World Applied Sciences Journal 25 (10): 1464-1468, 2013
ISSN 1818-4952
© IDOSI Publications, 2013
DOI: 10.5829/idosi.wasj.2013.25.10.13429

Application of Self Organizing Map (SOM) to Classify Treatments of the First Order Interaction: A comparison to Analysis of Variance

O.M. Ibrahim, A.T. Thalooth and Elham A. Badr

Field Crops Research Department,
National Research Centre, Dokki, Giza, Egypt

Submitted: Sep 26, 2013; **Accepted:** Oct 30, 2013; **Published:** Nov 3, 2013

Abstract: A self organizing map (SOM), an unsupervised artificial neural network (ANN) was applied to classify first order interaction treatments using normally distributed agronomic characters of corn, grain yield (g/plot), number of grains/row, 100 grains weight(g). The SOM showed its convenience to analyze differences among treatments. The SOM classified treatments into clusters which consisting of nodes where treatments in the same node are more similar than treatments in different nodes in the same cluster. However, treatments in the same cluster are more similar than treatments in different clusters. The results of SOM were similar to ANOVA in grain yield; however, in number of grains/row and 100 grains weights (g), ANOVA revealed no significant differences and SOM classified the treatments into five and four clusters, respectively. The SOM can perform analysis of higher degree of interaction such as fourth and fifth degree and higher. Also, SOM can perform analysis for four and five factors and more. The results suggested that using self organizing map is helpful to discriminate clearly the differences among factors and treatments.

Key words: Artificial neural networks · Corn · Salicylic acid · Salinity · Self organizing map

INTRODUCTION

Self-organizing map (SOM) was developed by Kohonen [1]. It is a data analysis, visualization and interpretation tool that is based on the principles of vector quantization and measures of vector similarity. It can be used to perform broad categories of operations such as, function fitting, prediction, estimation, clustering, pattern recognition or noise reduction and classification. In SOM analysis, each treatment is treated as an n-dimensional vector in a data space defined by its parameters. This treatment vector quantization approach means that both continuous and categorical parameters can be input, making the SOM technique ideal for the analysis of complex and disparate data. Because SOM is unsupervised, no prior knowledge is required as to the nature, or number, of "groupings" within the data set. These features are why the SOM technique has advantages over other more 'conventional' analysis methods such as clustering, factor analysis, principal components and traditional neural networks. In this paper we introduce self organizing maps [1] as an analysis technique for understanding relationships within and between treatments in comparison to analysis of variance (ANOVA) and to provide a means of analyzing and interpreting results in a meaningful fashion. Self-organizing map analysis procedures are widely used in fields such as finance, industrial control, speech analysis [2] and astronomy [3]. The network of SOM usually consists of two layers of neurons: an input layer and output layer. The neurons on input layer are fully connected to the neurons on output layer, whereas the neurons on each layer do not connect to the neurons in their layer regardless of their relative position. It is usually that neurons on the output layer are arranged in either a rectangular or hexagonal lattice [4].

The following Fig 1. shows the layers and connections of SOM.

Corresponding Author: O.M. Ibrahim, Field Crops Research Department,
National Research Centre, Dokki, Giza, Egypt.

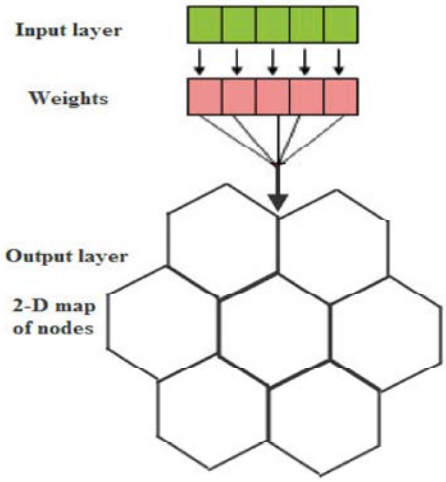

Fig. 1: Layers and connections of SOM

MATERIALS AND METHODS

An experiment was set up during the summer seasons of 2012 and 2013, Soil Salinity Laboratory, Alexandria, Egypt. Grains of maize cultivar (Gemmeiza 12) were obtained from Crops Research Institute, Agricultural Research Center, Giza, Egypt. The experiment was conducted to investigate the effect of three levels of salicylic acid (SA) (0, 100, 200 ppm) and water salinity (S) (0.5, 2.75, 5.5 dS/m) on yield and yield components of maize plants. Physical and chemical properties of the soil site (Table 1) were analyzed according to Chapman and Pratt [5].

The experimental design was split plot with four replicates where water salinity was located in the main plot, while Salicylic acid (SA) levels in the subplot, the experimental unit was a cemented plot with a dimension of 150 cm in long and 75 cm in wide with an area of 1.125 m^2. Every cemented plot contains four rows. Grains of maize cultivar (Gemmeiza 12) were sown in the 1st of May. Before sowing the cemented plot were prepared by adding calcium superphosphate (15.5% P_2O_5) at a rate of 100 kg/feddan (one feddan = 0.42 ha) and potassium sulfate (48% K_2O_4) at a rate of 50 kg/feddan while nitrogen fertilizer was added at the rate of 125 kg N/feddan in the form of ammonium sulfate (20.5% N) at three doses, the first at sowing, the second at the first irrigation and the third at the second irrigation. Salicylic acid (SA) was initially dissolved in a few drops of Dimethylsulfoxide and the final volume was reached by adding distilled water, then the pH was adjusted at 6-7 with NaOH (1.0N). A constant volume of solutions were sprayed twice on the leaves in the early morning when the plants had their fourth leaf and two weeks later.

At harvest time (120 days from sowing) the following characters were measured:

- Number of grains/row.
- 100 kernels weight (g).
- Grain yield (g/plot).

Statistical Analysis: The data was subjected to analysis of variance according to Snedecor and Cochran [6] using SAS computer software v.9.1.3 SP4 [7]. The mean values were compared with the least significant difference test.

RESULTS AND DISCUSSION

The SOM classified the treatments into clusters according to the differences in grain yield, number of grains/row and 100 grains weight, respectively (Fig. 2).

Table 1: Physical and chemical analyses of the used soil based on soil paste extract

Sand %	Silt %	Clay %	Soil texture	E.C. dS/m	pH	S.P%	SAR	Ca^{2+} meq/l	Mg^{2+} meq/l	Na^{+} meq/l	K^{+} meq/l	Cl^{-} meq/l	HCO_3 meq/l
74.0	10.4	15.6	Sandy loam	1.82	7.53	43.33	1.49	1.50	12.5	3.95	0.34	10.5	2.80

Table 2: Mean squares of grains yield (g/plot), number of grains/row, 100 grains weight (g) as affected by the interaction between salinity and salicylic acid

SOV	DF	Grain yield (g/plot)	Number of grains/row	100 grains weight(g)
Rep.	3	2674.03	38.39	18.78
S	2	220159.64*	684.69*	109.71*
Rep.*S	6	712.86	10.73	3.54
SA	2	18084.55*	388.19*	104.90*
S*SA	4	1334.24*	1.81 ns	3.01 ns
Error	18	287.69	17.98	2.95
Corrected Total	35	--	--	--
Coefficient of variation (CV %)	--	5.38	15.22	7.28

Rep=Replicates, S=Salinity, SA=Salicylic acid

Table 3: Means of grains yield (g/plot), number of grains/row, 100 grains weight (g) as affected by the interaction between salinity and salicylic acid

Treatments	Grain yield (g/plot)	Number of grains/row	100 grains weight (g)
S_1*SA_1	407.02 c	30.25 a	23.68 a
S_1*SA_2	457.58 b	37.75 a	27.41 a
S_1*SA_3	517.98 a	41.50 a	29.67 a
S_2*SA_1	262.03 e	19.25 a	20.40 a
S_2*SA_2	276.58 e	25.25 a	23.91 a
S_2*SA_3	337.54 d	29.25 a	24.60 a
S_3*SA_1	170.05 g	15.75 a	17.25 a
S_3*SA_2	194.36 fg	23.50 a	21.08 a
S_3*SA_3	214.50 f	28.25 a	24.60 a

$S_{1=0.5\ dS/m}$, $S_{2=2.75\ dS/m}$, $S_{3=5.5dS/m}$, $SA_{1=control}$, $SA_{2=100\ ppm}$, $SA_{3=200ppm}$

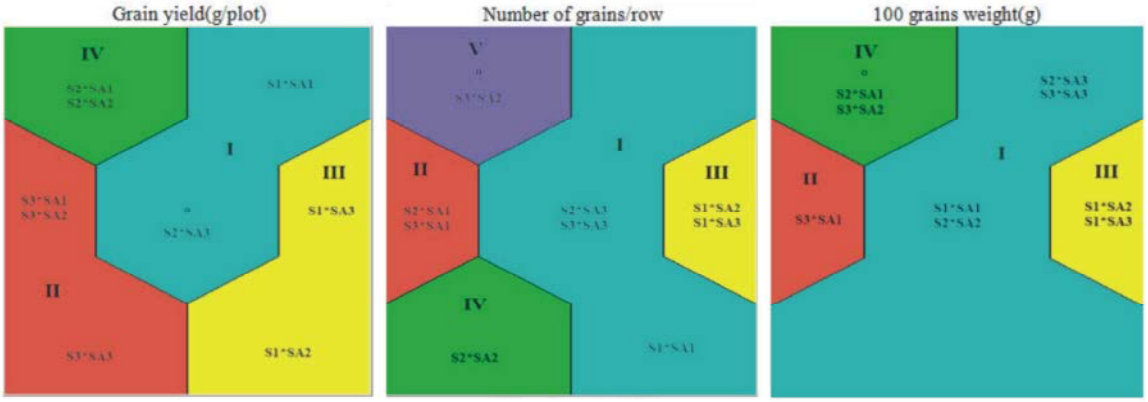

Fig. 2: Clustering of the trained SOM units. The U-matrix and Ward's method were applied to set boundaries on the SOM map. The Latin numbers (I-IV) display clusters and the codes in each unit of the map represent the interaction.

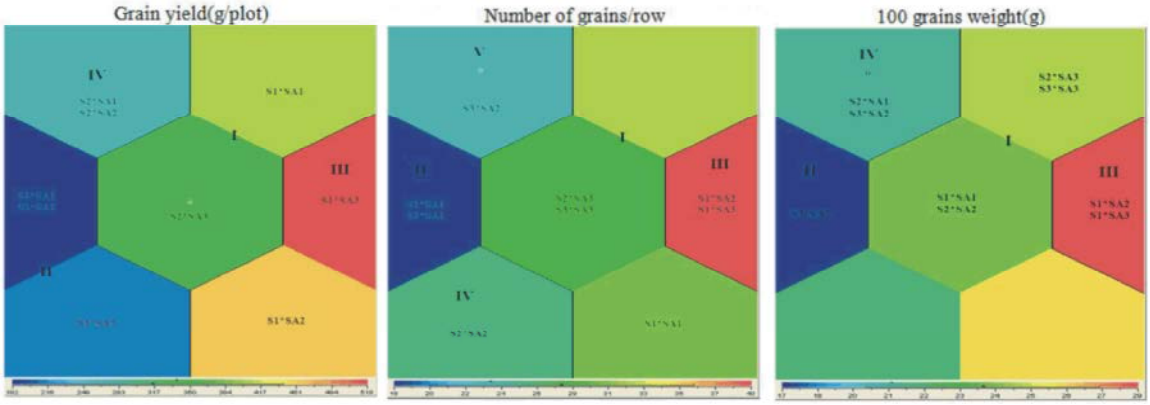

Fig. 3: Visualization of salinity and salicylic acid factors in the trained SOM in color scale. Red color represents high value, whereas blue is low value.

The units of the SOM map were divided into the four, five and four main clusters, respectively (I–V) based on a hierarchical cluster analysis of weight vectors of the trained SOM (Fig. 2). The clusters were indicated with solid and jagged lines. In each map unit the code inside the node represents the treatments of the first order interaction for each parameter. Fig. 3 showed the distribution of the connection weights for each parameter on the trained SOM map with a color scale, displaying the effect of each parameter on patterning of input datasets. Red represented high values, whereas blue indicated low values. With regard to grain yield, all treatments of third

level of salinity (S_3*SA_1, S_3*SA_2, S_3*SA_3) were located in the middle and lower left areas in the SOM map (clusters II), whereas two treatments of the first level of salinity (S_1*SA_2, S_1*SA_3) were placed in the middle and lower right areas (clusters III). Two treatments of the second level of salinity (S_2*SA_1, S_2*SA_2) were located in the upper left area in the SOM map (clusters IV). The rest of the treatments (S_1*SA_1, S_2*SA_3) were placed in the middle and upper right areas (clusters I). SOM clearly discriminated the treatments of salinity levels.

The results of SOM are in accordance with those obtained from ANOVA which reveal that significant differences were observed among treatments of grain yield (Tables 2 and 3). These results are in agreement with those reported by Hussein *et al.* [8], who mentioned that a negative relationship between vegetative growth of corn parameters and the increase in salt concentration in irrigation water, however, spraying plants with 200 ppm salicylic acid improved all growth characters. Farahbakhsha and Shamsaddin Saiidb [9] also added that plant vegetative growth traits were improved by increasing SA concentration up to 200 ppm. Mahdi *et al.* [10] revealed that the effect of salicylic acid spraying on the growth of morphological traits and increasing in the corn yield was considerable. Such enhancing effect of SA on maize salt tolerance in terms of improving the measured plant growth criteria may be attributed to the stimulating effect of SA by activating the photosynthetic process [11]. Regarding the number of grains/row, two treatments of first level of salinity (S_1*SA_2, S_1*SA_3) were located in the middle right area in the SOM map (clusters III), the treatments of (S_2*SA_3, S_3*SA_3, S_1*SA_1) were placed in the middle and lower right areas (clusters I). Treatments of (S_2*SA_1, S_3*SA_1) were located in the middle left area in the SOM map (clusters II). The treatments (S_2*SA_2), (S_3*SA_2) were placed in the lower and upper left areas, (clusters IV) and (clusters V), respectively.

Concerning the 100 grains weight, two treatments of first level of salinity (S_1*SA_2, S_1*SA_3) were located in the middle right area in the SOM map (clusters III), the treatments of (S_1*SA_1, S_2*SA_2), (S_2*SA_3, S_3*SA_3) were placed in the middle and upper right areas (clusters I). Treatment of (S_3*SA_1) was located in the middle left area in the SOM map (clusters II). The treatments (S_2*SA_1, S_3*SA_2) were placed in the upper left area (clusters IV). The results of SOM are in contradiction with those obtained from ANOVA which reveal that no significant differences were observed among treatments of number of grains/row and 100 grains weight, (Tables 2, 3).

CONCLUSION

On conclusion, we have presented a novel application of self-organizing maps by using them on agricultural yield data. After a statistical analysis of the available data sets, we briefly outlined the advantages of self-organizing maps in data visualization. A hypothesis on the differences among treatments could clearly be confirmed by using SOMs. The presented results are very promising and show that differences in the data sets can easily be assessed by visual inspection of the resulting component planes of the self-organizing map. The results are of immediate practical usefulness and demonstrate the advantage of using data mining techniques in agriculture.

REFERENCES

1. Kohonen, T., 2001. Self Organizing Maps. Springer, New York, USA.
2. Kaski, S., T. Honkela and T. Kohonen, 1998. Websom: Self-organizing maps of document collections. Neurocomputing, 21(1): 101-117.
3. Garcia B.E., T.S. Santiago and J. Isern, 2003. Using self-organizing maps to identify potential halo white dwarfs, Neural Networks, 16: 405-410.
4. Lansiluoto, A., 2004. Economic and Competitive Environment Analysis in the Formulation of Strategy a Decision-Oriented Study Utilizing Self-Organizing Maps, Publications of the Turku school of Economics and Business Administration, Turku.
5. Chapman, H.O. and P.E. Pratt, 1978. Methods of analysis for soil, plants and waters. Univ. Calif. Div. Agic. Sci.
6. Snedecor, G.W. and W.G. Cochran, 1982. Statistical Methods. 7th Ed., Iowa state Press Iowa, USA.
7. SAS Institute Inc., 2003. SAS Software v.9.1.3 sp4, Cary, NC: SAS Institute Inc.
8. Hussein, M.M., L.K. Balbaa and M.S. Gaballah, 2007. Salicylic acid and salinity effects on growth of maize plants. Research Journal of Agriculture and Biological Sciences, 3(4): 321-328.

9. Farahbakhsha, H. and M. ShamsaddinSaiidb, 2010. Effects of foliar application of salicylic acid on vegetative growth of maize under saline conditions. Plant Ecophysiology, 2: 27-30.
10. Mahdi, Z., K. Kh. Saeid, J.M. Matin and R.H. Ali, 2013. Effect of salicylic acid on morphological characteristics, yield and yield components of Corn (*Zea mays* L.) under drought condition. European Journal of Experimental Biology, 3(2): 153-161.
11. Khodary, S.E.A., 2004. Effect of salicylic acid on the growth, photosynthesis and carbohydrate metabolism in salt stressed maize plants. Int. J. Agric. Biol., 6(1): 1-8.