

APPLICATION OF BACKPROPAGATION NEURAL NETWORK TO ESTIMATE EVAPOTRANSPIRATION FOR CHIINAN IRRIGATED AREA, TAIWAN

Sheng-Feng Kuo¹
Ming-Hua Tsai²
Wei-Taw Lin³
Yi-Fong Ho⁴

ABSTRACT

Backpropagation Neural Network is applied to establish the relationship between meteorological factors and evapotranspiration, which is then used to predict the evapotranspiration in ChiaNan irrigated area, Taiwan. It takes the weather data from Irrigation Experiment Station of ChiaNan Irrigation Association as the input layer, which include the following weather factors: (1) the highest temperature; (2) the lowest temperature; (3) average temperature; (4) relative humidity; (5) wind speed; (6) sunlight hours; (7) solar radiation amount; (8) dew point; (9) forenoon ground temperature; (10) afternoon ground temperature.

From the result it can be known that the correlation coefficient reaches 0.993 between the evapotranspiration in 2004 calculated by FAO56 Penman–Monteith method and the one predicted by the neural network model with a hidden layer of 10 nodes. The actual evapotranspiration is 911.6cm and the prediction by the neural network is 864.4, between which the error ratio is 1.67%. The correlation coefficient is 0.708 between the actual evaporation in 2004 and the prediction by the neural network with a hidden layer of 10 nodes and an output layer with the pan evaporation as its target output. The pan evaporation is 1674.1cm, while the prediction by the neural network is 1451.7cm, between which the error ratio is 13.23%.

INTRODUCTION

Evapotranspiration refers to the amount of water needed for the normal growth of the crop and becomes the most basic data for the irrigation association to study out annual irrigation plan and estimate the water use amount for agriculture. According to the standard method recommended by Food and Agriculture Organization (FAO), the indirect estimate on the crop evapotranspiration can be divided into two steps: (1) to roughly estimate the potential evapotranspiration according to agricultural weather data and empirical formula; (2) to get the value of the evapotranspiration needed during the crop growth by multiplying the potential evapotranspiration by the crop factor of different crop during different growth phases.

¹ Professor, Dept. of Resource Environment, Leader University, Tainan 709, Taiwan. Member, Asian Regional Working Group, ICID; kuosf@mail.leader.edu.tw

² Director, Irrigation and Engineering Dept., Council of Agriculture, Taipei, Taiwan. Vice Chairman, Chinese Taipei Committee, ICID; mhtsai@mail.coa.gov.tw

³ Section Chief, Irrigation Management Sec., Council of Agriculture, Taipei, Taiwan. Deputy Secretary General, Chinese Taipei Committee, ICID; lwt@mail.coa.gov.tw

⁴ Specialist, Irrigation Management Sec., Council of Agriculture, Taiwan. Graduate Student, Dept. of Civil Engineering, National Taiwan University, Taipei, Taiwan; yifong@mail.coa.gov.tw

In recent years, several articles on the estimate of crop evapotranspiration have been published. Irmak et al. (2003a) study the important topic of how to efficiently apply Florida agricultural water resource to solve the increasingly rising demand on water resource, and therefore, evaluate 21 types of formula, with grass and alfalfa as norm, to estimate the crop evapotranspiration for further agricultural water use management. Irmak et al. (2003b) also adopts the empirical formula with net radiation as base to estimate the potential evapotranspiration in humid area. As known from above, since 1940, lots of empirical formulas have been developed internationally to estimate the non-linear relationship between potential evapotranspiration and crop weather data. Current empirical formulas can be classified into four types: (1) Blaney- Criddle Method; (2) Radiation Method; (3) FAO56 Penman–Monteith Method; (4) Pan Evaporation Method. The agricultural weather data they need include: (1) the highest and lowest temperatures; (2) wind speed; (3) sunlight hours; (4) relative humidity; (5) rainfall; (6) solar radiation, etc.

In recent years, artificial neural network has been widely applied to understand the non-linear issue of water resource and agricultural management. Yang et al. (1997) applies ANN to simulate the non-linear relationship between the ground temperature 10^{cm} , 50^{cm} and 150^{cm} below and the agricultural weather, among which the used weather data include rainfall, potential evaporation, maximum and minimum temperature. Han & Felker (1997) adopts ANN to estimate relationship among evapotranspiration and relative humidity of surface soil, wind speed and soil moisture content, and then compares the results with multiple linear regressions. Burks et al. (2000) applies backpropagation neural network to the comparison among the plant species. Liu et al. (2001) also uses the combination of ANN and genetic algorithm to estimate the corn yield, with the input factors: temperature, rainfall, soil texture and soil PH value. Drummond (2003) also applies ANN and multiple linear regressions to compare the forecast of corn and soybean. Kuo and Liu (2000a, 2000b) uses multi-variables factor analysis and backpropagation neural network respectively to analyze the groundwater quality change of Yun-lin area in Taiwan. The results show that such two factors as seawater salination and arsenic pollution represents 78% of the influence of all 13 groundwater quality items. Chang et al. (2000) adopts fuzzy ANN to predict the flow rate at the upstream Song-mao stream measurement station of Te-chi Reservoir to offer reference for the instant operation at Te-chi Reservoir.

In Taiwan, the agricultural water use occupies the most part of whole water resource, so the accurate estimation of crop evapotranspiration helps irrigation associations to efficiently manage the limited water resource. Different from the traditional method which uses empirical formula to estimate crop evaporation, this study applies Matlab software to establish backpropagation neural network models to analyze the non-linear relationship between crop evapotranspiration and agricultural weather factors, according to which then crop evapotranspiration at different period can be predicted.

METHODOLOGY

The present research is to take the agricultural factors collected by the agricultural weather stations at Irrigation Experiment Station of ChiaNan Irrigation Association as the input layer of artificial neural networks, and the potential evapotranspiration or the pan evaporation as the output layer. After the training and learning procedures, the neural networks can be used to predict the potential evapotranspiration and the pan evaporation at a period, based on the data from the input layer of this period, together with the weight coefficients of the neural networks.

Evapotranspiration Estimation

The Irrigation Experiment Station of ChiaNan Irrigation Association (23°13'N, 120°11'E) is about 4m in altitude and 10 km away from coast. Since this area lies within the ChiaNan plain which is of flat ground and consistent climate, the information and data acquired in this area can fully represent the agricultural and production environment in ChiaNan plain area. Figure 1 shows the deployment of the central testing field and the agricultural weather stations. The devices for weather observation include dry and wet-bulb thermometers, thermometer, sunshine recorder, pyrheliometer, evaporation pan, anemometer, ground temperature indicator and pluviometer. The present research utilizes the agricultural weather data in past years recorded by the agricultural weather station, and also takes the evapotranspiration estimated by the FAO56 Penman–Monteith formula and the actual pan evaporation as the data required by the output layer of ANN. Allen et al. (1998) described the FAO56 Penman–Monteith method as Eq.(1).

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

where ET_o denotes the crop reference evapotranspiration (mm day^{-1}); R_n denotes the net radiation at crop surface ($\text{MJm}^{-2} \text{day}^{-1}$); G represents the soil heat flux density ($\text{MJ m}^{-2} \text{day}^{-1}$); T is the mean daily air temperature at 2 m height ($^{\circ}\text{C}$); u_2 is the wind speed at 2 m height (m s^{-1}); e_s denotes the saturation vapour pressure (kPa); e_a represents the actual vapour pressure (kPa); $e_s - e_a$ denotes the saturation vapor pressure deficit (kpa); Δ represents the slope vapour pressure curve ($\text{kPa } ^{\circ}\text{C}^{-1}$); γ is the psychometric constant ($\text{kPa } ^{\circ}\text{C}^{-1}$).

As shown in Eq.(2), the difference between the water in need by referential plants and the estimate of pan evaporation can be adjusted by the evaporation pan coefficient.

$$ET_o = K_p \times E_{pan} \quad (2)$$

where ET_o denotes the crop reference evapotranspiration (mm day^{-1}); K_p represents the evaporation pan coefficient; E_{pan} is the evaporation of evaporation pan (mm/day).

Theoretical Analysis of Backpropagation Neural Networks

The backpropagation neural network is one of the most popularly used artificial neural networks. Since it has the capability of learning and memorizing, it can be used for training and prediction. A backpropagation neural network usually has three layers: (1) an input layer to receive external information, (2) an output layer to output information to the external environment, (3) a hidden layer to supervised learning networks, which are to retrieve the training cases from the problems under investigation, to derive the underlying corresponding rules between the input variables and output variables, through minimizing the error function via the gradient steepest descent method, and finally to determine the underlying rules and estimate the new testing cases to output variables based on the memorizing capability.

During the training and learning of backpropagation neural networks, the weighing factors that are used to connect the input, output and hidden layers will change, so as to establish the nonlinear relationship between the input and output variables. The present research will utilize this property, and take the agricultural weather factors automatically recorded in the agricultural weather station as inputs to the input layer, which include the following average monthly statistics data: (1) the highest temperature; (2) the lowest temperature; (3) average temperature; (4) relative humidity; (5) wind speed; (6) sunlight hours; (7) solar radiation amount; (8) dew point; (9) forenoon ground temperature; (10) afternoon ground temperature. And also the output layer needs the potential evapotranspiration and pan evaporation. After the complete training and learning, the potential evapotranspiration or the pan evaporation during this period can be predicted based on the data in the same duration to the input layer and the weighing factors in the trained neural network model. Figure 2 shows the sketch map of backpropagation artificial neural network.

Eq.(3) denotes that the data to the input layer are converted to the range between 0~1 by the sign function during the forward stage, which in turn are regarded as the inputs to the hidden layer. Based on the calculation of Eq.(3), the values (y) transmitted from the input layer to the hidden layer can be represented in Eq.(4), while the results of (y) by Eq.(4) are taken as the input from the hidden layer to the output layer, and Eq.(5) then is used to calculate the (v) in the output layer based on the y values, and then Eq.(6) again uses a nonlinear function to convert the outputs in the output layer (v) into estimates (Z) during the forward stage, which will be compared with the target values (t) in the output layer.

$$g(u) = \frac{1}{1 + e^{-u}} \quad (3)$$

$$u = a_0 + a_1 * x$$

where x denotes the input data of input layer; a_0 denotes the base weighing factor of input layer; a_1 represents the weighing factors between input and hidden layer.

$$y = g(u) = \frac{1}{1 + e^{-u}} \quad (4)$$

$$v = b_0 + \sum_{j=1}^J b_j y_j \quad (5)$$

$$Z = g(v) = \frac{1}{1 + e^{-v}} \tag{6}$$

where j denotes node numbers of hidden layer; b_j denotes the weighing factors between hidden and output layer; b_o represents the base weighing factor of output layer; y_j is the output values of j^{th} node within hidden layer.

As known from the flow chart of backward procedure, Eq.(7) is used to calculate the error between the outputs in the output layer (z) and the target values (t) during the backward stage. And then the training will continue till the weighing factors, which are changed accordingly, are less than the tolerance. After that, the model can be used for prediction.

$$E = \frac{\frac{1}{2} \sum_{n=1}^N \sum_{k=1}^K (Z_{kn} - t_{kn})^2}{NK} \tag{7}$$

where N denotes data numbers of input layer; K denotes the node numbers of output layer; b_o represents the base weighing factor of output layer; t_{kn} is the target values of the n^{th} data; Z_{kn} represents the prediction values of the n^{th} data.

The Analysis of Weighing Factors in Artificial Neural Networks

The present research uses the weighing factors to establish the weight indices of the artificial neural network, after finishing the training of the above artificial neural network model, so as to investigate the importance and influence of different input variables relative to the actual measurements. Howes and Crook (1999) proposed the use of the weighing factors in the network to measure the influence of different input variables, including three types of influence: (general influence, GI), (specific influence, SI) and (potential influence, PI), which can be used to analyze and explain the interrelation between input and output variables. In detail, GI is to quantify the different interrelation and influence between input variables/characteristics and output results, based on the whole learning and training samples; SI and PI are, instead, based on a specific prediction sample, to quantify the different interrelation and influence between input variables/characteristics and outputs. In the hypothesis in Howes and Crook (1999), as for a three-layer artificial neural network, which has an input layer of n nodes, a hidden layer of h nodes and an output layer of 1 node, the GI calculation is given in Eq.(8):

$$GI(x_i) = \frac{\sum_{j=1}^h \left| \frac{w_{ji}}{\sum_{i=0}^n |w_{ji}|} v_j \right|}{\sum_{j=0}^h |v_j|} \tag{8}$$

where h denotes node numbers of hidden layer; n denotes the node numbers of input layer; i, j represents the index; w_{ij} is the weighing factors between the i^{th} node of input layer and j^{th} node of hidden layer; v_j represents the weighing factor of the hidden layer's j^{th} node to output layer.

RESULTS AND DISCUSSION

Evapotranspiration Estimate by the A1, A2 Groups with Ten inputs

Group A1 has an input layer with ten agricultural weather factors as its inputs, and an output layer with the evapotranspiration estimate by FAO56 Penman–Monteith formula as its single output. To achieve the best training and prediction performance in group A1, the present research tests the models of 1 to 10 nodes in the hidden layer respectively. As for group A1 with an input layer of ten agricultural weather factors and an output layer of one evapotranspiration estimate, all models constructed are the following ten types: $10*1*1$, $10*2*1$, $10*3*1$, $10*4*1$, $10*5*1$, $10*6*1$, $10*7*1$, $10*8*1$, $10*9*1$, $10*10*1$. From the training result, it is known the one with a hidden layer of ten nodes is of the best performance. Figure 3 shows the backpropagation neural network with an input layer of ten nodes, a hidden layer of 10 nodes and an output layer of 1 node ($10*10*1$). Figure 4 shows the convergence of errors of the $10*10*1$ neural network model in group A1, which finally converges to 0.005. Figure 5 shows the linearly regressed values for the recorded weather data by Irrigation Experiment Station of ChiaNan Irrigation Association since 2004 based on the trained neural networks model, and the correlation coefficient reaches 0.993. Figure 6 shows the variation when the evapotranspiration predictions of ChiaNan irrigated area in 2004 by the $10*10*1$ model in group A1 are compared with those estimated through the FAO56 Penman–Monteith formula. As from the result analysis, the evapotranspiration in ChiaNan irrigated area in 2004 estimated by the FAO56 Penman–Monteith formula is 911.6cm, while the one predicted by the $10*10*1$ model in group A1 is 896.4cm, which shows an error ratio of 1.67% only.

Group A2 has an input layer with ten agricultural weather factors as its inputs, and an output layer with pan evaporation as its single output. To obtain the best training and prediction performance, the present research separately tests different models, with the number of nodes in the hidden layer within the range from 1 to 20. That is, group A2 forms the set of 14 models in the following with an input layer of ten agricultural weather factors and an output layer of pan evaporation: $10*1*1$, $10*2*1$, $10*3*1$, $10*4*1$, $10*5*1$, $10*6*1$, $10*7*1$, $10*8*1$, $10*9*1$, $10*10*1$, $10*11*1$, $10*12*1$, $10*15*1$, $10*20*1$. As learned from the training result, the model with 20 nodes in the hidden layer performs best. Figure 7 shows the backpropagation neural networks model ($10*20*1$) in group A2, with an input layer of 10 nodes, a hidden layer of 20 nodes and an output layer of 1 node. Figure 8 shows the convergence of errors during the training of the $10*20*1$ model in group A2, which converges to 0.7. Figure 9 shows the linearly regressed value for the recorded weather data by Irrigation Experiment Station of ChiaNan Irrigation Association in 2004 based on the model when the training is finished, which just has a correlation coefficient of 0.708. Figure 10 shows the variation when the evapotranspiration predictions about ChiaNan irrigated area in 2004 by the $10*20*1$ model in group A1 are compared with those pan evaporation. As from the result analysis, the pan evaporation in ChiaNan irrigated area in 2004 is 1673.1cm, while the one predicted by the $10*20*1$ model in group A2 is 1451.72cm, which shows an error ratio of 13.23%.

Analysis of Weighing Factors in Backpropagation Neural Network Models

After the training of backpropagation neural networks, all the weighing factors in input, hidden and output layers can be further utilized to calculate the GI values based on Eq.(8). The magnitude of GI values can be used to investigate the importance of the agricultural weather factors in the input layer to the evapotranspiration in the output layer. Since during both the training and prediction stages, group A1 shows rather high correlation, the weighing factors calculated based on Eq.(8) after the model training in group A1 has been finished are listed in Table 1. The analysis shows that wind speed has strong influence. Taking the group A1 for instance, which accepts 10 agricultural weather inputs and the output of the evapotranspiration estimate by the FAO56 Penman–Monteith formula, the influence rank in the descending order of importance according to the General Influence (GI) are: wind speed (GI=0.438)> average temperature (GI=0.204)>dew point(GI=0.162) >the highest temperature (GI=0.051)> the lowest temperature (GI=0.05)> relative humidity (GI=0.039)> forenoon ground temperature (GI=0.027)> afternoon ground temperature (GI=0.02)> solar radiation amount (GI=0.006)> sunlight hours (GI=0.002). It can be further concluded from Table 1 that the wind speed among all these agricultural weather factors affects the evapotranspiration the most significantly, followed by the average temperature, while the effects of solar radiation amount and the sunlight hours on evapotranspiration are least significant.

Table 1. The GI Values and Ranks of Ten Agricultural Weather Factors in Group A1

agricultural weather factors	GI values	Ranks
Highest temperature	0.0508	4
Lowest temperature	0.0503	5
Average temperature	0.2041	2
Relative humidity	0.0392	6
Wind speed	0.4383	1
Sunlight hours	0.0058	10
Solar radiation	0.0022	9
Dew point	0.1622	3
Forenoon ground temperature	0.0269	7
Afternoon ground temperature	0.0202	8

Comparison of Optimal Backpropagation Neural Networks in Different Groups

The correlations of training and prediction when using the evapotranspiration quantity calculated by the FAO56 Penman–Monteith formula as the target outputs during the training and prediction

are above 0.97, which indicates that the degree of consistency between the values by FAO56 Penman–Monteith formula and the predicted ones is very high; on the contrary, the correlation coefficient of training and prediction, when using the pan evaporation as the target outputs, are significantly lower than those with the target outputs from the empirical FAO56 Penman–Monteith formula. The correlation coefficients in training and prediction stage by the group A2 are 0.887 and 0.708, respectively.

As observed from the training mode, whatever the group is, the correlation coefficient will decrease and the mean square error will increase as the number of input factors decreases; and the correlation coefficient will increase when there are more nodes in the hidden layer. As observed from the prediction mode, the slope is most likely larger than 1, which means that most predicted values are less than the actual ones, that is to say, the predicted values underestimates the actual ones. The underlying reason may lie in the difference in the target outputs, which causes the difference in correlation coefficient. When the target outputs are values calculated through the FAO56 Penman–Monteith formula, the training mode is simulating the formula since it has a strong ability in fitting the data. After the model is constructed, it becomes a formula calculator, thus has large correlation coefficient; When the target outputs are values based on pan evaporation, the training performance is poor and the correlation coefficient are lower than those based on FAO56 Penman–Monteith formula, because the pan evaporation is a measurement rather than a calculated one, and is more complicated than those derived from the FAO56 Penman–Monteith formula.

The predictions to the evapotranspiration in ChiaNan irrigated area in 2004 based on the neural network models after training has been compared. The evapotranspiration quantity in ChiaNan irrigated area in 2004, as calculated through the FAO56 Penman–Monteith formula, is 911.63cm, while those predicted through the backpropagation neural networks is 896.41 cm in group A1, with the error ratios to the calculated ones being 1.67%. In addition, the evaporation in ChiaNan irrigated area in 2004 is 1673.11cm, as from the records based on pan evaporation, while those predicted by the backpropagation neural networks is 1451.72 cm in group A2, with the error ratios to the actual one being 13.23%.

The evapotranspiration in 2004 as estimated by the FAO56 Penman–Monteith formula based on the agricultural weather information recorded by the Irrigation Experiment Station of ChiaNan Irrigation Association is 911.63 cm, while the pan evaporation is 1673.11cm. The two have the ratio of 0.777, that is to say, the evaporating pan coefficient (K_p) in ChiaNan irrigated area is 0.777.

REFERENCES

Allen, R.G., et al., 1998. FAO56: Crop Evapotranspiration- Guidelines for Computing Crop Water Requirements. FAO of UN, Rome, Italy.

Burks, T.F. et al., 2000. Backpropagation Neural Network Design and Evaluation for Classifying Weed Species Using Color Image Texture. Transactions of the ASAE, 43(4), 1029-1037.

Chang, F.J. et al., 2000. Counterpropagation Fuzzy-Neural Network for Reservoir Inflow Prediction. Monthly Journal of Taipower's Engineering, 618, 7-19.

Drummond S.T. et al., 2003. Statistical and Neural Methods for Site-Specific Yield Prediction. Transactions of the ASAE, 46(1), 5-14.

Han, H., Felker, P., 1997. Estimation of Daily Water Evaporation using an Artificial Neural Network. J. of Arid Environments, 37, 251-260.

Irmak, S., Allen, R.G., Whitty, E.B., 2003a. Daily Grass and Alfalfa-Reference Evapotranspiration Estimates and Alfalfa-to-Grass Evapotranspiration Ratios in Florida. J. of Irrigation and Drainage Engineering of ASAE, 129(5), 360-369.

Irmak, S., Allen, R.G., Jones, J.W., 2003b. Solar and Net Radiation-Based Equations to Estimate Reference Evapotranspiration in Humid Climates. J. of Irrigation and Drainage Engineering of ASAE, 129(5), 336-347.

Kuo, Y.M., Liu, C.W., 2000a. Analysis on Variation of Groundwater Quality in Yun-Lin Coastal Area:(1) Multivariate Factor Analysis Method. Taiwan Water Conservancy, 48(1), 1-8

Kuo, Y.M., Liu, C.W., 2000b. Analysis on Variation of Groundwater Quality in Yun-Lin Coastal Area:(2) Back-Propagation Artificial Neural Network Method. Taiwan Water Conservancy, 48(1), 9-25.

Liu, J., Goering, C.E., Tian, L. 2001. A Neural Network for Setting Target Corn Yields. Transactions of the ASAE, 44(3), 705-713.

Yang, C.C. et al., 1997. Application of Artificial Neural Networks for Simulation of Soil Temperature. Transactions of the ASAE, 40(3), 649-656.

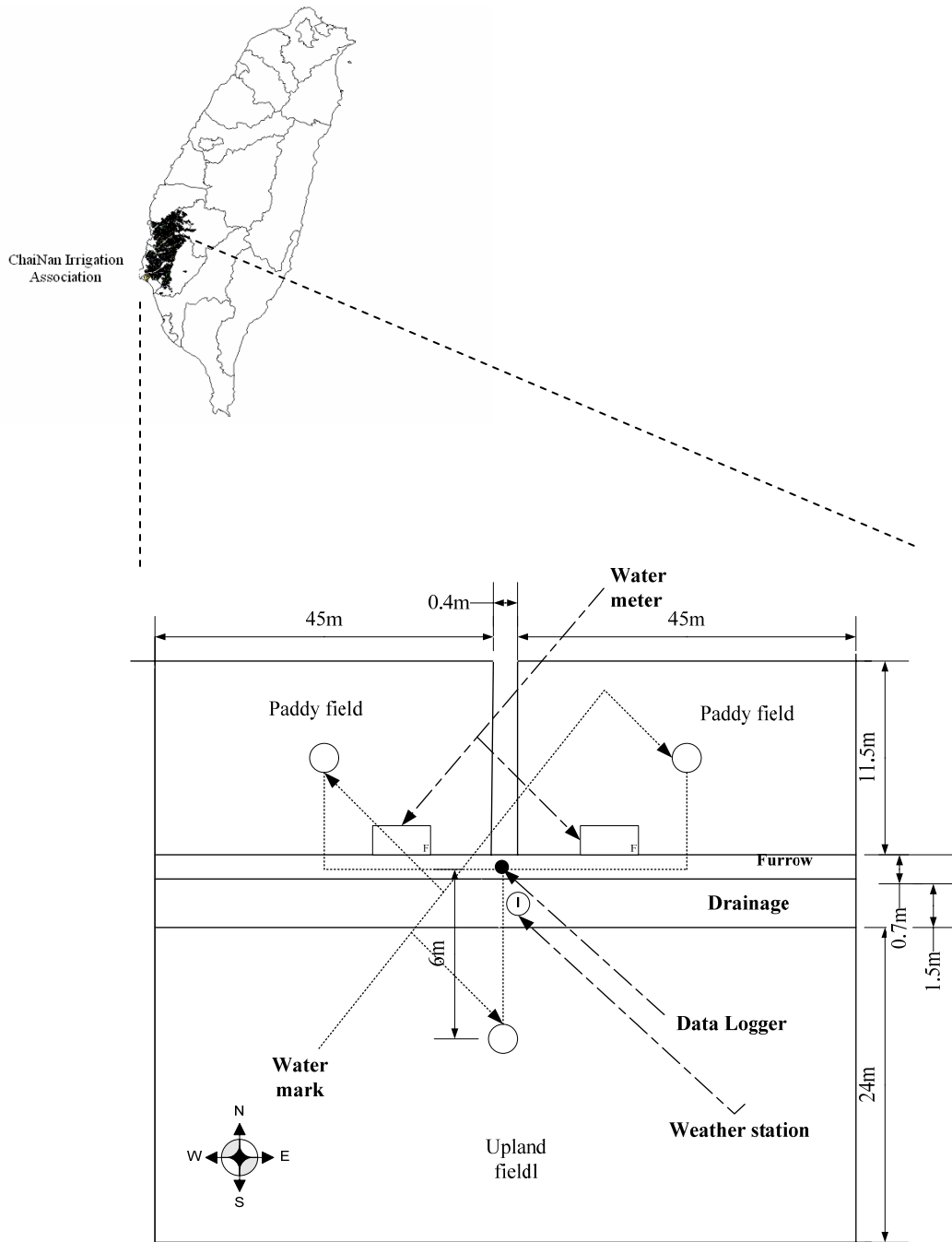


Figure 1. The Ichnography of the Irrigation Experiment Station of ChiaNan Irrigation Association

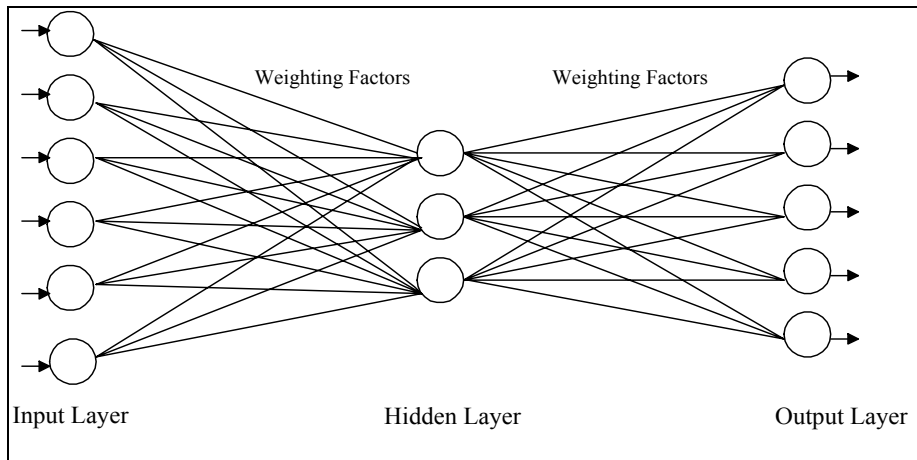


Figure 2. The Diagram of Bbackpropagation Artificial Neural Networks

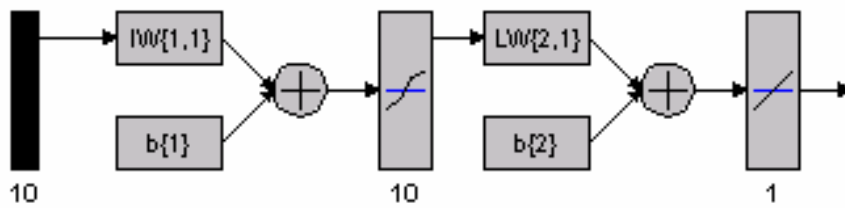


Figure 3. The Diagram of the Backpropagation Neural Network with an Input Layer of 10 Nodes, a Hidden Layer of 10 Nodes and an Output Layer of 1 Node

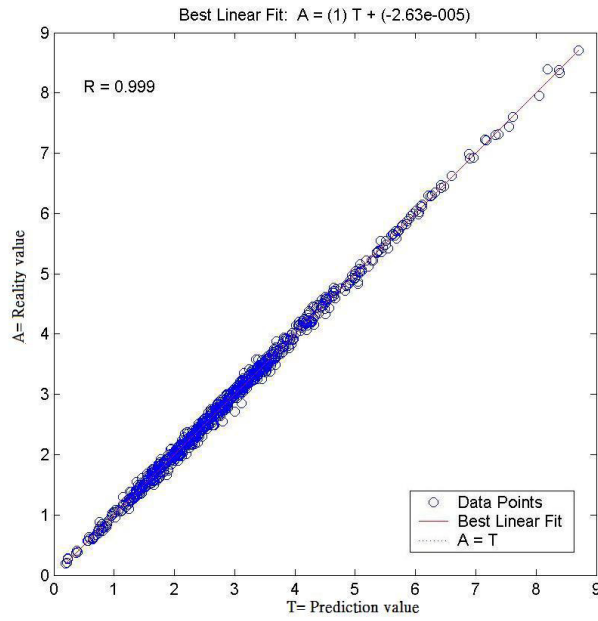


Figure 4. The Linear Regression Values after the Training of the 10*10*1 Model in Group A1

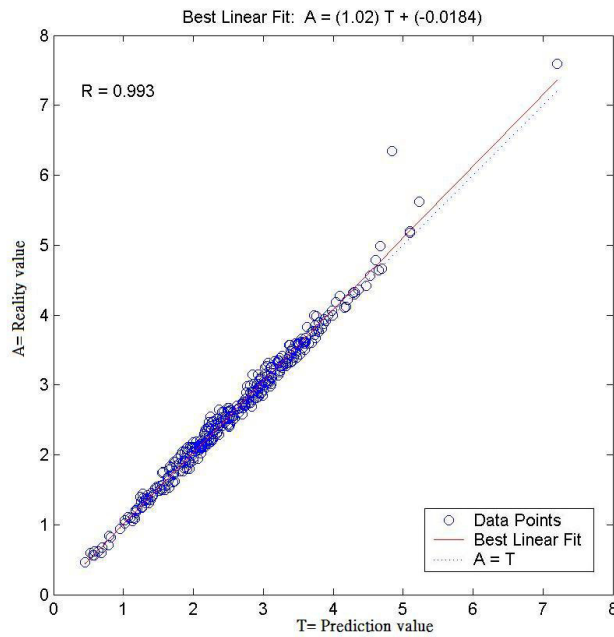


Figure 5. The Linearly Regressed Predictions to the Evapotranspiration in ChiaNan Irrigated Area in 2004 by the 10*10*1 Model in Group A1

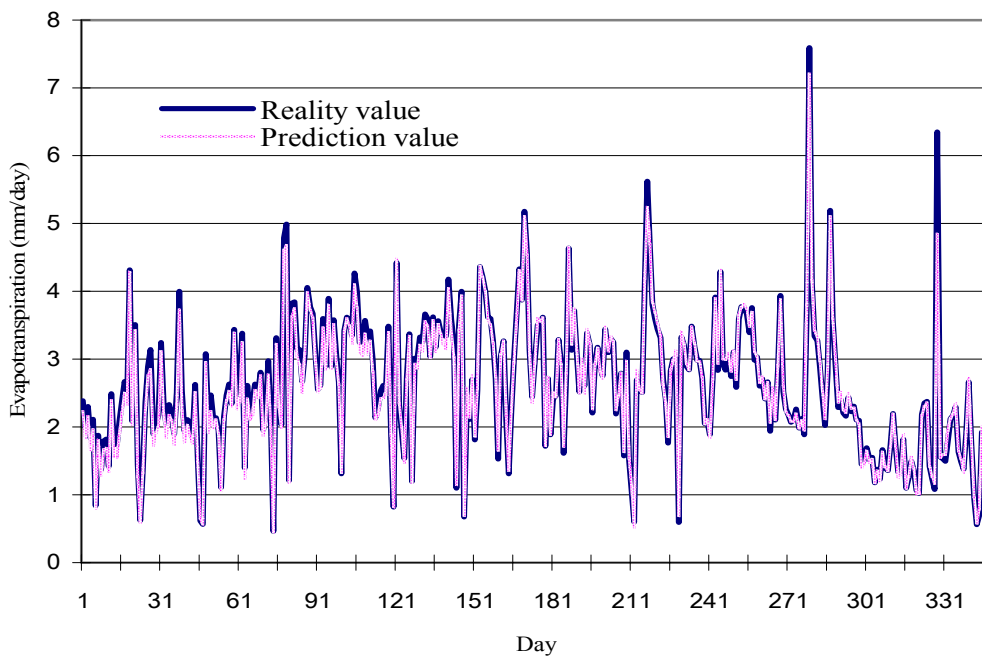


Figure 6. Comparison Between the Predicted Evapotranspiration in 2004 by the 10*10*1 Model in Group A1 and the One Estimated by FAO56 Penman-Monteith Method

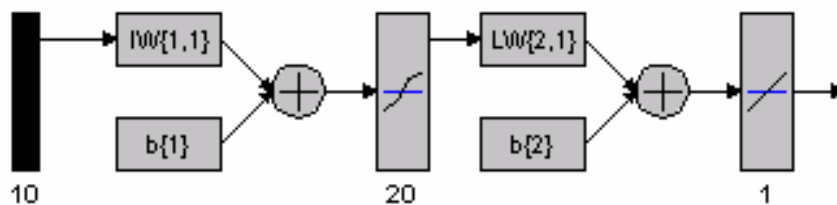


Figure 7. The Backpropagation Neural Network with an Input Layer of 10 Nodes, a Hidden Layer of 20 Nodes and an Output Layer of 1 Node in Group A2

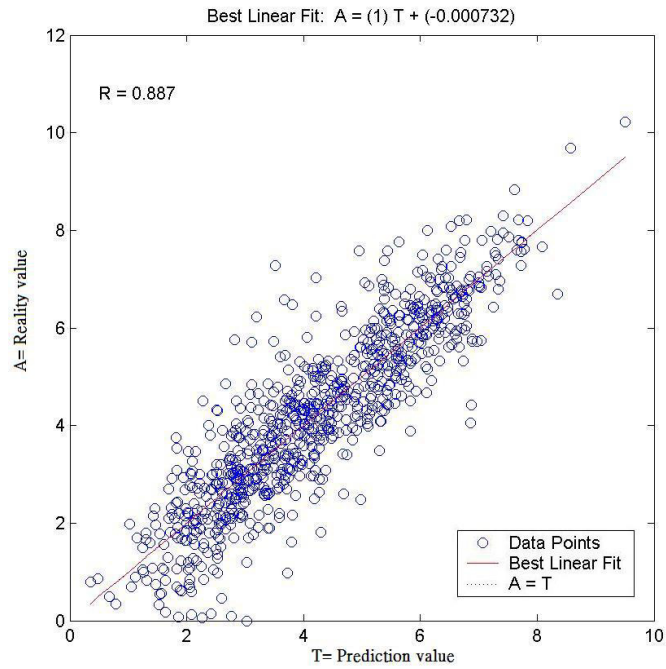


Figure 8. The Linear Regression Values after the Training of the 10*20*1 Model in Group A2

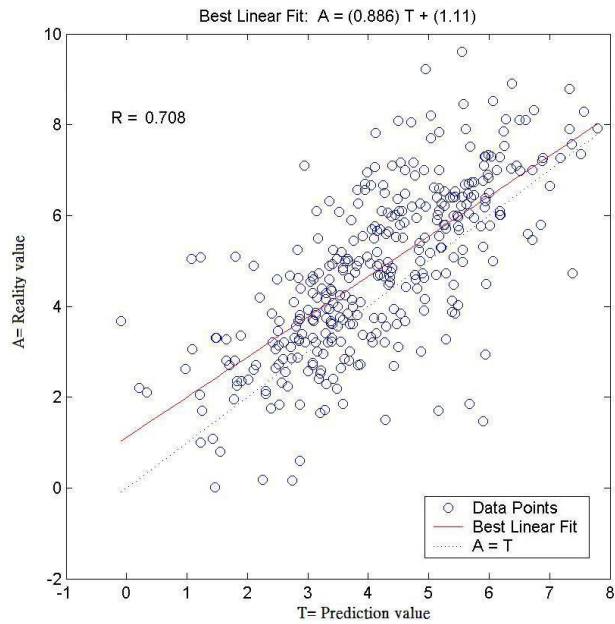


Figure 9. The Linearly Regressed Predictions to the Evaporation in ChiaNan Irrigated Area in 2004 by the 10*20*1 Model in Group A2

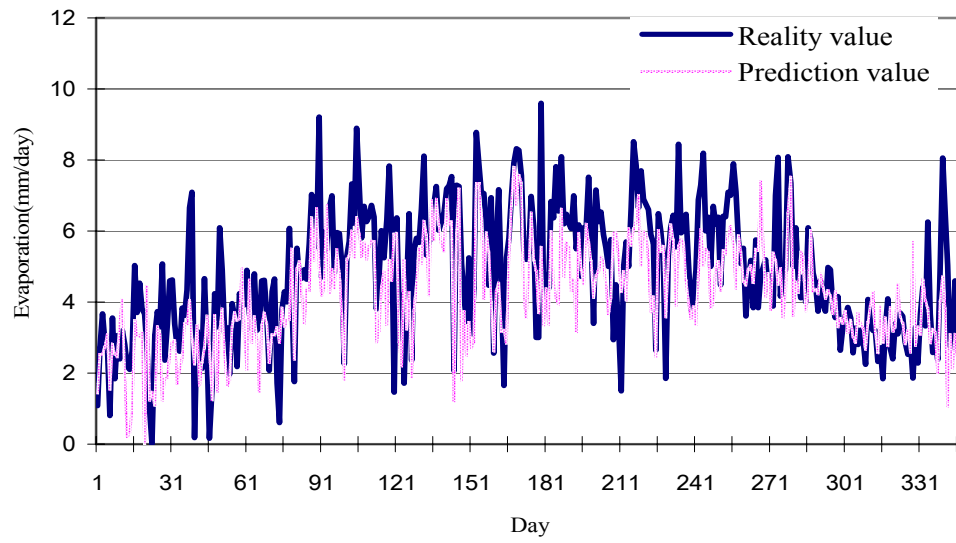


Figure 10. Comparison Between the Predicted Evaporation in 2004 by the 10*20*1 Model in Group A2 and the One Estimated by FAO56 Penman-Monteith method