University of Alberta

The Application of Artificial Neural Networks to Water Demand Modelling

By

Harold Lorenz Stark ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Master of Science**

In

Environmental Science

Department of Civil and Environmental Engineering

Edmonton, Alberta

Spring 2002

0-612-69765-7

University of Alberta

Library Release Form

Name of Author: Harold L. Stark

Title of Thesis: The Application of Artificial Neural Networks to Water Demand Modelling

Degree: Master of Science

Year this Degree Granted: 2002

Harold L. Stark
Box 1014
Wabamun, AB
T0E 2K0

Date: Jan.29/2002

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **The Application of Artificial Neural Networks to Water Demand Modelling** submitted by Harold Lorenz Stark in partial fulfillment of the requirements for the degree of **Master of Science** in **Environmental Science**.

Ian Buchanan
Dr. I.D. Buchanan

Dr. S.J. Stanley

Dr. J.J. Leonard

Dr. T. Yu

Jan. 28/2002

This work is dedicated to my family for their love and support, and to my wife Joanne for her endless love, friendship and encouragement.

Abstract

The cost of electricity for the pumping of water in water distribution systems, account for a large portion of the operating budgets of many water utilities. In North America there is currently a move towards the deregulation of the power industry that will change the rate structure for water utilities. It is therefore necessary for water utilities to better understand their power usage and pumping requirements to optimize there power usage to take advantage of the rate structure. An important component of this project is the accurate prediction of water demands. An artificial neural network model is presented which has been developed to predict the daily and 12-day water demands for the City of Edmonton. The developed daily model has an average error of 2.3%, while the 2 to 12 day model has an average error of 3.1%. An hourly prediction method that was developed has a 3.4% average error.

Acknowledgements

This work would not have been possible without the financial support of EPCOR Water Services. I would also like to recognize the following people:

Stephen Stanley, Ian Buchanan, Qing Zhang, and the staff of EPCOR Water Services for their assistance and help throughout this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

°C	degrees Celsius
%	percent
AI	artificial intelligence
ANN	artificial neural network
ANNs	artificial neural networks
CT	concentration time product
EDVAC	electronic discrete variable automatic computer
ML	million litres
ML/day	million litres per day
mm	millimetre
PC	personal computer
PLC-PC	programmable logic controller – personal computer
r	Pearson's Product Moment Correlation Coefficient
RAM	random access memory
SCADA	supervisory control and data acquisition
sd	standard deviation
XOR	exclusive or

The Application of Artificial Neural Networks to Water Demand Modelling

1 Introduction

1.1 Purpose of the Study

EPCOR Water Services currently spends over $3.5 million per year in power costs. These costs are distributed amongst two water treatment plants, 12 reservoir sites and two booster pump station sites. The recent changes in the power industry have resulted in a change in the way power is priced. It is therefore necessary to review the current practices for pump selection and to develop a tool to aid in the selection of which pumps to operate to meet a given condition. The first step in developing a pump optimization program is the development of an accurate water demand forecast; this will be carried out by using artificial neural networks (ANN) as a predictive model. This is to ensure that an adequate supply of water is available for use, as well as having an adequate emergency supply available. Having a demand forecast developed will allow EPCOR Water Services to evaluate the water demands of the city of Edmonton. By modeling the water distribution system and the electrical use of the system along with the water demand forecast, a pump schedule can be developed to minimize the electrical costs. This can be done by effectively using the storage (reservoir) space that is available, which allows shifting the time of pumping to a time of lower electrical costs and reducing the peak demands. The purpose of this study is to develop the first component of the power optimization program, which is the water demand component. This consists of a hourly, daily and twelve-day water demand forecast.

1.2 General Problem Description

The electricity cost associated with pumping water in the water distribution systems accounts for a large portion of the operating budget of a water utility. With the changes made by the Electric Utilities Act (1996) and complete deregulation of the power industry in Alberta set for the year 2001, it is important for a water utility to have a better understanding of its power use and pumping requirements to minimize their electrical costs. There are different pricing structure options available to the water utility, and the proper selection of the pricing structure can reduce electrical costs. Within most of the pricing structures available, certain constraints are to be met to reduce the electrical costs, such as a generation peak charge within a certain time of day. With the potential of obtaining electricity at cheaper rates during off peak consumption periods, and minimizing the water utility's peak demand in peak demand periods, these savings can be obtained. Potential problems with minimizing the electrical costs are ensuring an adequate water supply in the reservoirs at all times, and the proper operation of the pumps to minimize pumping during periods of peak electrical demand. The overall scope of this project is to develop a model system that would be used to set up a protocol and procedural system to aid operators in choosing the proper rate and time of pumping. It is expected that EPCOR Water Services can utilize cheaper electrical rates during off peak hours and that the peak electrical demand can be lowered for the pumping systems, thus leading to increased savings in electrical costs. Other options for increased savings may be realized after modeling the system, such as augmentation of the power supply by stand-by generators, for short periods of time, to avoid higher rates for the entire peak demand period. The power optimization program would greatly enhance the capability to assess options such as this in decreasing the overall cost of electricity for pumping.

The proposed model system must account for operational constraints, demand forecasting and electrical pricing structures. As a result development of a modeling system is a significant task involving a number of components. To facilitate timely completion of this large project it has been divided into a number of distinct components. By doing this, each component can be worked on relatively independently. Automation of the complete process will not be possible until all components are completed. However, as components are completed they can be used as part of the decision making process for pumping optimization. Components that are not completed will have to rely on the current methodology, which relies greatly on the experience of the staff of EPCOR Water Services. Although each component can be worked on independently it is also important to recognize how each component will affect the overall goal of pumping optimization, if at the end all components are to be integrated into an overall model. As a result the framework for the overall modelling system is presented in Figure 1.1 and the various components are explained below.

1.2.1 Maintenance

The maintenance component of the overall model is needed, as general maintenance and breakdowns change the options that are available for pumping at any given time. With scheduled maintenance, a new input would be entered in advance so that the pumping schedule can be adjusted, so that the desired goal can be reached. For instance the new goal might be to have a reservoir full at a certain time before the scheduled maintenance. The maintenance component should take into account EPCOR Water Services' maintenance schedule, as well as, EPCOR's Distribution and Transmission's maintenance schedule for the electrical distribution system.

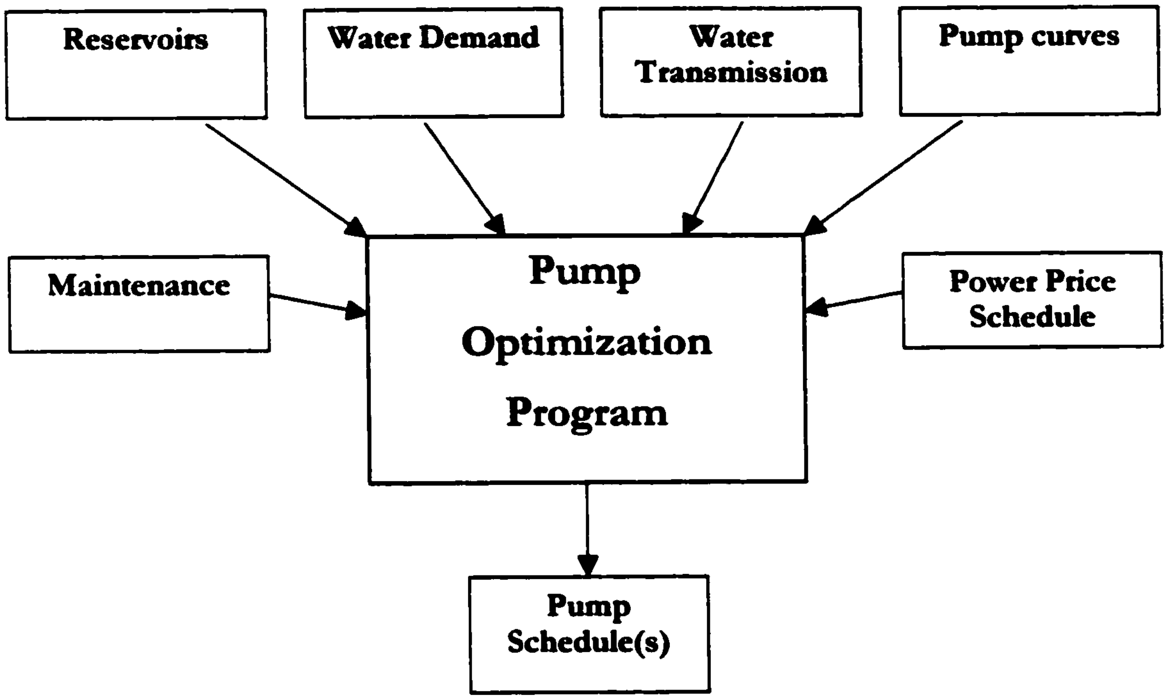

Figure 1.1 Overall Modelling Framework for Pump Optimization Program

1.2.2 Reservoirs

To use any type of pump schedule, the physical and safety constraints placed on the reservoirs must be identified. These include:

- pump capacity at each reservoir
- minimum reservoir levels (safety and performance measures)
- reservoir circulation (turnover)
- sub-station loading at each reservoir
- CT requirements

1.2.3 Water Demand Forecast

To determine the pump schedule, an accurate estimate of the water demand is needed as the pump schedule program needs to anticipate the future demand on the system. The water demand forecasts needed are the hourly, daily and 12-day demands. The daily and 12-day demands are to be estimated by using artificial neural networks to model the process. The hourly demand will be estimated using a different modelling technique where the hourly demand curves are normalized and superimposed on the daily demand. This is the area of focus for this research.

1.2.4 Water Transmission

How the water is transmitted through the distribution system is of importance to maintain the pressure in the system, and to minimize the energy needed to move the water. This would mostly likely be optimized using the Stoner model that EPCOR Water Services is currently using.

1.2.5 Pumping Curves

The pumping curve from each pump is needed, so as to know at which flowrates the pumps are most efficient. This is so that the final optimization program chooses the pumps for the desired pumping rate, which use the least power. It will identify the most efficient pumps and utilize them instead of the less efficient pumps. It will also identify any inefficient pumps that need maintenance or need replacing.

1.2.6 Power Price Schedule

The power price schedules also need to be input into the program, as the power price is the main driving factor in minimizing pumping costs. The price of power if preset and any peak demand charges can be entered in as an input. If the price is on a real-time schedule where the price is always fluctuating, then the predicted price of power (available from the power pool) will be needed. With the program broken down into different components, it should be possible to use the same program with the different price schedules available. This would allow EPCOR Water Services to evaluate which price schedule would best suit the needs in reducing the pumping costs even more. Within this component, looking into using an on-site power generator to reduce the use of electrical grid power during periods of peak demand may be cost beneficial.

The development of the proposed pumping optimization system should provide significant benefits for EPCOR Water Services. In addition to the cost savings, the proposed system will formalize the pumping optimization procedure that currently relies on the experience of the staff of EPCOR Water Services. The move by EPCOR Water Services to a multi-skilled workforce may result in operators not being able to specialize to the same degree as previously, resulting in a need to rely more on formalized procedures rather than specialized experience.

1.3 Organization of the Thesis

The balance of the thesis is organized into five sections: background information, methodology, results, applications and conclusions and recommendations. The background information section contains an overview of the water treatment and distribution system, water demand and artificial neural networks. The methodology section outlines the selection and collection of the data and model development. In the results section, the results from each of the 3 different models, daily, twelve-day and hourly water demand estimates are presented. The application section discusses the options that are available in the implementation of the models for a water utility. The conclusion and recommendation section provides a summary of the water demand models and recommendations for future work.

2 Background

2.1 Overview of EPCOR Water Services' Water Treatment and Distribution System

EPCOR Water Services currently owns and operates the E.L. Smith and Rossdale water treatment plants. The two water treatment plants are located in Edmonton, Alberta, Canada (Figure 2.1) and provide water for the city of Edmonton and the surrounding communities. EPCOR Water Services also owns and operates the distribution system.

2.1.1 E.L Smith Water Treatment Plant

The E.L. Smith water treatment plant was built in 1976 and was expanded in 1984. It is located on the North Saskatchewan River and is on the west edge of the City of Edmonton. The E.L. Smith water treatment plant currently is capable of treating 281ML/day and can be expanded to treat 800 ML/day.

2.1.2 Rossdale Water Treatment Plant

The first water treatment plant was built in downtown Edmonton on the banks of the North Saskatchewan River in 1903. The Rossdale Water treatment plant was built on the same site in 1947 to provide water to the rapidly increasing population of the city. The Rossdale water treatment facility can treat 239ML/day.

2.1.3 EPCOR Water Services' Distribution System

EPCOR Water Services has the third largest distribution system in Canada and supplies water to 40 communities within a 100-km radius of the City of Edmonton. The Distribution section of EPCOR Water Services is responsible for the maintenance of over 2,900 kilometres of water mains, more than 13,000 hydrants, and 39,000 valves. A regional distribution map can be found in Figure 2.2. Within the City of Edmonton, EPCOR Water Services has 12 reservoir sites with a total capacity of 808 ML or approximately a two day supply of water. The reservoir capacity of the system is of utmost importance as it determines the ability of a water utility to minimize its electrical costs with proper pump scheduling, while maintaining adequate storage to meet consumer demand. A summary of the reservoirs and their capacities is located in Table 2.1. The capacity of a reservoir is broken down into their gross, operating, fire and available storage. The gross storage is the total amount of storage capacity available in a reservoir. The operating storage is the storage that can be readily used to supply customers. The operating storage consists of the fire storage and available storage. The fire storage is the minimum water that must be stored in the reservoirs that needs to be readily available for fire protection. The available storage is the storage space in the reservoirs that can be used to supply water to the consumers. The City of Edmonton's overall water distribution system is illustrated in Figure 2.1.

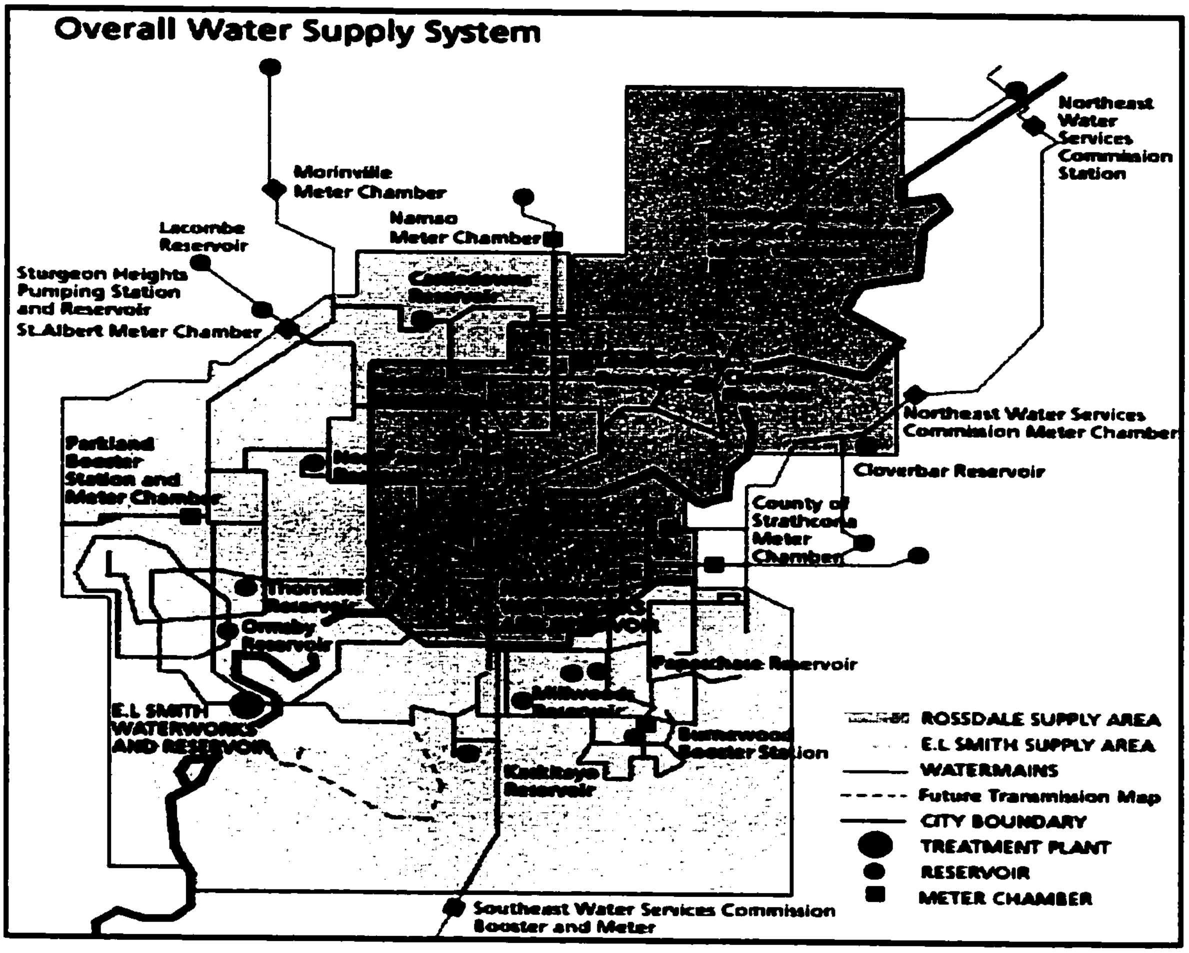

Figure 2.1 City of Edmonton's Water Distribution System (EPCOR Water Services Inc., 2000)

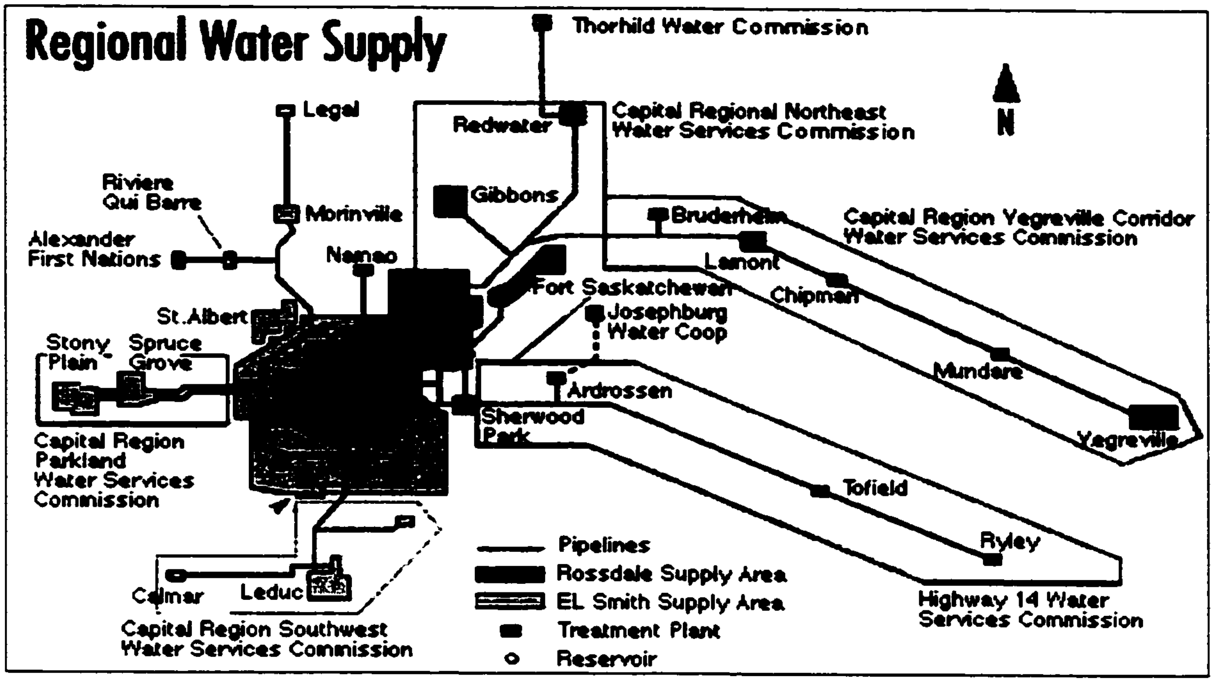

Figure 2.2 Regional Water Distribution System (EPCOR Water Services Inc., 2000)

Table 2.1 Storage Capacity of Existing Reservoirs Owned by EPCOR Water Services Inc.

Reservoir and W.T.P.	Gross Storage (ML)	Operating Storage (ML)	Fire Storage (ML)	Available Storage (ML)
Rosslyn	123.08	110.05	20.22	89.93
Papaschase	82.14	76.51	14.21	62.30
Londonderry	45.24	41.68	2.67	39.01
Thorncliff	45.35	40.04	4.71	35.33
Millwoods	56.23	52.90	7.13	45.77
N.J. Place	46.06	34.40	7.50	26.90
Ormsby	45.27	40.40	4.81	35.59
Clareview	64.60	53.46	4.75	48.71
Castledowns	34.04	25.74	2.51	22.61
Kaskitayo	28.94	25.74	4.77	20.97
Sub Total	570.95	500.92	73.28	427.12
Rossdale	100.00	100.00	0.00	100.00
E.L. Smith	137.47	125.60	0.00	125.60
Sub Total	237.47	225.60	0.00	225.60
Grand Total	808.42	726.52	73.28	652.72

2.2 *Water Demand*

2.2.1 Daily Water Demand

The daily water demand is an essential parameter that a water utility needs to evaluate for day to day operations. It indicates the volume that is required to be supplied for the entire day. This gives the operators a target demand that needs to be met at the end of the day. In 1998, EPCOR Water Services had an average water demand of 331 ML/day, with a maximum of 452 ML/day.

The water demand can be broken down in its use by type of users. There are approximately 176,200 residential dwellings consuming 50% of the water, 14,600 commercial and industrial consumers using 27% of the water and 7 wholesale/regional customers using 23% of the water.

2.2.1.1 Residential Water Use

The residential water use account for roughly 50% of the water used per year. The main residential interior use consists of shower/baths, toilets, washing machine, dishwasher, and faucet use. Residential exterior use consists mainly of lawn and garden watering and car washing. The interior use remains fairly constant for residential use, as it consists of mainly water for personal hygiene and cleaning. Conversely, the exterior use varies depending on the season and meteorological conditions. In the climate that the City of Edmonton is located in, there is very little to no exterior water use by a typical residential customer during the winter. In the summer, the exterior use can very dramatically depending on the meteorological conditions, as the main exterior residential water use is lawn. The fluctuation in the residential exterior water use is a major factor in the overall fluctuations in the water demand during the summer months.

2.2.1.2 Industrial and Commercial Water Demand

The commercial and industrial water demand accounts for approximately 27% of the total water demand. Commercial and industrial users cover a wider range of customers, they range from shopping malls, office buildings, food processing plants to breweries. The water use for commercial and industrial customers on average increases in water usage in the summer months. Most commercial and industrial customers water demand has little fluctuation, with some exceptions for the industries that have seasonal cycles, such as hotels.

2.2.1.3 Regional Water Demand

The regional water demand is composed of the other surrounding towns, cities and municipalities. Each of these customers are in turn composed of residential, commercial and industrial customers to varying degrees. The major problem with the regional customers is that they have their own reservoirs for storage. Thus they control when and at what rate they draw water into their system. With the storage that the regional customers have they can draw large quantities of water in a short period of time, which is than recorded on the SCADA system as a demand on EPCOR's distribution system. As the control of some of the regional customers' reservoirs have no set pattern and are not operated in conjunction with their customers real time use, there is no way to predict their operation; thus a portion of their resulting demand acts as noise in the modelling process. The problem with this is it is not real-time actual use by the residential, commercial and industrial customers. So the demand that EPCOR experiences from the regional customers is not the real time usage of the water. This is detrimental to the modelling process and adds uncertainty to any water prediction that is made.

2.2.2 Hourly Water Demand

The hourly water demand is of even more importance than the daily demand when considering distribution needs. Simply producing the volume of water required on a daily basis does not mean that the water demand will be met at all times of the day. This is due to the fluctuations in the water demand throughout the day. A typical hourly water demand curve and the ideal plant flow can be seen in Figure 2.3. The water treatment plant produces water at a constant rate that will satisfy the demand at the end of the period. However, water demands that are higher than the constant water production rate is satisfied by using stored operating water. The hourly water demand is also a very important component of developing a proper pump schedule. When the hourly demand exceed the water treatment rate, the reservoir will release water. The reservoir can fill when the hourly (instantaneous) demand is less than the water production rate. This eliminates the difficulty of trying to operate the water treatment plant to meet the instantaneous demand, which would be undesirable from an operational standpoint. In the process of treating the water it is desirable to keep the plant flow constant so that each treatment process can be controlled and maintained close to a steady-state conditions. This then allows the operators to optimize each water treatment process and maintain it. With fluctuating water flows through a water treatment plant, plants operators would be constantly varying chemical feeds, and other parameters, which would lead to lower water quality in the finished water. Using reservoirs for storage also allows for smaller water treatment facilities to be built as they only need to be designed to meet a maximum 5-day demand (can be shorter or longer period depending on the storage available). An hourly peak demand can easily exceed 600 ML/day while a 5-day demand would be approximately 400 ML/day. Thus, a water treatment facility with adequate storage can be two-thirds the size of a water treatment facility with very little storage, and still provide the same quality of water and still meet consumer demands.

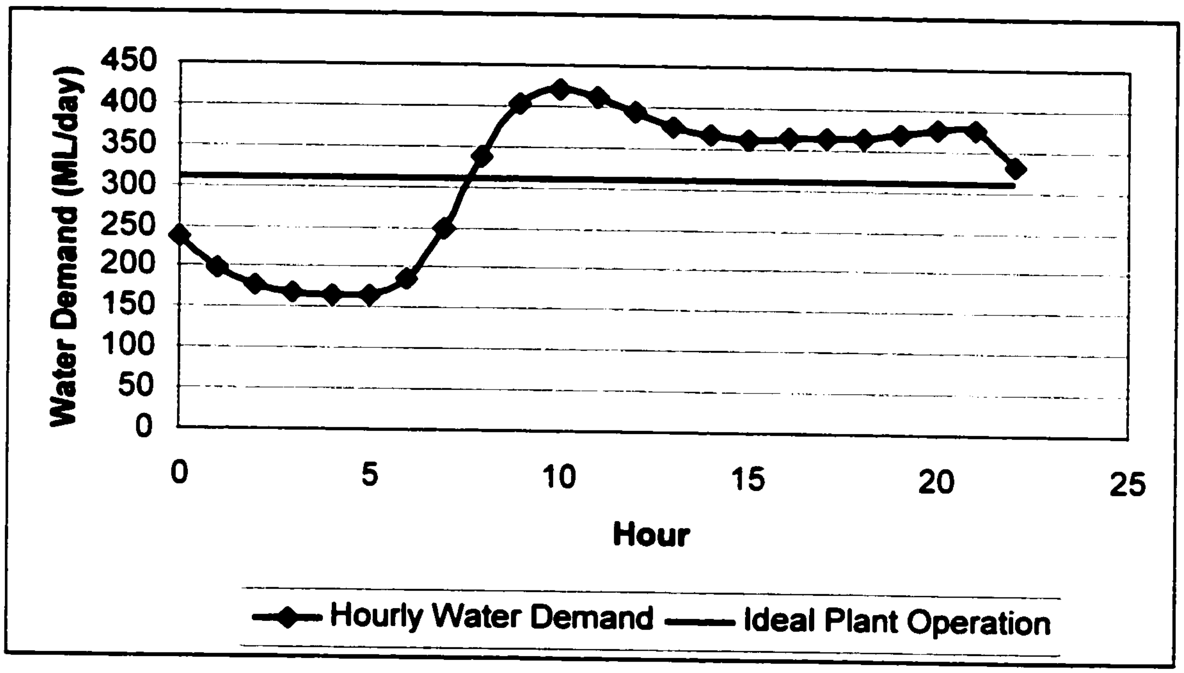

Figure 2.3 Typical Hourly Demand Curve and Ideal Plant Flow

2.2.3 Twelve Day Water Demand

The hourly and daily water demands are used for the short-term planning to meet the water demand of EPCOR Water Services' customers. Knowing the water demand for a number of days into the future is also of importance to a water utility because it allows the utility to slowly increase or decrease water production so that it can meet these demands and at the same time minimize its costs. EPCOR Water Services forecasts the current day and the following 11 days of water demand as a tool to plan the operations of the water treatment plants and storage reservoirs. This 12-day forecast allows EPCOR Water Services to foresee any potential shortages. If there is a potential shortage of water, then production can be increased to raise the reservoir levels so as to avert the shortage. Management also uses the 12-day forecast when scheduling maintenance that will interrupt the regular water supply. If at any time the maintenance will lead to the reservoirs dropping below 60% of their capacity, the maintenance is postponed to a later date, unless the maintenance is absolutely crucial. Maintaining a 60% volume in the reservoirs ensures an adequate and safe supply of water, in case of a breakdown in the water treatment plant or a major water main break. This gives EPCOR Water Services roughly a day, to a day and a half to remedy the problem without disrupting the transmission of water to its customers.

2.3 Conventional Modelling Methods for Water Demand

Many methodologies have been used in the past and are currently being used to model water demand. These include state - space and multiple regression methods (Billings and Agthe 1998), multi-linear regression, time series methods, and artificial neural networks (Fleming 1994), fuzzy logic and artificial neural network hybrid program (Lertpalangsunti and Chan Christine 1997), deterministic chaos method (Oshima and Kosuda 1998), pattern recognition (Shvartser et al. 1993), memory based learning in combination with neural networks (Tamada et al. 1993), as well as a variety of other methods and combinations of these methods. Artificial neural networks (ANNs) have been used as a component of a variety of these models. ANNs are valuable in modeling non-linear problems (Tamada et al. 1993) such as water demand forecasts. ANNs has also been shown to have superior modeling capabilities with fewer variables than multi-linear regression and time series methods (Fleming 1994).

Knowledge of the study domain is needed to use ANN for modelling water demand. Data relating to relevant input parameters are needed to train the model. It has been suggested that the some of the important parameters that affect water demand are the weather such as temperature, rainfall, humidity, sunshine hours (Shvartser et al. 1993; Ormsbee and Lansey 1994; Tamada et al. 1993; Hall and Maidment 1990; Hittle et al. 1996; Palmer et al. 1995; Jain and Ormsbee 1993; Fildes et al. 1997), season of the year (Ormsbee and Lansey 1994; Hall and Maidment 1990; Hittle et al. 1996; Kulshreshtha et al. 1996; Palmer et al. 1995), past water use trends (Ormsbee and Lansey 1994; Hall and Maidment 1990; Hittle et al. 1996; Jain and Ormsbee 1993; Fildes et al. 1997), and day of the week indicator (Hittle, et al. 1996; Shvartser et al. 1993).

Work has been done in the past to predict daily and hourly water consumption using ANNs (Crommelynck et al. 1992). Crommelynck et al. (1992) incorporated, various meteorological data, human behavioral data, seasonal, weekly and daily cycles into their ANN model, which utilized 54 different inputs in total. However, problems can arise when a large number of inputs are used. Increasing the number of inputs beyond the minimum number required to describe the process will diminish the model's capacity to differentiate between important and unimportant inputs.

In addition to having a sound knowledge of the study domain, the proper model architecture must be used. Generally, most time-series forecasting methods reported in the literature employ backpropagation architecture, as this tends to be one of the most robust architectures for forecasting problems.

2.4 Overview of Artificial Neural Networks

2.4.1 NeuralShell 2 Software

NeuralShell 2 is a general artificial neural network software that is made by Ward Systems Group Inc. and is the software used in this research. It requires a PC running Windows 95, 98, 2000 or NT operating system, with an 80486 or higher process, 16 megabytes of RAM, and 20 megabytes of hard drive space. NeuralShell 2 is designed to be used for classifying data or for predictions. The software is also designed to give liberal control and flexibility to allow for experimentation.

2.4.1.1 Architecture

2.4.1.1.1 Backpropagation Architectures

Backpropagation networks are used on the vast majority of neural network problems as they generalize well on a wide variety of problems. They are classified as a supervised type of network, as they are trained on both the inputs and outputs. In contrast a Kohonen network is an unsupervised network, as it doesn't use output in the training, but classifies and separates out the similar data patterns into a specified number of categories.

Backpropagation networks can vary in the way that the layers are interconnected and are outlined below.

2.4.1.1.1.1 Standard Connections

The standard connections in the backpropagation network are that each neuron is connected only to each neuron in the previous layer. They have one input layer, one output layer and from 1 to 3 hidden layers. A basic 3-layer backpropagation network can be seen in Figure 2.4.

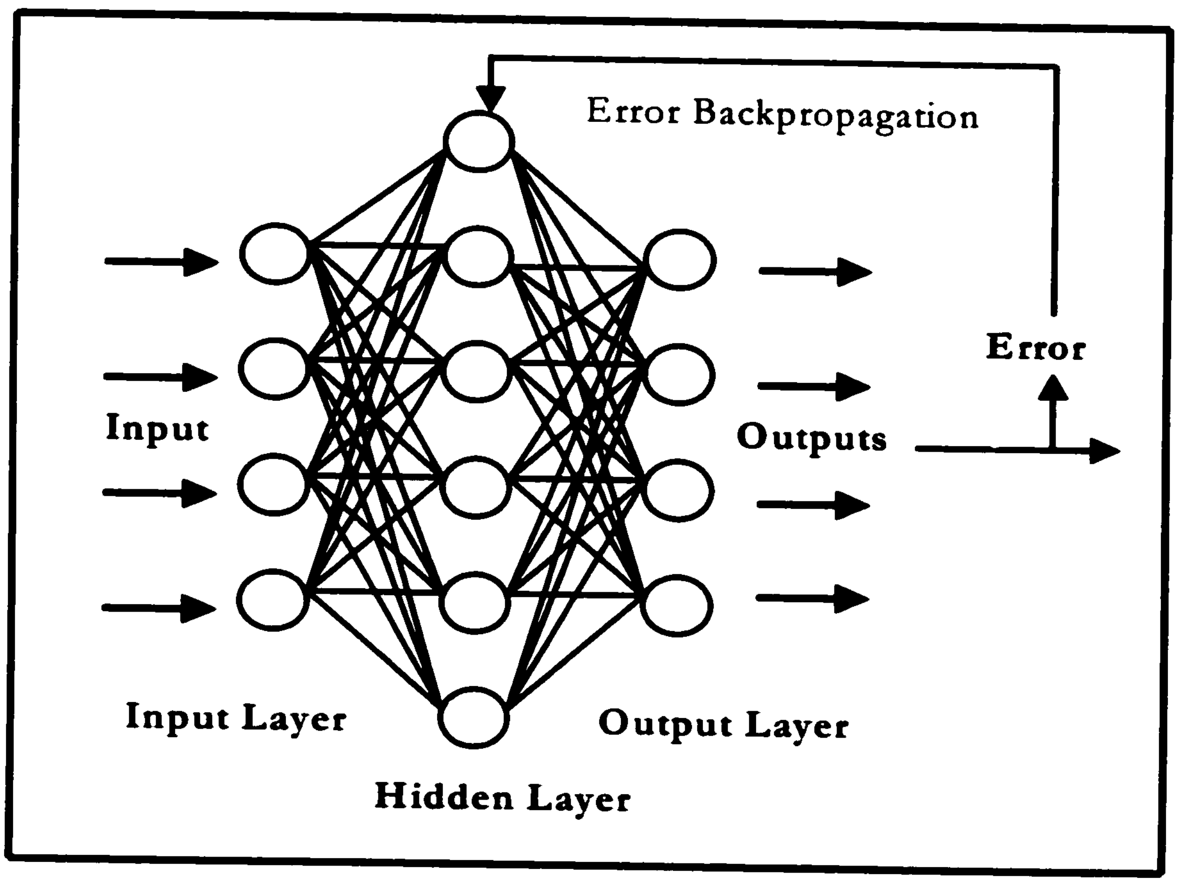

Figure 2.4 Three-layer Backpropagation Network

2.4.1.1.1.2 Jump Connections

The backpropagation network using jump connections means that every neuron is connected to each neuron in all the previous layers. They have one input layer, one output layer and from 1 to 3 hidden layers.

2.4.1.1.1.3 Recurrent Networks

Recurrent networks are trained in a similar way as standard backpropagation networks. The one major difference is that when using a recurrent network the data must be presented in the same order, this means that a rotational pattern selection is needed. The rotational pattern selection is needed because there is an extra layer identical to the input layer, which contains the previous inputs. The first hidden layer is then fed the current inputs as well as the previous set of inputs. This method is useful for time series data. However, if there is no temporal structure to the data, the extra layer will act as random noise to the network.

2.4.1.1.1.4 Ward Networks

The hidden layers in Ward networks can either have standard and/or jump connections. Also each hidden layer may be made up of multiple slabs. A slab is a group of neurons with the same activation function applied to them. In using standard connections each layer is one slab, as all the neurons in the layer have the same activation function. In a Ward network you can have multiple slabs in one hidden layer, and each slab having a different activation function. The reasoning for having different activation functions is that one activation function may detect the features in one range of data, while another activation function may yield better results for another range of data. This gives the network the unique ability to give the output layer different views of the data, which can lead to improved performance.

2.4.1.2 Scaling Function

A scaling function is needed when loading input variables into a neural network. It is needed to scale their numeric range into a range that the neural network can deal with efficiently. The NeuralShell 2 software supports both linear and non-linear scaling functions.

Typically the inputs are scaled between 0 and 1 or –1 and 1. If the linear scale is denoted as <<0,1>> or <<-1,1>>, larger or smaller values will not be clipped off. If the linear scale is denoted as [0,1] or [-1,1], then larger or smaller values will be clipped off. For instance, if the data have a range of 0 to 100 and a new input of 135 is entered, then it would be scaled to 1 if using [0,1] and 1.35 if using <<0,1>>.

For non-linear scaling there are the tanh and logistic scaling functions. The tanh function scales the data to (-1,1) using the formula f(value)=tanh((value-mean)/sd). The logistic scaling function scales the data to (0,1) using the formula $f(\text{value})=1/(1+e^{-(\text{value-mean})/\text{sd}})$. The tanh and logistic scaling function by definition exclude the possibility of having a value outside of the range of –1,1 and 0,1 respectively. By using these non-linear scaling functions, there are no out of range data as any out of range data are squeezed together at the high or low ends and are not clipped as when a linear function is used. Non-linear scaling functions are useful to minimize the effect of outliers, but in doing so also minimize the importance of extreme events, which may be of interest. In addition to the above main scaling functions, a scaling function can be customized to a specific problem.

2.4.1.3 Activation Functions

An activation function is needed in the hidden and output layers. The hidden and output layers produce outputs using the sum of weighted values that are fed into them. They produce their output by applying an activation function to the sum of the weighted values. The NeuralShell 2 software supports 8 activation functions. The logistic function tends to be the most widely used, but for each individual problem there is a specific activation function that will work best. A list of activation functions with their equations and mapping ranges, supported by NeuralShell 2 is presented in Table 2.2. The corresponding graphical representation of each of these functions can be found in Figure 2.5 to Figure 2.12.

Table 2.2 Activation functions

Activation Function	Equation	Range of Mapping
logistic	$f(x) = 1/(1+e^{-x})$	0,1
linear	$f(x)=x$	0,1 or -1,1
tanh	$f(x)=\tanh(x)$	-1,1
tanh15	$f(x)=\tanh(1.5x)$	-1,1
sine	$f(x)=\sin(x)$	-1,1
symmetric logistic	$f(x)=(2/(1+e^{-x}))-1$	-1,1
Gaussian	$f(x)=e^{-x^2}$	0,1
Gaussian complement	$f(x)=1- e^{-x^2}$	0,1

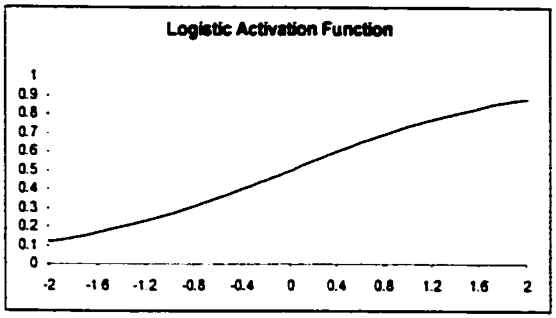

Figure 2.5 Logistic Activation Function

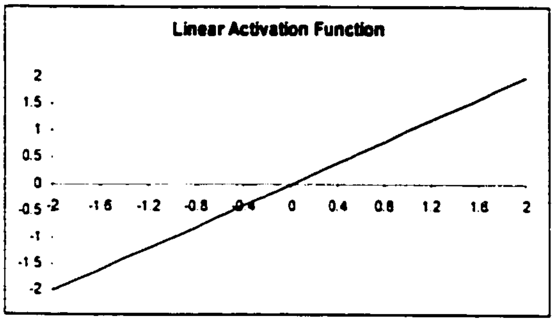

Figure 2.6 Linear Activation Function

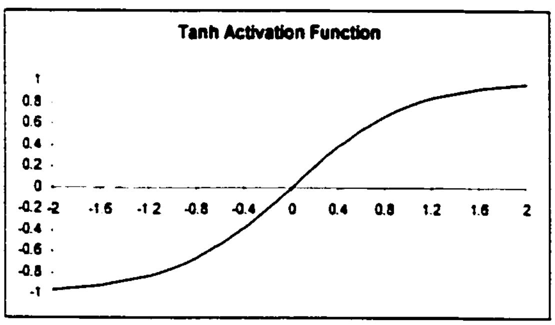

Figure 2.7 Tanh Activation Function

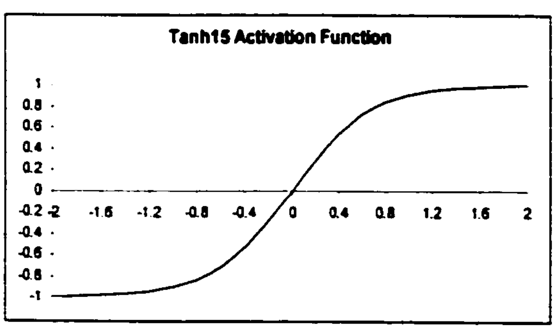

Figure 2.8 Tanh15 Activation Function

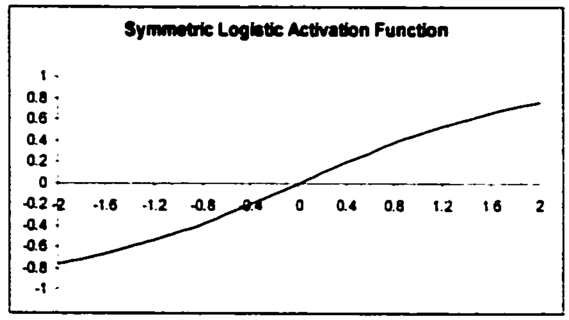

Figure 2.9 Symmetric Logistic Activation Function

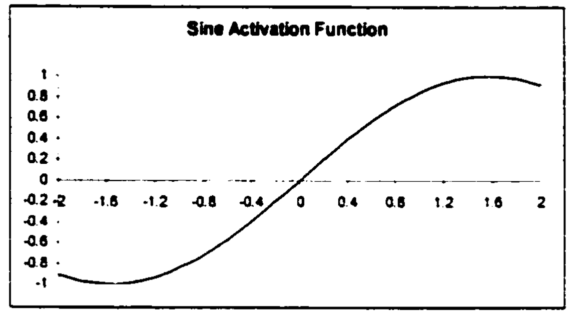

Figure 2.10 Sine Activation Function

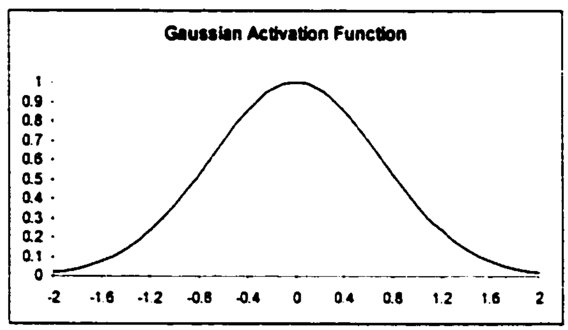

Figure 2.11 Gaussian Activation Function

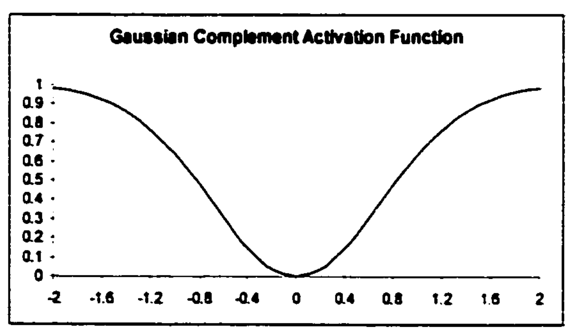

Figure 2.12 Gaussian Complement Activation Function

2.4.1.4 Learning Rate

Each time a pattern is presented in the training process, the weights leading to the output node are modified to produce a smaller error the next time the same pattern is presented. The weight modification is the error multiplied by the learning rate. Thus, for a learning rate of 0.1, the weight change is one-tenth the error. The larger the learning rate the larger the weight change that occurs, thus the learning will proceed faster. There is the risk when using a high learning rate that the model will oscillate and not converge to the desired outcome.

2.4.1.5 Momentum

Using large learning weights can lead to oscillation and nonconvergence, thus the model can converge to a solution that is not the optimum or the learning will not even complete. To allow for faster learning without the oscillation, the weight change is made a function of the previous weight change to keep the learning process in the right direction. The momentum factor controls the proportion of the previous weight change that is added into the new weight change thus providing a smoothing effect.

2.4.1.6 Weight Adjustment

The weights in the NeuroShell 2 program can be updated in 3 ways. It can use a slightly modified plain vanilla algorithm, a momentum update or a TurboProp™ update method. In using a modified vanilla algorithm only the learning rate is applied to the weight updates. A momentum update method uses the learning rate as well as a portion of the last weight change. The higher the momentum term, the greater the dampening effect will be on the weight fluctuations. This makes it ideal for problems with noisy data or where a high learning rate is desired. The last method to update the weights is by using TurboProp™. This method is not sensitive to the learning rate or momentum. The TurboProp™ method uses a separate weight update for each different weight, instead of using the same learning rate and momentum to adjust all the weights. It also differs in that when using Turboprop™ the weights are adjusted at the end of each epoch instead of after each learning event. TurboProp™ generally works better for recurrent networks, it is also recommended to use TurboProp™ when it is difficult to find the right values for the learning rate and momentum.

2.4.1.7 Pattern Selection

When training a backpropagation problem there are two different ways to present the data to the model. The first method is a rotational pattern selection. This method selects the training patterns in the order that they appear in the pattern or the training file. It is recommended to use rotational pattern selection when the training patterns are dispersed evenly throughout the data set or if training a recurrent network, as order is important. The second method is a random pattern selection. When using random pattern selection, the training pattern is randomly chosen. This also means that not every training pattern will be chosen an equal number of times. Random pattern selection is recommended for training sets that are cyclical such as data that contain seasonal variations and/or complicated problems with numerical outputs.

2.4.2 ANN Development

ANNs are computer programs, which convert one or more input signals to one or more output signals by means of an interconnected set of simple non-linear signal processors called neurons. ANNs are designed to simulate the way a simple biological nervous system is believed to operate. They are based on simulated nerve cells or neurons, which are joined together in a variety of ways to form a network. This network has the capacity to learn, memorize and recognize relationships amongst data.

2.4.2.1 Characteristics of ANNs

ANN modelling is an artificial intelligence method that mimics the human brain's problem solving ability. ANN modelling does not have pre-determined bias to the type of problem that is being modelled. Thus, the type of relationship between the input and the output is determined only from the data the model is presented with. ANNs are inherently fault tolerant, they can learn to recognize patterns, which are noisy, distorted or even incomplete. The reasoning for this is that in training, an ANN aquires redundant or distributed information encoding (storage) (Maren 1990). This is similar to the human brain. Brain cells are dying constantly with time (new ones are also being created), but our knowledge generally still remains intact. This gives ANNs the ability to produce correct or nearly correct responses when presented with partially incorrect or incomplete stimuli (Rajasekaran, et al. 1996). ANNs are unique in that they learn from the patterns (historical data) presented to them. They do not require complex mathematical formulas or algorithms. ANNs are useful for problems where the factors are known, but the interactions between the factors are not known. For example, in predicting water demand the major factors are meteorological factors and human behavior, but the exact influence of each and their interactions are not clearly known.

With the unique ability to learn from historical data, even if the historical data are noisy or incomplete, comes some disadvantages or limitations. The first limitation is that ANNs are generally considered a "black box" modelling approach; data are input and an output is produced with little knowledge of the rules governing the phenomena being modelled. There is some debate as to whether ANN models are true "black box" models as the algorithms governing the learning process are understood and defined. There is currently work going on in the ANN community to unravel the workings of the ANN black box. As a result of ANNs being considered a black box model, they are also prone to being misused in a variety of ways. It should be noted that in developing ANN models, that they are data intensive. Thus a large amount of relevant historical data is needed, as a general rule the number of training patterns needed at a minimum, is ten times the number of input parameters.

One common error is to use a data set that is missing a key factor, or containing factors unrelated to the actual problem at hand (this can lead to major deficiency in the model, especially if there is an actual correlation between the unrelated factor and the desired output). To remedy the first situation of having no historical data collected for a key input, a suitable surrogate may be used. In the case where there are no historical data collected or a suitable surrogate, the ANN technique should not be used, as it can lead to models with erroneous predictions. This is similar to giving a person only half the information in a question; there is a good chance s/he may come up with the correct or nearly correct answer to the question some of the time, but chances are s/he will not come up with the correct answer all the time. The second situation of using unrelated factors can be remedied by having knowledge in the area that is being modelled. To apply the ANN technique, an understanding of the problem and the factors affecting it, should be based on well researched information. One other concern in using ANNs is their ability to extrapolate beyond that data that they were trained on. It is not known to what extent an ANN model can extrapolate. Thus, caution should be taken with the predictions that are generated from input data that are out of the range the model was trained in.

2.4.2.2 Applications of ANN Modelling

ANN models work best in certain cases, but are not the best method in all cases. For appropriate applications of ANN modelling, the algorithm to solve the problem should be either unknown, too complex for conventional methods, or expensive to discover. Also, the problem should be data intensive, as ANNs require a large amount of data. ANNs are used in a wide variety of applications, such as stock market predictions, sales forecasts, quality control, cash flow forecasting, managerial decision making, drug screening, plus many other scientific and engineering applications.

Within the science and engineering realm, ANNs have become a popular modelling method, where applicable. They have been used successfully in areas such as ecosystem evaluation, polymer identification, chemical characterization, bacteria identification, water resource management, etc. There has also been some work in the application of ANNs that would be of interest to a water utility. This includes modelling of enhanced coagulation (Baxter, et al.

1998), automating water treatment plants (Zhang, et al. 1999), particulate removal (Chai and Andrews 1998).

ANNs have also been used for water demand forecasting in various forms. Longer-range water demand forecasting models were developed for predicting the long-term year to year forecasts (Zhang and Stanley 1997). Daniell (1991) used ANN modelling for predicting monthly water demands with great success. There has also been an application to predict chilled water demand for buildings (Hittle, et al. 1996). Chrommelynk et al. (1992) developed an ANN model for forecasting daily and hourly water demands with promising results. However, they used 54 different inputs, which is generally considered impractical.

2.4.2.3 Evolution of ANNs

The development and history of ANNs is relatively short. This is due to the fact that the computers required for the learning process for ANNs have not been available for a long period of time.

The development of the modern computer can be traced back to Charles Babbage. In around 1833, Babbage had begun designing Difference and Analytical Engines, which are considered to be the early ancestors of the modern day computers. But it wasn't until much later, that the modern type of computers became available to develop artificial neural network models.

In 1942, Norbert Weiner and his colleagues were developing what is now dubbed Cybernetics. Weiner defined what was later to be known as cybernetics, as the control and communication in the animal and the machine (Heims 1982). Indicating that biological mechanisms can be treated from an engineering and mathematical viewpoint. The most important aspect of Weiner's concept was that with the idea of viewing biological mechanisms from an engineering and mathematical perspective was the idea of feedback.

McCulloch and Pitts (1943) published what is believed to be the founding paper in the field of neural networks. McCulloch and Pitts modeled neurons as simple Binary Threshold Units (Threshold Logic Units) with fixed thresholds and uniform weights. They also concluded that any well-defined input-output relation could be implemented in a formal neural network (McCulloch and Pitts 1943).

In 1945, John von Neumann in his first draft of a report on the EDVAC (Electronic Discrete Variable Automatic Computer), made several comparisons between the proposed circuit elements and animal neurons which further contributed to the advent of ANNs.

During 1949, Donald Hebb suggested a mechanism by which real brains can learn from their experiences. Hebb described that as the synaptic strength changes; the change in strength reinforces any correspondence of activity level between the pre-synaptic and post-synaptic neurons within the brain (Hebb 1949). This is a key element of ANNs in that they can learn from historical data (experience) in a controlled training environment. In regards to ANNs, the weight on an input is increased to reflect a correlation between the input and the output neurons.

In 1958, Rosenblatt proposed a group of computational abstractions of neurons called perceptrons. Rosenblatt's perceptrons were more complex than McCulloch and Pitt's neurons, and more importantly, they were capable of learning. Rosenblatt found that there was proof that a simple training procedure (the perceptron training rule) would converge if a solution to the problem existed (Anderson and Rosenfeld 1988).

In 1969, the evolution of the ANN suffered a setback when Minsky and Papert published their book 'Perceptrons' (Minsky and Papert 1969). Minsky and Papert showed that there was an interesting type of problem, which was not linearly separable and that the single layer perceptron net could not solve, such as the XOR problem. An example of a simple XOR problem that consists of two inputs and one output can be seen in Table 2.3. They also held little hope that a multi-layered net would successfully deal with some of these types of problems. Minsky and Papert suggested that the previous work on perceptron development was without scientific value but had proceeded due to the romanticism surrounding the new idea of a machine learning. This position became the mainstream opinion and with that neural networks became an unfashionable research area in the shadows of other Artificial Intelligence areas. This was the case at least until the mid-1980s.

Table 2.3 Inputs for XOR problem

Input 1	Input 2	Output
1	1	0
0	0	0
0	1	1
1	0	1

The next two major breakthroughs occurred in 1982. First, Feldman and Ballard proposed a computationally sophisticated model of a neuron, characterized in terms of a potential. The Feldman and Ballard model also incorporated units with multiple input sites and conjunctive connections (Anderson and Rosenfeld 1988). The second major breakthrough was that John Hopfield showed that a highly interconnected network of threshold logic units could be analyzed by considering it to be a physical dynamic system (similar to an atom) possessing an 'energy'. The network is then started in some initial random state and goes on to some stable final state, this is similar to a system falling into a state of minimum energy.

One of the biggest breakthroughs occurred when Rumelhart, Hinton and Williams (1986) published the first well-known description of a back-propagation learning algorithm. It was actually first proposed by S.E. Dreyfus (1962). Then P.J. Werbos (1974) proposed a similar solution to the problem of learning in multilayer networks. Unfortunately, the solutions by Dreyfus and Werbos remained unknown to the research community and a well-known solution to the problem did not emerge until the mid-1980s. The solution was made known to the majority of the neural network community, when Parker (1985), Le Cun (1986), and Rumelhart, Hinton, and Williams (1986) independently described similar solutions to this problem. But in the end, it was Rumelhart, Hinton and Williams that reached the largest audience and popularized the solution of the backpropagation algorithm (Anderson and Rosenfeld 1988). Since then the backpropagation algorithm has been the most widely used algorithm for multilayer networks.

Error propagation uses the generalized idea of the delta rule in the back-propagation algorithm. Consider an ANN with an input layer, an output layer and one hidden layer

(Figure 2.4). The network is presented with an input that produces an output which then received by the middle or hidden layer. The output from the middle layer is received by the output layer where the final output is produced. This final output is then compared to the desired output. Since both the model output and the desired output are known, the generalized delta rule can be backpropagated to adjust the input weights in the neurons to minimize the error in the output layer.

2.4.2.4 ANN Learning Process

There are a variety of learning algorithms available of which the backpropagation algorithm is one of the most commonly used. The backpropagation algorithm is widely used and is used in this study, as it tends to learn and generalize well in most cases. As the backpropagation algorithm is the learning method used in this study, the following text will focus on the backpropagation algorithm.

A brief description of how an ANN learns follows, using a basic three-layer backpropagation network as an example of the learning process. The following description is also represented graphically in Figure 2.4. The input values in the input layer (or the first layer) are weighted and then transferred to the hidden (second) layer. The neurons in the hidden layer, then sum up the weighted values passed to them and produce an output value(s). The hidden layer then in turn passes these values to the output (third) layer. The output layer uses these values to predict the desired outcome. The predicted values are compared to the known values and an error value is computed. The output layer then backpropagates the error back into the hidden processing unit according to the learning rule being used, such as the generalized delta rule. The generalized delta rule modifies the strength of the input connections to reduce the squares of the differences between the predicted output value and the actual output. In the learning process, each individual pattern that is input into the network and the associated adjustment of the connection weights are considered one learning event. During the learning process, when each pattern has undergone a learning event it is called an epoch. The error backpropagation is repeated until the ANN produces an acceptable minimized error.

2.4.2.5 ANN Model Development

Three steps are involved in the ANN model development process:

1) Source Data Analysis
2) ANN Model Development
3) Model Evaluation

2.4.2.5.1 Source Data Analysis

There are two major components in the source data analysis. First is the familiarization with the problem. This involves an assessment of what factors are important in the process or problem that affect the output. Second is ensuring that the ANN modelling technique is applicable. This also means that the proper (quantity, quality and applicable) data is available.

2.4.2.5.2 ANN Model Development

The next step is the actual development of the models. This step involves the design and training of a variety of potential models. In the ANN model development step the data are generally divided into three data sets; training, testing and production sets. The training set of data is the data set that the ANN model actually trains on. The testing data set is used to prevent overtraining the network so that the model will generalize well on new data. During the learning process, the network is learning from the training set data, but at set intervals the network evaluates how well it can generalize on the testing set. This prevents the model from over-learning or memorization, and insures that it learns to generalize on the data. The last data set is the production set. The production set is used to evaluate the networks, as it is an independent data set that the network has not seen during the training process. The final step is to test and evaluate the models on the production set. The evaluation of the networks will be discussed in the next section.

2.4.2.5.3 Evaluation Of ANN Models

The evaluation of the models should be carried out in two steps. The first is to test the models on data that the ANN has not seen during the training process. If the model performs acceptably, then the second part of the test is to verify the final model on real-time data. A good indication of whether the ANN has generalized well, is to compare the results from the training, testing and production sets. They should be similar. If it is found that the training and/or testing sets result in significantly better results than the production set, there is a chance that the ANN has over learned or memorized the training data. There is also the possibility that the production set contains a portion of data that is out of the range in which the ANN was trained. This is unlikely if the production set was randomly chosen, but is a distinct possibility if it was chosen manually (such as when one specific year of data is used for the production set).

There are two methods of evaluating the models: quantitatively and qualitatively. The correct solutions and those produced by the network may be compared in a qualitative manner, such as with a visual comparison of plotted points or in a quantitative manner using a statistical test, such as the correlation coefficient (Flood and Kartam 1994). It is desirable to use both of these methods, to insure the best possible model is developed. It is also beneficial to use a variety of quantitative comparisons. The results of different statistical tests, test different areas of performance for the model developed.

3 Methodology

3.1 Data Collection

In using the ANN modelling method, one of the most important steps is obtaining the proper input and output data. Once the proper data are collected, the data then must be analyzed for any errors or missing data. Any data that appears to be in error then needs to be verified as such.

3.1.1 Water Demand Data

The water demand data were collected by EPCOR Water Services, using a PLC-PC based SCADA system. The water demand data is measured and collected every 15 minutes. The water demand is calculated by taking water produced in the 15-minute time frame and pumped into the distribution system minus the change in the storage in the distribution system. The volume of water in storage is measured every 15-minutes at each reservoir and sent into the SCADA system. The errors associated with the water demand measurements are 2.5% for 250 ML/day and 5% for a 500ML/day as measured by using a draw down test.

3.1.2 Meteorological Data

The meteorological data were obtained from Environment Canada. Environment Canada verifies the data before they are released in electronic form. Monthly summaries of the meteorological data are sent out at the end of every month. Weather forecasts can be obtained daily through text or phone recordings. It is also possible to get on-line real time weather data and forecast data.

3.2 Data Selection

Twenty-seven months of data were collected. These included daily water demand, minimum, mean and maximum air temperature, rainfall, sunshine hours, and an hourly water demand input. In the collection of data, there is always the potential for errors in the data collected. They can either be equipment or human type collection errors. In the data collected for this project, the only errors apparent were due to equipment type errors. There were no human type errors apparent in the data. This is mainly due to the fact that all the data were collected through on-line data collection systems, such as the EPCOR Water Services SCADA system. Any errors in the data collection process were apparent as the SCADA system recorded them with "??? ". Only one other type of error was found where the recorded water demand remained constant throughout the entire day. All erroneous data were removed from the data set.

3.3 Model Development

In developing the models for the both the daily water demand and the demand for the following 11 days, there are many different factors that can be adjusted to optimize the model. These include the inputs used, type of architecture, activation function, number of neurons, the ratio of training : testing: production that is used to divide the data, scaling function, learning rate, momentum, weight updates and the type of pattern selection that is used.

3.3.1 Input Selection

After the data are collected and selected, each possible input was analyzed to determine the relationship of the input to the water demand and its feasibility to be incorporated into the water demand model.

3.3.1.1 Meteorological Data

The meteorological conditions are considered to play a very important role in daily water demand. This is evident by the previous work carried out in forecasting water demand, information provided by operators, and the source data analysis that follows.

3.3.1.1.1 Temperature

The relationships between the minimum, mean and maximum temperature, with respect to the daily water demand, was very similar. Over the range of available data an exponential relationship between the daily water demand and the temperature is found in Figure 3.1. Whether it is best to use the maximum, mean, minimum or a combination of these temperatures for the day can not be determined at this point. In the past the maximum temperature has been included in many of the models developed for water demand forecasting, so it was used as an input into the model initially until, a later stage, where it can be determined which is the optimum temperature input or combination of inputs.

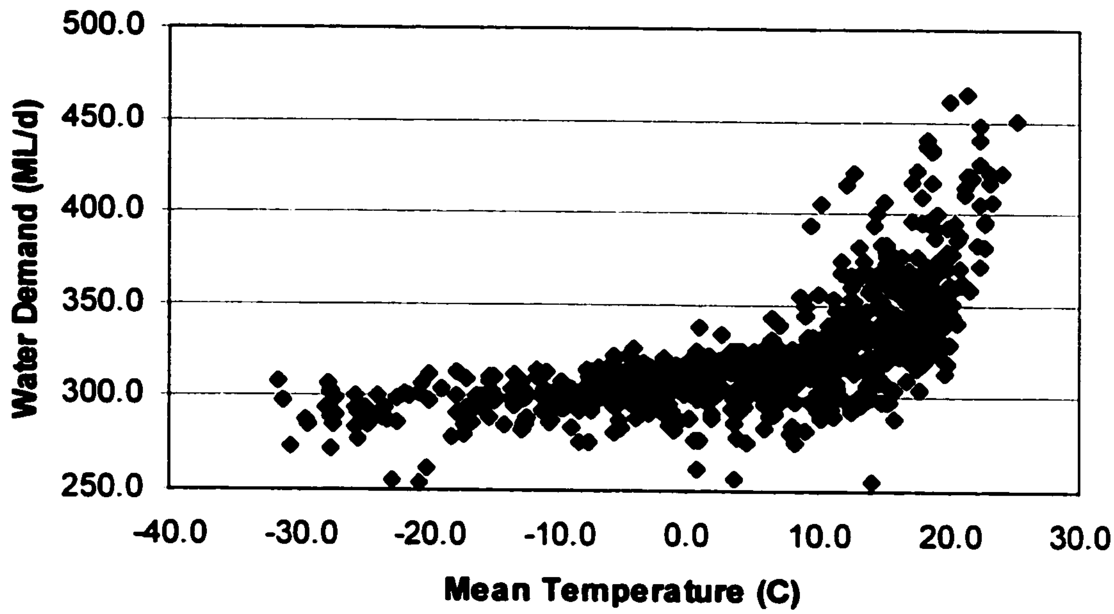

Figure 3.1 Comparison of Daily Water Demand vs. Temperature

3.3.1.1.2 Sunshine Hours

Sunshine hours are defined as the number of hours of "bright" sunshine, unhindered by fog or cloud cover. The number of sunshine hours has been used in the past to model water demand, but faces one major problem when used in forecasting water demand. Although sunshine hours can be measured, it is difficult to predict the number of sunshine hours in any given day. This is quite different to the temperature, which can be predicted generally within a few degrees. This factor was not used initially in the development of the model, but will be tested for its significance at a later stage to determine if, in the future, a surrogate measurement should be collected to be used in future model development. The most likely surrogate measurement would be using a cloudiness index of sorts, as weather forecasts include the cloudiness in descriptive terms already (i.e. clear, scattered clouds, mainly cloudy etc.).

3.3.1.1.3 Rainfall

The relationship between the rainfall data and the water demand is less evident, as during the winter months there is little or no rainfall. This then gives the general impression that when little or no rainfall occurs that the water demand is low. In reality the daily water demand in the non-summer months tends to have little variation and is generally lower than in the summer months. The opposite is generally true in the summer months, with less rainfall, the water demand tends to increase. The contrast between the summer and winter rainfall vs. daily water demand can be seen by comparing Figure 3.2 through Figure 3.7. Regardless of the relationship, the rainfall is also considered to be an important factor that influences the water demand. While rainfall is considered important, the duration since the last rainfall also has to be considered. The short term and long term rainfall plays an important role in determining the water demand, thus both must be included. A previous day rainfall (mm), a previous 5 days of rainfall (mm) and the previous 30 days of rainfall (mm) inputs were included in the model to reflect both the short and long term impact that rainfall has on water demand. While the short term inputs are to reflect the recent precipitation events, the long term input of 30 days may be taking into account an excessive period of time. To confirm that the 30-day period should be used as the long-term rainfall input, a 15-day and 45-day input will be tested against the 30-day input. It should be noted that the probability of precipitation (POP) was to be included as an input into the model and most likely should be. Environment Canada does not collect the probability of precipitation data and EPCOR

Water Services has only a partial record. Probability of precipitation data should be collected in the future to be included into future water demand models.

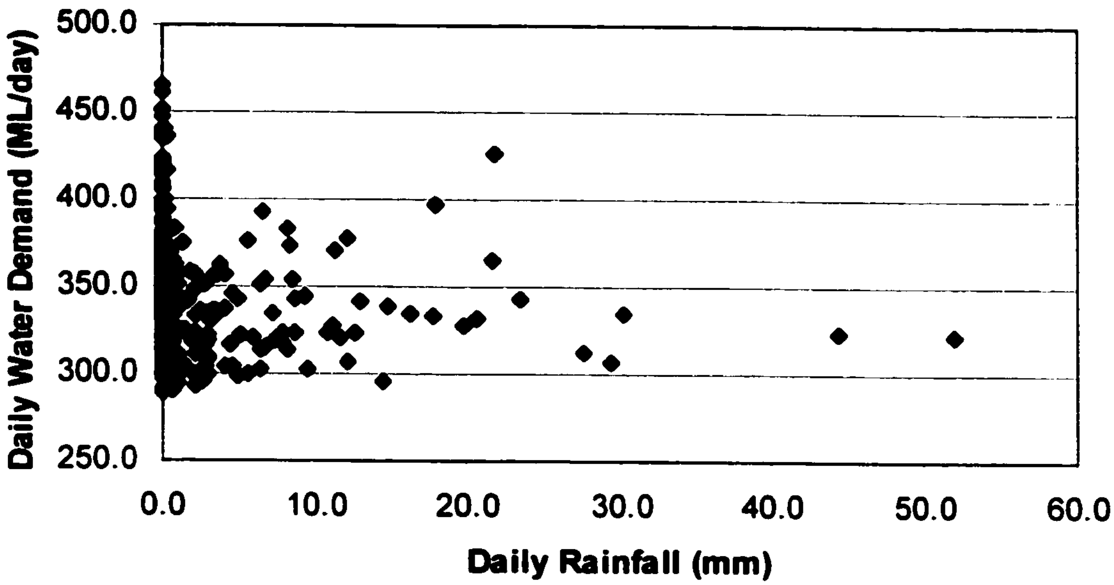

Figure 3.2 Daily Water Demand vs. Daily Rainfall (Summer)

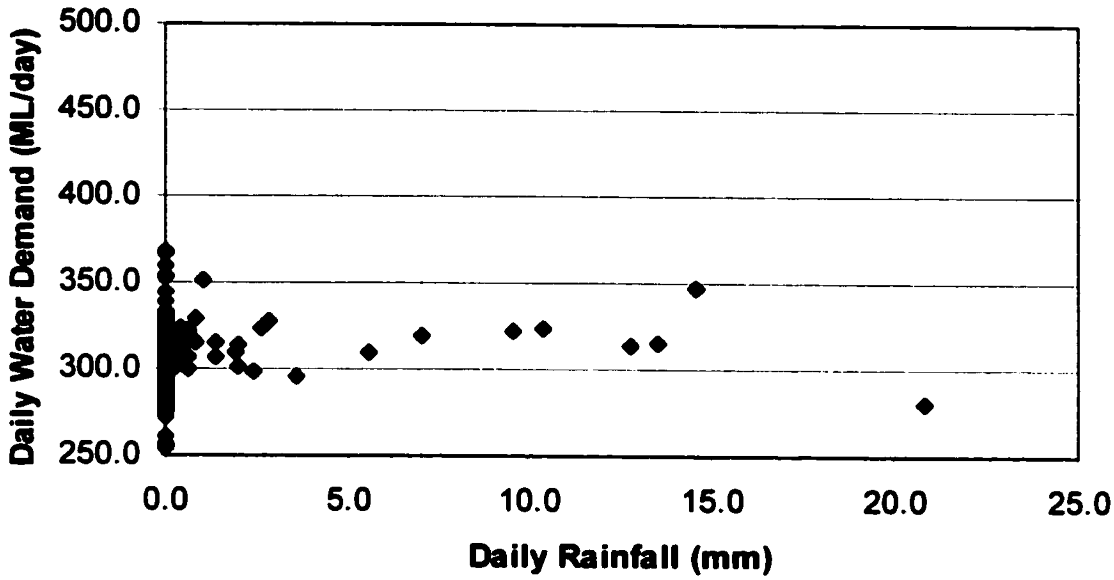

Figure 3.3 Daily Water Demand vs. Daily Rainfall (Winter)

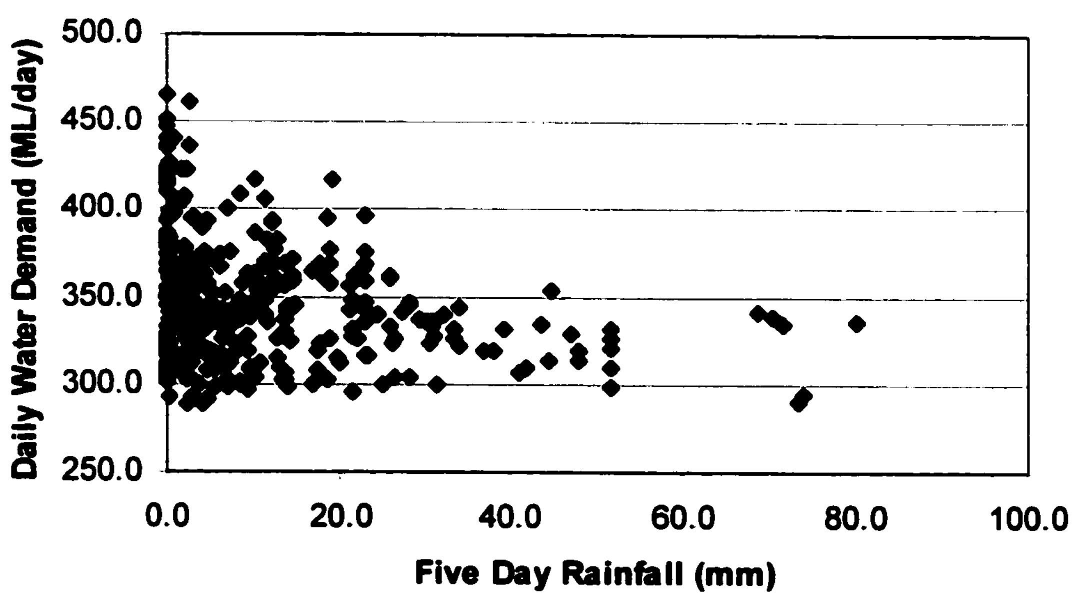

Figure 3.4 Daily Water Demand vs. Five Day Rainfall (Summer)

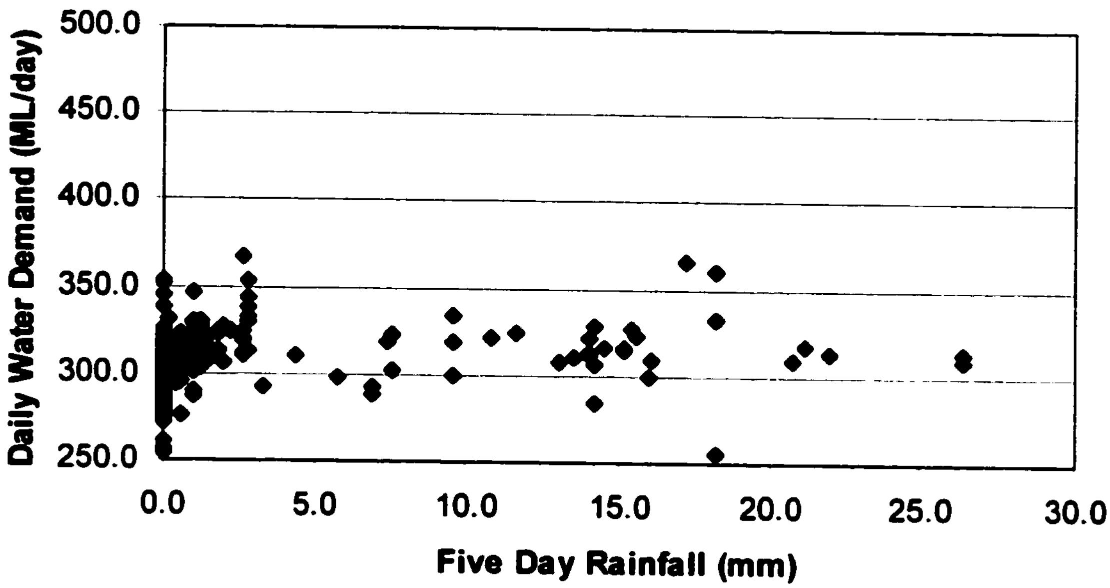

Figure 3.5 Daily Water Demand vs. Five Day Rainfall (Winter)

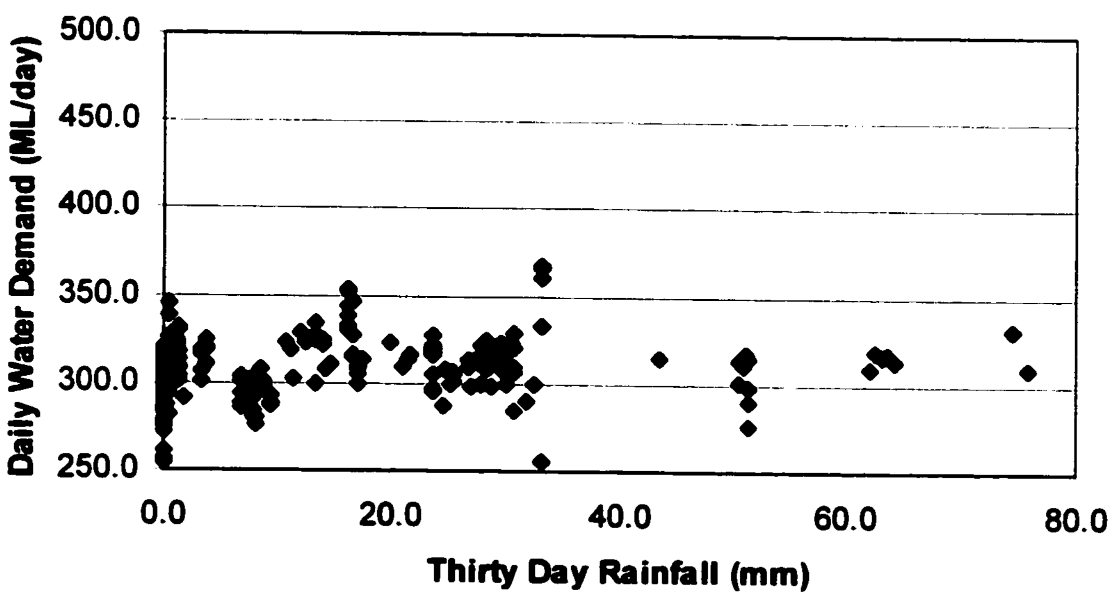

Figure 3.6 Daily Water Demand vs. Thirty Day Rainfall (Winter)

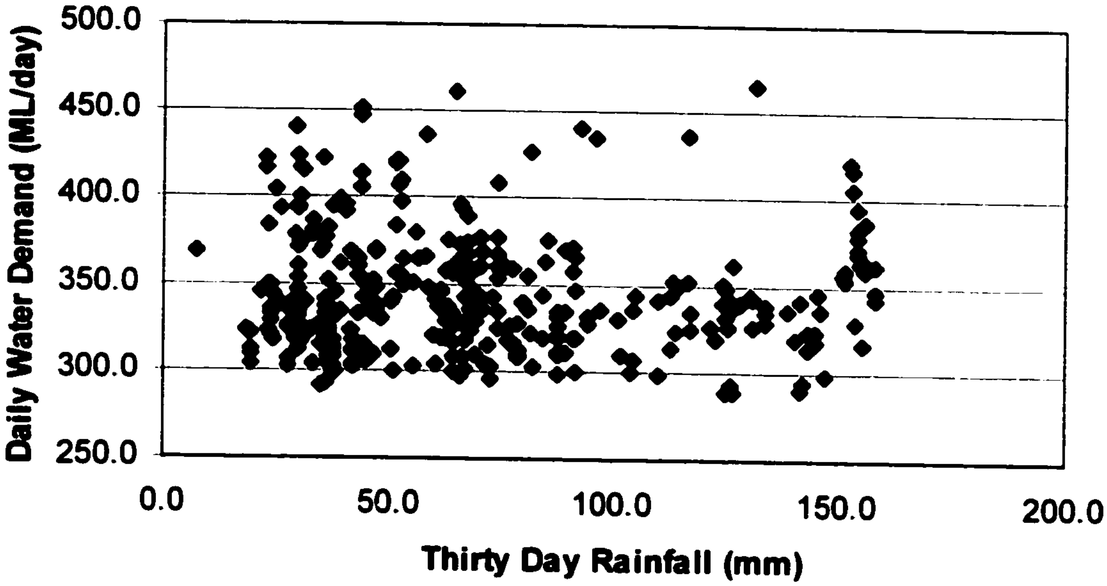

Figure 3.7 Daily Water Demand vs. Thirty Day Rainfall (Summer)

3.3.1.2 Human Behavioral Data

In predicting water demand, we must look at the different sources that create the water demand. The water demand is typically broken down into industrial, commercial and residential use. The water demand is dependent on these different uses. Each of these different sectors varies in their use of water in quantity and time. To account for these differences, a couple of different index inputs are used to discern the relationship for the seasonal variation and the weekend/weekday variation.

3.3.1.2.1 Seasonal Variations

The daily water demand was plotted as a function of the date, in Figure 3.8. The main feature evident in this figure is that the water demand is generally higher and more variable in the summer than in the winter. The winter months are generally characterized as having low water demand with little variation. This is mainly due to the different activities that the general public partakes in. Thus, this type of human behavior is still weather-driven. The seasonal variations not only affect the water demand for the entire day, but also affect the profile of the 24-hour demand period. Thus a season index was included as an input into the model initially, but two separate models will be tested, one for each 'season', to determine if these are a better predictor than the single model. The 'seasons' were broken in up into a May 1 – September 30 summer season and an October 1 – April 30 non-summer season. This reflects the higher demand that occurs during the summer season compared to the non-summer season as reflected in Figure 3.8.

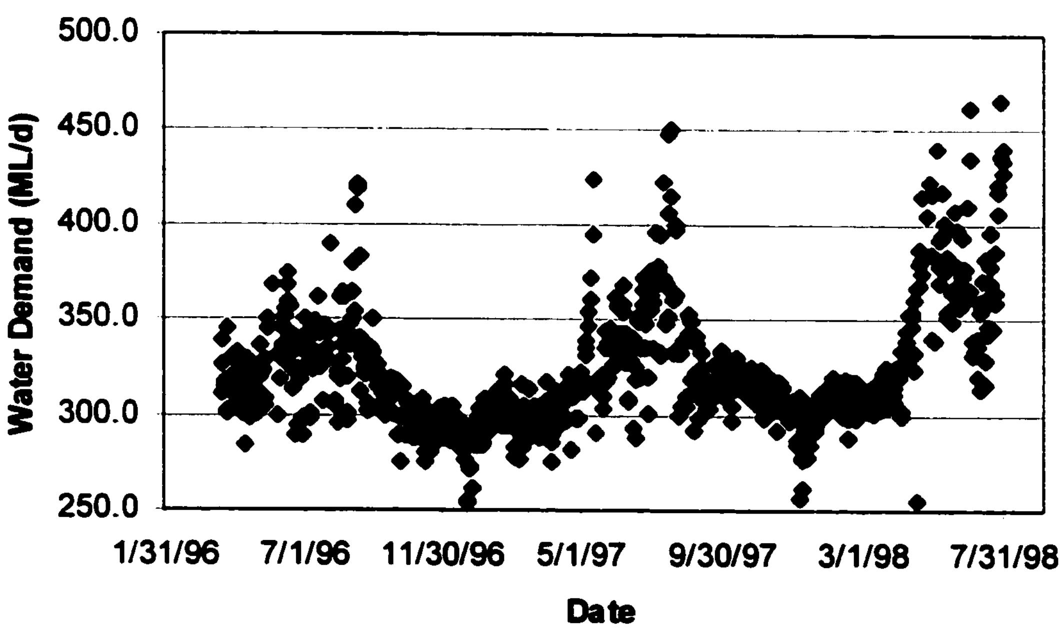

Figure 3.8 Daily Water Demand Vs. Date

3.3.1.2.2 Weekly Variations

Throughout the week, the water demand varies depending upon the day of the week. More specifically the water demand varies between weekdays and weekend/holidays both in quantity for the entire day and the profile of use for the 24-hour demand period. This weekend/weekday index is included as an input in the initial model development.

3.3.1.2.3 Reference Indicator

A reference indicator is used to give the model a general estimate of the water demand for the day. Either of two different references can be used. The first is the previous day's water demand. The previous day's water demand generally gives a good indication of the current day's water demand. The problem with using this as a reference is that it does not give a good reference point when the water demand changes dramatically from one day to the next. The other possible reference is to use the water demand for the preceding hour that the water forecast is being developed for. Since EPCOR Water Services develop their water demand forecast at 10:30 am, the water demand from 09:00 am to 10:00 am can be used as an indication or reference of the water demand for the rest of the day. The 9:00 to 10:00 water demand was used as an input initially, but was later tested to see if it is necessary. The previous day's water demand was also tested to see if it was a better reference point to use as an input.

3.3.2 Architecture Selection

Initially the data were divided into a basic training : testing : production ratio, such that no one set contains significantly more data than the others do. The model was then trained using the training and testing sets on each type of architecture using the default settings. Each model was then applied to the production set and the results were compared to see which architectures were suitable for development.

3.3.3 Scaling Function

The scaling function was simply chosen by training the model and varying the scaling function used on the inputs for each run. The scaling function that achieves the best results and met the needs of the problem was then chosen.

3.3.4 Activation Function

The model is then run while changing the activation function for each run. The results of using the different activation functions are then compared. The activation function that yields the best result is then used.

3.3.5 Number of Neurons

The number of neurons in each hidden layer can be adjusted to further improve the water demand model's performance. In a simple 3-layer backpropagation network, finding the optimum number of neurons is relatively easy as it contains only one hidden layer. In finding the optimum number of neurons for other architectures where 2 or more hidden layers exist, the task become more daunting, as the number of neurons in the first hidden layer affects the optimal number of neurons in the second and vice versa. The number of neurons in each hidden layer is not independent of each other. Thus, finding the optimal number of neurons can not be done by simply finding the optimal number of neurons in one layer and then the next. Here the optimum number of neurons should be found in conjunction with each other and not separately.

3.3.6 Pattern Selection

In the training process of backpropagation problems there are two different ways to present the data to the model. The data can either be presented randomly where the next pattern is chosen at random, or rotationally where the next pattern is chosen sequentially. To evaluate which type of pattern selection was best, each method was tried and the better result of the two was chosen. It was expected that the random pattern selection would yield the better result of the two as it is generally recommended for training sets that contain cyclical data that contain seasonal variations.

3.3.7 Learning Rate and Momentum

The learning rate and momentum control how the weights are adjusted during the ANN training process. Since both the learning rate and momentum are not mutually exclusive in the optimization process, the optimum combination of the two is obtained using a surface plot of the results.

4 Results

4.1 Daily Water Demand Model

4.1.1 Architecture Selection

To determine the architecture that is suitable for the daily water demand model, the daily water demand model was trained and tested on each architecture that is supported by the NeuroShell 2 software. The results of the architecture selection can be seen in Table 4.1. After the model was tested on each architecture the best possible architectures for the problem were determined to be the 4 layer backpropagation networks (Figure 4.1), as well as, the Ward net with 3 hidden slabs (Figure 4.2) and the Ward net with 2 hidden slabs and a jump connection (Figure 4.3). It is important to realize when assessing the architecture, or any other parameter, that more than one-selection criterion is needed. Different statistical tests, test different areas of performance for the models being developed. In using the mean absolute error, it is possible to determine which architecture gives the better predictions on average, but it does not give information regarding the general fit of the model. It also does not take into account whether the model is capable of hitting peaks and valleys of the water demand. The square of the Pearson's Product Moment Correlation Coefficient r^2 value is a measurement of the linear association between two variables. The equation for the Pearson Product Moment Correlation Coefficient can be found in Equation 4.1. It compares the accuracy of the model to the accuracy of a benchmark model where the prediction is just the mean of all of the samples. Where r^2 has a value between 0 and 1, with 0 being the value obtained if the average value of the output parameter was used for prediction. A perfect fit would have a r^2 value of 1, and very good fit model would have a value close to 1. The models with the architectures that were chosen to be developed further also predicted the water demand within 10% of the actual demand at all times on the production set. None of the other model architectures where able to accomplish this.

Table 4.1 Architecture Selections Results

Architecture	r^2	Mean Absolute Error (ML/day)
3 Layer Backpropagation	0.8949	8.037
4 Layer Backpropagation	**0.8952**	**7.994**
5 Layer Backpropagation	0.8927	8.059
Recurrent Net With Input Layer Feedback	0.8852	10.488
Recurrent Net With Hidden Layer Feedback	0.7793	13.930
Recurrent Net With Output Layer Feedback	0.7713	14.462
Ward Net 2 Hidden Slabs	0.8879	8.158
Ward Net 3 Hidden Slabs	**0.8949**	**7.925**
Ward Net 2 Hidden Slabs With Jump Connection	**0.8980**	**7.841**
3 Layer Backpropagation With Jump Connections	0.8700	8.886
4 Layer Backpropagation With Jump Connections	0.8726	8.812
5 Layer Backpropagation With Jump Connections	0.8687	8.994

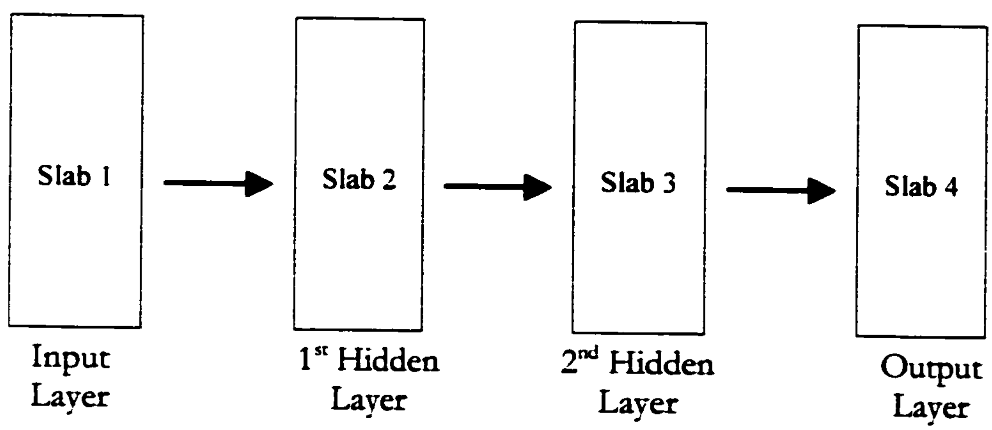

Figure 4.1 Four Layer Backpropagation Architecture

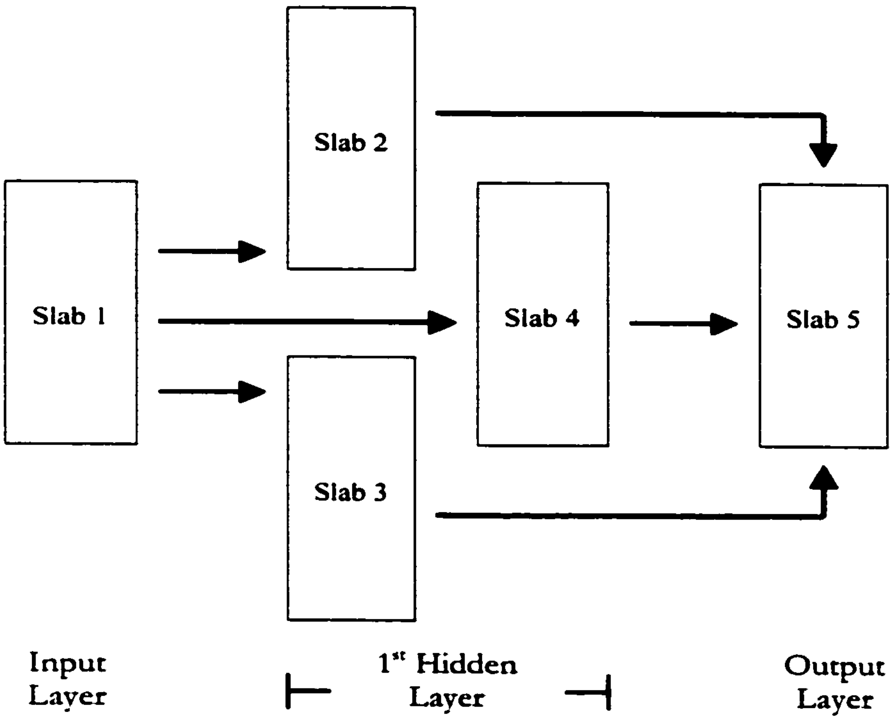

Figure 4.2 Ward Network with 3 Hidden Slabs

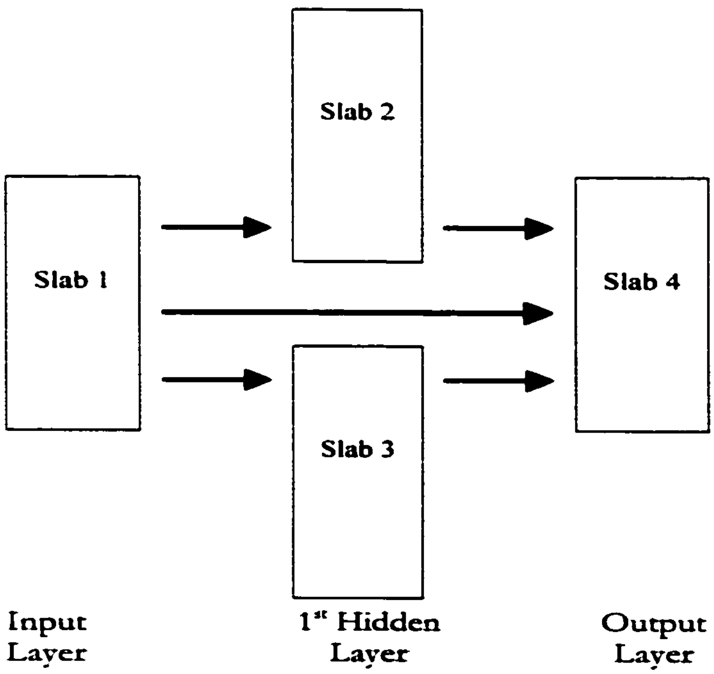

Figure 4.3 Ward Network 2 Hidden Slabs with a Jump Connection

Equation 4.1

$$r = \frac{nXY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$$

r = Pearson Product Moment Correlation Coefficient

n = numbered of paired observations

X = variable A

Y = variable B

4.1.2 Input Selection

Initially the models where trained using as inputs the maximum temperature, previous day rainfall, previous 5 days of rainfall, previous 30 days of rainfall, season index, weekend/weekday index and a reference indicator. These inputs into the model were used initially as they were deemed to be significant or their function was deemed necessary. As with the selection of possible architectures, the inputs can be further analyzed to determine their significance. Other inputs can also be included to see if they produce a better overall model. Also inputs can be removed or replaced by similar inputs to see if there are any improvements.

4.1.2.1 Temperature Input Selection

Initially, the maximum temperature was used as the initial input, as it has been shown to be an important parameter in other water demand forecasting models. At this stage the temperature input into the model was varied to see if a different input or combination of temperature inputs yielded a more robust model. Using the mean and minimum in place of the maximum input caused the model to be slightly less accurate. However, by using the minimum and maximum inputs in combination, the model tended to yield a slightly better result, as can be seen in Table 4.2. Thus the minimum and maximum daily temperatures were used for the subsequent model development.

Table 4.2 Temperature Input Selection Results

Architecture	Temperature Input	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	minimum/maximum	0.8953	7.993
	minimum	0.8869	8.253
	mean	0.8889	8.196
	maximum	0.8952	7.994
Ward Net 3 Hidden Slabs	minimum/maximum	0.8964	7.901
	minimum	0.8894	8.112
	mean	0.8947	7.928
	maximum	0.8949	7.925
Ward Net 2 Hidden Slabs With Jump Connection	minimum/maximum	0.8978	7.777
	minimum	0.8942	7.964
	mean	0.8980	7.847
	maximum	0.8980	7.841

4.1.2.2 Sunshine Hours Input Results

The use of sunshine hours in the model is limited by the fact that sunshine hours are not readily forecasted. Here, the actual sunshine hours are used as an input to see if there is an improvement in the model. If the model shows improvement, then there is the possibility to include a surrogate measure such as the cloud cover index as weather forecasts include cloud cover in descriptive terms. A cloud cover index would be tried here, but the weather forecasts are not archived by Environment Canada, only the actual weather for the day. Using the sunshine hours in the model leads to some improvement in the water demand prediction of all the models as is shown in Table 4.3. It is recommended that the cloud conditions be recorded for future water demand models.

Table 4.3 Sunshine Hours Input Selection

Architecture	Sunshine Hours	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	without	0.8953	7.993
	with	0.8983	7.911
Ward Net 3 Hidden Slabs	without	0.8964	7.901
	with	0.9031	7.542
Ward Net 2 Hidden Slabs With Jump Connection	without	0.8978	7.777
	with	0.9034	7.608

4.1.2.3 Rainfall Input Selection

Through initial work it was found that the previous day's rainfall, the previous 5 day rainfall and the previous 30 day rainfall were significant inputs, as they represent the short and long term precipitation conditions. The 30-day rainfall was replaced with a 15 and a 45-day rainfall input. This was to check that the 30-day rainfall is the most representative long-term rainfall input with regards to water demand. In analyzing the results, it can be seen in Table 4.4 that varying the long-term rainfall input did not to a large extent change the ANN utilizing the 4-layer backpropagation network. The 15 and 45 days of previous rainfall inputs improved the ANN that incorporates the Ward network with 3 hidden slabs over using the 30 days of previous rainfall input. Using the Ward network with two hidden slabs and a jump connection, the 15 days of previous rainfall input yielded the worst result while the model using the 45 days of previous rainfall had the best results as measured by the r^2 and the mean absolute error. Inclusion of the 45 days of previous rainfall improved the ANN model overall but it also increased the maximum error of prediction beyond 10%. While it is important to have as accurate a model on average, it is also important to minimize the maximum error. With the conflicting and negligible difference seen in Table 4.4, the significance of the long-term rainfall effect comes into question. To test the importance of the long term rainfall input, it was removed leaving the previous day rainfall and the previous 5 days of rainfall as the only rainfall inputs. In Table 4.4, it can be seen that the removal of the long term rainfall input has little effect on the model, in fact it causes a slight improvement when using the 4 layer backpropagation network. Therefore the long-term rainfall input was removed from the subsequent modeling.

Table 4.4 Long Term Rainfall Input Selection

Architecture	Long Term Input Previous Rainfall	r^2	Mean Absolute Error (ML/day)
	no long term input	0.9013	7.717
4 Layer	15 days	0.8942	7.983
Backpropagation	30 days	0.8953	7.993
	45 days	0.8933	7.999
	no long term input	0.9018	7.642
Ward Net 3 Hidden	15 days	0.8978	7.715
Slabs	30 days	0.8964	7.901
	45 days	0.9024	7.645
Ward Net 2 Hidden	no long term input	0.8996	7.693
Slabs With Jump	15 days	0.8957	7.893
Connection	30 days	0.8978	7.777
	45 days	0.9036	7.647

As mentioned in Section 3.3.1.1.3, the probability of precipitation should be considered as a possible input, but the data are not available. Instead of using the probability of precipitation, the actual rainfall for the day was used to see if it led to an overall improvement in the model. The other option was to use an input indicating whether or not rainfall occurred on the current day. Neither one of these options led to an improvement in the model, thus at this stage it appeared that the probability of precipitation would not make any significant improvement in the daily water model. It still can not be totally dismissed, as without actually using the probability of precipitation, the true significance of this input can not be known.

4.1.2.4 Reference Indicator Input Selection

A reference indicator gives a general indication of the water demand for that day. The need for a reference indicator in the daily water demand forecast model was investigated. In studying the possible reference indicators it was found that by using a simple linear relationship between the daily water demand and the reference indicator of the water demand from 9:00 to 10:00 am, or the previous day's water demand, an adequate approximation could be found. By using a linear relationship, obtained with previous day's water demand, the relationship can be used to predict the current day's water demand with a r^2 of 0.7532 and a mean absolute error of 11.709 ML/day. Using the same linear relationship, but using the water demand from 9:00 to 10:00 am instead of the previous day's water demand an r^2 of 0.7608 and a mean absolute error of 12.554 ML/day are obtained. Thus it is unclear at this time as to which reference indicator produces the best result. What is known is that the reference indicator is needed. If using the reference indicator alone, it is possible to obtain on average, a better water demand prediction for the current day then when using no reference indicator in the neural network models. It can be seen in Table 4.5 that the reference indicator is a very important input into the model. In comparing the results of using the different reference indicators and no reference indicator it can be seen that reference indicators led to an improvement in the modeling process. It also can be seen that using the 9:00 to 10:00 am water demand as the reference indicator yielded better results than using the previous day's water demand. One explanation is that using the previous day's water demand leads to less accurate water demand predictions on Saturday, when a weekday water demand would be used to predict a weekend demand and vice versa for Monday. In using the water demand from 9:00 to 10:00 am of the current day for the daily water demand forecast, the reference indicator being used is from the same day, as the one being predicted. Thus, it should be more accurate in that it does not contain the error of using a weekday reference indicator for a weekend demand and vice versa. This weekend/weekday error from using the previous day's water demand as an input would also be present in using a simple linear relationship between the previous day's water demand to forecast the present day water demand.

Table 4.5 Determination of the Reference Indicator Input

Architecture	Reference Indicator	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	no indicator	0.6725	12.491
	previous day's water demand	0.7997	9.901
	9:00 to 10:00 water demand	0.9013	7.717
Ward Net 3 Hidden Slabs	no indicator	0.6588	12.679
	previous day's water demand	0.8050	10.061
	9:00 to 10:00 water demand	0.9018	7.642
Ward Net 2 Hidden Slabs With Jump Connection	no indicator	0.6657	12.589
	previous day's water demand	0.8008	10.339
	9:00 to 10:00 water demand	0.8996	7.693

4.1.3 Activation Selection

A variety of activation functions where tried in the hidden layers. After trying a variety of combinations of activation functions it was found that the tanh15 function in the first and second hidden layers of the 4-Layer backpropagation network yielded the best results, the rest of the results for the other 2 architecture types can be found in Table 4.6. Ward networks are designed to have different slabs with different activation functions. This gives them the benefit of having the data presented in two different forms to detect different features in the data. What is interesting is that the Ward net with 2 hidden slabs and a jump connection yielded the best result when both the slabs had gaussian activation functions. This is contrary to the main design purpose of the Ward nets. The most likely explanation for this is that the gaussian activation function is designed more for the mid-range data while the gaussian complement is for the high and low range data. With more mid-range data presented than the extreme low and high range data the model may actually improve its mid-range prediction and decrease its performance at the low and high end of the data, but overall the performance would be better. Whether the decrease in prediction accuracy on the high and low data range was actually occurring was investigated. It was found that the model with both gaussian activation functions had the lowest maximum error associated with it and the distribution of error was very similar to that when using the other activation functions. Thus, the effect of the model increasing its performance due to the concentration of mid-range data is not the case, and is just the result that the model improved slightly at the high, mid and low range of the data.

Table 4.6 Activation Function Selection

Architecture	Activation Function 1[st] Hidden Layer	 2[nd] Hidden Layer	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	tanh15 →	tanh15	0.9056	7.570
Ward Net 3 Hidden Slabs	tanh15 → gaussian → sine →		0.9053	7.593
Ward Net 2 Hidden Slabs With Jump Connection	gaussian → gaussian →		0.9040	7.636

4.1.4 Neuron Selection

A grid system was used to determine the optimum number of neurons needed in the hidden layer to achieve optimum results. The grid system for the 4-layer Backpropagation network was from 6 to 16 neurons in each hidden layer and were adjusted in one neuron increments. The surface plots of the results (Figure 4.4 and Figure 4.5) show that there are three different areas of interest that yield the best results. Closer examination of the data shows that, the model using 8 neurons in both the first and second hidden layers yielded the best result with an r^2 of 0.9081 and a mean absolute error of 7.431 ML/day (2.29% error).

The r^2 and mean absolute error of the Ward network with 3 hidden slabs were set-up in a grid system similar to the 4-layer backpropagation network. All 3 slabs are considered to be one hidden layer as they receive their input from the input layer and their output is fed directly into the output layer (Figure 4.2). There was one change as there are 3 hidden slabs in this architectural set-up. An equal number of neurons were used in the first and second hidden layers while a different number of neurons were used in the third hidden slab. The resulting grid runs are represented graphically by surface plots in Figure 4.6 and Figure 4.7. These results show that the highest r^2 values occur when there are 8 neurons in both the second and third slabs and the 7 neurons in the fourth slab. Using these values for the number of neurons the model obtained a r^2 of 0.9069 and a mean absolute error of 7.493 ML/day (2.31% error).

The grid system for the Ward network with two hidden slabs and a jump connection was set up with from 9 to 15 neurons in each hidden slab and the number of neurons was adjusted in one neuron increments. The surface plots from the resulting grid system revealed that the optimal number of neurons was 12 neurons for each of slabs in the hidden layer (see Figure 4.8 and Figure 4.9). This configuration resulted in a r^2 value of 0.9040 and a mean absolute value of 7.636 ML/day (2.36% error).

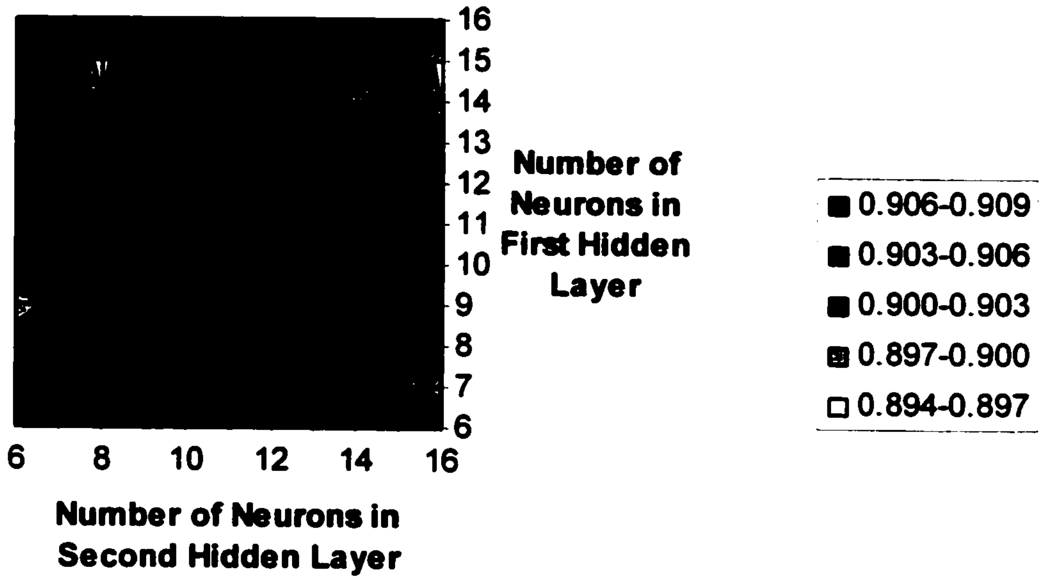

Figure 4.4 r^2 Surface Plot for Neuron Selection of 4-Layer Backpropagation Network

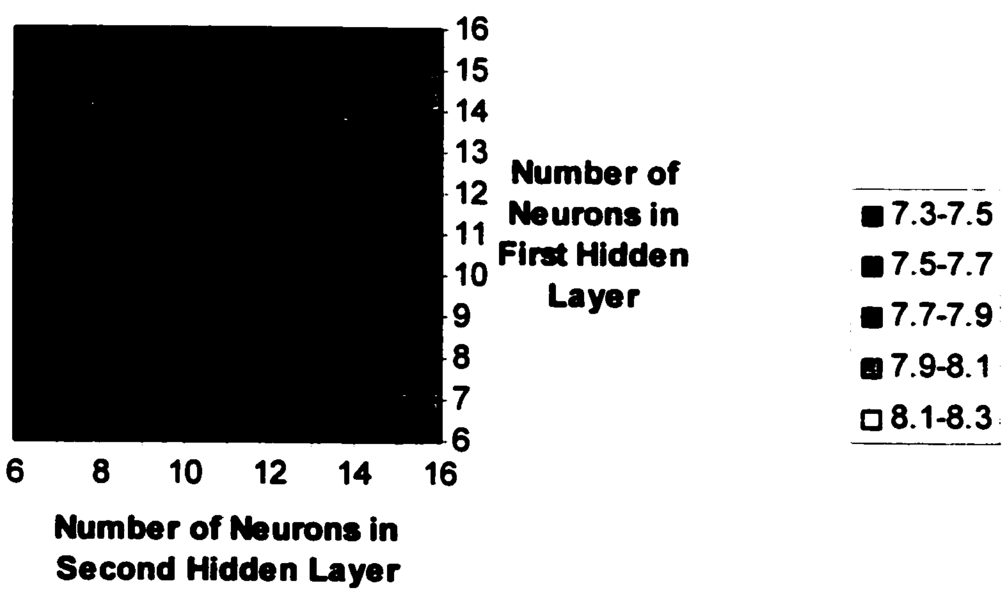

Figure 4.5 Mean Absolute Error Surface Plot for Neuron Selection of 4-Layer Backpropagation Network

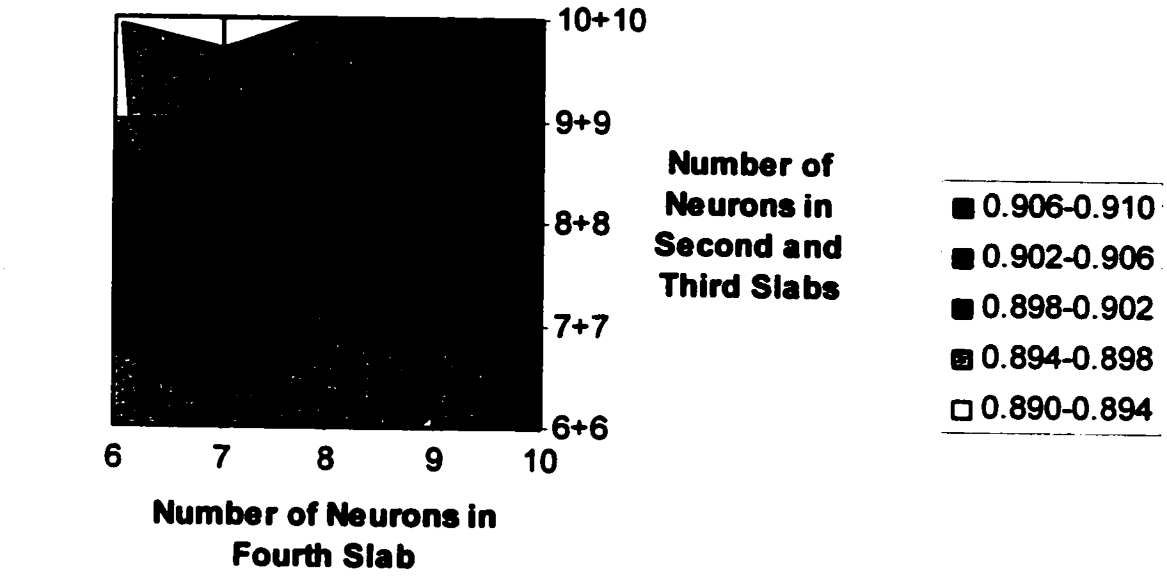

Figure 4.6 r^2 Surface Plot for Neuron Selection of Ward Net with 3 Hidden Slabs

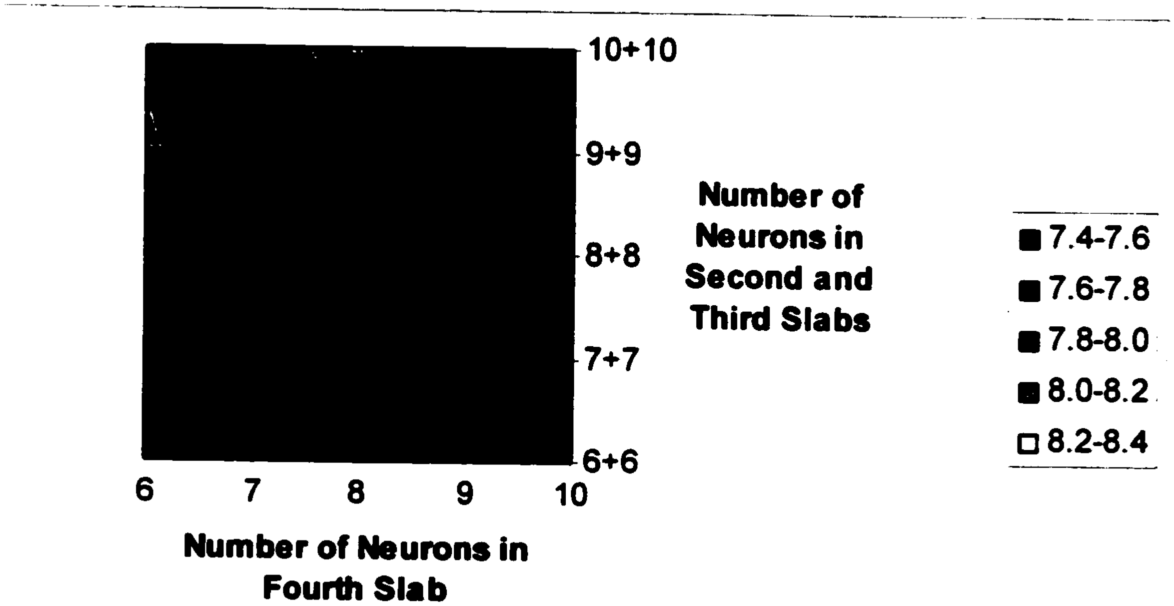

Figure 4.7 Mean Absolute Error Surface Plot for Neuron Selection of Ward Net with 3 Hidden Slabs

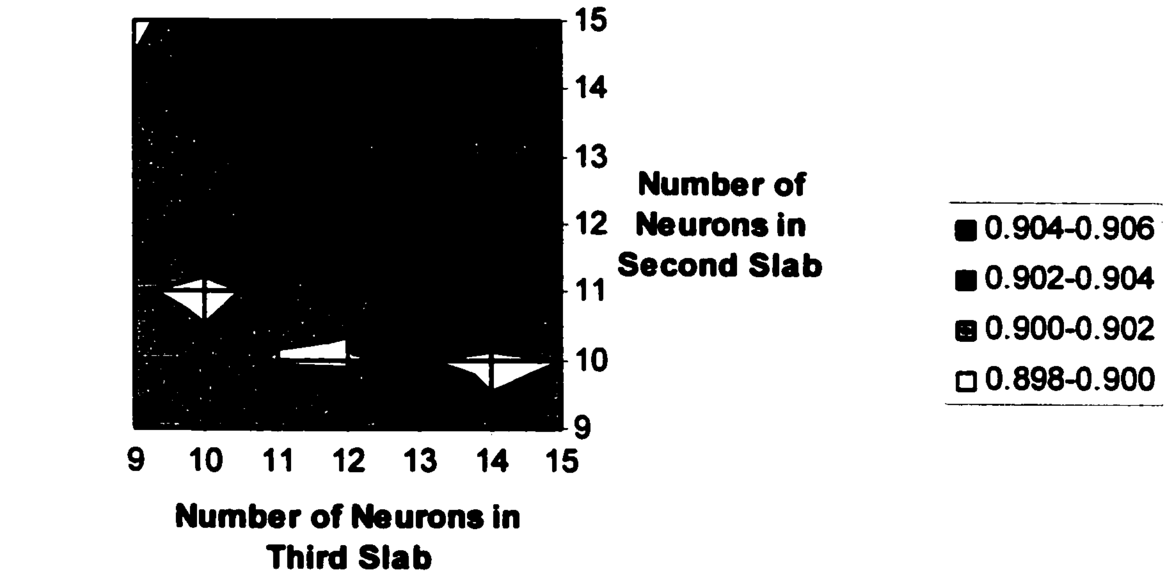

Figure 4.8 r^2 Surface Plot for Neuron Selection of Ward Net with 2 Hidden Slabs and a Jump Connection

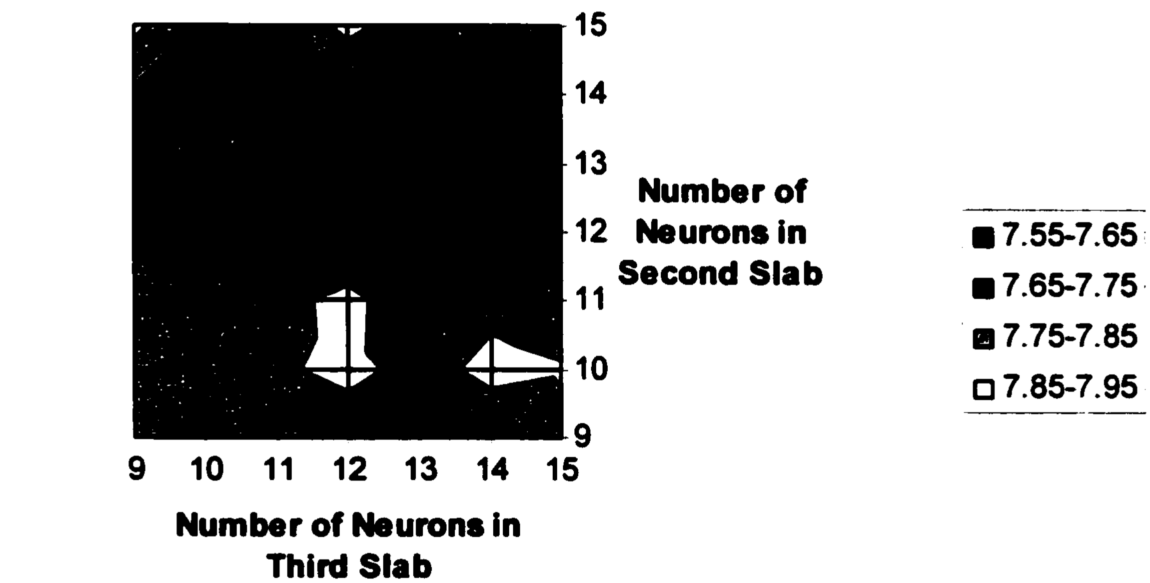

Figure 4.9 Mean Absolute Error Surface Plot for Neuron Selection of Ward Net with 2 Hidden Slabs and a Jump Connection

4.1.5 Learning Rate and Momentum Selection

The learning rate and momentum both determine the weight adjustment in the training process. To determine the optimum learning rate and momentum, a grid system was again used and the surface plot of this grid was used to graphically represent the results.

The initial grid was setup with the learning rate and momentum increasing from 0.1 to 0.9 in 0.2 increments. After the initial grid system was completed and analyzed, the area that showed the best results was then rerun using a 0.1 increment for both the learning rate and momentum terms. The resulting surface plots for each architecture can be found in Figure 4.10 to Figure 4.15. In Table 4.7 the optimal values for the learning rate and momentum for each architecture can be found along with their results.

Table 4.7 Optimum Learning Rate and Momentum

Architecture	Learning Rate	Momentum	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	0.4	0.2	0.9113	7.336
Ward Net 3 Hidden Slabs	0.1	0.1	0.9069	7.493
Ward Net 2 Hidden Slabs With Jump Connection	0.4	0.1	0.9062	7.576

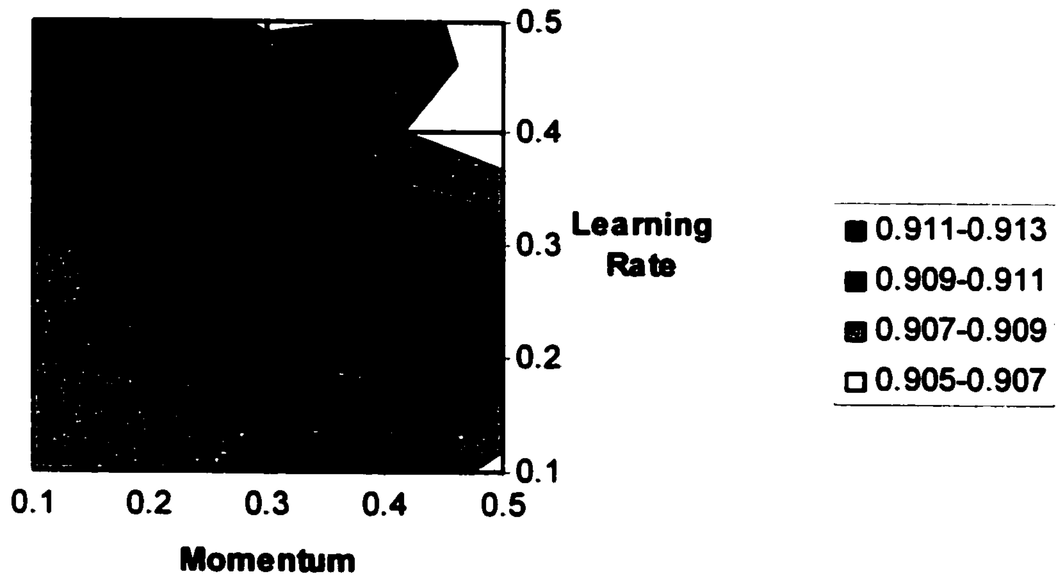

Figure 4.10 r^2 Surface Plot for Learning Rate and Momentum Selection of 4-Layer Backpropagation Network

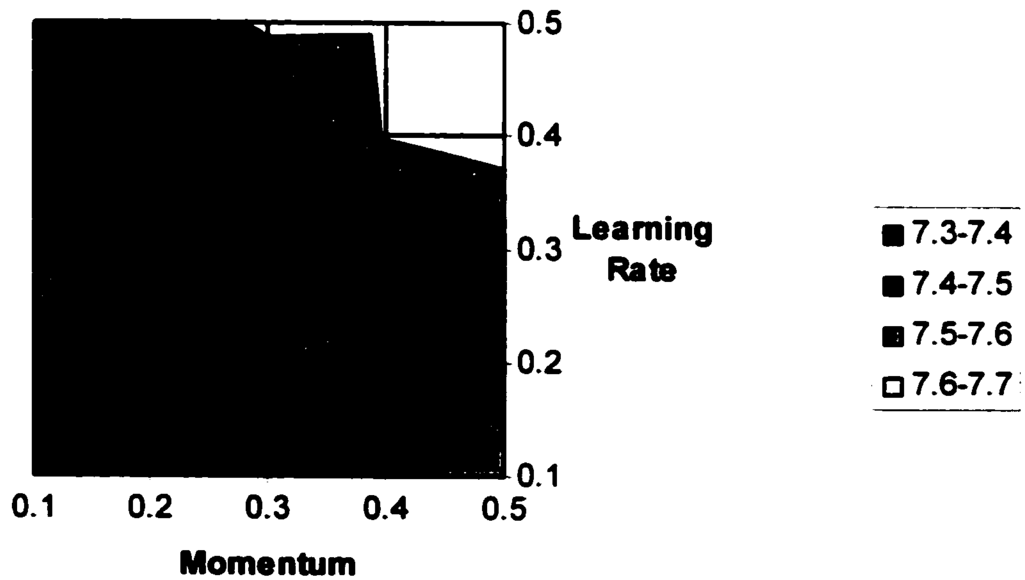

Figure 4.11 Mean Absolute Error Surface Plot for Learning Rate and Momentum Selection of 4-Layer Backpropagation Network

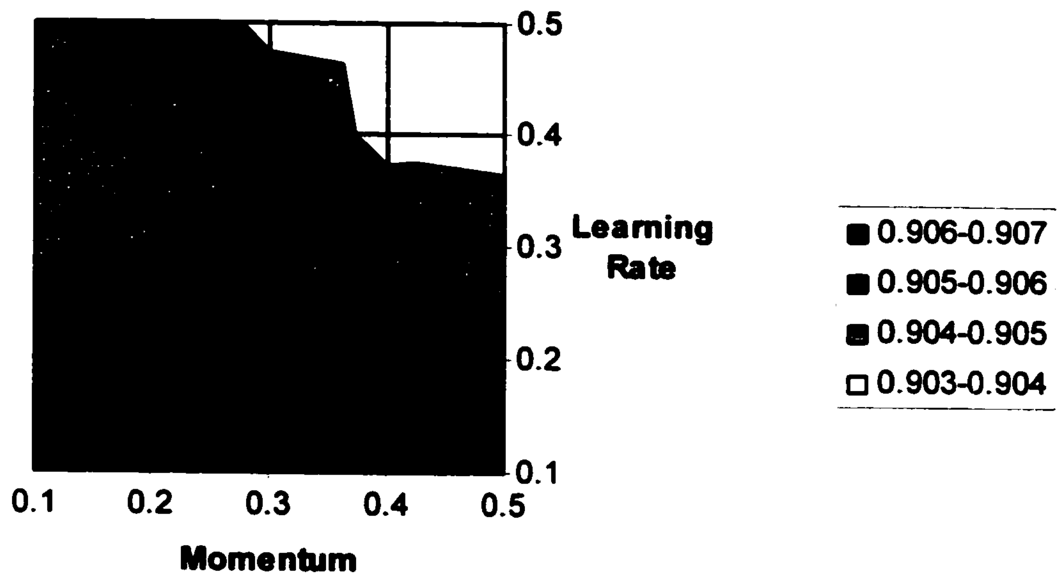

Figure 4.12 r^2 Surface Plot for Learning Rate and Momentum Selection of Ward Net with 3 Hidden Slabs

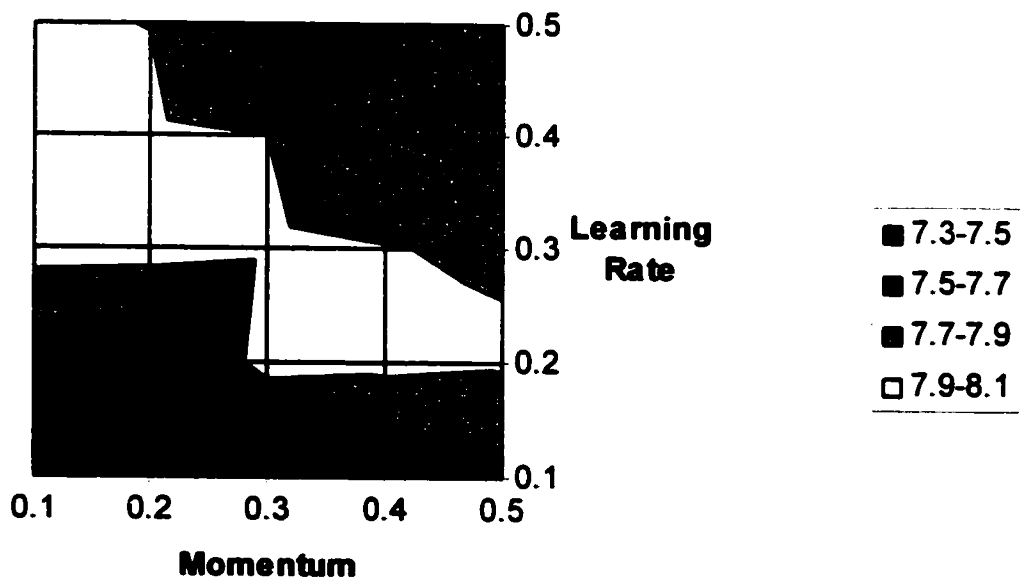

Figure 4.13 Mean Absolute Error Surface Plot for Learning Rate and Momentum Selection of Ward Net with 3 Hidden Slabs

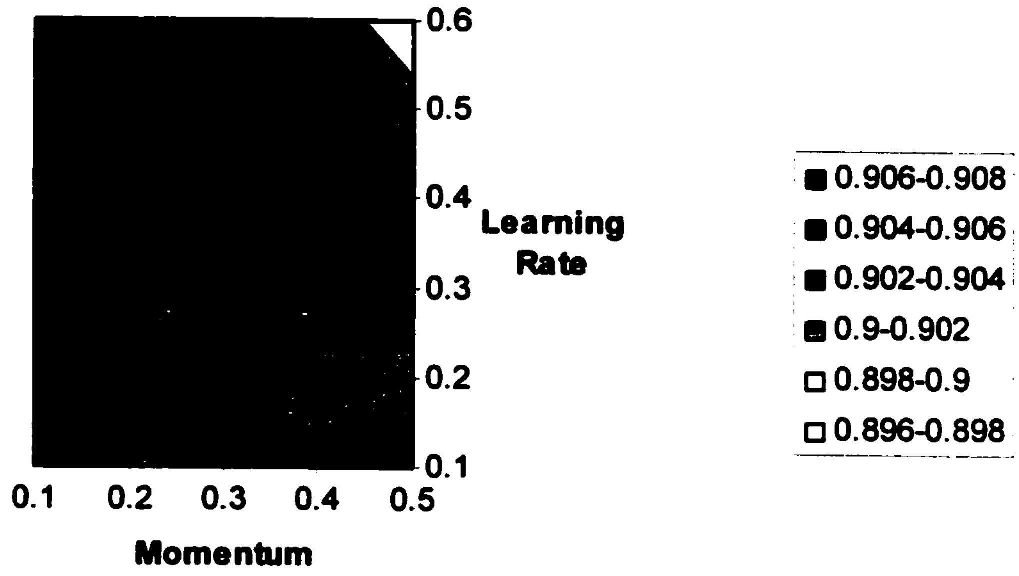

Figure 4.14 r^2 Surface Plot for Learning Rate and Momentum Selection of Ward Net with 2 Hidden Slabs and a Jump Connection

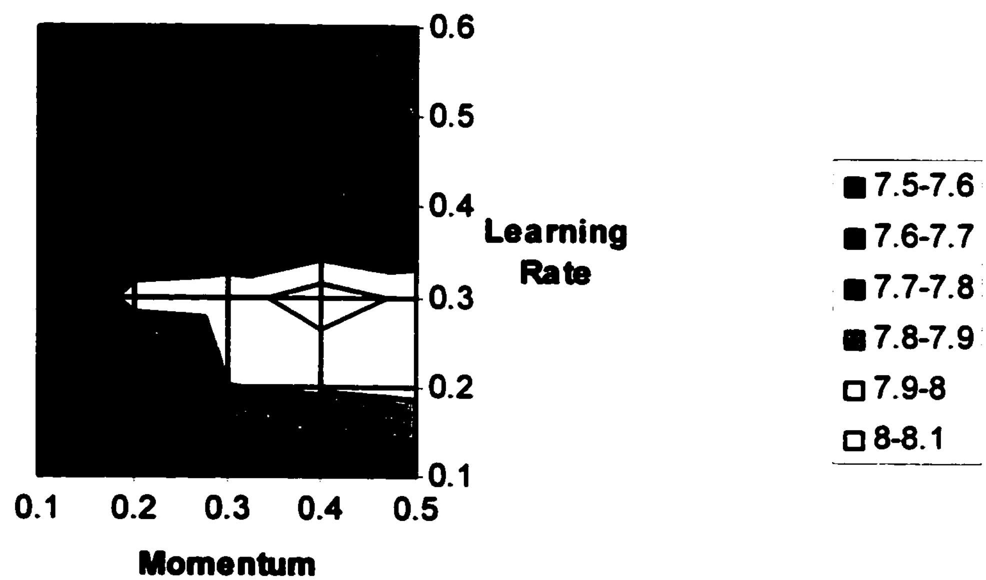

Figure 4.15 Mean Absolute Error Surface Plot for Learning Rate and Momentum Selection of Ward Net with 2 Hidden Slabs and a Jump Connection

4.1.6 Pattern Selection

In determining the type of pattern selection that best suits this problem each pattern selection was run and the pattern selection that yielded the best result was used. It can be seen that in Table 4.8 that random pattern selection produced the better result statistically for all three different architectures as the models are more robust. This further reinforces the idea that random pattern selection should be used in problems that are cyclical such as those that contain seasonal variations and/or complicated problems with numerical outputs.

Table 4.8 Pattern Selection

Architecture	Pattern Selection	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	random	0.9113	7.336
	rotational	0.8977	7.984
Ward Net 3 Hidden Slabs	random	0.9069	7.493
	rotational	0.8955	7.968
Ward Net 2 Hidden Slabs With Jump Connection	random	0.9062	7.576
	rotational	0.8722	8.877

4.1.7 Scaling Function

In selecting the scale function to use for each model, both linear and non-linear functions were tested. The linear functions out-performed the non-linear functions in their ability to correctly forecast the daily water demand (Table 4.9). The non-linear scaling functions squeeze together the high and low values of the original data range. This helps to minimize the effect of outliers, but also minimizes the importance of the extreme conditions. The non-linear functions in this case would minimize the importance of the extreme events, which are of the utmost importance in predicting daily water demand. This could be one possible explanation why the linear scaling functions achieve better results than the non-linear scaling functions. Within the linear scaling functions the data were scaled using <<-1,1>> and <<0,1>>. The <<>> brackets denote that any new data that is larger or smaller than the minimum and maximum values that the model was trained on, will not be clipped and the model will have to extrapolate. In the daily water demand problem, any new inputs into the model that would be outside of the range the model was trained on, would not be very far out of the range with the possible exception of the rainfall data. The temperature values in the data set had a minimum and maximum temperature of -28.5°C and 33.6°C respectively. Thus, the model will most likely have to only extrapolate a few degrees outside of what is was trained on. The best results were obtained using the linear <<-1,1>> scaling function.

Table 4.9 Scaling Function Selection

Architecture	Scaling Function	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	linear <<-1,1>>	0.9113	7.336
	linear <<0,1>>	0.9010	7.808
	logistic (0, 1)	0.8892	8.156
	tanh (-1, 1)	0.8670	8.385
Ward Net 3 Hidden Slabs	linear <<-1,1>>	0.9069	7.493
	linear <<0,1>>	0.8959	8.040
	logistic (0, 1)	0.8767	8.305
	tanh (-1, 1)	0.8510	8.924
Ward Net 2 Hidden Slabs With Jump Connection	linear <<-1,1>>	0.9062	7.576
	linear <<0,1>>	0.9003	7.875
	logistic (0, 1)	0.8793	8.297
	tanh (-1, 1)	0.8440	8.976

4.1.8 Final Daily Model Selection

In determining the final model for use in forecasting the daily water demand, each model was evaluated using its average prediction performance, performance on peak demand days and input to output generalization. Using these steps gives a more thorough analysis of each model than just using the average prediction performance to determine the models perform.

4.1.8.1 Average Prediction Performance

The average performance of the model quantifies how well the model predicts on average; this is what the model selection so far has been based on. The selection process also only considers the production data set, which is the independent data set that the model was not trained and tested on during the learning process. In Table 4.10, the models' performance on the training, testing and production sets of data, as well as on all the data combined from those three sets, is shown. A properly trained model will have similar errors on each data set. The main reason for this comparison is to ensure that the model has learned from the data, and not memorized them. If memorization of the data has occurred, the training set data and possibly the testing set will have a significantly higher r^2 and lower mean absolute error than the independent production set. As can be seen in Table 4.10, the models have similar statistics on the different data sets, with the production set values being slightly better than those of the training and testing sets. This indicates that the models have not memorized the data, and might indicate that the production set data contained fewer extreme or difficult to predict events than the other two data sets. The main points are that the models did not memorize the data sets and their performance on overall is good.

Table 4.10 Average Performance Indication

Architecture	Data Set	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	All Data	0.8911	8.201
	Training Set	0.8895	8.763
	Testing Set	0.8834	8.313
	Production Set	0.9113	7.336
Ward Net 3 Hidden Slabs	All Data	0.8876	8.391
	Training Set	0.8863	8.907
	Testing Set	0.8786	8.599
	Production Set	0.9069	7.493
Ward Net 2 Hidden Slabs With Jump Connection	All Data	0.8891	8.441
	Training Set	0.8845	8.905
	Testing Set	0.8853	8.686
	Production Set	0.9062	7.576

4.1.8.2 Peak Demand Performance

The average performance of the models is an important measure to gauge their ability to predict daily water demand, but is not the only item to be considered. One major consideration is the model's ability to predict the water demand during peak periods, as this is of major importance to a water utility. It is also important to see how the errors in the model's predictions are distributed with respect to the level of water demand.

In examining the error in the daily water demand prediction, it can be seen that the error in prediction generally increases as the daily water demand increases (Figure 4.16 to Figure 4.18). It can be seen in Figure 4.16 to Figure 4.18 that the error in prediction is smaller in the 290 – 320 ML/day range. This can be explained by the water demand in the winter months (which fall into this range), being generally not as variable as during the summer months, during which higher and more variable demands occur. With the water demand having less variability in this period, there are more data available for training and testing the model in this range. In conjunction with having more data in this particular water demand range because of the low variability in water demand during the October-April period, the lower variability also leads to improved learning for that specific range of data and smaller errors, as the water demand does not change drastically. Conversely, the error associated with the higher water demand tends to be larger. This is because the data for the model's training, testing and verification are spread out over a larger range. The one major fault of the 4-layer backpropagation and Ward net with 3 hidden slabs models is that their residuals have an apparent trend present up to approximately the 320 ML/day water demand. This trend is not present for demands greater than 320 ML/day. This trend indicates that the models may have some inadequacies. The most likely cause is that the inputs being used in the models are either too few or that an extra input is needed for lower water demands, where the trend is present. It was investigated whether splitting up the data into two sets using the season index and developing two separate models, one for each season would lead to an improved model. The best combination of the models led to an increase in the average water demand error by 0.96 ML/day or 0.3% and a r^2 that was reduced by 0.04. Thus, it did not improve the predictive ability by having two separate data sets based on the season instead of a season index, and did not remove the trend that was present from 290 ML/day to the 320 ML/day.

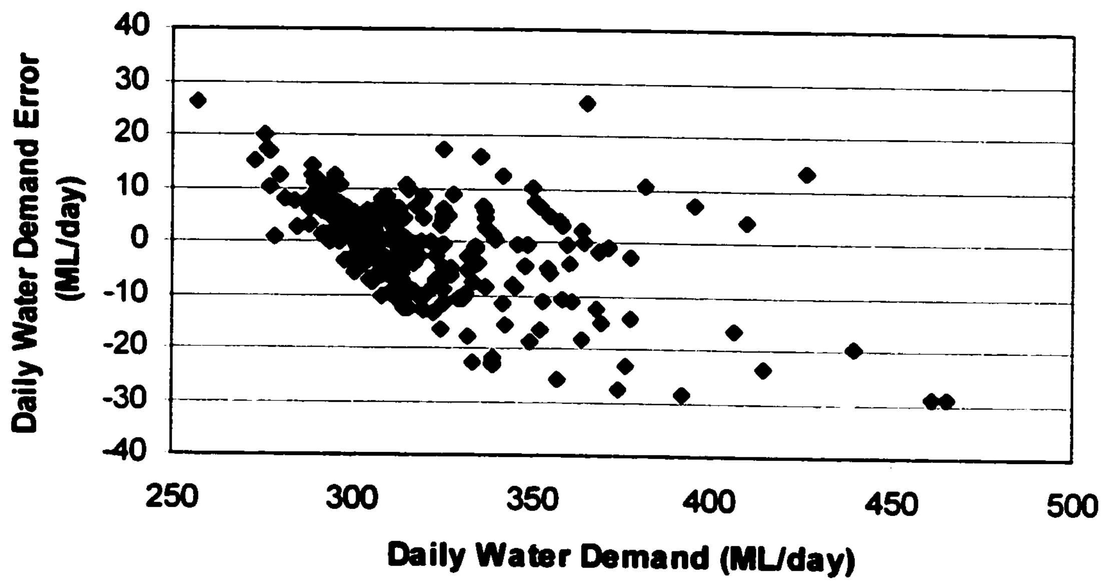

Figure 4.16 Daily Water Demand Error vs. Daily Water Demand for the 4-Layer Backpropagation Network Model (Production Set Data)

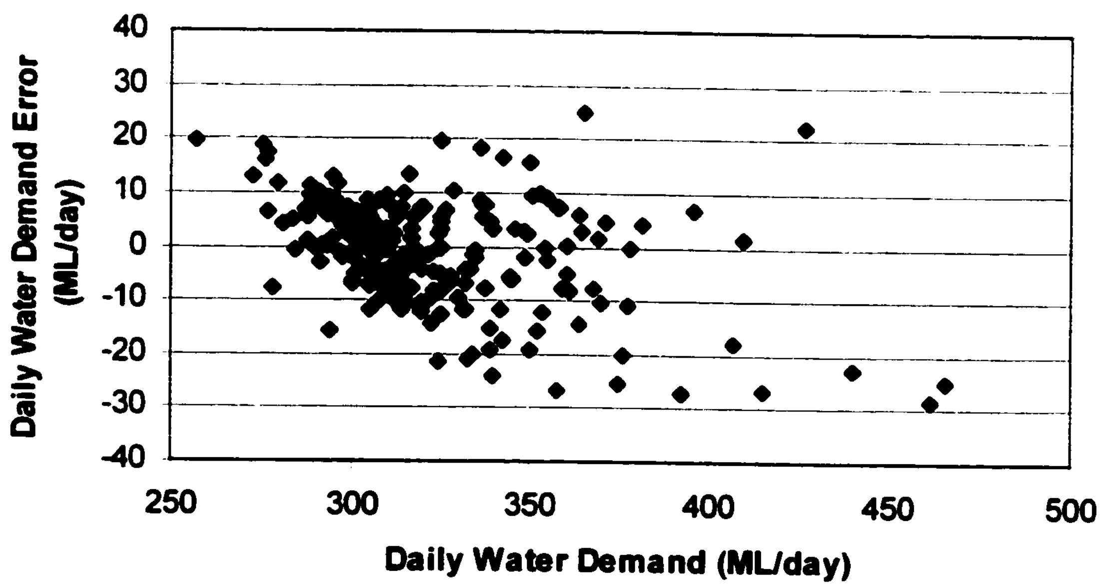

Figure 4.17 Daily Water Demand Error vs. Daily Water Demand for the Ward Net with 3 Hidden Slabs Model (Production Set Data)

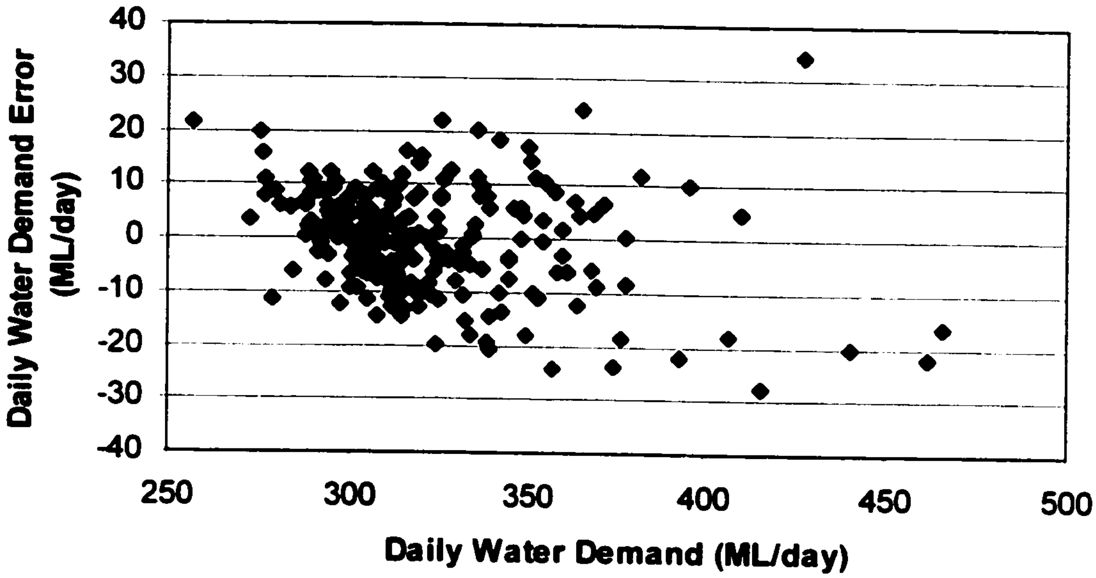

Figure 4.18 Daily Water Demand Error Vs Daily Water Demand for the Ward Net with 2 Hidden Slabs and a Jump Connection Model (Production Set Data)

In choosing a water demand model, an important component is the ability for the model to predict peak water demands. Peak water demands are of importance as this is the period where a high water demand can actually exceed the water production of a water utility. By forecasting the high demand in advance, water can be stockpiled in the reservoirs to offset any deficits that may occur. Also, with the foresight of a higher demand, the utility is able to produce and pump the extra forecasted water demand at times of lower power costs. It is also of importance to have a longer term forecast for this case as well; this will be discussed later. In analyzing the results of the final three models, the residual plots of all the data were used. The residual plots used were the percent error (i.e. [predicted-actual]/actual×100) vs. the actual water demand (Figure 4.19 to Figure 4.21). It can be seen that at all times each model is able to predict within 12% of the actual water demands. The distribution of the percent error can be found in Table 4.11. The general breakdown of the models performances on average is approximately that 90% of the time the models are able to predict within 5% of the actual demand, 99% of the time within 10% of the actual water demand and at all times within 12% of the actual demand. These results refer to using these models on the data from the specific time frame in which the models were trained, tested and verified. This does not mean that these models will always be capable of predicting within 12%, as new extreme or unique events may arise, or changes in the water use may occur over time. This can lead to potentially larger errors in the prediction of the models. The latter of the two problems can be remedied by retraining the model once a year, so as to include the latest trends in water use. While nothing can be done in the extreme or unique cases that may give rise to larger prediction errors, these new cases should be used in the annual retraining, to expand the domain of the models. In cases where new inputs are outside of the domain the model was trained on, the prediction can be flagged, to warn operators that the prediction may contain larger errors as the inputs are outside or approaching the domain limits of the models.

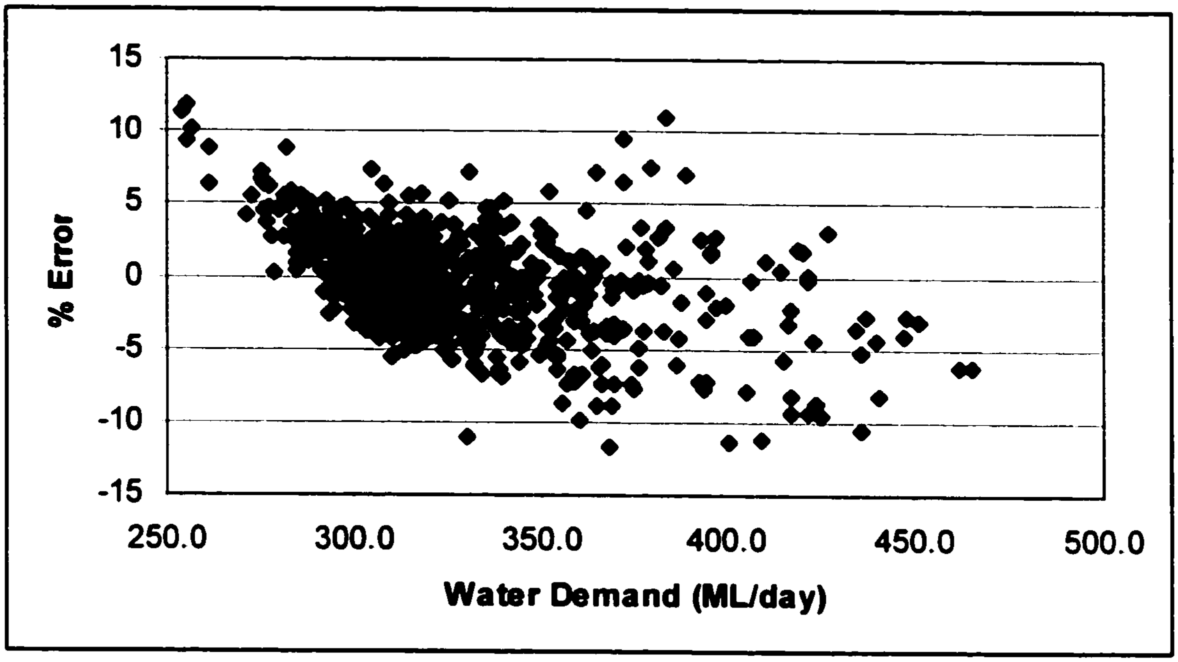

Figure 4.19 Daily Water Demand %Error vs. Daily Water Demand for the 4-Layer Backpropagation Network Model

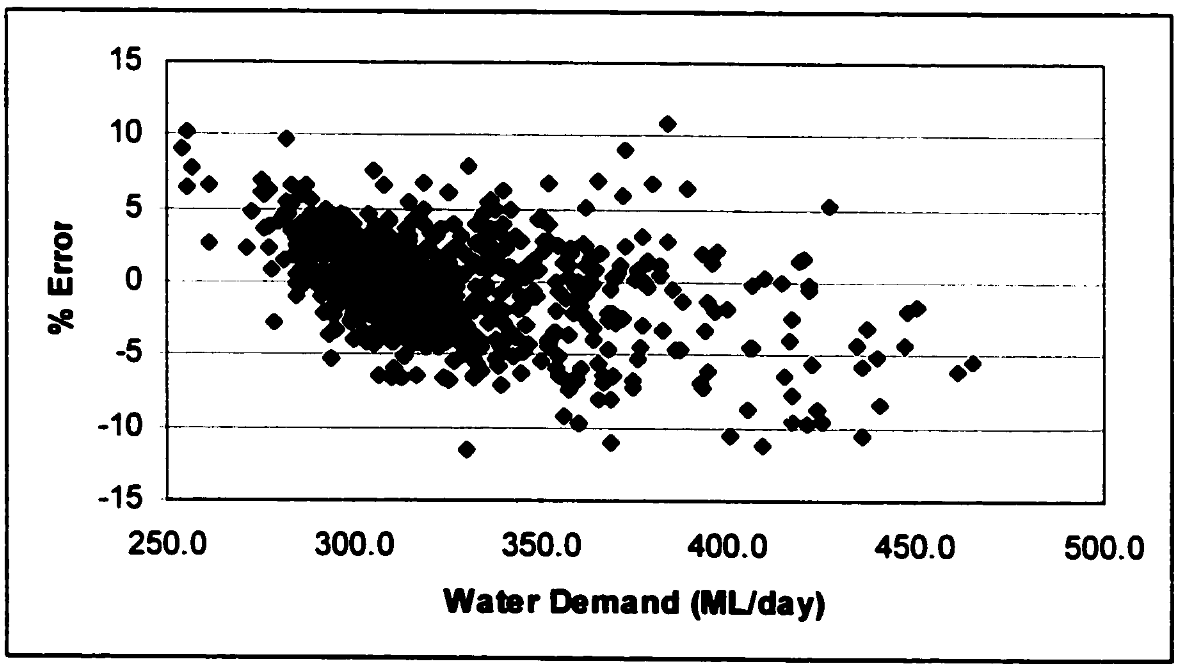

Figure 4.20 Daily Water Demand %Error vs. Daily Water Demand for the Ward Net with 3 Hidden Slabs Model

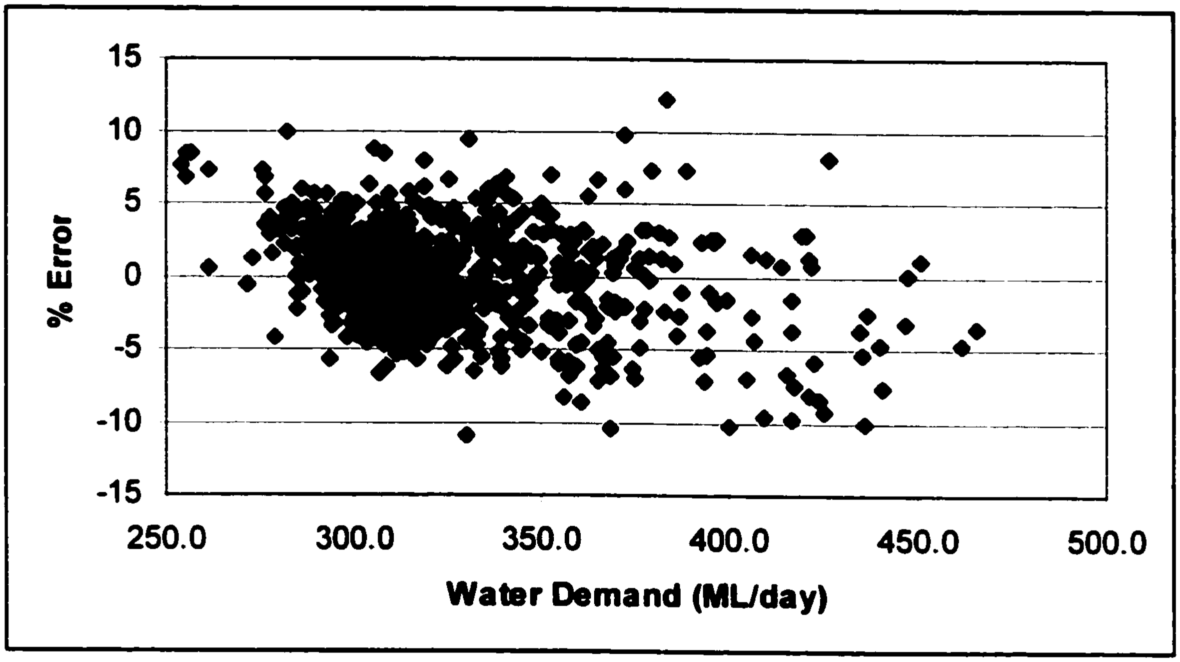

Figure 4.21 Daily Water Demand %Error vs. Daily Water Demand for the Ward Net with 2 Hidden Slabs and a Jump Connection Model

Table 4.11 Distribution of Percent Error for Daily Water Demand

Network	Data Set	Percentage error 0-5 %	5-10%	10-15%
4 Layer Backpropagation Network	Production	214 (90.7%)	21 (8.9%)	1 (0.4%)
	Pattern	700 (88.8%)	79 (10.0%)	9 (1.1%)
Ward Net with 3 Hidden Slabs	Production	212 (89.8%)	24 (10.2%)	0 (0.0%)
	Pattern	693 (87.9%)	88 (11.1%)	7 (0.9%)
Ward Net with 2 Hidden Slabs and a Jump Connection	Production	218 (92.4%)	18 (7.6%)	0 (0.0%)
	Pattern	698 (88.6%)	85 (10.8%)	5 (0.6%)

*may not add up to 100% due to rounding

4.1.8.3 Input to Output Generalization

When examining the models, one area of focus should be the input to output generalization. Studying the effect that each input has on the output is done to confirm that the model has learned the relationship between the input to the output as expected. For example, from the current literature with respect to water demand and from the utilities past experience, it is expected that the water demand should increase as the maximum temperature increases. When the relationship between the maximum temperature and the predicted daily water demand is plotted, there are two possible outcomes. First, the relationship between the daily water demand and the maximum daily temperature is as hypothesized. This then lends credibility to the past ideas on the relationship between the input and output. It also reinforces that the model has most likely learned the relationship between the input and output correctly (at least close to the correct generalization).

The second possibility is that the relationship between the input and output is not hypothesized based on past experience and the current literature on the subject. This then either indicates that the model has failed to learn the generalized relationship between the input and output or that the current literature and past experience may not be entirely correct.

4.1.8.3.1 Relationship of Temperature to Daily Water Demand

From the literature on the subject, the indication is that the daily water demand increases as the daily temperature increases (Shvartser, et al. 1993, Tamada, et al. 1993, Hall and Maidment 1990, Hittle, et al. 1996). The water utility also indicated that the relationship between the daily temperature and the daily water demand should follow this type of relationship with the temperature increasing in significance as the temperature increases. Thus an exponential relationship would be expected over the range of the data, as indicated when the actual daily water demand is plotted against the maximum daily temperature

(Figure 4.22). When examining the plots of the model relationship between the maximum daily temperature and the daily water demand found in Figure 4.22, it is found that the input-output relationship between the maximum daily temperature and the daily water demand is also an exponential relationship over the range of the data. The general model and actual relationship differ during periods of higher demand that occurs after 10°C. This occurs as the data for the model relationship were generated holding all the inputs constant with the exception of the minimum and maximum temperature. The minimum temperature was increased with the maximum temperature keeping a 10°C seperation between them. This was to more closely represent the real life situation that the models will be operating in. This was done to prevent the maximum temperature from being less than the minimum temperature, as this can not happen by the simple definition of minimum and maximum. It would also prevent unlikely scenerios where the minimum temperature would be -20°C and the maximum temperature would be 20°C, as this is unrealistic in Edmonton, Alberta. Thus, the relationship actually reflects the effect the minimum and maximum temperature has on the daily water demand. The other inputs that were used to generate the model output represent a summer weekday, no rain in the previous five days and a 10:00 am daily demand of 400 ML/day. In examining Figure 4.22, the model and actual daily water demand relationship with the maximum temperature, it should be noted that the other inputs affect the actual demand water demand. During periods of low water demand such as during the winter months there is little fluctuation in the actual water demand. During the summer period when water demand fluctuates and is higher that the generaliztion doesn't necessary hold true. This is because it is the combination of the individual inputs, as well as, their interactions that determine the daily water demand and are more pronounced during the summer months. This is where using ANNs are beneficial, as they learn the effect of each individual input, as well as, the interactions between the inputs.

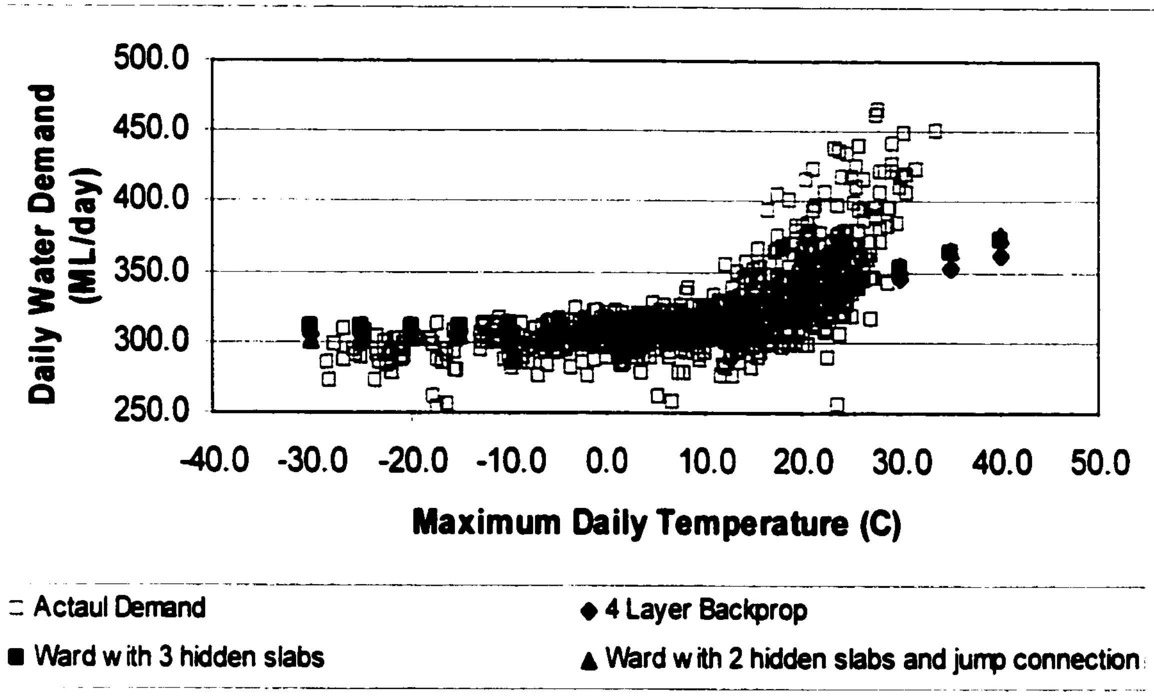

Figure 4.22 Actual and Model Relationship between Maximum Daily Temperature and Daily Water Demand

With holding the minimum and maximum temperature 10°C apart the effect that each individual temperature has on predicting the daily water demand is not apparent, only the overall effect that the temperature has. To differentiate between the effect that the minimum and maximum temperatures have on the daily water demand the minimum temperature was held constant at three different minimum temperatures while the maximum temperature was adjusted. It can be seen in Figure 4.23 that the lower the minimum temperature, the lower the daily water demand. In Figure 4.24 it can be seen that the greater the maximum temperature the greater the daily water demand.

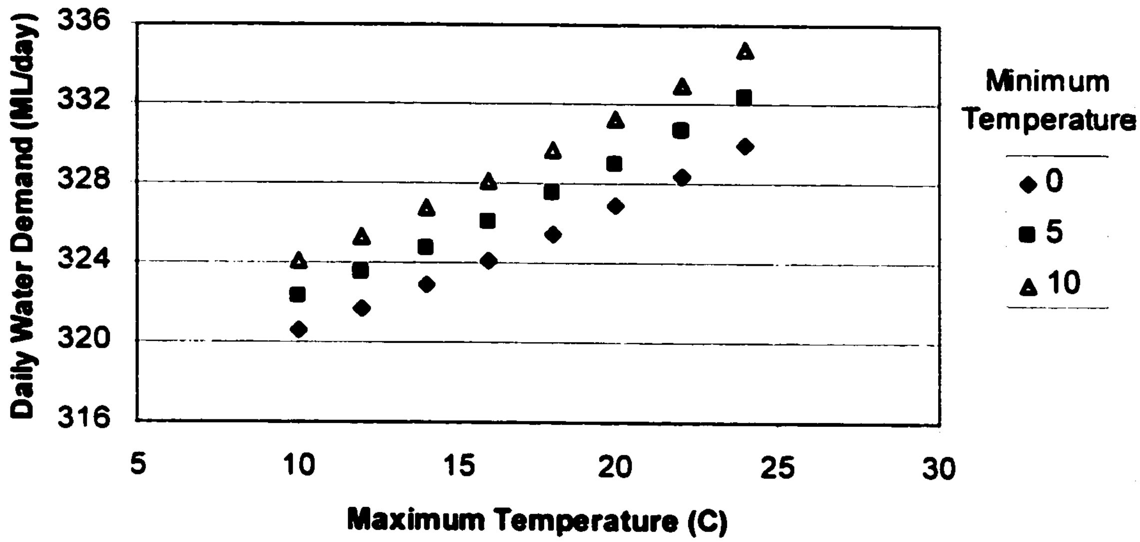

Figure 4.23 Minimum Temperature Effect With Respect to Maximum Temperature for Daily Water Demand Model

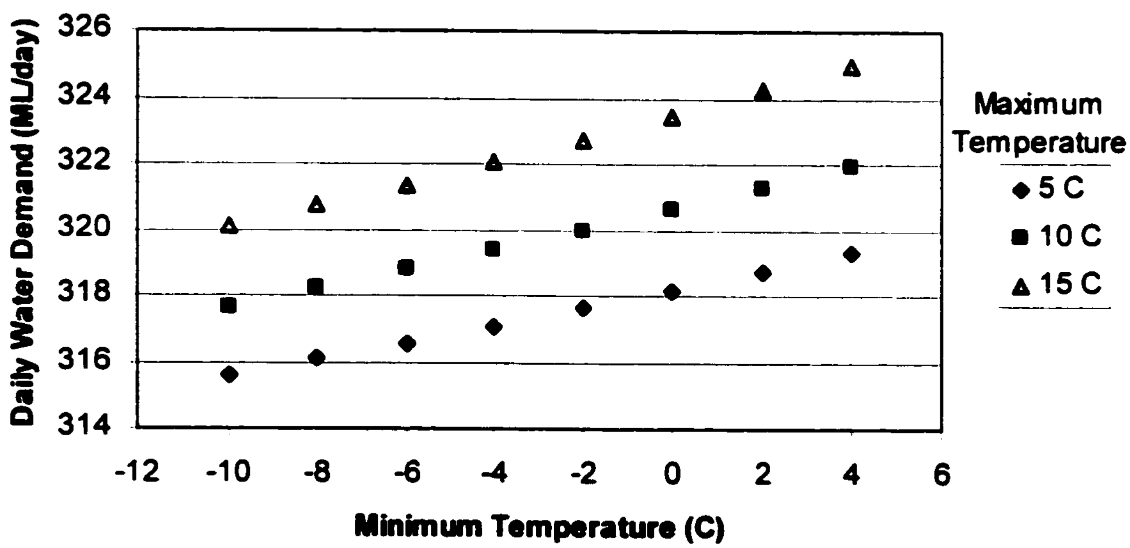

Figure 4.24 Maximum Temperature Effect With Respect to Minimum Temperature for Daily Water Demand Model

4.1.8.3.2 Relationship of Previous Day Rainfall to Daily Water Demand

The previous day rainfall input relationship with the daily water demand as indicated by recent literature on the subject indicated that the water demand should drop as the rainfall increases (Fleming 1994, Shvartser, et al. 1993, Tamada, et al. 1993, Fildes, et al. 1997). The water utility also indicated that this is the general case but depended on the other factors as well.

In Figure 4.25 it can be seen that there is a slight trend in that as the previous day rainfall increases the water demand decreases. It is also important to note in examining the data that only 3 of the data points that are less than 300 ML/day belong to the category of summer season and no previous day rainfall. Almost all of the remaining data that had a daily water demand less than 300 ML/day occurred in the winter season. The model generalization of the effect that the previous day's rainfall has on the daily water demand is found in Figure 4.26. It can be seen that the 4-layer backpropagation network and the Ward network with 3 hidden slabs produce similar generalization in regards to the daily water demand and previous day rainfall. Both models predicted a decrease in the daily water demand as the previous day rainfall increases. The daily water demand decreases, levels off, and then actually increases as the previous rainfall increases. This decrease and then the leveling off were expected. The slight increase after was not, and is due to the few data points that were available to train the models on. The Ward network with 2 hidden slabs and a jump connection model predicted that daily water demand decreases slightly and then increases as the previous day rainfall increases. Again, the initial decrease is as expected, the increase is not. The increase in daily water demand is contrary to the current literature, the water utilities past experience, and the other 2 models that were trained using two other types of architecture. The most probable explanation is that the Ward network with 2 hidden slabs and a jump connection incorrectly learned to generalize the relationship between the previous day rainfall and the daily water demand. The model generated data were obtained by inputting the other factors to represent a summer weekday with a minimum temperature of 10°C, a maximum temperature of 20°C, a 10:00 am daily demand of 420 ML/day. The rainfall in the previous 5 days was set equal to the previous day rainfall. This is simply saying that there was no rainfall in the previous 2-5 days.

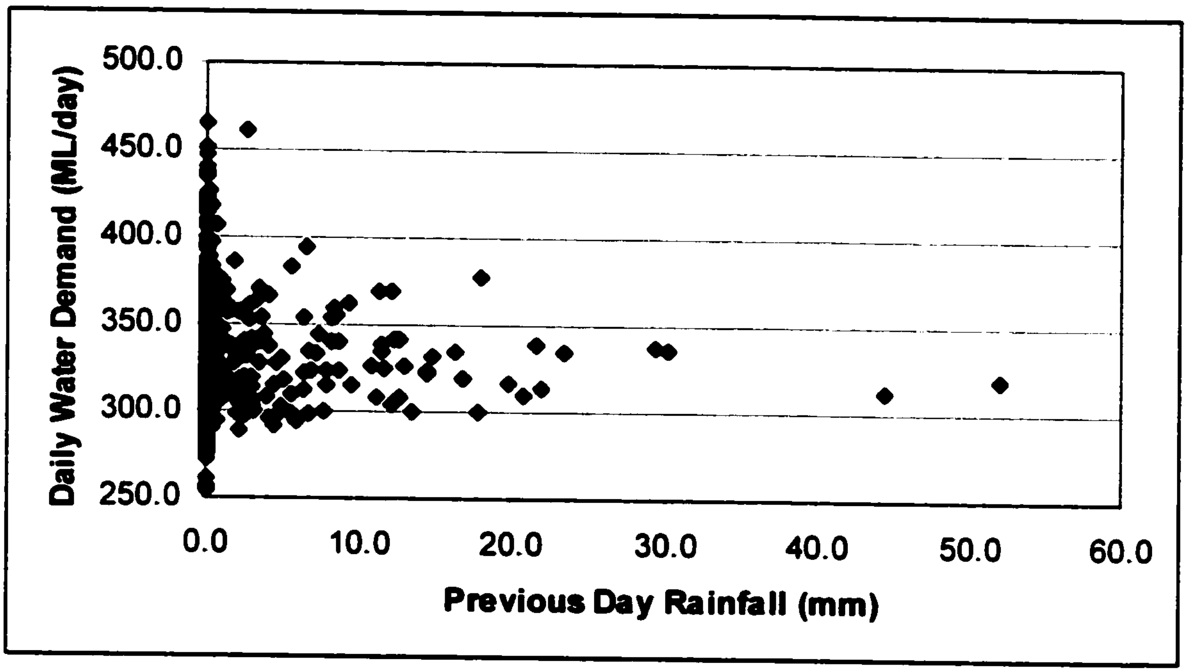

Figure 4.25 Actual Previous Day Rainfall vs. Daily Water Demand

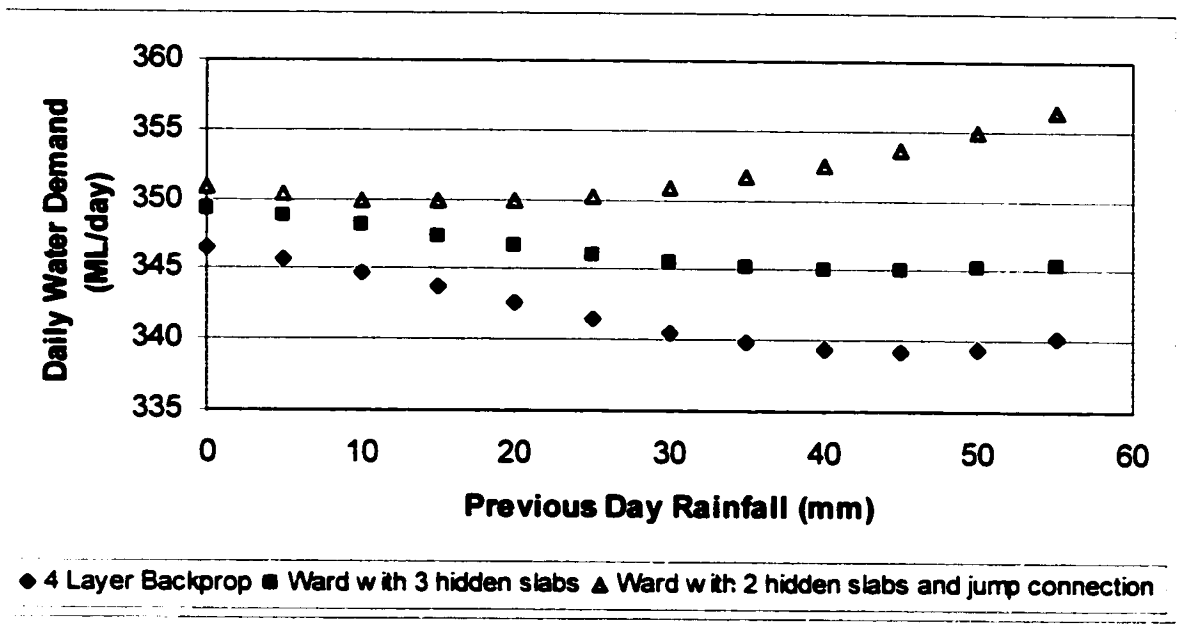

Figure 4.26 Model Generalization between Previous Day Rainfall and Daily Water Demand

4.1.8.3.3 Relationship of Previous 5-day Rainfall to Daily Water Demand

As mentioned in the previous section the previous rainfall is an important factor in the daily water demand. The importance of the previous rainfall is not only from the previous day, which acts as a short-term rainfall input, but also the importance of a longer-term rainfall. The previous 5-day rainfall is used for a mid to long-term rainfall input into the model as significant rainfall in this time period affects the daily water demand even if there is no rainfall the previous day. This is demonstrated in Figure 4.27. Similarly to the previous day rainfall, the cluster of data points that are around the 0-mm rainfall and under 300 ML/day are from the winter season. A 15-day, 30-day or 45-day longer-term rainfall input as discussed earlier was found to produce no improved performance. The general trend for the summer months is that the higher daily water demand occurs when there is little or no rain in the previous five days and decreases as the 5-day rainfall increases, as expected. In Figure

4.28 it can be seen that all three models developed have learned the generalization between the daily water demand similarly. The generalization is also similar to the relationship in Figure 4.27. It is important to remember that the model generated output data used for Figure 4.28 were produced by holding all the other inputs constant. The actual daily water demand in Figure 4.27 has all the inputs changing, as it is using the actual data for each day, so an exact fit is not expected, but the general trend is to be. The model generated data were obtained by inputting the other factors to represent a summer weekday with a minimum temperature of 10°C, a maximum temperature of 20°C, no rain in the previous day and a 10:00 am daily demand of 420 ML/day.

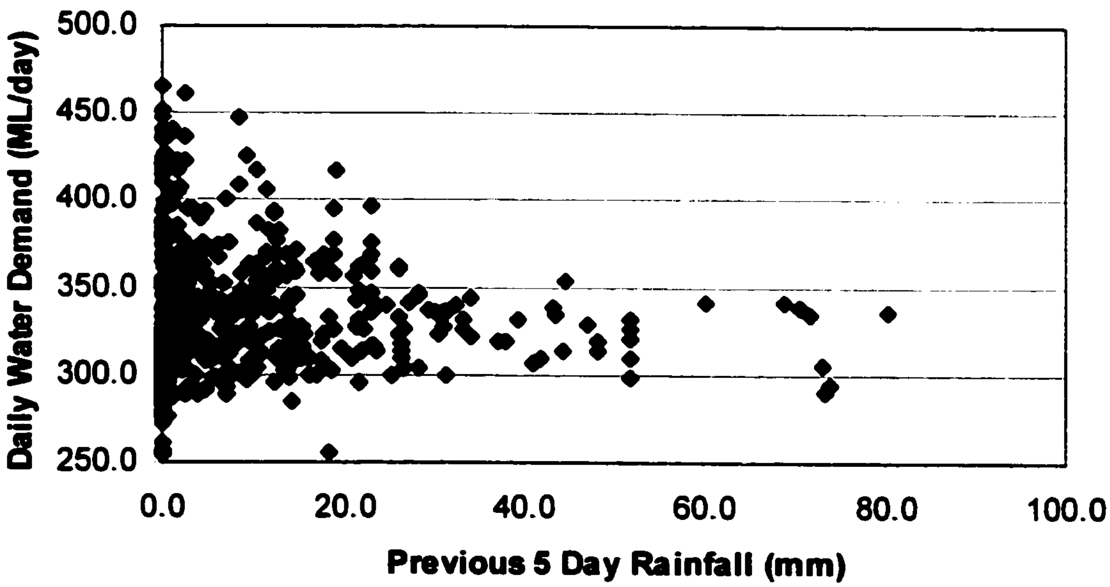

Figure 4.27 Actual Previous 5 Day Rainfall vs. Daily Water Demand

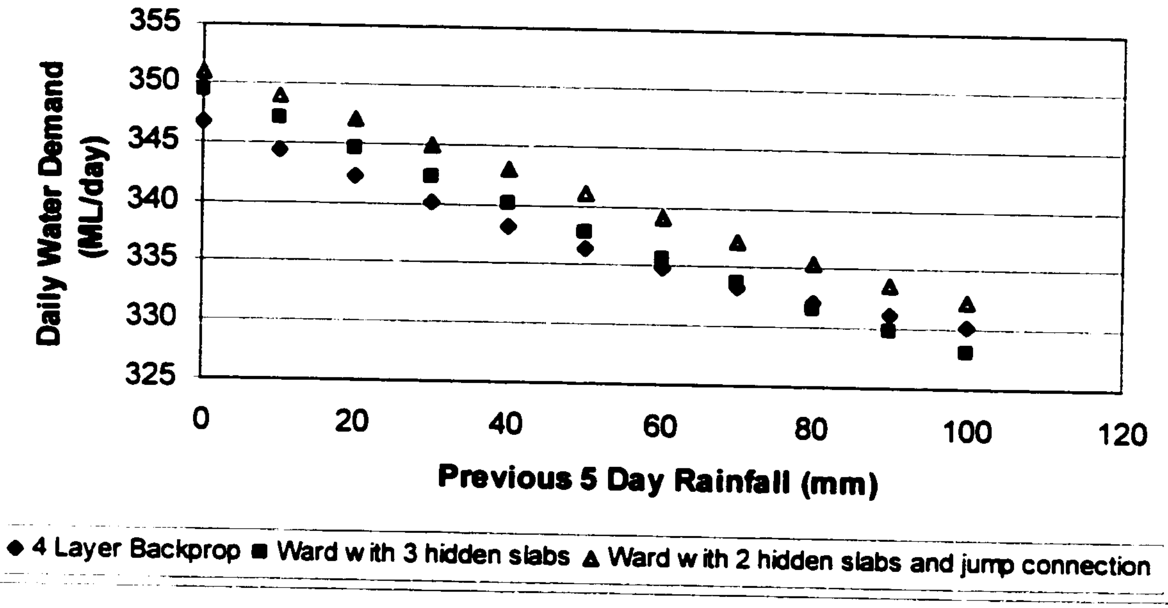

Figure 4.28 Model Generalization between Previous 5-Day Rainfall and Daily Water Demand

4.1.8.3.4 10:00 AM Water Demand to Daily Water Demand Relationship

The importance of using a reference indicator or past water use trends has been identified in predicting water demand (Ormsbee and Lansey 1994, Hall and Maidment 1990, Hittle, et al. 1996, Jain and Ormsbee 1993, Fildes, et al. 1997). The correlation between the 10 a.m. water demand and the daily water demand is expected in the sense that the 10 a.m. demand makes up part of the daily demand. In Figure 4.30, it is interesting to note that the relationship of the 10 a.m. water demand to the daily water demand is the most pronounced in the middle of the water demand range and tapers off at the high and low ends of the water demand. The leveling of the water demand coincides with the minimum and maximum water demand that the water utility experiences. This indicates that the ANN models have learned that even with new increasingly lower or higher 10 a.m. demands, the overall water demand has a minimum or maximum value that it will reach. This is to say that that there is a general minimum water requirement (approximately 225 ML/day) that the water utility customers require each and everyday. It also indicates that there is a maximum water demand (approximately 500 ML/day) that will generally not be exceeded, it will only approach this value. The minimum and maximum values will change as the number of customers, use of water-efficient appliances, and the climate change over time.

When comparing Figure 4.29 and Figure 4.30, it can be seen that the model trend is similar to the actual trend in the range of normal water demands. The only major difference is that of the slope. The actual plot of the data reveals an almost one to one correspondence between the 10:00 AM water demand and the daily water demand. The model generated data used to show the relationship that the model has learned between the 10:00 AM water demand and the daily water demand has a slope of less than one. The difference is due to the fact that the actual 10:00 AM demand is influenced by all the other inputs being used in the daily water demand as well. Thus the 10:00 AM water demand already has the major factors that affect the water demand built into it. In using only the 10:00 AM water demand to predict the daily water demand, it will give a rough prediction of the daily water demand. The model generated data in Figure 4.30 was generated by holding the minimum temperature at 10°C, maximum temperature at 20° and no rain in the previous 5-days on a summer weekeday. The resulting generalization between the 10:00 AM water demand and the daily water demand, does not give us the effect that the 10:00 AM demand has on the daily demand alone, even though all the other inputs are held constant. The 10:00 AM water demand is inherently linked to the other factors. By using the other factors in training the models, it allows the models to weight the importance of each of the other inputs independently, instead of grouping them together as when only the 10:00 AM water demand is used.

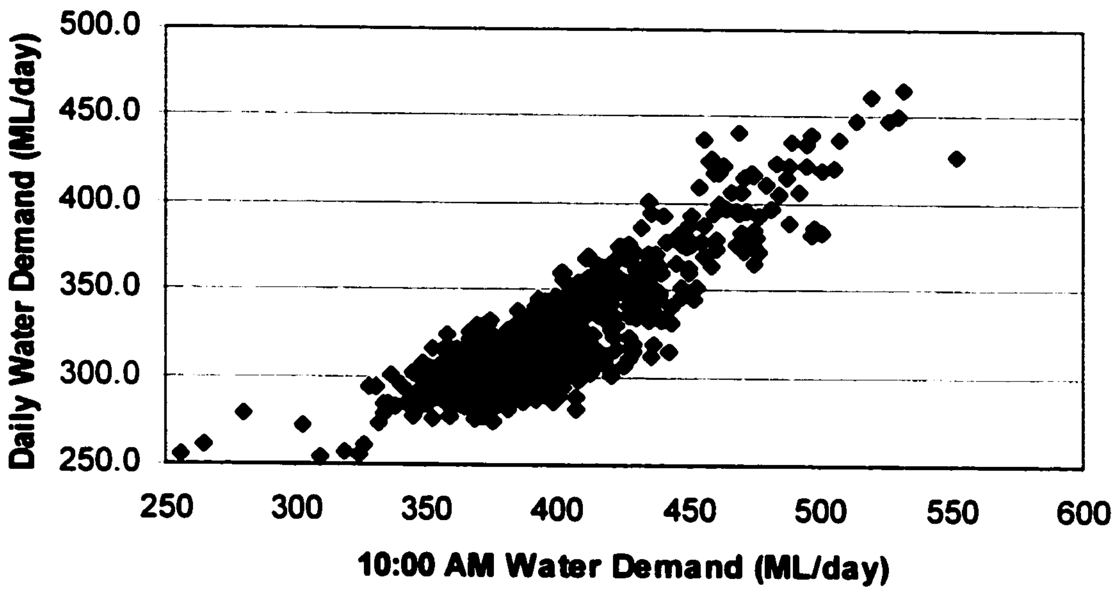

Figure 4.29 Actual 10:00 AM Water Demand Vs. Daily Water Demand

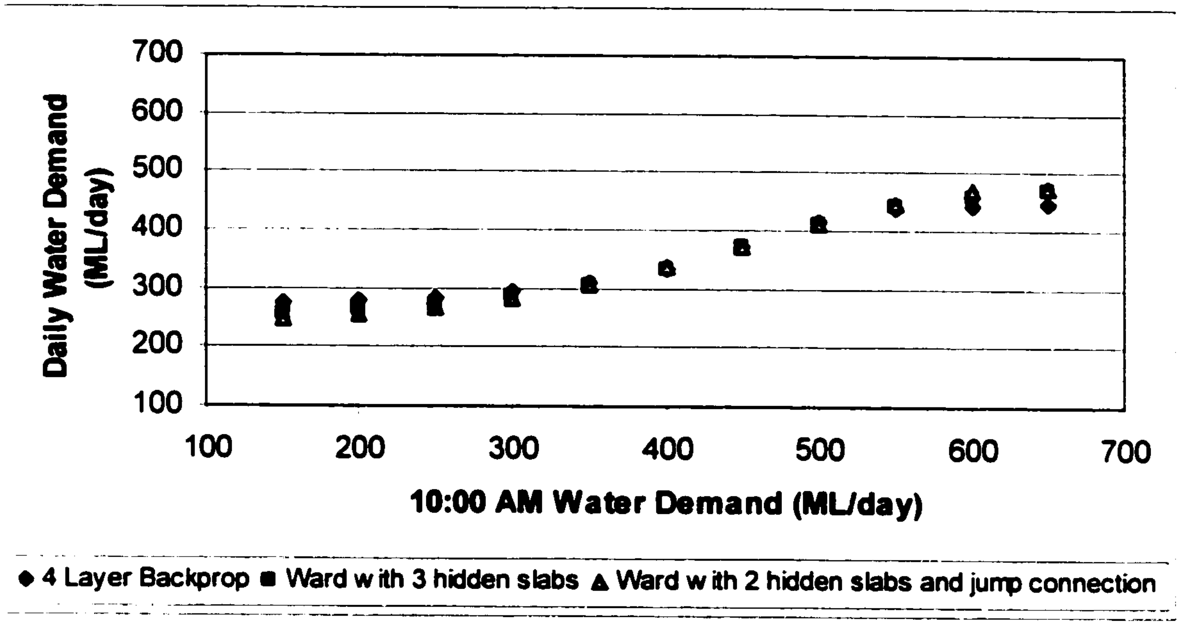

Figure 4.30 Model Generalization between 10:00 AM Water Demand and Daily Water Demand

4.1.8.3.5 Relationship of Weekend/Holiday and Weekday to Daily Water Demand

In forecasting water demand, one of the major contributing factor is human behavior. The activities of the customers have a major influence on the water demand. Thus, when examining the data and literature for the types of factors that affect the daily water demand, one factor that was found was the difference between weekend/holiday and weekday water demand use, both in quantity and the time of use throughout the day. The weekend and holiday water use is similar in quantity and time of use thus they were grouped together. It was found that, during the weekday period, the daily water demand was higher on average than it was for the weekend/holiday period as seen in Figure 4.31. The data from the model used to generate Figure 4.32, also show that the models predict that difference when presented with the weekend/holiday and the weekday scenarios.

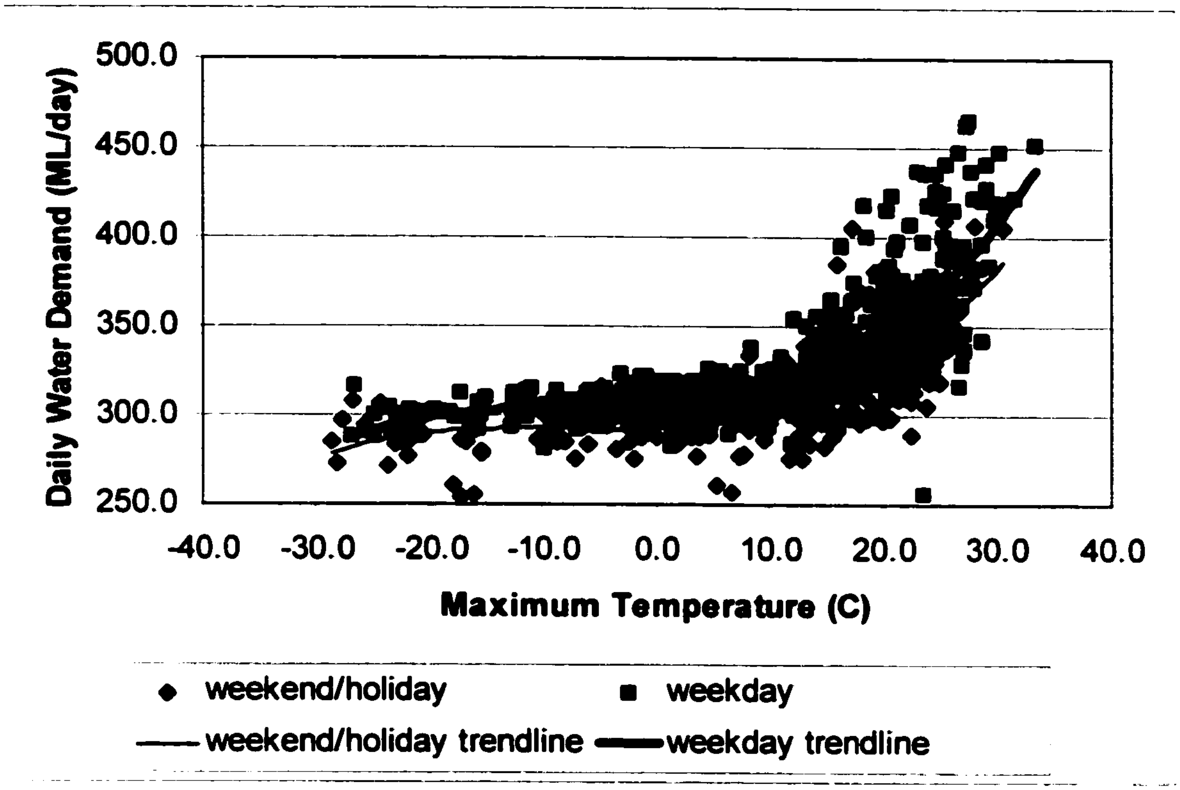

Figure 4.31 Actual Weekday and Weekend/Holiday Daily Water Demand

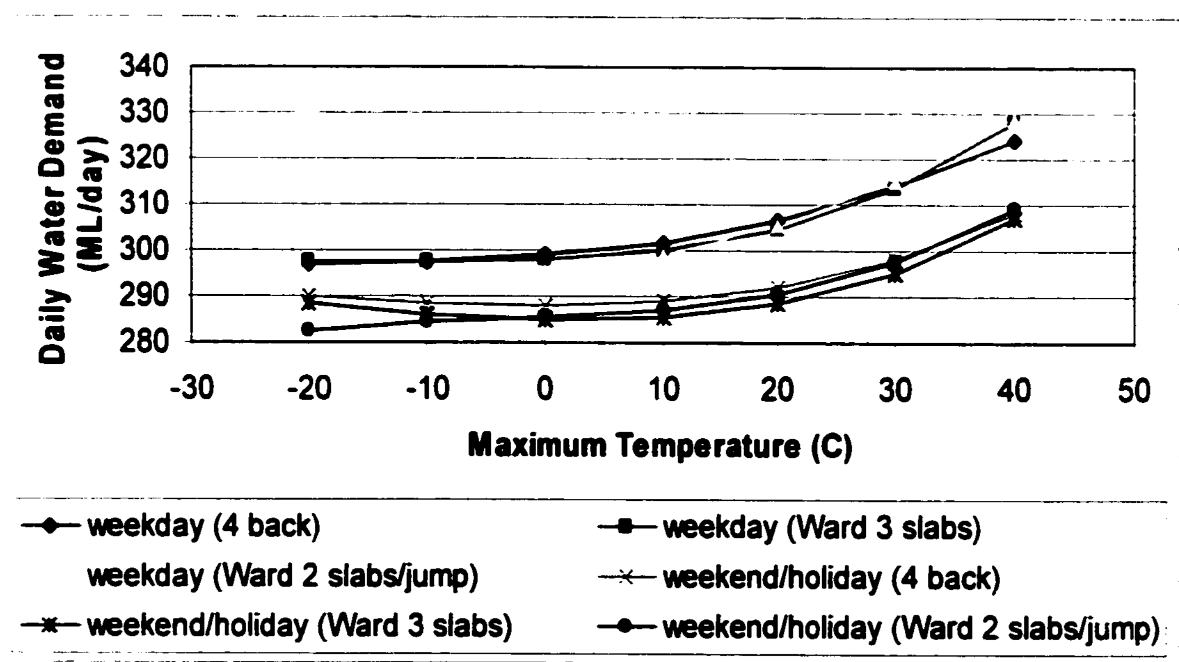

Figure 4.32 Comparison of Weekday and Weekend/Holiday Daily Water Demand

4.1.8.3.6 Summer and Winter Daily Water Demand Relationship

The activities of the water utility customers have a major influence on the water demand. With Edmonton, Alberta having a varying climate, were it is not uncommon to have temperatures drop below -30°C in the winter and to reach +30°C during the summer. This drastic change in temperature leads to different activities that the residents partake in, leading to a change in water usage. The ANN models indicate that the historical data presented to it have a trend that the water demand is slightly less during the October-April period than the May-September period at higher temperature, and little or no difference at lower temperatures as seen in Figure 4.34. This may be due to the fact that there is little overlap in the type of weather conditions that are similar between the two time periods. As during the October-April period there is generally little or no rainfall and cooler temperatures as compared to the May-September period. When comparing the data it has to be taken into account that the model has not been trained on data in that range and is extrapolating. As an example, for the May-September period no data were present that represented a maximum temperature of -15°C for the day thus the model is extrapolating outside of its domain to obtain the results displayed in Figure 4.34. The area of most interest is the data with the maximum temperature between 10°C and 20°C, as this is where the overlap of similar climate is going to occur between the two-season index. In between the maximum temperature of 10°C and the 20°C, it can be seen that the water demand is slightly higher for the May-September period. This is also apparent where the actual data are used in Figure 4.33, as the trend lines start to diverge between the 10°C and 20°C maximum temperature.

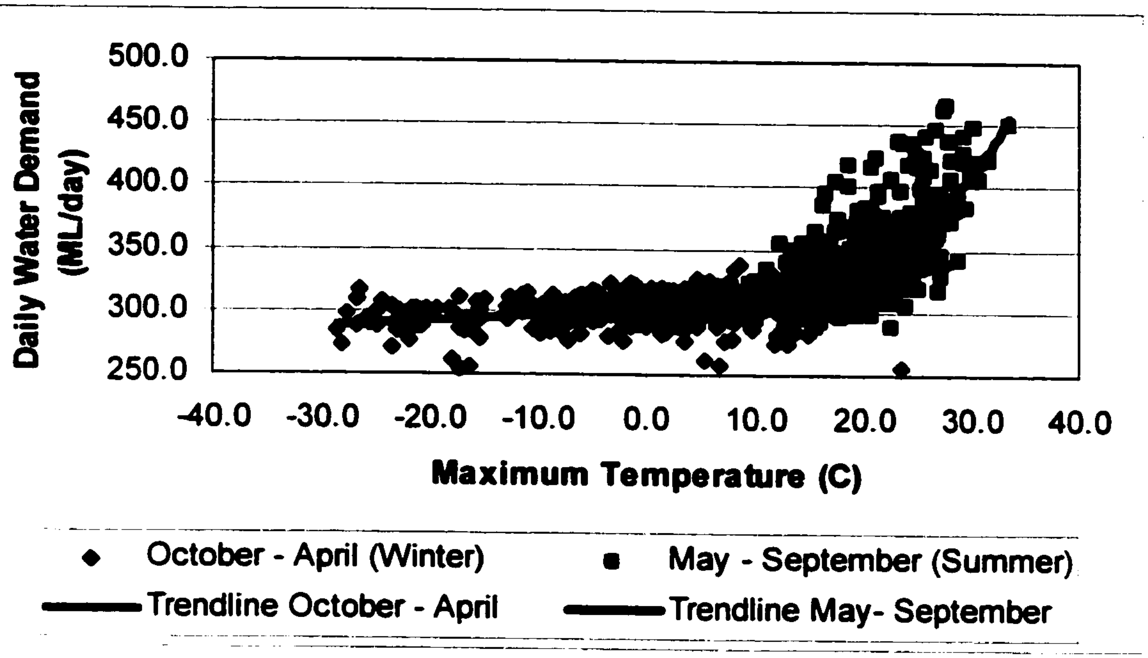

Figure 4.33 Actual May-September and October-April Daily Water Demand

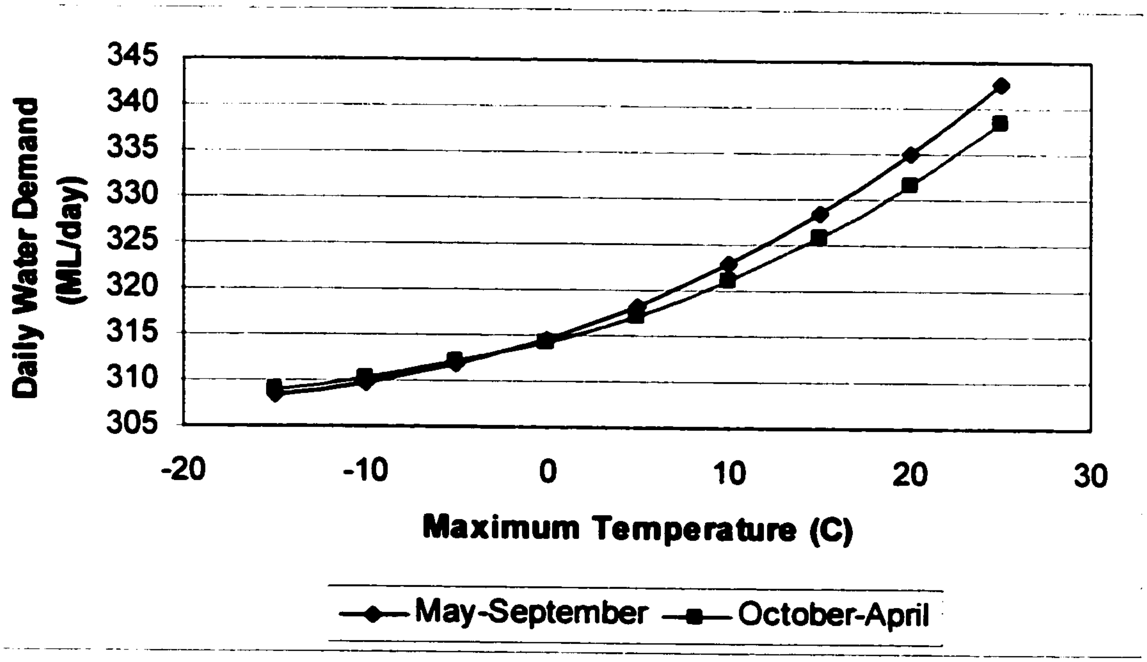

Figure 4.34 Comparison of May-September and October-April Daily Water Demand

4.1.9 Final Daily Water Demand Model

The final daily water model selected was a combination of the 4-layer backpropagation and the Ward net with 3 hidden slabs. The Ward net with 2 hidden slabs and a jump connection was excluded because the model input to output generalization for the previous day rainfall was contrary to what was expected from the literature and from the trend of the actual data. The combined model is simply the average of the two model predictions. The combined model has an r^2 of 0.911, mean absolute error of 7.35 ML/day (2.27% error), maximum absolute error of 28.53 ML/day (8.81% error), predicted within 5% of the actual water demand 90.7% of the time and within 10% of the actual water demand 100% of the time on the production set data. The combined model also follows the general trend of the data and predicts the peak demand as can be seen in Figure 4.35.

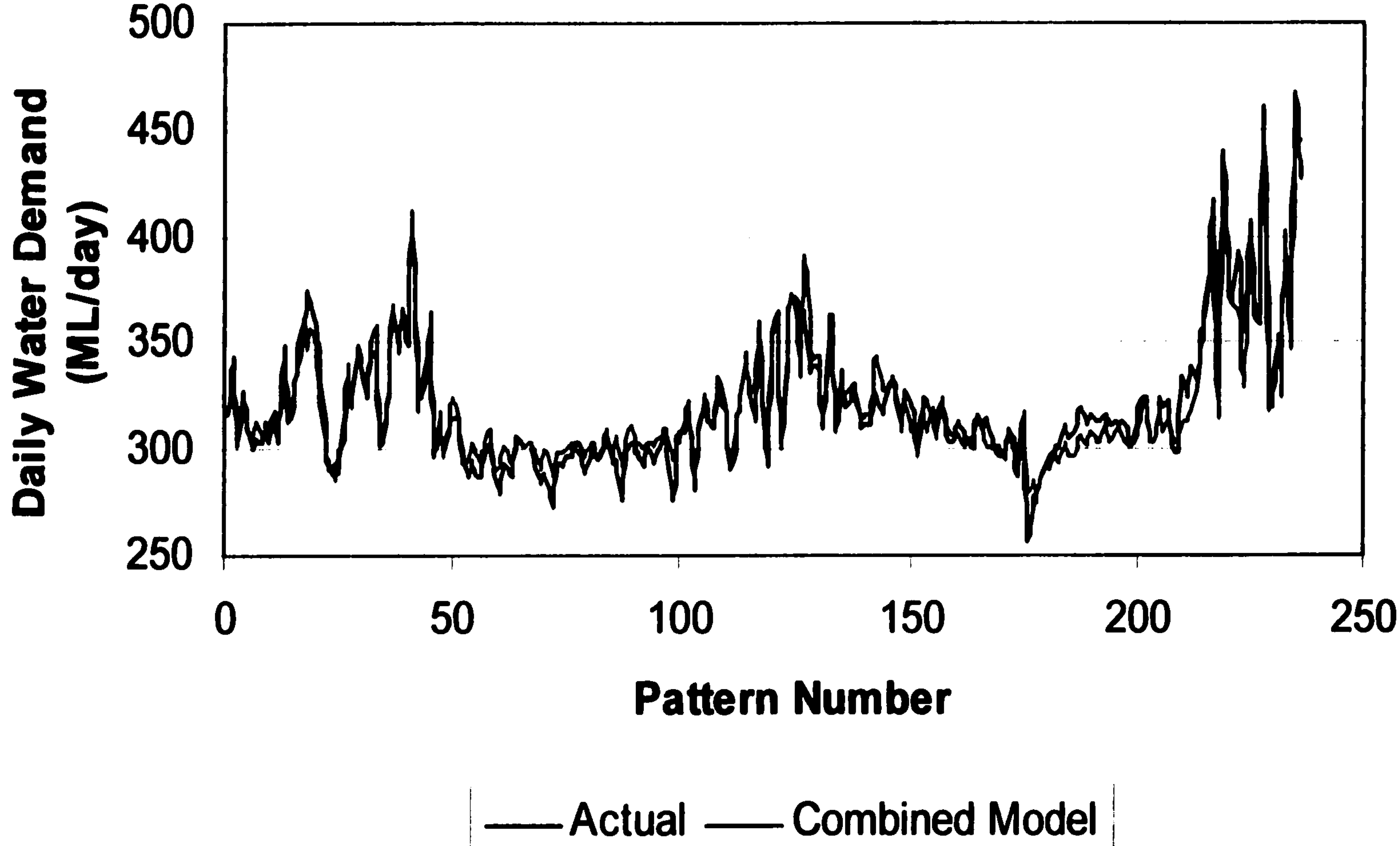

Figure 4.35 Actual vs. Predicted Daily Water Demand (Production Set)

4.2 Twelve Day Water Demand Forecast Model

The 12-day water demand forecast model is used to predict the water demand for each day from day 2 through day 12. It has similar inputs to the daily water demand with the exception that the previous day water demand is used as an input instead of the 10 A.M. water demand. Both of the aforementioned inputs are similar to each other, as they are reference indicators used in the model forecasts. With almost all the model inputs and data the same, as well as the output, with the one exception, the problem is virtually identical to the daily water demand forecast. Thus the same architecture, number of neurons, activation functions, etc. were used as the daily water demand for the 2-12 day model.

4.2.1 Final 2-12 Day Model Selection.

The 2-12 day model results are based on using the actual inputs for the data even though for real-time use, some of the inputs are obtained from forecasts. Thus on January 1st, the January 12th meteorological data were used as the inputs to predict the day 12 water demand. In real-time operation this would not be the case, as only the weather forecasts or 30-year average values will be available.

4.2.1.1 Average Prediction Performance 2-12 Day Model

As can be seen in Table 4.12 that the models have similar results on the different data sets, with the production set r^2 values only slightly lower than those of the training and testing sets. The different data sets having similar r^2 values indicate that the models have not memorized the data. The main points are that the models did not memorize the data sets and their performance on average are good with a mean absolute error of roughly 10

ML/day (3% error). With the water utility experiencing water demands that averaged 331 ML/day, the ANN models produced an average error of less than 3%, which is within the error of the flow meters.

Table 4.12 Average Performance Indication

Architecture	Data Set	r^2	Mean Absolute Error (ML/day)
4 Layer Backpropagation	All Data	0.8330	9.86
	Training Set	0.8429	9.48
	Testing Set	0.8384	10.20
	Production Set	0.8112	10.01
Ward Net 3 Hidden Slabs	All Data	0.8315	9.69
	Training Set	0.8340	9.49
	Testing Set	0.8467	9.91
	Production Set	0.8091	9.73
Ward Net 2 Hidden Slabs With Jump Connection	All Data	0.8547	9.23
	Training Set	0.8872	8.32
	Testing Set	0.8421	10.01
	Production Set	0.8188	9.66

4.2.1.2 Peak Demand Performance

In analyzing the results of the final three models, the residual plots of all the data were used. The residual plots used were the % error (i.e. [predicted-actual]/actual×100) vs. the actual water demand (Figure 4.36 to Figure 4.38). It can be seen that at all times each model is able to predict within 17% of the actual water demands. The distribution of the % error can be found in Table 4.13. The general breakdown of the models performances on average are approximately that 83% of the time the models are able to predict within 5% of the actual demand, 97% of the time within 10% of the actual water demand, 99.9% of the time within 15% of the actual water demand and at all times within 17% of the actual demand. That is when using these models on the data from the specific time frame in which the models were trained, tested and verified on. As discussed in the Daily Water Demand results, this does not mean that these models will always be capable of predicting within 17%. Thus, the models should be retrained once a year, to include the latest trends and any new extreme conditions in the water demand. In the same way as the daily water demand predictions are to be flagged when new inputs are outside of the domain that the model was trained on, the same should be done for the 2-12 day forecast models.

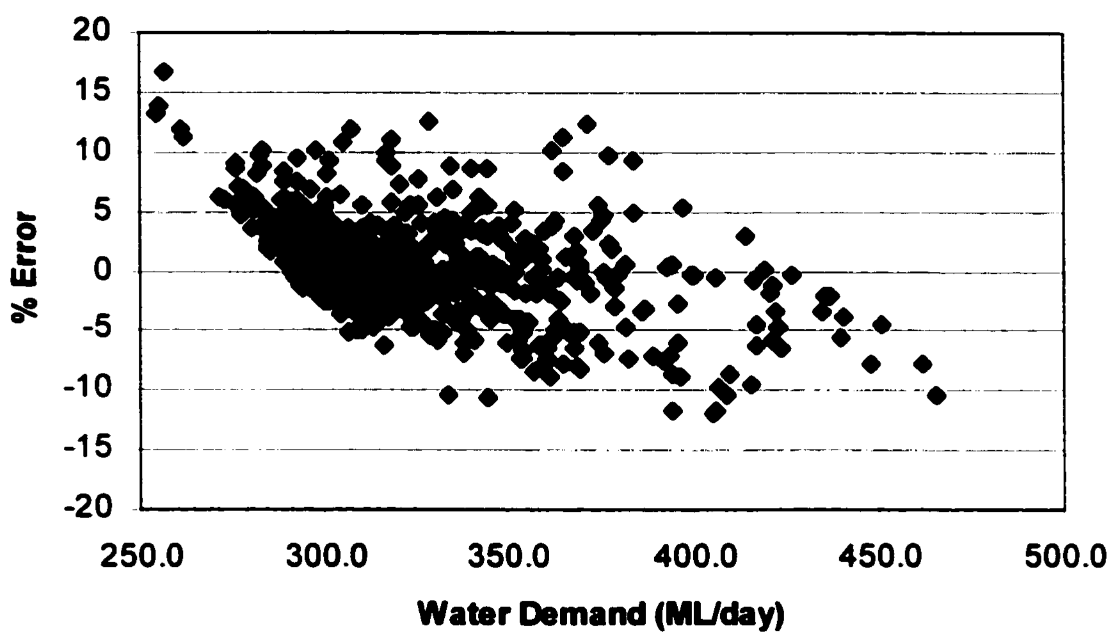

Figure 4.36 Daily Water Demand %Error vs. Daily Water Demand for the 4-Layer Backpropagation Network Model for 2-12 Day Model

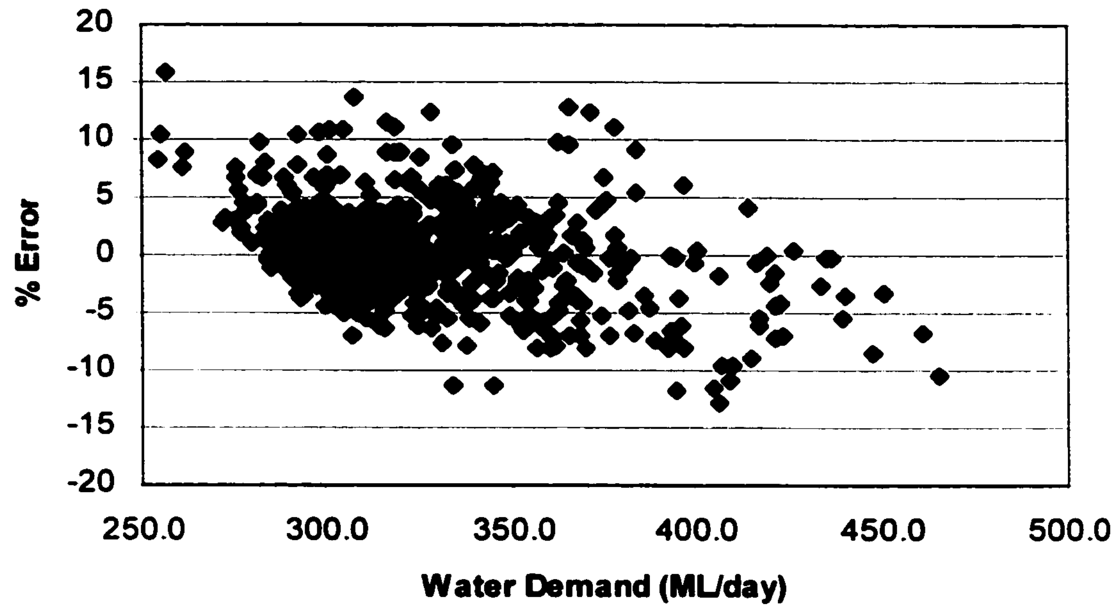

Figure 4.37 Daily Water Demand %Error vs. Daily Water Demand for the Ward Net with 3 Hidden Slabs Model for 2-12 Day Model

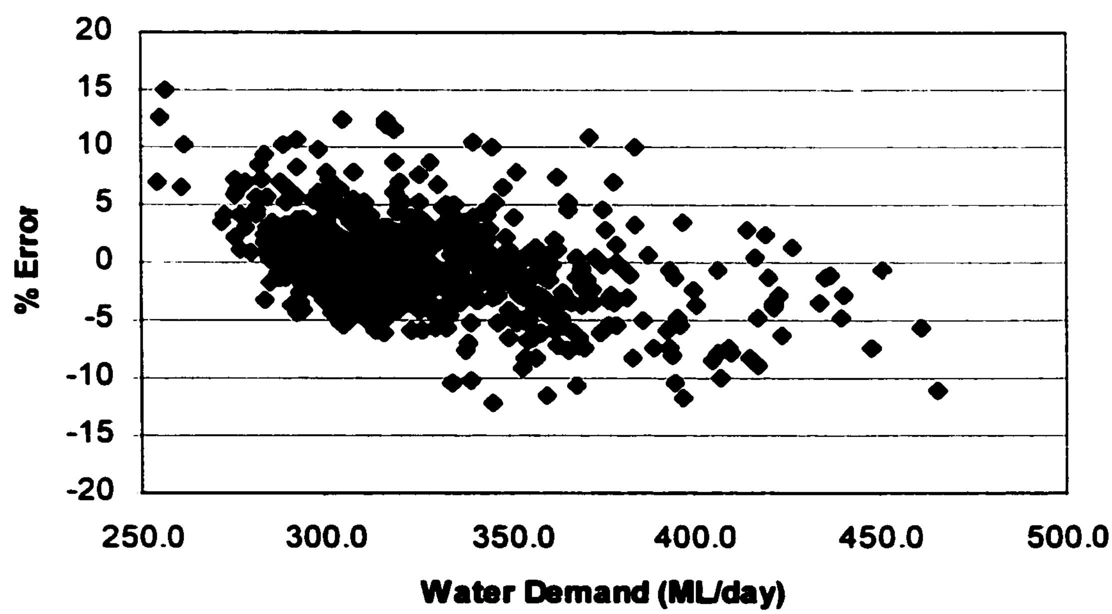

Figure 4.38 Daily Water Demand %Error vs. Daily Water Demand for the Ward Net with 2 Hidden Slabs and a Jump Connection Model for 2-12 Day Model

Table 4.13 Distribution of Percent Error for Daily Water Demand 2-12 Day Model

Network	Data Set	Percentage error			
		0-5 %	5-10%	10-15%	15%<
4 Layer Backpropagation Network	Production	192 (81.4%)	38 (16.1%)	6 (2.5%)	0 (0%)
	Pattern	649 (82.3%)	116 (14.7%)	22(2.8%)	1 (0.1%)
Ward Net with 3 Hidden Slabs	Production	196 (83.1%)	31 (13.1%)	9 (3.8%)	0 (0%)
	Pattern	654 (83.0%)	112 (14.2%)	21 (2.7%)	1 (0.1%)
Ward Net with 2 Hidden Slabs and a Jump Connection	Production	199 (84.3%)	26 (11.0%)	11 (4.7%)	0 (0%)
	Pattern	666 (84.5%)	100 (12.7%)	21 (2.7%)	1 (0.1%)

*may not add up to 100% due to rounding

4.2.1.3 Input to Output Generalization

The input to output generalization was studied to investigate the relationship between the inputs and the output. Studying the effect that each input has on the output is done to confirm that the model has learned the relationship between the input to the output as we had expected from the past literature and operator experience. The results will either give us the relationship we expected or an unexpected outcome. With the relationship being as what one generally expected, the ANN model has most likely learned the general relationship between the input and the output. If the relationship is unexpected, this can be either due to the fact the model has incorrectly learned the input to output relationship, or it can give us new insight to the input to output relationship.

4.2.1.3.1 Relationship of Temperature to Daily Water Demand

From the literature on the subject and the water utility's past experience, the indication is that the daily water demand increases as the daily temperature increases. The input to output generalization for the 2-12 day model should be same as the daily water demand model. Thus an exponential relationship would be expected as indicated when the actual daily water demand is plotted against the maximum daily temperature (Figure 4.22). Examining the plots of the model relationship between the maximum daily temperature and the daily water demand (Figure 4.39) it is found that the input-output relationship between the maximum daily temperature and the daily water demand is as was expected for the 4-layer backpropagation network and the Ward network with 3 hidden slabs. The Ward network with 2 hidden slabs and a jump connection follows the same trend as the other two models, but has a local maximum at -15°C. Comparing, the Ward network with 2 hidden slabs and a jump connection temperature to water demand relationship to the actual water demand to temperature relationship shown in Figure 4.22, the local maxima at -15°C is not present. With the increase in water demand not present in the actual data, and the fact that the other models developed do not experience this peak, this model has most likely not learned the generalization between the temperature and the water demand.

The data for Figure 4.39 were generated using the three models mentioned previously. When generating the output data all other inputs were held constant with the exception of the minimum temperature. The minimum temperature was increased with the maximum temperature with a 10°C seperation between them. This was to more closely represent the actual situation that the models will be operating in. This was done to prevent the maximum temperature from being less than the minimum temperature, as this can not happen by the simple definition of minimum and maximum. It would also prevent unlikely scenerios where the minimum temperature would be -20°C and the maximum temperature would be 20°C, as this is unrealistic. Thus, the realtionship actually reflects the effect the minimum and maximum temperature has on the daily water demand. The other inputs that were used to generate the model data represent a summer weekday with no rain in the previous five days and a previous day water demand of 320 ML/day. When examining the actual daily water demand relationship with the maximum temperature in Figure 4.22 it should be noted that the other inputs affect the actual demand but a general relationship is still seen.

Since a constant difference in the minimum and maximum temperature of 10°C was maintained, the effect that each individual temperature has on predicting the daily water demand is not apparent. To differentiate between the effect that the minimum and maximum temperatures have on the daily water demand the minimum temperature was held constant at three different minimum temperatures while the maximum temperature was adjusted. It can be seen in Figure 4.40 that with a lower minimum temperature, there is a slightly lower daily water demand. In Figure 4.41, it can be seen that the greater the maximum temperature the greater the daily water demand. For every 5°C increase in temperature there is a corresponding increase in the daily water demand of 7.5 ML/day.

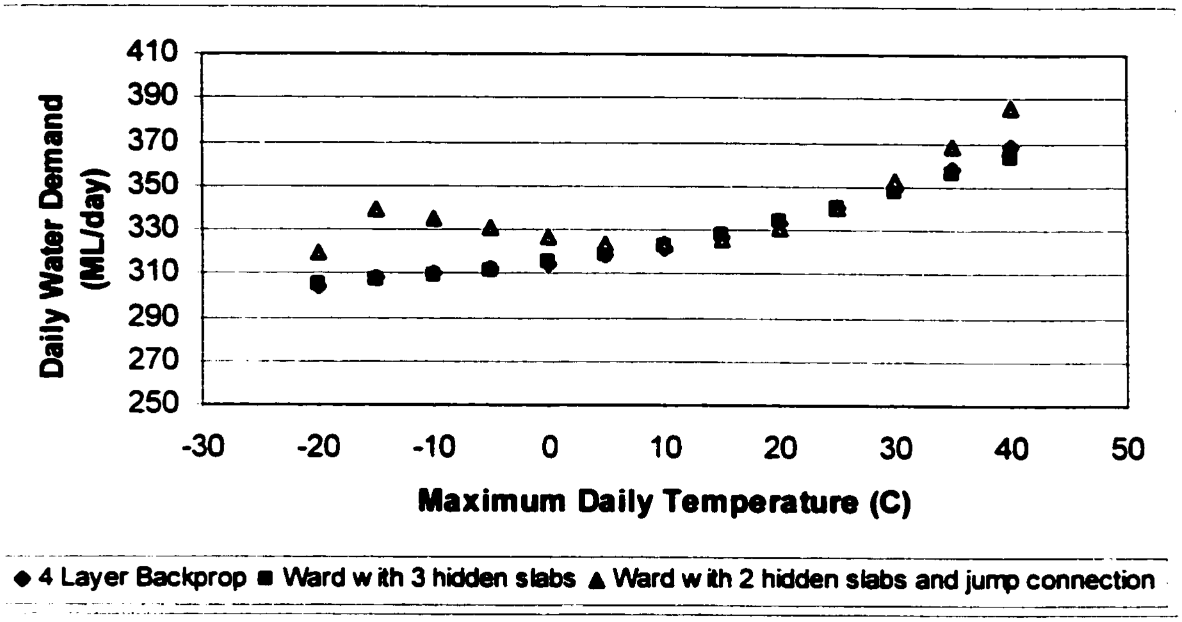

Figure 4.39 Model Generalization between Maximum/Minimum Daily Temperature Relationship and Daily Water Demand for 2-12 Day Model

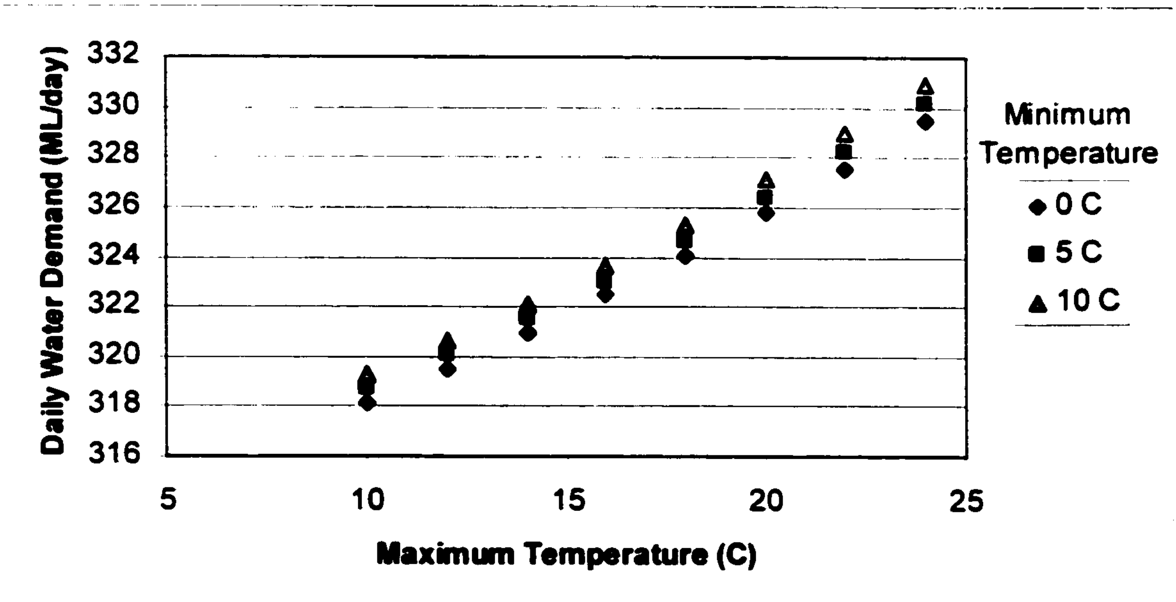

Figure 4.40 Minimum Temperature Effect With Respect to Maximum Temperature for Daily Water Demand for 2-12 Day Model

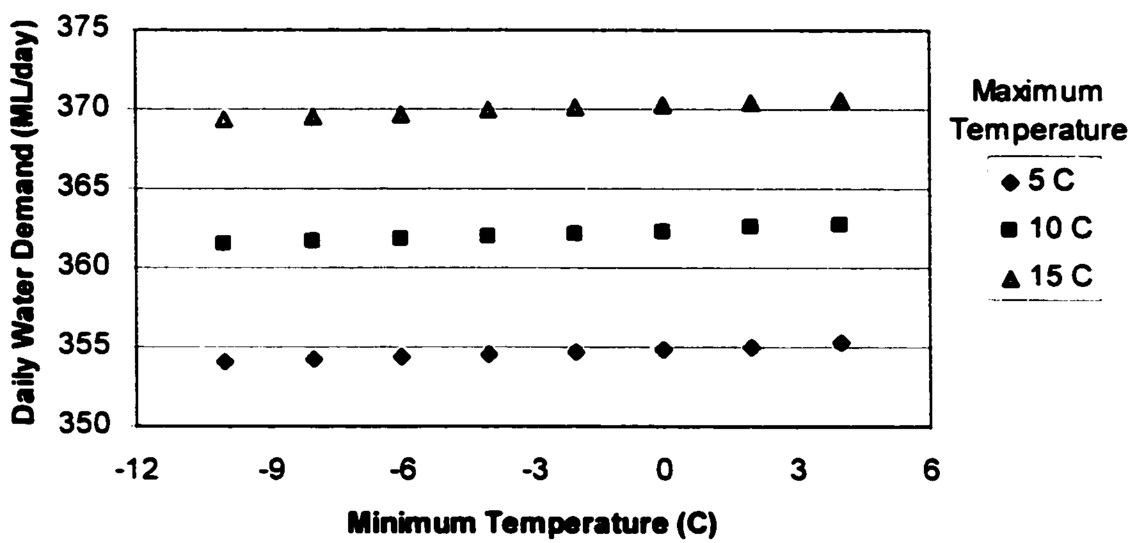

Figure 4.41 Maximum Temperature Effect With Respect to Minimum Temperature for Daily Water Demand for 2-12 Day Model

4.2.1.3.2 Relationship of Previous Day Rainfall to Daily Water Demand

The relationship between the previous day rainfall with the daily water demand as indicated by recent literature and the water utility, indicates that the water demand should drop as the rainfall increases in general, but is dependent on the other input interactions as well.

The model relationship of the previous day rainfall and the daily water demand is found in Figure 4.42. It can be seen that the 4-layer backpropagation network and the Ward network with 3 hidden slabs produce similar generalization in regards to the daily water demand and previous day rainfall. The daily water demand decreases with both models as the previous day rainfall increases. The Ward network with 2 hidden slabs and a jump connection model daily water increases as the previous day rainfall increases, which is contrary to the historical data, the current literature, and the other models developed. The most probable explanation is that the Ward network with 2 hidden slabs and a jump connection incorrectly learned to generalize the relationship between the previous day rainfall and the daily water demand. The model generated data was done by inputting the other factors to represent a summer weekday with a minimum temperature of 10°C, a maximum temperature of 20°C, and the previous day water demand of 320 ML/day. The rainfall in the previous 5 days was set equal to the previous day rainfall, this is simply saying that there was no rainfall in the previous 2-5 days.

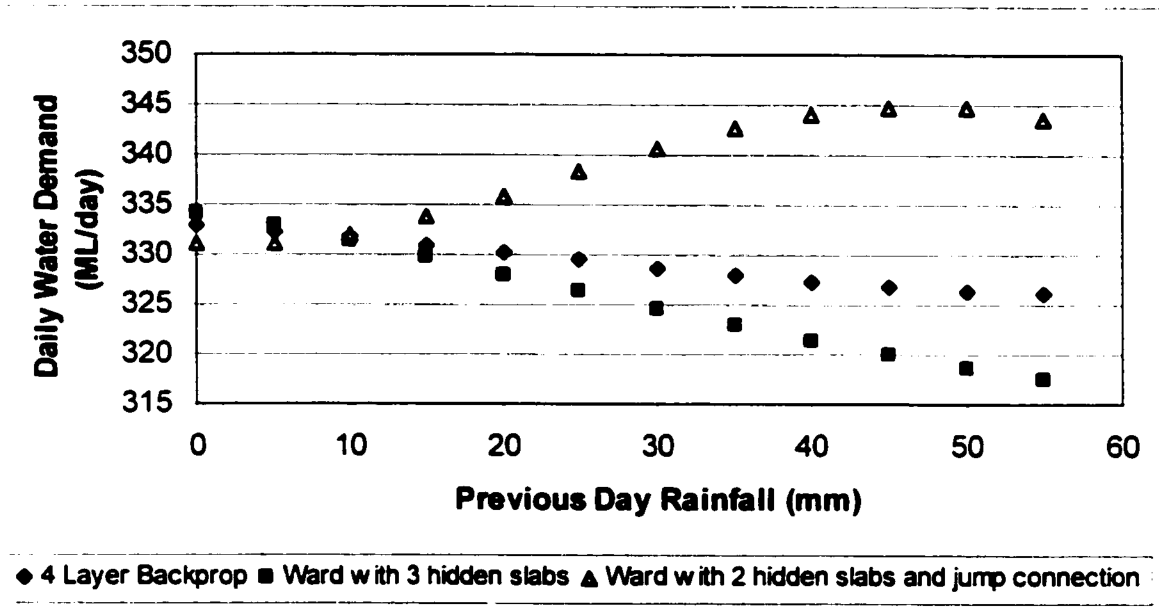

Figure 4.42 Model Generalization between Previous Day Rainfall and Daily Water Demand for 2-12 Day Model

4.2.1.3.3 Relationship of Previous 5-day Rainfall to Daily Water Demand

It can be seen in Figure 4.27, that previous 5-day rainfall is important as substantial rainfall in this time period effects the daily water demand. The general trend for the summer months is that when there is little or no rain in the previous five days, the higher the daily water demand is and decreases as the 5-day rainfall increases. In Figure 4.43 it can be seen that all three models developed, indicate that higher previous 5-day rainfall leads to lower daily water demands, but to varying degrees. The Ward network with 3 hidden slabs puts less significance on the importance of the previous 5 day rainfall input than the 4 layer backpropagation network. However, the Ward network with 3 hidden slabs puts more significance on the previous day rainfall than the 4 layer backpropagation network model does, thus both models have learnt the significance that rainfall has on the water demand, but differs on when the rainfall occurs. The Ward network with 2 hidden slabs and a jump connection has a slight increase and than a decrease in the water demand as the previous 5-day rainfall increases. The model generated data was done by inputting the other factors to represent a summer weekday with a minimum temperature of 10°C, a maximum temperature of 20°C, no rain in the previous day and a previous day water demand of 320 ML/day.

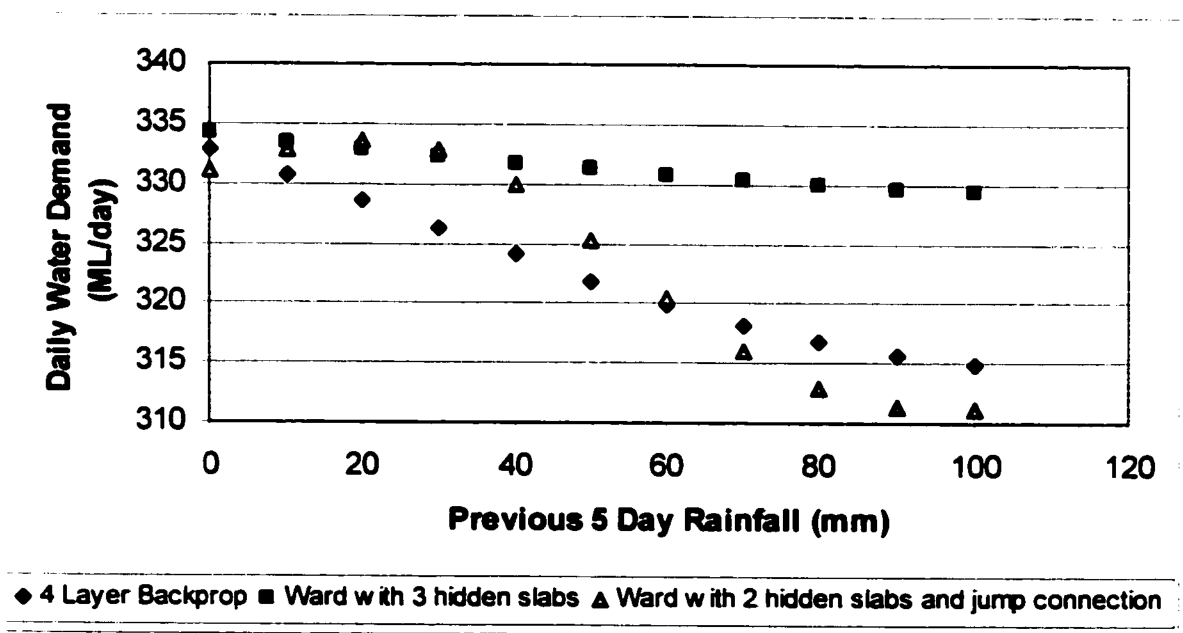

Figure 4.43 Model Generalization between Previous 5-Day Rainfall and Daily Water Demand for 2-12 Day Model

4.2.1.3.4 Relationship of Previous Day Water Demand to Daily Water Demand

The correlation between the previous day water demand and the daily water demand is expected in the same sense that the 10 a.m. water demand was used in the daily water demand model. When comparing Figure 4.44 and Figure 4.45, it can be seen that the model trend is similar to the trend that is seen in the historical data. The model generated data in Figure 4.45 was generated by holding the minimum temperature at 10°C, maximum temperature at 20° and no rain in the previous 5-days on a summer weekday.

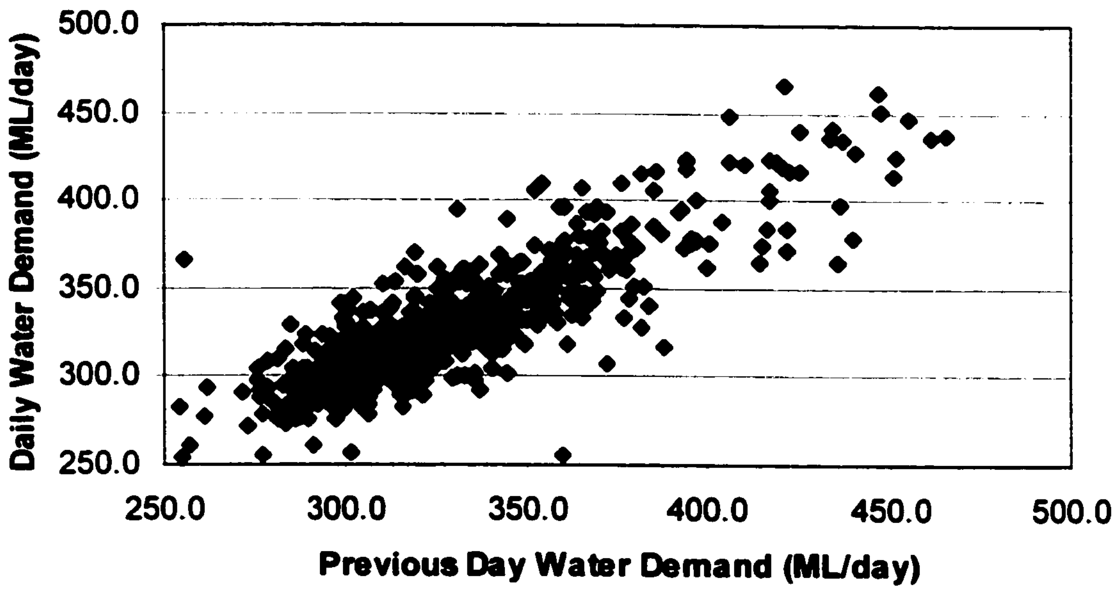

Figure 4.44 Actual Previous Day Water Demand Vs. Daily Water Demand

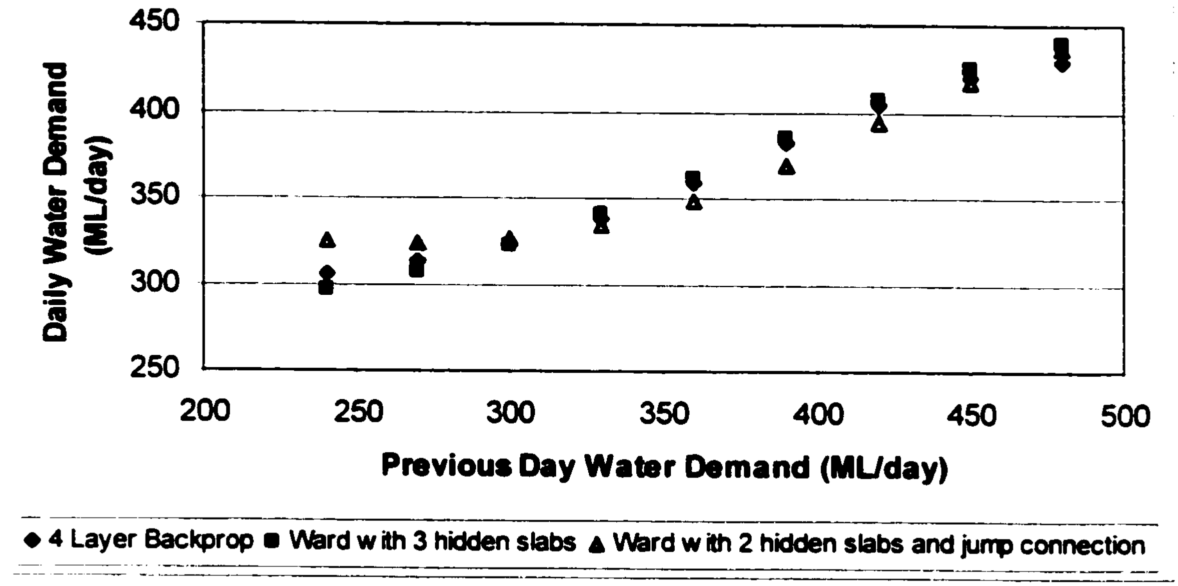

Figure 4.45 Model Generalization between Previous Day Water Demand and Daily Water Demand for 2-12 Day Model

4.2.1.3.5 Weekend/Holiday and Weekday Daily Water Demand Relationship

Human behavior needs to be taken into account when forecasting the water demand. One of these factors that were found was the difference between weekend/holiday and weekday water demand use, both in quantity and the time of use throughout the day. It was found that during the weekday that the daily water demand was higher on average than it was for the weekend/holiday period, as seen in Figure 4.31. The data generated from the model used for Figure 4.46, also show that the 2-12 day models predict that difference when presented with the weekend/holiday and the weekday scenarios. The Ward network with 2 hidden slabs and a jump connection is the only deviation from this. On the weekend it predicts the weekend water use similar to the other two models above 5°C, but the weekend predictions are similar to the weekday prediction below 5°C.

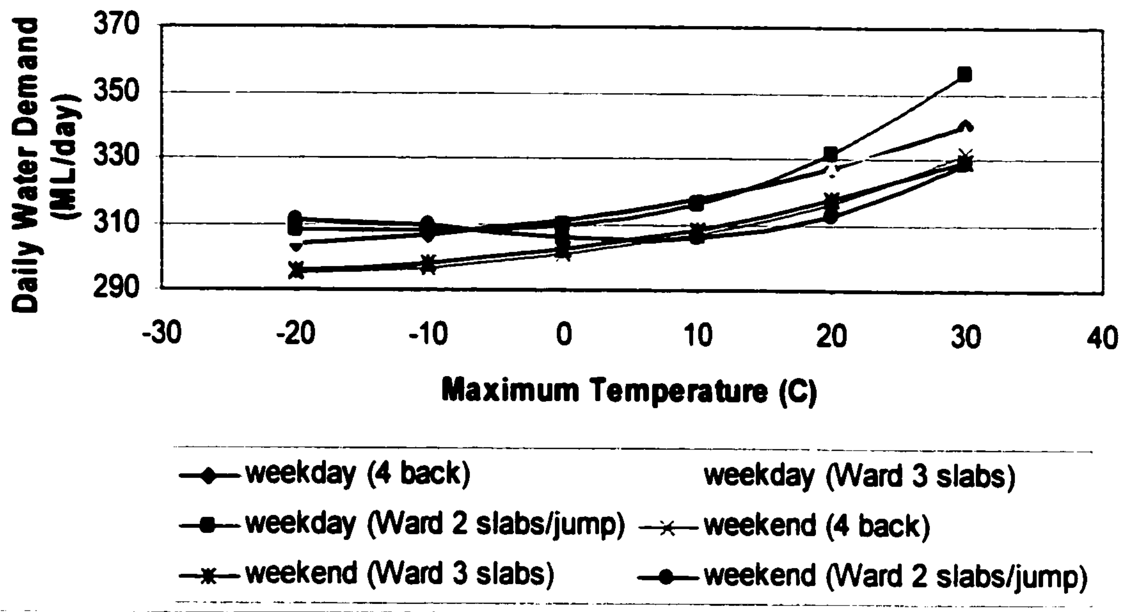

Figure 4.46 Comparison of Weekday and Weekend/Holiday 2-12 Day Water Demand

4.2.1.3.6 Summer and Winter Daily Water Demand Relationship

The ANN models indicate that the historical data have a trend in that the water demand is about 10 ML/day or 3% less during the October-April period than the May-September period under similar conditions as seen in Figure 4.47. This is slightly contradictory to the results obtained in the daily water demand model, where at higher temperatures there was about a 5 ML/day difference in water demand between the two seasons and at lower temperatures there were no difference. The difference is that the 10 AM water demand input in the daily water demand model has a greater weight than the day before water demand in the 2-12 day model. The use of the 10 am demand acts as a better reference indicator than the day before water demand, as it is part of the water demand for the day, and reflects the conditions and trends in the water use for that day. The day before water demand in general acts in a similar way, but is further away from the actual forecast in time.

This spatial difference then allows time for events to occur that would alter the water demand for the forecasting period. Thus, the day before water demand in the 2-12 day model is not weighted as heavily as an input parameter. With the day before water demand having less of a weighting, one or more of the other inputs would have to increase in their weighting to offset that difference.

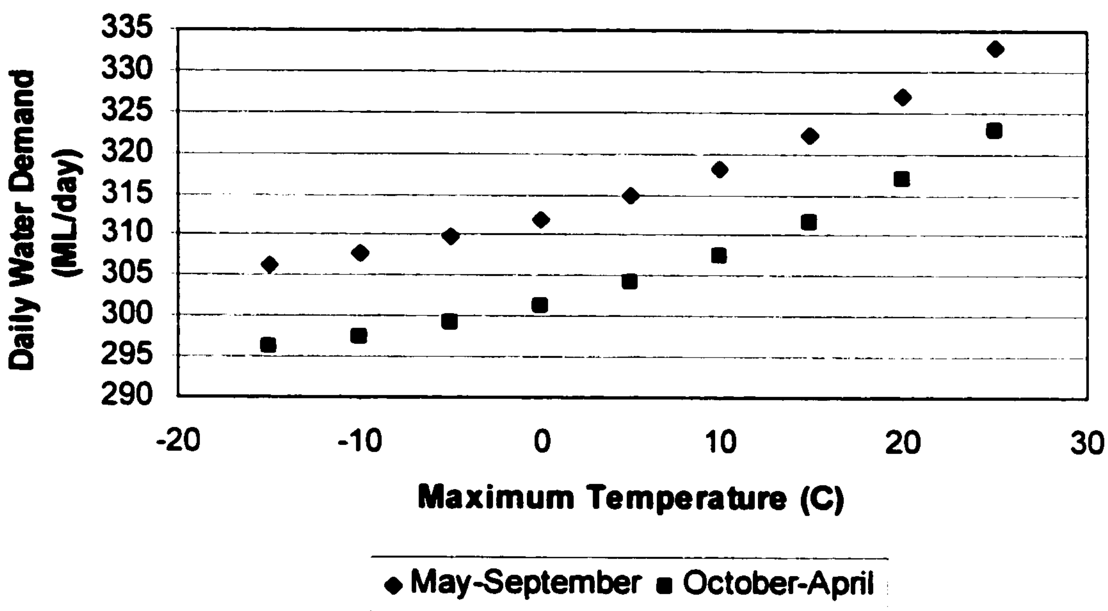

Figure 4.47 Comparison of May-September and October-April Daily Water Demand for 2- 12 Day Model

4.2.2 Final 2-12 Day Water Demand Model

The final 2-12 day water model selected was a combination of the 4-layer backpropagation and the Ward net with 3 hidden slabs. The Ward net with 2 hidden slabs and a jump connection was excluded in that the model input to output generalization for the maximum temperature, previous day rainfall and weekday/weekend were contrary to what was expected from the literature, as well as, the trend from the actual data. The combined model is the average of the 4-layer backpropagation model and the Ward network with 3 hidden slabs predictions. The combined model has a r^2 of 0.8448, mean absolute error of 9.96 ML/day (3.07% error), maximum absolute error of 47.50 ML/day (14.66% error) on the production set data. It also predicted within 5% of the actual water demand 82.2% of the time, within 10% of the actual water demand 96.6% of the time, and at all times within 12.4% of the on the production set data. The combined model also follows the general trend of the data and hits the peak demand as can be seen in Figure 4.48. The results in Figure 4.48 are based on using the actual temperature and rainfall inputs, as opposed to using the forecasted temperature and 30 year average values. The results for day 2-12 will decline as the weather forecasts are not as accurate the further it is predicted into the future, and as the model gets to day 6 it will be only using 30-year average values for the meteorological inputs. In using the 30-year average values the water demand forecast follows the general trend of the water demand as seen in Figure 4.49. The model can not predict the peaks in the water demand after day 5, as the inputs are average values only, and not the extreme events that are associated with the peaks in water demand. To examine the ability of the 2-12 day model, a real time simulation was run for the July of 98, which had an average demand of 376 ML/day. A summer month was used in the real time simulation as this is where the most variability in water demand is seen and thus is the most difficult to predict. Thus the results for the one summer month will be worse than if one full year is used, as during the winter months where there is little variability the model results will improve. The results of the real time simulation are outlined in Table 4.14. It can be seen that it goes from a r^2 of 0.66 with an average error of 21.5 ML/day (5.72% error) on day 2 to a r^2 of 0.09 and average error of 40.4 ML/day (10.74%) by day 12.

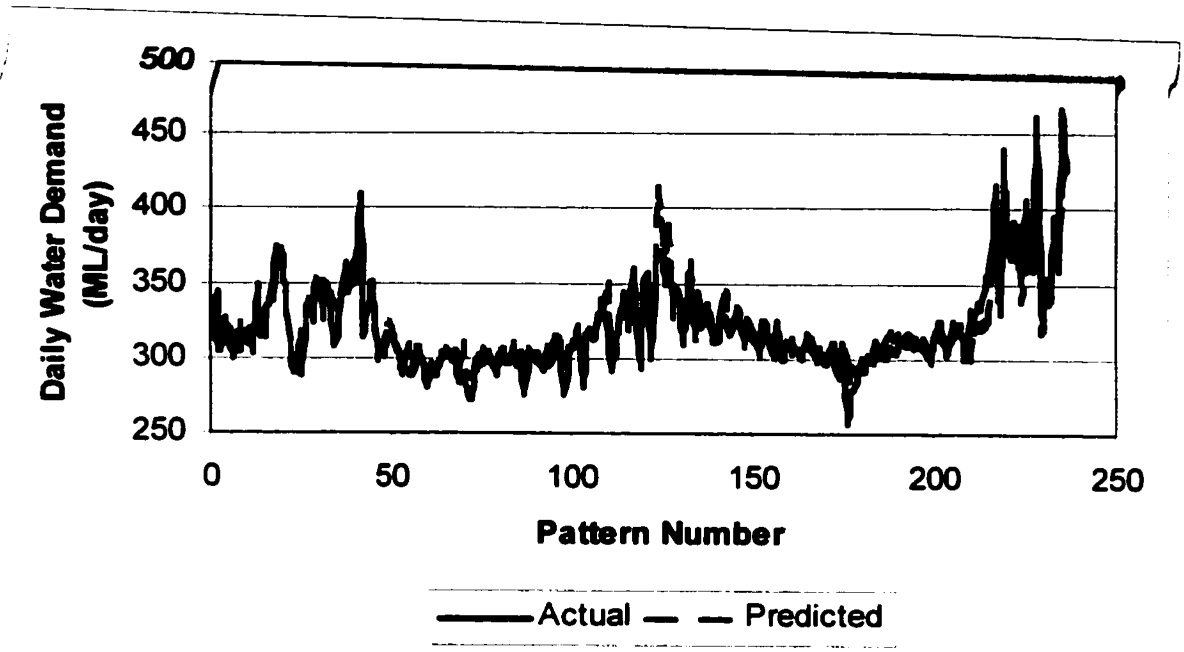

Figure 4.48 Actual vs. Predicted Water Demand for Days 2-12 (Production Set)

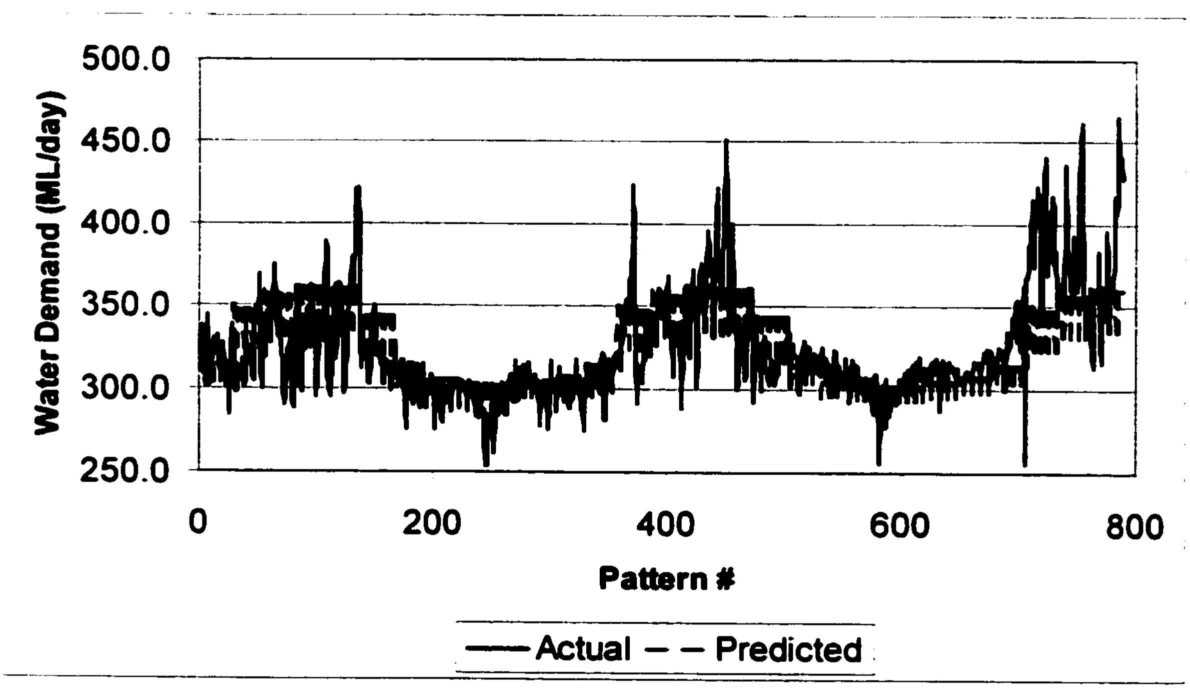

Figure 4.49 Day 12 Prediction Using 30-Year Average Inputs

Table 4.14 Real Time Simulation Results for July, 1998

Day	r^2	average mean absolute error (ML/day)	Percent Error
1	0.87	12.5	3.3
2	0.66	21.5	5.7
3	0.54	26.3	7.0
4	0.46	28.6	7.6
5	0.32	33.5	8.9
6	0.22	35.0	9.3
7	0.18	35.2	9.4
8	0.16	35.9	9.5
9	0.13	37.4	9.9
10	0.12	38.5	10.2
11	0.11	39.4	10.5
12	0.09	40.4	10.7

4.3 Hourly Water Demand

To model the hourly water demand a different approach was used. Since the water demand throughout the day is similar in its distribution to that of any other day, dimensionless (normalized) demand curves were used. The dimensionless demand curves were developed by dividing each hourly demand by the average of the daily demand for the specific curve being developed. As the water use pattern differs from weekday to weekend/holiday, seen in Figure 4.50, separate normalized demand curves where developed for weekdays and weekend/holidays. With the differences in water demand, different dimensionless demand curves were also developed to account for the slight difference in shape that was associated with the change in the water demand. This is demonstrated in Figure 4.51 to Figure 4.54. It can be seen that the major differences in the normalized demand curves occur at the two peaks in the water demand. At the first peak, which occurs at approximately at 10 am during the weekend/holiday curves and 7:00 am for the weekday curves, the lower the water demand the higher the resulting peak is on the normalized curves. Conversely the second peak, which occurs between 6:00 - 9:00 PM, the higher the daily water demand the higher the peak on normalized demand curve. Also with the second peak, the time that the peak occurs at also depends on the water demand. With low water demands the peak occurs around 6:00 PM, as the daily water demand increases, the time the peak occurs at increases up to a maximum of 9:00 PM.

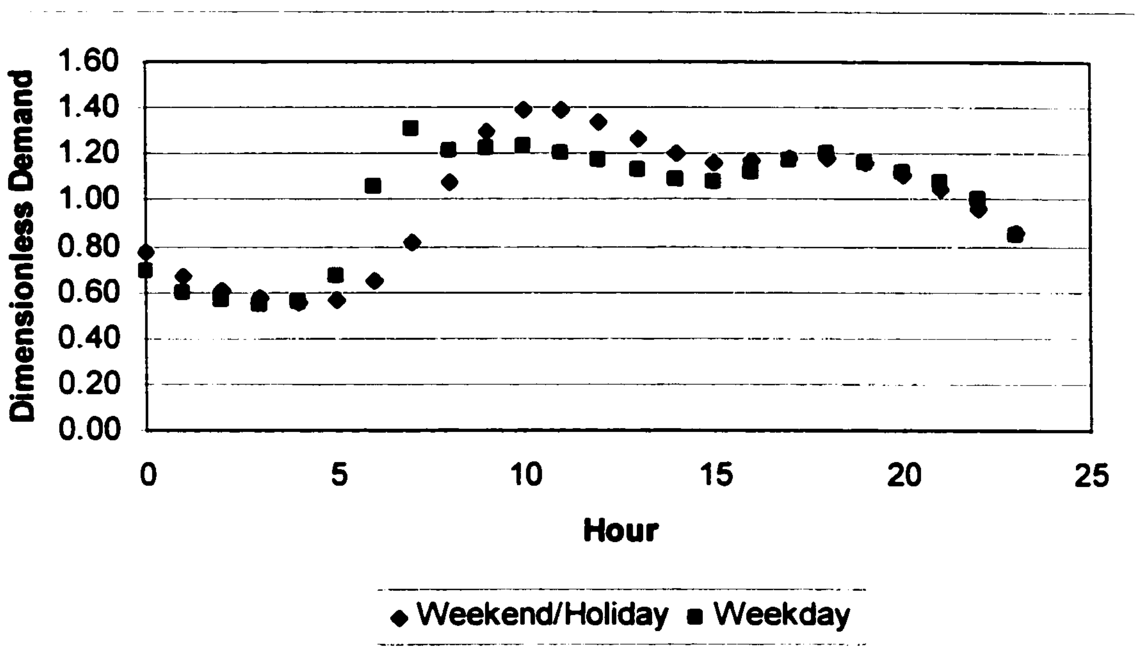

Figure 4.50 Normalized Demand Curves Weekend/Holiday and Weekday

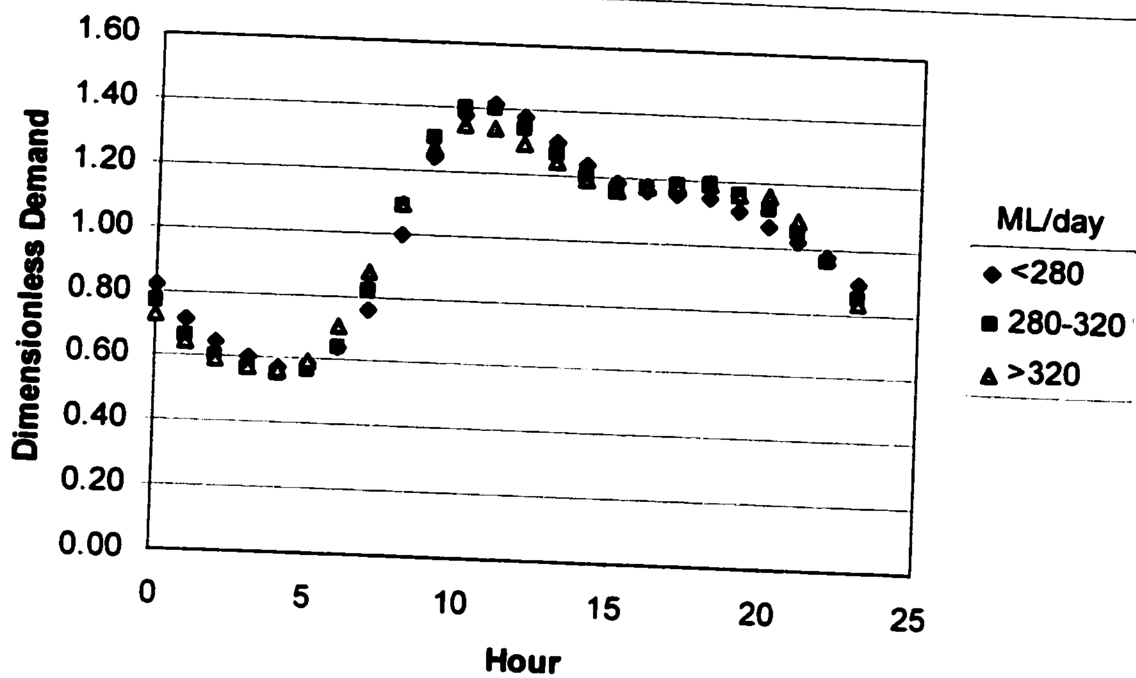

Figure 4.51 Normalized Demand Curves Weekend/Holiday October - April

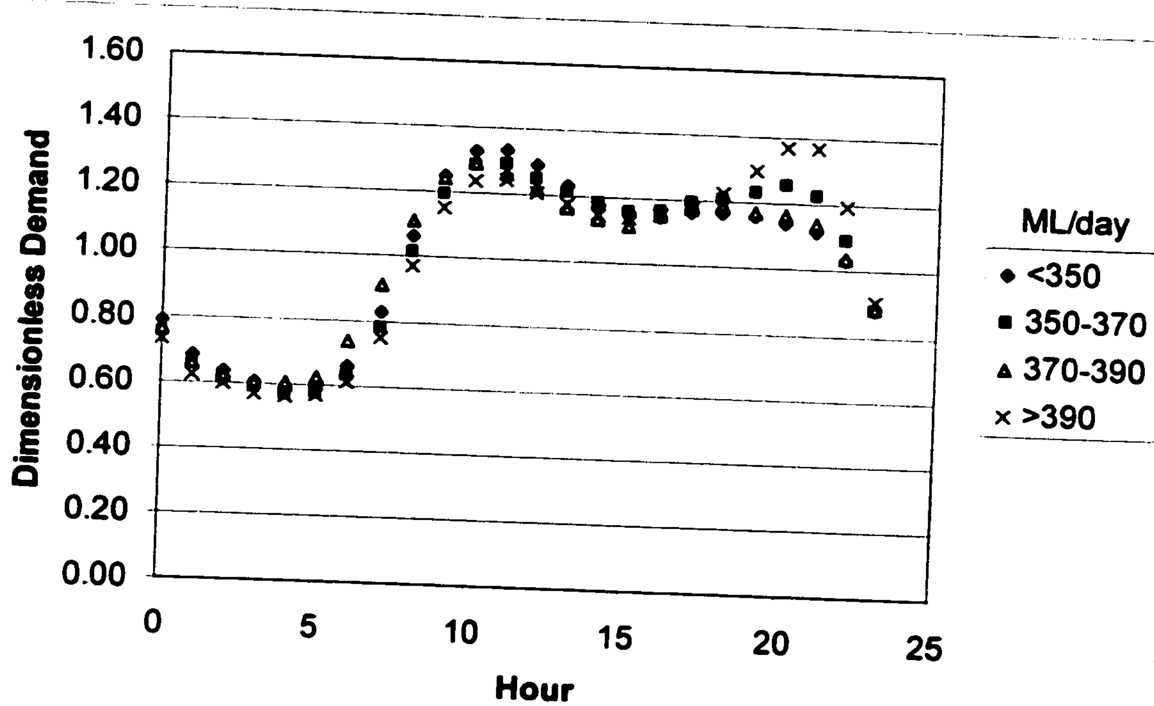

Figure 4.52 Normalized Demand Curves Weekend/Holiday May - September

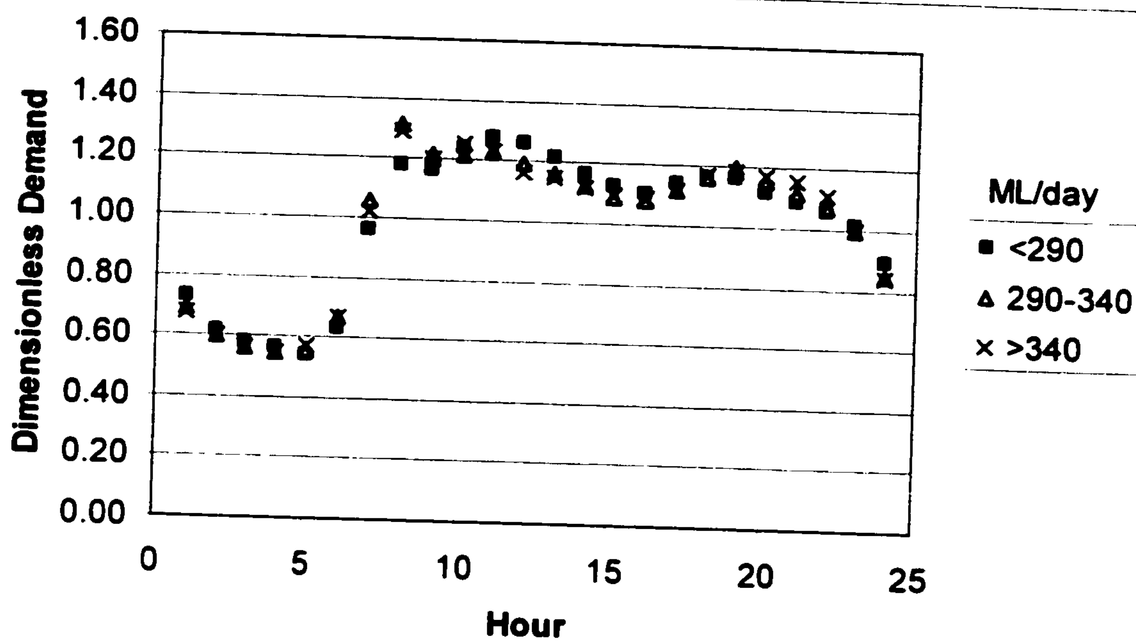

Figure 4.53 Normalized Demand Curves Weekday October – April

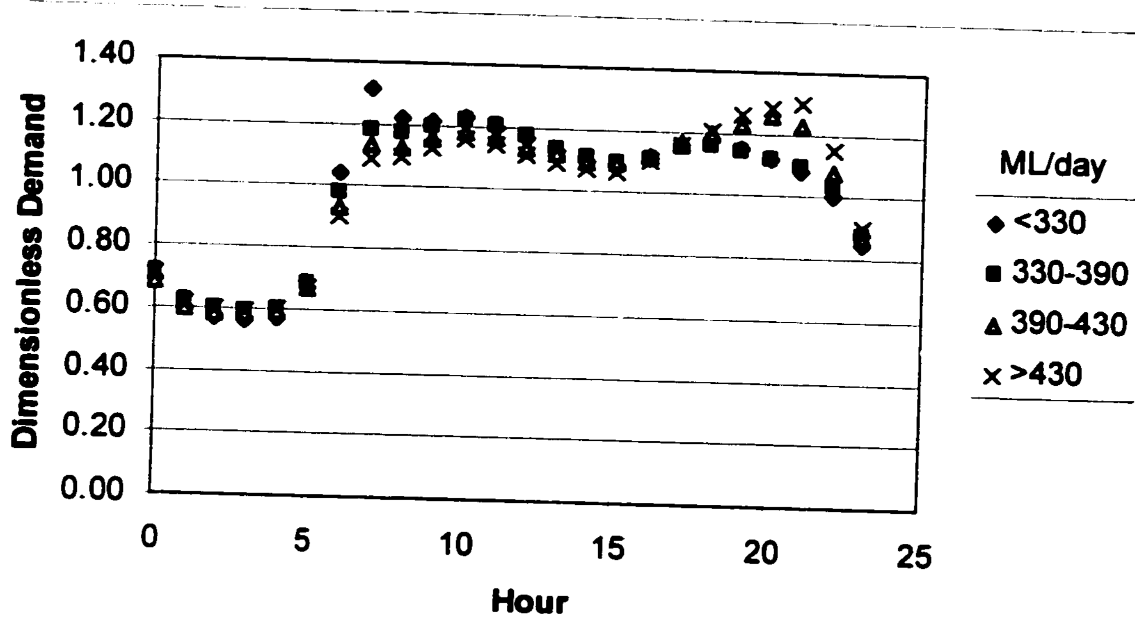

Figure 4.54 Normalized Demand Curves Weekday May - September

With the different dimensionless demand curves developed for the different seasons, day of the week and the daily water demand, their accuracy needed to be tested. They were tested by superimposing the appropriate normalized curve on each water demand in the data set and then comparing it to the actual water demand for that hour and condition. In Table 4.15, it can be seen that the normalized demand curves predictions on average have an error of 0.03 or 3% with a standard deviation of 0.03 and a r^2 0.96. The curves had a very high maximum error of 53%. To better understand the distribution of the hourly demand errors, the error was broken down into what percentage of the hourly prediction falls within different ranges of error in Table 4.16. Using the normalized demand curves to predict the hourly water demand, they predict within 10% of the hourly water demand 95% of the time and within 30% of the actual hourly demand over 99.9% of the time.

Table 4.15 Statistical Error Analysis of Dimensionless Demand Curves

	average error	minimum error	maximum error	standard deviation	r^2
Dimensionless Demand	0.03	0.00	0.53	0.03	0.96
Water Demand (ML/day)	11.34	0.00	181.04	11.17	0.96

Table 4.16 Error Distribution for Hourly Demand Predictions

	Percent Error in Hourly Prediction					
	<5%	**5-10%**	**10-15%**	**15-20%**	**20-30%**	**>30%**
Distribution of Errors	77.47%	18.02%	3.33%	0.79%	0.31%	0.08%

5 Applications

The final models can be used to predict the hourly, daily and 2-12 day water demands. Using these models as opposed to relying on operator experience is beneficial in a variety of ways. First, the models developed free up the operators, as they are no longer needed to develop the water demand forecast. Second, the operator developed water demand forecast is based on the operator's past experiences, while some operators may have the experience and expertise to develop an accurate forecast, others do not. Thus by using the ANN models and the normalized demand curves the accuracy and error associated with the forecasts are known, where as it varies with each operator. Third, by moving to a formalized water demand forecast, an accurate forecast isn't reliant on any one person having the ability to develop the forecast. When an experienced operator leaves the water treatment plant and a new operator is hired, the experience is lost and the ability to develop an accurate operator water demand forecast is compromised.

The primary use of a water demand forecast is to determine the amount of water production the utility needs to meet the demands of its customers. The daily demand forecast allows the water utility to ensure that the flow (production) rate through the plant is able to meet the demands for that day. The production rate is not solely set on the daily demand itself, it also needs to take into account other factors such as the reservoir levels, and the expected longer term demand. If the reservoir levels are low, the production level is adjusted so that it meets the daily demand, plus refills the reservoirs if the water treatment plant's capacity allows. The longer-term demand also needs to be considered. If the water treatment plant is coming into a predicted period of high water demand, it is able to increase production and fill the reservoirs so that it is able meet the high demands and avert a possible water shortage.

Being proactive in predicting the higher demands allows a smaller increase in production to be spread out over a longer period of time. This is to ensure that the water treatment plant's finished water quality remains as high as possible while still meeting the demands of its customers. The other major benefit of having the 12-day forecast is that it allows the water utility to evaluate any possible shortcomings in their ability to supply water, while scheduling maintenance that affects the plants treatment capacity. When planning scheduled maintenance and a possible water shortage is foreseen, the maintenance would be delayed to a later date when a lower demand is expected, providing the maintenance can be delayed.

The hourly water demand gives a more in-depth look at the water demand. It allows the utility to identify the water use closer to the instantaneous use of its customers. The hourly demand forecast can be used to aid in minimizing the cost of pumping the water to the reservoirs in the distribution system. This is essentially done by filling the reservoirs at night during periods of low demand and low energy costs. The idea is simple, but where it becomes more complex is that the reservoir space is limited, both at the water treatment plant and off-site throughout the distribution system. Using the hourly water demand forecast allows the utility to maximize the quantity of water that is pumped to the off-site reservoirs during periods of low electrical costs. During periods of high electrical costs, the quantity of water being pumped from the water treatment site can be reduced to a level such that the water is being drawn from the reservoirs and is only augmented from the water treatment plant's on-site storage. The hourly water demand forecast determines the amount of water that needs to be augmented, such that the levels within the reservoirs are not drawn down below a safe level. This is to ensure that there is sufficient water available in case of an emergency and for fire protection.

The NeuralShell 2 software supports dynamic link library (DLL) file types. This allows ANN models to be executed not only within the NeuralShell 2 software, but in programs written in Visual Basic, Access Basic, Pascal, C, Excel, plus a few other languages. The models can be used manually to fully automated for forecasting the water demands. The operator, to obtain the resulting water demand output can input the inputs manually into the models. The inputs also could be input automatically via a SCADA system, with the water demand output being fed back into the SCADA system to be input into a pump schedule, the water production schedule or any other system where it may be of use.

6 Conclusions and Recommendations

6.1 Conclusions

The purpose of this study was to develop artificial neural network models for forecasting the daily water demand for up to 12 days in advance for EPCOR Water Services in Edmonton, Alberta. In conjunction to the daily water demand a method was needed to breakdown the daily water demand in to its hourly demand. Both the daily and 2-12 day models were developed using historical data, and verifying them on a previously unseen data set. Both models developed were able to predict the water demand with a high degree of accuracy. The model characteristics are outlined in Table 6.1 and the input parameters are summarized in Table 6.2. The daily water demand out performed the 2-12 day model in its ability to predict the water demand as the further into the future one must peer, the larger the error in the prediction. The hourly water demand predictions were for the most part highly accurate as well, with only a few predictions that were significantly off. Even though the odd hourly prediction was off, the method used to develop the hourly prediction was such that the cumulative 24-hour prediction would not be.

Table 6.1 Daily and Twelve Day Water Demand Model Characteristics

Network Architecture	Learning rate	Momentum	Neurons	Scaling Function	Activation Selection	Pattern Selection
4-Layer Backpropagation	0.4	0.2	8:8	linear <<-1,1>>	tanh15 tanh15	random
Ward Net 3 Hidden Slabs	0.1	0.1	8 8 7	linear <<-1,1>>	tanh15 gaussian sine	random

Table 6.2 Classification of input parameters

Daily Input Parameter	12 Day Input Parameter	Parameter type
Minimum daily temperature	Minimum daily temperature	Meteorological
Maximum daily temperature	Maximum daily temperature	Meteorological
Previous day's rainfall	Previous day's rainfall	Meteorological
Previous 5-day's of rainfall	Previous 5-day's of rainfall	Meteorological
Weekday/weekday index	Weekday/weekday index	Human behavioral
Season index	Season index	Human behavioral
Water demand at 10:00 am*	Previous day's water demand	Reference indicator

***The water utility forecasts their water demand at 10:30 am**

6.1.1 Daily Model

The final daily water model selected was the average prediction of the 4-layer backpropagation and the Ward net with 3 hidden slabs. The Ward net with 2 hidden slabs and a jump connection was excluded as the model input to output generalization for the previous day rainfall was contrary to the trend of the actual data and what is expected from the literature. The combined model has a r^2 of 0.911, mean absolute error of 7.35 ML/day (2.27% error), maximum absolute error of 28.53 ML/day (8.81% error) on the production set data. The combined model also follows the general trend of the data and predicts the peak demands.

6.1.2 Two to Twelve Day Model

The final 2-12 day water model selected was the average prediction of the 4-layer backpropagation and the Ward net with 3 hidden slabs. The Ward net with 2 hidden slabs and a jump connection was excluded in that the model input to output generalization for the maximum temperature, previous day rainfall and weekday/weekend were contrary to the trend from the actual data and what was expected from the literature. The combined model has a r^2 of 0.8448, mean absolute error of 9.96 ML/day (3.07% error), maximum absolute error of 47.50 ML/day (14.66% error) on the production set data using the actual values as the inputs as opposed to the weather forecast and 30-year average values. The combined model also follows the general trend of the data and hits the peak demands. In using the 30-year average values, the water demand forecast follows the general trend of the water demand, but does not predict the peak demands. The real time simulation resulted in the combined model capable of an r^2 of 0.66 with an average error of 21.5 ML/day (5.72% error) on day 2 to an r^2 of 0.09 and average error of 40.4 ML/day (10.74% error) by day 12.

6.1.3 Hourly Water Demand

The normalized demand curves predictions have an average error of 0.03 or 3% with a standard deviation of 0.03 and a r^2 0.96 in predicting the hourly demand. The hourly demand had a very high maximum error of 53%. Using the normalized demand curves to predict the hourly water demand, they predict within 10% of the hourly water demand 95% of the time and within 30% of the actual hourly demand over 99.9% of the time.

6.2 Recommendations

The collection of historical data should be continued. The models should be retrained periodically. The models need to be retrained periodically so that any changes in water usage by the customers is learned and also to account for growth. The data collection should also be expanded to include other possible inputs that are not currently collected or periodically collected. Data recommended to be collected are the probability of precipitation and a sunshine hours surrogate input. The probability of precipitation is periodically recorded, but a more extensive database, which includes probability of precipitation, is needed. Even though the sunshine hours input is available, a forecast for this input is not available. A surrogate input could be collected in its place, such as cloud cover. Instead of collecting cloud cover information a system could also be developed to approximate the number of sunshine hours from the forecasted cloud cover. As this system does not need to predict the sunshine hours exactly, a ball park figure input to distinguish between clear, partially cloudy, mainly cloudy and over cast conditions may suffice. Developing a cloud cover to sunshine hour system is preferable to collecting cloud cover data as it could be implemented immediately as opposed to the time needed to collect the cloud cover data.

7 References

Anderson, J. A. and E. Rosenfeld. 1988. *Neurocomputing : foundations of research*. Cambridge, Mass.: MIT Press.

Baxter, C. W., S. J. Stanley, and Q. Zhang. 1998. Development of a Full-Scale Artificial Neural Network Model for the Removal of Natural Organic Matter by Enhanced Coagulation. *Aqua*, vol. 48, pp. 129-136.

Billings, R. B. and D. E. Agthe. 1998. State-space versus multiple regression for forecasting urban water demand. *Journal of Water Resources Planning and Management*, vol. 124, pp. 113-117.

Chai, A. and R. C. Andrews. 1998. Application of Neural Networks to Water Treatment Particulate Removal Processes. OWMA/OWWA Joint Annual Conference, Niagara Falls, Ont.

Crommelynck, V., C. Duquesne, M. Mercier, and C. Miniussi. 1992. Daily and Hourly Water Consumption Forecasting Tools Using Neural Networks. 1992 Computer Conference Proceedings, Nashville, TN. pp.665-676.

Daniell, T. M. 1991. Neural Networks - Applications in Hydrology and Water Resources Engineering. International Hydrology and Water Resources Symposium, Perth.

Fildes, R., A. Randall, and P. Stubbs. 1997. One day ahead demand forecasting in the utility industries: Two case studies. *Journal of the Operational Research Society. Jan*, vol. 48, pp. 15-24.

Fleming, N. 1994. Comparison of water consumption forecasting methods including artificial neural networks. *National Conference Publication Institution of Engineers, Australia*, vol. 3, pp. 115-120.

Flood, I. and N. Kartam. 1994. Neural Networks in Civil Engineering I: Principles and Understanding. *Journal of Computing in Civil Engineering*, vol. 8, pp. 131-163.

Hall, K. C. and D. R. Maidment. 1990. Analysis of the impact of weather variation on monthly water demand. Optimizing the Resources for Water Management - Proceedings of the ASCE 17th Annual National Conference. Fort Worth, TX, USA. pp. 1-6.

Hebb, D. O. 1949. *The Organization of Behavior*. John Wiley.

Heims, S. J. 1982. *John von Neumann and Norbert Weiner - From Mathematics to the Technologies of Life and Death*. Academic Press.

Hittle, D. C., P. A. Flocken, and C. W. Anderson. 1996. Neural networks for predicting chilled water demand in buildings. International Solar Energy Conference. pp. 387-394.

Jain, A. and L. Ormsbee. 1993. Comparison of short term forecast methods for municipal water use. *Water Resources Planning and Management and Urban Water Resources*, pp. 649-652.

Kulshreshtha, S. N., Z. Kos, and V. Priazhinskaya. 1996. Implications of climate change for future water demand. *Water Resources Management In The Face Of Climatic hydrologic Uncertainties*, vol. 18, pp. 106-167.

Lertpalangsunti, N. and Christine W. Chan. 1997. Implemented architecture for construction of hybrid intelligent forecasting systems. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics. pp. 934-939.

Maren, A. J. 1990. Introduction to Neural Networks. *Handbook of Neural Computing Applications*, A. J. Maren, C. T. Harston, and R. M. Pap, Eds. San Diego, CA: Academic Press, Inc., pp. 1-11.

McCulloch, W. S. and W. Pitts. 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115-132.

Minsky, M. L. and S. A. Papert. 1969. *Perceptrons: An Introduction to Computational Geometry.* Cambridge, MA: The MIT Press.

Ormsbee, L. E. and K. E. Lansey. 1994. Optimal control of water supply pumping systems," *Journal of Water Resources Planning and Management*, vol. 120, pp. 237-252.

Oshima, N. and T. Kosuda. 1998. Distribution reservoir control with demand prediction using deterministic chaos method. *Water Science and Technology*, vol. 37, pp. 389-395.

Palmer, R. N., J. M. Kersnar, and D. Choi. 1995. Estimating demand variability. Proceeding of the 22nd Annual Conference of Integrated Water Resource Planning in the 21 Century. ASCE. pp. 997-1000.

Rajasekaran, S., M. F. Febin, and J. V. Ramasamy. 1996. Artificial Fuzzy Neural Networks in Civil Engineering. *Computers-and-Structures*, vol. 61, pp. 291-302.

Shvartser, L., U. Shamir, and M. Feldman. 1993. Forecasting hourly water demands by pattern recognition approach. *Journal of Water Resources Planning and Management*, vol. 119, pp. 611-627.

Tamada, T., M. Maruyama, Y. Nakamura, S. Abe, and K. Maeda. 1993. Water demand forecasting by memory based learning. *Water Science and Technology*, vol. 28, pp. 11-12.

Zhang, Q., C. W. Baxter, and S. J. Stanley. 1999. Automating Water Treatment Plants Using Artificial Neural Network Models. Proceedings of the 1999 Engineering and Construction Conference, Orlando, FL.

Zhang, Q. and S. J. Stanley. 1997. The Artificial Neural Network Modeling Approach for Water Demand Forecasting in Edmonton. CSCE/ASCE 1997 Environmental Engineering Conference, Edmonton, Alberta.