



OPEN

A hybrid machine learning approach for estimating the water-use efficiency and yield in agriculture

Hossein Dehghanisani¹✉, Hojjat Emami², Somayeh Emami³ & Vahid Rezaverdinejad⁴

This paper introduces the narrow strip irrigation (NSI) method and aims to estimate water-use efficiency (WUE) and yield in apple orchards under NSI in the Miandoab region located southeast of Lake Urmia using a machine learning approach. To perform the estimation, a hybrid method based on an adaptive neuro-fuzzy inference system (ANFIS) and seasons optimization (SO) algorithm was proposed. According to the irrigation and climate factors, six different models have been proposed to combine the parameters in the SO-ANFIS. The proposed method is evaluated on a test data set that contains information about apple orchards in Miandoab city from 2019 to 2021. The NSI model was compared with two popular irrigation methods including two-sided furrow irrigation (TSFI) and basin irrigation (BI) on benchmark scenarios. The results justified that the NSI model increased WUE by 1.90 kg/m³ and 3.13 kg/m³, and yield by 8.57% and 14.30% compared to TSFI and BI methods, respectively. The experimental results show that the proposed SO-ANFIS has achieved the performance of 0.989 and 0.988 in terms of R^2 criterion in estimating WUE and yield of NSI irrigation method, respectively. The results confirmed that the SO-ANFIS outperformed the counterpart methods in terms of performance measures.

Water resources are declining in many regions of the world. Due to climate change, increased air temperatures, and reduced precipitation, we will face a decline in water resources in the future^{1,2}. Iran is an arid and semi-arid region in terms of climate, the amount of rainfall, and the limitations of water resources in this region. Optimal use of available water resources is an important goal of water conveyance and distribution systems. Surface irrigation is one of the most common irrigation methods in the world. More than 95% of the agricultural land in Iran is currently under the surface irrigation method. Despite the complexity of this irrigation method, researchers and users have not paid much attention to it. The current efficiency of surface irrigation in Iran is estimated at less than 35%^{1,3,4}. Surface irrigation is easy and needs inexpensive equipment to convey and distribute water in different areas. The maintenance and operation costs of the surface irrigation method are lower than other methods. Surface irrigation is performed according to topography and product type with different methods, including basin irrigation (BI) and two-sided furrow irrigation (TSFI). Due to the limited facilities for developing irrigated agriculture, increasing water-use efficiency by managing the irrigation and productivity of existing water and soil resources is necessary^{4,5}. Considering the excessive consumption of water resources, especially in the agricultural part of Lake Urmia, it is essential to precisely estimate the water-use efficiency (WUE) and yield by using optimal irrigation methods combined with artificial intelligence methods. WUE is an essential factor for identifying the adaptability of crops in water-limited regions under current climate conditions and future global changes^{6–13}. In addition, yield prediction, particularly strategic products is an interesting research topic for agricultural meteorologists due to the importance of national and international economic planning.

In the recent decade, researchers have evaluated the yield and WUE in orchards according to irrigation management and different surface irrigation methods^{14–18}. Osman¹⁹ stated that weak design and improper irrigation management in surface irrigation are the main reasons for low water-use efficiency. Lampinen *et al.*²⁰ investigated soil and plant data and evapotranspiration for irrigation management of walnut trees in California, USA. Fernandes-Silva²¹ by examining the effect of different irrigation regimes (dryland irrigation with 30% and 100%

¹Agricultural Research, Education and Extension Organization, Agricultural Engineering Research Institute, Post Box 31585-845, Karaj, Alborz, Iran. ²Department of Computer Engineering, University of Bonab, Bonab, Iran. ³Department of Water Engineering, University of Tabriz, Tabriz, Iran. ⁴Department of Water Engineering, Urmia University, Urmia, Iran. ✉email: h.dehghanisani@areeo.ac.ir

water requirement) on yield and WUE of olive, reported that crop evapotranspiration (ET_c) is the most influential parameter in changes in fruit yield. Dahikar and Rode²² proved that artificial neural networks (ANNs) have good efficiency in estimating crop yield. Dzikiti et al.²³ estimated the water requirement of young and productive apple orchards in South Africa using the evapotranspiration model. Their findings confirmed that this model offers reasonable estimates in mature and young gardens. Emami and Chooapan²⁴ estimated barley yield using radial basis function (RBF) and feed-forward neural (GFF) models in Torbat Heydariyeh, located in southeastern Iran. The results showed that the RBF model with the input parameter of irrigation water levels could better estimate the barley yield. Bang et al.²⁵ by comparing fuzzy logic, ARMA, SARIMA, and ARMAX methods, concluded that the fuzzy logic method has an excellent ability to predict crop yield. Kumar et al.²⁶ used the random forest (RF) model to estimate crop yield. The results showed that the RF algorithm has a high capability in estimating crop yield by considering the minimum number of parameters. Sharifi²⁷ proved that the Gaussian process regression algorithm has the best performance in estimating barley yield. Prasad et al.²⁸ used the random forest algorithm for estimating cotton crops for the regional level. The results indicated that the RF model had a high potential in predicting crop yield. Dehghanisani et al.²⁹ reported irrigation-fertilizer and crop variety parameters are the most effective parameters in estimating yield and water productivity of tomato crops.

According to the above-mentioned studies, climatic factors are effective parameters affecting the yield and WUE. Given the complexity of the impact of climatic factors on plant growth, it is necessary to estimate its effects. In this paper, a new irrigation method referred to as narrow strip irrigation (NSI) was introduced to reduce applied water in apple orchards. Then, a hybrid predictive method based on adaptive neuro-fuzzy inference system (ANFIS) and season optimization (SO) algorithm is proposed to estimate WUE and yield. To summarize, the contributions of this paper are as follows:

- Introducing the narrow strip irrigation (NSI) method for the first time and estimate its WUE and yield parameters. The NSI method reduces the growth of weeds and prevents the penetration of water outside the shade of the tree.
- Introducing the hybrid SO-ANFIS method to estimate the WUE and yield parameters of the NSI irrigation method. The SO-ANFIS takes the advantages of both SO algorithm and ANFIS methods.
- Evaluating the SO-ANFIS method on a benchmark dataset and compared it with state-of-the-art WUE and yield estimation methods. The results justify that the proposed SO-ANFIS outperformed its counterparts in terms of performance measures.

The remaining part of this paper is organized as follows. Section “[Materials and methods](#)” describes the test case and the working principle of the proposed approach. In section “[Results and discussion](#)”, the results and discussions are presented. Section “[Conclusion](#)” concludes the paper and presents some suggestions for future work.

Materials and methods

Test case. This study was conducted in the agricultural lands of Dolatabad village located in Miandoab region. Miandoab is a city in the northwest of Iran located in the southeast of Lake Urmia. The geographical coordinates of Miandoab are 46° 2' N and 36° 58' E at 1314 m above sea level (Fig. 1). In this region, the weather is variable, with relatively hot summers and cold winters. Miandoab is a significant agricultural region in West Azerbaijan province. The main crops in Miandoab are wheat, barley, sugar beet, corn, and apple orchards.

Field studies and sampling. In this study, a total of 120 field data from two farms under study (M₁ and M₂ farms in Fig. 1) were collected. This data set was randomly divided into two parts; 80% of the data was used for model training and the remaining 20% for tests. Soil sampling was performed from the end of the tree shading surface and three depths of 0–30 cm, 30–60 cm, and 60–90 cm. Three types of irrigation methods including BI, NSI, and TSFI, were considered. The cultivars studied were *Golden Delicious*. The distances of the trees were 6 × 6 m². The dimensions of the control and treatment strips were 3.6 × 6 m² and 6 × 6 m², respectively. The irrigation interval was considered equal to 10–15 days based on the climate condition to ensure an optimal outcome. The crop was harvested on September 30, 2021.

Water-use efficiency. WUE can be defined as (Eq. 1)³⁰.

$$WUE = \frac{Y \text{ (usually economical yield)}}{\text{applied water } (P_e + I + SW)} \quad (1)$$

where Y denotes the economical yield was measured base on the delivered product to the market, I is irrigation water measured using a WSC flume, P_e is effective rainfall and SW indicates soil water depletion from the root zone during the growing season. The SW is estimated based on the water balance at the selected farm.

Improving the economic water use efficiency at the farm level requires better adaptation and coordination of water use according to the needs of products at the time and amount of its use, which ultimately improves crop yield. This is possible by using new emerging technology and applying better management methods. Applying new management methods in planning for planting, irrigation, and using other inputs plays an effective role in achieving high WUE. Chemical and physical analyses of soil and fertilizers used are presented in Tables 1, 2 and 3.

Application efficiency. Application efficiency (AE) indicates losses in the farm in the form of deep infiltration and runoff at the end of the farm. At each irrigation interval, AE is calculated as follows⁴³:



Figure 1. Study area on the map, the graph is plotted in ESRI ArcGIS Desktop 10.8.2.28388 software [https://www.esri.com/en-us/arcgis/about-arcgis/overview].

B.D (gr/cm ³)	PWP (cm ³ /cm ³)	FC (cm ³ /cm ³)	OC (%)	TNV (%)	θ _s	pH	Texture	EC (ds/m)	Silt	Sand	Clay	Depth (cm)
1.25	0.14	0.328	1.35	10.4	52	8.3	Sic	0.875	52	9	39	0–30
1.25	0.149	0.331	0.82	14.9	57	8.11	Sic	1.45	46	11	43	30–60
1.25	0.144	0.328	-	17.8	55	8.05	Sic	2.78	48	11	41	60–90

Table 1. Physical and chemical analysis of soil at a depth of 0–90 cm. EC: Electrical Conductivity; pH: Acidity of water; θ_s: Saturated moisture (volumetric); TNV: Lime; OC: Organic carbon; FC: Field capacity; PWP: Permanent wilting point; B.D: Soil Bulk Density.

K _R (kg/ha)	P _R (kg/ha)	N _R (kg/ha)	K _A (ppm)	P _A (ppm)	N _T (%)	Date of harvest	Date of planting	Variety	Cultivation pattern	Area (ha)
100	75	50	158	4.8	8	2021/09/30	23 year	Red/golden	Apple	0.8

Table 2. Physical and chemical characteristics of the soil. N_T: Total nitrogen; P_A: Absorbable phosphorus; K_A: Absorbable potassium; N_R: Nitrogen requirement; P_R: Phosphorus requirement; K_R: Potassium requirement.

Fertilizer	Amount	Date
Di-ammonium phosphate	60 kg	2020/03/26
Ammonium sulfate	50 kg	2020/03/26
Potassium sulfate	100 kg	2020/03/26
Zinc sulfate	30 kg	2020/03/26
Iron sulfate	50 kg	2020/03/26
Agricultural sulfur	100 kg	2020/03/26
Thiobacillus bacteria	5 kg	2020/03/26
Manure	5 ton	2020/03/26

Table 3. Amounts of fertilizer used in an apple orchard.

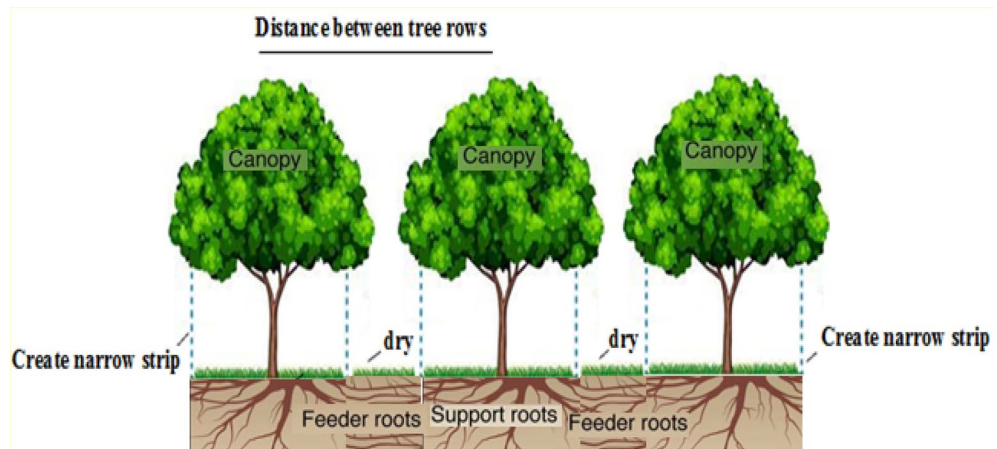


Figure 2. Schematic of NSI method.

$$AE = \frac{D_z}{D_{app}} \times 100 \quad (2)$$

where D_z shows the average water storage in the root zone depth (mm) and D_{app} is the average depth of water entering the irrigated area. D_{app} is defined as

$$D_{app} = \frac{V}{A} \quad (3)$$

where V is the volume of inlet flow to the irrigated area (lit) and A indicates the irrigated area (m^2).

Irrigation methods. *Narrow strip irrigation.* In the NSI method, the entire orchard surface is not irrigated, and evaporation losses are minimized. Therefore, the daily water requirement of the tree is mainly limited to the amount of transpiration from the aerial parts exposed to sunlight. Plant shading level is one of the critical factors in calculating the water requirement of trees. This parameter is determined experimentally in terms of the type and age of the plant (between 50 and 70%)³¹. Overall, in the NSI, the main area where transpiration occurs is the shading level. In the NSI method (Fig. 2), a space is created in the middle of the trees row in orchards. This area is dry during irrigation, and applied water is reduced due to the lack of weed growth and water evaporation from this area. In the NSI, the following equation is used to calculate the daily water requirement of the plant³¹:

$$R_r = ET_c = (h_s + 0.15(1 - h_s)) \quad (4)$$

where R_r is the maximum daily water requirement (mm/day), ET_c is the maximum daily evapotranspiration (mm/day), and h_s is the maximum shading level (%). According to the studies, the shading level for trees is 50–70% in the optimal state³¹. In this study, the shading level of apple trees was determined based on age and crown environment.

Two-sided furrow irrigation. In the TSFI method, water moves inside the furrow on both sides of the trees and deep in the soil irrigates the root development area vertically and laterally (Fig. 3). This method tries to wet the soil surface less. The water is directed by two furrows created on either side of the rows of trees. The distance of the furrows from the rows of trees varies depending on the distance between the rows of trees, soil texture, and age of the trees. By performing furrow irrigation, two dry parts are created in the orchard. These arid areas form one along the rows of trees and the other between the furrows in the middle of the rows of trees. Try to prevent weeds from growing in dry areas as much as possible with tools such as garden tractors, cultivators, or retractors. In fact, the existence of these arid areas and the lack of weed growth and water evaporation from these arid areas, which play a role in the real reduction of water consumption. If these two arid areas are full of weeds, water consumption will not be really saved and only the irrigation efficiency will increase due to the movement of water in the furrows³².

Basin irrigation. BI is a method in which water penetrates the soil permanently or intermittently, and the soil is permanently submerged (Fig. 4). In basin irrigation, water penetrates the crown area of the plant, and the problem of clogging heavy soils and reducing soil aeration occurs³².

In general, in the NSI method, compared to the TSFI and BI methods, the water in the shade of the tree travels in a straight path at the same width and travels to the next tree. In this method, water penetration is prevented outside the shade of the tree, and conditions for weed growth will not be provided.

Season's optimization (SO) algorithm. The SO algorithm is a population-based optimization meta-heuristic³³. It models the growing process of trees in four seasons of a year. Figure 5 illustrates the flowchart of

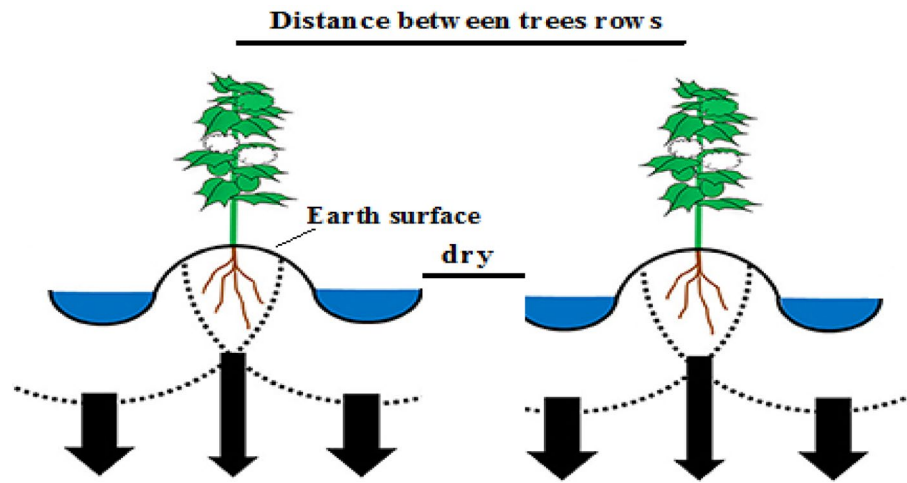


Figure 3. Schematic of TSFI method.

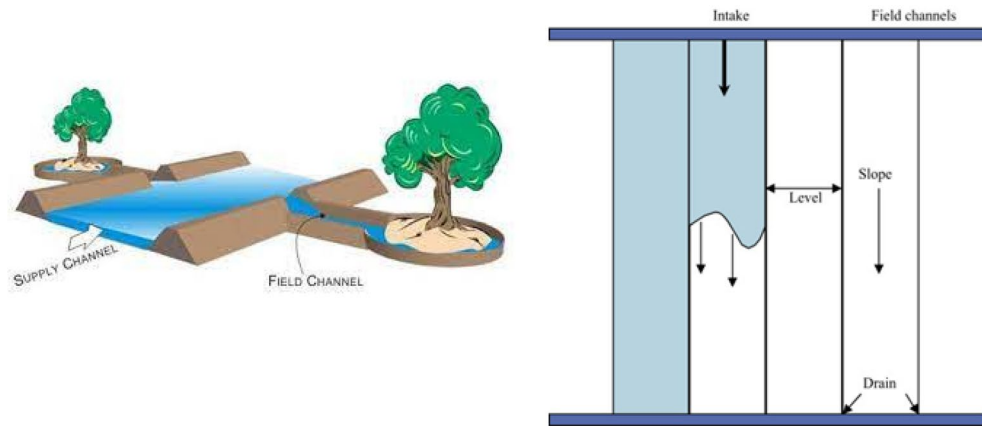


Figure 4. Schematic of BI method.

the SO algorithm. The SO is an iterative algorithm in which each agent is called a tree. For solving an optimization problem, the algorithm starts its process with a population referred to as a forest. Each member of the population is called a tree which denotes a potential solution for the given problem. For an optimization problem $f(X) = f(x_1, x_2, \dots, x_D)$ with D dimensions, the initial forest F is initialized as follows³³:

$$\begin{aligned}
 F &= [T_1, T_2, \dots, T_N] \\
 T_i &= [t_{i1}, t_{i2}, \dots, t_{iD}] \\
 t_{ij} &= l_{ij} + r_{ij} \cdot (u_{ij} - l_{ij})
 \end{aligned}
 \tag{5}$$

where r_{ij} is a random number in the interval $[0, 1]$ generated by the uniform distribution u_{ij} and l_{ij} are the upper and lower bounds of t_{ij} , respectively. The fitness of each tree is evaluated by a strength function.

$$S_i = S(T_i) = S(t_{i1}, t_{i2}, \dots, t_{iD})
 \tag{6}$$

The algorithm updates the trees using four operators, including renew, competition, seeding, and resistance. The renew phase models the impact of the spring on the growth of trees. The following equations are defined to model the renew phase mathematically:

$$\begin{aligned}
 F^{y+1} &= \{F^y\} + \{R\} \\
 R &= \Phi(p_r \times A^y)
 \end{aligned}
 \tag{7}$$

where R indicates the set of new seedlings, F^y shows the forest at the y th iteration, A^y is the number of seeds generated in the previous autumn and p_r is the renew rate. The function Φ randomly produces some seedlings in various locations of the forest. The algorithm does not execute the renew phase in the generation $y = 0$.

The competition phase models the growth of trees in the summer. In this phase, the trees compete with their neighbor trees on shared resources, including nutrients, water, light, and other resources. To simulate the

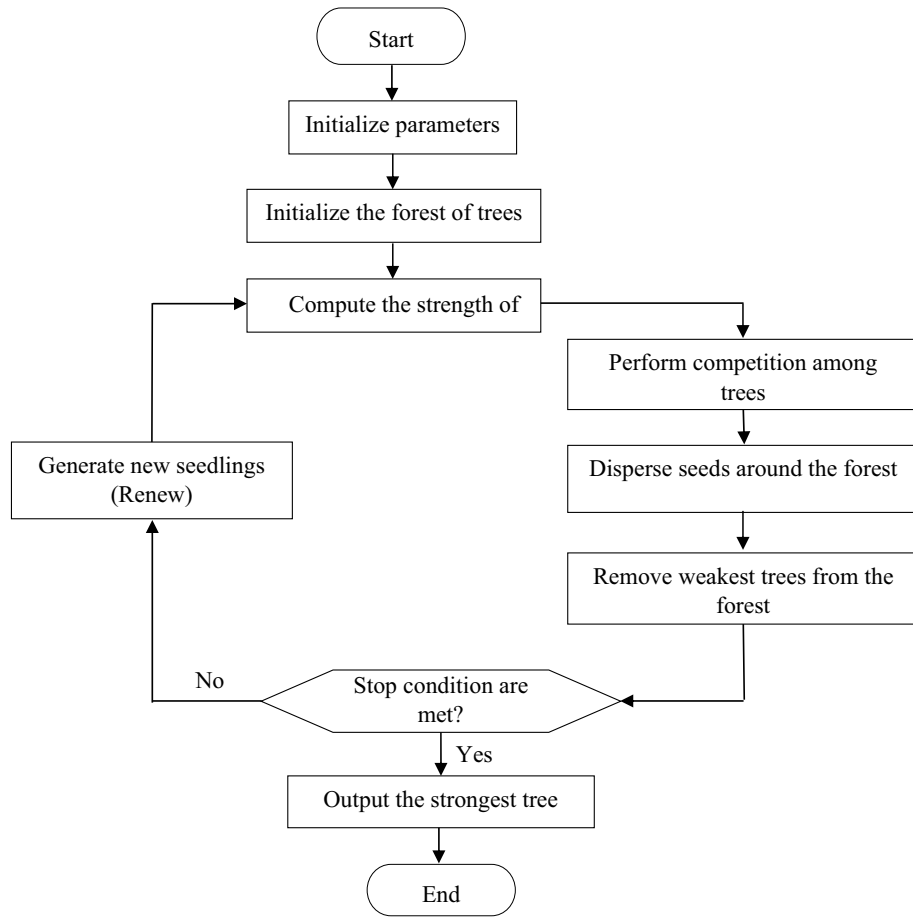


Figure 5. Flowchart of the SO algorithm.

competition process, first N_c most vital trees are identified. The number of neighbors of a cored tree is calculated as follows³³:

$$\begin{aligned}
 N_c &= \lceil p_c \times N \rceil \\
 Z_i &= \lceil \tau_i \times N_g \rceil \\
 N_g &= (N - N_c)
 \end{aligned}
 \tag{8}$$

τ_i is the normalized fitness of T_i , which is calculated as follows:

$$\tau_i = (S_i - \min(I)) / \left(\sum_{k=1}^N S_k \right), \quad I = \{S_k | k = 1, 2, \dots, N\}
 \tag{9}$$

Then, Z_i neighbors are elected to create the neighborhood zone. To simulate the impact of the competition on a neighbor T_i , the below relationship is defined:

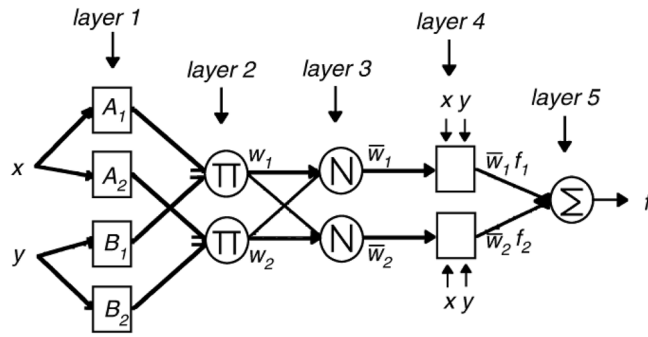


Figure 6. A big picture of the ANFIS system with two inputs.

$$T_j^{y+1} = \frac{1}{\Lambda_j + 1} \times \varphi(T_j^y)$$

where

$$\varphi(T_j^y) = T_j^y + \theta$$

$$\Lambda_j = \sum_{k=1}^{Z_i} S_k \times \Delta_{j,k}^{-2} \times \lambda_{j,k} \tag{10}$$

$$\Delta_{j,k} = \sqrt{\sum_{z=1}^D (T_{jz} - T_{kz})^2}$$

$$\lambda_{j,k} = \begin{cases} 1 & \text{if } (S_k \geq S_j) \\ 1 - \gamma & \text{else} \end{cases}$$

where T_j^y is the location of T_j in the generation y . Λ_j is the value of competition index or crowdedness, which computes the effect of the neighbors on T_j . D shows the number of variables of trees. The function $\varphi(\cdot)$ calculates the growth of T_j in the same environment when its neighbors are ignored. S_k indicates the strength/ fitness of the k th neighbor tree, $\Delta_{j,k}$ is the distance between T_j and the k th neighbor, the variable $\lambda_{j,k}$ is the effect of the neighbor on the growth of the tree T_j . The parameter $\gamma \in [0, 1]$ is a random asymmetry index, which shows the value to which the impact of relatively weak neighbor is decreased³³.

The new location of the cored tree T_i is calculated as

$$T_i^{y+1} = \begin{cases} T^* & \text{if } S(T_i^y) \leq S(T^*) \\ T_i^y & \text{if } S(T_i^y) > S(T^*) \end{cases} \tag{11}$$

where T^* shows the strongest neighbor tree around T_i .

The seeding phase is inspired by the seeding mechanism of trees in the autumn. In this phase, several trees are randomly selected and participate in the seeding phase. The number of seeds (A) at each generation is calculated as

$$A = \psi(p_s \times N) \tag{12}$$

where p_s indicates the seeding rate, which is a uniform random number. The ψ function identifies the fittest trees from the population. From each tree T_i selected in the seeding phase, several elements are randomly identified, and their current values are updated with new random deals in the boundary of search space. Let m be a random number, and $\{t_{i1}, t_{i2}, \dots, t_{im}\}$ are the elements selected from T_i , where $m < D$. Each component $t_{ij} \in T_i$ is calculated as

$$t'_j = t_j + \ell \times r \tag{13}$$

ℓ is a two-valued variable, either 1 or -1 , and $r \in [l_j, u_j]$ is a random number.

The resistance phase simulates the resistance of the trees against harsh winter cold. The resistance operator removes the least-strength trees from the population. This operator is mathematically modeled as follows:

$$\begin{aligned} F^{y+1} &= \{F^y\} - \{W\} \\ W &= \chi(p_w \times N) \\ p_w &= 1 - (1 - p_s) \end{aligned} \tag{14}$$

where W is the collection of weak trees. $\chi(\cdot)$, removes $p_w \times N$ trees from the forest, p_s is the resistance rate.

When the stopping measures are met, the algorithm updates the trees in the population by iteratively applying to renew, competition, seeding, and resistance operators. Finally, the fittest tree is identified as the optimal solution³³.

Adaptive neuro-fuzzy inference system (ANFIS). The ANFIS integrates the artificial neural networks (ANNs) and fuzzy inference system (FIS)³⁴. The ANFIS combines the advantages of both FIS and ANNs. The ANFIS system has high adaptation and fast learning capacity, captures the non-linear structure of processes, and causes less memorization. These characteristics make the ANFIS the best choice for predictive problems such as WUE and yield estimation problems. The ANFIS has been used successfully in various fields including mechanical design problems, chemical processes, data mining applications, communications, economics, geotechnical engineering problems, scheduling problems, and many other engineering problems.

In ANFIS, the relationship between inputs and outputs and the best values for the parameters related to the membership functions are identified by the fuzzy section and ANNs, respectively. The structure of ANFIS is determined considering the input data, rules, functions of the output membership function, and the membership degree. The ANFIS system with five layers is shown in Fig. 6. In the first layer, the level of dependence of each input data on different fuzzy domains is determined. The weight of the rules is obtained by multiplying the input values of each node in the second layer. The computation of the importance of regulations is carried out in the third layer. The rules layer is created by performing operations on the input signals described by the fourth layer. The network output is indicated by the fifth layer. ANFIS has n rules and m input components. Each rule R_j is represented as follows:

$$R_j: \text{if } (x_1 \text{ is } q_{j1}) \text{ and } (x_2 \text{ is } q_{j2}) \text{ and } \dots (x_m \text{ is } q_{jm}) \text{ then output} = f_i \quad (15)$$

where x_j indicates the j th input q_{ij} indicates the membership function of the rule on x_j , f_i is the output of rule, The output of the network is presented as follows:

$$f(x) = \frac{\sum_{i=1}^n \mu_i f_i}{\sum_{i=1}^n \mu_i} \quad (16)$$

where μ_i indicates the activation degree of the rule. Each node has a function with adjustable parameters. μ_i is defined as follows:

$$\mu_i = \prod_{j=1}^m q_{ij}(x_j) \quad (17)$$

In the current implementation of ANFIS, we used the Gaussian membership functions, which is defined as follows:

$$q_{ij}(x) = \exp \left[-\frac{1}{2} \left[\frac{x - c_{ij}}{\sigma_{ij}} \right]^2 \right] \quad (18)$$

where c_{ij} and σ_{ij} are the center and standard deviation of the Gaussian membership function, respectively. Gaussian membership function is a popular method for specifying fuzzy sets because of its smoothness and concise notation. Five factors should be determined in designing ANFIS, the number and type of input and output fuzzy sets, the number of iterations, and the optimization method. The SO algorithm was used to optimize the parameters of the ANFIS membership function. In this paper, the fuzzy c -means clustering (FCM) is used to create fuzzy inference system which obtained superior results in the literature.

SO-ANFIS model. Two structural parameters of the ANFIS system are antecedent and consequent parameters. For tuning these parameters, researchers often used gradient-based techniques. The main drawback of the gradient-based methods is the low convergence rate and trapping in local optima. Meta-heuristic algorithms can be used as efficient alternatives to overcome the limitations of gradient-based methods in training the ANFIS model. To train the ANFIS system using the SO algorithm, two issues need to be determined: strength function and the boundary of variables. In this study, root means square error (RMSE) is used as a strength function for evaluating the performance of the ANFIS system. Assume the following relationship:

$$\begin{aligned} R_i: \text{if } I_r \text{ is } q_1(\sigma_{1i}, c_{1i}) \text{ and } T_{emp} \text{ is } q_2(\sigma_{2j}, c_{2j}) \text{ and } P_e \text{ is } q_3(\sigma_{3l}, c_{3l}) \text{ and} \\ RH_{avg} \text{ is } q_4(\sigma_{4k}, c_{4k}) \text{ and } S_{sh} \text{ is } q_5(\sigma_{5t}, c_{5t}) \text{ then} \\ y_i = s_1 I_r + s_2 T_{emp} + s_3 P_e + s_4 RH_{avg} + s_5 S_{sh} + s_6 \end{aligned} \quad (19)$$

The input variables are water consumption during the growing season (I_r), temperature (T_{emp}), average relative humidity (RH_{avg}), the amount of solar radiation in terms of sunshine hours (S_{sh}), and the rainfall (P_e) of each month of the growing season. The model parameters that need to be configured are σ , c , s_1 , s_2 , s_3 , s_4 , s_5 , s_6 . The variables s_1 , s_2 , s_3 , s_4 , s_5 , s_6 are consequent parameters, which should be measured during the ANFIS training process. The optimal values for parameters c and σ is measured by the SO algorithm. To identify the value of parameters c and σ , first, a forest composed of several trees is initiated. Each tree contains candidate values for

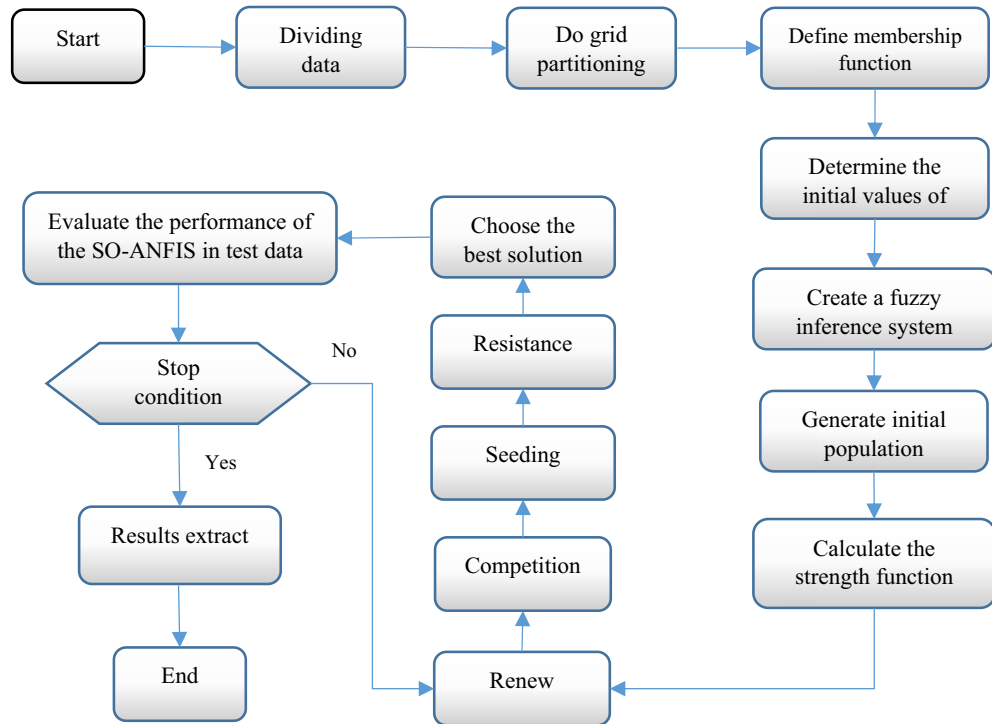


Figure 7. Flowchart of the SO-ANFIS.

Irrigation	Average inlet flow (L/s)			Cut-off time (min)			Irrigation depth (I_p) (mm)			Net irrigation requirement (I_n) (mm)			Application efficiency (AE) %		
	BI	NSI	TSFI	BI	NSI	TSFI	BI	NSI	TSFI	BI	NSI	TSFI	BI	NSI	TSFI
1	8.71	8.61	7.25	150	93	110	95	74.3	129	29.9	57.5	78.9	31.2	77.4	66.7
2	10.30	8.62	9.26	200	82	94	149.8	86.7	123.1	27.8	64.4	79.5	26.2	74.3	61.6
3	10.63	8.65	9.23	220	80	90	170.2	91.2	108.4	33.3	65	73	28.1	69.3	51.2
4	10.53	8.15	9.15	250	68	80	205.7	93.3	100.6	44.8	71.6	55.5	21.8	64.8	45

Table 4. Mean values of irrigation characteristics in research treatments.

the ANFIS parameters. The trees are updated iteratively using four operators (renew, competition, seeding, and resistance). This process iterates for a pre-determined number of generations (Fig. 7).

Performance Criteria. In the present study, R^2 , RMSE, SI, δ , and NSE indices were applied to appraise the ability of the introduced hybrid method³⁵:

$$R^2 = \left[\frac{\sum_{i=1}^n (w_i - \bar{w})(z_i - \bar{z})}{\sum_{i=1}^n \sqrt{(w_i - \bar{w})^2} \sum_{i=1}^n \sqrt{(z_i - \bar{z})^2}} \right]^2 \tag{20}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (w_i - z_i)^2} \tag{21}$$

$$SI = \frac{RMSE}{\bar{w}} \tag{22}$$

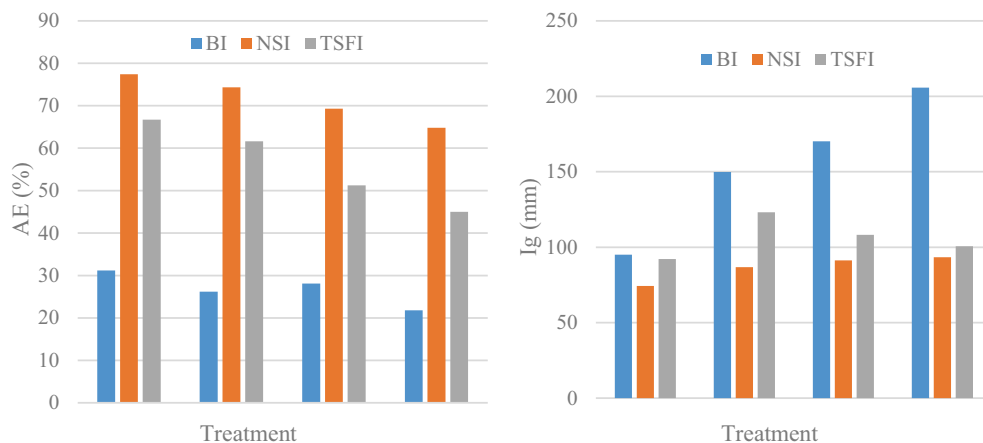


Figure 8. Comparison of application efficiency (AE) and Irrigation depth (Ig) in different treatments BI, NSI, and TSFI.

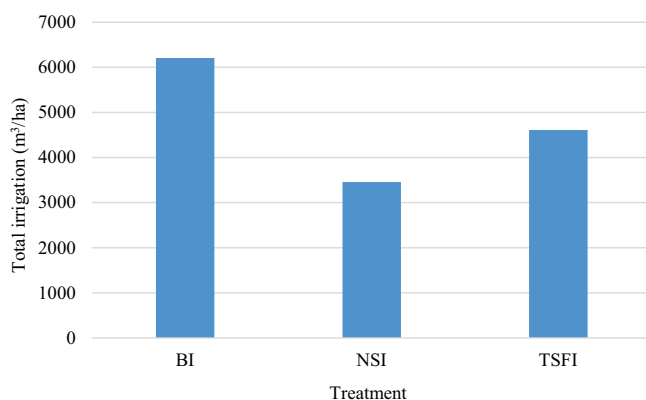


Figure 9. Comparison of total applied water in different treatments.

Reduction of applied water (%)			WUE (kg/m³)			Yield (kg/ha)		
BI	NSI	TSFI	BI	NSI	TSFI	BI	NSI	TSFI
-	42.8	22.7	4.01	7.14	5.24	30,000	35,000	32,000

Table 5. Yield and WUE in research treatments.

$$\delta\% = \frac{\sum_{i=1}^n |(z_i - w_i)|}{\sum_{i=1}^n w_i} \times 100 \tag{23}$$

$$NSE = 1 - \frac{\sum_{i=1}^n (w_i - z_i)^2}{\sum_{i=1}^n (w_i - \bar{z})^2} \tag{24}$$

w_i and z_i are the observed and predicted values of yield and WUE, respectively. \bar{w} and \bar{z} are average observed and predicted values of yield and WUE, respectively.

Variable	Unit	Min	Max	Avg	SD*	CV
Yield	kg/ha	25,000	35,000	30,000	5.89	1767.20
EC ₀₋₃₀	dS m ⁻¹	0.850	2.80	1.32	2.75	0.03
pH ₀₋₃₀	-	8.00	8.50	8.10	0.53	0.042
K ₀₋₃₀	ppm	120.00	158.00	132.00	3.75	4.80
P ₀₋₃₀	ppm	4.00	4.80	4.30	2.82	0.11
Clay ₀₋₃₀	-	30.00	39.00	33.00	3.22	1.10
Silt ₀₋₃₀	-	46.00	52.00	48.00	1.72	0.78
Sand ₀₋₃₀	-	7.50	9.00	8.10	3.15	0.25
EC ₃₀₋₆₀	dS m ⁻¹	1.35	1.45	1.38	0.82	0.011
pH ₃₀₋₆₀	-	7.80	8.11	8.00	0.92	0.075
K ₃₀₋₆₀	ppm	78.30	168.00	137.20	15.83	21.30
P ₃₀₋₆₀	ppm	3.30	4.10	3.70	1.72	0.058
Clay ₃₀₋₆₀	-	39.00	43.00	40.40	1.36	0.52
Silt ₃₀₋₆₀	-	40.00	46.00	42.00	1.86	0.74
Sand ₃₀₋₆₀	-	9.00	11.00	10.10	4.81	0.45
EC ₆₀₋₉₀	dS m ⁻¹	2.10	2.78	2.35	4.42	0.098
pH ₆₀₋₉₀	-	7.80	8.05	7.90	0.51	0.038
K ₆₀₋₉₀	ppm	123.00	179.30	131.40	2.40	3.02
P ₆₀₋₉₀	ppm	2.80	3.92	3.36	6.05	0.202
Clay ₆₀₋₉₀	-	38.00	41.00	39.00	0.96	0.37
Silt ₆₀₋₉₀	-	42.00	48.00	44.00	2.15	0.77
Sand ₆₀₋₉₀	-	9.00	11.00	10.10	4.81	0.45
EC _{water}	dS m ⁻¹	0.98	1.12	1.00	0.82	0.008
Q**	L/s	8.09	8.78	8.17	0.40	0.03
I [']	day	10.00	25.00	15.00	0.13	1.81
Area	ha	0.80	0.80	0.80	-	-
I ^{''}	m ³ /ha	3950.00	4945.20	4202.20	2.72	89.35
EL. [^]	m	1300.00	1314.00	1306.00	0.19	2.42
Tree age	year	15.00	25.00	20.00	9.00	1.80

Table 6. The results of descriptive statistics. *Standard deviation, **Water flow, 'Irrigation interval, ''Amount of water consumed, ^Elevation.

Model	Inputs parameters
ω1	I _r , P _e , RH _{avg} , T _{emp} , S _{sh}
ω2	I _r , P _e , RH _{avg} , S _{sh}
ω3	I _r , P _e , T _{emp} , S _{sh}
ω4	I _r , P _e , RH _{avg} , T _{emp}
ω5	I _r , T _{emp} , RH _{avg} , S _{sh}
ω6	T _{emp} , P _e , S _{sh}

Table 7. Effective input combination in estimating WUE and yield.

Results and discussion

Field monitoring. Table 4 and Fig. 8 present the results of irrigation depth, inlet flow, net irrigation requirement, and AE in research treatments. The first irrigation has the lowest WUE due to the dryer soil surface and impacts of tillage operations. Deep penetration losses are primarily due to the excellent permeability of the soil. The application efficiency increased to 21.4% due to the NSI method and irrigation time management compared to the two-sided irrigation method. Irrigation depth increased in treatment BI from the first to fourth irrigation event because of the loss of large amounts of water as deep infiltration and the scarcity of soil moisture increases at the point of root access. By increasing the irrigation depth, application efficiency decreased accordingly.

Based on the results, the average increase in application efficiency in NSI compared to BI is about 62.40%. The amount of applied water in the NSI method was 3455 m³/ha, which indicates a reduction of 42.80% and 22.70% of applied water in NSI treatment compared to BI and TSFI treatments (Fig. 9). The decreases were mainly attributed to the less soil surface wetted area, which was minimum in NSI and NSI < TSFI < BI.

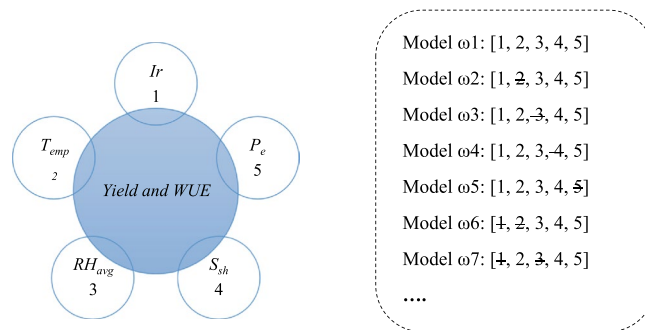


Figure 10. Some combinations of input parameters to estimate yield and WUE parameters.

Model	Train					Test					Membership functions	
	R ²	RMSE	SI	δ	NSE	R ²	RMSE	SI	δ	NSE	c	σ
ω1	0.980	0.008	0.010	0.878	0.962	0.960	0.009	0.012	1.160	0.915	0.4	0.2
ω2	0.992	0.004	0.006	0.836	0.987	0.988	0.006	0.007	0.860	0.982	0.5	0.1
ω3	0.975	0.010	0.014	1.125	0.918	0.972	0.012	0.018	1.415	0.906	0.7	0.2
ω4	0.890	0.012	0.016	1.142	0.846	0.851	0.020	0.033	1.650	0.705	0.5	0.1
ω5	0.905	0.013	0.018	1.325	0.823	0.890	0.017	0.025	1.480	0.740	0.3	0.2
ω6	0.880	0.015	0.026	1.589	0.725	0.868	0.019	0.031	1.620	0.710	0.1	0.1

Table 8. The efficiency of the proposed model in yield estimation. Bold numbers indicate better results compared with others.

The estimated yield in BI, NSI, and TSFI treatments was 30,000 kg/ha, 35,000 kg/ha, and 32,000 kg/ha, respectively. Higher yield in NSI and TSFI attributed to the soil moisture condition. The yield of NSI and TSFI treatments compared to the BI treatment was 14.30% and 6.25%, respectively, and in comparison with each other, increased by 8.57% (treatment NSI compared to treatment TSFI) (Table 5). The estimated WUE in NSI and TSFI treatments was 7.14 kg/m³ and 5.24 kg/m³, respectively (Table 5).

Effect of soil properties. The results of descriptive statistics are shown in Table 6. Based on³⁶ classification, the variation coefficient (CV), less than 15% shows low changes, between 15 and 35% moderate changes, and more than 35% great changes. According to this classification, soil sand content, tree age, and irrigation interval have moderate changes, soil acidity has low changes, and other variables have great changes (due to management factors). Absorbable phosphorus concentration at a depth of 30–60 cm in the soil is the most influential parameter for crop yield. The results of the present study are consistent with^{37, 38} studies. Elimination of tree irrigation at different growth periods reduces the quality and quantity of crop yield. Sedaghati et al.^{28, 37} concluded that increasing the irrigation interval from 25 to 45 days increased the percentage of porosity. According to studies, water flow and irrigation interval have a positive effect on crop yield³⁸. Therefore, reducing the irrigation interval with methods such narrow strip³⁸ can be considered as one of the management methods. Increasing the percentage of sand reduces the soil's ability to retain water and nutrients used by the plant. The age of the trees in this area is high, and with the increase of the tree's age, its ability to grow and produce its production gradually decreases.

Modeling results. *Investigating the effect of input combinations.* Yield and WUE of apple trees depend on various factors, including water consumption during the growing season (I_r), climatic factors including temperature (T_{emp}), average relative humidity (RH_{avg}), the amount of solar radiation in terms of sunshine hours (S_{sh}), and the rainfall (P_e) of each month of the growing season [<http://tatweather.areeo.ac.ir/?LRef=52c6c899-7597-412d-83d7-c4cd2d05204b>] (Table 7)^{39, 40}.

To examine the most appropriate input parameters, different input combinations of parameters were evaluated. To select the most effective input parameters, first, all input combinations are considered to train the ANFIS model, and then the effective input combinations are selected. Next, ignore the remaining parameters one by one from the input combinations and train the model with the same structure and the rest ignored. This approach is also used by other researchers as given in references^{39, 40}. Figure 10 shows 7 of the best-performing models.

The results obtained by the proposed SO-ANFIS method using different input combinations are shown in Table 8. Psize and fitness function evaluations (FEs) in the SO-ANFIS are considered 50 and 3000, respectively. According to the results in the observed data, the model ω2 obtained the most accurate results. The irrigation parameter (I_r) was proposed as the influential input variable in estimating yield and WUE. Then, rainfall and

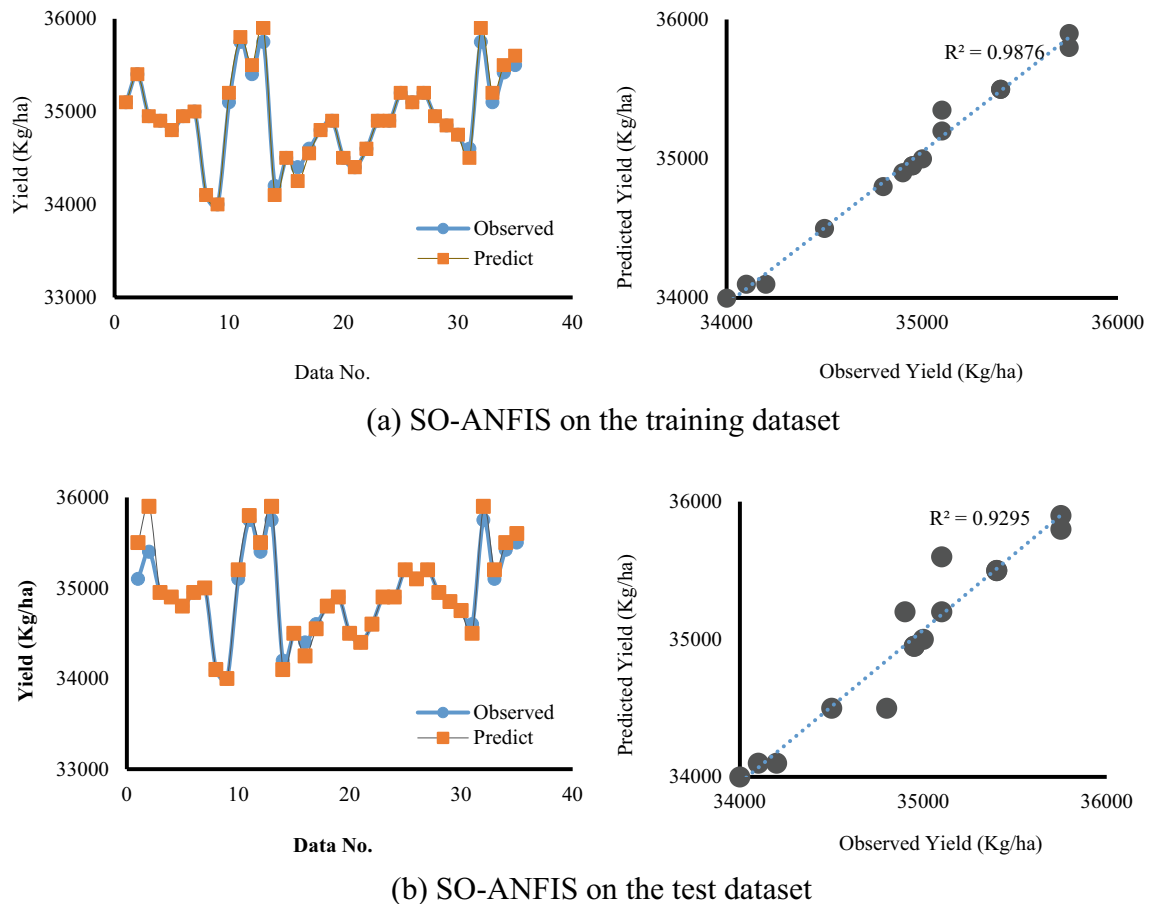


Figure 11. Comparison of predicted yield with observed results in the training and test stages.

sunshine hours are essential, respectively. Sensitivity analysis showed that after irrigation and rainfall parameters, which affect leaves and plant reproductive growth, sunshine hours also play an important role in estimating yield. Montazer et al.³⁸, Zeinadini et al.⁴⁰, and Emami and Choopan²⁴. Also stated that the amount of water consumed has an influential effect on crop yield.

According to Figs. 11 and 12, it is clear that the yield using the SO-ANFIS hybrid method is estimated with high accuracy and is in good agreement with the observed values. Also, ω_2 modeled the yield with lower error ($RMSE = 0.006$) according to the irrigation parameter.

The SO-ANFIS and ANFIS error distribution diagrams on the test stage are shown in Fig. 13. The results show that about 80% of the yield values estimated utilizing the SO-ANFIS have an error of less than 2%.

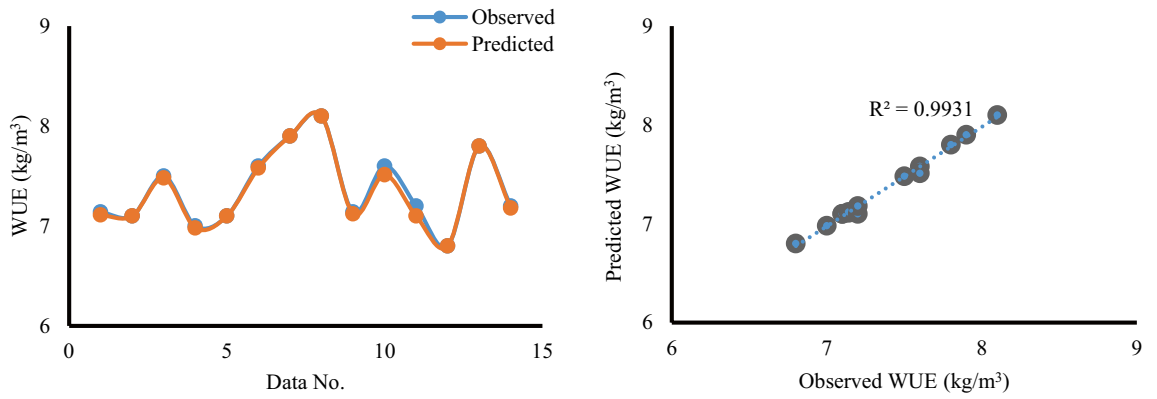
Comparison of SO-ANFIS with other methods. Table 9 compares the results generated by the proposed SO-ANFIS and other counterparts. The results confirm that the proposed SO-ANFIS outperformed its counterparts in estimating yield and WUE. Comparison of the results of the present study with other works shows acceptable accuracy ($R^2 = 0.988$ in test stage). Compared to similar studies such as Sharif²⁶ and Prasad et al.²⁷, which have evaluated the crop yield and WUE using the random forest (RF) and Gaussian process regression (GPR), the SO-ANFIS with $R^2 = 0.988$ and $RMSE = 0.006$ has a better performance than the mentioned methods and can be used as a powerful method in estimating the yield and WUE.

Conclusion

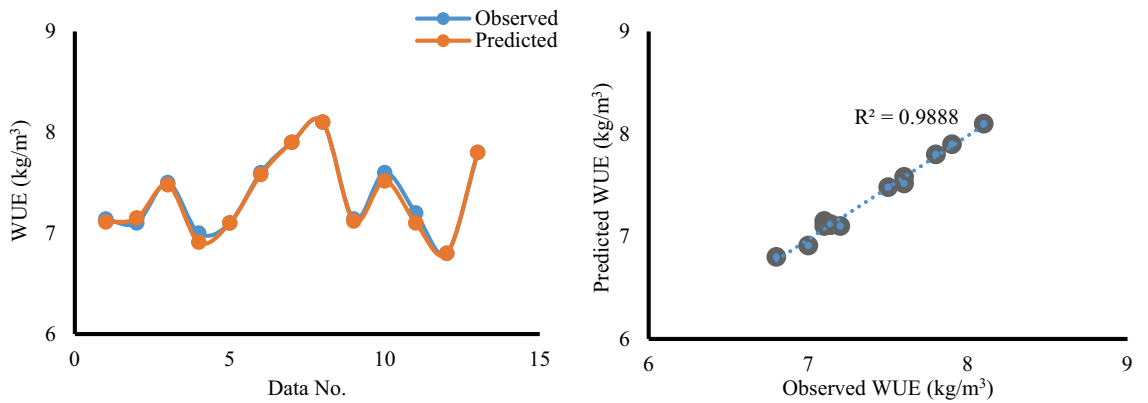
In this study, the effect of NSI method on yield and WUE in apple orchards was investigated. The SO-ANFIS method was proposed to estimate WUE and yield in the NSI model. In the SO-ANFIS, six models were created to determine the most effective parameters in estimating WUE and yield of NSI method. The SO-ANFIS with model ω_6 generated the superior results with $R^2 = 0.988$, $RMSE = 0.006$, $SI = 0.007$, $\delta = 0.860$, and $NSE = 0.982$, respectively. One of the future works is to apply the SO-ANFIS method to other engineering problems to identify its strengths and weaknesses.

Guidelines statement

All measurements and laboratory tests performed in this study are following scientific and international standards, such as soil texture determination⁴¹, volumetric soil moisture monitoring⁴², and water quality analysis (EPA).



(a) SO-ANFIS on the training dataset



(b) ANFIS on the test dataset

Figure 12. (a) SO-ANFIS on the training dataset, (b) SO- ANFIS on the test dataset. Comparison of predicted WUE with observed results in the training and test stages.

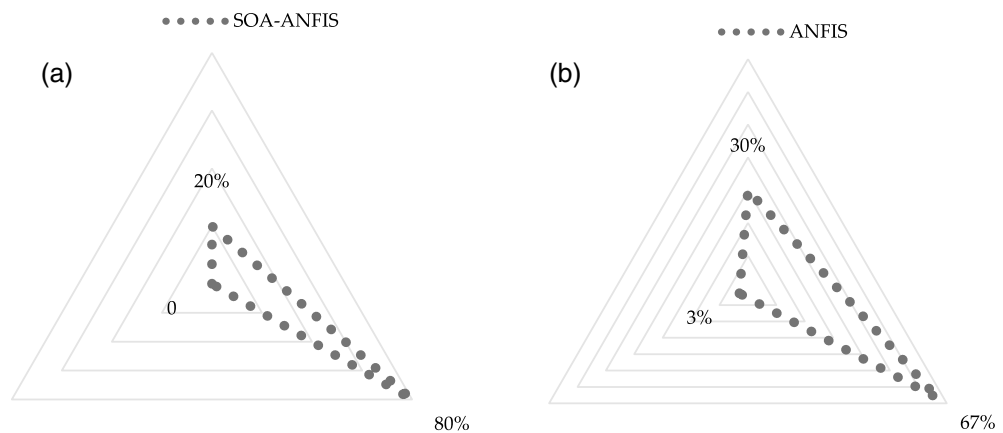


Figure 13. Error distribution of (a) SO-ANFIS and (b) ANFIS on the test stage.

Model	R ²	RMSE	SI	δ	NSE
GPR	0.840	0.055	–	–	0.835
RF	0.690	0.045	–	–	0.687
ANFIS	0.965	0.009	0.802	0.898	0.958
SO-ANFIS	0.988	0.006	0.007	0.860	0.982

Table 9. Comparison of the efficiency of SO-ANFIS model with similar methods.

Data availability

The data that support the findings of this study are openly available.

Received: 28 December 2021; Accepted: 4 April 2022

Published online: 25 April 2022

References

- Zahedpour, H., Rezaverdinejad, R. & Dehghanisanij, H. Evaluation of water application efficiency and productivity of surface irrigation systems in fields of nazloo-chay region, urmia. *J. Water Res. Agric.* **31**(4), 685–698 (2018).
- Mirzaie-Nodoushan, F., Morid, S. & Dehghanisanij, H. Reducing water footprints through healthy and reasonable changes in diet and imported products. *Sustain. Prod. Consumpt.* **23**, 30–41 (2020).
- Hamdi Ahmadabad, Y., Liaghat, A., Sohrabi, T., Rasoulzadeh, A. & Ebrahimian, H. Improving the performance of furrow irrigation systems using simulation modeling in the Moghan plain Iran. *Irrig. Drain.* **70**(1), 131–149 (2021).
- Gholami, Z., Ebrahimian, H. & Noory, H. Investigation of irrigation water productivity in sprinkler and surface irrigation systems (case study: Qazvin plain). *J. Irrig. Sci. Eng.* **39**(3), 135–146 (2012).
- Dehghanisanij, H., Naseri, A., Anyoji, H. & Eneji, A. E. Effects of deficit irrigation and fertilizer use on vegetative growth of drip-irrigated cherry trees. *J. Plant Nutr.* **30**(3), 411–425 (2007).
- Mubiru, J. Predicting total solar irradiation values using artificial neural networks. *Renew. Energy* **33**(10), 2329–2332 (2008).
- Norouzi, M., Ayoubi, S., Jalalian, A., Khademi, H. & Dehghani, A. A. Predicting rainfed wheat quality and quantity by artificial neural network using terrain and soil characteristics. *Acta Agric. Scand. Sect. B Soil Plant Sci.* **60**, 341–352 (2010).
- Onder, S., Caliskan, M. E., Onder, D. & Caliskan, S. Different irrigation methods and water stress effects on potato yield and yield components. *Agri. Water Manag.* **73**(1), 73–86 (2005).
- Palanivel, K. & Surianarayanan, C. An approach for prediction of crop yield using machine learning and big data techniques. *Int. J. Comput. Eng. Technol.* **10**(3), 110–118 (2019).
- Piri, J. *et al.* Daily pan evaporation modeling in a hot and dry climate. *J. Hydrol. Eng.* **14**, 803–811 (2009).
- Qin, S., Zhang, J., Dai, H., Wang, D. & Li, D. Effect of ridge-furrow and plastic-mulching planting patterns on yield formation and water movement of potato in a semi-arid area. *Agric. Water Manag.* **131**, 87–94 (2014).
- Ravichandran, G. & Koteeshwari, R.S. Agricultural crop predictor and advisor using ANN for smartphones. *1st Int. Conf. Emerg. Trends Eng. Technol. Sci. ICETETS.* 2–7 (2016).
- Rodrigues, M. A. *et al.* Nitrogen use efficiency and crop yield in four successive crops following application of biochar and zeolites. *J. Soil Sci. Plant Nutr.* **21**(2), 1053–1065 (2021).
- Zhang, B., Valentine, I. & Kemp, P. Modelling the productivity of naturalised pasture in the North Island, New Zealand: A decision tree approach. *Ecol. Modell.* **186**(3), 299–311 (2005).
- Alvarez, R. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. *Eur. J. Agron.* **30**(2), 70–77 (2009).
- Gonzalez-Sanchez, A., Frausto-Solis, J. & Ojeda-Bustamante, W. Attribute selection impact on linear and nonlinear regression models for crop yield prediction. *Sci. World J.* 2014.
- Kaul, M., Hill, R. L. & Walthall, C. Artificial neural networks for corn and soybean yield prediction. *Agric. Syst.* **85**(1), 1–18 (2005).
- Matsumura, K., Gaitan, C. F., Sugimoto, K., Cannon, A. J. & Hsieh, W. W. Maize yield forecasting by linear regression and artificial neural networks in Jilin, China. *J. Agric. Sci.* **153**(3), 399–410 (2015).
- Osman, H.E., Mehawed, H.S. & Abdel-Aziz, A.A. using modified surface irrigation in old mango farm in Nile valley, in watershed management to meet water quality standards and emerging tmdl (total maximum daily load). *Proceedings of the Third Conference 5–9 March (Atlanta, Georgia USA) Publication.* (2005).
- Lampinen B., Buchner, R., Fulton, A., Grant, J., Mills, N., Prichard, T., Schwankl, L., Shackel, K., Gilles, C., Little, C. & Metcalf, S. Irrigation management in walnut using evapotranspiration, soil and plant based data. *Report to the California Walnut Board.* 113–136 (2004).
- Fernandes-Silva, A. A., Ferreira, T. C., Correia, C. M., Malheiro, A. C. & Villalobos, F. G. Influence of different irrigation regimes on crop yield and water use efficiency of olive. *Plant Soil.* **333**(1), 35–47 (2010).
- Dahikar, S.S., & Rode. S.V. Agricultural crop yield prediction using artificial neural network approach. *Int. J. Innov. Res. Electr. Electron. Instrument. control Eng.* **2**(1), 683–686 (2014).
- Dzikiti, S. *et al.* Estimating the water requirements of high yielding and young apple orchards in the winter rainfall areas of south Africa using a dual source evapotranspiration model. *Agric. Water Manag.* **208**, 152–162 (2018).
- Emami, S. & Choopan, Y. Estimation of barley yield under irrigation with wastewater using RBF and GFF models of artificial neural network. *J. Appl. Res. Water Wastewater* **6**(1), 73–79 (2019).
- Bang, S., Bishnoi, R., Chauhan, A.S., Dixit, A.K. Chawla, I. fuzzy logic based crop yield prediction using temperature and rainfall parameters predicted through arma, sarima, and armax models. *12th Int. Conf. Contemp. Comput.* 1–6 (2019).
- Kumar, Y. J. N., Spandana, V., Vaishnavi, V. S., Neha, K., & Devi, V. G. R. R. Supervised machine learning approach for crop yield prediction in agriculture sector. *In 2020 5th International Conference on Communication and Electronics Systems (ICCES).* 736–741 IEEE, (2020).
- Sharifi, A. Yield prediction with machine learning algorithms and satellite images. *J. Sci. Food Agric.* **101**(3), 891–896 (2021).
- Prasad, N. R., Patel, N. R. & Danodia, A. Crop yield prediction in cotton for regional level using random forest approach. *Spat. Inf. Res.* **29**(2), 195–206 (2021).
- Dehghanisanij, H., Emami, S., Achite, M., Nguyen Linh, T.T. & Quoc B.P. Estimating yield and water productivity of tomato using a novel hybrid approach. *Water.* **13**(24), 3615 (2021).
- Howell, T. A. Enhancing water use efficiency in irrigated agriculture. *Agron. J.* **93**, 281–289 (2001).
- Wright, G.C. Irrigating Citrus Trees. *Coop. Ext.* 1–5 (2000).

32. Brouwer, C., Prins, K., Kay, M. & Heibloem, M. Irrigation water management: Irrigation methods. *Train. Manual* **9**(5), 5–7 (1988).
33. Emami, H. Seasons optimization algorithm. *Eng. Comput.* 1–21 (2020).
34. Jang, J. S. R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **23**(3), 665–685 (1993).
35. Emami, S., Parsa, J., H. Emami, & Abbaspour, A. An ISaDE algorithm combined with support vector regression for estimating discharge coefficient of W-planform weirs. *Water Supply* **21**(7), 3459–3476 (2021).
36. Pourmohamadali, B., Salehi, M. H., Hosseinifard, S. J., Shirani, H. & Esfandiarpour Borujeni, I. Pistachio yield prediction using multiple linear regression and artificial neural network (a case study: rafsanjan and anar regions, kerman province). *J. Agric. Eng. Soil Sci. Agric. Mech. (Sci. J. Agric.)* **40**(2), 71–87 (2018).
37. Sedaghati, N., Hosseinifard, S.J., Hokmabadi, H. & Mohammadi Mohammadabadi, A. Study the possibility of alternate irrigation borders in order to reducing irrigation intervals in pistachio orchards. *Res. Construct.* **21**, 153–162 (2009).
38. Montazer, A., Azadeghan, B. & Shahraki, M. Assessing the efficiency of artificial neural network models to predict wheat yield and water productivity based on climatic data and seasonal water-nitrogen variables. *Iran. Water Res. J.* **3**(2), 17–29 (2010).
39. Shirdeli, A. & Tavassoli, A. Predicting yield and water use efficiency in saffron using models of artificial neural network based on climate factors and water. *J. Saffron Agron. Technol.* **3**(2), 121–131 (2015).
40. Zeinadini, M., Shirani, H., Mozafari, V. & Esfandiarpour, I. Variability of soil electrical conductivity and sodium adsorption ratio and their impact on growth pistachio. *J. Water Soil Conserv.* **20**(6), 165–181 (2013).
41. ASTM, D 422. Standard test method for particle-size analysis of soils. (2007).
42. Devices, D.-T. *User manual for the profile probe: Type PR2* (Delta-T Devices Ltd., 2008).
43. Kanooni, A. Evaluation of furrow irrigation efficiency under different managements in Moghan region. *J. Agric. Res.* **8**(2), 17–32 (2008).

Acknowledgements

We would like to thank the Agricultural Engineering Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Alborz, Iran, for assisting to conduct this study. This research was carried out with the support of the Conservation of Iranian wetlands project (CIWP) and in the framework of “Modeling local community participation in Lake Urmia restoration via the establishment of sustainable agriculture”.

Author contributions

H. D, S.E and V.R. wrote the main manuscript text and H.E. and S.E. prepared Figs. 1–13. S.E. and H.E. prepared the hybrid proposed model/materials/equipment’s/irrigation system. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022