ELSEVIER

Agricultural and Forest Meteorology 71 (1994) 115–132

AGRICULTURAL AND FOREST METEOROLOGY

Development of a neural network model to predict daily solar radiation

David Elizondo[a], Gerrit Hoogenboom*[,b], R.W. McClendon[c]

[a]*Artificial Intelligence Programs, University of Georgia, Athens, GA 30602, USA*

[b]*Department of Biological and Agricultural Engineering, University of Georgia, Georgia Station, Griffin, GA 30223-1797, USA*

[c]*Department of Biological and Agricultural Engineering, University of Georgia, Athens, GA 30602, USA*

Received 3 September 1993; revision accepted 29 December 1993

Abstract

Many computer simulation models which predict growth, development, and yield of agronomic and horticultural crops require daily weather data as input. One of these inputs is daily total solar radiation, which in many cases is not available owing to the high cost and complexity of the instrumentation needed to record it. The aim of this study was to develop a neural network model which can predict solar radiation as a function of readily available weather data and other environmental variables. Four sites in the southeastern USA, i.e. Tifton, GA, Clayton, NC, Gainesville, FL, and Quincy, FL, were selected because of the existence of long-term daily weather data sets which included solar radiation. A combined total of 23 complete years of weather data sets were available, and these data sets were separated into 11 years for the training data set and 12 years for the testing data set. Daily observed values of minimum and maximum air temperature and precipitation, together with daily calculated values for daylength and clear sky radiation, were used as inputs for the neural network model. Daylength and clear sky radiation were calculated as a function of latitude, day of year, solar angle, and solar constant. An optimum momentum, learning rate, and number of hidden nodes were determined for further use in the development of the neural network model. After model development, the neural network model was tested against the independent data set. Root mean square error varied from 2.92 to 3.64 MJ m^{-2} and the coefficient of determination varied from 0.52 to 0.74 for the individual years used to test the accuracy of the model. Although this neural network model was developed and tested for a limited number of sites, the results suggest that it can be used to estimate daily solar radiation when measurements of only daily maximum and minimum air temperature and precipitation are available.

*Corresponding author.

SSDI 0168-1923(94)02139-B

1. Introduction

Detailed crop growth simulation models have been developed for many major crops. These mechanistic and deterministic models are used as research tools to help understand the basic physiology of crop growth and development, and as decision support tools to optimize crop management strategies. Most crop simulation models require daily weather data such as temperature, precipitation, and solar radiation as input (Jones, 1991; Hoogenboom et al., 1992). At many sites, only precipitation and daily maximum and minimum temperatures are recorded. Owing to the high cost of instrumentation, and maintenance and calibration of pyranometers and other solar radiation sensors, accurate solar radiation measurements are available for only a few sites (Vives, 1986). Solar radiation has a significant effect on plant growth through photosynthesis. The interception of solar radiation by the plant is a function of canopy structure, angular distribution of the incident radiation, and the spectral properties of the leaves (Jones, 1992). Therefore, in attempting to predict crop yield for a particular location using computer simulation models, the availability of solar radiation data is as important as temperature and precipitation data (International Benchmark Sites Network for Agrotechnology Transfer Project (IBSNAT), 1990).

Attempts to develop algorithmic models to predict weather variables such as solar radiation have had limited success. An example of an algorithmic approach is the weather generator WGEN, developed by Richardson and Wright (1984). The WGEN model generates daily values for precipitation, maximum and minimum temperature, and solar radiation for any location in the USA. However, these data are generated stochastically and do not include a year effect; also, WGEN is unable to estimate solar radiation, given other environmental variables. Hodges et al. (1985) evaluated five algorithms created to estimate daily solar radiation using daily maximum and minimum temperatures and rainfall as inputs. The algorithm based on the Richardson (1981) weather generator seemed to perform best and caused the least bias when solar radiation values were used as inputs in crop simulation models.

Bristow and Campbell (1984) developed a relationship between atmospheric transmittance and the daily range of air temperature to predict incoming solar radiation. For the three sites for which solar radiation and other weather variables were used for model development, the Bristow and Campbell (1984) model accounted for 70–90% of the daily solar radiation. Seco et al. (1993) found that the phase lag between average monthly solar radiation and average monthly temperature varied from 15 to 30 days for two sites in Chile. Hook and McClendon (1992) showed that a solar radiation model which included daily open pan evaporation, precipitation, and daily maximum and minimum air temperature from Tifton, GA, as inputs, had an overall coefficient of determination of 78% with recorded solar radiation data. When pan evaporation was removed as an input variable to the model, the coefficient of determination between recorded and predicted solar radiation dropped to 64% for a 100 day summer period in 1990.

Other techniques which have been used to estimate solar radiation or related variables include the use of actual weather forecasts to predict solar irradiance transmission (Fynn and Roller, 1989), time series analysis to generate missing monthly solar radiation data (Shih and Cheng, 1989), hemispherical photographs (Smith and Somers, 1993), and sunshine duration (Troyo-Diéguez, 1992).

1.2. Neural network applications

A neural network can be viewed as a computer system that is made up of several simple and highly interconnected processing elements (McClelland et al., 1986a,b) which process information by their dynamic state response to inputs. Problems not solvable by traditional algorithmic approaches frequently can be addressed with neural networks (Davidson and Lee, 1991). They are suitable for investigations requiring the interpretation of large data sets and for problems in which the inputs and corresponding output values are known but the relationship between the inputs and the outputs is not well understood. These conditions are commonly found in many agricultural applications, owing to the interrelationships between physical, chemical, and biological processes in agricultural and environmental systems.

Neural networks have been applied successfully to various agricultural systems. Bolte (1989), Zhuang and Engel (1990), and Muttiah and Engel (1991) presented introductions to neural networks with applications in agriculture and natural resources. Whittaker et al. (1991) implemented an ultrasonic signal classified for beef quality grading. Their study showed that a statistical model and a neural network model generated similar results. Thai et al. (1991) presented a neural network for determining the maturity level of green tomatoes, at harvest, using X-ray computed tomography images. Thai and Shewfelt (1991) used a neural network and statistical regression to find mathematical relationships linking human sensory judgments to color of tomato and peach. A milking robot developed by Marchal et al. (1991) included a neural network model as an adaptive control to enhance the precision in the placement of the arms of the robot.

In terms of forecasting and prediction, two different applications were presented by Neural Computer Sciences (1992), in which neural network models were used in time-series predictions for financial forecasting. Chen et al. (1992) used hourly and average daily temperatures as inputs for a neural network to predict short-term electrical loads in an energy management system. Uhrig et al. (1992) used a neural network to predict corn yield, using weekly maximum and minimum temperatures, soil moisture, and cumulative degree days as inputs.

Rogers (1991) presented a weather predictor using genetic memory, which is a hybrid model that uses genetic algorithms and neural networks. The system was developed and trained with rainfall data that included samples taken every 4 h for a 25 year period from one site in Australia. Cook and Wolfe (1991) developed a neural network to predict average air temperatures. In this model, the monthly average of daily maximum temperatures for 3 months in advance was predicted. The network was trained with 9 years of data and tested against 1 year of average monthly maximum temperatures recorded at Abilene, TX.

The aim of the research presented in this paper was the development of a neural network model which could predict daily solar radiation as a function of both

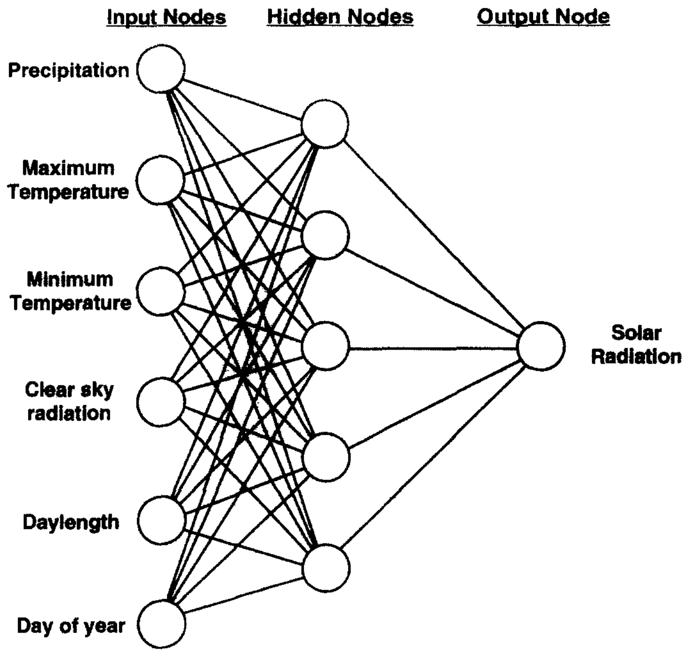

Fig. 1. Neural network connections and structure for a solar radiation model, shown with five hidden nodes.

observed weather and calculated environmental variables. The input variables for the neural network were observed weather variables, including daily precipitation and maximum and minimum air temperatures, and calculated environmental variables, including daylength and daily total clear sky radiation. The neural network model was evaluated by comparing the predicted daily solar radiation results with observations that were not used for model development.

2. Methods

2.1. Neural networks

The back-propagation neural network algorithm was selected for this investigation because neural networks using this algorithm can learn from examples and accept non-binary values as inputs (McClelland et al., 1986a,b; Vemuri, 1988). The back-propagation training algorithm is based on gradient descent and is designed to minimize the mean square error between the actual output of a multi-layer feed-forward network and the desired output. The network is normally made up of nodes in the input, hidden and output layers. Usually the network is fully connected, in that each input node is connected to each hidden layer node and each hidden layer node is connected to each output layer node (Fig. 1). The connection strengths (or weights) between the nodes are used as a multiplicative factor with the activation level of the sending node. A summation is performed over the sending nodes for each node in the receiving layer to obtain the net input,

$$\mathrm{net}_j = \sum_{i=1}^{n} w_{ij} a_i + \mathrm{bias}_j \tag{1}$$

where net_j is the net input to receiving node j, w_{ij} is the weight or connection strength between the receiving node j and the sending node i in the previous layer, a_i is the activation of sending node i, bias_j is the bias for the receiving node j, and n is the number of sending nodes in the previous layer.

The bias term for a node behaves as an additional weight in which the sending node always has an activation of unity. This term, in effect, serves as a threshold for each node. The net input to each node in the hidden layer and output layer is calculated using this method. This net input is then applied to an activation function. The activations of the output nodes are then compared with the known or target output values. The summed squared error E is defined as follows:

$$E = \sum_p E_p = \sum_p \sum_j (t_{pj} - a_{pj})^2 \tag{2}$$

where E is the summed squared error, p is the index for the set of patterns consisting of known inputs and outputs, j is the index for the set of output nodes, t_{pj} is the target output for node j and pattern p, and a_{pj} is the output node activation for node j and pattern p.

The objective of a least mean square error method is to find the set of weights w_{ij} which minimizes the error. The back-propagation method uses gradient descent to change the weights proportional to the negative of the derivative of the error function with respect to each weight. A thorough discussion of the mathematics of the back-propagation method has been given by McClelland et al. (1986a,b).

During training, a pattern consisting of known inputs and outputs is presented to the network. The inputs are fed forward through the network and an output is calculated. The error is calculated and the weights between the hidden layer and output layer are adjusted. Similarly, the weights between the input layer and hidden layer are also adjusted. This procedure is repeated for all patterns in the training set. The patterns are continuously presented and weights are adjusted until the error is sufficiently low.

NeuroShell™ (Ward Systems Group, Frederic, MD) was selected as the neural network software for this research because it contains a user-friendly back-propagation implementation. NeuroShell™ is a menu-driven program for developing neural network models, and includes several utilities for data manipulation, model development, graphical options, and a runtime option to generate source code. One of the most important features of NeuroShell™ is the optimal network option. This feature can prevent, to a certain degree, over-learning. Over-learning becomes critical in prediction applications because the neural network should be capable of generalizing over a training set as opposed to learning the specific aspects

Table 1
Location and year of weather data used in the training and testing phases of the solar radiation neural network

Location	Latitude (deg)	Longitude (deg)	Elevation (m)	Annual average rainfall (mm)	Year of weather data used	
					Training	Testing
Gainesville, FL	29.65	−82.35	52	1276	1976	1978
					1980	1979
					1982	1981
					1984	1985
Quincy, FL	30.58	−84.58	76	1426	-	1979
Tifton, GA	31.48	−85.53	113	1208	1981	1980
					1982	1983
					1986	1984
					1987	1985
					1988	1989
					1991	1990
Clayton, NC	35.65	−87.50	101	1193	1985	1986

of the training data set. The user can define the interval at which the NeuroShell™ will check the current network with a testing data set that is independent of the training data set. If the average error of the prediction of the neural network improves over the previous optimal network, the older optimal network is replaced with the latest model developed. Otherwise, the previous optimal network is saved until an improved optimal network is reached, or the training is completed.

The criterion for defining when a neural network model is trained varies with each implementation of the back-propagation algorithm. In NeuroShell™, a learning threshold is set by the user. This is the error between the actual and predicted output for each pattern presented. If the error is greater than this learning threshold, the weights are adjusted by applying the back-propagation algorithm. Once the error between the actual and predicted value is below the given learning threshold for each pattern in the training set, the neural network is considered to be trained.

2.2. Model development

Climatological data from Tifton, GA, Clayton, NC, Gainesville, FL, and Quincy, FL, were used in the development of the solar radiation model. The total combined number of data sets was 23, and each set contained a complete year of 365 days of daily weather data for one of the selected locations (Table 1). The weather data were stored in ASCII files, using the IBSNAT format, which includes daily weather data for precipitation, maximum and minimum air temperature, and solar radiation (IBSNAT, 1988).

The next step was to determine the input combination that produced the most accurate neural network model for predicting solar radiation. The weather data

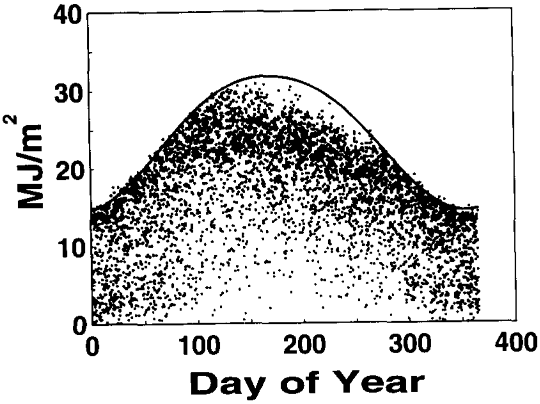

Fig. 2. Calculated daily global radiation (line), using a maximum clear sky atmospheric transmissivity of 0.75 and daily observed solar radiation (points) as a function of day of year for Tifton, GA. Independent test data for years 1975–1989 are shown.

consisted of daily maximum air temperature, daily minimum air temperature and daily total rainfall. Daily clear sky radiation and daylength, which can be calculated in terms of latitude and day of year, were also included in each input data set in the proposed neural network models. Daily clear sky radiation was calculated using the equations presented by Spitters et al. (1986). A maximum clear sky atmospheric transmission coefficient of 0.75 was applied, similar to the value found in the weather generator WGEN (Richardson and Wright, 1984). Fig. 2 shows all daily total solar radiation data recorded from 1975 to 1989 in Tifton. The maximum clear sky atmospheric transmission coefficient was based on fitting an 'envelope' of maximum potential solar radiation for each day, using this set of 15 years of daily historical solar radiation data. We assumed that maximum clear sky atmospheric transmissivity was constant for the entire year. The daylength was calculated using the formulae implemented in the SUNRIS routine of the dry bean model BEANGRO (Hoogenboom et al., 1991).

An example of a neural network model is shown in Fig. 1, with six input nodes, five hidden nodes, and one output node. The input nodes shown are day of year, daily maximum air temperature, daily minimum air temperature, daily total precipitation, daylength, and daily clear sky radiation. The output node is daily total solar radiation.

Complete years of weather data from Tifton, Gainesville, and Clayton, were randomly selected for training and testing the neural network model as listed in Table 1. The one available weather year from Quincy was included in the testing set and was not used for training. A total of 11 years of data were selected for the training set and 12 years for the testing set. As part of the data preparation for the neural network, the data were converted from an ASCII file into a LOTUS spreadsheet file, and clear sky radiation and daylength were calculated for each site and day of year. The LOTUS data files

Table 2
Results obtained from the various sets of internal model parameters used in developing the solar radiation neural network model; results shown include only the independent test data sets[1]

Momentum	Learning rate	No hidden nodes	R^2 Observed and predicted solar radiation
0.9	0.6	5	0.615
0.9	0.6	10	0.619
0.9	0.6	15	0.626
0.9	0.6	25	0.617
0.9	0.6	30	0.597
0.9	0.6	50	0.560
0.9	0.6	100	0.575
0.3	0.3	15	0.622
0.1	0.1	5	0.616
0.1	0.1	10	0.634
0.1	0.1	15	0.623
0.0	0.1	5	0.610
0.0	0.05	10	0.635
0.0	0.05	15	0.612
0.0	0.05	20	0.630

[1]The learning threshold was 0.0001. Model inputs were day of year, daily maximum and minimum air temperature, daily total precipitation, daylength and daily total clear sky radiation. R^2 is the coefficient of determination.

were then input to the neural network package NeuroShell™ for development of the solar radiation model.

Each neural network configuration considered was trained with the 11 years of data consisting of over 4000 patterns. Once the neural network training was completed for each model, the testing set was presented to the neural network to predict daily solar radiation for the 12 years of the test set. The predicted daily solar radiation values were compared with the observed daily solar radiation values for these 12 years. The accuracy of the neural network model was evaluated based on the root mean square error of the predictions of the independent test data set and on the R^2 value between predicted and observed solar radiation of the independent test data set.

Some of the input variables were removed individually and the process described above was repeated. This was done to determine if eliminating an input variable could improve the accuracy of the solar radiation neural network prediction. We found that the most accurate neural network used all of the six inputs described previously. Several experiments were conducted to find the combination of number of hidden nodes, learning rate, momentum and learning threshold which gave the greatest accuracy in predicting the test data set. A typical training session on an IBM- compatible 386SX 16 MHz computer with an 80387SX math co-processor with six inputs and 10 hidden nodes required approximately 18 h. The number of learning events required for the neural network to converge was approximately 1.5×10^6.

Table 3
Statistical comparison between observed and predicted solar radiation for each individual data set; results shown include only the independent test data sets

Location	Year	N	Average daily solar radiation ($MJ\,m^{-2}$)		R^2	CV	Av. error ($MJ\,m^{-2}$)	Absolute error ($MJ\,m^{-2}$)	RMSE ($MJ\,m^{-2}$)
			Observed	Predicted					
Gainesville, FL	1978	365	16.22	16.74	0.70	17.79	+0.52	2.71	2.98
	1979	365	16.23	16.49	0.71	17.73	+0.25	2.54	2.92
	1981	365	16.67	17.43	0.65	18.28	+0.76	2.67	3.19
	1985	365	16.21	16.33	0.68	18.68	+0.12	2.78	3.05
Quincy, GL	1979	365	17.12	16.21	0.71	18.76	+0.91	3.92	3.04
Tifton, GA	1980	365	15.58	16.37	0.57	21.21	+0.79	3.30	3.47
	1983	365	16.63	16.43	0.60	21.84	−0.20	3.80	3.59
	1984	365	16.96	16.47	0.58	20.83	−0.49	3.60	3.43
	1985	365	15.97	16.20	0.56	21.65	+0.23	3.50	3.51
	1989	365	16.08	16.10	0.52	22.59	+0.02	4.13	3.64
	1990	365	18.02	16.95	0.64	20.04	−1.07	3.48	3.40
Clayton, NC	1986	365	13.49	14.67	0.74	24.60	+1.18	3.33	3.61

N, number of data points; R^2, coefficient of determination; CV, coefficient of variation; average error = $[\sum(\text{predicted} - \text{observed})/N]$; absolute error = $[\sum|(\text{predicted} - \text{observed})|/N]$; RMSE, a root mean square error.

3. Results and discussion

Table 2 shows the results obtained with various combinations of parameters in developing the neural network model for predicting solar radiation. The initial values for the momentum and the learning rate were set at 0.9 and 0.6, respectively. These are the default values used by the NeuroShell™ package, and they were kept constant while searching for the preferred number of hidden nodes. As the number of hidden nodes increased, the R^2 value of the predicted vs. observed solar radiation increased until the number of hidden nodes reached the value of 15. For values higher than 15, the number of hidden units did not seem to improve the model's predictive capability, and tended to extend the time required to train the neural network. For each experiment, the time and learning events lapsed were similar to those described previously. After experimenting with the number of hidden nodes, the momentum and learning rate values were reduced and additional values for the number of hidden nodes were selected. Some complex and unstable problems can actually train faster when a low learning rate and momentum are used. Modifying both momentum and learning rates caused some improvement in the prediction accuracy of the neural network. The highest R^2 value, i.e. 0.635, was obtained with a momentum value of 0.0, a learning rate of 0.05, and 10 hidden nodes. These values for momentum, learning rate, and number of hidden nodes were used to develop the neural network for predicting solar radiation. An exhaustive search of all combinations of values for momentum, learning rate, and hidden nodes was not performed owing to excessive training times, as discussed above.

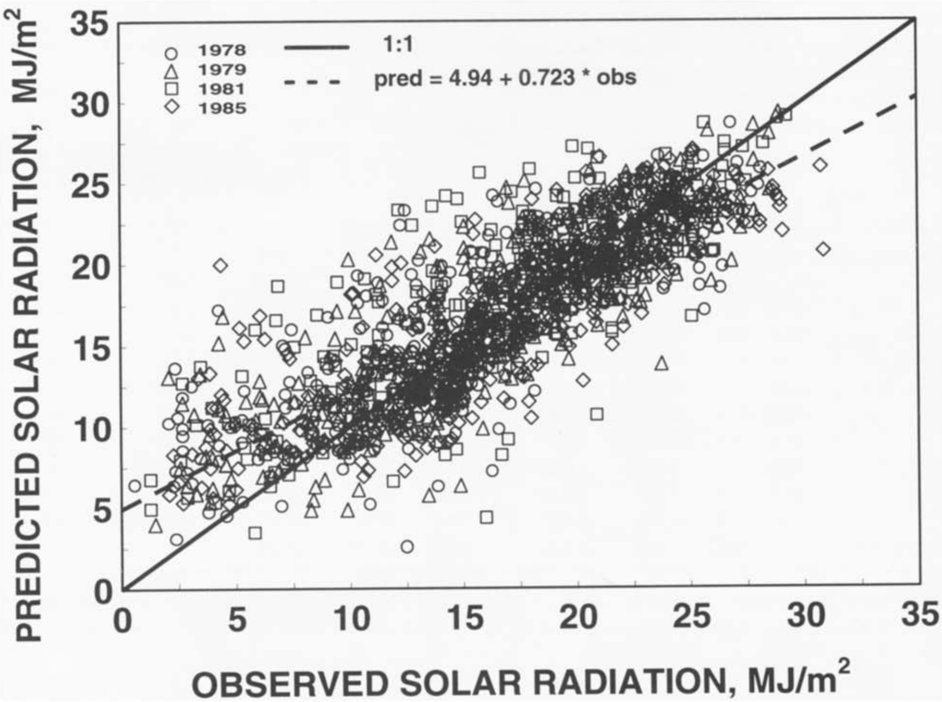

Fig. 3. Predicted daily solar radiation for Gainesville, FL, as a function of observed daily solar radiation. Independent test data for years 1978, 1979, 1981, and 1985 are shown; R^2 is 0.68.

On the basis of the best model described in the previous paragraph (Table 2), the model was applied using the individual test data sets listed in Table 1. Each individual test data set contained a total of 365 observed solar radiation data. The coefficient of determination for each individual data set varied from 0.52 for Tifton (1989) to 0.74 for Clayton (1986). The coefficient of variation varied from 17.73 for Gainesville (1979) to 22.59 for Tifton (1989). The average error varied from - 1.07 MJ m^{-2} for Tifton (1990) to +0.91 MJ m^{-2} for Quincy (1979). Absolute error varied from 2.54 MJ m^{-2} for Gainesville (1979) to 4.13 MJ m^{-2} for Tifton (1989). Root mean square error varied from 2.92 MJ m^{-2} for Gainesville (1979) to 3.64 MJ m^{-2} for Tifton (1989) (Table 3). Overall, the neural network model had a tendency to overpredict solar radiation, as shown by three negative average errors and eight positive average errors. It should also be noted that the model performed very well for the Quincy data set, although this data set was not used for training. Quincy had one of the highest coefficients of determination, and one of the lowest coefficients of variation and root mean square error values.

A plot of all daily predicted and observed solar radiation data for our most southern site, i.e. Gainesville, for years 1978, 1979, 1981, and 1985 is presented in Fig. 3. These individual weather years were not used in the training phase of the neural network. The R^2 value for the combined Gainesville solar radiation data

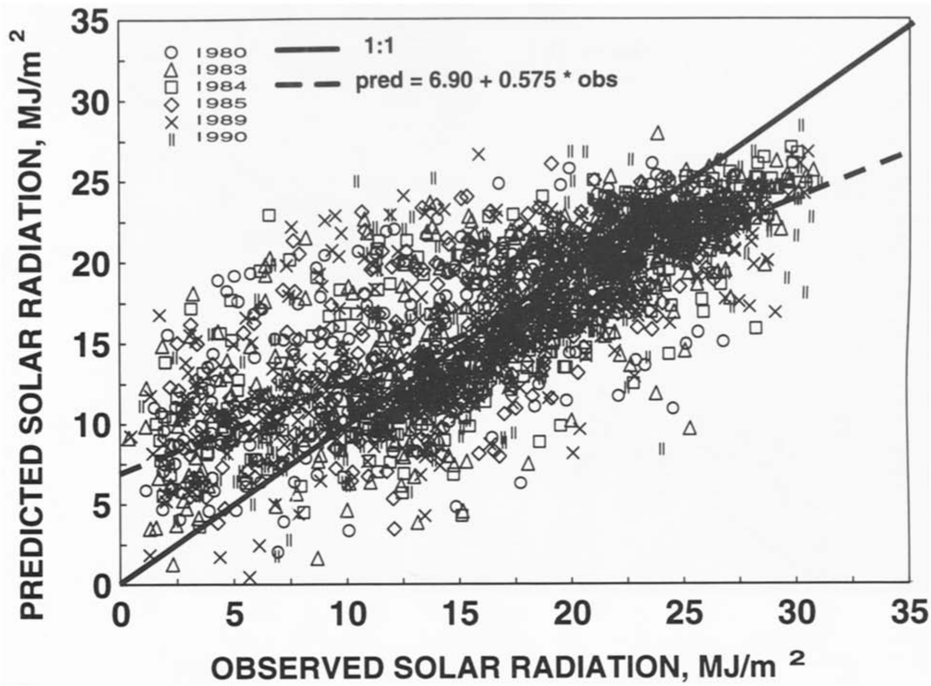

Fig. 4. Predicted daily solar radiation for Tifton, GA, as a function of observed daily solar radiation. Independent test data for years 1980, 1983, 1984, 1985, 1989, and 1990 are shown; R^2 is 0.57.

used for testing the neural network model was 0.68. Fig. 4 presents the daily predicted and observed solar radiation data for Tifton, for years 1980, 1983, 1984, 1985, 1989, and 1990. The R^2 value for the combined Tifton solar radiation data used for testing the network model was 0.57. The R^2 value for Tifton is relatively low compared with those for the other locations, owing to lower R^2 values for various individual weather years (Table 3). This could be partially due to the fact that our neural network model uses only a limited set of input data and therefore is unable to account for all physical processes. However, the results from the other three sites are better than those from Tifton, showing that the model accounts for the physical processes at those sites. Fig. 5 presents the predicted and observed solar radiation for Clayton in 1986. The R^2 value for the Clayton solar radiation data used for testing the network model was 0.74. For the three sites presented in Figs. 3–5, the neural network model had a tendency to overpredict solar radiation at low values and to underpredict solar radiation at high values. Most of the values, however, are located close to the 1:1 line, represented by the very dark areas.

In the plots shown in Figs. 6, 7 and 8, the years were divided into 4 month periods or thirds of a year to add to the resolution and to the clarity of presentation. A plot of predicted and observed daily solar radiation as a function of day of year is shown for Gainesville (1979) in Fig. 6. The average error for this data set was $+0.25$ MJ m^{-2}, and the average absolute error was 2.54 MJ m^{-2}. Predicted and observed solar

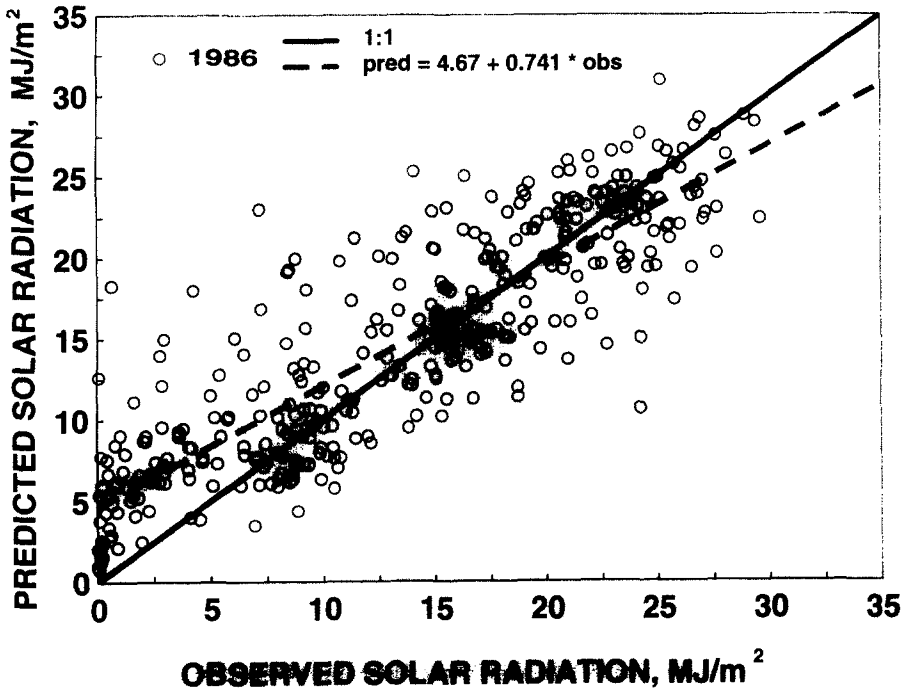

Fig. 5. Predicted daily solar radiation for Clayton, NC, as a function of observed daily solar radiation. Independent test data for year 1986 are shown. R^2 is 0.74.

radiation data for Tifton (1990) are shown in Fig. 7. Predicted solar radiation shows the same day-to-day variation as the observed data, especially during the first (Fig. 7(A)) and third 4 months (Fig. 7(C)), whereas during the second 4 months not as much variation is found (Fig. 7(B)). The average error for this Tifton 1990 data set was -1.07 MJ m^{-2}, and the average absolute error was 3.48 MJ m^{-2}. Hook and McClendon (1992) used the same data set to test their solar radiation model. The R^2 was 0.64 for a 100 day period during the summer, i.e. Days 100–200, using daily clear sky radiation, precipitation, average temperature, and the difference between maximum and minimum temperature as inputs. The R^2 for the neural network model for the entire year of 1990, i.e. Days 1–365, was also 0.64. Hook and McClendon's model was only developed for one site, i.e. Tifton, and was not applied to other locations. The neural network model presented in our study was developed for three sites and tested for four sites. Predicted and observed solar radiation data for Clayton (1986) are displayed in Fig. 8. From the data presented in Figs. 6–8, it can be seen that the solar radiation predicted by the neural network has a similar day-to-day variation to the observed solar radiation.

The model also seems to be able to predict the extremes at both the high and low end during the same period as found for the observed solar radiation data. The neural network model accurately predicted solar radiation maxima, as shown, for instance, on Day 145 (25 May) for Gainesville (Fig. 6(B)). It also accurately predicted solar

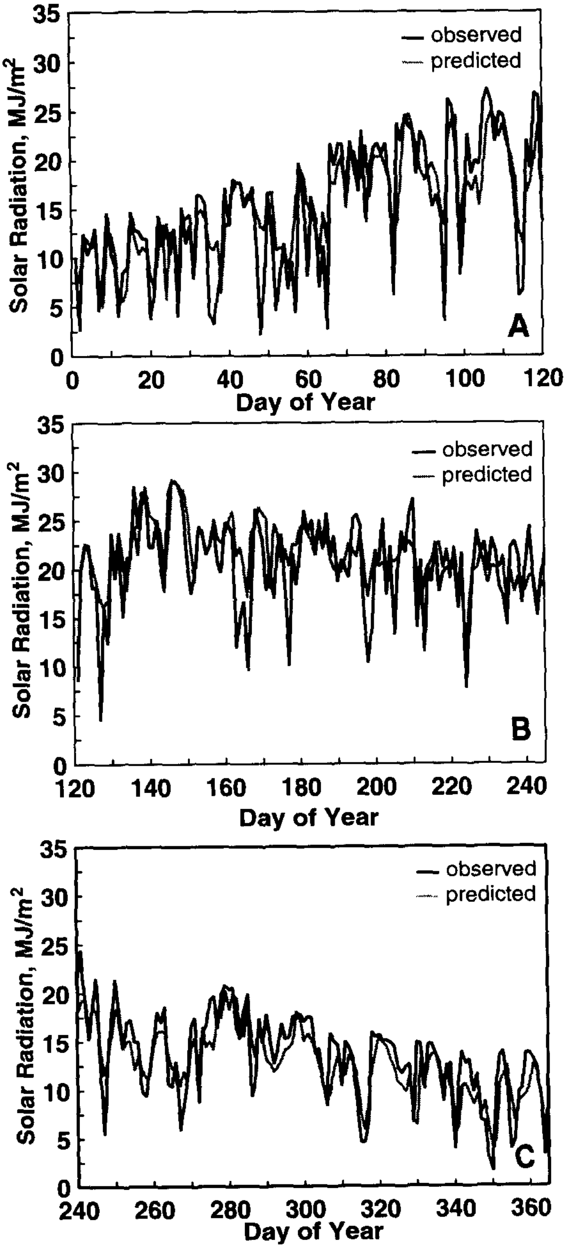

Fig. 6. Daily observed and predicted solar radiation for Gainesville, FL, as a function of Days 1–120 (A), Days 120–245 (B), and Days 245–365 (C). Independent test data for year 1979 are shown.

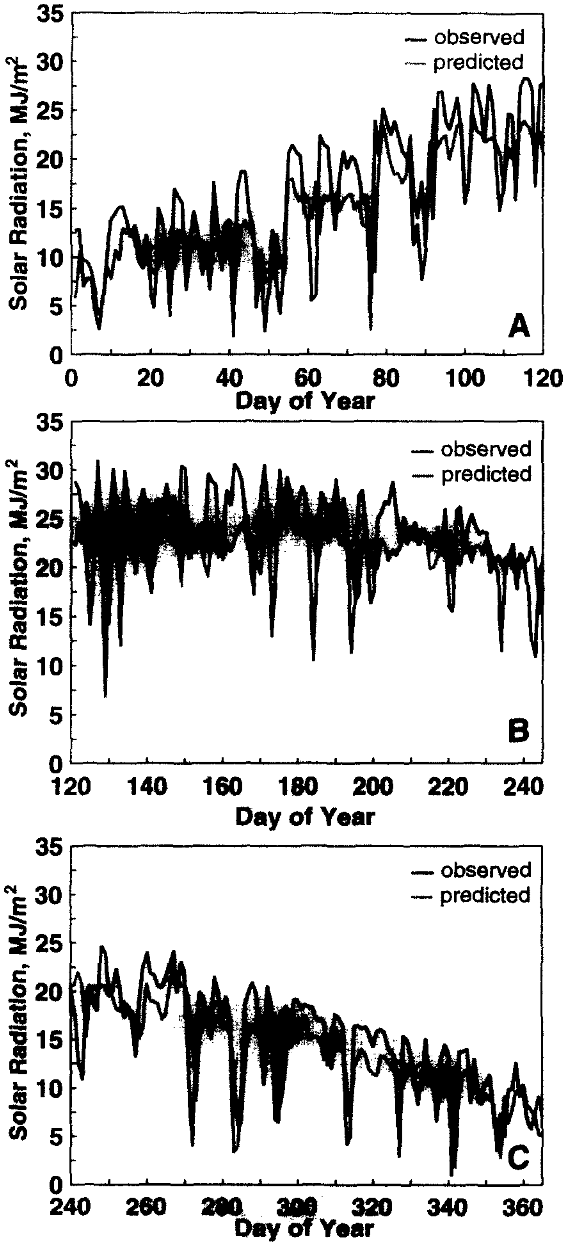

Fig. 7. Daily observed and predicted solar radiation for Tifton, GA, as a function of Days 1–120 (A), Days 120–245 (B), and Days 245–365 (C). Independent test data for year 1990 are shown.

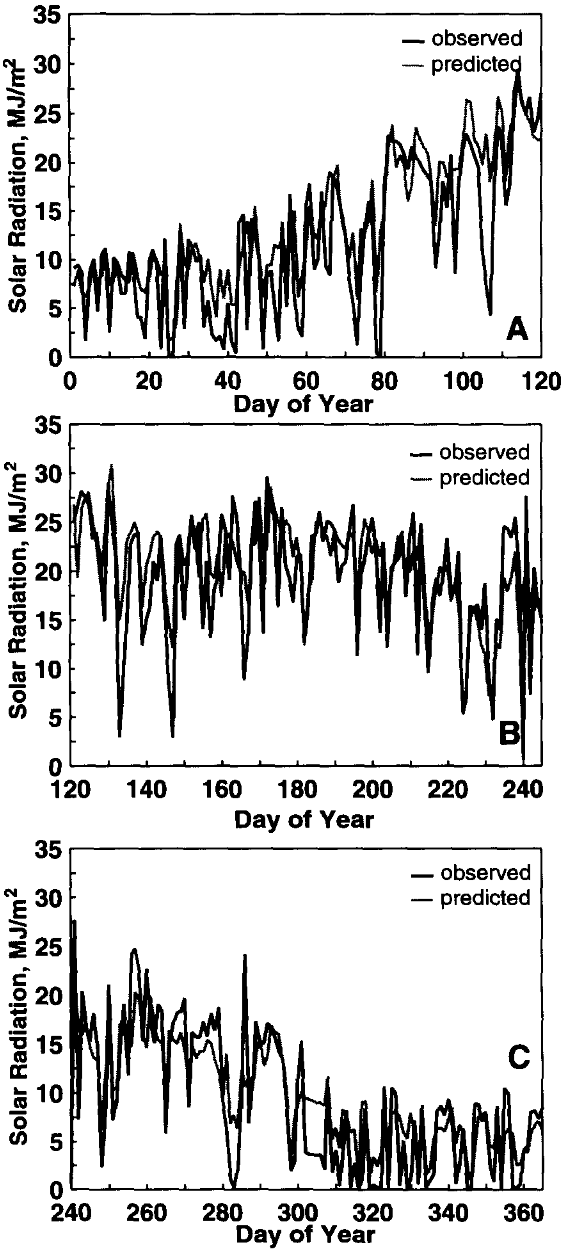

Fig. 8. Daily observed and predicted solar radiation for Clayton, NC, as a function of Days 1–120 (A), Days 120–245 (B), and Days 245–365 (C). Independent test data for year 1986 are shown.

radiation minima, as shown, for instance, on Day 340 (6 December) for Tifton (Fig. 7(C)) or Day 350 (16 December) for Clayton (Fig. 8(C)). Although one could argue that this is expected because maximum and minimum temperatures were either high or low on these example days, the neural network model also included other weather variables as input. One of the negative aspects of using neural networks is that they cannot explicitly present the predictive equation. However, this negative aspect of neural network model development has not prevented its success in other disciplines (Cook and Wolfe, 1991; Davidson and Lee, 1991; Muttiah and Engel, 1991). The advantages of using a neural network approach compared with a multi-linear regression approach are that the former has greater flexibility in model development and it assigns relative weights to each connection of the input variables.

4. Conclusions

Results presented in this paper show that a neural network model can be used to predict daily solar radiation as a function of daily maximum and minimum air temperature, daily total precipitation, daily clear sky radiation, and daylength. However, the model was tested at only a limited number of sites and for a relatively small number of years. Although the final neural network model had a tendency to overpredict solar radiation at low values and underpredict it at high values, average errors and coefficients of variation were low and the coefficients of determination were high. It is likely that the accuracy of the neural network for predicting solar radiation could be improved by adding other field-measured daily parameters such as per cent cloud cover, solar brightness or pan evaporation, as discussed in the literature review of other methods used to estimate solar radiation. However, our goal was to develop a model to predict solar radiation using only daily maximum and minimum temperature and precipitation as observed data. With these limited sets of inputs, the neural network model gives reasonably good predictions of daily solar radiation, with R^2 as high as 0.74 and the root mean square error as low as 2.92 MJ m^{-2}. Because many locations only have daily maximum and minimum air temperatures and daily total precipitation available as climatological variables, this neural network model can potentially be used to predict daily total solar radiation for these sites. Further research is needed to test this approach at locations with latitudes higher than 35.87°N (Clayton, NC) or lower than 29.63°N (Gainesville, FL).

5. Acknowledgments

This work was funded in part by Corporacion Suiza para el Desarrollo (COSUDE) through a research assistantship for David Elizondo, and by State and Federal funds allocated to Georgia Agricultural Experiment Stations Hatch Projects GEO00526 and GEO01446. The weather data for Tifton, GA, were kindly provided by Dr.

J.E. Hook, Department of Crop and Soil Science, University of Georgia, Tifton. The weather data for Clayton, NC, were kindly provided by Dr. G.G. Wilkerson, Department of Crop Science, North Carolina State University, Raleigh. The weather data from Gainesville, FL, and Quincy, FL, were obtained from the distribution version of the soybean model SOYGRO (Jones et al., 1989).

Trade names and company names are included for the benefit of the reader, and do not imply any endorsement or preferential treatment of the product by the University of Georgia, Corporacion Suiza para el Desarrollo (COSUDE), or Centro Agronómico Tropical de Investigación y Enseñanza (CATIE).

6. References

Bolte, J.P., 1989. Application of neural networks in agriculture. ASAE Paper 89-7591, American Society of Agricultural Engineers, St. Joseph, MI.

Bristow, K.L. and Campbell, G.S., 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agric. For. Meteorol., 31: 159–166.

Chen, S., Yu, D.C. and Moghaddamjo, A.R., 1992. Weather sensitive short-term load forecasting using non-fully connected artificial neural network. Trans. Power Syst., 7: 1098–1105.

Cook, D.F. and Wolfe, M.L., 1991. A back-propagation neural network to predict average air temperatures. AI Appl., 5(1): 40–46.

Davidson, C.S. and Lee, R.H., 1991. Artificial neural networks for automated agriculture. Proc. 1991 Symp. on Automated Agriculture for the 21st Century. American Society of Agricultural Engineers, St. Joseph, MI, pp. 106–115.

Fynn, R.P. and Roller, W.L., 1989. A method of estimating solar irradiance probabilities. ASAE Paper 89-4561, American Society of Agricultural Engineers, St. Joseph, MI.

Hodges, T., French, V. and LeDuc, S.K., 1995. Estimating solar radiation for plant simulation models. AgRISTARS Rep. JSC-20329; YM-15-00403. NOAA–NESDIS–AISC, Columbia, MO.

Hoogenboom, G., White, J.W., Jones, J.W. and Boote, K.J., 1991. BEANGRO V1.01. Dry bean crop simulation model. User's guide. Florida Agricultural Experiment Station Journal N-00379. University of Florida, Gainesville, 122 pp.

Hoogenboom, G., Jones, J.W. and Boote, K.J., 1992. Modeling growth, development and yield of grain legumes using SOYGRO, PNUTGRO, and BEANGRO: a review. Trans. ASAE, 35(6): 2043–2056.

Hook, J.E. and McClendon, R.W., 1992. Estimation of solar radiation data missing from long-term meteorological records. Agron. J., 84(4): 739–742.

International Benchmark Sites Network for Agrotechnology Transfer Project (IBSNAT), 1988. Technical report 1, experimental design and data collection procedures for IBSNAT. The minimum data sets for systems analysis and crop simulation, 3rd (revised) edn. Department of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources, University of Hawaii, Honolulu.

International Benchmark Sites Network for Agrotechnology Transfer Project (IBSNAT), 1990. Technical report 2. Field and laboratory methods for the collection of the IBSNAT minimum data set for the decision support system for agrotechnology transfer (DSSAT V2.1). Department of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources, University of Hawaii, Honolulu.

Jones, H.G., 1992. Plants and Microclimate. A Quantitative Approach to Environmental Plant Physiology. Cambridge University Press, Cambridge.

Jones, J.W., 1991. Crop growth, development and production modeling. Proc. 1991 Symp. on Automated Agriculture for the 21st Century. American Society of Agricultural Engineers, St. Joseph, MI, pp. 447–457.

Jones, J.W., Boote, K.J., Hoogenboom, G., Jagtap, S.S. and Wilkerson, G.G., 1989. SOYGRO V5.42. Soybean crop growth simulation model. User's guide. Florida Agricultural Experiment Station Journal 8304. University of Florida, Gainesville, 75 pp.

Marchal, P., Collewet, C. and Rault, G., 1991. Using a neural network for adaptive control of a milking robot. Proc. 1991 Symp. on Automated Agriculture for the 21st Century. American Society of Agricultural Engineers, St. Joseph, MI, pp. 155–162.

McClelland, J.L., Rumelhart, D.E. and the PDP Research Group, 1986a. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. I: Foundations. MIT Press, Cambridge, MA.

McClelland, J.L., Rumelhart, D.E. and the PDP Research Group, 1986b. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. II: Psychological and Biological Models. MIT Press, Cambridge, MA.

Muttiah, R.S. and Engel, B.A., 1991. Neural network methodology in agriculture and natural resources. ASAE Paper 91-7018, American Society of Agricultural Engineers, St. Joseph, MI.

Neural Computer Sciences, 1992. Time dependent data prediction using neural nets with application to financial forecasting. Application Note F001. Neural Computer Sciences, Southampton, UK.

Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour. Res., 17(1): 182–190.

Richardson, C.W. and Wright, D.A., 1984. WGEN: a model for generating daily weather variables. US Department of Agriculture, Agricultural Research Service, ARS-8, 83 pp.

Rogers, D., 1991. Weather predicting using a genetic memory. In: P. Antognetti and V. Milutinovic (Editors), Neural Networks: Concepts, Applications and Implementations. Prentice–Hall, Englewood Cliffs, NJ, pp. 275–289.

Seco, J., Calvo, A., Egido, M. and Egido, A., 1993. Solar radiation and air temperature: a study of Santiago de Chile and Salamanca. Atmos. Environ., 27A(3): 457–462.

Shih, S.F. and Cheng, K.S., 1989. Generation of synthetic and missing climatic data for Puerto Rico. Water Resour. Bull., 25(4): 829–836.

Smith, W.R. and Somers, G.L., 1993. A system for estimating direct and diffuse photosynthetically active radiation from hemispherical photographs. Comput. Electron. Agric., 8: 181–193.

Spitters, C.J.T., Toussaint, H.A.J.M. and Goudriaan, J., 1986. Separating the diffuse and direct component of global radiation and its implications for modeling canopy photosynthesis. I. Components of incoming radiation. Agric. For. Meteorol., 38: 217–229.

Thai, C.N. and Shewfelt, R.L., 1991. Modeling sensory color quality of tomato and peach: neural networks and statistical regression. Trans. ASAE, 34(3): 950–955.

Thai, C.N., Pease, J.N. and Tollner, E.W., 1991. Determination of green tomato maturity from X-ray computed tomography images. Proc. 1991 Symp. on Automated Agriculture for the 21st Century. American Society of Agricultural Engineers, St. Joseph, MI, pp. 134–143.

Troyo-Diéguez, E., 1992. SUNLIGHT, A program to calculate sunshine duration parameters for evapotranspiration models. Agron. J., 83: 542–545.

Uhrig, J.W., Engel, B.A. and Baker, W.L., 1992. An application of neural networks: predicting corn yields. In: D.G. Watson, F.S. Zazueta and A.B. Bottcher (Editors), Computers in Agricultural Extension Programs. Proc. 4th Int. Conf. American Society of Agricultural Engineers, 28–31 January 1992, Orlando FL, St. Joseph, MI, pp. 348–353.

Vemuri, V., 1988. Artificial Neural Networks: Theoretical Concepts. Computer Society Press of IEEE, Washington, DC, 145 pp.

Vives, L., 1986. Agrometeorologia. In: Agroambientes. Centro Agronómico Tropical de Investigación y Enseñanza, Turrialba, Costa Rica, pp. 11–133.

Whittaker, A.D., Park, B.S. and McCauley, J.D., 1991. Ultrasonic signal classification for beef quality grading through neural networks. Proc. 1991 Symp. on Automated Agriculture for the 21st Century. American Society of Agricultural Engineers, St. Joseph, MI, pp. 116–125.

Zhuang, X. and Engel, B.A., 1990. Neural networks for applications in agriculture. ASAE Paper 90-7024, American Society of Agricultural Engineers, St. Joseph, MI.