Australian Journal of Crop Science AJCS
AJCS 10(4):539-545 (2016) ISSN:1835-2707
DOI: 10.21475/ajcs.2016.10.04.p7336x

The self-organizing map for determination of main features related to biological yield and yield of wheat

Ehsan Bijanzadeh*[1], Marzieh Mokarram[2]

[1]Department of Agroecology, College of Agriculture and Natural Resources of Darab, Shiraz University, Shiraz, Iran
[2]Department of Range and Watershed Management, College of Agriculture and Natural Resources of Darab, Shiraz University, Shiraz, Iran

***Corresponding author: bijanzd@shirazu.ac.ir**

Abstract

Among various methods of artificial neural networks (ANNs) and learning algorithms, self-organizing map (SOM) is one of the most popular models. The aim of this study is to classify features influencing the biological yield and yield of wheat using SOM algorithm. In SOM, according to qualitative data, the clustering tendency of yield and biological yield of wheat were investigated using 11142 data from 16 features. Data was collected from the literatures on the subject of wheat in Iran that was existed in http://sid.ir website. Results showed that when biological yield was as output, K with soil pH, irrigation regime with 1000-kernel weight and organic content (OC) with grain/spike were related to each other closely. Moreover, grain/spike and OC had closer relationship to biological yield. In contrast, negative relationship was observed between soil pH (r= -0.47) and HI (r= -0.61) with biological yield. When wheat grain yield was output of SOM model, K with soil pH, and P with OC was related to each other closely. Overall, grain/spike, P and OC were much closer related to crop yield than other parameters. Similar to biological yield, labels map showed that data classified in three classes for wheat yield and the top four rows of U-matrix were placed in class A. A clear separation was observed among class A with B and C. The characteristics of each group in the study area showed that group 2 with 0.784 (kg/m^2) had the highest yield than group 1 (0.241 kg/m^2) and group 3 (0.401 kg/m^2) so that in group 2, the amount of P (0.003 kg/m^2), OC (0.47%), pH (7.78), rainfall (492.45 mm), grain/spike (43.71) and spike/m^2 (668.21) and HI (37.53%) were higher than the other groups and related to yield directly. Our results showed that among the yield components, grain/spike was the most important features contributing to grain yield than spike/m^2 and 1000-kernel weight using SOM.

Keywords: Artificial neural network, component layers, self-organizing maps, wheat yield.
Abbreviations: ANNs_Artificial neural networks; HI_harvest index; OC_organic content; SOM_self-organizing map.

Introduction

Artificial neural networks (ANNs) are similar to biological neural networks in performing functions. It can get solutions with ameliorated performance compared with traditional methods. It usually refers to models applied in statistics and artificial intelligence. Neural network models which emulate the central nervous system are part of theoretical neuroscience and computational neuroscience (Rumelhart and McClelland, 1986). Among various methods of ANNs and learning algorithms, self-organizing map (SOM) is one of the most popular neural network models. It belongs to the category of competitive learning networks that it is trained using unsupervised learning to produce a low-dimensional (typically two-dimensional) and discretized representation of the input space of the training samples, called a map (Vesanto et al., 1999). SOM offers a solution to apply a number of visualizations linked together (Buza et al., 1991). When several visualizations are linked together, scanning through them is very efficient because they are interpreted in a similar way. The U-matrix produced from SOM visualizes distances between neighboring map units and thus shows the cluster structure of the map. Samples within the same cluster will be the most similar according to the variables considered (Dhubkarya et al., 2010). Topologically, preserved mapping from input to output space can provide by the SOM algorithm. The SOM algorithm is optimal for vector quantization. It is applicable to many applications such as clustering, classification, and data visualization. SOM have been applied as a clustering and projection algorithm for high dimensional data (Kohonen, 1995). Ferentinou and Sakellariou (2005 and 2007) applied SOM in order to rate slope stability controlling variables in natural slopes, while Ferentinou et al. (2010) used SOM to classify marine sediments. Olawoyin et al. (2013) used SOM for the categorization of water, soil and sediment quality in petrochemical regions. Their results showed valuable assessment using the SOM visualization capabilities and highlighted zones of priority that might require additional investigations and also provided productive pathway for effective decision making and remedial actions. Also, Wang et al. (2009) applied SOM to identify functional groups. For their study, quantitative traits and distributional information on 127 invasive plants in 28 provinces of China were collected to form the matrices for their study. The results indicated that Jiangsu was the top province with the highest number of invasive species, while Ningxia was the lowest. Klobucar and Subasic (2012) used SOM in the visualization and analysis of forest inventory and showed that SOM performed a nonlinear dimensionality reduction and good clustering, which is a good basis for data visualization results. Mokarram et al. (2014) used SOM for relationships between geomorphological features of the fans and their drainage basins. The results of the analysis showed that several morphologically different fan types were recognized based on their geomorphological characteristics in the study area.

Wheat is the most important food crop in Iran and many other countries (Emam, 2007). Approximately, 32% of wheat-growing areas in developing countries experience some types of limiting factors such as abiotic stresses, nutrients deficiency and etc. (FAO, 2009). Traditionally, agricultural research has focused primarily on maximizing total production. However, in recent years, focus has shifted to the limiting factors in crop production systems (Katerji et al., 2008). Up to now, researchers have only considered a limited number of characteristics under wheat field conditions that related to biological yield and crop yield. It has now become obvious that analyzing a large number of factors under different field conditions can provide a comprehensive overview of important features responsible for wheat yield improvement (Bijanzadeh et al., 2010). Understanding the importance of attributes among a large dataset of features can play a key role in wheat yield improvement. The aim of this study was determination of main features related to biological yield and wheat yield, located in Iran, using SOM. It seems that this is the first report about classification of the effective parameters in the biological and crop yield of wheat by SOM method.

Results and Discussion

SOM for biological yield

The visualization in Fig. 1 for biological yield as output consisted of 16 hexagonal grids, with the U-matrix in the upper left, along with the 16 component layers, so one layer for each morphometric parameter examined in this study. According to Fig. 1, the seventeen figures were linked by position. In each figure, the hexagon in a certain position corresponded to the same map unit. The legend for each of the hexagons showed degree of color compared to each other. In SOM method, similar colors showed direct relationship between the parameters. It could be seen that K with soil pH, Irrigation regime with 1000-kernel weight and OC with grain/spike were related to each other closely. Also, grain/spike and OC had closer relationship to biological yield than other parameters. In contrast, soil pH, spike/m^2 and HI had negative relationships to biological yield (Fig. 1).

As was shown in Fig. 2, number written in hexagons were data that absorbed by each of nodes in the neural network (Venna and Kaski, 2001). According to Fig. 2a, the maximum number of hexagons was 7, indicating that the maximum data for these areas was 7. On the other hand, the minimum number of hexagons was 0, indicating that these areas had no data. Also, Principal Component Projection (PCP) showed that study data had high density (Fig. 2b). In fact data had good distribution. Labels map (Fig. 2c) showed that data classified in three classes for biological yield and the top four rows of U-matrix was as class A. The other classes were B and C form and there was clear separation between class A and B in the U-matrix of labels.

The characteristics of each group related to biological yield that determined by label map (Fig. 2c) provided in the Table 1. Considering the data from group 1 to group 3 showed that by increasing N from 0.0008 to 0.0104 (kg/m^2), K from 0.0006 to 0.0024 (kg/m^2), plant density from 178.18 to 301.52 and growing season from 216.44 to 231.39 (d) the highest biological yield (1.207 kg/m^2) was observed in group 3. Also, increasing the rainfall amount, plant height, 1000-kernel weight and grain/spike related to biological yield improvement in group 3 (Table 1). In contrast, negative relationship was observed between soil pH and HI with biological yield.

Bijanzadeh et al. (2012) reported that field water status, such as irrigation regime or rainfall, was another important features related to biological yield, and K and P applied to the soil value) was not found to be important using attribute-weighting models. In our study, no strong relationship was observed between K, P, and irrigation regimes with biological yield. On the other hand, Malakoti (2003) found that southern soils of Iran were rich in available potassium ions, and farmers often did not apply potassium fertilizer in these areas. Similar to our results, Ghodsi et al. (2005) also reported a positive relationship between spike number/m^2 and biological yield. Bijanzadeh et al. (2012) reported that when biological yield was as output, nitrogen and grain yield had a strong relationship with biological yield, with values of 0.5 to 1.0 in various attribute-weighting algorithms. Farahani and Arzani (2007) found that grain/spike and plant height was correlated to biological yield, positively. Recently, Emam et al. (2009) showed that nitrogen applied to the soil, a key element in crop nutrition, had an important role in increasing biological yield and wheat grain yield. Interestingly, no strong relationships were observed between N applied and plant height with biological yield while increasing OC related to biological yield improvement (Figure 1 and Table 1).

SOM for wheat yield

According to Fig. 3, it could be seen that K with soil pH, P with OC, and OC with grain/spike were related to each other closely. Also, negative relationships were observed between K and soil pH with yield. Interestingly, no positive relationship was observed between HI and 1000-kernel weight with yield. Generally, grain/spike, P and OC were much closer related to crop yield than other parameters.

In Fig. 4a, hexagons of 7 showed that 7 data were absorb in the place or number of 0 showed that in these places there was no data. Also, PCP showed that study data had high density and good distribution (Fig. 4b). Label map determine that study data classify in three classes for crop yields (Fig. 4c). Similar to biological yield, labels map (Fig. 4c) showed that data classified in three classes for wheat yield and the top four rows of U-matrix was as class A. Likewise, a clear separation was observed between class A with B and C in the U-matrix.

Data characteristics of each group for determination of SOM algorithm showed in the Table 2. It is clear that group 2 with 0.784 kg/m^2 had the highest yield than group1 (0.241 kg/m^2) and group 3 (0.401 kg/m^2). As shown in Table 3, in group 2 the amount of P (0.003 kg/m^2), OC (0.47%), pH (7.78), rainfall (492.45 mm), grain/spike (43.71), spike/m^2 (668.21) and HI (37.53%) were higher than the other groups and related to yield directly. Also, wheat yield was not affected by plant density, growing season, soil pH, and 1000-kernel weight. Bijanzadeh et al. (2010) reported that based on supervised feature selection model, OC and rainfall amount affected wheat grain yield, whereas soil pH had a marginal effect on wheat grain yield. They also demonstrated that factors classification using feature selection algorithm may be

Table 1. Characteristicsof each group in the study area for biological yield.

Parameters	Group 1	Group 2	Group 3
Irrigation regime(according to FC)	52.73	92.05	92.99
N (kg/m^2)	0.0008	0.0006	0.0104
P (kg/m^2)	0.0003	0.0034	0.0034
K (kg/m^2)	0.00006	0.0005	0.0024
Plant density (plant/m^2)	178.18	190.4	301.52
Growing season (day)	216.44	228.6	231.39
OC %	0.34	0.49	0.56
Soil pH	7.81	7.64	6.18
Rianfall amount(mm)	321.21	465.05	483.02
Plant height (cm)	64.17	76.16	81.05
Grain/spike	31.32	44.45	38.98
Spike/m^2	337.11	714.73	489.84
1000 kernel weight(gr)	31.04	33.41	36.86
HI %	38.11	34.04	32.13
BY (kg)	3868.89	7519.73	12074.16

Table 2. Characteristics of each group in the study area for wheat yield.

Parameters	Group 1	Group 2	Group 3
Irrigation regime (according to FC)	100.00	89.75	56.60
N (kg/m^2)	0.01	0.0077	0.0012
P (kg/m^2)	0.0025	0.0034	0.0007
K (kg/m^2)	0.00	0.0006	0.00
Plant density (plant/m^2)	100.00	171.70	175.47
Growing Season (day)	223.00	223.21	223.34
OC %	0.36	0.47	0.35
Soil pH	7.50	7.69	7.78
Rianfall amount(mm)	423.00	492.45	415.56
Plant height (cm)	85.60	75.70	63.07
Grain/spike	42.50	42.71	44.71
Spike/m^2	232.00	668.21	590.17
1000 kernel weight(gr)	54.00	32.71	29.66
HI %	31.50	37.53	31.63
Yield (kg)	2413.00	7841.25	4017.43

Table 3. Maximum, minimum and average of each data to find the main features related to biological yield and wheat yield.

Features	Data properties		
	Maximum	Minimum	Average
Irrigation regime(according to FC)	125.00	40.00	86.10
N(kg/m^2)	0.0138	0.00	0.0072
P(kg/m^2)	0.0090	0.00	0.0027
K(kg/m^2)	0.0050	0.00	0.0009
Plant density(plant/m^2)	500.00	100.00	226.75
Growing Season (d)	263.00	152.00	226.42
OC %	1.10	0.30	0.48
Soil pH	7.90	7.10	7.65
Rainfall amount(mm)	751.00	140.00	420.05
Plant height (cm)	142.30	48.00	77.85
Grain/Spike	76.27	22.70	39.38
Spike/m^2	1925.28	118.00	510.73
1000-kernel weight(g)	63.80	22.79	36.92
HI(%)	50.00	19.03	33.64
Biological yield(kg/m^2)	29553.26	2412.70	10077.01
Yield (kg/m^2)	9300.00	460.00	3310.33

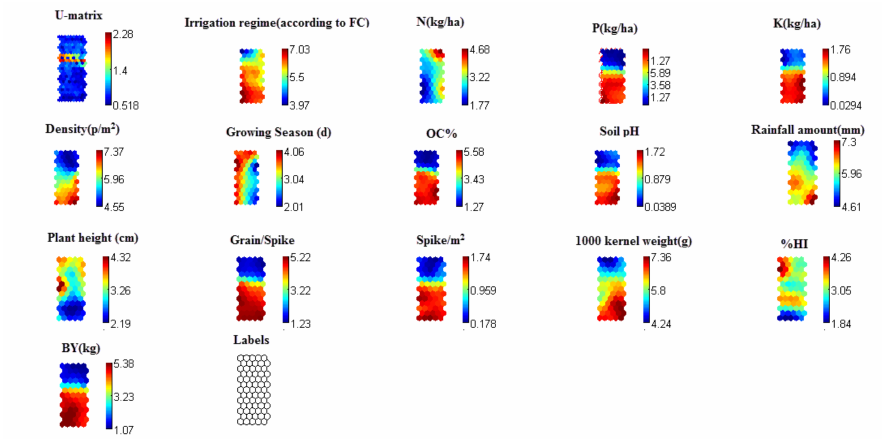
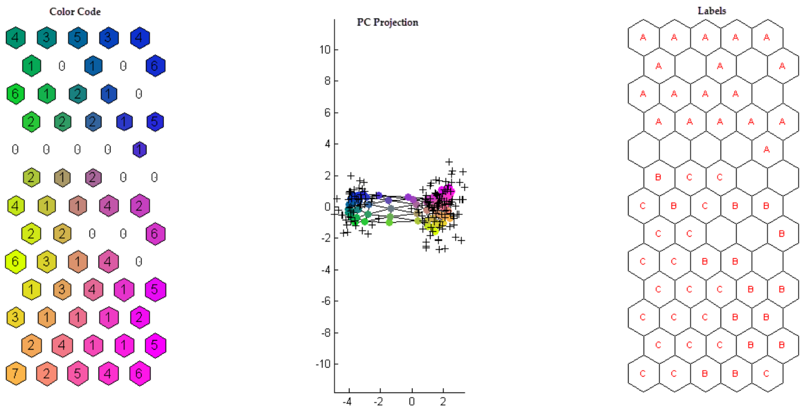


Fig 1. Self-organizing maps visualization through U-matrix (top left) and 16 component layers for biological yield.



a) *(b)* *(c)*

Fig 2. Different visualizations of the clusters obtained from the classification of the morphological variation through SOM. Color code (a); Principal component projection (b); Label map with the names of biological yield (c).

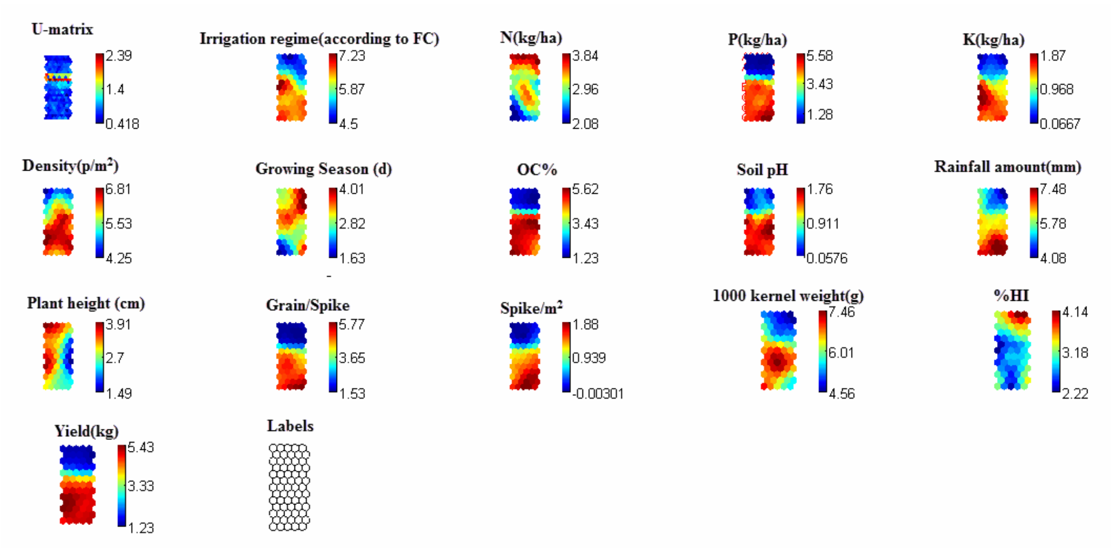
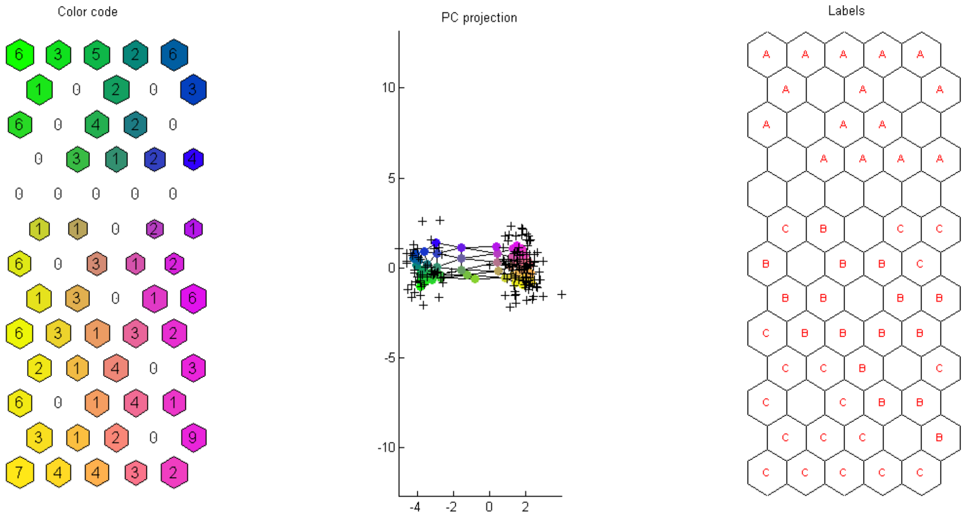


Fig 3. Self-organizing maps visualization through U-matrix (top left) and 16 component layers for yield.



(a) (b) (c)

Fig 4. Different visualizations of the clusters obtained from the classification of the morphological variation through SOM. Color code (a); Principal component projection (b); Label map with the names of wheat yield (c).

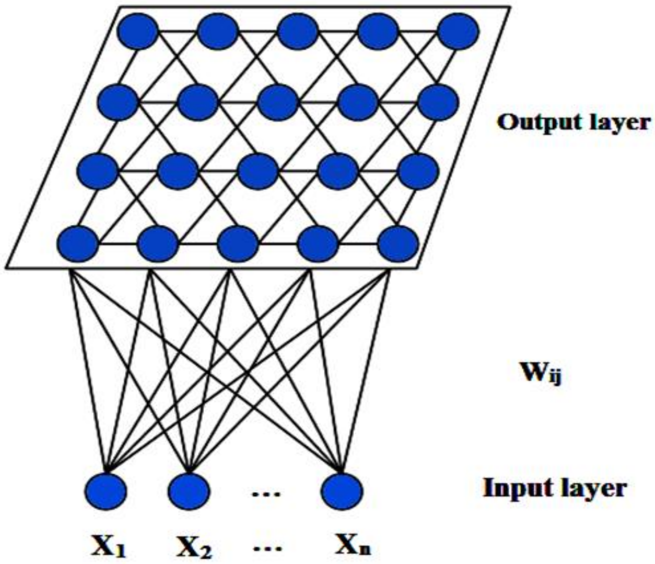


Fig 5. The structure of a SOM network (adapted from Dykes et al., 2005).

a suitable option for determining the important factors contributing to wheat grain yield, and for providing a comprehensive view of different traits. Using unsupervised weighting algorithms, Bijanzadeh et al. (2012) reported that when grain yield was as output, nitrogen and rainfall amount had a strong relationship with grain yield, with values of 0.5 to 1.0. In various attribute-weighting algorithms, harvest index was less important in modern wheat genotypes and was only selected by the Relief model when biological yield or grain yield were the outputs. Emam (2007) declared that an alternative for grain yield improvement was increasing the yield components including grain/spike and spike/m^2. On the other hand, Sharma and Smith (1996) found that wheat grain yield may be increased by improving biomass at a given level of harvest index in three winter wheat populations. Farid et al. (1996) reported that improving harvest index appears to be difficult, and recent increases in wheat grain yield have been attributed to increases in grain/spike. Our results showed that among the yield components, grain/spike was the most important features contributing to grain yield than spike/m^2 and 1000-kernel weight (Fig. 3 and Table 2).

In similar study on barley, soil organic content (0.941 values), electrical conductivity of water (0.911), harvest index (0.905), and plant density (0.904) had the marginal effect on barley grain yield. The rest of the features including 1000-kernel weight, soil texture, plant height, soil pH, and potassium and phosphorus applied to the soil were recognized to be unimportant (Bijanzadeh and Naderi, 2014). Opposite to our results, Bijanzadeh and Mokarram (2013) used fuzzy- AHP methods to assess fertility classes for wheat and its relationship with soil salinity and declared that soil salinity had a high correlation (R^2=0.82) with wheat yield and significant relationship was observed between soil salinity and fertility and saline areas had low fertility compared to non-saline areas.

Materials and methods

Data collection

Data presented in this study was collected from the literatures on the subject of wheat in Iran that was existed in http://sid.ir website. A total of 11142 data from 16 features, including irrigation regime (according to FC), nitrogen (N), phosphorus (P) and potassium (K) applied to the soil (kg/m^2), plant density (plant/m^2), growing season length (days), soil organic content (OC,%), soil pH, rainfall amount (mm), plant height (cm), grain/spike, spike/m^2, 1000 kernel weight (g), harvest index (HI%), biological yield (BY, kg/m^2), and grain yield (kg/m^2) were prepared in Excel software sheets. The amount of each feature including maximum, minimum, and average were shown in Table 3.

Self-organizing map (SOM)

Artificial neural networks (ANNs) are as non-linear mapping structures and powerful tools for modeling when the underlying data relationship based on the function of the human brain. ANNs can determine and learn correlated patterns between input data sets and corresponding target values, after which can be used to predict the outcome of new independent input data. ANNs are similar the learning process of the human brain and complex data even if the data is imprecise and therefore, are ideally suited for the modeling of crop plant, which are known to be often non-linear. ANNs have great capacity in predictive modeling, whereby all the parameters describing the unknown situation can be presented to the trained ANNs (Dhubkarya et al., 2010).
Self-Organizing Map is a type of neural network. It was developed in 1995 by Kohonen, a professor emeritus of the Academy of Finland. SOM are unsupervised ANNs formed from neurons located on a regular, two-dimensional regular planar array grid (Fig. 5). In fact SOM is based on unsupervised learning, which means that no human intervention is needed during the learning and little needs to be known about the characteristics of the input data (Lee et al., 2007; Mokarram et al., 2010). SOM offers a solution to apply a number of visualizations linked together (Buza et al., 1991). When several visualizations are linked together, scanning through them is very efficient because they are interpreted in a similar way. The U-matrix produced from SOM visualizes distances between neighboring map units and thus shows the cluster structure of the map. Samples within the same cluster will be the most similar according to the variables considered (Dhubkarya et al., 2010).
SOM algorithm is comprised of two layers (Lee et al., 2007). In the input layer, neurons represent the inputs. In the second layer, the competitive process is done and the weight of connection is updated to choose a winner neuron (Ghadamyari and Safavi, 2011).
In the input layer, the output of each neurons, x_i for i=1, 2, 3, …, is connected to all neurons of competitive layer and each connection is assigned a variable weight, w_{ij} for i=1, 2, 3, ….
SOM algorithm operates as follow:
1) Initialization: in the first step a random weight shall be assigned to each connection.
2) Sampling: one member of the input space is chosen.
3) Matching: the winning neuron is chosen when the weight vector of this neuron is 1.
4) Updating: the weight update law is applied.
5) Continuation: this process is repeated until the ultimate goal is achieved.

Chosen input will be compared with all weight of connections according to the following equation (Lee et al., 2007; Ghadamyari and Safavi, 2011):

$$d_j = \sqrt{\sum_{i=1}^{d} (x_i - w_{ji})^2} \qquad (1)$$

$$d_j^* = \min d_j$$

Where the winner neuron x_i, is specified when the weight vector is closest to input space and $d_j^*(x)$ is minimized.
Afterward, by Kohonen weight update law, weights are updated according to following equation (Kohonen, 1995)

$$w_{i+1} = \begin{cases} (1-\alpha)w_i + \alpha x_i & \text{winner neuron} \\ w_i & \text{other neuron} \end{cases} \qquad (2)$$

Where α is the learning rate (0<α<1) (Lee et al., 2007).
Statistical analyses for SOM algorithm were performed using Matlab software (2014). First, data were transported from Excel to Matlab and biological yield and wheat grain yield was set as output variables and the other variables were set as input.

Conclusion

The results showed that the SOM is an excellent tool in the visualization of high dimensional data. As such SOM method is most suitable for data understanding phase of the knowledge discovery process. SOM method consisted of U-matrix, PC projection and label. Using of the U-matrix can show almost clear cluster that in the labels by rows color. U-matrix showed that some of the data have closely related to each other. Additional using PC projection can find density of data. In our study PC projection showed that study data had high density for biological and crop yields. Finally, label map in the SOM method determined three classes for biological yield and crop yield so that the group 3 and group 2 features were better intervals than other groups for biological yield and crop yield, respectively. Overall, the important features related to biological yield and grain yield improvement were grain/spike and OC. For the first time, our results showed that SOM can provide a comprehensive view of important features contribute to wheat grain yield improvement and there might be a scope by selecting cultivars with a higher grain/spike than spike/m^2 or 1000-kernel weight. This study opened a new vista in wheat production and finding the main factors contributing to biological yield and wheat grain yield from different wheat field conditions by SOM.

Acknowledgments

We would like to thank the Research Council of Shiraz University in Iran for financial support of this research.

References

Bijanzadeh E, Mokarram M (2013) The use of fuzzy- AHP methods to assess fertility classes for wheat and its relationship with soil salinity: east of shiraz, Iran: a case study. Aust J Crop Sci. 7:1699-1706.
Bijanzadeh E, Emam Y, Ebrahimie E (2010) Determining the most important features contributing to wheat grain yield using supervised feature selection model. Aust J Crop Sci. 4: 402-407.

Bijanzadeh E, Emam Y, Ebrahimie E, Ebrahimie M (2012) Application of unsupervised weighting algorithms for identifying important attributes and factors contributing to grain and biological yields of wheat. Crop Breed J. 2:111-117.

Bijanzadeh E, Naderi R (2014) Application of supervised and unsupervised algorithms to find the important features related to barley (*Hurdeum vulgare* L.) grain yield: a new vista in data mining. Aust J Crop Sci. 8:1590-1596.

Buza A, Mcdonald JA, Michalak J, Stuetzle W (1991) Interactive data visualization using focusing and linking. In: Proceedings of IEEE conference of visualization, P. 156–163.

Dhubkarya DC, Nagariya D, Kapoor R (2010) Implementation of a radial basis function using by VHDL. Global J Com Sci Tech. 21:16-19.

Dykes J, Mac Eachren AM, Kraak MJ (2005) Exploring geovisualization. In: Dykes J, Mac Eachren AM, Kraak MJ (ed) Amsterdam : Elsevier. P. 3-19.

Emam Y (2007) Cereal production. Shiraz University Press. 190 pp. (In Persian).

Emam Y, Salimi S, Shekofa A (2009) Effect of nitrogen level on grain yield and yield components of wheat (*Triticum aestivum* L.) under irrigation and rainfed conditions. Iranian J Field Crops Res. 7: 323-334.

FAO (2009) Statistical yearbook. Rome (Italy): FAO. Available at: http://faostat.fao.org/.

Ferentinou M, Sakellariou M (2005) Assessing land slide hazard on medium and large scales, using SOM. In: Landslide Risk Management, Hungr O, Fell R, Couture R, Eberhardt E. Taylor and Francis. 639-648.

Ferentinou M, Sakellariou, M (2007) Computational intelligence tools for the prediction of slope performance. Computat Geosci. 34: 362-384.

Ferentinou M, Hasiotis T, Sakellariou M (2010) Clustering of geotechnical properties of marine sediments through self-organizing maps: An example from the Zakynthos Canyon valley system, Greece. Adv Nat Tech Haz Res. 28: 43-54.

Farahani E, Arzani A (2007) Study of genetic diversity of cultivars and F1 hybrids of durum wheat using agronomic and morphological traits. J Sci Tech Nat Res. 10: 341-355.

Farid M, Niknejad M, Emam Y (1996) Analysis of phytomass in 36 Iranian winter wheat (*Triticum aestivum* L.) genotypes. Iran Agric Res. 15: 69-77.

Ghadamyari A, Safavi AA (2011) Self organizing map (SOM) neural network based on novel fuzzy wavelet for nonlinear function approximation. 2nd International Conference on Control, Instrumentation and Automation (ICCIA).

Ghodsi M, Chaeichi MR, Mazaheri D, Jalal Kamali MR (2005) Determination of susceptibility of developmental stages in bread wheat to water stress and its effect on yield and yield components. Seed Plant. 20: 489-509.

Katerji N, Marcello M, Gianfranco R (2008) Water use efficiency of crops cultivated in the Mediterranean region: Review and analysis. Eur J Agron. 28:493-507.

Klobucar D, Subasic M (2012) Using self-organizing maps in the visualization and analysis of forest inventory. Ital Soc Sil For Eco. 5: 216-223.

Kohonen T (1995) Self organizing maps. Antiquity. 70: 588-611.

Lee KC, Cho HR, Kim JS (2007) A self-organizing feature map driven approach to fuzzy approximate reasoning. Expert Syst Appl. 33: 509-521.

Malakoti MJ (2003) Recommendations and applications of chemical fertilizers in fields of Iran. Tarbiat Modarres University Press. (In Persian).

Mokarram M, Seif A, Sathyamoorthy D (2014) Use of morphometric analysis and self-organizing maps for alluvial fan classification: case study on Oshtorankooh altitudes, Iran. IOP Conference Series: Earth Envi Sci. 6:22-41.

Olawoyin R, Nieto A, Grayson RL, Hardisty F, Oyewole S (2013) Application of artificial neural network (ANN)-self-organizing map (SOM) for the categorization of water, soil and sediment quality in petrochemical regions. Expert Syst Appl. 40: 3634-3648.

Rumelhart DE, Mc Clelland J (1986) Parallel distributed processing: explorations in the microstructure of cognition. Cambridge: MIT Press. 244 pp.

Sharma RC, Smith EL (1996) Selection for high and low harvest index in three winter wheat population. Crop Sci. 26: 1147-1150.

Venna J, Kaski S (2001) Neighborhood preservation in nonlinear projection methods: an experimental study. LN Ser Computing. 2130: 485-491.

Vesanto J, Himberg, J, Alhoniemi E, Parhankangas J (1999) Self-organizing map in matlab: the SOM toolbox. Proceedings of the Matlab DSP Conference 1999, Espoo, Finland, November 16-17, p. 35-40.

Wang ZB, You-Fang C, You-Hua C (2009) Functional grouping and establishment of distribution patterns of invasive plants in china using self-organizing maps and indicator species analysis. Arch Biol Sci. 61: 71-78.