P1.39

ARTIFICIAL NEURAL NETWORKS APPLICATION TO PREDICT WHEAT YIELD USING CLIMATIC DATA

B. Safa*
Iranian Meteorological Organization, Tehran, Iran
A. Khalili
Department of Irrigation and Reclamation Engineering, College of Agriculture, University of Tehran, Karaj, Iran
M. Teshnehlab
Department of Electrical Engineering, K.N. Toosi Technology University, Tehran, Iran
A. Liaghat
Department of Irrigation and Reclamation Engineering, College of Agriculture, University of Tehran, Karaj, Iran

Abstract:

Prediction of crop yield mainly strategic plants such as, wheat, corn and rice has since long been an interesting research area to agrometeorologists, as it is important in national and international economic programming.

The main purpose of such studies is to estimate the crop production a few days or months before harvesting, using meteorological data.

Recently, the application of Artificial Neural Networks (ANNs) has developed into a powerful tool that can compute most complicated equations and numerical analyses to the best approximation. The goal* of this study was to apply the ANNs to predict dry farming wheat yield. According to the available data and information from different areas in Iran, this research was accomplished using Sararood station data in Kermanshah Province which has the most complete homogeneous statistics. In this study, the results of climatology for the period (1990-99) for each of the eleven phenological stages of wheat including sowing, germination, emergence, third leaves, tillerng, stem formation, heading, flowering, milk maturity, wax maturity, full maturity and also eleven meteorological factors including: mean daily minimum temperature, extreme daily minimum temperature, mean daily maximum temperature, extreme daily maximum temperature, total daily rainfall, number of rainfall days, sum of sun hours, mean daily wind speed, extreme daily wind speed, mean daily relative humidity and sum of water requirement were collected separately for each farming year and arranged in two matrices:

A matrix whose rows are repetitions of the statistical years (i) at each phenological stages of wheat (j) and the columns are meteorological factors (k).

A matrix whose rows form each of the statistical years (i) and the columns are meteorological factors (k) at each phenological stage (j).

In fact, statistical years (i), phonological stages (j) and meteorological factors (k) are the basic elements of 3-D matrix (M ijk) arranged as above and the best answer was derived only from the second matrix.

Finally, different networks were made for each stage and the optimum values of network parameters were obtained by trial and error. It should be reminded that two of the eight- year farming were randomly excluded from network training computations and that the comparison of the estimated data with the real data for these two years were used to test the accuracy of the models. The model that obtained has the following capabilities:

1. Prediction of wheat yield with maximum errors of 45-60 kg/ha at least two months before full maturity stage (end-of-stem formation stage).
2. Determination of the sensitivity of each phenological stage with respect to meteorological factors.
3. Determination of the priority order and importance of each meteorological factor effective in plant growth and crop yield.

1- Introduction:

Dry farming crop production, apart from its relationship to the genetic of cultivator, soil conditions, effect of pests and pathology and weeds, the management and control quality during the growing season etc. heavily depends on climatic events. Among them, the nature of rainfall interval, temperature variation during growth, speed and direction of wind and evapotranspiration are very important. Therefore, it is not very unlikely to achieve relations or systems that can predict the yield with higher accuracy using meteorological data. Nowadays, there are a variety of yield prediction models, generally falling into two classes: Statistical Models, Crop Simulation Models.

Statistical Models are chiefly based on using various regressions that compute the crop yield empirically. The need for long-term climatic data is among the leading essentials of this model. Also among the disadvantages of these models is the

Corresponding author address: **Babak Safa,** Afshin Lane, Vila Lane, Motahhari st., Isfahan 81848, Iran; Email: babaksafa@hotmail.com

explicit and clear description of the mechanism or the effect of each input on crop yield, and further the neglect of the effect of hierarchical structure of the principal phenological stages of each product on the computations. (Monteith, 1977; Katz, 1981; Franch & Thornly, 1984; Toure et al., 1994).

The structure of the second-group models is based on the identification of each physiological stage of the plants and also knowing their dynamic mechanistic and fitting them in mathematical models to describe the crop growth. On these lines, CERES (Crop Estimation through Resource and Environment Synthesis) – WHEAT model is the most important and popular. The need for the great variety of data in all the areas affecting the yield constitutes the essentials of this class.

Recently, the application of Artificial Intelligence (AI) such as Artificial Neural Networks (ANNs), Fuzzy Systems and Genetic Algorithm has shown positive efficiency in such areas. Their application can model the complex natural processes more conveniently and with greater accuracy.

This study attempts to study the possibility of the complex process modeling of prediction of yield using ANNs. If we design a network that correctly encompasses the relations of climatic factors affecting crop yield, it can be used to as a good model estimate the production of wheat and with sufficient useful data test the effect of each area, a good ANNs model can be developed for the area. Furthermore, using ANNs can test the most effective factors on crop yield and then identify and select the most effective ones. Therefore, some factors whose measurements are difficult and costly can be ignored. In this study it must be noticed that the effect of climatic and meteorological factors on wheat yield have been measured and modeled.

2- Methodology:

a) Fundamentals of Artificial Neural Networks (ANNs):

ANNs are a free-model intelligent dynamic system, which can compute empirical data and find out the hidden rule for them, and then make the network structure. Also they learn the relations or rules for them or for other instances. These systems endeavor to make a model that is similar to neuro-synaptic structure of the brain (Menhaj, 1998). A neuron is the smallest computation unit of data and is the basis of ANNs work. Figure (1) shows a neuron with one input. P, a, are the scalar input and output of the network. The degree of effect (p) on (a) is computed by scalar (w). The other input is constant (1). It is multiplied by term of bias (b) and then added to (wp), which will be the net input for transfer function (f). Thus, the output of neuron is expressed by the equation:
$a = f(wp + b)$

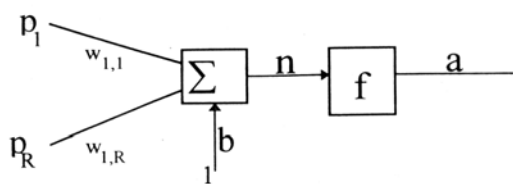The term bias (b) can be taken as weight (w). Since the quantity of constant (1) is reflected by

Figure (1) : Model of single - input neuron

the input of neuron. The parameters (w, b, f) are set by the designer. The transfer function can be used as linear or non-linear. It is selected on the basis of the specified requirement for dissolving a function. Transfer function must also be differentiable. The parameters (w, b) are regulated on the basis of selection of transfer function (f) and the type of learning rule (Algorithm). Learning, i.e. weight and bias (w, b) changes until the input and output relations of each neuron conform to our main purpose. The learning rule is generally expressed by different equations. The goal of learning rule is to train ANNs to perform a specified function. In other words, in the process of training, after each repetition of the learning rule, ANNs are informed further of their environs, conditions and learning type is specified by regulating the trend of the network parameters (Menhaj, 1998).

It must be reminded that even one neuron with many inputs is not sufficient for solving various problems. In such cases we will have a mono-layer or multi-layer network which has been formed from a combination of neurons. In multi-layer networks, each layer has a weight matrix (w), bias vector (b), net input vector (n) and output vector (a) for itself only, and also different layers may be formed by different transfer functions. Each layer whose output is the output of ANNs is named an output layer, and the other layers (except the input layer) are called hidden layers. In the input layer using logistic function and tangential hyperbolic function or other non-linear functions, it has generally been recommended to use the linear function, because it has been proved that using linear function is better than the non-linear function aimed at solving problems with a non-linear trend. The best learning method for multi-layer-perceptron (MLP) is the learning rule of Steepest Descent Back – Propagation organizing two main paths (Figure 2). The first path is called forward path. In this path, the input vector is applied to MLP network and its effects via the hidden layer are distributed onto the output layer. The output vector formed in the output layer is the actual solution to MLP network. In this path, the network parameters are constant. The second path is called the backward path. In this path, unlike the first path, all network parameters change and are regulated. This is accomplished by the error correction rule. Error signal is formed in the output layer. Error vector is equal to the difference between the actual (desired) answer
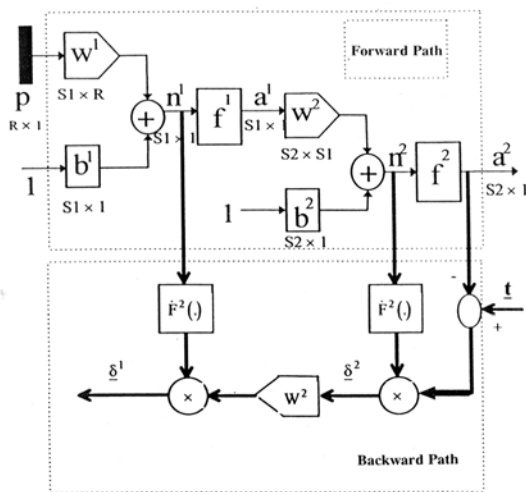
Figure (2):Steepest Descent Back - Propagation

Algoritm at a MLP network with one hidden layer

and the estimate of the answer network. The amount of error, after computation, in the backward path, from the out put layer and via hidden layers is distributed in the entire network. The network parameters are extremely regulated such that, the actual (desired) answer and estimate of answer become similar. Selection of the number of fit neurons for each hidden layer depends on the complexity of function so that whatever the function was getting many of inflection points, the number of neurons in hidden layers must be more. But, it must be reminded that the number of network regulation parameters are fewer than the number of learning data, to the extent that the network does not learn any more than is possible, because, the network instead of searching for mathematics or logical relations in the data, memorizes all of them (Abareshi, 1996). But, the number of fit hidden-layer neurons is generally determined by trial and error (Menhaj, 1998). Steepest Descent Back – Propagation algorithm has various regulation parameters of which the most important is briefly explained:

1) Learning rate:
The learning rate determines the length of scales in each repetition of network parameters optimization, and recommends that the learning rate in hidden layers count more in the output layer (Neural ware, 1993).
If the less the learning rate is selected the changes in the network parameters will be much less after each repetition a case which will help to smooth the movement path of parameters toward the optimum quantities and will slow down learning. Inversely, when the learning rate is increased, although the learning speed is increased too, but great changes are made from one repetition to the next, which occasionally will bring instability in the network generally referred to

as divergent network parameters (Ashrafzadeh, 1999).

2) Momentum:
The amount of inertia that is increased in each of the network parameters is called momentum, so that each one changes in the way that decreases the amount of energy.
Momentum generally is utilized to increase and improve the learning rate and also prevents instability in network (Salehi and et al., 1998).

3) Epoch:
Each input vector which is given to the network at each learning cycle and updating of weight is called epoch. It is suggested that the same number of hidden layer neurons is the best trial and error for epoch selection (Abareshi, 1998). The specifications of ANNs include:
I) Training ability:
Learning ability means the ability to regulate the network parameters (synaptic weights), in the course of time when the networks environment changes, with this purpose in mind that if the network were taught for a specified condition and a small change were made in the environment, it would be efficient for the new conditions with minor training.
II) Dispersion of information:
There is not one-to-one correspondence between the input and the synaptic weights, because each synaptic neuron can be said to be connected to all inputs. In other words, each neuron is affected by a network of the activity of other neurons.
III) Generalization:
After the primary examples have been taught to the network, it can yield a fit output for each untaught input. Put more clearly, the network learns the function, teaches the algorithm and yields an appropriate analytic relation for a number of points in the space.
IV) Parallel Processing:
When ANNs are organized as hardware, each cell of the same level can respond simultaneously to the inputs that resulting in increasing the speed of processing.
V) Robustness:
In ANNs, each cell functions independently, and total activities of the network is resultant of the many local cells activity. This characteristic causes cells to correct each other's local errors in their functioning and increases the robustness of the system (Menhaj, 1998). But, the disadvantages of ANNs can be referred to their architectural complexity which is due to lack of a stable design and standard for solution of various problems and also need to powerful high-speed computer which impose heavy expenses (Abareshi, 1996).

b) Area selection and a brief description of its climate conditions:
There is a lot of difference between the various cultivators for purposes sowing dates, dates of

appearance at each phenological stage, and the amount of yield and so on. Sararood agrometeorological station in Kermanshah Province is the only area of Iran where one variety of wheat has been continuously sowed over the past ten years. In this area, Sardari (white) variety has been sowed since (1988-89), but unfortunately there were not reliable statistics for the years: (1988-89), (1989-90), (1993-94); therefore, using eight years data this research was conducted (from 1990-91 till 1998-99). Although there are few year statistics and if there were a good number of years with data the accuracy of work would be increased; but with the increase in the number of meteorological variables, statistics expanded and the deficiency to some extent decreased. Some characteristics of Sararood dry farming station include:

Latitude: 34°; Longitude: 47°; Altitude of sea surface: 1352 meters, this location is situated between the relatively high Zagros mountains, the soil texture is clay-loamy, minimum, mean and maximum rainfall is 241, 461 and 783 mm respectively, rainfall time ranges from October to June, rarely observed in other months. The mean temperature in January is from (0°c) to (5°c), and also the period of freezing cold ranges from November to April. The best sowing time for dry farming wheat, on the basis of the results of area Ambrotic curve, is in the middle of October, which is about the beginning of rainfall time. About 75 days of growing season (April, May and a few days of June) in some years, dry farming crops need irrigation, because intervals of rainfall are irregular and the soil humidity is much less than the water requirement for wheat (Ezadi, 1995).

c) Selection of input vector elements:

One of the important stages in the planning of ANNs is the selection of the input vector elements. The most important factor for element selection is the physical basis of the system intended for modeling by ANNs. Because our purpose of this study is wheat yield prediction, the input vector elements must be selected by factors affecting it. The most important of these elements are meteorological factors such as: air temperature, wind speed, rainfall quantity, interval rainfall, sun hours, air relative humidity and evapotranspiration. Of course, the effect of radiation factors (SSR, TSR, RSR) is very important too. But, due to lack of correct and complete statistics, it was not included in the input matrix.

d) Structure of data matrix:

To make a data matrix, at first, for each year (8 years) the dates of the beginning and the end of each phenological stage (11 stages) of wheat were collected: sowing, germination, emergence, third leaves, tillering, stem formation, heading, flowering, milk maturity, wax maturity, full maturity. For each stage in each year meteorological factors were selected:

Table 1: Matrix Figure For any Stage

Years \ Parameters	T mean (min)(°c)	T abs (min)(°c)	T mean (max)(°c)	T abs (max)(°c)	P (total) (mm)	P (day)	Sun (hr)	FF (mean) (m/s)	FF abs (max) (m/s)	%RH (mean)	W.R. (total) (mm)	Y. (Kg/ha)
90/91												
91/92												
92/93												
94/95												
95/96												
96/97												
97/98												
98/99												

4

Tmean (min) = mean daily minimum temperature (for each stage in each year)
Tabs (min) = extreme daily minimum temperature (for each stage in each year)
Tmean (max) = mean daily maximum temperature (for each stage in each year)
Tabs (max) = extreme daily maximum temperature (for each stage in each year)
P (total) = sum of daily rainfall (for each stage in each year)
P (day) = number of rainfall days (for each stage in each year)
Sun hours = sum of sun hours (for each stage in each year)
FF (mean) = mean daily wind speed (for each stage in each year)
FFabs (max) = extreme daily wind speed (for each stage in each year)
%RH (mean) = mean daily relative humidity (for each stage in each year)
W.R. (total) = sum of daily water requirement (for each stage in each year).
W.R. has been obtained with FAO Penman–Montieth evapotranspiration at FAO technical paper in 1998). The reason for selection of this factor was the observation of reciprocal effects for other elements (net elements) the synthetic form.

Y. = The amount of wheat yield at the end of the growing season, because this amount for all the stages in a year is constant, so it is repeated at the end of each stage matrix. Finally, data matrix was made the same as that in Table (1).

e) Presentation of data to the network and making input files:
After preparing the data matrix, at each wheat phenological stage, the two year-data (%20) was used for making the test file and the six year-data (%80) was used for learning file. These files (test and learn) were made for two cases:
In the first case, all the stages were set in rows (Table 2). Each file contains a matrix whose rows were repetitions of the years at each wheat phenological stage and its columns were meteorological factors. Totally eleven training files and eleven test files were made. But, these data structure during training did not bring a good and correct result, as with the selection of different primary amounts for each of network parameters, even, minimum amounts, the network quickly became divergent. In other words, the network was saturated. Thus, this type of data structure was put by.

File1	File2		File9	File10	File11
Stage 1 + column of Yield	Stage 1 + column of Yield	….	Stage1+ column of Yield	Stage 1 + column of Yield	Stage 1 + column of Yield
	Stage 2 + column of Yield		Stage 2 + column of Yield	Stage 2 + column of Yield	Stage 2 + column of Yield
			Stage 3 + column of Yield	Stage 3 + column of Yield	Stage 3 + column of Yield
			Stage 4 + column of Yield	Stage 4 + column of Yield	Stage 4 + column of Yield
			Stage 5 + column of Yield	Stage 5 + column of Yield	Stage 5 + column of Yield
			Stage 6 + column of Yield	Stage 6 + column of Yield	Stage 6 + column of Yield
			Stage 7 + column of Yield	Stage 7 + column of Yield	Stage 7 + column of Yield
			Stage 8 + column of Yield	Stage 8 + column of Yield	Stage 8 + column of Yield
			Stage 9 + column of Yield	Stage 9 + column of Yield	Stage 9 + column of Yield
				Stage 10 + column of Yield	Stage 10 + column of Yield
					Stage 11 + column of Yield

Table 2: Method (1)

In the second case, all the stages were set in columns (Table 3). Each file contains a matrix whose rows were years and whose columns were meteorological factors at each phenological stage. But, the main disadvantages of this case was the fact as the number of neurons into the input layer on the basis of selection of the file type, regularly changed, planning of network and its parameters were different from file to file and this required that each training file designed a network with a special identification and work, application of model made difficulties for other areas.

Table 1 in small Form

File 1:	Stage1 + column of Yield

File 2:	Stage 1	Stage2 + column of Yield

File 3:	Stage 1	Stage 2	Stage 3 + column of Yield

.
.
.

File 9:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9 + column of Yield

File 10:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9	Stage10+ column of Yield

File 11:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9	Stage 10	Stage11 + column of Yield

Table 3: Method (2)

Thus, to solve these problems, instead of the data matrix for each stage which was omitted in the input files (except the eleventh file) a matrix with a constant number one (1) was set (Table 4). Now instead of planning the various networks, only one network was designed or in other words, it was made a constant structure for different files and also the accuracy of model estimation improved too. But, for two reasons network planning for yield prediction was accomplished up to the end of stem formation stage:

1. Movement to the primary stage was not necessary because, it created long time distance till the end of full maturity stage and many effective stages on crop yield such as heading and flowering did not develop yet.

2. At the beginning of stem formation stage, the amount of error (e) estimation increased rapidly so that, even with minimum amounts no appropriate answer was obtained.

f) Identification software and planned network:

In this study two boxes of software were used for network planning:

MATLAB   Ver. 5.3

Neural works professional II/plus   Ver. 5.23

Two hidden layers were used for network planning and the learning rule was Normalize Cumulative Delta (N.C.D.) which is similar to the Delta rule. In this rule, weights change and their updates are stored at the end of each epoch and the learning rate is independent of the epoch size. Transfer function used in hidden layers was logistic function whose amounts range between [0, 1]. But transfer functions at the input and output layers were selected linear functions. The best answer was obtained when the epoch size was (1) for all files.

File										
File 1:	Stage 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1 + column of Yield
File 2:	Stage 1	Stage 2	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1	Matrix 1 + column of Yield

.
.
.

File										
File 9:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9	Matrix 1	Matrix 1 + column of Yield
File 10:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9	Stage 10	Matrix 1 + column of Yield
File 11:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9	Stage 10	Stage 11 + column of Yield

Table 4: Method (2) with Matrix 1

The input data were fixed between [-1, 1] that are idiomatically called Bipolar. Meanwhile, the primary weight amounts were selected to be between [-0.1, +0.1], the amount of weight set at the first hidden layer is in the range of [-0.91, +0.91], at the second hidden layer in the range of [-0.141, +0.141] and in the output layer in the range of [-0.707, +0.707]. Also to achieve optimum values, the number of hidden layer neurons and primary amounts of momentum and learning rate were obtained by trial and error. The learning rate at the second hidden layer was set (0.2) and also the repetition number of network training for all cases was 100,000.

g) Model evaluation method:
Apart from the usual quantitative measures used at the evaluating performance of a model, our study used the Root Means Square Error (RMSE) in which the accuracy of model was evaluated on the basis of the difference between the actual and estimates. RMSE was computed by this equation:

$$RMSE = \sqrt{\frac{\sum (P_i - O_i)^2}{n}}$$

"n" is number of observations.

3- Results and discussion:

a) Results of the comparison of the second case using a matrix (1):
Table(5) shows that the learning rate parallel with the diminishing processing elements (PE) in out put layer is gradually decreasing from the end of full maturity stage to the beginning of the heading stage. The main reason for it can be the data disband.

Stages	#PE in Input Layer	#PE in Hidden 1 Layer	#PE in Hidden 2 Layer	#PE in Output Layer	Learning Coefficient Hidden 1 & output	Momentum	RMSE	Estimate Of Yield	Actual Yield	e max
Full Maturity	121	50	2	1	0.080	0.700	0.0000	2197	2192	5
								2702	2706	
Wax Maturity	110	73	2	1	0.005	0.655	From 0.0002 To 0.0001	2179	2192	17
								2723	2706	
Milk Maturity	99	14	2	1	0.003	0.755	From 0.0007 To 0.0001	2224	2192	34
								2740	2706	
Flowering	88	16	2	1	0.002	0.100	From 0.009 To 0.0003	2138	2192	57
								2763	2706	
Heading	77	14	2	1	0.002	0.777	From 0.0030 To 0.0001	2156	2192	36
								2741	2706	
Stem Formation	66	50	2	1	0.001	0.100	From 0.0090 To 0.0010	2147	2192	45
								2729	2706	

$e = |$ (Estimate of Yield) – (Actual Yield)$|$

Table 5: Results of Method (2)

Stages	#PE in Input Layer	#PE in Hidden 1 Layer	#PE in Hidden 2 Layer	#PE in Output Layer	Learning Coefficient Hidden 1 & output	Momentum	RMSE	Estimate Of Yield	Actual Yield	e max
Full Maturity	121	50	2	1	0.080	0.700	0.0000	2197	2192	5
								2702	2706	
Wax Maturity	121	50	2	1	0.017	0.200	0.0000	2198	2192	6
								2704	2706	
Milk Maturity	121	50	2	1	0.004	0.555	From 0.0002 To 0.0001	2191	2192	1
								2707	2706	
Flowering	121	50	2	1	0.004	0.600	From 0.0030 To 0.0001	2172	2192	20
								2687	2706	
Heading	121	50	2	1	0.003	0.200	From 0.0050 To 0.0001	2173	2192	19
								2719	2706	
Stem Formation	121	50	2	1	0.001	0.200	From 0.0050 To 0.0001	2144	2192	54
								2759	2706	

$e = |$ (Estimate of Yield) – (Actual Yield)$|$

Table 6: Results of Method (2) with Matrix1

**Table 7: Sensitivity Analysis To End Of Full Maturity Stage**

Stages	Sowing	Germination	Emergence	Third Leaves	Tillering	Stem Formation	Heading	Flowering	Milk Maturity	Wax Maturity	Full Maturity
W.R.(total)	-0.207871	0.357211	0.419587	-0.833988	0.755191	0.445813	-0.680268	-0.154376	-0.020206	0.113904	-0.315189
%RH(mean)	0.281781	0.708669	0.178546	-0.954807	0.58648	0.226736	0.381738	0.746727	0.149876	0.093222	0.14323
FFabs(max)	0.112057	0.137448	-0.279725	0.231802	0.587583	-0.272334	0.019282	0.013888	0.203669	0.180006	-0.5714
FF(mean)	0.270396	0.745654	-0.093073	0.228077	-0.104785	-0.030488	-0.900269	-0.154287	0.516802	0.113726	0.386775
Sun(hr)	-0.515014	-0.454247	-0.894785	0.430971	-0.011474	-0.053495	0.967294	0.412434	0.241607	-0.162512	0.25323
P(day)	-0.103116	0.100166	0.156552	-0.604898	0.718147	0.333697	0.670463	0.529736	-0.167012	0.381857	0.066102
P(total)	-0.487179	0.184625	0.140876	-0.530809	0.846118	0.685573	0.905722	0.958174	-0.260621	-0.205815	-0.435889
Tabs(max)	-0.596017	-0.8156	-0.364542	0.561267	-0.488997	0.220269	0.157028	0.01052	-0.605226	-0.246793	-0.025272
Tmean(max)	-0.200421	-0.306755	-0.022411	0.39807	-0.727594	-0.223319	-0.562608	-0.158697	-0.617802	-0.070095	0.222892
Tabs(min)	-0.157386	0.484765	0.606447	0.171244	-0.640243	-0.663757	-0.641167	-0.249028	0.343531	-0.105947	-0.109375
Tmean(min)	0.158668	0.51862	0.371784	-0.041515	-0.394851	0.0446644	-0.369132	0.184476	-0.151217	-0.020206	-0.723839

Because, with the omission of each stage, the noise increases, this problem can be solved, by decreasing the learning rate, diminishes the length of scales during movement on the function curve. But, this is not observed as a constant trend in the case of the number of hidden layers neurons and specially the momentum amount. On the other hand, RMSE and (e) are continuously increasing from the end of full maturity stage till the beginning of the heading stage. Only, one abnormal result was seen at the flowering stage in comparison with the other stages that can be explained in two cases:

Disband and noise in trend of data arrangement and their relations with each other at this stage is more than the other stages.

Confluence of the occurrence days at this stage in comparison with those of the preceding and following stages occurs earlier than other wheat phenological stages; which itself supports the first case.

However, the results of using matrix (1) instead of data omission were seen at table (6). Here, the same as it was before, from the beginning of full maturity stage till the beginning of the heading stage, the index of learning rate and RMSE is gradually increased, and the momentum index also varies and does not follow a stable trend. On the basis of Figure (3) the difference between the actual and estimate yield is constantly ascending which on the basis of the evidence for the decrease in the index of learning rate is a natural occurrence. But, as mentioned earlier, the first advantage of this case is stabilization of the number of neurons at each layer and the stabilization of the network structure from each stage to the other stage. The second advantage is the significant error (e) decrease for all wheat phonological stages indicated in the two methods. Meanwhile, Figure (4) indicates the more normal pattern of the error changes as compared to the first method.


Figure (3): Comparison of Estimate of Yield & Actual Yield (Method 2 with Matrix 1)
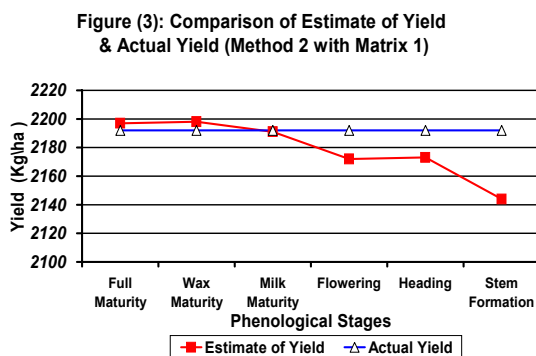
Figure (4): Comparison of emax (Method 2) & (Method 2 with Matrix1)
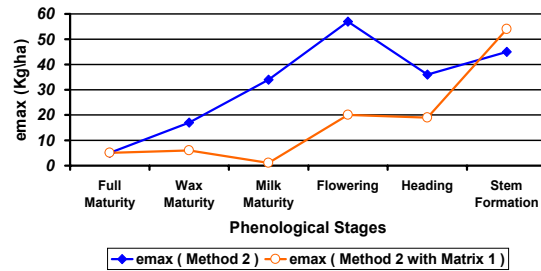
b) Distribution of actual data and estimate of wheat yield:

Figure (5) shows the histogram and normal curve of the wheat yield for eight years, in Sararood of Kermanshah. As is observed the curve of data distribution is right skewed and the data frequency is more striking in the environs of lower yield. Accordingly, if the two years' data used to test model are selected with in the mean of data in an area with the most data frequency; the accuracy of model estimation increases and farther away we go from the mean point and move to the two curve angles, specially to the right where maximum skew is seen. But, the error of yield estimation increases because:

Figure (5): Histogram with Normal Curve (YIELD)

Yield (Kg\ha)
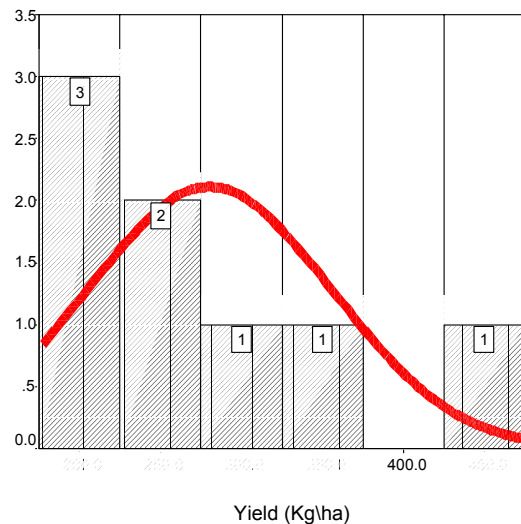
1. Since training in ANNs specially, those made in this project are based on the data quality and quantity, the more the data accessible, the learning ability increases in the model and gets closer to reality. So if the data bordering extreme amounts are omitted from the training file and are transferred to the test file, the ability to estimate the model rapidly increases. In fact, we cannot

expect an accurate and exact beyond what we have taught.

2. On the basis of what was referred to above, the greater the frequency of the data and the more symmetrical and even their distribution around the mean the lighter the extreme amounts and the less than accuracy of the responsive models would be less expected to change much. Therefore, the greater the number of the statistical years available, obviously the change in error for the statistical years would approach zero. In this study the results mentioned, the first two years would make the test file and the next six years would make the learning file.

c) The rate of sensitivity for each phonological stage with respect to meteorological factors:

On the basis of Table (7) the changes in sensitivity of the phenological stages for wheat at the end of maturity stages has been shown for each meteorological factor.

As is seen the greatest sensitivity relative to Tmean(min) are germination, emergence and flowering stages which are most effective in the yield and growth of wheat.

The Maximum sensitivity index with respect to Tabs(min) appears in Sararood area at germination, emergence, flowering and milk maturity stages and the minimum sensitivity starts with the tillering stage to the beginning of flowering. So complete appearing of tillering stage up to before winter's drowsiness is very important for crop production.

Maximum wheat sensitivity with respect to Tmean(max) definitely occurs at each stage of developing into third leaves, wax maturity and full maturity stages and relatively at stem formation and flowering stages.

The greatest reaction to Tabs(max) occurs changing into third leaves and stem formation stage. Meanwhile, during the last growing stages the sensitivity rate is gradually increased over this factor.

The most important effective meteorological factors on crop yield, is the quantity of rainfall which makes the greatest contribution to growth, but the most sensitive phonological stages relative to this stage are flowering and heading afterwards, are the primary stages of planting seeds. Therefore, the rainfall quantity after sowing the seeds and also during the first two months of spring is very important to the crop production (Figure 6).

Tillering, heading and flowering are the most sensitive stages in wheat in terms of time and distribution of rainfall in Sararood.

Sun hour is absolutely important at the heading stage and relatively important at third leaves stage. After the heading stage the sensitivity is gradually decreased, each stage with respect to it. But at two periods of time, one after emergence stage and at the beginning of heading stage, the ascending pattern of the effect of this factor on the growth of plant is very striking (Figure 7).

Also, the greatest sensitivity index relative to FF(mean) are the speed of wind first at tillering and at the formation of third leaves, germination and milk maturity.

Figure (6): Sensitivity Analaysis for P(total)
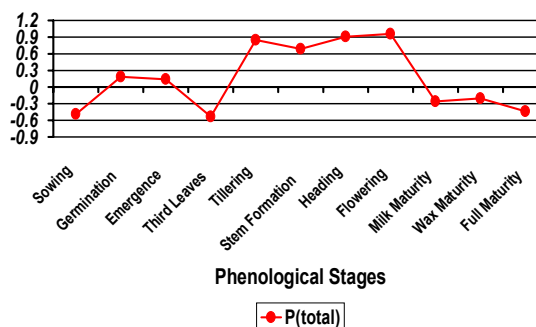
Maximum wheat sensitivity relative to FFabs occurs at first at heading stage and then in arrangement in third leaves, milk maturity and wax maturity stages.

Figure (7): Sensitivity Analysis for Sun(hr)
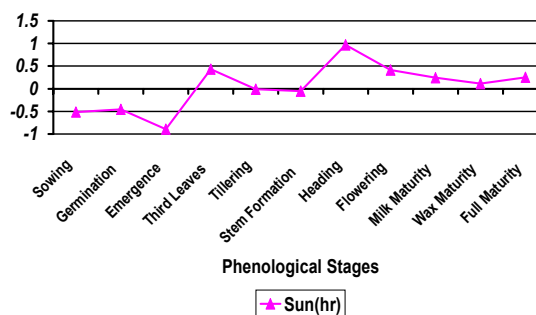
Flowering, germination and heading are the most sensitive stages in wheat relative to the mean relative humidity %RH(mean).

Two maximal points in tillering and emergence stages and two minimal points in chaining into third leaves and heading are the answers to the sensitivity analysis relative to W.R. (total) in Sararood area.

Of course, it is important that the sensitivity analysis results are specific to the Sararood area in kermanshah, with its particular climate. Therefore, it can be generalized to the other regions for wheat plantation although there might be similarities among various regions.

d) The intensity of the effect of each meteorological factor on yield:

Since the smallest matrix size for data speeds up both when collecting the data, designing and training attempts, though qualitative were made to achieve a classification due to the significance and role of each element during the growth stage and

also due to the yield in the respective region. Hence, the elements were divided into four classes in order of their importance:

Class I: P(total), Tmean(max), P(day)

Class II: Sun(hr), %RH(mean)

Class III: W.R.(total), FF(mean), FFabs(mean), Tmean(min)

Class IV: Tabs(min), Tabs(max)

Looking quickly at the four classes listed above it can be concluded that P(total) in the region under study, the pattern of rain distribution ,P(day), and the Tmean(max) particularly at mid and terminal stages are specially important. Further, temperature extremes have a slight or insignificant effect on the yield.

Conclusions:

The results of this study are summarized as follows:

1. Presentation of a new method for wheat yield prediction.

2. The high speed, accuracy and efficiency of this method in comparison with the other methods with respect to the low bulk of the data for the region.

3. Prediction of wheat yield with maximum error (45-60 kg/ha) at least two months before crop ripening (end of stem formation stage).

4. Achieving the sensitivity for each phenological stage with respect to particular meteorological factors that helps to understand the effect of decreasing and increasing factors at each stage. Thus, enabling us to minimize and control the harmful effect of each factor at different stages.

5. Recognizing the order of priority and importance of each meteorological factor, in the plant growth and yield, enabling us to eliminate a number of the elements in the data matrix in order to speed up and facilitate the stages for the preparation of the input data and to speed up network responses without affecting the estimates and accuracy significantly.

References:

1. Abareshi, B., 1996; Sensible heat flux estimation over a prairie grassland by neural networks, M.S. thesis, 114 pp., McGill university, Montreal.

2. Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998 (E); Crop evapotranspiration-Guidelines for computing crop water requirement. Food and Agriculture Organization of the United Nations, Rome, 300 pp.

3. Ashrafzadeh, A., 1999; Application of artificial neural networks for prediction of evaporation from evaporative ponds, M.S. thesis, 131 pp., University of Tehran, Tehran.(in Persian)

4. Burks, T.F., Shearer, S.A., Sobolik, C.J., Fulton, J.P., 2000; Yield prediction using a neural network classifier trained using soil landscape features and soil fertility data, *ASAE paper No. 001084*, Annual International Meeting, Midwest Express Center, Milwaukee, Wisconsin, July 9-12, 10 pp.

5. Chipanshi, A.C., Ripley, E.A., Lawford, R.G., 1997; Early prediction of spring wheat yields in Saskatchewan from current and historical weather data using the CERES-WHEAT model, *Agriculture and Forest Meteorology* **84**(1997): 223 - 232.

6. Dehghan, A., 1998; Aggregation levels in yield forecasting system experience of Hamadan irrigated wheat variety Navid, M.S. thesis, 74 pp., Wageningen Agricultural University, Wageningen.

7. Ezadi, M., 1995; Agroclimatic studying of dry farming lands in Kermanshah province, Iranian Meteorological Organization Publishers, 50 pp. (in Persian)

8. Harrison, P.A., Porter, J.R., Downing, T.E., 2000; Scaling up the AFRCWHEAT2 model to assess phenological development for wheat in Europe. *Agriculture and Forest Meteorology* **101**(2000): 167-186.

9. Landau, S., Mitchel, R.A.C., Barnett, V., Colls, J.J., Craigon, J., Payne, R.W., 2000; A parsimonious multiple – regression model of wheat yield response to environment, *Agriculture and Forest Meteorology* **101**(2000): |5| - 166.

10. Malek, E., 1993; Knowledge and measurement of climatic factors affecting agriculture, First Edition, Shiraz University Publishers, 239 pp. (in Persian)

11. Menhaj, M.B., 1998; Fundamentals of neural networks, First Edition, Professor Hesabi Publishers, 502 pp. (in Persian)

12. Menhaj, M.B., Seifipour, N., 1999; Application of computational intelligence in control, First Edition, Professor Hesabi Publishers, 236pp. (in Persian)

13. Salehi, F., Lacroix, R., Wade, K.M., 1998; Effects of learning parameters and data presentation on the performance of Back-Propagation networks for milk yield prediction. *Transaction of the ASAE*. **41**(1):253 - 259.

14. Teshnehlab, M., Watanabe, K., 1999; Intelligent control based on lexible neural

networks, First Edition, Kluwer Academic Publishers, 235 pp.

15-WWW.neuralware.com/FrequentlyAsked Questions.html