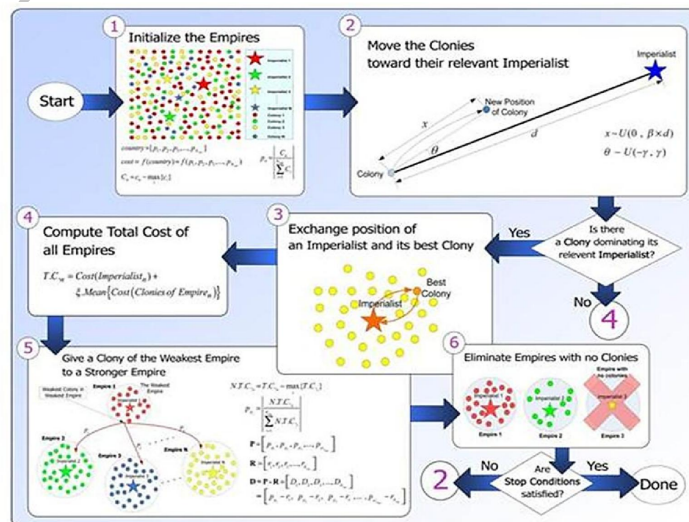
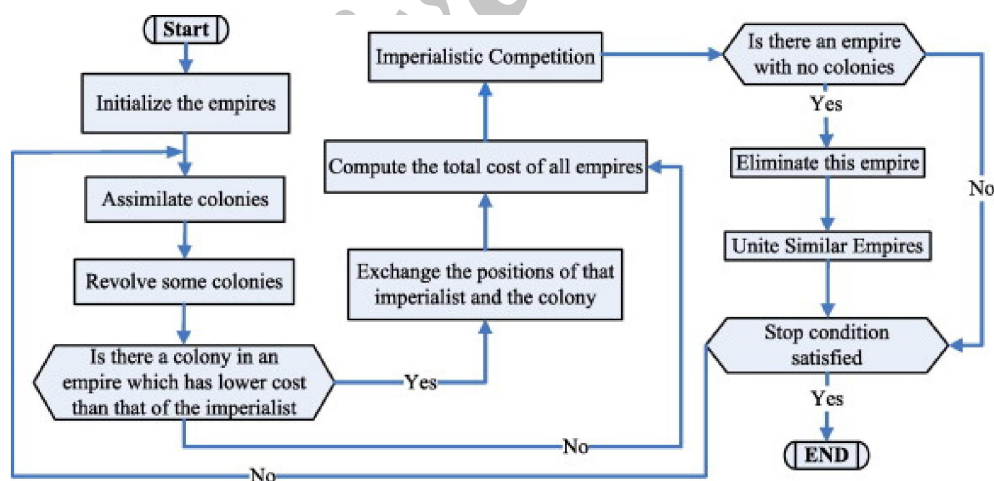


Figure 3. Scheme of Imperialist Competitive Algorithms (Atashpaz-Gargari 2009).

As Figure 3 shows, just like other evolutionary optimization methods, the developed algorithm begins with a number of initial populations (Part 1 of Fig. 3).

In this algorithm, each element of the population is called country. Countries are divided into two groups namely colony and imperial countries. Depending on their power, the imperial countries dominate over a number of colon and bring them under their control. Here the stronger imperialist country shown with a larger star, has the largest number of countries shown by circles with similar colors and the weakest imperialist country shown with a small star, has the smallest number of countries under its dominion. The next stage is known as the imperialistic competition and assimilation policy which constitutes the core of the algorithm (Part 2 of Fig. 3). In the next stage, the costs are calculated and in case the colonial costs are lower than the costs incurred by the imperialist country, the colony and the imperialist state exchange positions by applying the revolution function (Part 3 of Fig. 3). This invokes the colonial countries to assimilate to the imperialist country. This would result in colonized country moving towards the imperialist countries. In other words, each of the colonies follow a specific procedure to come closer to the imperialist countries. At this stage, the normal procedure may lead to revolutions in some colonies and finally seizure of empire power by them (revolution operator). After assimilation and revolution functions are implemented, the objective function of the total cost of Empire is calculated in the third stage (Part 4 of Fig. 3). In any of the empires, in case the best objective function of each colony is of higher value compared to the objective function of imperialist country, the two countries would exchange positions. In addition, after this procedure is followed for all empires, the best solution and value of all imperialist countries will be saved as the best current solution and value in the algorithm iteration. In the next phase, colonies leave the weakened empires and join more powerful ones (Part 5 of Fig. 3). Finally empires become weakened and this algorithm cycle continues until one empire that is actually the optimal solution remains (Part 6 of Fig. 3). Flowchart of this algorithm is shown in Figure 4 (Moradi and Zandieh 2013).



**Figure 4.** Flowchart of Imperialist Competitive Algorithm (Atashpaz-Gargari 2009).

#### *Mathematical trend of imperialist competitive algorithm*

Taking into account the function  $f(x)$ , in Optimization problems attempts are made to find arguments  $x$  in such a way that its corresponding costs, i.e. the cost of a country, is measured by evaluating  $f$  function for variables  $[P_1, P_2, P_3, \dots, P_{nvar}]$ . Therefore:

$$\text{Cost}_i = f(\text{country } i) = f(P_1, P_2, P_3, \dots, P_{nvar})$$

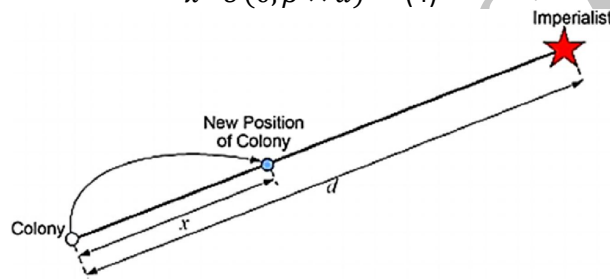
In this algorithm, first  $N_{country}$  Initial countries are created and  $N_{imp}$  best members of the population (the countries with the lowest cost function) as selected as imperialist countries. The remaining  $N_{col}$  countries form colonies, each of which belong to an empire.

By applying the assimilation policy, the imperialistic countries attract the colonial countries in line with different optimization axes. Depending on their power, the imperialist countries attract these colonies towards themselves via (3) connection. The total power of an empire is determined by calculating the power of its both constituents, i.e. the power of the entire imperialistic countries plus a percentage of their colonies' average power (Atashpaz-Gargari *et al.* 2008).

$$T.C_n = Cost(imperialist_n) + \alpha \text{mean}\{Cost(colonies\ of\ empire_n)\} \tag{3}$$

The Colonial country moves a distance of  $x$  units along the line towards the imperialist country and is dragged to a new position. In Figure 5, the distance between the imperialist country and the colonial country is shown by  $d$  and  $x$  is a random number with uniform distribution (or any other appropriate distribution). This means that for  $x$ , we have:

$$x \sim U(0, \beta \times d) \tag{4}$$

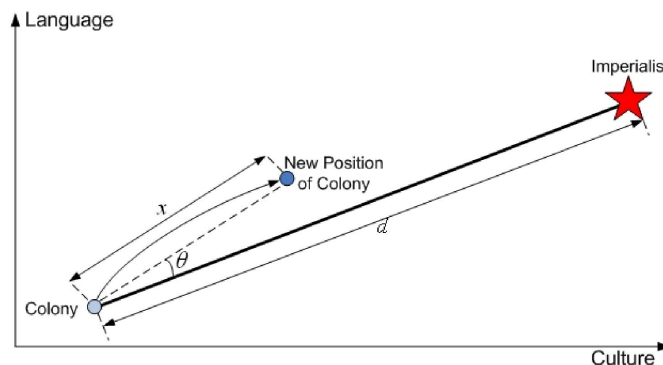


**Figure 5.** Moving colonies to imperialist in culture and language axes (Atashpaz-Gargari *et al.* 2008).

Where  $\beta$  is a number greater than 1 and close to 2. An appropriate choice can be  $\beta = 2$ . The movement angle is considered to be in the following uniform distribution (Atashpaz-Gargari 2009):

$$\phi \sim U(-\gamma, \gamma) \tag{5}$$

In this algorithm, with a potential deviation, the colony moves closer to the imperialist country. This deviation angle is denoted by  $\phi$  the value of which is selected randomly and through uniform distribution (Fig. 6).



**Figure 6.** Motion of colonies toward their relevant imperialist (Atashpaz-Gargari 2009).

In this equation,  $\gamma$  is an optional parameter and any increase in the value of this parameter will increase the search space around the imperialists and reduction in its value causes the colonies to move as close as possible to the vector that connects colonial countries to the imperialist countries. Considering the radian unit for  $\theta$ , a number close to  $\pi/4$  is a perfect choice in most implementations (Atashpaz-Gargari 2009). During the movement of colonies towards the imperialist country, some of these colonies may reach a position better than that of the imperialist country. In this case, the imperialist country and the colony may exchange positions. For modeling the competition, first the possibility of colonies being dominated by any empire (taking into account the total cost of the empire) is calculated as follows (Atashpaz-Gargari 2009):

$$N.T.C_n = \max_i \{T.C_i\} - T.C_n \quad (6)$$

In this equation  $T.C_n$  is the total cost of the  $n$ th empire and  $N.T.C_n$  is the total normalization cost of the empire. The possibility of each colony being dominated by any empire is calculated as follows:

$$P_{pn} = \left| \frac{N.T.C_n}{\sum_{i=1}^{N_{imp}} N.T.C_n} \right| \quad (7)$$

In this study, the ICA algorithm is used as back-propagation neural network training algorithm to estimate the distribution of the pest. One of the shortcomings of artificial neural network is associated with finding a suitable network structure for estimation, which was usually obtained through trial and error in previous studies. In this study, ICA is used as a network training algorithm. This algorithm is able to determine the best network state for estimation purposes, and this process is conducted through determination of optimal network weight and bias modes (Enayatifar *et al.* 2013). In this study, the MSE function is used as the network cost function. The main objective of using this algorithm was to minimize the cost function.

#### Stop condition

Number of all iterations of imperialist competitive algorithm was considered equal to 600 and if an improvement in the amount of fitness doesn't make after 150 iterations, algorithm stops.

#### Statistical Analysis

Normality test of the data associated with sampling was evaluated using SPSS software version 19 and Kolmogorov-Smirnov test, and due to lack of normality in the data, the data were normalized by Cox box. Data analysis for the geostatistics method was conducted by GS + software version 5.1.1. Computer code of the neural network with imperialist competitive algorithm was prepared by Matlab 8.1 software.

## RESULTS AND DISCUSSION

According to Table 1, the results of data fitting in geostatistics method shows that 3 out of the total of four sampling stages match the spherical model and one matches the exponential model. Based on this result, it can be safely concluded that the distribution of two-spotted spider mite is cumulative. According to the results, the coefficient of determination in all of cases more than 0.50 and the degree of spatial dependence in all cases, exceeds half the variograms threshold and ranged from 0.510 to 0.556 that is actually an inappropriate degree of spatial dependence. This shows the fact that only 0.50 percent of the data variances have spatial structure and these results can't be trusted.



**Table 1.** Geostatistical characteristics of the infected plants to *T. urticae* in the cucumber farm.

Sampling date	Model	Coefficient of Determination	MSE	Degree of spatial dependence	Range of spatial dependence	Nugget
01.09.2016	Spherical	0.540	0.245	0.510	125.45	0.0665
08.09.2016	Exponential	0.576	0.229	0.556	185.90	0.0453
19.09.2016	Spherical	0.697	0.189	0.514	150.4	0.0457
22.09.2016	Spherical	0.751	0.163	0.532	205.11	0.0367

In order to select optimum imperialist competitive algorithm parameters like Population, Initial Empire and Revolutionary Rate, algorithms were conducted several times. According to table 2, results of trial and error showed that the best imperialist competitive algorithm output becomes possible with a population of 200, Initial Empire 80 and Revolutionary Rate 0.2 for mentioned parameters, mean square error becomes the minimum for recommended neural network output component.

**Table 2.** Mean squared error of proposed neural network with different parameters of imperialist competitive algorithms

MSE	Revolutionary Rate	Initial Empire	Population
0.090	0.5	50	100
0.088	0.4	60	100
0.085	0.3	70	100
0.058	0.2	80	100
0.031	0.5	50	150
0.023	0.4	60	150
0.023	0.3	70	150
0.015	0.2	80	150
0.005	0.5	50	200
0.004	0.4	60	200
0.004	0.3	70	200
0.004	0.2	80	200

**Table 3.** Statistical comparisons between the observed and estimated to *T. urticae* density by MLP neural networks.

Sampling date	Utilization phase	Comparisons of means	Comparisons of variance	Comparisons of distribution
01.09.2016	Training	0.624	0.886	0.82
	Test	0.923	0.993	0.92
08.09.2016	Training	0.868	.0783	0.74
	Test	0.996	0.887	0.56
19.09.2016	Training	0.701	0.865	1.00
	Test	0.967	0.921	1.00
22.09.2016	Training	0.756	0.921	1.00
	Test	0.932	0.933	1.00

The results of this evaluation are presented in Table 3. In order to ensure learning in the ANN that was trained to predict the distribution pattern of two-spotted spider mite, the actual data and the data predicted by the network were compared statistically. Here the null hypothesis implies equality of mean, variance, and statistical distribution. Each hypothesis was tested at 95% probability using the p parameter. The t test, F test and Kolmogorov–Smirnov test was used to compare mean, variance and statistical distribution of the data respectively. The calculated P values for each case are shown in Table 2. The results show that there is no significant difference between mean, variance of ANN and the actual and predicted statistical distribution values ( $p > 0.3$ ). The p value more than 0.74 for

the statistical distribution between the predicted and actual values of the two-spotted spider mite density in the field, indicates the high precision and potential of ANN.

Determination coefficients and linear regression relationship between actual values of each sampling vs. the values were predicted by the ANN are shown in Table 4. The results showed greater generalize ability of ANN in estimating the density of two-spotted spider mite. Therefore, the output of this approach was used to develop the density maps. Coefficient of determination  $R^2$  and Mean square error (that were equal to 0.751 and 0.163 in the kriging method and 0.9801 and 0.0044 in the neural network method, respectively) were used to evaluate the effectiveness of the two methods. The results suggest greater precision of ANN in estimation of the pest density. In a study to determine the distribution patterns of insects in a pasture through the neural network it was found that the MLP, LVQ and linear neural networks can properly identify the distribution patterns of insects. However, MLP proved to have the strongest algorithm for pattern recognition (Zhang *et al.* 2008). The results obtained by Young *et al.* (2000) show high efficiency of neural network in prediction of the Diptera (Cecidomyiidae) population dynamics in the coniferous forests in America. The results of these two studies are consistent with the results of the present study.

**Table 4.** Linear regression relationship and coefficient of determination between dv (actual value) and pv (predicted value by model).

Sampling date	Network Utilization phase	Linear regression relationship	$R^2$	MSE
01.09.2016	Training Phase	pv= 0.8104 av +0.0324	0.8545	0.0189
08.09.2016	Training Phase	pv= 0.9200 av +0.0122	0.9299	0.0030
19.09.2016	Training Phase	pv= 0.9368 av +0.0109	0.9413	0.0038
22.09.2016	Training Phase	pv= 0.9646 av +0.019	0.9801	0.0044

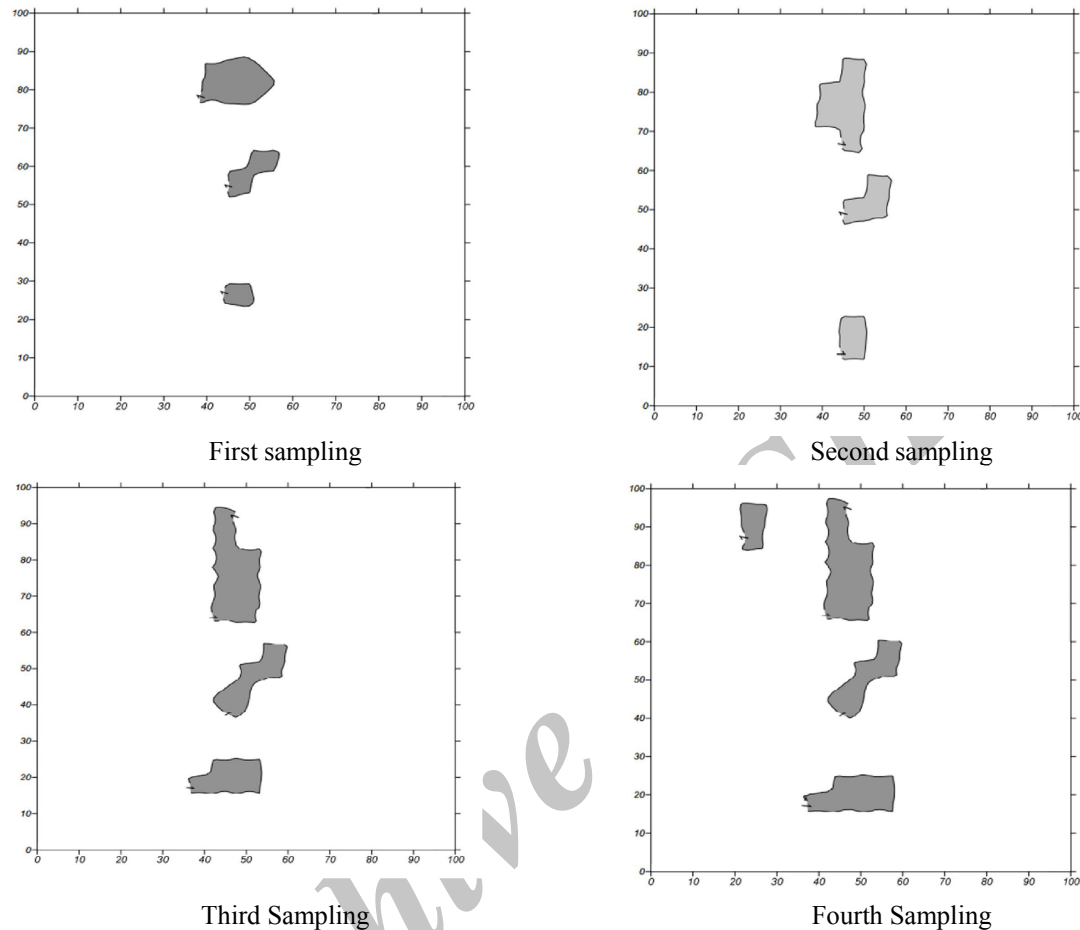
#### *Two-spotted spider mite spatial distribution maps*

According to the guideline of Figure 7, the black areas are the most infected areas in the farm. According to different samplings, first there was only a single center of infection in the north of the farm, but it gradually spread towards the midpoint of the farm and finally towards its seat. Therefore, the infected areas can be locally sprayed in order to avoid overall farm spraying and prevent further contamination of the environment. On the other hand, natural control agents such as predatory mite (*Phytoseiulus persimilis* Athias-Henriot) can be concentrated in the infected areas to improve the efficiency of the above-mentioned measures.

## CONCLUSION

Biology-inspired algorithms have important rule in computational sciences, which are essential to many branches of knowledge. Many computational methods are inspired from natural and biological activities. Some of these biologically inspired computations include genetic algorithms, neural networks, cellular automata, and other algorithms. These nature-inspired algorithms are popular in recent years to face real world problems and solve complex optimization functions whose actual solution doesn't exist. In this research, an ANN combined with imperialist competitive algorithm could predict and draw dispersion of two-spotted spider mite with high accuracy. Comparison of ANN and geostatistical showed that ANN capability is more than kriging method so that the ANN predict distribution of this pest dispersion with 0.98 coefficient of determination and 0.0038 mean squares errors lower than the geostatistical methods. In general, it can be concluded that the ANN with imperialist Competitive Algorithm approach with combining latitude and longitude can forecast pest density with sufficient accuracy. Our map showed that patchy pest distribution offers large potential for using site-specific pest control on this field. The resulted map reveals cumulative dispersion of this pest. Therefore, only by spraying points with high density, we can achieve good

management of farm and decrease of using poisons and this will reduce environmental damage and it will save money because it prevents the pulling out of money from the country.



**Figure 7.** Distribution of *T. urticae* in different stages of sampling.

## REFERENCES


- Anonymous (2011) Agricultural statistics. Department of Planning and Economy, The office of Statistics and Information Technology, Tehran, 360 pp.
- Atashpaz-Gargari, E. (2009) Imperialist Competitive Algorithm development and its applications. M. Sc. Thesis, University of Tehran, 110 pp. (In Persian with English summary).
- Atashpaz-Gargari, E., Hashemzadeh, F., Rajabioun, R. & Lucas, C. (2008) Colonial competitive algorithm: A novel approach for PID controller design in MIMO distillation column process. *International Journal of Intelligent Computing and Cybernetics*, 1(3): 337–355.
- Azadeh, A., Ghaderi, S.F. & Sohrabkhani, S. (2006) Forecasting electrical consumption by integration of Neural Network, time series and ANOVA. *Applied Mathematics and Computation* 186(12): 1753–1761.
- Choudhury, S.K. & Bartarya, G. (2003) Role of temperature and surface finish in predicting tool wear using neural network and design of experiments. *International Journal of Machine Tools & Manufacture*, 10(2): 747–753.
- De Alves, M.C., Da Silva, F.M., Moraes, J.C., Pozza, E.A., De Oliveira, M.S., Souza, J.C.S & Alves, L.S. (2011) Geostatistical analysis of the spatial variation of the berry borer and leaf miner in a coffee agroecosystem. *Precision Agriculture*, 12(2): 18–31.

- Enayatifar, R., Sadaei, H.J., Abdullah, A.H. and Gangi, A. (2013) Imperialist competitive algorithm combined with refined high-order weighted fuzzy time series (RHWFTS-ICA) for short term load forecasting. *Energy Conversion and Management*, 76(3): 1104–1116.
- Goovaets, P. (1997) *Geostatistics for Natural Resources Evaluation*. Oxford University Press, Oxford, 512 pp.
- Gorman, K., Hewitt, F., Denholm, L. & Devine, G.J. (2001) New developments in insecticide resistance in the glasshouse whitefly (*Trialeurodes vaporariorum*) and the two-spotted spider mite (*Tetranychus urticae*) in the UK. *Pest Management Science*, 58(6) 123–130.
- Gressie, N. (1993) *Statistics for spatial data*. John Wiley, New York, 550 pp.
- Hassani Pak, A. (2005) *Geostatistics*. University of Tehran Press, Tehran, 538 pp.
- Habashi, H., Hosseini, M., Mohammadi, J. & Rahmani, R. (2007) Geostatistic applied in forest soil studing process. *Journal of Agricultural Science and natural Resources*, 14(3): 1–10.
- Isman, M. (2000) Plant essential oils for pest and disease management. *Crop Protection*, 18(10): 603–608.
- Journel, A.G. & Huijbregts, C.J. (1978). *Mining Geostatistics*. Academic Press, 599 pp.
- Kim, K. (2006) Artificial Neural Network with evolutionary instance selection for financial forecasting. *Expert systems with application*, 30(3): 519–526.
- Katherine, A.R. (2001) *Geostatistic using SAS software*. Owen Analytic Inc., Deep. River, CT., 6 pp.
- Castera, I. & Boyd, M. (1996) Designing a Artificial Neural Network for forecasting financial and economic time series. *Neurocomputing*, 12(5): 13–19.
- Kaul M, Hill RL & Walthall C. (2005) Artificial neural networks for corn and soybean yield prediction. *Agriculture System*, 85(2): 1–18.
- Krige, D.G. & Magri, E.J. (1982) Studies of the effects of outliers and data transformation on variogram estimates for a base metal and a gold ore body. *Mathematical Geology*, 14(2): 557–567.
- Liebhold, A.M., Zhang, X., Hohn, M.E., Elkinton, J.S., Ticehurst, M., Benzon, C.L. & Campbell, R.W. (1991) Geostatistical analysis of Gypsy moth (Lepidoptera: Lymantridae) egg mass population. *Environmental Entomology*, 20(5): 1407–1417.
- Latifian, M. & Soleymannejadian, E. (2009) Study of the Lesser moth *Batrachedra amydraula* (Lep.: Batrachedridae) distribution based on geostatistical models in Khuzestan province. *Journal of Entomological Research*, 1(1): 43–55.
- Makarjian, H., Rashed Mohassel, M.H., Bannayan, M. & Nassiri, M. (2007) Soil seed bank and seedling populations of *Hordeum murinum* and *Cardaria draba* in saffron fields. *Agriculture Ecosystems and Environment*, 120(3): 307–312.
- Moradi, H. & Zandieh, M. (2013) An imperialist competitive algorithm for a mixed-model assembly line sequencing problem. *Journal of Manufacturing Systems*, 32(6):46–56.
- Nario, L.S., Oliver-Verel, J. & Stashenko, E.E. (2010) Repellent activity of essential oils: A review. *Bioresource Technology*, 101(5): 372–378.
- Ribes-Dasi, M., Almacellas, J., Sió, J., Torà, R., Planas, S. & Avilla, J. (2005) The use of Geostatistics and GIS to optimise pest control practices in precision farming systems. *Information and Technology for Sustainable Fruit and Vegetable Production*, 10(2):583–590.
- Story, M. & Congalton, R.G. (1994) Accuracy assessment: A user's perspective: L.K. Fenster maleer. Remote sensing thematic assessment. *American Society for Photogrammetry and Remote Sensing*, 12(4): 257–259.
- ShafieeNasab, B., Shakarami, J., Mohiseni, A. & Jafari, S.H. (2015) Geostatistical characteristics of the spatial distribution of the infestation pods by the pod borer, *Heliothis virescens* Huf. (Lep.: Noctuidae) in rain-fed chickpea (*Cicer arietinum* L.) fields in Delfan (Lorestan province). *Plant Pests Research*, 5(2): 49–59.
- Sciarretta, A., Trematerra, P. & Baumgärtner, J. (2001) Geostatistical analysis of *Cydia funebrana* (Lepidoptera: Tortricidae) pheromone trap catches at two spatial scales. *American Entomologist*,

47(3): 174–184.

- Torrecilla, J.S., Otero, L. & Sanz, P.D. (2004) A neural network approach for thermal/pressure food processing. *Food Engineer*, 62(9): 89–95.
- Vakil-Baghmisheh, M.T. & Pavešić, N. (2003) Premature clustering phenomenon and new training algorithms for LVQ. *Pattern recognition*, 36(5): 1901–1921.
- Wright, R.J., Devries, T.A., Young, L.J., Jarvi, K.J. & Seymout, R.C. (2002) Geostatistical analysis of small-scale distribution of European corn borer (Lepidoptera: Crambidae) larvae and damage in whorl stage corn. *Environmental Entomology*, 31(3): 160–167.
- Young-S.P., Ja-Myung, K., Buom-Young, L., Yeong-Jin, L. & YooShin, K. (2000) Use of an artificial neural network to predict population dynamics of the forest–pest pine needle gall midge (Diptera: Cecidomyiida). *Environmental Entomology*, 29(6): 1208–1215.
- Yuxin, M., Mulla, D.J. & Pierre, C.R. (2006) Identifying important factors influencing corn yield and grain quality variability using artificial neural networks. *Precision Agriculture*, 7(2) 117–135.
- Zhang, Y.F. & Fuh, J.Y.H. (1998) A neural network approach for early cost estimation of packaging products. *Computers & Industrial Engineering*, 34(4): 433–50.
- Zhang, W.J., Zhong, X.Q. & Liu, G.H. (2008) Recognizing spatial distribution patterns of grassland insects: neural network approaches. *Stochastic Environmental Research and Risk Assessment*, 22(8): 207–216.
- Zhao, J., Zheng, F.Q., Wang, Y.J., Ye, B.H., Zhao, X.I., Mu, H.Y., Hao, L.U. (2011) Geostatistical analysis of spatial patterns of *Bemisia tabaci* (Homoptera: Aleyrodidae) adults in Tobacco field. *6th IEEE Conference on Industrial Electronics and Applications*, Beijing, China Pp. 2394 – 2398.

#### **COPYRIGHT**

 Shabaninejad *et al.* Persian Journal of Acarology is under free license. This open-access article is distributed under the terms of the Creative Commons-BY-NC-ND which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ارزیابی روش‌های زمین‌آمار و شبکه عصبی مصنوعی بهینه شده با الگوریتم رقابت استعماری در تعیین پراکنش مکانی کنه تارتن دولکه‌ای (*Tetranychus urticae*) (Acari: Tetranychidae) در مزرعه خیار شهرستان بهبهان

علیرضا شعبانی نژاد<sup>۱</sup>، بهرام تفقدی نیا<sup>۲\*</sup> و نوشین زندی سوهانی<sup>۳</sup>

۱. گروه گیاهپزشکی، دانشکده کشاورزی، دانشگاه شاهرود، شاهرود، ایران؛ رایانامه: [shabanialireza565@gmail.com](mailto:shabanialireza565@gmail.com)

۲. گروه گیاهپزشکی و کشاورزی پایدار، سازمان پژوهش‌های علمی و صنعتی ایران، تهران، ایران؛ رایانامه: [tafaghodinia@gmail.com](mailto:tafaghodinia@gmail.com)

۳. گروه گیاهپزشکی، دانشکده کشاورزی، دانشگاه کشاورزی و منابع طبیعی رامین، اهواز، ایران؛ رایانامه: [nzandisohani@yahoo.com](mailto:nzandisohani@yahoo.com)

\* نویسنده مسئول

### چکیده

پژوهش حاضر با هدف پیش‌بینی تراکم کنه تارتن دولکه‌ای با روش‌های زمین‌آمار کریجینگ و شبکه عصبی مصنوعی بهینه شده با الگوریتم رقابت استعماری در کشتزار خیار شهرستان بهبهان استان خوزستان انجام گرفت. داده‌های مربوط به تراکم جمعیت کنه تارتن دولکه‌ای با نمونه‌برداری روی یک شبکه علامت‌گذاری شده مربعی با ابعاد ۱۰×۱۰ متر و در مجموع از ۱۰۰ نقطه از سطح کشتزار به دست آمد. مختصات طول و عرض نقاط علامت‌گذاری شده سطح کشتزار به عنوان ورودی‌های هر دو روش تعریف شد. خروجی هر روش نیز تعداد این آفت در آن نقاط بود. در بخش زمین‌آمار از روش کریجینگ معمولی و در بخش شبکه عصبی مصنوعی، ساختار پرسپترون سه لایه با الگوریتم رقابت استعماری به عنوان بهینه‌یاب اوزان شبکه، مورد ارزیابی قرار گرفت. مقایسه نتایج زمین‌آمار و شبکه عصبی مصنوعی بیانگر توانایی زیاد شبکه عصبی در مقابل روش زمین‌آمار کریجینگ بود، به طوری که در شبکه عصبی مصنوعی با ضریب تبیین 0.98 و مجموع مربعات خطا 0.0038 خطای کم‌تر نسبت به زمین‌آمار داشت. در مجموع می‌توان چنین نتیجه گرفت که روش شبکه عصبی مصنوعی بهینه شده با الگوریتم رقابت استعماری با تلفیق دو عامل طول و عرض جغرافیایی، قادر به پیش‌بینی تراکم آفت با دقت مناسب بود.

**واژگان کلیدی:** الگوریتم؛ کریجینگ؛ توزیع آفت؛ روش‌های آماری؛ تغییر نما.

**اطلاعات مقاله:** تاریخ دریافت: ۱۳۹۶/۲/۳، تاریخ پذیرش: ۱۳۹۶/۷/۲، تاریخ چاپ: ۱۳۹۶/۷/۲۳