ABSTRACT

HAN, JUN. Application of Artificial Neural Networks for Flood Warning Systems. (Under the direction of Dr. Robert C. Borden and Dr. Jy S. Wu).

Artificial neural networks (ANN) are extremely useful for solving problems without existing algorithmic solutions or with algorithms that are too complex to implement. ANN-based models were developed to study hydrologic processes including the forecasts of watershed runoff, derivation of radar-based rainfall estimation, and development of radar-based flood warning systems.

In the first phase of this study, ANN was applied to the study of rainfall-runoff processes on two urbanized watersheds in North Carolina. Rainfall and discharge/gage height data were employed to train and test the ANN-hydrologic models. Stream flows or gage heights can be accurately forecasted, with forecasting periods ranging from 15 minutes to 2 hours, at the watershed outlet and for downstream locations. Based on downstream gauging stations, it was also possible to generate missing historic stream flow data of an upstream station.

The second phase of this study involved the comparison of empirical and ANN approaches to radar-based rainfall estimation. The results indicated that although z-R relationships provided fairly accurate rainfall estimates on the average, these relationships tended to overestimate the rainfall amount of low intensity storms and underestimate the rainfall amount of high intensity storms. Two ANN radar-rainfall models have been developed for radar-based rainfall estimation. Weather radar provides detailed, in both time and space, information about precipitation patterns. With proper

training and testing, ANN models have exhibited their ability of rainfall estimation from the radar data. In comparison to the empirical z-R relationship, ANN models provide more accurate estimates of rainfall for both low and high intensity storms.

The third phase of research included the development of two radar-based stream flow forecast models: a radar-rainfall-runoff (3R) model and a radar-runoff (2R) model. In the 3R-model, radar data was employed to provide rainfall estimates which, in turn, served as input to a hydrologic stream forecast model for stream flow forecasts. The 2R-model employs historic radar and stream-flow data as inputs to the ANN model which directly produce predication of stream flow at one-hour lead time. Better performance was obtained from the ANN-based 2R model in terms of accuracy, efficiency and cost.

APPLICATION OF ARTIFICIAL NEURAL NETWORKS FOR FLOOD WARNING SYSTEMS

by

JUN HAN

A dissertation submitted to the graduate faculty of
the North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

CIVIL AND ENVIRONMENTAL ENGINEERING

Raleigh, North Carolina

2002

APPROVED BY:

Dr. James D. Gregory	Dr. Margery F. Overton
Dr. Robert C. Borden Chair of Advisory Committee	Dr. Jy S. Wu Co-chair of Advisory Committee

BIOGRAPHY

Jun Han was born in Guangxi, P.R. China. In August, 1990, Jun graduated from Sichuan University with a Bachelors degree in Biochemistry. From 1990 to 1993, he worked as research associate in Guangxi Agricultural University, China. Mr. Han completed his Master of Environmental Engineering from Asian Institute of Technology, Bangkok, Thailand, in 1995. He also received his Master degree in Computer Science from University of Carolina at Charlotte, USA in 2001. Finally he defended his dissertation in December of 2002 and will graduate with his Ph.D in Civil/Environmental Engineer from North Carolina State University, Raleigh, in May of 2003.

ACKNOWLEDGEMENTS

I am greatly thankful to my advisors Dr. Robert C. Borden and Dr. Jy S. Wu for providing supervision and supporting me in many ways during the study over the last couple of years. I would also like to acknowledge Margery F. Overton and James D. Gregory, my committee members, who helped me develop a fine paper.

This research was partially supported by cities of Greensboro and Raleigh, and to them I am grateful. I also would like to thank the people from U.S. Geological Survey, who provided large amount of precipitation and stream flow data. They made this project possible.

Most of all, I am eternally grateful to my wife Yuxin Gan, who worked hard to support the family all these years. It was because of her selfless contribution that I could concentrate on my study. I also want to acknowledge my lovely son, Luke. My family means the whole world to me and without their love, support, encouragement and understanding I could not have completed my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I: Application of Artificial Neural Network in Hydrological Process

1.1 Introduction to Artificial Neural Networks (ANN)

ANNs are based on the present understanding of the biological nervous systems. An ANN is a massively parallel-distributed information processing system that has certain performance characteristics resembling biological neural networks of the human brain (Haykin 1994). The network consists of layers of parallel processing elements, called neurons. In most networks, the input layer receives the input variables for the problem at hand. This consists of all quantities that can influence the output. The output layer consists of values predicted by the network and thus represents model output. Between the input layer and output layer there may be one or more hidden layer. The neurons in each layer are connected to the neurons in a proceeding layer by a weight, w, which can be adjusted during training. The networks are organized by training methods, which greatly simplify the development of specific applications. Classical logic in ordinary artificial intelligence (AI) systems is replaced by vague conclusions and associative recall. This is a big advantage in all situations where no clear set of logical rules can be given. Figure 1.1 illustrates a three-layer neural network consisting of four neurons in input layer, four neurons in hidden layer and two neurons in output layer, with the interconnection weights between layers of neurons.

1.2 ANNs Training:

Training network is a procedure during which ANN processes a training set (input-output data pairs) repeatedly, changing the values of its weights, according to a predetermined algorithm, to improve its performance.

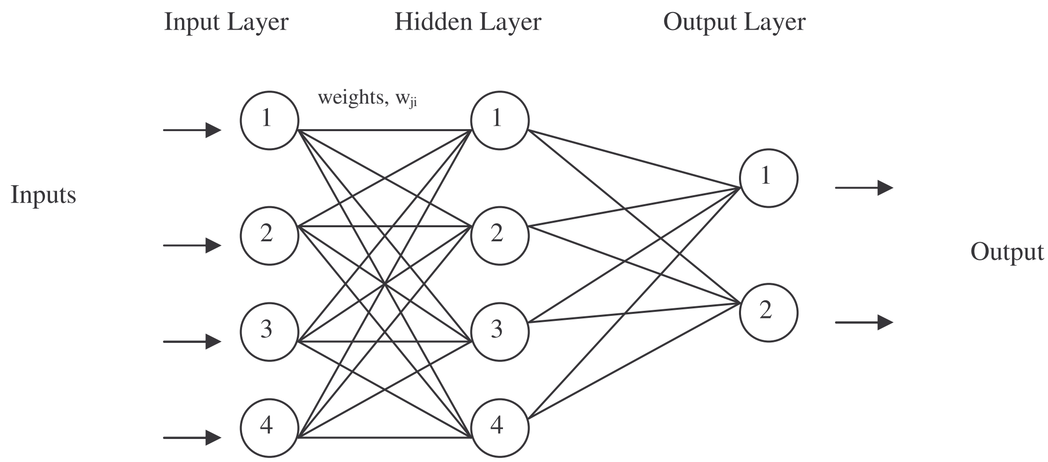


Figure 1. 1 Configuration of Three-layer Neural Network

Back-propagation is perhaps the most popular algorithm for training ANNs. The back-propagation algorithm gives a prescription for changing the weights, w_{ji}, in any feed-forward network to learn a training vector of input-output pairs. It is a supervised learning method in which an output error is fed back through the network, altering connection weights so as to minimize the error between the network output and the target output. The following equation is used for the connection weights adjustment.

$$\Delta w_{ij}(n) = -\varepsilon^*(\partial E/\partial w_{ij}) + \alpha^*\Delta w_{ij}(n-1) \qquad (1.1)$$

Where

$\Delta w_{ij}(n)$ and $\Delta w_{ij}(n-1)$ = weight increments between node i and j during the nth and (n-1)th pass, or epoch.

ε = learning rate.

α = momentum.

The momentum factor can speed up training in very flat regions of the error surface and help prevent oscillations in the weights. A learning rate is used to increase the chance of avoiding the training process being trapped in local minima instead of global minima.

Back-propagation is a first-order method based on the steepest gradient decent, with the direction vector being set equal to the negative of the gradient vector. Consequently, the convergence may progress slowly and may show oscillatory behavior. It is also possible for the training process to be trapped in the local minimum despite the use of learning rate. The network architecture is required to be prefixed by trials to do so. Figure 1.2 shows the flowchart of the Backpropagation training algorithm.

1.3 Advantages of ANN Models

ANNs offer valuable characteristics unavailable together elsewhere (Zealand et al. 1999): ANN models infer solutions from data without prior knowledge of the regularities in the data; they extract the regularities empirically. ANN networks learn the similarities among patterns directly from instances or examples of them. ANNs can modify their behavior in response to the environment (i.e. shown a set of inputs with corresponding desired outputs, they self adjust to produce consistent responses). ANNs can generalize from previous examples to new ones. Generalization is useful because real-world data are noisy, distorted, and often incomplete. ANNs are also very good at the abstraction of essential characteristics from inputs containing irrelevant data. ANN models are non-linear; that is, they can solve complex problems more accurately than linear techniques do. ANN models can provide predications of output parameters in real time in response to simultaneous and independent fluctuations of the values of model input parameters. Finally, ANNs are highly parallel. They contain many identical, independent operations that can be executed simultaneously, often making them faster than alternative methods.

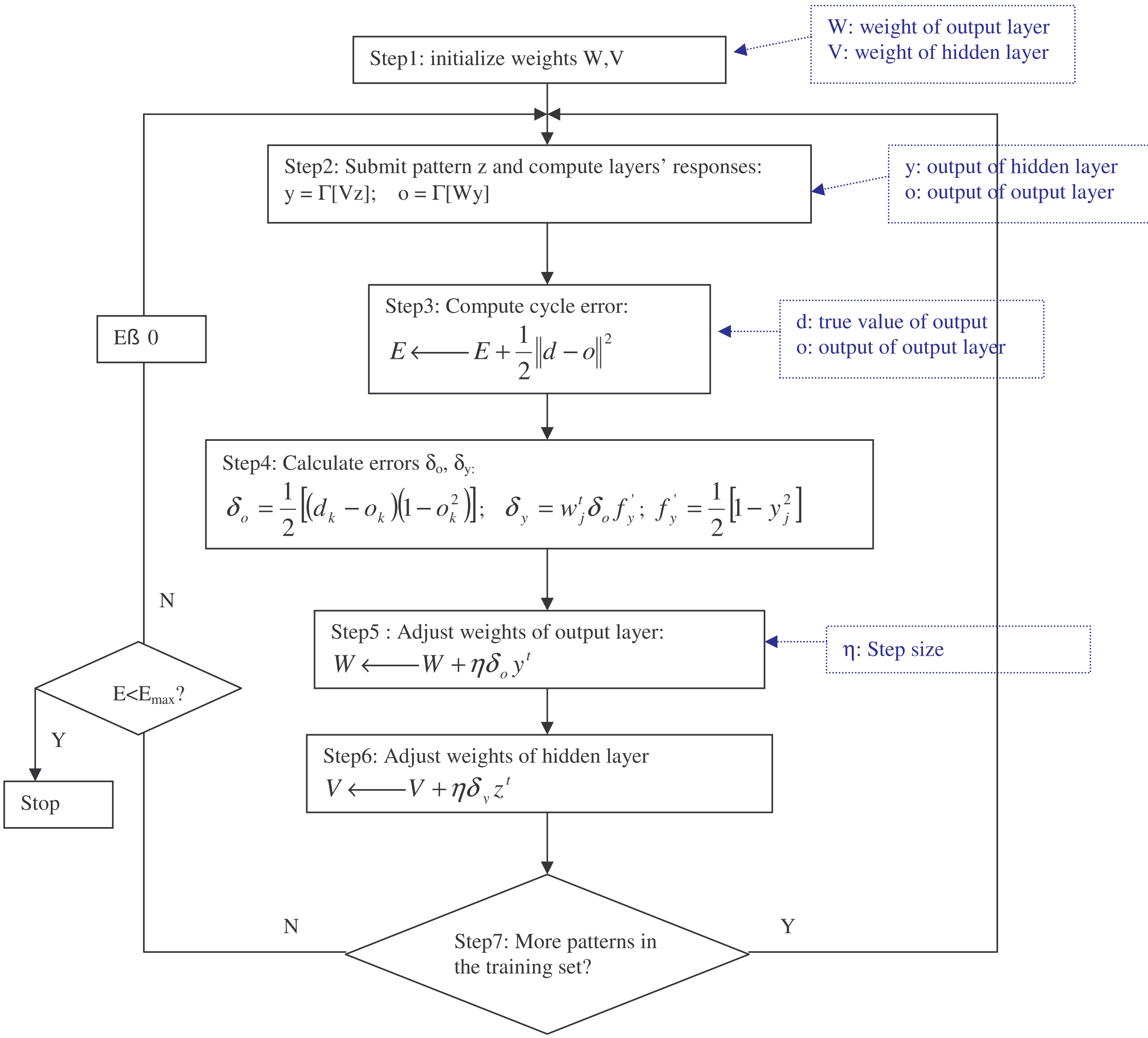


Figure 1.2 Back-propagation training algorithm flowchart

1.4 Development of ANN Models

1.4.1 Selection of Model Inputs and Outputs

The selection of appropriate input variables for proper mapping desired output variables is a very important step to ensure successful application of ANN models in hydrologic processes. Normally, it starts with a set of inputs that are KNOWN to affect the process, then add other inputs that are suspected of having a relationship in the process one at a time. A good understanding of the hydrologic system can lead to better choice of input variables for proper mapping. This will help in avoiding loss of information that may result if key variables are omitted, and also prevent inclusion of spurious inputs that tend to confuse the training process. A sensitivity analysis can be used to determine the relative importance of a variable (Maier and Dandy 1996) when sufficient data is available. The input variables that do not have a significant effect on the performance of an ANN can be trimmed from the input vector, resulting a more compact network. (Liong, Lim and Paudyal 2000).

1.4.2 Hidden Layer Size

The number of neurons in the input and output layers is determined by the number of input and output variables for a given system. The size of a hidden layer is one of the most important considerations when solving actual problems using multilayer feedforward networks. If there are fewer hidden layer neurons, there may not be enough opportunity for the neural network to capture the intricate relationships between indicator parameters and the nature of contaminating source. Too many hidden layer neurons not only require a large computational time for accurate training, but may also result in ‘overtraining’ (Brion et. al, 1999). A neural network is said to be ‘overtrained’ when the network focuses on the characteristics of individual data points rather than just capturing the general patterns present in the entire training set.

The following function can be recommended calculating the number of neurons in hidden layer: (http://www.neuralware.com/frequent.htm)

N = (Number of input + output)*(2/3) (1.2)

Where

N = the number of neurons in hidden layer.

1.4.3 Data Initialization

The contribution of an input/output will depend heavily on its variability relative to other inputs/outputs. If one input/output has a range of 0 to 1, while another input/output has a range of 0 to 1,000,000, then the contribution of the first input/output to the distance will be swamped by the second input/output. So it is essential to rescale the inputs/outputs so that their variability reflects their importance, or at least is not in inverse relation to their importance. One of the most useful ways to standardize inputs is mean 0 and standard deviation 1 method, which is shown as following:

$$mean_i = \frac{\sum X_i}{N} \tag{1.3}$$

$$Std = \sqrt{\frac{\sum \left(X_i^{\,2} - mean_i\right)}{N-1}} \tag{1.4}$$

$$S_i = \frac{X_i - mean_i}{Std} \tag{1.5}$$

X_i = value of the raw input/output variable X for the ith training case.

S_i = standardized value corresponding to X_i.

N = number of training cases

Once the output values are obtained from an ANN model, the actual values are obtained by the inverse transformation of the equation:

$$T = O \times Std + mean \tag{1.6}$$

T = Actual Output Value

O = output from ANN model.

1.4.4 Some Other Inspects

Initialization of weights and threshold values is an important consideration. The closer the initial guess is to the optimum weight space, the faster the training process. However, there is no way of making a good initial guess of weights, and they are initialized in a random fashion.

Learning rate affect the speed of convergence. If it is large, the weights will be changed more drastically, but this may cause the optimum to be overshot. If it is small, the weights will be changed in smaller increments, thus causing the system to converge more slowly with little oscillation. Normally it is in the range of 0.2 to 0.5.

1.4.5 Software Support

The software used for this study is NeuroSolutions (version 4.0). NeuroSolutions is developed by NeuroDimension Incorporated. The Microsoft Excel add-in in this version gives the ability to visually tag the data as Training, Cross Validation, Testing, and

Production. It also can train a neural network, and test the neural network's performance directly from within a Microsoft Excel worksheet.

1.5 Comparison of ANNs and Statistical Regression Methods

ANNs and statistical methods have similarities in many aspects. Both approaches are used to model a relationship between the dependent and independent variables. It has been observed that any generalized linear model can be mapped onto an equivalent single-layer neural network (Warner and Misra, 1996). For example, given the linear equation $y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n$, the independent variable (x_i) correspond to the inputs of the neural network and the response variable y to the output. The coefficients, β_i , correspond to the weights in the neural network.

There are differences between ANNs and statistical regression models. In regression models a functional form is imposed on the data. In the case of multiple linear regression this assumption is that the outcome is related to a linear combination of the independent variable. If this assumed model is not correct, it will lead to error in the prediction. So traditional linear models are simply inadequate when it comes to modeling data that contains non-linear characteristics. While ANN models do not assume any functional relationship and let the data define the functional form. Thus ANNs is extremely useful when there is no idea of the functional relationship between the dependent and independent variables. In effect, a neural network remains a “black box” that may produce useful results, but cannot be precisely understood. If you had an idea of the functional relationship, you are better of using a regression model.

Several authors have done comparison studies between ANNs against traditional linear and non-linear regression methods in rainfall-runoff processes (Anmala et al. 2000, Tokar and Johnson 1999, Zhang et al. 2000, Elshorbagy et al 2000). The results showed that the artificial neural network model generally performs better than other regression models in terms of accuracy and consistency.

1.6 Reference:

1. Anmala, J., Zhang, B., and Govindaraju, R.S (2000). "Comparison of ANNs and Empirical Approaches for Predicting Watershed Runoff" *Journal of Water Resources Planning and Management,* Vol. 126, No.3 May/June, 2000, page 156-166.
2. Atiya, A.F, El-Shoura, S. M, Shaheen, S.I., and El-Sherif, M. S (1999). "A comparison Between Neural-Network Forecasting Techniques – Case Study: River Flow Forecasting" *IEEE Transactions on Neural Network*, Vol. 10, No. 2, March 1999.
3. Brion, Gail Montgomery and Lingireddy, Srinivasa (1999). "A Neural Network Approach to Identifying Non-Point Sources of Microbial Contamination*." Water Resources.* Vol. 33 No. 14 pp. 3099-3106. 1999.
4. Elshorbagy A, Simonovic S.P and Panu U.S. (2000). "Performance Evaluation of Artificial Neural Networks for Runoff Predication" *Journal of Hydrologic Engineering. Vol.5,* No. 4, October, 2000. 424-427.
5. Haykin, S. (1994). Neural networks: a comprehensive foundation. Mac-Millan. New York.
6. Hsu, K., Gupta, H.V., and Sorooshian, S. (1995). "Artificial neural network modeling of the rainfall-runoff process." *Water Resources Research*, vol. 31, No. 10, Page 2517-2530.

7. Kuligowski, R.J., and Barros A.P (1998). "Experiments in short-term precipitation forecasting using artificial neural networks." *Mon. Wea. Rev*., 126, 470-482.
8. Liong, Shie-yui, Lim, W. and Paudyal, G. N., (2000). "River stage forecasting in Bangladesh: Neural network approach." *Journal of Computing in Civil Engineering*. Vol. 14, No. 1. January, 2000. Page 1-8.
9. Maier, H. R., and Dandy, G. C (1996). "The use of artificial neural networks for the prediction of water quality parameters." *Water Resources Research*, vol. 32, No. 4, 1013-1022 pp.
10. Rogers, L.L. and Dowla, F.U. (1994). "Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling." *Water Resources Research.* 30, 2: 457-481.
11. Sajikumar, N., and Thandaveswara, B.S. (1999). "A non-linear rainfall-runoff model using an artificial neural network." *Journal of Hydrology,* 216 (1999) 32-55.
12. Smith, J. and Eli, R.N. (1995). "Neural-network models of rainfall-runoff process." *Journal of Water Resources Planning and Management,* Vol. 121, No.6 November/December, 1995, page 499-508.
13. Tokar, A.S, and Johnson, P.A (1999). "Rainfall-Runoff Modeling Using Artificial Neural Networks". *Journal of Hydrologic Engineering.* Vol.4, No. 3, July, 1999. 232-239.
14. Warner, B., and Misra, M. (1996). "Understanding Neural Networks aas Statistical Tools" *The American Statistician,* 50, 284-293.
15. Zealand, C.M., Burn, D.H., and Simonovic S.P. (1999). "Short term streamflow forecasting using artificial neural networks." *Journal of Hydrology,* 214 (1999) 32-48.

16. Zhang, Bin and Rao S. Govindaraju.(2000) “Prediction of Watershed Runoff Using Bayesian Concepts and Modular neural networks.” *Water Resources Research*, vol. 36, No. 3, 753-762 pp, March 2000.

Chapter II: Application of Artificial Neural Networks for Rainfall-Runoff Processes

2.1 Abstract

An application of artificial neural network (ANN) was investigated for forecasting watershed runoff and stream flows/gage heights in two urban watersheds in Greensboro and Raleigh, North Carolina. Three ANN models, a Watershed Runoff Prediction (WRP)/Watershed Gage Height (WGH) model, a Stream flow Forecasting (SFF)/Stream Gage Height (SGH) model and a Historic Data Generation (HDG) model, were developed for forecasting watershed runoff at the collection point, stream flow/gage height at two downstream locations on the same stream, and generating the missing hydrologic data respectively. Available rainfall and stream flow/gage height data, including current and a preceding period of records, were employed as inputs to the WRP/WGH model. The model then performs forecasting of watershed runoff at the collection point with 15-minute increments. Similarly, available precipitation and stream flow/gage height data at the collection point, the upstream station, were used in the SFF/SGH model to forecast stream flows/gage height at two downstream locations of the same stream. The SFF/SGH model differs from a hydrologic routing model in such a way that it has been trained to recognize additional watershed runoff contributing to the downstream stations. ANN-hydrologic models are thus expected to provide real-time prediction of stream flows, which is particularly useful in urban watersheds for providing timely and early warning of flash floods. The HDG model was useful for generating missing historic stream flow record based on downstream flow data. ANN model performs better than conceptual/empirical models in terms of both low and high stream flow forecasts.

2.2 Introduction

In order to issue reliable and timely flood warnings, river flow forecasting must be performed on a real-time basis. Such timely warnings are particularly critical for urbanized areas confronting frequent occurrences of flash floods. Forecasting river flow after rainfall is a complex process that can be influenced by many factors including both watershed characteristics and precipitation patterns. There are two major modeling approaches for forecasting river flow: conceptual models and black-box models. Conceptual models are developed according to the physical processes in the hydrologic cycle. A large amount of precipitation information, basin characteristics and meteorological data would be required to process and synthesize the stream flows. For example, the SAC-SMA model, which is used in the National Weather Service river forecast system, requires 20 model parameters and 6 state variables (Tokar and Markus, 2000). Because of the high cost of data collection, and the complexity of the model calibration for the conceptual models, the black-box models become more attractive. Black-box models are developed from input-output measurements. Artificial Neural Networks (ANN) is one of the black-box models.

Another type of model has emerged from the application of artificial neural networks (ANN). The advantages of ANN models can be attributed to their resilience against distortion in input data and their capability of learning. They are extremely useful for solving problems without existing algorithmic solutions or with algorithms that are too complex to implement. Recent research on hydrologic applications of ANN range from the prediction of peak discharge and time-to-peak from a single rainfall event to the forecast of hourly or daily river stages or discharges (Smith 1995; Thirumalaiah and Deo 1998, 2000; Imrie et al. 2000; Liong et al. 2000). Zealand et al. (1999) concluded that the ANN approach provides a superior alternative to the time-series approach for developing input-output simulations and hydrologic models that do not require modeling of the internal processes of a watershed. Campolo et al. (1999) developed a neural network model to analyze and forecast the behavior of the river Tagliamento in Italy. The model was credited with the advantage of simplicity relative to those complex, physically based models.

It is noted that none of the aforementioned research attempted to generate river flows for advanced prediction with lead times in the order of minutes. This type of near-term or instantaneous forecasting of river flows is particularly important for small urbanized watersheds requiring timely warnings for flood mitigation and preparedness. Two urbanized watersheds in North Carolina were employed in this study. Rainfall data or the discharge data from an upstream rain gage station was serve as input to the ANN model and the stream flows for the stations along the stream could be predicated in 15-minute interval. The result shows the capability of ANN models in warning flood for small watershed on 15-min incremental intervals.

Research in the past has shown the successful application of ANN in the area of rainfall-runoff modeling (Smith and Eli, 1995, Hsu, K ,1995, and Anmala, J., 2000) and streamflows modeling. In order to develop an ANN model in a specific watershed for the purpose of stream flow forecast, significant amount of data are required for ANN training and testing processes. However, real-world data are noisy, distorted, and often incomplete. It is not uncommon to find that stream flow records at the point of interest may be too short for this application. The advantages of ANN learning capability and the tolerance of distorted data provide its potential application in generating or recreating the missing and /or incomplete historical data.

The overall objectives of this chapter are to demonstrate the applications of ANN models for river-flow predication and historical data generating. The following applications of ANN have been discussed:

(1) Stream flow/gage height forecasting at 15-min intervals up to a lead time of one hour, based on current and past rainfall and flow information at same station.

(2) Downstream flow forecasting at 15-min intervals up to a lead time of three hours, based on current and past rainfall and flow information at an upstream gauging station.

(3) Generating upstream historical stream flow based on the stream flow at a downstream gauging station.

2.3 ANN Applications on Stream Flow Predication

Normally there are two approaches used to forecast stream flow. The first one is obtained by modeling rainfall-runoff process; the other one is based on the hydrograph observed at upstream gauging site. Compared to these two approaches, rainfall-runoff process may provide greater lead time in some circumstances, the second approach tends to give more accurate results. (Bertoni et al, 1992).

Nachimuthu Karunanithi (1994) who tried to use neural network to do river flow predication, found that ANN model can give the prediction of daily flows at down stream station based on the flows observed at upstream and downstream stations. Empirical comparisons are performed between the predictive capability of the neural network models and the most commonly used analytic nonlinear power model in terms of accuracy and convenience of use; the result shows that the neural network models are more accurate predictors than the power model.

Cameron M. Zealand et al. (1999) studied the Winnipeg River System (20,000 km^2) in Northwest Ontario, Canada, and concluded that the ANN approach may provide a superior alternative to the time-series approach for developing input-output simulations and forecasting models in situations that do not require modeling of the internal structure of the watershed.

Thirumalaiah (1998) employed an ANN, which was trained by using three algorithms, namely error back propagation, cascade correlation, and conjugate gradient, to give daily stream flow forecasting based on data at both the given and upstream locations. Two years later (2000), the author introduced another options of ANN, which could give real-time forecasting of hourly flood at given stream flow gauge station with a warning time of a few hours based on on-line observations of rainfall at different rain gauge stations located in the same catchment.

Campolo et al. (1999) developed a neural network model to analyze and forecast the behavior of the river Tagliamento, in Italy. The model was found with the advantages of low cost and simplicity with respect to complex physical based models. The predictions of the model are accurate when the model is used with a 1-hour time horizon.

2.4 Description of the Study Sites

2.4.1 North Buffalo Creek in Greensboro, NC

The City of Greensboro, with a current population of 225,000, is located in North Carolina's Piedmont physiographic region. The North Buffalo creek originates in and flows northeastward through the central part of Greensboro. It also serves as a headwater urban stream of the Cape Fear River basin. There are three U.S. Geological Survey (USGS) gage stations along North Buffalo creek within the central Greensboro area. All of these stations include precipitation and stream flow measurements. USGS station number 02095181 is most upstream station of the North Buffalo creek. USGS 02095271 and USGS 02095500 are about 2.2 miles and 6.5 miles, respectively, downstream of USGS 02095181, see Figure 2.1. The time of concentration of each watershed was calculated by Kirpich's equation (equation 2.1), The general watershed information is given in Table 2.1.

Kirpich's Equation:

$$t_c = 0.0078(L^{0.77} / S^{0.385}) \tag{2. 1}$$

Where:

t_c = time of concentration (min.)

L = length of travel (ft)

S = slope (ft/ft)

Table 2.1 Watershed Characteristics

Parameter	Station-I	Station-II	Station-III
Station Number	USGS 02095181	USGS 02095271	USGS 02095500
Latitude	36°04'45''	36°05'52''	36°07'13''
Longitude	79°48'48"	79°46'55"	79°42'30''
Drainage area (mi^2)	9.55	14.2	37.1
Length of travel (ft)	20000	35000	70000
Slope (ft/ft)	0.005	0.005	0.005
Time of Concentration (Min.)	123	189	322

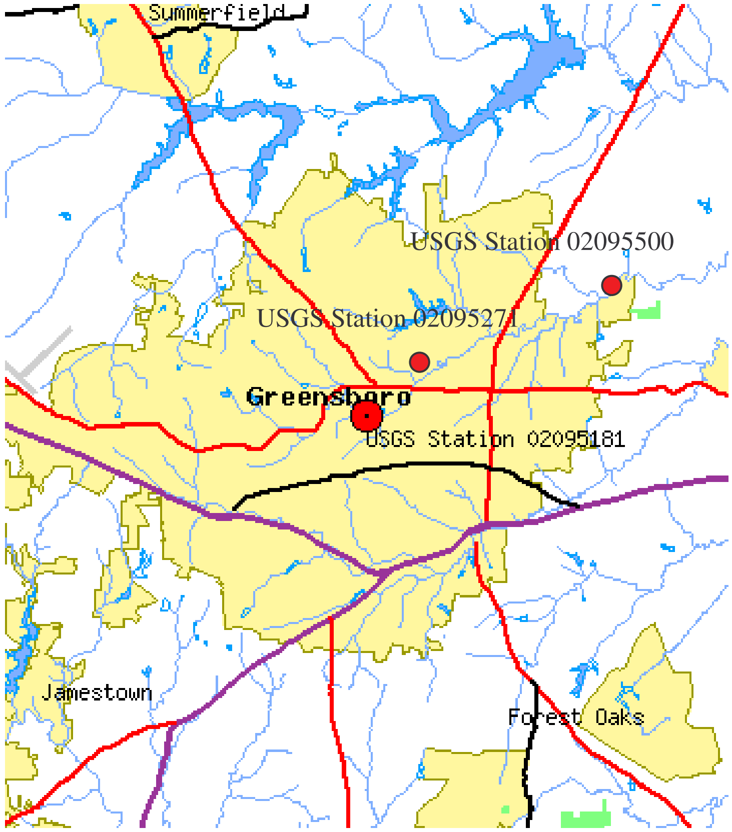


Figure 2.1 City of Greensboro, NC

2.4.2 Crabtree Creek in Raleigh, NC

There were four USGS stations selected along Crabtree Creek for this case study. The gage height data was obtained from USGS stations along Crabtree Creek. Station 0208731190(Crabtree-S1) was located at upstream of Crabtree Creek, Raleigh, as shown in Figure 2.2. The other two stations were less than 1 mile downstream of Crabtree-S1. The general information is shown in Table 2.2. The precipitation was obtained from USGS station 0208732885, which is 1 mile Northeast from USGS station 02087324.

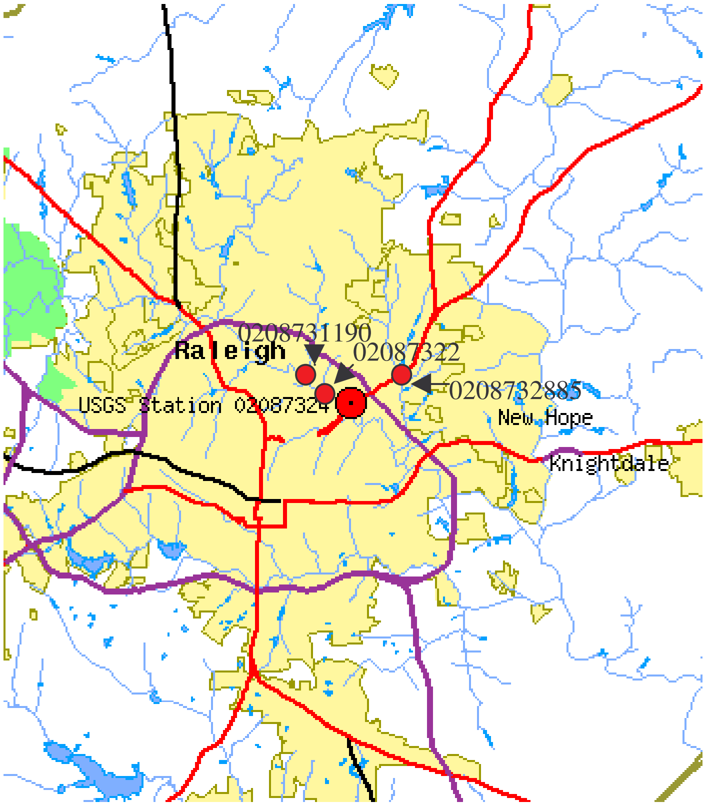


Figure 2.2 City of Raleigh, NC

Table 2.2 The Locations of Gage Stations at Crabtree Creek, Raleigh

Station Name	USGS Number	Latitude	Longitude	Distance from Upstream Station (mi.)	Drainage area (mi^2)
Crabtree-S1	0208731190	35°49'16"	78°37'34"	0	111
Crabtree-S2	02087322	35°48'57"	78°37'33"	0.23	119
Crabtree-S3	02087324	35°48'40"	78°36'43"	0.96	121

2.5 Data Record Classification

2.5.1 Data Selection at Greensboro Sites

Station-I started to record precipitation and stream flow in 6/99, station-II and III started in 8/98. During the beginning of their period of record, due to mechanical or equipment failure, some of the gage height and stream flow data were skewed or not available in their first year. Table 2.3 shows the record periods for these three stations.

Table 2.3 USGS Stations Record Periods in Greensboro, NC

Station	Precipitation	Stream Flow
Station-I	07/99 – 12/01	01/01 – 12/01
Station-II	09/98 – 12/01	09/98 – 12/01: some high flow missing in 99, 00
Station-III	09/98 – 12/01	09/98 – 12/01

While the intention of the neural network setting up is to forecast the downstream flow based on the upstream information, we need to have both upstream and downstream historical data available for ANN training and testing. Therefore ten storms in 2001 were selected for Greensboro, NC. Seven of them were used for ANN training and three of them were used for testing. The frequency of these storms were analyzed based on the Depth-duration-frequency table (see Appendix 1.17) developed by Dr. H.R. Malcom, Department of Civil Engineering, NC State University. The summary of these storms was list in Appendix-1.1a.

2.5.2 Data Selection at Raleigh Sites

The data record in Raleigh sites was summarized in Table 2.4. Twenty-five storms in the period of 1998 to 2001 were selected for Raleigh, NC. The frequency of these storms were analyzed based on the Depth-duration-frequency table, see Appendix 1.17. The summary of these storms was list in Appendix-1.1b.

Table 2.4 USGS Stations Record Periods in Raleigh, NC

Station	Precipitation	Gage Height
Crabtree-S1	-	6/97 – 12/2001
Crabtree-S2	-	6/97 – 12/2001
Crabtree-S3	-	1/96 – 12/2001
USGS0208732885	1/98 – 12/2001	-

2.6 Neural-Network Models Setting Up

2.6.1 ANN Model Development in Greensboro

Three types of ANN models were investigated in city of Greensboro. Initially, an ANN WRP model was employed to investigate the rainfall-runoff relationship at all three stations along North Buffalo Creek. These three watersheds have time-of-concentration, Tc, of approximately 2, 3, and 5 hours respectively. Storms occurring within a period equal to Tc would exhibit the greatest influence on stream flows. Consequently, precipitation, P, of the preceding 2, 3 and 4 hours (at 15-minute interval) and stream flows, Q, of the preceding 1hours (at 15-minute interval) at each station were taken as inputs to the input layer. Stream flows from the preceding one hour help provide the base flow information prior to the onset of a storm event. The output neurons yield predicted watershed runoff or stream flows, Q', for station-I, II and III at 15-min interval up to 1 hour, 1.5 hour and 3 hours lead times respectively, as summarized in Table 2.5. This approach is therefore significantly more useful than conventional hydrologic modeling in which current stream flows are calculated from the rainfall records available after the occurrence of the storm. Training of the WRP model included employing rainfall and runoff data from historical storm events and treating the future flow data at 15-min intervals as predicted outputs for determining the optimal weighting factors.

A second ANN SFF model was developed for forecasting stream flows at downstream locations based on precipitation and stream flows of the upstream station (Station I). The upstream station was chosen as the reference station to generate forecasted flows for the two downstream stations (Station II and Station III). The average travel times of peak flow from the reference station to the first and second downstream stations are approximately one hour and three hours, respectively. Preliminary investigation indicated the existence of a reasonable correlation between stream flows at the upstream station and those at either one of the downstream stations when the upstream data was lagged by appropriate time intervals. If the time lags were too long, part of the upstream data would exhibit no relation to the downstream flows. It would also prolong training time and reduce model performance. At each time step, data input to ANN include the preceding 6-hr record of precipitation and stream flows at the upstream station. A 15-minute time step was chosen to be consistent with the USGS data format. The predicted stream flows at the first and second downstream stations will be one hour and three hours ahead in real time, see Table 2.6. The SFF model differs from a hydrologic routing model in such a way that it has been trained to recognize additional watershed runoff contributing to the downstream stations.

Due to mechanical or equipment failure, the stream flow historical data at upstream stations (Station-I and II) is missed and/or incomplete. A Third ANN HDG model was developed for generating the upstream incomplete historical data using stream flow data at downstream station (Station-III). A sensitivity analysis has been made to select time lags of downstream flow as inputs. Different periods of stream flow at Station III were tested to identify its correlation with upstream Station-I, see Table 2.8. The results of analysis showed the 4 hours lag time gave the best correlation of the stream flow at the downstream and upstream stations, see Figure 2.3. The input and output selection of HDG model was shown in Table 2.7.

Table 2.5 ANN Application of Watershed Runoff Prediction (WRP) Model

Input	Station I: $P(t)$, $P(t - k\Delta t)$; $k = 1$ to 8 Station II: $P(t)$, $P(t - k\Delta t)$; $k = 1$ to 12 Station III: $P(t)$, $P(t - k\Delta t)$; $k = 1$ to 16 $Q(t)$, $Q(t - j\Delta t)$; $j = 1$ to 3
Output	Station I : $Q'(t + j\Delta t)$; $j = 1$ to 4 Station II : $Q'(t + j\Delta t)$; $j = 1$ to 6 Station III : $Q'(t + j\Delta t)$; $j = 1$ to 12
Computational time step	$\Delta t = 15$ minutes
Network Configuration	*Station I: 8-6-1 *Station II: 12-8-1 *Station III: 16-12-1
Iterations	5000
Training Time (min)	10-20 min

Number of neurons in input layer – Number of neurons in hidden layer – Number of neurons in output layer

Table 2.6 ANN Application of Stream Flow Forecasting (SFF) Model

Input: upstream station	P(t), P(t – jΔt); j = 1 to 23 Q(t), Q(t – jΔt); j = 1 to 23
Output 1st down stream station 2nd down stream station	 Q'(t + 60) Q'(t + 180)
Computational time step	Δt = 15 minutes
Network Configuration	*48-32-1
Iterations	5000
Training Time (min)	40 min

Number of neurons in input layer – Number of neurons in hidden layer – Number of neurons in output layer

Table 2.7 ANN Application of Watershed Runoff Prediction (WRP) Model

Input:	S_{III}(t+Δt),…S_{III}(t+iΔt); i= 1 to 16. S_{III}: Stream flow at Station III.
Output:	S_{II}(t) : Stream Flow at Station II. S_I(t): Stream Flow at Station I.
Computational time step	Δt = 15 minutes
Network Configuration	*16-12-1
Iterations	5000
Training Time (min)	15 min

Number of neurons in input layer – Number of neurons in hidden layer – Number of neurons in output layer

Table 2.8 Sensitivity Analysis of ANN model

Time Lag (hr)	Input (Stream Flow at Station III)	Output (Stream Flow at Station I)
1	S(t+Δt),…S(t+iΔt); i = 1 to 4	S(t)
2	S(t+Δt),…S(t+iΔt); i = 1 to 8	S(t)
3	S(t+Δt),…S(t+iΔt); i = 1 to 12	S(t)
4	S(t+Δt),…S(t+iΔt); i = 1 to 16	S(t)
5	S(t+Δt),…S(t+iΔt); i = 1 to 20	S(t)
6	S(t+Δt),…S(t+iΔt); i = 1 to 24	S(t)

Δt = 15 minutes.

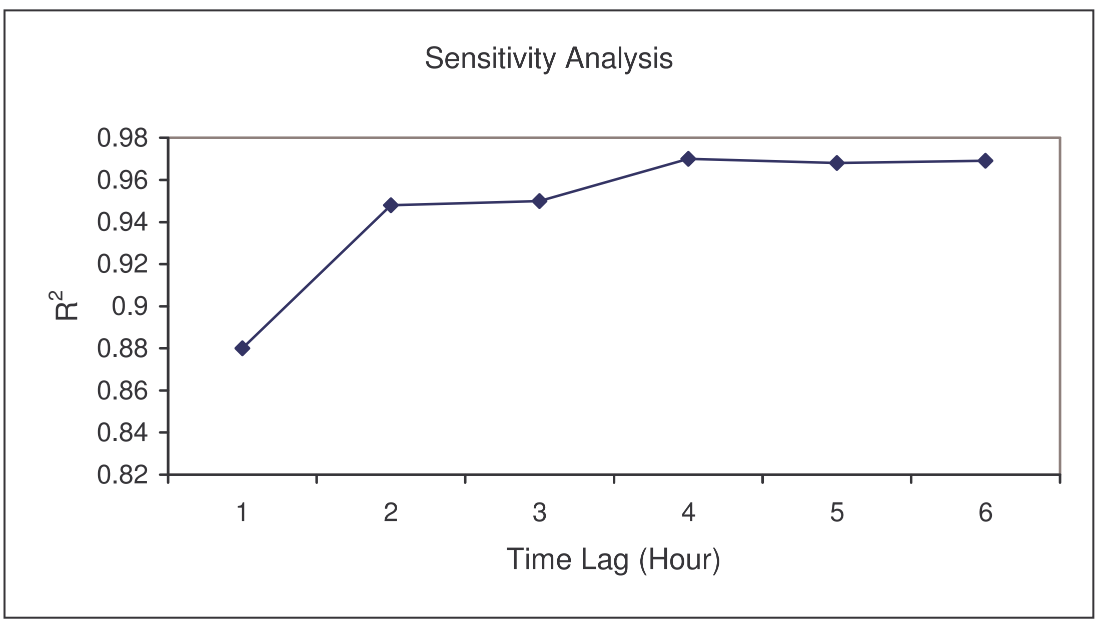


Figure 2.3 R^2 Values Using Different Downstream Flow as Input

2.6.2 ANN Model Development in Raleigh

Two applications of ANN-hydrologic models have been investigated for predicting stream gage heights. An ANN Watershed Gage Height (WGH) model was developed to predict stream gage heights at the watershed outlet. This model was initially trained to determine the optimal duration of preceding rainfall periods as model input. After training, results of model testing at Crabtree-S1, as shown in Figure 2.4, implying a 4-hour preceding period or lag time yield acceptable R^2 between predicted and observed gage heights. A four-hour lag in precipitation data as input was considered acceptable for subsequent model runs. A one-hour lag time of gage height data, including the current reading, was employed as additional input to the ANN model neurons. Crabtree-S2 and S3 are only 0.23 and 0.96 mile downstream of Crabtree-S1; similar lag time requirements for precipitation and gage height input data were applied. With proper training, the WGH model can predict gage height readings at the watershed outlet with a forecasting period ranging from 15 minutes to 2 hours from the present time. Details of input and output selections are provided in Table 2.9.

A second Stream Gage Height Forecasting (SGH) model was developed as an extension of the WGH model for forecasting gage heights at downstream locations. At each time step, input to the SGH model include precipitation data (4-hour lag time) at USGS 0208732885 and gage height data (2-hour lag time) taken from a reference upstream station (Crabtree-S1). A 15-minute time step was chosen to be consistent with the USGS data format. The predicted gage heights at the first and second downstream stations will have a maximum forecast period of 2 hours ahead in real time at 15 min intervals, see Table 2.9. The SGH model differs from a hydrologic routing model in such a way that it has been trained to recognize additional watershed runoff contributing to the downstream stations.

Table 2.9 Inputs and Outputs Selections for ANN Models at Crabtree Creek

Watershed Gage Height (WGH) Model	
Input	P(t), P(t - iΔt); i = 1 to 15 ; Q(t), Q(t - jΔt); j = 1 to 3
Output	Q'(t + kΔt); k = 1 to 8
Computational time step	Δt = 15 minutes
Network Configuration	20-14-1
Iterations	5000
Training Time (min)	20 min
Stream Gage Height (SGH) Model	
Input: upstream station	P(t), P(t – iΔt); i = 1 to 15 Q(t), Q(t – jΔt); j = 1 to 7
Output 1st down stream station 2nd down stream station	 Q'(t + kΔt); k = 1 to 8 Q'(t + kΔt); k = 1 to 8
Computational time step	Δt = 15 minutes
Network Configuration	24-16-1
Iterations	5000
Training Time (min)	20 min

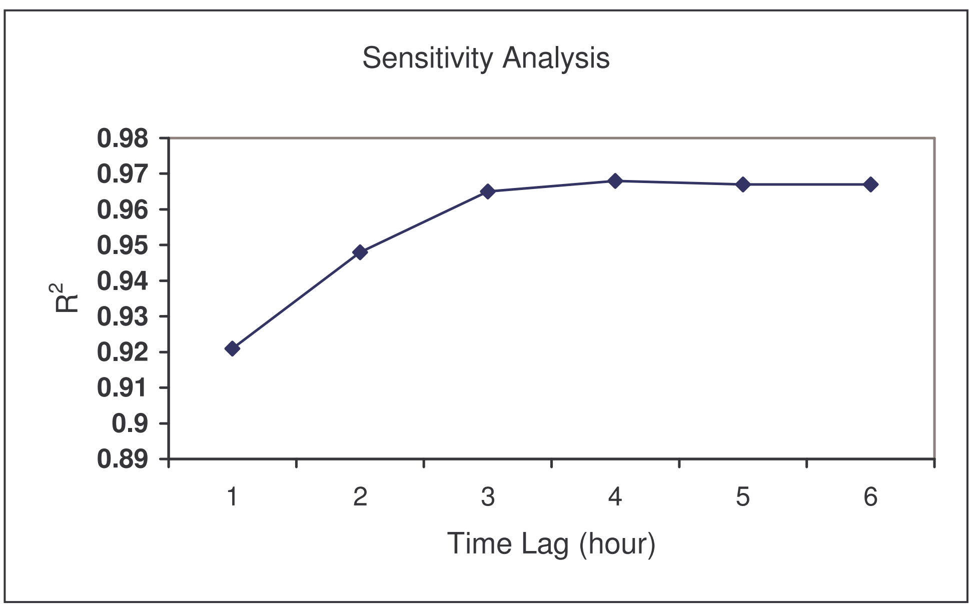


Figure 2.4 Sensitivity Analysis of ANN Gage Height Predication Model in Raleigh

2.6.3 Application of HEC-HMS model

A HEC-HMS model is developed for Station II at North Buffalo Creek in Greensboro, NC. The detail information of North Buffalo watershed characteristics was not available; the best estimates of the parameters were made based on four storms on 1/18/01, 2/17/01, 3/3/10 and 3/15/01. Table 2.10 shows the parameters used for simulation. The HEC-HMS model was tested using the storms on 3/20/01 and 3/30/01. The summary of these storms was shown in Appendix 1.8.

Table 2.10 Watershed Parameters used for HEC-HMS Simulation

Parameters	
Method	SCS Curve No. Method
Initial Loss (in.)	0.2
% Impervious	25
SCS Curve No.	70
SCS lag time (min.)	150
Base Flow (cfs)	7
Drainage Area (sq. mi)	14.2

2.6.4 Application of Empirical Models

There are two empirical equations that are selected to develop rainfall-runoff models.

Empirical model 1:

$$S(t+i\Delta t)= a + b*S(t) + c*R(t) + d*R(t-\Delta t) +\ldots+g*R(t-j\Delta t) \quad (2.\ 2)$$

Empirical model 2:

$$S(t+i\Delta t) = a + b*S(t)^c + d*R(t)^e + f*R(t-\Delta t)^g +\ldots + l*R(t-j\Delta t)^m \quad (2.\ 3)$$

Where:

Δt = 15 minutes.

S(t+iΔt) : Stream flow predication at iΔt lead time. i = 1 to 4.

S(t): Current stream flow.

R(t-jΔt): Previous precipitation at lag time of jΔt. j = 1 to 12.

The parameters of the models are calibrated by minimizing MSE (mean squared error) between observed and predicated stream flow.

2.6.5 ANN Model Performances Tests

The accuracy of model results is evaluated based on the correlation coefficient (R^2), the mean squared error (MSE) and mean relative error (MRE) and coefficient. R^2 measures the degree to which two variables are linearly related. The MSE and MRE provide different types of information about the predictive capabilities of the model. MSE and MRE are defined by equation 2.4 & 2.5. During flow simulation, the higher of the flow rate, the bigger absolute error it is likely to be given by the model; the major part of MSE is attributed to the top 20% biggest absolute errors. It indicates that the MSE is a good measure for indicating goodness of fit at the high flows. The MSE is an absolute value and not referenced to any benchmark point, not comprehended according to a predesigned threshold to indicate the significance of the MSE value or even the significance of the difference between two calculated MSEs of different models. (Elshorbagy 2000). During a storm, most part of flow is moderate flow, which gives the major contribute to MRE, and the MRE provides a more balanced perspective of the goodness of fit at moderate flows. (Nachimuthu Karunanithi, 1994).

$$MSE = \frac{\sum_{i=1}^{N}(D_i - D'_i)^2}{N} \qquad (2.\ 4)$$

$$MRE = \frac{\sum_{i=1}^{N} \left| \frac{D_i - D'_i}{D_i} \right|}{N} \quad (2.5)$$

Where

D_i = the measured (actual) value;

D_i' = the estimated value;

N = number of observations.

2.6.6 Hypothesis Test: Chi-Squared Goodness-of-Fit Test

The chi-squared test is a test of significance based on the chi-squared statistic. The statistic is derived by the sum of squares of independent standard normal varieties. The details are shown as following:

Definition The chi-square test is defined for the hypothesis:

H_0: The data follow a specified distribution.

H_a: The data do not follow the specified distribution.

Test Statistic: For the chi-square goodness-of-fit computation, the data are divided into *k* bins and the test statistic is defined as:

$$X^2 = \sum_{i=1}^{l} \frac{(O_i - E_i)^2}{E_i} \quad (2.6)$$

where O_i is the observed frequency and E_i is the expected frequency. A large value of this statistic indicates a poor fit.

2.7 Results and Discussion

2.7.1 ANN Application in North Buffalo Creek, NC

2.7.1.1 ANN WRP Model: Watershed Runoff Predication

The ANN WRP model was trained with seven storm events and tested by three storm events that occurred in the year of 2001, see Appendix 1.1a. One major features of the ANN model was that the model could give stream flow predication at different lead times based on current and past stream and precipitation information. Figure 2.5 to Figure 2.7 showed the results of stream forecasting at these stations for the storms on October 14, 2001. The precipitation and stream flow at different lag time before 15:30 were served as the inputs to ANN model. The points afterward are the outputs from ANN model, which give the stream prediction for station-I, II and III at 15-min interval up to 1 hour, 1.5 hour and 3 hours lead times respectively.

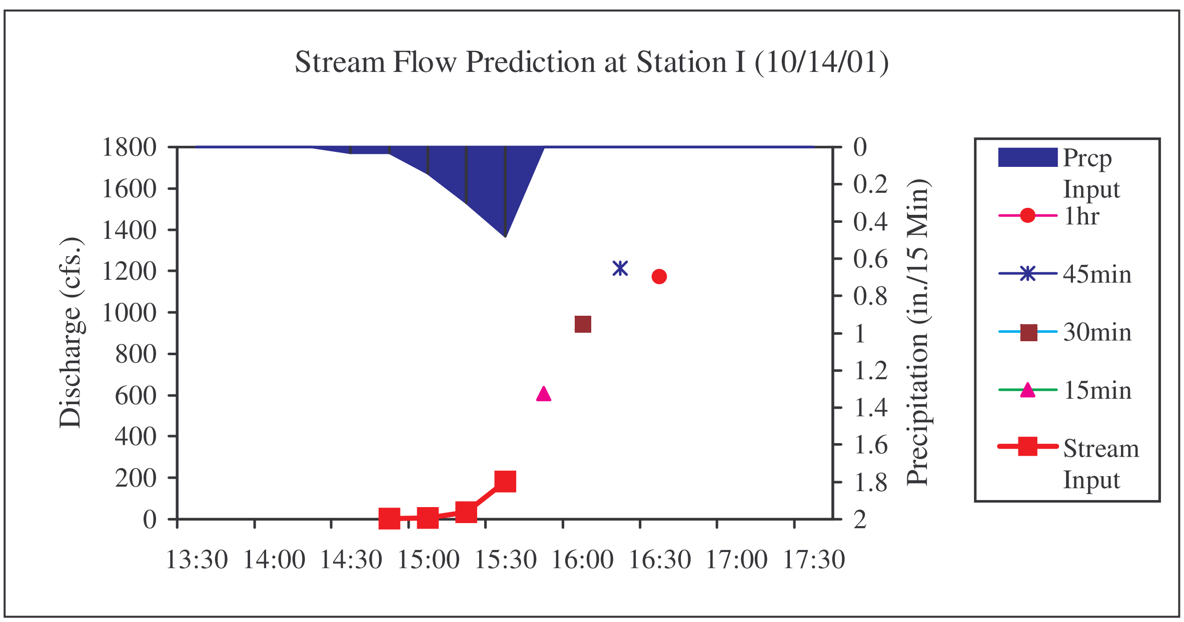


Figure 2.5 Stream Flow Prediction at Station-I Using ANN Model

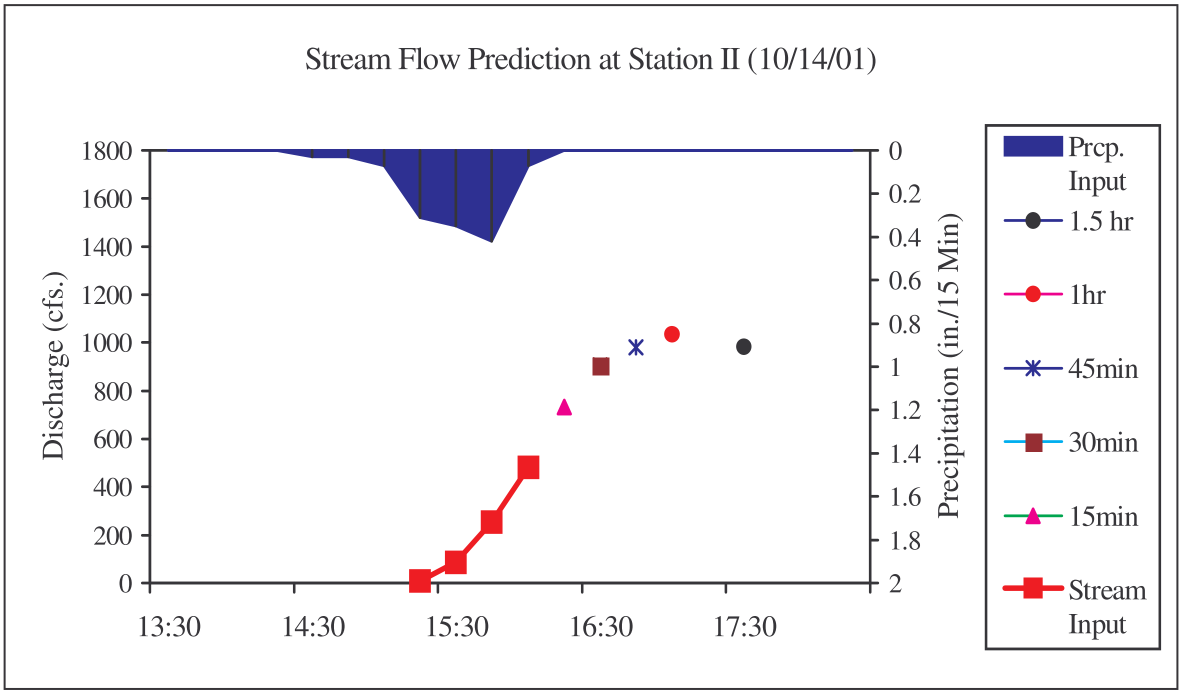


Figure 2.6 Stream Flow Prediction at Station-II Using ANN Model

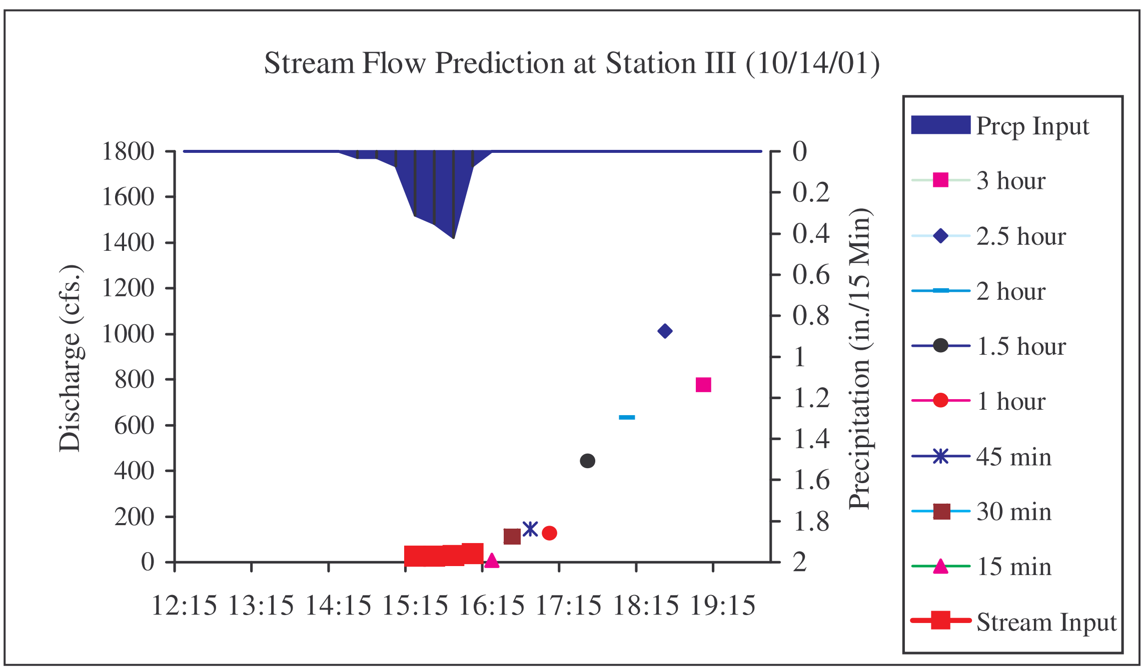


Figure 2.7 Stream Flow Prediction at Station-III Using ANN Model

The model output obtained from testing, as compared with observed discharges, were shown in Figure 2.8, 2.9 and 2.10. More results are given in appendix 1.2, 1.3 and 1.4 based on the studies on other storms in 2001. Comparison of observed peak flow and ANN output in Station-I was shown in Figure 2.11; Chi-squared goodness-of-fit analysis of both training and testing data in Table 2.11 indicated that the peak flow estimation had the linear distribution compared with observed peak flow at significance level of 0.95. Similar observations were given for Station-II and III, see Appendix 1.5. Figure 2.12, 2.13 and 2.14 compared model performance with respect to MRE, MSE and R^2 values. It can be seen from these figures that hydrographs generated by employing the 15-minute future time-step scenario exhibit the highest R^2 and least MRE and MSE values. In general, using a shorter time increment of future prediction should yield more reliable results for both low and high stream flows. When the prediction time-increment is reduced to zero, the ANN WRP model would be equivalent to a physically based watershed simulation model. A prolonged time-increment of prediction may result in reduced model reliability. However, the relaxation of model reliability within a tolerance limit is acceptable provided the objectives of early warning can be reasonably achieved. The current study was carried out with a small number of data sets for training and testing. It nevertheless indicates the ANN's model ability to predict watershed runoff from the North Buffalo creek watershed with an early warning period commencing from the current time and extending to 1, 1.5 and 3 hours in advance for Station I, II and III respectively. Furthermore, the predicted outputs can be continuously improved and the entire model can be retrained when additional or new data become available. Nowadays, many USGS stations are equipped with telecommunication capability for instantaneous retrieval of field recorded rainfall or stream stages. These real-time data can be obtained in real time and at a very fast rate to facilitate running an ANN model.

Table 2.11 Chi-Squared Goodness-of-Fit Analysis of Peak Flow Estimation at Station-I, Greensboro

Null Hypothesis H_0:	Null hypothesis Ho : The peak flow estimation has linear distribution compared with observed peak flow.
Alternate Hypothesis H_1:	The peak flow estimation does not have the specific distribution
Level of Significance α:	0.05
Critical Region:	$X^2 \geq X^2_{9,0.05} = 16.9$
Result:	$X^2 = 1.5 < X^2_{9,0.05} = 16.9$
Decision:	The null hypothesis is not rejected. It can be concluded that the peak flow estimation has linear distribution compared with observed peak flow.

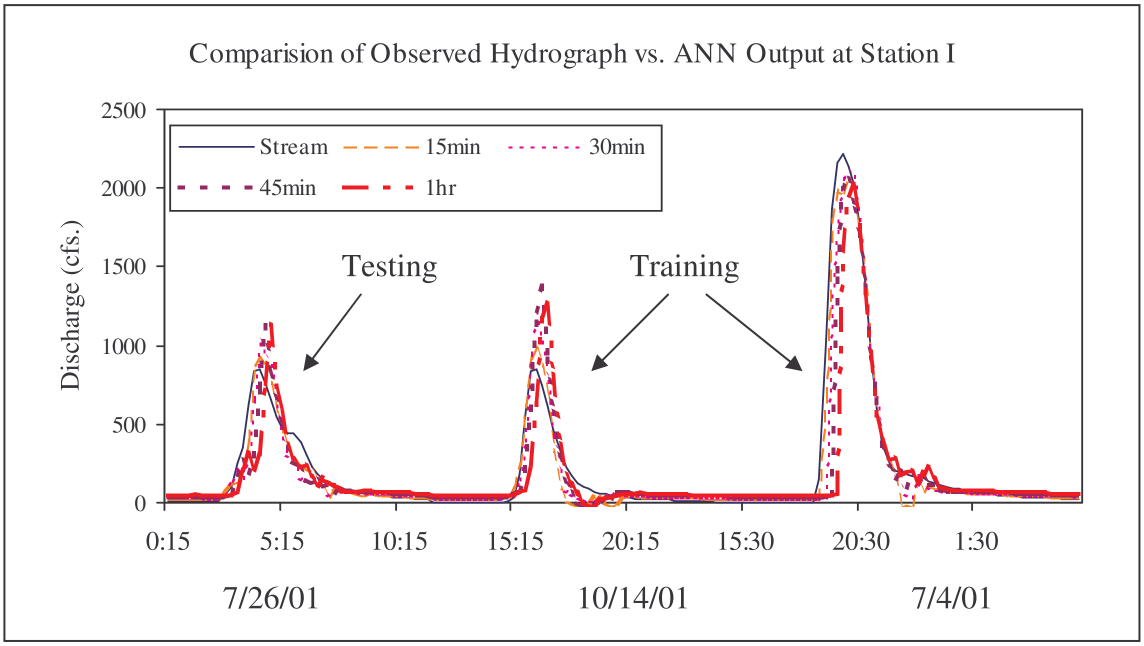


Figure 2.8 Output Hydrographs vs. Observed Discharge at Station-I

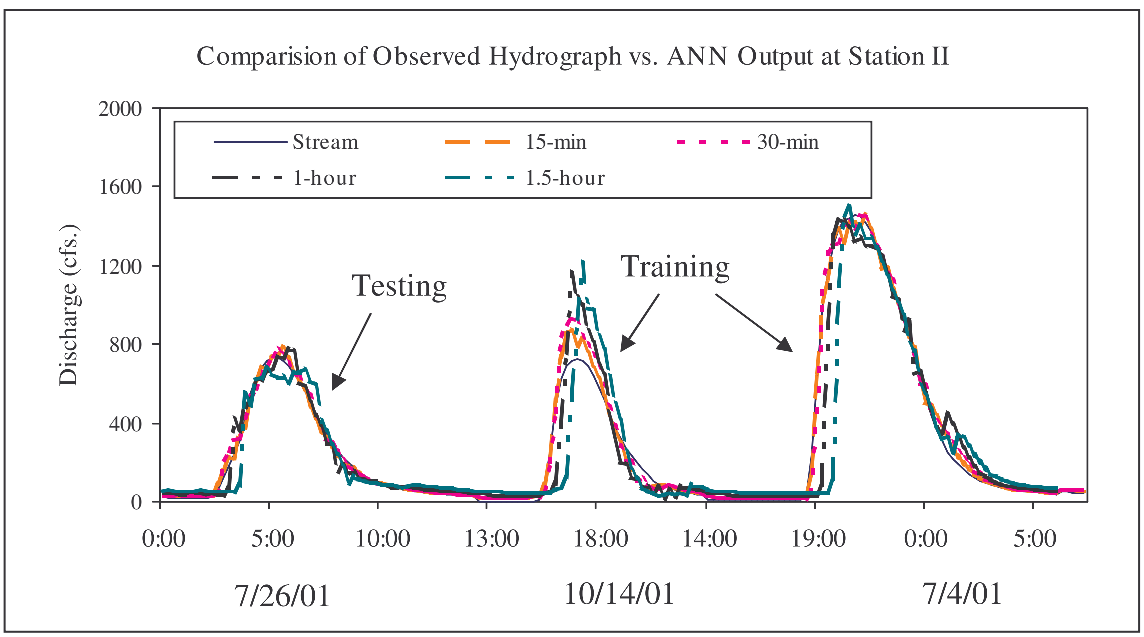


Figure 2.9 Output Hydrographs vs. Observed Discharge at Station-II

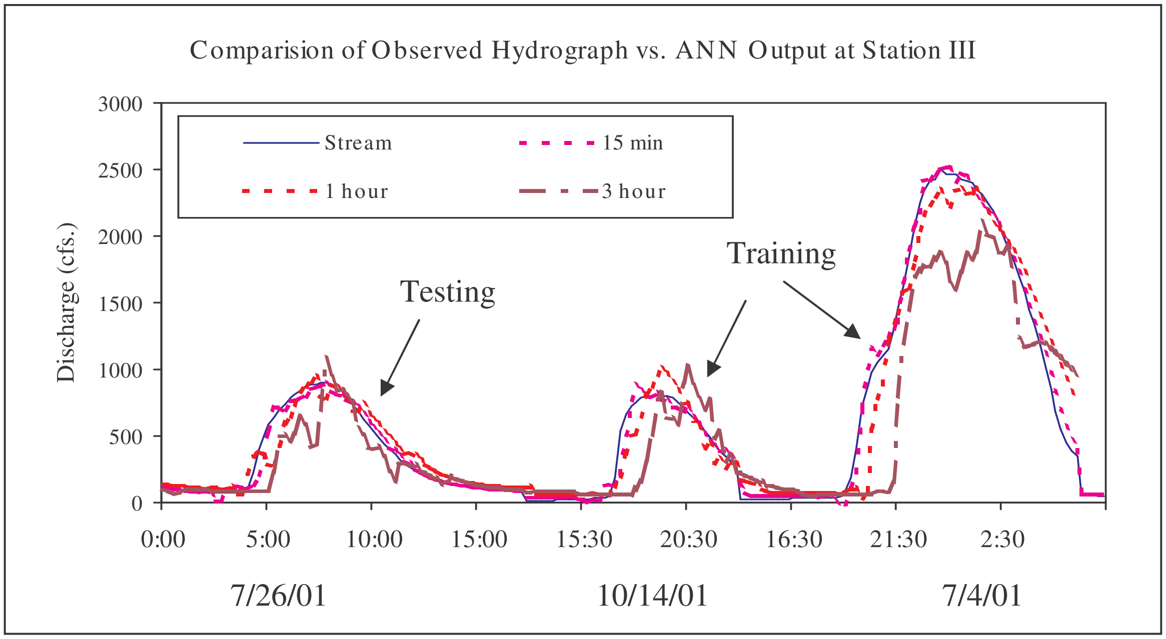


Figure 2.10 Output Hydrographs vs. Observed Discharge at Station-III

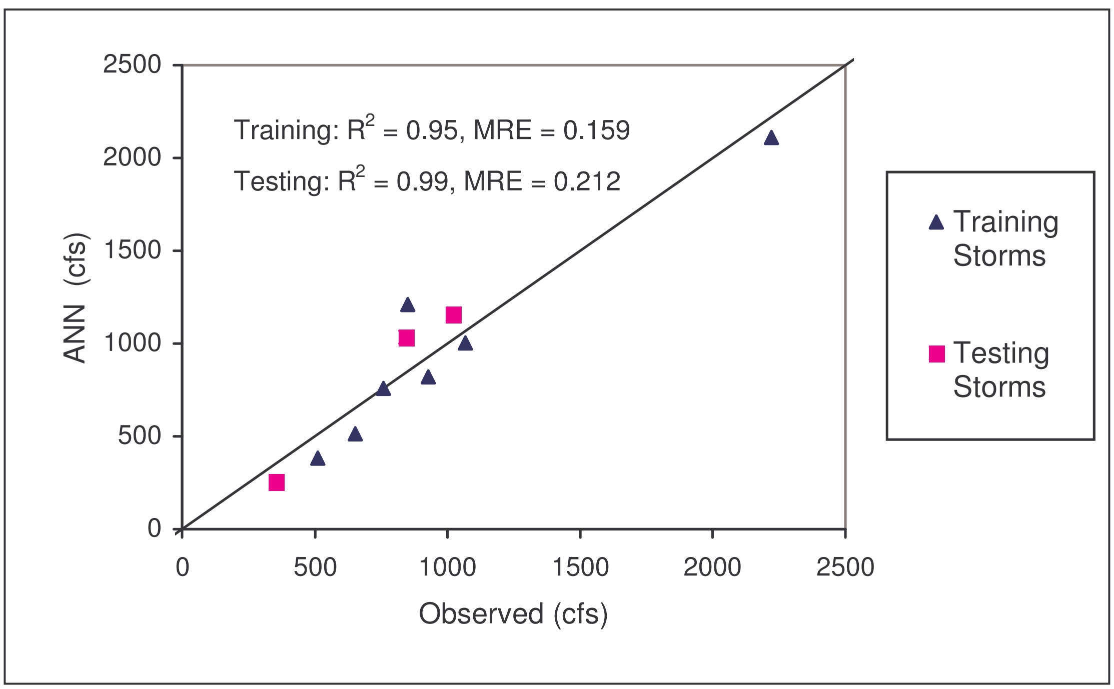
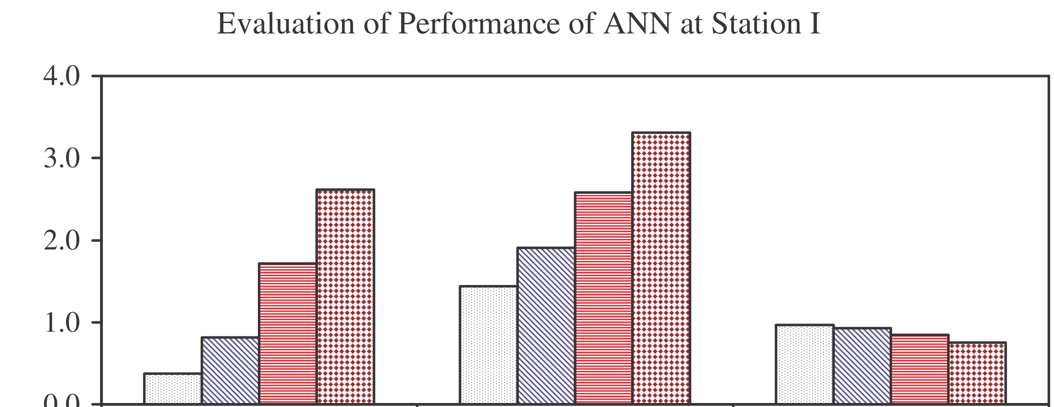
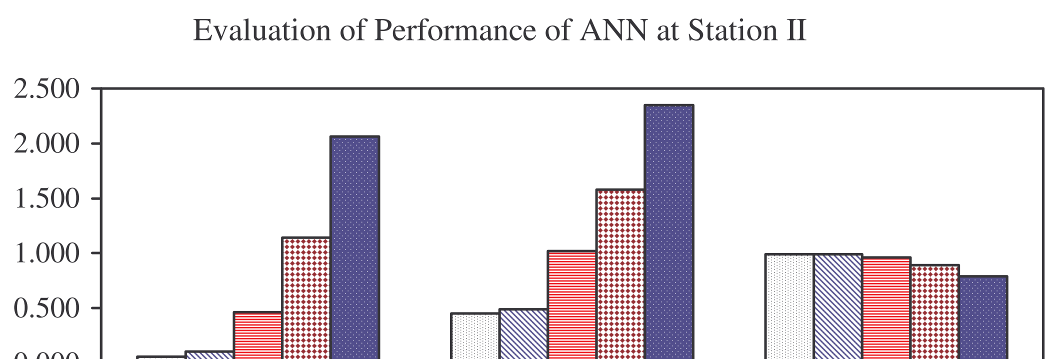


Figure 2.11 Observed-ANN Output Peak Flow Scatter plots at Station-I



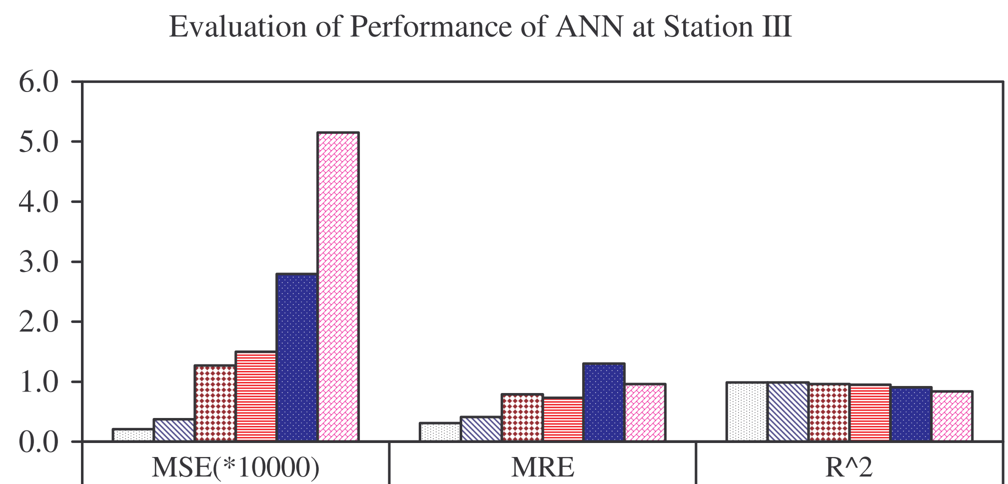
	MSE (*10000)	MRE	R^2
15-min.	0.377	1.440	0.969
30-min.	0.817	1.910	0.931
45-min.	1.713	2.580	0.848
1-hour	2.615	3.310	0.756

Figure 2.12 Evaluation of the Performances of ANN model at Station-I



	MSE (*10000)	MRE	R^2
15 min.	0.058	0.450	0.990
30 min.	0.106	0.490	0.990
1 hour	0.464	1.020	0.960
1.5 hour	1.141	1.580	0.890
2 hour	2.063	2.350	0.790

Figure 2.13 Evaluation of the Performances of ANN model at Station-II



	MSE(*10000)	MRE	R^2
15 min.	0.206	0.310	0.990
30 min.	0.372	0.410	0.990
1 hour	1.269	0.790	0.960
1.5 hour	1.498	0.730	0.950
2 hour	2.796	1.300	0.910
3 hour	5.152	0.960	0.840

Figure 2.14 Evaluation of the Performances of ANN model at Station-III

2.7.1.2 ANN SFF Model: Stream Flow Forecast

The same storms used in the previous watershed runoff study were employed in this phase of study. It can be seen from Figures 2.15 and 2.16 that a reasonably good match of the predicted and observed peak flows were obtained for both downstream stations. More hydrographs are given in appendix 1.6 and 1.7. The chi-squared goodness-of-fit analysis for both training and testing data in Table 2.12 indicated that the peak flow estimation had the linear distribution compared with observed peak flow at significance level of 0.95 for both downstream stations. At station-III, the forecast discharge hydrograph did not match the observed hydrograph as well as station-II. It may because that station-III is farther away from upstream station than station-II, and therefore there are some other factors that have potential effects on discharges instead of upstream precipitation and stream flow. But it still gave satisfactorily forecast at high flows. In both cases, ANN tends to give overestimate of low-flow period or base flow. But it is not critical for the purpose of flood warning system. The prediction reliability can be improved with more data sets to be employed for training when the model will be further tested for practical application. The ANN SFF model can also be continuously updated and retrained as additional meteorological and hydrologic data are made available. It is important to choose a lead time of prediction that is equal to the average travel time of flood peaks.

Table 2. 12 Chi-Squared Goodness-of-Fit Analysis of Peak Flow Estimation of SFF Model, Greensboro

Null Hypothesis H_0:	Null hypothesis Ho : The peak flow estimation has linear distribution compared with observed peak flow.
Alternate Hypothesis H_1:	The peak flow estimation does not have the specific distribution
Level of Significance α:	0.05
Critical Region:	$X^2 \geq X^2_{9,0.05} = 16.9$
Result:	Station II: $X^2 = 1.5 < X^2_{9,0.05} = 16.9$ Station III: $X^2 = 2 < X^2_{9,0.05} = 16.9$
Decision:	The null hypothesis is not rejected. It can be concluded that the peak flow estimation has linear distribution compared with observed peak flow.

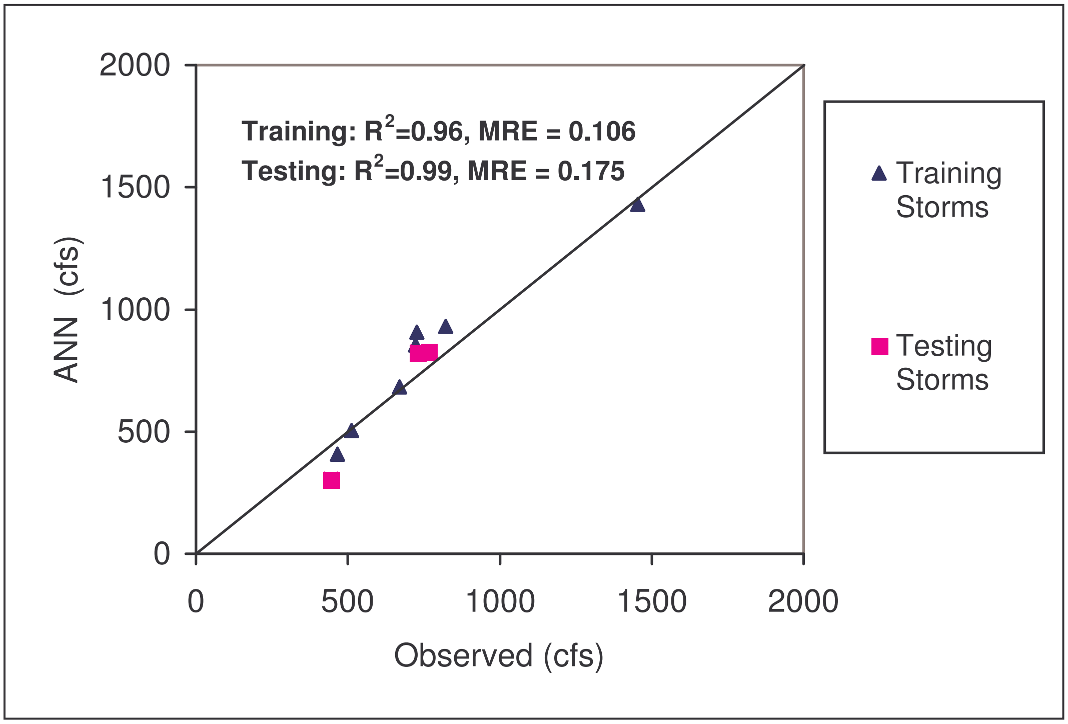


Figure 2.15 Observed-ANN Output Peak Flow Scatter plots at Station-II

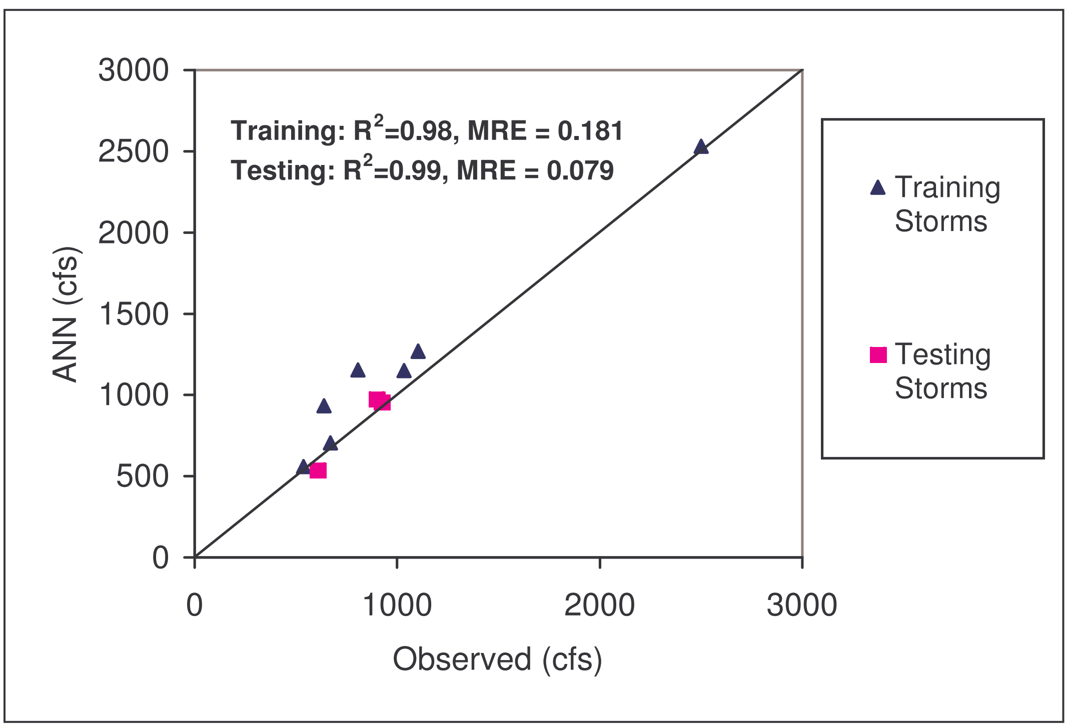


Figure 2.16 Observed-ANN Output Peak Flow Scatter plots at Station-III

2.7.1.3 ANN HDG Model: Historical Data Generation Model

This model was trained and tested using the same storm as listed in Appendix 1.1a. The output hydrographs from ANN model were compared with observed discharge hydrographs shown in Figure 2.17 and 2.18. The graphical comparisons show that the ANN outputs and observed discharges match very well at both low and high flows. ANN model can be expected to give reliable missing historical data recovered at upstream stations based on downstream information.

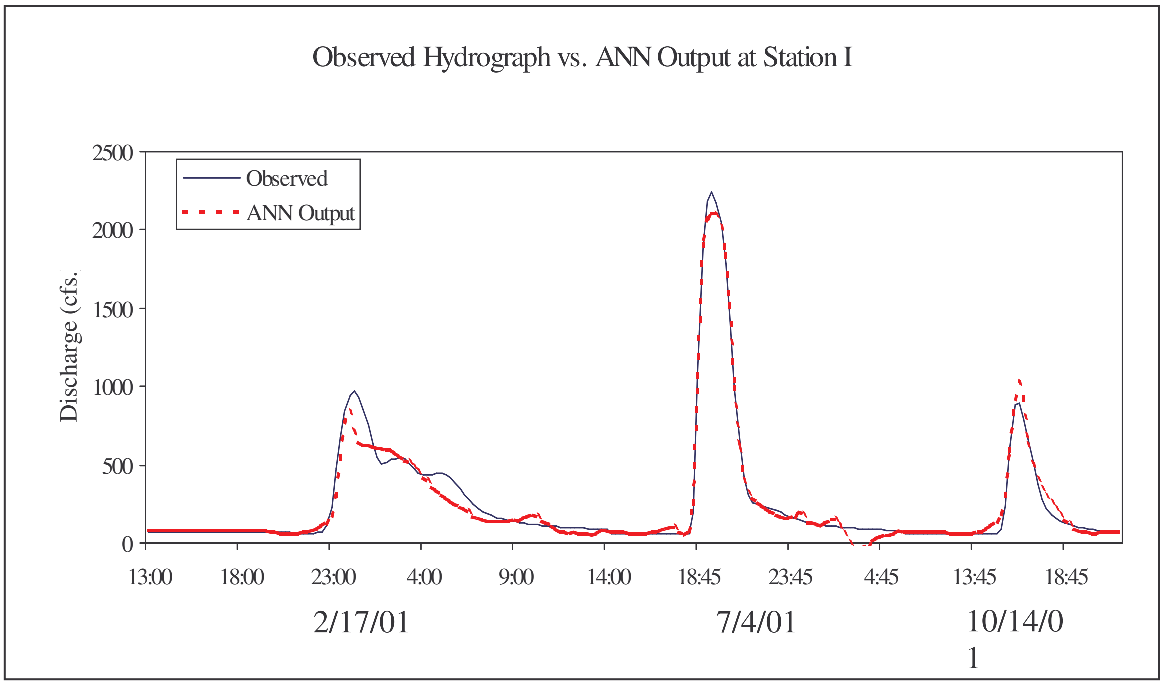


Figure 2.17 Observed Hydrographs vs. ANN Output at Station-I

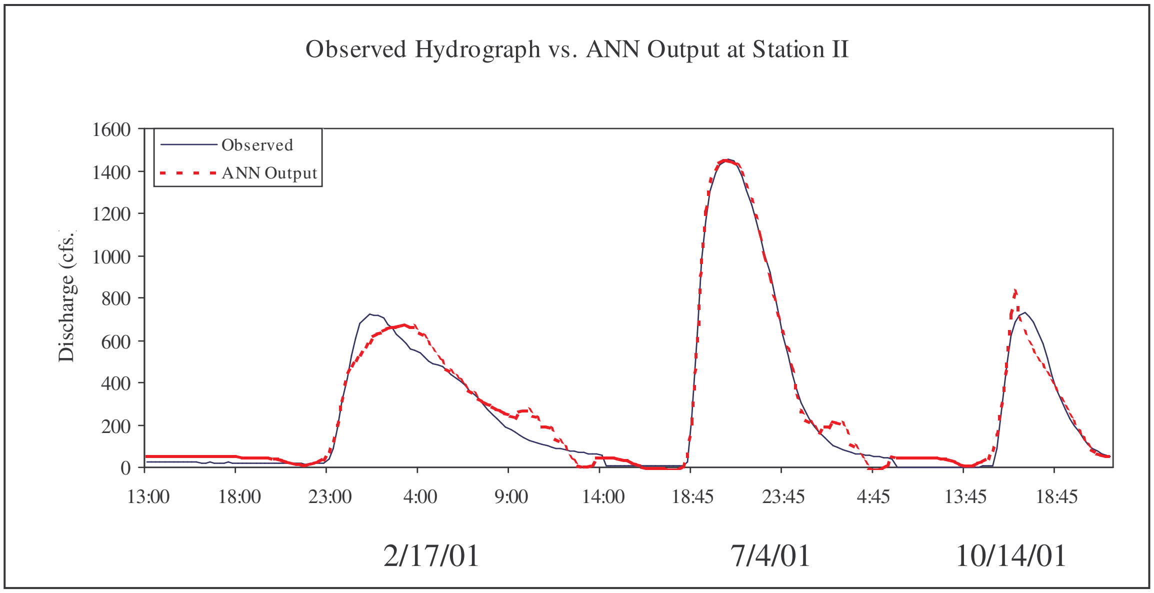


Figure 2.18 Observed Hydrographs vs. ANN Output at Station-II

2.7.1.4 Comparison of ANN model and Conceptual/Empirical models

Six selected hourly storms, shown in Appendix 1.8, were tested by both ANN and HEC-HMS models at Station-II in Greensboro area. Table 2.13 showed the input and output selection of ANN model. ANN model was trained by the storms on 1/18/01, 2/17/01, 3/3/10 and 3/15/01. The same storms were applied for HEC-HMS calibrations. Both models were tested using the storms on 3/20/01 and 3/30/01. The comparison of the peak flow estimation from ANN and HEC-HMS models was shown in Figure 2.19; smaller MRE values from ANN models indicated the better estimation of peak flow of ANN models. The computed hydrographs were shown in Appendix 1.9. The results showed the output hydrographs from ANN model matched the observed hydrographs better than HEC-HMS for both calibration and testing cases. MRE, MSE and R^2 values of ANN and HEC-HMS were calculated based on both calibration and testing storms and showed in Figure 2.20. The results demonstrated that ANN model gave much less MSE and MRE compared to HEC-HMS model. It indicated that ANN model provides more accurate stream flow prediction than HEC-HMS model at both high and low stream flow conditions. Table 2.14 listed the Chi-squared analysis of ANN and HEC-HMS models. The results showed that the hydrographs from both ANN and HEC-HMS models gave same distribution as observed hydrographs at significance level of 0.95. The smaller X^2 value from ANN model implied the better performance of ANN model compared with HEC-HMS model.

Table 2.13 Input and Output Selection of Hourly ANN Stream Forecast Model

Hourly ANN Stream Forecast Model	
Input	$P(t)$, $P(t - j\Delta t)$; $j = 1$ to 4 $Q(t)$, $Q(t - \Delta t)$;
Output	$Q'(t + \Delta t)$
Computational time step	$\Delta t = 1$ hour

Table 2.14 Chi-squared analysis of ANN and HEC-HMS models

Null Hypothesis H_0:	The Output Hydrograph from the model have same distribution as observed hydrographs.
Alternate Hypothesis H_1:	The random variable does not have the specific distribution
Level of Significance α:	0.05
Critical Region:	$X^2 \geq X^2_{\propto,0.05} = 34.8$
Result:	ANN: $X^2 = 5.285$ HEC-HMS: $X^2 = 32.11$
Decision:	$X^2 \leq X^2_{\propto,0.05} = 34.8$, the null hypothesis is not rejected. It can be concluded that the hydographs from both models have a distribution as specified

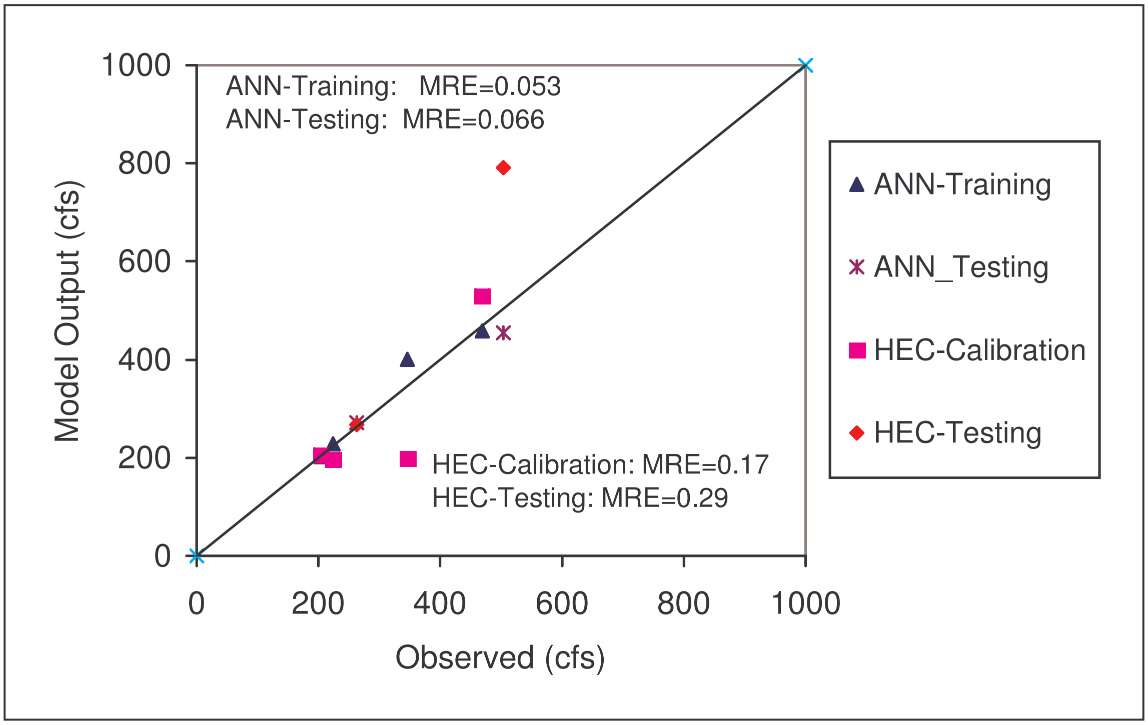


Figure 2.19 Comparison of ANN and HEC-HMS Peak Flow Estimation

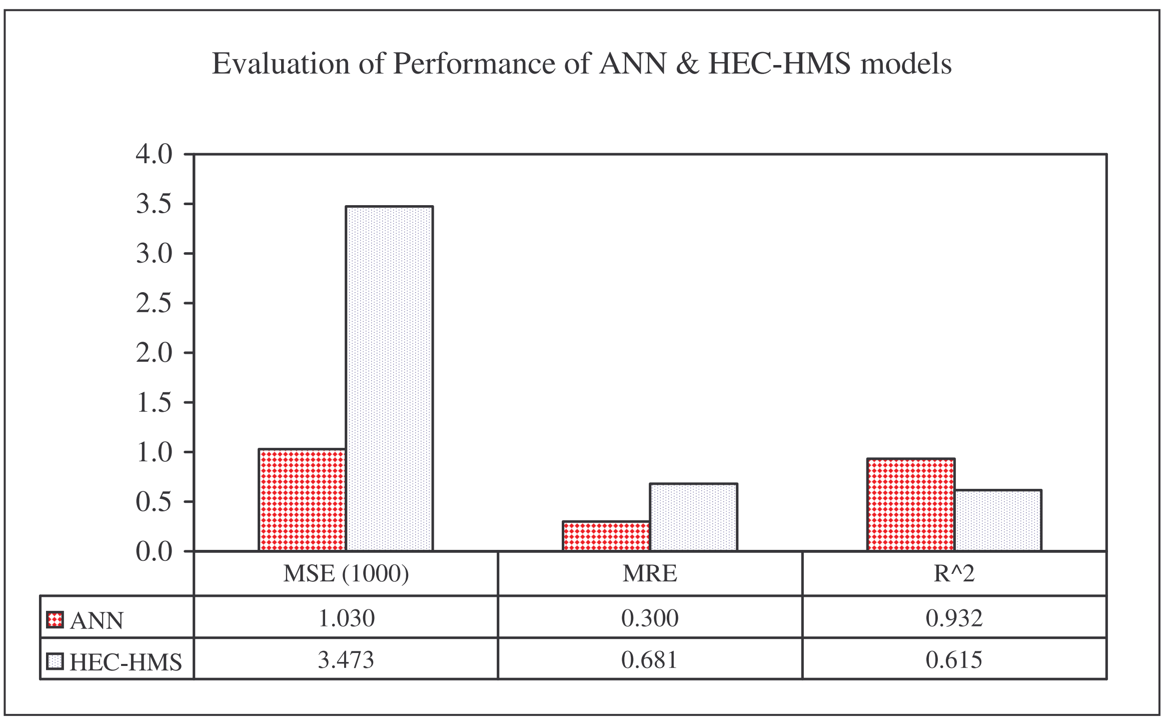


Figure 2.20 Comparison of the Performances of ANN and HEC-HMS models

2.7.1.5 Comparison of ANN and Empirical Approaches

According to the number of the variables, different empirical approaches (Equation 1.6 & 1.7) and ANN structures were selected. The details of ANN and empirical approaches were given in appendix 1.10. Both models were tested by the storms in 2001, see Appendix 1.1. The parameters of empirical models were calibrated by minimizing MSE between observed and predicated stream flow, and the results were tabulated in appendix 1.11.

To compare the performance of ANN and empirical models, the MSE and R^2 were shown in Figure 2.21 & 2.22. The results showed that the increase of the number of independent variables gave better performance of both ANN and empirical models. But when there were more than four variables in the model, the increase of the number of independent variables did not give significant improvement of the model performance. It because that as the number of independent variables increased, more precipitation information was included in the model input, it tended to give more accurate runoff estimates. Since the precipitation within one hour period had biggest effect on the runoff in the watershed, further addition of independent variables, which providing the precipitation beyond the 1-hour lag time, would have little effect on the performance of rainfall-runoff models. Figure 2.23 indicated that with proper selection of ANN structure, ANN could give better performance than empirical models. For the best-fit ANN model with four input neurons and 5 hidden neuron, the MSE and R^2 are 6459 and 0.956, respectively. For the best-fit empirical model 1 with four independent variables, MSE and R^2 are 8231 and 0.944, respectively. For the best-fit empirical model 2 with four independent variables, MSE and R^2 are 8076 and 0.945, respectively. The R^2 value obtained from ANN model is slightly better than empirical models. The MSE from ANN is about 20% less compared with the MSE from empirical models, which means that ANN model is expected to give better medium and high stream flow estimates than empirical models.

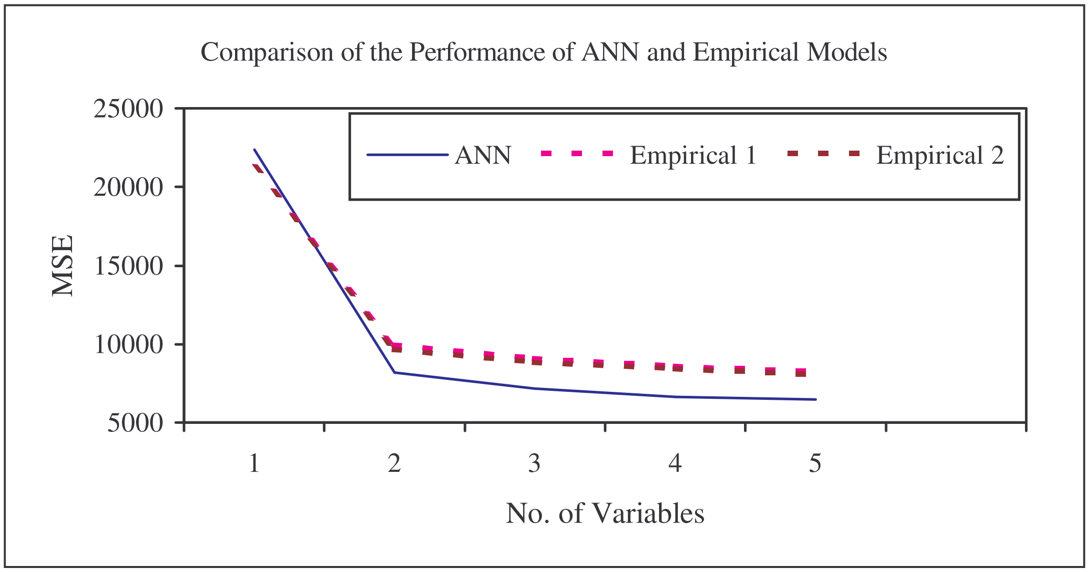


Figure 2.21 Comparison of the Performance of ANN and Empirical Models

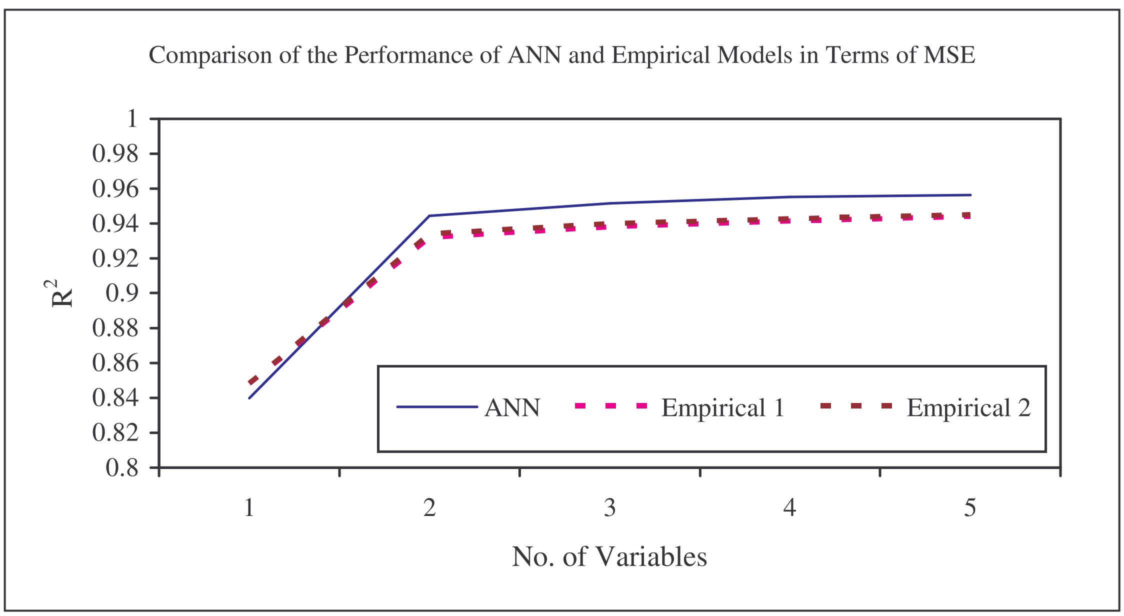


Figure 2.22 Comparison of the Performance of ANN and Empirical Models in Terms of MSE

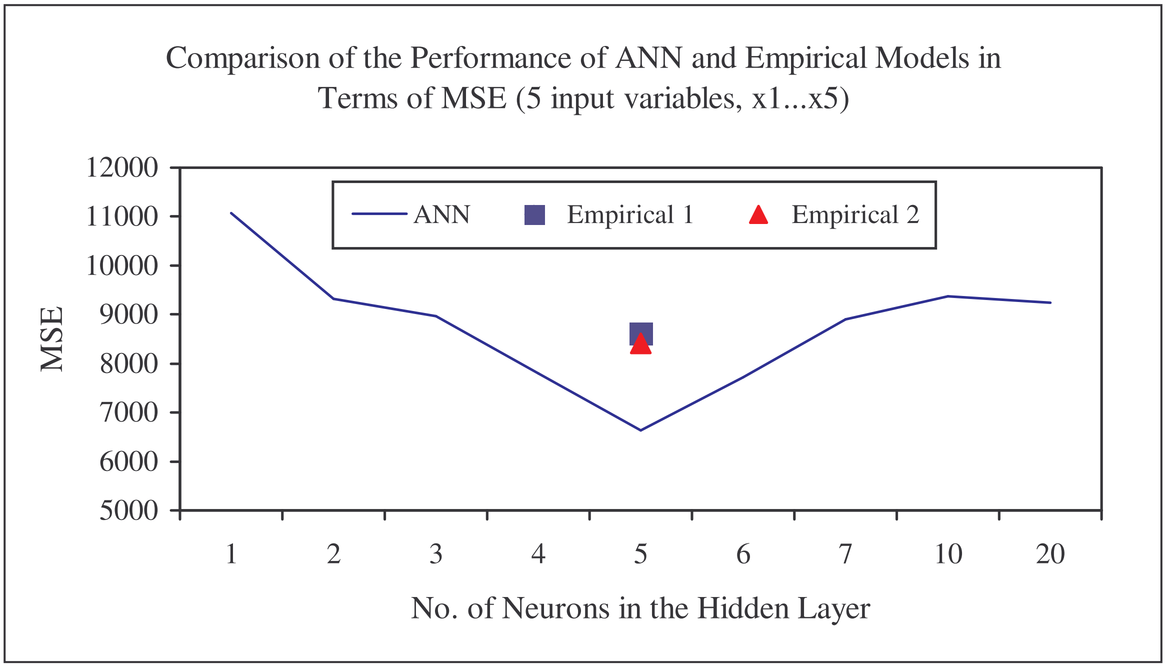


Figure 2.23 Comparison of the Performance of ANN with Different Hidden Neurons

2.7.2 ANN Application in Crabtree Creek, Raleigh NC

2.7.2.1 ANN WGH Model

The ANN WGH model was trained using nineteen storms events from the period of 1998 to 2000, and tested by five storm events in the year of 2001, see Appendix 1.1b. After training, the model can be expected to yield gage-high predications with a forecasting period of starting from 15 minutes and up to 2 hours from the present time as shown in Figure 2.24.

Figure 2.25 compares the output and observed gage heights at Crabtree-S1. Since all these three stations were nearby (less than one mile away), similar observations were obtained at station Crabtree-S2 and S3, which given in appendix 1.12 and 1.13. Figure 2.26 showed the comparison of the peak gage height estimates with the observation from gage stations at Crabtree-S1; Chi-squared goodness-of-fit analysis of testing data in Table 2.15

demonstrated that the peak stream height estimation had the linear distribution compared with peak stream height observation at significance level of 0.95. More observations were given in appendix 1.15. Performance of the ANN WGH model as evaluated by MRE, MSE and R^2 at the reference station is shown in Figure 2.27. It appears that gage height predictions improve with shorter lead-time prediction (smaller MRE and MSE values and higher R^2). Shorter lead-time prediction renders more accurate estimates of gage height during both low stream flow and high stream flow periods. However, a 2-hour lead time for gage height prediction was found to be fairly accurate and could provide reliable early warnings for small urban watersheds.

Table 2.15 Chi-Squared Goodness-of-Fit Analysis of Peak Flow Estimation of WGH Model, Raleigh

Null Hypothesis H_0:	Null hypothesis Ho: The peak flow estimation has linear distribution compared with observed peak flow.
Alternate Hypothesis H_1:	The peak flow estimation does not have the specific distribution
Level of Significance α:	0.05
Critical Region:	$X^2 \geq X^2_{3,0.05} = 7.81$
Result:	$X^2 = 0.502 < X^2_{3,0.05} = 7.81$
Decision:	The null hypothesis is not rejected. It can be concluded that the peak flow estimation has linear distribution compared with observed peak flow.

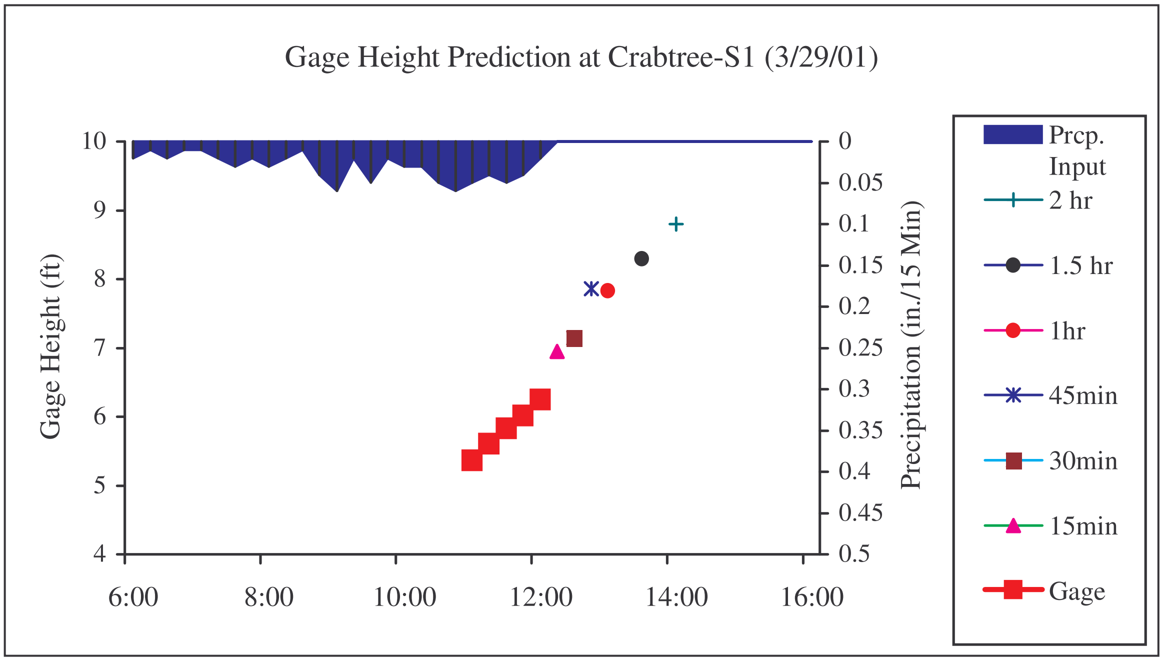


Figure 2.24 Gage Height Prediction at Crabtree-S1

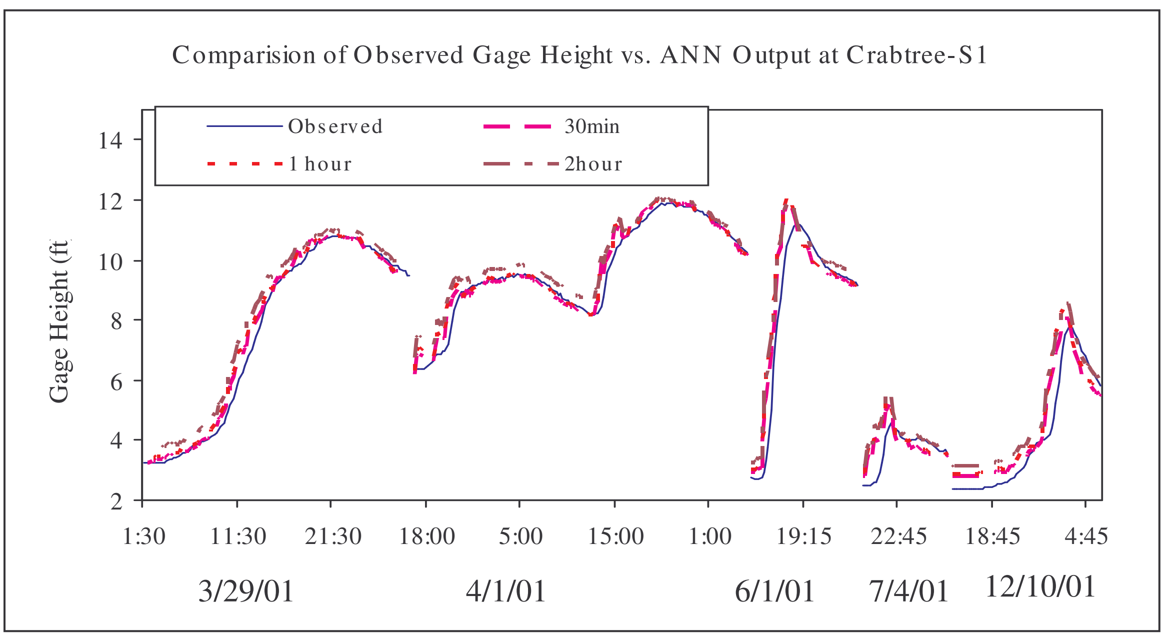


Figure 2.25 Comparison of Observed Gage Height and ANN Output at Crabtree-S1 (Testing Storms)

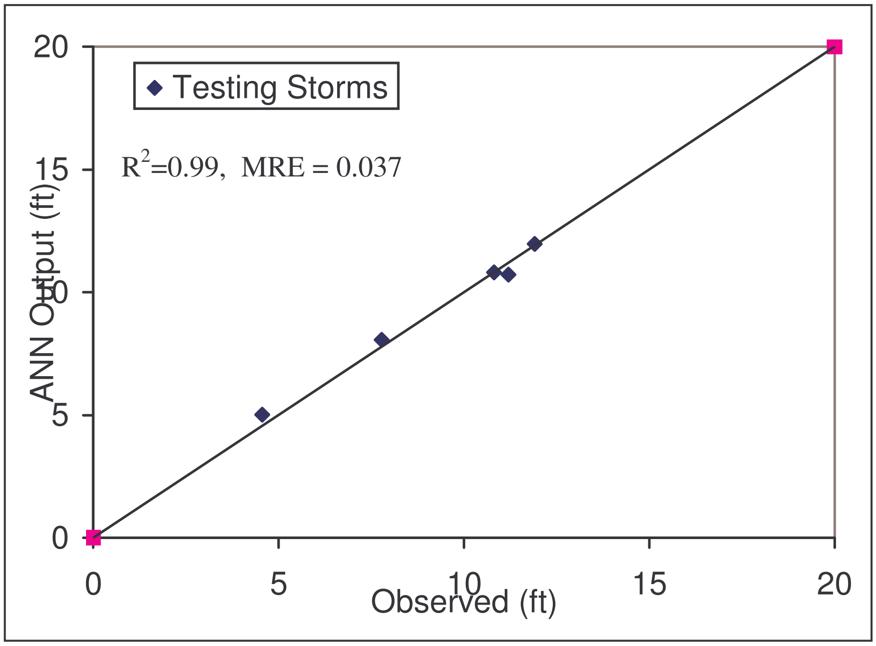
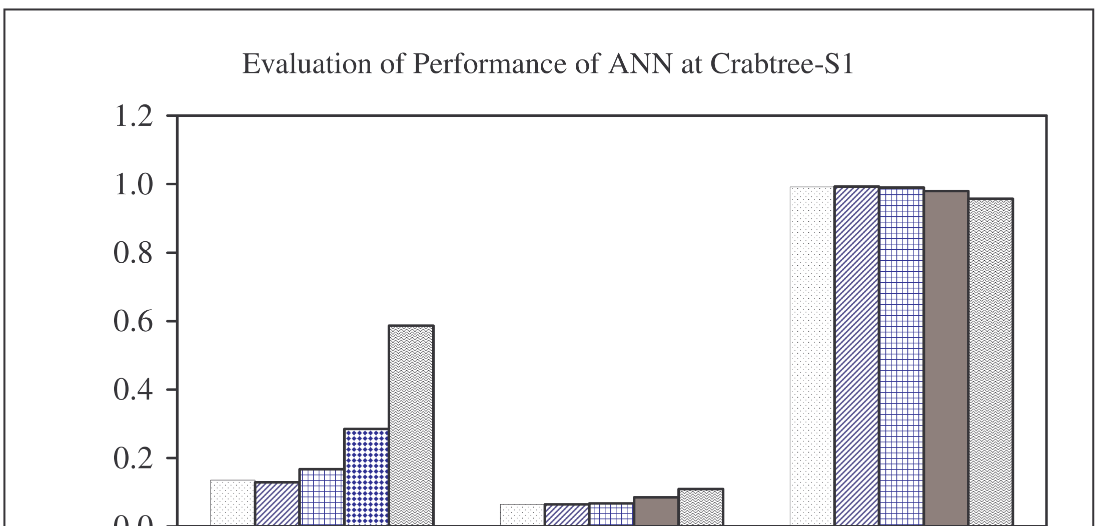


Figure 2.26 Observed-ANN Output Peak Gage Height Scatter plots at Crabtree-s1 (Testing Storms)



	MSE	MRE	R^2
15-min.	0.135	0.064	0.992
30-min.	0.129	0.064	0.993
1-hour	0.167	0.067	0.990
1.5 hr	0.286	0.085	0.980
2 hr	0.587	0.109	0.958

Figure 2.27 Evaluation of Performance of ANN at Crabtree-S1

2.7.2.2 ANN SGH Model

The same storms used for the WGH model were employed in this phase of study. It can be seen from Figures 2.28 and 2.29 that a reasonably good match of the predicted and observed gage heights was obtained for both downstream stations.

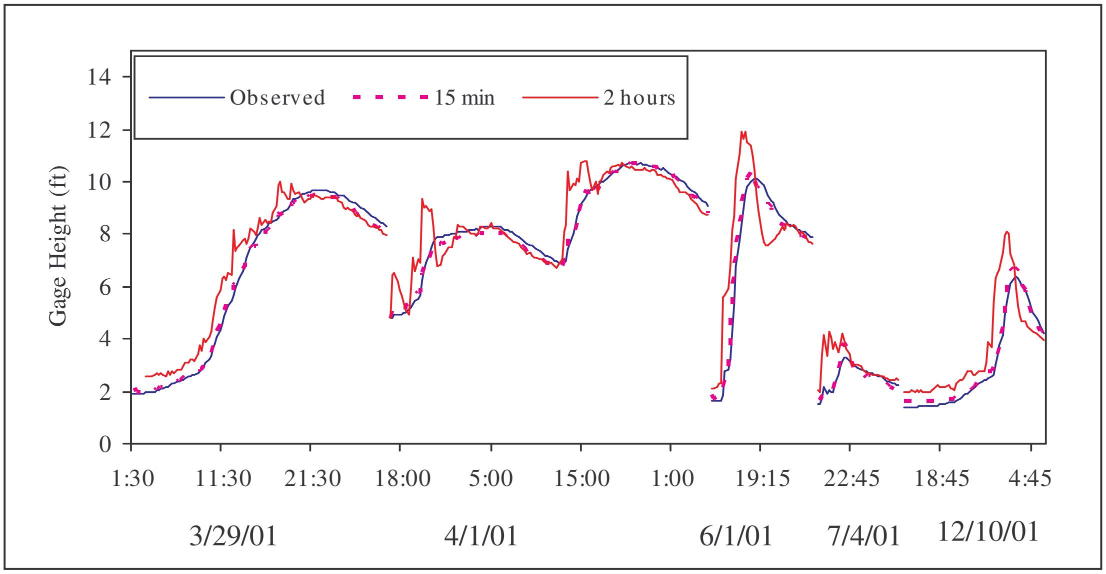


Figure 2.28 Observed Gage Height vs. ANN Output at Crabtree-S2 (Testing Storms)

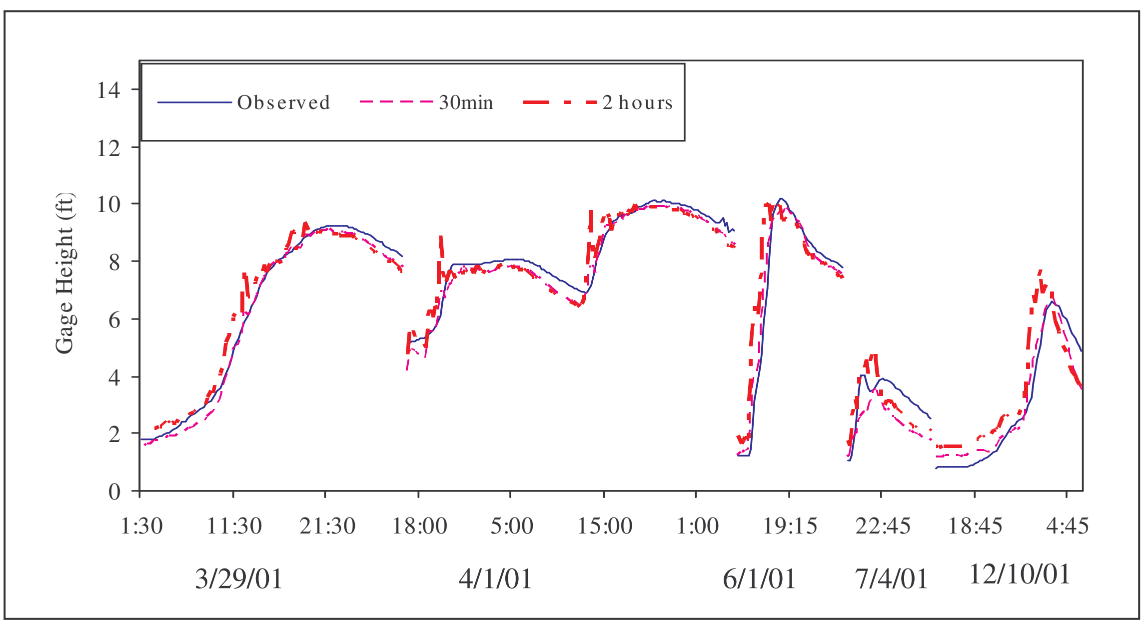


Figure 2.29 Observed Gage Height vs. ANN Output at Crabtree-S3 (Testing Storms)

2.8 Conclusions and Recommendations

Three ANN type models have been developed for the North Buffalo Creek in Greensboro, and one ANN model has been introduced for Crabtree Creek in Raleigh.

The ANN Watershed Runoff Prediction (WRP) model and Watershed Gage Height (WGH) model can be used to predict stream flow/gage height, including both base flow and direct runoff, at the outlet of a watershed. The model can be used to study rainfall-runoff relationships using historical precipitation and stream flow/gage height data. If rainfall data can be obtained in real time, the model can be operated to forecast stream flows at 15-min interval. The ANN Stream Flow Forecast (SFF) model and Stream Gage Height (SGH) provide advanced forecasting of downstream flows using meteorological and stream flow records of an upstream station. In applying the ANN SFF/SGH model, it is necessary to determine the time-of-travel to the desirable downstream station. Wherever the downstream flow/gage height records are unavailable, simultaneous measurement of upstream and downstream flows/gage heights for a few storm events need to be made to enable training of the ANN model. In addition, both WRP/SFF and WGH/SGH models can be linked in sequence and operated in real-time mode for predicting runoff and downstream stream. Further applications of ANN type models to precipitation data from the Doppler radar measurements are being investigated. The ANN History Data Generation (HDG) model is useful for generating upstream historical stream flow based on the stream flow at a downstream gauging station.

This study also gave the comparison of ANN model with one conceptual model (HEC-HMS model) and other empirical regression models. The results demonstrated the better performance of ANN model in terms of both low and high stream flow forecast. Other advantages of ANN application compared with conceptual model included (1) Easy application. After neural network models were set up, the operators can run the models and get expected results without having much knowledge of the principle of the models. (2) Low cost. The time involved to apply neural network models was significant less than conceptual models. And the operators are expected to require less training process.

In addition to the inherent superiority and efficiency of ANN models in solving complex problem, ANN-hydrologic models are particularly useful for real time prediction of stream flows and watershed modeling, which is particularly useful in urban watersheds for providing timely and early warning of flash floods.

The vast amount of meteorological and stream flow data which are routinely being collected by the USGS or other federal, state, and local government agencies, can be organized and synthesized for ANN model applications. On the other hand, the selection of training data plays an important role to enhance the ANN model output. The selected training data should be representative of the watershed or flow conditions and the training process should be repeated periodically to reflect the changing watershed and stream characteristics. The USGS record periods are less than five years in this study, and most of the storms, which used for ANN training were supposed less than 10-year storm. Before ANN models are applied for big flood forecast, such as 50-year or 100-year flood, there is more stream flow data, which are necessary to be collected from USGS stations, especially heavy storms. These data will be used for training and testing ANN models.

2.9 References

1. Anmala, J., Zhang, B., and Govindaraju, R.S., (2000). "Comparison of ANNs and empirical approaches for predicting watershed runoff." *Journal of Water Resources Planning and Management,* Vol. 126, No.3 May/June, 2000, page 156-166.
2. Bertoni, Juan Carlos, Tucci, Carlos Eduardo and Clarke, Robin Thomas., (1992). " Rainfall-Based Real-time Flood Forecasting." *Journal of Hydrology,* 131: 313-339.
3. Campolo, M., Andreussi, P., and Soldati, A. (1999). "River flood forecasting with a neural network model." *Water Resources Research,* vol. 35, No. 4, Page 1191-1197, April 1999.
4. Elshorbagy A, Simonovic S.P and Panu U.S. (2000). "Performance Evaluation of Artificial Neural Networks for Runoff Predication" *Journal of Hydrologic Engineering. Vol.5,* No. 4, October, 2000. 424-427.
5. Hsu, K., Gupta, H.V., and Sorooshian, S. (1995). "Artificial neural network modeling of the rainfall-runoff process." *Water Resources Research*, vol. 31, No. 10, Page 2517-2530.
6. Imrie, C.E, Durucan, S. and Korre, A., (2000). "River flow prediction using artificial neural networks: generalization beyond the calibration range." *Journal of Hydrology,* 233: 138-153.
7. Karunanithi, N., Grenney, J.W., Whitley, D., and Bovee, K., (1994). "Neural networks for river flow prediction." *Journal of Computing in Civil Engineering*, Vol. 8. No. 2, April, 1994, page 201-219.
8. Liong, Shie-yui, Lim, W. and Paudyal, G. N., (2000). "River stage forecasting in Bangladesh: Neural network approach." *Journal of Computing in Civil Engineering.* Vol. 14, No. 1. January, 2000. Page 1-8.

9. Smith, J. and Eli, R.N., (1995). "Neural-network models of rainfall-runoff process." *Journal of Water Resources Planning and Management,* Vol. 121, No.6 November/December, 1995, page 499-508.

10. Thirumalaiah, Konda, and Deo, M.C., (1998). "River stage forecasting using artificial neural networks." *Journal of Hydrologic Engineering,* vol. 3, No. 1, January, 1998, page 26-31.

11. Thirumalaiah, Konda, and Deo, M.C., (2000). "Hydrological forecasting using artificial neural networks." *Journal of Hydrologic Engineering*, vol. 5, No. 2, April, 2000, page 180-189.

12. Tokar, A. Sezin and Markus, Momcilo, (2000). "Precipitation-Runoff Modeling Using Artificial Neural Networks and Conceptual Models." *Journal of Hydrologic Engineering*, vol. 5, No. 2, April, 2000, page 156-161.

13. Zealand, C.M., Burn, D.H., and Simonovic S.P., (1999). "Short term streamflow forecasting using artificial neural networks." *Journal of Hydrology,* 214 : 32-48.

Chapter III: Application of Artificial Neural Network for Radar-based Rainfall Estimation

3.1 Abstract

The reliability of real-time hydrologic forecasting depends on good estimates of rainfall intensity as input to rainfall-runoff model. Weather radar is an important source of rainfall estimation. While radar does not measure rainfall directly, the accuracy of the measurements may not be adequate; many sources of error can affect radar rainfall estimation. People believe the selection of a valid z-R relationship can provide the most significant improvement in precipitation estimates. Two algorithms, which are empirical relationship and ANN, were discussed for radar-based rainfall estimation. Two z-R relationships, which were $z = 26.3R^{2.38}$ and $z = 160R^{1.77}$, were employed for Greensboro and Raleigh respectively based on PRMSE analysis. The results indicated that although these z-R relationships gave fairly rainfall estimates on average, they tended to overestimate of the rainfall amount of low intensity storms and underestimate of the rainfall amount of high intensity storms. Two ANN radar-rainfall models were developed for radar-based rainfall estimation. With proper training and testing, ANN model showed their ability of rainfall estimation using radar data. Compared with empirical z-R relationship, ANN models gave more accurate estimates of rainfall for both low and high intensity storms.

3.2 Introduction

The reliability of real-time hydrologic forecasting depends on good estimates of rainfall intensity as input to rainfall-runoff model. Rainfall intensity can be obtained by three techniques: (1) rain gage data, (2) radar data, and (3) radar data combined with rain gage measurements.

Rain gages are generally accurate and reliable for ground-level, point precipitation measurements. It is not unusual to find insufficient number of gages to cover the entire drainage area. Rain gage data from point measurements are not adequate for distributed hydrologic flood forecasting models. Weather radar (such as the Doppler type radar) provides detailed, in both time and space, information about precipitation patterns.

Spatially distributed rainfall data can be developed by weather radar measurement. The relation between rainfall intensity (R) and radar reflectivity (z) is given by:

$$z = aR^b \tag{3. 1}$$

where z = reflectivity factor (mm^6/m^3).

R = rainfall intensity(mm/h).

a and b = empirical coefficients.

While radar does not measure rainfall directly, many sources of error can affect the accuracy of radar rainfall estimation (Zawadzki, 1984). The z-R relationship is believed to be nonlinear and complex, it can vary over geographical areas and among storms (Battan 1975). The difference of radar estimates of rainfall and rain gage measurements made at ground level can be as high as a factor of two on some occasions (i.e., 100% error) (Wilson and Brandes 1979). Sensitivity analysis of rain rate estimates from radar data by the OSF (Operational Support Facility) and WSR-88D Adaptable Parameter Working Group showed that the selection of a valid z-R relationship can provide the most significant improvement in precipitation estimates (http://www.roc.noaa.gov/ops/z2r_osf5.htm). It is extremely important to use an appropriate z-R relationship when using radar to estimate rainfall rates.

With the advantages of solving problems that are too complex for conventional technologies, ANNs have been successfully applied in nonlinear system in hydrology-related areas such as rainfall forecasting (French, et. al. 1992), rainfall-runoff process (Smith and Eli, 1995), river flow forecast (Karunanithi, 1994) and more. It appears that a proper designed ANN model can be a potential approach for radar estimates of rainfall process. There was no known previous research in area of ANN applications in z-R relationships; the present study examined the potential applications of ANNs in estimating rainfall based on weather radar applications.

The goal of this chapter is to develop radar-based rainfall estimation algorithms using empirical and ANN approaches. This study is employed in cities of Greensboro and Raleigh, NC. This study also evaluates the performance of ANN approach compared to empirical method.

3.3 Literature Review

3.3.1 Introduction to Weather Radar System

The word radar is an acronym for "**RA**dio **D**etection **A**nd **R**anging". Radar transmits microwave signals into the atmosphere and then listens for return signals. If the transmitted signal intercepts an object, most of the energy is scattered, but some will be reflected back to the radar receiver. The quantity that a radar measures is the returned energy, which is converted to a quantity called reflectivity. The amount of reflected energy can be used to estimate the number or density of targets in the atmosphere. Because microwaves travel through air at a known velocity, time between transmission and reception of the reflected energy can be used to estimate the distance of a target from the radar. Reflectivity is often represented by the symbol “z”. Reflectivity is calculated by a complex empirical equation which involves the amount of power received (in Watts), the distance and nature of the target, and other factors. z is in units of mm^6/m^3 and can actually be defined in terms of the density of water droplets which would return a similar amount of power. z covers a wide range of values from very weak to very strong signals, so to produce a more convenient number for calculations and comparison, a decibel scale is often used. Mathematically:

$$Z = 10 \times \log\left(\frac{z}{z_0}\right) \tag{3. 2}$$

Where

Z = the logarithmic radar reflectivity factor measured in units of dBZ

z = reflectivity factor in mm^6/m^3

$z_0 = 1\ mm^6/m^3$

Originally, radar was designed to detect military targets. By the early 1950’s radar was designed specifically for weather detection. WSR-88D weather surveillance radar or NEXRAD (**NEX**t generation **RAD**ar) is a Doppler radar system that has greatly improved the detection of meteorological events such as thunderstorms, tornadoes and hurricanes. NEXRAD displays measures of reflectivity in Z or decibels that are represented as different colors on the radar image. An extensive network of NEXRAD stations provides almost complete radar coverage of the continental United States, Alaska, and Hawaii as shown in Figure 3.1. The range of each NEXRAD is 150miles.

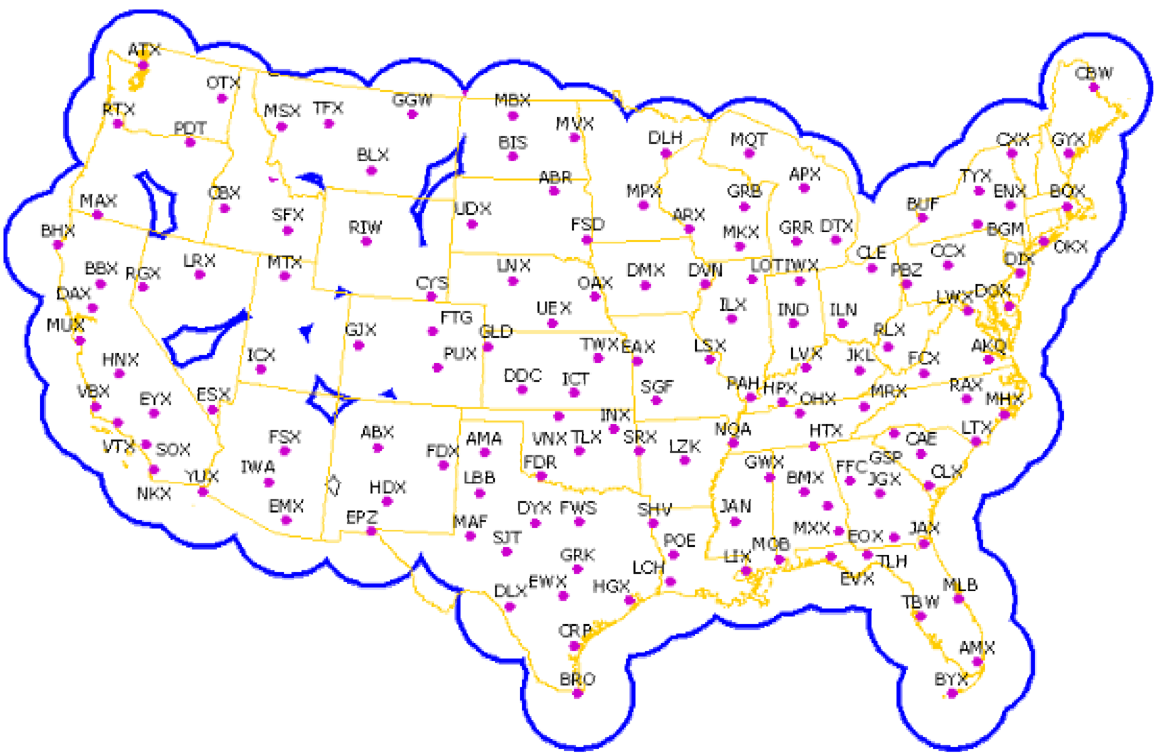


Figure 3.1 Network of NEXRAD Stations in US

Rain, snow, and hail can reflect considerable amounts of energy and thus are detectable by radar. Additionally, aerial biota such as insects, birds and bats, and particulate such as smoke, dust, and pollen also reflect radar signals. These reflected signals produce radar images.

NEXRAD is a Doppler radar, and it can track the velocity and direction of targets moving relative to the radar station.

Ground clutter on radar is detected when nearby buildings, trees, and towers reflect energy back to the radar antenna. Because NEXRAD is a doppler radar, many of these targets do not appear on radar. Immobile objects are pre-filtered from the reflectivity scan.

3.3.2 Introduction to WSR-88D Algorithm Testing And Display System (WATADS) and WSR-88D Precipitation Processing Subsystem (PPS)

WATADS is a product of the National Severe Storm Laboratory (NSSL) and consists of many special meteorological computer programs (“algorithms”) which process, analyze, and display WSR-88D data. WATADS extracts base data from 8-mm radar data tapes (in Level II format) and convert it to RADS format, which produces high-resolution images.

WSR-88D Precipitation Processing Subsystem (PPS) was executed to compute maps of one-hour, two-hour, three-hour and storm-total precipitation accumulations. The PPS has been fairly well documented (FMH-11; WSR-88D OSF Interactive Training Module Volume 1). It is divided into five parts: 1) preprocessing, 2) rainfall rate, 3) rainfall accumulation, 4) rainfall adjustment, and 5) precipitation products.

In the first step, the base data are quality controlled to correct problems due to radar beam blockage, spurious noise, reflectivity outliners, and ground returns. In the precipitation algorithm, only the base reflectivity data from the four lowest elevation scans (currently 0.4°, 1.4°, 2.4° and 3.3°) are used. The hybrid scan (1^{o} X 2 km) is built in the step.

The second step involves calculation of a precipitation rate. The hybrid scan of reflectivity data is converted to rainfall rate using z-R relationship followed by two quality control steps: 1) a time continuity step and 2) a range correction. The later corrects for underestimation of precipitation rates at long ranges due to signal degradation and partial beam filling.

The third step integrates volume scans of rainfall rate to determine precipitation accumulations.

The fourth step involves calibrations of radar using rain gauges. Currently this algorithm is bypassed.

Finally, precipitation products are generated, which include one-hour, two-hour, three-hour and storm-total precipitation accumulations.

3.3.3 Weather Radar Application in Rainfall Estimation

Forecast of flooding for a watershed due to rainfall requires a quantitative estimate of precipitation. Rain gauges are most commonly and widely used. The accuracy of areal time rainfall estimation is important when these estimates are used as inputs to flood forecasting models. However, rain gauges are generally accurate for making point precipitation measurement and may not be suitable for areal rainfall measurements. Usually, mean areal rainfall intensity must be estimated by averaging available rain gage readings. These estimates are subject to various sources of error: (1) inadequate spatial coverage or configuration, (2) inadequate gauge density and (3) instrument errors. (M. Borga, 2000).

Weather radar is another important source of rainfall estimation. Radar covers extended areas and provides spatial and temporal measurements. It allows rapid access for real-time hydrological applications. The NEXRAD system of WSR-88D radar is known to provide a wide spatial and temporal coverage of data collection and expected to provide high-quality, high-resolution precipitation data for the United States. A single high-resolution radar station can monitor a total area of 200,000 km^2 with a less than 5-km^2 resolution. The radar is a volume-scanning device employing successive tilt angles to cover an entire volume of the atmosphere and can measure reflectivity up to a range of 460 km. A very dense rain gauge network is needed to achieve this kind of coverage that a radar station can accomplish.

There are more than 160 WSR-88D radars manufactured and installed in the conterminous United States. The use of NEXRAD data provides a new approach of rainfall-runoff modeling, storm forecast, and, most importantly, near real-time flood prediction. Based on the study of more than 1 year of WSR-88D data and rain gauge data from the southern plains, Smith J. A. (1996) found that (1) the WSR-88D tends to underestimate the precipitation compared to the rain gauge observations, (2) accuracy and stability of radar calibration are among the most important prerequisites for the NEXRAD precipitation algorithms, and (3) the WSR-88D had far superior capability for monitoring heavy rainfall than rain gauge networks. Johnson D. and Smith M. (1999) compared over 4,000 pairs of mean areal precipitation values from NEXRAD and precipitation gauge network, the results showed that mean areal estimates derived from NEXRAD generally are 5-10% below gauge-derived estimates. They also found that by applying the mean areal precipitation values to hydrologic models, this difference can result significant differences in runoff volume.

Radar does not measure rainfall directly and many sources of error can affect radar rainfall estimation, which included (1) High variation in raindrop size distribution, (2) Bright band caused by the scattering of radar waves due to ice particles present in the higher levels of some clouds, (3) False echo caused by anomalous propagation of radar waves, (5) Miscalibration of radar electronic instruments. (Zawadzki, 1984).

Weather radar measurements of precipitation can be improved by calibrating the results from radar data with the data from rain gauge network. Collier (1986) reported that a less dense gauge network with a radar system calibrated using the data from a few of rain gauges is capable of producing measurements which have the same or better accuracy as sparse gauge network over a large area. James (1993) showed that both the modified Brandes and Kriging calibration techniques could be used for weather radar calibration and calibrated weather radar was shown to be effective for flood forecasting. The relation between radar reflectivity (z) and rainfall intensity (R) is based on the theory that back scattered radar power form the precipitation particles is proportional to the summation of the particle diameter to the sixth power times the number of particles in that size classification. Reflectivity is a meteorological parameter that is determined by the number and size of the particles present in a sample volume. Most of the reflectivity comes from the largest droplets. Moreover, different drop size distributions (DSD's) can yield the same z but different R. The z-R relationship has been studied for more than 50 years. Many empirical z-R relationships have been developed and used to convert radar measurements into rainfall estimates. Battan (1973) listed more than 60 experimentally determined z-R relationships, and many more have been determined since 1973! It was not unusual to have radar estimates of rainfall and rain gage measurements made at ground level to differ by a factor of two on some occasions (i.e., 100% error), making calibration with rain gage measurements necessary (Wilson and Brandes 1979). Seed (2000) analyzed z-R relationships by using several statistic methods, concluded that it will be difficult if not impossible to detect systematic shifts in the z-R relationship in real time based on a small number of observations. Perhaps the best that can be done operationally is to have a limited number of climatological z-R relationships, for convective and widespread rainfall say, and attempt to classify the rainfall prior to rainfall estimation. The most commonly used z-R relationship is given by Equation 3.1.

3.3.4 Sensitivity of Rain Rate Estimates to Different z-R relationship

The OSF (Operational Support Facility) and WSR-88D Adaptable Parameter Working Group encourage sites to optimize the performance of the PPS by selecting z-R relationships based on season, geographic location, and weather type. They believe the selection of a valid z-R relationship can provide the most significant improvement in precipitation estimates. (http://www.roc.noaa.gov/ops/z2r_osf5.htm)

The default WSR-88D Convective z-R relationship ($z=300R^{1.4}$) is most accurate during heavy rainfall associated with deep convection. In 1997, the OSF authorized sites to use a Tropical z-R relationship ($z=250R^{1.2}$) developed by Rosenfeld, et al. (1993). The Tropical z-R relationship should be used to improve PPS estimates in tropical convective systems, particularly during land falling hurricanes and tropical storms.

The OSF, the NWS Office of Hydrology (OH), and NWS field offices have continued to investigate radar rainfall estimates during non-convective events. They recommend that the Marshall-Palmer relationship ($z=200R^{1.6}$), developed by Marshall, et al.(1955), be used to provide the best PPS estimates during general stratiform rainfall events. Several studies [Super and Holroyd (1998), Cairns, et al. (1998), Huggins and Kingsmill (1998), and Quinlan and Sinsabaugh (1999)] have shown that the best z-R relationship in cool season stratiform rainfall events depends partly on geographic location. Based on these studies, we are authorizing sites to select from two cool season stratiform relationships, $z=130R^{2.0}$ for sites east of the continental divide and $z=75R^{2.0}$ for sites west of the continental divide.

Different z-R relationships give difference rainfall estimates. Table 3.1 shows the comparison of the rainfall rate calculation from different z-R relations. The result indicates that it is extremely important to use an appropriate z-R relationship when using radar to estimate rainfall rates.

Table 3.1 Rainfall Rate Comparison

Reflectivity (dBZ)	Marshall-Palmer ($z=200R^{1.6}$) (in/hr)	East-Cool Stratiform ($z=130R^{2.0}$) (in/hr)	West-CoolStratiform ($z=75R^{2.0}$) (in/hr)	WSR-88D Convective ($z=300R^{1.4}$) (in/hr)	Rosenfeld Tropical ($z=250R^{1.2}$) (in/hr)
15	0.01	0.02	0.03	<0.01	<0.01
20	0.03	0.04	0.05	0.02	0.02
25	0.05	0.06	0.08	0.04	0.05
30	0.11	0.11	0.14	0.09	0.13
35	0.22	0.19	0.26	0.21	0.33
40	0.45	0.35	0.46	0.48	0.85
45	0.93	0.61	0.81	1.10	2.22
50	1.91	1.09	1.44	2.50	5.80
55	3.93	1.94	2.56	5.68	15.14
60	8.07	3.45	4.55	12.93	39.53

3.4 NEXRAD Data Acquisition and Processing

3.4.1 Data Acquisition

WSR-88D Level II data were purchased from National Climatic Data Center (NCDC) for the station of KRAX. KRAX is located Raleigh/Durham NC. This station covers the whole study areas, which including Greensboro and Raleigh, NC. Figure 3.2 displays the study areas covered by KRAX station. City of Greensboro is about 80 miles away from KRAX station.

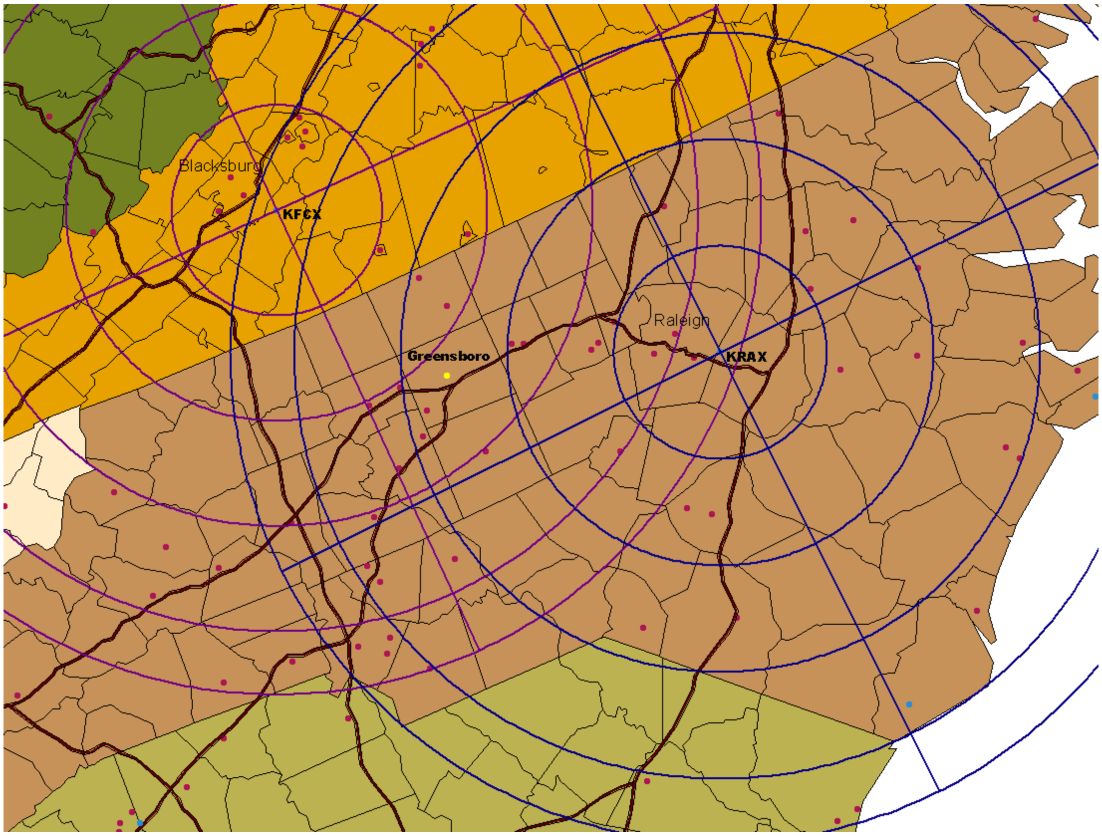


Figure 3.2 The Coverage of KRAX Station

WSR-88D Level II data are the high-resolution base data (reflectivity, mean radial velocity and spectrum width) produced at the full spatial and temporal resolution of the WSR-88D system.

A number of USGS gage stations in Greensboro and Raleigh areas were selected for developing z-R relationship. Hourly rainfall data was obtained from the USGS. Six and eleven storm events were used to examine the z-R relationship for Greensboro and Raleigh respectively. Two-third of the selected storms were used to calibrate the relationship, while one-third of the storm data were for verification. The details of these storms were given in Appendix 2.1a and 2.1b. The details of KRAX and selected USGS stations are shown in

Table 3.2. Figure 3.3 shows the locations of the selected gage stations under KRAX radar umbrella.

Table 3.2 The Locations of KRAX and USGS stations

Location	USGS Station Number	latitude(N)	longitude(W)	Distance from KRAX(mile)	Precipitation Record Period
Raleigh	KRAX station	35.67	78.49		
Raleigh	0208732534	35°47'14"	78°39'17"	12.49	2/1/00 - 12/31/01
Raleigh	0208732885	35°48'59"	78°35'37"	11.9	5/11/00 - 12/31/01
Raleigh	02087359	35°45'30"	78°34'58"	8.24	7/24/96 - 12/31/01
Raleigh	02087182	35°56'27"	78°34'57"	19.65	2/1/98 - 12/31/01
Greensboro	02094770	36 02'16"	79 48'01"	77.9	1/1/00 - 12/31/01
Greensboro	02095271	36 05'52"	79 46'55"	78.38	1/1/00 - 12/31/01
Greensboro	02095181	36 04'45"	79 48'48"	79.55	6/2/99 - 12/31/01
Greensboro	02095500	36 07'13"	79 42'30"	75.2	1/1/00 - 12/31/01
Greensboro	02093800	36°10'22"	79°57'12"	89.21	11/18/99 - 12/31/01
Greensboro	0209399200	36°08'12"	79°51'40"	83.49	8/1/99 - 12/31/01
Greensboro	02094412	36°10'46"	79°38'51"	73.95	6/1/99 - 12/31/01
Greensboro	02094659	36°02'59"	79°51'22"	81.12	1/1/00 - 12/31/01
Greensboro	02094775	36°01'51"	79°47'56"	77.67	8/1/99 - 12/31/01
Greensboro	02095000	36°03'36"	79°43'33"	74.5	8/1/98 - 12/31/01
Greensboro	0209553650	36°07'40"	79°39'43"	73.06	8/1/98 - 12/31/01

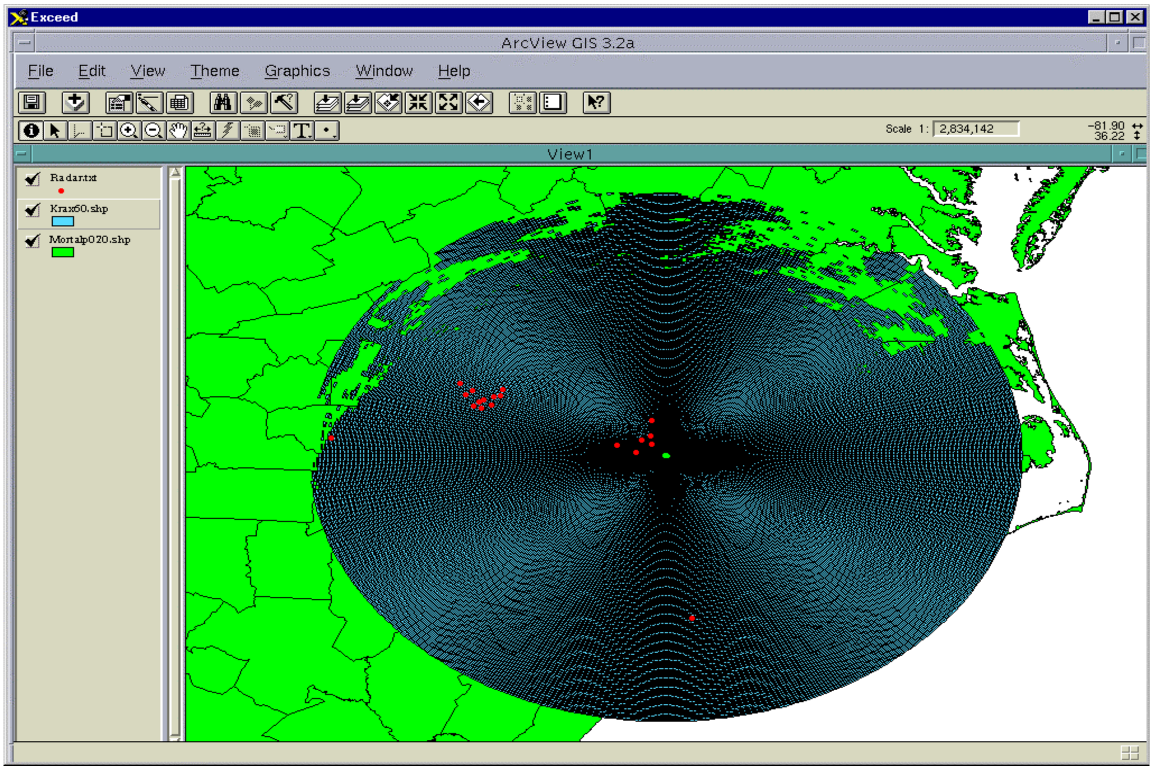


Figure 3.3 The Selection of USGS Stations in Greensboro and Raleigh

3.4.2 Data Processing

The WATADS (WSR-88D Algorithm Testing And Display System) software (free from NCDC) was installed on a SUN Workstation at the University of North Carolina at Charlotte. This software was used to read the NEXRAD data, and convert it into hourly precipitation totals.

3.5 z-R Relationship Setting Up

3.5.1 Empirical Relationship

The most commonly used deterministic model is the empirical relationship, which is given by Equation 3.1. The coefficients a and b can be obtained by a least-squares regression analysis of the reflectivity factors on the rain gage measurements.

3.5.2 Neural Network Applications

3.5.2.1 ANN architectures

Three-layer feed-forward ANN models, which include one input layer, one hidden layer and one output layer, were selected for z-R relationship applications.

The number of neurons in the input and output layers was determined by the number of input and output variables for a given system. The output of the model is rainfall intensity at target point. The inputs of the model were the reflectivity factors in a 5*5 polar matrix surrounding the target point as shown in Figure 3.4 based on the assumption that these reflectivity factors gave highest correlation of the precipitation at interested point. The grid resolution of the data was 2 km in range by 1 degree in azimuth.

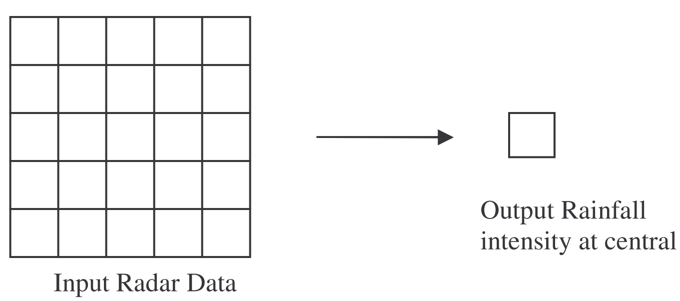


Figure 3.4 Input and Outputs Selection for ANN Application

3.5.3 Analysis of Performances of z-R Relationships

The accuracy of model results was evaluated based on the percent mean error (PME) and the percent root mean square error (PRMSE). The PME measures how well the unbiasedness property of the z-R relationships is realized. The PRMSE is a measure of performance in the minimum-error-variance sense (Seo et al., 1990). The PME and PRMSE are defined by:

$$PME = \frac{100}{m_x}\frac{1}{n}\sum_{i=1}^{n}\left(x'_i - x_i\right) \tag{3. 3}$$

$$PRMSE = \frac{100}{\sigma_x}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x'_i - x_i\right)^2} \tag{3. 4}$$

Where

PME = The percentage mean error;

PRMSE = The percent root mean square error

m_x = The mean of the observed gage data;

n = The total number of gage data;

x_i = the observed gage data;

$x_{i'}$ = the radar estimation of rainfall;

σ_x = the standard deviation of the observed gage data.

3.6 Results and Discussion:

3.6.1 Rainfall Estimation Using Empirical z-R Relationship

One of the major objectives of this paper was to find out the optimum z-R parameters for Greensboro and Raleigh areas. The most commonly used deterministic model is the empirical relationship, which is given by Equation 3.1.

The default z-R relation used in WSR-88D is $z = 300R^{1.4}$. The optimum coefficients a and b, which were (26.3, 2.38) for Greensboro and (160, 1.77) for Raleigh respectively, were obtained by a least-squares regression analysis of the reflectivity factors on the rain gage measurements. Figure 3.5 shows the comparison of the performance of z-R relationships with defaults (a, b) selection in WSR-88D and the optimum combinations of selecting (a, b). The comparison of PME and PRMSE values indicated that the optimum z-R relationship showed slightly improvement of performance against WSR-88D default z-R, but the improvement was not significant. Default z-R relationship was supposed to give fairly radar-based rainfall estimate for both areas.

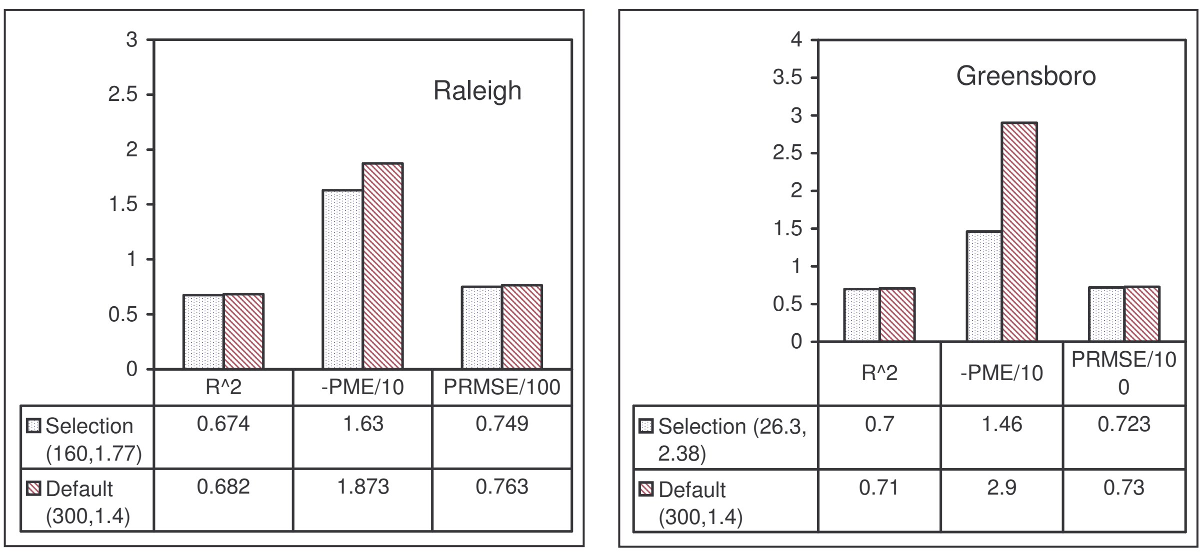


Figure 3.5 Comparison of the Performance of z-R Relationships with Defaults and Selected Coefficients in Greensboro and Raleigh Area

Table 3.3 showed the comparison of the computed total rainfall estimates from z-R relationship and the observation from rain gage for selected storms in Greensboro and Raleigh, see more details in Appendix 2.2a &b. Figure 3.6 showed the comparison of the radar-based hourly rainfall estimates with the observation from rain gage stations. The results indicated that although these z-R relationships gave fairly accurate rainfall estimates on average, they tended to overestimate the amount of precipitation for low intensity storms and underestimate the amount of precipitation for high intensity storms. The negative PME values for low intensity storms and positive PME values for high intensity storms in Table 3.4 also confirmed the same trend.

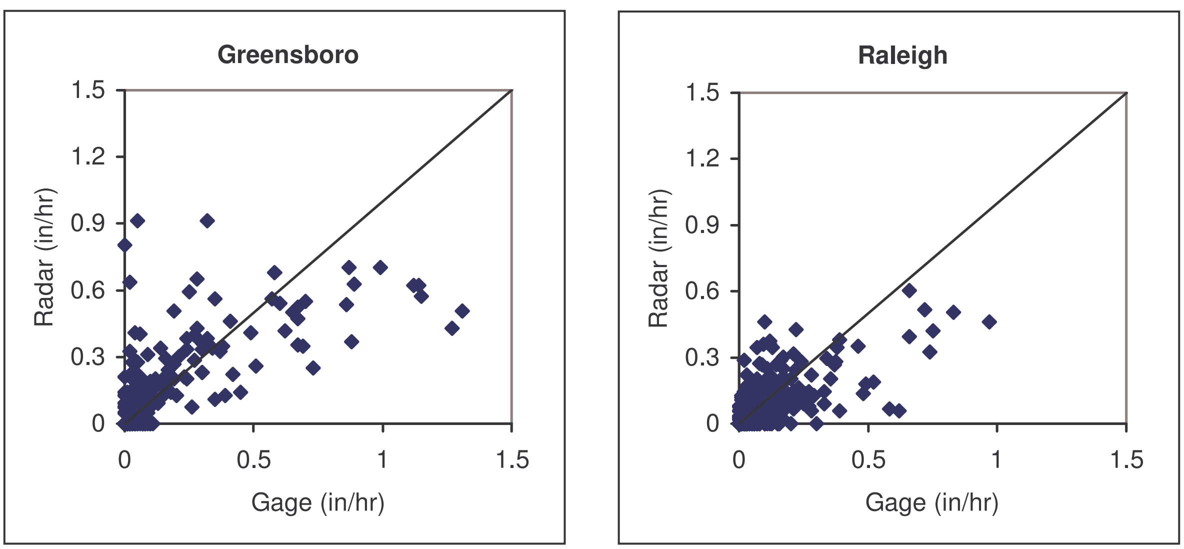


Figure 3.6 Rain gage-Radar Scatter plots in Greensboro and Raleigh

Table 3.3 Total Precipitation Estimates by z-R relationship vs. Rain Gage Observation

Greensboro (a = 26.3, b = 2.38)				
No.	Date	Duration (hr)	z-R Prcp. (in.)	Gage (in.)
1	9/29/99	8	0.70	0.60
2	9/25/00	8	0.53	0.43
3	6/13/01	2	0.87	0.58
4	7/26/01	4	0.79	0.90
5	8/11/01	3	0.83	0.61
6	10/14/01	3	0.55	1.16
Raleigh (a = 160, b = 1.77)				
No.	Date	Duration (hr)	z-R Prcp. (in.)	Gage (in.)
1	09/28/99	10.00	0.64	1.06
2	09/30/99	5.00	0.34	0.39
3	09/18/00	10.00	0.25	0.40
4	09/22/00	5.67	0.11	0.11
5	09/25/00	9.25	1.28	1.42
6	06/13/01	8.50	0.84	0.89
7	07/26/01	8.25	1.17	1.32
8	08/12/01	7.00	0.73	0.55
9	08/13/01	3.00	0.34	0.16
10	09/24/01	7.50	0.80	0.84
11	10/14/01	2.00	0.64	1.03

Table 3.4 Performance of z-R Relationship for High-intensity and Low-intensity Storms

	PME	
	Greensboro (26.3, 2.38)	Raleigh (160, 1.77)
High-intensity storm (>0.3 in/hr)	-40.74	-47.97
High-intensity storm (>0.3 in/hr)	39.56	17.03

3.6.2 Rainfall Estimation Using ANN

A sensitivity analysis was performed to make the selection of ANN inputs. Three different combinations of input selections were tested as shown in Table 3.5; and the results were demonstrated in Figure 3.7. The results indicated that Model I provided highest R and lowest PRMSE value; the reflectivity factors in a 5*5 polar matrix surrounding the target point was selected as the best-fit inputs to ANN model. It pointed out that the precipitation at the target point was not only effected by the reflectivity factor above the target point, but also the reflectivity factors around the target point.

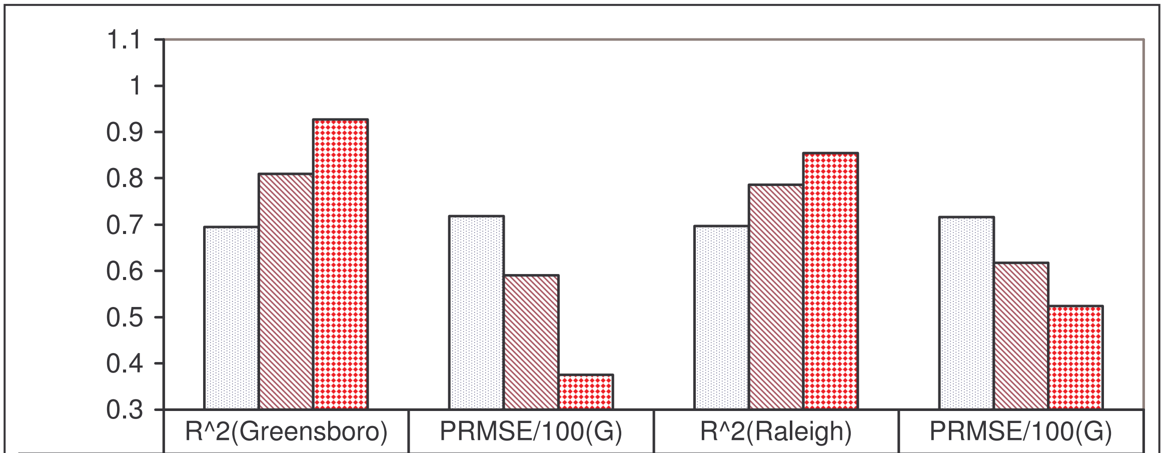
Table 3.6 illustrated the computed total rainfall estimates ANN model versus the rainfall observation from rain gage in Greensboro and Raleigh, see more details in Appendix 2.2a &b. The comparisons of hourly precipitation between gage data and radar-based rainfall estimates by ANN models were shown in Figure 3.8. The results indicated that with proper training and testing, ANN model gave fairly rainfall estimates for both low and high intensity storms.

Table 3.5The Structures of ANN Models

Model	Input	Output	Note
I	$z_{3,3}$	$R_{3,3}$	The input of the model is the reflectivity factor at the point just above the target point.
II	$z_{i,j}$, where i = 2,3,4 and j = 2,3,4	$R_{3,3}$	The inputs of the model are the reflectivity factors in a 3*3 polar matrix surrounding the target point.
III	$z_{i,j}$, where i = 1,2,3,4,5 and j = 1,2,3,4,5	$R_{3,3}$	The inputs of the model are the reflectivity factors in a 5*5 polar matrix surrounding the target point.

$z_{i,j}$: Current reflectivity factor at point of (i, j).

$R_{3,3}$: Current rainfall intensity at central point.



	R^2(Greensboro)	PRMSE/100(G)	R^2(Raleigh)	PRMSE/100(G)
Model-I	0.695	0.718	0.697	0.716
Model-II	0.81	0.59	0.786	0.617
Model-III	0.927	0.375	0.855	0.524

Figure 3.7 The Performance of ANN Models for Radar-based Rainfall Estimation

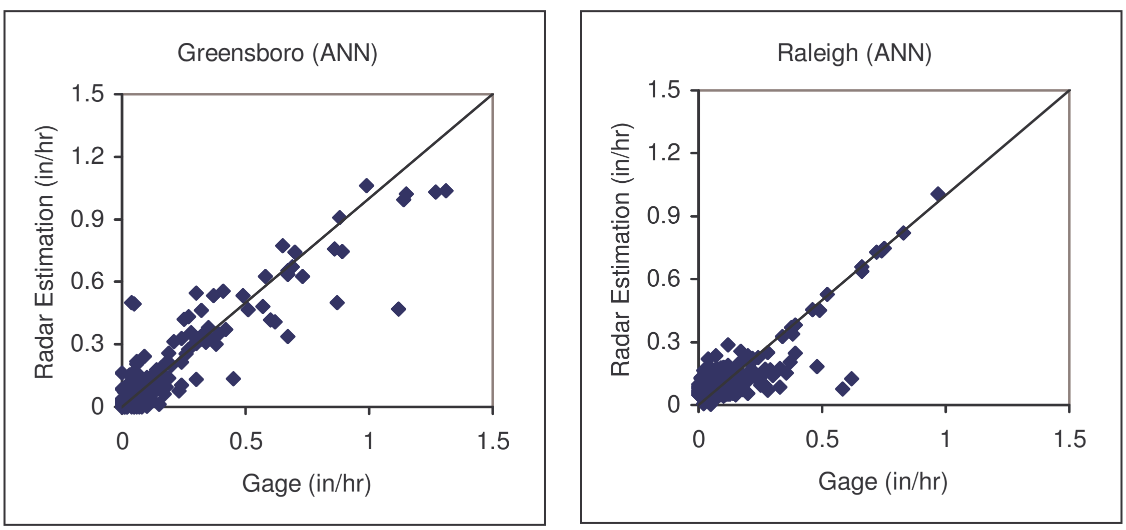


Figure 3.8 Rain gage-Radar Scatter Plots in Greensboro and Raleigh Area

Table 3.6 Total Precipitation Estimates by ANN vs. Rain Gage Observation

No.	Date	Duration (hr)	ANN	Gage (in.)
Greensboro				
1	9/29/99	8	0.53	0.60
2	9/25/00	8	0.46	0.43
3	6/13/01	2	0.69	0.58
4	7/26/01	4	0.98	0.90
5	8/11/01	3	0.73	0.61
6	10/14/01	3	0.89	1.16
Raleigh				
No.	Date	Duration (hr)	ANN	Gage (in.)
1	09/28/99	10.00	0.93	1.06
2	09/30/99	5.00	0.39	0.39
3	09/18/00	10.00	0.67	0.40
4	09/22/00	5.67	0.32	0.11
5	09/25/00	9.25	1.70	1.42
6	06/13/01	8.50	0.99	0.89
7	07/26/01	8.25	1.29	1.32
8	08/12/01	7.00	0.77	0.55
9	08/13/01	3.00	0.36	0.16
10	09/24/01	7.50	0.94	0.84
11	10/14/01	2.00	0.90	1.03

3.6.3 Comparison of Empirical z-R Relationship and ANN models

Comparison of the results in Figure 3.6 and 3.8 showed the better performance of ANN models over z-R relationship. Table 3.7 presented the comparison of the performance of ANN and empirical z-R relationship on high-intensity storms; the lower values of PME and PRMSE indicated that ANN provided more accurate rainfall estimates in high-intensity storms. It was very important when these estimates are used for modeling the flood warning system.

Table 3.7 Comparison of Performance of ANN and z-R Empirical Models for High-intensity Storm (>0.3 in/hr)

	Raleigh		Greensboro	
	PME	PRMSE	PME	PRMSE
z-R Relationship	-47.97	157.84	-40.74	138.06
ANN	-20.64	99.13	-11.14	61.63

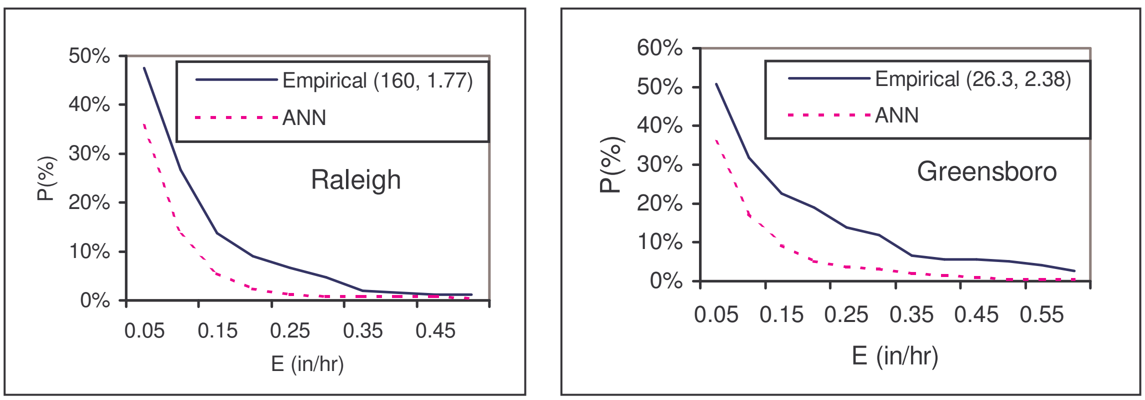


Figure 3.9 the percentage of radar data (P) with errors greater than the stated values (E) for the Greensboro and Raleigh area.

Figure 3.9 showed the percentage of radar estimates in Greensboro/Raleigh area with error greater than particular values with different radar-rainfall estimation approaches. For rainfall of more than 0.2 inch in one hour, 41% of the z-R radar-based rainfall estimates in Raleigh and 51% in Greensboro had errors greater than 0.2 inch; whereas when the data were based on ANN radar estimates, it was reduced to about 15% and 16%, respectively. It was recognized that the rainfall estimates from ANN model were less biased than z-R relationship. The results signified that ANN models presented more accuracy of hourly radar estimates of rainfall for both high and low intensity storms.

3.7 Conclusion and Recommendation

A series of storm events were analyzed and processed. Weather radar with appropriate radar-rainfall algorithm could be effective for flood forecasting. Two potential radar-rainfall estimation models were introduced for the selected areas: empirical z-R relationship model and ANN radar-rainfall model.

- Empirical z-R relationship model: Two z-R relationships, which are $z = 26.3R^{2.38}$ and $z = 160R^{1.77}$, were developed for Greensboro and Raleigh respectively based on PRMSE analysis. The result indicated that although these z-R relationships gave fairly rainfall estimates on average, they tended to overestimate the rainfall amounts for low intensity storms and underestimate the rainfall amounts for high intensity storms.
- ANN radar-rainfall model: with proper training with represent rainfall data, ANN models showed their ability of rainfall estimation using radar data. Compared with empirical z-R relationships, ANN models gave better estimates of rainfall in terms of both low and high intensity storms.

It was the first attempt to develop an ANN model for the purpose of radar application to rainfall estimation. The ANN models were considered to be a primary phase, there was still large room to give the improvement of the model. By adding other information that was available in NexRad level II data, such as wind speed and direction, ANN model could be expected to give more accurate rainfall estimate and longer predication time. It is very useful when the rainfall estimate will be introduced as an input to flood forecasting model.

3.8 Reference:

1. Battan, L.J., (1973). "Radar Observation of the Atmosphere." *The University of Chicago Press*, 324pp.
2. Borga, M, D. Da Ros, S. Fattorelli and A. Vizzaccaro., (2000). "Influence of various weather radar correction procedures on mean areal rainfall estimation and rainfall-runoff simulation." *Proc., 3rd Int. Symp. on Hydro. applications of weather Radar*, ABRH.
3. Cairns, M., A. Huggins, and S. Vasiloff, (1998). "Precipitation Algorithm improvements in the Eastern Sierra*." NWS Western Region Technical Attachment No. 98-08, Salt Lake City, UT*, 5 pp.
4. Collier, C.G., (1986). "Accuracy of Rainfall Estimates by Radar, Part II: Comparison with Raingauge Network. " *Journal of Hydrology*, 83: 225-235.
5. Federal Meteorological Handbook (FMH) No. 11.
6. French, M.N., Krajewski, W.F., and Cuykendall R.R., (1992). "Rainfall Forecasting in Space and Time Using Neural Network." *Journal of Hydrology,* 137: 1-31.
7. Huggins, A., and D. Kingsmill, (1998). "Improvements of WSR-88D Algorithms in the Intermountain West with Applications to Flash Flood Forecasts and Wintertime QPFs." *CIASTA Annual Report on Progress under Task II: Weather Research. Desert Research Institute, Dandini Research Park, Reno, NV,* 26 pp.
8. James, W. P, Bell, J.F, (1993). "Radar-Assisted Real-time Flood Forecasting." *Journal of Water Resources Planning and Management*, Vol. 119, No. 1, January/February, 1993, 32-44.

9. Johnson D., Smith M., Koren V. and Finnerty B., (1999) "Comparing Mean Areal Precipitation Estimates From NEXRAD and Rain Gauge Networks*." Journal of Hydrologic Engineering,* April 1999. 117-124pp.

10. Karunanithi, N., Grenney, J.W., Whitley, D., and Bovee, K., (1994). "Neural networks for river flow prediction." *Journal of Computing in Civil Engineering*, Vol. 8. No. 2, April, 1994, page 201-219.

11. Marshall, J. S., W. Hitschfeld, and K. L. S. Gunn, (1955). "Advances in radar weather." *Adv. Geophys*. 2, 1-56.

12. Quinlan, J. S. and E. J. Sinsabaugh, (1999). "An evaluation of the performance of the Snow Algorithm at NWFO Albany, NY during the 1997-98 Winter Season*." 29th Conf. on Radar Meteor., Montreal, Quebec, Canada, Amer. Meteor. Soc.,* 794-797.

13. Rosenfeld, D., D. B.Wolff, and D. Atlas, (1993). "General probability-matched relations between radar reflectivity and rain rate." J. *Appl. Meteor.,* 32: 50-72.

14. Seo, D.J, Krajewski, W.F, Azimi-Zonooz, A, and Bowles, D.S. (1990). "Stochastic Interpolation of Rainfall Data From Rain Gages and Radar Using Cokriging 2. Results." *Water Resources Research*, vol. 26, No. 5, 915-924 pp, MAY 1990.

15. Smith, J. and Eli, R.N., (1995). "Neural-network models of rainfall-runoff process." *Journal of Water Resources Planning and Management*, Vol. 121, No.6 November/December, 1995, page 499-508.

16. Smith, J. A., Seo, D. J., Baeck, M. L. and Hudlow M.D., (1996). "An Intercomparison study of NEXRAD precipitation estimates." *Water Resources Research*, vol. 32, No. 7, 2035-2045 pp, JULY 1996.

17. Super, A. and E. Holroyd III, (1998). "Snow accumulation algorithm for the WSR-88D radar, Final Report." *Bureau of Reclamation Report R-98-05, Denver, CO,* 75 pp.

18. Wilson, J.W., and Brandes, E.A., (1979). "Radar measurement of rainfall- A summary." *Bull. Amer. Meteor. Soc.,* 60(Sept.), 1048-1058.

19. Zawadzki I., (1984). "Factors affecting the precision of radar measurements of rain." *Preprints of Conf. Of Radar Meteorology*, 10-13 Sept. 1984, Zurich, AMS, 251-256.

Chapter IV: Radar-based Flood Warning Systems

4.1 Abstract

People have been looking for the flood forecasting model utilizing radar data for many years. This chapter describes the utilization of radar data in flood forecasting systems. Two radar-based stream flow forecast models, which were radar-rainfall-runoff (3R) model and radar-runoff (2R) model, were introduced. In the 3R model, radar data was employed to provide rainfall estimates, which served as input to a hydrologic stream forecast model to give the predication of stream flow. Even through the radar data showed its superiority of application in non-uniform storms, it had the tendency to underestimate peak flows at high intensity storms and overestimate peak flows at low intensity storm. For the 2R model, historical radar data and stream flow were served as direct input to ANN model; and the model provided stream flow predication at one hour lead time. The results indicated the better performance of ANN-based 2R model in terms of predication accuracy, efficiency and cost.

4.2 Introduction

A flash flood is typically defined as flooding that occurs within six hours of a heavy rain event. Most flood deaths are due to flash floods**.** Because flash floods are local, short lived events; it makes long-term flood warning very difficult. Sometimes there was no warning for these deadly, sudden floods.

The NEXRAD (Next Generation Weather Radar) system of WSR-88D (Weather Surveillance Radar-1988 Doppler) radar is known to provide a wide spatial and temporal coverage of data collection and expected to provide high-quality, high-resolution precipitation data for the United States. The use of NEXRAD data provides a new approach of rainfall-runoff modeling, storm forecast, and, most importantly, real-time flood prediction. Even people have been looking for the flood forecasting model utilizing radar data for many years; there was no direct relationship of stream flow and radar data being developed. The reasons are obviously; it is believed that both of the radar vs. rainfall and rainfall vs. runoff relationships are highly non-linear, it will be very difficult if not impossible to setup the mathematical functions for radar-stream flow relationships. Most research has focused on rainfall estimation using radar, and then the estimated rainfall information was used as an input to rainfall-runoff/streamflow model to get stream flow forecast (James, W. P., 1993, Shidawara, M, 2000, Baltas, E.A. and Mimikou, M.A., 2000). Both radar-rainfall and rainfall-runoff models needed to be calibrated and tested. It not only increased the possibilities of flow forecasting errors, but also prolonged the data process time.

Engineers are still looking for the effective ways of weather radar usage to elongate lead time of the runoff forecasting. Artificial neural networks (ANN) have been widely used as models for a variety of hydrologic processes in recent years, and shown the advantages when the underlying problem is either poorly defined or not clearly understood. This chapter demonstrated the radar applications in flood forecasting system. The main purposes of this chapter include (1) discussing different approach of radar-based flood forecasting systems. (2) examining the potential application of ANNs in directly radar-stream flow forecasting process. The whole paper is organized as follows:

- Setting up non-directly radar-based flood forecasting systems, which including radar-rainfall and rainfall-runoff processes.
- Developing directly radar-based flood forecasting systems using ANN applications.
- Comprising the performance of directly and non-directly radar-based flood forecasting systems.

4.3 Literature Review

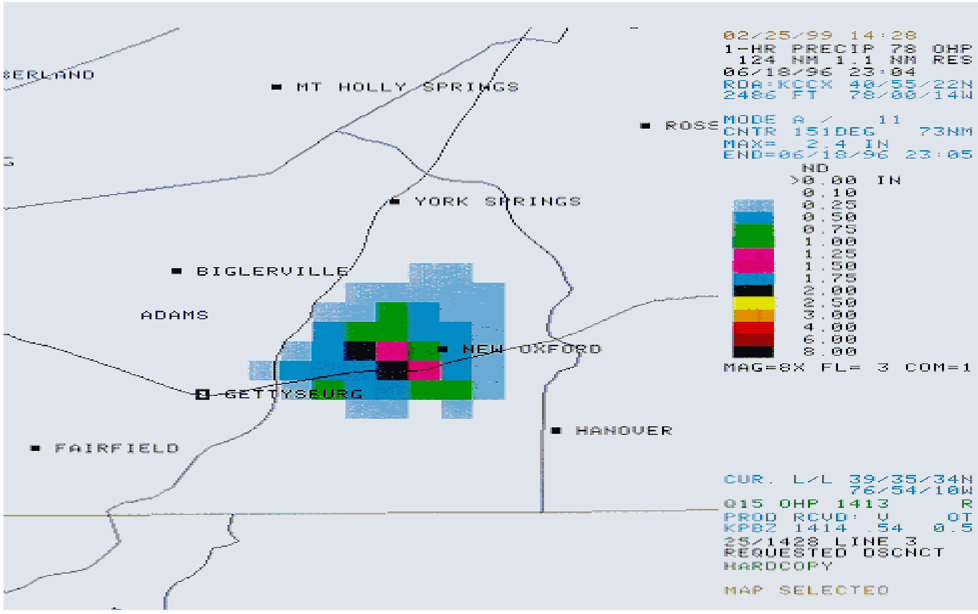
4.3.1 Weather Radar Application in Stream Flow Forecast

A flash flood is typically defined as flooding that occurs within six hours of a heavy rain event. Most flood deaths are due to flash floods**.** Because flash floods are local, short lived events; it makes long-term flood warning very difficult. Sometimes there was no warning for these deadly, sudden floods. Urbanization of flood plains presents additional challenges in flood forecasting and warning. Increasing urbanization reduces the infiltration capacity of watershed and increases the flood potential; cities have greater runoff rates than rural areas. Population growth increases and greater recreational use of mountainous regions put more and more people at risk.

The NEXRAD (Next Generation Weather Radar) system of WSR-88D (Weather Surveillance Radar-1988 Doppler) radar is known to provide a wide spatial and temporal coverage of data collection and expected to provide high-quality, high-resolution precipitation data for the United States. The use of NEXRAD data provides a new approach of rainfall-runoff modeling, storm forecast, and, most importantly, real-time flood prediction.

4.3.2 Are New Radars Helping Identify Floods?

The following example shows the successful application of radar in flash flood warnings in PA. New radars estimate rain that falls in one-hour time, in 3 hours, and throughout the life of a storm. Timely warnings for recent summer floods in Hollidaysburg, Bedford, York and Gettysburg were based on rainfall estimates from radar. The following Figure shows State College Radar Estimate of Rainfall at 704 pm on June 18, 1996. The Radar photo estimates 2.4 inches of rainfall has fallen northeast of Gettysburg in one hours time at 704 pm on June 18th of 1996 (see line 4 in the upper right corner of photo, time is Greenwich Mean Time). Just minutes after the time of this picture, Flash Flood Warnings were issued for Adams County. County Emergency Management of Adams County maintains an automatic rain gage on the roof of the County Court House in Gettysburg. The following chart shows how rain in that gage accumulated in Gettysburg. Rain began around 9:00 p.m. on June 18th and continued through the early morning hours of the 19th. Flash Flood Warnings for Adams County were issued at 7:11 p.m. based on the radar rainfall estimates. Radar continued to show the heavy precipitation that moved west and remained over the Gettysburg area as the evening progressed. Final totals were close to 11 inches overnight, an incredible amount for Pennsylvania!



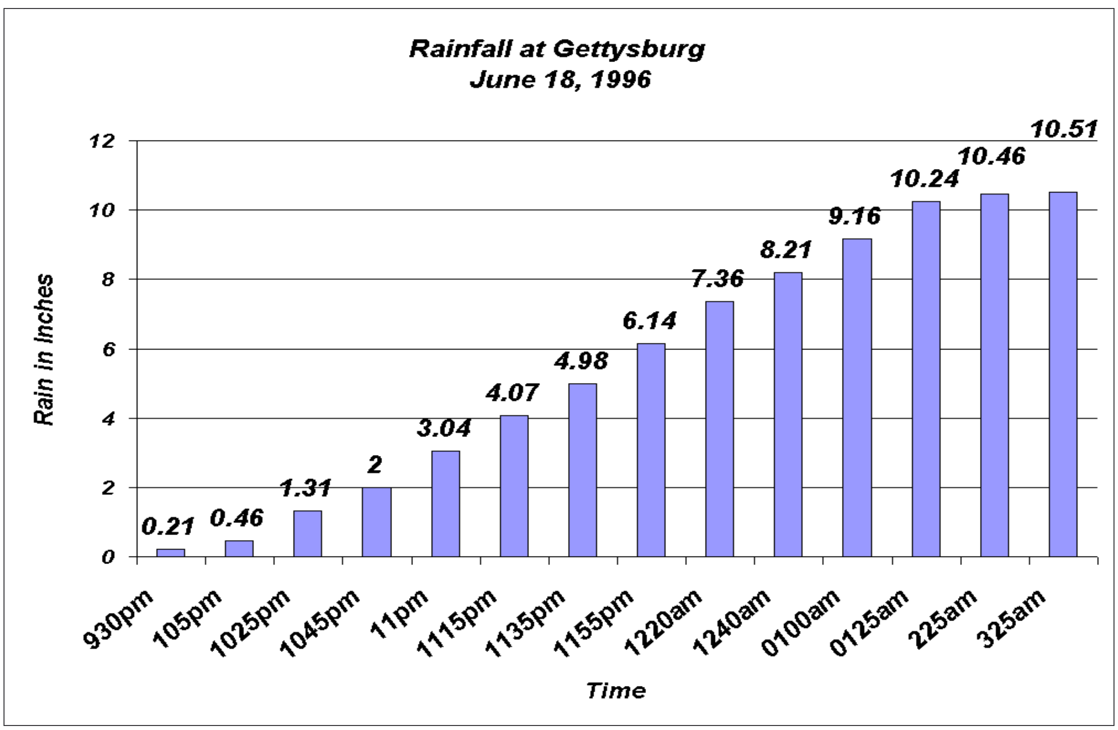


Figure 4.1 Application of Radar in Flash Flood Warnings in PA

4.3.3 Recent Research in the Area of Flood Forecast

Potential gains from using weather radar in flood forecasting have been studied for many years. Most of the research focused on rainfall estimation using radar, and then the estimated rainfall information was used as an input to rainfall-runoff/stream flow model to get stream flow forecast (James, W. P., 1993, Shidawara, M, 2000, Baltas, E.A. and Mimikou, M.A., 2000).

James (1993) showed that both the modified Brandes and Kriging calibration techniques could be used for weather radar calibration and calibrated weather radar was shown to be effective for flood forecasting. A distributed rainfall-runoff model was applied to the Yockanookany watershed in Mississippi equipped with two rain gauges and covered by a radar. Data recorded in October, November, and December of 1986 provided inputs for computing three flood hydrographs from (1) rainfall recorded by rain gauges, (2) radar estimates of rainfall, and (3) combined rain gauge measurements and radar estimates. The hydrograph computed from the combined input was the closest to the observed hydrograph.

Shidawara (2000) made a case study in central Japan to elongate the lead time of runoff forecasting utilizing radar data. The author used the radar data not directly in runoff calculation, but in finding the general tendency of the rain area movement which was utilized to decide the runoff calculation system. A neural network is developed for short-term runoff prediction, making use of the characteristics of rain area movement derived from radar data analysis.

An adaptive grid-square-based flood-forecasting model was set up by Baltas (2000) in the Portaikos river basin with a drainage area of 135 km^2. The entire basin was divided in grid squires of dimensions 2*2 km^2. The raw radar data measured in each grid area was first converted into rainfall estimates by z-R exponential relationship, and then calibrated with rain gage measurement based on gage/radar (G/R) ratio. The radar rainfall estimates in grid squared were used as inputs to the flood-forecasting model; the simulated hydrographs from the model gave relatively close agreement to the observed ones; also the model provided acceptable forecasts with a maximum lead time of 4 hours.

While the weather radar provides superior information about the spatial and temporal resolution of the rainfall event, there are still some disadvantages related to the application of radar rainfall estimation in the field of flood forecasting.

First, the results from past research showed the complexity of the radar application in stream flow forecasting. In order to get stream flow forecasting using radar data; there are two subsequent models needed to be developed: one for radar-rainfall estimation, another for rainfall-streamflow forecasting. Both models need to be calibrated and tested. The complexity of the application not only increases the possibilities of flow forecasting errors, but also prolongs the data process time.

Second, radar does not measure rainfall directly and many sources of error can affect radar rainfall estimation (Zawadzki, 1984). The z-R relationship is believed to be nonlinear and complex, and vary over geographical areas and among storms (Battan 1973). It was not unusual to have radar estimates of rainfall and rain gage measurements made at ground level to differ by a factor of two on some occasions (i.e., 100% error), which making calibration with rain gage measurements necessary (Wilson and Brandes 1979).

Last, rain gauges may be too sparse to properly debias the radar estimates. It is not uncommon that the rain gage networks do not cover the whole drainage area of interest. It will increase the radar rainfall estimates errors, and further affect the accuracy of the flow forecast.

One way to solve the problem is to set up directly relationship of radar and stream flow. There is no such previous research found at present. The reasons are obviously; it is believed that both of the radar vs. rainfall and rainfall vs. runoff relationships are highly non-linear, so it will be very difficult if not possible to setup the mathematical functions for radar-stream flow relationship.

4.4 Methodology

4.4.1 Description of the area and data selection

North Buffalo Creek watershed in Greensboro was selected as study area, see Figure 2.1. The stream flow and precipitation data at Station 02095271, as well as the radar data from National Climatic Data Center (NCDC) were used for radar-based flood-forecasting model setting up. Total six storms were chosen in 2001. The storm summary was shown in Appendix 3.1.

4.4.2 Radar-based Runoff Forecast models Development

Two radar-based flood-forecasting models were investigated. The first one, which did not use radar data directly, was called Radar-Rainfall-Runoff (3R) model, see Figure 4.2. The second one, which provided stream flow forecast using radar data directly, was named Radar-Runoff model (2R), see Figure 4.3.

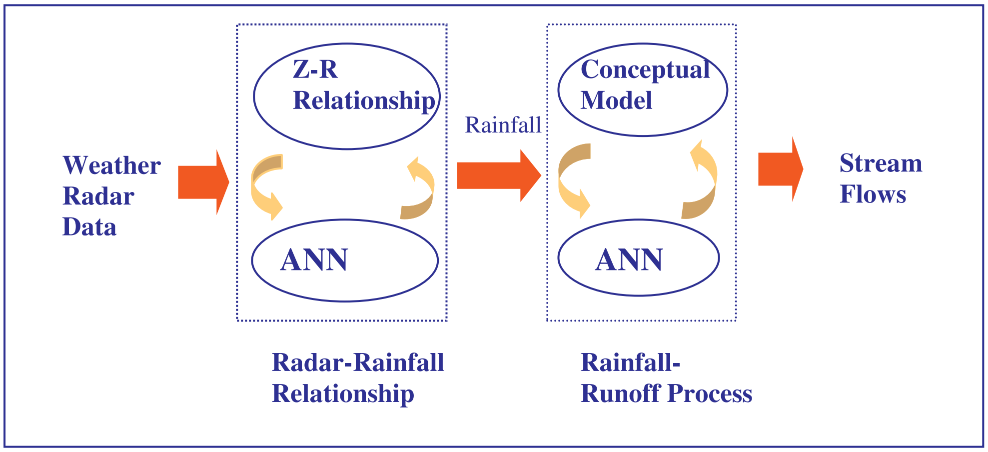


Figure 4.2 Radar-Rainfall-Runoff model (3R Model)

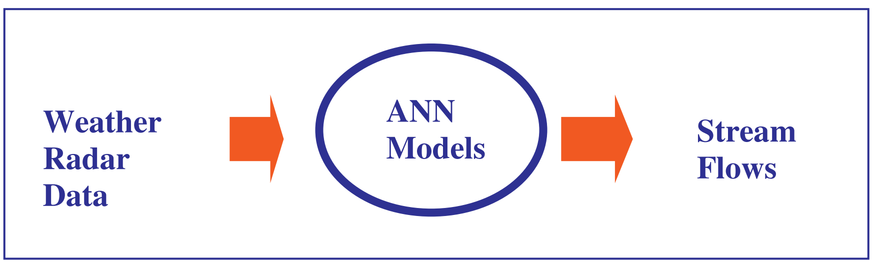


Figure 4.3 Radar-Runoff model (2R Model)

4.4.2.1 Radar-Rainfall-Runoff (3R) Model

There were two steps involved in 3R model: (1) rainfall estimation based on radar, (2) stream flow forecast from radar-based rainfall estimates.

Empirical z-R relationship, which was most widely used approach to provide radar-based rainfall estimates, was given by equation 3.1. The coefficients (a, b) were selected as (26.3, 2.38) for Greensboro based on the previous study.

A HEC-HMS model was proposed as rainfall-runoff model in the study area. The upstream watershed was treated as single basin. The watershed parameters were estimated based on the study of the storms on 6/13/01, 8/11/01 and 9/24/01 and listed in Table 4.1.

Table 4.1 Watershed Parameters in HEC-HMS Model

Parameters	
Method	SCS Curve No. Method
Initial Loss (in.)	0.2
% Impervious	25
SCS Curve No.	70
SCS lag time (min.)	150
Base Flow (cfs)	10
Drainage Area (sq. mi)	14.2

4.4.2.2 Radar-Runoff (2R) Model

Three-layer feed-forward ANN model, which included one input layer, one hidden layer and one output layer was developed for stream flow forecast based on weather radar data.

The design phase of the ANN model was used to estimate the stream flow at the outlet of watershed. The inputs of the model were the reflectivity factors in a 3*3 polar matrix covering the watershed area. The grid resolution of the data is 2 km in range by 1 degree in azimuth, see Figure 4.4. This 14.2-mi^2 watershed has a time-of-concentration, Tc, of approximately 4 hours. Storms occurring within a period equal to Tc would exhibit the greatest influence on stream flows. Consequently, reflectivity, R, of the preceding 4 hours and stream flows, Q, of the preceding 1 hour were taken as inputs to the input layer. Stream flows from the preceding one hour help provide the base flow information prior to the onset of a storm event. The output neurons yield predicted watershed runoff or stream flows, Q' at 1 hour ahead in real time, as summarized in Table 4.2.

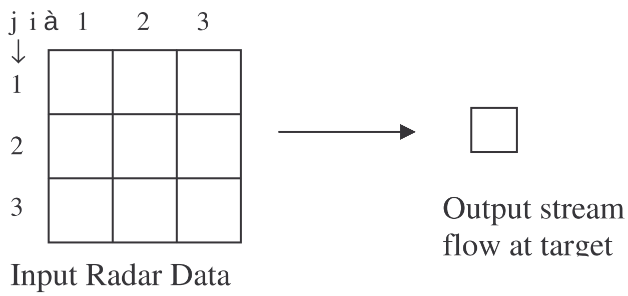


Figure 4.4 Input and Output Selections of ANN Application

Table 4.2 Structure ANN Radar-Runoff (2R) Models

Input	$R_{i,j}(t)$, $R_{i,j}(t - k\Delta t)$; i = 1 to 3, j = 1 to 3, k = 1 to 3 $Q(t)$, $Q(t - \Delta t)$;
Output	$Q'(t + \Delta t)$
Computational time step	Δt = 1 hr

4.5 Results and Discussion

4.5.1 3R Model: Stream Flow Forecast Using Radar Estimates of Rainfall

There are two steps involved in this approach: First, rainfall estimates are computed from radar data by z-R relationship. Second, stream flow predication can be obtained by HEC-HMS model using radar estimates of rainfall.

4.5.1.1 Rainfall Estimating

Figure 4.5 shows the radar-computed and observed hourly precipitation at Station 02095271 in Greensboro. The radar-based rainfall estimation is based on z-R relationship created at previous chapter. The previous study indicated that the precipitation at high intensity storms was underestimated by z-R relationship.

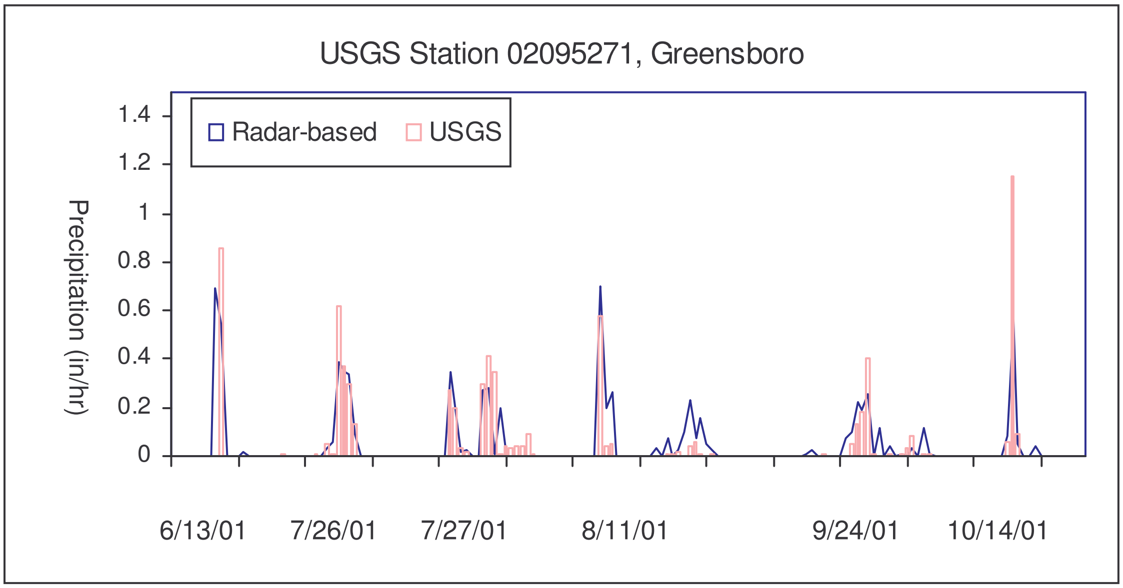


Figure 4.5 Radar-computed Rainfall vs. Rainfall Observation in Greensboro

4.5.1.2 Flood Forecasting

The hydrographs produced using rainfall observation from gage station and radar-computed rainfall estimates were shown in Appendix 3.2.

The results indicated that radar tended to underestimate the rainfall for high intensity storms and overestimate low intensity storms, and accordingly result in an underestimation or overestimate of the peak flow. Figure 4.6 &7 showed such examples. The radar estimate of rainfall for the storm on Oct. 14, 2001 was 0.72 inch, which was 55% of the observation (1.3 inch) from gage station. The hydrograph produced using radar estimate of rainfall gave 42% underestimate of peak flow, and the hydrograph obtained from rain gage observation provided 35% overestimate of peak flow. For the storm on Aug. 11, 2001, the observed storm total is 0.67 and 0.16 inch in the morning and evening, the radar estimate is 1.16 and

0.77, which gave 73% and 381% overestimate respectively; consequently, radar-based hydrograph presented 120% and 236% overestimate of peak flow at these periods. These results indicated that accurate rainfall estimation was critical for peak flow forecasting in rainfall-runoff model. Radar may underestimate or overestimate rainfall amount for different types of storms; consequently it may cause the inaccuracies in stream flow forecast when these results were used as input to rainfall-runoff model.

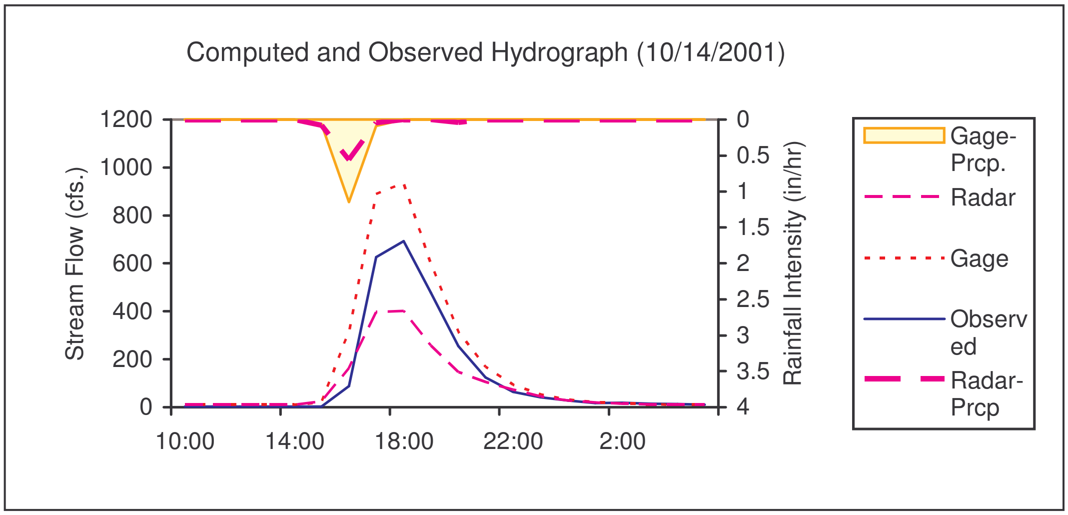


Figure 4.6 Hydrographs generated using gage and radar-based rainfall data (10/14/2001)

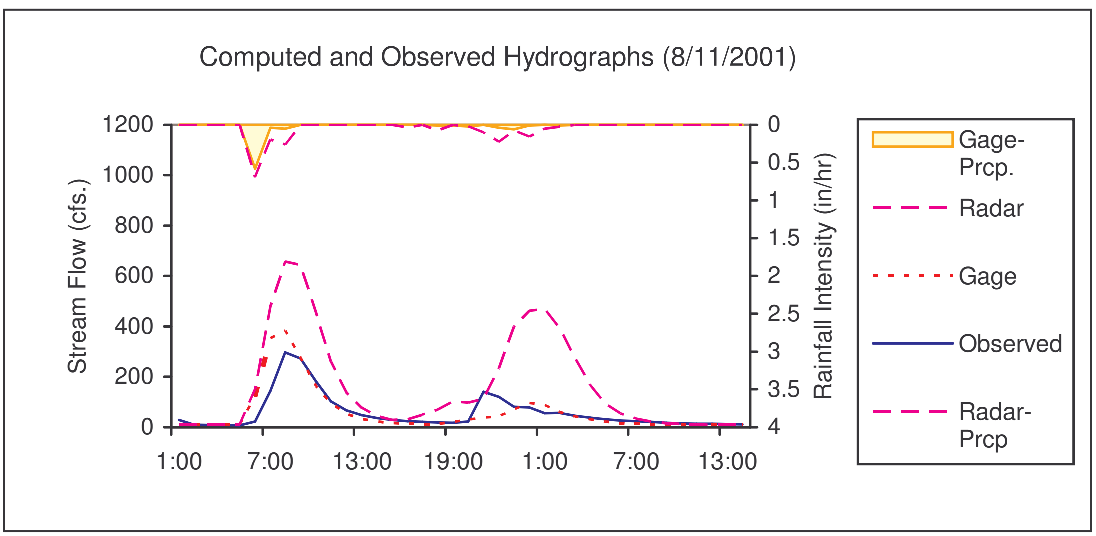


Figure 4.7 Hydrographs generated using gage and radar-based rainfall data (8/11/2001)

Rain gage and radar provide different rainfall information. Rain gage is generally accurate and good for point precipitation measurements; and radar provides spatially distributed rainfall data for whole watershed. While the storm was not uniform in the watershed, the rainfall observation from rain gage might not be satisfied when it was used as input to flood forecasting model. Such an example was shown in Figure 4.8. According to the records from radar, rainfall intensity, which varied from 0.13 to 0.35 in/hr at 5:00 am on July 26, 2001, was more uniform compared with the storm on July 27, 2001, which intensity varied from 0.09 to 0.72 in/hr at 9:00 am. The hydrograph computed from gage rainfall observation gave more accurate peak flow estimation on July 26 than July 27. Comparison of the radar-based hydrograph against gage-based hydrograph showed that radar estimate of rainfall was superior to gage observation while the radar provided average areal rainfall information and the observation from gage gave overestimates of areal precipitation for this non-uniform storm.

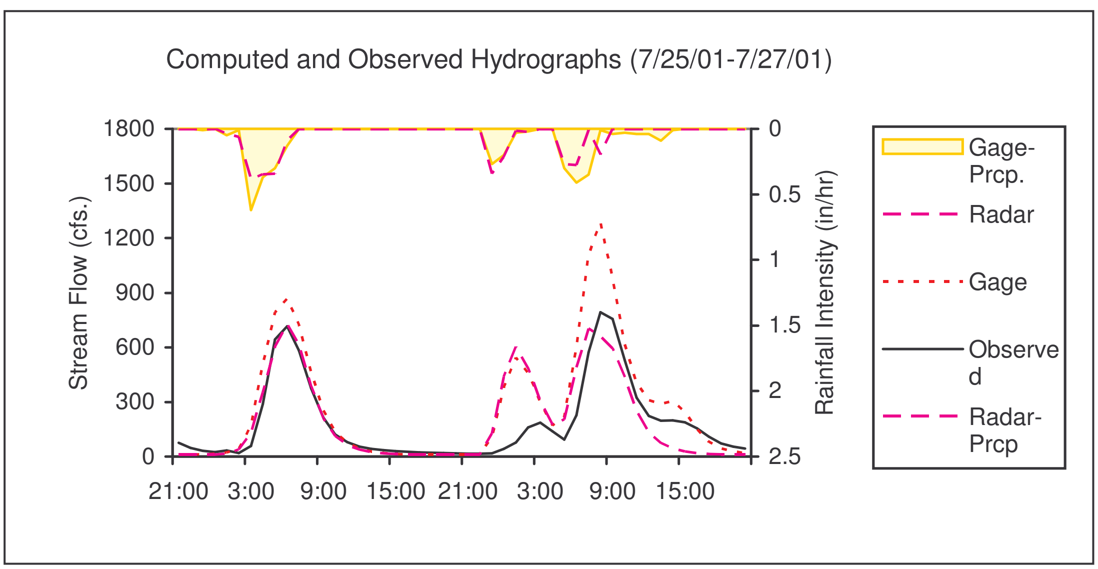


Figure 4.8 Hydrographs generated using gage and radar-based rainfall data (7/25/2001-7/27/2001)

A new radar-based ANN stream flow forecast model, 2R model, has been investigated to provide stream flow forecast using radar data directly. The historical stream flow data and radar reflectivity around the drainage area were served as inputs of the model. The output of the model was stream flow at one hour lead time. The output hydrograph in Figure 4.9 from Radar-Runoff model matched the observed graphs very well for both low and high stream flows. The results showed that with proper design, the ANN model were able to forecast the stream very well using weather radar. More graphs are shown in Appendix 3.3.

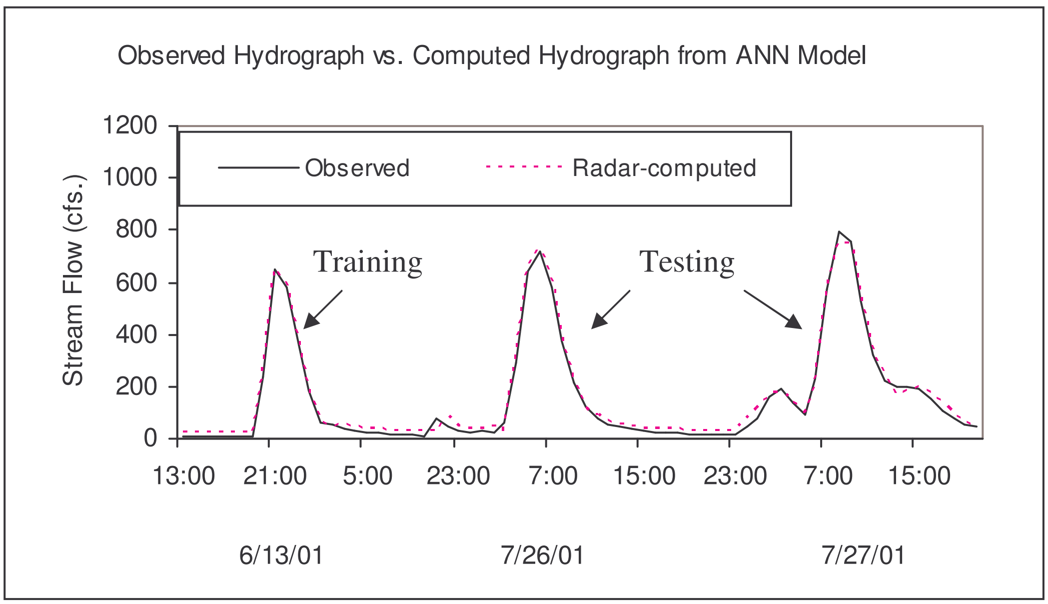


Figure 4.9 Stream Flow Predication by Radar Observation

Comparison of 2R and 3R model showed the advantages of application of 2R model in stream flow forecast, which including (1) Accuracy: It gave more accurate predication of stream flow. (2) Efficiency: It set up radar-stream flow relationship directly. (3) Time Consuming: Shorter computing time and longer predication time can be expected. (4) Cost:

The time involved to 2R model would be significant less; and the operators were expected to require less training process. (5) Data Requirement: No rain gage data was required in 2R model; the model can be applied to the area without rain gage network.

4.6 Conclusion and Recommendation

Two radar-based flood-forecasting models, which were radar-rainfall-runoff (3R) model and radar-runoff (2R) model, were discussed in this paper.

3R model did not use radar data directly. Weather radar was initially employed to compute the rainfall estimates; the estimates of rainfall were then utilized as inputs to conceptual rainfall-runoff model, which provided stream flow forecasts. The research showed some limitations of radar application in stream forecast. For example, the use of radar-based rainfall estimates, computed using z-R relationship, as inputs to hydrological models, may lead to incorrect flow forecasts due to its trend of underestimating rainfall amounts in high intensity storms and overestimating rainfall amounts in low intensity storms. The results also demonstrated the superiority of radar data to gage observation while these data were used to give stream flow predication during non-uniform storms. However, before the radar data can be widely applied for stream flow forecast, many investments are necessary to be made to improve the accuracy of the radar estimates and generate a reliable radar-based stream flow predication model.

A new radar-based stream flow predication algorithm was developed in this paper. The direct relationship of radar and stream flow was setup using ANN model. Historical hourly radar reflectivity and stream flow were introduced as direct input to the model; the model can operated to provide stream flow predication at one hour lead time. Comparison of observed hydrographs against computed hydrographs indicated ANN model was capable of producing accurate stream flow predication. Direct application of radar data gave the advantages of forecasting stream flow more accurate, faster and more efficiently. Even though the results of ANN model indicated the direct relationships of radar data and stream discharge, the model itself did not help understand the principle of the radar and stream flow process. The ANN model developed in this study was just the first phase of radar-runoff model; there are still big opportunities for model improvement. By adding other NexRad level II data, such as wind speed and direction, ANN model can be expected to give more accurate estimates of stream discharge and provide longer predication time.

4.7 Reference:

1. Baltas, E.A. and Mimikou, M.A.,(2000) "A Grid Based Rainfall-Runoff Model for Flood Forecasting" *Proc., 3rd Int. Symp. on Hydro. applications of weather Radar*, ABRH.
2. Battan, L.J., (1973). Radar Observation of the Atmosphere. The University of Chicago Press, 324pp.
3. James, W. P, Bell, J.F, (1993). "Radar-Assisted Real-time Flood Forecasting." *Journal of Water Resources Planning and Management*, Vol. 119, No. 1, January/February, 1993, 32-44.
4. Shidwara, M (2000) "Classification of Rain Area Movements by Radar Observation and Its Uitlization in Runoff Prediction" *Proc., 3rd Int. Symp. on Hydro. applications of weather Radar*, ABRH.
5. Wilson, J.W., and Brandes, E.A., (1979). "Radar measurement of rainfall- A summary." *Bull. Amer. Meteor. Soc.,* 60(Sept.), 1048-1058.

6. Zawadzki I., (1984). “Factors affecting the precision of radar measurements of rain.” *Preprints of Conf. Of Radar Meteorology*, 10-13 Sept. 1984, Zurich, AMS, 251-256.

Chapter V: Summary of ANN Applications and Recommendations

In this study ANN models were applied to three studies of hydrologic processes, including rainfall-runoff process, radar-rainfall process, and radar-runoff process.

First, ANN models have been developed for watershed runoff and stream flow prediction based on the studies on two urbanized areas in North Carolina. The results demonstrated that ANNs could provide better runoff prediction than conceptual models and standard regression models. ANN models were developed directly from data, therefore significant amount of historical rainfall and runoff data were required for the neural network models training and testing processes. The vast amount of meteorological and stream flow data which are routinely being collected by the USGS or other federal, state, and local government agencies, can be organized and synthesized for ANN model applications. The real-time stream flow and precipitation data can be obtained quickly by calling USGS stations or accessed by internet connection. It makes it possible of application of ANN models for real-time stream flow prediction at watershed outlet.

Second, ANN models have shown the abilities to estimate the precipitation on the ground according to the radar information. Weather radar is an important source of rainfall estimation. More than 160 WRS-88D radars were installed and covered whole area of United States. But the quality of weather radar measurements of precipitation was always questioned especially when these measurements were used for rainfall-runoff modeling. Currently, most NEXRAD data reading software, such as WATADS, used empirical z-R relationships to derive precipitation reading. The results of this study demonstrated that ANN models could provide significant improvement in precipitation estimation compared to the widely used empirical z-R relationships. It is particularly useful for application of NEXRAD data to real-time flood prediction. The weather radar can be accessed in real-time easily; the accurate real-time areal precipitation estimation provided by ANNs is particularly useful in timely and early warning of flash floods in urban areas.

Finally, a new radar-runoff (2R) model has been developed using ANN. By introducing radar data, this 2R model was capable of producing accurate stream flow predication. Because of the accessibility of the NEXRAD radar data, this 2R model was only trained and tested by six storms. In the future work, more storms will be tested to verify the reliability of the model. NEXRAD II data are high-resolution base data (reflectivity, mean radial velocity and spectrum width) produced at the full spatial and temporal resolution of WSR-88D system. It provides the information of precipitation, temperature, wind speed and direction, etc. All these information can be accessed in real-time easily. All these information combined with the flow information obtained from USGS stations can be used as inputs to ANN radar-runoff model, which provides real-time flood warning of the watershed.

Appendix

Appendix 1.1 a: The summary of analyzed storms in Greensboro, NC

Storms For Training								
NO	Date	Frequency (Year)	STATION 02095181		STATION 02095271		STATION 02095500	
			Rainfall (in.)	Duration (hr)	Rainfall (in.)	Duration (hr)	Rainfall (in.)	Duration (hr)
1	1/19/01	<1	1.1	34	1.37	33	1.19	33
2	2/16/01	1	2.07	46	2.19	47	2.06	46
3	6/13/01	1	0.92	3	0.86	3	0.05	2
4	7/4/01	3	2.63	5	2.92	5	2.46	5
5	7/27/01	<1	0.73	15	1.84	16	1.22	16
6	9/24/01	<1	0.97	15	0.92	15	0.77	15
7	10/14/01	1	1.38	2	1.3	2	1.07	2

Storms For Testing								
NO	Date	Frequency (Year)	STATION 02095181		STATION 02095271		STATION 02095500	
			Rainfall (in.)	Duration (hr)	Rainfall (in.)	Duration (hr)	Rainfall (in.)	Duration (hr)
1	5/26/01	<1	1.32	5	1	7	2.3	7
2	7/26/01	<1	1.62	8	1.8	8	1.6	7
3	12/10/01	<1	1.3	15	1.45	15	1.29	14

Appendix 1.1 b: The summary of analyzed storms in Raleigh, NC

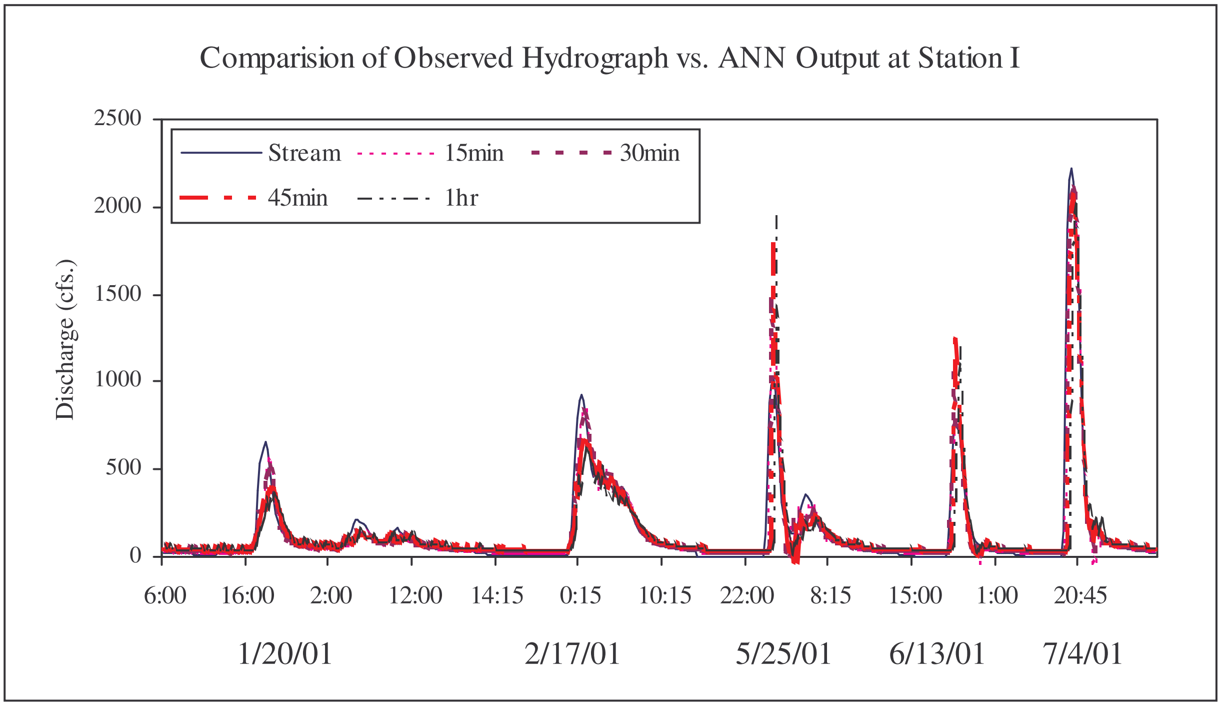
Storm	Date	Precipitation (in)	Duration (hr)	Frequency (year)	Max. Peak Gage Height (ft.)		
		0208732885			Crabtree-S1	Crabtree-S2	Crabtree-S3
1	1/15/98	2.54	46	<1	15	13.95	11.75
2	1/23/98	1.49	35	<1	11.34	10.05	8.86
3	1/27/98	1.63	6	<1	14.72	13.53	11.55
4	2/16/98	1.78	24	<1	15.15	13.95	11.66
5	3/9/98	2.00	41	<1	14.59	13.35	11.2
6	3/17/98	3.35	42	1	18.58	17.62	14.08
7	7/4/98	2.86	5	2	12.22	11.37	9.82
8	9/3/98	2.38	18	<1	7.2	6.25	5.72
9	3/21/99	1.51	16	<1	7.8	6.72	5.74
10	8/21/99	2.22	2	2	7.63	7.01	4.16
11	9/4/99	7.86	56	5	16.17	15.24	13.18
12	9/15/99	7.21	34	25	21.04	19.93	16.88
13	9/27/99	5.35	14	10	14.69	13.54	12.69
14	10/17/99	2.64	17	<1	7.65	6.28	7.01
15	12/13/99	1.71	8	<1	10.05	8.48	8.67
16	1/10/00	2.29	7	1	10.18	8.85	8.79
17	7/15/00	1.77	2	1	7.29	6.1	7.17
18	9/3/00	1.15	6	<1	6.67	5.35	5.79
19	9/26/00	1.50	45	<1	11.32	10.29	9.87

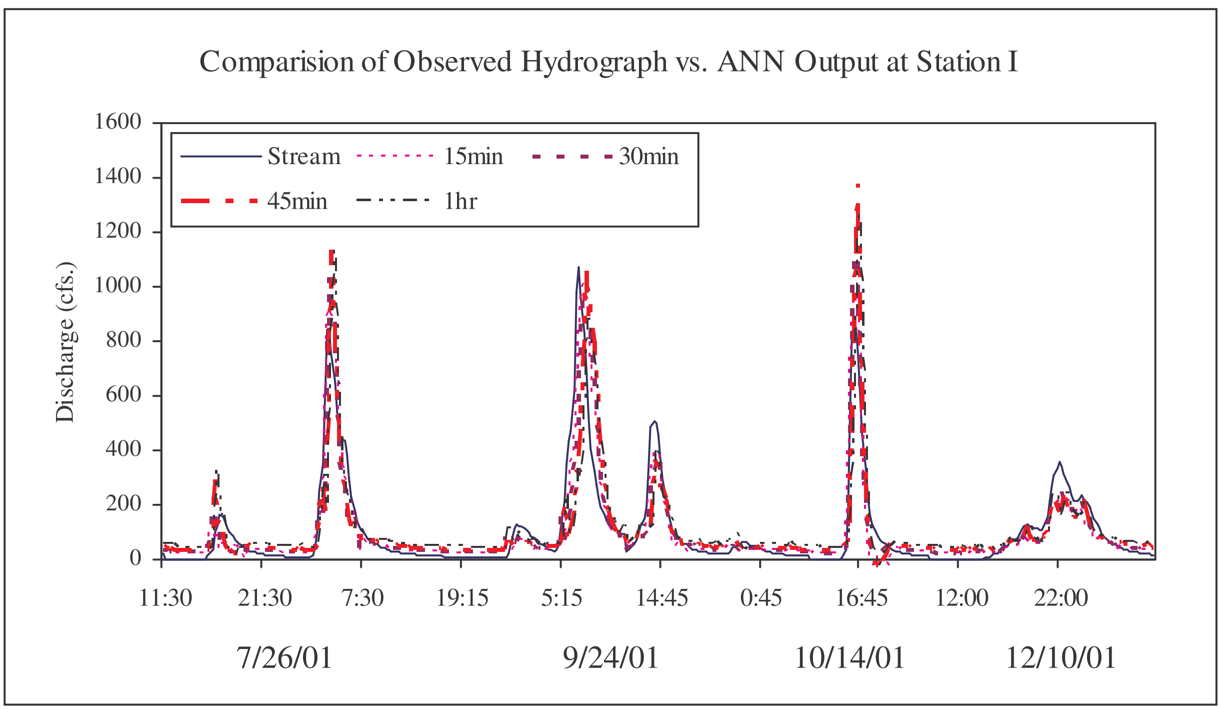
Appendix 1b: The summary of testing storms in Raleigh, NC

Storm	Date	Precipitation (in)	Duration (hr)	Frequency (yr)	Max. Peak Gage Height (ft.)					
		0208732885			Crabtree-S1		Crabtree-S2		Crabtree-S3	
					Observed	*ANN Model	Observed	*ANN Model	Observed	*ANN Model
1	3/29/01	1.78	19	<1	10.81	10.8	9.68	9.55	9.26	9.34
2	3/31/01	1.77	27	<1	11.91	11.97	10.74	10.63	10.12	10.14
3	6/1/01	1.91	10	1	11.2	10.72	10.16	10.51	10.17	10.83
4	7/4/01	1.11	4.5	<1	4.56	5.02	3.26	3.51	4.05	5.21
5	12/10/01	1.55	17	<1	7.78	8.06	6.39	6.8	6.58	6.93

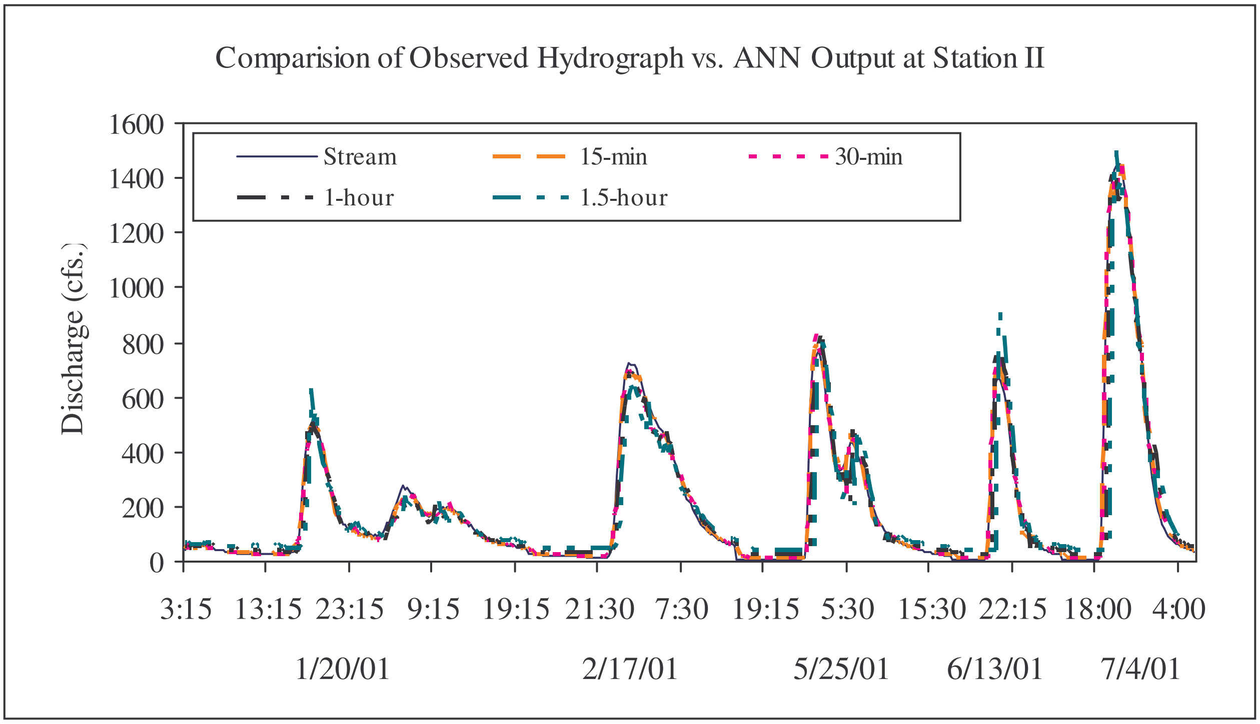
* 30 min Predication from ANN WGH Model

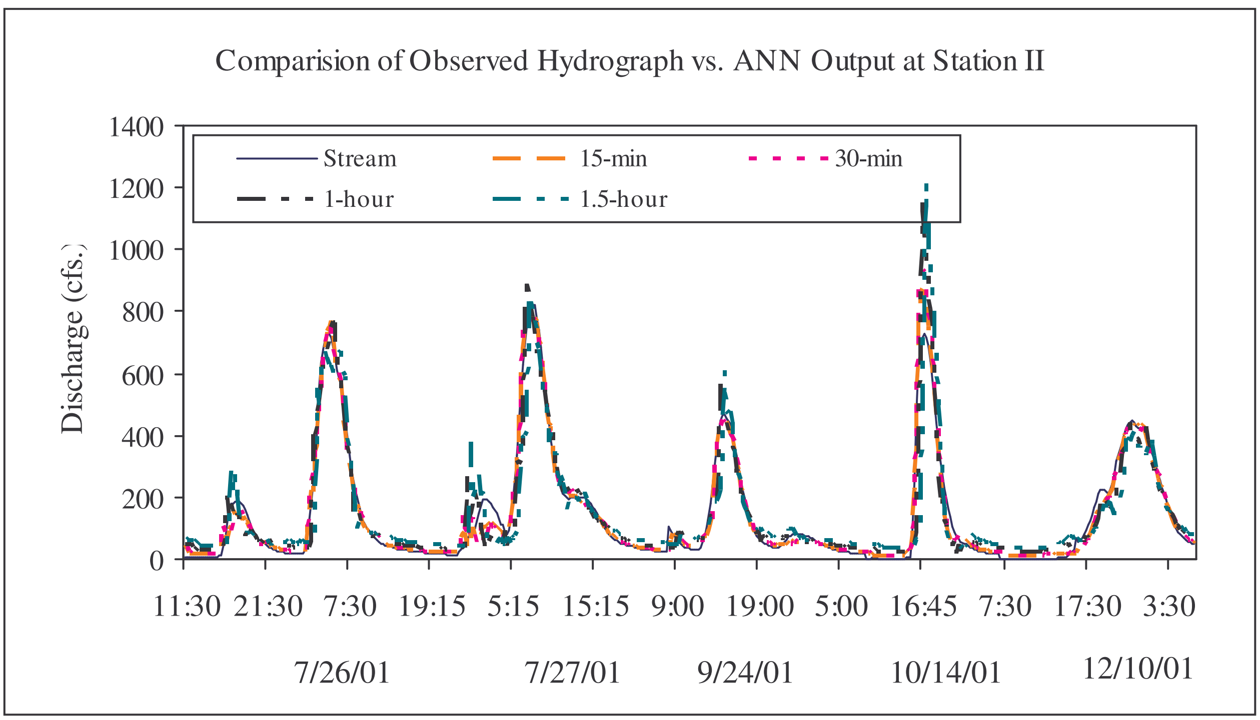
Appendix 1.2: ANN Output Hydrographs at Station-I in Greensboro, NC



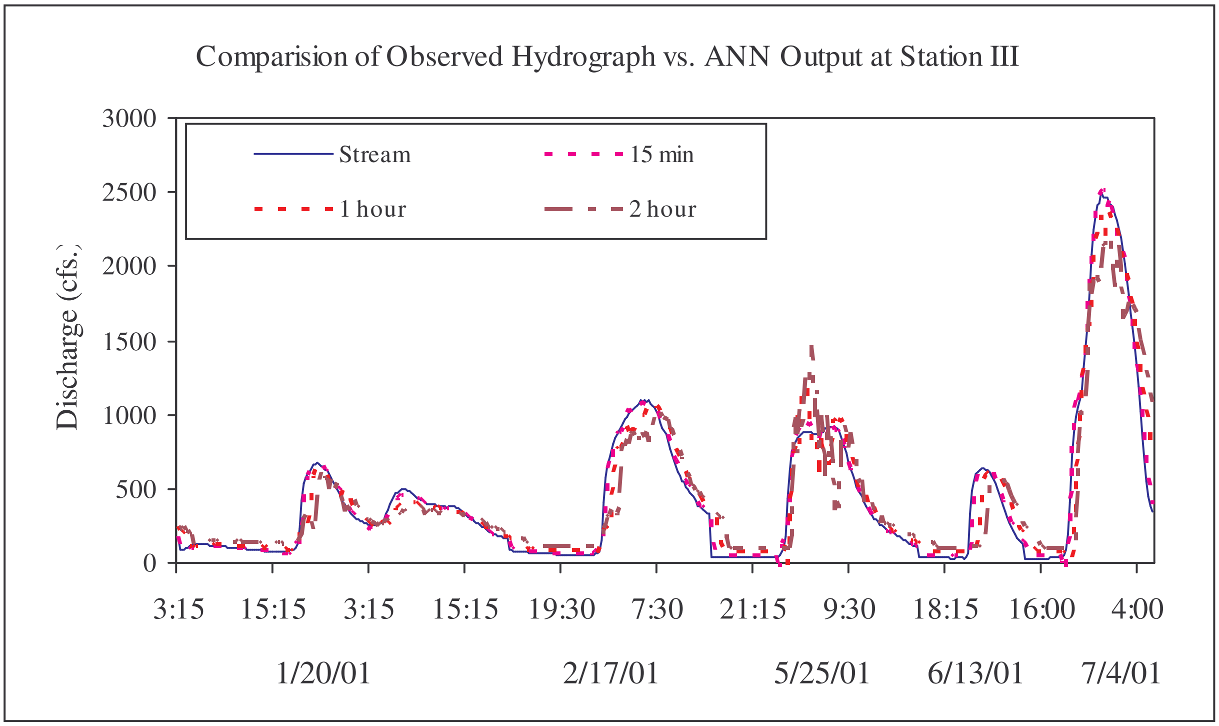


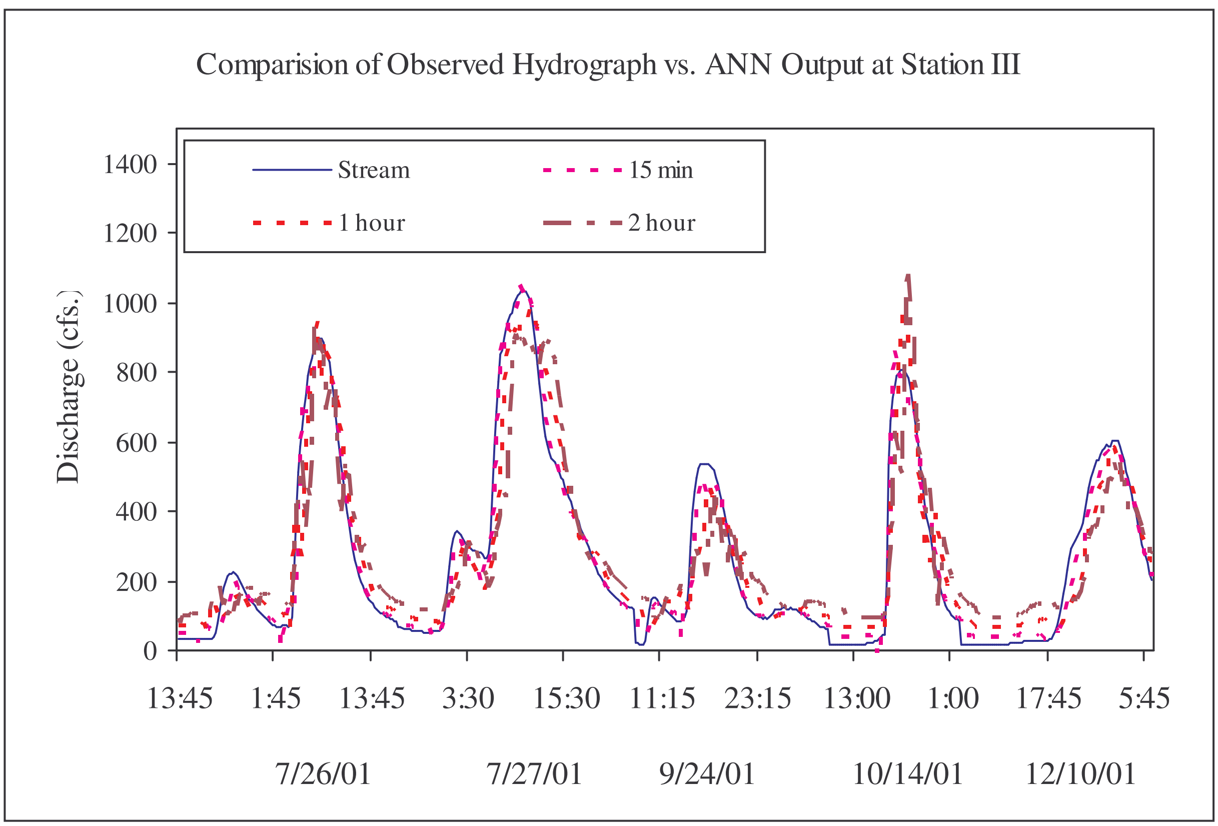
Appendix 1.3: ANN Output Hydrographs at Station-II in Greensboro, NC



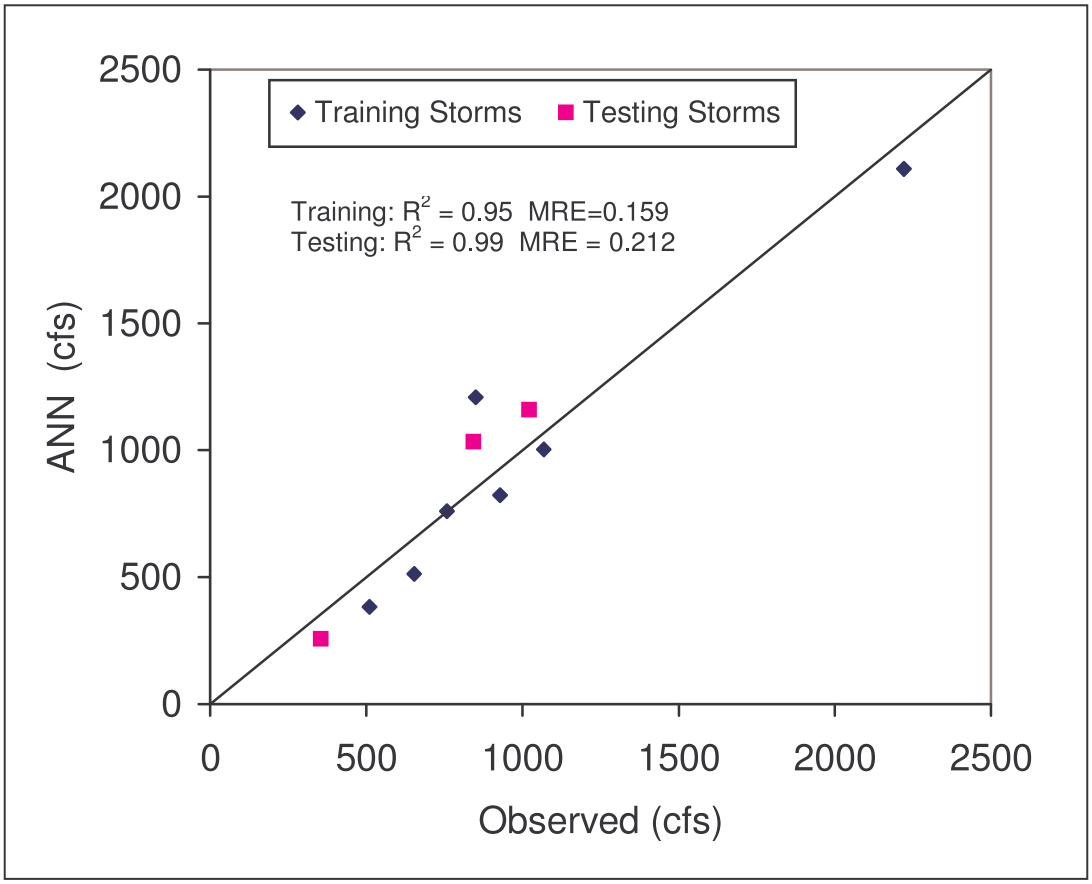


Appendix 1.4: ANN Output Hydrographs at Station-III in Greensboro, NC

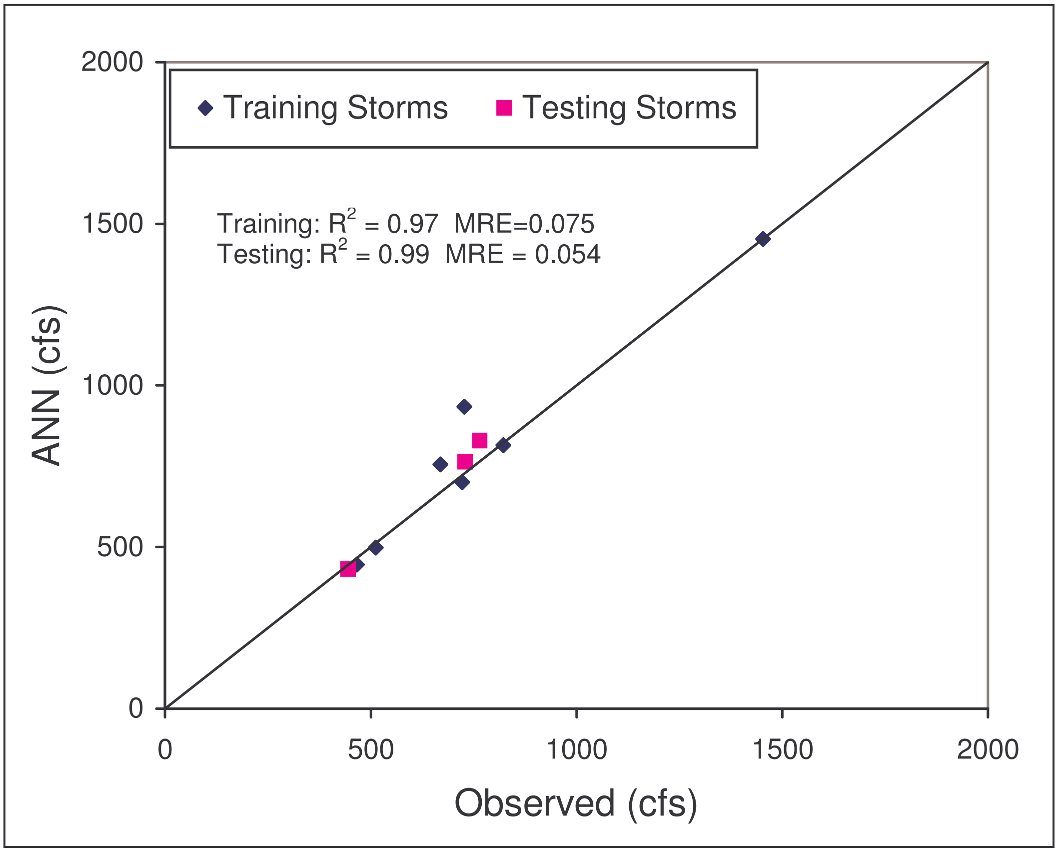




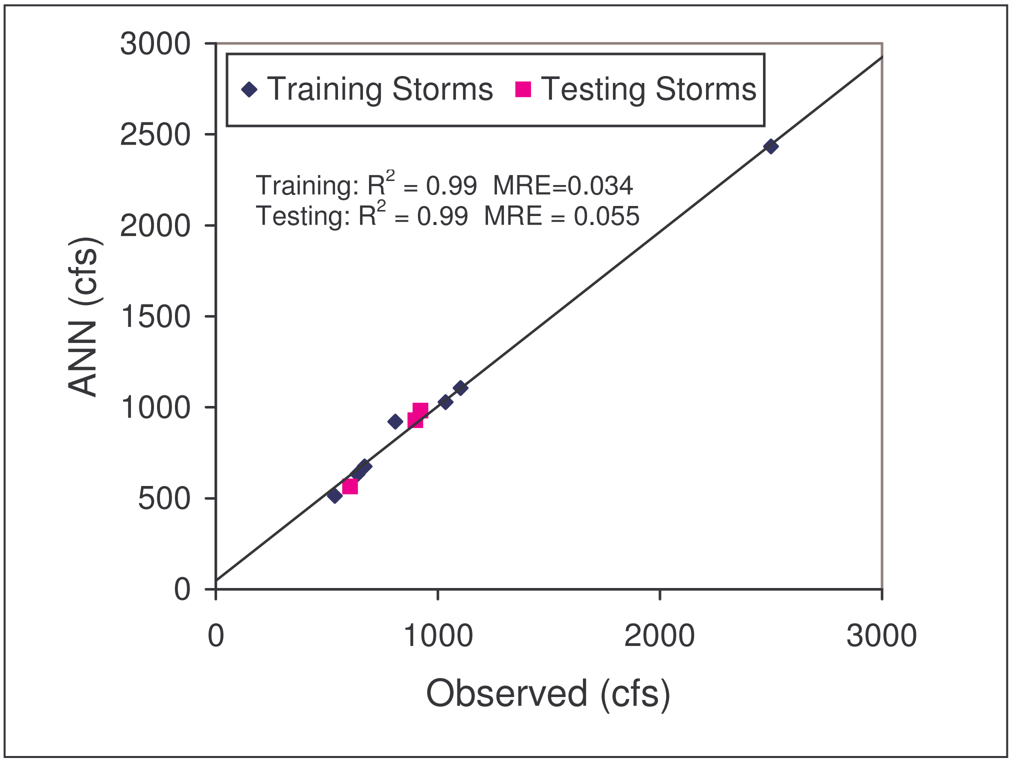
Appendix 1.5a: Comparison of Observed and ANN Peak Flow in Station-I



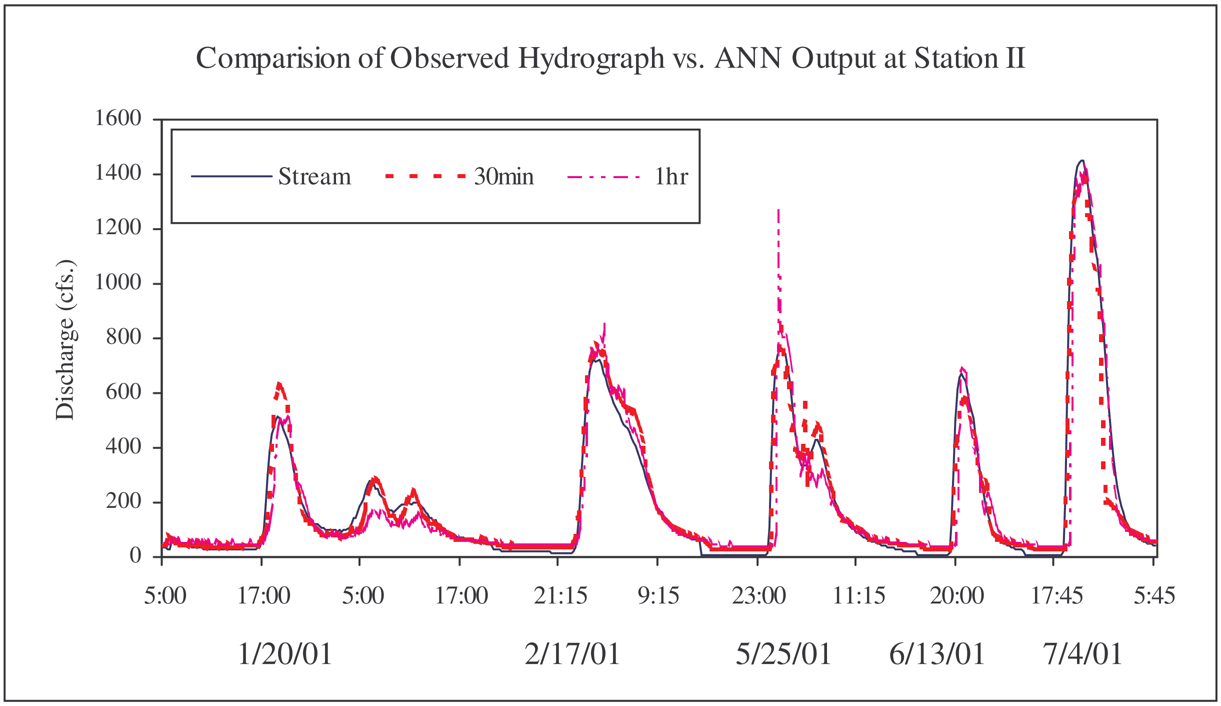
Appendix 1.5b: Comparison of Observed and ANN Peak Flow in Station-II

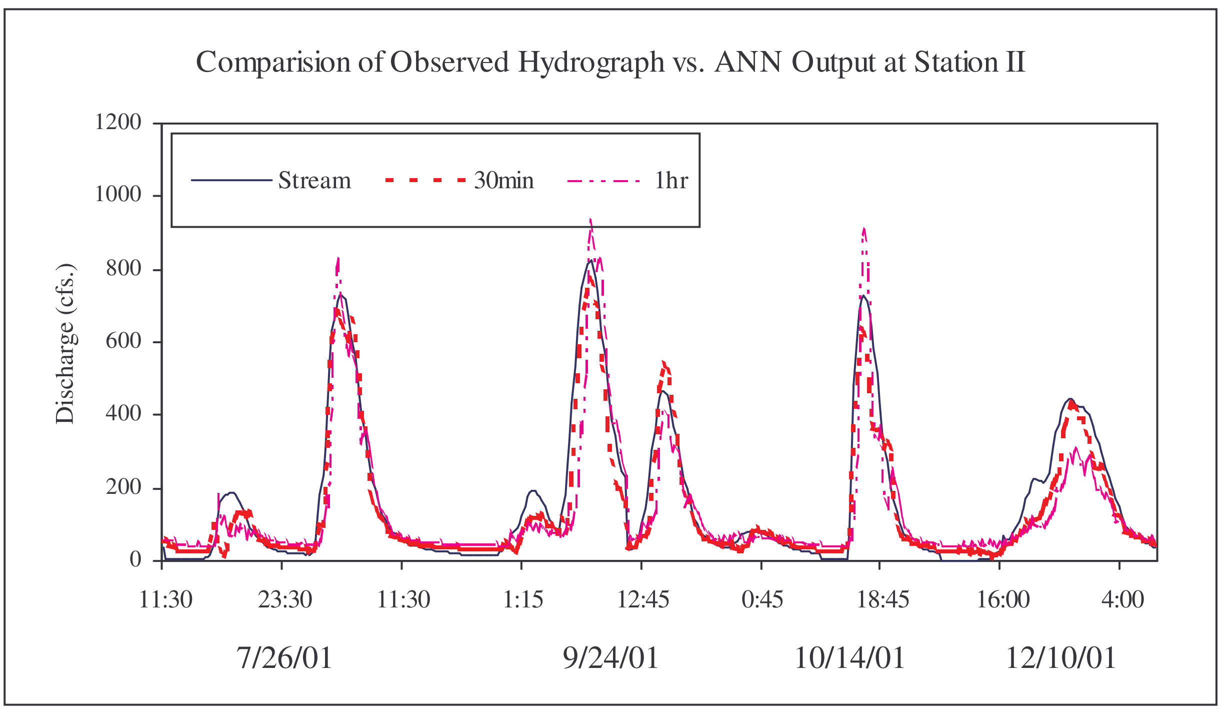


Appendix 1.5c: Comparison of Observed and ANN Peak Flow in Station-III

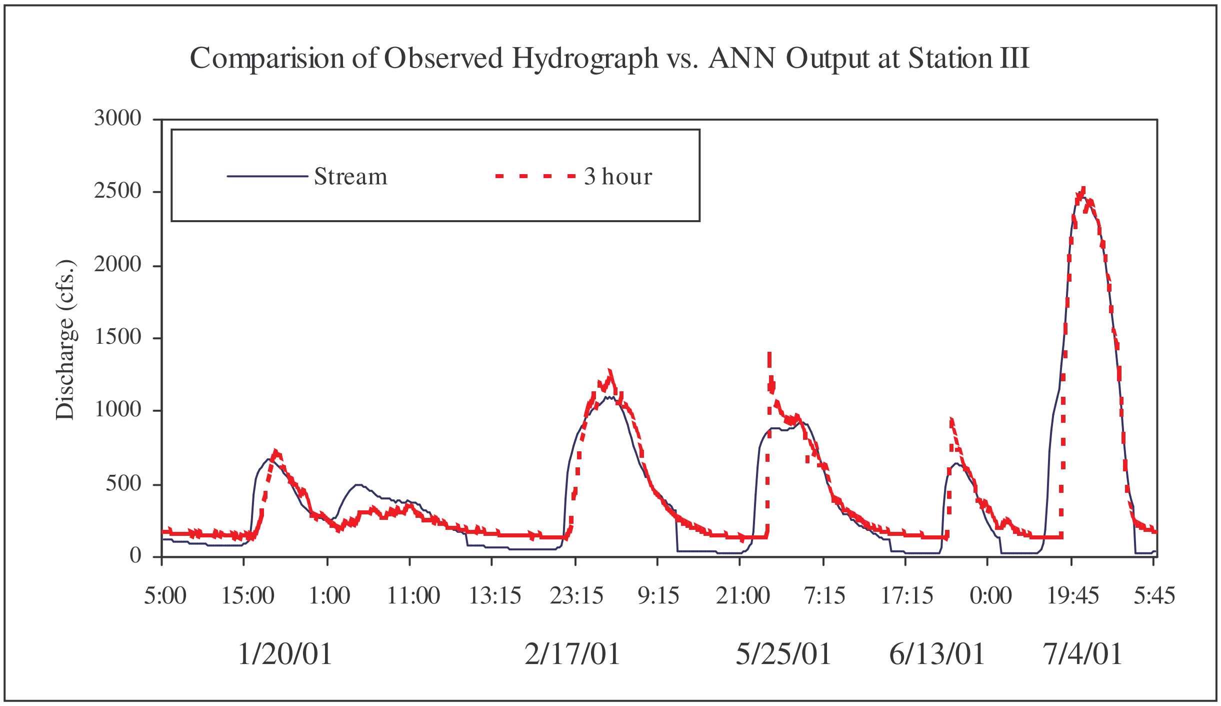


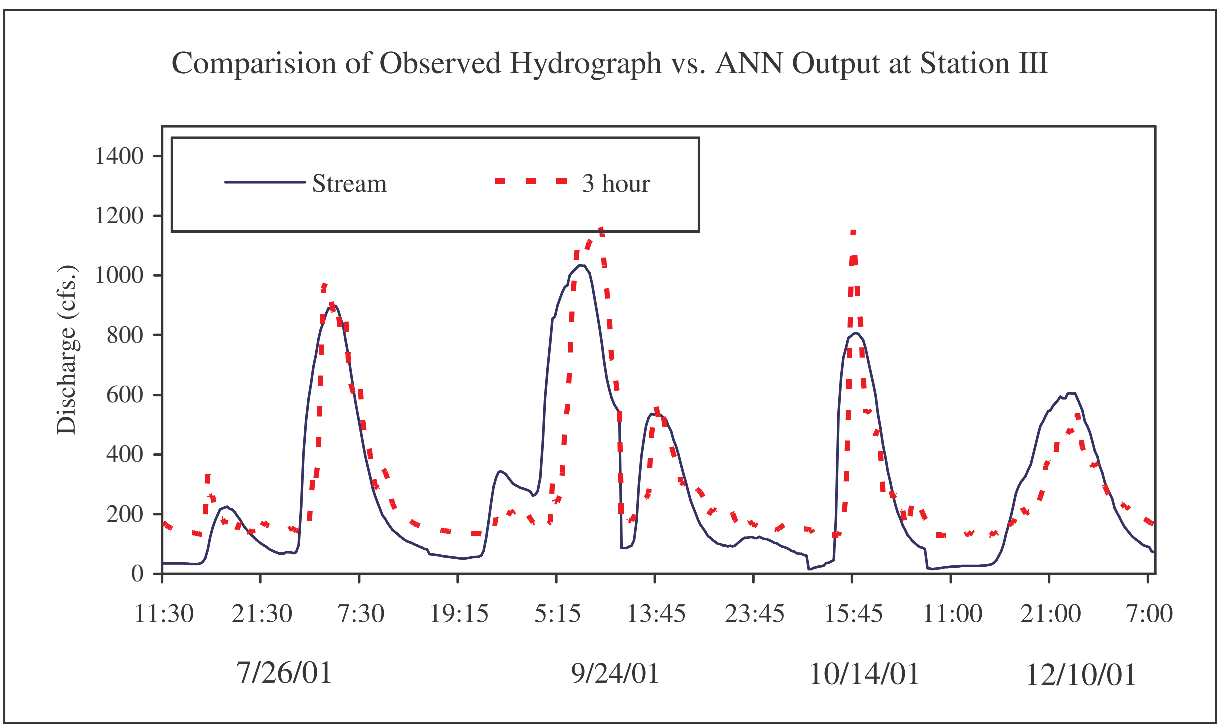
Appendix 1.6: Downstream Flow Prediction Results at Station II





Appendix 1.7: Downstream Flow Prediction Results at Station III

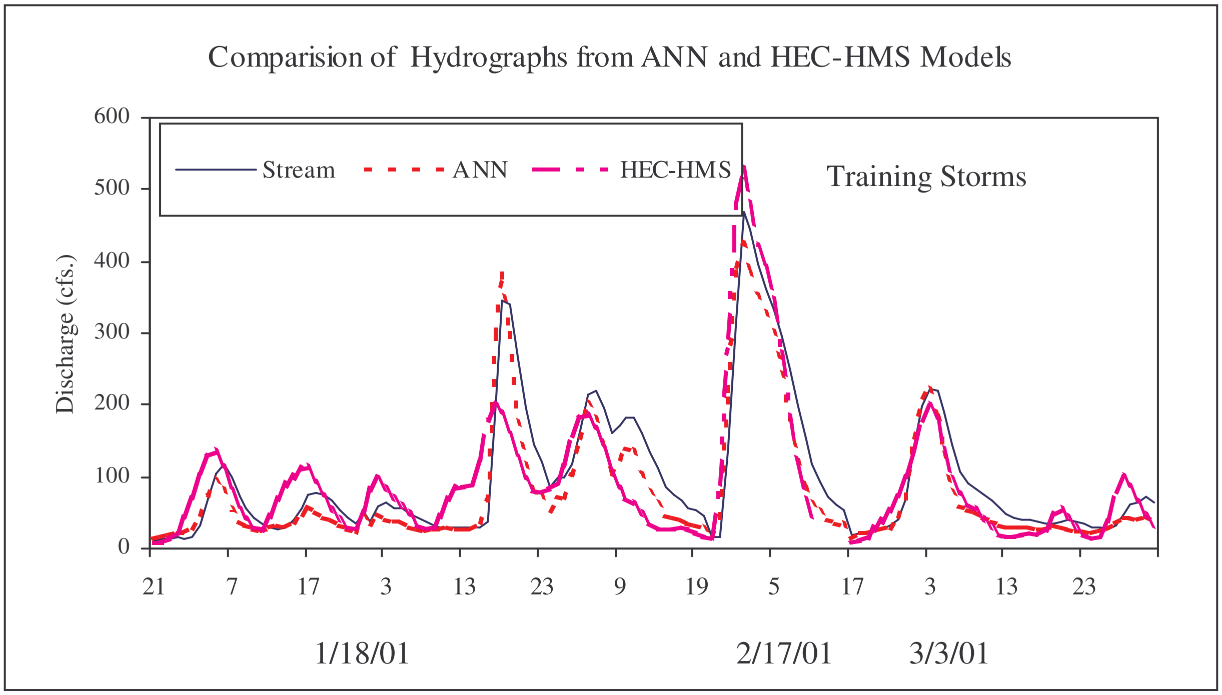


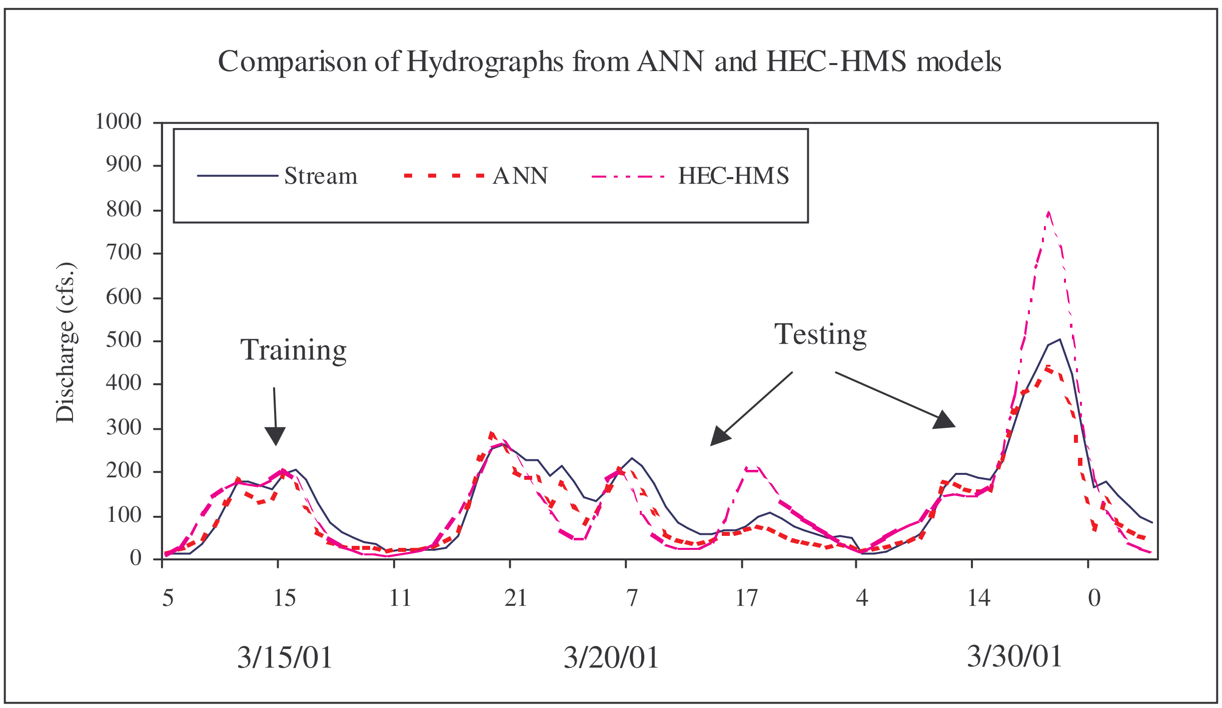


Appendix 1.8: Selected Storms for ANN and HEC-HMS Models Comparison

Date	Duration (hr)	Total Precipitation (in.)
1/18/01	66	1.78
2/17/01	8	1.12
3/3/01	36	0.73
3/15/01	10	0.66
3/20/01	36	1.2
3/30/01	17	1.52

Appendix 1.9: Comparison of Hydrographs from ANN and HEC-HMS models





Appendix 1.10: The Selection of ANN and Empirical Models with Different Variables

ANN Model:

No. of Variables	Input	Output	No. of neurons in Hidden Layer
1	S	S+Δt	3
2	S, P	S+Δt	3
3	S, P, P-1	S+Δt	4
4	S, P, P-1, P-2	S+Δt	5
5	S, P, P-1, P-2, P-3	S+Δt	6

Empirical Model 1:

No. of Variable	Functions
1	$S(i+\Delta t) = a + b*S(i)$
2	$S(i+\Delta t) = a + b*S(i) + c*R(i)$
3	$S(i+\Delta t) = a + b*S(i) + c*R(i) + d*R(i-1)$
4	$S(i+\Delta t) = a + b*S(i) + c*R(i) + d*R(i-1) +e*R(i-2)$
5	$S(i+\Delta t) = a + b*S(i) + c*R(i) + d*R(i-1) +e*R(i-2) + f*R(i-3)$

Empirical Model 2:

No. of Variable	Functions
1	$S(i+\Delta t) = a + b*S(i)^c$
2	$S(i+\Delta t) = a + b*S(i)^c + d*R(i)^e$
3	$S(i+\Delta t) = a + b*S(i)^c + d*R(i)^e + f*R(i-1)^g$
4	$S(i+\Delta t) = a + b*S(i)^c + d*R(i)^e + f*R(i-1)^g +h*R(i-2)^i$
5	$S(i+\Delta t) = a + b*S(i)^c + d*R(i)^e + f*R(i-1)^g +h*R(i-2)^i + j*R(i-3)^k$

S(i+Δt): Stream flow predication at Δt lead time.
S(i): Current stream flow.
S(i-k): Previous stream flow at lag time of k.

Appendix 1.11: The Calibration Results of Empirical Models

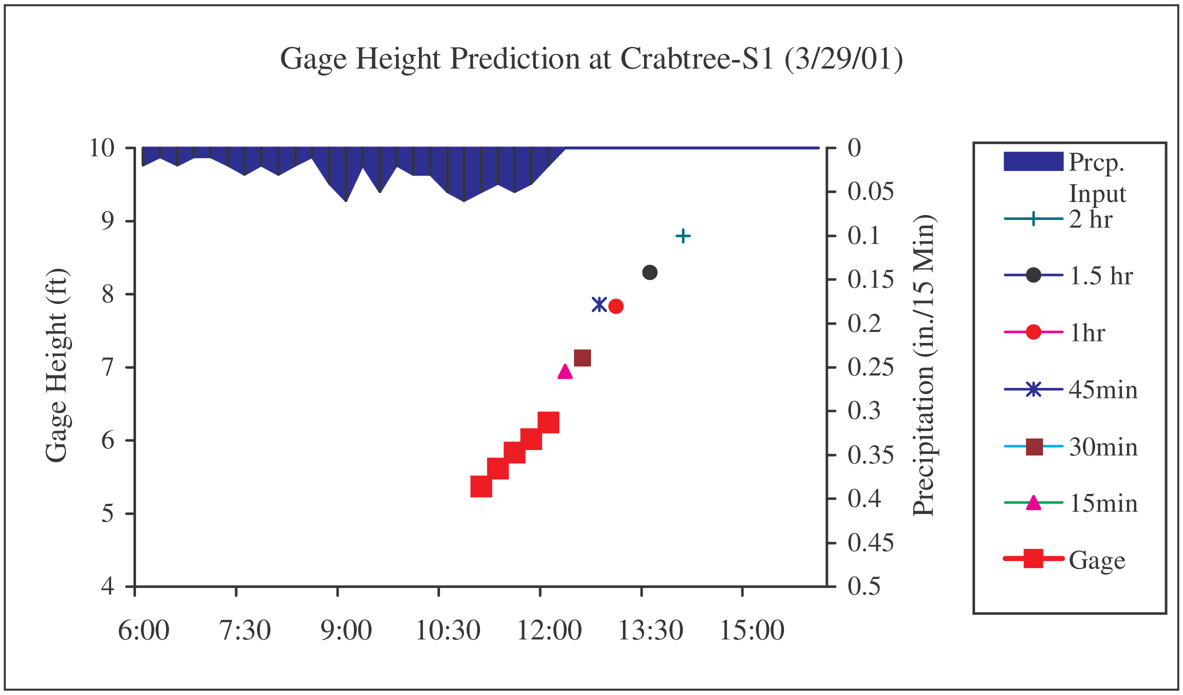
Empirical Model 1: $S(i+t) = a + b*S(i) + c*R(i) + d*R(i-1) + e*R(i-2) + f*R(i-3)$

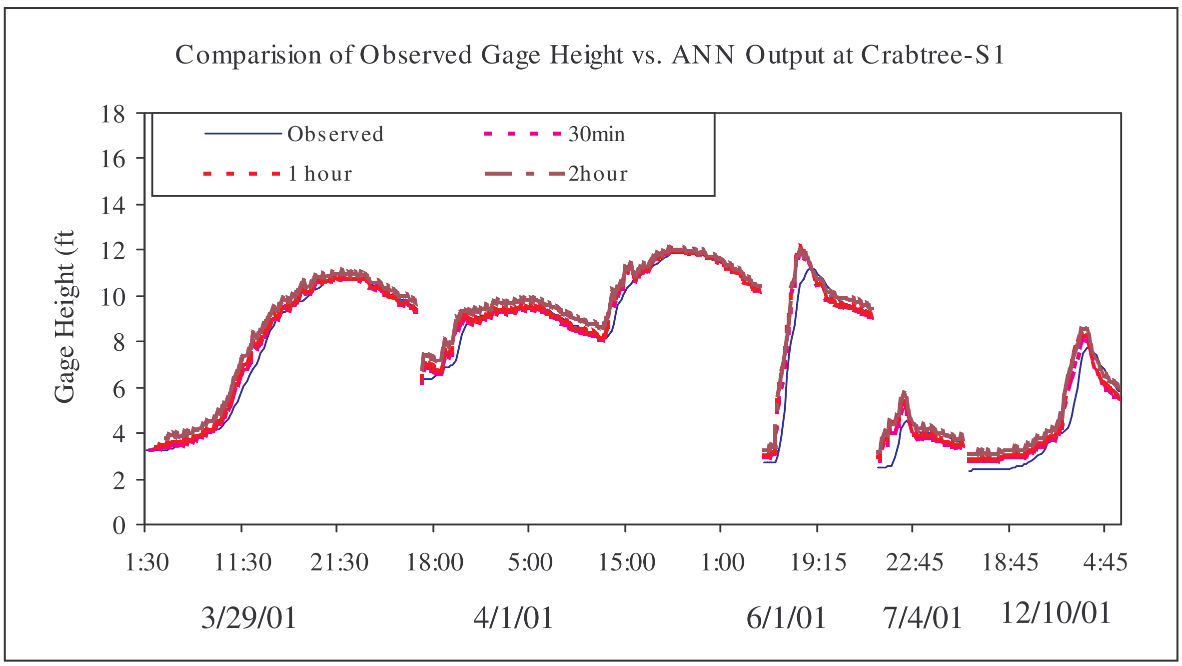
Variable	1	2	3	4	5
a	23.90475	12.34981	14.42631	16.52943	18.5909
b	0.847619	0.761512	0.717111	0.673306	0.624819
c		1317.507	1130.277	1152.717	1178.599
d			439.3532	319.5532	357.6721
e				342.9606	254.25
f					310.673

Empirical Model 2: $S(i+t) = a + b*S(i)^c + d*R(i)^e + f*R(i-1)^g + h*R(i-2)^i + j*R(i-3)^k$

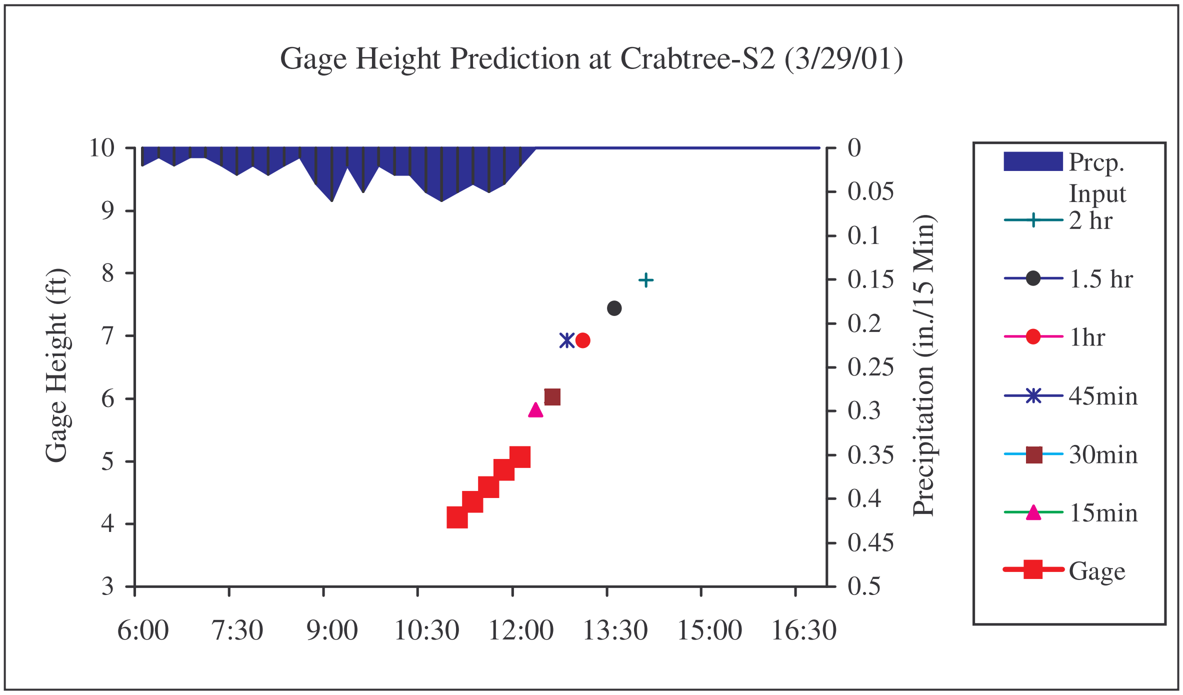
Variable	1	2	3	4	5
a	18.6216	10.89211	11.88696	11.35294	10.37157
b	1.05119	0.566744	0.532421	0.594687	0.703761
c	0.97017	1.039678	1.039822	1.016176	0.982315
d		1234.019	1081.19	1112.555	1146.65
e		0.823176	0.923836	0.948396	0.961806
f			390.1757	288.418	332.1094
g			0.668604	0.734057	0.803257
h				300.0666	235.4385
i				0.744038	0.904545
j					286.862
k					0.778771

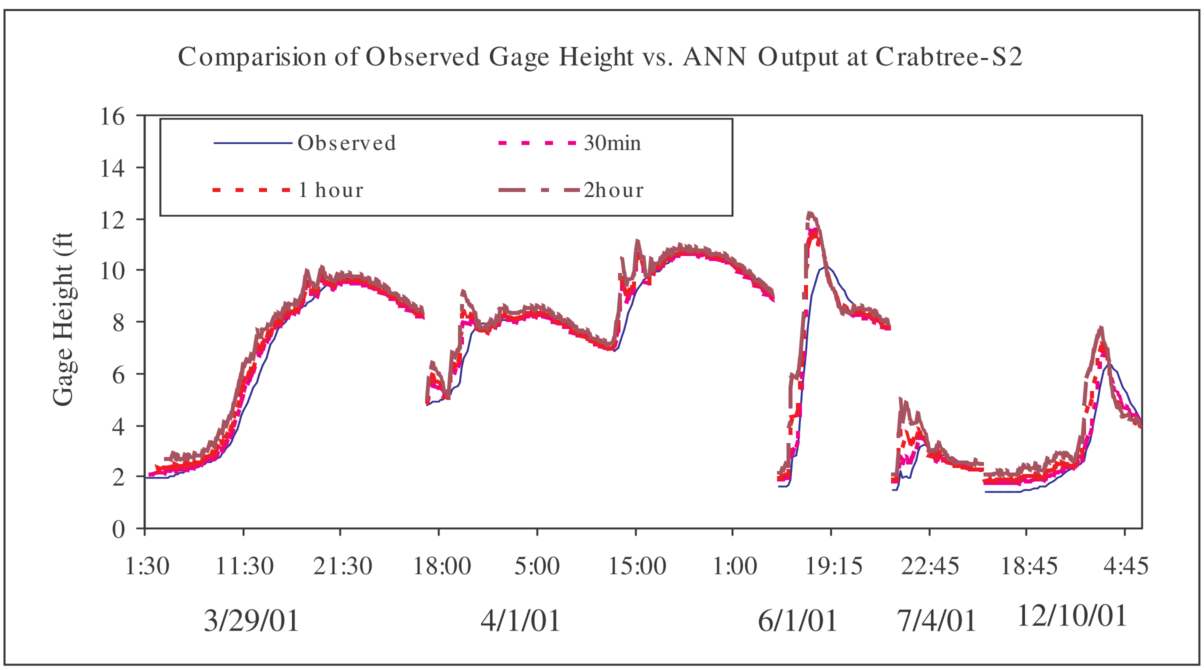
Appendix 1.12: Gage Height Prediction at Station Crabtree-S1, Raleigh NC



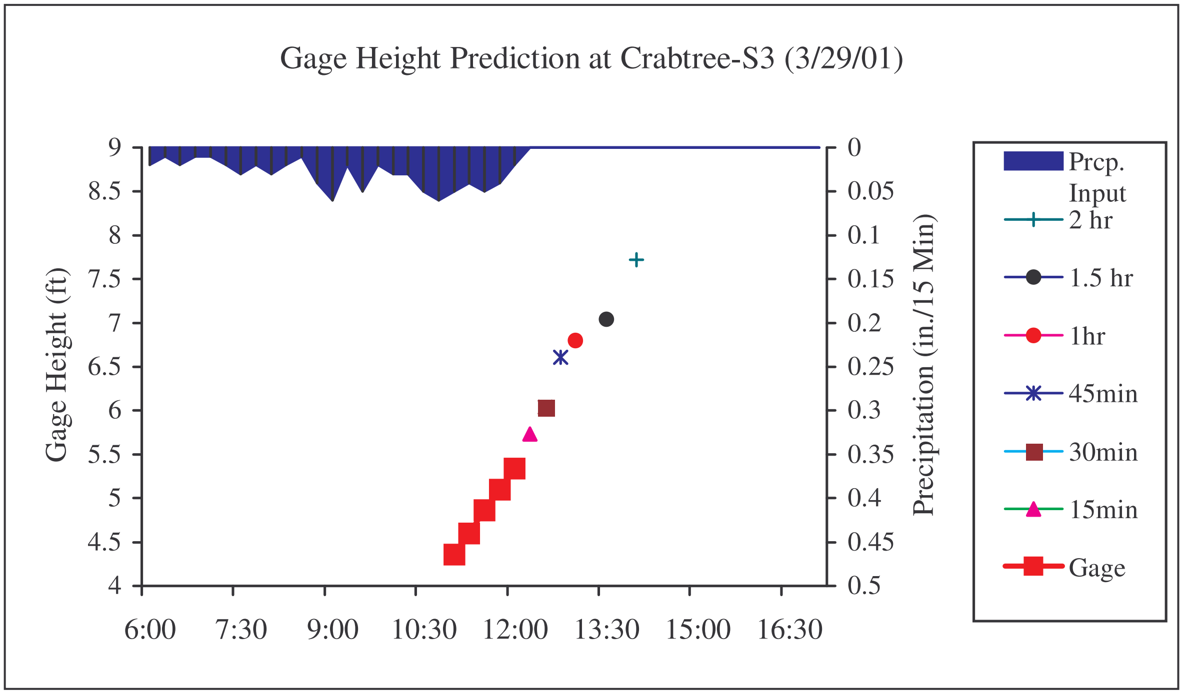


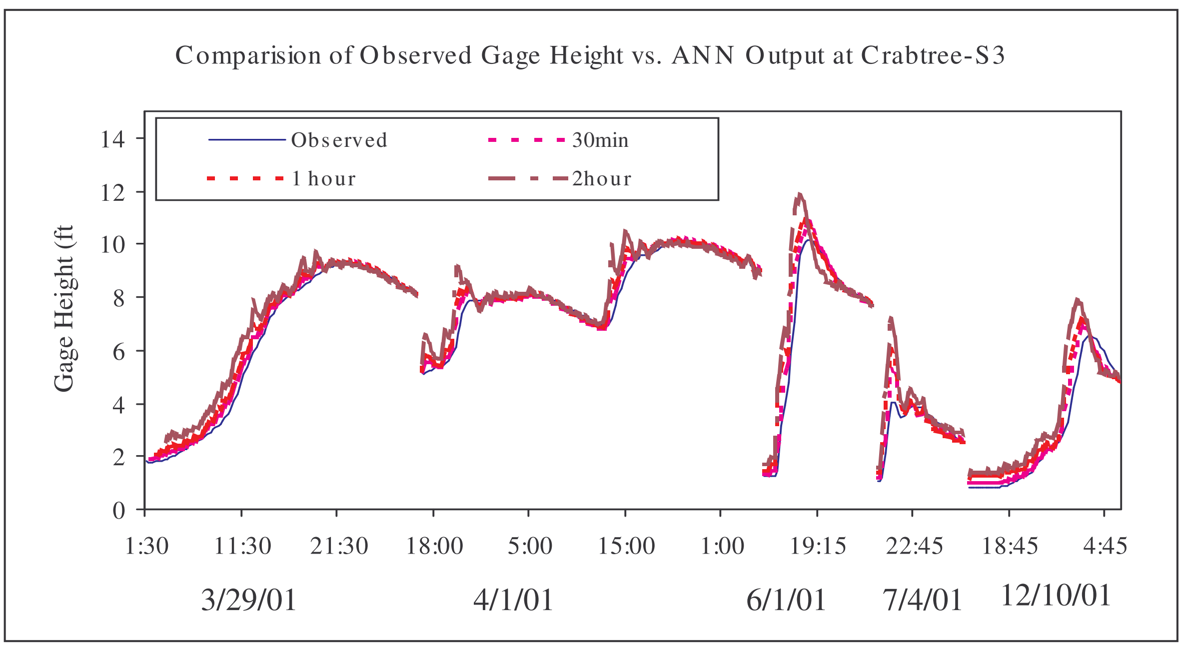
Appendix 1.13: Gage Height Prediction at Station Crabtree-S2, Raleigh NC



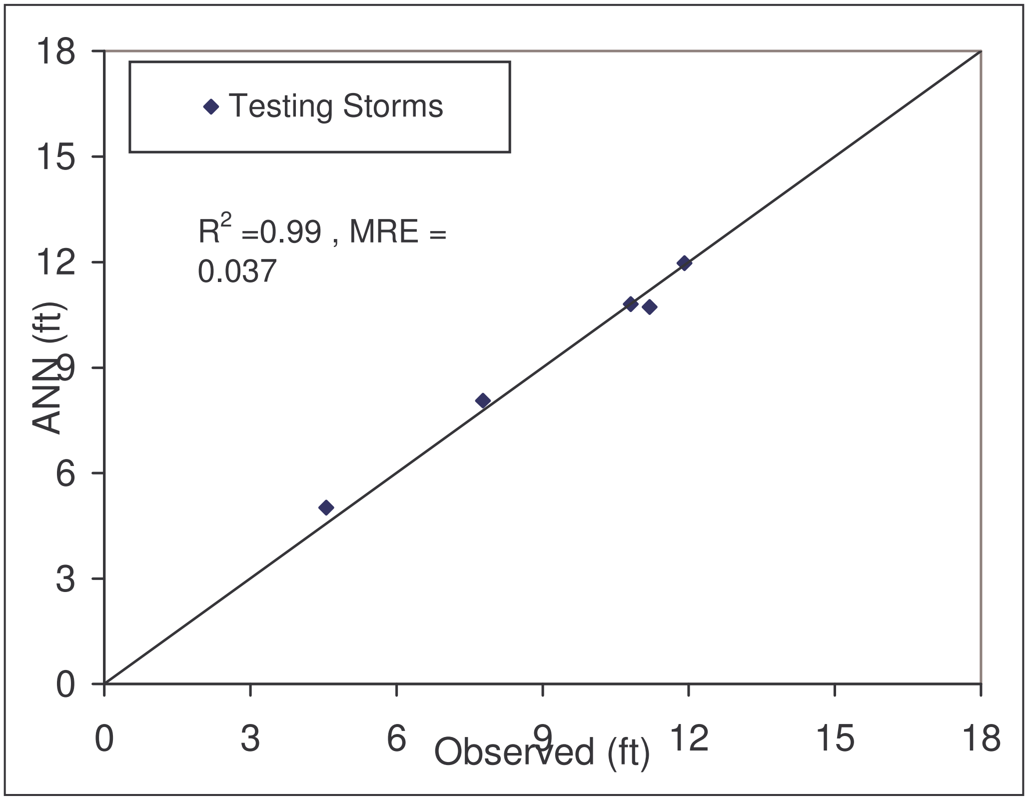


Appendix 1.14: Gage Height Prediction at Station Crabtree-S3, Raleigh NC

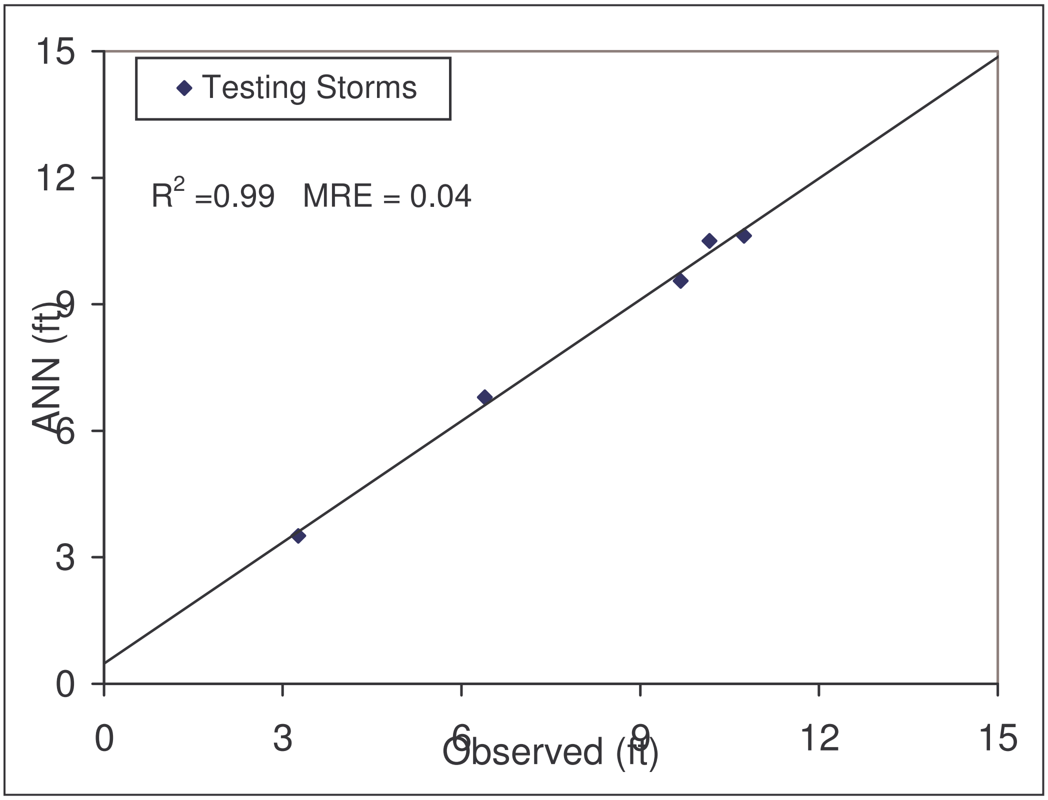




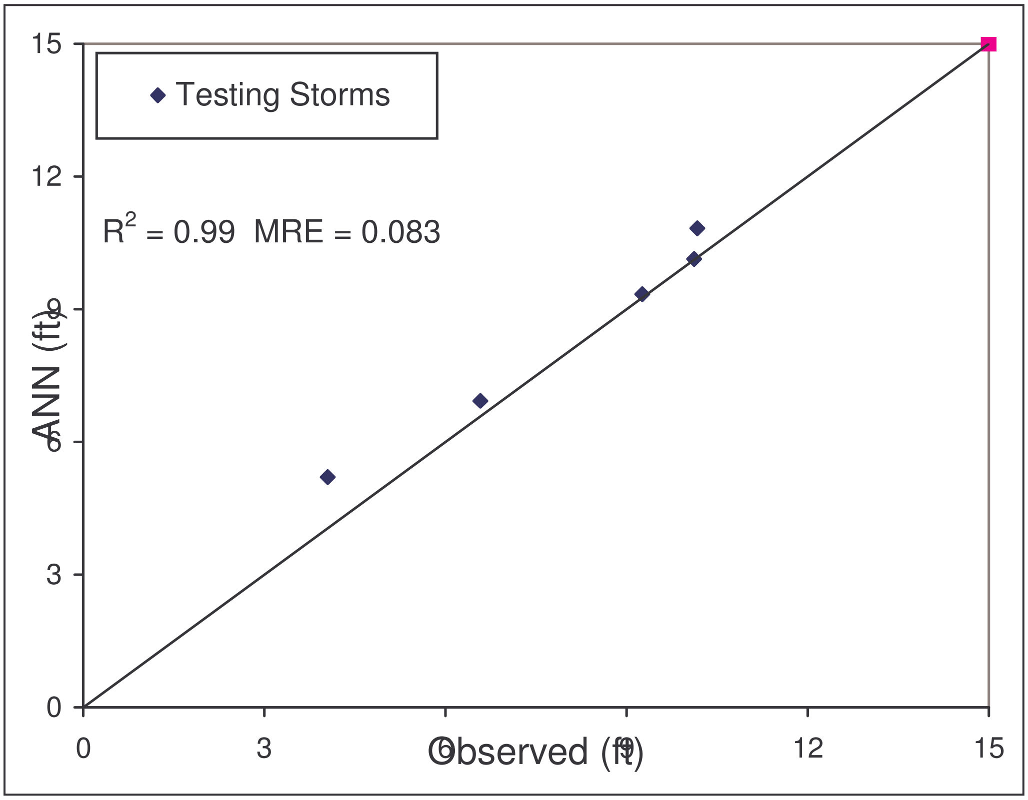
Appendix 1.15a: Observed-ANN Output Peak Flow Scatter plots at Crabtree-S1



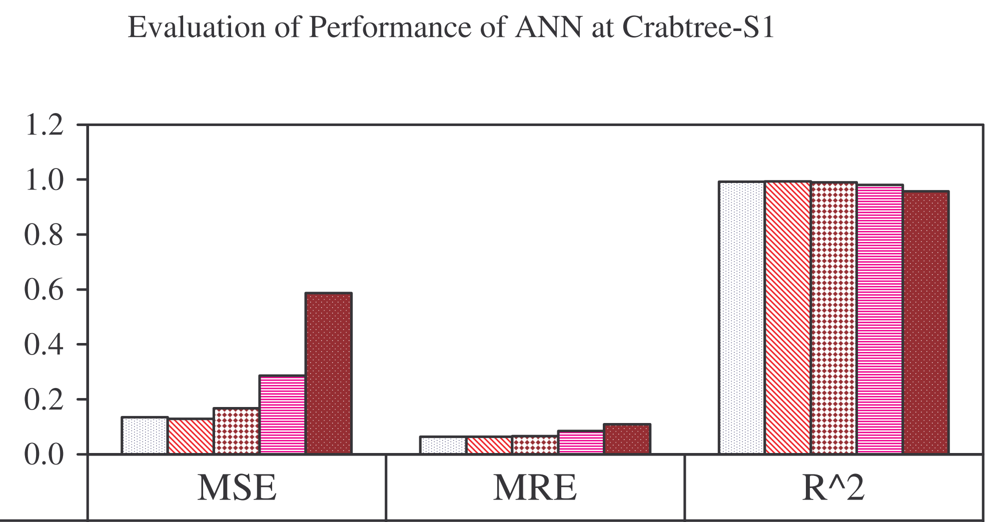
Appendix 1.15b: Observed-ANN Output Peak Flow Scatter plots at Crabtree-S2



Appendix 1.15c: Observed-ANN Output Peak Flow Scatter plots at Crabtree-S3



Appendix 1.16: Evaluation of Performance of ANN at Crabtree-S1



	MSE	MRE	R^2
15-min.	0.135	0.064	0.992
30-min.	0.129	0.064	0.993
1-hour	0.167	0.067	0.990
1.5 hr	0.286	0.085	0.980
2 hr	0.587	0.109	0.958

Appendix 1.17: Depth-Duration-Frequency Tables

Location: Raleigh/Durham, NC

Duration	Return Period (yr) --->					
	2	5	10	25	50	100
5 min	0.48	0.55	0.60	0.68	0.75	0.81
10 min	0.79	0.92	1.02	1.17	1.28	1.40
15 min	1.01	1.18	1.31	1.51	1.66	1.81
30 min	1.35	1.64	1.85	2.16	2.40	2.64
60 min	1.70	2.12	2.41	2.84	3.17	3.50
2 hr	1.91	2.40	2.74	3.23	3.61	4.00
3 hr	2.12	2.68	3.07	3.62	4.06	4.49
6 hr	2.65	3.38	3.90	4.62	5.19	5.75
12 hr	3.13	4.02	4.64	5.52	6.20	6.88
24 hr	3.60	4.65	5.38	6.41	7.21	8.00

Location: Greensboro, NC

Duration	Return Period (yr) --->					
	2	5	10	25	50	100
5 min	0.48	0.55	0.6	0.68	0.75	0.81
10 min	0.74	0.87	0.97	1.11	1.22	1.34
15 min	0.92	1.09	1.22	1.41	1.55	1.7
30 min	1.25	1.55	1.76	2.06	2.3	2.53
60 min	1.6	2.02	2.32	2.74	3.07	3.4
2 hr	1.81	2.28	2.61	3.08	3.45	3.82
3 hr	2.02	2.54	2.9	3.43	3.83	4.24
6 hr	2.55	3.19	3.64	4.29	4.8	5.3
12 hr	3.03	3.78	4.31	5.07	5.66	6.25
24 hr	3.5	4.36	4.97	5.84	6.52	7.2

Developed by Dr. H.R. Malcom, Department of Civil Engineering, NC State University, and the authors based on NOAA HDRY-35 and USWB TP-40

Appendix 2.1a: The Selection of Storms in Greensboro Area

Storm No.	Date	Total precipitation in USGS Stations (in.)										
		02095181	02095271	02095500	02094770	02093800	020939 9200	02094412	02094659	02094775	02095000	020955 3650
1	9/30/99	1.69	-	1.46	-	-	1.25	1.84	-	-	1.11	1.34
2	9/26/00	0.3	0.47	0.52	0.71	0.01	0.44	0.46	0.48	-	-	0.55
3	6/13/01	0.92	0.86	-	0.63	0.01	0.27	-	1.07	0.67	0.38	-
4	7/26/01	1.28	1.48	1.45	0.77	1.08	1.28	2.05	-	0.75	0.91	1.96
5	8/11/01	0.34	0.67	0.24	0.94	-	4.03	0.14	1.41	1.06	0.37	0.3
6	10/14/01	1.38	1.3	1.07	1.34	-	0.5	-	0.96	1.39	1.17	1.36

Appendix 2.1b: The Selection of Storms in Raleigh Area

Storm No.	Date	**Total precipitation in USGS Stations (in.)**			
		0208732534	0208732885	02087359	02087182
1	9/28/99	-	1.03	0.86	0.65
2	9/30/99	-	0.5	0.26	0.45
3	9/18/00	0.19	0.59	0.59	0.33
4	9/22/00	0.09	0.14	0.62	0.03
5	9/25/00	1.17	1.61	1.54	1.35
6	6/13/01	1.05	0.9	1.22	0.41
7	7/26/01	1.04	0.67	1.50	2.05
8	8/12/01	0.58	0.34	0.86	0.42
9	8/13/01	-	-	0.26	0.06
10	9/24/01	0.74	0.87	1.23	0.55
11	10/14/01	-	1.16	1.14	0.8

Appendix 2.2a: Total Precipitation Estimates by Z-R relationship vs. Rain Gage Observation in Greensboro, NC

Storm No.	Date	Station 02095181				Station 02095271			
		Duration (hr)	Total Precipitation (in.)			Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN		Gage	z-R	ANN
1	9/29/99	10	0.54	0.86	0.53	-	-	-	-
2	9/25/00	7	0.30	0.66	0.35	10	0.47	0.65	0.48
3	6/13/01	3	0.92	1.31	0.83	1	0.86	1.27	0.87
4	7/26/01	4	1.08	0.81	0.98	4	1.05	0.82	1.08
5	8/11/01	4	0.34	1.17	0.64	3	0.67	1.15	0.77
6	10/14/01	3	1.38	0.64	1.06	3	1.30	0.69	1.14

Storm No.	Date	Station 02095500				Station 02094770			
		Duration (hr)	Total Precipitation (in.)			Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN		Gage	z-R	ANN
1	9/29/99	4	0.32	0.37	0.23	-	-	-	-
2	9/25/00	10	0.52	0.61	0.34	5	0.57	0.75	0.63
3	6/13/01	-	-	-	-	3	0.63	1.35	1.00
4	7/26/01	4	1.01	0.95	1.31	2	0.62	0.68	0.81
5	8/11/01	3	0.24	0.88	0.39	3	0.94	1.14	1.10
6	10/14/01	3	1.07	0.69	1.02	3	1.34	0.72	0.58

Storm No.	Date	Station 02093800				Station 020939220			
		Duration (hr)	Total Precipitation (in.)			Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN		Gage	z-R	ANN
1	9/29/99	-	-	-	-	4	0.41	0.45	0.25
2	9/25/00	-	-	-	-	9	0.44	0.53	0.48
3	6/13/01	-	-	-	-	-	-	-	-
4	7/26/01	4	0.77	1.00	0.83	6	0.84	0.99	0.81
5	8/11/01	-	-	-	-	-	-	-	-
6	10/14/01	-	-	-	-	3	0.50	0.27	0.48

Storm No.	Date	Station 02094412				Station 02094659			
		Duration (hr)	Total Precipitation (in.)			Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN		Gage	z-R	ANN
1	9/29/99	10	0.75	0.82	0.76	-	-	-	-
2	9/25/00	11	0.46	0.72	0.68	4	0.38	0.51	0.37
3	6/13/01	-	0.00	0.60	0.04	-	-	-	-
4	7/26/01	4	1.23	0.97	1.44	-	-	-	-
5	8/11/01	4	0.16	0.55	0.13	4	1.43	1.18	1.28
6	10/14/01	-	-	-	-	2	0.96	0.54	0.93

Storm No.	Date	Station 02094775				Station 02095000			
		Duration (hr)	Total Precipitation (in.)			Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN		Gage	z-R	ANN
1	9/29/99	-	-	-	-	8	0.65	0.77	0.61
2	9/25/00	-	-	-	-	-	-	-	-
3	6/13/01	2	0.67	1.35	0.87	2	0.38	0.71	0.52
4	7/26/01	2	0.60	0.68	0.68	4	0.68	0.72	0.73
5	8/11/01	3	1.06	1.14	1.38	3	0.37	0.90	0.54
6	10/14/01	3	1.39	0.72	1.17	4	1.17	0.63	0.50

Storm No.	Date	Station 0209553650			
		Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN
1	9/29/99	12	0.92	0.93	0.83
2	9/25/00	4	0.34	0.58	0.33
3	6/13/01	-	-	-	-
4	7/26/01	5	1.12	0.92	1.10
5	8/11/01	3	0.30	0.54	0.34
6	10/14/01	4	1.36	0.72	1.09

Storm No.	Date	Greensboro Area (Average)			
		Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN
1	9/29/99	8	0.60	0.70	0.53
2	9/25/00	8	0.43	0.63	0.46
3	6/13/01	2	0.58	1.10	0.69
4	7/26/01	4	0.90	0.85	0.98
5	8/11/01	3	0.61	0.96	0.73
6	10/14/01	3	1.16	0.63	0.89

Appendix 2.2b: Total Precipitation Estimates by Z-R relationship vs. Rain Gage Observation in Raleigh, NC

Storm No.	Date	Station 0208732534				Station 0208732885			
		Duration (hr)	Total Precipitation (in.)			Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN		Gage	z-R	ANN
1	09/28/99	-	-	-	-	7	1.01	0.48	0.72
2	09/30/99	-	-	-	-	7	0.50	0.25	0.48
3	09/18/00	9	0.18	0.19	0.55	11	0.55	0.26	0.78
4	09/22/00	7	0.10	0.11	0.43	4	0.12	0.11	0.27
5	09/25/00	10	1.17	1.21	1.56	12	1.61	1.03	1.63
6	06/13/01	11	1.05	0.86	0.99	10	0.90	0.81	0.76
7	07/26/01	7	1.04	1.15	0.88	9	0.67	1.27	0.91
8	08/12/01	6	0.58	1.12	0.89	10	0.34	0.89	0.89
9	08/13/01	-	-	-	-	-	-	-	-
10	09/24/01	6	0.73	0.74	0.79	7	0.86	0.81	0.94
11	10/14/01	-	-	-	-	2	1.16	0.56	0.91

Storm No.	Date	Station 02087359				Station 02087182			
		Duration (hr)	Total Precipitation (in.)			Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN		Gage	z-R	ANN
1	09/28/99	12	1.28	0.88	1.23	11	0.89	0.52	0.84
2	09/30/99	3	0.22	0.23	0.25	5	0.44	0.52	0.45
3	09/18/00	10	0.56	0.37	0.77	10	0.33	0.18	0.59
4	09/22/00	6	0.12	0.11	0.27	-	-	-	-
5	09/25/00	8	1.54	1.60	1.79	7	1.35	1.23	1.83
6	06/13/01	10	1.22	1.22	1.76	3	0.41	0.45	0.45
7	07/26/01	8	1.50	1.16	1.69	9	2.05	1.07	1.70
8	08/12/01	6	0.86	0.49	0.79	6	0.42	0.36	0.53
9	08/13/01	3	0.26	0.33	0.41	3	0.06	0.36	0.32
10	09/24/01	9	1.22	1.01	1.23	8	0.54	0.62	0.81
11	10/14/01	2	1.14	0.78	0.94	2	0.80	0.60	0.87

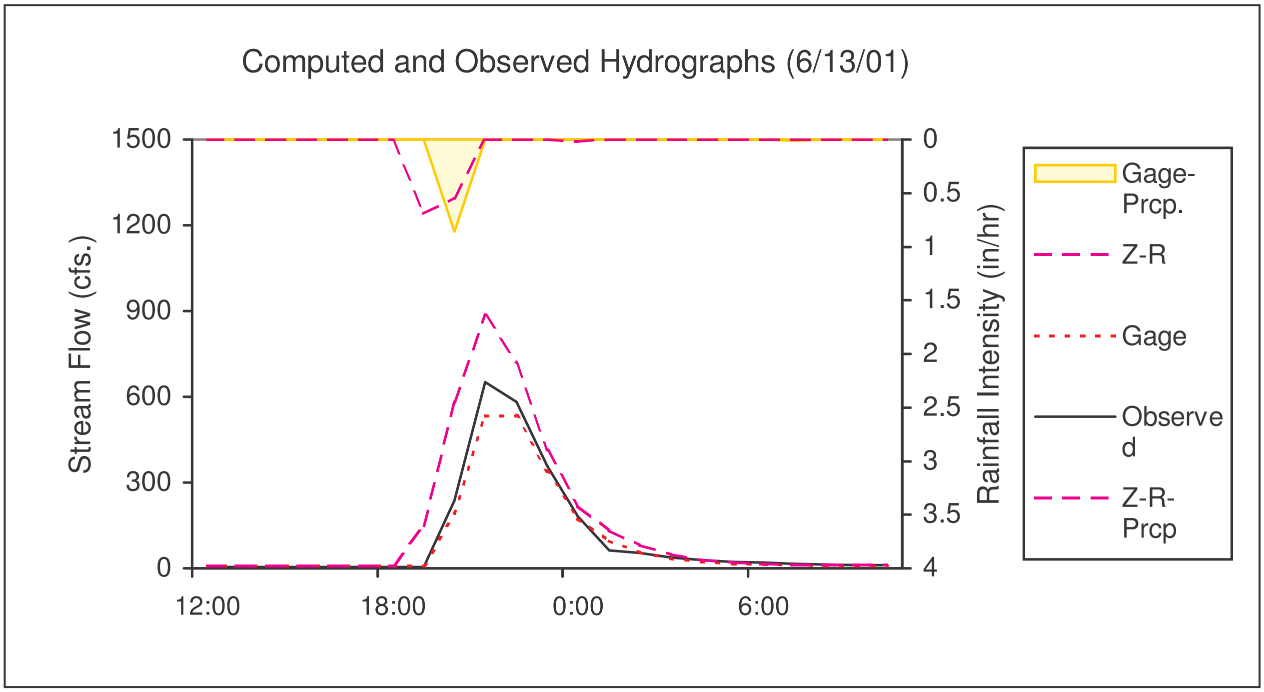
Storm No.	Date	Average Total in Raleigh			
		Duration (hr)	Total Precipitation (in.)		
			Gage	z-R	ANN
1	09/28/99	10	1.06	0.63	0.93
2	09/30/99	5	0.39	0.33	0.39
3	09/18/00	10	0.40	0.25	0.67
4	09/22/00	5.67	0.11	0.11	0.32
5	09/25/00	9.25	1.42	1.27	1.70
6	06/13/01	8.50	0.89	0.84	0.99
7	07/26/01	8.25	1.32	1.16	1.29
8	08/12/01	7	0.55	0.71	0.77
9	08/13/01	3	0.16	0.34	0.36
10	09/24/01	7.5	0.84	0.79	0.94
11	10/14/01	2	1.03	0.65	0.90

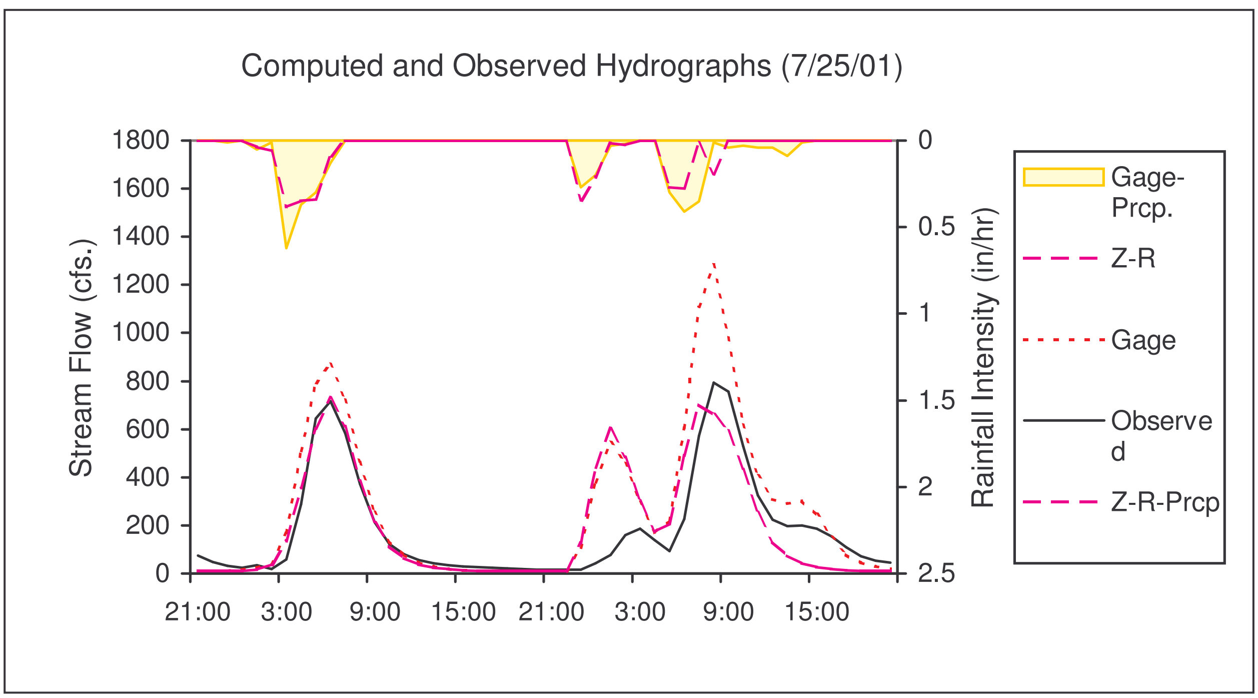
Appendix 3.1: Summary of the Selected Storms at Station 02095271 in Greensboro

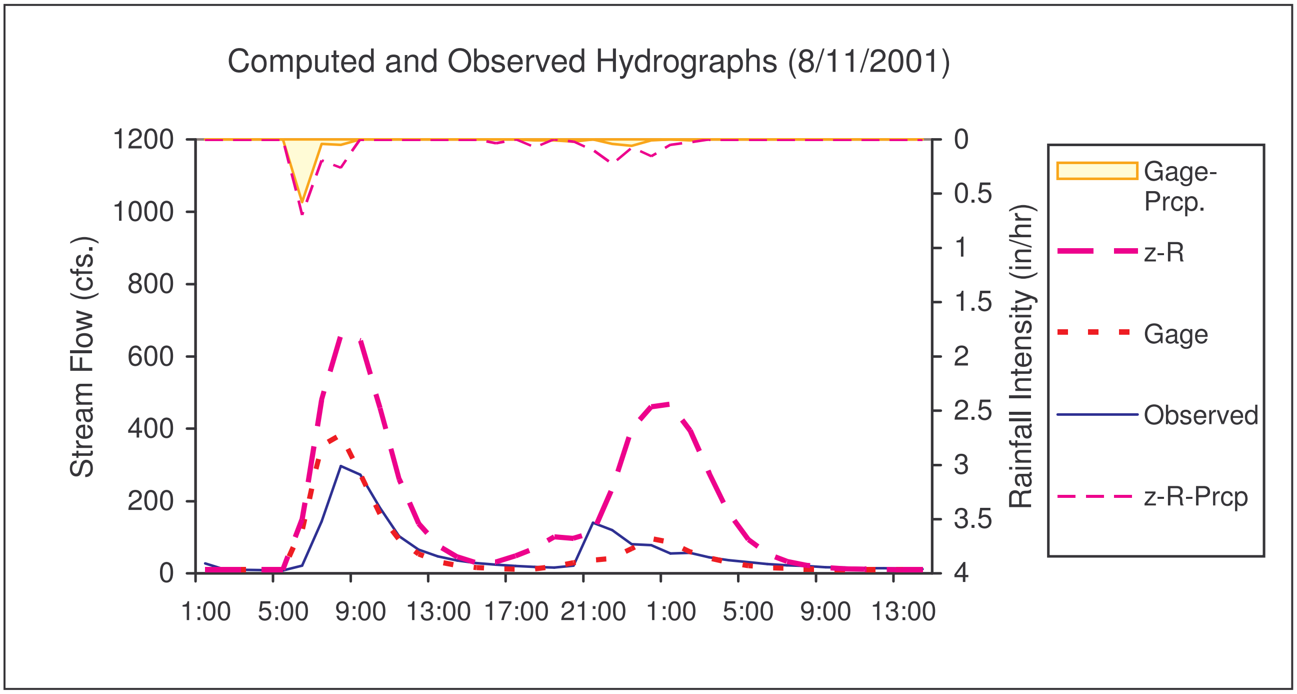
Storm No.	Date	Total Precipitation (in.)	Peak flow (cfs.)
1	6/13/2001	0.86	652
2	7/26/2001	1.49	716
3	7/27/2001	1.84	794
4	8/11/2001	0.83	297
5	9/24/2001	0.93	440
6	10/14/2001	1.3	692

- The storms on 6/13, 8/11, and 9/24 were used for training and calibration purposes.
- The storms on 7/26, 7/27, and 10/14 were used for testing.

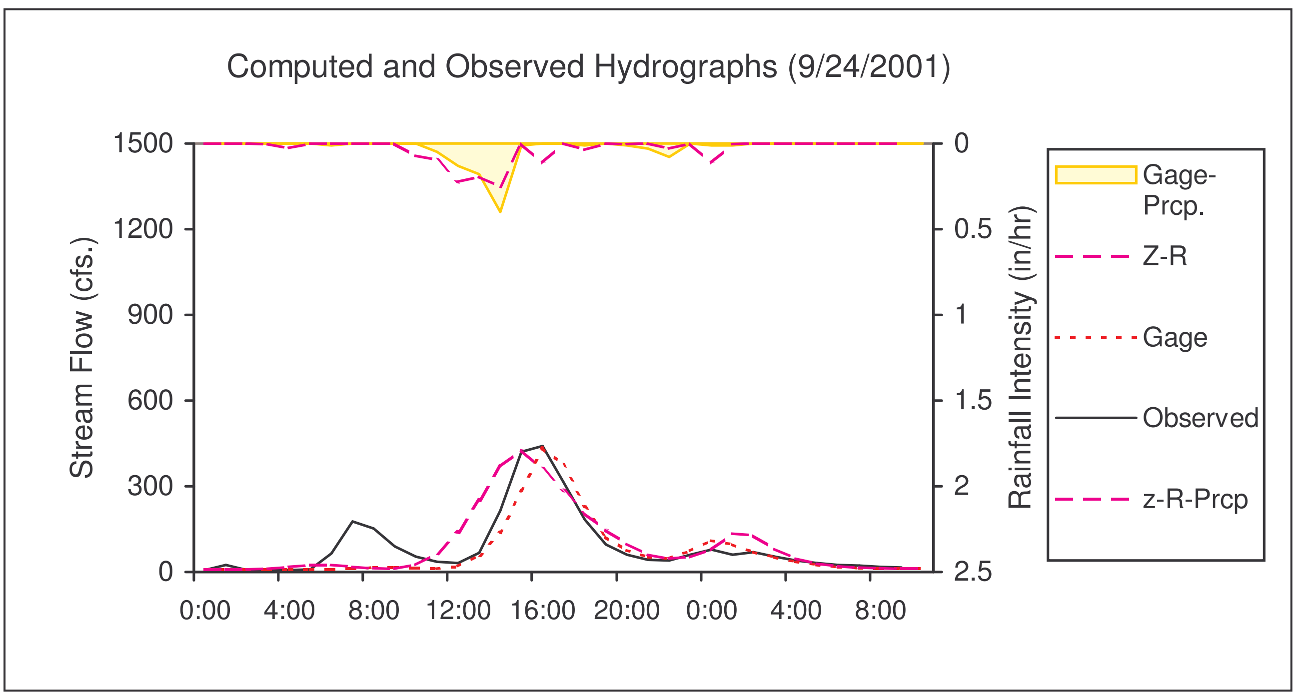
Appendix 3.2: Stream Flow Predication by Rainfall Observation and Radar-based Rainfall Estimation



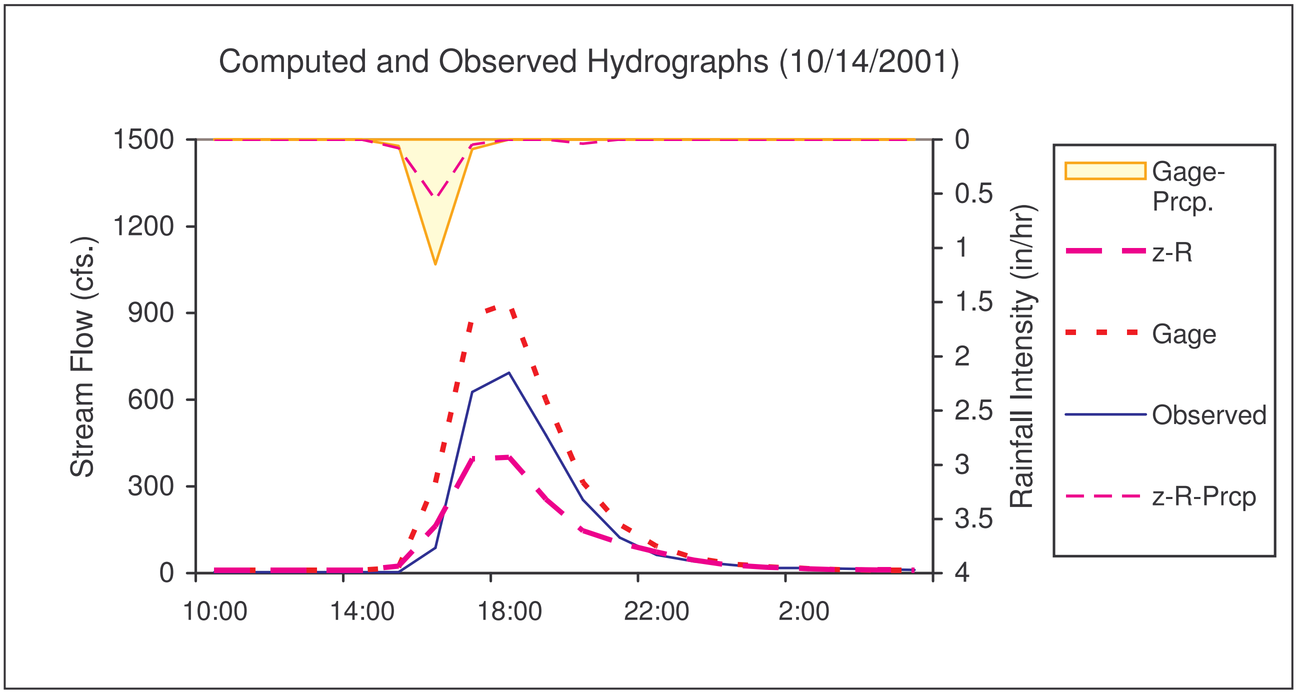




Computed and Observed Hydrographs (8/11/2001)
Stream Flow (cfs.)
Rainfall Intensity (in/hr)
1200
1000
800
600
400
200
0
0
0.5
1
1.5
2
2.5
3
3.5
4
1:00
5:00
9:00
13:00
17:00
21:00
1:00
5:00
9:00
13:00
Gage-Prcp.
z-R
Gage
Observed
z-R-Prcp

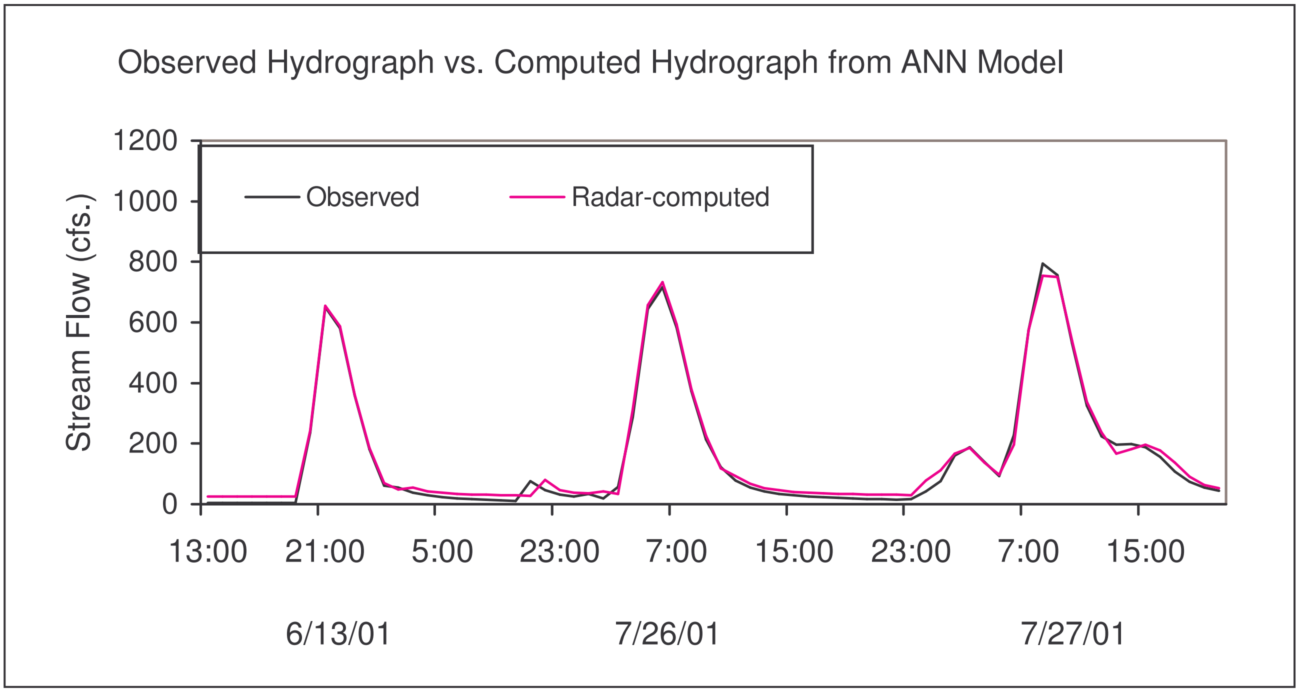


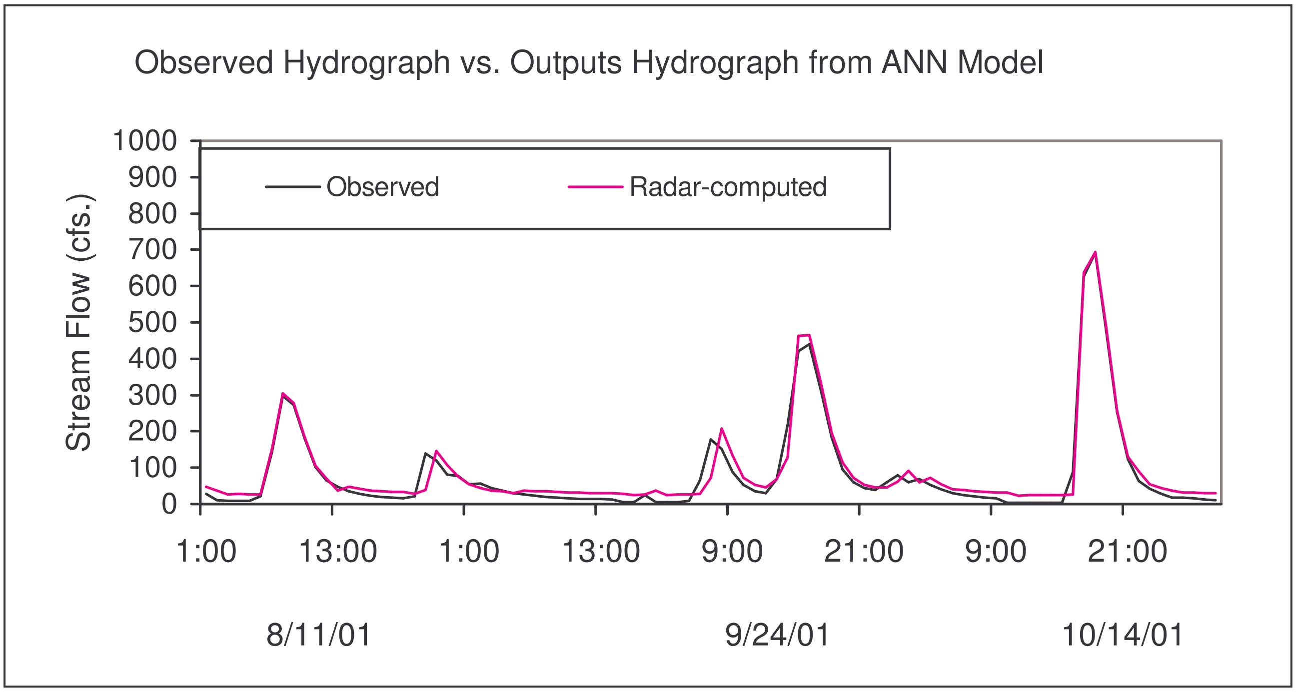
Computed and Observed Hydrographs (9/24/2001)
Stream Flow (cfs.)
Rainfall Intensity (in/hr)
1500
1200
900
600
300
0
0
0.5
1
1.5
2
2.5
0:00
4:00
8:00
12:00
16:00
20:00
0:00
4:00
8:00
Gage-Prcp.
Z-R
Gage
Observed
z-R-Prcp



Computed and Observed Hydrographs (10/14/2001)
Stream Flow (cfs.)
1500
1200
900
600
300
0
10:00
14:00
18:00
22:00
2:00
Rainfall Intensity (in/hr)
0
0.5
1
1.5
2
2.5
3
3.5
4
Gage-Prcp.
z-R
Gage
Observed
z-R-Prcp

Appendix 3.3: Computed Hydrograph by Radar-Stream flow Model





Appendix 4: List of Equations