PRECIPITATION PREDICTION USING ARTIFICIAL NEURAL NETWORKS

by

KEVIN L. CROWELL

(Under the direction of Gerrit Hoogenboom)

ABSTRACT

Precipitation, in meteorology, is defined as any product, liquid or solid, of atmospheric water vapor that is accumulated onto the earth's surface. Water, and thus precipitation, has a major impact on our daily livelihood. As such, the uncertainty of both the future occurrence and amount of precipitation can have a negative impact on many sectors of our economy, especially agriculture. There is, therefore, a need to use innovative computer technologies such as artificial intelligence to improve the accuracy of precipitation predictions. Artificial neural networks have been shown to be useful as an aid for the prediction of weather variables. The goal of this study was to develop artificial neural network models for the purpose of predicting both the Probability of Precipitation and quantitative precipitation over a 24-hour period beginning and ending at midnight.

INDEX WORDS: Artificial Neural Networks; Probabilistic Neural Network; Precipitation; Probability of Precipitation; Quantitative Precipitation; Brier Score

PRECIPITATION PREDICTION USING ARTIFICIAL NEURAL NETWORKS

By

KEVIN L. CROWELL

B.A., The University of Georgia, 2005

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

2008

PRECIPITATION PREDICTION USING ARTIFICIAL NEURAL NETWORKS

by

KEVIN L. CROWELL

Major Professor: Gerrit Hoogenboom

Committee: Ron W. McClendon
Walter D. Potter
Joel O. Paz

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2008

DEDICATION

I dedicate this thesis to my parents, John and Sherry Crowell, who have always pushed me strive farther than I could imagine.

ACKNOWLEDGEMENTS

I would like to thank Dr. Potter for originally getting me started in the Artificial Intelligence program and for finding the opportunity for me to work on the research conducted herein. I would also like to thank my committee members who were very involved in my research process: Dr. McClendon for his continuing guidance even after his retirement, Dr. Hoogenboom for his vast domain knowledge he contributed to this research, and Dr. Paz who initially peaked my interest in precipitation prediction.

In addition to my committee, I would like to thank everyone who has contributed to this research in any way. Specifically, Daniel Shank, Bob Chevalier, and Max Martin were of great help with their comments during our regular group meetings. Brian Smith was especially of help with his experienced knowledge and for providing the initial java code used for the Artificial Neural Network model development conducted throughout this research.

This work was funded in part by a partnership between the USDA-Federal Crop Insurance Corporation through the Risk Management Agency (RMA) and the University of Georgia, by a grant from the Southeastern Peanut Research Initiative (SPRI) 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

Precipitation, in meteorology, is defined as any product, liquid or solid, of atmospheric water vapor that is accumulated onto the earth's surface. Precipitation can take on a variety of forms, including, but not excluded to rain, snow, and hail. A sudden large accumulation of precipitation or a long period of precipitation can cause flooding, especially in areas near large bodies of waters such as rivers or oceans. This flooding can be detrimental to crops, livestock, sewer systems, roadways, bridges, homes, and any type of structure in the flood area. Even a moderate amount of precipitation that is unexpected can be damaging the crops if the proper preparations were not taken. It is important to have accurate models for predicting both the Probability of Precipitation (PoP) and quantitative precipitation for any area.

The PoP is an estimation of a measurable amount of rain or snow falling anywhere in the predicted region during any of the prediction time period. Quantitative Precipitation Prediction (QPP) is an estimation of the amount of liquid precipitation to be accumulated at any point in the predicted region during the prediction time period. Currently, the National Weather Service (NWS) has a model package developed for PoP prediction (Carrol, 2004). This model package was developed by the NWS Meteorological Development Laboratory and is an Eta-based Model Output Statistics (MOS) (Glahn, 1972) guidance package which runs twice daily at midnight and noon UTC. At each run, the package generates both PoP predictions and QPPs covering every 6, 12, and 24 hour period between the time of the model package execution and 60 hours beyond the time of execution. The MOS technique is used with multiple linear regression to develop predictive equations. Although the NWS does not provide detailed measures of the accuracies of the model package, improving on the models would be beneficial to a wide array of groups.

It has been shown that artificial neural network models are able to accurately predict weather variables, including precipitation. Hall et al. (1999) used ANNs to predict PoP for a 24 hour period. Their ANN model had a Brier score (Brier, 1950) of 0.25 and a skill score (Wilks, 1995) of 0.73 using 34.2% as the precipitation occurrence of the sample climatology. Kuligowski and Barros (1998) had mixed results when comparing the use of ANNs to linear regression for quantitative precipitation forecasts on a 24-hour prediction period. In comparison to linear regression, they found that the ANN model had a lower threat score (Schaefer, 1990) when the observed precipitation amount was below 15 mm and a higher threat score when the observed amount was above 15 mm. Using spatial and temporal data of recent rainfall, Luk et al. (2000) developed an ANN model for predicting rainfall amounts for 15 minutes ahead for several areas of western Sydney, Australia.

The goal of the research described in this thesis is the development of ANN models for the purpose of predicting both the Probability of Precipitation and quantitative precipitation over a 24-hour period beginning and ending at midnight. The prediction is to be generated at 6:00 PM, six hours prior to the start of the prediction period. To accomplish goal, weather data for the development of the ANN models were obtained from the University of Georgia's Automated Environmental Monitoring Network (AEMN) (Hoogenboom, 2000). The AEMN was established in 1991 and currently consists of over 75 automated weather stations throughout the state of Georgia. Each AEMN station collects weather data such as temperature, humidity, dew point, wind speed, wind direction, and precipitation. The weather data and associated predictors can be found at www.georgiaweather.net.

Chapter 2 describes the development of Artificial Neural Network (ANN) models for the purpose of PoP prediction. The description provided in Chapter 2 includes introducing the problem of predicting PoP, describing the methodology involved in the development of the ANN models, presenting the experiment results, and stating research conclusions and possible future work. Chapter 3 describes the development of ANN models for the purpose of quantitative precipitation prediction. The description provided in Chapter 3 also includes introducing quantitative precipitation predictions, outlining the methodology used for ANN model development, and presenting the results produced, the drawn conclusions, and possible future work. Chapter 4 provides a summary and conclusion to the topics discussed in the thesis.

CHAPTER 2

PROBABILITY OF PRECIPITATION PREDICTION USING ARTIFICIAL NEURAL NETWORKS[1]

[1] Crowell, K.L., McClendon, R.W., Paz, J.O., and Hoogenboom, G. To be submitted to Weather and Forecasting.

ABSTRACT

The benefits of improving Probability of Precipitation (PoP) predictions include minimizing losses and maximizing gains for a variety of sectors that include agriculture, forestry, and natural resources management. The goal of this study was to develop artificial neural network (ANN) models for predicting the PoP for a 24-hour period beginning and ending at midnight. The prediction is to be generated at 6:00 PM, six hours prior to the start of the prediction period. The specific objectives included determining the preferred ANN parameters, selection of important weather related inputs, and the comparison of year-round vs. seasonal models. Prediction errors were minimized when seasonal models were used for predicting the fall, winter, and spring seasons. However, a year-round model produced a smaller error than the seasonal model when predicting the summer season. An iterative search for weather input variables during the development of each model showed that the minimum Brier score occurred when using two weather input variables. The preferred variables for winter, spring, summer, fall, and year-round were barometric pressure and temperature, humidity and barometric pressure, barometric pressure and precipitation amount, humidity and barometric pressure, and humidity and barometric pressure respectively. It was also found that barometric pressure was selected for inclusion in the development of every model. The evaluation of the final models developed showed that the Brier scores for winter, spring, summer and fall were 0.1778, 0.1866, 0.2113, and 0.1544, respectively and the Brier skill scores were 0.3456, 0.2455, 0.2950, and 0.3382, respectively. All of the models included a large number of predictions below a 30% PoP and very few predictions above a 70% PoP. It is important for the models to have a larger number of predictions below 30% PoP and above 70% PoP to have a large degree of certainty.

INTRODUCTION

The Probability of Precipitation (PoP) is an estimation of a measurable amount of rain or snow falling anywhere within a specific region during any of the prediction time period. Currently, the National Weather Service (NWS) has a model package developed for PoP prediction (Carrol, 2004). This model package was developed by the NWS Meteorological Development Laboratory and is an Eta-based Model Output Statistics (MOS) (Glahn, 1972) guidance package which runs twice daily at midnight and noon UTC. At each run, the software program generates PoP predictions covering every six, 12, and 24 hour period between the time of the model execution and 60 hours beyond the time of model execution. The MOS technique is used with multiple linear regression to develop predictive equations for weather variables, including precipitation.

Artificial Neural Network (ANN) models have been used to predict atmospheric variables including precipitation. Hall et al. (1999) used ANNs to predict PoP for a 24 hour period. Their ANN model had a Brier score (Brier, 1950) of 0.25 and a Brier skill score (Wilks, 1995) of 0.73 using 0.342 as the reference score. Kuligowski and Barros (1998) had mixed results when comparing the use of ANNs to linear regression for quantitative precipitation forecasts for a 24-hour prediction period. In comparison to linear regression, they found that the ANN model had a lower threat score (Schaefer, 1990) when the observed precipitation amount was below 15 mm and a higher threat score when the observed amount was above 15 mm. A threat score measures the accuracy of positive forecasts and ranges from zero to one with one being a perfect score. Using spatial and temporal data of recent rainfall, Luk et al. (2000) developed an ANN model for predicting flash flood rainfall amounts for 15 minutes ahead for several areas of western Sydney, Australia.

The University of Georgia's Automated Environmental Monitoring Network (AEMN) was established in 1991 and currently consists of over 75 automated weather stations throughout the state of Georgia. These stations are primarily located in rural areas and cover the breadth of the state's geographic diversity (Hoogenboom, 2000). The stations are strategically placed to cover the various climatic zones of Georgia, including the Coastal Plain, the Piedmont, and the Blue Ridge Mountains. Every second, each station scans the sensors for temperature, relative humidity, dew point, wind speed, wind direction, and precipitation. These observations are aggregated into 15-minute averages, totals, or extremes. For example, the temperature values for the last 15 minutes are averaged to calculate the aggregated value. Precipitation is aggregated as a sum, which is the total accumulation of precipitation for the last 15 minutes.

The AEMN has been useful in aiding the development of ANN models for modeling a range of atmospheric variables. Bruton et al. (2000) used the data from the AEMN to develop ANN models for the estimation of daily pan evaporation. The results of these models were shown to have a slight improvement over regression models. Jain et al. (2003) developed an ANN model to predict hourly air temperature for up to twelve hours during the winter. Smith et al. (2006) improved on the models created by Jain et al. and developed ANNs for year-round air temperature prediction for the entire AEMN domain. A large part of the improvement over previous models was a result of the ANN network parameter search using multiple instantiations of the ANNs. Shank et al. (2008a) used the same data set to create ANN models for dew point temperature predictions of up to twelve hours in advance. The research of Shank et al. (2008b) included improving ANN models by showing the value of seasonal models compared to year-round models. They also searched for the preferred duration of prior data to be used for each of the twelve prediction periods. They found that the prediction periods greater than two hours in advance required at least 18 hours of prior data for the most accurate predictions.

The goal of the study presented herein was to develop ANN models for predicting PoP for a 24-hour period beginning and ending at midnight. The prediction is to be generated at 6:00 PM, six hours prior to the start of the prediction period, with specific models for a single site. The specific objectives to achieve this goal were: a) to determine the preferred ANN parameters such as the number of hidden nodes in the hidden layer and the learning rate, b) to determine the best combination of weather variables to be used as inputs for the ANN model, and c) to compare the accuracy of the year-round model vs. the four seasonal models.

METHODOLOGY

A. Data Sets

The weather data used for model development were obtained from the Dempsey Research Farm of the University of Georgia (UGA) Campus, located in Griffin, Georgia. This location was selected because it had data starting as early as 1992. However, the sensor for measuring barometric pressure was added in 1999. The annual weather data files were left intact during partitioning into model development and evaluation. The weather data from the years 2000, 2002, 2003, 2005, 2006, and 2007 were used for model development for a total of 2190 data patterns. The weather data from 2001 and 2004 were used for model evaluation for a total of 731 data patterns (Figure 1). Two-thirds of the development data set, taken at random, was used as the training data set for a total of 1460 data patterns. The remaining third of the development data set was used as the selection data set for a total of 730 data patterns. The selection data set was used to compare the accuracy of the different ANN models, select the preferred values of the

ANN parameters, and to decide when to stop the training process of a particular ANN instantiation.

A data pattern consists of a set of inputs and outputs for presentation to the ANN. Each pattern contained variables based on 24 hours of weather data to be used as the input vector for the ANN. The input vector for the ANN included different combinations of weather variables. The available weather variables being considered as inputs to the model were temperature, relative humidity, barometric pressure, wind speed, wind direction, solar radiation, and precipitation amount. The output of the data pattern was either the occurrence or lack of occurrence of precipitation. A training epoch was defined as the period for which the ANN completes an iteration of all of the data patterns in the training data set, including weight updates, after each data pattern has been presented.

Three different ways of representing the weather variables temperature, relative humidity, barometric pressure, wind speed, and solar radiation were included. First, the observed values at the time of prediction and values in 15 minute intervals for 24 hours of prior data for each of these variables were included. Secondly, the difference between each consecutive 15-minute interval was calculated and then multiplied by four to obtain the hourly rate of change for the input vector. Finally, the difference between each 15 minute value and the value at the time of prediction was calculated and used as an overall change. Seasonality was represented by four input variables which were created by using fuzzy logic-type membership functions to capture its cyclic nature. These variables were based on the solstices and equinoxes (Figure 2) and expressed as the degree to which it is one of the four seasons. For example, the summer solstice is the first day of summer. This day was represented by having a value of 0.5 as the degree to which that day is spring and a value of 0.5 as the degree to which that day is summer. This day is directly between the middle of spring and the middle of summer. For this day, the fall and winter values were set to zero. Wind direction was presented to the ANN in a similar manner as seasonality, using the cardinal directions of north, south, east, and west instead of the four solstices and equinoxes. Precipitation amount was presented as a single value as the accumulation of precipitation over the last 15 minutes.

The data used as the target output for the ANN was the occurrence of precipitation for an entire 24-hour period. This occurrence of precipitation was represented as a binary event, with a zero representing no precipitation events during the 24-h period and a one representing a precipitation event occurred. The smallest level of rainfall measurement is a single tip of the automated rain gage system or 0.254 mm. During the initial phase of this research in which the input variables were selected, this amount during a 24-hour period was considered a precipitation event. During the latter phase of the research in which the final ANN models were developed, a single tip during a 24-hour period was considered a trace. As a result, it was not included for model development and assumed to be a zero precipitation event.

All inputs and outputs introduced to the ANN were transformed into a range between 0.1 and 0.9. Using the minimum and maximum values for each specific weather variable of the training data set, a simple scaling function was used to transform the input vector so that all inputs were inside this range. The target output used the minimum and maximum values of the transformation range to represent the binary event of the occurrence of precipitation. A non-precipitation event was represented as 0.1 and a precipitation event was represented as 0.9.

When a data pattern is presented to a trained ANN in feed-forward mode only, the output is the ANN’s prediction of precipitation for that specific data pattern and will be in the range of 0.1 to 0.9. For interpretation, this value was then transformed to a range between 0.0 and 1.0 and multiplied by 100. This output was interpreted as the percent chance of precipitation for that data pattern, which was used as the PoP value. This probability value was used instead of a binary prediction of precipitation or non-precipitation.

B. Model Development

The ANN models were created using an error backpropagation (EBP) algorithm as described by Haykin (1999). The input, hidden, and output layers of the ANN were fully-connected. The architecture selected for the ANN was the Ward architecture. The Ward architecture has been used for predictions of other weather variables such as air temperature (Jain et al., 2003; Smith et al., 2006) and dew point temperature (Shank et al., 2008a, b). This architecture consists of three slabs of nodes in the hidden layer with each slab using a different activation function: Gaussian, Gaussian complement, and hyperbolic tangent (Ward System Group, 1993). The output node used a logistic activation function. The Ward architecture requires the selection of the number of hidden nodes to be used in each slab of the hidden layer and a learning rate to determine how much of the calculated weight change to apply. A method for stopping the training process must also be selected to avoid over fitting.

An instantiation of a model is defined as a unique set of initial weights to begin the ANN training process. For each experiment, 30 instantiations of the same model were created by randomly generating initial weights. The instantiation that produced the lowest error for the selection data set was used to represent that particular model. This process was used to search for the optimal set of parameters and combination of weather variables. To determine the best number of hidden nodes per slab to be used for all of the experiments, a preliminary test was performed. The results showed that 60 hidden nodes per slab for a total of 180 hidden nodes would suffice and that increasing the number of hidden nodes had a minimal effect on the error.

A preliminary experiment to choose the learning rate of the model was also performed in which a learning rate of 0.01 was selected. For all subsequent experiments during this research, 60 hidden nodes per slab and a learning rate of 0.1 were used.

ANN models were developed using 30 Dell Pentium IV workstations in a UGA computer laboratory. Each workstation was used to train one instantiation at a time. A single instantiation of a year-round model required an average of 48 hours to be trained. A single instantiation of a seasonal model required an average of 24 hours to be trained.

C. Experiments

To decrease the likelihood of overtraining, the selection data set as described earlier was used to decide when to stop the training process. After each epoch of training, the selection data set was presented to the ANN in a feed-forward mode only. The prediction was evaluated by calculating the Brier score, which is essentially the mean squared error of a binary event (Brier, 1950). The Brier score was calculated as follows:

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2 \quad \text{, where} \qquad (1)$$

BS is the Brier score, p_i is the predicted PoP, o_i is the observation of precipitation (o_i is zero if there was no observed precipitation and o_i is one if there was observed precipitation), and N is the number of predictions. A perfect forecast has a Brier score of zero, while a forecast of no skill has a Brier score of one. As the Brier score of the selection data set decreased, the closer the forecast of this data set was to being considered perfect. If the Brier score of the selection data set increased for five training epochs in a row, the training process of the ANN was stopped. The values of the weights of the ANN during the training epoch before the Brier score began to increase were saved as the final weight values to be used for that particular training process. The training process was arbitrarily stopped if the number of training epochs exceeded 300 without being stopped by an increase in the Brier score.

While the Brier score measures the magnitude of the PoP prediction error, reliability diagrams (Hartmann et al. 2002) can be used to show how well the predicted probabilities of an event correspond to their observed frequencies. These reliability diagrams depict the observed frequency of an event plotted against the forecast probability of an event. Reliability diagrams are often coupled with frequency diagrams. An excellent prediction model would have a close to perfect accuracy on the reliability diagram and have a large frequency of predictions near 0% and 100%. This makes the model both accurate and certain.

The evaluation of the final models was calculated using the Brier skill score (Wilks, 1995). The Brier skill score is calculated as follows:

$$BSS = \frac{BS - BS_{reference}}{0 - BS_{reference}} = 1 - \frac{BS}{BS_{reference}} \text{, where} \qquad (2)$$

BSS is the Brier skill score, BS is the Brier score, and $BS_{reference}$ is the reference Brier score. The reference Brier score used was the average percentage of days that had precipitation events during each season for the entire data set (Table 1). The $BS_{reference}$ for the winter season was 0.2717, the $BS_{reference}$ for the fall season was 0.2473, the $BS_{reference}$ for the summer season was 0.2997, and the $BS_{reference}$ for the fall season was 0.2333. The Brier skill score measures the improvement of the PoP predictions made by the models when compared to the reference prediction. The Brier skill score ranges from negative infinity to a perfect score of one. A negative value for the Brier skill score represents a loss in skill when compared to the reference prediction while a positive Brier skill score represents a gain in skill when compared to the reference prediction.

The first objective of these experiments was to find the best combination of weather variables as inputs to the ANN. The weather variables considered during this experiment were temperature, relative humidity, barometric pressure, wind speed, wind direction, solar radiation, and precipitation amount. Using an iterative approach, each stage of the input variable determination experiment would find the input variable that, when added, would produce the lowest Brier score for the selection data set. The process was started by finding the first input variable that would produce the lowest Brier score. Subsequent stages of the input variable determination experiment were performed until there was no input variable that could be added that would produce a lower Brier score for the selection data set. This provides an iterative method for finding the best combination of input variables one at a time by adding the next best input variable at each stage of the experiment.

Rainfall data patterns in terms of amount, intensity, and distribution vary by season (Figure 3). Since the occurrence of precipitation fluctuates as the seasons change, four different seasonal models were created, corresponding to the winter, spring, fall, and summer seasons. A single year-round ANN model incorporating data from all seasons was also developed. The use of seasonal models creates season-specific prediction models. These seasonal models were trained with data from a particular season as well as with the data from 30 days before and after the season to account for the similarity between the end of one season and the beginning of the next season. When testing the seasonal models with the selection data set or the evaluation data set to calculate the Brier score, only data patterns from that season were used. For example, if the spring season is from day 79 to day 171, the spring model would be trained on data patterns from days 49 to 201 and tested on data patterns from days 79 to 171. Day 79, for example, corresponds to the day of year and can either be March 19th for a leap year or March 20th for a common year. The year-round model was trained with all the available data of the training data set for the entire year and tested with all the available data of the selection data set or the evaluation data set for the entire year. When comparing the use of the year-round model for a specific season to the use of the seasonal model for that season, the year-round model would only be evaluated with data patterns from that particular season.

Different weather input variables and ANN parameters might be more useful to one specific seasonal model or the year-round model than it is to another model. To explore this possible difference, each seasonal model and the year-round model went through the input variable determination test and the ANN parameter test independently.

After the input variable determination test and the ANN parameter test were completed for all of the models, the final model development was performed on each model using an enlarged training data set. The training and selection data sets were merged to form a larger training data set. A larger training data set provides more information to the ANN and should improve how well it performs. Since one-third of the development data set was originally used for the selection data set, this combined data set increased the number of data patterns in the training data set by 50%. In this case, there was no selection data set to help decide when to stop the training process to avoid overtraining. The training process of each model was stopped after a specific number of epochs. The number of epochs used was the average number of epochs that each model required during the original training process. Each of the models was trained 30 different times using the best configuration gathered from the previous experiments and the instantiation with the lowest Brier score was selected to represent each model. For final model development experiments, the training data set acted as the selection data set. These models were then evaluated over the evaluation data set only once to determine results for a final evaluation.

RESULTS AND DISCUSSION

A. Weather Variables

The weather variables considered during this experiment were temperature, relative humidity, barometric pressure, wind speed, wind direction, solar radiation, and precipitation amount. The Brier score was used to determine the accuracy of the models for the selection data set (Table 2). The best winter model was based on barometric pressure and temperature as input variables and had a Brier score of 0.1797. The best spring model was based on relative humidity and barometric pressure as input variables and had a Brier score of 0.1508. The best summer model was based on barometric pressure and precipitation amount as input variables and had a Brier score of 0.2157. The best fall model was based on relative humidity and barometric pressure as input variables and had a Brier score of 0.1586. The best year-round model was based on relative humidity and barometric pressure as input variables and had a Brier score of 0.1820. In all cases, the inclusion of a third input variable resulted in an increase of the Brier score, which meant a lower performance of the model. Barometric pressure was an input that was selected either first or second for all models. Barometric pressure was the first input variable to be included for both the winter model and summer model. Relative humidity was an input variable that was selected for the spring model, fall model, and year-round model and it was the first input variable to be included for all three of these models. Temperature was the second input variable to be included in the winter model. Precipitation was the second input variable to be included in the summer model. Wind speed, wind direction, and solar radiation were never selected as input variables for any of the models.

B. Seasonal vs. Year-round

The results of the seasonal models were compared to the use of the year-round model for the prediction of PoP for each individual season (Table 3). For instance, for the evaluation of the winter season, the year-round model was evaluated using the selection data set from the winter data set only. The resulting Brier score for the year-round model was 0.2007, which was higher than the seasonal model Brier score of 0.1797. The Brier score for the year-round model applied to the spring season was 0.1578, which was also higher than the seasonal model, which had a Brier score of 0.1508. The Brier score for the year-round model applied to the summer season was 0.2219, which was also higher than the seasonal model, which had a Brier score of 0.2157. The Brier score of the year-round model applied to the fall season was 0.1603, which was also higher than the seasonal model, which had a Brier score of 0.1586.

Using only the Brier Score, it would seem that the year-round model was less accurate than the respective seasonal models. However, considering the reliability diagrams and the frequency histograms of the summer season, it can be seen that the year-round model improves on the prediction of the summer season when compared to the summer seasonal model. The frequency histogram of the summer seasonal model (Figure 4) showed that there were only a few predictions that were associated with the 10%-20% range and the 50%-60% range. These ranges are the lowest and highest ranges for any predictions using the summer seasonal model. The frequency histogram of the year-round model when predicting the summer season (Figure 4) showed a larger number of predictions that were associated with the 10%-20% range as well as a few predictions above 70%. When the reliability diagrams for these two models were compared, the accuracy of the year-round model was not compromised by the larger frequency of predictions closer to 0% and 100% (Figure 5). The reliability diagrams and frequency histograms of the seasonal and year-round models were also compared for the winter, spring, and fall seasons, but the reliability diagrams and frequency histograms of the seasonal model were favored in all cases.

C. Final Model

The final models were created using the expanded training data set, which included the selection data set as part of the training data set. It was now assumed that the summer season would be predicted with the year-round model and the other three seasons would be predicted using their respective seasonal models based on the input variable determination experiment. These final models were trained with the larger training data set and then evaluated with an independent evaluation data set using data patterns that were not used for model development. The Brier score of the final winter season model was 0.1778, the Brier score of the final spring season model was 0.1866, the Brier score of the final summer season model was 0.2113, and the Brier score of the final fall season model was 0.1544 (Table 3). The Brier skill score of the final winter season model was 0.3456, the Brier skill score of the final spring season model was 0.2455, the Brier skill score of the final summer season model was 0.2950, and the Brier skill score of the final fall season model was 0.3382 (Table 4).

The frequency histogram of the final winter model shows that there were a large number of predictions below 30% and there were no predictions above 70% (Figure 6). The reliability diagram of the final winter model shows that it was accurate over all predictions bins (Figure 7). The frequency histogram of the final spring model shows that the largest amount of predictions occurred below 30% and that there were no predictions above 80% (Figure 8). The reliability diagram of the final spring model shows that the least accurate predictions occurred in the prediction bins of 30%-40% and 70%-80% (Figure 9). The frequency histogram of the final summer model shows that the largest amount of predictions occurred between 20% and 50%, there were only a few predictions below 10%, and there were no predictions above 80% (Figure 10). The reliability diagram of the final summer model shows that all predictions bins contained accurate predictions except for the 70%-80% prediction bin (Figure 11). The 70%-80% prediction bin contained only two predictions and precipitation was observed for both of the prediction days. The frequency histogram of the final fall model shows that there were a large number of predictions below 20% and there were a few predictions above 80% (Figure 12). The reliability diagram of the final model shows that accuracy was off for the 20%-30% prediction bin, the 40%-50% prediction bin, and all of the prediction bins above 70% (Figure 13). There were only seven predictions for the 70%-80% prediction bin and two predictions for both the 80%-90% prediction bin and the 90%-100% prediction bin.

The final models of each season were combined to create an ensemble ANN model. Each season uses a different final model to predict the days of that particular season. Together they can predict rainfall for every day of the year as an ensemble model. The results of each season for the final ensemble model were grouped together and a reliability diagram and a frequency histogram were created for the final ensemble model. The majority of the predictions occurred in between the 10% and 50% range, as shown by the frequency histogram of the final ensemble model (Figure 14). Seventy-seven of the total 713 predictions were inside the 0%-10% range, but only 14 predictions were above 70%. The reliability diagram of the final ensemble model showed that for all predictions below 70%, the observed frequencies were almost directly on the line of a perfect prediction (Figure 15). For the predictions above 70%, there was a larger error, but the sample size consisted of only 14 predictions. The low number of predictions above 70% made it difficult to show the accuracy of the models for predicting a high chance of precipitation.

D. Future Research

The results of this work leave open the possibility of further study. There frequency histogram of the final ensemble model shows that there were very few predictions made above 70%. Providing the ANN model with duplicate precipitation data patterns could result in the model increasing the number of predictions above 70%. Additional research should also focus on improving the accuracy of predictions of high PoP. One such improvement could involve predicting an amount of precipitation along with the PoP. A prediction of both a high PoP and a large amount of precipitation may lead to a greater confidence of the accuracy of PoP predictions above 70%. Predictions made during the summer season had a Brier score that was much higher than any other season, but did not have the lowest Brier skill score. There was also less of an improvement in the Brier score for the summer season between the first and second tiers of the input variable determination experiments than with the other seasons. This showed how the Brier score of the summer model was raised by its lack of certainty with its predictions. Further analysis of the weather data during the summer could lead to new ideas for the implementation of the ANN model and improving the ability of the summer model to increase the number predictions of PoP that are less than 10% or greater than 60%. A larger amount of historical weather data might be needed for the ANN model to fully capture the tendencies of the summer season. The weather data that were used in this study were from a single location. The ANN model may benefit from receiving information from multiple locations. The predictions of the ANN model could also be applied to other locations or applied to areas that include multiple locations. Adding weather data from other locations would also increase the number of data patterns available for the development of the ANN model.

SUMMARY AND CONCLUSION

In this study models were developed for predicting the PoP for a 24-hour period with the prediction generated 6 hours prior to the start of the period. It can be concluded from the preliminary experiments that the ANN model for PoP prediction should be based on 60 hidden nodes per hidden layer for a total of 180 hidden nodes in all three layers and a learning rate of 0.01. The determination of the best combination of input weather variables involved a different search for the year-round model and the four seasonal models. Although different combinations of weather variables were shown to be useful depending on the applicable time of the year, barometric pressure was selected for each model. When the application of the year-round model versus the individual seasonal models was compared, it was found that the summer season was predicted more accurately using year round data. The other three seasons were predicted more accurately only data from that particular season. The year-round model was only used to predict the summer season.

The final models for each season improved over a prediction of no skill since the Brier skill scores were all above zero. The lowest Brier skill score was for the spring season prediction and the highest Brier skill score was for the fall season prediction, with values of 0.2455 and 0.3382, respectively. Even though the final fall model had the highest Brier skill score, its reliability diagram shows that it could have been more accurate for the 20%-30% and 40%-50% prediction bins (Figure 13). The Brier skill score of the final fall model was higher due to the large number of accurate predictions below 30%. The final summer model was accurate for all prediction bins where predictions occurred (Figure 11). Although the final summer model did not have any predictions below 10% for the selection data set, there were four predictions below

10% for the evaluation data set. The Brier skill score of the summer model was lower due to the low number of predictions below 20%.

The reliability diagram and frequency histogram of the final ensemble model showed that the ensemble model was more reliable at predicting a low chance of precipitation than a high chance of precipitation. There were a large number of predictions below 50% which were extremely accurate, while there were only a few predictions above 70% and precipitation was observed during only 50% of those predictions.

REFERENCES

Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. Monthly Weather Review, 78(1): 1-3.

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

Carrol, K.L. and Maloney, J.C., 2004. Improvements in extended-range temperature and probability of precipitation guidance. Symposium on the 50th Anniversary of Operational Numerical Weather Prediction, Am. Meteor. Soc., College Park, MD.

Glahn, H.R. and Lowry, D.A., 1972. The use of model output statistics (MOS) in objective weather forecasting. Journal of Applied Meterology, 11(8): 1203-1211.

Hall, T., Brooks, H.E. and Doswell, C.A., 1999. Precipitation forecasting using a neural network. Weather and Forecasting, 14(3): 338-345.

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

Kuligowski, R.J. and Barros, A.P., 1998. Localized Precipitation forecasts from a numerical weather prediction model using artificial neural networks. Weather and Forecasting, 13(4): 1194-1204.

Luk, K.C., Ball, J.E. and Sharma, A., 2000. A study of optimal lag and spatial inputs for artificial neural network for rainfall forecasting. Journal of Hydrology, 227: 56-65.

Schaefer, J.T., 1990. The critical success index as an indicator of warning skill. Weather and Forecasting, 5(4): 570-575.

Shank, D.B., G. Hoogenboom, and R.W. McClendon. 2008a. Dew point temperature prediction using artificial neural networks. Journal of Applied Meteorology and Climatology, 47(6):1757- 1769.

Shank, D.B., McClendon, R.W., Paz, J.O., and Hoogenboom, G. 2008b. Ensemble artificial neural networks for prediction of dew point temperature. Applied Artificial Intelligence, 22(6): 523-542.

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

Wilks, D.S., 1995. Statistical methods in the atmospheric sciences. London, United Kingdom: Academic Press: 467.

Table 1: Values used for calculating the reference score for each season, entire data set (2000 – 2007).

Season	Precipitation Events	Number of Days	Reference Score
Winter	194	714	0.2717
Spring	184	744	0.2473
Summer	223	744	0.2997
Fall	168	720	0.2333

Table 2: Results of the input variable determination experiment, selection data set.

Season / Tier	Barometric Pressure	Relative Humidity	Temperature	Precipitation Amount	Wind Speed	Wind Direction	Solar Radiation
Winter / 1	**0.1971**[a]	0.2076	0.2060	0.2148	0.2128	0.2034	0.2101
Winter / 2	--[b]	0.1980	**0.1797**	0.2006	0.1901	0.1938	0.1925
Winter / 3	--	0.1971	--	0.1957	0.1853	0.1922	0.1840
Spring / 1	0.1913	**0.1621**	0.1793	0.2003	0.1978	0.1798	0.1655
Spring / 2	**0.1508**	--	0.1625	0.1610	0.1694	0.1656	0.1603
Spring / 3	--	--	0.1605	0.1524	0.1665	0.1653	0.1545
Summer / 1	**0.2161**	0.2242	0.2264	0.2252	0.2297	0.2185	0.2352
Summer / 2	--	0.2172	0.2181	**0.2157**	0.2225	0.2233	0.2308
Summer / 3	--	0.2197	0.2281	--	0.2289	0.2302	0.2286
Fall / 1	0.1936	**0.1656**	0.1837	0.2072	0.1926	0.1833	0.1797
Fall / 2	**0.1586**	--	0.1616	0.1647	0.1668	0.1628	0.1670
Fall / 3	--	--	0.1597	0.1604	0.1630	0.1642	0.1609
YR / 1	0.1983	**0.1896**	0.2007	0.2059	0.2075	0.2005	0.2016
YR / 2	**0.1820**	--	0.1885	0.1913	0.1905	0.1890	0.1926
YR / 3	--	--	0.1900	0.1883	0.1845	0.1887	0.1869

[a]Bold values indicate chosen weather input variable. [b]Not applicable because variable was already chosen

Table 3: Comparison of the Brier score for the seasonal models, year-round models, and final ensemble model.

Season	Seasonal Model (Selection Data Set)	Year-Round Model (Selection Data Set)	Final Ensemble Model (Evaluation Data Set)
Winter	0.1797	0.2007	0.1778
Spring	0.1508	0.1578	0.1886
Summer	0.2157	0.2219	0.2113
Fall	0.1586	0.1603	0.1544

Table 4: Brier score, reference score, and Brier skill score for the final seasonal models, evaluation data set.

Season	Brier Score	Reference Score	Brier Skill Score
Winter	0.1778	0.2717	0.3456
Spring	0.1886	0.2473	0.2455
Summer	0.2113	0.2997	0.2950
Fall	0.1544	0.2333	0.3382

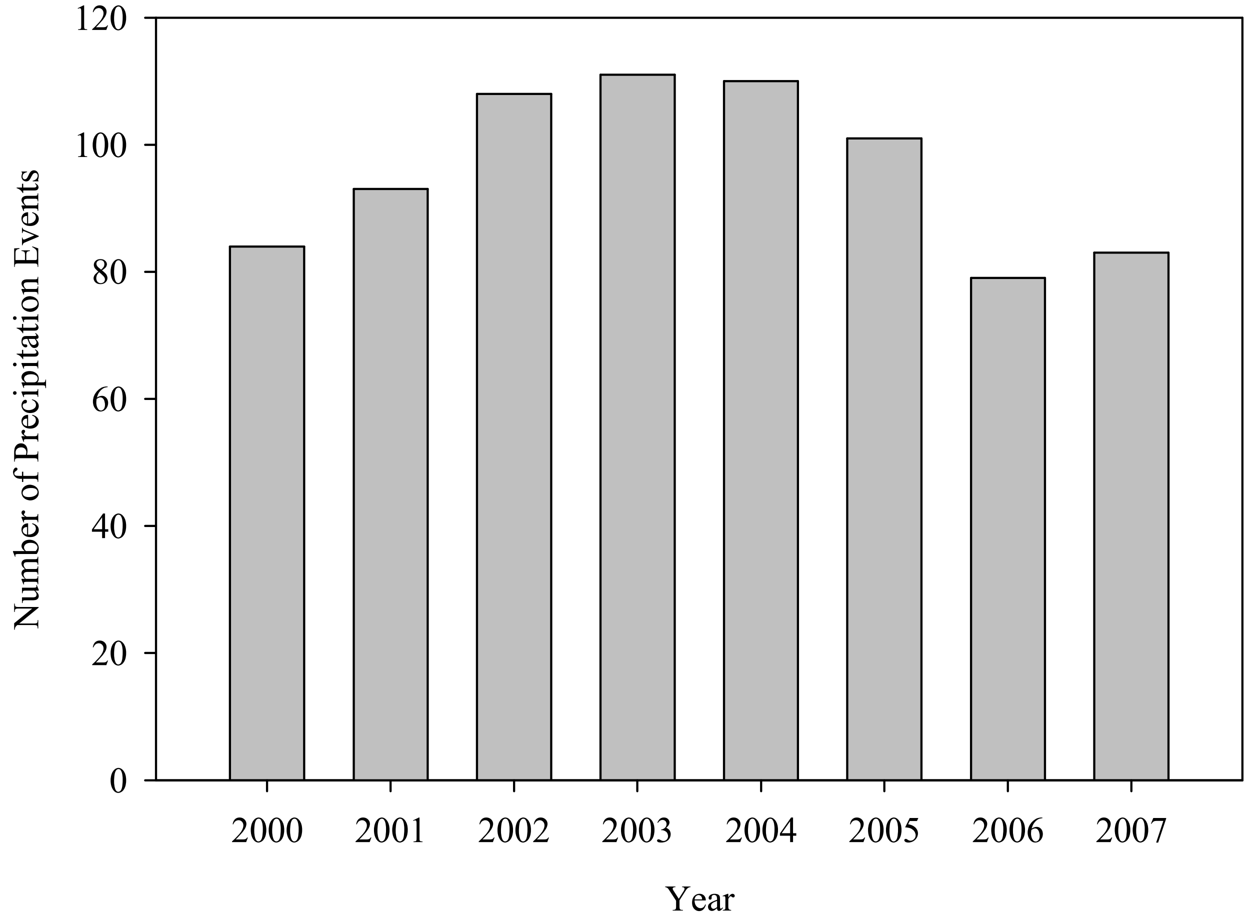

Figure 1: Number of Precipitation Events by year, entire data set (2000-2007).

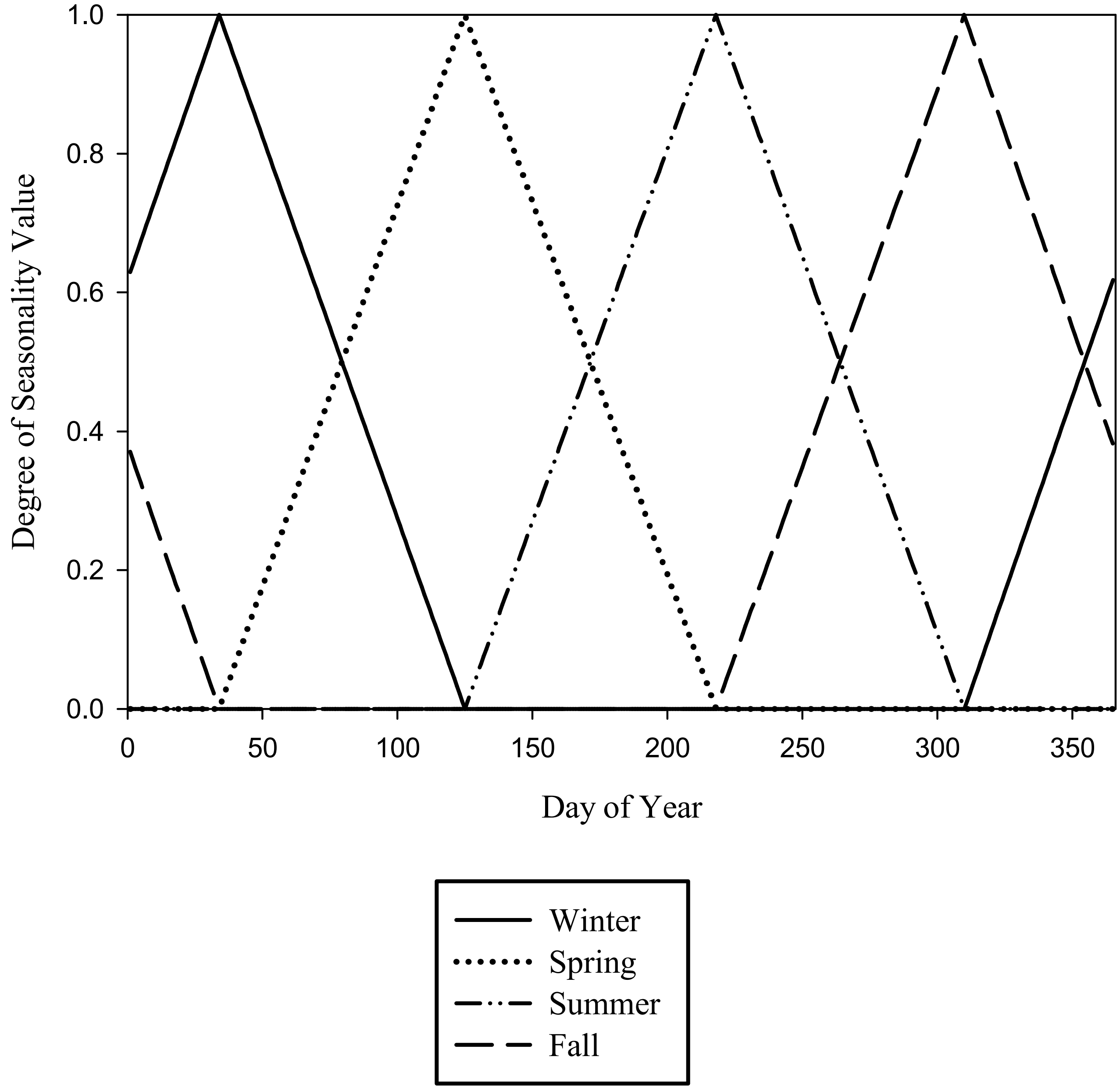

Figure 2: Inputs representing seasonal values of the four seasons, fuzzy logic type membership functions.

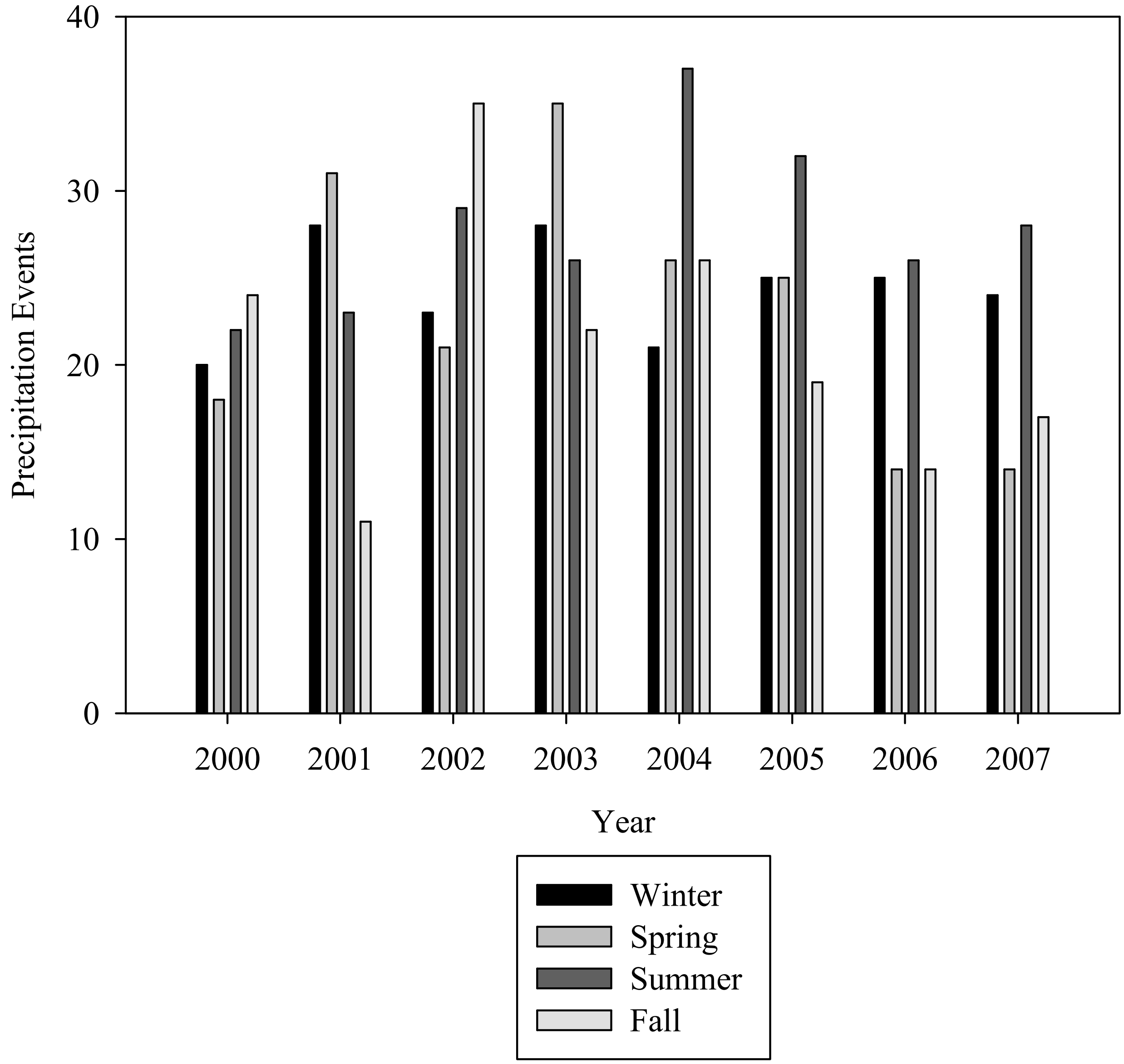

Figure 3: Number of precipitation events grouped by season and year, Dempsey Research Farm, UGA Campus, Griffin, Georgia.

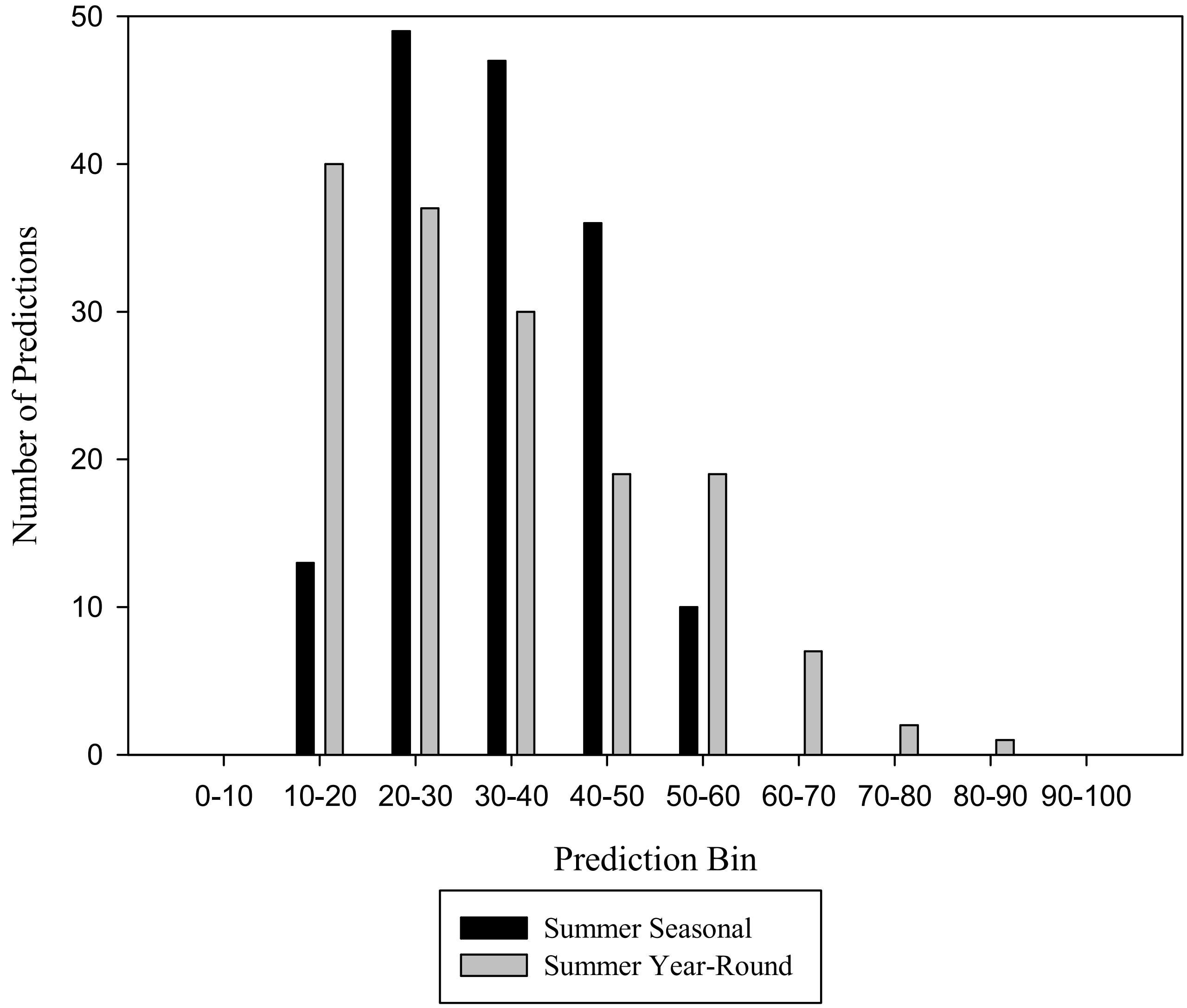

Figure 4: Frequency histogram for the summer seasonal model and the summer year-round model, selection data set.

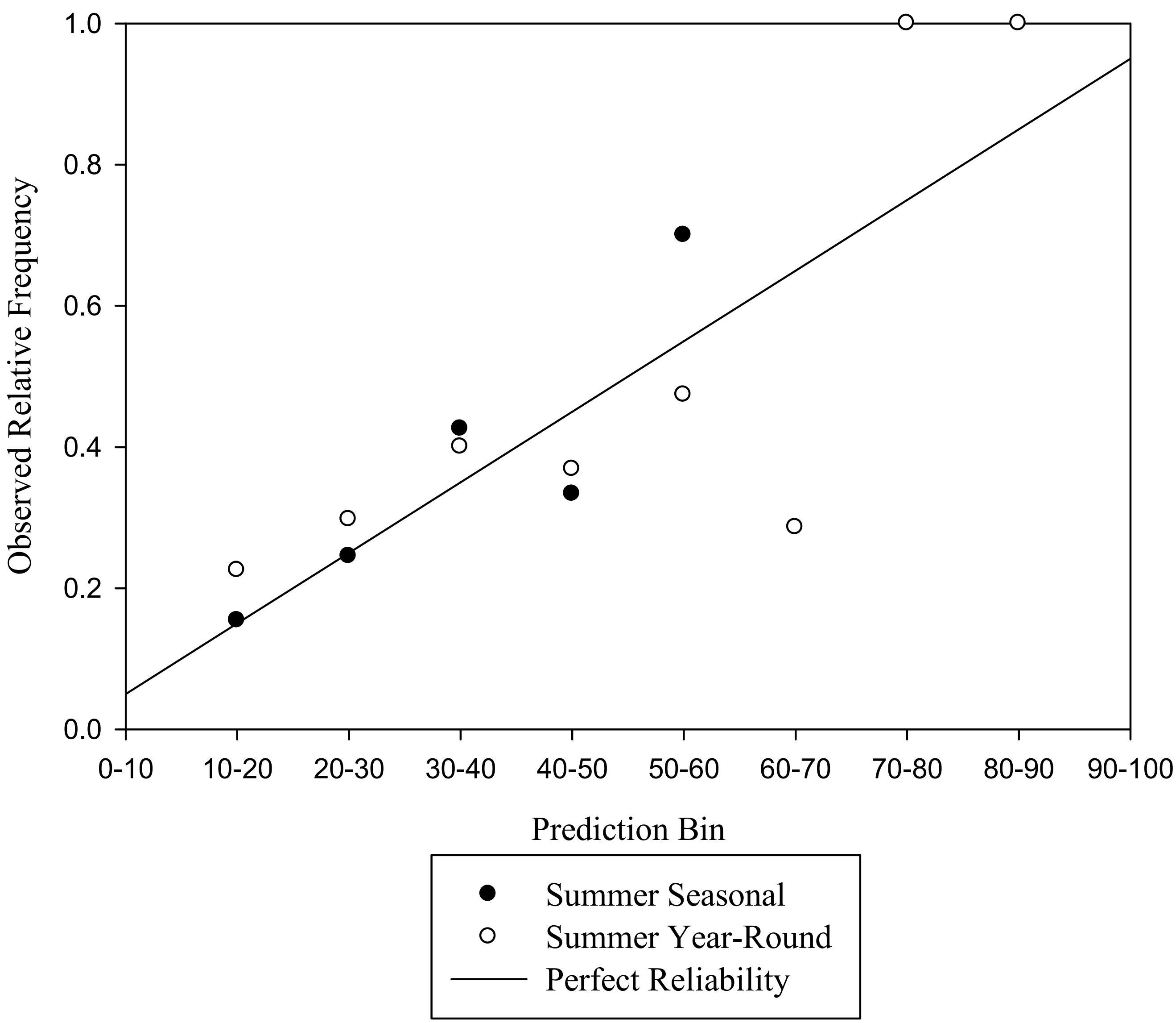

Figure 5: Reliability diagram for the summer seasonal model and the summer year-round model, selection data set.

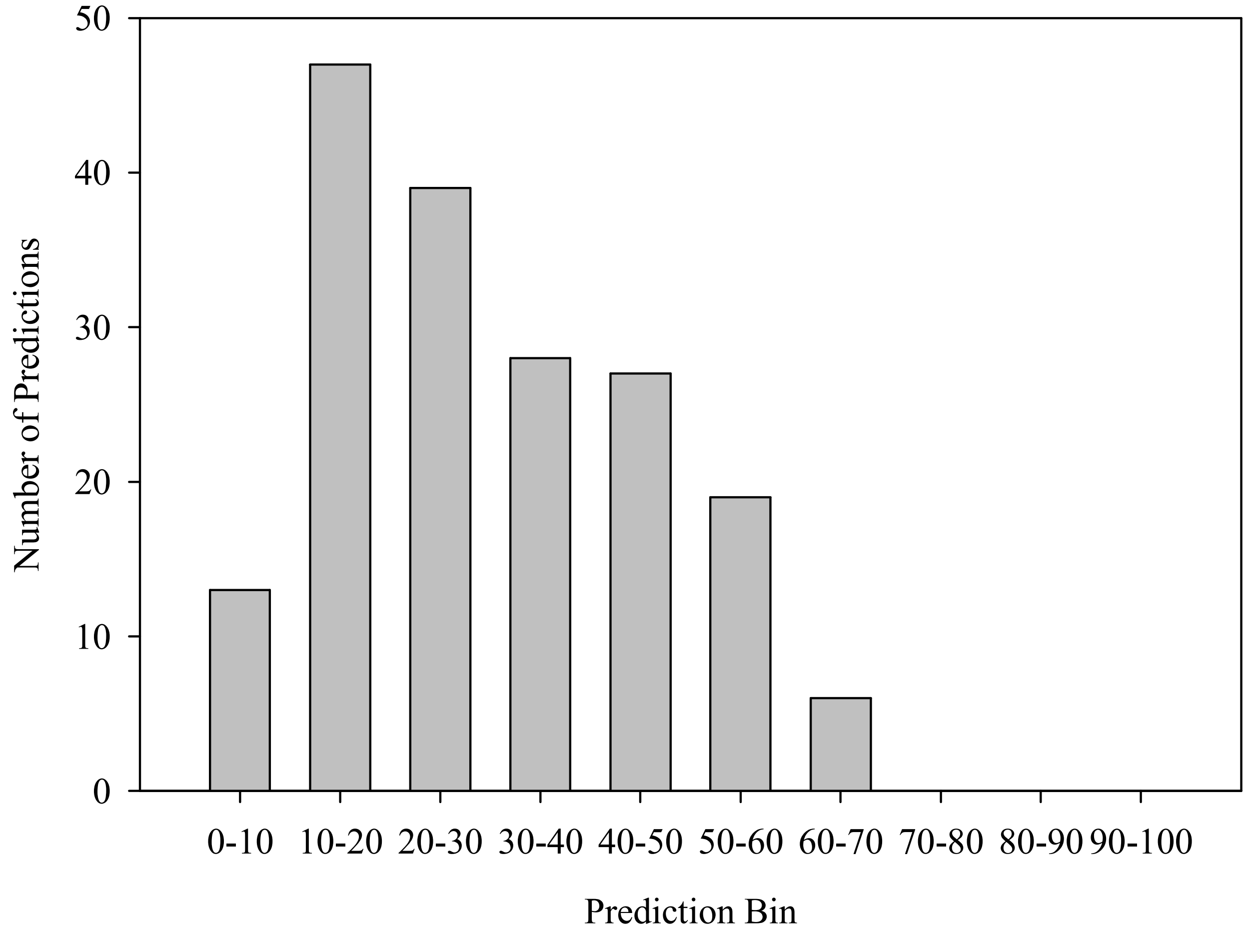

Figure 6: Frequency histogram for the final winter model, evaluation data set.

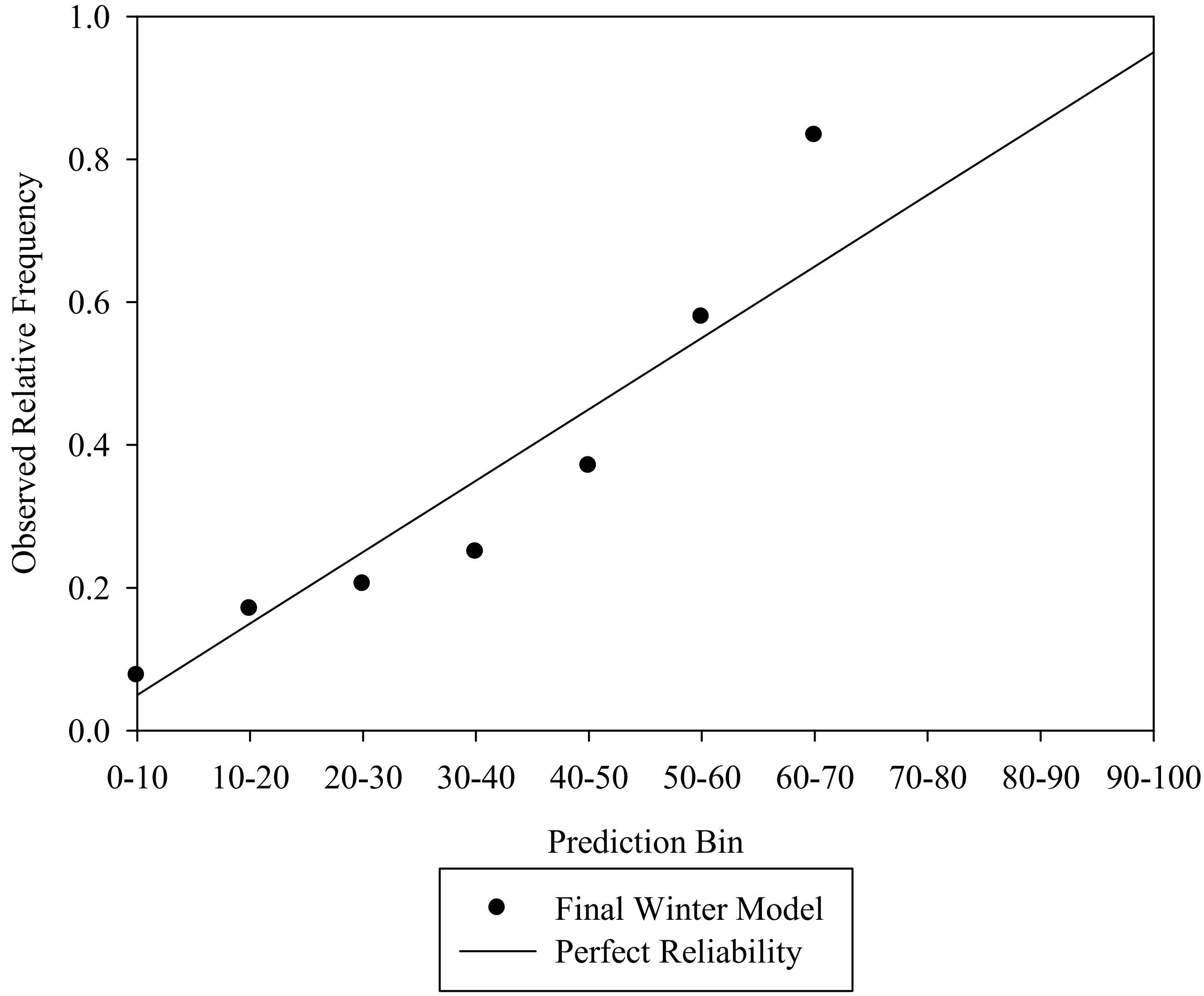

Figure 7: The reliability diagram for the final winter model, evaluation data set.

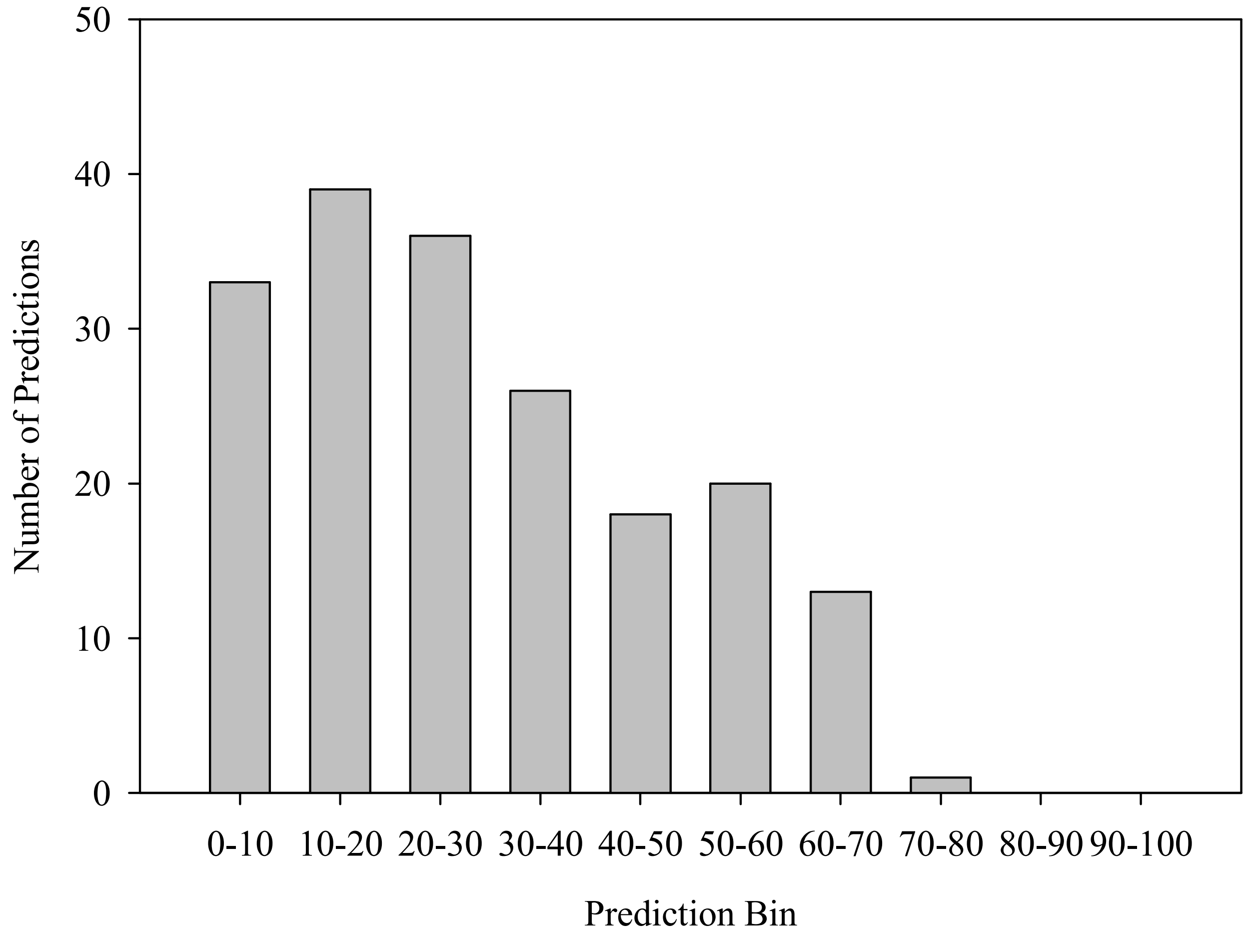

Figure 8: Frequency histogram for the final spring model, evaluation data set.

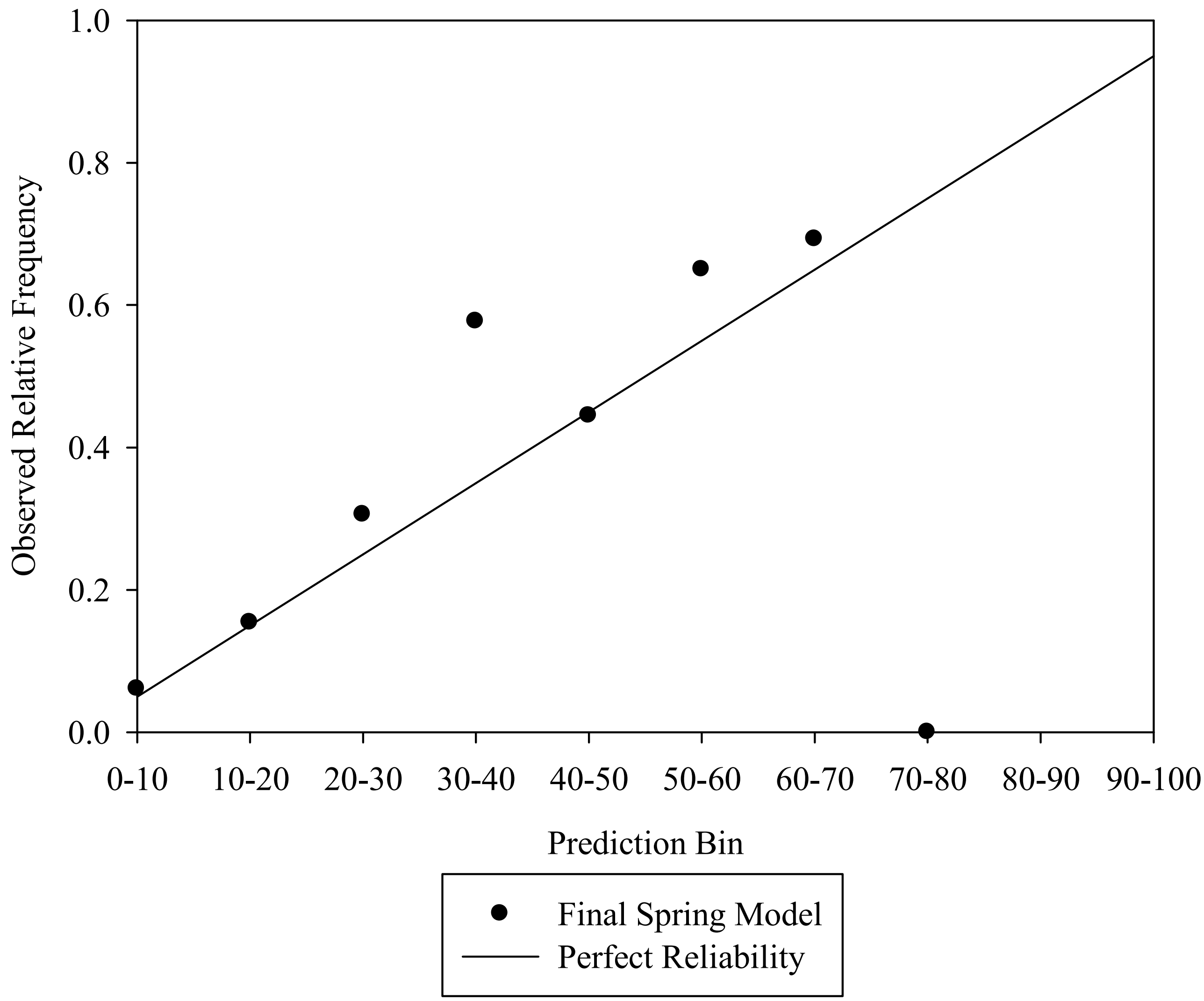

Figure 9: The reliability diagram for the final spring model, evaluation data set.

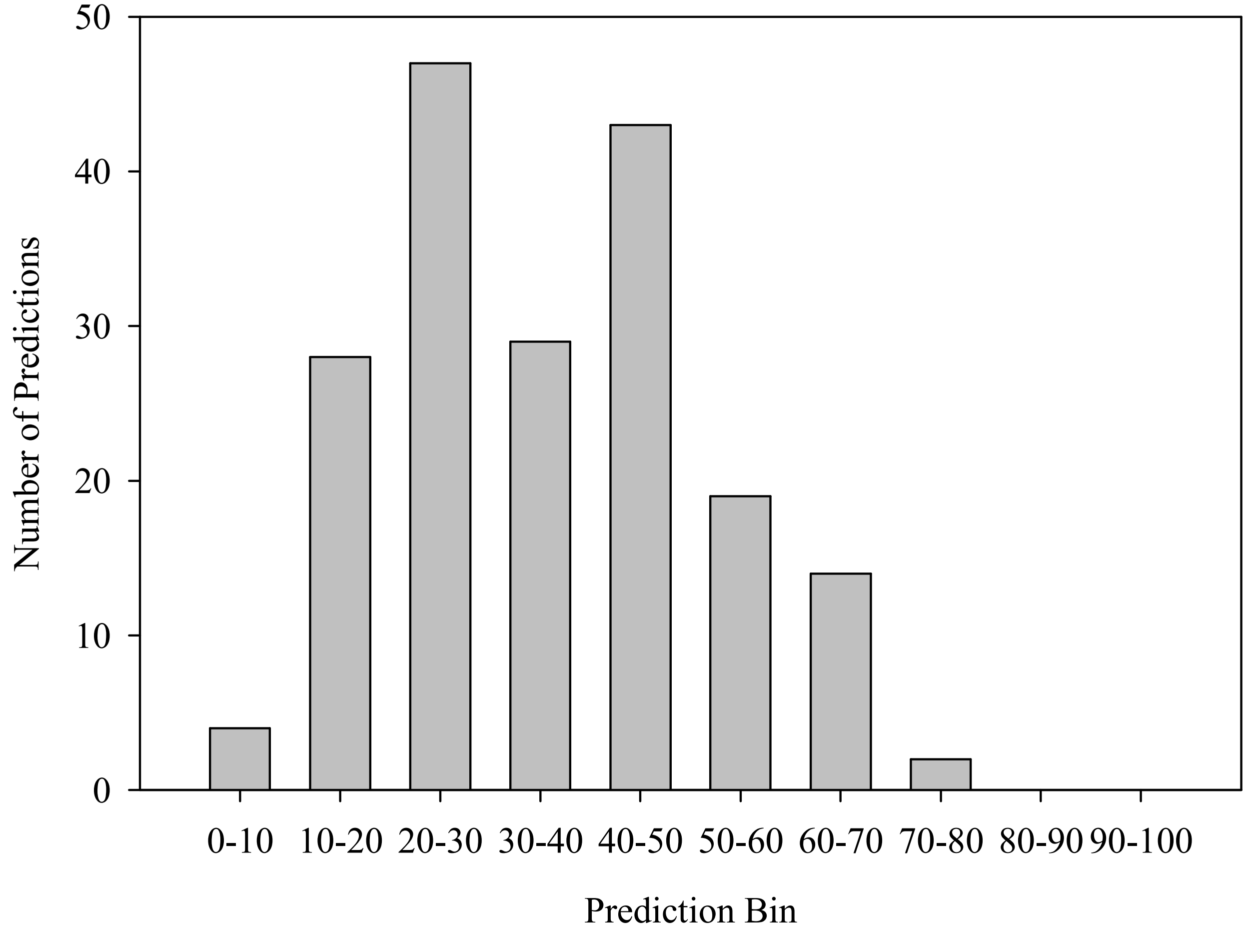

Figure 10: Frequency histogram for the final summer model, evaluation data set.

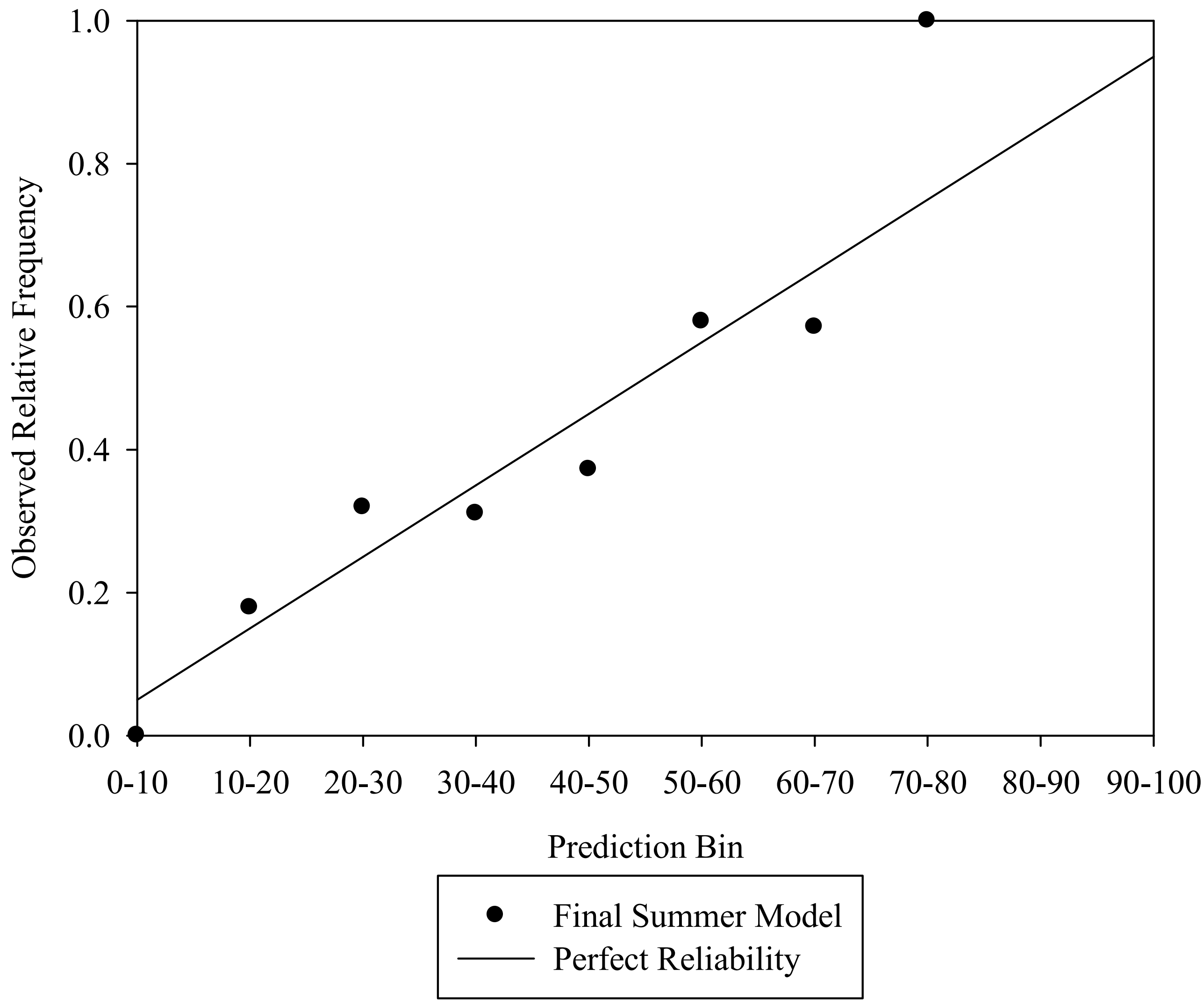

Figure 11: The reliability diagram for the final summer model, evaluation data set.

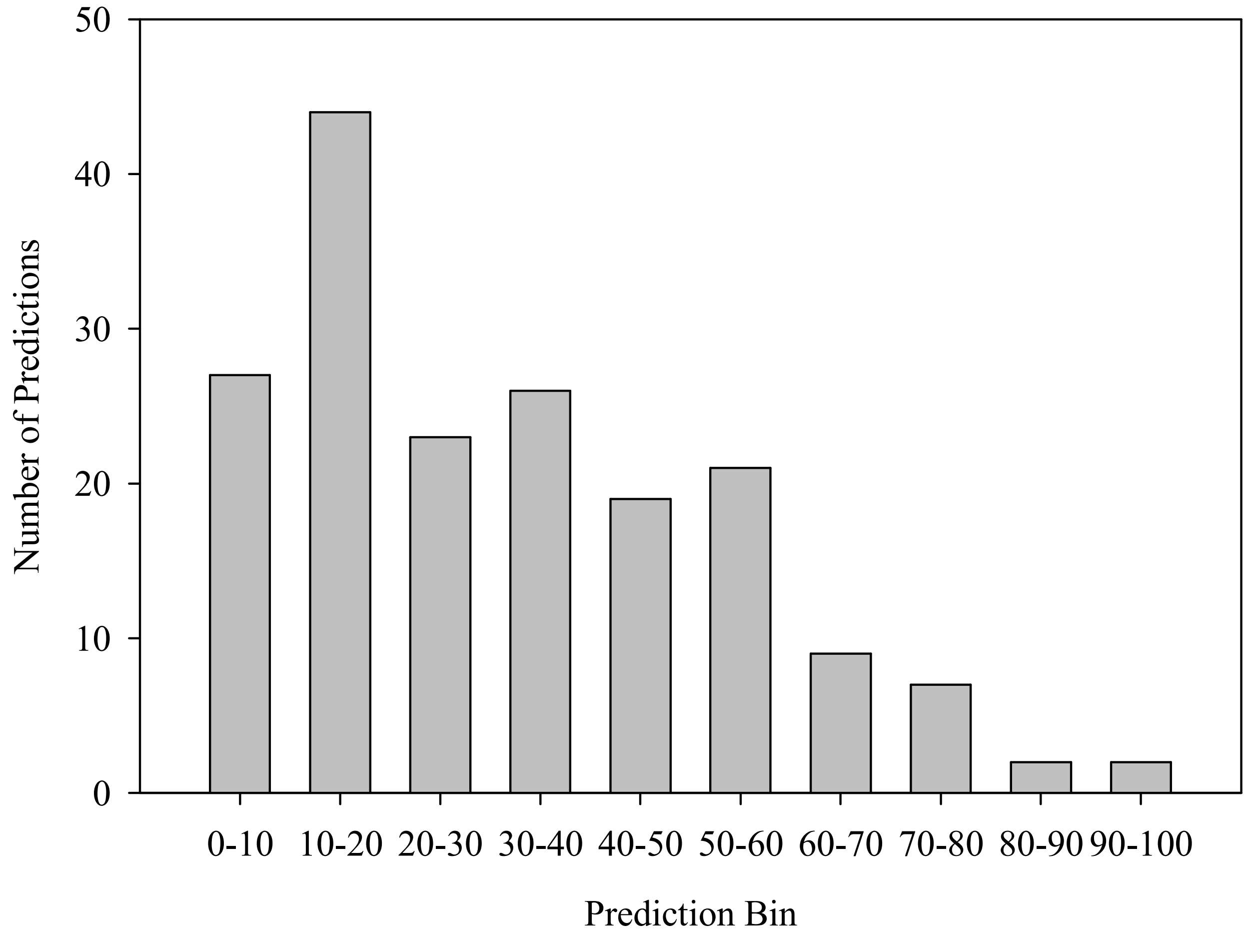

Figure 12: Frequency histogram for the final fall model, evaluation data set.

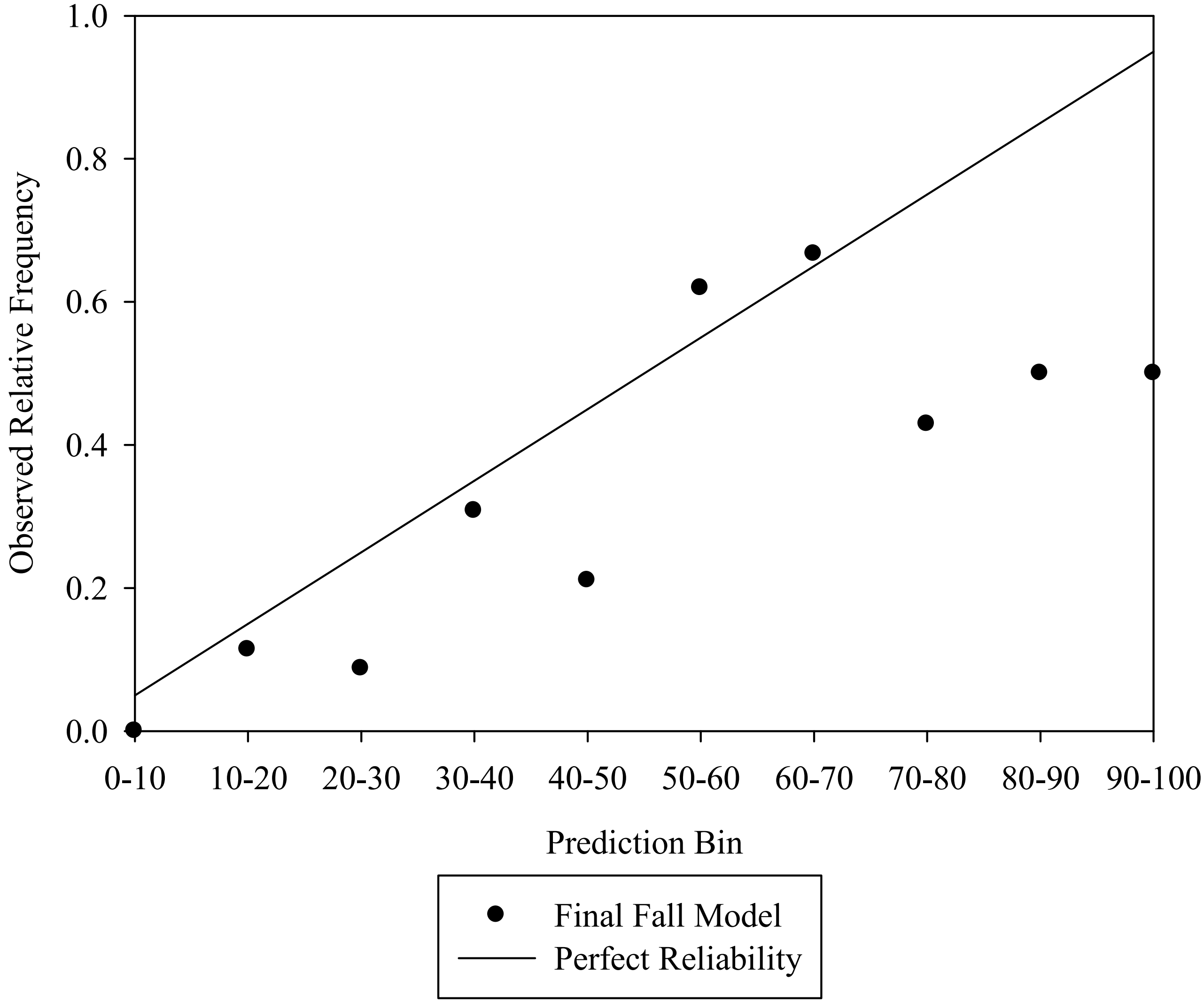

Figure 13: The reliability diagram for the final fall model, evaluation data set.

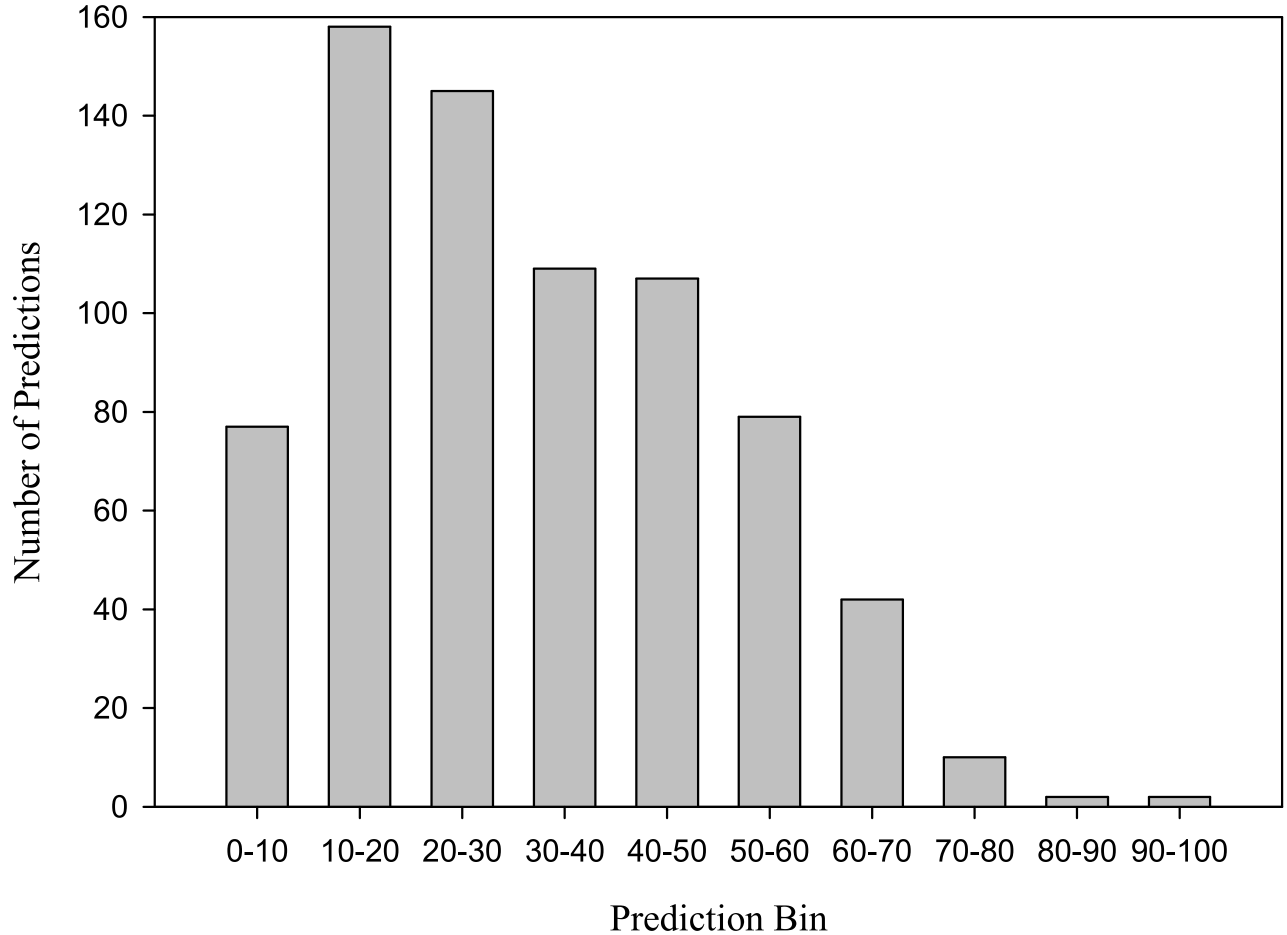

Figure 14: Frequency histogram for the final ensemble model, evaluation data set.

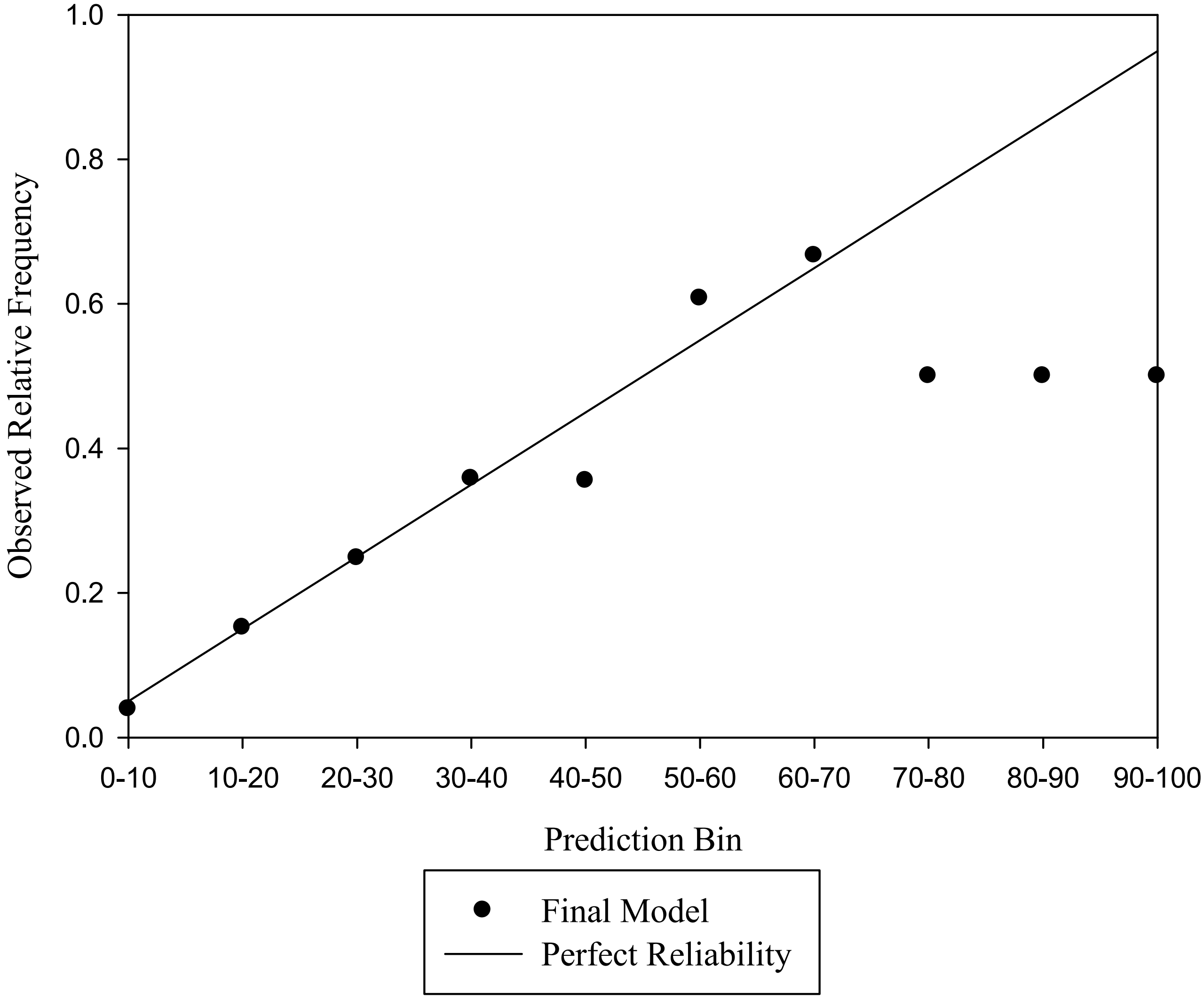

Figure 15: Reliability diagram for the final ensemble model, evaluation data set.

CHAPTER 3

QUANTITATIVE PRECIPITATION PREDICTION USING ARTIFICIAL NEURAL NETWORKS[1]

[1] Crowell, K.L., McClendon, R.W., Paz, J.O., and Hoogenboom, G. To be submitted to Weather and Forecasting.

ABSTRACT

Accurate quantitative precipitation prediction is important for minimizing crop loss, controlling river floods, and improving the weather forecast for the general public. The goal of this study was to develop artificial neural network (ANN) models for predicting quantitative precipitation for a 24-hour period beginning and ending at midnight. The prediction is to be generated at 6:00 PM, six hours prior to the start of the prediction period. Specific objectives to achieve this goal included selection of important weather variables, removal of less useful inputs that were associated with the selected weather variables, and evaluating the use of model calibration using a selection data set. An iterative search for weather variables showed that temperature and solar radiation should be chosen for the model development. The model created after removing less useful inputs using calibration was able to correctly classify 62.4% of the selection data set. When this model was applied to the evaluation data set, which had not previously been presented to the developed model in any way, the overall accuracy dropped to 39.2%. The model created without removing inputs and without using calibration was able to correctly classify 40.7% of the evaluation data set.

INTRODUCTION

A quantitative precipitation prediction (QPP) is the expected amount of liquid precipitation to be accumulated over a certain timeframe. Improving on current QPPs is important for a variety of fields. Crop producers benefit by being able to schedule irrigation based on forecasts and by being able to prepare for flood situations. The operation of surface water systems, which is crucial during flood events, could improve with more accurate QPPs. The general public would also benefit by management of municipal water supply from reservoirs and rivers, power generation by hydroelectricity, and even becoming more aware of flooding, especially in areas prone to river or coastal flooding.

Currently, the National Weather Service (NWS) has developed a model and software system for the prediction of QPP (Carrol, 2004). This system was developed by the NWS Meteorological Development Laboratory and is a Model Output Statistics (MOS) (Glahn and Lowry, 1972) guidance package, based on the National Centers for Environmental Prediction's (NCEP) Eta model output (Black, 1994). The Eta model is run twice daily at midnight and at noon UTC. At each run, the system generates QPPs covering every six, 12, and 24 hour period between the time of execution and 60 hours beyond the time of execution. The MOS approach has been used with multiple linear regression to develop predictive equations for weather variables, including precipitation, by statistically relating observed data with model output data.

Artificial Neural Network (ANN) models have been used for the prediction of precipitation. Hall et al. (1999) used ANNs to predict both Probability of Precipitation (PoP) predictions and QPP. The results of their research showed that the QPPs for all forecasts greater than 2 mm had a probability of detection of precipitation greater than 0.75, a false alarm rate less than 0.20, and a threat score greater than 0.63. A threat score (Schaefer, 1990) measures the accuracy of positive forecasts and ranges between zero and one with a perfect score being one. Kuligowski and Barros (1998) had mixed results when comparing the use of ANNs to linear regression for quantitative precipitation forecasts for a 24-hour prediction period. In comparison to linear regression, they found that the ANN model had a lower threat score when the observed precipitation amount was below 15 mm and a higher threat score when the observed amount was above 15 mm. Kim and Barros (2001) used ANN models to improve predictions of peak streamflow of four watersheds located on the leeward side of the Appalachian mountains in the mid-Atlantic region. These models reduced the r-squared for the predictions by up to 60% for forecasts made 24 hours in advance. Using spatial and temporal data of recent rainfall, Luk et al. (2000) developed an ANN model for predicting flash flood rainfall amounts for 15 minutes ahead for several areas of western Sydney, Australia.

The University of Georgia's Automated Environmental Monitoring Network (AEMN) was established in 1991 and currently consists of over 75 automated weather stations throughout the state of Georgia. These stations are primarily located in rural areas and cover the breadth of the state's geographic diversity (Hoogenboom, 2000) and the different climatic zones of Georgia, including the Coastal Plains, the Piedmont, and the Blue Ridge Mountains. Each station collects weather data such as temperature, relative humidity, dew point, wind speed, wind direction, and precipitation at a one second frequency. These observations are aggregated into 15-minute averages, totals, or extremes. For example, the temperature values for the last 15 minutes are averaged to calculate the aggregated value. Precipitation is aggregated as a total, which is the total accumulation of precipitation for the last 15 minutes.

The AEMN has been useful in aiding the development of ANNs to model atmospheric variables. Bruton et al. (2000) used data from the AEMN to develop ANN models for the estimation of daily pan evaporation. The results of these models were shown to have a slight improvement over regression models. Jain et al. (2003) developed an ANN model to predict hourly air temperature for the winter months for up to twelve hours. Smith et al. (2006) improved on the models created by Jain et al. and developed ANNs for year-round air temperature prediction for the entire AEMN domain. A large part of the improvement over previous models was a result of the ANN network parameter search using multiple instantiations of the ANNs. Shank et al. (2008a) used the same data set to create ANN models for dew point temperature predictions of up to twelve hours in advance. The research of Shank et al. (2008b) included improving ANN models by showing the value of seasonal models compared to year-round models. Shank et al. also searched for the preferred duration of prior data to be used for each of the 12 prediction periods. They found that the prediction periods greater than two hours in advance required at least 18 hours of prior data for the most accurate predictions.

The goal of the study was to develop ANN models for predicting the amount of precipitation for a 24-hour period beginning and ending at midnight. The specifics of the model include for the prediction to be generated at 6:00 PM, six hours prior to the start of the prediction period. The models will use weather data from a single site. The specific objectives are a) to determine the best combination of weather variables to be used as inputs for the ANN models, b) to optimize model performance by searching for and removing less useful inputs that are associated with the selected weather variables, and c) to evaluate the use of model calibration using a selection data set.

METHODOLOGY

A. Data Sets

The weather data used for model development were obtained from the AEMN weather station located at the Dempsey Research Farm of the UGA campus in Griffin, Georgia. This was the same location selected for the ANN models developed for PoP prediction by Crowell et al. (2008). This location had data available from 1996 to 2008. The collection of barometric pressure started in 1999 for this location, thus these data were available for a complete year starting in 2000. The same years of weather data were partitioned into model development and evaluation as used by Crowell et al. This partitioning included only complete years of weather data from 2000 to 2007. For all of the data sets discussed herein, only data patterns which resulted in a precipitation event were considered. Weather data from the years 2000, 2002, 2003, 2005, 2006, and 2007 were used for model development for a total of 676 data patterns, while the data from 2001 and 2004 were used for model evaluation data set for a total of 199 data patterns. For the training data set, 70% of the model development data set was used, which equals 474 data patterns. The remaining 30% of the development data set was used as the selection data set for a total of 202 data patterns. The selection data set was used to compare the overall accuracy of different ANN models and for the calibration of the network model.

Each of the data sets consisted of multiple data patterns. A data pattern is a set of inputs and outputs for presentation to the ANN. Each data pattern contained variables based on 24 hours of prior weather data to be used as the input vector for the ANN. The input vector for the ANN included different combinations of weather variables. The available weather variables being considered as inputs to the model were observed hourly values for temperature, relative humidity, barometric pressure, wind speed, wind direction, solar radiation, and the amount of precipitation. The output or target of the data pattern was the quantitative precipitation observed for the 24-hour period from midnight to midnight beginning six hours after the time of prediction at 6:00 PM.

Three different ways of representing the weather variables temperature, relative humidity, barometric pressure, wind speed, and solar radiation were included in the set of inputs. First, the observed values at the time of prediction and values in one hour intervals for 24 hours of prior data for each of these variables were included. Secondly, the difference between each consecutive one hour interval was determined as the hourly rate of change for the input vector. Finally, the difference between each hourly value and the value at the time of prediction was calculated and used as an overall change. Seasonality was represented by four input variables, created using fuzzy logic-type membership functions to capture its cyclic nature. These variables were based on the solstices and equinoxes (Figure 1). For example, the summer solstice is the first day of summer. This day was represented as having a value of 0.5 as the degree to which it is spring and a value of 0.5 as the degree to which that day is summer, since it is directly in between the middle of spring and the middle of summer. For this day, the fall and winter values would be set to 0. Each data pattern contained all four seasonality values. Wind direction was presented to the network in a similar manner as seasonality, using the cardinal directions of north, south, east, and west in place of the four solstices and equinoxes. Each data pattern contained four values for each of the 25 data points of wind direction for a total of 100 input values. Precipitation amount was presented as a single value which is the accumulation of precipitation over the last one hour period.

Each data pattern contained 25 values for the observed values of temperature, relative humidity, barometric pressure, wind speed, and solar radiation; there were 24 values for each previous hour and value for the current value. Each data pattern also contained 24 values for each hourly rate of change and 24 values for the overall change. Seasonality was represented in each data pattern with all four seasonality values for the day of the year for that data pattern. There were four values for each hourly data point and the current data point of wind direction for a total of 100 values. Each data pattern contained a value for each of the hourly precipitation amounts for a total of 24 values. There were a total of 493 inputs included in each data pattern.

The data used as the target output for the ANN was the accumulated precipitation for an entire 24-hour period. The amount or quantitative precipitation was represented as one of three classes. Each class represented a different range for the amount of precipitation. The AEMN weather stations use a tipping bucket rain gauge to measure liquid precipitation. A tipping bucket rain gauge measures liquid precipitation in multiples of 0.254 mm or one tip of the tipping bucket. A measure of 0.254 mm for a 24-hour period was considered a trace amount of precipitation and was ignored as actual accumulation. The only data patterns used for model development and evaluation were data patterns that included an output which represented a quantitative precipitation greater than 0.254 mm, or more than one tip of the tipping bucket rain gauge. Therefore, no separate class was needed for precipitation amounts that were 0.254 mm or less, including zero. The range of the three classes was defined as follows: a) 0.508 mm to 2.286 mm, b) 2.287 mm to 12.446 mm, and c) greater than 12.446 mm. These ranges were designed to have a relatively equal distribution of the three classes across the entire data set (Figure 2). After a precipitation amount was assigned to a precipitation range class, it was represented as an output using three output nodes. Each output node represented one of the three classes. The output node which represented the correct classification was set to have a value of one and the other two output nodes were set to have a value of zero.

When a data pattern is presented to a trained ANN in feed-forward mode, the output is the ANN's prediction for that specific data pattern. Since three classes of quantitative precipitation ranges were used in the output vector during training, the ANN's prediction would also include three values that represented the prediction for each of the three classes. For interpretation, the value for each output node represented the percent chance for the class it was assigned to would be observed. Using this interpretation, the model output presented the percent chance that each precipitation range would be observed assuming that precipitation occurred.

B. Model Development

The ANN models were created using a data pattern classification algorithm as introduced by Specht (1990) known as a Probabilistic Neural Network (PNN). The PNN uses a supervised training data set to estimate the likelihood of an input vector being part of one of the classes in the output layer. The input layer has a node for each element in an input vector, or data pattern. The pattern layer, also referred to as the hidden layer, contains a node for each data pattern in the training data set. This layer is used to calculate how similar the inputs of a "new" data pattern are represented by each data pattern in the training data set. The summation layer, also referred to as the output layer, contains a node for each class to be predicted. The values of these nodes are the network's estimation of the likelihood that the given input vector matches the class each node represents. The sum of the values of the output nodes of a PNN always add up to 100%.

The software used for training and evaluation of the PNN models was NeuroShell 2 (Ward, 1993). NeuroShell 2 has a special optional feature for PNN models. The software allows the use of a built in genetic algorithm to aid the training of the PNN models by calibrating using a data set. The genetic algorithm is used to optimize the network structure of the PNN model and to eliminate inputs that are found to be less useful. Over the course of calibrating the PNN model using the selection data set, the genetic algorithm calculates contribution factors for all of the input variables. If the contribution factor is low enough, the algorithm will automatically remove that input variable from being used. In addition, the contribution factor is used to increase or decrease the weight of particular input variables. Lastly, the genetic algorithm constantly updates a calculation for a smoothing factor by fine-tuning it to the selection data set. The smoothing factor that is calculated using the selection data set is applied to the model during model evaluation. The smoothing factor is calculated by how well the input vector can be used to generalize the target output vector.

The genetic algorithm requires the user to select a breeding pool size which is equivalent to the population size of a genetic algorithm. The maximum breeding pool size allowed by the software is 300. For maximum performance, the breeding pool size should be less than or equal to the training data set size. Since the training data set size is greater than 300, the maximum size of 300 was selected for the genetic breeding pool size. NeuroShell 2, by default, sets the stopping criteria for calibration to be 20 generations of the genetic algorithm without an improvement on the selection data set.

ANN models were developed using four machines located at UGA's Institute for Artificial Intelligence (IAI). Each of these machines has two Intel Zeon 3-GHz CPUs each with two instruction pipelines. The machines were available to use inside the UGA IAI or by remote login through a virtual private network (VPN). A single instantiation of a model, which is a unique set of initial weights to begin the ANN training process, required about four hours to be trained using NeuroShell 2.

C. Experiments

The metric by which each of the PNN models was evaluated was the overall accuracy when applied to a data set. After a model is trained, each of the data patterns in the selection data set were presented to the trained model using a feed-forward approach and the three outputs were calculated. The three outputs were then interpreted as a percent chance that a particular range of quantitative precipitation was predicted. The predicted range of quantitative precipitation was chosen by a winner-take-all approach for the three output classes. To determine if the prediction was correct or incorrect, the predicted range of quantitative precipitation was compared to the observed value of quantitative precipitation. If the observed value was inside the predicted range, that data pattern was considered as a correctly predicted data pattern. If the observed value was outside the predicted range, that data pattern was considered as an incorrect classification. The number of correct classifications was divided by the total number of predictions and multiplied by 100 to determine the overall accuracy.

The first objective of these experiments was to find the best combination of weather variables as inputs for the ANN. The weather variables considered during this experiment were temperature, relative humidity, barometric pressure, wind speed, wind direction, solar radiation, and precipitation amount. As discussed previously, temperature, relative humidity, barometric pressure, wind speed, and solar radiation were represented in three ways. For the variable selection study, two approaches of presenting these variables were considered. In the first approach only the observed hourly values were included. In the second approach, the observed values and both rates of change were used. Wind direction and precipitation amount only used the first type of representation or observed values. Seasonality was included in every experiment. Using an iterative approach, each stage of the input variable determination experiment would find the input variable that, when added, would produce the highest percentage that was classified correct for the selection data set. The process was started by finding the input variable that would produce the highest percentage classified correctly for the selection data set. Subsequent stages of the input variable determination experiment were performed until there were no input variable could be added that would produce a higher percentage that was classified correctly for the selection data set. This provided an iterative method for finding the best combination of input variables one at a time by adding the next best input variable at each stage of the experiment.

The second objective of these experiments was to use the individual contribution factors calculated by the genetic algorithm to remove certain inputs that were associated with the selected weather variables that were below a threshold. As discussed previously, each weather input variable was represented with a number of variables that were included in each data pattern. The genetic algorithm calculates a contribution factor for each of the individual input values of a data pattern. The contribution factors ranged from zero to three and provided a way of measuring the importance of an individual input value. Any input value with a contribution factor between 0 and 0.3 was deemed as being not useful enough and removed from the model. After the removal of these inputs, the model was trained and calibrated again. This process was repeated three times to create a model that contained a larger percentage of input values that were more useful for the model.

The final objective of these experiments was to compare the model described in the second objective with a model that does not use calibration. The input vector of all of the models for this experiment was comprised of the seasonality values and the weather variables that were selected in the first objective. Only the model from the second objective utilized the removal of less useful input variables. The smoothing factor used for the models from the second objective was calculated using the genetic algorithm during the calibration of the model. The overall smoothing factor used for the model that did not include calibration was left as the default smoothing factor defined by NeuroShell 2, which was 0.6. Each model was applied to the evaluation data set to determine which model was better at predicting the evaluation data set.

RESULTS AND DISCUSSION

A. Weather Variable Determination Experiments

The following results are based on a winner-take-all approach when deciding which output is selected for a particular model and are determined using the selection data set. The best model was based on the inclusion of temperature and solar radiation, both including the rates of change (Table 1). The model which included temperature only, with the rates of change, predicted

53.0% of the data patterns in the selection data set correctly. The model which included both temperature and solar radiation, with the rates of change, correctly classified 125 out of 202 data patterns for an overall accuracy of 61.9%. Both of the weather variables chosen included the rates of change for the respective weather variable. For the chosen weather variables, excluding the rates of change produced a lower overall accuracy on the selection data set (Table 2).

B. Input Removal Experiments

Barometric pressure was not chosen as a weather variable during the input variable determination experiments. Without the inclusion of barometric pressure, more years of data became available for model development. In this case, the complete years of 1996, 1997, 1998, and 1999 become available for use. Barometric pressure did not become available in a complete year until 2000. All of the complete years of 1996 to 1999 are added to the training data set of all of the subsequent models. With a new training set, the overall accuracy of the models may differ from accuracies described earlier.

During each of the tiers of the input removal experiment, the overall accuracy of the model when applied to the selection data set increased. The number of inputs removed either remained static or decreased after each of the tiers. The first model created, which was created without removing any inputs, correctly classified 107 out of 202 data patterns for an overall accuracy of 53.0% (Table 3). The last model created, which was created after four iterations of removing inputs, correctly classified 127 out of 202 data patterns for an overall accuracy of 62.4%. A total of 62 of the original 153 inputs were removed.

C. Comparison of Models

The first model used for the comparison is the final model created from the input removal experiments. This model uses temperature and solar radiation both including the rates of change but is optimized by the removal of less useful inputs used to represent temperature and solar radiation. When applied to the evaluation data set, this model correctly classified 78 out of 199 data patterns for an overall accuracy of 39.2%. The second model used for comparison is a model created without using calibration. In this case, a selection data set is unnecessary. The selection data set is included as part of the training data set for this model. When applied to the evaluation data set, this model correctly classified 81 out of 199 data patterns correctly for an overall accuracy of 40.7%.

D. Future Research

Further research in the area of applying ANNs to quantitative precipitation predictions, especially PNNs, could improve the outcomes of this study. Improvement could be made by concentrating on eliminating the possibility of over-fitting. Data for a selection data set can be taken from years that are not included in the training data set. Also, a larger training data set might provide more diversity, especially considering the behavior of a PNN model. Increasing the number of sites that data is selected from could help increase the training data set size, but it may be beneficial to still predict a single location for a single model. Localized quantitative precipitation forecasts are more useful than regional forecasts. Previous research by Crowell et al. (2008) and Shank et al. (2007) show that the use of seasonal models can be beneficial when predicting weather variables. The use of PNN seasonal models may improve the overall accuracy of the predictions. Also, other classification neural network architectures might be considered when conducting further research. An architecture that is less prone to over-fitting would be a good place to begin.

SUMMARY AND CONCLUSION

The research presented in this paper explored the development of an ANN model for predicting the amount of precipitation for a 24-hour period with the prediction generated 6 hours prior to the start of the period. A weather variable determination experiment was preformed to search for the best weather variables that should be included in the input vector of the model. Temperature and solar radiation, with their respective rates of change, were found to be the most useful weather variables. Introducing a third weather variable to the model resulted in a decrease in overall accuracy when applied to the selection data set. After the weather variable determination experiment was complete, two different types of models were created. The first model was created by removing inputs associated with the selected weather variables that were found to be less useful for the model. Using a genetic algorithm, the model was calibrated for the selection data set in order to determine contribution factors for each of the inputs. This contribution factor was used to determine which inputs to remove. The second model was created without using a selection data set for calibration. For this model, there was no calibration performed.

Even though the overall accuracy of the first model when applied to the selection data set was higher than the overall accuracy of the second model when applied to the selection set, the second model had a higher overall accuracy when applied to the evaluation data set. In this case, while the calibration on the selection data set was improving the overall accuracy of the first model when applied the selection data set, it was decreasing the overall accuracy when applied to the evaluation data set. Both models had an unacceptable overall accuracy when applied to the evaluation data set. The overall accuracy of both models was not much higher than a prediction of no skill which would be expected to have an overall accuracy of 33.3%.

REFERENCES

Black, T.L., 1994. The new NMC mesoscale Eta model: Description and forecast examples. Weather and Forecasting, 9(2): 265-278.

Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. Monthly Weather Review, 78(1): 1-3.

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

Carrol, K.L. and Maloney, J.C., 2004. Improvements in extended-range temperature and probability of precipitation guidance. Symposium on the 50th Anniversary of Operational Numerical Weather Prediction, Am. Meteor. Soc., College Park, MD.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37: 35-46.

Glahn, H.R. and Lowry, D.A., 1972. The use of model output statistics (MOS) in objective weather forecasting. Journal of Applied Meterology, 11(8): 1203-1211.

Hall, T., Brooks, H.E. and Doswell, C.A., 1999. Precipitation forecasting using a neural network. Weather and Forecasting, 14(3): 338-345.

Kuligowski, R.J. and Barros, A.P., 1998. Localized Precipitation forecasts from a numerical weather prediction model using artificial neural networks. Weather and Forecasting, 13(4): 1194-1204.

Luk, K.C., Ball, J.E. and Sharma, A., 2000. A study of optimal lag and spatial inputs for artificial neural network for rainfall forecasting. Journal of Hydrology, 227: 56-65.

Schaefer, J.T., 1990. The critical success index as an indicator of warning skill. Weather and Forecasting, 5(4): 570-575.

Shank, D.B., G. Hoogenboom, and R.W. McClendon. 2008a. Dew point temperature prediction using artificial neural networks. Journal of Applied Meteorology and Climatology, 47(6):1757- 1769.

Shank, D.B., McClendon, R.W., Paz, J.O., and Hoogenboom, G. 2008b. Ensemble artificial neural networks for prediction of dew point temperature. Applied Artificial Intelligence, 22(6): 523-542.

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

Specht, D.F., 1990. Probabilistic Neural Networks. Neural Networks, 3(1): 109-118.

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

Table 1: Weather variable determination results for variables represented by both values and rates of change by overall accuracy, selection data set.

Tier	Temperature	Solar Radiation	Barometric Pressure	Relative Humidity	Precipitation Amount	Wind Speed	Wind Direction
1	**53.0%**	52.5%	44.1%	49.5%	48.0%	54.0%	49.5%
2	--	**61.9%**	50.0%	48.0%	53.5%	57.9%	50.5%
3	--	--	54.5%	54.5%	56.9%	58.4%	50.0%

Table 2: Results of each tier of the input removal experiment using the increased training data set.

Tier	Input Values Removed	Overall Accuracy
1	17	53.0%
2	17	55.0%
3	16	56.9%
4	12	60.4%
5	--	62.4%

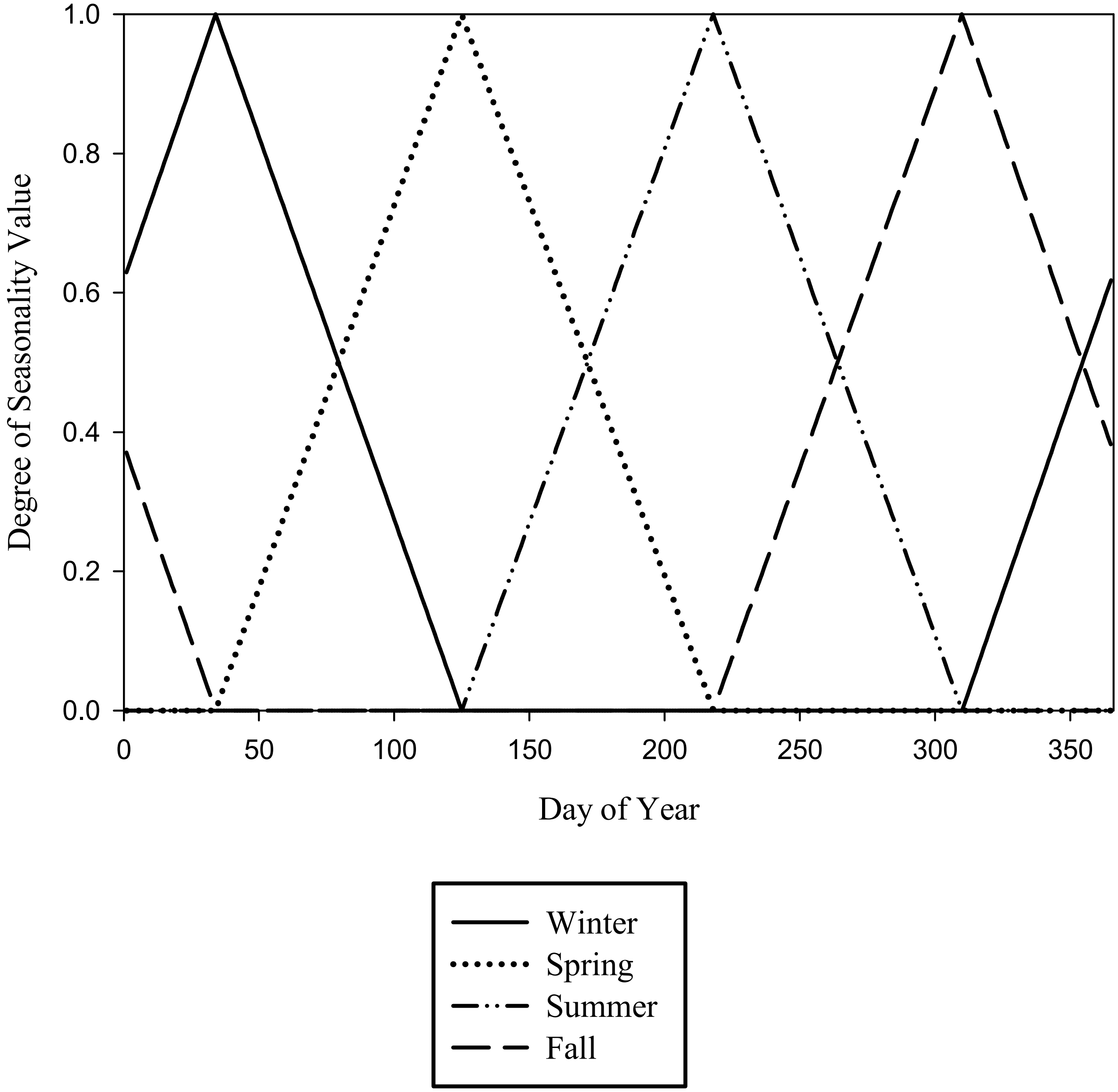

Figure 1: Inputs representing seasonal values of the four seasons, fuzzy logic type membership functions.

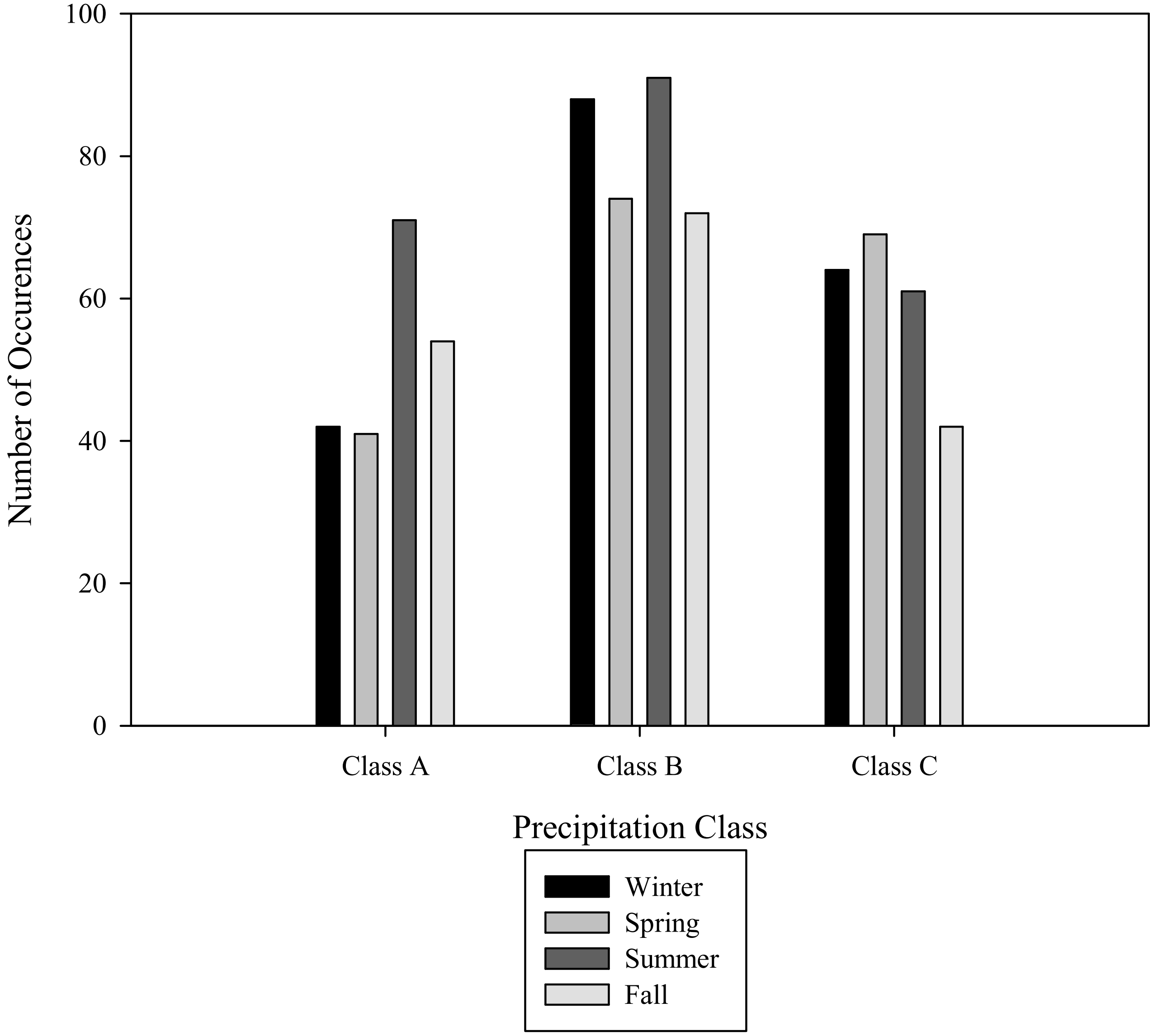

Figure 2: Distribution of precipitation classes by season. Class A = 0.508 mm to 2.286 mm, B = 2.287 mm to 12.446 mm, and Class C = greater than 12.446 mm.

CHAPTER 4

SUMMARY AND CONCLUSION

The goal of this thesis was to develop ANN models for the purpose of predicting both the Probability of Precipitation and quantitative precipitation over a 24-hour period beginning and ending at midnight. The prediction is to be generated at 6:00 PM, six hours prior to the start of the prediction period. In Chapter 2, the goal of developing an ANN model for predicting the PoP was accomplished. This part of the overall goal was accomplished by determining the preferred number of hidden nodes in the hidden layer and the preferred learning rate, searching for the preferred combination of weather input variables to be used as inputs to the ANN model, and comparing the use of a year-round model vs. a seasonal model to predict each season. Barometric pressure was found to be an important variable in the year-round model and all four seasonal models. The year-round model was used for predicting summer while the other three seasons were predicting using their respective seasonal models. The final models developed for predicting each of the four seasons all had a Brier skill score (Wilks, 1995) greater than 0.24. The resulting combination of models was merged to create a final ensemble model. The final ensemble model showed to be useful for the prediction of PoP for all seasons.

In Chapter 3, the goal of developing an ANN model for quantitative precipitation predictions was accomplished. This part of the overall goal was accomplished by using a Probabilistic Neural Network as the ANN architecture to classify the precipitation amount as three classes each representing a different amount range. Also, a search was conducted to determine the preferred combination of weather variables to be given to the input vector, inputs associated with the selected weather variables were removed based on a calculated contribution factor for each input, and an evaluation of the use of model calibration was completed. The search weather variables found that temperature and solar radiation, including the rates of change, should be used for model development. A total of 62 of the original 153 inputs were removed. The resulting models did not show to be useful for quantitative precipitation prediction when applied to the evaluation set. The models developed without using model calibration and input removal were also found not to be useful for quantitative precipitation prediction when applied to the evaluation set.

This research confirms that it is more difficult to predict quantitative precipitation than it is to predict the PoP. Further work could enhance both types of predictions, especially for quantitative precipitation. Both types of predictions could be aided by increasing the size of the development set, specifically by adding in data from surrounding areas. Research involving the use of different types of ANN architectures may also be able to produce more accurate results. It is possible that a different classification neural network, besides the probabilistic neural network, may be able to produce comparable results across the selection and evaluation sets.

REFERENCES

Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. Monthly Weather Review, 78(1): 1-3.

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

Carrol, K.L. and Maloney, J.C., 2004. Improvements in extended-range temperature and probability of precipitation guidance. Symposium on the 50th Anniversary of Operational Numerical Weather Prediction, Am. Meteor. Soc., College Park, MD.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37: 35-46.

Glahn, H.R. and Lowry, D.A., 1972. The use of model output statistics (MOS) in objective weather forecasting. Journal of Applied Meterology, 11(8): 1203-1211.

Hall, T., Brooks, H.E. and Doswell, C.A., 1999. Precipitation forecasting using a neural network. Weather and Forecasting, 14(3): 338-345.

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

Kuligowski, R.J. and Barros, A.P., 1998. Localized Precipitation forecasts from a numerical weather prediction model using artificial neural networks. Weather and Forecasting, 13(4): 1194- 1204.

Luk, K.C., Ball, J.E. and Sharma, A., 2000. A study of optimal lag and spatial inputs for artificial neural network for rainfall forecasting. Journal of Hydrology, 227: 56-65.

Raible, C.C., Bischof, G., Fraedrich, K., Kirk, E., 1999. Statistical single-station short-term forecasting of temperature and probability of precipitation: area interpolation and NWP combination. Weather and Forecasting, 14(2): 203-214.

Schaefer, J.T., 1990. The critical success index as an indicator of warning skill. Weather and Forecasting, 5(4): 570-575.

Shank, D.B., G. Hoogenboom, and R.W. McClendon. 2008a. Dew point temperature prediction using artificial neural networks. Journal of Applied Meteorology and Climatology, 47(6):1757- 1769.

Shank, D.B., McClendon, R.W., Paz, J.O., and Hoogenboom, G. 2008b. Ensemble artificial neural networks for prediction of dew point temperature. Applied Artificial Intelligence, 22(6): 523-542.

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

Specht, D.F., 1990. Probabilistic Neural Networks. Neural Networks, 3(1): 109-118.

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

Wilks, D.S., 1995. Statistical methods in the atmospheric sciences. London, United Kingdom: Academic Press: 467.