J6.4

Artificial Neural Networks (ANNs) Application for the Occurrence Date Prediction of Each Phenological Stage in Wheat Using Climatic Data

Babak Safa*
Sepahan Kara Company, Isfahan, Iran
A. Liaghat
Department of Irrigation and Reclamation Engineering, College of Agriculture, University of Tehran, Karaj, Iran

Abstract:
Prediction of the time of crop phenological stages occurrence mainly strategic plants such as, wheat, corn and rice help us to achieve to the exact time in order to control pests, weeds and pathogens and also, to the best time for operating such as, fertilizing and irrigation.
The main purpose of such study is to estimate the occurrence of phenological stages in dry farming wheat at the time interval of short duration before their occurrence 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. 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 four meteorological factors in period (1990-99) including degree days (heat units), total daily rainfall, sum of sun hours and sum of water requirement for each of eleven phenological stages in wheat including sowing, germination, emergence, third leaves, tillering, stem formation, heading, flowering, milk maturity, wax maturity and full maturity 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 are each of the statistical years (i) and the columns are meteorological factors (k) at each phenological stage (j). In fact, statistical years (i), phenological stages (j) and meteorological factors (k) are the basic elements of 3-D matrix (M ijk) arranged as above. 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 network training computations and 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 which obtained has the following capabilities:

1. Prediction of the date of phenological stages occurrence (From stem formation till full maturity) with maximum error 3 to 6 days at least five days before the occurrence of each stage.
2. Determination of the sensitivity of each phenological stage with respect to meteorological factors.

1- Introduction:

The time of crop phenological stages occurrence, 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, and evapotranspiration are very important. Therefore, it is not very unlikely to achieve relations or systems that can predict the occurrence date of each phonological stage with higher accuracy using meteorological data.
Recently, the application of Artificial Intelligence (AI) such as Artificial Neural Networks (ANNs), Fuzzy Systems 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 the occurrence date prediction of each phonological stage in wheat using ANNs. If we design a network that correctly encompasses the relations of climatic factors affecting on development, it can be used to as a good model estimate the occurrence date prediction of each phonological stages.

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)

* *Corresponding author address:* **Babak Safa**, No. 1/58, Afshin Lane, Vila Lane, Motahhari St., Isfahan 81848, Iran;
Email: babaksafa@hotmail.com

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

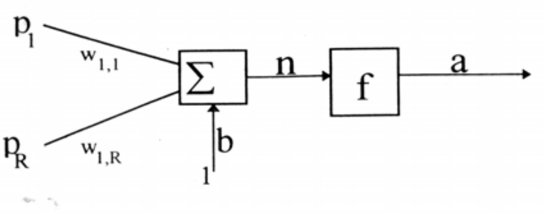


Figure (1) : Model of single - input neuron

number of neurons. 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

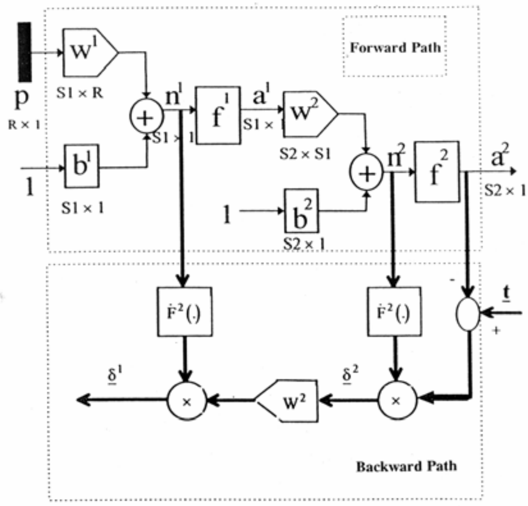


Figure (2):Steepest Descent Back - Propagation

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 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. 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 the occurrence date of each phonological stages in wheat, the input vector elements must be selected by factors affecting it. The most important of these elements are meteorological factors such as: air temperature, rainfall quantity, interval rainfall, sun hours, 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, these meteorological factors were selected:

Parameters / Years	*D.D.* (°c)	*P.* (mm)	*Sun* (hr)	*W.R.* (mm)	*N.D.* (Day)
1990/91					
1991/92					
1992/93					
1994/95					
1995/96					
1996/97					
1997/98					
1998/99					

Table 1: Matrix Figure for Any Stage

D.D. = degree day (heat unit) accumulations (from the sowing date to the end of each phonological stage)
P. = sum of daily rainfall (from the sowing date to the end of each phonological stage).
Sun hours = sum of sun hours (from the sowing date to the end of each phonological stage).
W.R. = sum of daily water requirement (from the sowing date to the end of each phonological stage). W.R. has been obtained with FAO Penman–Montieth evapotranspiration at FAO technical paper in 1998.
N.D. = number of days (from the sowing date to the end of each phonological stage). Finally, data matrix was made the same as that in Table (1).

e) Presentation of data to 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 as shows in (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.
In the second case, all the stages were set as shows in (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. But, the reason why network planning for the date occurrence prediction was accomplished up to the end of stem formation stage was that which 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 and sensitive against pests and pathology such as heading and flowering have not developed yet.

File1	File2	…	File9	File10	File11
Stage 1	Stage 1		Stage1	Stage 1	Stage 1
	Stage 2		Stage 2	Stage 2	Stage 2
			Stage 3	Stage 3	Stage 3
			Stage 4	Stage 4	Stage 4
			Stage 5	Stage 5	Stage 5
			Stage 6	Stage 6	Stage 6
			Stage 7	Stage 7	Stage 7
			Stage 8	Stage 8	Stage 8
			Stage 9	Stage 9	Stage 9
				Stage 10	Stage 10
					Stage 11

Table 2: Method (1)

Table (1) in small Form

File 1:	Stage1 + N.D. *(Day)*										
File 2:	Stage 1	Stage2 + N.D. *(Day)*									
File 3:	Stage 1	Stage 2	Stage 3 + N.D. *(Day)*								
.											
.											
.											
File 9:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9 + N.D. *(Day)*		
File 10:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9	Stage10+ N.D. *(Day)*	
File 11:	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9	Stage 10	Stage11+ N.D. *(Day)*

Table 3: Method (2)

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 tangent hyperbolic 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.

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 20,000.

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:

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

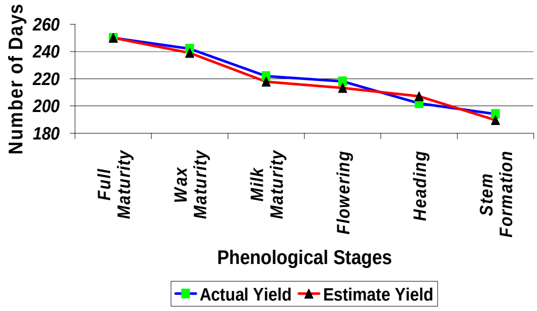
“n” is number of observations.

3- Results and discussion:

a) The results of second case:

As has been showed in table (4), no constant trend whatsoever at no one of the network parameter such as learning coefficient, Momentum and even the number of PEs(processing elements) at each of phonological stage is observed. But, the amount of RMSE in two stages (milk maturity & heading) is not zero; because of confluence of the occurrence days at these stages in comparison with those of the following stages occurs earlier than other wheat phenological stages. Also, on the basis of figure (1) by moving to the primary stage for the reason that disband and noise in trend of data arrangement and their relation with each other the emax values continuously is increasing.

Figure (1): Comparison of Actual & Estimate Values of the Occurrence Date at each Phenological Stages



Meanwhile 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. 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.

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

On the basis of Table (5) the changes in sensitivity of the phenological stages for wheat at the end of full maturity stage has been shown for each meteorological factor.

As is seen the greatest sensitivity relative to (D.D.) are full maturity, third leaves and then flowering stages which are most effective stage in the yield and growth in wheat. (Figure 2)

The Maximum sensitivity index with respect to (P.) appears in Sararood area at sowing and emergence stages. But, on the other hand, from heading to wax maturity stage an incredible trend in sensitivity index is observed. (Figure 3)

Maximum sensitivity for length of stages in wheat with respect to sun hour occurs in primary stages such as, sowing, germination and emergence and then in tillering and milk maturity stages. (Figure 4)

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	44	30	2	1	0.005	0.700	0.0000	250	250.04	**3**
								242	245.18	
Wax Maturity	40	40	2	1	0.001	0.800	0.0000	242	239.00	**3**
								232	233.20	
Milk Maturity	36	30	2	1	0.005	0.755	From 0.0004 To 0.0001	229	230.18	**5**
								222	217.91	
Flowering	32	24	2	1	0.008	0.800	0.0000	218	213.31	**4**
								200	203.61	
Heading	28	10	2	1	0.005	0.700	From 0.0007 To 0.0001	206	203.57	**5**
								202	207.15	
Stem Formation	24	14	2	1	0.002	0.755	0.0000	194	189.12	**6**
								175	175.34	

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

Table 4: Results of Method (2)

Parameters	D.D. (°c)	P. (mm)	Sun (hr)	W.R. (mm)
Sowing	-0.01565	1.697689	1.029074	-0.52065
Germination	0.502616	0.657439	1.080781	-0.56446
Emergence	0.050664	1.94639	0.749677	-0.56163
Third Leaves	1.070052	1.003146	-0.17077	-0.20877
Tillering	0.159889	0.096411	0.40248	-0.5126
Stem Formation	0.458658	-0.02548	-0.15259	-0.0754
Heading	0.075698	0.246912	-0.20966	-0.27582
Flowering	0.719428	0.37685	-0.73597	-0.60797
Milk Maturity	0.166297	0.431091	-0.05916	-0.75638
Wax Maturity	0.153184	0.707954	-0.46775	-0.2569
Full Maturity	1.36286	0.210106	-1.48147	1.341999

Table 5: Sensitivity Analysis to End of Full Maturity Stage

Figure (2): Sensitivity Analysis for D.D.(°c)

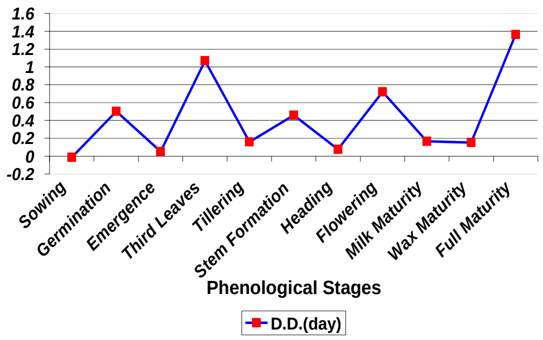


Figure (3): Sensitivity Analysis for P.(mm)

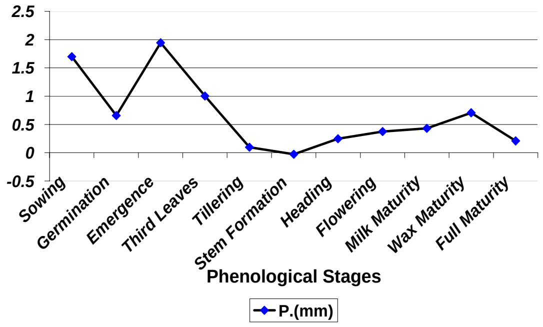


Figure (4): Sensitivity Analysis for Sun(hr)

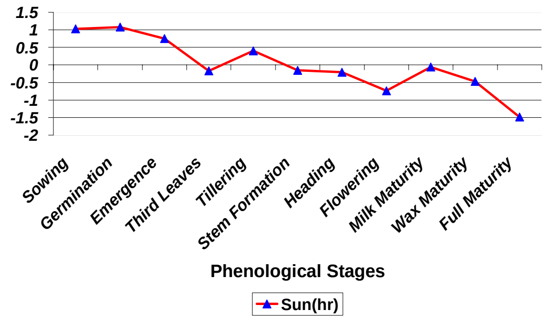
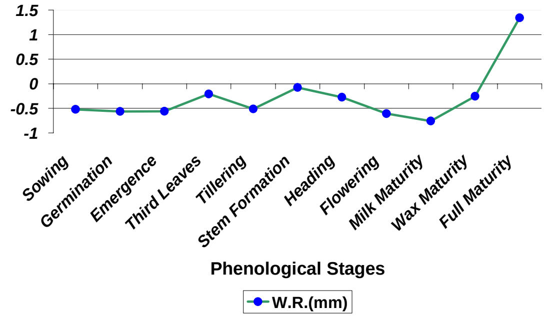


Figure (5): Sensitivity Analysis for W.R.(mm)



Full maturity stage has the greatest reaction to (W.R.). Also, during the middle growing stages the sensitivity rate is to some extent great over the factor. (Figure 5)

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

Conclusions:

The results of this study are summarized as follows:

1. 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.
2. The occurrence date prediction of each phonological stage in wheat with maximum error (3-6days) at least five days before observing of the following stage (From full maturity to stem formation stage).
3. 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.

Acknowledgments:

I wish to acknowledge the support provided by the Iranian Meteorological Organization in preparing and making available the data and statistics for this study and also to acknowledge the technical and scientific contributions made by Tehran university.

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 & 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. Elizondo, D.A., McClendon, R.W., Hoogenboom, G., 1994; Neural Network Model for Predicting Flowering and Physiological Maturity of Soybean. *Transaction of the ASAE* **37**(3): 981-988.

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

6. 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.

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

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

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

10. 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.

11. Snyder, R.L., 1985; Hand Calculating Degree Days. *Agriculture and Forest Meteorology* **35**(1985): 353-358.

12. Teshnehlab, M., Watanabe, K., 1999; Intelligent control based on lexible neural networks, First Edition, Kluwer Academic Publishers, 235 pp.

13.WWW.neuralware.com/FrequentlyAsked Questions.html