SPATIAL INTERPOLATION OF WEATHER VARIABLES USING ARTIFICIAL NEURAL NETWORKS

by

Bin Li

Under the direction of Dr. Ron W. McClendon

ABSTRACT

Crop growth simulation models use weather data such as temperature, solar radiation, and rainfall to simulate crop development and yield. The crop models are often needed for locations with missing or incomplete observed weather data. An accurate estimation of these weather variables has thus become necessary. Artificial neural network (ANN) models could be used to accurately estimate these weather variables. In this study, ANN-based methods were developed to estimate daily maximum and minimum air temperature and total solar radiation for locations in Georgia. Observed weather data from 1996 to 1998 were used for model development, and data from 1999 to 2000 were used for final ANN model evaluation. In the ANN model development, the preferred number of input weather stations and the input variables for estimating each weather variable were determined. The ANNs provided higher accuracy than the traditional average, inverse distance, and multi-linear regression methods.

INDEX WORDS: Artificial neural network, ANN, Maximum air temperature, Minimum air temperature, Solar radiation, Estimation, Georgia

SPATIAL INTERPOLATION OF WEATHER VARIABLES USING ARTIFICIAL NEURAL NETWORKS

by

BIN LI

B.ENG., The University of Wuhan, China, 1987

M.ENG., Asian Institute of Technology, Thailand, 1994

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2002

SPATIAL INTERPOLATION OF WEATHER VARIABLES USING ARTIFICIAL NEURAL NETWORKS

by

BIN LI

Approved:

Major Professor: Ron W. McClendon

Committee: Gerrit Hoogenboom
Suchi Bhandarkar

Electronic Version Approved:

Gordhan L. Patel
Dean of the Graduate School
The University of Georgia
August 2002

ACKNOWLEDGMENTS

I would like to acknowledge my committee, without their support, this research would not have been possible. I would like to thank Dr. Ron McClendon, my major professor, and Dr. Gerrit Hoogenboom for their guidance in the research and writing; I would also like to thank Dr. Suchi Bhandarkar for taking the time to serve on my graduate committee.

I would like to thank Dr. Don Potter for his support throughout my graduate study. I would also like to thank Ms. Angie Paul for her support during my graduate study in Artificial Intelligence Center.

I would like to express my gratitude to Dr. Edward Kanemasu and Ms. Karen Kanemasu for their support and encouragement throughout my graduate study.

I would like to thank my family for their love, encouragement, and support in all of my endeavors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Computer simulation models are becoming increasingly important tools in research and decision making related to agricultural production. Most crop growth simulation models use weather data inputs such as daily maximum and minimum temperatures, solar radiation, and rainfall to simulate crop yields (Amir and Sinclair 1991; Hoogenboom et al., 1992; Chapman et al., 1993). Air temperature is an important input to a variety of spatially distributed hydrological and ecological models (Cook and Wolfe, 1991; Dodson and Marks, 1997; Bolstad et al., 1998; Scheifinger and Kromp-Kolb, 2000). In addition, air temperature has been used to analyze climate change and the effects of the change (Robeson and Janis 1998; Michaels et al., 1998; Goodale et al., 1998; Price et al., 2000). Moreover, most processes in the atmosphere and biosphere, such as evaporation, sensible heat flux, soil heat flux, are driven directly or indirectly by solar radiation (Bruton et al., 2000; Scheifinger and Kromp-Kolb, 2000).

In crop-growth simulation model development, it is assumed that the future climate will fit the same distribution as the historical climate used in the analyses. Most current models require long-term daily weather records with a high spatial resolution with matched temporal resolution. However, for some areas weather measurements are not available due to the high cost of instrumentation, maintenance and calibration. For some areas, only a limited period of records is available. Therefore it is highly desirable to develop accurate weather data estimation models for use in simulation, weather analysis, and decision support applications.

Traditional interpolation methods include Thiessen polygons, inverse distance interpolations, kriging, splines, and regression model. The Thiessen polygon or Voronoi polygon has been widely used in climatological rainfall and precipitation estimations (Goovaests, 2000; Wilk and Andersson, 2000; Pardo-Iguzquiza, 1998; Dirks et al, 1998; Panagoulia, 1992). Dirk et al. (1998) compared the Thiessen polygon with kriging spatial interpolation method to estimate spatially continuous rainfall. They found Thiessen polygon method was comparable with kriging. Goovaerts (2000) pointed out that the Thiessen polygon method did not consider the elevation effects and rainfall records from surrounding stations, therefore, large prediction errors could occur in the prediction of rainfall.

Inverse-distance is a weighting interpolation method. The number of neighbors necessary in the weighting function is important in terms of reducing computation time while maintaining a smooth surface. Dodson and Marks (1997) have suggested that with inverse-squared-distance interpolation using eight nearest neighbors is reasonable. Robeson (1993) investigated three methods of spatially interpolating temperature anomaly data. He found that the inverse-distance method gave about the same results as triangulated surface patches and thin-plate splines. In order to consider the elevation effects on climate, gradient plus inverse-distance-squared (GIDS) interpolation technique was derived (Nalder and Wein, 1998, Price et al., 2000) from the inverse-distance-square method. Price et al. (2000) used gradient plus inverse-distance-squared method to interpolate Canadian monthly mean climate data. It was suggested that this method is attractively simple and appears to give results adequate for modeling long term forest ecosystem responses to climate in relatively flat terrain.

Regression has been used successfully in weather data estimation (Ollinger et al. 1993, 1995). Bolstad et al. (1998) used a regression approach as a main method to predict air temperature and compared regression approach with local lapse models or kriging method. They stated that the regression approach provided an accurate estimate of station temperature. Christine et al. (1998) used a regression technique to predict the monthly precipitation, monthly averaged maximum and minimum temperature, and monthly averaged sunshine hours and compared the regression approach with a modified inverse-distance-square interpolation. They reported that the prediction accuracy did not differ between these two methods.

Kriging has been implemented in analysis of climatologic factors, such as the analysis of precipitation (Dingman et al. 1988), evapo-transpiration (Martinez-Cob & Cuenca 1992), and temperature (Holdaway 1996, Bolstad et al. 1998). Holdaway (1996) has applied kriging to the spatial interpolation of monthly temperature. In his research, monthly empirical variograms, averaged over 90 years, were modeled with Gaussian or linear models in the ordinary Kriging method. He concluded that anisotropies were found in the winter months, suggesting the presence of a large-scale regional trend. Bolstad et al. (1998) used kriging (co-kriging) to predicate air temperatures. They found that Kriging and co-kriging may be particularly appropriate for temperature predictions in regions with little topographic relief, but not useful where temperature measurement stations were sparse or high terrain effects were existing.

Splining has been discussed by some researchers (Eckstein 1989; Hutchinson and Gessler 1994). Hutchinson et al. (1994) used multi-dimensional thin plate splines to fit temperature surfaces by minimizing the roughness of the interpolated surface. He reported that the thin plate spline method worked as good as Kriging while requiring less parameterization, however thin plate splines are computationally demanding and complicated to implement. Price et al. (2000) employed the thin-plate smoothing splines to interpolate Canadian monthly mean climate data. The GIDS was used to compare with the thin-plate smoothing splines. They found that thin plate smoothing splines produced better results for the west region where predicting precipitation is difficult.

Artificial neural networks (ANNs) are potential alternative to estimate such weather data. ANNs are computer models that mimic the structure and functioning of the human brain (Ward Systems Group Inc., Frederick, MD, 1993). ANNs can determine the relationships among the independent variables to predict or estimate dependent variables. Back propagation (BP) ANNs are known for their ability to generalize well on a wide variety of problems and are well suited for prediction applications. Unlike statistical methods, ANN models do not make dependency assumptions among input variables and solves multivariate problem with nonlinear relationship among input variables. This technique has been used in a wide range of applications, such as classification, pattern recognition, automatic control and function approximation (McAvoy et al., 1989; Leonard et al., 1992, Rao & Gupta, 1993). Han and Felker (1997) implemented an ANN to estimate daily soil water evaporation from average relative air humidity, air temperature, wind speed, and soil water content in a cactus field study. They found that the ANN achieved a good agreement between predicted and measured values. They concluded that the ANN technique appeared to be an improvement over the multi-linear regression technique for estimating soil evaporation. Elizondo et al. (1994) used an ANN to estimate daily solar radiation for locations in the southeastern US based on daily maximum and minimum air temperature, daily total precipitation daily clear sky radiation and day length for that location. They did not include weather data from other locations as inputs. They found r^2 of 0.74 and a root mean square error of 2.92 MJ/m^2. Cook and Wolfe (1991) developed a neural network to predict average air temperatures. In their study, the monthly average of daily maximum temperatures for three months in advance was predicted. Bruton et al. (2000) developed ANN models for estimating daily pan evaporation. The results were compared with those of multiple linear regression and Priestly-Taylor model and they found that the ANN model provided the highest accuracy.

The goal of this research was to develop ANN-based methods to estimate daily weather data for locations in Georgia based on daily weather data from neighboring weather stations as inputs. The specific objectives were: 1) develop localized ANN models to estimate daily maximum and minimum air temperature and total solar radiation specifically for Tifton (south Georgia) and Griffin (north Georgia), 2) develop general ANN models to estimate daily maximum and minimum air temperature and total solar radiation for locations throughout Georgia, 3) determine the number of known weather stations required as inputs for estimating each weather variable, 4) determine which inputs are required for each weather variable, and 5) compare these ANN models with traditional methods such as averaging, multi-linear regression and inverse distance weighting interpolation methods.

CHAPTER 2

SPATIAL INTERPOLATION OF WEATHER VARIABLES FOR SINGLE LOCATIONS USING ARTIFICIAL NEURAL NETWORKS

Li, B., R.W. McClendon, G. Hoogenboom, to be submitted to Computers and Electronics in Agriculture.

ABSTRACT

Crop growth simulation models use weather data such as temperature, solar radiation, and rainfall to simulate crop development and yield. The crop models are often needed for locations with missing or incomplete observed weather data. An accurate estimation of these weather variables has thus become necessary. Artificial neural network (ANN) models can be used to accurately interpolate these weather variables, based on neighboring weather stations. The goal of this study was to develop artificial neural network models for estimating and interpolating daily maximum air temperature, minimum air temperature, and total solar radiation for Tifton (south Georgia) and Griffin (north Georgia). Historical daily weather data from 1996 to 1998 were used for model development, and data from 1999 to 2000 were used for final model evaluation.

The results of the study indicated that the preferred input variables were straight line distance (Δs) and the elevation difference (Δz) between the target location and input weather stations as well as the values of the variable being estimated at the input stations. Maximum temperature was also found to improve the accurate in estimating solar radiation. The optimum number of input weather stations for estimating each weather variable for both target locations was also determined. The best models for estimating these weather variables were compared with other spatial interpolation techniques, including inverse distance, average, and multi-linear regression methods. The results showed that ANN and regression models provided superior accuracy over inverse distance and average methods. ANN models and regression models were comparable in estimating maximum temperature. The ANN model was clearly more accurate than

regression in estimating minimum temperature at Griffin and comparable at Tifton. The ANN models were superior in estimating solar radiation at both locations. Using the evaluation data set, the Tifton models had mean absolute error (MAE) values as follows: maximum temperature, 0.61ºC; minimum temperature, 0.74ºC; and solar radiation, 1.24 MJ/m^2. The Griffin models had MAE values as follows: maximum temperature, 0.36ºC; minimum temperature, 0.82ºC; and solar radiation, 1.51 MJ/m^2. ANN models thus provided an accurate approach for estimating daily weather variables using data from neighboring weather stations for a particular location.

INTRODUCTION

Computer simulation models are becoming increasingly important tools in research and decision making related to agricultural production. Most crop growth simulation models use weather data inputs such as daily maximum and minimum temperatures, solar radiation, and rainfall to simulate crop yields (Amir and Sinclair 1991; Hoogenboom et al., 1992; Chapman et al., 1993). Air temperature is an important input to a variety of spatially distributed hydrological and ecological models (Cook and Wolfe, 1991; Dodson and Marks, 1997; Bolstad et al., 1998; Scheifinger and Kromp-Kolb, 2000). In addition, air temperature has been used to analyze climate change and the effects of the change (Robeson and Janis 1998; Michaels et al., 1998; Goodale et al., 1998; Price et al., 2000). Moreover, most processes in the atmosphere and biosphere, such as evaporation, sensible heat flux, soil heat flux, are driven directly or indirectly by solar radiation (Bruton et al., 2000; Scheifinger and Kromp-Kolb, 2000).

In crop growth simulation model development, it is assumed that the future climate will fit the same distribution as the historical climate used in the analyses. Most current models require long-term daily weather records with a high spatial resolution with matched temporal resolution. For some areas, however, weather measurements are not available due to the high cost of instrumentation, maintenance and calibration. Therefore it is highly desirable to develop accurate weather data estimation models for use in simulation, weather analysis, and decision support applications.

Traditional interpolation methods to estimate weather data include inverse distance interpolations and regression models. Inverse-distance is a weighting interpolation method. The number of neighbors necessary in the weighting function is important in terms of reducing computation time while maintaining a smooth surface. Dodson and Marks (1997) have suggested that with inverse-squared-distance interpolation using eight nearest neighbors is reasonable. Robeson (1993) investigated three methods of spatially interpolating temperature anomaly data. He found that the inverse-distance method gave about the same results as triangulated surface patches and thin-plate splines. In order to consider the elevation effects on climate, gradient plus inverse-distance-squared (GIDS) interpolation technique was derived (Nalder and Wein, 1998, Price et al., 2000) from the inverse-distance-square method. Price et al. (2000) used gradient plus inverse-distance-squared method to interpolate Canadian monthly mean climate data. It was suggested that this method is attractively simple and appears to give results adequate for modeling long term forest ecosystem responses to climate in relatively flat terrain.

Regression has been used successfully in weather data estimation (Ollinger et al. 1993, 1995). Bolstad et al. (1998) used regression to predict air temperature and compared their approach with local lapse models or kriging methods. They stated that the regression approach provided an more accurate estimate of station temperature. Christine et al. (1998) used a regression technique to predict the monthly precipitation, monthly averaged maximum and minimum temperature, and monthly averaged sunshine hours and compared the regression approach with a modified inverse-distance-square interpolation. They reported that the prediction accuracy did not differ between these two methods.

Artificial neural networks (ANNs) provide a potential alternative to estimating weather data. ANNs are computer models that mimic the structure and functioning of the human brain (Ward Systems Group Inc., Frederick, MD, 1993). ANNs can determine the relationships among the independent variables to predict or estimate dependent variables. Back propagation ANNs are known for their ability to generalize well for a wide variety of problems and are well suited for prediction applications. Unlike statistical methods, ANN models do not make dependency assumptions among input variables and solve multivariate problem with nonlinear relationship among input variables. This technique has been used in a wide range of applications, such as classification, pattern recognition, automatic control and function approximation (McAvoy et al., 1989; Leonard et al., 1992, Rao & Gupta, 1993). Han and Felker (1997) implemented an ANN to estimate daily soil water evaporation from average relative air humidity, air temperature, wind speed, and soil water content in a cactus field study. They found that the ANN achieved a good agreement between predicted and measured values. They concluded that the ANN technique appeared to be an improvement over the multi-linear regression technique for estimating soil evaporation. Cook and Wolfe (1991) developed a neural network to predict average air temperatures for a single location. In their study, the monthly average of daily maximum temperatures for three months in advance was predicted. Bruton et al. (2000) developed ANN models for estimating daily pan evaporation. The results were compared with those of multi-linear regression and Priestly-Taylor model and they found that the ANN model provided the highest accuracy.

The goal of this research was to develop ANN models to estimate the spatial interpolation of daily weather variables for two specific locations in Georgia. The objectives were: 1) to develop ANN models for estimating daily maximum and minimum temperature and solar radiation specifically for Tifton, i.e. south Georgia and Griffin, i.e. north Georgia, 2) to determine the number of neighboring weather stations required as input, 3) to determine the inputs required for each estimated weather variable, and 4) to compare these ANN models with traditional multi-linear regression, averaging, and inverse distance weighting interpolation methods.

MATERIALS AND METHODS

Weather data for this study were obtained from the Georgia Automated Environmental Monitoring Network (AEMN) (Hoogenboom, 1996, 2000a, 2000b; Hoogenboom et al., 2000). The AEMN is a network that consists of more than 45 automated weather stations, located across the state of Georgia (www.Georgiaweather.net). Sensors in the AEMN are polled with a one-second frequency and averages or totals are calculated every 15 minutes. At midnight daily extremes and totals are determined for each weather variable. The data are downloaded to a central computer in Griffin twice a day. The weather variables that are measured include air temperature, rainfall, soil temperature at three depths, relative humidity, wind speed and direction, solar radiation and open pan evaporation. The locations of the weather stations near Tifton, Georgia and Griffin, Georgia used in this study for developing ANN models are listed in Table 2.1.

Daily weather data from 1996 to 2000 were divided into two separate overall data sets for model development and model evaluation. The first dataset consisted of the weather data collected from 1996 to 1998 (1096 daily weather observations) and was used for model development. The second dataset consisted of data collected from 1999 to 2000 (731 daily weather observations) and was kept separate as an independent dataset. It was used only for evaluation of the final ANN models. Computer programs were developed in the Java programming language to create input data files for the ANN models. The potential input variables that were considered included daily maximum air temperature, daily minimum air temperature, daily total solar radiation, and the fractions that a day of year belongs to each season (spring, summer, fall, and winter). Other inputs that were considered included the difference in elevation (Δz), the difference in east-west direction (Δx), and the difference in north-south direction (Δy). The straight line distance between the target and known location (Δs) was also considered as a substitution for Δx and Δy. The values for the daily maximum and minimum temperatures and solar radiation were the observed data with the day of the year as an index. The other values were computed data from the topographic data and day of year. During preprocessing, the calculated data were computed first, based on the appropriate equations. The daily observed data were then combined with the calculated data. Finally, the dataset was randomized by day of the year.

The computer program also included a routine to select neighboring stations for estimating the weather data at the target location. The stations were selected based on the shortest distance from the target location. Data from these weather stations were organized into each pattern in the data set, but the order of weather stations for each pattern was randomly distributed. This was done to maintain some generality by location. Each pattern could be reproduced by altering the order of the weather station data. Therefore, more patterns could be generated for developing the ANN models thus increasing the ability of the model to generalize. If for each pattern, there is another reproduced pattern created, the total number of patterns for developing the ANN models would be doubled. After data preprocessing, each pattern consisted of the three daily weather variables per input location and target location and the calculated variables of Δx, Δy, Δs, Δz and four seasonal effect terms.

The back propagation approach is based on gradient descent and is designed to minimize the mean square error between the observed and estimated outputs. In this study, the two back propagation ANNs that were considered included standard back propagation and Ward ANNs. Standard back propagation ANN is a three-layer back propagation ANN consisting of an input layer, a hidden layer, and an output layer. Ward ANN is also a three-layer back propagation, except that the Ward ANN has multiple slabs in the hidden layer (Ward System Group Inc., Frederick, MD, 1993). Ward ANN allows the user to select different activation functions such as the Gaussian function, hyperbolic tangent function, and Gaussian-complement function for each slab. With different activation functions, these networks could possibly detect different features of the input vectors and are therefore well suited for prediction.

The NeuroShell™ (Ward System Group Inc., Frederic, MD, 1993) is an ANN software package with a user friendly menu-driven infrastructure. It has several utilities for data manipulation, model development, graphical options, and a runtime option to generate source code. One of the important options of NeuroShell™ is its optimal network option which helps to prevent over-training. Over-training is critical issue in prediction applications because the neural network should be capable of generalizing over a similar data set as opposed to only mimicing the training set. NeuroShell™ allows the user to set an interval for the program to check the accuracy of the current ANN with a separate testing data set. If the average error of the prediction of the ANNs improves over the previous optimal ANN, the older optimal ANN is replaced with the model developed with the least error. Otherwise, the previous optimal ANN can be saved until an improved optimal ANN is reached, or training is completed.

In our study, the model development data set consisted of data from year 1996 to 1998 for a total of 1096 daily observations. It was divided into three data sets i.e., training, testing, and production with a distribution of 50%, 25%, and 25%, respectively. We duplicated each pattern by altering the order of weather stations, thus the actual number of patterns in model development was 2192. Therefore, the training, testing, and production data sets consisted of 1096, 548, and 548 patterns, respectively. The NeuroShell™ software package was used to develop the ANN models to estimate the three daily weather variables at the two target locations. Computer experiments were conducted to select the optimum architecture and parameters, including number of hidden nodes, learning rate, momentum, stopping criteria, and calibration interval for the ANN models. These parameters and architecture were selected based on the coefficient of determination (r^2), mean absolute error (MAE), and root mean square error (RMSE) of model results using production data sets.

Computer experiments were conducted to determine which input variables were important for the ANN models. Different combinations of input variables were examined to determine the ANN model that had the lowest estimation error for the production dataset. This was accomplished by adding an input variable for the model, retraining and analyzing the effects on r^2, MAE, and RMSE. Two different types of input variables were considered, e.g., the daily weather variables and variables which could be calculated knowing the locations of the weather stations and day of year.

For the development of the ANN model to estimate maximum temperature, the following approach was taken to determine the importance of the different inputs. An ANN model was initially developed consisting of only maximum temperature inputs at surrounding weather station locations. The accuracy of this model was compared to one developed with the additional input of Δs. The more accurate of these two models was then compared to one with Δx and Δy added as inputs. This process was repeated with the inclusion of minimum daily temperature, the seasonal effect terms and Δz added as inputs A similar approach was taken to develop the minimum temperature and solar radiation models. As the number of input weather stations was varied, a search was performed to determine the preferred number of hidden modes for each model.

The number of weather stations required as input was also an important factor in developing the optimal ANN models. Files were created with an increasing number of input weather stations for each pattern and different ANN models were developed based on these input files. The number of nearest weather stations required was chosen based upon the accuracy of the developed models. All decisions regarding the preferred ANN architecture, inputs, and number of input weather station locations were made using the model accuracy on the production data set.

Final model evaluation was based on an independent dataset of daily weather data from year 1999 to 2000 that included a total of 731 daily observations. This data set was prepared similarly to the model development data set, except that it did not include any reproduced patterns. In addition, all patterns were arranged in the order of day of year. For model evaluation, a one-time feed forward mode was used for the ANN models with the highest accuracy on the production dataset. This final evaluation was conducted to determine if the localized models developed for Tifton and Griffin were able to estimate the weather variables for that particular location. A crosscheck was also performed to determine if a model developed for maximum temperature, minimum temperature, or solar radiation for one location could accurately estimate the same variable at the other location. In this evaluation, the Tifton models were used with the Griffin evaluation dataset and Griffin models were used with the Tifton data set.

The results of the ANN models were also compared with the results of other spatial analysis techniques that included average, inverse distance, and multi-linear regression methods on the evaluation dataset. The purpose of this comparison was to determine if ANN models could provide a higher accuracy for estimating maximum temperature, minimum temperature, and solar radiation compared to the more traditional methods.

RESULTS AND DISCUSSION

ANN models were initially developed with both the standard three-layer back propagation architecture and Ward network architecture. For the same conditions, the models with Ward network architecture had a higher r^2 and smaller MAE and RMSE than the standard back propagation architecture. Based on the results of this preliminary study, the Ward network architecture was selected for all further ANN model development. In this preliminary study, we also found that a variation in the number of hidden nodes was the only ANN parameter that had a significant effect on model accuracy. For all subsequent model development, a learning rate of 0.1, momentum of 0.1, stopping criteria of 20,000 events past the minimum test set error, and a test interval of 200 learning events were used.

A heuristic search was performed to determine the best values for the number of input weather stations, input variables, and number of hidden nodes for each weather variable ANN at the two target locations. The approach used to develop the maximum temperature ANN at Tifton is discussed as an example, but applies to all variables for both locations. The inputs consisted of maximum temperatures from five known weather stations surrounding the Tifton target location (Table 2.1). The number of hidden nodes was then varied to maximize r^2 and minimize MAE and RMSE. Subsequently the number of closest weather station was reduced by one and new models were developed. The optimal number of hidden nodes was determined for each value of the number of weather stations. Once the best model was determined using only maximum temperature as an input, additional inputs were considered. The same search was repeated in terms of varying the number of weather stations and number of hidden nodes.

From the search results it was determined that the preferred inputs for a particular weather variable ANN at Tifton were the same for Griffin. In each case, the weather variable being estimated was important as an input from the known stations. For example, when developing a maximum temperature ANN, maximum temperature from the surrounding weather stations was always an important input. Also the inclusion of Δs and Δz improved the accuracy in the individual estimation of the three weather variables. For the solar radiation ANN, it was determined that maximum temperature at the known weather stations was also important to help increase the accuracy.

The number of nearest weather stations required to maximize the accuracy varied by weather variable as well as by target location. For example, Table 2.2 shows the results for the maximum temperature ANN for Tifton in which the search was varied from one to five input weather stations. Two input weather stations provided the highest overall accuracy for estimating maximum temperature at Tifton using maximum temperature, Δs, and Δz as inputs. Results are shown for the training, testing, and production data sets, although only the production data set was used in selecting the preferred value. Table 2.3 shows the results for the search for the preferred number of hidden nodes for the case of two input weather stations. Eight hidden nodes was the value which produced the highest accuracy on the production set. Similar searches were performed for all three weather variables at both target locations. Table 2.4 summarizes the results of the search for the number of input weather stations and hidden nodes for each of the weather variable models by location. The MAE was analyzed for the three weather variables as a function of the number of input weather stations for the Tifton location (Figure 2.1 A). The minimum MAE for maximum temperature and solar radiation was clearly at two input weather stations. For minimum temperature there was a slight reduction in error at five stations. The results for the search for the Griffin location were somewhat different (Figure 2.1 B). Maximum temperature and solar radiation had a minimum MAE with only one nearest station as input, whereas minimum temperature had a minimum MAE for two nearest stations as input. From this analysis it was determined that more weather station locations were needed as input for estimating minimum temperature than for estimating maximum temperature or solar radiation.

The nearest weather station for the Griffin location was Williamson, with a straight line distance of only 14.92 km. The results indicated that this station was the only one needed as an input for estimating maximum temperature and solar radiation. The nearest weather station to Tifton was Dawson at a distance of 91.03 km. Two weather stations were required for estimating maximum temperature and solar radiation and five for minimum temperature at Tifton. Fewer stations were consistently needed for the Griffin target location due to the close proximity of the Williamson weather station. When only one weather station was needed, Δs and Δz were eliminated as inputs because they were then constants and did not contribute to model accuracy. Therefore, only maximum temperature was used as an input variable for estimating maximum temperature at Griffin. Similarly, only solar radiation and maximum temperature were used as input variables for estimating solar radiation at Griffin.

The ANN models for the three weather variables developed for each of the two locations were then used with the evaluation dataset with the results shown in Table 2.5. This dataset had not been used in model development or in the selection of the preferred network architecture or parameters. We found that the ANN models produced estimations for maximum temperature, minimum temperature, and solar radiation which were comparable to the results obtained for the model development dataset. For Tifton, the model for estimating maximum temperature had an r^2 of 0.987, MAE of 0.61ºC, and RMSE of 0.84ºC The estimated maximum temperature versus observed maximum temperature for Tifton for the evaluation dataset is shown in Figure 2.2 A. A linear regression was performed to determine how well the ANN results matched the observed maximum temperature. The ANN model estimated maximum temperature well for the entire temperature range. The model for estimating minimum temperature had an r^2 of 0.987, a MAE of 0.74ºC, and a RMSE of 0.93ºC. The estimated minimum temperature versus observed minimum temperature for Tifton for the evaluation dataset is shown in Figure 2.2 B. The ANN model tended to slightly overestimate minimum temperature at lower minimum temperatures. However, for all other temperature ranges, the ANN model estimated minimum temperature well. The intercept of the regression equation was significantly different from 0 at 95% confidence level. The model for estimating solar radiation had an r^2 of 0.944, a MAE of 1.24 MJ/m^2, and a RMSE of 1.74 MJ/m^2. The estimated solar radiation versus observed solar radiation for Tifton for the evaluation data set is shown in Figure 2.2 C. The ANN model tended to overestimate solar radiation for low observed values and underestimate solar radiation for high values. For intermediate values the ANN worked well. The intercept of the regression equation was significantly different from 0 at the 95% confidence level (Table 2.5).

The model for estimating maximum temperature at Griffin had an r^2 of 0.997, a MAE of 0.36ºC, and a RMSE of 0.51ºC. The estimated maximum temperature as a function of observed maximum temperature of Griffin for the evaluation data set is shown in Figure

2.3 A. The Griffin ANN model estimated maximum temperature well for the entire temperature range. The model for estimating minimum temperature had an r^2 of 0.984, a MAE of 0.82ºC, and a RMSE of 1.07ºC. The estimated minimum temperature as a function of observed minimum temperature of Griffin for the evaluation data set is shown in Figure 2.3 B. The ANN model for estimating minimum temperature at Griffin performed well for all observed values, although the variation was greater than for the maximum temperature model. The model for estimating solar radiation had an r^2 of 0.958, a MAE of 1.51 MJ/m^2, and a RMSE of 2.09 MJ/m^2. The estimated solar radiation versus observed solar radiation of Griffin for the evaluation data set is shown in Figure 2.3 C. The ANN model slightly overestimated solar radiation at extremely low values and underestimated solar radiation at higher values. The intercept of the regression equation was significantly different from 0 at 95% confidence level (Table 2.5). In general, the ANN models accurately estimated maximum temperature, minimum temperature, and solar radiation for Tifton and Griffin (Table 2.5).

In the crosscheck evaluation it was determined that, a model developed specifically for one location did not accurately estimate the weather variable at another location (Table 2.5). For example, when Tifton data was applied to Griffin models, the ANN model had an r^2 of 0.545, a MAE of 9.62ºC, and a RMSE of 10.83ºC for estimating maximum temperature. The Griffin ANN models using Tifton data had an r^2 of 0.859, a MAE of 3.67ºC, and a RMSE of 4.32ºC for estimating minimum temperature and an r^2 of 0.680, a MAE of 12.65 MJ/m^2, and a RMSE of 13.31 MJ/m^2 for estimating solar radiation. Therefore Griffin models were not useful for estimating the weather variables for Tifton. Similarly, when Griffin data was applied to Tifton models, the ANN models had an r^2 of 0.930, a MAE of 4.75ºC, and a RMSE of 5.24ºC for estimating maximum temperature, an r^2 of 0.964, a MAE of 3.40ºC, and a RMSE of 3.75ºC for estimating minimum temperature, and an r^2 of 0.907, a MAE of 3.15 MJ/m^2, and a RMSE of 4.03 MJ/m^2 for estimating solar radiation.

Table 2.6 presents a comparison between ANN, and other traditional spatial interpolation techniques, including averaging, inverse distance, and multi-linear regression for estimating the three weather variables. We found that the r^2 values were comparable for all four methods. The ANN and regression models were consistently more accurate than the average or inverse distance methods for all comparisons. The ANN and regression models were generally comparable in estimating maximum temperature for both Tifton and Griffin. The ANN model was clearly more accurate than regression in estimating minimum temperature at Griffin and comparable to regression for the Tifton location. The ANN models were clearly more accurate than the regression models for estimating solar radiation at both locations. Overall the ANN models were comparable to or more accurate than regression method for the three weather variables at both locations.

SUMMARY AND CONCLUSION

From the results of this study it can be concluded that the ANN models produced the highest overall accuracy in estimating maximum and minimum temperatures and as solar radiation for a single location. The ANN models for solar radiation were consistently more accurate than regression method for both locations that were tested. For estimating maximum temperature, both the ANN model and the regression method were generally comparable. For estimating minimum temperature, the ANN model was clearly more accurate than the regression for the Griffin location and comparable to regression for Tifton.

The preferred inputs for each weather variable were the same for both locations. In each case, Δs and Δz were important along with the corresponding weather variable being estimated. For the solar radiation models, the addition of maximum temperature also improved the accuracy. Although the models developed for a specific location performed well in estimating daily weather at that location, they were inaccurate when attempting to estimate the results at the other location.

For further research, the development of a generalized model to estimate the weather variables throughout the state could be considered. This model would be developed using multiple target locations in both the model development and evaluation data sets. The optimal number of input stations and input variables could be determined in a similar manner to the current study. The accuracy of the generalized models could then be compared to the localized models of this study.

REFERENCES

Amir, J. and Sinclair, T.R., 1991. A model of the temperature and solar-radiation effects on spring wheat growth and yield. Field Crops Res. 28, 47-58.

Bolstad, P. V., Swift, L., Collins, F., Regniere, J., 1998. Measured and predicted air temperatures at basin to regional scales in the southern Appalachian mountains. Agric. & For. Met. 91, 161-176

Bruton, J.M., McClendon, R.W., Hoogenboom, G., 2000. Estimating daily Pan Evaporation with Artificial neural networks. Trans. ASAE 43(2), 491-6.

Chapman, S.C., Hammer, G.K. and Meinke, H., 1993. A crop simulation model for sunflower. I. Model development. Agron. J. 85, 725-735.

Christine L. Goodale, Jhon D. Aber, Scott V. Ollinger, 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. Clim. Res. 10, 35-49.

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

Dirks, K.N., Hay, J.E., Stow, C.D., Harris, D., 1998. High-resolution studies of rainfall on Norfolk Island. J. Hydrol. 208(3-4), 187-193.

Dodson, R. and Marks, D., 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. Clim. Res. 8, 1-20.

Goodale, C.L., Aber, J.D., Olllinger, A.V., 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. Clim. Res. 10, 35-49.

Han, H. and Felker, P., 1997. Estimation of daily soil water evaporation using an artificial neural network. Journal of Arid Environments 37, 251-260.

Hoogenboom, G. 1996. The Georgia Automated Environmental Monitoring Network. In: Preprints of the 22nd Conf. On Agricultural and Forest Meteorology, 343-346. Boston, Mass., American Meteorological Society.

Hoogenboom, G. 2000a. Contribution of agrometeorology to the simulation of crop production and its applications. Agric. & For. Met. 103(1-2), 137-157.

Hoogenboom, G. 2000b. The Georgia Automated Environmental Monitoring Network. p. 24-25. *In*: Preprints 24rd Conference on Agricultural and Forest Meteorology. American Meteorological Society, Boston, Massachusetts.

Hoogenboom, G., Georgiev, G.A., and Gresham, D.D., 2000. Development of weather based products for agricultural and environmental applications. p. 66-67. *In*: Preprints 24rd Conference on Agricultural and Forest Meteorology. American Meteorological Society, Boston, Massachusetts.

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.

Leonard, J.A., Kramer, M.A. & Ungar, L.H., 1992. Using radial basis functions to approximate function and its error bonds, IEEE Trans. on Neural Networks 3, 624-627.

McAvoy, T.J., Wang, N.S., Naidu, S., Bhar, N., Gunter, J. and Simmons, M., 1989. Interpreting biosensor data via backpropagation. IEEE International Neural Network Conference in 1989, 8, 227-233.

Michaels, P.J., Balling Jr. R. C., Vose, R.S., Knappenberger, P. C., 1998. Analysis of trends I the variability of daily and monthly historical temperature measurements. Clim. Res. 10, 27-23.

Nalder, I.A., Wein, R.W., 1998. Spatial interpolation of climatic normal: test of a new methods in the Canadian boreal forest. Agric. & For. Met. 9, 211-225.

Ollinger, S.V., Aber, J.D., Lovett, G. M., Millham, S.E., Lathrop, E.G., Ellis, J.M., 1993. A Spatial model of atmospheric deposition for the northeastern U.S. Ecol. Appl. 3, 459-472.

Panagoulia, D., 1992. Hydrological modeling of a medium-size mountainous catchment from incomplete meteorological data. J. Hydrol 137(1-4), 279-310.

Parmar, R.S., McClendon, R.W., Hoogenboom, G., Blankenship, P.D., Cole, R.J., Forner, J.W., 1997. Estimation of aflatoxin contamination in preharvest peanuts using neural networks. Trans. ASAE 40(3), 809-813.

Pardo-Iguzquiza, E., 1998. Comparison of geostatistical methods for estimating the areal climatological rainfall mean using data on precipitation and topography. Int. J. Climatol. 18, 1031-1047.

Price, D.T., McKenney, D.W., Nalder, I. A., Hutchinson, M. F., Kesteven, J. L., 2000. A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agric. & For. Met. 101, 81-94.

Rao, D.H. and Gupta, M.M., 1993. Dynamic neural controller with somatic adaptation. IEEE International Conference on Neural Networks in 1993, 8, 558-563.

Rich, E. and Knight, K., 1991. Artificial intelligence. New York: McGraw-Hill.

Robeson, S.M., 1993. Spatial interpolation, network bias, and terrestrial air temperature variability. Publ Climatol 46(1), 1-51

Robeson, S.M. and Janis, M. J., 1998. Comparison of temporal and unresolved spatial variability in multiyear time-averages of air temperature. Clim. Res. 10, 15-26.

Schaap, M.G. and Bouten, W., 1996. Modeling water retention curves of sandy soils using neural networks. Water Resour. Res. 32(10), 3033-3040.

Scheifinger, H. and Kromp-Kolb, H., 2000. Modeling global radiation in complex terrain: comparing two statistical approaches. Agric. For. Met. 100: 127-136.

Ward System Group, Frederic, MD, 1993. Manual of NeuroShell 2. Frederic, MD: Ward Systems Group.

Wilk, J. and Adersson, L., 2000. GIS-supported modeling of areal rainfall in a mountainous river basin with monsoon climate in southern India. Hydrol Sci. J. 45(2), 185-202.

Table 2.1. AEMN Weather stations in Georgia used in model development and evaluation

Target location	Input weather stations	Latitude (º)	Longitude (º)	Elevation (m)	Δx (km)	Δy (km)	Δs (km)	Δz (m)
Tifton	Tifton	31.483	-83.533	113	0	0	0	0
	Dawson	31.759	-84.436	102	-86.54	28.24	91.03	-11
	Alma	31.560	-82.510	65	96.92	12.08	97.67	-48
	Plains	32.047	-84.371	162	-81.16	60.33	101.12	49
	Fort Valley	32.531	-83.890	154	-37.28	115.22	121.10	41
	Attapulgus	30.761	-84.485	83	-88.66	-82.52	121.12	-30
	Pine Mountain	32.839	-84.859	277	-129.01	147.13	195.68	164
Griffin	Griffin	33.262	-84.284	283	0	0	0	0
	Williamson	33.176	-84.407	266	-11.22	-9.83	14.92	-17
	Pine Mountain	32.839	-84.859	277	-52.62	-48.15	71.33	-6
	Roopville	33.422	-85.055	266	-72.17	16.14	73.95	-17
	Eatonton	33.397	-83.488	167	73.68	17.19	75.66	-116
	Fort Valley	32.531	-83.890	154	39.11	-80.06	89.10	-129

Table 2.2. Effects of varying the number of nearest locations on the coefficient of determination (r^2), mean absolute error (MAE), and root mean square error (RMSE) on ANN models for estimating maximum temperature at Tifton. Input variables consisted of maximum temperature, Δs, and Δz.

		Production dataset			Testing dataset			Training dataset		
# of weather stations	# of hidden nodes per slab	r^2	MAE (ºC)	RMSE (ºC)	r^2	MAE (ºC)	RMSE (ºC)	r^2	MAE (ºC)	RMSE (ºC)
1	7	0.973	0.83	1.19	0.981	0.77	1.02	0.974	0.86	1.18
2*	8	0.990	0.56	0.72	0.990	0.53	0.71	0.990	0.56	0.74
3	10	0.989	0.56	0.73	0.989	0.57	0.76	0.991	0.55	0.72
4	18	0.989	0.60	0.79	0.989	0.58	0.75	0.990	0.55	0.73
5	15	0.986	0.64	0.84	0.987	0.62	0.84	0.991	0.52	0.68

* preferred model.

Table 2.3. The effects of varying number of hidden nodes on the coefficient of determination (r^2), mean absolute error (MAE), and root mean square error (RMSE) on ANN models for estimating maximum temperature at Tifton. Input variables consisted of maximum temperature, Δs, and Δz of two input weather stations.

# of hidden nodes per slab	Production dataset r^2	MAE (ºC)	RMSE (ºC)	Testing dataset r^2	MAE	RMSE (ºC)	(ºC)	Training dataset r^2	MAE	RMSE (ºC)	(ºC)
5	0.986	0.65	0.86	0.985	0.64	0.86		0.986	0.66	0.89	
7	0.986	0.64	0.85	0.986	0.63	0.84		0.986	0.66	0.88	
8*	0.990	0.56	0.72	0.990	0.53	0.71		0.990	0.55	0.74	
9	0.980	0.66	0.88	0.985	0.68	0.90		0.985	0.68	0.90	
10	0.990	0.56	0.72	0.990	0.53	0.71		0.990	0.56	0.74	
11	0.990	0.56	0.72	0.990	0.53	0.71		0.990	0.56	0.74	
12	0.986	0.65	0.88	0.986	0.63	0.84		0.986	0.66	0.88	
15	0.986	0.65	0.87	0.986	0.63	0.85		0.986	0.67	0.89	
20	0.986	0.65	0.86	0.986	0.64	0.85		0.986	0.67	0.88	

* preferred model.

Table 2.4. Search results for the preferred number of input weather stations and hidden nodes for the production data set.

				Production dataset		
Weather* Variable	target location	# of weather stations	# of hidden nodes per slab	r^2	MAE	RMSE
Tmax	Tifton	2	8	0.990	0.56	0.72
(ºC)	Griffin	1	1	0.997	0.31	0.43
Tmin	Tifton	5	5	0.986	0.69	0.89
(ºC)	Griffin	2	5	0.989	0.66	0.86
SR	Tifton	2	5	0.944	1.20	1.73
(MJ/m^2)	Griffin	1	1	0.978	0.78	1.12

* Maximum temperature (Tmax), minimum temperature (Tmin), solar radiation (SRad)

Table 2.5. Comparison of the Tifton and Griffin models for the evaluation data sets.

Weather** Variable	Model Location	Data	r^2	MAE	RMSE	slope[1]	intercept[2]
Tmax (ºC)	Tifton	Tifton	0.987	0.61	0.84	1.004	0.0184
		Griffin	0.930	4.75	5.24		
	Griffin	Tifton	0.545	9.62	10.83		
		Griffin	0.997	0.36	0.51	0.9817	0.3092*
Tmin (ºC)	Tifton	Tifton	0.987	0.74	0.93	0.968	0.7097*
		Griffin	0.964	3.40	3.75		
	Griffin	Tifton	0.859	3.67	4.32		
		Griffin	0.984	0.82	1.07	0.9908	-0.0941
SR (MJ/m^2)	Tifton	Tifton	0.944	1.24	1.74	0.9178	1.0953*
		Griffin	0.907	3.15	4.03		
	Griffin	Tifton	0.680	12.65	13.31		
		Griffin	0.958	1.51	2.09	0.8786	0.8627*

* significantly different from 0 at 95% of confident level.

** Maximum temperature (Tmax), minimum temperature (Tmin), solar radiation (SRad)

[1, 2] Linear Regression analysis: Estimated = intercept + slope* observed

Table 2.6. Comparison of ANN, averaging, inverse distance, and multi-linear regression methods for the evaluation data set.

Location	Model*	Method	r^2	MAE	RMSE
Tifton	Tmax (ºC)	ANN	0.987	0.61	0.84
		Average	0.987	0.80	1.02
		Inverse distance	0.987	0.80	1.03
		Regression	0.987	0.61	0.70
Tifton	Tmin (ºC)	ANN	0.987	0.74	0.93
		Average	0.901	0.75	1.03
		Inverse distance	0.901	0.75	1.03
		Regression	0.988	0.73	0.81
Tifton	SR (MJ/m^2)	ANN	0.944	1.24	1.74
		Average	0.946	1.81	2.28
		Inverse distance	0.946	1.79	2.26
		Regression	0.944	1.23	2.93
Griffin	Tmax (ºC)	ANN	0.997	0.36	0.51
		Average	0.997	0.89	0.94
		Inverse distance	0.997	0.89	0.94
		Regression	0.997	0.35	0.54
Griffin	Tmin (ºC)	ANN	0.984	0.82	1.07
		Average	0.981	1.09	1.85
		Inverse distance	0.984	1.11	1.99
		Regression	0.981	1.18	2.23
Griffin	SR (MJ/m^2)	ANN	0.958	1.51	2.09
		Average	0.957	2.84	3.34
		Inverse distance	0.957	2.84	3.34
		Regression	0.957	1.61	4.55

* Maximum temperature (Tmax), minimum temperature (Tmin), solar radiation (SRad)

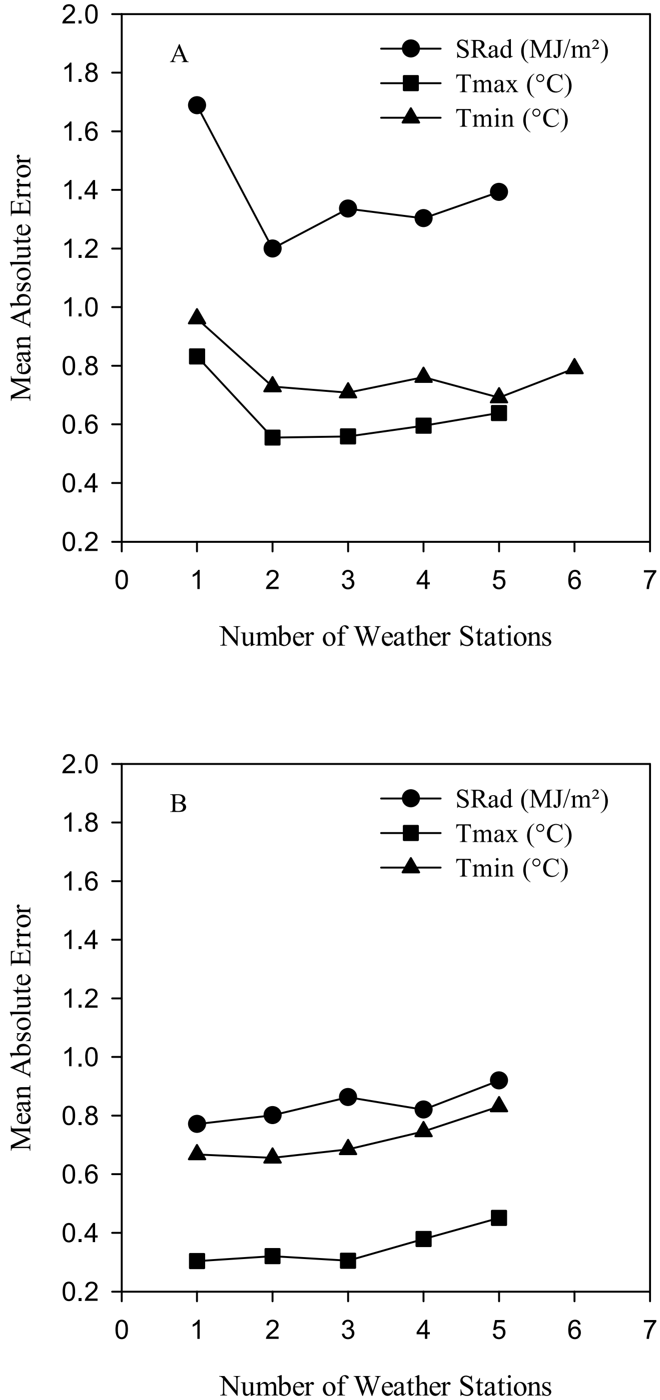
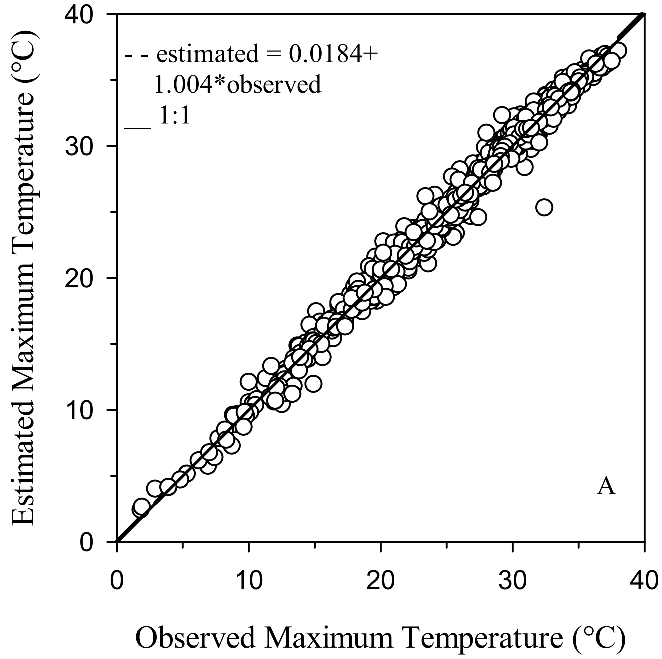


Figure 2.1 Mean absolute error as a function of number of input weather stations for estimating maximum temperature, minimum temperature, and solar radiation for Tifton GA (A) and Griffin GA (B), production dataset.



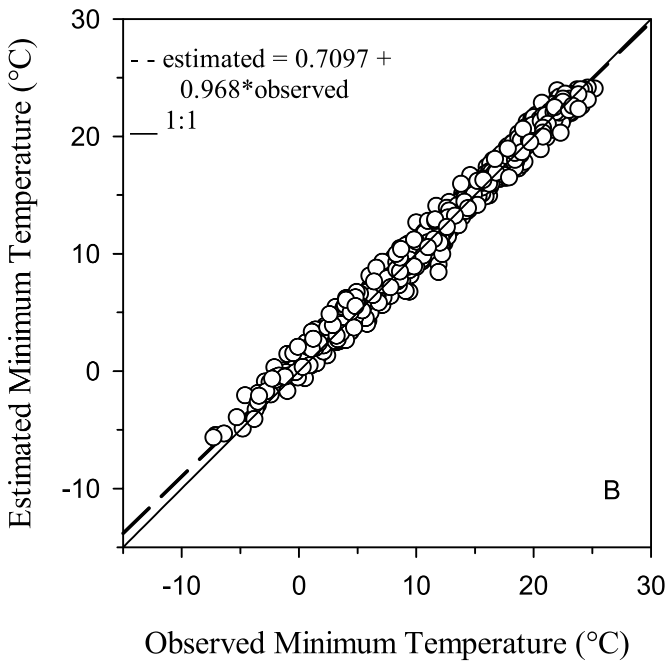


Figure 2.2 Estimated vs. observed maximum temperature (A) and minimum temperature (B), Tifton GA, evaluation dataset.

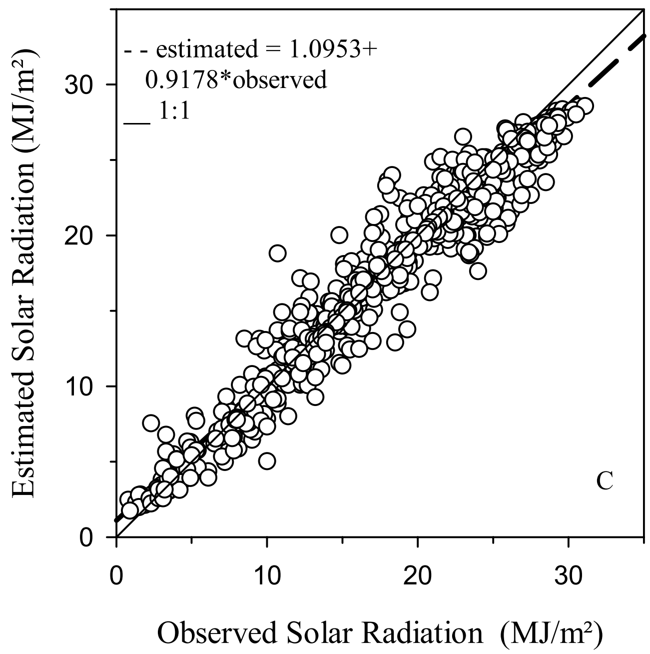


Figure 2.2 Estimated vs. observed solar radiation (C), Tifton GA, evaluation dataset.

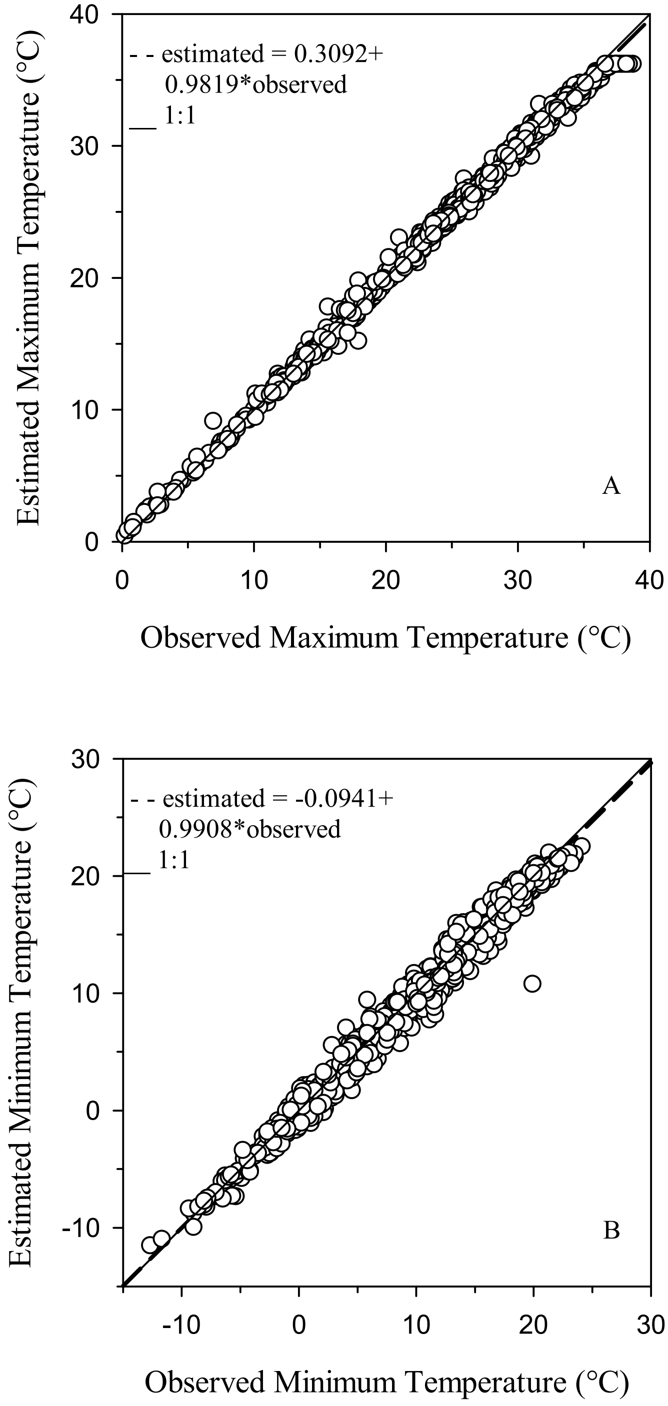


Figure 2.3. Estimated vs. observed maximum temperature (A) and minimum temperature (B), Griffin GA, evaluation dataset.

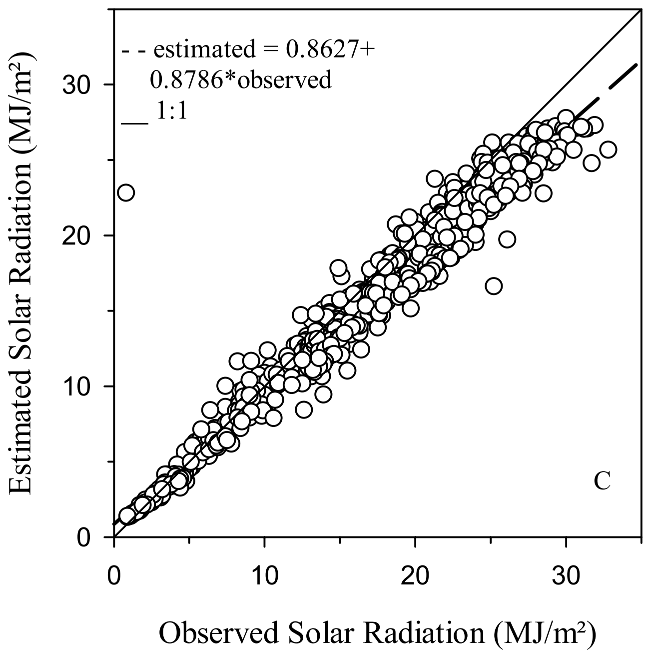


Figure 2.3. Estimated vs. observed solar radiation (C), Griffin GA, evaluation dataset.

CHAPTER 3

SPATIAL INTERPOLATION OF WEATHER DATA FOR MULTIPLE LOCATIONS USING ARTIFICIAL NEURAL NETWORKS

Li, B., R.W. McClendon, G. Hoogenboom. To be submitted to Transactions of the ASAE.

ABSTRACT

Daily weather data such as temperature, solar radiation, and rainfall are inputs to crop growth simulation models for decision support. An accurate estimation of these weather variables has thus become necessary in order to use these models for locations without complete historical weather data. Artificial neural networks (ANNs) are relative new technique for accurately estimating these weather variables. In this study, artificial neural network (ANN) models were developed to estimate daily maximum air temperature, minimum air temperature, and total solar radiation throughout Georgia. Weather data from fourteen automated weather stations for the period of 1996 to 1998 were used for model development, and 1999 to 2000 data were used for final model evaluation. For the development of the ANN model, daily weather data were used to create patterns for the fourteen target locations. Based on the results of the study, the best input variables were determined to be straight line distance (Δs) and the elevation difference (Δz) between the target location and neighboring weather stations as well as the values of the weather variable being modeled from the input weather stations. Maximum temperature was also found to be important in estimating solar radiation.

It was determined that only two closest weather stations provided the highest accuracy for estimating the three weather variables. The ANN models for estimating these weather variables were compared with inverse distance, average, and multi-linear regression methods. The results showed that ANN models provided superior accuracy over the other methods. With the overall evaluation dataset, the general ANN model for estimating maximum temperature had an r^2 of 0.985, MAE of 0.71°C, and RMSE of 0.99°C. The general ANN model for estimating minimum temperature had an r^2 of 0.977,

MAE of 0.94°C, and RMSE of 1.24°C. The general model for estimating solar radiation had an r^2 of 0.907, MAE of 1.53 MJ/m^2, and RMSE of 2.15 MJ/m^2. The ANN general models for estimating the three weather variables were compared to localized models developed in a previous study specifically for two of the fourteen locations and the localized models were found to be more accurate.

INTRODUCTION

Computer simulation models are becoming increasingly important tools in research and decision making related to agricultural production. Most crop growth simulation models use weather data inputs such as daily maximum and minimum temperatures, solar radiation, and rainfall to simulate crop yields (Amir and Sinclair 1991; Hoogenboom et al., 1992; Chapman et al., 1993). Air temperature is an important input to a variety of spatially distributed hydrological and ecological models (Cook and Wolfe, 1991; Dodson and Marks, 1997; Bolstad et al., 1998; Scheifinger and Kromp-Kolb, 2000). In addition, air temperature has been used to analyze climate change and the effects of the change (Robeson and Janis 1998; Michaels et al., 1998; Goodale et al., 1998; Price et al., 2000). Moreover, most processes in the atmosphere and biosphere, such as evaporation, sensible heat flux, soil heat flux, are driven directly or indirectly by solar radiation (Bruton et al., 2000; Scheifinger and Kromp-Kolb, 2000).

In crop growth simulation model development, it is assumed that the future climate will fit the same distribution as the historical climate used in the analyses. Most current models require long-term daily weather records with a high spatial resolution with matched temporal resolution. However, for some areas weather measurements are not available due to the high cost of instrumentation, maintenance and calibration. For some areas, only a limited period of records is available. Therefore it is highly desirable to develop accurate weather data estimation models for use in simulation, weather analysis, and decision support applications.

Traditional interpolation methods include Thiessen polygons, inverse distance interpolations, kriging, splines, and regression model. The Thiessen polygon or Voronoi polygon has been widely used in climatological rainfall and precipitation estimations (Goovaests, 2000; Wilk and Andersson, 2000; Pardo-Iguzquiza, 1998; Dirks et al, 1998; Panagoulia, 1992). Dirk et al. (1998) compared the Thiessen polygon with kriging spatial interpolation method to estimate spatially continuous rainfall. They found the Thiessen polygon method was comparable with kriging. Goovaerts (2000) pointed out that the Thiessen polygon method did not consider the elevation effects and rainfall records from surrounding stations. Therefore, large prediction errors could occur in the prediction of rainfall.

Inverse-distance is a weighting interpolation method. The number of neighbors necessary in the weighting function is important in terms of reducing computation time while maintaining a smooth surface. Dodson and Marks (1997) have suggested that with inverse-squared-distance interpolation using eight nearest neighbors is reasonable. Robeson (1993) investigated three methods of spatially interpolating temperature anomaly data. He found that the inverse-distance method gave about the same results as triangulated surface patches and thin-plate splines. In order to consider the elevation effects on climate, gradient plus inverse-distance-squared (GIDS) interpolation technique was derived (Nalder and Wein, 1998, Price et al., 2000) from the inverse-distance-square method. Price et al. (2000) used gradient plus inverse-distance-squared method to interpolate Canadian monthly mean climate data. It was suggested that this method is attractively simple and appears to give results adequate for modeling long term forest ecosystem responses to climate in relatively flat terrain.

Regression has been used successfully in weather data estimation (Ollinger et al. 1993, 1995). Bolstad et al. (1998) used a regression approach as a main method to predict daily air temperature and compared regression approach with local lapse models or the kriging method. They stated that the regression approach provided an accurate estimate of station temperature. Christine et al. (1998) used a regression technique to predict the monthly precipitation, monthly averaged maximum and minimum temperature, and monthly averaged sunshine hours and compared the regression approach with a modified inverse-distance-square interpolation. They reported that the prediction accuracy did not differ between these two methods.

Kriging has been implemented in analysis of climatologic factors, such as the analysis of precipitation (Dingman et al. 1988), evapo-transpiration (Martinez-Cob & Cuenca 1992), and temperature (Holdaway 1996, Bolstad et al. 1998). Holdaway (1996) has applied kriging to the spatial interpolation of monthly temperature. In his research, monthly empirical variograms, averaged over 90 years, were modeled with Gaussian or linear models in the ordinary kriging method. He concluded that anisotropies were found in the winter months, suggesting the presence of a large-scale regional trend. Bolstad et al. (1998) used kriging (co-kriging) to predicate daily air temperatures. They found that kriging and co-kriging may be particularly appropriate for temperature predictions in regions with little topographic relief, but not useful where temperature measurement stations were sparse or high terrain effects were existing.

Splining has been discussed by some researchers (Eckstein 1989; Hutchinson and Gessler 1994). Hutchinson (1989) used multi-dimensional thin plate splines to fit temperature surfaces by minimizing the roughness of the interpolated surface. He reported that the thin plate spline method worked as good as kriging while requiring less parameterization, however thin plate splines are computationally demanding and complicated to implement. Price et al. (2000) employed the thin-plate smoothing splines to interpolate Canadian monthly mean climate data. The GIDS was used to compare with the thin-plate smoothing splines. They found that thin plate smoothing splines produced better results for the west region of Canada where predicting precipitation is difficult.

Artificial neural networks (ANNs) offer an alternative to estimate such weather data. ANNs are computer models that mimic the structure and functioning of the human brain (Ward Systems Group Inc., Frederick, MD, 1993). ANNs can determine the relationships among the independent variables to predict or estimate dependent variables. Back propagation (BP) ANNs are known for their ability to generalize well on a wide variety of problems and are well suited for prediction applications. Unlike statistical methods, ANN models do not make dependency assumptions among input variables and solves multivariate problem with nonlinear relationship among input variables. This technique has been used in a wide range of applications, such as classification, pattern recognition, automatic control and function approximation (McAvoy et al., 1989; Leonard et al., 1992, Rao & Gupta, 1993). Han and Felker (1997) implemented an ANN to estimate daily soil water evaporation from average relative air humidity, air temperature, wind speed, and soil water content in a cactus field study. They found that the ANN achieved a good agreement between predicted and measured values. They concluded that the ANN technique appeared to be an improvement over the multi-linear regression technique for estimating soil evaporation. Elizondo et al. (1994) used an ANN to estimate daily solar radiation for locations in the southeastern US based on daily maximum and minimum air temperature, daily total precipitation, daily clear sky radiation and day length for that location. They did not include weather data from other locations as inputs. They found r^2 of 0.74 and a root mean square error of 2.92 MJ/m^2. In their study, the monthly average of daily maximum temperatures for three months in advance was predicted. Bruton et al. (2000) developed ANN models for estimating daily pan evaporation. The results were compared with those of multiple linear regression and Priestly-Taylor model and they found that the ANN model provided the highest accuracy.

Li et al. (2002) developed ANN models for estimating daily maximum air temperature, minimum air temperature, and total solar radiation for a specific site in south Georgia and one in north Georgia. Weather data collected at the Georgia Automated Environmental Monitoring Network (AEMN) from 1996 to 2000 were used. The dataset for model development consisted of data from 1996 to 1998 (1096 daily weather observations). The evaluation dataset consisted of data from 1999 to 2000 (731 daily weather observations). For each of two target locations and three weather variables, they determined the preferred number of input weather stations and the preferred input variables. The ANN models performed well when estimating the weather data at the location for which it was developed. However when one of the models was used to estimate the data from the other location, large errors occurred. Each of these ANN models was developed with only one particular location as the target location thus was localized to that particular location. Therefore, the application of these models for other locations out of the region was questionable.

The goal of this research was to develop general ANN models to estimate daily weather data which would be applicable throughout the state of Georgia. The specific objectives were: 1) develop general ANN models using datasets with fourteen target locations to estimate daily maximum and minimum temperature and solar radiation, 2) determine the number of weather stations required to provide inputs, 3) determine which inputs are required for each weather variable, 4) compare these ANN models with traditional averaging, multi-linear regression, and inverse distance weighting interpolation methods, and 5) compare the general ANN models with localized models by Li et al. (2002).

MATERIALS AND METHODS

Daily weather data collected by the Georgia Automated Environmental Monitoring Network (AEMN) were used in this study (Hoogenboom, 1996, 2000b; Hoogenboom et al., 2000). The AEMN is a network that consists of over 47 automated weather stations located across the state of Georgia. The sensor of each weather station is polled with a one-second frequency and averages or totals are logged every 15 minutes. All recorded data are saved to a centrally located computer on a daily basis. Weather variables measured in the AEMN include air temperature, rainfall, soil temperature at three different depths, relative humidity, wind speed and direction, solar radiation and open pan evaporation. Daily weather data collected from 1996 to 2000 were used for this study. The locations of the 14 weather stations used are shown in Figure 3.1. These locations roughly depict a study area from south of Atlanta, Georgia to the border with Florida in the south and reaching to the border with South Carolina in the east and the border with Alabama in the west.

The weather data from 1996 to 2000 were divided into two overall data sets: one for model development and one for model evaluation. The data from 1996 to 1998 were used to develop the ANN models. This data set included a total of 15,246 daily observations for each variable. The data from 1999 to 2000 were used to evaluate the final ANN modesl. This overall evaluation dataset included a total of 10,220 daily observations. Unlike the ANN model for localized models (Li et al., 2002), there were 14 target locations for each day for both model development and evaluation. For each of the 14 target locations, the remaining 13 weather stations were used as potential input locations. The ANN models were thus generalized for Georgia as represented by these weather stations (Figure 3.1). A general ANN model was developed for each weather variable consisting maximum temperature, minimum temperature, and solar radiation. The overall goal was that the models could estimate the weather variables at any location in Georgia and for any day of year. The overall evaluation dataset using 1999 and 2000 weather data contained patterns with the 14 target locations. This dataset was applied to the final ANN models to determine the overall accuracy of the general models. Evaluation datasets for the 14 individual locations were also presented to the final models to determine the accuracy of the general model by location.

Computer programs were developed in the Java programming language to create data files for ANN models from the raw data of the 14 weather stations. The possible inputs considered were daily maximum air temperature, minimum air temperature, daily total solar radiation, and the fractions that a day of year belongs to each season (spring, summer, fall, and winter). Other inputs that were considered included the difference in elevation (Δz), the difference in east-west direction (Δx), and the difference in north-south direction (Δy) between the target location and the neighboring stations. The straight line distance between the target and input station location (Δs) was also considered as a substitute for Δx and Δy. The values for the daily temperatures and solar radiation were observed data with day of year as index, whereas the other values were computed data. During preprocessing, the calculated data were computed first, based on the corresponding equations. The daily observed data were then combined with the calculated data, and the dataset was randomized by day of year.

The computer program also included a routine to select the closest neighboring stations for all target locations. For each of the 14 target locations, the stations from the 13 remaining weather stations were selected based on the shortest distance from the target location. Data from these weather stations were organized into a pattern in the data set based upon target and day of year, but the order of weather stations for each pattern was randomly distributed. This was done to maintain some generality by location and prevent the ANN from memorizing data for a particular location. After data preprocessing, each pattern consisted of the three daily weather variables per input weather station and target location and the calculated variables of Δx, Δy, Δs, Δz per weather station and four seasonal effect terms. In addition, there were fourteen patterns for each day of the year with fourteen different locations as targets.

The back propagation ANN approach is based on gradient descent and is designed to minimize the mean square error between the observed and estimated outputs. Ward ANN is a three-layer back propagation network consisting of an input layer, a hidden layer, and an output layer. Ward ANN has multiple slabs in the hidden layer (Ward System Group Inc., Frederick, MD, 1993). A Ward ANN also allows the user to select different activation functions such as Gaussian function, hyperbolic tangent function, and Gaussian-complement function for each slab. With different activation functions, these networks could possibly detect different features of the input vectors and are well suited for prediction. Ward neural networks were selected in this study based on our preliminary comparisons with the standard back propagation architecture.

The NeuroShell™ (Ward System Group Inc., Frederic, MD, 1993) is an ANN software package with a user friendly menu-driven infrastructure. It has several utilities for data manipulation, model development, graphical options, and a runtime option to generate source code. One of the important options of NeuroShell™ is the optimal network option which helps to prevent over-training. NeuroShell™ allows the user to set an interval for the program to check the accuracy of the current ANN with a separate testing data set. If the average error of prediction of the ANN improves compared to the previous optimal ANN, the older optimal ANN is replaced with the current model. Otherwise, the previous optimal ANN can be saved until an improved optimal ANN is reached, or training is completed.

In our study, the model development dataset consisted of 15,246 daily observations and was divided into three separate datasets i.e., training, testing, and production with a distribution of 50%, 25%, and 25% of the data, respectively. Preliminary analyses showed that the number of hidden nodes was the only ANN parameter that had a significant effect on model accuracy. Therefore, the learning rate was set at 0.1, momentum was set at 0.1, stopping criteria was set at 20,000 events past the minimum test set error, and a test interval of 200 learning events was used. The best value for the number of hidden nodes was determined based on the coefficient of determination (r^2), mean absolute error (MAE), and root mean square error (RMSE) of model results using the training, testing, and production datasets.

Computer experiments were conducted to determine which input variables were important for each ANN model. Different combinations of input variables were examined to determine the ANN model that had the lowest estimation error for the production dataset. One of the objectives of this study was to identify which inputs were needed to estimate each of the three weather variables. This was accomplished by adding an input variable for the model, retraining and analyzing the effects on r^2, MAE, and RMSE. Daily weather variables that were evaluated included observed daily maximum temperature and minimum temperature, and solar radiation. Calculated variables were determined from the topographic data and day of year.

In developing a general ANN model to estimate maximum temperature for the region, the following approach was taken to determine the importance of the different inputs. An ANN model was initially developed consisting of only maximum temperature inputs at surrounding weather stations for each of the 14 weather stations. The accuracy of this model was compared to one developed by adding the input Δs. The more accurate of these two models was then compared to one with Δx and Δy added as inputs. This process was repeated with the inclusion of minimum daily temperature, the seasonal effect terms and Δz as inputs. A similar approach was used to develop the general ANN models for estimating minimum temperature and solar radiation.

The number of closest weather stations required as input for each of the 14 different target locations was also an important factor in developing the optimal ANN models. Different files were created with an increasing number of input weather stations for all 14 target locations. Different ANN models were developed based on these input files. The number of nearest weather stations required for each of fourteen different target locations was determined based upon the accuracy of the developed models. All decisions regarding the preferred ANN inputs and number of closest weather station locations for all 14 target locations were made using the model accuracy on the production data set.

Final model evaluation was based on the daily weather data at 14 weather stations for 1999 and 2000 weather data. This overall evaluation dataset was prepared similarly to the model development dataset. For model evaluation, a one-time feed forward mode was used with the final ANN models. The accuracy of the ANN models for estimating each of the three weather variables for an overall evaluation dataset with 14 target locations was obtained. The evaluation was also carried out to determine how well the ANN general models performed for estimating the three weather variables for each of the fourteen target locations. Another evaluation was carried out to determine if the general ANN models developed with 14 different target locations were able to estimate the three weather variables at Tifton or Griffin as well as the localized models of Li et al. (2002).

The results of ANN models were compared with the results of multi-linear regression, averaging, and inverse distance methods on the overall evaluation dataset. Based on the literature review, the Thiessen polygon method assumes a uniform weather area, which is not the case for our study. Therefore, the Thiessen polygon method was not selected for comparison with ANN models. Bolstad et al. (1998) compared regression with local lapse models and the kriging method. They found that the regression approach provided a more accurate estimate of air temperature. Therefore, in our study, regression was selected for comparison. In addition, Robeson (1993) investigated inverse distance method, triangulated surface patches and thin-plate splines. He found that the inverse-distance method gave about the same results as other two methods. Therefore, in our study, inverse distance method was selected for comparison. The main purpose of this comparison was to determine if ANN models could provide a higher accuracy for estimating maximum temperature, minimum temperature, and solar radiation than the traditional statistical interpolation methods.

RESULTS AND DISCUSSION

A heuristic search was performed to determine the preferred number of input weather stations, input variables, and number of hidden nodes for each weather variable ANN. The approach taken to develop the maximum temperature ANN is discussed below as an example. Inputs consisted of maximum temperatures from six known weather stations surrounding the target location. The number of hidden nodes was then varied to maximize r^2 and minimize MAE and RMSE. The number of closest weather stations was then reduced by one and ANN models were developed again. The optimal number of hidden nodes was determined for each value of the number of weather stations. Once the best model was determined using only maximum temperature as input, additional inputs were considered. The same search was continued by varying the number of weather stations and again determining the optimal number of hidden nodes.

As expected, it was found that the weather variable being modeled was always an important input. For example, when developing a maximum temperature ANN, maximum temperature from the surrounding weather stations was always an important input. Also the inclusion of Δs and Δz improved the accuracy for the estimation of all three weather variables. For the solar radiation ANN, it was determined that maximum temperature from the neighboring input weather stations was also important in improving the accuracy of the ANN.

The number of neighboring weather stations required to maximize the accuracy did not vary by weather variable. For example, Table 3.1 shows the results for maximum temperature ANN search from one to six input weather stations. Two input weather stations resulted in the highest overall accuracy (MAE of 0.68ºC) for estimating maximum temperature using maximum temperature, Δs, and Δz as inputs. Results are shown for the training, testing, and production data sets, although only the production data set was used in selecting the preferred value. Table 3.2 shows the results of the search for the preferred number of hidden nodes for the case of two input weather stations. Twenty-nine hidden nodes produced the highest accuracy for the production set. Similar searches were also performed for minimum temperature and solar radiation. Table 3.1 also shows the results for the minimum temperature ANN search from one to six input weather stations. Two weather stations again resulted in the highest overall accuracy (MAE of 0.84ºC) for estimating minimum temperature using minimum temperature, Δs, and Δz as inputs. Table 3.1 shows the results for solar radiation ANN search from one to five input weather stations as well. Two weather stations resulted in the highest overall accuracy (MAE of 1.37 MJ/m^2) for estimating solar radiation using solar radiation, maximum temperature, Δs, and Δz as inputs. Table 3.3 summarizes the results of the search for the number of input weather stations, hidden nodes, and preferred inputs for each of the weather variable ANN models. The MAE is plotted for the three weather variables for the production dataset as a function of number of input weather stations in Figure 3.2. The minimum error for all three weather variables is clearly at two input weather stations. Table 3.4 identifies the input weather stations for each of the 14 target locations and gives the Δs and Δz values.

The overall evaluation dataset consisting of data for the 14 locations for year 1999 and 2000 was then presented to the final ANN models for the three weather variables. This dataset had not been used either in model development or in the selection of the preferred network architecture or parameters. The evaluation results for estimating maximum temperature are shown in Table 3.5. We found that the ANN maximum temperature model produced estimations for the overall evaluation dataset with an accuracy which was comparable to the results for the model development dataset. This model had an r^2 of 0.985, a MAE of 0.71ºC, and a RMSE of 0.99ºC. Figure 3.3 shows the estimated maximum temperature versus observed maximum temperature for the overall evaluation dataset. A linear regression was performed to determine how well the ANN results matched the observed maximum temperature. The ANN model for estimating maximum temperature slightly overestimated at lower values and slightly underestimated at higher values of measured maximum temperature. However, at all other temperature ranges, the ANN model estimated maximum temperature well. The intercept of the linear regression equation was significantly different from 0 at the 95% confidence level.

Table 3.5 also presents the results of the general ANN model for maximum temperature for evaluation dataset for each of the 14 target locations. The results indicted that the eight of the 14 target locations had a higher accuracy than the overall accuracy on the overall evaluation dataset, thus six target locations had lower accuracy than the overall accuracy. Attapulgus had the least accurate estimation with an r^2 of 0.962, a MAE of 0.98ºC, and a MASE of 1.39ºC. This site used weather stations at Dawson (110.79 km) and Tifton (121.12 km) as input. These two sites were both located north of the Attapulgus site and had longer distances compared to other target locations. Williamson had the most accurate estimation with an r^2 of 0.996, a MAE of 0.44ºC, and a RMSE of 0.59ºC. The Williamson site used the weather stations in Griffin (14.92 km) and Pine Mountain (54.41 km) as input. These two sites were close to the target location and were located on either side of the target location along the latitude direction.

The evaluation results for estimating minimum temperature are shown in Table 3.6. The ANN minimum temperature model produced estimations for the overall evaluation dataset with an accuracy that was comparable to the results for the model development dataset. The general ANN model for estimating minimum temperature had an r^2 of 0.977, a MAE of 0.94ºC, and a RMSE of 1.24ºC. Figure 3.4 shows the estimated minimum temperature versus observed minimum temperature for the overall evaluation data set. A linear regression was also performed to determine how well the ANN results matched the observed minimum temperature. The figure indicates that the ANN model had a tendency to slightly overestimate minimum temperature at lower values and slightly underestimate minimum temperature at higher values. However, for all other temperature ranges the ANN model estimated minimum temperature well. The intercept of the linear regression equation was significantly different from 0 at the 95% confidence level.

Table 3.6 also presents results of the general ANN model of minimum temperature for evaluation datasets for each of the 14 target locations. The results indicated that half of the 14 target locations had a higher accuracy than the overall accuracy. Eatonton had the least accurate estimation with an r^2 of 0.956, a MAE of 1.40ºC, and a RMSE of 1.80ºC. This site used weather stations at Griffin (75.66 km) and Williamson (89.10 km) as input and they were both located southwest of the target location. Williamson again had the highest accuracy with an r^2 of 0.987, a MAE of 0.70ºC, and a RMSE of 0.95ºC.

The evaluation results for estimating solar radiation are shown in Table 3.7. We again found that the ANN solar radiation model produced estimations for the overall evaluation dataset with an accuracy that was comparable to the results for the model development dataset. The model for estimating solar radiation had an r^2 of 0.907, a MAE of 1.53 MJ/m^2, and a RMSE of 2.15 MJ/m^2. Figure 3.5 shows the estimated solar radiation versus observed solar radiation for the overall evaluation data set. A linear regression was also performed to determine how well the ANN results matched the observed solar radiation. It indicated that the ANN model had the tendency to overestimate solar radiation at low values and to underestimate solar radiation at high values. The intermediate values were well predicted. The plot shows more variability than the maximum and minimum temperature plots. The intercept of the regression equation was 1.3979 that was significantly different from 0 at the 95% confidence level.

Table 3.7 also presents the results of the general ANN model of solar radiation for evaluation datasets for each of the 14 target locations. The results indicated that 10 of the 14 target locations had a higher accuracy than the overall accuracy. Attapulgus again had the least accuracy with an r^2 of 0.835, a MAE of 3.37 MJ/m^2, and a RMSE of 4.03 MJ/m^2. Plains had the highest accuracy with an r^2 of 0.965, a MAE of 0.97 MJ/m^2, and a MASE of 1.33 MJ/m^2. This site used weather stations at Dawson (32.53 km) and Fort Valley (70.27 km) as input. These two sites were close to the target location and were located on either side of the target location along the latitude direction.

Four of the 14 target locations, i.e. Dawson, Midville, Tifton, and Williamson, had a higher accuracy than the overall accuracy for estimating all three weather variables. These four sites are located between two input weather stations along the latitudinal or longitudinal directions. Three of the 14 target locations, i.e. Attapulgus, Savannah, and Roopville, had a lower accuracy than the overall accuracy for estimating all three weather variables. These three sites did not have neighboring weather stations on both sides of the target. Moreover, Attapulgus was located far away from the closest input weather stations, and Savannah was the only station that was located close to the ocean.

A comparison was also performed between the general model prediction and a localized model for a particular location from a previous study (Li et al., 2002). The general ANN models were used with the evaluation dataset for Tifton (south GA) and Griffin (north GA). For example, the maximum temperature model developed specifically for Tifton provided an r^2 of 0.987, a MAE of 0.61ºC, and a RMSE of 0.84ºC (Li et al. 2002). This was approximately equal to the results of the general model. The general ANN model for maximum temperature using Tifton weather data produced an r^2

of 0.987, a MAE of 0.61ºC, and a RMSE of 0.84ºC. The maximum temperature model developed specifically for Griffin provided an r^2 of 0.997, a MAE of 0.36ºC, and a RMSE of 0.51ºC (Li et al. 2002). This was higher accuracy than the general model using the Griffin evaluation dataset. With respect to minimum temperature, the minimum temperature model developed specifically for Tifton provided an r^2 of 0.987, a MAE of 0.74ºC, and a RMSE of 0.93ºC (Li et al. 2002). This was slightly higher accuracy than the general model using the Tifton evaluation dataset. The minimum temperature model developed specifically for Griffin provided an r^2 of 0.984, a MAE of 0.82ºC, and a RMSE of 1.07ºC (Li et al. 2002). This was slightly higher accuracy than the general model using the Griffin evaluation dataset. Similarly, the solar radiation model developed specifically for Tifton provided an r^2 of 0.944, a MAE of 1.24 MJ/m^2, and a RMSE of 1.74 MJ/m^2 (Li et al. 2002). This was a higher accuracy than the general model using Tifton evaluation dataset. When using Griffin weather data, the solar radiation model developed specifically for Griffin provided an r^2 of 0.958, a MAE of 1.51 MJ/m^2, and a RMSE of 2.09 MJ/m^2 (Li et al. 2002). This was slightly higher accuracy than the general model using Griffin evaluation dataset.

Li et al. (2002) showed that an ANN model for estimating daily weather data for a particular site in South Georgia was not accurate for estimating weather data for a site in north Georgia, and vice versa. Our results indicate that the general model provided an acceptable accuracy for all locations in both regions. However, the general model accuracy for a particular location was less than the accuracy of a model developed specifically for that location.

Elizondo et al. (1994) developed an ANN solar radiation model for the southeastern USA. They used three locations for model development and additional location (Tifton) for testing. They did not use any neighboring weather stations for model development and they also did not consider the variation of distance between the weather station and elevation effects from any surrounding locations. Using 1999 data they obtained an r^2 of 0.64, a MAE of 3.48 MJ/m^2, and a RMSE of 3.40 MJ/m^2. Our results for the general ANN model for solar radiation for the Tifton evaluation dataset provided a higher accuracy (Table 3.8).

Table 3.9 provides a comparison among ANN, average, inverse distance, and multi-linear regression for estimating the three weather variables. We found that the ANN was consistently the most accurate approach compared to averaging, inverse distance, and multi-linear regression methods for all three weather variables. Multi-linear regression models were more accurate than the average and inverse distance methods for estimating maximum temperature and solar radiation. However, multi-linear regression model had a same accuracy as the inverse distance method for estimating minimum temperature. The inverse distance method was more accurate than the average method for estimating maximum temperature and minimum temperature, whereas the average method was more accurate for estimating solar radiation.

SUMMARY AND CONCLUSION

General ANN models were developed to estimate daily maximum air temperature, minimum air temperature, and total solar radiation using datasets with 14 target locations across Georgia. The number of input weather stations needed to provide accurate estimations was examined. For all three weather variables, two weather station locations were determined as the preferred number of input locations. The ANN model for estimating maximum temperature was slightly more accurate than the model for estimating minimum temperature. Using r^2 value for comparison, the ANN model for estimating solar radiation was the least accurate. Although the ANN models for estimating all three weather variables overestimated at lower values and underestimated at higher values, the magnitude of overestimation or underestimation was small especially for maximum temperature and minimum temperature. For each weather variable, the ANN model estimated well at all other ranges.

The preferred inputs for each weather variable were also determined using a heuristic search. In each of the three weather variables, Δs and Δz were found to be important along with the corresponding weather variable being estimated. For the solar radiation model, the addition of maximum temperature also improved the accuracy of the ANN model.

The general models provided estimates that were slightly less accurate than the localized models for the three weather variables at two locations from a previous study (Li et al., 2002). The general models, however, provide a means for estimating weather variables at any location in Georgia with reasonable accuracy. The localized models, while not being accurate for other locations, provide higher accuracy at locations where historical weather data are available. From the results of the study it was determined that the ANNs were consistently the most accurate methods among other average, inverse distance, and multi-linear regression methods for estimating maximum temperature, minimum temperature, and solar radiation.

REFERENCES

Amir, J. and Sinclair, T.R., 1991. A model of the temperature and solar-radiation effects on spring wheat growth and yield. Field Crops Res. 28, 47-58.

Bolstad, P. V., Swift, L., Collins, F., Regniere, J., 1998. Measured and predicted air temperatures at basin to regional scales in the southern Appalachian mountains. Agric. & For. Met. 91, 161-176

Bruton, J.M., McClendon, R.W., Hoogenboom, G., 2000. Estimating daily Pan Evaporation with Artificial neural networks. Trans. ASAE 43(2), 491-6.

Chapman, S.C., Hammer, G.K. and Meinke, H., 1993. A crop simulation model for sunflower. I. Model development. Agron. J. 85, 725-735.

Christine L. Goodale, Jhon D. Aber, Scott V. Ollinger, 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. Clim. Res. 10, 35-49.

Dingman, S.L.. Seely-Reynolds, D.M., Reynolds, R.C. III, 1988. Application of kriging to estimating mean annual precipitation in a region of orographic influences. Water Resour Bull. 24:329-339

Dirks, K.N., Hay, J.E., Stow, C.D., Harris, D., 1998. High-resolution studies of rainfall on Norfolk Island. J. Hydrol. 208(3-4), 187-193.

Dodson, R. and Marks, D., 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. Clim. Res. 8, 1-20.

Eckstein, B.A. 1989. Evaluation of spline and weighted average interpolation algorithms. Computers and Geoscience 15, 79-94.

Elizondo, D., Hoogenboom, G., McClendon, R.W., 1994. Development of a neural network model to predict daily solar radiation. Agric. & For. Met. 71, 115-132.

Goodale, C.L., Aber, J.D., Olllinger, A.V., 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. Clim. Res. 10, 35-49.

Han, H. and Felker, P., 1997. Estimation of daily soil water evaporation using an artificial neural network. Journal of Arid Environments 37, 251-260.

Holdaway, M. 1996. Spatial Modeling and interpolation of monthly temperature using kriging. Clim. Res. 6, 215-25.

Hoogenboom, G. 1996. The Georgia Automated Environmental Monitoring Network. In: Preprints of the 22nd Conf. On Agricultural and Forest Meteorology, 343-346. Boston, Mass., American Meteorological Society.

Hoogenboom, G. 2000a. Contribution of agrometeorology to the simulation of crop production and its applications. Agric. & For. Met. 103(1-2), 137-157.

Hoogenboom, G. 2000b. The Georgia Automated Environmental Monitoring Network. p. 24-25. *In*: Preprints 24rd Conference on Agricultural and Forest Meteorology. American Meteorological Society, Boston, Massachusetts.

Hoogenboom, G., Georgiev, G.A., and Gresham, D.D., 2000. Development of weather based products for agricultural and environmental applications. p. 66-67. *In*: Preprints 24rd Conference on Agricultural and Forest Meteorology. American Meteorological Society, Boston, Massachusetts.

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.

Hutchinson, M.F. and P.E. Gessler 1994. Splines—more than just a smooth interpolator. Geoderma 62, 45-67.

Leonard, J.A., Kramer, M.A. & Ungar, L.H., 1992. Using radial basis functions to approximate function and its error bonds, IEEE Trans. on Neural Networks 3, 624-627.

Li, B., McClendon, R. W., and Hoogenboom, G., 2002. Spatial interpolation of weather variables for single locations using artificial neural networks, to be submitted to Computer and Electronics in Agriculture.

Martinez-cob, A. and Curnca, R.H., 1992. Influence of elevation on regional evapotranspiration using multivariate geostatistics for various climatic regimes in Oregon. J. Hydrol. 136, 353-380

McAvoy, T.J., Wang, N.S., Naidu, S., Bhar, N., Gunter, J. and Simmons, M., 1989. Interpreting biosensor data via backpropagation. IEEE International Neural Network Conference in 1989, 8, 227-233.

Michaels, P.J., Balling Jr. R. C., Vose, R.S., Knappenberger, P. C., 1998. Analysis of trends I the variability of daily and monthly historical temperature measurements. Clim. Res. 10, 27-23.

Nalder, I.A., Wein, R.W., 1998. Spatial interpolation of climatic normal: test of a new methods in the Canadian boreal forest. Agric. & For. Met. 9, 211-225.

Ollinger, S.V., Aber, J.D., Lovett, G. M., Millham, S.E., Lathrop, E.G., Ellis, J.M., 1993. A Spatial model of atmospheric deposition for the northeastern U.S. Ecol. Appl. 3, 459-472.

Pardo-Iguzquiza, E., 1998. Comparison of geostatistical methods for estimating the areal climatological rainfall mean using data on precipitation and topography. Int. J. Climatol. 18, 1031-1047.

Panagoulia, D., 1992. Hydrological modeling of a medium-size mountainous catchment from incomplete meteorological data. J. Hydrol 137(1-4), 279-310.

Parmar, R.S., McClendon, R.W., Hoogenboom, G., Blankenship, P.D., Cole, R.J., Forner, J.W., 1997. Estimation of aflatoxin contamination in preharvest peanuts using neural networks. Trans. ASAE 40(3), 809-813.

Price, D.T., McKenney, D.W., Nalder, I. A., Hutchinson, M. F., Kesteven, J. L., 2000. A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agric. & For. Met. 101, 81-94.

Rao, D.H. and Gupta, M.M., 1993. Dynamic neural controller with somatic adaptation. IEEE International Conference on Neural Networks in 1993, 8, 558-563.

Rich, E. and Knight, K., 1991. Artificial intelligence. New York: McGraw-Hill.

Robeson, S.M., 1993. Spatial interpolation, network bias, and terrestrial air temperature variability. Publ Climatol 46(1), 1-51

Robeson, S.M. and Janis, M. J., 1998. Comparison of temporal and unresolved spatial variability in multiyear time-averages of air temperature. Clim. Res. 10, 15-26.

Schaap, M.G. and Bouten, W., 1996. Modeling water retention curves of sandy soils using neural networks. Water Resour. Res. 32(10), 3033-3040.

Scheifinger, H. and Kromp-Kolb, H., 2000. Modeling global radiation in complex terrain: comparing two statistical approaches. Agric. For. Met. 100: 127-136.

Ward System Group, Frederic, MD, 1993. Manual of NeuroShell 2. Frederic, MD: Ward Systems Group.

Wilk, J. and Adersson, L., 2000. GIS-supported modeling of areal rainfall in a mountainous river basin with monsoon climate in southern India. Hydrol Sci. J. 45(2), 185-202.

Table 3.1. Effects of varying the number of the input weather stations on the coefficient of determination (r^2), mean absolute error (MAE), and root mean square error (RMSE) on ANN models of estimating maximum temperature, minimum temperature, and solar radiation.

				Production dataset			Testing dataset			Training dataset		
Weather** Variable	input variables	# of weather stations	# of hidden nodes per slab	r^2	MAE	RMSE	r^2	MAE	RMSE	r^2	MAE	RMSE
Tmax	Tmax,	1	19	0.976	0.91	1.25	0.975	0.92	1.24	0.976	0.92	1.24
(ºC)	Δs,	2*	29	0.986	0.68	0.95	0.986	0.67	0.95	0.986	0.67	0.95
	and Δz	3	29	0.985	0.72	0.99	0.984	0.71	1.00	0.984	0.69	0.99
		4	10	0.983	0.76	1.04	0.982	0.76	1.05	0.982	0.75	1.05
		5	30	0.983	0.76	1.05	0.982	0.77	1.06	0.982	0.75	1.06
		6	35	0.982	0.79	1.08	0.981	0.83	1.14	0.981	0.76	1.15
Tmin	Tmin,	1	15	0.963	1.16	1.55	0.962	1.19	1.59	0.963	1.17	1.55
(ºC)	Δs,	2*	13	0.981	0.84	1.13	0.979	0.88	1.19	0.980	0.85	1.15
	and Δz	3	25	0.980	0.85	1.14	0.979	0.88	1.18	0.981	0.84	1.12
		4	5	0.978	0.92	1.21	0.976	0.97	1.26	0.978	0.91	1.20
		5	15	0.978	0.91	1.21	0.977	0.94	1.25	0.979	0.89	1.18
		6	10	0.976	0.96	1.27	0.974	1.01	1.33	0.977	0.94	1.24
SRad	SRad	1	10	0.890	1.62	2.40	0.903	1.59	2.26	0.905	1.58	2.23
(MJ/m^2)	Tmax	2*	24	0.912	1.37	2.15	0.928	1.31	1.95	0.928	1.30	1.94
	Δs,	3	30	0.907	1.43	2.21	0.924	1.36	2.01	0.920	1.40	2.03
	and Δz	4	25	0.906	1.47	2.22	0.921	1.42	2.04	0.917	1.47	2.09
		5	25	0.904	1.51	2.25	0.919	1.45	2.07	0.915	1.49	2.11

* indicates the preferred model.
** Maximum temperature (Tmax), minimum temperature (Tmin), solar radiation (SRad)

Table 3.2. The effects of varying number of hidden nodes on the coefficient of determination (r^2), mean absolute error (MAE), and root mean square error (RMSE) on ANN models of estimating maximum temperature. Input variables consisted of maximum temperature, Δs, and Δz for two input weather stations.

	Production dataset			Testing dataset			Training dataset		
# of hidden nodes per slab	r^2	MAE (°C)	RMSE (°C)	r^2	MAE (°C)	RMSE (°C)	r^2 (°C)	MAE (°C)	RMSE
5	0.984	0.74	1.02	0.983	0.74	1.02	0.983	0.73	1.03
10	0.984	0.75	1.03	0.983	0.75	1.03	0.983	0.74	1.04
15	0.984	0.73	1.01	0.984	0.72	1.00	0.984	0.72	1.01
20	0.984	0.75	1.03	0.983	0.75	1.03	0.983	0.74	1.04
23	0.985	0.72	0.99	0.984	0.71	0.99	0.984	0.71	1.00
25	0.986	0.69	0.96	0.985	0.67	0.96	0.985	0.68	0.97
28	0.986	0.70	0.96	0.985	0.69	0.97	0.985	0.68	0.97
29*	0.986	0.68	0.95	0.976	0.67	0.95	0.986	0.67	0.95
30	0.983	0.76	1.04	0.983	0.76	1.04	0.983	0.75	1.05
34	0.983	0.77	1.05	0.982	0.77	1.05	0.982	0.77	1.06

* indicates the preferred model.

Table 3.3. Search results for preferred number of input weather stations and hidden nodes, production data set.

Weather* Variable	# of weather stations	# of hidden nodes per slab	input variables	Production dataset r^2	MAE	RMSE
Tmax (ºC)	2	29	*Tmax, Δs, Δz*	0.986	0.68	0.95
Tmin (ºC)	2	13	*Tmin, Δs, Δz*	0.981	0.84	1.13
S.Rad (MJ/m^2)	2	24	*S.Rad, Tmax, Δs, Δz*	0.912	1.37	2.15

* Maximum temperature (Tmax), minimum temperature (Tmin), solar radiation (SRad)

Table 3.4. The input weather stations for each of the fourteen target locations in general ANN models

Target location	input weather station 1			input weather station 2		
Name	Name	Δs (km)	Δz (m)	Name	Δs (km)	Δz (m)
Alma	Tifton	97.67	48	Statesboro	122.74	-1
Attapulgus	Dawson	110.79	19	Tifton	121.12	30
Dawson	Plain	32.53	60	Tifton	91.03	11
Eatonton	Griffin	75.66	116	Williamson	89.10	99
Fort Valley	Plains	70.27	8	Williamson	86.41	112
Griffin	Williamson	14.92	-17	Pine Mountain	71.33	-6
Midville	Statesboro	62.20	-17	Eatonton	132.20	86
Pine Mountain	Williamson	56.41	-11	Roopville	67.20	-11
Plains	Dawson	32.53	-60	Fort Valley	70.27	-8
Roopville	Williamson	66.24	0	Pine Mountain	67.20	11
Savannah	Statesboro	70.10	57	Alma	127.50	58
Statesboro	Midville	62.20	17	Savannah	70.10	-57
Tifton	Dawson	91.03	-11	Alma	97.67	-48
Williamson	Griffin	14.92	17	Pine Mountain	56.41	11

Table 3.5. Evaluation of general ANN models for estimating maximum temperature using overall evaluation dataset and fourteen evaluation datasets for each of the fourteen target locations.

Data	r^2	MAE (ºC)	RMSE (ºC)	*slope*[1]	*intercept*[2]
Overall	0.985	0.71	0.99	0.9812	0.4785*
Alma	0.984	0.69	0.93		
Attapulgus	0.962	0.98	1.39		
Dawson	0.993	0.70	0.87		
Eatonton	0.983	0.91	1.21		
Fort Valley	0.988	0.71	0.97		
Griffin	0.994	0.49	0.69		
Midville	0.989	0.67	0.93		
Pine Mountain	0.987	0.69	1.05		
Plains	0.994	0.75	0.89		
Roopville	0.986	0.74	0.99		
Savannah	0.970	0.96	1.28		
Statesboro	0.989	0.63	0.85		
Tifton	0.987	0.61	0.84		
Williamson	0.996	0.44	0.59		

* significantly different from 0 at 95% of confident level.

[1, 2] estimated variable regression analysis coefficient. *Estimated* = *intercept* + *slope***observed*

Table 3.6. Evaluation of general ANN models for estimating minimum temperature using overall evaluation dataset and fourteen evaluation datasets for each of the fourteen target locations.

Data	r^2	MAE (ºC)	RMSE (ºC)	*slope*[1]	*intercept*[2]
Overall	0.977	0.94	1.24	0.9778	0.1875*
Alma	0.977	1.09	1.33		
Attapulgus	0.972	1.05	1.35		
Dawson	0.989	0.73	0.96		
Eatonton	0.956	1.40	1.80		
Fort Valley	0.979	0.92	1.17		
Griffin	0.982	0.88	1.13		
Midville	0.988	0.81	1.02		
Pine Mountain	0.973	1.02	1.46		
Plains	0.990	0.79	1.00		
Roopville	0.972	1.07	1.46		
Savannah	0.977	1.00	1.28		
Statesboro	0.984	1.04	1.23		
Tifton	0.985	0.74	0.94		
Williamson	0.987	0.70	0.95		

* significantly different from 0 at 95% of confident level.

[1, 2] estimated variable regression analysis coefficient. *Estimated* = *intercept* + *slope***observed*

Table 3.7. Evaluation of general ANN models for estimating solar radiation using overall evaluation dataset and fourteen evaluation datasets for each of the fourteen target locations.

Data	r^2	MAE (MJ/m^2)	RMSE (MJ/m^2)	*slope*[1]	*intercept*[2]
Overall	0.907	1.53	2.15	0.9163	1.3979*
Alma	0.931	1.16	1.66		
Attapulgus	0.835	3.37	4.03		
Dawson	0.949	1.16	1.60		
Eatonton	0.931	1.37	1.99		
Fort Valley	0.926	1.37	1.97		
Griffin	0.884	2.11	2.63		
Midville	0.950	1.18	1.64		
Pine Mountain	0.936	1.35	1.80		
Plains	0.965	0.97	1.33		
Roopville	0.878	2.21	2.66		
Savannah	0.898	1.67	2.25		
Statesboro	0.949	1.01	1.56		
Tifton	0.940	1.32	1.81		
Williamson	0.969	1.19	1.63		

* significantly different from 0 at 95% of confident level.

[1, 2] estimated variable regression analysis coefficient. *Estimated* = *intercept* + *slope***observed*

Table 3.8. Comparison of ANN, average, inverse distance, and multi-linear regression methods, overall evaluation data set.

Model	Method	r^2	MAE	RMSE
Maximum Temperature (ºC)	ANN	0.985	0.71	0.99
	Average	0.978	0.93	1.46
	Inverse distance	0.979	0.93	1.42
	Regression	0.981	0.83	1.10
Minimum Temperature (ºC)	ANN	0.977	0.94	1.24
	Average	0.962	1.26	1.61
	Inverse distance	0.965	1.20	1.54
	Regression	0.965	1.20	1.54
Solar Radiation (MJ/m^2)	ANN	0.907	1.53	2.15
	Average	0.880	1.84	2.45
	Inverse distance	0.877	1.88	2.49
	Regression	0.881	1.84	2.44

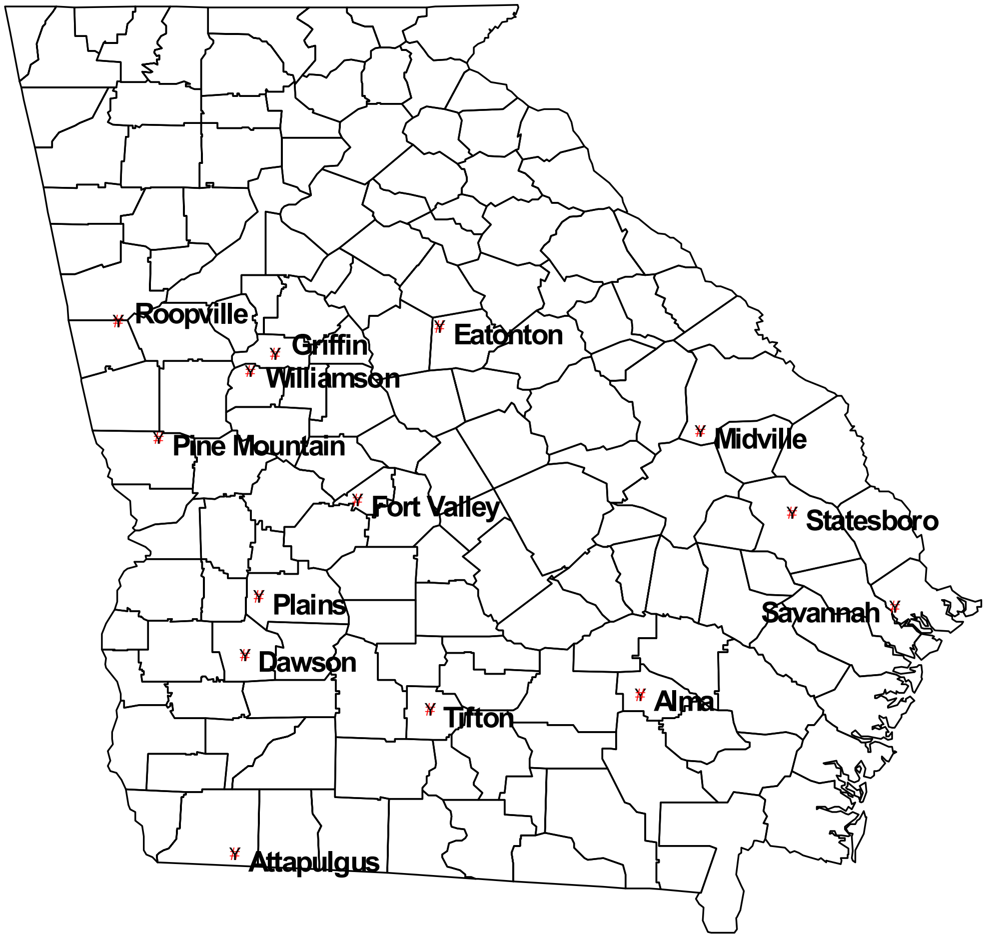


Figure 3.1. Fourteen AEMN weather stations in Georgia used in model development and evaluation.

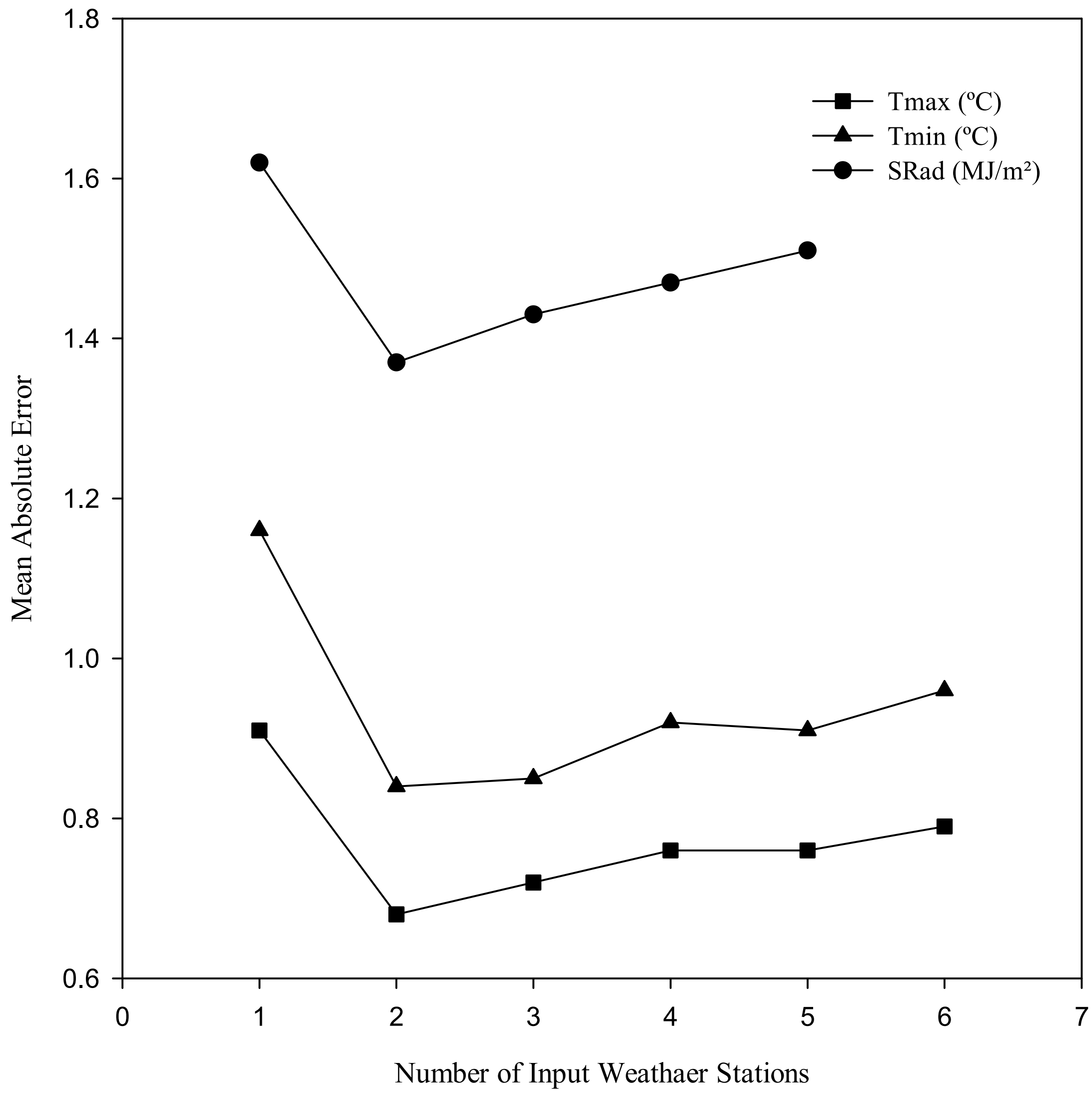


Figure 3.2. Mean absolute error as a function of number of input weather stations for estimating maximum temperature, minimum temperature, and solar radiation, production dataset.

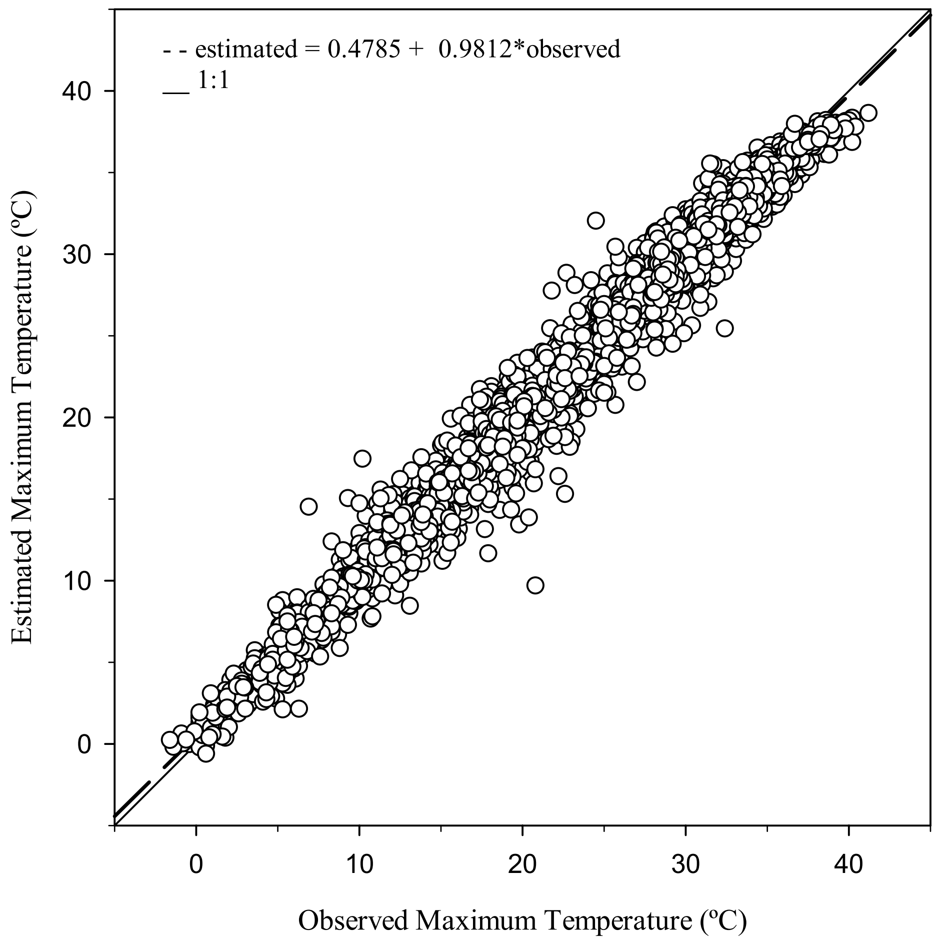


Figure 3.3. General ANN model of maximum temperature vs. observed maximum temperature, overall evaluation dataset

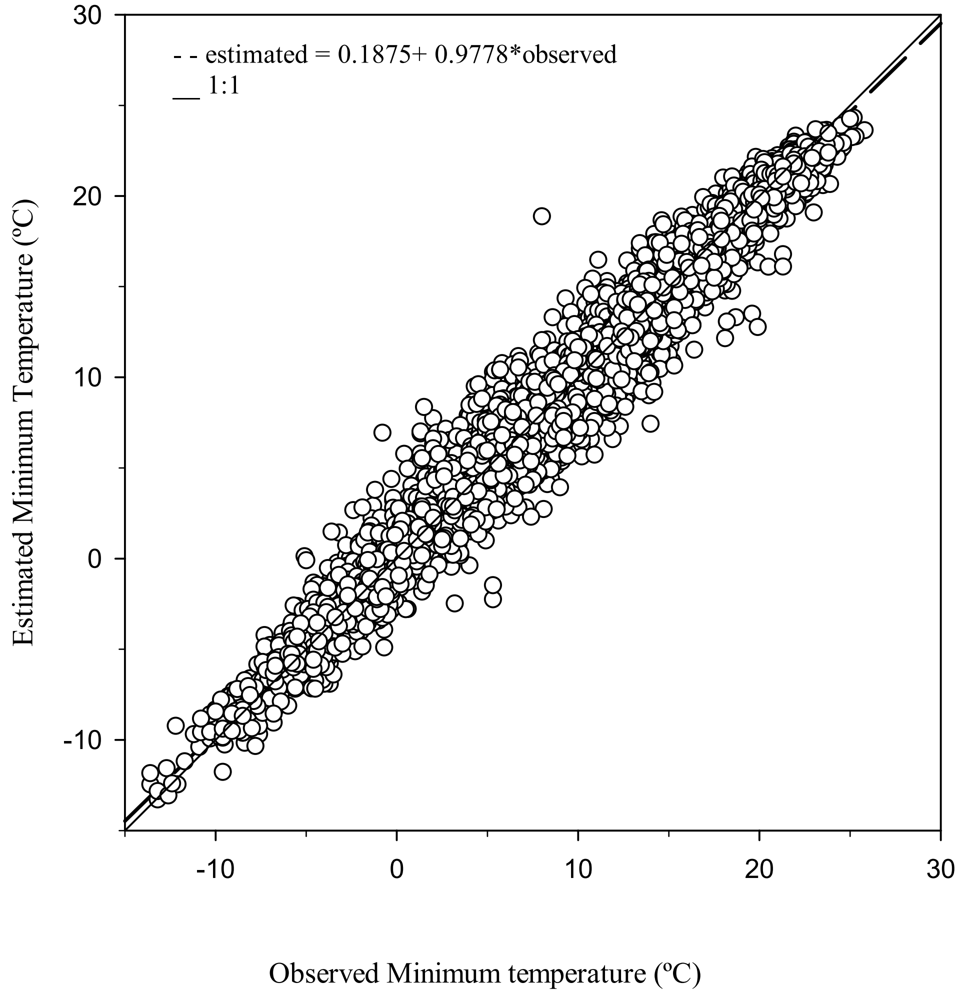


Figure 3.4. General ANN model of minimum temperature vs. observed minimum temperature, overall evaluation dataset.

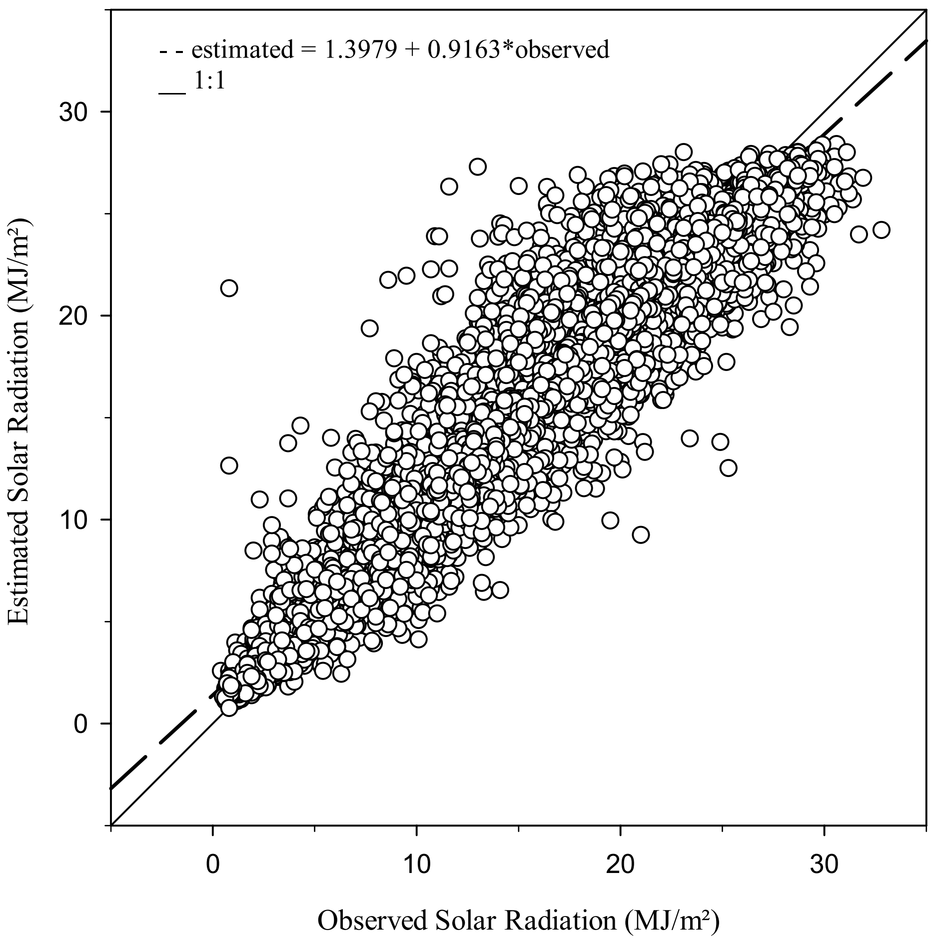


Figure 3.5. General ANN model of solar radiation vs. observed solar radiation, overall evaluation dataset

CHAPTER 4

SUMMARY AND CONCLUSIONS

In this study, two approaches were taken to develop ANN models, localized model and general model, were developed to estimate daily maximum and minimum air temperature and total solar radiation in Georgia. In the first study, localized ANN models were developed to estimate the three weather variables for the specific locations of Tifton and Griffin Georgia. In the second study, general ANN models were developed to estimate the three weather variables at locations throughout Georgia. Observed weather data from 1996-1998 was used in all model development and weather data from 1999-2000 was used in model evaluation. The measures of accuracy used were r^2, MAE, and RMSE.

In the development of localized and general ANN models, the number of input weather stations needed to provide accurate estimations was examined. For all three weather variables, two weather station locations were determined as the input locations for general ANN models, whereas for the localized ANN models, the number of input weather stations required varied with the particular weather variable and the target location.

The preferred inputs for each weather variable model were also determined using a heuristic search. In each of the three weather variable models, Δs, Δz, and the corresponding weather variable being estimated were important in improving the accuracy. For the solar radiation models, the additional input of maximum temperature also improved the accuracy.

The three general ANN models were developed using data from fourteen weather stations and performed well in estimating daily weather data throughout Georgia. The general models were compared to the three localized models for Tifton and the three models for Griffin. The accuracy of the general ANN models for all three weather variables was less than the localized model for that specific location. The localized models developed specifically for Tifton and Griffin were found to be inaccurate when used at a location other than the one for which they were developed.

In the comparison with alternative approaches of averaging, inverse distance, and multi-linear regression, the general ANN models clearly provided the highest accuracy when estimating weather variables for the overall evaluation dataset (data from fourteen weather stations). When the alternative approaches were compared to the localized ANN models, the localized ANN models were equal to or slightly better than multi-linear regression, depending on the variables and location. The localized ANN models were clearly more accurate than averaging or inverse distance methods.

For future research, effort could be directed to estimate other weather variables such as rainfall. A user interface could also be developed to allow the user to input a location in terms of latitude and longitude and the date and the system would respond with the estimated weather data. The system would automatically determine the two closest weather stations and present the weather data and calculated inputs to the appropriate general ANN model.

REFERENCES

Amir, J. and Sinclair, T.R., 1991. A model of the temperature and solar-radiation effects on spring wheat growth and yield. Field Crops Res. 28, 47-58.

Bolstad, P. V., Swift, L., Collins, F., Regniere, J., 1998. Measured and predicted air temperatures at basin to regional scales in the southern Appalachian mountains. Agric. & For. Met. 91, 161-176

Bruton, J.M., McClendon, R.W., Hoogenboom, G., 2000. Estimating daily Pan Evaporation with Artificial neural networks. Trans. ASAE 43(2), 491-6.

Chapman, S.C., Hammer, G.K. and Meinke, H., 1993. A crop simulation model for sunflower. I. Model development. Agron. J. 85, 725-735.

Christine L. Goodale, Jhon D. Aber, Scott V. Ollinger, 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. Clim. Res. 10, 35-49.

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

Dingman, S.L.. Seely-Reynolds, D.M., Reynolds, R.C. III, 1988. Application of kriging to estimating mean annual precipitation in a region of orographic influences. Water Resour Bull. 24:329-339

Dirks, K.N., Hay, J.E., Stow, C.D., Harris, D., 1998. High-resolution studies of rainfall on Norfolk Island. J. Hydrol. 208(3-4), 187-193.

Dodson, R. and Marks, D., 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. Clim. Res. 8, 1-20.

Eckstein, B.A. 1989. Evaluation of spline and weighted average interpolation algorithms. Computers and Geoscience 15, 79-94.

Elizondo, D., Hoogenboom, G., McClendon, R.W., 1994. Development of a neural network model to predict daily solar radiation. Agric. & For. Met. 71, 115-132.

Goodale, C.L., Aber, J.D., Olllinger, A.V., 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. Clim. Res. 10, 35-49.

Han, H. and Felker, P., 1997. Estimation of daily soil water evaporation using an artificial neural network. Journal of Arid Environments 37, 251-260.

Holdaway, M. 1996. Spatial Modeling and interpolation of monthly temperature using kriging. Clim. Res. 6, 215-25.

Hoogenboom, G. 1996. The Georgia Automated Environmental Monitoring Network. In: Preprints of the 22nd Conf. On Agricultural and Forest Meteorology, 343-346. Boston, Mass., American Meteorological Society.

Hoogenboom, G. 2000a. Contribution of agrometeorology to the simulation of crop production and its applications. Agric. & For. Met. 103(1-2), 137-157.

Hoogenboom, G. 2000b. The Georgia Automated Environmental Monitoring Network. p. 24-25. *In*: Preprints 24rd Conference on Agricultural and Forest Meteorology. American Meteorological Society, Boston, Massachusetts.

Hoogenboom, G., Georgiev, G.A., and Gresham, D.D., 2000. Development of weather based products for agricultural and environmental applications. p. 66-67. *In*: Preprints 24rd Conference on Agricultural and Forest Meteorology. American Meteorological Society, Boston, Massachusetts.

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.

Hutchinson, M.F. and P.E. Gessler 1994. Splines—more than just a smooth interpolator. Geoderma 62, 45-67.

Leonard, J.A., Kramer, M.A. & Ungar, L.H., 1992. Using radial basis functions to approximate function and its error bonds, IEEE Trans. on Neural Networks 3, 624-627.

Martinez-cob, A. and Curnca, R.H., 1992. Influence of elevation on regional evapotranspiration using multivariate geostatistics for various climatic regimes in Oregon. J. Hydrol. 136, 353-380

McAvoy, T.J., Wang, N.S., Naidu, S., Bhar, N., Gunter, J. and Simmons, M., 1989. Interpreting biosensor data via backpropagation. IEEE International Neural Network Conference in 1989, 8, 227-233.

Michaels, P.J., Balling Jr. R. C., Vose, R.S., Knappenberger, P. C., 1998. Analysis of trends I the variability of daily and monthly historical temperature measurements. Clim. Res. 10, 27-23.

Nalder, I.A., Wein, R.W., 1998. Spatial interpolation of climatic normal: test of a new methods in the Canadian boreal forest. Agric. & For. Met. 9, 211-225.

Ollinger, S.V., Aber, J.D., Lovett, G. M., Millham, S.E., Lathrop, E.G., Ellis, J.M., 1993. A Spatial model of atmospheric deposition for the northeastern U.S. Ecol. Appl. 3, 459-472.

Pardo-Iguzquiza, E., 1998. Comparison of geostatistical methods for estimating the areal climatological rainfall mean using data on precipitation and topography. Int. J. Climatol. 18, 1031-1047.

Panagoulia, D., 1992. Hydrological modeling of a medium-size mountainous catchment from incomplete meteorological data. J. Hydrol 137(1-4), 279-310.

Parmar, R.S., McClendon, R.W., Hoogenboom, G., Blankenship, P.D., Cole, R.J., Forner, J.W., 1997. Estimation of aflatoxin contamination in preharvest peanuts using neural networks. Trans. ASAE 40(3), 809-813.

Price, D.T., McKenney, D.W., Nalder, I. A., Hutchinson, M. F., Kesteven, J. L., 2000. A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agric. & For. Met. 101, 81-94.

Rao, D.H. and Gupta, M.M., 1993. Dynamic neural controller with somatic adaptation. IEEE International Conference on Neural Networks in 1993, 8, 558-563.

Rich, E. and Knight, K., 1991. Artificial intelligence. New York: McGraw-Hill.

Robeson, S.M., 1993. Spatial interpolation, network bias, and terrestrial air temperature variability. Publ Climatol 46(1), 1-51

Robeson, S.M. and Janis, M. J., 1998. Comparison of temporal and unresolved spatial variability in multiyear time-averages of air temperature. Clim. Res. 10, 15-26.

Schaap, M.G. and Bouten, W., 1996. Modeling water retention curves of sandy soils using neural networks. Water Resour. Res. 32(10), 3033-3040.

Scheifinger, H. and Kromp-Kolb, H., 2000. Modeling global radiation in complex terrain: comparing two statistical approaches. Agric. For. Met. 100: 127-136.

Ward System Group, Frederic, MD, 1993. Manual of NeuroShell 2. Frederic, MD: Ward Systems Group.

Wilk, J. and Adersson, L., 2000. GIS-supported modeling of areal rainfall in a mountainous river basin with monsoon climate in southern India. Hydrol Sci. J. 45(2), 185-202.