J. Agr. Sci. Tech. (2011) Vol. 13: 627-640

Predicting Dryland Wheat Yield from Meteorological Data Using Expert System, Khorasan Province, Iran

A. Khashei-Siuki[1], M. Kouchakzadeh[1*], and B. Ghahraman[2]

ABSTRACT

Khorasan Province is one of the most important provinces of Iran, especially as regards agricultural products. The prediction of crop yield with available data has important effects on socio-economic and political decisions at the regional scale. This study shows the ability of Artificial Neural Network (ANN) technology and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for the prediction of dryland wheat (*Triticum aestivum*) yield, based on the available daily weather and yearly agricultural data. The study area is located in Khorasan Province, north-east of Iran which has different climate zones. Evapotranspiration, temperature (max, min, and dew temperature), precipitation, net radiation, and daily average relative humidity for twenty-two years at nine synoptic stations were the weather data used. The potential of ANN and Multi-Layered Preceptron (MLP) methods were examined to predict wheat yield. ANFIS and MLP models were compared by statistical test indices. Based on these results, ANFIS model consistently produced more accurate statistical indices (R^2= 0.67, RMSE= 151.9 kg ha^{-1}, MAE= 130.7 kg ha^{-1}), when temperature (max, min, and dew temperature) data were used as independent variables for prediction of dryland wheat yield.

Keywords: ANFIS, Artificial neural network, Dryland wheat yield, Khorasan, Multi-layered preceptron, Prediction.

INTRODUCTION

Wheat is man's food stable stuff, planted in most parts of the world. Wheat generally grows between 30-50 degree north and 25-40 degree south where dryland wheat constitutes 66% of the total area. Climatic conditions on the other hand from some of the most the important factors in fluctuation in dryland wheat production and yield. The prediction of wheat yield is a country's major concern for scheduling, food security, distribution appraisal, as well as for import/export issues. It is indispensable to schedule and plan the production of food industries. Many import/export industries provide food (such as baking flour and macaroni) presently, materials from wheat.

Presently, many proposed prediction models of crop yield have been divided into two categories of mechanistic and empirical approaches (Poluektov and Topaj, 2001). The mechanistic models use mathematical functions to represent physical, biological, and chemical processes (Whisler *et al.,* 1986). However, these models are suitable for areas outside the data range used for development. They tend to be complex and require many input parameters (Wang *et al.,* 2002; Basso *et al.,* 2001; Bolte, 1997). The empirical models are based on correlative factors between variables, which are relatively simple and require less data; but

[1] Department of Irrigation and Drainage Engineering, College of Agriculture, Tarbiat Modares University, Tehran, Islamic Republic of Iran.
[*] Corresponding author, e-mail: kouchakm@modares.ac.ir
[2] Department of Irrigation, College of Agriculture, Ferdowsi University of Mashhad, Mashhad, Islamic Republic of Iran.

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

such models cannot be used in the areas outside data range that they have been created for.

In early 1920s, simple descriptive approaches of the relationship between weather and crop growth were advented (Landau *et al.,* 2000). Many such factors as changes in the weather conditions, soil moisture, topology, plant root water uptake, temperature-related stress, and the degree of nutrient consumption will affect crop yield. Predictive yield models needed many input data, the collecting of which was difficult.

With increasing knowledge on plant growing processes and the way to express them by mathematical formulations, deterministic models have reached a high complexity. Simulations consider full crop developmental stages and/or durations between them as well as the plant reactions in different phenological stages to different environmental conditions by using empirical approximation functions. The underlying assumptions, *viz* the response of plant growth to temperature and other environmental parameters during developmental stages are sometimes linear or even constant (Jame *et al.,* 1999; Porter and Jame, 1985). There are many studies on modeling crop environment relationships and developing operational yield-prediction systems. Furthermore, requirement for detailed meteorological, soil, and management inputs, not always available everywhere, has to be emphasized. CropSyst (Cropping Systems Simulation Model) and MEDIWY (Model for Estimation of Dryland and Irrigated Wheat Yield) are two important simulation models (Sepaskhah *et al.,* 2006; Claudio *et al.,* 2003). CropSyst a multi-year, multi-crop, and daily time step cropping system simulation model was developed to serve as an analytical tool to study the effect of climate, soil, and management on cropping system productivity as well as on environment. CropSyst simulates soil water, nitrogen budgets, crop growth and development, crop yield, residue production and decomposition, and soil erosion through water and salinity (Claudio *et al.,* 2003). Sepaskhah *et al.* (2006) evaluated MEDIWY model and modified it for simulation of Sabalan winter wheat under irrigated and rainfed conditions in Maragheh area (Eastern Azarbayejan Province, I. R. Iran) for three consecutive crop years. They compared the simulated and the obtained grain yields. It was found out that the simulated grain yield, under irrigated conditions was satisfactory, but they could not satisfactorily simulate grain yield under rainfed conditions.

As a result, a large number of later approaches, models, algorithms, and statistical tools have been proposed and used for assessing yield prediction in agriculture. Many researchers used simple linear correlation of yield with soil properties, but the results are different from field to field and year to year (Gemtos *et al.,* 2004; Khakural *et al.,* 1999; Drummond *et al.,* 1995). Many studies have adopted complex linear methods such as multiple linear regressions, which consist of similar results (Kravchenko and Bullock, 2000; Khakural *et al.,* 1999; Drummond *et al.,* 1995). Some scientists proposed non linear statistical methods to investigate the yield response (Adams *et al.,* 1999; Wendroth *et al.,* 1999). The important dilemma is how the independent variables are coupled to each other. To alleviate this difficulty, ANN models have come to play a role. Expert systems and artificial intelligent algorithms are a relatively new subjet of nonlinear techniques. Through ANN models one is able to solve highly nonlinear problems and to approximate virtually any smooth and measurable functions. In comparison to the state of the art of crop models, the requirements concerning the number of input parameters are less.

Some researchers used the systems of ANN and Adaptive-Neuro Fuzzy Inference System (ANFIS) for precision agriculture. They examined the applicability of ANN and ANFIS for development of yield mapping and forecasting systems by using satellite images *vs.* soil parameters. (Uno *et*

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

al., 2005; Park *et al.*, 2005; Stathakis *et al.*, 2006) compared adaptive techniques with general linear models to predict crop yield response under different soil and land management conditions. Input data for models were soil physical and chemical properties.

Heinzow and Richard (2002) discussed the applicability of ANN models for predicting crop yield under climate change conditions. Input data were daily temperature and precipitation for different growth stages in four German zones. They developed a four-layer ANN. Hosseini *et al.* (2007) adopted ANN and multi-variable regression models for dryland wheat yield in a moderate climate in Ghorve of Kordestan Province, Iran. They showed that ANN model can estimate the crop yield with acceptable accuracy. Maximum and minimum air temperature, daily mean relative humidity, net radiation, precipitation, dew point temperature, and wind velocity were included as input data in their ANN models. Kaul *et al.* (2005) developed ANN for corn and soybean yield predictions. They used the historical yield data at numerous locations in Maryland, USA. The results indicated ANN models as consistently producing more accurate results than others.

Due to the importance of prediction models for crop yield, paying attention due to possible differences between MLP and ANFIS models, specifically under dryland conditions in the arid and semi-arid climates, is essential. The objectives of this study are: (1) to_evaluate ANN and ANFIS models for wheat yield estimation in an arid and semi-arid climate by using meteororogical data, and (2) to compare the results obtained from the two models ANN and ANFIS.

MATERIALS AND METHODS

The study was conducated over the set of the three Khorasan Provinces as one of the important regions of agricultural production in Iran. Khorasan Provinces rank first in farming, and second in wheat production in Iran. The data in this study were taken from nine synoptic stations (Figure 1) and as well from the Agricultural Organization of Khorasan Province recorded for a period of twenty two years (1984 to 2006). The meteorological data employed in this study consisted of the daily observations of maximum and minimum air temperatures (Tmax and Tmin), daily average relative humidity (RHmean), net radiation (Rn), and

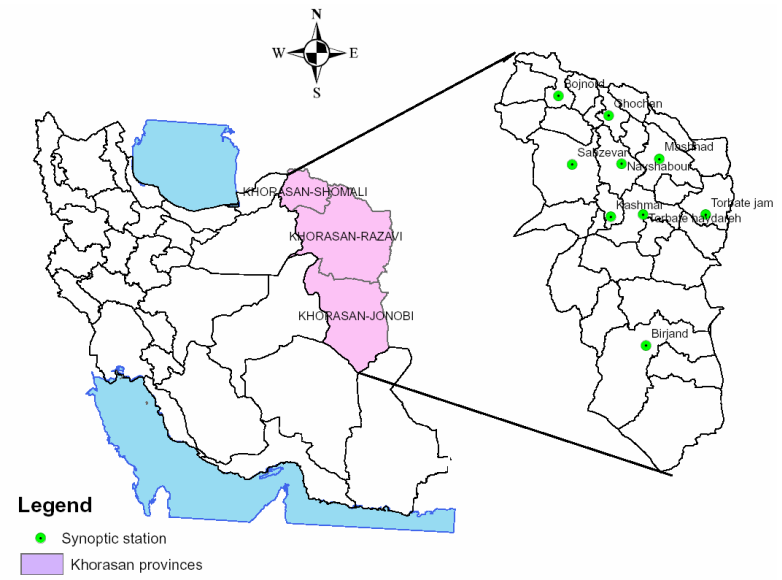

Figure 1. The location of the study area (Khorasan Province, Iran) and synoptic stations.

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

precipitation (P), as well as dew point temperature (Tdew).

The data related to mean air temperature were obtained by averaging the maximum and minimum values. Thermo hydrograph, dry and wet bulb thermometers placed in Stevenson Screen, were used to provide relative humidity values. The mean relative humidity data were obtained through averaging the maximum and minimum values. Evapotranspiration (ET) was calculated using FAO penman Montith equation (Allen *et al.,* 1998). The whole parameters were considered from 23 September to 21 June for each year, considered as an average for wheat growth period in this region. Dryland wheat yields in kilograms per hectare are depicted in Table 1. Models or inputs were chosen based on type and available parameters and also on their influence on crop yield. Six different models were employed for the analysis (Table 2).

The data were divided into three partitions, 60 percent for training, 20 percent for testing, and the remainder for validation. The structural design for MLP models was done by Sigmoueid function and hidden layer. The optimum number of hidden neurons was selected by trial and error. The number of membership functions and parameters of models were also determined through trial and error procedures. One hundred out of a total of one hundred and sixty data items were used for training the ANFIS model and the rest used for test and validation of the models.

To avoid overfitting, an early stopping criterion was adopted to improve the network training speed and efficiency. For this criterion, the data sets were divided into three as corresponded to validation-checking of ANN, and ANFIS toolbox of MATLAB (Rahimi khob, 2008). Two toolboxes of ANN and ANFIS in MATLAB software were used for the research. To obtain the consistency of the the model, all data sets were normalized first in the range of 0 to 1 and then returned to the original values after simulation by using the equation of (Doğan, 2008):

$$X_{norm} = \left[\frac{X - X_{min}}{X_{max} - X_{min}} \right] \times 0.8 + 0.1 \qquad (1$$

Where X is the original value, X_{min} and X_{max} are minimum and maximum values in

Table 1. Range and statistical characteristics of collected data set.

Row	Parameter	STDEV	Average	Max	Min
1	ET(mm)	97.14584	716.106	930.0401	364.7413
2	ET mean (mm day^{-1})	0.328748	2.653832	3.40674	1.981985
3	Rn	1436.379	2479.702	7470.921	865.9323
4	Rn mean	5.215586	9.168492	27.26613	5.378461
5	P(mm)	72.21752	219.1228	464.6	60
6	P mean (mm day^{-1})	0.938252	3.956598	7.039394	0.686996
7	RH %	2359.842	14179.69	19402.7	9052.4
8	Average RH %	8.667092	52.64687	71.07216	33.28088
9	Tdew ºC	399.7992	562.893	1590.65	-497.75
10	Tdew mean ºC	1.501469	2.103749	5.847978	-1.84352
11	Tmax ºC	494.3111	4530.087	5580.5	2564.75
12	Tmax(mean) c^{o}	1.549894	16.8201	20.44139	13.43476
13	Tmin ºC	427.5177	1076.949	2091.75	205.5
14	Tmin (Mean) ºC	1.575802	4.036637	7.634124	0.769663
15	Wheat yield of dryland (kg ha^{-1})	231.2015	415.1355	1183.6	25

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

Table 2. Performance statistics of the MLP models for dryland wheat yield estimation.

Model Types	Input parameters	Best number of hidden node	Epoch
(a)	ET mean-P-RH	1	25
(b)	P-RH mean	4	14
(c)	Tdew mean-Tmax mean-Tmin mean	37	12
(d)	ET mean-Rn mean-P mean-RH mean-Tdew mean -Tmax mean -Tmin mean	3	18
(e)	ET-Rn-P-RH-Tdew-Tmax-Tmin	8	18
(f)	ET mean- Rnmean -P mean- RH mean -Tdew mean - Tmax mean- Tmin mean- ET -Rn -P –RH- Tdew- Tmax-Tmin	4	19

the series, respectively, X_{norm} is the normalized value, while 0.8 and 0.1 are the scaling factors. Different values may be assigned for the scaling factors. However, there is no proposed rule on standardization approach that can be used in particular circumstances (Dawson and Wilby, 1998). In this study, the scaling factors were selected as 0.8 and 0.1, respectively.

The data sets , used for testing of neutral network section (the daily weather data), were used for a comparison of selected ANN, ANFIS, as well as for the observation values. This comparison was performed by using three statistical indices: Determination Coefficient (R^2), Root Mean Square Error (RMSE), and MAE (Mean-Absolute Errors), as follows:

$$R^2 = \left[\sum_{i=1}^{N}(P_i - \overline{P})(O_i - \overline{O})\right]^2 \left[\sum_{i=1}^{N}(P_i - \overline{P})^2 (O_i - \overline{O})^2\right]^{-1} \quad (2$$

$$RMSE = \left[N^{-1}\sum_{i=1}^{N}(P_i - O_i)^2\right]^{0.5} \quad (3$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|O_i - P_i\right| \quad (4$$

Where N is the number of observations, P_i the estimated values (using ANN and ANFIS), O_i is the observed data, while $\overline{P}$ and $\overline{O}$ the mean values for P_i and O_i, respectively. For finding the best results, the data sets used for testing of ANN and ANFIS, were also used for a comparison of selected ANN as well as for the conventional method. This comparison was performed by using three statistical parameters, categorized through their sorting and then selecting the most appropriate degree.

The statistical comparison of the regression coefficients were investigated through student *t*-test and conducted between predicted *vs.* observed in the form of $Y= a+bX$. The statistic indicators for each ANFIS model in *t*-test period were compared with the ANN model**.**

Artificial Neural Network (ANN) and Adaptive-Neuro Fuzzy Inference System (ANFIS)

A neural network is a form of artificial intelligence that imitates some functions of the human brain. ANN is a relatively new nonlinear statistical technique. It can solve the problems, which do fit the conventional statistical methods. A neural network consists of simple synchronous processing elements, called neurons which are inspired by biological nerve system (Malinova and Guo, 2004). The network comprises of a large number of simple processing elements which are connected to each other by the weighted connections, according to the required specified architecture. These networks learn from_the training data by adjusting the connection weights (Bishop, 1995). The neural network approaches have been successfully applied in various number of such diverse

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

fields as water resources. The neural networks are general-purpose computing tools, able to solve complex non-linear problems (Khashei-Siuki *et al.,* 2009).

In modern modeling methods, fuzzy system and fuzzy logic have their peculiar places (Zadeh, 1965). The characteristics of these methods are the ability of implementing human knowledge through tongue label and fuzzy rules, nonlinearity of these systems and their adaptability (Tashnehlab and Menhaj, 2001). A fuzzy system is a logical system based on "if-then" rules. Initial point of building and developing a new fuzzy system is the derivation of a set of if-then fuzzy rules by knowledge of expert person or modeling field (Nayak *et al.,* 2004). Estimation of a method or tool to achieve fuzzy rules from numerical, statistical tongue information is a suitable and simple method for modeling (Dezfoli, 2003). The other modern modeling method is the artificial neural network. The most important ability of these methods is the training ability from train sets (proper input and output pairs). These methods use several training algorithms to extract the relationships between input and output parameters (Jang, 1993). Combining fuzzy systems with ANNs, which extract information from numerical processes can develop models, which simultaneously use these numerical information and tongue statements. This combination of artificial neural network and fuzzy systems were named adaptive neuro-fuzzy inference system (Gopakumar and Mujumdar, 2007; Kisi *et al.,* 2006; Sen and Altunkaynak, 2006). A fuzzy system is a system, which is based on logical rules of "if-then" statements. This system portrays the input variable space to output variable space by using tongue statement and a fuzzy decision making procedure (Jang and Gulley, 1995). The fuzzy rule sets are a set of logical rules, which describes the relationships between fuzzy variables and the most important component of a fuzzy system (Karamouz *et al.,* 2004). Due to the inherent uncertainty in real and field data set, a fuzzification transition was employed to transform deterministic values of fuzzy steps, and a de-ffuzification transition that was used to transform fuzzy values (Riahi-madvar *et al.,* 2009; Maier and Dandy, 1996).

Based on a combination of dependent variables six different models were adopted. The number of dependent variables varied from 2 for model (b), and 14 for model (f) (Table 2). Mean and the total number of parameters for the wheat growing period were employed.

RESULTS AND DISCUSSION

In spite of variation from place to place as

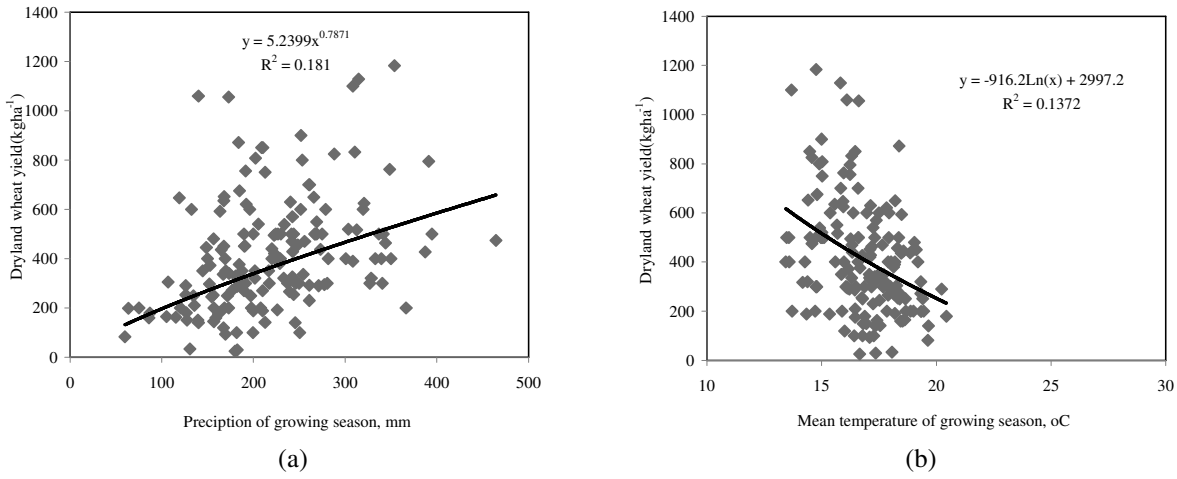

Figure 2. dependency of dry wheat yield upon rainfal (a) and temperature (b), considered for the growth period of wheat in the region.

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

well as from year to year, dryland wheat yield demonstrated a highlighted dependency on climate parameters. Figure 2 shows such a dependency on precipitation and temperature (as an average over the wheat growing period) respectively. The yield was increased by increasing rainfall, which indicated the importance of water in the arid climate. On the other hand, the increasing of yield with decreasing temperature was due to avoiding thermal stress under inadequate moisture. The variability of these variables (meteorological and wheat yield) for this study is summurized in Table 1.

ANN Models

In this section, the statistical parameters for accuracy of ANN model were presented. Coulibaly *et al.* (1999) confirmed that for many experimental results, one hidden layer may suffice most of the prediction problems. Therefore, in this study only one hidden layer was used for MLP model (ANN). The multi-layered preceptor was trained by using 1 to 50 hidden nodes and after each training, RMSE, R^2, and MAE were calculated using only the test data set to find the optimal number of hidden nodes for input layer.

Table 2 shows the best number of nodes for hidden layer on the network accuracy. These nodes are varied from 1 in model (a) to 37 in model (c). The most suitable RMSE and MAE (152 and 119, respectively) are related to model (f) (Figure 3) with the highest number of independent variables. The results show that MLP model with its total input layers (model f) performed the best, where the slope of the fitted line between the actual data and predicted ones, was close to one (lying on 1:1 line) at 5% level of significance. Model (c) (with dependent variables of Tdew, Tmean, Tmax, Tmin) was the second best one. The temperature data set is recorded in many meteorological stations. The present results of ANN model are different in their structure section from the model reported by (Hosseini *et al.,* 2007). The possible reason is due to different climatic conditions in these two study sites. All statistical performance criteria are included in Figure 3 for all the MLP models, listed in Table 2. The selected model performs well in comparison to the observational data.

Heinzow and Richard (2002) calculated the regional correlation coefficients from 0.68 to 0.97 for different crops by pruning in Germany through two aggregated meteorological parameters (Tmean and P). They demonstrated different combinations of meteorological inputs (Tmean, Tdew and P for oat; Tmean, Tmax and P for spring-barley; and Tmean and P for silage maize) for their best performance. However, here it was found that the most suitable meteorological input combination is Tmax, Tmin, and Tdew.

The results of these models, the parameters of which come from annual time scale (model e), indicate that the prediction of yield was as good as that for the models with parameters which correspond to wheat growth period for 23 September to 21 June (Figuers 3 and 4). Also, Kaul *et al.* (2005) confirmed that the parameter for different time scales will lead to different performances.

ANFIS Model

Different membership functions, fuzzy rules, and epoch numbers were considered to attain the best combination of independent variables (list of modes are presented in Table 2). The performances of all models were traced accordingly to find the best model for predicting dryland wheat yield in Khorasan province of Iran as based on ANFIS methodology.

The fuzzy membership functions are able to take many forms, but simple straight line triangular and Gaussian functions are the most common (Kisi and Oztork, 2007). In this study, the triangular and Gaussian membership functions were tried, but the results showed that they wouldn't satisfactorily predict the yield (Table 3),

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

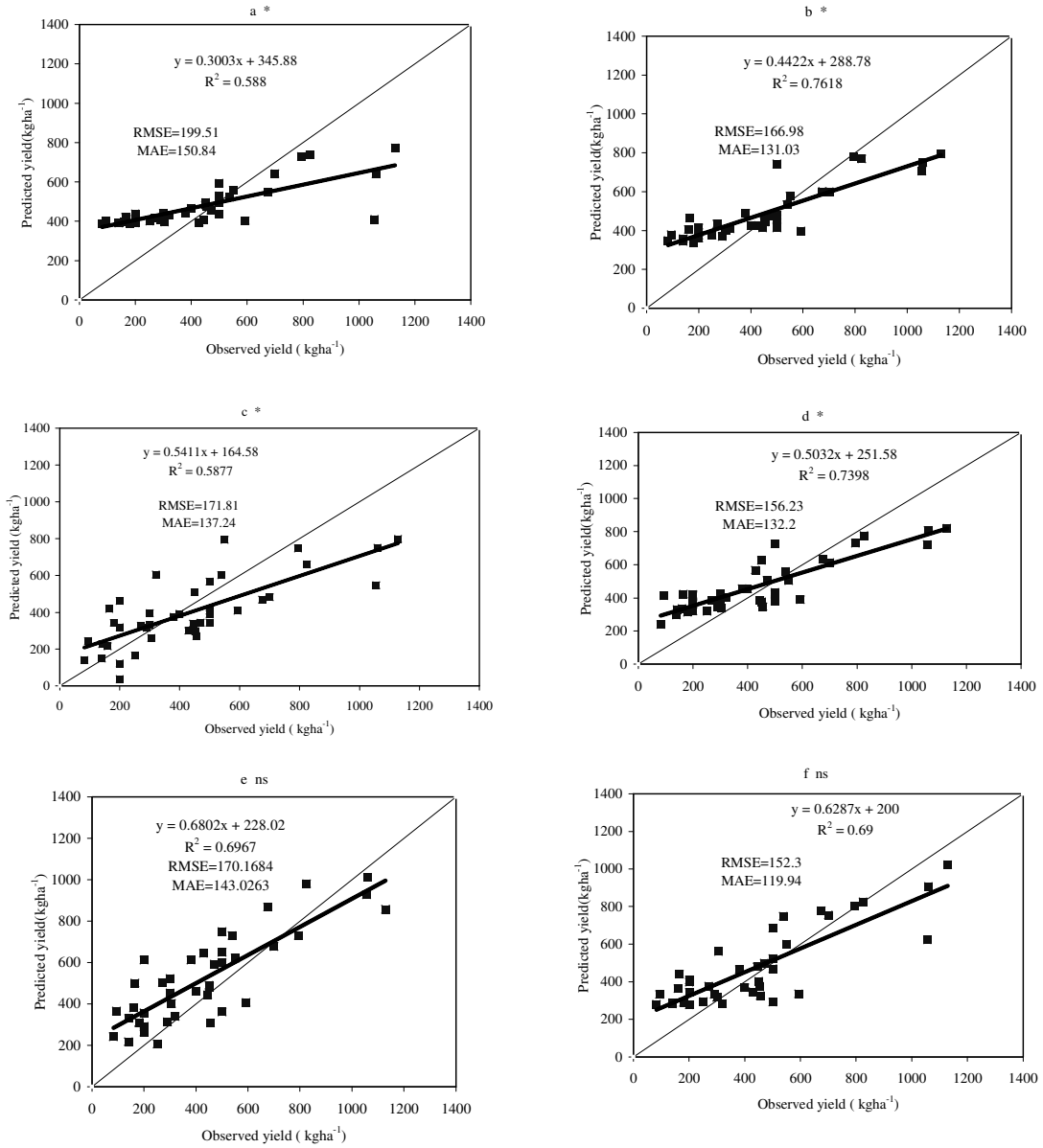

Figure 3. Scatter plot of observed versus estimated values of crop yield for the testing data set of MLP for different models of (a), (b), (c), (d), (e), and (f). Dependent variables for the models are described in Table 2. The thick line is related to regression line and the thin one is for the best fit between observed (X) and predicted (Y) yield. The statistical inferences on *a* and *b* parameters of the equation $Y= a+bX$ are: ns for non significant, * for significant at the 0.05 probability level, respectively.

therefore the clustering option in ANFIS options was checked before generating the FIS. By this, one can suppose that the performance of the model will be increased. This subtractive clustering method, seperates the data set into different groups called clusters, and also generates FIS with the minimum number of rules, required to distinguish the fuzzy qualities associated with each cluster. The structure of ANFIS, used is presented in Figure 5. The results of

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

Table 3. Performance statistics of the membership function models of ANFIS for wheat yield estimation.

Moldel	Type of membership functions	Optimal model	Number of membership function	R^2 test	RMSE test (Kg ha^{-1})	AME test (Kg ha^{-1})
(a)	Gus2mf	Backpropagation	3 3 3	0.5641	189.6993	131.5536
(b)	Trimf	Hybrid	3 3 3	0.2893	397.2745	205.7042
(c)	Trimf	Hybrid	3 3 3	0.3049	233.4089	179.126
(d)	Trimf	Hybrid	3 3 3 3 3 3	0.0785	1088.027	758.7483
(f)	Gus2mf	Backpropagation	3 3 3 3 3 3 3	0.1425	505.0128	437.8468
(e)	Gus2mf	Hybrid	3 3 3 3 3 3 3	0.1343	838.4182	694.9826

this new model are presented in Table 4 and also in Figures 6 and 7.

Table 4 and Figures 6 and 7 are created because of testing period of the data set and also the statistical performances of these models, corresponding to the different input layers. The results indicate that the ANFIS model with clustering option provides the most accurate dryland wheat yield estimation. The degree of improvement was 350% and 173% for RMSE and MAE, respectively.

As mentioned before, the increasing of the sub-clustering parameters always increase the accuracy, but sometimes the results for poor estimations are due to the high number of parameters, which need to be optimized. The sub-clustering parameters for prediction, including the fuzzy rule-based for the model, are presented in Table 3. Following 400 epochs, the desired parameters were achieved.

It is clear from the scatterplots (Figures 6 and 7) that the models (c) and (d) are more closely matched with the corresponding observed yield values than those of the others. As seen from the fitted line equations (Y= a+bX), the *a* and *b* coefficients for the models are closer to 0 and 1, respectively, than the others. The results of *a* and *b* parameters of student *t*-test in Figure 7 show that the regression coefficients of model (c) and (d) are not significantly different from 0 and 1 at the 5% level of significance.

Comparing the Results of ANN and ANFIS Models

The scatter plots of the observed and predicted values, including the best fit line, are presented in Figures 3 and 7. The results show that ANFIS model is more accurate than ANN for (a), (b) and (c) models where ANFIS benefits from the most accuracy with temperature inputs (model c). The results also indicate that the predictive capability of ANFIS model is poor in comparson with ANN model as the number of input layers increase. Obviously the three-parameter

Table 4. Performance statistics of MLP models for wheat yield estimation.

Model	Reject ratio	Accept ratio	Squash factor	Range of influence	Optimal model	R^2 test	RMSE test (Kg ha^{-1})	AME test (Kg ha^{-1})
(a)	0.3	0.5	5	0.3	Hybrid	0.7378	170.3721	123.8827
(b)	0.15	0.5	4	0.2	Hybrid	0.772	161.3242	128.8192
(c)	0.15	0.5	1.25	0.5	Backpropagation	0.6698	151.968	130.702
(d)	0.3	0.5	1.75	0.4	Hybrid	0.484	188.6661	154.922
(f)	0.4	0.5	2	0.3	Hybrid	0.3351	1776.623	150.0755
(e)	0.4	0.5	2	0.3	Hybrid	0.2707	213.0884	1202.578

models performed better than the four-parameter solutions for ANFIS model. The four-parameter models of ANN predicted better results than the others.

The expert system analyzed the experimental data while deducing their latent rules. Another important characteristic of these models is that they wouldn't need all the information that are active in this phenomena, and can significantly save time in the data collection. The most appropriate ANFIS model was (c) with temperature inputs. Temperature is an effective parameter which controls evapotranspiration

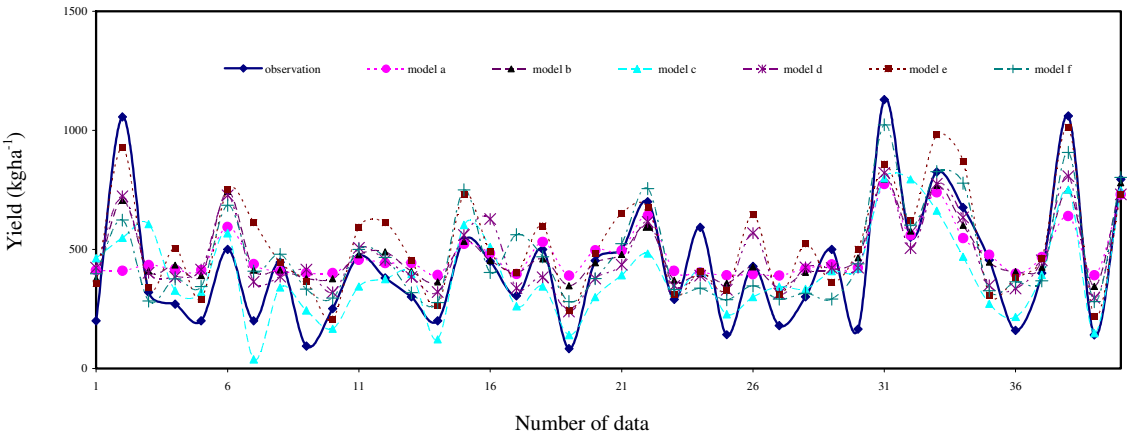

Figure 4. Performance of MLP model results during test period.

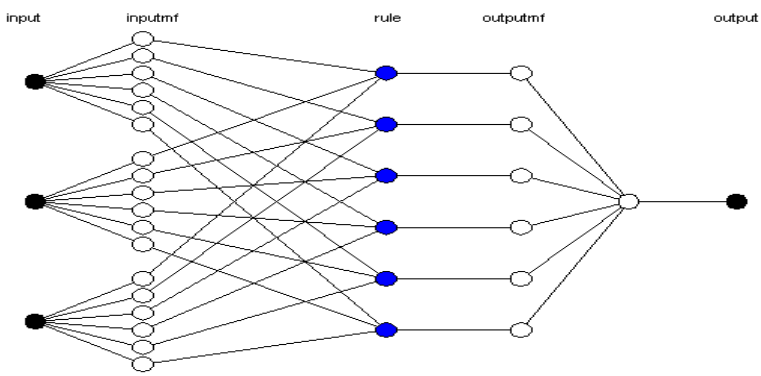

Figure 5. ANFIS architecture for model c. Not all the connections and nodes are shown. Rule operator is AND in logic rule.

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

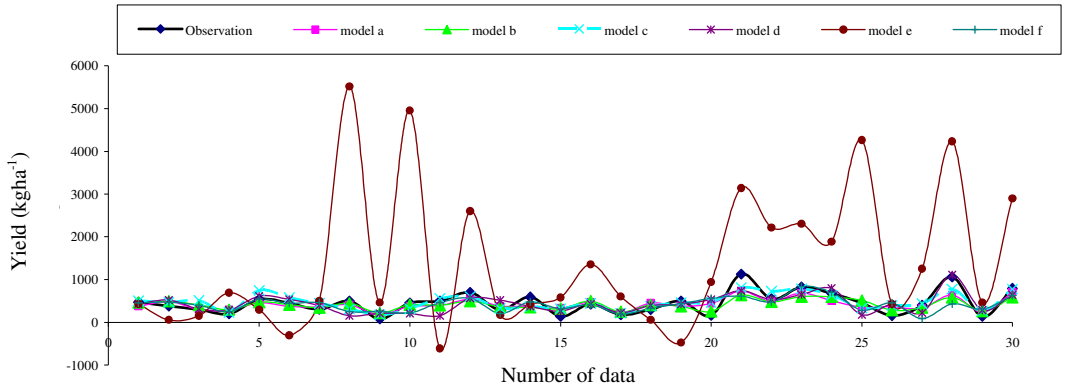

Figure 6. Performance of ANFIS models in testing step.

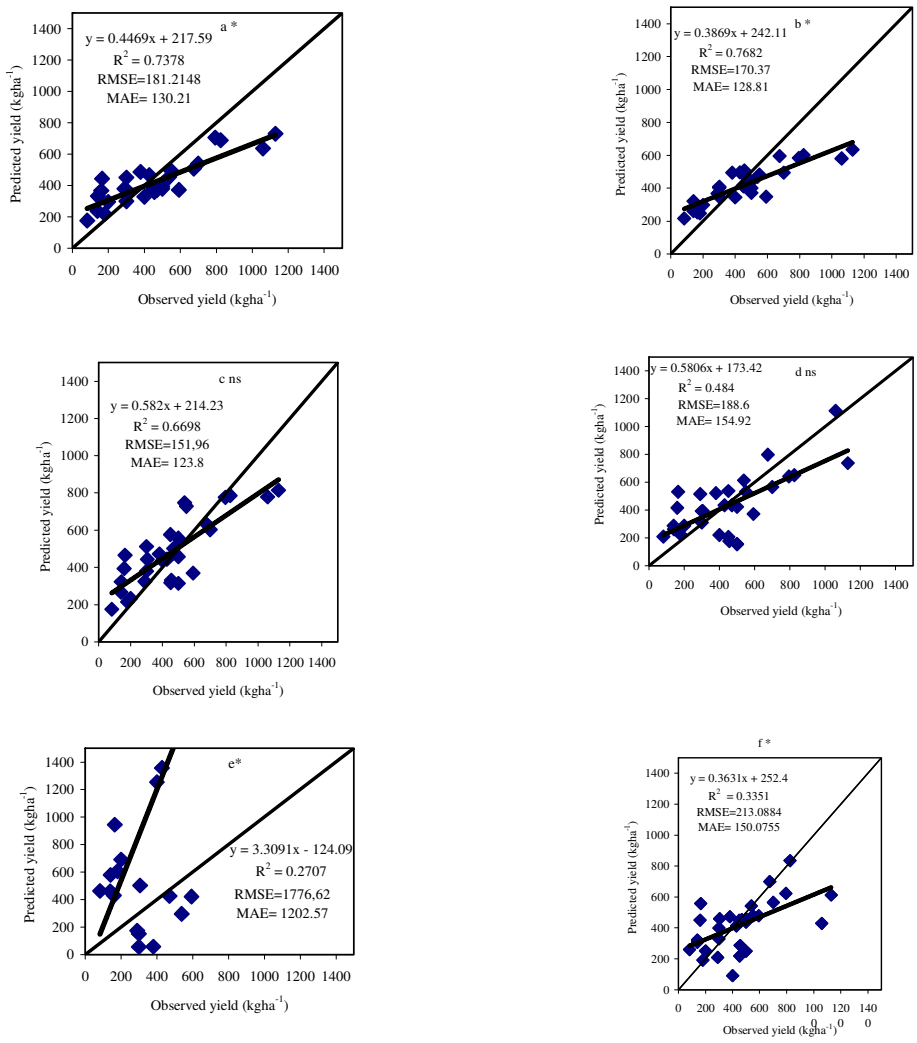

Figure 7. Scatter plot of observed versus estimated values of crop yield for the testing data set of ANFIS for different models of (a), (b), (c), (d), (e), and (f). Dependent variables for the models are described in Table 2 The thick line is related to regression line and the thin one represents the best fit between observed (X) and predicted (Y) yield. The statistical inferences on *a* and *b* parameters of the equation $Y= a+bX$ are: ns for non significant, * for significant at the 0.05 probability level, respectively.

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

and is affected by rainfall. Evapotranspiration also depends on many other parameters. However, considering all parameters greatly involves the measuring errors. On the other hand, the correlation between crop yield, as the dependent variable, and temperature or precipitation, as the independent one, do not differ markedly (their R^2 was about 13 to 18 percent; see Figure 2). This reason is not considered as an important factor, considering the performance of the model, whilst the model performed better with different temperature parameters. On the other hand, ANFIS model was dependent on the number of input parameters, because the accuracy of these models was reduced with increasing the number of input parameters. The accuracy of ANN model remains acceptable when the number of parameters increase. These models benefit from high accuracy when input parameter is temperature. This might be due to the high relationship among temperature parameters, which could be accurately measured.

CONCLUSIONS

In this study a method was explained to estimate the dryland wheat yield from meteorological data sets with several nonlinear modeling techniques for an arid and semi-arid climate. In this study, we used seven meteorological variables (precipitations, humidity, evapotranspiration, net radiation, maximum and minimum temperature, and dew temperature) were employed for an estimation of dryland wheat yield. The use of such nonlinear modeling methods as MLP and ANFIS was demonstrated. ANFIS method provided a general framework for the combination of ANN and fuzzy systems' capabilities. The performance of ANFIS model was more pronounced than MLP in testing period. In general, it can be concluded that ANFIS model has the ability for precise estimation of dryland wheat yield, while MLP being the most suitable model for this study area. Despite many applications of nonlinear models, which needed many input data, there is a lack of comparative studies of different models. Most research studies published so far, have employed a limited number of nonlinear models, while in this study different expert nonlinear models have been used for wheat yield estimations, which could be a valuable source of information for other researchers. ANN and ANFIS techniques could be used in many fields including scheduling, politics, design, and various other analyses. These models can also be integrated into modules for application in general economic models.

REFERENCES

1. Adams, M. L., Cook, S. E., Caccetta, P. A. and Pringle, M. J. 1999. Machine Learning Methods in Site–Specific Management Research: An Australian Case Study. *In Proc. 4th Int. Conf. on Precision Agriculture*, Madison, USA, PP. 1321–1333.
2. Allen, R. G., Preira, L. S., Raes, D. and Smith, M. 1998. Crop Evapotranspiration Guidelines for Computing Crop Water Requirement. FAO Irrigation and Drainage Paper, No. 56, Rome, Italy.
3. Basso, B., Ritchie, J. T., Pierce, F. J., Braga, R. P. and Jones, J. W. 2001. Spatial Validation of Crop Models for Precision Agriculture. *Agric. Sys.*, **68**: 97–112.
4. Bishop, C. M. 1995. *Neural Networks for Pattern Recognition.* Clarendon Press, Oxford. PP45.
5. Bolte, J., 1997. Biosystem Modeling Techniques. Biosystems Analysis Group, Oregon State University. [Online]Available at ttp://biosys.bre.orst.edu/BRE571/intro/intro_ concepts.htm (Verifed 29th April, 2004).
6. Claudio, O., Marcello, D. and Roger, N. 2003. CropSyst, a Cropping Systems Simulation Model, *Eur. J. Agron.*, **18(3-4)**: 289-307.
7. Dawson, W. C. and Wilby, R. 1998. An Artificial Neural Network Approach to Rainfall-runoff Modelling. *Hydrol. Sci. J.*, **43(1)**: 47–66.
8. Coulibaly, P., Anctil, F. and Bobee, B. 1999. Provision Hydrologique par Reseaux de Neurons Artificials. *Can. J. Civil Eng.*, **26(3)**: 293–304.
9. Dezfoli, K. A. 2003. Principles of Fuzzy Theory and Its Application on Water Engineering Problems. Jihad Press, Tehran, Iran, 227 PP.
10. Doğan, E. 2008. Reference Evapotranspiration Estimation Using Adaptive Neuro-Fuzzy

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

Inference System. *J. Irrig. Drain. Eng., ASCE,* **58**:617-628
11. Drummond, S. T., Sudduth, K. A. and Birrell, S. J. 1995. Analysis and Correlation Methods for Spatial Data. *ASAE,* **95**-1335.
12. Gemtos, A. T., Markinos, A., Toulios, L., Pateras, D. and Zerva, G. 2004. Precision Farming Applications in Cotton Fields of Greece. *In CIGR International Conference*, 11-14 October, Beijing, China, (CD ROM).
13. Gopakumar, R. P. and Mujumdar, P. 2007. A Fuzzy Dynamic Wave Routing Model. *Hydrol. Proc*., **21**: 458–467.
14. Heinzow, T. and Richard, T. 2002. Prediction of Crop Yields Across Four Climate Zones in Germany: an Artificial Neural Network Approach. Working Paper FNU-34 Centre for Marine and Climate Research, Hamburg University, Germany.
15. Hosseini, S. M., Seyosemorde, M., Fathi, P and Seyosemorde, A. 2007. Application of ANN and Multi-variable Regression Models on Yield of Dryland Wheat in Gorve of Kordestan. *Agri. Res. J.,* **7(1)**: 41-54. (In Persian)
16. Jame, Y. W., Cutforth, H. W. and Ritchie, J. 1999. Temperature Response Function for Leaf Appearance Rate in Wheat and Corn. *Can. J. Plant Sci*., **79**: 1-10.
17. Jang, J. S. R. 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE, Trans. Sys. Man. Cyb.,* **23(3)**: 665–685.
18. Jang, J. S. R. and Gulley, N. 1995. *The Fuzzy Logic Toolbox for Use with MATLAB*. The Mathworks Inc., Natick, MA.
19. Karamouz, M. M., Tabehs, A. and Nazief, A. 2004. Prediction of Pressure in Water Networks Using ANFIS Models. *Isfahan J. Water Wastewater,* 53: 3–14. (In Persian)
20. Kaul, M., Hill, R. L. and Walthall C. 2005. Artificial Neural Networks for Corn and Soybean Yield Prediction. *Agric. Sys*., **85**: 1–18.
21. Khakural, B. R., Robert, P. C. and Huggins, D. R. 1999. Variability of Corn/Soybean Yield and Soil/Landscape Properties Across a Southwestern Minnesota Landscape. *In Proc. 4th Int. Conf. on Precision Agriculture*, 19-20July, St. Paul, MN, Madison WI, USA., PP.573–579.
22. Khashei-Siuki, A., Kouchakzadeh, M. and Ahmadzadeh, K. 2009. Developing Expert Systems for Prediction of Evaporation by MLP, RBF and ANFIS Models. *XXXIII CIOSTA Conference*, 17-19 June 2009, Reggio Calabria, Italy, PP. 20-25
23. Kisi, O., Karahan, M. E. and Sen, Z. 2006. River Suspended Sediment Modeling Using a Fuzzy Logic Approach. *Hydrol. Proc*., **20**: 4351–4362.
24. Kisi, O. and Oztork. O. 2007. Adaptive Neurofuzzy Computing Technique for Evapotranspiration Estimation. *J. Irrig. Drain. Eng, ASCE*, **133(4)**: 368-379
25. Kravchenko, A. N. and Bullock, D. G. 2000. Correlation of Corn and Soybean Grain Yield with Topography and Soil Properties., *Agron. J.* **92 (1)**: 75–83.
26. Landau, S., Mitchell, R. A. C., Barnett, V., Colls, J. J., Craigon, J. and Payne, R. W. 2000. A Parsimonious, Multiple-Regression Model of Wheat Yield Response to Environment. *Agric. For. Meteorol*., **101**: 151-166.
27. Maier, H. R. and Dandy, G. C. 1996. The Use of Artificial Neural Networks for the Prediction of Water Quality Parameters. *Water Resour. Res.,* **32(4)**: 1013–1022.
28. Malinova, T. and Guo, Z. X. 2004. Artificial Neural Network Modeling of Hydrogen Storage Properties of Mg-based Alloys. *Mater. Sci. Eng. A.*, **365**: 219–227.
29. Nayak, P. C., Sudheer, K. P., Rangan, D. M. and Ramasastri, K. S. 2004. A Neuro-Fuzzy Computing Technique for Modeling Hydrological Time Series. *J. Hydrol*., **291**: 52–66.
30. Park, S. J., Hwang, C. S. and Vlek, P. L. G. 2005. Comparison of Adaptive Techniques to Predict Crop Yield Response Under Varying Soil and Land Management Conditions. *Agric. Sys*., **85**: 59–81.
31. Poluektov, R. A. and Topaj, A. G. 2001. Crop Modeling: Nostalgia about Present or Reminiscence about Future. *Agron. J.*, **93**: 653–659.
32. Porter, I. R. and Jame, R. 1985. Approaches to Modelling Canopy Development in Wheat. In: "*Wheat Growth and Modelling*". (Eds.): Day, W. and Atkin, R. K., Plenum Press, New York; London, PP. 69-82.
33. Rahimi-Khoob, A. 2008. Artificial Neural Network Estimation of Reference Evapotranspiration from Pan Evaporation in a Semi-Arid Environment. *Irrig. Sci.,* **27**:35–39.
34. Riahi-Madvar, H., Ayyoubzadeh, S. A., Khadangi, E. and Ebadzadeh, M. M. 2009. An Expert System for Predicting Longitudinal Dispersion Coefficient in Natural Streams by Using ANFIS. *Expet Sys. Appl.*, **36**: 589–8596.
35. Sen, Z. and Altunkaynak, A. 2006. A Comparative Fuzzy Logic Approach to Runoff Coefficient and Runoff Estimation. *Hydrol. Proc.,* **20**: 1993–2009.

[DOR: 20.1001.1.16807073.2011.13.4.7.6]

36. Sepaskhah A. R., Aghayari, F. and Tavakoli A. R. 2006. Evalution and Modification of MEDIWY Model for of Irrigated and Rainfed Wheat in Maraghe Area. *J. Agric. Eng. Res.,* **7(28)**:133-150. (In Persian)
37. Stathakis, D., Savina, I. and Nègrea, T. 2006. Neuro-fuzzy Modeling for Cop Yield Prediction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Part XXX, 34.
38. Tashnehlab, M. and Menhaj, S. 2001. Modeling Trip Tours Using ANFIS Modeling. *Eng. J. Tehran Univ*., **31(3)**: 361–370. (In Persian)
39. Uno, Y., Prasher, S. O., Lacroix, R. Goel, P. K., Karimi, Y., Viau, A. and Patel, R. M. 2005. Artificial Neural Networks to Predict Corn Yield from Compact Airborne Spectrographic Imager Data. *Comput. Electron. Agric*., **47**: 149–161.
40. Wang, F., Fraisse, C. W., Kitchen, N. R. and Sudduth, K. A. 2002. Site-specic Evaluation of the CROPGROW-Soybean Model on Missouri Claypan Soils. *Agric. Sys.,***76**: 985–1005.
41. Wendroth, O., Jurschik, P. and Nielsen, D. R. 1999. Spatial Crop Yield Prediction from Soil and Land Surface State Variables Using an Autoregressive State–space Approach. *Precision Agric.*, **99**: 419–428.
42. Whisler, F. D., Acock, B., Baker, D. N., Fye, R. E., Hodges, H. F., Lambert, J. R., Lemmon, H. E., McKinion, J. M. and Reddy, V. R. 1986. Crop Simulation Models in Agronomic Systems. *Adv. Agron.,* **40**: 141–208.
43. Zadeh, L. A. 1965. Fuzzy Sets. *Information Control*, **8(3)**:338–353.

پیش بینی عملکرد گندم دیم با استفاده از داده های هواشناسی بوسیله سیستم‌های هوشمند در استان خراسان، ایران

ع. خاشعی سیوکی، م. کوچک زاده و ب. قهرمان

چکیده

استان خراسان یکی ازاستانهای مهم ایران در تولید محصولات کشاورزی است. پیش بینی عملکرد محصولات با استفاده از داده های موجود تاثیرات مهمی در مسایل اجتماعی–اقتصادی و تصمیم گیریهای سیاسی در مقیاس منطقه ای دارد. این مقاله توانایی تکنولوژی شبکه عصبی مصنوعی (ANN) و سیستم استنتاج فازی (ANFIS) را برای پیش بینی عملکرد گندم دیم *(Triticum aestivum)* بر اساس اطلاعات روزانه هواشناسی و داده های سالانه کشاورزی نشان می دهد. استان خراسان در شمال شرق ایران واقع شده و دارای اقلیم های متفاوتی است. داده های هواشناسی مورد استفاده شامل دادهای ۲۲ ساله ۹ ایستگاه سینوپتیک استان و شامل تبخیر–تعرق، دما (حداکثر، حداقل و نقطه شبنم)، میانگین رطوبت نسبی، تشعشعات و بارندگی می باشد. پتانسیل مدل چند لایه پرسپترن (MLP) شبکه عصبی، برای پیش بینی محصول ارزیابی شد. مدل‌های ANFIS و MLP با استفاده از شاخص‌های آماری مقایسه شدند. بر اساس نتایج حاصله، در مدل ANFIS هنگامی‌که از داده های دمایی (حداکثر، حداقل و نقطه شبنم) به-عنوان متغیّرهای مستقل برای پیش‌بینی گندم دیم استفاده می شود (R^2= 0.67, RMSE= 151.9 kg ha^{-1}, MAE=130.7 kg ha^{-1})، بیشترین کارآیی حاصل می‌گردد.

[DOR: 20.1001.1.16807073.2011.13.4.7.6]