AIR TEMPERATURE PREDICTION USING ARTIFICIAL NEURAL NETWORKS

by

BRIAN A. SMITH

(Under the Direction of Ron W. McClendon)

ABSTRACT

Extreme air temperatures are responsible for economic losses in crops and livestock of agricultural producers. Freezing temperature during the growing season damage floral buds of fruit trees and extreme heat can wither plants and lead to heat stress in livestock. Suitable air temperature predictions can provide farmers and producers with valuable information when they face decisions regarding the use of mitigating technologies such as orchard heaters or irrigation. The research presented in this thesis developed artificial neural networks models for the prediction of air temperature up to 12 hours ahead. The predictions of the final models are now available year-round for all sites in the University of Georgia's Automated Environmental Monitoring Network via the network's website, www.georgiaweather.net.

INDEX WORDS: Decision support systems, frost protection, fruit, heat stress, time-series prediction, weather modeling.

AIR TEMPERATURE PREDICTION USING ARTIFICIAL NEURAL NETWORKS

by

BRIAN A. SMITH

B.A. & B.S., The University of Toledo, 2001

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

2006

AIR TEMPERATURE PREDICTION USING ARTIFICIAL NEURAL NETWORKS

by

BRIAN A. SMITH

Major Professor:	Dr. Ron W. McClendon
Committee:	Dr. Gerrit Hoogenboom Dr. Walter D. Potter

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2006

DEDICATION

I would like to dedicate this thesis to the wonderful teachers who have fueled my love of learning throughout my life. Especially worthy of gratitude are Mr. Hutton, Mr. Adams, Mr. Stoner, Coach Cloud, Mr. Thomas, Professor Dowd, Professor Keith, Professor LeSage, Dr. Potter, Dr. McClendon, and, of course, my parents, my first teachers.

A very special thank you is owed to Dr. David Hoch, who passed away this March. I was fortunate enough to have him as a professor, an advisor, and, when I needed it, a friendly ear.

ACKNOWLEDGEMENTS

I would like to thank my all of my committee members for their time, patience, knowledge, and suggestions during the course of my research and my time in the Artificial Intelligence Center. I also owe a great debt to Abhishek Jain whose research made mine possible. Kevin Crowell and Dr. Paz have also contributed with their comments during our group meetings. Daniel Shank has been an enormous help and an excellent sounding board during his time on the Automated Environmental Monitoring Network project, helping me refine my ideas and suggesting avenues of research. I would also like to acknowledge Julian Bishop, Darren Cassella, Srigopika Radhakrishnan, and Daniel Touhy for their thoughts during the course of my research.

This work was funded in part by a partnership between the USDA-Federal Crop Insurance Corporation through the Risk Management Agency and the University of Georgia and by state and federal funds allocated to Georgia Agricultural Experiment Stations Hatch projects GEO00877 and GEO01654.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Damage to plants caused by extremely high or low temperatures is a serious concern for farmers in the state of Georgia and elsewhere in the southeastern United States. When bud formation and flowering occur during late-winter and early-spring, frost damage is a significant problem. For example, in early 1996 unseasonably cold temperatures damaged floral buds of peaches, leading to reduced fruit harvests (Okie, et al., 1998). Orchard heaters or irrigation can protect fruit trees and bushes from the worst frost damage provided that growers are given adequate warning of freezing conditions. Extreme heat can not only damage plants, but cause heat exhaustion in both livestock and farmers themselves (National Agricultural Statistical Service, 2005). The goal of the research described in this thesis was the development, using artificial neural networks (ANNs), of air temperature prediction models suitable for frost protection and the development of similar models for inclusion in general, year-round decision support aides.

The University of Georgia's Automated Environmental Monitoring Network (AEMN) was initiated in 1991 and has grown to more than 70 weather stations throughout the state of Georgia. The stations cover the breadth of the state's geographic diversity, from the coastal plain in the southeast, through the Piedmont, and into the Blue Ridge Mountains in the north (Hoogenboom, 2000). The automated, solar-powered stations are primarily situated in rural areas where the National Weather Service does not provide detailed local observations. A number of environmental variables are monitored, including air temperature, relative humidity, wind speed, wind direction, solar radiation, and rainfall. Since March 1996 these observations have been aggregated into 15-minute summaries that include averages, totals, and extremes, depending on the nature of the variable. Prior to March 1996, these summaries were aggregated hourly. The whether data and associated predictors and tools are disseminated via the AEMN website, www.georgiaweather.net.

The data provided by the AEMN have been used to develop a number of ANN models for the prediction and estimation of atmospheric variables. Along with three out-of-state sites, observations from the Tifton were used in the development of ANNs to estimate daily solar radiation (Elizondo et al., 1994). Bruton et al. (2000) developed ANN models to estimate daily pan evaporation using AEMN data, improving slightly over the accuracy of predictions resulting from regression models. Li (2002) used AEMN observations to train both site-specific and general-use ANNs to estimate daily maximum and minimum air temperature as well as total solar radiation.

Several studies have used AEMN data to develop ANN models to aid in frost protection decision support. Ramyaa (2004) developed ANNs to classify an upcoming 12-hour period as a freeze, near-freeze, or non-freeze event. These classification networks included current and prior observations for air temperature, solar radiation, wind speed, relative humidity, and rainfall. Dew point predictions can help assess the severity of frost and freeze events when coupled with accurate air temperature predictions. Ward-style ANN models to predict dew point temperatures up to twelve hours in advance were developed by Shank (2006).

Among the online decision support tools available on the AEMN are short-term air temperature predictions. Prior to July 2006, winter and early spring air temperature prediction ANNs developed by Jain et al. (2003) and Jain (2003) were implemented to provide these predictions. That work had found that air temperature, solar radiation, wind speed, and relative humidity were suitable meteorological inputs to such models. The final networks predicted air temperature from one to 12 hours ahead and included not only current observations, but up to six hours of prior observations for each input series. Jain et al. (2003) also used four cyclic input variables to encode the time of day at the point of prediction. They also faced software constraints that limited the number of patterns used in model development to 32,000. The work also relied on preliminary experiments that trained and evaluated a single network to determine the effects of changes to model inputs and parameters.

Chapter 1 introduces the problem of air temperature prediction via ANN models and describes the AEMN system from which development and evaluation data were obtained. The introduction also describes a number of studies that used AEMN data to develop ANN predictors and outlines the structure and organization of the thesis.

Chapter 2 describes the development of temperature prediction models using more advanced and flexible neural network technologies than those used by Jain et al. (2003). The research in chapter 2 resulted in general, i.e., not site-specific, air temperature prediction models for use one to 12 hours ahead. The models were developed for the winter and early spring months and were more accurate than those of Jain et al. (2003).

Chapter 3 presents research leading to the development of a set of general, year-round ANN models for air temperature prediction for one to 12 hours ahead. The research is based on the techniques developed in Chapter 2 and examines the effects of applying bootstrapping techniques from the field of machine learning. In addition, an analysis of prediction errors based on the day of year and the time of day was performed. The resulting models have been implemented on the AEMN website, www.georgiaweather.net.

Chapter 4 summarizes the research conducted in this study and presents final conclusions. It also suggests avenues of future research to improve the performance of air temperature prediction models with ANNs.

CHAPTER 2

IMPROVING AIR TEMPERATURE PREDICTION WITH ARTIFICIAL NEURAL NETWORKS[1]

[1] Smith, B.A., R. W. McClendon, and G. Hoogenboom. Published in International Journal of Computational Intelligence, 3(3): 179-186.

ABSTRACT

The mitigation of crop loss due to damaging freezes requires accurate air temperature prediction models. Previous work established that the Ward-style artificial neural network (ANN) is a suitable tool for developing such models. The current research focused on developing ANN models with reduced average prediction error by increasing the number of distinct observations used in training, adding additional input terms that describe the date of an observation, increasing the duration of prior weather data included in each observation, and reexamining the number of hidden nodes used in the network. Models were created to predict air temperature at hourly intervals from one to 12 hours ahead. Each ANN model, consisting of a network architecture and set of associated parameters, was evaluated by instantiating and training 30 networks and calculating the mean absolute error (MAE) of the resulting networks for some set of input patterns. The inclusion of seasonal input terms, up to 24 hours of prior weather information, and a larger number of processing nodes were some of the improvements that reduced average prediction error compared to previous research across all horizons. For example, the four-hour MAE of 1.40°C was 0.20°C, or 12.5%, less than the previous model. Prediction MAEs eight and 12 hours ahead improved by 0.17°C and 0.16°C, respectively, improvements of 7.4% and 5.9% over the existing model at these horizons. Networks instantiating the same model but with different initial random weights often led to different prediction errors. These results strongly suggest that ANN model developers should consider instantiating and training multiple networks with different initial weights to establish preferred model parameters.

INTRODUCTION

Frost damage is a significant concern for horticultural producers in Georgia and elsewhere in the southeastern United States, especially when bud formation and flowering occur during late-winter and early-spring. For example, unseasonably cold temperatures during early 1996 and 2002 damaged floral buds and were responsible for reduced fruit harvests (Salehi et al., 1998). Growers can take steps to lessen the effects of frost by using orchard heaters or irrigation to protect their trees and bushes from the worst damage, but these methods require advance warning of freezing conditions.

The University of Georgia's Automated Environmental Monitoring Network (AEMN) was created in 1991 and currently consists of 68 automated weather stations throughout the state of Georgia. The stations cover the breadth of the state's geographic diversity, from the coastal plain in the southeast, through the Piedmont, and into the Blue Ridge Mountains in the north (Hoogenboom, 2000). The solar-powered stations are primarily situated in rural areas where the National Weather Service does not provide detailed local observations. The monitoring stations collect weather data such as air temperature, relative humidity, wind speed, wind direction, solar radiation, and rainfall at one-second intervals. Since March 1996 these observations have been aggregated into 15-minute averages, totals, and extremes, depending on the nature of the variable. Previous observations were aggregated hourly.

Among the online decision support tools made available by the AEMN are short-term air temperature predictions. These hourly predictions range from one to 12 hours ahead and are available on the AEMN website, www.georgiaweather.net, during winter and early spring. The temperature predictions are generated by artificial neural network

(ANN) models developed by Jain et al. (2003) and Jain (2003). To predict temperature for a location, the ANNs use as inputs up to six hours of prior weather observations from the site. The models incorporate the time of day, as well as measurements of air temperature, humidity, wind speed, and solar radiation, and were developed for use from January through April. Classification models using ANNs to predict freeze events were developed by Ramyaa (2004). These networks classify observations into one of three classes depending on whether the model predicts freezing, near-freezing, or non-freezing conditions over a 12-hour prediction period. For the classification problem, the addition of recent rainfall observations as input variables was found to improve performance. ANN models have also been used to predict inputs to a special frost deposition model in order to more accurately predict frost and ice on roads and bridges (Temeyer et al., 2003).

The previous temperature prediction and classification networks faced software constraints limiting the number of patterns used in model development to 32,000 (Jain et al., 2003, Jain, 2003, Ramyaa, 2004). These studies also relied on preliminary experiments that trained and evaluated a single network to determine the effects of altering model inputs or parameters. The goal of the current research is to improve these temperature prediction models using more advanced and flexible neural network technologies. Specifically, this research explores four possible methods of improving prediction accuracy: (1) increasing the number of training patterns, (2) including input variables encoding seasonal information, (3) extending the duration of the prior data used as inputs, and (4) varying the number of nodes in the hidden layer.

METHODOLOGY

A. Data Sets

The previous temperature prediction work in the AEMN domain by Jain (2003), trained networks using a development set drawn from sites which were selected so as to encompass a broad range of conditions. Model evaluation was performed using a data set composed of sites collectively representative of the southern and central growing regions of Georgia. The same sites and years were used herein allowing for a comparison of these new results with the previous study. The model development sites included Alma, Arlington, Attapulgus, Blairsville, Fort Valley, Griffin, Midville, Plains, and Savannah, which have relatively long histories of weather data. For these nine stations the data up to and including the year 2000 were included in the development set. Model evaluation data were from 2001 to 2005, and included patterns from Brunswick, Byron, Cairo, Camilla, Cordele, Dearing, Dixie, Dublin, Homerville, Nahunta, Newton, Valdosta, and Vidalia. The previous work used the same locations for the years 2001-2003 for evaluation (Jain 2003). To allow for a direct comparison to this previous work, the evaluation data in this study was divided into two sets: the first composed of the data from 2001-2003 and the second composed of the 2004-2005 patterns. The development and evaluation sets were restricted to patterns from the first 100 days of the year, through April 9 or 10 for leap and non-leap years, respectively. This range includes winter observations and the early growing season. The data sets were restricted to "low-temperature" patterns, those with current temperature measurements below 20°C. Temperatures above 20°C were found not to be associated with freeze events within a 12-hour prediction horizon, the longest such horizon considered in this research.

Model inputs included five weather variables: temperature, relative humidity, wind speed, solar radiation, and rainfall. In addition to the current values for each observation on record, prior data, spaced at one hour intervals, were also included in each training pattern. Hourly first-difference terms for the current and prior weather variables were also included. Note that the information contained in the first-difference variables is implicit in the current and prior data, but providing this information explicitly was found to improve model performance.

Each training and evaluation pattern contained two sets of cyclic variables associated with the time and the date of the observation. Because the time of day and year are periodic variables, simply representing each with a single input fails to capture all information inherent in a measurement. To overcome this limitation, cyclic variables were constructed using fuzzy logic membership functions. For the time variable, four such triangular functions with an output range of 0 to 1 were used over the domain 0000 to 2400 hours (Figure 2.1). Note that one of the variables, corresponding to the concept midnight, “wraps around” the domain’s upper and lower bounds. An analogous approach was taken to convert the day-of-year for each observation to four seasonal variables.

B. Model Development

Software constraints restricted the previous AEMN temperature prediction models to 32,000 development patterns. To overcome this limitation, a neural network suite was written in the Java programming language. This suite placed no limits on the size of the sets used in the training or evaluation process. All networks were trained via the well-known error backpropagation (EBP) algorithm as described by Haykin (1999). EBP training was successfully applied in previous ANN research involving temperature prediction using AEMN data (Jain et al., 2003, Jain, 2003, Ramyaa, 2004).

Throughout this paper, the term model is understood to be an ANN architecture and a set of associated parameters. A model is instantiated as a network by using a random seed to assign initial weight values and a training set order and subsequently training the network. That is, a model is a description of a group of potential networks differing only in the set of initial weights assigned before training and the order in which training patterns are presented. All models explored in this research were based on the Ward-style network architecture used in previous research by Jain et al. (2003) and Jain (2003). The Ward network is an ANN with multiple node types that implement multiple activation functions (Ward Systems Group, 1993). The models used a linear input layer, three equally-sized, parallel "slabs" in the hidden layer, and a single, logistic output node, interpreted as the temperature at some prediction horizon.

A linear transformation carried out by the input layer was determined for the entire model development set. Each data series used as an input was transformed to the range 0.1 to 0.9. As the transformation made use of the maximum and minimum values of each series in the development set, this range may not hold when an evaluation pattern is presented. The hidden layer contained three slabs using the Gaussian, Gaussian complement, and hyperbolic tangent activation functions. Fully connected, biased weight matrices connect the input layer to the hidden layer and the hidden layer to the output node.

The networks used in this research provide a mapping of a vector of I real-valued inputs, x, ranging over the values [0.1, 0.9] onto a real value z such that

$$z = g(\beta_0 + \sum_{j=1}^{J} \beta_j \cdot y_j) \text{ , where}$$

$$g(n) = \frac{1}{1+\exp(-n)} \ . \qquad (1)$$

y is a vector of length J containing the signals of the nodes in the ANN's hidden layer.

$$y_j = f_j\left(\alpha_{j0} + \sum_{i=1}^{I} \alpha_{ji} \cdot x_i\right) \text{ , where}$$

$$f_j(n) = \begin{cases} \tanh(n), & \text{for } 0 < j \le j_1 \\ \exp(-n^2), & \text{for } j_1 < j \le j_2 \\ 1-\exp(-n^2), & \text{for } j_2 < j \le j_3 \end{cases} \qquad (2)$$

and $0 < j_1 < j_2 < j_3$.

Instantiating a Ward-style architecture requires specifying a number of network parameters including the learning rate and momentum, initial weight range, size of the training and testing sets, number of hidden nodes in each slab, and the included input series. Variations in the learning rate, momentum, and the initial weight range were considered in preliminary studies, but these parameters were found to have a relatively small effect on model accuracy. For all models considered in this research, a learning rate of 0.1 and an initial weight range of -0.1 to 0.1 were used. A momentum term was not included.

ANN models are typically evaluated by instantiating a single network and measuring the resulting accuracy of the trained network for a set of patterns. Such an evaluation scheme assumes that the performance of a single network is an accurate measure of any network that may instantiate the model. However, due to the random nature of the initial weights and the training pattern ordering, there is no guarantee that two networks instantiating the same model will converge to the same final state from distinct starting points in the multidimensional weight space (Salehi et al., 1998). This suggests that another method of model evaluation, involving multiple networks, is warranted.

The previous temperature prediction models developed by Jain et al. (2003) and Jain (2003) relied on single-network evaluation. An alternative approach was taken in this research whereby multiple ANNs were trained for identical model configurations. These networks, referred to as instantiations, differed only in the initial random weights and the order of the patterns presented. Each network was trained on a set of patterns independently constructed from all available development patterns via random selection without replacement for four million learning events prior to evaluation. Preliminary work using this approach showed that the use of a testing set to determine when to stop training was not helpful. Test set accuracies mirrored those of the training sets and it was rare for an instantiated network's accuracy to decrease. In addition, rare occurrences of increasing error for the testing set during training corresponded to the presence of increasing error for the training set as well. This phenomenon was also associated with poorly performing networks. After training, the mean absolute error (MAE) associated with each network's temperature prediction was calculated for the entire development set. Because the goal of the research was to develop a single, highly accurate ANN, the network with the minimum MAE of this group was selected as the appropriate performance measure for a model.

The error for the development set was used to decide between models so that comparisons between the final model and the previous research would not be biased in favor of the current work. A retrospective evaluation indicated that network MAEs for the development set are highly predictive of performance for the evaluation set.

Network training took place using 30 Dell Pentium IV workstations in a University of Georgia computer laboratory. Training was stopped after four million events because preliminary work suggested that epoch-by-epoch improvements were generally inconsequential by this time. Processing time was also a factor in the determination of the number of learning events. Using the threshold of four million learning events allowed the fastest of the machines used to train and evaluate two instantiated networks in a typical 12-hour run.

C. Experiments

To explore the effects of increased training set sizes on model performance, six models, differing only in the number of training patterns used, were instantiated by thirty networks each. Training set sizes of 10K, 25K, 50K, 100K, 200K, and 400K were considered. All weather variables and related first-difference series, as well as the four diurnal variables, were used as inputs. A six-hour duration of prior data was used for this experiment. Next, to determine the effect of adding time-of-year information to the input vector, these models were compared to a second group, modified to include the four seasonal variables. All other inputs and parameters were the same, including the six hours of prior data. A third experiment explored the effect of variations in the number of prior hourly observations for the environmental inputs by instantiating multiple models with seasonal variables for durations of two, four, six, 12, 18, 24, 30, 36, and 48 hours to determine if increasing the duration beyond six hours improved prediction accuracy. A fourth experiment was conducted comparing the accuracy of models with seasonal inputs and hidden layer sizes of six, 15, 30, 45, 60, 75, 105, 120, 150, 180, and 225. To allow a single parameter to represent the number of nodes, the three slabs were constrained to be of equal size, so that the hidden layer sizes considered ranged between two and 75 nodes per slab.

Finally, the best-performing model was instantiated thirty times for each prediction horizon from one to 12 hours ahead. The instantiation with the lowest MAE for the development set was selected to represent the model. These final models were then run over an evaluation set consisting of all cold-weather patterns from 2001 through 2003. The relationships between target temperature, predicted temperature, and prediction error for the Ft. Valley site were evaluated for these ANNs. Additionally, the performance of the models over the damaging freeze events of late-February and early-March 2002 was examined. Each model was also evaluated over a final set that consisted of all of the low-temperature patterns from the evaluation sites during winter 2004 and 2005 (those with a temperature no greater than 20°C at the time of the prediction).

RESULTS AND DISCUSSION

The results discussed here are for experiments with four-hour prediction models. Results for other horizons were qualitatively similar. Overfitting was exceedingly rare and occurred only during runs with poor prediction accuracy. An MAE for the development data was calculated for networks associated with six different models, corresponding to training set sizes that ranged from 10K and 400K unique patterns. Each model was instantiated by thirty networks (Figure 2.2). The most accurate network was trained over 50K patterns and had an MAE of 1.51°C. However, the minimum MAEs associated with the most accurate instantiations of the 50K and 200K-pattern models differed by less than 0.006°C. These training set sizes were capable of yielding similar minimum MAEs over 30 network instantiations. Furthermore, there was no clear relationship between minimum network MAE and training set size for large sets. The use of single-network evaluation allows for the possibility of misleading approximations of model accuracy. In general, each model MAE would be approximated by making a single draw from the distribution of MAEs associated with the model. However, the range of MAE values for each model is sizable. When drawing a single MAE value for each model, any combination of values is possible, many of which could suggest markedly different interpretations of the results.

The second experiment evaluated six additional models with seasonal input terms, corresponding to the six distinct training set sizes. These models were more accurate than those without seasonal inputs for all the different sizes of the training set (Figure 2.3). The most accurate model with seasonal inputs had an MAE of 1.48°C for the entire development set. This was an improvement of more than 0.03°C compared to the best model that did not consider seasonal inputs. Again there was no convincing evidence for a relationship between training set size and performance for sufficiently large sets. The difference between the most and least accurate model MAE was 0.02°C. Similar to the non-seasonal case, the 50K-pattern model using seasonal data exhibited higher accuracy than the other models. To explore whether this was typical, the seasonal experiment was repeated. This involved the instantiation of 30 new networks for each of the six models. In the second trial, the 50K-pattern model using seasonal data was slightly outperformed by the three seasonal models using larger training sets. With no evidence to suggest a preference for training set size, subsequent experiments made use of all available development patterns. All subsequent experiments continued to employ four million learning events during training.

Six hours was the preferred duration of prior data for prediction horizons of four hours or more in the prior temperature prediction study (Jain, 2003). The current research compared various models with seasonal terms that differed only in the number of hours of prior data included as inputs. The results of the experiment indicate that a prior duration of six hours is clearly suboptimal for this horizon (Figure 2.4). In fact, with an MAE of 1.48°C, the six-hour model was outperformed by all of the longer-duration models considered. The inclusion of 24 hours of prior data resulted in an MAE of 1.38°C, the lowest observed in the experiment. Models with data from more than 24 prior hourly observations were less accurate. The success of the 24-hour model makes intuitive sense as such a history is capable of generalizing over trends associated with the familiar daily cycle. The decision in previous research to use six hours of prior data was likely due to the method of increasing the duration by short increments until evaluation errors began to increase (Jain, 2003). Because that work relied on single-network evaluations and found that a network with eight hours of prior information was less accurate than a six-hour network, the experiment was stopped before exploring longer durations of prior data. The results of the research reported herein suggest that the use of multiple-network evaluation can avoid such errors.

The final experiment instantiated networks for models with seasonal inputs that differed in hidden layer size. Even if a 75-node hidden layer (25 nodes per slab) was optimal for a model with six hours of prior data, there was no guarantee that it would be best for a model with 24 hours of prior data and seasonal inputs. The results of the experiment, which evaluated each model over 30 instantiations, revealed that for models with 24 hours of prior data, a larger network with 120 hidden nodes (40 per slab) led to an instantiation with an MAE of 1.35°C, the smallest of the models considered (Figure 2.5). Increasing the total number of hidden nodes beyond this level did not reduce average prediction error for any of the models considered. The time-consuming nature of the training process precluded the possibility of evaluating all possible models.

To establish a direct comparison between the models developed here and those obtained by Jain (2003), 30 networks were instantiated for each prediction period between one and 12 hours. For each horizon, the network having the lowest MAE for the development set was selected to represent the model. The selected networks were evaluated with the same 2001-2003 weather data for the Brunswick, Byron, Cairo, Camilla, Cordele, Dearing, Dixie, Dublin, Homerville, Nahunta, Newton, Valdosta, and Vidalia sites used by Jain (2003). These data did not include input patterns with a temperature greater than 20°C at the time of prediction.

The prediction accuracies of the best ANN models developed in this study are compared to those obtained by Jain (2003) in Table 2.1. The models developed in the current research made use of seasonal input terms, 24 hours of prior observations, and 120 hidden nodes and led to an improvement in model MAE over all horizons. For example, the four-hour prediction MAE of 1.40°C is an improvement of 0.20°C, or 12.5%, compared to the previous model. The MAE improvements at the one-, eight-, and 12-hour horizons of 0.09°C, 0.17°C, and 0.16°C respectively, do not provide a clear pattern relative to forecast horizon. However, the percent improvement in the MAE compared to the previous model decreases as the prediction period increases, from a more than 14% improvement at the one-hour horizon to less than 6% at the 12-hour horizon. The new networks were also evaluated for a data set consisting of the same sites with patterns from 2004-2005. For this set the magnitudes of the errors were, in general, slightly smaller than those associated with the 2001-2003 period.

The distribution of prediction errors across all horizons is centered near zero, while the variance of these error distributions increases relative to horizon length. The increased divergence between predicted and observed temperatures at longer horizons is apparent in the plots of Figure 2.6. As prediction horizon increases, so does deviation from the line of perfect fit. The trend holds, specifically, in cases where a model fails to predict freezing temperatures. At the other extreme, the use of a logistic activation function in the output node, and the inverse of the scaling function to convert the output to a temperature, placed an upper bound on the model predictions. Because the scaling range was smaller than the output range of a logistic node, this bound was several degrees higher than the 20°C threshold used to select observations for the development and evaluation sets. As a result, models were constrained from predicting temperatures above 25°C. At temperatures near 25°C, models are more likely to under-predict. As the prediction horizon increases, the number of observed temperatures above this threshold increases. The February 25 – March 1, 2002 time frame for the Fort Valley site provides an illustration of the relative performance of the final models. This period included three freeze events during the mornings of February 27 – March 1. The first of these freezes occurred shortly after 0200 on the 27th. This freeze, however, was not predicted by the 12-hour model (Figure 2.7a). Instead, the 12-hour model predicted a near freeze shortly after the observed temperature dropped below freezing. The model performed much better over the second freeze period: both the time of onset and the severity of the freeze were accurately predicted. While the predicted onset of the third freeze was off by several hours, it still indicated an approaching, sustained freeze more than six hours prior to the temperature falling below zero.

The eight-hour model predicted a brief freeze during the morning of February 27th (Figure 2.7b). Though the time of onset and severity of the freeze were not perfect, the model predicted several hours of near-freezing temperatures, a noticeable improvement over the 12-hour model. The predictions for the second and third freeze events were similar to the 12-hour model. As a practical matter, the prediction of a near-freeze event by a long-horizon model would alert the user to the possibility of damaging temperatures.

The four-hour model predicted the first freeze event, though the time of onset was off by nearly three hours (Figure 2.7c). Subsequently, however, the model's prediction of the severity of the first freeze event was quite close to the true minimum temperature. The four-hour model also correctly predicted the time of onset of the second freeze, which began later that evening and lasted well into the 28th. The freeze event ending March 1st was predicted with much better accuracy than either the 12- or eight-hour models managed, though time of onset was two hours late.

The most useful measure of model performance, however, comes from evaluating the sequence of 12 predictions leading to severe freeze events such as those in February and March 2002. Such a sequence is comprised of a chain of predictions generated at the same time for all 12 prediction horizons. The observed temperatures for Fort Valley during the period from 1400, February 28 to 1000, March 1, 2002 and the series of chained predictions generated at 1600 on February 28 are shown in Figure 2.8. These predictions suggest a shallow freeze beginning sometime between 0300 and 0400 the following morning. In fact, overnight temperatures would dip below freezing by 2200 and bottom out below -4°C. This early, imprecise, series of predictions is subsequently refined in the presence of new data. The user, already alerted to the potential of damaging temperatures, could receive a much more accurate sequence of predictions four hours later. The predictions made using the data available at 2000 on February 28 correctly indicate a sustained freeze lasting at least until the end of the 12-hour horizon (Figure 2.8). These sequences of predictions show that the model was able to provide useful and actionable information to its users, even when early predictions were imperfect. The retrospective application of the final temperature prediction models to patterns from outside the development set suggests that users, once made aware of freezing or near-freezing temperature predictions, would be well served by checking for updated predictions throughout the day.

SUMMARY AND CONCLUSIONS

The research presented in this paper explored improvements for the ANN models that are currently used to predict temperature for the Georgia AEMN data. Improvements included larger training set sizes, seasonal input terms, an increased duration of prior observations, and varying the size of the hidden layer. Increases to the size of the training set slightly reduced the prediction errors. However, the inclusion of seasonal variables corresponding to membership in the fuzzy sets winter, spring, summer, and fall did improve model accuracy, even though all observations were from the January-April period. Similar improvements resulted from extending the duration of historical data in the input vector from six to 24 hours. Models with a hidden layer with 40 nodes per slab were more accurate than other models over repeated instantiations.

The results of this work suggest avenues for further study. The introduction of seasonal terms may provide a means of implementing an accurate year-round temperature prediction model. Likewise, ensemble network approaches are worth investigating, as networks with similar MAEs over the same prediction horizon often make different predictions. Finally, when applied to data-rich environments, a clear distinction should be maintained between abstract neural network models and actual instantiations of these models. The performance of a single instantiated network is not likely to be a valid measure of model performance. In this study, model evaluation over multiple instantiations led to better parameter selection by presenting more accurate comparisons of distinct models than those afforded by single-network evaluation.

ACKNOWLEDGEMENTS

This work was funded in part by a partnership between the USDA-Federal Crop Insurance Corporation through the Risk Management Agency and the University of Georgia and by state and federal funds allocated to Georgia Agricultural Experiment Stations Hatch projects GEO00877 and GEO01654.

REFERENCES

Haykin, S., 1999. Neural networks: a comprehensive foundation, 2nd edition. Upper Saddle River, NJ: Prentice Hall: 161-175.

Hoogenboom, G., 2000. The Georgia automated environmental monitoring network. Preprints of the 24th Conference On Agricultural and Forest Meteorology, American Meteorological Society: 24-25.

Jain, A., R. W. McClendon, G. Hoogenboom, and R. Ramyaa, 2003. Prediction of frost for fruit protection using artificial neural networks. American Society of Agricultural Engineers, St. Joseph, MI, ASAE Paper 03-3075.

Jain, A., 2003. Frost prediction using artificial neural networks: A temperature prediction approach. M.S. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA.

Okie, W. R., G. L. Reighard, W. C. Newall, Jr., C. J. Graham, D. J. Werner, A. Powell, G. Krewer, and T. G. Beckman, 1998. Spring freeze damage to the 1996 peach and nectarine crop in the southeastern United States. HortTechnology, vol. 8, pp. 381-386.

Ramyaa, 2004. Frost prediction using artificial neural networks: A classification approach. M.S. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA.

Salehi, F., R. Lacroix, and K. M. Wade, 1998. Effects of learning parameters and data presentation on the performance of backpropagation networks for milk yield prediction. Transactions of the ASAE, 41: 253-259.

Temeyer, B. R., W. A. Gallus, Jr., K. A. Jungbluth, D. Burkheimer, and D. McCauley, 2003. Using an artificial neural network to predict parameters for frost deposition on Iowa bridgeways.Proceedings of the 2003 Mid-Continent Transportation Research Symposium, Iowa State University, Ames, IA.

Ward Systems Group, 1993. Manual of NeuroShell 2, Frederick, MD.

Table 2.1
Comparison of model prediction accuracy over the evaluation dataset

Horizon length, hours	Previous model* 2001-3 MAE, °C	Current model 2001-3 MAE, °C	Improvement, °C	Improvement, %	Current model 2004-5 MAE, °C
1	0.62	0.53	0.09	14.5%	0.53
2		0.88			0.86
3		1.17			1.12
4	1.60	1.40	0.20	12.5%	1.34
5		1.62			1.55
6		1.81			1.72
7		1.99			1.87
8	2.30	2.13	0.17	7.4%	2.01
9		2.24			2.09
10		2.36			2.19
11		2.44			2.25
12	2.69	2.53	0.16	5.9%	2.33

*Jain [4]

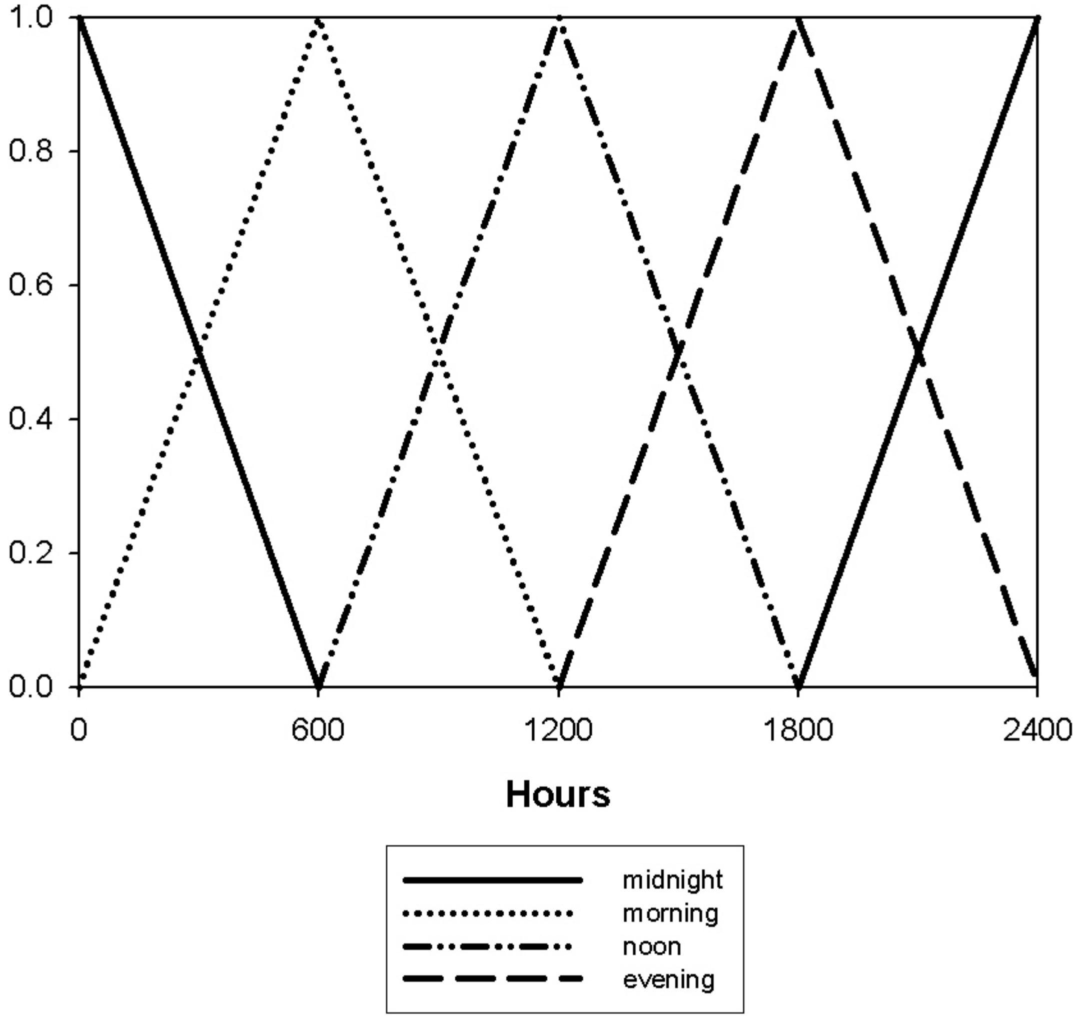

Figure 2.1: 1 Four fuzzy logic membership functions ranging over the time of day

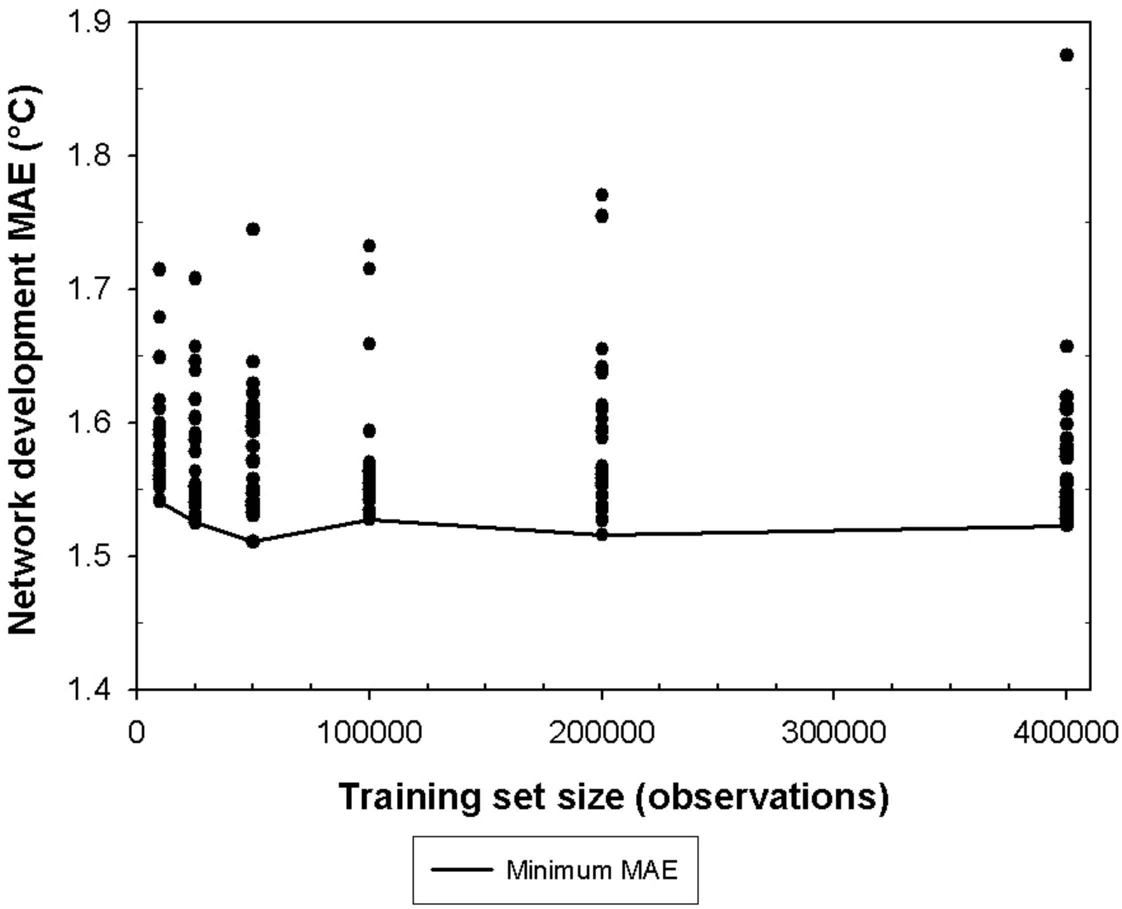

Figure 2.2: Multiple-network evaluation for four-hour prediction models distinguished by training set size. Each point corresponds to the MAE, calculated for all patterns in the development set, of one of the 30 networks instantiating each model.

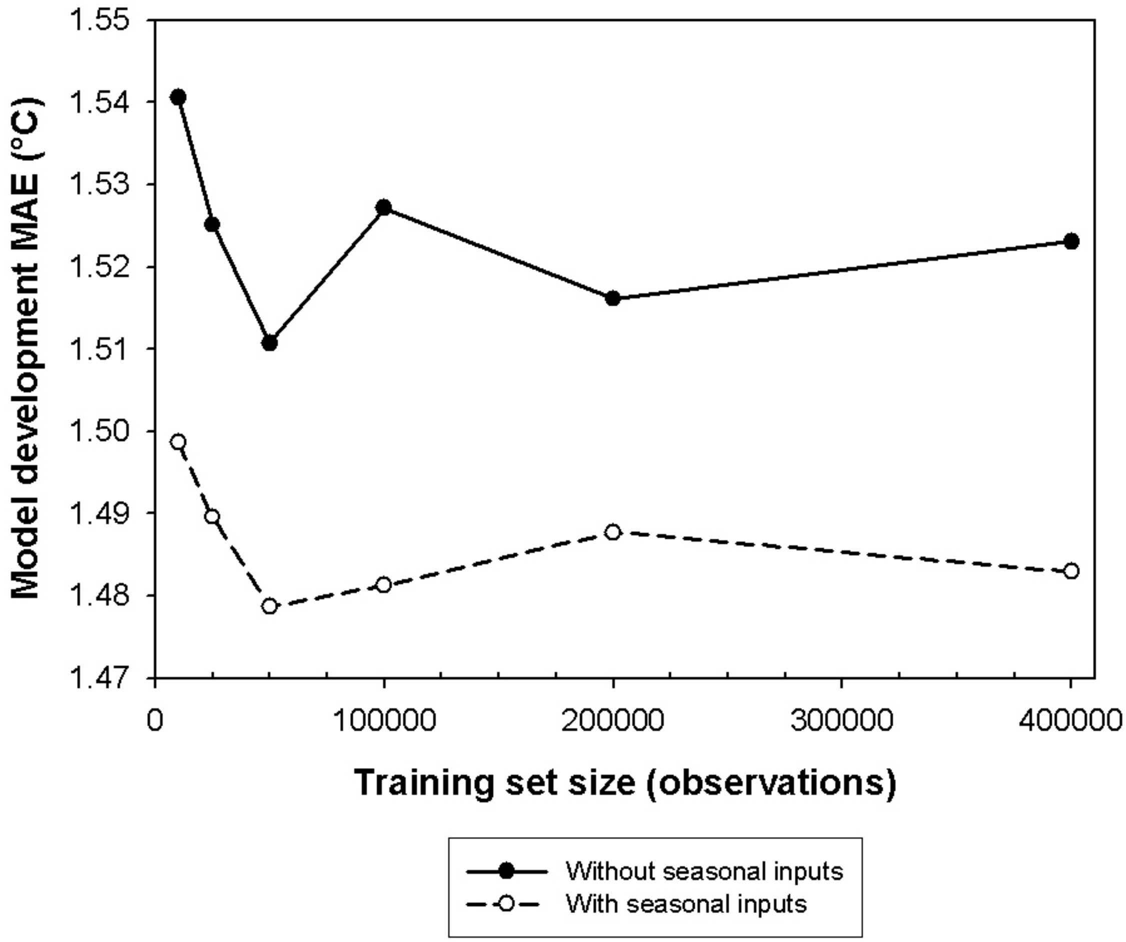

Figure 2.3: A comparison of four-hour prediction models with and without seasonal input terms using minimum-error, multiple-network evaluation. Each point corresponds to the minimum MAE obtained over 30 networks instantiating each model.

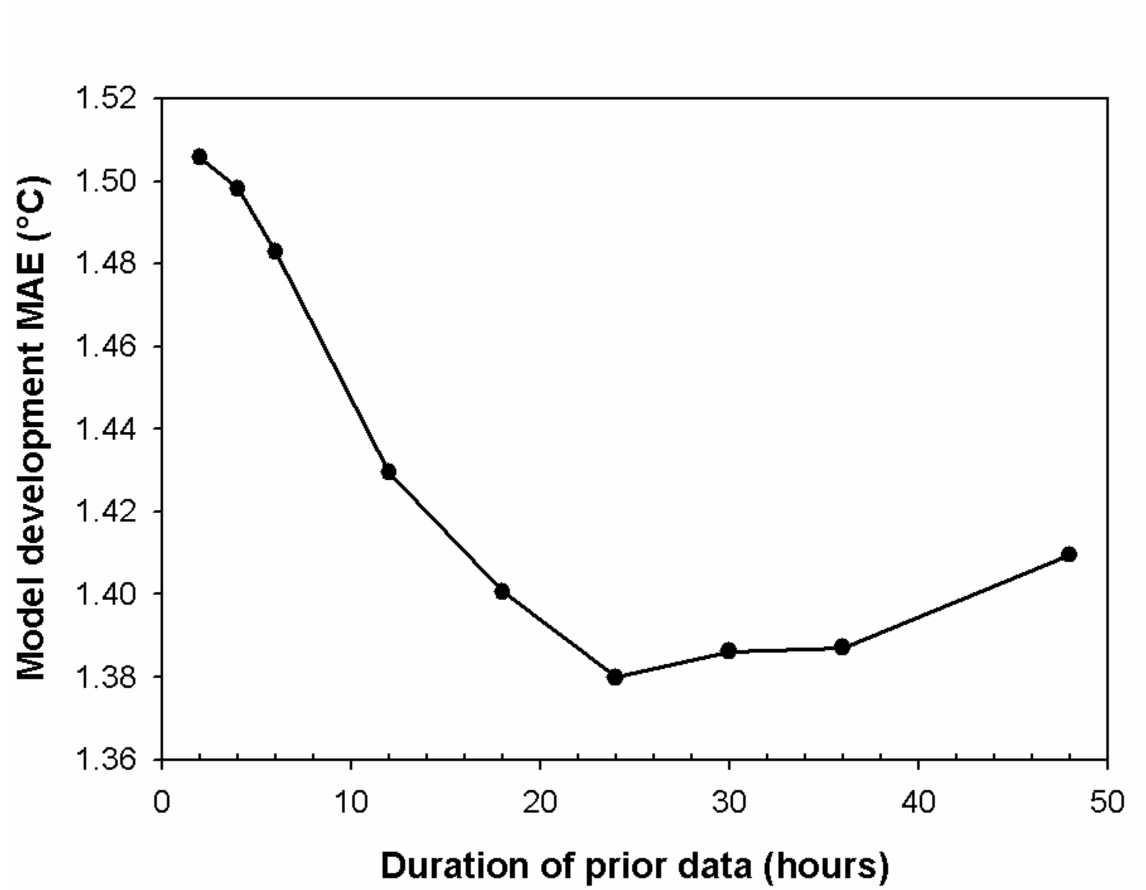

Figure 2.4: A comparison of four-hour prediction models distinguished by the duration of prior data using minimum-error, multiple-network evaluation. Each point corresponds to the minimum MAE obtained over 30 networks instantiating each model.

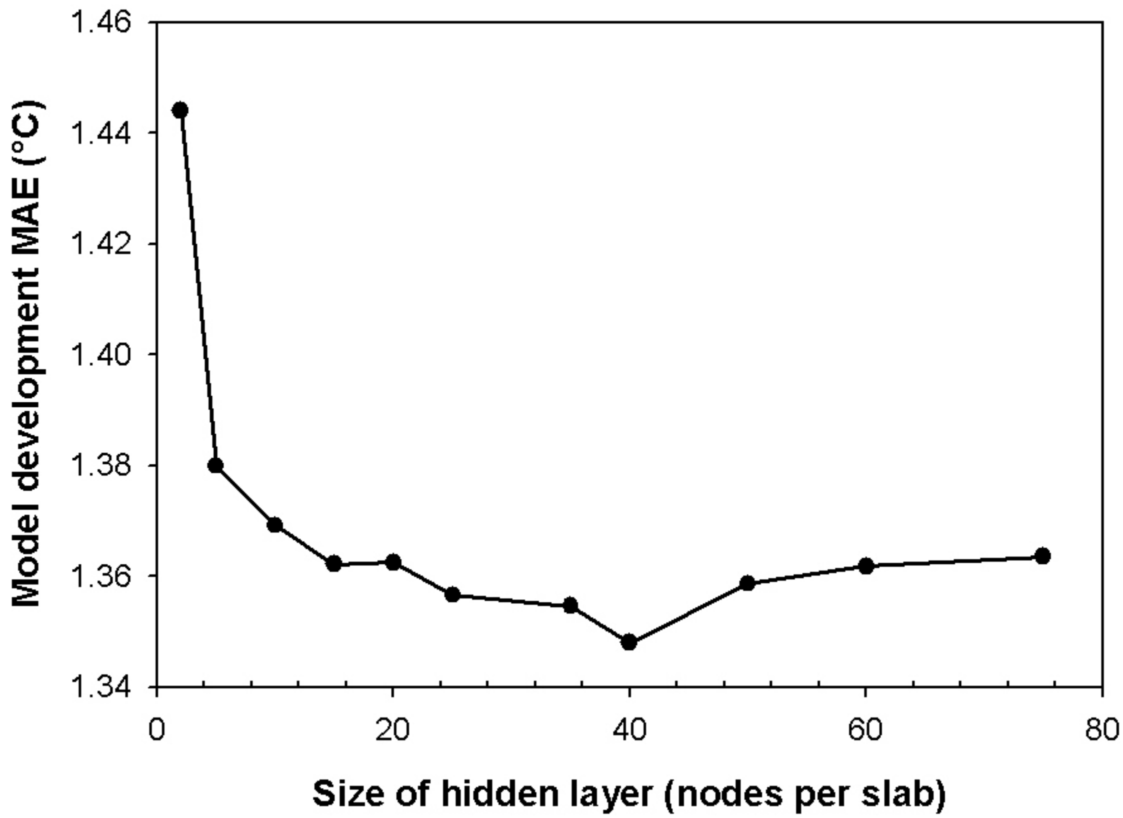

Figure 2.5: A comparison of four-hour prediction models distinguished by hidden layer size using minimum-error, multiple-network evaluation. Each point corresponds to the minimum MAE obtained over 30 networks instantiating each model. All models use three equally-sized slabs per hidden layer.

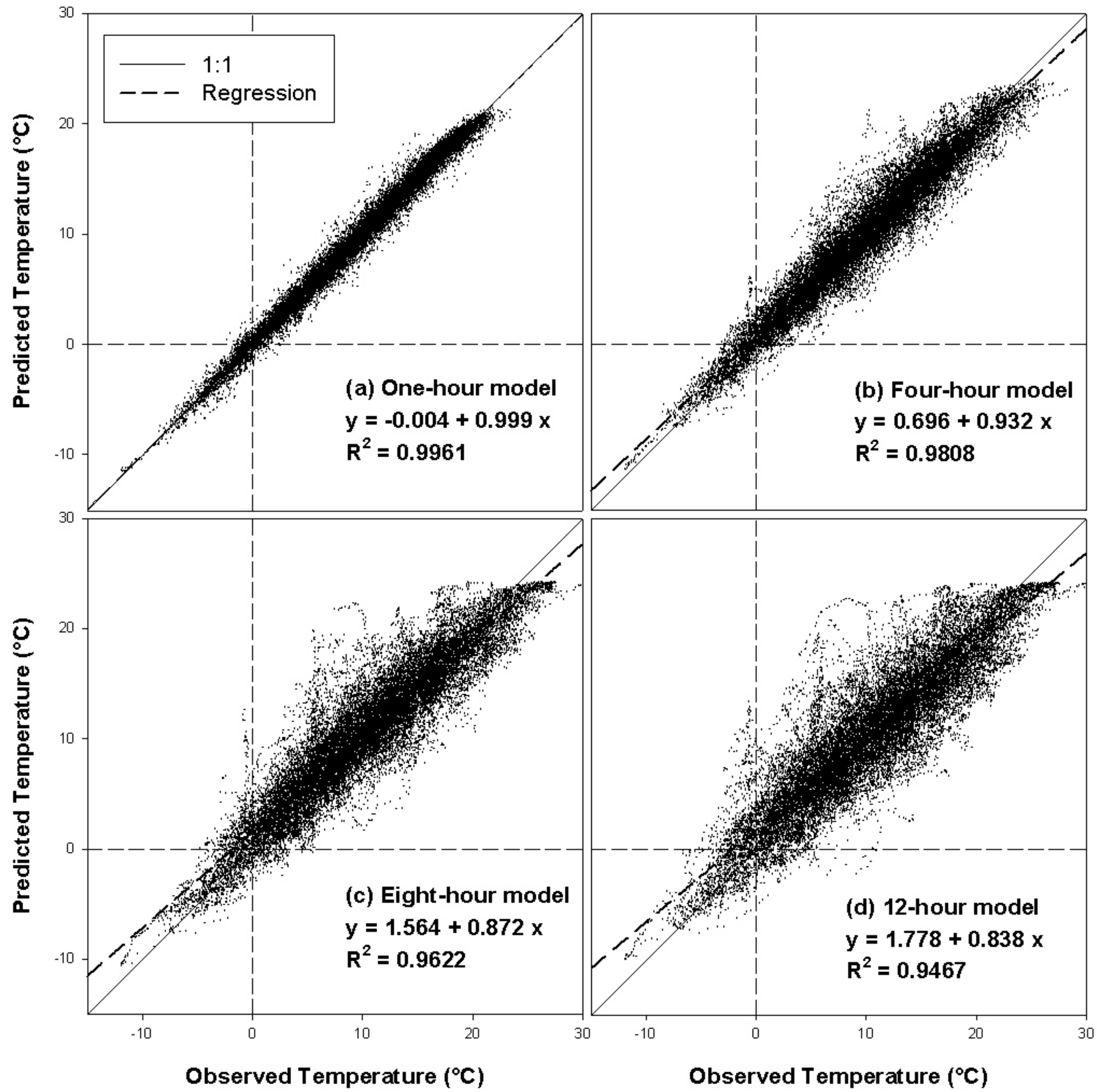

Figure 2.6: A comparison of predicted and observed temperatures for the 2001-2003 evaluation set for the final (a) one-hour model, (b) four-hour model, (c) eight-hour model, and (d) 12-hour model. A solid diagonal line indicates a hypothetical perfect model.

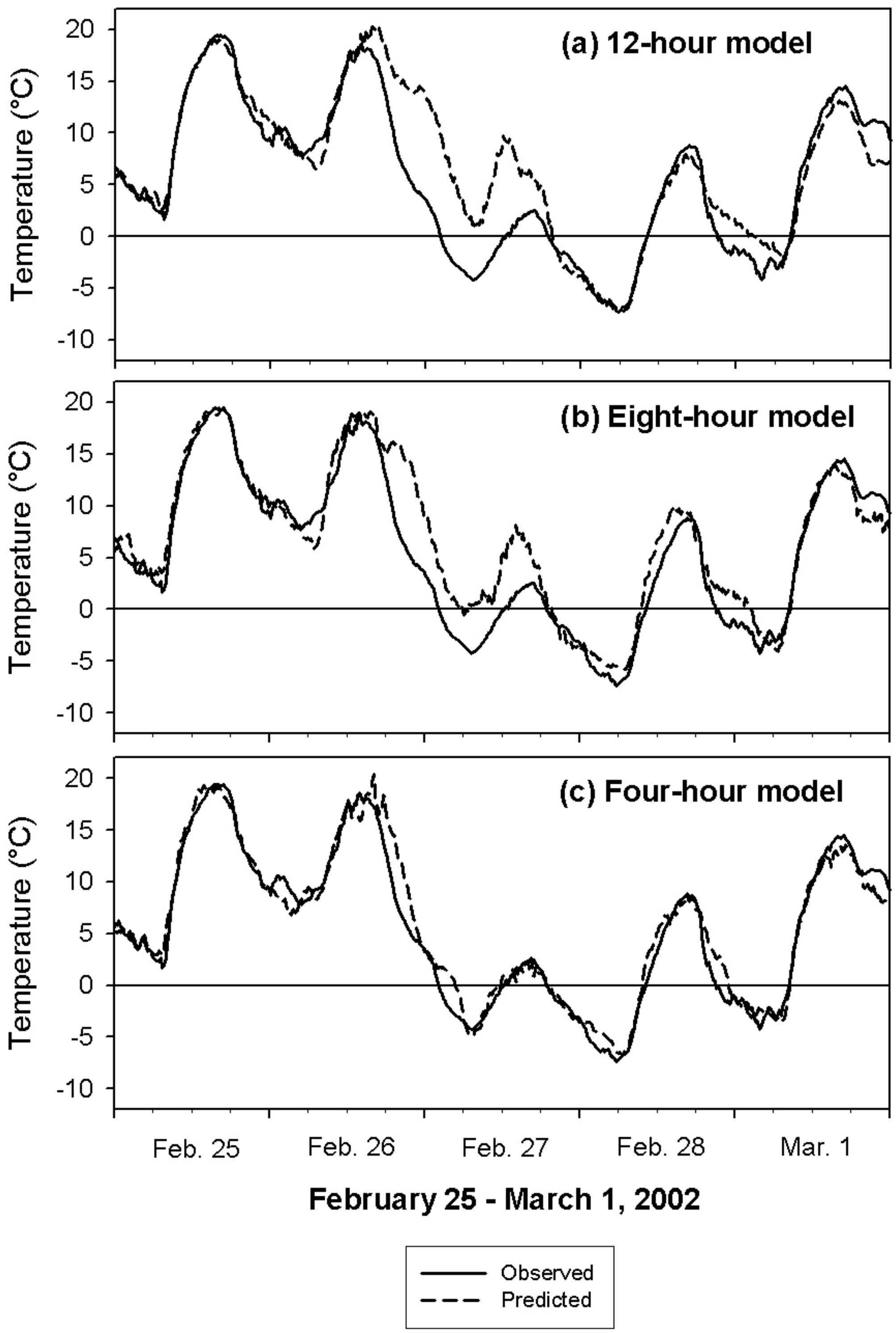

Figure 2.7: Time-series plots of observed and predicted temperatures from the final (a) one-hour model, (b) four-hour model, (c) eight-hour model, and (d) 12-hour model for the period of February 25-March 1, 2002.

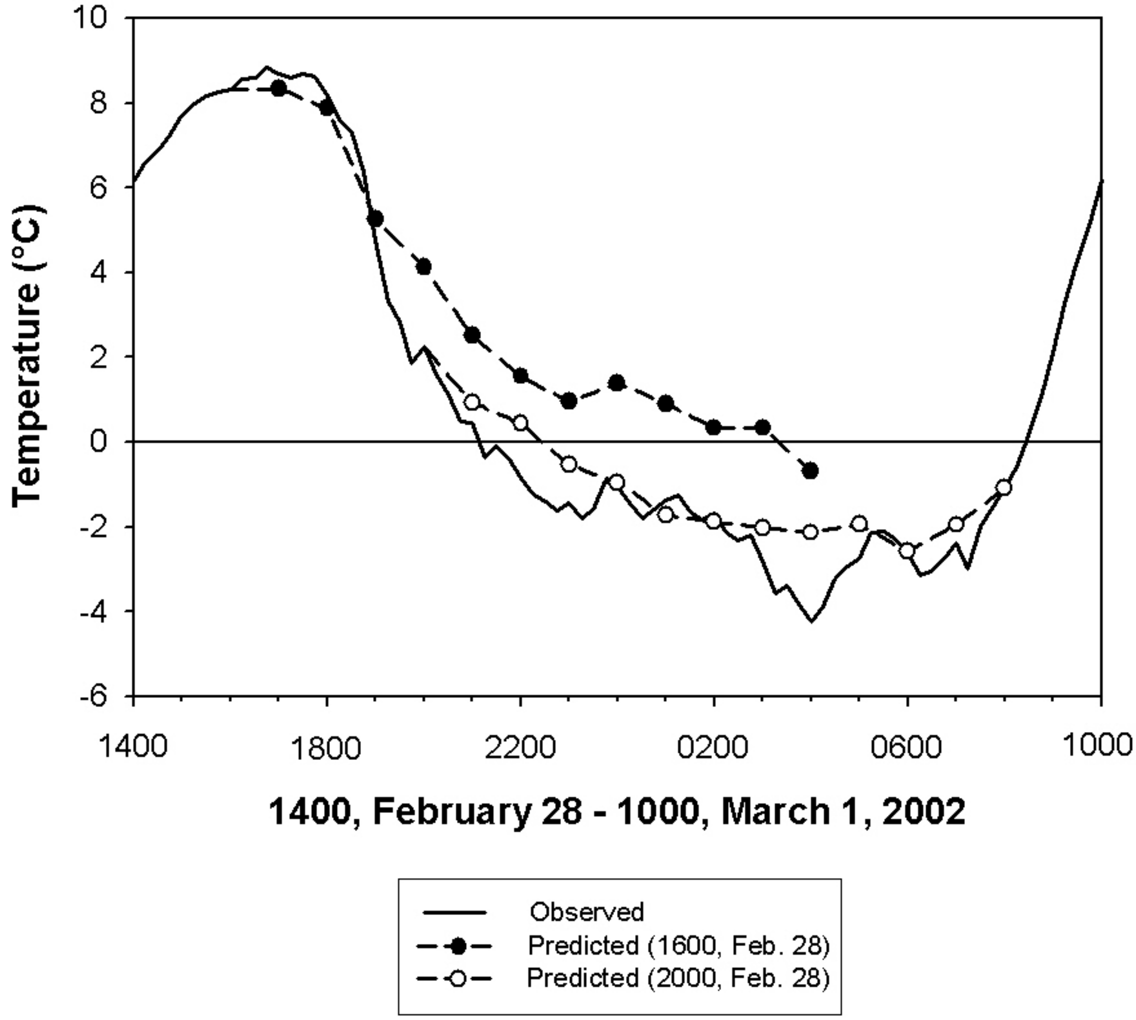

Figure 2.8: A time-series plot of observed temperatures and 12-hour prediction tracks during February 28-March 1, 2002.

CHAPTER 3

ARTIFICIAL NEURAL NETWORKS FOR YEAR-ROUND TEMPERATURE PREDICTION[2]

[2] Smith, B. A., G. Hoogenboom, and R. W. McClendon. To be submitted to Agricultural and Forest Meteorology.

ABSTRACT

Crops and livestock of farmers in much of the southeastern United States are susceptible to damage from extreme cold and heat. Given suitable warning, agricultural and horticultural producers can mitigate the damage of extreme temperature events through appropriate techniques such as orchard heating or irrigation. To provide such warning, air-temperature prediction models were developed for use as general, year-round decision support tools in the state of Georgia using a Ward-style artificial neural network (ANN) architecture. The final models were applied to prediction horizons of one to 12 hours ahead. The year-round ANN models reduced mean absolute error (MAE) for winter observations relative to an existing winter-specific model. Prediction MAEs for a final, year-round evaluation set ranged from 0.516°C at the one-hour horizon to 1.873°C at the 12-hour horizon. MAEs of the final models during the winter months were less than those resulting from the application of previously-developed winter-specific models (Smith et al., 2006).

A detailed analysis of MAE by time of year and time of day was performed. A tendency to over-predict temperatures during summer afternoons was associated with localized cloud cover during that period. The results from this study suggest that accurate cloud cover predictions could be used to reduce the errors of the air temperature predictions. The inclusion of rainfall variables as inputs to the model was shown to improve prediction accuracy. Bootstrapping techniques were explored and neither bagging nor boosting was found to reduce prediction errors.

INTRODUCTION

Artificial neural networks (ANNs) have been used in a number of prediction studies involving atmospheric time series. Yi and Prybutok (1996) predicted daily maximum ozone levels in Texas metropolitan areas with a simple three-layer ANN model with nine inputs and four hidden nodes and found it to be superior to statistical methods. A three-layer ANN model with 17 inputs was developed by Jiang et al. (2004) to predict the air pollution levels of cities in China. Inputs to the models were not site-specific, allowing the model to be applied to a number of locations across China. Air temperature, wind spend, and relative humidity in Saskatchewan, Canada were predicted 24 hours in advance by ANN models developed and applied by Maqsood et al. (2004). They found that combining the outputs of a standard feed-forward ANN, a recurrent ANN, a radial basis function network, and a Hopfield network into a simple "winner-take-all" ensemble led to more accurate predictions of wind speed, relative humidity, and air temperature than any of the component networks. ANNs have also been used to predict *indoor* air temperature. Ruano et al. (2005) used a multi-objective genetic algorithm to develop a radial basis function neural network model for the prediction of building temperature in a secondary school in Portugal. Air-conditioning control scheme simulations indicated that temperatures could be more consistently managed and that air conditioner run times could be reduced using the model.

In 1991, the University of Georgia initiated the Georgia Automated Environmental Monitoring Network (AEMN) to collect weather data from sites across the state (Hoogenboom, 2000). This network has expanded to more than 70 sites that gather local information on a variety of environmental variables. The data are obtained at one-second intervals and aggregated into 15-minute summaries. Each site consists of an automated, solar-powered station that periodically downloads the data to a central server located at the University of Georgia Griffin Campus. The weather data are gathered primarily from locations for which detailed observations from the National Weather Service are unavailable. The data are disseminated via the AEMN website, www.georgiaweather.net. The website also provides a number of calculators, maps, and decision support tools to assist agricultural producers and other users. Twelve-hour air temperature predictors are available during the winter and early spring. Accurate temperature predictions during this time of year can provide advance warning of upcoming freezes, allowing fruit growers to employ mitigation techniques, such as irrigation or orchard heating, to reduce the damage to the developing flowers and fruits.

The data provided by the AEMN have been used to develop ANN models for the prediction or estimation of atmospheric variables. Weather observations from the Tifton site along with three out-of-state sites were used to develop ANNs to estimate daily solar radiation based on daily minimum and maximum air temperatures, rainfall, and calculated values of clear-sky radiation and length of day (Elizondo et al., 1994). Data from three AEMN sites were used by Bruton et al. (2000) to develop ANN-based models to estimate daily pan evaporation. The predictions of the final model were a slight improvement over those of regression models. Li et al. (2004) developed ANNs to estimate daily maximum and minimum air temperature as well as total solar radiation for sites in Tifton and Griffin, Georgia using AEMN data from nearby sites.

Several studies have used AEMN data to develop ANN models to aid in frost protection decision support. Air temperature, solar radiation, wind speed, and relative humidity were found to be suitable meteorological inputs for air temperature prediction ANNs during the winter and early spring (Jain et al., 2003; Jain 2003). Inputs to the final networks, which predicted air temperature from one to 12 hours ahead, included up to six hours of prior observations for each input series. In addition, the work encoded the time of day at the point of prediction using four cyclic input variables. Ramyaa (2004) developed classification ANNs to predict freeze events rather than air temperature. The networks were trained to classify an upcoming 12-hour period as a freeze, near-freeze, or non-freeze event. In addition to the inputs identified by Jain (2003), these classification networks included current and prior rainfall observations. Ward-style ANN models to predict dew point temperatures up to twelve hours in advance were developed by Shank (2006). Predictions of dew point temperature can help assess the severity of frost and freeze events when coupled with accurate air temperature predictions.

Smith et al. (2006) found that instantiating each model with multiple networks increased the likelihood that ANN design alternatives were evaluated accurately. This approach was applied to the development winter-only air temperature prediction models with the goal of developing more accurate models than those of Jain et al. (2003). The predictors were developed with prediction horizons from one to 12 hours ahead using the Ward-style ANN architecture and AEMN weather data. Observations from January through early April, at which the temperature at the time of prediction was less than or equal to 20°C, were used for ANN development and evaluation. Rainfall observations were included as inputs based on the results of Ramyaa's (2004) freeze classification approach. Smith et al. (2006) found that models which included cyclic, day-of-year variables and 24 hours of prior data as inputs produced more accurate predictions, as measured by mean absolute error (MAE), than models without such inputs. The inclusion of day-of-year variables suggested the possibility of extending the methodology to a year-round prediction scheme.

The goal of the research presented herein was to develop a set of year-round ANN models for air temperature prediction for one to 12 hours ahead for inclusion in general, year-round decision-support aids. The objectives related to this goal were: (1) a comparison of the accuracy of year-round models to that of existing winter models, (2) a determination of the effect on accuracy of including rainfall input terms, (3) an examination of bootstrapping techniques from the field of machine learning in order to improve prediction accuracy, and (4) an analysis of prediction errors based on the day of year and the time of day to assess the suitability of the models for general decision making.

METHODOLOGY

A. Data sets

Raw data observations delivered by the automated stations of the AEMN system were formatted and scaled into input-target patterns. Three sets of patterns were created: a development set, a selection set, and an evaluation set. The development set was used to train multiple networks for each model and select a preferred network from among those instantiating the same model. Distinct instantiations of the same model differed in the initial random weights assigned to the network and the order of presentation of training patterns. The development set was created from observations recorded during the years 1997-2000 and consisted of approximately 1.25 million patterns. The patterns in the development set were drawn from the stations located in Alma, Arlington, Attapulgus, Blairsville, Fort Valley, Griffin, Midville, Plains, and Savannah. These sites are located across the state and represent both geographical and agricultural diversity. For example, Arlington and Attapulgus are in the Georgia Coastal Plain, Blairsville is located in northern Georgia's Appalachian Mountains, Griffin is in the central Piedmont, and Savannah is located on the Atlantic coast.

Observations from Brunswick, Byron, Cairo, Camilla, Cordele, Dearing, Dixie, Dublin, Homerville, Nahunta, Newton, Valdosta, and Vidalia were used for model selection and evaluation. Collectively, these stations were drawn from important agricultural production areas in the southern and central parts of the state of Georgia. The selection set patterns were drawn from 2001-2003 and numbered approximately 1.25 million. The MAE for this set was used to select between competing ANN design alternatives. The evaluation set from the years 2004 and 2005 consisted of approximately 800,000 patterns. This data set was used to evaluate the accuracy of the final model. Both the model development and evaluation sites were the same as those used by Smith et al. (2006).

Previous models developed by Jain (2003) and Smith et al. (2006) predicted air temperature to aid in frost prediction and were trained on patterns generated from observations that occurred during the first 100 days of the year. The observations used were also restricted to those in which the temperature at the time of the prediction was less than or equal to 20°C. As a result of these constraints, the models would not be expected to perform well at higher temperatures or during other times of the year. In fact, the models developed by Smith et al. (2006) were incapable of predicting temperatures above 25°C.

The goal of the current research was to develop models that perform well over the entire year without sacrificing accuracy over winter observations. Therefore, winter-period subsets of the development, selection, and evaluation sets were created. Decisions regarding preferred networks and models were based on the MAE for year-round data sets, but the MAE calculated using the winter subsets was considered as an indication of model accuracy during the winter months.

Current values and a 24-hour duration of prior observations for air temperature, solar radiation, wind speed, rainfall, and relative humidity from the time of prediction were used as inputs for the ANN models. The hourly rate of change in each of the five weather variables at the prediction point and at one-hour intervals over the previous day were also included as inputs for the models. These 250 inputs were rescaled using a linear transformation such that all observations used in model development were in the range [0.1, 0.9]. The scaled value, x_{scaled}, of an observation, x, was given by

$$x_{\text{scaled}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \cdot 0.8 + 0.1, \tag{1}$$

where $x_{\min}$ and $x_{\max}$ were the minimum and maximum values of the variable found in the development data set. Because the output of the network was restricted to the domain [0, 1], it was necessary to map the output signal back to the range of expected temperatures. The inverse of the linear function used to scale the input temperature was used for this purpose.

The time of day and the day of year were each encoded as four cyclic variables using triangular fuzzy-logic membership functions in the range [0, 1]. The four time-of-day membership functions are presented in Figure 3.1. The variable *midnight* "wraps around" the day, with a maximum value at 2400 hours. The figure is a smooth, continuous representation of the membership functions. In practice AEMN observations are aggregated at fifteen-minute intervals beginning at the top of the hour. The values of the time-of-day membership functions were determined by the hour in which the observation occurred. Day-of-year variables were treated in a similar manner, using four membership functions to represent seasonality. With the inclusion of these eight cyclic variables, each model that was considered had a total of 258 inputs.

B. Model development

Ward-style ANNs were used for model development, as in previous studies of air temperature prediction with AEMN data (Jain et al., 2003, Ramyaa, 2004, and Smith et al., 2006). These are feed-forward, backpropagation networks that implement multiple activation functions in a single hidden layer. In the Ward-style networks used in this study, the signal of the kth output node, z_k, iss given by

$$z_k = g(\sum_{j=0}^{J} \beta_{kj} \cdot y_j), \text{ where } g(n) = \frac{1}{1+\exp(-n)}. \quad (2)$$

The term y_0 is set to a constant value of 1 and the coefficient β_{k0} is the *bias* of the kth output node. The value of the summation over the J hidden nodes is referred to as the *induced local field* of the node (Haykin, 1999). Some ANN architectures use the same activation function, g, for each node in the hidden layer. The Ward-style architecture, in contrast, makes use of multiple activation functions in the hidden layer. The value of y_j, the jth hidden node in a Ward-style network, is given by

$$y_j = f_j\left(\sum_{i=0}^{I} \alpha_{ji} \cdot x_i\right), \text{ where}$$

$$f_j(n) = \begin{cases} \tanh(n), & \text{for } 0 < j \le j_1 \\ \exp(-n^2), & \text{for } j_1 < j \le j_2 \\ 1-\exp(-n^2), & \text{for } j_2 < j \le J \end{cases} \text{ and } 0 < j_1 < j_2 < J \ . \tag{3}$$

As in Equation 2, $x_0 = 1$, so that each α_{j0} serves as a bias for its respective hidden node. All "slabs" of nodes with the same activation function were constrained to be of an equal size. As the models were used to predict a single output, temperature, there was only a one output node z.

The α and β terms in Equations 2 and 3 are adjustable weight coefficients. Training a neural network via error back propagation (EBP) is an attempt to identify a set of weights that reduces the mean squared error (MSE) of the output node over a development set. EBP performs a local gradient search in the space of possible weights. The algorithm was applied after the presentation of each training pattern and a single feed-forward pass followed by a weight adjustment constituted a learning event. For a training pattern n with target $t(n)$, the prediction error and error energy of a network with output $z(n)$, are $e(n)$ and $E(n)$, respectively, where

$$e(n) = t(n) - z(n) \ , \tag{4}$$

and

$$E(n) = \frac{1}{2} \cdot e(n)^2 \ . \tag{5}$$

The change to a weight w in light of a training pattern, n, and a learning rate, η, was given by the delta rule:

$$\Delta w = -\eta \cdot \frac{\partial E(n)}{\partial w} \ . \tag{6}$$

The value of the last term in Equation 6, the partial derivative of the network error energy with respect to a selected weight w, depends on whether the weight in question feeds a signal into the output layer or the hidden layer. In the former case, w is β_{kj}. Differentiating Equation 2 yields

$$\Delta\beta_{kj} = \eta \cdot e(n) \cdot g'(\sum_{j=0}^{J} \beta_{kj} y_j) \cdot y_j \quad . \tag{7}$$

For weights leading into the hidden layer,

$$\Delta\alpha_{ji} = \eta \cdot e(n) \cdot g'(\sum_{j=0}^{J} \beta_{kj} \cdot y_j) \cdot \beta_{kj} \cdot f_j'(\sum_{i=0}^{I} \alpha_{ji} \cdot x_i) \cdot x_i \quad . \tag{8}$$

Such weight adjustment equations can be calculated for a network with any number of layers. In practice, the networks in this research implemented a *local gradient* signal which was propagated back through the network during learning. A local gradient of a node is the partial derivative of error energy, $E(n)$, with respect to the induced local field of a node. Haykin (1999) provides a concise presentation of local gradient calculation.

The temperature-prediction models made use of several constant parameters selected from Smith et al. (2006). All networks had a hidden layer consisting of 120 nodes that were distributed equally among the three slabs comprising the layer. All initial weights were in the range [-0.1, 0.1]. The learning rate, η, was set to 0.1 and no momentum term was used.

Smith et al. (2006) established that comparing alternative ANN models by instantiating and training multiple networks for each model led to improved parameter selection and a reduction in prediction error when compared to previous air temperature prediction research (Jain, 2003). Because the EBP algorithm attempts to minimize prediction error by performing gradient descent in the space of possible weights, networks tend to converge towards local optima. Making use of repeated instantiations of the same model, differing only in the random seed used to generate initial weights and the order of training pattern presentation, is analogous to making repeated draws from the distribution of possible local optima.

C. Experiments

Each network created during the course of this research required between six to 20 hours to train and evaluate using one of 30 Dell Pentium IV workstations in a University of Georgia computer laboratory. When evaluating the accuracy of competing ANN models, each alternative was instantiated 30 times for the four-hour prediction horizon. The most accurate network for each competing model, measured in terms of MAE over the development set, was used to represent the performance of that model.

Once a representative instantiation was assigned to a model, it was evaluated over the selection set. When comparing among alternatives, the model with the lowest MAE over the selection set was preferred. All model comparisons were made on the basis of the four-hour horizon and then implemented across all other horizons. While it was possible that other horizons might have benefited from alternative design decisions, the amount of processing time necessary to investigate all horizons was prohibitively large and beyond the scope of this study. Prior work had shown that selecting model parameters on the basis of the four-hour horizon reduced MAE across all horizons (Smith et al., 2006).

One aspect of this research involved comparing the accuracy of winter and year-round models. Networks instantiating the year-round model were trained over a data set of 300,000 patterns for 15 epochs, or 4.5 million learning events. For each instantiation, the year-round model randomly selected a different subset from the more than 1.25 million training patterns available. Year-round instantiations, therefore, differed from one another not only in the initial weights and the order of the training patterns, but in the subset of training patterns itself. This constraint on the number of training patterns was used to allow for a fair comparison between the year-round and winter models. Winter networks made use of all of the approximately 300,000 patterns in the winter-specific subset of development patterns for training.

Jain (2003) found that the inclusion of rainfall variables did not increase the accuracy of temperature prediction, while Ramyaa (2004) concluded that rainfall was helpful in classifying upcoming temperature events as a *freeze*, *near-freeze*, or *non-freeze* event. Both studies relied on single-network evaluation to select inputs. Smith et al. (2006) did not address this issue, but arbitrarily included rainfall data as inputs. An experiment was conducted herein to determine the effects of using rainfall variables as inputs to the year-round model. Thirty networks were trained over randomly-selected, 300,000-observation subsets of the development set for 15 epochs. These ANNs instantiated a model without rainfall variables as inputs.

Several additional experiments were conducted to determine the utility of changes to the input vector and the output range. A model was instantiated that included additional values and rates of change for observations 15, 30, 45, 75, 90, and 105 minutes prior to the point of prediction. Because of the additional temperature, solar radiation, wind speed, humidity, and rainfall variables, the model used an input vector with 318 values. Networks developed with these additional inputs did not provide more accurate predictions. Likewise, no improvement in accuracy was found in a model with the values and rates of change corresponding to 13, 15, 17, 19, 21, and 23 hours prior to the prediction point removed from the input vector. A model was also developed with the hyperbolic tangent function, rather than the logistic function, used for the activation function of the output node. This change, which doubled the output range of the network from [0, 1] to [-1, 1] did not improve accuracy.

Two common bootstrapping techniques from machine learning, boosting and bagging (Mitchell, 1997), were also investigated. Boosting involves improving the output of a single model by training a second model on the output of the first and combining their outputs in some fashion. Bagging involves the development of a meta-model that aggregates the outputs of several primary models. Boosting was implemented by training a network to predict the errors of the most accurate year-round ANN, as measured by development set MAE. Thirty instantiations of the four-hour boosting model were trained using randomly-selected, 400,000-observation subsets of the development set for 10 epochs, or 4 million learning events. An investigation of the effect of bagging, in which the outputs of the five most accurate existing networks were given as additional inputs to a new model, was also conducted. Thirty networks instantiating a four-hour bagging model were trained for eight epochs over the entire development set, approximately 10 million learning events per network.

RESULTS AND DISCUSSION

The first experiment compared the accuracy of year-round and seasonal models, each of which was instantiated by 30 distinct initial networks differing only in the initial random weights and the order of presentation of training patterns. The representative network for any model was the instantiation that minimized MAE over its development set. The winter model differed from other models in that its development set was restricted to the winter subset of development patterns. As expected, the winter model was less accurate than the year-round model over the year-round selection set. When evaluated over the year-round selection set, the winter model MAE was 1.518°C, compared to 1.239°C for the year-round model. Over the winter subset, the winter model MAE was 1.414°C, only slightly lower than the 1.416°C winter selection MAE for the year-round model. This result suggested that expanding the coverage of ANN models to the entire year would not sacrifice the accuracy of temperature prediction in the winter.

In the second experiment, the year-round model was compared to a similarly trained model that did not include rainfall as an input. Removing rainfall variables reduced the number of inputs from 258 to 208 and the number of weights from 31,210 to 25,201. The year-round model including rainfall as an input produced a selection data set MAE of 1.239°C. The model without rainfall inputs produced a selection data set MAE of 1.259°C. The no-rainfall model was also less accurate for the winter selection subset, generating an MAE of 1.428°C compared to 1.416°C. It is likely that the use of single-network evaluation by Jain (2003) led to the exclusion of rainfall variables as inputs in that study.

In the first experiment, ANNs were trained over a randomly-chosen subset of 300,000 development patterns of the more than 1.25 million patterns available. Such a restriction was necessary for an accurate comparison of year-round models with winter models. Having shown that the year-round model was a suitable replacement for the winter model, training was performed with no restriction on the number of training observations. The accuracy improvements of the year-round model were consolidated by training networks over the 1.25 million available development patterns for eight epochs.

For the four-hour prediction horizon, 30 year-round networks were, therefore, trained using the entire development set for approximately 10 million learning events. The selection data set MAE of the representative network, referred to hereafter as the standard network, was 1.226°C, an improvement of 0.013°C over the results of the first experiment.

In the first of two bootstrapping experiments conducted for the four-hour horizon, the standard network was *boosted*. A group of 30 networks were trained to predict the errors of the standard network. The boosting model made use of 259 inputs: the original 258 as well as the standard network's output. By combining the boosting network's output with that of the standard network, accuracy would be improved to the extent that the boosting model was successful in predicting the standard network's prediction errors. Following 10 epochs of training over a randomly-selected subset of 400,000 development patterns (4 million learning events), the selection MAE decreased by less than 0.004°C relative to the standard network. Though errors did decrease slightly, the reduction in error was negligible while the training was time-consuming. As such, boosting was not deemed to be a useful method of improving prediction accuracy for this problem.

An experiment to explore the second technique, *bagging*, was implemented by training a four-hour prediction model with 263 inputs: the 258 provided to the year-round model as well as the output of the five most accurate year-round networks from the consolidation phase, as measured by development MAE. After training was complete, the final networks had undergone more than 10 million learning events. The 1.220°C selection MAE of the representative network for the bagging model improved upon the four-hour standard network by approximately 0.006°C. These improvements did not merit the additional complexity and computational time required to implement the bagging model across all horizons.

Final models were then developed for the other prediction horizons from one to 12 hours ahead. These models made use of 258 inputs, including air temperature, wind speed, relative humidity, solar radiation, and rainfall and hourly rates of change at the time of prediction as well as the history of prior observations at one-hour intervals going back 24 hours. Also among the models' inputs were four cyclic time-of-day and four cyclic day-of-year terms. For each model, 30 networks were trained over all available development observations for eight epochs. Due to the computational time required to train and evaluate so many networks over such a large number of patterns, poorly performing networks were discarded at two stages of the training process. Following two epochs of training (approximately 2.5 million learning events), each network was evaluated over a 100,000-pattern subset of development patterns and the 15 networks with the highest MAEs were discarded. Following another two epochs the process was repeated and five of the remaining 15 networks were discarded. The remaining networks were trained for an additional four epochs, for a total of more than 10 million learning events each.

The MAEs of the final models at all prediction horizons for both the selection and evaluation sets are presented in Table 3.1. The MAE of the selection data set increased monotonically from 0.525°C at the one-hour horizon to 1.908°C at the 12-hour horizon. The trend was also apparent for the MAEs of the evaluation data set calculated for the same sites during 2004 and 2005. The MAEs associated with the models during this period were slightly lower than those for the selection set, from 0.516°C for the one-hour model to 1.873°C for the 12-hour model. The four-hour horizon used as the basis for experimentation in this research resulted in an MAE of 1.187°C for the evaluation data set.

The final year-round models were also evaluated over the winter selection and winter evaluation subsets (Table 3.2), allowing for a direct comparison with the winter models developed by Smith et al. (2006). When compared to these winter models, the year-round ANN errors were less than or equal to winter selection and winter evaluation MAEs at all horizons. Winter MAEs proved to be higher than those for the entire year across all horizons, increasing along with horizon length. The winter selection and winter evaluation MAEs for the one-hour model were 0.529°C and 0.522°C, respectively. At the 12-hour horizon they were 2.495°C and 2.299°C. These results indicate that the year-round air temperature prediction models, developed to be included in general decision support aids, would be suitable for use during the winter months.

The plots in Figure 3.2 present the changes in prediction error for four prediction horizons. Differences between predicted and observed temperatures at the one-hour horizon remain small, increasing along with horizon length. The higher MAEs are reflected in the greater dispersion about the 1:1 line of a hypothetical, perfect model. Likewise, as horizon length increases, the value for R^2, denoting accuracy of a simple linear regression of predicted on observed temperatures, decreases. The regression analysis also suggested that models for longer horizons have a tendency to over-predict low-temperature observations and under-predict high temperature observations.

The accuracy of the final model was also evaluated in relation to the day of year and time of day at three different prediction horizons: four, eight, and 12 hours. Figure

3.3 presents a contour plot of MAE for the combined selection and evaluation sets partitioned by week of year and hour of day. Both data sets were combined to reduce the impact of weather trends from any single year. Each point corresponds to the mean of all absolute errors associated with a particular week of the year and hour of the day. For example, the MAE during the third week at 1200 hours was calculated over all predictions made during January 15-21 at 1200, 1215, 1230, and 1245 hours for the years 2001-2005 at each of the 13 sites in the combined data set. At the four-hour horizon (Figure 3.3a), the final model was most accurate when predicting overnight/early-morning temperatures, especially during the summer. Most of the MAEs for summer nights and early mornings were less than 1°C. Prediction MAEs were highest for the four-hour model between 0900 and 1445 hours from fall through early spring and between 1400 and 1945 hours during the summer.

The magnitudes of the prediction errors during summer nights and mornings were also smaller, relative to the rest of the day, for the eight and 12-hour horizons shown in Figure 3.3b and 3.3c. During the summer, the eight and 12-hour horizons also showed the highest MAEs of the day during the afternoons. Behavior of the final eight- and 12-hour model differed substantially from the four-hour horizon during the rest of the year. Mornings during fall and winter had noticeably larger errors for the eight-hour horizon than for the four-hour horizon. Errors for the 12-hour horizon were of even greater magnitude. Periods of higher-than-average MAE for the four-hour horizon occurred in the morning and early-afternoon outside of the summer. At the eight- and 12-hour horizons, these periods of lower accuracy persisted until early evening.

The bias of the final models as measured by mean error partitioned by week and hour is presented in Figure 3.4. A negative mean error indicates a tendency to under-predict, while a positive mean error is evidence of over-prediction. Figure 3.4a shows that the two areas of greatest MAEs identified at the four-hour horizon, fall/winter midday and summer afternoon, were associated with at least two distinct phenomena. During fall and winter at midday, the largest errors showed a negative bias, suggesting that the unanticipated weather events were warming events. During the summer afternoons, bias was positive, suggesting the presence of unpredicted cooling events. For the eight and 12-hour horizons (Figures 3.4b and 3.4c) the summer afternoons were also associated with a positive bias and periods of maximum bias were closely associated with high MAEs. Biases associated with winter errors for each of these horizons were also largely positive.

The unanticipated cooling events that occurred during the summer afternoons across all horizons are likely associated with rain showers and thunderstorms. Figure 3.5 presents mean rainfall partitioned by week and hour for the same sites and periods. Summer afternoons and evenings clearly accounted for the most active periods of rainfall. Moreover, these times corresponded to the weeks and hours associated with relatively higher MAEs and biases in the four-, eight-, and 12-hour horizons. This strongly suggests that the unanticipated summer cooling events were associated with rainfall or the corresponding cloud cover.

The air temperature predictions provided on the AEMN website (www.georgiaweather.net) are presented to users as a sequence of 12 temperatures generated at the same time, each corresponding to a different horizon. Two such sequences are presented alongside observed air temperature in Figure 3.6 for March 1 and

2, 2005 at Dearing. A sustained freeze began at 2300 hours on March 1 and lasted until 0830 hours the following day. A minimum temperature of -2.564°C occurred at 0700 on March 2. The first of the two prediction sequences was generated at 1800 hours on March 1, five hours before the first freezing temperature was recorded. This sequence accurately predicted the time of onset and had absolute errors less than 1°C for the first ten hours of prediction. Such a sequence of predictions would provide time for fruit growers who use the AEMN website to obtain temperature predictions for their local area and take steps to mitigate crop damage. The second prediction sequence was generated at midnight, one hour after the beginning of the freeze, and was remarkably consistent with the first. Both sequences under-predicted temperatures during the early morning hours. The prediction sequence generated at midnight continued to under-predict observed air temperature by 1°C to 2°C from 0500 to 1200 hours on March 2.

The observed air temperatures and prediction sequences for July 28 and 29, 2005 at Homerville are displayed in Figure 3.7. During this period, high temperatures caused heat exhaustion in several farmers and degraded pesticides in the south-central Georgia Agricultural Statistical District that includes Homerville (NASS, 2005). On July 28th in Homerville, temperatures climbed above 35°C with two periods of cooling during the afternoon, the first between 1445 and 1545 hours and a second, more rapid cooling event beginning at 1645 hours. Each of these cooling periods was associated with a brief decrease in solar radiation not typical of clear-sky conditions. The cause of this reduction in solar radiation was presumably cloud cover. A brief rainfall of more than 5mm occurred between 2130 and 2200 hours on the night of the 28th. The first prediction sequence, generated at 0600 hours on the 28th, accurately predicted observed temperatures within ±1.1°C for all but the final prediction at 1800 hours. The second sequence was generated at noon and also closely predicted observed temperatures. Prediction errors were within ±0.5°C for the five temperature predictions prior to 1800 hours. Following the second cooling event, the prediction error at 1800 hours was more that 3.4°C, with subsequent, late-night observed air temperatures more closely reflecting the 1200 prediction sequence.

SUMMARY AND CONCLUSIONS

Year-round air temperature prediction models were developed for prediction horizons of one to 12 hours using Ward-style ANNs. These models were intended for use in general decision support and are currently implemented on the AEMN website, www.georgiaweather.net. Suitable ANN design modifications made it possible to meet or exceed the accuracy of previously-developed, winter-specific models during the winter period. It was shown that models that included rainfall terms in the input vector were more accurate than those that did not. Applying the bootstrapping techniques of boosting and bagging to single-network models was not found to be useful for this problem domain, as the very modest improvements in prediction accuracies came with a heavy computational cost. The accuracy of the final four, eight, and 12-hour models was analyzed, showing that unanticipated cooling events were the most significant obstacle faced, especially at longer horizons.

The results suggest that accurate cloud cover predictions might aid in the prediction of associated cooling events, especially during the summer. Further study might also determine if models specifically tailored to the periods of greater-than-average prediction errors might be useful in an ensemble approach. Because of the focus on developing models applicable to a broad range of locations, temperature prediction work in the AEMN domain has not made use of geographic information. Future work could focus on the possibility of adding the information in such a manner as to preserve the general applicability of the model.

ACKNOWLEDGEMENTS

This work was funded in part by a partnership between the USDA-Federal Crop Insurance Corporation through the Risk Management Agency and the University of Georgia and by state and federal funds allocated to Georgia Agricultural Experiment Stations Hatch projects GEO00877 and GEO01654.

The authors would also like to thank the other members of the AEMN research team, Daniel Shank and Kevin Crowell, for their input and suggestions.

REFERENCES

Bruton, J.M., R.W. McClendon, and G. Hoogenboom, 2000. Estimating daily pan evaporation with artificial neural networks. Transactions of ASAE 43(2): 491-496.

Elizondo, D., G. Hoogenboom, and R.W. McClendon, 1994. Development of a neural network to predict daily solar radiation. Agriculture and Forest Meteorology 71: 115-132.

Haykin, S., 1999. Neural networks: a comprehensive foundation, 2nd edition. Upper Saddle River, NJ: Prentice Hall.

Hoogenboom, G., 2000. The Georgia automated environmental monitoring network. Preprints of the 24th Conference On Agricultural and Forest Meteorology: 24-25.

Jain, A., 2003. Frost prediction using artificial neural networks: A temperature prediction approach. M.S. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA.

Jain, A., R.W. McClendon, G. Hoogenboom, and R. Ramyaa, 2003. Prediction of frost for fruit protection using artificial neural networks. American Society of Agricultural Engineers, St. Joseph, MI, ASAE Paper 03-3075.

Jiang, D., Y. Zhang, X. Hu, Y. Zeng, J. Tan, and D. Shao, 2004. Progress in developing an ANN model for air pollution index forecast, Atmospheric Environment 38(40): 7055-7064.

Li, B., R.W. McClendon, and G. Hoogenboom, 2004. Spatial Interpolation of Weather Variables for Single Locations Using Artificial Neural Networks, Transactions of the ASAE 47(2): 629-637.

Maqsood, I., M.R. Khan, and A. Abraham, 2004. An ensemble of neural networks for weather forecasting. Neural Computing and Applications 13: 112-122.

National Agricultural Statistical Service, 2005. Crop Progress and Condition Report. July 31, 2005.

Patel, V.C., R.W. McClendon, and J.W. Goodrum, 1998. Color computer vision and artificial neural networks for the detection of defects in poultry eggs. Artificial Intelligence Review 12(1):163-176.

Pydipati, R., T.F. Burks, and W.S. Lee, 2005. Statistical and neural network classifiers for citrus disease detection using machine vision. Transactions of the ASAE 48(5): 2007-2014.

Ramyaa, 2004. Frost prediction using artificial neural networks: A classification approach. M.S. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA.

Ruano, A.E., E.M. Crispim, E.Z.E. Conceição, and M.M.J.R. Lúcio, 2006. Prediction of building's temperature using neural networks models. Energy and Buildings 38: 682-694.

Shank, D., 2006. Dew point temperature prediction using artificial neural networks. M.S. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA.

Smith, B.A., R.W. McClendon, and G. Hoogenboom, 2006. Improving air temperature prediction with artificial neural networks. International Journal of Computational Intelligence 3(3):179-186.

Ward Systems Group, 1993. Manual of NeuroShell 2, Frederick, MD.

Yi, J. and R. Prybutok, 1996. A neural network model forcasting for prediction of daily maximum ozone concentration in an industrialized urban area. Environmental Pollution 92(3): 349-357.

Table 3.1
Final year-round model prediction accuracies over the selection and evaluation datasets

Horizon length (hours)	Mean absolute error (°C)	
	Selection data set, 2001-3	Evaluation data set, 2004-5
1	0.525	0.516
2	0.834	0.814
3	1.046	1.015
4	1.226	1.187
5	1.404	1.356
6	1.483	1.432
7	1.577	1.532
8	1.669	1.623
9	1.734	1.686
10	1.801	1.755
11	1.865	1.815
12	1.908	1.873

Table 3.2
Comparison of model prediction accuracies over the winter selection and evaluation subset

Horizon length, (hours)	Winter Selection MAE, 2001-3 (°C)		Winter Evaluation MAE, 2004-5 (°C)	
	Winter model*	Year-round model	Winter model*	Year-round model
1	0.534	0.529	0.527	0.522
2	0.884	0.883	0.864	0.860
3	1.167	1.167	1.118	1.117
4	1.401	1.398	1.338	1.331
5	1.624	1.615	1.546	1.535
6	1.811	1.793	1.715	1.695
7	1.987	1.930	1.874	1.831
8	2.126	2.081	2.007	1.958
9	2.243	2.213	2.091	2.055
10	2.362	2.311	2.191	2.161
11	2.443	2.417	2.250	2.239
12	2.526	2.495	2.333	2.299

* Smith et al. (2006)

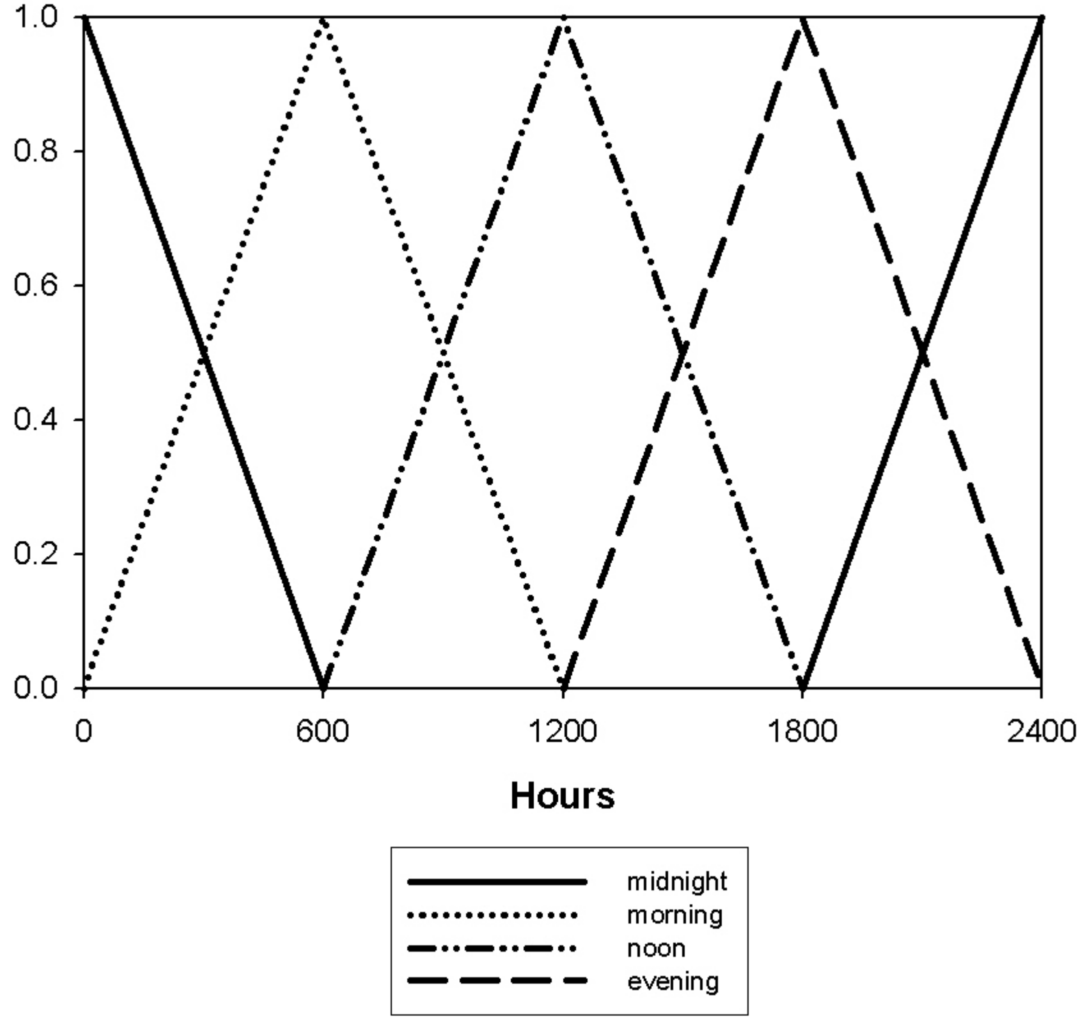

Figure 3.1: Triangular fuzzy logic membership functions for time-of-day that are used as inputs to the models

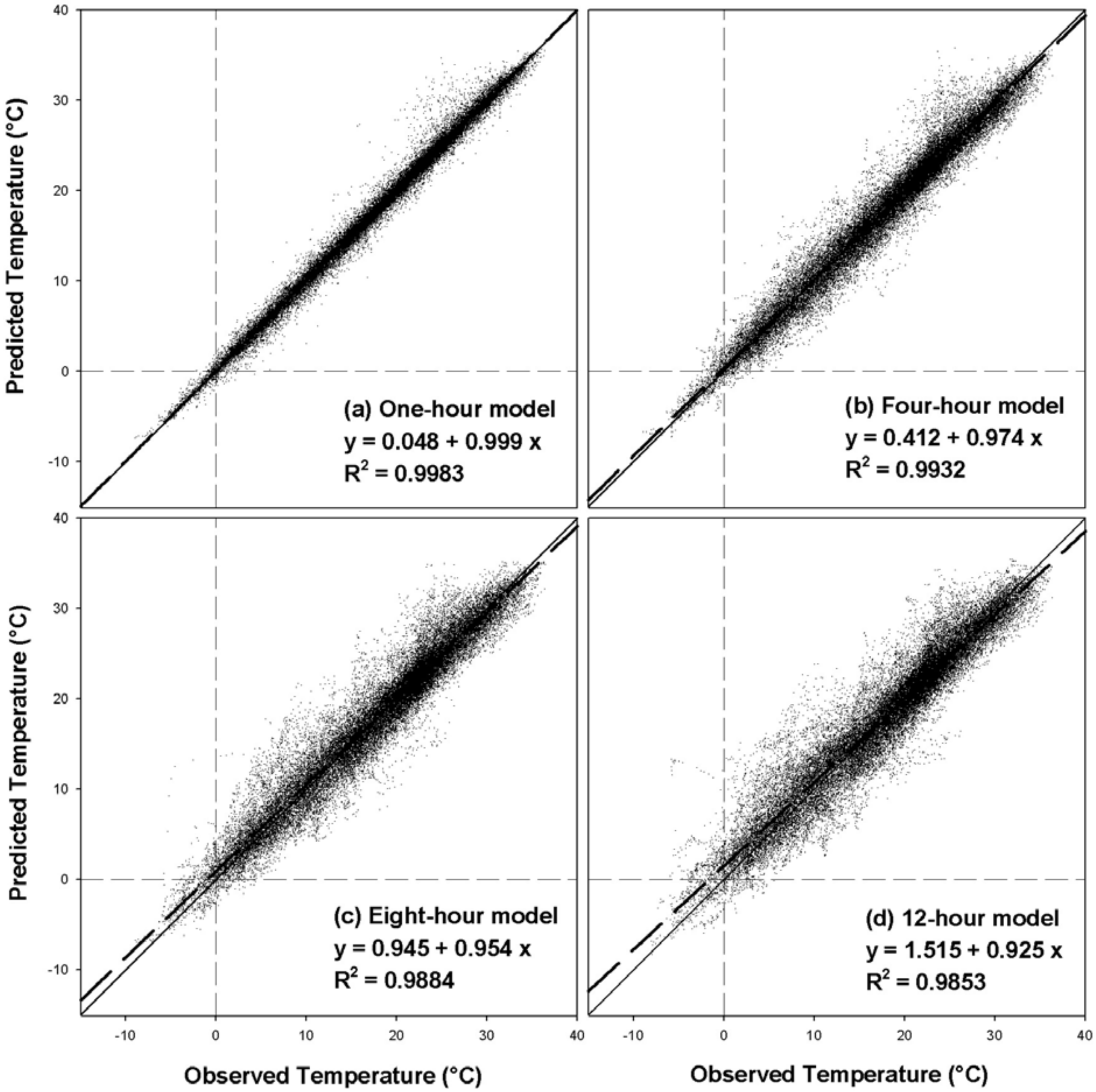

Figure 3.2: Predicted and observed temperatures for the Byron site during 2004-2005 for the final model at the (a) one-hour horizon, (b) four-hour horizon, (8) eight-hour horizon, and (d) 12-hour horizon. A solid diagonal line indicates a hypothetical perfect model.

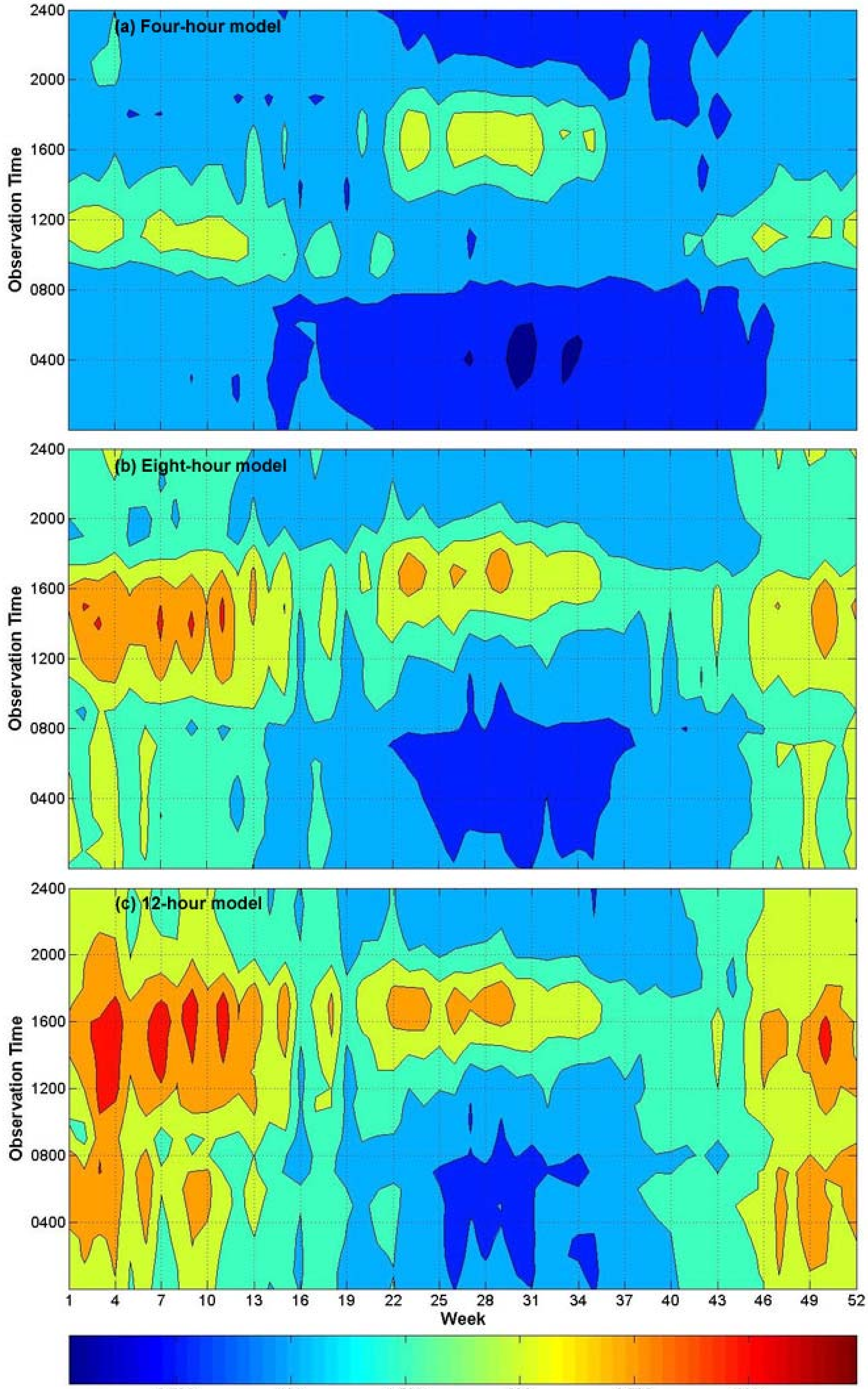

Figure 3.3: Prediction MAE across all evaluation sites during 2001-2005. Errors partitioned by week and hour.

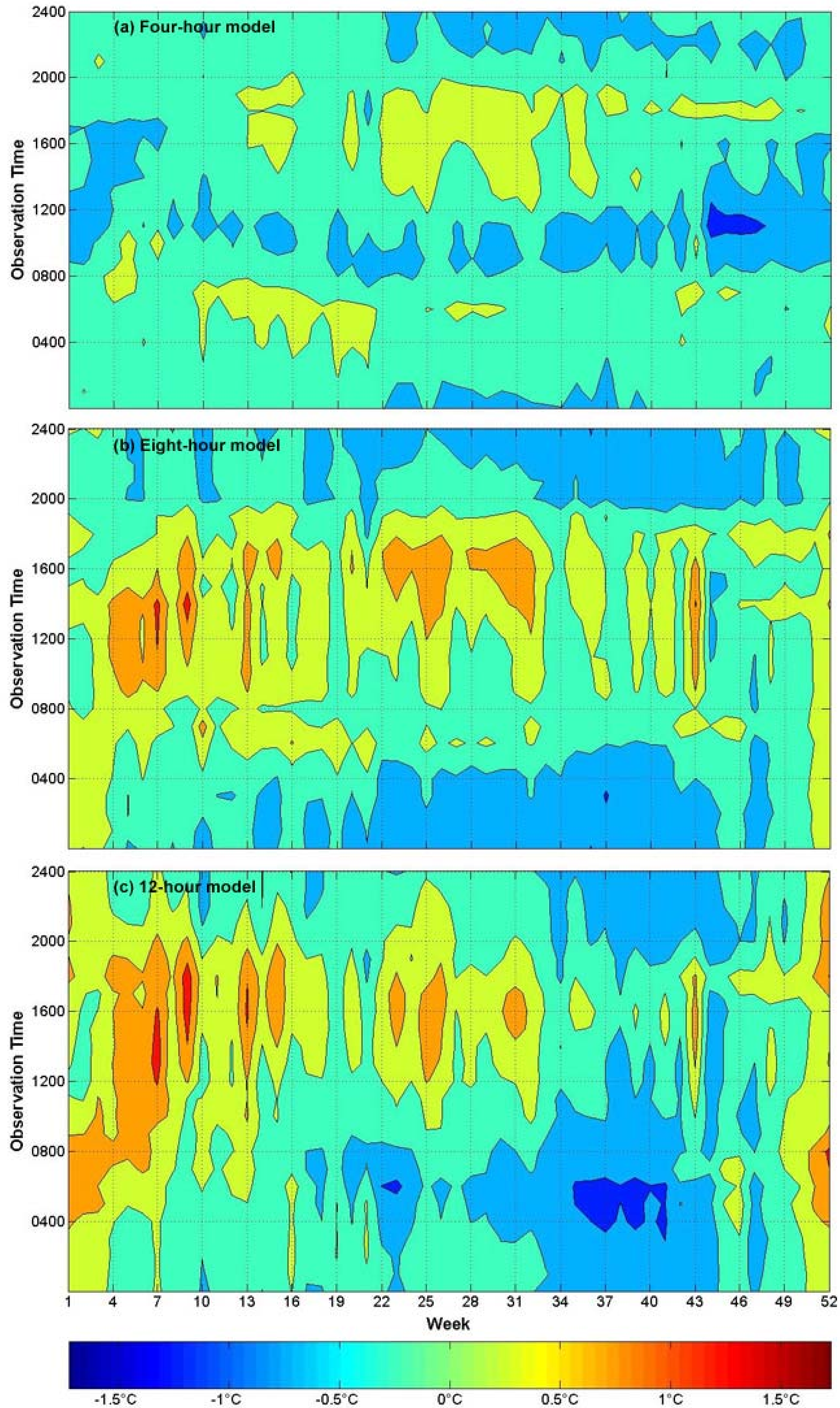

Figure 3.4: Prediction mean errors across all evaluation sites during 2001-2005. Errors partitioned by week and hour.

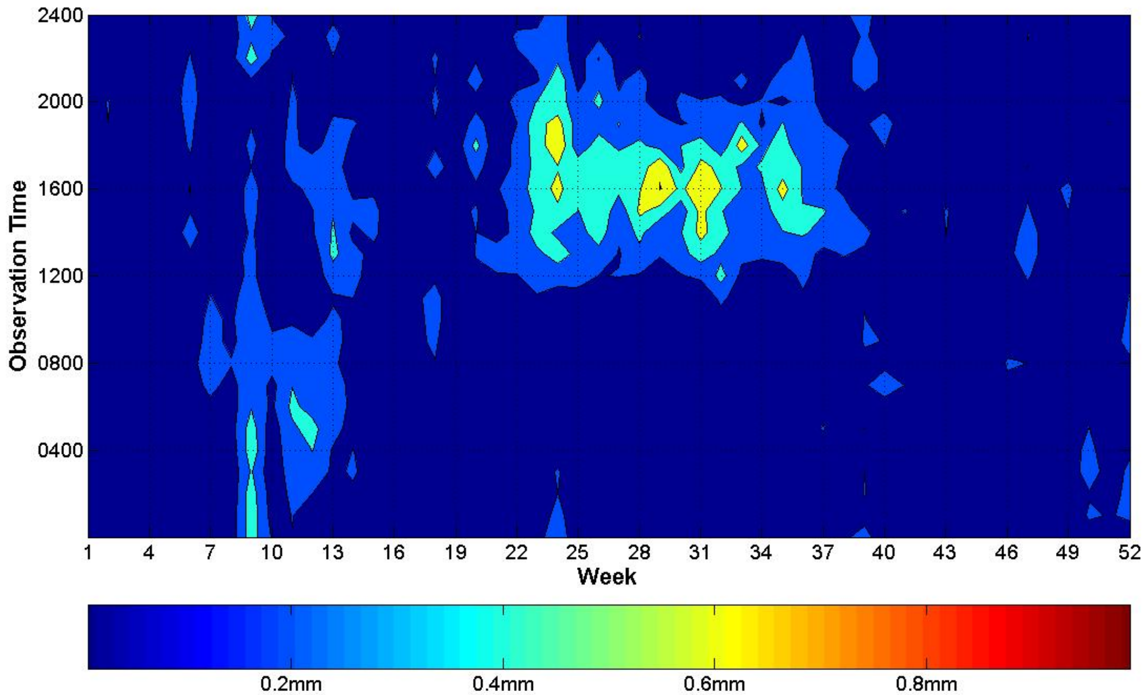

Figure 3.5: Mean rainfall across all evaluation sites during 2001-2005 partitioned by week and hour

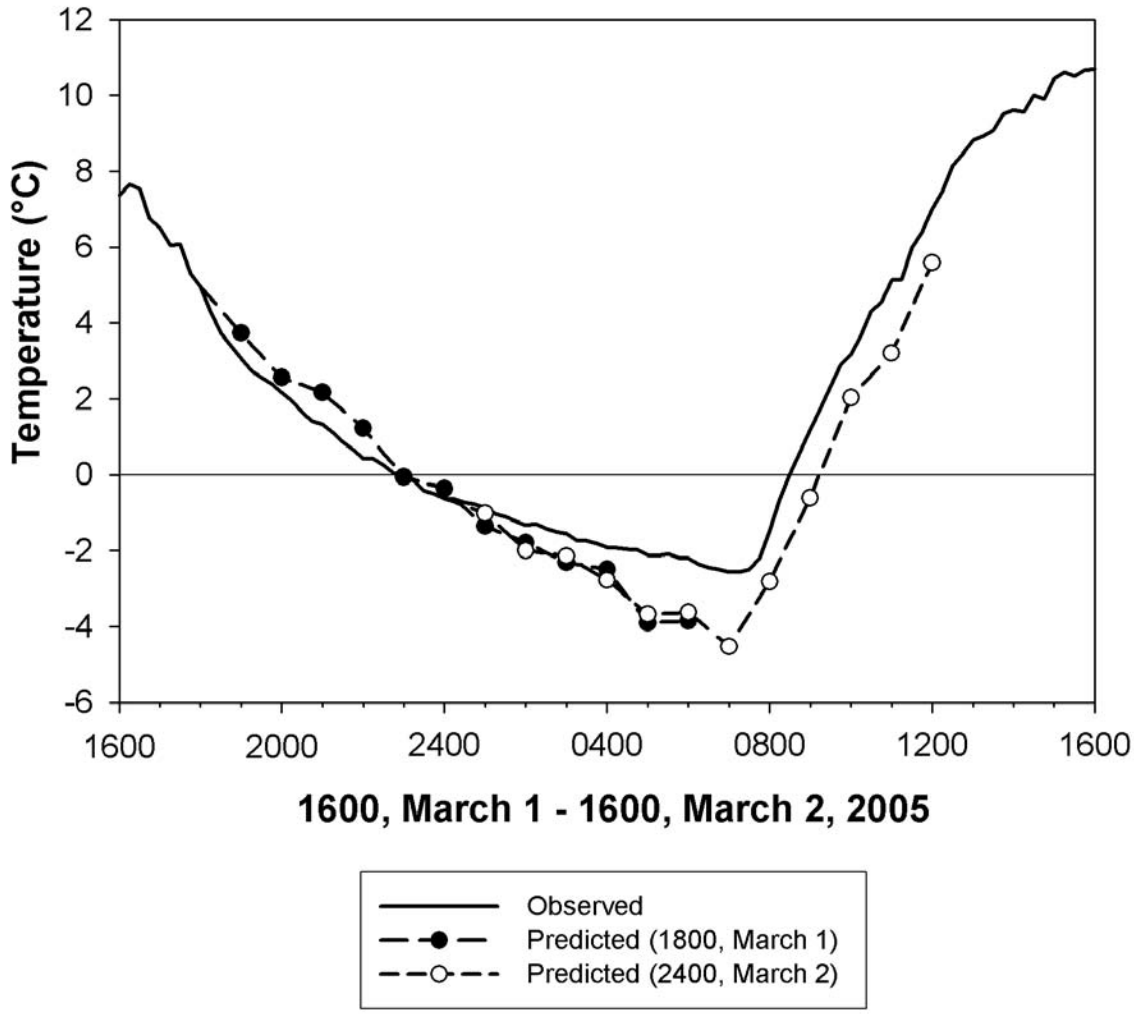

Figure 3.6: Observed and 12-hour prediction sequences of temperature during March 1-2, 2005, for Dearing, Georgia

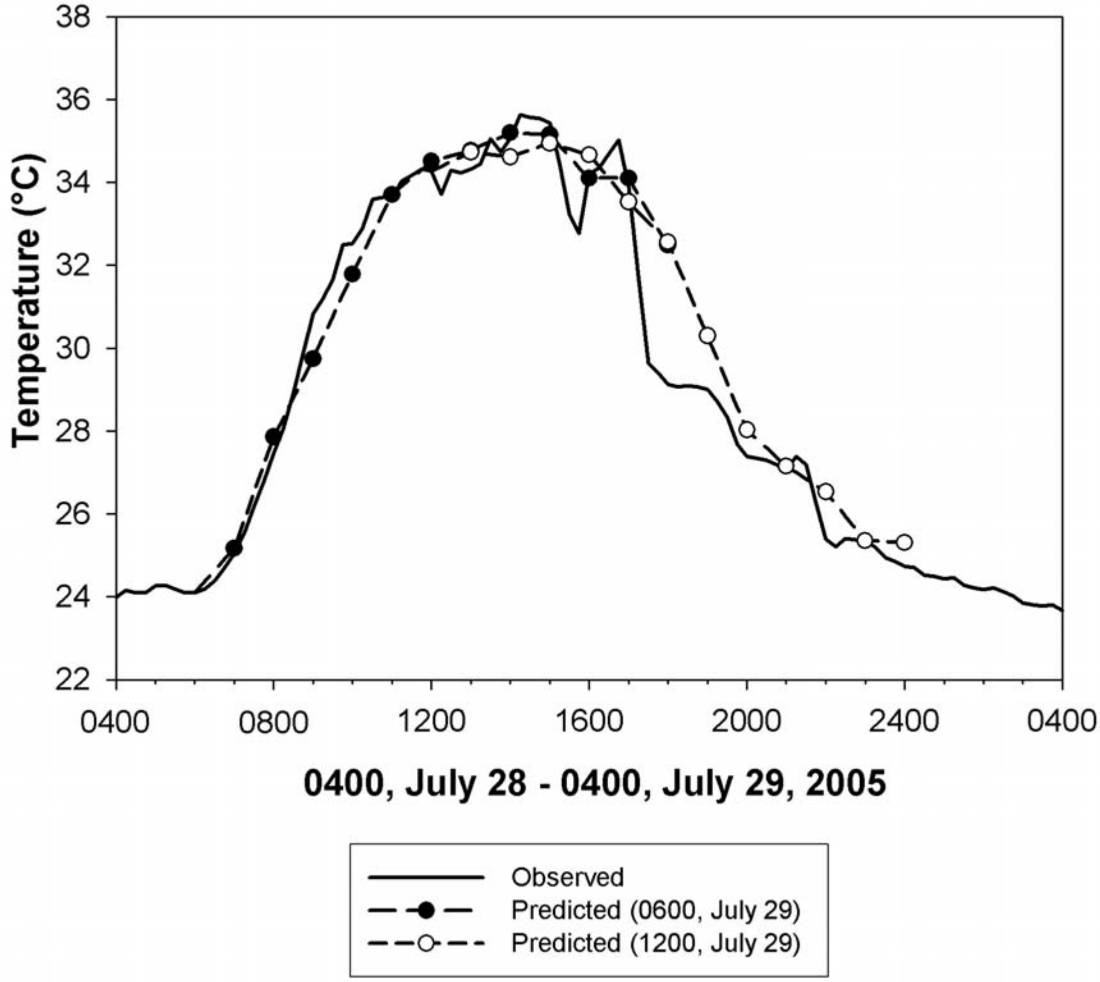

Figure 3.7: Observed and 12-hour prediction sequences of temperature during July 28-29, 2005, for Homerville, Georgia

CHAPTER 4

SUMMARY AND CONCLUSION

This thesis was focused on the development of ANN models for air temperature prediction over the AEMN domain. Though primarily oriented towards agricultural users, the predictions of the final, year-round models may be useful to a broad range of users. The research in chapter 2 built upon the work of Jain et al. (2003) and Jain (2003). It explored ANN model improvements including larger training set sizes, the inclusion of seasonal input terms, an increased duration of prior observations as inputs, and varying the size of the hidden layer. Slight reductions in prediction errors were obtained by increasing the size of the training set. The inclusion of seasonal variables corresponding to membership in the fuzzy sets winter, spring, summer, and fall improved model accuracy. Extending the duration of historical data in the input vector from six to 24 hours led to similar improvements. Hidden layers with 120 nodes, or 40 nodes per slab, were more found to be more accurate than other models. Finally, when applied to data-rich environments, a clear distinction should be maintained between abstract neural network models and actual instantiations of these models. The performance of a single instantiated network is not likely to be a valid measure of model performance. Model evaluation over multiple instantiations led to better parameter selection by presenting more accurate comparisons of distinct models than those afforded by single-network evaluation.

In chapter 3, year-round air temperature prediction models were developed using Ward-style ANNs for prediction horizons of one to 12 hours. These models were intended for use in general decision support and are currently implemented on the AEMN website, www.georgiaweather.net. These year-round networks met or exceed the accuracy of previously-developed, winter-specific models during the winter period. The inclusion of rainfall terms in the input vector led to more accurate ANN models without increasing computational cost, while the application of bootstrapping techniques to single-network models did not. The most significant obstacle faced, especially at longer horizons, were unanticipated cooling events.

While the simple ensemble approach examined in chapter 3 did not lead to useful reductions in model error, networks that have been specifically trained to predict air temperature during periods of greater-than-average error might be successfully combined in an ensemble approach. The results of chapter 3 also suggest that a cloud cover predictor could increase the accuracy of air temperature predictors if used as an input to such models. Since this thesis developed general prediction models that were not specific to a particular site, neither geographic information nor observations from multiple sites were used. Future work could focus on the possibility of adding the information in such a manner as to preserve the general applicability of the developed models.

REFERENCES

Bruton, J.M., R.W. McClendon, and G. Hoogenboom, 2000. Estimating daily pan evaporation with artificial neural networks. Transactions of ASAE 43(2): 491-496.

Elizondo, D., G. Hoogenboom, and R.W. McClendon, 1994. Development of a neural network to predict daily solar radiation. Agriculture and Forest Meteorology 71: 115-132.

Haykin, S., 1999. Neural networks: a comprehensive foundation, 2nd edition. Upper Saddle River, NJ: Prentice Hall.

Hoogenboom, G., 2000. The Georgia automated environmental monitoring network. Preprints of the 24th Conference On Agricultural and Forest Meteorology: 24-25.

Jain, A., 2003. Frost prediction using artificial neural networks: A temperature prediction approach. M.S. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA.

Jain, A., R.W. McClendon, G. Hoogenboom, and R. Ramyaa, 2003. Prediction of frost for fruit protection using artificial neural networks. American Society of Agricultural Engineers, St. Joseph, MI, ASAE Paper 03-3075.

Jiang, D., Y. Zhang, X. Hu, Y. Zeng, J. Tan, and D. Shao, 2004. Progress in developing an ANN model for air pollution index forecast, Atmospheric Environment 38(40): 7055-7064.

Li, B., R.W. McClendon, and G. Hoogenboom, 2004. Spatial Interpolation of Weather Variables for Single Locations Using Artificial Neural Networks, Transactions of the ASAE 47(2): 629-637.

Maqsood, I., M.R. Khan, and A. Abraham, 2004. An ensemble of neural networks for weather forecasting. Neural Computing and Applications 13: 112-122.

National Agricultural Statistical Service, 2005. Crop Progress and Condition Report. July 31, 2005.

Okie, W. R., G. L. Reighard, W. C. Newall, Jr., C. J. Graham, D. J. Werner, A. Powell, G. Krewer, and T. G. Beckman, 1998. Spring freeze damage to the 1996 peach and nectarine crop in the southeastern United States. HortTechnology, vol. 8, pp. 381-386.

Patel, V.C., R.W. McClendon, and J.W. Goodrum, 1998. Color computer vision and artificial neural networks for the detection of defects in poultry eggs. Artificial Intelligence Review 12(1):163-176.

Pydipati, R., T.F. Burks, and W.S. Lee, 2005. Statistical and neural network classifiers for citrus disease detection using machine vision. Transactions of the ASAE 48(5): 2007-2014.

Ramyaa, 2004. Frost prediction using artificial neural networks: A classification approach. M.S. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA.

Ruano, A.E., E.M. Crispim, E.Z.E. Conceição, and M.M.J.R. Lúcio, 2006. Prediction of building's temperature using neural networks models. Energy and Buildings 38: 682-694.

Salehi, F., R. Lacroix, and K. M. Wade, 1998. Effects of learning parameters and data presentation on the performance of backpropagation networks for milk yield prediction. Transactions of the ASAE, 41: 253-259.

Shank, D., 2006. Dew point temperature prediction using artificial neural networks. M.S. thesis, Artificial Intelligence Center, University of Georgia, Athens, GA.

Smith, B.A., R.W. McClendon, and G. Hoogenboom, 2006. Improving air temperature prediction with artificial neural networks. International Journal of Computational Intelligence 3(3):179-186.

Temeyer, B. R., W. A. Gallus, Jr., K. A. Jungbluth, D. Burkheimer, and D. McCauley, 2003. Using an artificial neural network to predict parameters for frost deposition on Iowa bridgeways.Proceedings of the 2003 Mid-Continent Transportation Research Symposium, Iowa State University, Ames, IA.

Ward Systems Group, 1993. Manual of NeuroShell 2, Frederick, MD.

Yi, J. and R. Prybutok, 1996. A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. Environmental Pollution 92(3): 349-357.