# IMPROVED QUANTITATIVE ESTIMATION OF RAINFALL BY RADAR

by

Md Rashedul Islam

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Civil Engineering
University of Manitoba
Winnipeg

# Executive Summary

This study investigates procedures to get higher correlations between gauge and radar measurements of precipitation at high time resolution (e.g. 10-minute), explores the performance of artificial neural network models compared to the Z-R relationship, and generates surface precipitation maps from the atmospheric precipitation measured by the radar. Although higher correlation between gauge and radar at hourly or daily accumulations are reported by many authors, it is rarely observed at higher time resolution (e.g. 10 -minute). This study investigates six major rainfall events in year 2000 in the greater Winnipeg area with durations varying from four to nine hours. The analyzed weather maps provided by the Environment Canada shows that these rainfalls are the result of synoptic scale frontal systems. The correlation between gauge and radar measurements of precipitation is found to be only 0.3 at 10-minute resolution and 0.55 at hourly resolution using Marshall-Palmer's Z-R relationship ($Z=200R^{1.6}$). Other reflectivity-rain rate relationship are tested (e.g. $Z=300R^{1.4}$), but they generally show similar or lower correlation. The rainfalls are classified into convective and stratiform regions using Steiner et al. (1995)'s algorithm and two different Z-R relationships are tested to minimize the error associated with the variability of drop-size-distribution, however no improvement is observed.

The performance of the artificial neural network is explored as a reflectivity-rainfall mapping function. Three different types of neural networks are explored: the back propagation network, the radial basis function network, and the generalized regression neural network. It is observed that the neural network's performance is better than the Z-

R relationship to estimate the rainfall events which was used for training and validation (correlation 0.67). When this network is tested on a new rainfall its performance is found quite similar to that obtained from the Z-R relationship (correlation 0.33). Based on this observation neural network may be recommended as a post-processing tool but may not be very useful for operational purposes - at least as used in this study. Variability in weather and precipitation scenarios and uncertainties associated with radar affects the gauge and radar measurements which apparently makes it impossible for the neural network or the Z-R relationship to show consistent performance at every rainfall event.

It appears necessary to account for variability in weather and rainfall scenarios to get a better correlation between gauge and radar measurements. Hence, conventional correction schemes for attenuation and hail contamination are applied and a trajectory model is developed to account for rainfall advection due to wind drift. Reflectivity from the July 7$^{th}$, 2000, rainfall appears significantly attenuated due to its larger spatial extent of convective rain showers. The weather observers reported hail occurrence during three of the six storms; among them it appears that the July 23$^{rd}$ rainfall was severely affected by hail (reflectivity beyond 60 dBZ). The June 10$^{th}$ rainfall appeared to be significantly affected by wind shear as the mean wind speed is computed as 14 m/s from the Velocity Azimuth Display (VAD) profile. The trajectory model uses velocity obtained from the single-doppler observation and the quality of the VAD data is ensured by comparing the direction obtained from the VAD profile with the direction obtained from a correlation based storm tracking algorithm at the precipitation generation level (4 km assumed). The near-surface direction is compared with the anemometer data located at Winnipeg International Airport. A space-time interpolation technique is applied to generate

reflectivity maps at one-minute resolution based on the direction obtained from the correlation based tracking algorithm. The trajectory model uses the generated reflectivity maps having one-minute resolution which help to account for the travel time by the rainfall mass to reach to the ground. It is observed that the correlation between the gauge and the radar measurement is 0.81 after applying the various algorithms. The precipitation processing model developed in this study generates more accurate rainfall estimates at the surface from radar observations and may be a better choice for rainfall-runoff modellers.

It was found that the attenuation correction algorithm adversely increases the reflectivity which is observed in the case of the June 10$^{th}$ and the July 23$^{rd}$ rainfall events. This study assumes that the higher reflectivity caused by hail contaminated regions is one reason for the overestimation in the attenuation correction process. It was observed that the hail capping method applied prior to the attenuation correction algorithm helps to improve the situation. An attempt is made to develop an expression to account for radome attenuation. It is assumed that the underestimation observed even after applying the attenuation correction algorithm is due to radome attenuation.

Although Marshall-Palmer's relationship is recommended for stratiform precipitation only, this study found it suitable for both convective and stratiform precipitation when attenuation is properly taken into account. It is expected that the high time resolution structure of the model will make it more compatible for flash-flood forecasting and for now casting purposes.

# Acknowledgement

I would like to express my gratitude to my supervisor Dr. Peter Rasmussen, Associate Professor, University of Manitoba, for his informative and financial support to carry out my thesis and graduate studies. I would also like to thank Dr. Ken Snelgrove, former Asst. Professor, University of Manitoba, for his kind support on getting data and information.

I am indebted to Mr. Dave Patrick, Application Meteorologist, Environment Canada, for his continuous support for getting information and data. I am also thanking Dr. Norman Donaldson, Research Scientist, Environment Canada for his valuable information.

I am thanking Mr. Zeljko Bodiroga for providing processed rain gauge data. I am also thanking my friends and fellow graduate students for their valuable support on programming and related issues.

I am also thanking the Faculty of Graduate Studies, University of Manitoba, for partially financing my graduate studies at the University of Manitoba.

Finally, I am thanking the Almighty for making all these activities possible.

# Table of Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 PROBLEM STATEMENT

Radar estimated precipitation is found very useful for flood and rainfall forecasting, for now casting, for detecting tornadoes, etc. The underlying reason for its usefulness is its ability to provide spatial distribution of rainfall. Ensuring the quality of radar data is deemed necessary for getting accurate results from the applications that uses radar data. This is an issue on which researchers have made relentless efforts since its advent in the early 1940s.

Radar became a long range detection tool after the development of a powerful transmitting device in 1940 (Doviak and Zrnic, 1993). During the Second World War, radar was deployed in many parts of the world to detect enemy aircrafts. It is documented that the potential for detecting precipitation using radar was identified during its application in the war (Maynard, 1945). Although radar is efficient in detecting precipitation regions, its ability to make quantitative measurement of precipitation is still a concern. Variability in atmospheric conditions and rainfall characteristics influence radar quantification of precipitation. Attenuation, anomalous propagation, hail and bright band contamination, variability of drop-size-distribution, etc. are but a few of the

conditions that affect radar quality. It is then necessary to correct radar derived precipitation before applying it in a rainfall-runoff model; otherwise severe attenuation may suppress peak values and hail may show erroneously high peaks.

Reliable correction algorithms for several problems are devised and available in the literature. The Next Generation Weather Program (NEXRAD) assimilates up-to-date correction algorithms to improve the quality of radar precipitation products so that it can be directly used by the flood and weather forecasters (Klazura and Imy, 1993). The algorithm is known as the 'Precipitation Processing System' (PPS) (Fulton et al., 1998). In Canada, such processed precipitation products are not yet available (Donaldson 2005, personal communication) for operational purposes. One objective of the current study is to develop an integrated model that embeds several corrections in the radar product applicable for the summer rainfalls (typically May to August) within a-100 km radial distance from the radar site. For this purpose, six major summer rainfalls occurring in year 2000 in the Winnipeg area are selected. The gauge data is used to validate the correction schemes.

There is some concern regarding the use of gauge data to validate radar products. The reason is that radar measures the average precipitation intensity of a volume at a certain distance aloft and the gauge measures the precipitation accumulation at a point on the earth surface. Strong wind may advect rainfall mass up to several kilometres and it may also lead to under-catch at the gauge data due to turbulence. These are some of the reasons that gauge and radar data do not show better correlation at high temporal resolution and that radar precipitation in the atmosphere is not always representative of the precipitation at the surface. To solve the intensity-accumulation conflict, the proposed

model uses tracking and a space-time interpolation algorithm to generate reflectivity maps at one minute resolution which should help to compute accumulation in an efficient way. To get the reflectivity map at the surface this model uses a trajectory algorithm to track the precipitation mass from the atmosphere to the earth surface by using the velocity profile obtained from the doppler scanning. This trajectory algorithm shifts both the space and the time coordinates of a rainfall mass from the atmosphere to its arrival at the earth surface. In this way, the model attempts to generate reflectivity maps at the earth surface at one minute resolution. Thereafter accumulation can be computed at any spatial location for a chosen time interval and can be compared with the gauge data.

The NEXRAD precipitation processing system generates hourly rainfall accumulations (Fulton et al., 1998) which may not always be sufficient for now-casting or flash-flood forecasting purposes. It has been found that the gauge-radar correlation significantly increases when radar data is integrated over time and space (Doviak and Zrnic, 1993, p. 226). Hence, hourly accumulations are usually used to determine bias with respect to gauge or other quality checks of the radar data. One reason for the higher correlations at the hourly time step may be that hourly (or daily) accumulations smoothen the error associated with the time shifting during its travel from the atmosphere to the earth surface and partially eliminates the need for accurate time synchronization between gauge and radar sampling. The trajectory model coupled with space-time interpolation techniques may be useful for obtaining better correlations with gauges at high time resolutions and may be more useful for now-casting and flash-flood forecasting in small urban watersheds.

Radar incorporates numerous sources of errors mainly caused by the variability and complexity of the atmospheric conditions. Proper correction schemes for all possible errors are not available due to the lack of physical knowledge or the lack of necessary atmospheric data. For example, the drop-size-distribution of a rainfall mass in the atmosphere may change due to wind shear or due to coalescence of rainfall masses and the real Z-R relationship will change. It is not possible to measure rainfall intensity accurately during a hail storm as hail causes an unusual increase in reflectivity. It is reported that pixel-by-pixel attenuation correction algorithms often create unusual high reflectivity at the far edge of the rainfall area (Li and Illingworth, 2001). Hence, performance of a 'black box' model, which can efficiently map any input-target pairs should be explored. Artificial neural networks are widely believed to be capable of mapping any complex, non-linear relationship between input-target data sets (Hsu et al., 1995). It is also believed that the neural network is relatively more insensitive to noisy data (Zealand 1997, p. 11). The interest of exploring the possibility of the neural network is partly motivated by the results of Liu et al. (2001) who obtained a correlation between gauge and radar equal to 0.95 in an hourly accumulations comparison using the adaptive radial basis function neural network. It is of interest to explore the capability of neural networks with higher time resolution data (e.g. 10-minute) and to compare the performance with the trajectory model described above.

## 1.2 RESEARCH OBJECTIVES

The main objective of the research is to develop a precipitation processing system applicable to summer rainfall (typically May to August). The case study area is Winnipeg

located at 50-70 km radial distance from the radar site. The performance of the system will be compared to a 'black-box' model and to Z-R relationships.

The performance evaluation of the model is based on high temporal resolution data (e.g. 10-minute). The precipitation processing model generates rainfall at the earth surface at one-minute resolution. The performance is assessed based on 10-minute accumulations.

It is expected that the research will help to assess the potential usefulness of the correction schemes and the 'black-box' models ability to handle the uncertainties in reflectivity-rainfall relationships due to the variability of atmospheric conditions. It is also expected that the precipitation processing algorithm will be helpful in now-casting and in flash-flood forecasting for urban watershed due to its high time resolution structure.

The other objectives of the study are:

- to develop a Z-R relationship using hourly accumulations data;
- to assess the performance of a rainfall classification algorithm for minimizing the error associated with the variability in drop-size-distribution;
- to develop a 'climatological' Z-R relationship using probability matching techniques;
- to assess the quality of the velocity data obtained from single doppler scanning in synoptic scale precipitations;
- to develop a statistical expression for attenuation due to radome wetting; and
- to check the performance of a smoothing algorithm on radar-derived precipitation.

## 1.3 SCOPE OF THE RESEARCH

This study will develop a precipitation processing model for summer rainfall in Winnipeg. The proposed model generates reflectivity maps at the 1.5 km CAPPI at one-minute resolution. This is possible by applying a correlation based tracking algorithm to assess the speed and direction of the precipitation system and then applying space-time interpolation techniques to generate intermediate reflectivity maps within the available time resolution. A trajectory scheme transports each pixel to the ground using velocity data obtained from single doppler scanning and thereafter generates reflectivity maps at the earth surface. Both space and time coordinates of a pixel in the atmosphere change when it reaches the earth surface. An optimum attenuation and radome-wetting correction scheme is also developed and incorporated in the model. The model also accounts for unusually higher echoes generated by hail. Bright band and anomalous propagation issues are not considered assuming these factors are not a serious issue in convective rainfalls in the summer season in Winnipeg.

The performance of an artificial neural network is explored for mapping the reflectivity to rainfall. The performance of three different neural networks is explored; these are the back-propagation network, the radial basis function network, and the generalized regression neural network. The performance of each network is further evaluated by using two different input vectors, one consisting of a short time series of reflectivity at a particular location and the other consisting of reflectivities of nine surrounding pixels at a fixed time coordinate.

Chapter 2 provides a brief literature review of research work carried out on different aspects of radar meteorology. Attention is focused on the work carried out to

improve the quality of radar data and to reduce the discrepancy between gauge and radar measurements. Chapter 3 provides a brief description of the storm events employed in the study and the discrepancy of gauge and radar measurement observed when radar data is interpreted by a Z-R relationship. Chapter 4 discusses the algorithm of the neural network and the results obtained from the analysis. Chapter 5 provides a description and output from the precipitation processing model and finally Chapter 6 summarizes results and describes some future research directions.

# Chapter 2

# Literature Review

## 2.1 INTRODUCTION

Radar has been used for measuring precipitation for over 40 years. Despite significant progress in radar technology, there is still much uncertainty associated with quantitative radar rainfall estimates. This chapter discusses radar principles, reasons for discrepancies between radar and rain gauge measurements and improvement initiatives to reduce the discrepancy.

## 2.2 OVERVIEW OF WEATHER RADARS

### 2.2.1 Working Principle

A radar transmits pulses of energy towards the atmosphere. If the energy is obstructed by hydrometeors, a portion of the energy is backscattered and the rest is transmitted through the hydrometeor and may be backscattered by other hydrometeors. This pulse of energy is usually termed a signal. The radar measures the strength of the backscattered signal. By measuring the time interval between the transmitted and the received signal, it is possible to estimate the position of the hydrometeor. The higher the rainfall intensity, the greater will be the strength of the backscattered signal. The strength of the back-scattered signal

is usually expressed by 'reflectivity' ($Z$). From the reflectivity, it is possible to estimate the intensity of the rainfall at different points. After transmitting a pulse, the radar waits for a while to 'listen' to the returned signal and then it transmits another pulse. If the hydrometeors are not stationary, there will be a phase difference between the transmitted signal and the received signal. By measuring the phase shift, it is possible to estimate the radial velocity of the hydrometeor. This phenomenon is known as doppler effect.

A weather radar will usually make a series of rotations at different elevation angles. After completing one $360^{o}$ rotation, it increases its elevation angle and completes another $360^{o}$ rotation. In this way, it completes scanning in several elevation angles and a 3D volume scan of the atmosphere is done. It is thus possible to extract reflectivity map at any horizontal, vertical or inclined 'slice'. When a reflectivity map is displayed on a horizontal surface at a particular elevation, it is usually termed a Constant Altitude Plan Position Indicator (CAPPI). Figure 1 shows 1.5 km CAPPI for the Woodlands radar. The '1.5 km' indicates the elevation of CAPPI at the radar site. Although 'CAPPI' considers the effect of earth curvature and standard refraction, it does not consider local topography (Donaldson 2005, personal communication).

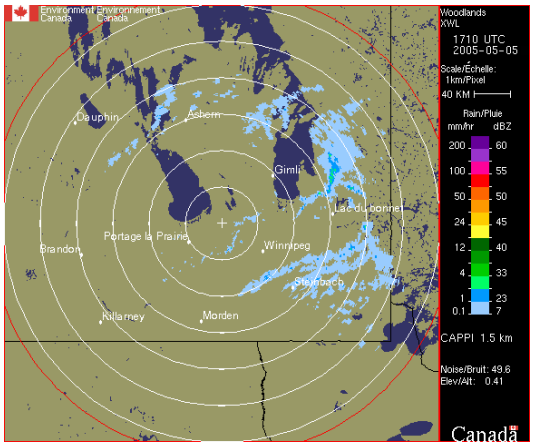

Figure 1: The 1.5 km CAPPI (Source: Environment Canada).

### 2.2.2 Radar Classification

Weather radars are usually classified according to their wavelength and frequency and termed C-band radar, S-band radar, etc. Table 1 shows a classification system.

Current weather radars can be divided into three broad categories, the conventional radar, the doppler radar, and the polarimetric radar. Polarization is the direction of the wave crest. The wave crest may lie either in a horizontal plane or in a vertical plane. The conventional doppler radar sends and receives only horizontal pulses of energy whereas the polarimetric radar transmits and receives both horizontal and vertical pulses. Hence, the doppler radar is a 'single polarization radar' and the polarimetric radar is a 'dual polarization radar'.

By analyzing both horizontal reflectivity ($Z_{HH}$) and vertical reflectivity ($Z_{VV}$), one can compute some other useful parameters such as differential reflectivity ($Z_{DR}$) and differential phase ($K_{DP}$) (Goddard et al., 1991). These parameters help to assess the size and shape of precipitation particles and thereby improve the quantitative measurements of rainfall.

Space-borne radar adds a new dimension in radar meteorology. The Tropical

Table 1: Classification of weather radars.

| Band | Wavelength (cm) | Frequency (GHz) | Remark |
|---|---|---|---|
| L | 15-30 | 1-2 | suitable for clear air turbulence study |
| S | 8-15 | 2-4 | not susceptible attenuation, suitable for long range weather study |
| C | 4-8 | 4-8 | susceptible to attenuation |
| X | 2.5-4 | 8-12 | suitable for cloud study |
| K | 1.7-2.5 | 12-18 | Similar to X band but more sensitive |

Rainfall Measuring Mission (TRMM) is jointly being carried out by the National Space Development Agency of Japan (NASDA) and the National Aeronautics and Space Administration (NASA). This mission carried the first space-borne precipitation radar to get the three dimensional structure of the precipitation for tropical and sub-tropical regions (Hiroshima, 1999).

### 2.2.3 The Canadian Radar Network

In Canada, a total of 31 radars have been installed so far mainly in the southern portion of the country. Figure 2 shows the location of the radars. These radars perform both 'conventional' scans and 'doppler' scans. They complete one scanning schedule in ten minute during which the first five minutes are used for the conventional scan and the last five minute are used for the Doppler scan. In conventional mode, no radial velocity is computed; in Doppler mode radial velocity is recorded. The conventional mode performs scanning in 24 elevation angles and the Doppler mode performs scanning in 4 elevation angles. All these radars are C-band radars except McGill University's S-band dual polarization radar.

Figure 2: The Canadian radar network (Source: Environment Canada).

## 2.3 Z-R RELATIONSHIP

During World War II, radar systems were extensively used for the first time to detect enemy aircrafts. It was observed that precipitation caused unwanted echoes in the radar scanning. Commander Maynard reported these weather echoes in a scientific journal in 1945 (Maynard, 1945). Bent (1946) undertook a comprehensive study of precipitation detection by radar and reported that 10-cm wavelength radar may be the most suitable for detecting precipitation. Wexler (1947) proposed an equation to get a quantitative estimation of rainfall intensity from radar echoes. Marshall et al. (1947) carried out extensive research on this topic and proposed the Z-R relationship. This relationship has been extensively used to estimate rainfall rate from radar reflectivity over the last four decades. Attempts have been made by researchers to develop other Z-R relationships. Battan (1973) compiled and reported 69 different Z-R relationships. The key steps of Marshall's derivation are described hereafter.

A radar measures the strength of the backscattered signal which is proportional to the reflectivity ($Z$). The reflectivity is a function of the drop-size distribution (DSD) of the rainfall event. Rain is made up by raindrops of different diameters and typical raindrop diameters vary from 0.2 to 5 mm (Ahrens, 1994). The DSD of a rainfall event is the histogram of raindrop diameters of which the X-axis contains the drop diameters ($D$) and the Y-axis contains the number of drops corresponding to that diameter. Using a Disdrometer, one can collect a rainfall sample from the ground and can calculate the drop-size distribution of the rainfall at a particular time. The disdrometer used by Marshall et al. (1947) was dyed paper exposed to the rainfall for a very short period. Raindrops when dropped on the paper produce a stain and provide information on the

drop-diameters and number of drops. Much more sophisticated disdrometers are available today. From the DSD, one can compute the reflectivity $Z$ using the following relationship,

$$Z = \sum_{i=1}^{n} N_i D_i^6 \qquad \text{Eq 2.1}$$

where $n$ is the number of drop-size bins in a sample volume, $N_i$ is the number of drops in the $i^{th}$ bin, $D_i$ is the diameter of the drops in the $i^{th}$ bin, and $Z$ is the measured reflectivity in $mm^6/m^3$.

Rainfall rate is a function of the drop-size-distribution and fall velocity, and can be computed as follows,

$$R = \tfrac{1}{6}\pi \sum_{i=1}^{n} N_i V_i D_i^3 \qquad \text{Eq 2.2}$$

where $R$ is the rainfall rate and $V_i$ is the fall velocity of the drop-sizes of the $i^{th}$ bin. Marshall et al. (1947) determined the DSDs for 135 samples and for each sample they calculated $Z$ and $R$ using Eq 2.1 and Eq 2.2, respectively. They got fall velocity information from the meteorological literature available at that time. Fall velocity is usually given as a function of the drop-diameter in the form $V=aD^b$, where $a$, $b$ are empirically derived coefficients, $V$ is the fall velocity, and $D$ is the diameter. For example, Atlas and Ulbrich (1977) reported the following equation to compute fall velocity,

$$V = 386.6D^{0.67} \qquad \text{Eq 2.3}$$

where $D$ is in meter and $V$ is in m/s. Curve-fitting of observed $Z$ and $R$ values produced the following relationship:

$$Z = 190R^{1.72} \qquad \text{Eq 2.4}$$

One year later, Marshall and Palmer (1948) proposed a statistical relationship to get the drop-size-distribution when the rainfall rate is known, based on the drop-size data they collected earlier for developing the Z-R relationship. This relationship is exponential and takes the form,

$$N = N_0 e^{-\Lambda D} \qquad \text{Eq 2.5}$$

where $N$ is the number of drops with diameter $D$ in a unit volume,

$N_o = 0.08$ cm$^{-4}$,

$\Lambda = 41R^{-0.21}$ cm$^{-1}$.

where $R$ is the rainfall rate in mm/hr. Using this fitted drop-size-distribution, Marshall and Palmer (1948) revised their work and found the following Z-R relationship,

$$Z = 220R^{1.60} \qquad \text{Eq 2.6}$$

Gunn and Marshall (1955) later revised their equation as follows,

$$Z = 200R^{1.60} \qquad \text{Eq 2.7}$$

The Z-R relationship expressed in Eq 2.7 has been extensively used for the last four decades.

In practice, one does not compute $Z$ from the DSD, rather it is measured by the radar. The power $P$ received by the radar is related to the reflectivity, $Z$, in the following way (Marshall et al., 1947):

$$P = P_o \frac{\pi^4 Ah(\eta^2 - 1)^2 Z}{8r^2 \lambda^4 \left(\eta^2 + 2\right)^2} \qquad \text{Eq 2.8}$$

where $P$ is the returned power, $P_o$ is the power transmitted by radar, $A$ is the effective area of the antenna, $h$ is the length of the wave train, $r$ is the range, $\eta$ is the refractive index, and $\lambda$ is the wavelength. As $P$ and other parameters are known, one can obtain $Z$ from Eq 2.8. Marshall et al. (1947) found fairly good agreement between $Z$ and $R$ measured from disdrometer data (Figure 3a). A time series of $P$ obtained from the radar and Z obtained from the disdrometer showed fairly close agreement (Figure 3b).

An important issue that should be mentioned is that Marshall et al. (1947) collected DSD samples at a range of 8.8 km from the radar location, which allowed beam elevations from 120 m to 350 m above the ground. This small elevation range should minimize the movement of rainfall mass due to wind and travel time required by the rainfall mass. Marshall et al. (1947) in their paper calculated that if the beam is located at an elevation of 2 km above the ground, then the rainfall mass should take 6 minutes to reach the ground in calm air and should horizontally drift 3 km to 9 km if a shower moves at a speed of 60 km/hr. They recommended taking 50 $km^2$ to 100 $km^2$ aerial averages when comparing radar signals and rate of rainfall at a range of 100 km or above

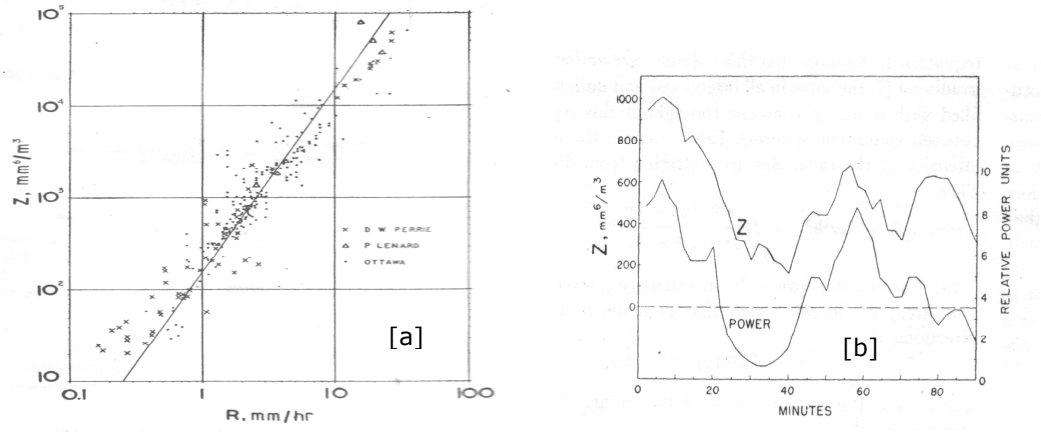


Figure 3: [a] *logZ-logR* association [b] *Z-P* association (Marshall et al. 1947).

Table 2: Z-R relationships.

| Relationship | Optimum for | Also recommended for |
|---|---|---|
| Marshall-Palmer ($Z=200R^{1.6}$) | General stratiform precipitation | |
| East-Cool Stratiform ($Z=130R^{2.0}$) | Winter stratiform precipitation - east of continental divide | Orographic rain – East |
| West-Cool Stratiform ($Z=75R^{2.0}$) | Winter stratiform precipitation - west of continental divide | Orographic rain – West |
| WSR-88D Convective ($Z=300R^{1.4}$) | Summer deep convection | Other non-tropical convection |
| Rosenfeld Tropical ($Z=250R^{1.2}$) | Tropical convective systems | |

Source: http://www.roc.noaa.gov/ops/z2r_osf5.asp

to compensate for the drift of rainfall.

The CAPPI images are usually presented in terms of *dBZ* or 'decibel of reflectivity' which is the logarithmic transformation of *Z* and can be obtained using the following equation,

$$dBZ = 10 \log_{10} Z \qquad \text{Eq 2.9}$$

Environment Canada is currently using this relationship for light rain to thunderstorm. Different Z-R relationships for stratiform and convective and other types of rainfall events are suggested. Table 2 provides a list of Z-R relationships recommended by NOAA. The equation $Z=300\ R^{1.4}$ is termed the 'standard NEXRAD' *Z-R* relationship as it is believed that this relationship is a compromise between convective and stratiform rainfall (Krajewski and Vignal, 2000).

While determining the drop-size-distribution in Marshall et al. (1947)'s experiment, the maximum measured rainfall intensity was 35 mm/hr and this equation produced 100 mm/hr rainfall at a *dBZ* equal to 55. However, in such intense rainfall the radar signal may be highly attenuated (especially in case of C-band radars) and

reflectivity may be much lower than that. In such cases, a Z-R relationship which measures intense rainfall at relatively low *dBZ* may be useful.

Figure 4 shows the reflectivity and corresponding rainfall rate of three widely used Z-R relationships (Marshall-Palmer, WSR-88D convective, and Rosenfeld tropical). This figure shows that all three relationships yield the same rainfall intensity up to 35 dBZ (about 5 mm/hr). The Marshall-Palmer and WSR-88D convective measures the same rainfall up to 42 dBZ (15 mm/hr). Hence, for light to moderate rainfall, there is no real difference between these equations. Several classification systems of rainfall events based on reflectivity is available in the literature. One such classification is available on the link http://grappa.meteo.mcgill.ca/mcgill_img.html.

## 2.4 ERRORS AND UNCERTAINTIES

### 2.4.1 General

Radar derived rainfall usually does not provide good correlation with raingauge measurements. Seo et al. (2000) compared radar and raingauge data for 11 convective

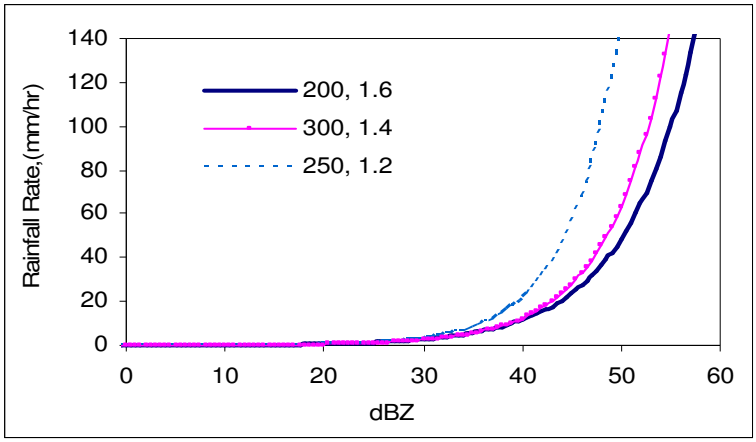


Figure 4: Rainfall rate and corresponding reflectivity.

storms and 12 stratiform events. Their study concluded that radar overestimates convective rainfall by 72% on the average and underestimate by 91% in case of stratiform events. Their study used $Z=300R^{1.4}$ for conversion from reflectivity to rainfall rate. Legates (2000) reported that the standard *Z-R* relationship overestimates light precipitation and underestimate heavy precipitation. Austin (1987) found that radar underestimates light widespread rainfall and overestimates intense convective storms. Smith and Krajewski (1991) worked on the hourly radar data of the Norman, Oklahama, radar and found the bias varies within a range of 1.6 to 2.5 for all hours. The word 'bias' means the ratio of the summation of the gauge rainfall to the summation of the radar estimated rainfall. Anagnostou et al. (1998) compared the Tulsa, Oklahama, radar data with 240 raingauges and found that correlation coefficients between radar and raingauges vary within a range of 0.3 to 0.95. Baeck and Smith (1998) reported that significant underestimation of rainfall occurs at ranges greater than 100 km and at ranges smaller than 40 km and overestimation occurs within the intermediate range.

It is interesting to mention that the discrepancy reported by various authors also shows wide variations. While some authors reported that radar overestimates convective precipitation, others have reported underestimation in such case. Krajewski and Smith (2002) concluded that "....there is much that we do not understand about the instrument that has been in use for over 40 years". The reason for the inconsistencies in conclusions may be due to the difference in data processing and experimental scenarios carried out by different researchers. For example, if one does not apply correction for attenuation, radars may underestimate peak thunderstorms whereas a different result may be obtained if a correction for attenuation is employed for C-band radars. If one works with gauges

located close to the radar station and use the lowest elevation angle, one may get better correlation due to the reduction in the error caused by the wind drift.

In general, three factors may be held responsible for the discrepancies between gauge and radar measurement, (i) errors in radar estimation (e.g. anomalous propagation, attenuation, ground clutter, updraft-downdraft, variations in the drop-size-distribution etc.), (ii) errors in the gauge estimation (e.g. wind), and (iii) errors while comparing the two sensors (difference in temporal and spatial sampling, wind-drift, etc.). Even if one can eliminate all possible errors in the radar and the gauge measurements, one should not expect perfect correlation between gauge and radar measurement due to the third factor mentioned above. These factors are discussed briefly below.

### 2.4.2 Sampling Error

Radar measures the aerial average rainfall over several square kilometres whereas a gauge measures point rainfall. Radar takes a snap shot of precipitation whereas a gauge measures the rainfall accumulation. These sampling issues may hinder good correlation between radar and raingauge estimation (Jayakrishnan et al., 2004; Ciach and Krajewski, 1999; Seed et al., 1996). Figure 5 shows 24 revolutions at different elevation angles in a complete volume scan produced in five minutes. It takes approximately 10 seconds to complete one $360^0$ revolution at a particular elevation angle. If a raingauge is located 50 km from the radar location, it may be assumed that the seventh revolution provides the rainfall information over this raingauge location for the 1.5 km CAPPI. Hence, the 1.5 km CAPPI data provides rainfall information after $1.1^{th}$ minute (10 seconds multiplied by 7) of each 5-minute cycle. On the other hand, a raingauge provides the total rainfall that occurred in the 5-minute time period at that point. If the rainfall intensity varies

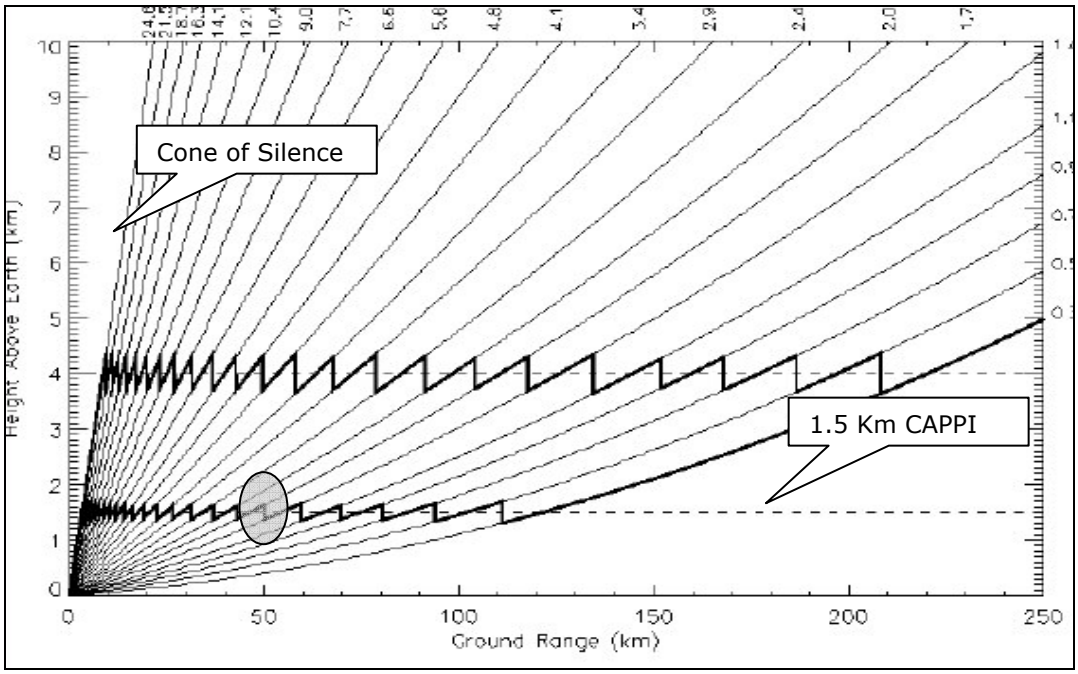

Figure 5: Radar beam elevation angles (Source: Environment Canada).

appreciably within the 5-minute time bin, it should result in a discrepancy between gauge and radar data even if both sensors measure rainfall 100% accurately. In fact each region of a CAPPI reflects rainfall rate at a different time, so this time difference may vary from several seconds to several minutes. Krajewski (2004) concluded that "one should not expect high correlation between raingauge observed rainfall and radar-estimated rainfall at short time scales".

Kitchen and Blackball (1992) computed the error structure between two raingauges located below the same radar pixel having 3 kmx3 km spatial resolution. They found that the standard deviation of *log(r/g)* for a gauge located at the center of the radar pixel and another gauge located at 1.2 km from the center is 0.23 and 0.3, respectively, where *r* and *g* are the two-minute rainfall accumulations of radar and gauge, respectively. When the same analysis was carried out with one-hour rainfall accumulations, the

standard deviation for these gauges was found to be 0.14 and 0.24, respectively. This implies that the variance is reduced when accumulation time is increased.

Time-space synchronization often creates problems as radars measure aerially-averaged rainfall 1-2 km aloft and there may be significant time displacement (up to several minutes) between radar and gauge measurement (Krajewski, 2004).

### 2.4.3 Radar Hardware Errors

The signal received by the radar is associated with white noises due to thermally induced random motion of electrons in the detectors of the receivers and due to atmospheric and cosmic radiation (Doviak and Zrnic, 1993, pp 97). Although algorithms based on Fourier transformation and other empirical relationships are used to account for these errors in radar hardware, a precision error of 1 dBZ is accepted for WSR-88D radars. Hunter (2005) pointed out that due to the exponential formulation of Z-R relationships, this small error in reflectivity should cause significant error in rainfall estimation. Figure 4 shows rainfall rate obtained for reflectivity in the range of 10 to 60 dBZ using three common Z-R relationships. This figure shows that rainfall doubles for an increment of reflectivity of only 5 dBZ with the $Z=200R^{1.6}$ relationship. This implies that one dBZ error in radar hardware is not negligible.

### 2.4.4 Drop-Size Distribution

It has been noted that two different drop-size-distributions may produce the same reflectivity (Hunter, 2005). For example, in a 1 $m^3$ volume, 4096 drops each having 1 mm diameter will produce the same reflectivity as produce by 1 drop having 4 mm diameter (NOAA website). This reflectivity should be equal to 4096 $mm^6/mm^3$ (using Eq

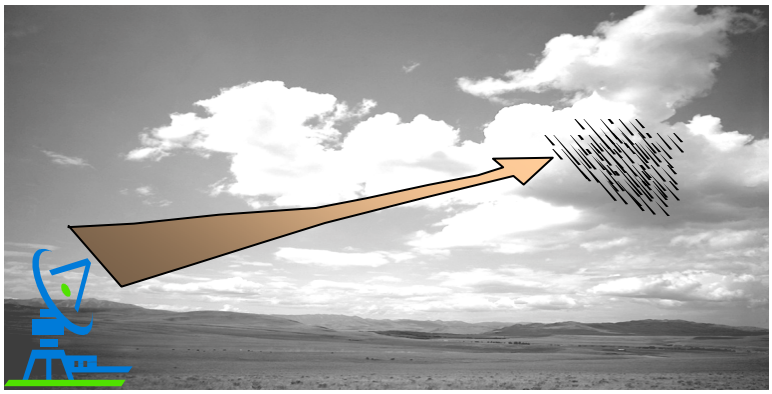

Figure 6: Demonstration of Virga.

2.1) or 36 dBZ. Fall velocity and the rainfall rate for these two DSDs can be computed using Eq 2.2 and Eq 2.3. It is found that rainfall rate for these two cases are equal to 29.15 mm/hr and 1.15 mm/hr, respectively.

Gunn and Marshall (1955) showed that the drop-size-distribution may also vary along a vertical column of the atmosphere due to the wind shear. This should also lead to different rainfall measurements in the atmosphere and on the ground.

### 2.4.5 Virga

Sometimes atmospheric precipitation evaporates before reaching the ground. This rainfall will be recorded by the radar, but will not be recorded by the raingauge. This is a potential source of error known as 'virga'. This is more likely to occur in the case of light precipitation associated with low humidity in the atmosphere (Austin, 1987). This phenomenon is graphically shown in Figure 6.

### 2.4.6 Anomalous Propagation

Due to temperature or humidity inversion in the atmosphere, radar signals may be bent down and reflected from the ground (MSC, 2004). This type of refraction is known as 'super-refraction'. This might erroneously be interpreted as intense precipitation and is

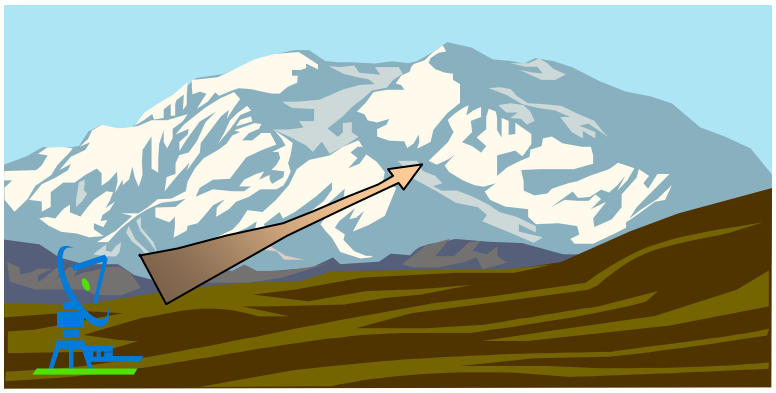

Figure 7: Demonstration of beam blockage.

also known as 'anomalous propagation'. Alberoni et al. (2001) reported that anomalous propagation may cause *dBZ* values up to 50-60. It is also well-understood that during thunderstorms, updraft-downdraft hinders the possibility of developing a stable temperature/moisture gradient in the atmosphere, making it susceptible for anomalous propagation (Raghaban, 2003, p. 59).

However, in the wake of the thunderstorm, due to downdraft cold air may spread in the lower atmosphere keeping warm air in the higher altitude. This situation may cause anomalous propagation (Austin, 1987).

### 2.4.7 Ground Clutter and Beam Blockage

Radar signals may be partially intercepted by ground targets (e.g. buildings, towers) and this source of error is known as 'ground clutter' (MSC, 2004). If the beam is completely blocked by mountains, it is known as 'beam blockage' as shown in Figure 7.

### 2.4.8 Attenuation

Attenuation is the weakening of the radar beam when it passes through hydrometeors. Radar calibration system possesses correction for gaseous attenuation, but it does not provide any correction for attenuation due to rainfall or wet radome. Austin (1987)

reported that for radars with 10 cm wavelength (S-band), attenuation caused by rainfall or wet radome should be negligible. For example, a 50 dBZ rainfall located at a range of 110 km will be attenuated 1 dBZ in case of an S-band radar (Doviak and Zrnik, 1984). Gunn and East (1954) provided the following relationship to quantify attenuation,

$$k_R = 2\int_{r=0}^{r_2} cR_r^d dr \qquad \text{Eq 2.10}$$

where $k_R$ is the two-way attenuation in dBZ and $R_r$ is the rainfall rate in mm/hr at a range $r$. For C-band radars, $c$=0.0022 and $d$=1.17. To compute attenuation at a range $r_2$ one needs the spatial distribution of rainfall along the radar beam up to the range $r_2$ which makes it computationally intensive. The computed attenuation at a range should be added to the measured reflectivity at that range.

### 2.4.9 Bright Band Contamination

Melting snow below the zero degree isotherm altitude causes an enhancement of reflectivity. The line of enhanced reflectivity in the vertical profile is commonly known as the bright band. Fabry and Zawadzki (1987) reported that the bright band can cause 5-10 dB enhancement of reflectivity which translates to rainfall rates five times greater than the actual one.

### 2.4.10 Beam Spreading

The width of a $1^o$ beam increases with range. The width of a $1^o$ beam is 0.9 km at a range of 50 km and 3.5 km at a range of 200 km. Hence, at far range the radar resolution gradually decreases. This spreading may weaken the bright band (Fabry et al., 1992) and may cause problems in the detection of severe weather (Hunter et al., 1993).

### 2.4.11 Higher Altitude Beam

Far away from the radar (>110 km), even the lowest elevation angle may overshoot precipitation or may intercept snows or snow-rain mixtures. In such cases, the radar may produce erroneous precipitation. The earth's curvature may worsen this problem.

Kitchen and Jackson (1993) considered overshooting as the major cause of radar estimation error at far range. They reported that radars underestimate rainfall by a factor 10 at far range. Chumchean et al. (2003) computed error variances as a function of range and found that beyond a distance of 55 km from the radar location, error variance gradually increases. This increase in error variance occurs not only for overshooting, but also for bright band contamination, beam spreading, and other range dependent causes.

Faisal et al. (2004) computed accumulated rainfall for radar scans at two different CAPPIs (the 1.5 km CAPPI and the 4.5 km CAPPI). Their study concluded that the radar hyetograph for the 1.5 km CAPPI showed much closer agreement with the gauge hyetograph compared to the 4.5 km CAPPI. The authors recommended using radar scans as close to the ground as possible. One reason for getting inferior result with higher altitude scans may be that the higher altitude CAPPI may exceed the zero degree isotherm or is affected by the brightband.

### 2.4.12 Extreme Rainfall

Blanchard and Spenchar (1975) found that the median diameter of raindrops remains constant for rainfall events varying from 100 mm/hr to 700 mm/hr. In such cases, drop growth is balanced by the drop brake up. List (1991) theoretically proved that in such cases a linear Z-R relationship will exist, that is, the power of $R$ should be one in the Z-R

relationship. If it is true then conventional Z-R relationship may estimate erroneous rainfall rate at this range.

### 2.4.13 Wind Drift

Radar meteorologists usually compare gauge rainfall with the radar pixel located above the gauge location. However, wind may cause horizontal displacement of rainfall mass and the radar pixel located directly above the gauge may not represent the corresponding gauge rainfall. Gunn and Marshall (1954) reported that larger particles will be exposed to less horizontal displacement than smaller particles. They proposed a simplified approach for computing particle displacement due to wind drag which is as follows:

$$x = \int_0^z \frac{u(z)}{w} dz \qquad \text{Eq 2.11}$$

where $u(z)$ is the x-component of the wind velocity at elevation $z$, and $w$ is the vertical velocity. The y-component of the drift can be computed using a similar approach.

This procedure assumes that the drift of hydrometeors at any level is equal to the wind velocity at that level. This may be true for snow flakes or very tiny hydrometeors, but may not be true for raindrops having larger diameter. Another limitation of Eq 2.11 is that the vertical velocity is assumed constant while in reality it may vary due to up-draft and down-draft.

Harrold et al. (1974) measured the advection distance to be about 1-2 km if rain drop has a fall speed of 5 $ms^{-1}$ and radar elevation is 500 m above the ground level. Dalezios and Kouwen (1990) found the advection distance close to the distance measured by Harrold et al. (1974). If one works with the 1.5 km CAPPI, the advection distance

could be higher than 1-2 km. Collier (1999) worked on wind drift in a small town in the UK where strong secondary flows were observed due to a large difference in sensible heat fluxes between the city and the surrounding rural areas. Due to the secondary flow system, surface winds could be in the opposite horizontal direction of winds 500 m aloft. The author stressed the need for real time velocity measurement in urban areas and the necessity to account for wind drift in high resolution radar data especially in urban areas. Ignoring the complexity of the wind profile, the authors proposed some simplified graphs to measure wind drift assuming wind velocity varies linearly or remain constant with increasing elevation. The author concluded that "it appears that for very small urban catchments real-time measurements of wind-profile are necessary to identify the wind drift, and if necessary, allow for it". This statement appears reasonable as wind changes speed and direction in the vertical due to various meteorological factors and estimating drift from a simplified graph may be misleading.

Lack and Fox (2004) concluded that "wind drift has been identified as a problem, but it is never addressed as something that leads to large errors in estimating surface rainfall fields or as an error that could be corrected."

It should be noted that, according to Mittermaier et al. (2004), "Although such movements are important for gauge calibrations, they are less significant for operational radar which produces rainfall estimates with a resolution of 1-2 km for use in hydrology."

### 2.4.14 Cone of Silence

The cone of silence is usually defined as the area very close to the radar site (Figure 5). A radar scans to a certain elevation angle but does not usually reach $90^{o}$. As its elevation

angle does not reach $90^0$, regions vertically above the radar will not be available at CAPPI (Mizzell, 1999).

### 2.4.15 Updraft-downdraft

Battan (1976) reported that updraft causes overestimation of radar measured rainfall, whereas downdraft underestimates rainfall. Austin (1987) reported that downdraft underestimates rainfall by 15-50%. Mizzell (1999) worked with three convective rainfall events and found that two of them overestimated gauge rainfall whereas one underestimated. He reported updraft-downdraft as the prime reason behind this discrepancy. Updraft causes an upward movement of hydrometeors and thereby increases the concentration of hydrometeors which may lead to enhanced reflectivity. On the other hand, downdraft will increase the fall velocity. Rainfall rate measured at the ground is a function of drop-size-distribution and fall velocity of hydrometeors whereas radar reflectivity is a function of the drop-size-distribution only (not the fall velocity) (Eq 2.1 and Eq 2.2). For this reason updraft will cause the radar to overestimate rainfall rates whereas downdraft will cause underestimation by radar.

### 2.4.16 Hail Contamination

Hail is often the cause of enhanced reflectivity. A basic assumption of Z-R relationships is that the wavelength of the radar signal is far greater than the diameter of the hydrometeors (Rayleigh approximation). This assumption is valid for raindrops which have typical diameters of 2 mm, much less than the radar wavelength (5-10 cm), but is not valid for hail which may have diameters as large as 2 cm. This and a number of other factors are the reason for getting higher echoes from hail storms. Hunter (2005) suggested

to discard reflectivities above 53 dBZ, considering that reflectivity beyond this would be due to hail.

### 2.4.17 Gauge Errors

Gauge readings are not considered accurate when strong winds are present. The error associated with wind/turbulence is usually not more that 5%, but during severe thunderstorms, errors may be as high as 40% (Wilson and Brandes, 1979).

The tipping bucket gauge may also underestimate heavy rainfall due to its tipping process and underestimate light rainfall due to evaporation (http://www.criacc.qc.ca/villes/term_e.html).

## 2.5 IMPROVEMENT STATEGIES

### 2.5.1 General

Research has been carried out for the last four decades to improve the quality of radar data products. This research may be divided into several categories which are discussed below.

### 2.5.2 Removal of Systematic and Random Errors

Jordan (2000) and Viux (2002) attempted to classify radar-gauge errors in the two broad categories (i) systematic errors and (ii) random errors. Systematic errors are associated with specific problems such as incorrect Z-R relationships, which causes systematic underestimation/overestimation by radar and which can be removed by adding a multiplicative constant in the Z-R relationship. For example, if one measures length with a wrongly calibrated measuring tape, it will systematically underestimate/overestimate

measurements but one can easily correct errors in measured lengths after identifying the error in the measuring tape. To eliminate the systematic error, Wilson and Brandes (1979) proposed to compute multiplicative 'mean field bias' using radar and gauge measured rainfall. The 'mean field bias' ($B$) is computed by the following equation:

$$B = \frac{\sum_{i=1}^{n} G_i}{\sum_{i=1}^{n} R_i} \qquad \text{Eq 2.12}$$

where $G$ is the gauge rainfall, $R$ is the radar estimated rainfall, and $n$ is the number of observations. The calculated bias is then multiplied with Z-R relationship as follows:

$$R_b = B\left(\frac{Z}{200}\right)^{1/1.6} \qquad \text{Eq 2.13}$$

where $R_b$ is the bias corrected rainfall rate.

Borga and Tonolli (2002) suggested taking daily accumulations when computing bias to avoid uncertainty associated with the temporal and the spatial mismatch.

Hydrologists have also tried to update the bias in real time by assuming it follows an Autoregressive (AR1) process. The estimated time-varying bias can then be corrected using a Kalman filtering approach. Chumchean et al. (2004) proposed a correction scheme of real time radar rainfall bias using the Kalman filtering approach. In their model, the bias is estimated using an AR1 model and if new raingauge data become available, the model process variance is updated using Kalman filter techniques. The lag-one correlation coefficient of the mean field bias and the stationary variance of the logarithmic bias process is estimated through trial and error in the calibration process. It is observed that Chumchean's model is not effective in predicting biases of large (greater

than 3) or small (less than 0.8) magnitudes. It is also noteworthy that the authors discarded biases greater than 5 and smaller than 0.2 during their calibration considering these to be outliers. Earlier attempts with Kalman filtering were done by Krajewski and Smith (2002), Anagnostau et al. (1998), and others.

Rosenfeld at al. (1994) attempted to apply probability matching techniques to correct radar data using gauge data. This method assumes that the probability density function of radar derived rainfall is the same as that of the gauge derived rainfall. The parameters of the *Z-R* relationship are determined so as to satisfy this criterion.

Even after removing systematic errors, radar measured rainfall may contain residual random errors due to updraft-downdraft, real time variation of the drop-size-distribution, sampling errors, etc. Vieux (2002) suggested quantifying the residual random errors using the following relationships:

$$\varepsilon = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{G_i - Rb_i}{G_i} \right| \qquad \text{Eq 2.14}$$

where $G$ is the gauge rainfall, $R_b$ is the bias corrected radar rainfall, and $\varepsilon$ is a parameter measuring the residual error.

### 2.5.3 Time Integrations

As mentioned earlier, radars take a snap shot of precipitation while gauges measure accumulation of precipitation. When comparing the two sensor types, development of a suitable algorithm to assess accumulated rainfall from instantaneous rainfall is important. The usual practice is to assume rainfall intensity to be constant throughout the time step (say, 10-minute) in case of radar data (Hannesen, 2002). This may lead to a discrepancy

between radar and gauge data in case of fast moving convective cores. To solve this problem, it is necessary to estimate the spatial rainfall patterns in between two scans. This requires tracking a convective core between two consecutive scans. Once one is able to identify the spatial movement of a convective core, it is easy to estimate its position and shape at an intermediate time. For tracking, a cross-correlation based technique has been described by Li et al. (1995) and Bellon and Austin (1978). Austin (1987) tracked the convective cores by manual observation.

One problem in the tracking of convective cells is the formation of new cells in the vicinity of existing cells and disappearing of existing cells. Downdraft from existing cells in their mature stage triggers new cells in the surrounding areas in a preferred direction (Raghaban, 2003, p. 189). This behaviour may impose erroneous result in the cross-correlation algorithm.

### 2.5.4 Merging Techniques

There has been a significant interest in combining radar and raingauge estimated rainfall optimally using statistical techniques such as Bayesian methods, Co-kriging, and multivariate regression. The objective of such techniques is to take radar estimated precipitation and gauge data and generate a new precipitation field combining information obtained from the two sources.

Co-kriging based radar and gauge data merging techniques was explored by Krajewski (1987). Co-kriging is an interpolation tool that combines two sources of data in a way that minimizes the error variance. It determines the interpolation weights based on the spatial correlation. As rainfall has a high degree of spatial correlation, co-kriging

appears to be an appropriate tool for merging radar and gauge data. Todini (2001) attempted to develop a model based on combining block kriging and Kalman filtering in a Bayesian framework. The author added a known noise to signal ratio of 30% as the true error structure is not known. The author reported that his approach could work well in eliminating noise. This approach was developed under the framework of the MUSIC (Multi Sensor Precipitation Measurements Integration Calibration and Flood Forecasting) project funded by the EU. Lee (2004) attempted to combine gauge and radar data for the City of Winnipeg with a co-kriging approach. In this city, 24 raingauges are located in an area of approximately 18 km x 22 km. Lee (2004)'s result showed that with this dense network of gauges, the performance of gauge interpolation is close to the co-kriging based radar-gauge merging approach.

Sokol (2003) proposed a regression model to combine radar and rainfall estimates. His model interpolates raingauge data over the catchment and takes radar rainfall and interpolated gauge rainfall as two input vectors in the regression equation. Gauge data is interpolated based on the Euclidian distance from gauge locations.

Bayesian methods for combining radar and gauge observations have been considered by several authors. These methods attempt to combine different sources of information optimally by quantifying the uncertainties in terms of probability distributions (Harouche and Rasmussen, 2002).

One common problem with the various merging techniques is that they will not work well far from the gauge locations. In such cases, the model output will reflect the

radar estimation since weights from the gauges may be insignificant at far distance. Despite this fact, merging is an appealing way to get improved spatial precipitation data.

### 2.5.5 Hydrograph Simulation

As neither radar nor raingauge can simulate the true rainfall structure, hydrograph generation based on radar, raingauge, or a combination of the two has been attempted by many researchers. The model generated hydrograph is then compared with the actual measured hydrograph. The underlying assumption behind this approach is that the most accurate estimation of the precipitation system will generate the most accurate hydrograph.

Todini (2001) attempted to generate hydrographs for a gauging site of the Reno river. His simulation showed that during part of the simulation period the radar significantly underestimated peak discharge whereas a merging approach could estimate the peak discharge more successfully.

Borga (2002) simulated hydrographs based on radar-derived data at different elevation angles. His study showed that the simulated hydrograph is more accurate for lower scans compared to that of higher scans. He also obtained better results with bias adjusted radar data than with unadjusted radar data.

Sun et al. (2000) carried out hydrograph simulation based on radar derived rainfall, gauge interpolated rainfall using kriging, and a combination of gauge and radar data based on co-kriging. The authors reported that radar precipitation causes large amplitude error and gauge interpolated precipitation failed to simulate one peak in the hydrograph. The overall best agreement was obtained with the co-kriging approach.

However, hydrologic rainfall-runoff models also have inherent errors and uncertainties and results will depend on model calibration.

### 2.5.6 Other Reflectivity-Rainfall Relationships

Some researchers have explored Artificial Neural Networks (ANN) and Fuzzy logic to estimate rainfall rates from radar reflectivity. Xiao and Chandrasekar (1997) may have been the first to explore the performance of neural networks in this area. They emphasized the inherent robust nature of ANN's, in particular that errors at a few input neurons will not undermine the overall performance of the model. As radar data have different sources of errors which are not easy to remove, the application of neural networks in this field seems appealing. Xiao and Chandrasekar (1997) used a multilayer feed forward back propagation neural network in their study. Out of 20 raingauge stations, they used data of six raingauge stations for training the network while the remaining 14 gauges were used for validation. Their experiment showed a 17% error for two-day rainfall accumulations and a 12% error for four-day accumulations. However, a good agreement for two or four day accumulations does not guarantee a good agreement for high resolution data (e.g. 10-minute accumulations). Another limitation is that they did not report model output for rainfall events that were not used for the training. If the ANN model does not work well for a new rainfall event, the network will need to be trained further using both old and new data sets for several raingauges.

Xiao and Chandrasekar (1998) developed a radial basis function (RBF) neural network for snowfall estimation. This work employs a training data set from one gauge and validates the model using data from two other gauges for the same snowfall events. Liu et al. (2001) developed an adaptive radial basis function (RBF) neural network

algorithm for estimating rainfall from radar. The network is trained by the initial data and thereafter can be further trained any time new raingauge data become available. This implies that when new gauge data become available, the RBF neural network model will not be retrained with the whole data set (combining both old and new data), rather it will add or replace the neurons and change the center vector to account for the new data only. The author mentioned that "This procedure ensures that the new data have higher priority in the determination of weights of the modified network". However, if the new data are associated with a high degree of error, it may deteriorate the model performance for estimating a new rainfall event.

For example, if radar data are contaminated with serious attenuation in an intense thunderstorm (for C-band radars), and a model fitted with this attenuated radar data is used to estimate another storm in which attenuation is not significant, the model is likely to appreciably overestimate the rainfall. An adaptive scheme that updates the model using gauge information for each and every rainfall should be useful. Alternatively one could train the model frequently with a huge ensemble of data.

Hessami et al. (2003) proposed an adaptive neuro-fuzzy system for post-calibration of weather radar using CAPPIs of one-hour rainfall accumulations and corresponding gauge data. The author used subtractive clustering to generate the initial fuzzy inference system. However, the need to improve 1-hour rainfall accumulations using ANN may not be critical for hydrologists because radar and gauge show sufficiently better correlation with hourly accumulation data using conventional Z-R relationship.

Hessami et al. (2004) compared the performance of a feed forward error back-propagation network and a radial basis network. For the back-propagation network, the authors tried five different training algorithms, quasi Newton, one-step secant, resilient back propagation, Levenberg-Marquardt method and Levenberg-Marquardt algorithm with Bayesian regularization. They found the back-propagation network with the Levenberg-Marquardt algorithm using Bayesian regularization most suitable for post-calibration of radar data. However, the authors used all their 15 rain gauge data sets for training and they compared the nets based on the quantitative performance (correlation and rmse) in the training sessions. They did not carry out any quantitative comparison of model performance on unused raingauge data or on new rainfall events.

Other related work associated with neural networks includes the development of the detection of rain/no-rain conditions based on a radial basis function neural network (Chandrasekar and Gorgucci, 2001), and on Kohenen neural network (Xiao et al., 1998).

### 2.5.7 Clutter and AP Correction Algorithms

Several correction algorithms are available in the literature to remove errors such as ground clutter and anomalous propagation.

Moszkowicz et al. (1993) proposed a statistical method to detect and eliminate anomalous propagation from radar reflectivity. Their study included statistical pattern classification techniques for detecting such echoes. Bech et al. (1998) used mesoscale Numerical Weather Prediction (NWP) model data to determine the atmospheric condition likely to produce anomalous propagation. Grecu and Krajewski (2000) developed a neural network based algorithm to detect anomalous propagation. Pamment and Conway

(1998) used Bayesian statistics to determine the probability of an echo being due to anomalous propagation.

Torres and Zrnic (1999) developed an algorithm based on regression filters to remove ground clutter. Sachidanada and Zrnic (2000) developed an algorithm that use Fourier transform and magnitude deconvolution procedures to remove clutter. Steiner and Smith (2002) used three dimensional reflectivity profiles to determine the presence of anomalous propagation. The vertical extent, vertical gradient, and spatial variability of reflectivity are assessed by means of a 'decision tree' in their algorithm.

Cheng and Collier (1993) and Kitchen et al. (1994) proposed an algorithm to correct bright band signals. This algorithm detects the bright band on the basis of a significant increase in reflectivity in the vertical profile and then replaces it with background reflectivity. However, Seo et al. (2000) reported that a thick bright band with mild enhancement in reflectivity will not be detected. A detailed description of these algorithms is beyond the scope of this study.

### 2.5.8 Multiparameter Estimation

Dual polarization radars provide reflectivity in both a horizontal ($Z_H$) and a vertical plane ($Z_V$), as discussed in Section 2.2.2. Based on these variables, a new precipitation classification system has been proposed where rainfall is classified as drizzle, rain, snow, graupel, etc. if several atmospheric criteria are met.

Table 3 shows a microphysical classification system proposed by May and Keenan (2003). In this classification system, $Z_{DR}$ is the ratio of the horizontal and the vertical reflectivity, often termed the differential reflectivity, $K_{Dp}$ is the rate of change of

differential phase in propagation, and $\rho_{HV}$ is the correlation coefficient. Fuzzy logic is usually applied to combine polarimetric estimators.

### 2.5.9 Convective/Stratiform Separation

As reflectivity is a function of the DSD (drop size distribution) and DSD depends on rainfall type, the classification of rainfall and the associated Z-R relationships may improve radar measurements. Uijlenhoet et al. (2003) measured drop-size-distributions while a squall line passed over the Northern Mississippi region and reported significant variation in mean drop diameters, standard deviation of drop-sizes, and drop concentrations between convective, stratiform and transition part of the storm and reported different Z-R relationships suitable for these conditions. The classification algorithms are based on the analysis of the vertical profile of reflectivity (VPR), except for the algorithm proposed by Steiner et al. (1995). To calibrate/validate an algorithm researchers often use the TRMM space borne radar precipitation products.

The main focus of this classification system is to separate convective and stratiform region in a radar scan. Stratiform rainfall may results from mid-latitudes frontal systems, especially when a warm front moves from the south/south-west direction. It may also be caused by wide-spread updraft due to upslope flow. Convective systems may develop due to the movement of a cold front from the west/north-west direction (frontal convection) or updraft due to any local uneven heating of the earth's surface due to solar radiation (Stull 1995).

Recent studies show that convective cells may be found in stratiform rainfall because updraft destabilizes the atmosphere and stratiform rainfall may also occur in a

Table 3: Microphysical precipitation classification system (May and Keenan, 2003)

| | $Z_{HH}$ (dBZ) | $Z_{DR}$ (dB) | $\rho_{HV}(0)$ | $K_{DP}$ (deg $km^{-1}$) | Temperature (°C) | Classification Number in File |
|---|---|---|---|---|---|---|
| Drizzle | 10-25 | 0.2 to 0.7 | > 0.97 | 0 to 0.06 | > -10 | 1 |
| Rain | 25 to 60 | 0.5 to 4 | > 0.95 | 0 to 20 | > -10 | 2 |
| Snow (Dry, Low Density) | -10 to 35 | -0.5 to 0.5 | > 0.95 | -1 to 1 | < 0 | 3 |
| Snow[(a)] (Dry, High Density) | -10 to 35 | 0.0 to 1 | > 0.95 | 0 to 0.4 | < 0 | 4 |
| Snow (Wet, Melting) | 20 to 45 | 0.5 to 3 | 0.5 to 0.9 | 0 to 1 | 0 to 5 | 5 |
| Graupel, Dry | 20 to 35 | -0.5 to 1 | > 0.95 | 0 to 1 | < 0 | 6 |
| Graupel, Wet | 30 to 50 | -0.5 to 2 | > 0.95 | 0 to 3 | -15 to 5 | 7 |
| Hail, Small < 2cm Wet | 50 to 60 | -0.5 to 0.5 | 0.92 to 0.95 -0.95 | -1 to 1 | -15 to 5 | 8 |
| Hail, Large > 2 cm Wet | 55 to 65 | -1 to 0.5 | 0.90 to 0.92 | -1 to 2 | -25 to 5 | 9 |
| Rain and Hail | 45 to 80 | -1 to 6 | > 0.9 | 0 to 20 | -10 to 10 | 10 |
| (a) Rimed and aggregated snow. | | | | | | |

mesoscale convective system in its mature and decaying stage (Houze, 1993). In the mature stage of the mesoscale convective system (MCS), stratiform rainfall may cause 10-50% of the total precipitation amount.

Parker and Johnson (2000) classify mid-latitudes MCS in three broad categories: (i) leading stratiform where the stratiform umbrella is observed in front of the convective line, (ii) trailing stratiform, where the convective line leads the stratiform umbrella and (iii) parallel stratiform where both regions lie in a line. Their study reveals the co-existence of convective and stratiform regions in mid-latitudes precipitation systems.

Figure 8 shows the movement of a cold front and warm front over Winnipeg. The cold frontal precipitation shows a parallel stratiform system. It is observed that both frontal systems show a wide range of echoes.

Stratiform rainfall usually contains smaller drop diameters and convective rainfall usually consists of larger drop diameters. As noted earlier, convective rainfall exhibits sharp spatial and temporal gradients of precipitation. Based on these observations, Steiner et al. (1995) proposed an algorithm to separate convective and stratiform regions. His algorithm is also known as the SHY algorithm.

This algorithm works with 2D PPI's. It employs three criteria to determine whether a point is convective or not. A pixel is considered convective if at least one of the following criteria is satisfied: (i) reflectivity greater than 40 dBZ, (ii) the surrounding pixels are convective, (iii) the difference between reflectivity of a pixel and reflectivity of its background exceeds a given threshold. This background reflectivity is determined by the linear average of the nonzero reflectivity of an area of 11 km radius with the center located at the pixel. If none of the above criteria is satisfied, the pixel is considered stratiform.

The TRMM classifies rainfall based on the existence of a bright band. If a bright

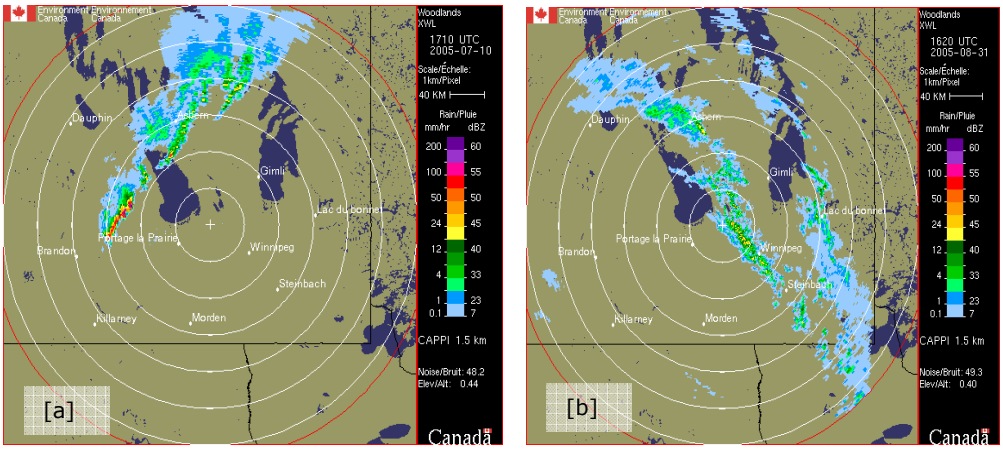

Figure 8: Movement of a [a] cold front and [b] warm front over Winnipeg

band is detected in the vertical profile, rain is considered stratiform; if not and if reflectivity exceeds a threshold, it is considered convective (http://trmm.gsfc.nasa.gov/). This nature of reflectivity is also reported by Houze (1993) and Steiner et al. (1995). Classifying rainfall solely based on the bright band criteria has been criticized by Steiner et al. (1995). These authors reported that as the vertical thickness of the bright band is typically 0.5 km, it may not be captured by radar at ranges greater than 100 km. The reason is that at high range, both the horizontal and the vertical resolution decreases appreciably due to beam spreading. Steiner et al. (1995) also reported that stratiform rainfall may exhibit distinct bright band during its emerging and decaying stage. Some recent methods for classifying rainfall are using both the horizontal and the vertical reflectivity profile. Johnson et al. (1998) developed an algorithm to classify rainfall based on the 3D analysis of reflectivity. His algorithm is known as "Storm Cell Identification and Tracking" (SCIT). Rigo and Llsat (2003) classified rainfall as mesoscale convective systems, multicell systems, isolated convection, convection embedded in stratiform rainfall, and stratiform rainfall based on horizontal and vertical profile. They used classification from the TRMM algorithm for calibrating their model. Anagnostou (2004) developed a neural network based algorithm for rainfall classification. He selected storm height, reflectivity at 2 km elevation, standard deviation of the reflectivity in pixels within an 11 km radius, difference in height between the rain column top and its maximum reflectivity level, vertical gradient of reflectivity, and the product of rain column top and the 2 km reflectivity value as predictors. Seo et al. (2000) identified three very useful parameters for separating convective/stratiform regions: the maximum apparent rain rate in the vertical, the local spatial correlation coefficient of the maximum

apparent rain rate in the vertical, and the local spatial correlation coefficient of the height of the top of the apparent convective core. They recommended classifying rainfall based on the conditional probability of these three variables and a set of threshold values.

Al-Sowayan and Chandrasekar (2001) used wavelet analysis for classifying precipitation obtained from the TRMM space-borne radar. They performed a 1-D multi-resolution analysis using the dyadic wavelet transform on the vertical profile of reflectivity ranging from 1.25 to 8.5 km elevation. They calculated a decision index for each vertical profile vector and judged the profile to be convective if the decision index was greater than a certain threshold. The threshold value is determined experimentally. Anagnostou (2004) developed a neural network based precipitation classification system. For training and validating the algorithm, precipitation information from the TRMM space borne radar and ground based radars was used.

The above algorithms do not make use of synoptic scale weather charts and are solely based on the morphology of the reflectivity map. In mid-latitudes, precipitation systems are often governed by frontal movements. A movement of a cold front may cause convection in a wide area (frontal convection) and can cause thunderstorm associated with cumulonimbus clouds, whereas a movement of a warm front may cause stratiform rainfall associated with nimbostratus clouds. Convective rainfalls may also be associated with local convection due to uplift of air masses as a result of unequal heating of the earth surface (Lutgens and Tarbuck, 2001). Mizzel (1999) used synoptic scale weather charts for classifying rainfall events based on the analysis of frontal movement. He classified rainfall events in three categories, namely convective, stratiform, and

tropical. Austin (1987) classified rainfall synoptically and suggested different Z-R relationships for different systems.

### 2.5.10 Satellite Based Initiatives

Precipitation estimation using satellite data is an emerging technology. The University of Arizona is carrying out a long term research project known as "Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network" (PERSIANN) (Hsu et al., 1997). This project developed an adaptive ANN based model to incorporate satellite IR image and ground data to estimate precipitation. The ground data is mainly used for updating the network parameters and are applied when available. The spatial resolution is $0.25^{o}$x$0.25^{o}$ and the temporal resolution is half-an-hour. This product covers $50^{o}$N to $50^{o}$S globally and uses data from the GOES-8, GOES 9-10, GMS-5, Metsat-6, and Metsat-7 geostationary satellites. Bellerby et al. (2000) developed a neural network based algorithm to estimate precipitation using the TRMM space-borne precipitation radar and multispectral Geostationary Operational Environment Satellite (GOES) imagery. The algorithm produces half hourly rainfall data at $0.12^{o}$ resolution.

Several satellite based initiatives are also focusing on rainfall classification. Anagstou and Kummarow (1996) and Hong et al. (1998) developed an algorithm for classifying convective and stratiform areas over oceans using microwave brightness temperature considering the differences in latent heat profiles in convective and stratiform clouds. Due to the low spatial resolution, satellite based initiatives are attractive mainly for ocean and ungauged areas of the earth surface. However, for radar covered areas, satellite based initiatives may have limited practical interest.

### 2.5.11 Wind drift

Lack and Fox (2004) proposed an algorithm to correct rainfall advection due to wind drift. In this algorithm, the vertical velocity is assumed to vary linearly from the ground to the CAPPI. The *x* and *y* component of wind velocity is obtained from a single-doppler scan and the fall speed is computed using Lacy's equation relating the rainfall rate to the fall velocity. The fall time is computed by dividing the fall speed by the CAPPI altitude. Using the velocity obtained from the doppler scan, it is then easy to locate the pixel on the ground whereto the rainfall mass from a pixel in the atmosphere will be advected.

A major concern with this algorithm is the fact that wind direction may change with altitude due to the Coriolis force or due to movement of fronts. In addition, wind velocity may not follow a linear variation in the vertical. The authors also did not make any validation of their scheme based on gauge data and it is not clear whether the corrected radar data at ground level is really improving.

## 2.6 SUMMARY AND CONCLUDING REMARKS

There have been considerable efforts to obtain a satisfactory agreement between radar derived rainfall estimates and raingauges. As these two methods measure rainfall at two different spatio-temporal scales and as numerous sources of errors are still unresolved, one should not expect a high degree of correlation between the two sources. Merging data from two systems is an appealing option that has showed promising results when hydrographs are generated based on the merging schemes. One drawback of merging schemes is that they require a fairly dense gauge network.

Extensive research is being carried out to improve the quality of radar data and validation of these schemes is sometimes done using rain gauge data. The fact that radar and gauge measures rainfall at different temporal and spatial scales leads to problems in validating the correction techniques. Even if a correction scheme applied to radar shows improvement in correlation between gauge and radar data, it is valid to ask whether it improves in reality as radar and gauge measures the 'real' rainfall from two different perspectives. It would be of significant interest to compare radar data and gauge data at the same temporal and spatial scale. For example, radar measures precipitation several kilometres aloft and hence one should try to locate the position where the rainfall mass will reach the ground as well as the required travel time. When comparing gauge and radar data, this time and space displacement should be taken into account especially in the case of high resolution data. Unfortunately, very few studies have addressed these issues. If the mean field bias is determined without considering the time-space displacement between gauge and radar data and applied to develop a Z-R relationship, the resulting relationship may not be optimal. As reliable wind profiles may be available from dual-doppler scanning, it should be possible to increase the accuracy of radar precipitation accumulation on the ground. To eliminate residual random errors from CAPPI images, one could consider the use of wavelet denoising schemes. Although this scheme has been used for denoising images for a decade, no application to images of CAPPI reflectivity could be found in the literature. The performance of neural networks in estimating high resolution rainfall accumulation is yet to be explored as previous work has been based on at least one-hour accumulations.

# Chapter 3

# Data Processing and Errors

## 3.1 INTRODUCTON

This chapter discusses sources of radar and gauge data as well as the basic preprocessing carried out on these data. The chapter also comments on error statistics and the quality of the data.

## 3.2 SOURCE OF DATA

The radar data used in this study is collected from the Woodlands radar located approximately 50 km to Winnipeg, Manitoba. The latitude and longitude of the radar site is 50.15377$^{o}$N and 97.77820$^{o}$W, respectively. The scanning area of the radar is divided into 360 azimuth radials. Hence, the angular distance between each two consecutive radials is 1$^{o}$. The horizontal length of a radial is 256 km and it is divided into 256 range bins. Each bin contains reflectivity data of that area. In this study, a data window located over Winnipeg is used. The window is located from azimuth 100 to 150 and range bin 30 to 75. A total of 2346 (=51x46) points are covered by this window. Figure 9 shows the Woodland radar's scanning area and the study window.

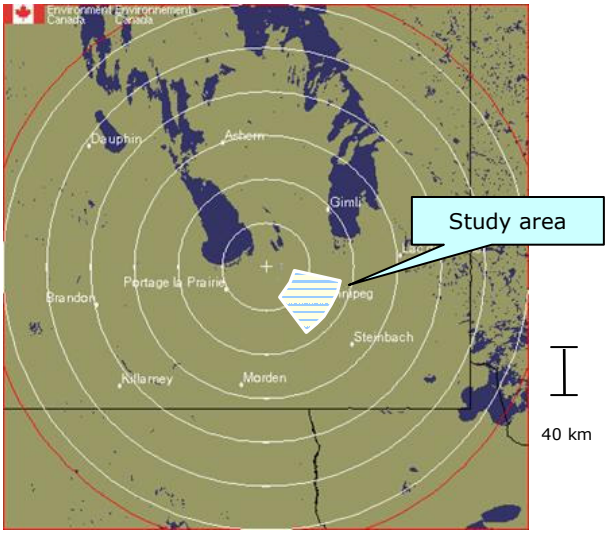


Figure 9: Woodlands radar's scanning area and the study area (Environment Canada).

This study considered the 1.5 km CAPPI (Constant Altitude Plan Position Indicator). Beyond a certain range (around 110 km), reflectivity at this level is not available even with the lowest elevation angle. Hence, the 1.5 km CAPPI is not available at far range (MSC, 2004).The study area contains 24 raingauges. Figure 10 shows the location of these raingauges and location of radar bins. The raingauges are of the tipping bucket type with a resolution of 0.2 mm. The time resolution of the radar data is 10 minutes. Bodiroga (2004) processed Winnipeg's raingauge data to get rainfall accumulation at each 10 minute interval to compare with the radar data. This is done by summing the number of recorded tips within 10-minute interval and then multiplied by 0.2 which produce 10-minute accumulation in millimetre. This processed rainfall data set is used here. In this study, raingauges are numbered from 1 to 24 to facilitate the programming. The gauges are administered by the City of Winnipeg who uses alphabetic identifiers for each gauge. Table 4 shows the actual identifier, the locations of raingauges, and the corresponding numerical numbers assigned in this study.

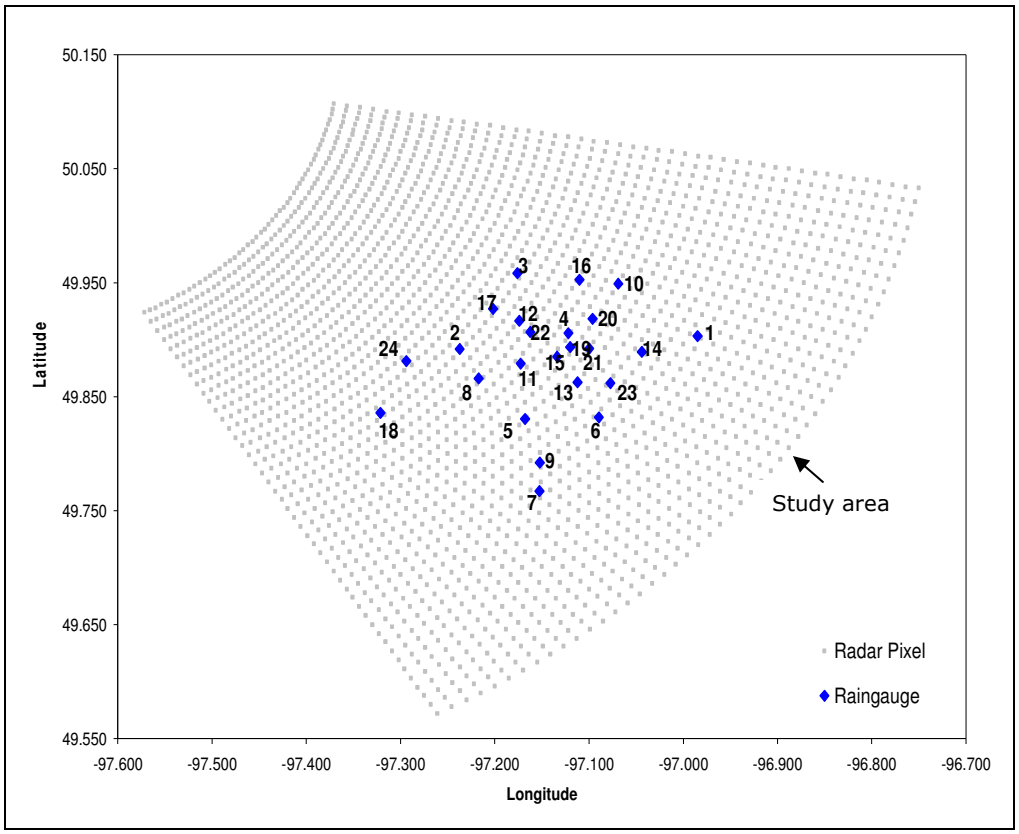


Figure 10: Location of radar bins and raingauges in the study area.

### 3.3 RAINFALL EVENTS

Several major rainfall events occurring in the year 2000 are selected for this study. These events are shown in Table 5. The column of 'Interpolated gauges' contains the ID number of the inoperative gauges. The rainfall at these gauges has been interpolated using the 'inverse distance' method. The last column of the table shows the rainfall intensity at gauge #15. The rainfall intensity is computed by dividing the rainfall accumulation by the duration of the rainfall event. The total rainfall accumulation at gauge #15 is 190.13 mm. The statistics of gauge #15 is provided because the station is located close to the center point of the study area.

Table 4: Raingauge locations, identifiers and numbering.

| ID | No | Location | Latitude | Longitude |
|---|---|---|---|---|
| AD | 1 | 43 Whitehall Blvd | 49°54' 11.318544" | -96° 59' 5.198676" |
| AP | 2 | West of Hangerline Rd | 49° 53' 30.402816" | -97° 14' 13.948584" |
| AW | 3 | 1520 Jefferson Ave | 49° 57' 29.68704" | -97° 10' 33.406212" |
| CG | 4 | Corner of Rover Ave and Anabella St | 49° 54' 20.49228" | -97° 7' 18.676884" |
| CL | 5 | Corner of Clarence Ave and Hamelin Ave | 49° 49' 49.152396" | -97° 10' 4.725588" |
| D5 | 6 | Corner of Granby Ave and Wales St | 49° 49' 54.693156" | -97° 5' 22.186752" |
| EN | 7 | 45 Ave De LaDigue | 49° 46' 1.462692" | -97° 9' 9.186156" |
| ET | 8 | 2300 Corydon Ave | 49° 51' 57.57912" | -97° 13' 1.18758" |
| FR | 9 | 99 Killerney Ave | 49° 47' 31.704612" | -97° 9' 8.654256" |
| FS | 10 | Corner of Mclvor Ave and Rothesay St. | 49° 56' 56.767344" | -97° 4' 8.081364" |
| LS | 11 | 960 Wolsey Ave | 49° 52' 44.7222" | -97° 10' 21.00162" |
| MC | 12 | 360 McPhillips St | 49° 54' 59.386284" | -97° 10' 26.140836" |
| MG | 13 | Corner of Mager Dr and St. Mary's Rd | 49° 51' 45.0027" | -97° 6' 43.668468" |
| MN | 14 | West side of Bournais Dr. | 49° 53' 20.986908" | -97° 2' 37.813092" |
| MY | 15 | Intersection of Mayfair Ave and Main St | 49° 53' 5.758188" | -97° 8' 1.209876" |
| NE | 16 | 2230 Main St | 49° 57' 9.05094" | -97° 6' 36.247752" |
| PL | 17 | Between Parklane Ave and King Edward St. | 49° 55' 37.209576" | -97° 12' 6.122952" |
| PM | 18 | Corner of Perimeter and Wilkes Ave. | 49° 50' 8.89836" | -97° 19' 16.842324" |
| PP | 19 | 227 Provencher Ave | 49° 53' 36.737376" | -97° 7' 12.043812" |
| PR | 20 | 545, Watt St | 49° 55' 5.990052" | -97° 5' 46.19454" |
| SB | 21 | 598, Plinguet St | 49° 53' 32.198208" | -97° 5' 58.752492" |
| TW | 22 | Corner of Tecumseh St and William Ave | 49° 54' 24.167196" | -97° 9' 44.5104" |
| WN | 23 | Corner of Cottonwood Rd and Autumnwood Dr. | 49° 51' 42.527556" | -97° 4' 38.374644" |
| WW | 24 | Corner of Portage Ave and Westwood Dr. | 49° 52' 52.73094" | -97° 17' 37.86828" |

Table 5: Rainfall events used in the study.

| Date | Time (CDT*) | Duration (hh:mm) | Interpolated gauges | Avg. intensity at gauge #15 (mm/hr) | Peak intensity at gauge #15 (mm/hr) |
|---|---|---|---|---|---|
| June 10, 2000 | 3:00 to 12:00 | 09:00 | 12,14,17,21 | 4.31 | 27.6 |
| June 13, 2000 | 10:00 to 14:30 | 04:30 | 7,14,17 | 3.64 | 10.8 |
| July 4, 2000 | 18:00 to 22:00 | 04:00 | 9 | 5.4 | 22.8 |
| July 7, 2000 | 3:00 to 8:00 | 05:00 | 13 | 11.64 | 80.4 |
| July 23, 2000 | 22:00 to 5:00 | 07:00 | 13 | 5.96 | 37.2 |
| Aug 6, 2000 | 15:00 to 18:50 | 03:50 | 9 | 5.3 | 36 |

*Central Daylight Time

The analyzed weather charts provided by Environment Canada shows that the June 13$^{th}$ and the July 4$^{th}$ rainfall is the result of the passing of a warm front, the June 10$^{th}$ and July 7$^{th}$ rainfall is the result of the passing of an occluded front, and the remaining rainfalls are due to the presence of a cold front.

Figure 11 shows the cumulative probability distribution of the gauge rainfall and radar reflectivity over the gauge location. The 10-minute gauge accumulations are converted to the equivalent rainfall intensity in mm/hr by multiplying by six. This figure shows that in around 36% of cases, the equivalent rainfall intensity measured by the gauge is less than or equal to 2 mm/hr, in around 27% of cases the intensity ranges from 2 to 5 mm/hr, in around 29% of cases the intensity ranges from 5 to 20 mm/hr, in around 7% of cases the intensity varies from 20 to 50 mm/hr, and in only 2% of cases the intensity is above 50 mm/hr. Hence, according to the classification provided in the link http://grappa.meteo.mcgill.ca/mcgill_img.html, these rainfall events consists of almost equal portions of light rain (36%), moderate precipitation (27%), and heavy rain (29%). The portion of very heavy rain is significantly less (8%).

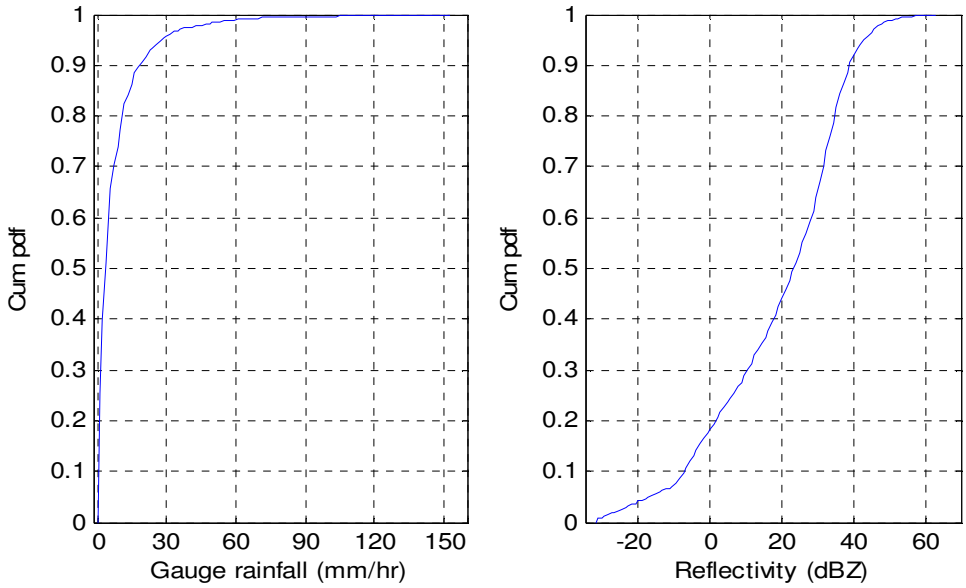


Figure 11: Cumulative probability distribution of equivalent gauge intensity and reflectivity over the gauge locations.

### 3.4 DATA QUALITY

An animation of rainfall events from the lowest elevation angle scanning was created using MATLAB in order to detect ground clutter. Any reflectivity persistent over a prolonged duration at a given location may be considered as ground clutter. However, the animation did not show any evidence of ground clutter. The animation also helped to appreciate visually the direction of rainfall movement, its lateral extent and distribution at each time step. The conventional scan data does not contain any correction for attenuation, anomalous propagation, virga, or any other sources of errors (Patrick 2004, personal communication). Furthermore, during summer thunderstorms the bright band should be well above 1.5 km and hence the 1.5 km CAPPI should not be contaminated by bright band enhancement.

As the study area is located within a 50 km to 70 km radial distance from the radar location, it may be assumed that range dependent errors (e.g. overshooting beam) is not affecting the data and it will not be within the 'cone of silence' region. The weather observer reported the occurrence of hail on July 10$^{th}$, July 7$^{th}$, and July 23$^{rd}$ (Patrick 2005, personal communication). Hence, these rainfall events are affected by enhancement of reflectivity associated with hail. The correlation coefficient between two closely located raingauges (gauge #15 and #19) is estimated as 0.9 for these rainfall events. The geographic distance between these two gauges is 1.37 km. As the radar resolution is only 1 km, we may assume that the spatial variability of rainfall within a 1 km distance is not significant for the considered rainfall events. Hence, aerial average sampling by the radar and point sampling by the gauge should not be a major reason for radar-gauge discrepancy, at least for these rainfall events. Huff and Ship (1969)'s study also shows that gauge-to-gauge correlation is about 0.95 for geographic distances up to 5 km in case of thunderstorms and rain showers for one-minute rain-rates. For steady rains gauge-to-gauge correlation for such distances may be even greater than 0.95. As the rainfall events selected for this study are thunderstorms or rain showers, it may be assumed that anomalous propagation is not a serious issue except in the cases of the wake of thunderstorm. The animation of rainfall shows movements of precipitation consistent with the Velocity Azimuth Display (VAD) data at 3 km to 6 km altitude which would not be the case if AP echoes were the source of the signal.

In summary, the main reason for the discrepancy between the radar and the gauge measurements should be wind drift, temporal sampling issues, attenuation, and hail.

### 3.5 DATA PROCESSING

The radar data (*N*) provided by Environment Canada varies within a range of 0 to 255. The following equation is used to obtain the logarithmic reflectivity (*dBZ*),

$$dBZ = 0.5 * N - 32 \quad \text{Eq 3.1}$$

Rainfall rates are computed using Marshall-Palmer's equation in the following form,

$$Z = 10^{(dBZ/10)} \quad \text{Eq 3.2}$$

$$R = (Z/200)^{(1/1.6)} \quad \text{Eq 3.3}$$

where $R$ is the radar estimated rainfall rate in mm/hr and $Z$ is the reflectivity in $mm^6/m^3$.

A few rain gauges occasionally did not provide data. Rainfall values for inoperative gauges were interpolated from other gauges using the inverse distance method. The equation for the inverse distance method is

$$G_j = \frac{\sum_{i=1}^{n} \frac{G_i}{d_{i,j}^2}}{\sum_{i=1}^{n} \frac{1}{d_{i,j}^2}} \quad \text{Eq 3.4}$$

where $d_{i,j}$ is the Euclidian distance between interpolated points and other gauges, $G_i$ is the rainfall value at the surrounding gauges, and $G_j$ is the interpolated rainfall accumulation. The power of the inverse distance is chosen as two as it is considered the most common choice (Issaks and Srivasta, 1989). The location of raingauges is provided by their global coordinates (latitude and longitude) and hence to calculate the Euclidian distance between gauges, the study area is re-plotted in a coordinate system with kilometre as scale unit. A $1^o$ latitude difference is equivalent to 111.15 km and $1^o$ longitude difference is equivalent to 72.08 km in this area. This scale is used to calculate the Euclidian distance between gauges. The gauges subjected to interpolation are listed in Table 5. A

cross-validation scheme was undertaken by removing a gauge (gauge #10) from the network for a rainfall event and then comparing its interpolated rainfall with the actual rainfall. The correlation coefficient between the interpolated and the actual rainfall was found to be 0.96 which may be considered satisfactory.

It was observed that in many cases the radar is inoperative for a prolonged period (more than 1 hour). The periods with missing radar data were not included in the study.

It is necessary to extract the radar reflectivity at the point located above or at some distance away from the gauge location. To accomplish this, a triangular based linear interpolation method is applied. This method computes the reflectivity at a geographic point from three neighbouring radar pixels by solving the linear equations

$$\begin{aligned} ax_1 + by_1 + c &= R_1 \\ ax_2 + by_2 + c &= R_2 \\ ax_3 + by_3 + c &= R_3 \end{aligned} \qquad \text{Eq 3.5}$$

where $\{(x_1,y_1), (x_2,y_2), (x_3,y_3)\}$ is the set of spatial coordinates of three vertices of the triangle, and $R_1$, $R_2$, and $R_3$ are the radar pixel values corresponding to the corner points. The value of the parameters *a, b, c* can be determined by solving the system of linear equations. Once solved, the radar echo at any intermediate point in the triangle can be determined using the following expression:

$$R_j = ax_j + by_j + c \qquad \text{Eq 3.6}$$

where $(x_j,y_j)$ is the coordinate of the point to be interpolated and $R_j$ is the interpolated radar echo.

In some cases (e.g. for rainfall classification, storm tracking, etc.) it is necessary to rearrange radar echoes in a Cartesian coordinate system. To do this, a Cartesian grid

with1 km grid spacing in both *x* and *y* directions is constructed. The radar echoes are then interpolated onto the Cartesian grid using the same interpolation scheme (triangular based linear interpolation).

## 3.6 ERROR STRUCTURE

### 3.6.1 General

In this section, the error structure of the radar and gauge measurements is analyzed. It is assumed that the radar pixel located over the gauge represents the radar rainfall corresponding to that gauge location. It is also assumed that the rainfall intensity within each 10-minute cycle is constant. Hence, in the case of radar the rainfall accumulation for a 10-minute period will be equal to the rainfall intensity (mm/hr) divided by six. In the same way, rainfall intensity from gauge measurements will be equal to the gauge accumulation multiplied by six. The radar and gauge data corresponding to the same time point are compared.

### 3.6.2 Z-R Association

The Z-R relationships are usually constructed by fitting a straight line to the scatter points of *logZ* and *logR,* where Z is in $mm^6/m^3$ and *R* is in mm/hr. Both variables can be measured from the disdrometer data. A Z-R relationship will be significant only when these two variables show a specific trend. As an alternative to disdrometer data, it is possible to check the association of *logZ* and *logR* by taking *Z* from radar observations and *R* from gauge observations. Rainfall rates computed from gauge data, are referred to as *G* instead of *R* to distinguish the difference. Figure 12 shows the *logZ-logG* association

of the six rainfall events used in the study. The rainfall rate is computed by multiplying the 10-minute gauge accumulations by six.

The scatter plots of these rainfalls (Figure 12) do not look encouraging for developing a relationship. Even if a Z-R relationship works well in one rainfall period, it may not work in other rainfall events. For example, in the case of the June 10$^{th}$ rainfall, a value of $Z$ 10$^5$ mm$^6$/m$^3$ is found at a gauge rainfall rate of 20 mm/hr, whereas in the case of the July 7$^{th}$ rainfall, a value of similar magnitude has a gauge rainfall rate of 100 mm/hr.

One possible reason for the underestimation of $Z$ in the case of the peak rainfall of July 7$^{th}$ may be attenuation as attenuation is a function of rainfall rate. Hence, an equation that provides a good fit for the June 10$^{th}$ rainfall may not produce acceptable results in the case of the July 7$^{th}$ rainfall. On the other hand, it is observed that for a rainfall rate equal to 10 mm/hr, $Z$ varies within a range of 10$^1$ to 10$^5$ mm$^6$/mm$^3$. One possible reason behind this may be wind drift which advects the rainfall mass at 1.5 km CAPPI to a different location. It is observed that the June 13$^{th}$ and July 4$^{th}$ rainfall shows relatively better association between the two variables, and that the July 23$^{rd}$ and August 6$^{th}$ rainfall shows no association at all. No Z-R relationship will work for these two rainfall events unless possible errors are eliminated and a better association is identified. Despite this fact, the following section explores the performance of two common Z-R relationships.

### 3.6.3 Z-R Relationships

From Table 2, it is observed that $Z=200R^{1.6}$ is recommended for stratiform precipitation and $Z=300R^{1.4}$ is recommended for non-tropical convection. Hence, the performance of

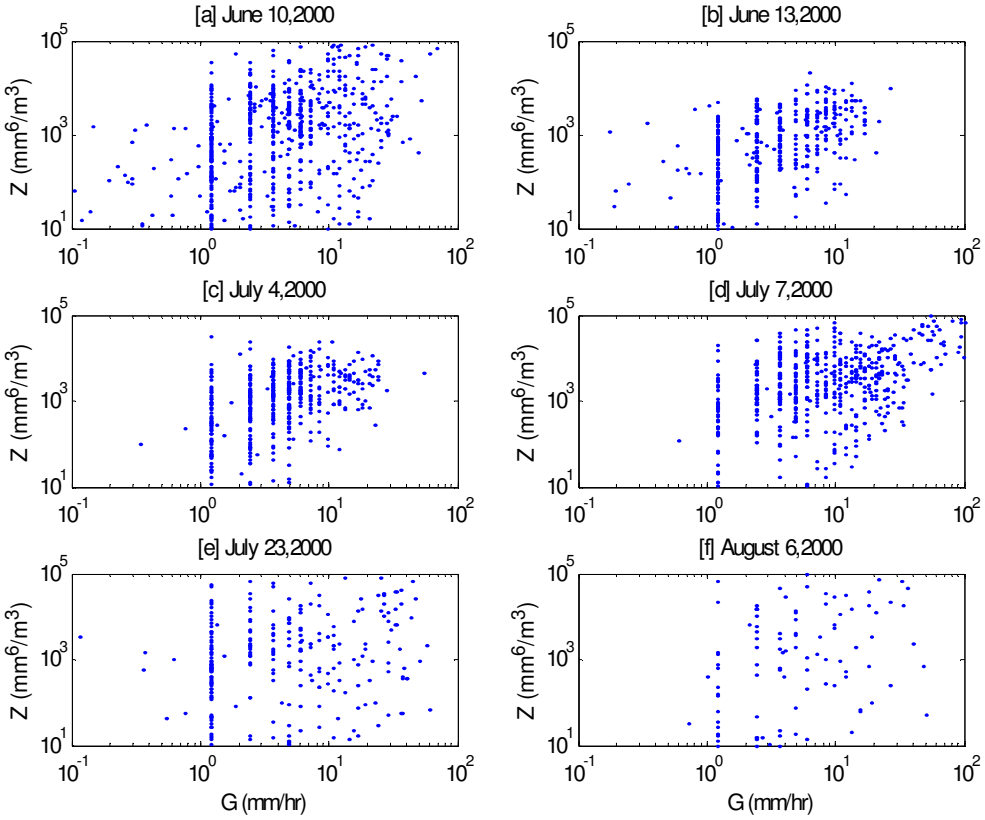


Figure 12: *logZ-logG* association in six rainfall events used in the study.

these two relationships is explored. It was mentioned in Section 2.5.9 that rainfall systems are often composed of convective and stratiform regions and that raindrop characteristics are different within each region. Rainfall regions used in this study is classified using Steiner et al. (1995)'s algorithm. This algorithm is described in brief below:

1. Any pixel is considered convective if its reflectivity is greater than forty dBZ.
2. If the difference between reflectivity at a point and its background reflectivity exceeds a threshold, the point is considered convective. The background reflectivity is the average reflectivity of nonzero echoes in a circle of 11 km

radius centered at the pixel of interest. The following three criteria are proposed as threshold for deciding that a pixel is convective:

$$\Delta dBZ > \begin{cases} 10 & dBZ' < 0 \\ 10 - dBZ'^2/180 & 0 \le dBZ' \le 42.43 \\ 0 & dBZ' \ge 42.43 \end{cases} \quad \text{Eq 3.7}$$

where $dBZ'$ is the background reflectivity and $\Delta dBZ$ is the difference between the reflectivity at a pixel and its background reflectivity.

3. The surrounding areas of any pixel identified as convective from these two criteria is also convective. The radius of the surrounding area to be convective depends on the background reflectivity at that point and is provided in the following equation,

$$r_s = \begin{cases} 1 & dBZ' < 25 \\ (dBZ' - 20)/5 & 25 \le dBZ' \le 45 \\ 5 & dBZ' \ge 45 \end{cases} \quad \text{Eq 3.8}$$

where $r_s$ is the radius of the surrounding area in kilometres.

Figure 13 shows the rainfall classification using this algorithm for a scan of the study area on June 10, 2000, 8:00 UTC. Inside the study area, the white color indicates areas with no precipitation, the grey color is areas of stratiform precipitation, and the black color is convective precipitation.

For the six events considered in this study, it is found that 55% pixels are identified as convective pixels, and the rest (45%) are stratiform pixels. Figure 13 is consistent with the gauge data where in 36% of cases, the rainfall intensity is found to be in the range 0-2 mm/hr. If precipitation echoes for the whole study area are considered,

21% of the pixels are found to be convective and the rest are stratiform pixels. Although the selected precipitations are all intense convective storms, they also include widespread stratiform regions.

After the classification, convective pixels are converted to rainfall rate using $Z=250R^{1.2}$ and stratiform pixels are converted using $Z=200R^{1.6}$. The reason for not using $Z=300R^{1.4}$ for convective pixels is that this equation and Marshall-Palmer's equation produce the same rainfall rate up to 42 dBZ (Figure 4) and reflectivity above 40 dBZ is considered convective according to Steiner et al. (1995)'s algorithm. Hence, classification using $Z=300R^{1.4}$ and $Z=200R^{1.6}$ will produce the same performance as obtained by using solely $Z=300R^{1.4}$. The rainfall rate is divided by six to get the equivalent 10-minute rainfall accumulations to compare with gauge accumulations. Table 6 shows the performance of the two Z-R relationships and the classification system for the six rainfall events.

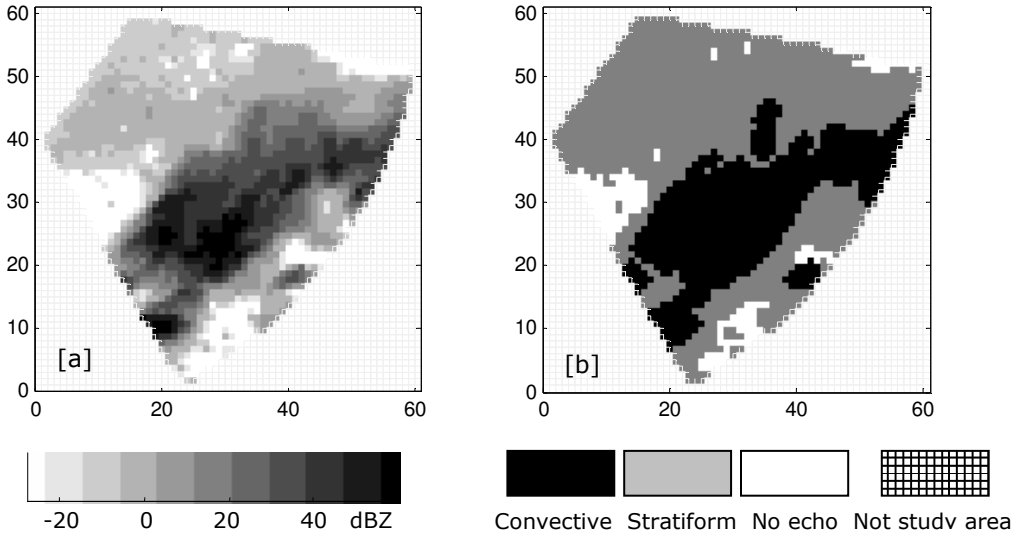

Figure 13: Rainfall classification using Steiner et al. (1995)'s algorithm, June 10, 2000, 8:00 UTC.

Table 6: Performance of Z-R relationships.

| Date of Rainfall | Correlation Coefficient | | | RMSE (mm) | | |
|---|---|---|---|---|---|---|
| | (200,1.6) | (300,1.4) | Classification | (200,1.6) | (300,1.4) | Classification |
| June 10, 2000 | 0.25 | 0.21 | 0.16 | 2.29 | 3.30 | 9.22 |
| June 13, 2000 | 0.58 | 0.56 | 0.54 | 0.56 | 0.56 | 0.64 |
| July 4, 2000 | 0.40 | 0.38 | 0.35 | 0.86 | 0.89 | 1.18 |
| July 7, 2000 | 0.66 | 0.66 | 0.65 | 2.75 | 2.66 | 3.56 |
| July 23, 2000 | 0.22 | 0.21 | 0.19 | 4.64 | 7.34 | 22.74 |
| Aug 6, 2000 | 0.28 | 0.27 | 0.25 | 1.32 | 1.54 | 3.2 |

It may be observed from Table 6 that using $Z=200R^{1.6}$ for all pixels irrespective of their classification as convective or stratiform shows relatively better result. Indeed, the magnitude of the correlation coefficient for classified precipitation is quite poor in all rainfall events. Figure 14 shows scatter plots of 10-minute rainfall accumulations measured from the gauge and the radar data. The radar rainfall is obtained using $Z=200R^{1.6}$. This figure shows that the radar underestimates rainfall in cases where gauge accumulations are greater than 5 mm. Significant overestimation is also observed when gauge accumulations are less than 5 mm. Hence, if one attempts to fix the underestimation of peak rainfall using a multiplicative bias, this correction will cause excessive overestimation for the points corresponding to gauge accumulations less than 5 mm. Absence of systematic underestimation or overestimation prevents a simple application of bias corrections.

### 3.6.4 Bias and Error Variance

The multiplicative bias, mean error, and error variance are computed for each rainfall event using the $Z=200R^{1.6}$ relationship and 10-minute rainfall accumulations. The multiplicative bias correction factor varies from 0.5 to 1.8 in the different rainfall events,

Table 7: Bias, mean error, and error variance.

| Date of Rainfall | Multiplicative Bias | Mean Error (mm) | Variance of Error ($mm^2$) |
|---|---|---|---|
| June 10, 2000 | 0.84 | -0.12 | 5.25 |
| June 13, 2000 | 1.62 | 0.19 | 0.28 |
| July 4, 2000 | 1.26 | 0.15 | 0.72 |
| July 7, 2000 | 1.80 | 0.87 | 6.83 |
| July 23, 2000 | 0.51 | -0.59 | 21.2 |
| Aug 6, 2000 | 1.02 | 0.01 | 1.74 |

as shown in Table 7. This implies that the radar is overestimating (on average) in some rainfall events and underestimating in others. For example, if one calculates the bias using the July 7$^{th}$ rainfall (bias=1.8) where the radar is underestimating and applies this bias correction to the July 23$^{rd}$ rainfall (bias=0.5) where the radar is overestimating, one should expect results that are worse than if no bias correction had been used at all. The overall bias for the six rainfall events is 1.05 which implies that there is no long term systematic underestimation or overestimation by the radar when using $Z=200R^{1.6}$. The error variance is particularly high in the case of the June 10$^{th}$, the July 7$^{th}$, and the July 23$^{rd}$ rainfall events. When the multiplicative bias is greater than 1, the mean error is positive.

It is of interest to explore the variation of error statistics as a function of range. Six gauges are selected and are grouped into three pairs. The first pair (gauges #3 and #18) is the closest to the radar, and the third pair (gauges #1 and #7) is the farthest from the radar among the 24 gauges. The second pair (gauges #10 and #15) is located halfway between the other two pairs. Table 8 shows error statistics for these six gauges. This table shows no particular trend in error statistics as a function of range. Probably the distance between gauges is insufficient to capture any range dependent variation of error statistics.

Table 8: Error statistics at some selected gauges.

| Gauge # | Multiplicative bias | Correlation coefficient | Variance of error ($mm^2$) |
|---|---|---|---|
| 3 | 0.82 | 0.26 | 11.4 |
| 18 | 1.75 | 0.43 | 4.33 |
| 10 | 0.85 | 0.06 | 13.1 |
| 15 | 1.28 | 0.48 | 2.95 |
| 1 | 1.07 | 0.33 | 3.39 |
| 7 | 1.39 | 0.44 | 3.54 |

### 3.6.5 Hourly Accumulation

The hourly accumulations show a significantly improved correlation between gauge and radar data. Figure 15 shows scatter plots of the gauge and the radar data in the case of 10-

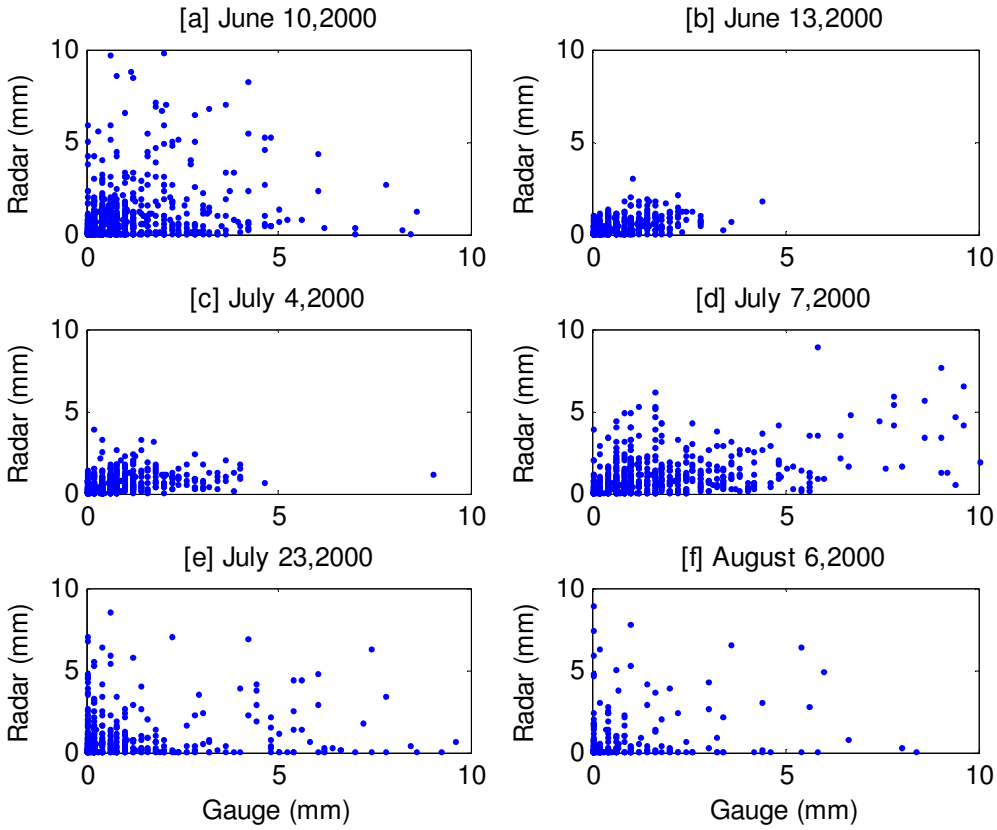


Figure 14: The 10-minute rainfall accumulations from gauge and radar data.

minute accumulations and hourly accumulations, respectively.

The correlation coefficient for the two cases is 0.3 and 0.55, respectively, which is a significant improvement. An hourly time resolution apparently smoothens out sampling errors due to the temporal mismatch between the two datasets.

Despite the improvement in the hourly accumulations, overestimation is still observed in the case of the July 23rd rainfall event. Overall the data show a systematic underestimation. In particular, an excessive underestimation occurred in the peak rainfall of the July 7th rainfall. The systematic underestimation can be removed by applying a multiplicative bias correction factor. The multiplicative bias changes the multiplicative coefficient of the Z-R relationship and essentially creates a new Z-R relationship. It does not change the exponent of the Z-R relationship. Hence, a bias correction is equivalent to new Z-R relationship.

A new Z-R relationship that reduces the underestimation can be developed from the data. Admittedly, the application of a bias correction will further deteriorate the

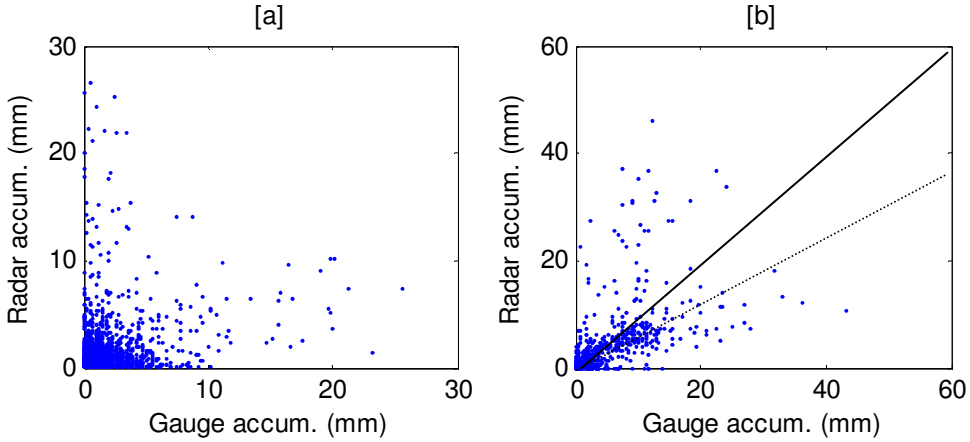


Figure 15: Comparison of gauge and radar data [a] 10-minute accumulations, [b] hourly accumulations.

overestimated points of the July 23rd rainfall. Considering these points as a result of error, the following Z-R relationship is found as optimum:

$$Z = 150R^{1.5} \qquad \text{Eq 3.9}$$

where $R$ is the rainfall rate in mm/hr and $Z$ is the reflectivity in $mm^6/m^3$. Figure 16 shows the comparison of hourly gauge and radar accumulations using this relationship.

From Figure 11 it is possible to extract rainfall intensity and reflectivity associated with similar cumulative probability (e.g. 0.1, 0.2, 0.3,...,0.9,1). It is also possible to plot reflectivity and rainfall rate having the same cumulative probability. Figure 17 shows the probability matching curve and the $Z=150R^{1.5}$ relationship. It can be seen that the two curves match fairly closely up to 50 dBZ. It was found that the probability matching curve agreed perfectly with $Z=30R^2$.

It is observed that the gauge-radar correlation at an hourly time step increases to 0.6 using the Z-R relationship obtained from the probability matching techniques. However, this relationship increases bias to 0.71 from 1.05.

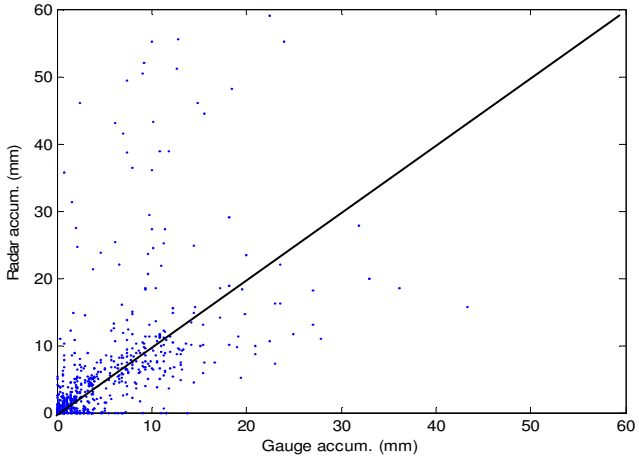


Figure 16: Hourly accumulations using $Z=150R^{1.5}$ relationship.

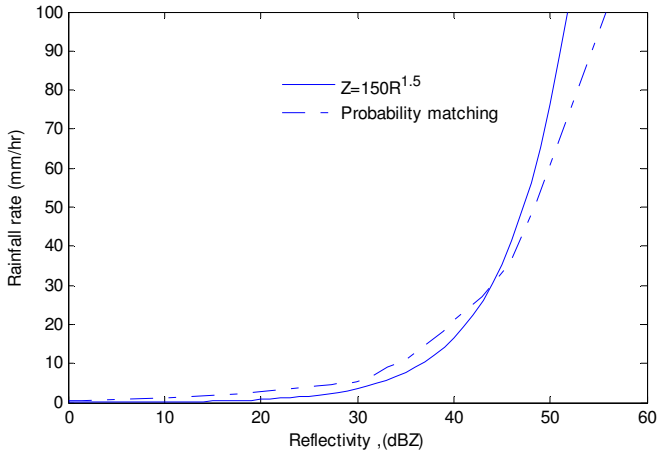

Figure 17: Comparison with the probability matching curve.

## 3.7 SUMMARY

The objective of this study is to investigate methods for achieving a better correlation in the case of 10-minute accumulations. The 10-minute accumulations show a wide range of variations between radar and gauge measurements which leads to questions regarding the use of high resolution rainfall data in flood forecasting models or in urban drainage system management. The displacement of rainfall mass due to wind drift and the time taken to reach the ground may be an important reason for the mismatch. It would seem that a proper correction algorithm is necessary when working with high resolution radar data. A statistical approach based on neural networks is explored in the next chapter with the objective to obtain a better correlation between gauge and radar 10-minute accumulations. In Chapter 5, a trajectory model is developed to assess the impact of wind drift and time shifting on the 10-minute accumulations. This model also accounts for attenuation and hail contamination.

# Chapter 4

# Exploration of Neural Networks

## 4.1 INTRODUCTION

Artificial Neural Networks (ANN) can fit any type of nonlinear input-output relationships (Hsu et al., 1995) and also considered a robust/error tolerant method of estimation. Hence, using neural networks in the case of radar rainfall estimation is an appealing idea. A brief summary of earlier work on neural networks in radar-rainfall estimation was given in Section 2.5.6. These attempts used spatial reflectivity patterns to estimate rainfall rate at the centre. The performance of the trained ANN was tested with the same rainfall event used for training but for different locations, and time resolutions were at least one hour. It is of interest to explore the performance of the ANN model compared to the conventional Z-R relationships at high time resolution and for different rainfall events that was not included in the training data.

This chapter explores the performance of ANN for rainfall estimation. A comparative study on the performance of back propagation network, radial basis function network, and the generalized regression network is presented. All simulations are carried out for a temporal resolution of 10 minute. Section 4.2 provides a short description of the algorithm of the ANN's used in the study and Section 4.3 presents the results obtained.

## 4.2 ALGORITHM DESCRIPTION

### 4.2.1 The Back propagation Network

Figure 18a shows the system behind the back propagation network. When input is provided to the network, the network generates output which is compared to the target data. The error is the difference between the target and the output. Based on the error, the weights and biases in the network are then adjusted. This adjustment is called the training of the network. In this study, the input is the radar reflectivity and the target is the rain gauge rainfall.

Figure 18b shows the schematic of the network, consisting of neurons, connections, and layers. Each connection is associated with a weight and each neuron is associated with a bias. The weights and biases are updated during the training stage.

Figure 18c shows the data processing inside a neuron. Data provided to the input neuron is multiplied by weights and added up to produce the data inside a neuron of the $1^{st}$ hidden layer. A bias is added to the summed value and the result is then passed through a transfer function. Three common transfer functions are the log-sigmoid, the tan-sigmoid, and the linear. The log-sigmoid function is used in this study.

A network is trained by a series of dataset. If the network has three input neurons, it takes three input data at a time from the series and generates output. This small block of input data is usually termed ‘pattern’. The output error may be computed for each pattern and weights may be updated based on this error. This process may be repeated for the whole data series and after completing the data series, it may retrain the data series in the same way. This type of training is called incremental training (Demuth and Beale, 1994)

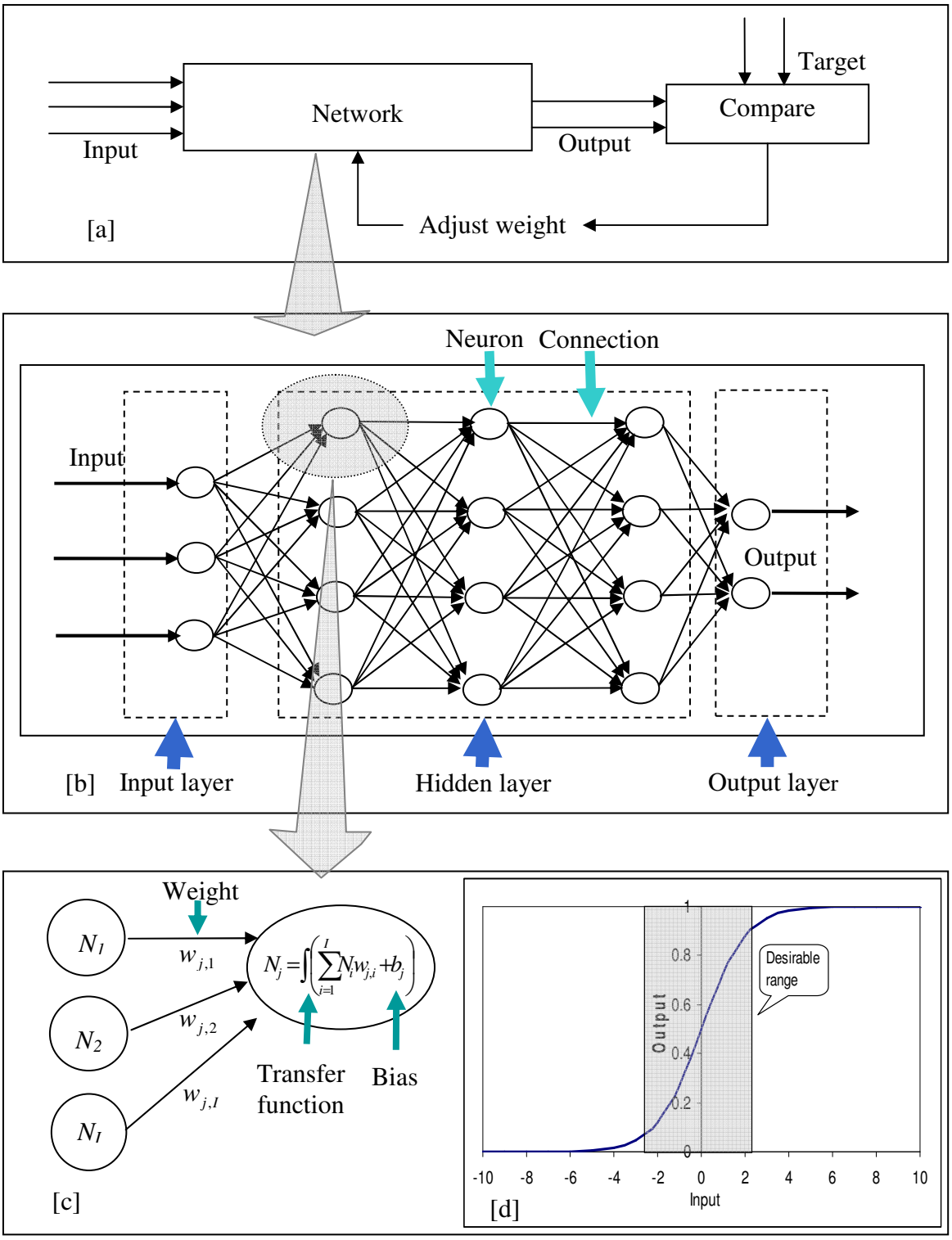


Figure 18: [a] The ANN system, [b] the network [c] a neuron [d] the log-sigmoid transfer function.

and the training of one cycle of the data series is usually referred to as an 'epoch'. For example, if the model trains the data series one hundred times, it will complete one hundred epochs. One problem of incremental training is that the network may be skewed to the last pattern presented in the network (Zurada, 2000).

Alternatively, weights may be updated after all patterns have been processed by the network. This type of training process is usually called 'batch' training. The 'batch' training process is employed in this study.

When the log-sigmoid transfer function is used, the equation for computing the generated data inside a neuron is as follows:

$$N_j^p(l) = \frac{1}{1+\exp\left\{-\left(\sum_{i=1}^{I} N_i^p(l-1).w_{j,i}^p(l-1) + b_j^p(l)\right)\right\}} \qquad \text{Eq 4.1}$$

where $N$ is the value of a neuron, $w$ is the weight connecting two neurons, and $b$ is the bias. The subscript $j$ refers to the neurons of the $l^{th}$ layer and subscript $i$ refers to the neurons of the layer preceding the $l^{th}$ layer. The superscript $p$ refers to the pattern.

The advantage of using transfer functions is that it allows for nonlinearity in the data and yields output in a certain range of values whatever the input is (Zealand, 1997). For example, the log-sigmoid transfer function provides output between 0 and 1. The advantage of adding a bias lies in the fact that it allows to shift the data in a desirable range other than [0, 1]. The desirable range is shown in Figure 18d by the shaded region.

Eq 4.1 is used to compute values of neurons in the first hidden layer from the neurons of the input layer. When data inside all neurons in the first hidden layer have been computed, the program proceeds to the next layer and generates the value inside

each neuron in the same fashion. Eq 4.1 is used to compute the value inside each neuron of the second hidden layer. In this case, subscript *j* refers to the neurons in the second hidden layer and subscript *i* refers to the neurons in the first hidden layer. Finally, the model output is generated. The program then computes the error between the output and the target and based on the error, weights and biases are updated. Weights and biases are usually updated using a gradient descent algorithm. The objective of the gradient descent algorithm is to minimize the error by updating the weights and biases. Initially, weights are assigned randomly. There is no guarantee that the global minimum will be reached. A detailed derivation of gradient descent algorithms is available in textbooks on neural networks (e.g. Zurada, 2000). The derived equations are provided below.

For the output layer, the error function can be determined from the following equation:

$$\delta_j^p(L) = 2[T_j^p - N_j^p(L)]N_j^p(L)[1 - N_j^p(L)] \qquad \text{Eq 4.2}$$

where subscript *j* refers to the neurons of the output layer and superscript *p* refers to the pattern. The notation *(L)* refers to the output layer. The notation *T* refers to the target and *N(L)* refers to the values produced in the neurons of the output layer. This error function is back-propagated to the hidden layers and may be computed as follows:

$$\delta_i^p(l-1) = N_i^p(l-1)[1 - N_i^p(l-1)]\sum_{j=1}^{J}\delta_j^p(l)w_{j,i}(l) \qquad l = L, L-1,...,1 \qquad \text{Eq 4.3}$$

where subscript *j* refers to the neurons in the $l^{th}$ layer and subscript *i* refers to the neurons in the layer preceding the $l^{th}$ layer. The notation $w_{j,i}(l)$ refers to the weights connecting the $j^{th}$ neuron of the $l^{th}$ layer and the $i^{th}$ neuron of the $(l\text{-}1)^{th}$ layer. Eq 4.3 is used iteratively to compute error functions for neurons in all hidden layers. In each iteration,

the notation (*l-1)* refers to the layer for which the error function is to be calculated and the notation (*l)* refers to the layer following the $(l\text{-}1)^{th}$ layer. After computing error functions for each neuron and for each pattern, the gradient of error can be computed for each weight and for each pattern as,

$$\frac{\partial e^p}{\partial w_{j,i}^l} = N_i^p(l-1)\delta_j^p(l) \qquad \text{Eq 4.4}$$

where subscript *j* refers to the neurons of the $l^{th}$ layer and subscript *i* refers to the neurons of the $(l\text{-}1)^{th}$ layer. The term $\partial e^p / \partial w_{j,i}^l$ refers to the gradient of error corresponding to the weight connecting the $j^{th}$ neuron of the $l^{th}$ layer to the $i^{th}$ neuron of the $(l\text{-}1)^{th}$ layer. Each pattern generates a set of error gradients.

The error gradient for biases is determined in a similar fashion (Mukto, 2001) and is given by

$$\frac{\partial e^p}{\partial b_j^l} = \delta_j^p(l) \qquad \text{Eq 4.5}$$

In case of batch training, the average error gradient is computed by averaging error gradients for a weight or bias for all patterns. This averaging is expressed mathematically as follows:

$$\nabla \overline{E} w_{j,i}(l) = \frac{1}{P}\sum_{p=1}^{P}\frac{\partial e^p}{\partial w_{j,i}^l} \qquad \text{Eq 4.6}$$

The average gradient for biases is computed in the same fashion. The change of each weight in the network is computed by multiplying the corresponding error gradient with a constant, known as the learning constant, which in fact determines the step size at

each epoch. The smaller the learning constant, the slower will be the rate of convergence. Mathematically, the change of weights is computed as

$$\Delta w_{j,i}(l) = \eta \nabla \overline{E} w_{j,i}(l) \qquad \text{Eq 4.7}$$

where $\eta$ is the learning constant. The change of bias is computed in the same fashion. It is customary to add a small fraction of the changes of weights obtained in the previous training cycle or epoch. This is known as the momentum method which helps to accelerate convergence. In that case Eq 4.7 is modified as follows:

$$[\Delta w_{j,i}(l)]^t = \eta(1-\alpha)[\nabla \overline{E} w_{j,i}(l)]^t + \alpha[\Delta w_{j,i}(l)]^{t-1} \qquad \text{Eq 4.8}$$

where superscript *t* refers to the current epoch and superscript *(t-1)* refers to the previous epoch. The added term is known as the momentum term and the parameter $\alpha$ is known as the momentum constant. The optimum value is selected by experimentation. The same procedure is applied for the bias updating.

All these processes complete one training cycle or epoch. The changes of weights should be added to the initially randomly generated weights to get the weights for the next epoch. Mathematically, this is expressed as

$$[w_{j,i}(l)]^{t+1} = [w_{j,i}(l)]^t + [\Delta w_{j,i}(l)]^t \qquad \text{Eq 4.9}$$

The next training cycle or epoch uses the weights and biases obtained from Eq 4.9 and determines another set of changes of weights and biases. After completion of a training cycle or epoch, the sum of squared errors for all patterns used in the training cycle is computed. If the error is smaller than a selected threshold, the training is terminated. The excessive computational requirements of the gradient descent method have prompted researchers to look for faster training algorithms using variable learning

and/or momentum parameters. Another research direction has focused on applying numerical optimization techniques to accelerate convergence, leading to the application of quasi-Newton methods, nonlinear least squares methods, etc. as training algorithms (Hagan and Menhaj, 1994).

Hagan and Menhaj (1994) attempted to apply Levenberg-Marquardt optimization techniques as a training algorithm. This method computes error gradients for all weights and biases using Eq 4.4 and Eq 4.5. However, the error function for the outer layer (Eq 4.2) is computed as follows:

$$\delta_j^p(L) = -N_j^p(L)[1 - N_j^p(L)] \qquad \text{Eq 4.10}$$

A Jacobian matrix is constructed combining the error gradients for all weights and biases for all patterns in an epoch (Castro, 2004, p. 55). The Jacobian matrix is given by

$$J(x) = \begin{bmatrix} \frac{\partial e_1^1}{\partial w_{1,1}^1} & \frac{\partial e_1^1}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_1^1}{\partial w_{2,1}^1} & \frac{\partial e_1^1}{\partial w_{2,2}^1} & \cdots & \frac{\partial e_1^1}{\partial b_1^1} & \frac{\partial e_1^1}{\partial b_2^1} & \cdots & \frac{\partial e_1^1}{\partial w_{1,1}^2} & \frac{\partial e_1^1}{\partial w_{1,2}^2} & \cdots \\ \frac{\partial e_2^1}{\partial w_{1,1}^1} & \frac{\partial e_2^1}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_2^1}{\partial w_{2,1}^1} & \frac{\partial e_2^1}{\partial w_{2,2}^1} & \cdots & \frac{\partial e_2^1}{\partial b_1^1} & \frac{\partial e_2^1}{\partial b_2^1} & \cdots & \frac{\partial e_2^1}{\partial w_{1,1}^2} & \frac{\partial e_1^1}{\partial w_{1,2}^2} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \\ \frac{\partial e_q^1}{\partial w_{1,1}^1} & \frac{\partial e_q^1}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_q^1}{\partial w_{2,1}^1} & \frac{\partial e_q^1}{\partial w_{2,2}^1} & \cdots & \frac{\partial e_q^1}{\partial b_1^1} & \frac{\partial e_q^1}{\partial b_2^1} & \cdots & \frac{\partial e_q^1}{\partial w_{1,1}^2} & \frac{\partial e_1^1}{\partial w_{1,2}^2} & \cdots \\ \frac{\partial e_1^2}{\partial w_{1,1}^1} & \frac{\partial e_1^2}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_1^2}{\partial w_{2,1}^1} & \frac{\partial e_1^2}{\partial w_{2,2}^1} & \cdots & \frac{\partial e_1^2}{\partial b_1^1} & \frac{\partial e_1^2}{\partial b_2^1} & \cdots & \frac{\partial e_1^2}{\partial w_{1,1}^2} & \frac{\partial e_1^1}{\partial w_{1,2}^2} & \cdots \\ \frac{\partial e_2^2}{\partial w_{1,1}^1} & \frac{\partial e_2^2}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_2^2}{\partial w_{2,1}^1} & \frac{\partial e_2^2}{\partial w_{2,2}^1} & \cdots & \frac{\partial e_2^2}{\partial b_1^1} & \frac{\partial e_2^2}{\partial b_2^1} & \cdots & \frac{\partial e_2^2}{\partial w_{1,1}^2} & \frac{\partial e_2^2}{\partial w_{1,2}^2} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \\ \frac{\partial e_q^p}{\partial w_{1,1}^1} & \frac{\partial e_q^p}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_q^p}{\partial w_{2,1}^1} & \frac{\partial e_q^p}{\partial w_{2,2}^1} & \cdots & \frac{\partial e_q^p}{\partial b_1^1} & \frac{\partial e_q^p}{\partial b_2^1} & \cdots & \frac{\partial e_q^p}{\partial w_{1,1}^2} & \frac{\partial e_1^1}{\partial w_{1,2}^2} & \cdots \end{bmatrix} \qquad \text{Eq 4.11}$$

where the superscript $p$ refers to the number of patterns and the subscript $q$ refers to the number of neurons in the output layer.

The change of weights and biases is calculated using the following expression:

$$\Delta \mathbf{x} = [\mathbf{J}^{\mathbf{T}}\mathbf{J} + \mu\mathbf{I}]^{-1}\mathbf{J}^{\mathbf{T}}\mathbf{e} \qquad \text{Eq 4.12}$$

where $\mu$ is a scalar and $\mathbf{I}$ is the identity matrix. The vector $\mathbf{e}$ contains errors for all neurons in the output layer and for all patterns. Hence, its size should be equal to [*nx1*], where *n* is the product of the number of neurons in the output layer and the number of patterns. This error is computed by subtracting the target vector from the model output vector. If the total number of weights and biases in a network is *m*, the dimension of $\mathbf{J}$ is [*nxm*]. The dimension of $\mathbf{e}$ is [*nx1*], the dimension of $\Delta\mathbf{x}$ is [*mx1*] and the dimension of $\mathbf{I}$ is [*mxm*]. The elements of $\Delta\mathbf{x}$ represent the change of weights/biases. The weights/biases for the next iteration are

$$[\mathbf{x}]^{t+1} = [\mathbf{x}]^{t} + [\Delta\mathbf{x}]^{t} \qquad \text{Eq 4.13}$$

where superscript *t* stands for epoch.

To avoid over-fitting, the available data set is usually divided into three categories: the training data, the validation data, and the test data. After completing one training cycle or epoch, the mean square error in the validation data set is computed. It is usually observed that the error in the training data decreases gradually with the number of epochs, however the error in the validation dataset may decrease initially but after a certain number of epochs begin to increase. This is an indication of over-fitting and the training is usually terminated at this point.

The trained network may then be tested using the test data set. Figure 19 shows the point where training should be stopped.

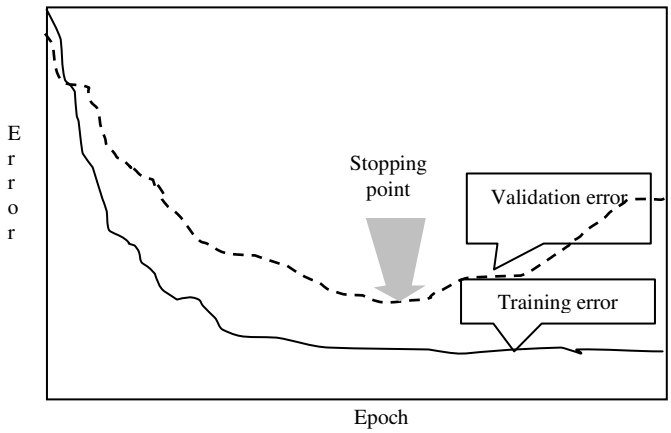


Figure 19: Demonstration of the stopping criteria.

### 4.2.2 The Radial Basis Function Network

The radial basis function networks usually involve lower computational time than the error back propagation networks. This type of neural net consists of one input layer, one hidden layer, and one output layer. Usually it possesses a large number of neurons in the hidden layer. Values of neurons in the hidden layer is computed using the following expression:

$$N_j = \exp\left( - \frac{\sum_{i=1}^{I} (N_i - w_{j,i})^2}{1.44\sigma^2} \right) \qquad \text{Eq 4.14}$$

where $N_i$ are the neurons of the input layer, $N_j$ are the computed values of the neurons of the hidden layer, $w_{j,i}$ are weights (also known as center), and the parameter $\sigma$ is known as the 'width' (or radius or spread).

The term $\sum_{i=1}^{I} (N_i - w_{j,i})^2$ is the weighted input to the $j^{th}$ neuron of the hidden layer. The weighted input to the $j^{th}$ neuron of the hidden layer is the square of the Euclidian distance between the input vector and its corresponding weights connecting the input neurons to the $j^{th}$ neuron of the hidden layer.

The output produced from Eq 4.14 is a Gaussian bell-shaped curve. If the weighted input to a neuron in the hidden layer is zero, the output of the function is one. The higher the weighted input to a neuron in the hidden layer, the lower is the output. The width parameter determines the spread of the 'bell'. The lower the width, the steeper is the peak of the 'bell'.

The output of the network is the output of the neurons of the output layer. These are computed from the output of the neurons of the hidden layer. Like back-propagation networks, radial basis network may also have several outputs. In that case, the network will have several neurons in the output layer instead of just one neuron. The output neurons have no activation function but a bias is added. The output of neurons in the output layer is usually computed by the following expression,

$$N_j = \sum_{i=1}^{I} w_{j,i} N_i + b_j \qquad \text{Eq 4.15}$$

where *i* refers to the neurons in the hidden layer and *j* refers to the neurons in the output layer. The term *b* refers to the added bias.

To develop the radial basis function model, one needs to determine the value of the weights for the hidden layer and the output layer and a reasonable value for the 'width'. A detailed description of the radial basis function network is available in the textbook "Neural Networks" (Phil, 2000). This study uses MATLAB's built-in functions to construct radial basis network.

The number of neurons in the hidden layer is kept the same as the number of input patterns presented in the network during the training session. The weights connecting the

input neurons to the $j^{th}$ neuron in the hidden layer is equal to the $j^{th}$ pattern presented in the network. Thus the network attempts to memorize the whole training data series resulting in a huge number of neurons in the hidden layer. When an input pattern equal to the weight vectors connecting a hidden neuron is presented, the output of that neuron should be one, and for other hidden neurons, whose weights are not close to the input vector, the output should be close to zero. Hence, the output from different hidden neurons is different, ranging from zero to one due to the application of a Gaussian activation function in the hidden neurons. The optimum value for width should be chosen by trial and error and may be kept the same for all neurons in the hidden layer.

Weights and biases for the output layer are determined from the target data for different input patterns. If the total number of neurons in the hidden layer is $I$ and the total number of neurons in the output layer is $J$ and the total number of patterns presented in the network is $P$, weights and biases for the output layer can be determined by solving the following system of equations,

$$\begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,I} & b_1 \\ w_{2,1} & w_{2,2} & \cdots & w_{2,I} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ w_{J,1} & w_{J,2} & \cdots & w_{J,I} & b_J \end{bmatrix} * \begin{bmatrix} N_{1,1} & N_{1,2} & \cdots & N_{1,P} \\ N_{2,1} & N_{2,2} & \cdots & N_{2,P} \\ \vdots & \vdots & \ddots & \vdots \\ N_{I,1} & N_{I,2} & \cdots & N_{I,P} \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,p} \\ T_{2,1} & T_{2,2} & \cdots & T_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ T_{J,1} & T_{J,2} & \cdots & T_{J,p} \end{bmatrix} \quad \text{Eq 4.16}$$

where $w_{j,i}$ refers to weights, $b_j$ refers to biases, $N_{i,p}$ refers to neurons in the hidden layer, and $T_{j,p}$ refers to the target. The elements of **N** and **T** are known. The value of $I$ and $P$ should be equal.

These weights and biases yields an exact solution, that is, if an input vector from the training data series is presented in the network, the output will be exactly the same as the target vector provided for that input vector during the training of the network.

In order to avoid large number of hidden neurons, clustering the input patterns is often recommended. This study applies a *k*-means clustering algorithm. This algorithm starts by generating *k* patterns randomly. The distances from each input pattern vector to the randomly generated pattern vectors are calculated and the input patterns are divided into *k* clusters based on the minimum distance criterion. The mean of each cluster is calculated. In the next iteration, the distance from each input pattern to the mean-pattern of each cluster is re-calculated. This iterative process continues until the mean patterns between successive iterations do not change appreciably. The number of neurons in the hidden layer is equal to *k* and the mean value of each cluster is then used as the weights connecting the input neuron to the hidden neuron. The mean value of the target of each cluster is computed and used as the target for training the network.

### 4.2.3 The Generalized Regression Neural Networks

The algorithm for the Generalized Regression Neural Network is quite similar to the algorithm for the radial basis function neural network. Like radial basis function networks, it has three layers, one input layer, one hidden layer, and one output layer. The hidden layer neurons also have a Gaussian activation function. The determination of weights and width for the hidden layer is the same as that of the radial basis function networks and can be determined using Eq 4.14. However, the determination of weights for the output layer is quite different. A description of the generalized regression network

is available in the reference, “Advanced Method in Neural Computing” (Wasserman, 1993).

Like radial basis function networks, the number of patterns presented in the network during training is equal to the number of neurons in the hidden layer. The only difference is that the target provided for an input pattern is directly used as weight connecting the corresponding hidden neuron to the output neuron. Mathematically, if the number of neurons in the hidden layer is $I$ and the number of neurons in the output layer is $J$, and the number of patterns presented to the network is $P$, the weights connecting the hidden layer to the output layer can be determined from the following expression:

$$\begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,I} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,I} \\ \vdots & \vdots & \ddots & \vdots \\ w_{J,1} & w_{J,2} & \cdots & w_{J,I} \end{bmatrix} * \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,P} \\ T_{2,1} & T_{2,2} & \cdots & T_{2,P} \\ \vdots & \vdots & \ddots & \vdots \\ T_{J,1} & T_{J,2} & \cdots & T_{J,P} \end{bmatrix} \quad \text{Eq 4.17}$$

where $T_{j,p}$ refers to the target elements. The values of $I$ and $P$ are equal. Using these weights, the output is computed as,

$$N_j = \frac{\sum_{i=1}^{I} w_{j,i} N_i}{\sum_{i=1}^{I} N_i} \quad \text{Eq 4.18}$$

where $N_i$ refers to the neurons in the hidden layer and $N_j$ refers to the neurons of the output layer. Unlike radial basis function networks, no bias term is added in the output layer.

As the underlying principle of radial basis networks and generalized regression networks is based on regression and not inspired by the biological neural system, these models are also known as statistical neural networks (Phil, 2000).

## 4.3 RESULTS AND DISCUSSIONS

### 4.3.1 General

Three different types of neural network are explored: (i) the back propagation network, (ii) the radial basis function network, and (iii) the generalized regression network. In each case, the input data for the network is the radar reflectivity in *dBZ* and the output is the 10-minute rainfall accumulations in mm. The raingauge data of 10-minute accumulations is used as the target while training the network. It was mentioned earlier that there are 24 rain gauges in the study area and six rainfall events are extracted for use in the study. The location of gauges is shown in Figure 10. Out of these 24 gauges, data from 16 gauges (gauge # 1 to 16) are used for training the network and the rest are used for validation (gauge #17 to 20) and testing (gauge #21 to 24) purposes. The process of training, test, and validation is carried out using the rainfall data of the first five rainfall events listed in Table 5. The calibrated networks are tested using the sixth rainfall event for all gauges. This is done to explore the capability of the neural networks to estimate rainfall accumulations for new rainfall events with different spatial distribution of rain. Hence, a total of four data sets are used in this study. It is of interest to look at how the gauge and radar data sets correlate with each other when using Marshall-Palmer's Z-R relationship ($Z=200R^{1.6}$). Figure 20 shows scatter plots of gauge and radar for the four data sets. The root mean square error (RMSE) and the correlation coefficient (CC) are indicated in the figures. The correlation coefficients are quite small in all four data sets. This error statistic will be used to compare the neural networks output with that of a traditional Z-R relationship. As the Z-R relationship provides rainfall rate in mm/hr, the rainfall rate is

divided by six to get 10-minute rainfall accumulations in mm. This should work if the rainfall intensity within the 10 minute interval remains fairly uniform.

Each of the neural networks is tested for two different input patterns: (i) a temporal pattern and (ii) a spatial pattern. The temporal input data contains reflectivity for the past several time bins of the radar pixels located above the raingauges. The spatial input data contains reflectivity located above and in the neighbourhood of the raingauge locations for the same time bin. The radar pixel value located above or at a certain distance from a gauge is determined using the triangular based linear interpolation method described in Section 3.5.

In summary, a total of six (=3x2) different scenarios are tested. The results of the six scenarios are presented in Table 9 and the best network is selected based on the minimum RMSE of the validation dataset.

### 4.3.2 Performance of the Back-Propagation Network

To determine "the best" back propagation network is an impossible task for two reasons: (i) it is impossible to know if a global minima has been reached, (ii) the back-propagation network can be designed an indefinite number of ways. Hence, the optimum back propagation network is reported based on the limited number of trials carried out.

The excessive time requirement of the gradient descent algorithm prevents a comprehensive study using this training algorithm. Hence, the Levenberg-Marquardt algorithm will be employed. However, as this algorithm converges too fast, it causes problems if $\mu$ (described in Eq 4.12) is taken too small and the stopping criteria described

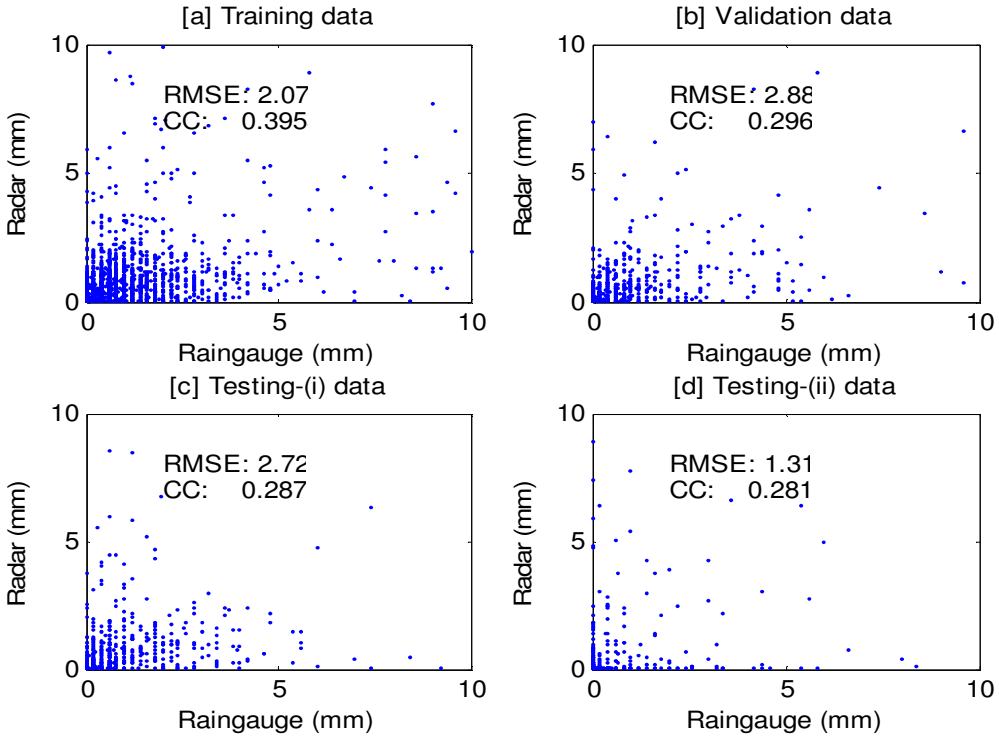


Figure 20: Comparison of the gauge and the radar accumulations using Z-R relationship (200, 1.6) at four data sets used in the study.

in Figure 19 are used. In such cases, it is recommended to employ a larger value of $\mu$. Hence, in this study $\mu$ is set to 0.9.

A total of 50 networks are tested, varying the number of input neurons, number of neurons in the hidden layer, and the number of hidden layers. Figure 21 shows the performance of the network that produced optimum result from the 50 networks tested. The parameters of this network are:

| | |
|---|---|
| No of input neuron: | 5 |
| No of hidden layers: | 2 |
| No of neurons in hidden layer: | 12 |
| No of output neurons: | 1 |
| Transfer function between layers: | log-sigmoid |

This network uses reflectivity in the five previous time bins and hence it uses a temporal pattern as input data. The stopping criterion shown in Figure 19 is applied here. To implement this, 16 raingauges are used for training (gauge #1 to 16), 4 raingauges are used for validation (gauge #17 to 20), and the remaining 4 gauges are used for test (gauge #21 to 24). Training is stopped when the mean-square-error of the validation data begins to increase. Training, validation, and test are carried out for first five rainfall events listed in Table 5. The trained network is tested using the sixth rainfall event. The purpose of this procedure was discussed in the Section 4.3.1. The input and the target data are linearly scaled. The input data are divided by 68, while the target data are divided by 15. Figure 21 shows scatter plots between gauge and radar accumulations in the training, validation, and test data sets. Two test data sets are shown in the figure. One (Figure 21c) is for gauges 21 to 24 for the first 5 rainfall events listed in Table 5 and the other (Figure 21d) is for the all gauge data of the sixth rainfall event.

The neural network works fairly well for the first test data set (Figure 21c) compared to the Z-R relationship, however its performance for the second test data set is not satisfactory (Figure 21d). In the case of the second test data set, the performance of the neural network is quite similar to that observed for the Z-R relationship (Figure 20d). This issue is discussed in detail later in this chapter.

It was observed that with increasing neurons in the input layer, the correlation coefficient in the validation and the test data set increased. Despite this, more than 5 input neurons were not tested because in such case it may not work well in the case of short-lived thunderstorms or rain showers. However, the correlation coefficient in the second set of test data was found to decrease except when there was only one input neuron.

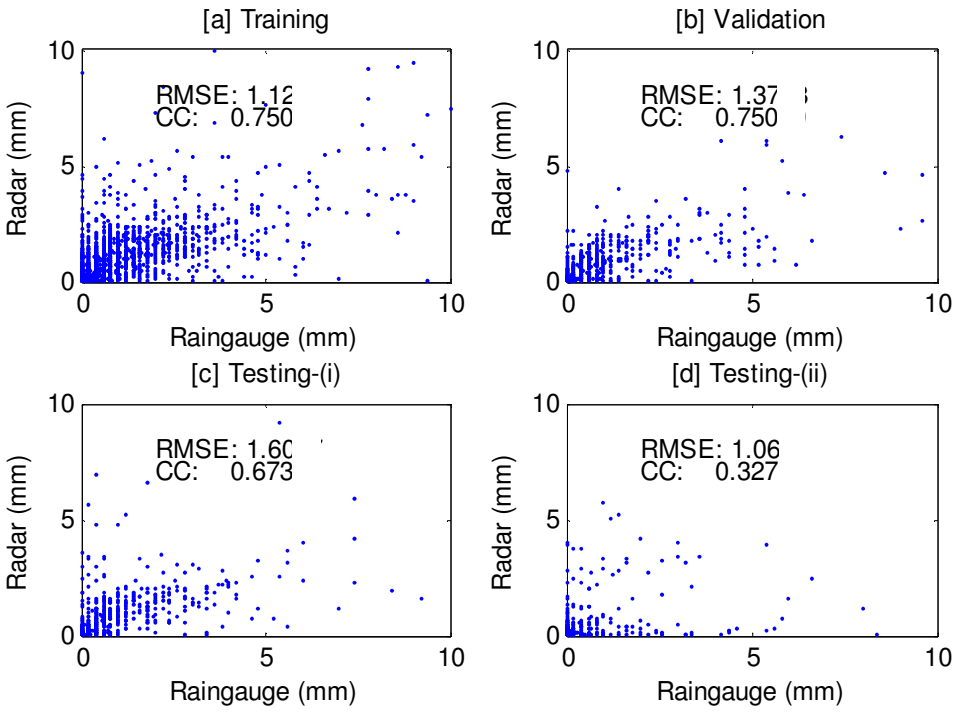


Figure 21: Comparison of 10-minute accumulations of gauge rainfall and radar estimation with back-propagation network for temporal input pattern.

Increasing the number of hidden neurons while keeping the other parameters the same does not show any systematic increase or decrease in the performance functions. The relative performance of networks having one and two hidden layers with a constant total number of hidden neurons was examined. For example, a total of 16 hidden neurons can be arranged in two hidden layers each having 8 neurons or in one hidden layer having 16 neurons. A comparison was made between these two arrangements. It was observed that two hidden layers works better in 56% of cases in case of the first test data and 52% cases in case of the second test data. However, as the ratio is close to 50%, it may be concluded that the arrangement of hidden neurons has relatively insignificant impact on model performance.

Table A.1 of Appendix A shows the weight matrix connecting the input layer and the first hidden layer. The sixth column of the table contains the bias vector of the first

hidden layer neurons and is shaded in grey to separate it from the weight matrix. The weight matrix is a [12x5] matrix and is shown in the first five columns of the table. This is a [12x5] matrix because the network has 12 neurons in the first hidden layer and 5 neurons in the input layer. The $[m \times n]^{th}$ element of the matrix refers to the weights connecting the $m^{th}$ element of the hidden layer to the $n^{th}$ element of the input layer. The bias vector is a [12x1] vector because there are 12 hidden neurons. Table A.2 shows the weights connecting to the first hidden layer to the second hidden layer and the bias for the second hidden layer. The weight matrix is a [12x12] matrix as there are 12 neurons in the second hidden layer and 12 neurons in the first hidden layer. The bias vector for the output layer contains 12 elements because there are 12 neurons in the second hidden layer. Table A.3 shows the weights and bias for the output layer.

Table B.1 of the Appendix B shows outputs of 50 test runs carried out to determine an optimum network. Each test run creates a network which is listed in the first column of the table under the 'net ID' heading. The fifth column refers to the number of epoch required by the program to achieve the minimum error in the validation data set. Beyond this epoch, the mean square error of the validation data set started to increase.

The performance of the back-propagation network in case of the spatial data is now explored. In this case, nine surrounding radar pixels are considered as input. The gauge is located at the centre of the nine radar pixels. The training, the test, and the validation data sets are used as in the previous case. A total number of 10 networks are tested, varying the number of neurons in the hidden layers and the number of hidden layers. The optimum network obtained has the following characteristics:

| | |
|---|---|
| No of input neuron: | 9 |
| No of hidden layers: | 2 |
| No of neurons in hidden layer: | 8 |
| No of output neurons: | 1 |
| Transfer function between layers: | log-sigmoid |

The same scaling functions as in the previous case are used. Figure 22 shows the comparison of the gauge and the radar accumulation for the training, validation, and the two sets of test data. The error statistics (the RMSE and the correlation coefficient) are indicated on the figure. The overall error statistics are inferior to the previous case (temporal input), except for the new rainfall event (Figure 22d). Although performance improved statistically, the visual observation of the scatter plot is not satisfactory at all. This model is unable to estimate rainfall greater than 5 mm (10-minute accumulations) in all four data sets. Increasing the number of hidden neurons does not show any systematic improvement or deterioration of the model performance.

The early stopping criterion is used in this case too, that is, the training ceases when the error statistics in the validation data set begins to increase. Table A.4 shows the weight matrix connecting the input layer with the first hidden layer and the bias vector for the first hidden layer.

Table A.5 shows the weight matrix connecting the first hidden layer to the second layer and the bias vector for the second hidden layer. Table A.6 shows weights and biases for the output layer. Table B.2 shows the performance of all the networks tested in this category.

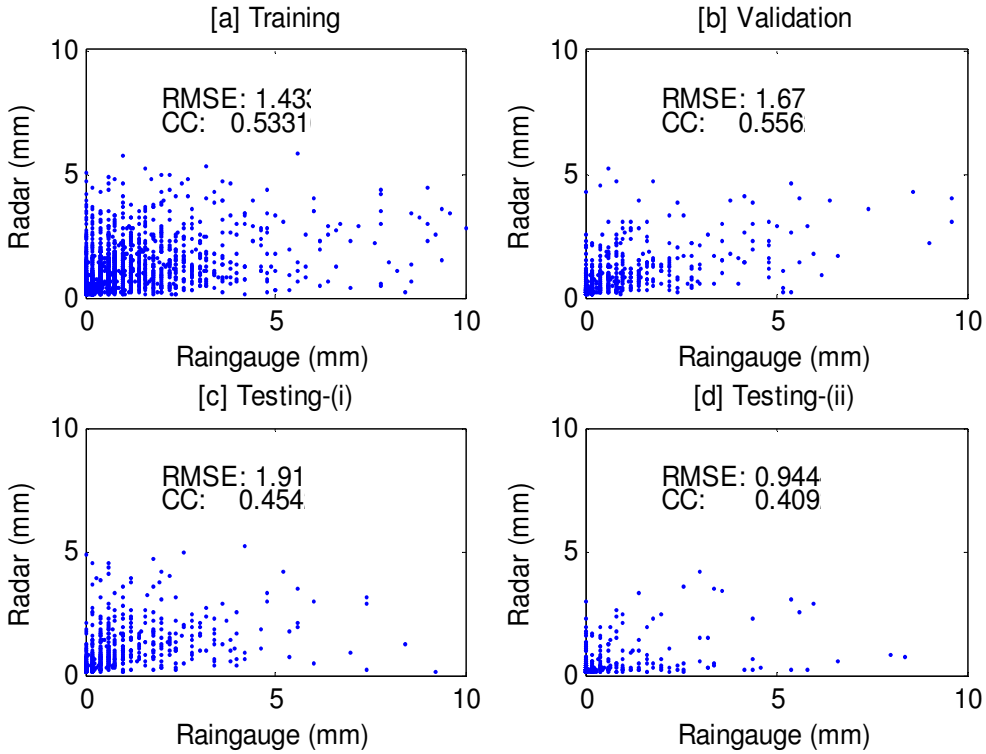


Figure 22: Comparison of 10-minute accumulations of gauge rainfall and radar estimation with back-propagation network for spatial input pattern.

### 4.3.3 Performance of the Radial Basis Network

The width or radius is an important model parameter for the radial basis network that affects the model performance. Hence different widths have been tested. As the model creates one neuron in the hidden layer for each input pattern, a clustering of the input patterns was done. Hence, the model is tested for different numbers of clusters. A total of 150 networks were tested and the optimum model parameters are shown below:

| | |
|---|---|
| No of input neuron: | 3 |
| No of clusters: | 50 |
| Width: | 0.2 |
| No of output neuron: | 1 |

The scaling of the input and the target is done as in the previous cases. The scatter plots of the four data sets are shown in Figure 23. Like the other cases, this network also

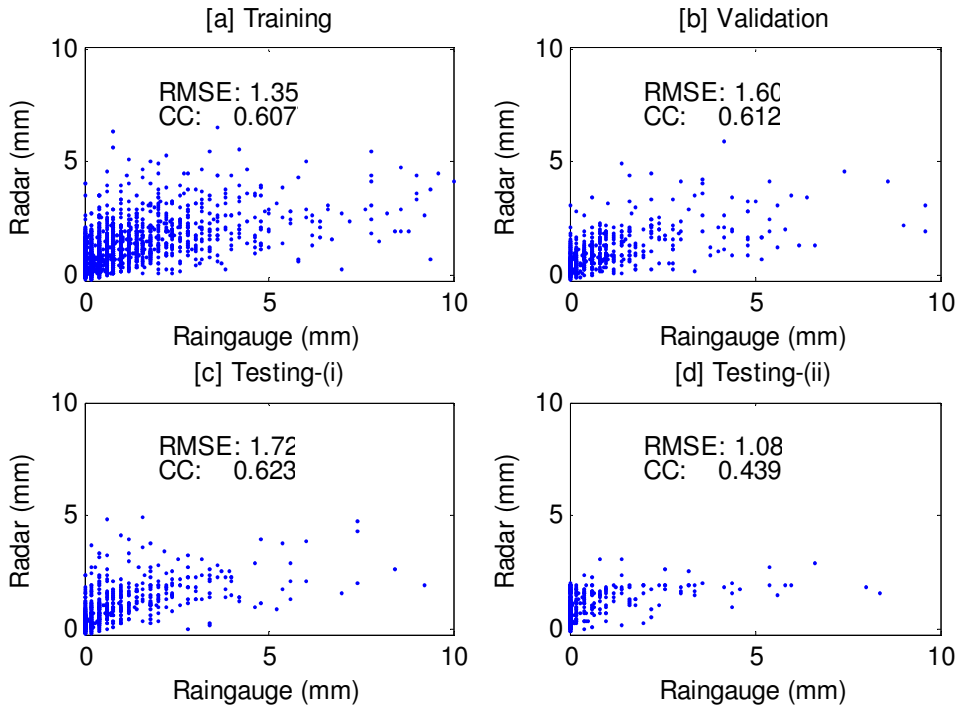


Figure 23: Comparison of 10-minute accumulations of gauge rainfall and radar estimation with the radial-basis network for the temporal input pattern.

failed to show satisfactory performance for the new rainfall event (Figure 23d). In case of the test data (Figure 23c) and the validation data (Figure 23b), it is found that the model overestimates when gauge accumulations (10-minute) are less than 2.5 mm and underestimates when gauge accumulations are greater than 2.5 mm.

The correlation for the validation data is lower than that of the back propagation network. However, correlation in the second test data set is improved in this model although the scatter plot shows serious underestimation of rainfall.

The clustering is important in radial basis networks, as otherwise the solution of the matrix system for determining the weights of the output layer becomes ill-conditioned. Ill-conditioned matrices were encountered when the number of clusters was

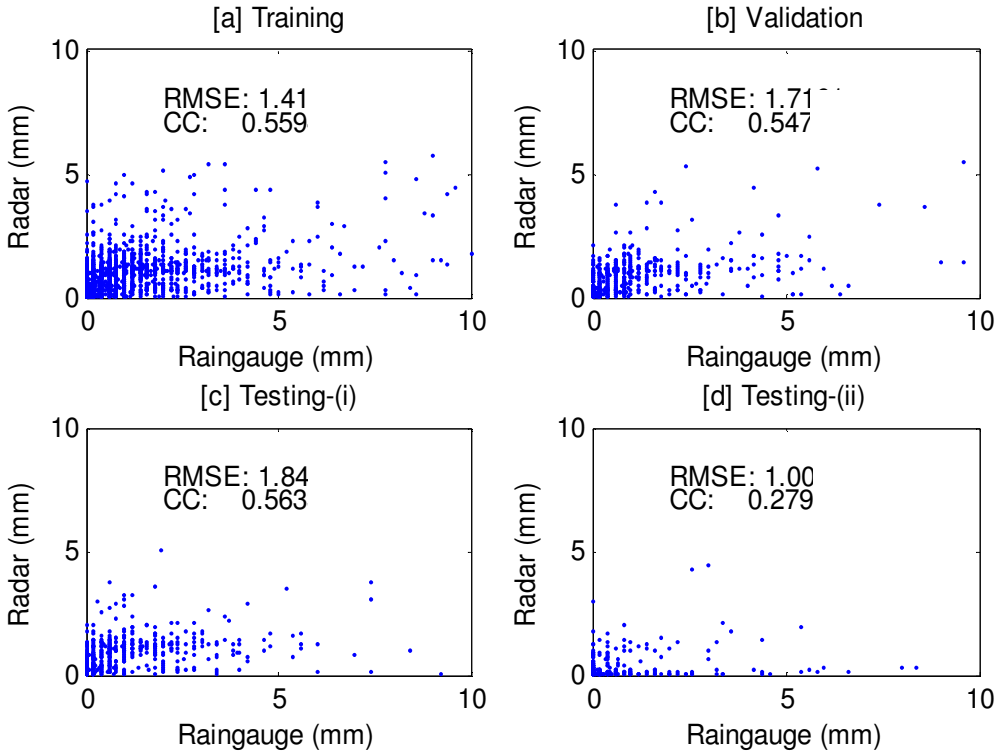


Figure 24: Comparison of 10-minute accumulations of gauge rainfall and radar estimation with the radial-basis network for the spatial input pattern.

greater than 50 and/or the number of input neurons is one. The optimal width was found to be 0.2. Table B.3 shows the performance of the networks tested.

Ill-conditioned matrices are not found frequently in the spatial input data probably because of the larger size of the input vector. A total of 30 networks are tested by varying widths and *k* values. The optimum model parameters obtained is shown below:

| | |
|---|---|
| No of input neurons: | 9 |
| No of clusters: | 50 |
| Width: | 0.2 |
| No of output neurons: | 1 |

The scatter plots for the four data sets are shown in Figure 24. The performance of all the networks tested in this category is shown in Table B.4.

### 4.3.4 Performance of the Generalized Regression Network

Like the radial basis network, the generalized regression network is also tested for different widths and different numbers of clusters. Unlike the radial basis function network, the training algorithm of the generalized regression network is free from the ill-conditioning problem. A total of 150 networks are tested and the optimum network obtained is

| | |
|---|---|
| No of input neuron: | 5 |
| No of clusters: | N/A |
| Width: | 0.1 |
| No of output neuron: | 1 |

The optimum network has no clustering of the input patterns. Hence, the total number of hidden neurons is equal to the total number of training patterns (=2812). The dimension of the weight matrix connecting the input layer to the hidden layer is [2812x5] and the dimension of the weight matrix connecting the hidden layer to the output layer is [1x2812]. In fact the first weight matrix is the input data and the second weight matrix is the target data. If **I** is the input vector having 2812 elements and **T** is the target vector having 2812 elements, the first weight matrix is [ I(1:5) ; I(2:6);…….; I(2807:2812)] and the second weight matrix is [T(1) T(2) …….T(2812)].

The scatter plots of the four data sets are shown in Figure 25. Like the other networks, this network also fails to show good correlation in the second test data set (Figure 25d). The performance of the network tested is shown in Table B.5.

A total of 30 networks are tested with the spatial input data, varying widths and *k* values. The optimum network obtained is

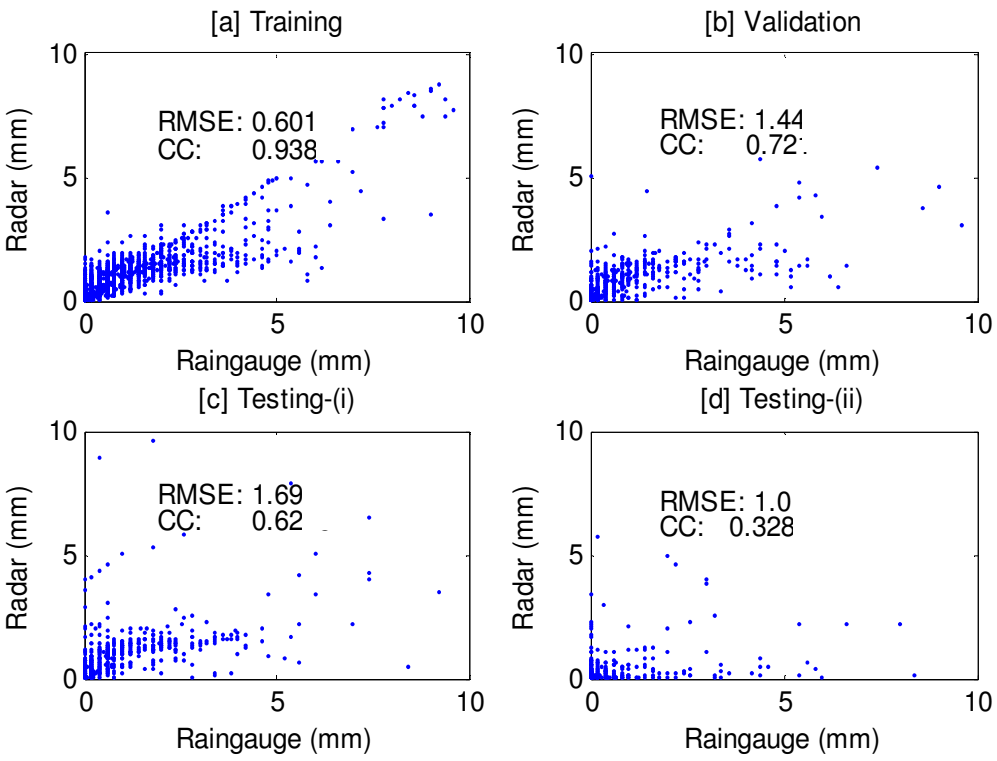


Figure 25: Comparison of 10-minute accumulations of gauge rainfall and radar estimation with the generalized regression network for the temporal input pattern.

| | |
|---|---|
| No of input neuron: | 9 |
| No of clusters: | 500 |
| Width: | 0.05 |
| No of output neuron: | 1 |

The width is varied from 0.05 to 1. It was observed that with increasing width, the model performance in terms of the correlation coefficient initially increases and then decreases. The best performance is observed at a width equal to 0.2 in 50% of cases, and equal to 0.1 in 27% of cases. Increasing the number of clusters has little impact on the model performance at widths greater than or equal to 0.2. At widths equal to 0.05 and 0.1, it is observed that increasing the number of clusters results in better performance in the validation and first test data set and worse performance in the second test data set.

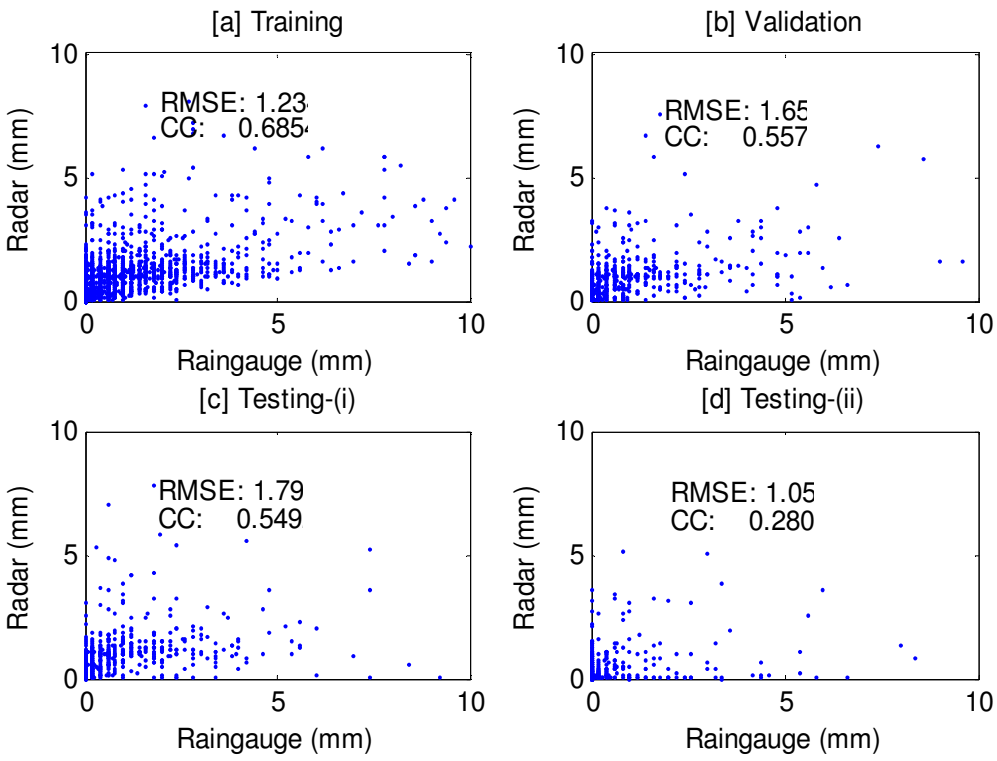


Figure 26: Comparison of 10-minute accumulations of gauge rainfall and radar estimation with the generalized regression network for the spatial input pattern.

As the optimal number of clusters is found to be 500, the dimension of the weight matrix connecting the input to the hidden layer is [500x9] and the dimension of the weight vector connecting the hidden layer to the output layer is [1x500]. The performance of the network tested is shown in Table B.6. The scatter plot of the four data sets is shown in Figure 26.

### 4.3.5 Overall Performance

So far optimum networks for three different types of neural networks with two different input data have been selected. At this point, the best overall network among these optimum networks is sought. The performance figures for the different neural networks in the case of the validation and two sets of test data are listed in Table 9.

Table 9: Performance of optimum networks.

| Net | Input | Validation | | Test set (i) | | Test set (ii) | |
|---|---|---|---|---|---|---|---|
| | | RMSE | CC | RMSE | CC | RMSE | CC |
| Back-propagation network | Temporal | 1.37 | 0.75 | 1.60 | 0.67 | 1.06 | 0.33 |
| | Spatial | 1.68 | 0.56 | 1.91 | 0.45 | 0.94 | 0.41 |
| Radial basis network | Temporal | 1.61 | 0.61 | 1.73 | 0.62 | 1.09 | 0.44 |
| | Spatial | 1.72 | 0.55 | 1.85 | 0.56 | 1.01 | 0.28 |
| Generalized regression network | Temporal | 1.44 | 0.72 | 1.69 | 0.62 | 1.01 | 0.33 |
| | Spatial | 1.66 | 0.56 | 1.80 | 0.55 | 1.05 | 0.28 |

From Table 9, it is observed that the performance of the temporal input pattern is consistently better than the spatial input pattern with the exception of the second test data set of the back-propagation network, where the spatial input pattern shows better performance.

The best network is the back propagation network with a temporal input pattern as it has the minimum RMSE in the validation data set. Table 10 shows a comparison of this model with the Z-R relationship.

From Table 10, it is observed that the selected BPNN shows improvement in the case of first test data (120%); however, the improvement in the case of second test data is not significant, suggesting the models inability to estimate new rainfall events.

In general, it is observed that the models that show better performance in the

Table 10: Comparison of the performance of the Z-R relationship and the BPNN.

| Model | Validation data | | Test set (i) | | Test set (ii) | |
|---|---|---|---|---|---|---|
| | RMSE | CC | RMSE | CC | RMSE | CC |
| Z-R (200,1.6) | 3.26 | 0.23 | 2.53 | 0.31 | 1.32 | 0.28 |
| BPNN | 1.37 | 0.75 | 1.60 | 0.67 | 1.07 | 0.33 |
| % Improvement | 57.8 | 223.6 | 36.6 | 120.0 | 19.25 | 16.63 |

validation data set, show worse performance in the second test data set and better performance in the first test data set. For example, in case of the back-propagation network with temporal input data, the correlation coefficient in the validation data set is 0.75, but the correlation in case of the second test data set is only 0.32. On the other hand, the radial basis network with temporal input data has a correlation coefficient equal to 0.61 in the case of the validation data and 0.44 in the case of the second test data set. This trend is also observed in the case of the performance of the individual groups of networks.

For example, Figure 27a shows a scatter plot of the correlation coefficients between the validation data and the first test data, and Figure 27b shows the correlations between the validation data and the second test data for different back propagation networks tested with temporal input pattern. It is observed that the correlation for the first test data is increasing with increase in the correlation of the validation data whereas the correlation of the second test data is decreasing with increasing correlation of the validation data. The x-values of these two plots are column 9 of Table B.1 and the y values of these two plots are column 11 and column 13 of the same table, respectively.

It is also observed that when the correlation of the first test data set is the maximum (=0.71), the correlation of the second test data set is only 0.38, whereas when the correlation of the second test data set is the maximum (=0.65), the correlation of the first test data set is 0.64 which may be an acceptable value. This trend is also observed for other networks.

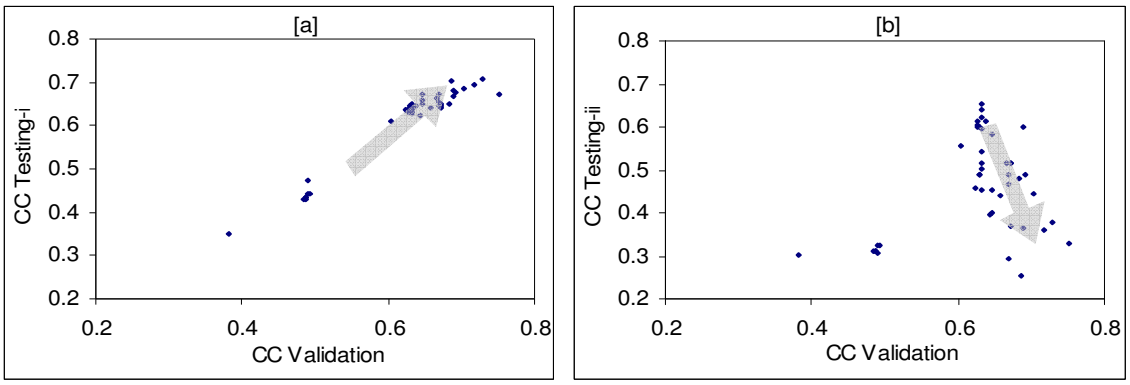


Figure 27: Comparison of correlations of validation data set with [a] test data (i) and [b] test data (ii) for back propagation network with temporal input.

It is of interest to explore the probable reason for this trend. One may recall that these networks are tested with the data from 16 gauges and five rainfall events. The validation data is from four gauges of the same rainfall events, and the first test data is another four gauge data of the same rainfall events. Hence, these data sets represent the same rainfall events at different spatial locations. On the other hand, the second test data set is a completely new rainfall event. It is well known that rainfall shows a high degree of spatial correlation up to certain geographical distance. Figure 28 shows correlation of gauge #15 with the surrounding gauges as a function of distance. This figure shows that correlation is at least 0.6 for distances up to 8 km.

Due to the high degree of spatial correlation between gauges, training, validation, and the first test data sets represent quite similar time-series of rainfall. Hence, the model may try to over-fit, thereby deteriorating its ability to estimate new rainfall events. It is also understood that a similar trend may not be observed if the experiment is carried out on another set of rainfall events.

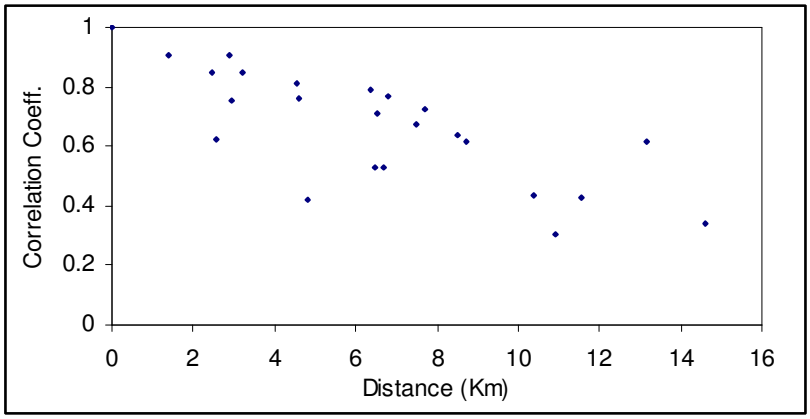


Figure 28: Correlation with Gauge #15 with other gauges as a function of distance.

Earlier works on neural network on radar rainfall estimation used different gauge locations for the same rainfall events as training, validation, and test data (Hessami et al., 2004; Hessami et al., 2003; Liu et al., 2001; Xiao et al., 1998; Xiao and Chandrasekar, 1997). This approach to training may be useful if one uses the model to obtain rainfall for other spatial locations for the same precipitation events used in the training. However, it may not be wise to use such models to obtain rainfall maps for new rainfall events without proper test with the new rainfall events.

It is of interest to check the model performance when it is validated with the new rainfall event. Hence, the neural network models are validated with the rainfall event that occurred on August 6, 2000 (second test data set) and tested with another new rainfall event that occurred on June 9, 2002. Table 11 shows the performance of the optimum models in each category for this particular event. The optimum model is selected based on the minimum RMSE in the validation data set. In this table, the first test data set is the same as previously and the third test data set is the new rainfall event that occurred on June 9, 2002.

It can be observed that increasing the performance of the validation data set ensures better performance in the case of the third test data set as well as the first test data set. However, the performance in the case of the first test data set is slightly worse than the previous validation set (Table 9). This suggests that if a new rainfall event is used as validation data set, the resulting model may be useful for getting rainfall maps of the rainfall used in the training purposes as well as of new rainfall events. It is observed from Table 11 that the temporal input pattern yields better result for all networks in this case too.

The best network according to Table 11 is the generalized regression network using the rainfall that occurred on August 6, 2000 as validation data set. The parameters of this network is:

| | |
|---|---|
| No of input neurons: | 2 |
| No of clusters: | 50 |
| Width: | 0.1 |
| No of output neurons: | 1 |

The weights of the hidden layer and the output layer for this network are shown in Table A.7.

Table 11: Reconstruction of Table 9 using August 6$^{th}$ rainfall as the validation data set.

| Net | Input | Validation | | Test set (i) | | Test set (iii) | |
|---|---|---|---|---|---|---|---|
| | | RMSE | CC | RMSE | CC | RMSE | CC |
| Back-propagation network | Temporal | 0.81 | 0.62 | 1.70 | 0.63 | 0.45 | 0.80 |
| | Spatial | 0.93 | 0.44 | 1.89 | 0.48 | 0.54 | 0.70 |
| Radial basis network | Temporal | 0.84 | 0.60 | 1.67 | 0.64 | 0.47 | 0.79 |
| | Spatial | 1.00 | 0.28 | 1.85 | 0.56 | 0.51 | 0.74 |
| Generalized regression network | Temporal | 0.80 | 0.63 | 1.74 | 0.62 | 0.45 | 0.80 |
| | Spatial | 0.96 | 0.36 | 1.98 | 0.42 | 0.57 | 0.69 |

Table 12: Comparison of the performance of the Z-R relationship and the GRNN.

| Model | Test set (i) | | Test set (ii) | | Test set (iii) | |
|---|---|---|---|---|---|---|
| | RMSE | CC | RMSE | CC | RMSE | CC |
| Z-R (200,1.6) | 2.53 | 0.31 | 1.32 | 0.28 | 0.59 | 0.74 |
| GRNN | 1.74 | 0.62 | 0.80 | 0.63 | 0.45 | 0.80 |
| % Improvement | 31.34 | 101.5 | 39.3 | 123.2 | 23.8 | 9.09 |

It should be emphasized, however, that the performance of the radial basis function network and the back-propagation network for temporal input pattern is quite similar to the generalized regression network. It is of interest to check the performance of this network with the performance of the Z-R relationship for the three test data sets.

Table 12 shows 9% improvement in the third test data set. The improvement is significant in the case of the first and the second test data sets, but this is not surprising as the first test data contains the rainfall used in the training and the second data is the validation data itself. Hence, the improvement over the Z-R relationship is not significant in the case of estimation of a new rainfall event. Figure 29 shows the rainfall accumulation map (10-minute) of a scan from the back-propagation network and the Marshall-Palmer's Z-R relationship. This rainfall is appreciably underestimated by radar compared to the gauge data.

Neural network may be an appealing choice if one trains a model using several gauges for a rainfall event and wants to obtain rainfall estimate for other geographic locations of the same rainfall event. This conclusion is based on the fact that the networks shown in Table 10 and Table 12 show significant improvement in the first test data set (120% and 101%, respectively). Hence, this study recommends using neural networks as

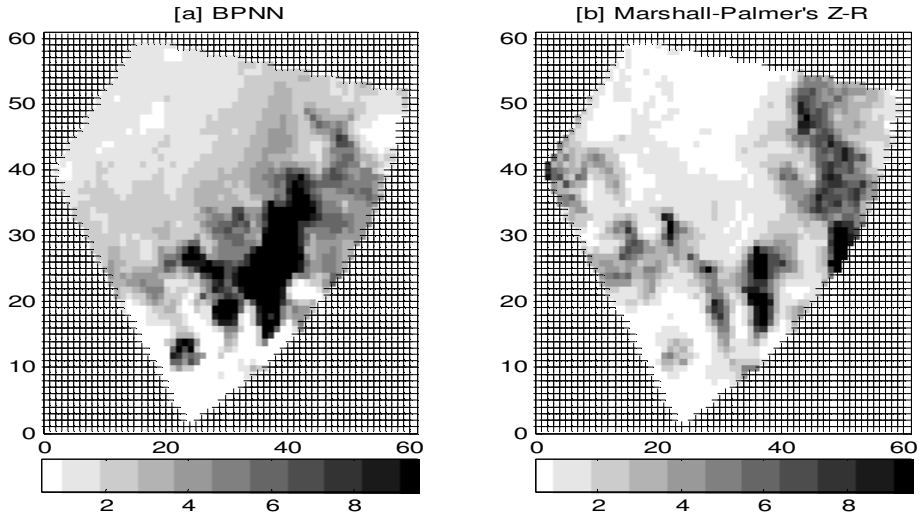

Figure 29: The rainfall accumulation map (10-minute) obtained from [a] BPNN, [b] The Marshall-Palmer's Z-R relationship, (July 7, 2000, 9:00 UTC).

a post-processing tool to obtain rainfall maps of a precipitation event, using gauge data for that event to train and test the network.

### 4.4 SUMMARY

This chapter has investigated the use of neural networks to improve the interpretation of reflectivity. It was found that neural networks may be useful as a post-processing tool, that is, a trained network should be used only for the same rainfall event as used in the training. Its performance is not much better than the Z-R relationship for interpreting new rainfall events. However, this conclusion is drawn based on the high temporal resolution mapping (10-minute accumulations). Another recommendation is to use time series as input data as this data type yields better result compared to input of concurrent reflectivities at surrounding spatial points.

# Chapter 5

# Correction Schemes

## 5.1 INTRODUCTION

The objective of this chapter is to apply various correction schemes to the Z-R relationship in order to achieve a better correlation between gauge and radar accumulations at high time resolution (10-minute). The sensitivity of each scheme is analyzed. It is also attempted to generate precipitation maps at one minute time resolution to reduce the uncertainty in the computation of rainfall accumulation associated with small scale convective cells. The generated maps are passed through a trajectory model which derives the precipitation map at the earth surface.

## 5.2 TIME INTEGRATION

### 5.2.1 Motivation

Whenever one tries to compare gauge data with radar data, a problem pertaining to the temporal structure of the data of these two sensors arises. The gauge provides rainfall accumulation on the ground while the radar provides rainfall intensity at a certain distance above the ground. It is often of interest to estimate rainfall accumulation from radar data. This is usually done by assuming the rainfall intensity is constant within the

radar's temporal resolution. This assumption may be valid for stratiform rainfall which shows small spatial gradients of echoes. However, for thunderstorms this assumption may not be valid for 10-minute resolution. The reason is that the single-cell thunderstorms usually last for 20-30 minutes (http://ww2010.atmos.uiuc.edu/(Gh)/wwhlpr/single). Although multicell thunderstorms developed in a frontal passage system may last for several hours, an individual cell usually decays quickly and a new cell is developed in a preferred direction. The cluster of cells forms a convective line which may persist for several hours or even days. Such storms exhibit steep spatial gradients of reflectivity which introduces uncertainties due to the assumption of uniform intensity within 10-minute cycles. This is further explained with an example.

Figure 30 shows two consecutive reflectivity scans over gauge #1. The first scan (June 10, 2000, 8:00 UTC) shows a cell south-west of the gauge location and the second scan (June 10, 2000, 8:10 UTC) shows that the cell has moved to the north-east direction of the gauge location. These cells are marked by a white ellipse and the gauge is shown with the symbol 'o'. The reflectivities above the gauge location in the two scans are found to be 28 and 34.4 dBZ, respectively, from the nearest neighbour interpolation. The peak intensity of the cell passing over the gauge is found to be 50.2 dBZ. Hence, the intensity at any time between the two scans, e.g. at 8:05 UTC, may be quite different from the value of 28 dBZ and 34.4 dBZ or linear average of the two; it may be as high as 50 dBZ. The gauge measures the true accumulation at this point, so one should expect significant differences from the radar measurement.

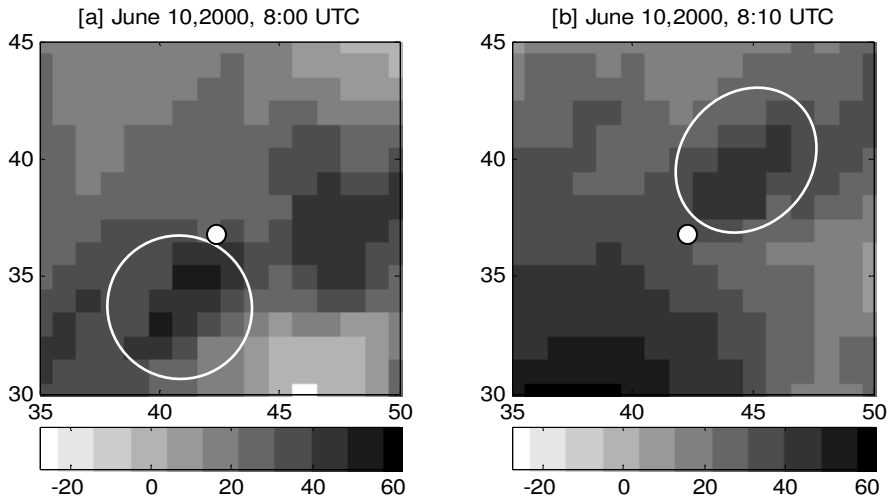


Figure 30: Two consecutive reflectivity scans over gauge #1 (shown with 'o' symbol, latitude 49° 54' 11.38", longitude -96° 59' 5.19"). The axes are distance in kilometre.

A coarser time resolution may increase the correlation between the gauge and the radar measurement. The time resolution of radar data may also be increased by tracking the convective cells and locating their geographic position at intermediate times between successive scans. Such an algorithm is described by Hannesen (2002).

Tracking convective cells for precipitation forecasting of short lead time (also known as now-casting) and for estimating wind motion has been practiced since 1960 (Li et al., 1995) and is often used by airport authorities (Wolfson et al., 1995). Hence, conventional tracking algorithms may be used to increase time resolution of the radar data.

### 5.2.2 The Tracking Algorithm

Rinehart and Garvey (1978) describe a method to track the precipitation based on correlation to derive the motion of the small scale patterns. This algorithm is closely related to algorithms used for identifying patterns in images. In this method, a 10km×10km area is selected and this area is translated over a region of the radar scan of

the next time step. This area is termed the 'window'. For each position of the window over the radar scan during the translation, the correlation coefficient is determined between the window and the 10 km×10 km area of the radar scan over which the window is placed is determined. The position of the maximum correlation coefficient is determined. It is assumed that the storm cell has moved to the position where the correlation coefficient attains its maximum.

It is noteworthy that during multi-cell thunderstorms, convective cells may continuously grow on one side of the convective line and decay on the other side which creates an apparent motion which may be different from the large scale wind direction. This apparent movement is known as 'propagation' (Li and Schmid, 1995). Raghabhan (2003) explained that when a convective cell reaches its mature stage, the downdraft from it generates new cells. Thus old cells decay and new cell emerge creating an apparent motion. The tracking algorithm described above is suitable for determining the apparent direction of the convective cells but may not be suitable for deriving the actual direction of the precipitation system.

Browning (1979) developed an algorithm to track the 'storm envelope' instead of individual cells. He degraded the reflectivity map to 20 km×20 km scale from the original 5 km×5 km scale. This smoothing operation helps to track the large scale movement. Wolfson et al. (1999) introduced an elliptical filter to track the large scale features.

This study assumes that the movement of the large scale features is more representative of the mean motion of the precipitation system. To track the large scale features Rinehart and Garvey (1978)'s procedure is followed, except that a 75 km×75 km

window is selected. The rainfall events selected for this study are all synoptic scale in nature and covers the whole radar umbrella. This 75 km×75 km window is translated over a 100 km×100 km area of the next radar scan. It is assumed that storms will not move more than 15 km in a 10-minute time step. Mathematically, the cross-correlation is calculated as,

$$C(p,q) = \sum_{i=-\frac{I-1}{2}}^{\frac{I-1}{2}} \sum_{j=-\frac{J-1}{2}}^{\frac{J-1}{2}} W(i,j) S(p+i, q+j) \qquad \text{Eq 5.1}$$

where $W$ is the window matrix, $S$ is the radar scan of the next time step, and $C$ is the cross-correlation matrix. The size of the window is [$I$x$J$]. Both $W$ and $S$ are reflectivity matrices in dBZ.

Figure 31 shows the time series of movement vectors of precipitation. The length of an arrow at a time reflects the movement of the precipitation system at that time and its angle from the vertical reflects the angle from the north direction. This vector plotting shows that the change in movement vectors is gradual and consistent with the movement vector of the previous time step. The exception is the July 4$^{th}$ rainfall during which two vectors show an abrupt change in direction (00:20 and 01:20 UTC). Such anomalous movement may be corrected by comparing the current direction with the direction obtained in the previous time step, and if the change of direction is greater than a threshold (e.g. 45$^{O}$), the direction obtained in the previous time step may be used. This simplified filtering approach is used in this study. Alternatively, a Kalman filtering approach may be recommended to predict and update the direction vector at each time step. Another observation is that the precipitation system is not changing its direction more than 90$^{o}$ within a rainfall.

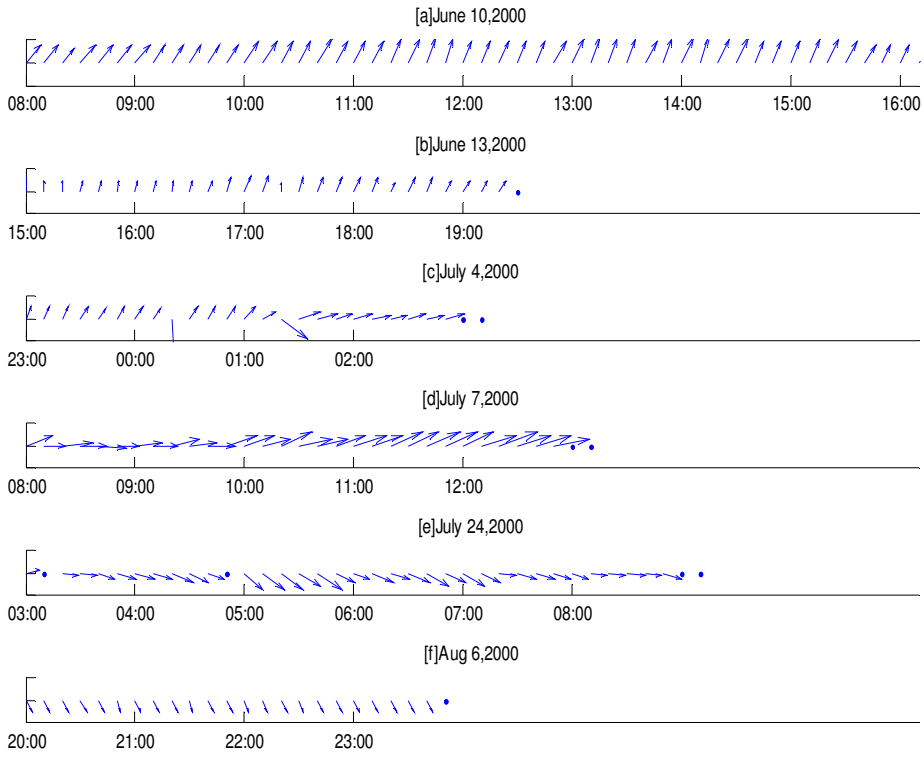


Figure 31: Time series of movement vector of rainfall. Time is in UTC. The first vector represents 7.5 km and $45^0$ from North.

The movement of precipitation can also be derived from the Velocity Azimuth Display (VAD) of the radar data. VAD provides velocity at given elevation and the displacement in a 10-minute time period can be computed by multiplying by an appropriate factor.

Figure 32 shows the scatter plot of the magnitude and direction of precipitation movement derived from the correlation based tracking and from the VAD data corresponding to 4 km altitude. It is assumed that precipitation is generated from the 4 km altitude. The linear association suggests consistency between the two methods. The variations in the direction measured from the two procedures are found within $30^o$.

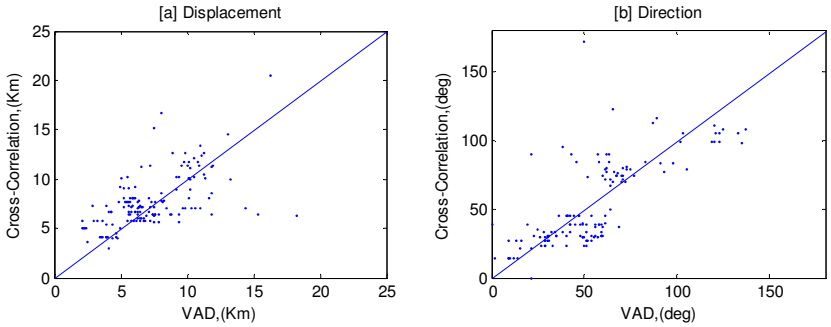


Figure 32: Comparison of precipitation [a] displacement and [b] direction obtained from correlation tracking (without filtering) and from VAD data.

### 5.2.3 Space-Time Interpolation

Once the movement of precipitation has been determined by applying the tracking algorithm, it is easy to generate the reflectivity maps within each 10-minute time period. The reflectivity map at an intermediate time can be generated using the method described by Hannesen (2002). If the *X* and *Y* component of the movement vector of the precipitation system at any time *t* is *dx* and *dy* and if the time interval between two successive scan is *dt,* the precipitation map at any intermediate time ($t+\eta dt$) can be generated using the following expression,

$$dBZ^{t+\eta dt}_{x+\eta dx,\, y+\eta dy} = (1-\eta)dBZ^{t}_{x,y} + \eta dBZ^{t+dt}_{x+dx,\, y+dy} \qquad \text{Eq 5.2}$$

where $\eta$ is a coefficient whose value varies from 0 to 1. Using this equation, nine intermediate reflectivity maps have been generated for each 10-minute cycle. Here, $dBZ^{t}_{x,y}$ is the reflectivity at a location ($x,y$) and time *t* is associated with the 1.5 km CAPPI.

Figure 33 is an extension of Figure 30 with two additional scans generated by space-time interpolation. The generated scans are shown in Figure 33c and Figure 33d. It is observed that the convective cell is passing over the gauge (marked with white circle)

which suggests that the interpolation should provide more accurate rainfall accumulations. Figure 34 shows one-minute rainfall accumulations over gauge #1, within the 10-minute time interval. This figure shows significant increase in rainfall within the time range due to the passing of the cell.

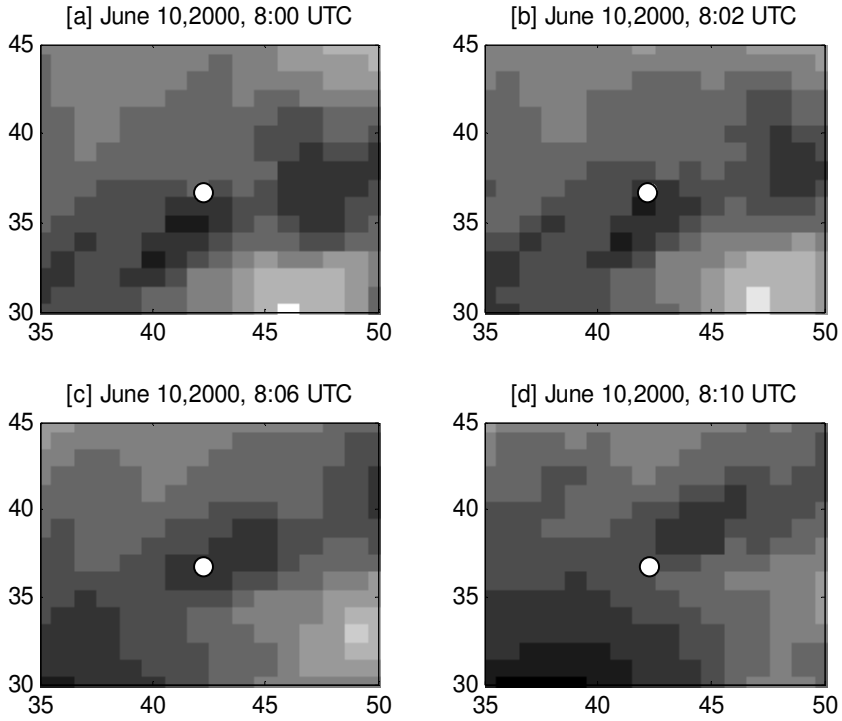


Figure 33: Convective cell is passing over the gauge #1, [a] and [d] is from conventional scan, [b] and [c] is generated scan from space-time interpolation. The axes show distance in km. The location of gauge is marked with white circle.

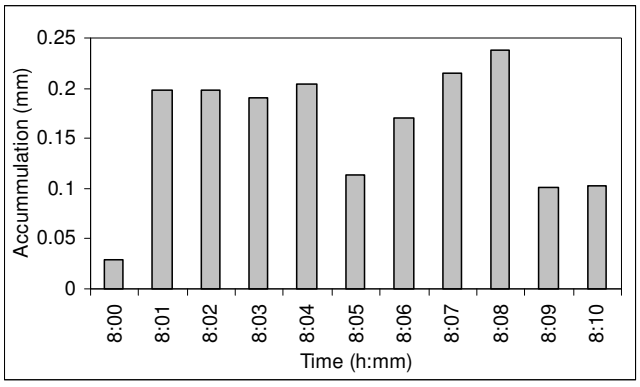


Figure 34: The one minute radar rainfall accumulations over gauge #1 for the June 10th rainfall from 8:00 to 8:10 UTC.

### 5.2.4 Time Synchronization

Accurate determination of the time coordinates of pixels in the CAPPI appears important in order to establish a good correlation for 10-minute accumulations. Although a CAPPI is stamped with a time, it does not reflect the rainfall rate at that time. For example, the radar used in this study takes 10-minute to complete a cycle. During the first-half of this time span, it operates in conventional mode and in the second-half it operates in doppler mode (Patrick 2005, personal communication). It completes scanning the 24 elevation angles in the conventional mode and the four elevation angles in the Doppler mode. These scannings are stamped at the median time of each 10-minute cycle. For example, if a scanning product shows time stamped at 10:20 am, it means that the conventional scannings lasted from 10:15 to 10:20 am and the doppler scan lasted from 10:20 to 10:25. As the conventional scanning schedule is started from the highest elevation angle (Donaldson 2005, personal communication), rainfall intensity at locations closest to the radar site is five minutes older than the time stamped at the CAPPI and the rainfall intensity at locations farthest from the radar site is the nominal.

The Winnipeg city raingauges are located within a 50 km to 70 km radial distance from the Woodland radar. Out of 24 complete rotations, the 1.5 km CAPPI for this study area will be captured from the $19^{th}$ and $20^{th}$ rotations (Figure 5). This rotation will occur at the beginning of the $5^{th}$ minute of each 10-minute cycle. If the time stamp of a CAPPI of this study area shows 10:20 am (hh:mm), it can be calculated that this portion of the CAPPI shows rainfall intensity in the time interval ranging from 10:18:57 to 10:19:22 (hh:mm:ss). For simplicity, it may be assumed that the CAPPI of the study area shows the rainfall echo at time 10:19 am when time is stamped as 10:20 am.

It was mentioned earlier that Bodiroga (2004) prepared 10-minute gauge accumulations for Winnipeg's rain gauges which are used in this study. To synchronize the time between gauge accumulations and radar accumulations, it is necessary to compute radar accumulations from the second scan, which is a generated scan from the space-time interpolation. By summing up radar accumulations from 10 consecutive 1-minute rain fall at a geographic location, it is possible to get 10-minute rainfall accumulations from the radar.

### 5.2.5 Results and Discussions

It is of interest to check the correlation of 10-minute gauge accumulations and 10-minute accumulations computed from the radar using the aforementioned procedures. Figure 35 shows the comparison between the radar and the gauge accumulations (10 minute). The performance is measured in terms of the correlation coefficient (CC) and the bias. The bias is the ratio of the summation of all gauge accumulations to the summation of all radar accumulations. Obviously, radar accumulations are computed only at the gauge locations. The value of the correlation coefficient and the bias is shown on the respective figures.

The correlation coefficient measures the strength of linear association; its value may be close to 1 even if the radar appreciably underestimates or overestimates the precipitation. Hence, a good agreement between the gauge and radar should be ensured if both of these parameters have values close to 1. Comparing these values with the tabulated values in Table 6, it is observed that the correlation coefficient is significantly increased in all cases. Figure 36 shows the bar diagram of the correlation coefficients of two methods of time integration. Here method 1 refers to the method adopted in Chapter

3, that is, rainfall intensity obtained from radar is assumed constant within 10-minute cycles and the accumulation is computed from this intensity, and method 2 is the computed accumulations following the tracking algorithm, space-time interpolation, and accurate time synchronization. The correlation coefficient for the whole ensemble of six rainfalls is 0.3 and 0.57, respectively, for the two methods which implies a 90% improvement. It should be mentioned that Marshall-Palmer's relationship is used in both estimations.

It is observed that radar underestimates the June 13$^{th}$, the July 4$^{th}$ and the July 7$^{th}$ rainfall and overestimates the July 23$^{rd}$ rainfall event.

The space time interpolation is used to get the rainfall intensity at each minute of a 10-minute cycles. It is of interest to explore which intensity is the most representative relative to the gauge accumulation. Figure 37 shows the correlation between gauge accumulations and radar accumulations. Nine scans are generated for each 10-minute cycle using the space-time interpolation. Each bar shows the correlation between the gauge accumulation and the set of each generated radar scan. The 10$^{th}$ bar shows the correlation of the original radar scan with the gauges. It appears that the first seven scans show better correlation with the gauge accumulations. Thereafter, the correlation begins to fall sharply. One possible reason behind this trend may be the time lag for a rainfall mass from the atmosphere to reach the ground. This study attempts to compute travel time of a rainfall mass from 1.5 km CAPPI to the ground. Several sources are available to estimate fall velocity of hydrometeors and once fall velocity has been computed, the travel time can be computed by dividing the 1.5 km distance by the fall velocity.

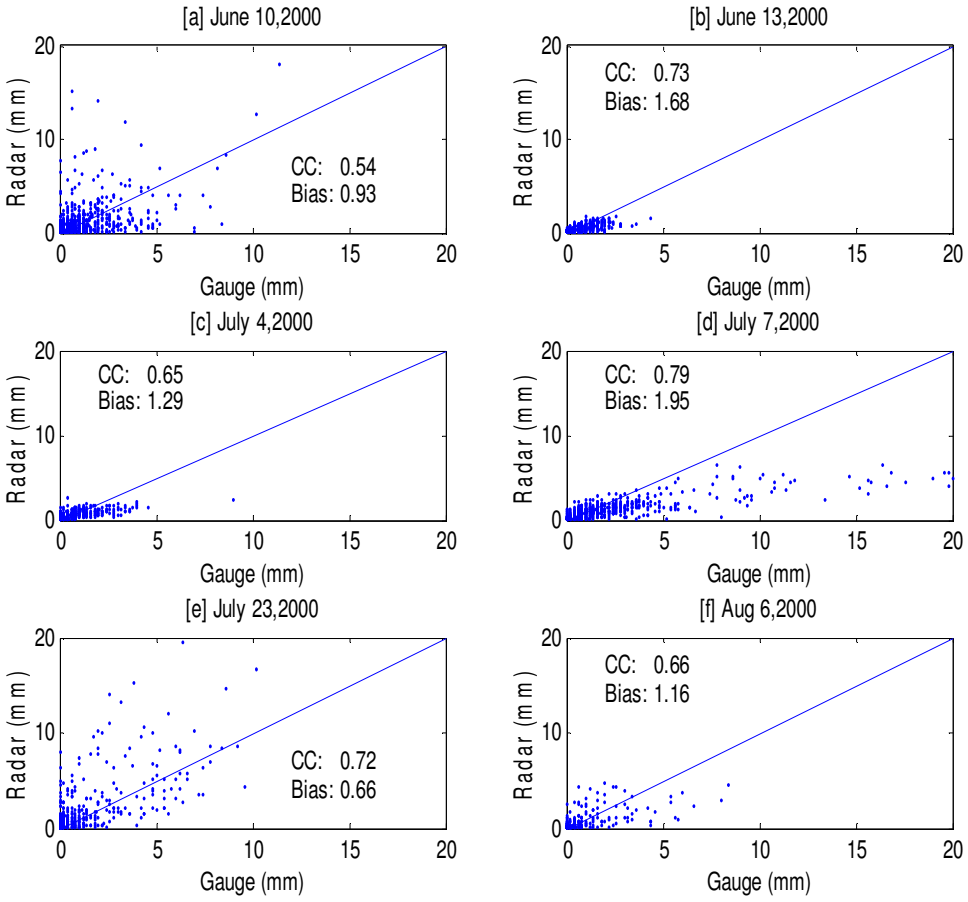


Figure 35: Comparison of 10-minute gauge and radar accumulations for the six rainfall events. Radar accumulation is computed using generated scan from space-time interpolation.

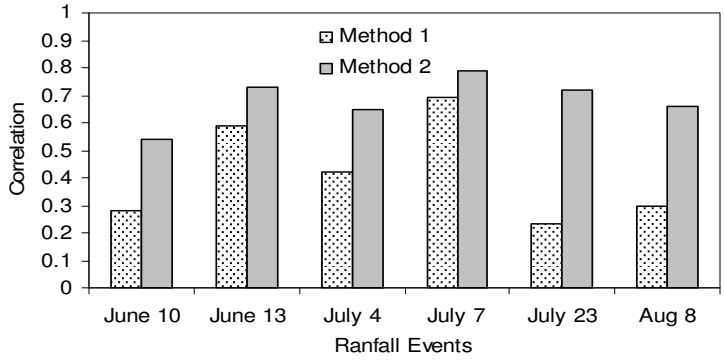


Figure 36: Comparison between two methods of time integration.

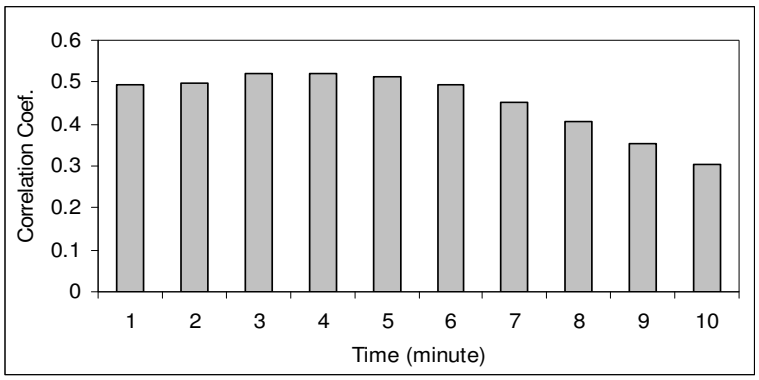


Figure 37: Correlation between gauge accumulation and radar intensity. Each bar shows correlation of a set of radar scan with gauge accumulation.

Ahrens (1994) tabulated fall velocities of raindrops having different diameters. This table suggests that the travel time for small raindrops, typical raindrops, and large raindrops should be 6.26 minutes, 3.83 minutes, and 2.78 minutes, respectively. The fall velocity may also be computed using the equation proposed by Atlas and Ulbrich (1977) as shown in Eq 2.3. It is noteworthy that most sources for estimating fall velocity of raindrops do not consider the effect of updraft and downdraft. Based on the literature, it may be concluded that a rainfall mass at the 1.5 km CAPPI may take 2 to 6 minutes to reach the ground. Hence, rainfall at the eighth scan and onward of each 10-minute cycle contributes to the next 10-minute accumulations. This may be the reason that the correlation sharply falls after the seventh scan of 10-minute cycles. Obviously, in the case of hourly accumulations this should not be a serious issue. If the timing of the radar scan falls on the last minute of a 10-minute cycle, only one scan out of six scans should contribute to the gauge accumulation of the next hour. Hence, a better correlation should be expected. This may be the reason for getting better correlations with hourly accumulations as described in Section 3.6.5. It is further observed that gauge to radar correlations for hourly accumulations using Marshall-Palmer's relationship is 0.55 which

is fairly close to the correlation obtained for the first several scans of each 10-minute cycle.

It is also observed that the correlation for the ensemble of the first six generated scans of each 10-minute cycle is close to 0.5 whereas the correlation for the accumulation algorithm using space time interpolation techniques is 0.57. Obviously, one may question the necessity of applying this complicated algorithm when accumulations computed from a single scan shows fairly close correlation. The answer is that the necessity of this algorithm may be case specific. For example, this algorithm should be useful for fast moving isolated cell. If a 50 km radius cell moves 10 km in 10 minutes, this algorithm may not be useful. On the other hand, if a 5 km radius cell moves 10 km in 10 minutes, this algorithm appears essential for computing accurate rainfall accumulations.

It may be concluded that if one wants to avoid the use of a complicated space-time interpolation algorithm without significantly loss of accuracy, one should (i) exactly measure the time when the study area is scanned by the radar, and (ii) determine the time cycle/interval for gauge accumulations so that the radar scan times falls within the first half of the time interval used for the gauge accumulation.

## 5.3 TIME SHIFTING

### 5.3.1 Motivation

Instead of comparing radar and gauge accumulations at the same time point, one may be interested in keeping a time lag between the two accumulations to account for the travel time from the atmosphere to the ground. This travel time will generally be affected by localized updraft/downdraft and accurate determination is not possible without dual-

doppler scanning. Despite this fact, one may attempt to determine a 'climatological' travel time. In this study, the correlation between gauge and radar accumulation is determined by examining time lags from zero to 10 minutes. The generated scans from the space-time interpolation are used for this purpose. In Figure 38, each bar shows the gauge and radar correlation with time lags shown on the *x* axis. This figure shows that the correlation at two minutes lag is 0.6 whereas correlation at zero lag is 0.57. However, absence of a sharp peak probably indicates that travel time is varying in a wide range.

It should be expected that allowing for travel time may improve the rainfall estimation at the earth's surface from radar observation. This study makes an attempt to generate the reflectivity/rainfall map at the surface from the atmospheric radar observation considering time shifting. The travel time should vary depending on the drop-diameters and the vertical velocity of the air due to updraft/downdraft. Roger's (1964) provides an expression for computing fall velocity based on reflectivity which is

$$V_f = 3.8Z^{0.072} \qquad \text{Eq 5.3}$$

where $Z$ is in $mm^6/m^3$ and $V_f$ is the fall velocity in m/s. The travel time ($t_f$) for a pixel is computed by dividing the elevation of the CAPPI by the fall velocity ($V_f$).

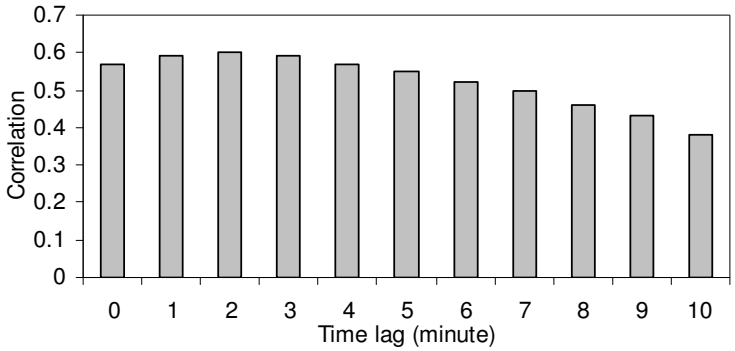


Figure 38: Correlation between gauge and radar accumulations by keeping time lag.

This study uses Roger's (1964) expression to compute travel time. Each pixel will have a particular $t_f$ value, reflecting the fact that different Z-values will yield different $t_f$. Hence, an algorithm is developed that takes the generated CAPPI at one minute resolution as input, computes the $t_f$ for each pixel, transfers each pixel to the earth surface by changing its time coordinate and keeping the space coordinate the same and finally computes the rainfall accumulation at any space coordinate for any suitable time interval. In fact, this algorithm generates a new set of CAPPI's which represents the reflectivity at the earth surface. It appears that as the travel time is in the order of 2 to 6 minutes, the original radar scans having 10-minute resolution are not useful for this model. Hence, the generated scans having one minute resolution are used. The generated scans at the earth surface can be represented by reflectivity (*dBZ*) or by rainfall rate using a Z-R relationship or by rainfall accumulations at one minute resolution. This model can be expressed as follows,

$$R^{t+t_f}_{x,y,z_o} = R^{t}_{x,y,z_a} \qquad \text{Eq 5.4}$$

where $t_f$ is the travel time taken by the rainfall mass to reach the ground, $z_a$ stands for elevation of CAPPI, and $z_o$ stands for earth surface. It is worth noting that a situation may arise where two reflectivities having the same space-coordinate but two different time coordinates may reach the ground at the same space-time coordinate. In such case reflectivity may not be additive. Hence, each reflectivity is converted to rainfall accumulation and the accumulations are added together.

### 5.3.2 Results and Discussions

Figure 39 shows the scatter plot of the 10-minute gauge and radar accumulations after applying the time-shifting algorithm integrated with the space-time interpolation

algorithm. This figure still shows significant underestimation in case of the June 13th, the July 4th, and the July 7th rainfall events and overestimation in case of the July 23rd rainfall event. Indeed this correction scheme should not change the bias as the total accumulation measured by the gauge and radar will not change.

Figure 40 shows the correlation coefficients of six rainfall events with and without consideration of time shifting. It is observed that the correlation coefficient increases in four rainfall events and decreases in two events. This behaviour is not unexpected as computation of travel time does not consider the updraft/downdraft which may alter the fall velocity. Despite this limitation, the correlation of the ensemble is increased to 0.59 from 0.57 which implies a 3.5% improvement.

## 5.4 WIND DRIFT

### 5.4.1 Motivation

The usual practice is to compare the gauge accumulation with the radar intensity located vertically above the gauge location. The underlying assumption of this practice is that wind is not advected more than the radar's spatial resolution. For high spatial resolution radars (e.g. 1 km×1$^{o}$), this assumption may not be valid. If it is valid, the correlation between gauge and radar pixel located overhead should be higher than the correlation obtained from surrounding radar pixels.

Figure 41 shows contour of correlations between gauges and surrounding radar pixels for the first two hour of the rainfall occurring on June 10, 2000. For example, the value of a pixel with coordinates (2, 4) is the correlation between gauges and radar pixels located at 2 km East and 4 km North from the gauge location.

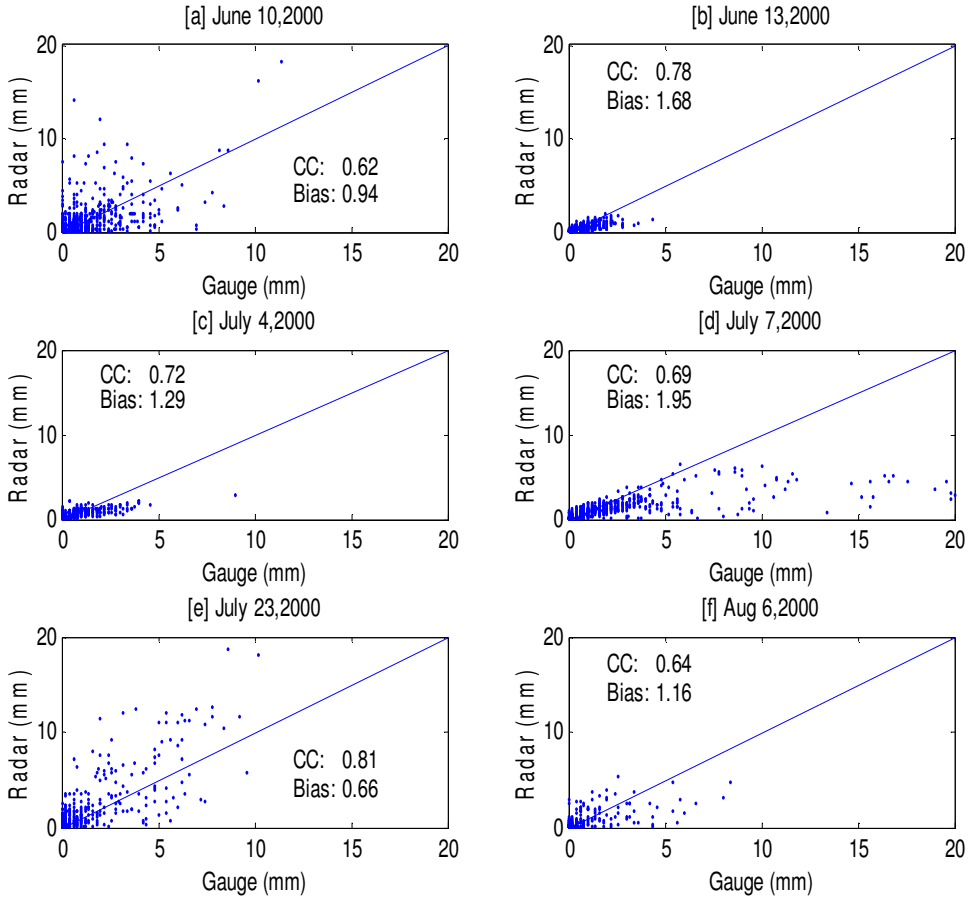


Figure 39: The comparison of 10-minute accumulations of gauge and radar data after considering travel time of the rainfall mass.

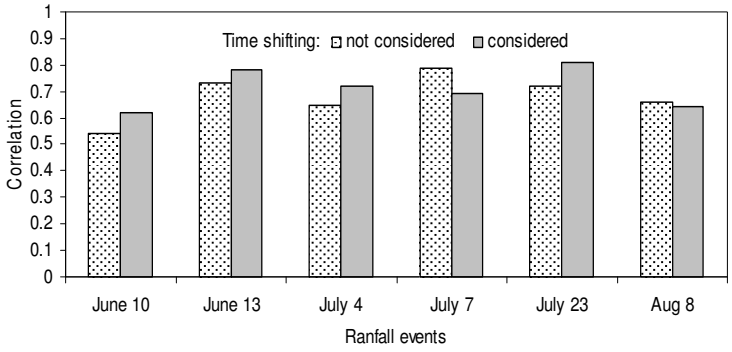


Figure 40: The correlation coefficients of the six rainfall events with and without considering travel time.

The radar value at a particular position is determined using triangular based linear interpolation. The correlation is calculated using all gauges. The figure shows that the correlation between gauge and overhead radar pixel (0, 0) is 0.5 whereas correlation at pixel (0, 3) is 0.8. This may imply that rainfall is advected 3 km to the west during this time period. The white spaces in this figure refer to correlations below 0.4. It is observed that regions having correlation above 0.4 takes an elliptical shape aligned along a $45^{o}$ line. The reason for this shape may be that rainfall passed over the study area from the south-west direction. This direction was already observed in Figure 31a derived from the correlation tracking algorithm. The shape of the convective line passing over the area is elongated and narrow in width (Figure 13a). Rainfalls show spatial correlation between surrounding pixels to a considerable extent. Hence, pixels in the neighbourhood of the maximum correlation also show higher correlations creating a 'shadow' of the convective line passing over the region.

It is observed from the VAD profile that the wind changes its direction from the north-east to the east from the earth's surface to 1.5 km elevation. This change of direction with elevation may also be responsible for spreading the high correlation regions. It is also observed that the wind speed is the maximum (around 20 m/s) at 0.5 km elevation and wind direction at this elevation is from the east. Above and below 0.5 km, the wind speed gradually decreased. This may be the reason for getting the peak of correlations from east. If the average wind speed is 15 m/s, it is possible that the precipitation may travel 2.7 km in 3 minutes. Hence, it is reasonable to get the peak of correlation at a distance of 3 km. Other reasons that may affect the advection is the radial outflow of wind during downbursts, change of wind speed and direction with time, etc.,

which may also have some contribution to the spread of the maximum correlation region. If wind speed and direction were constant at all space and time coordinates, one should expect a sharper peak.

### 5.4.2 The Trajectory Model

This study attempts to develop a trajectory model in order to get the locus of the precipitation mass from 1.5 km CAPPI to the ground. Trajectory models are widely used in the atmospheric sciences to track air pollutants, smoke, etc. (Draxler, 1996). The proposed model is similar to pollutant trajectory models except dispersion is not accounted for as it is not appropriate for precipitation. The first attempt in this direction was probably made by Lack and Fox (2004). A description of this algorithm is provided in Section 2.5.11. The algorithm developed in this study needs three dimensional gridded velocity components and can be described in three steps:

Step 1: A suitable time step *dt* is chosen. The *x, y,* and *z* components of displacement for the first iteration is computed using the following expressions:

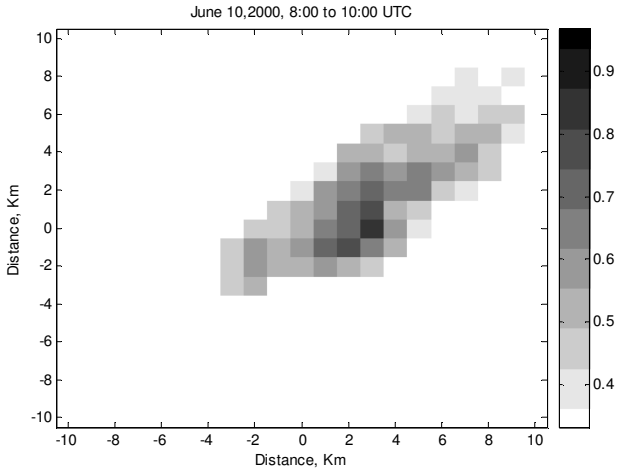


Figure 41: Contour of correlation between gauges and surrounding radar pixels.

$$\begin{aligned} dx^1 &= u^t_{x,y,z} dt \\ dy^1 &= v^t_{x,y,z} dt \\ dz^1 &= w^t_{x,y,z} dt \end{aligned} \qquad \text{Eq 5.5}$$

where $u$ and $v$ are the two horizontal components and $w$ is the vertical component of the precipitation movement vector. Here $w$ refers the net vertical movement per unit time due to the fall speed of the hydrometeors and due to updraft/downdraft. It is also assumed that the velocity components remain constant within the distance traveled by the precipitation mass in time step $dt$. This assumption should be valid if $dt$ is small.

Step 2: The space-time coordinates of the rainfall mass after first iteration are as follows:

$$R^{t+dt}_{x+dx^1, y+dy^1, z-dz^1} = R^t_{x,y,z} \qquad \text{Eq 5.6}$$

where $R$ is the radar rainfall rate. $dz^1$ is subtracted from $z$ to indicate decreasing elevation. At this stage, a radar pixel at 1.5 km CAPPI is getting a new space-time coordinate after time $dt$ without changing its value.

Step 3: The velocity components at this new space-time coordinate is then retrieved from the 3D gridded velocity components and used to compute horizontal and vertical displacement for the next time step. This process is repeated until the pixel reaches the ground (e.g. $z$-$\sum dz^i$ =0).

In other words, the whole system can be represented in a single equation as follows,

$$R^{t+ndt}_{x+\sum_{i=1}^{n} dx(i),\, y+\sum_{i=1}^{n} dy(i),\, z-\sum_{i=1}^{n} dz(i),} = R^{t}_{x,y,z} \qquad \text{Eq 5.7}$$

where $n$ is the number of iterations needed for the rainfall mass to reach the ground. The expressions to compute $dx(i)$, $dy(i)$, and $dz(i)$ are as follows:

$$\begin{aligned} dx(i) &= u^{t+(i-1)dt}_{x+dx(i-1),\, y+dy(i-1),\, x-dz(i-1)}\, dt \\ dy(i) &= v^{t+(i-1)dt}_{x+dx(i-1),\, y+dy(i-1),\, x-dz(i-1)}\, dt \\ dz(i) &= w^{t+(i-1)dt}_{x+dx(i-1),\, y+dy(i-1),\, x-dz(i-1)}\, dt \end{aligned} \qquad \text{Eq 5.8}$$

The initial values for $dx$, $dy$, and $dz$ (at $i$=1) is set to zero, that is, $dx(0)$=0, $dy(0)$=0 and $dz(0)$=0. The objective of this model is to generate rainfall accumulation map at the ground level based on the reflectivity at the 1.5 km CAPPI.

### 5.4.3 The Velocity from Radar

To simulate this model, one needs to have the 3D gridded velocity components. Doppler radars measure only the radial velocity. The $u$, $v$ and $w$ components of velocity of a rainfall mass in the atmosphere can be related to the radial velocity $V_r$ using the following relationship (Lhermitte, 1970),

$$V_r = u\sin\varphi\cos\theta + v\cos\varphi\cos\theta + w\sin\theta \qquad \text{Eq 5.9}$$

where $\varphi$ and $\theta$ are azimuth and elevation angle, respectively. The azimuth is computed from north and clockwise direction.

In Eq 5.9, $u$, $v$ and $w$ are unknowns. Hence, three equations are required to determine the values of $u$, $v$, and $w$ at a space coordinate. This can be achieved by scanning a space by three doppler radars concurrently. It is also possible to retrieve velocities using two doppler radars by using the equation of continuity in the following form:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \qquad \text{Eq 5.10}$$

The boundary condition to solve this system of equations is that the vertical velocity at storm top or at ground level is zero. The velocity components cannot be retrieved from single doppler radial velocity. Despite this fact, Lhermitte and Atlas (1961) described a method to get velocity components assuming uniform flow at an elevation, which is known as Velocity Azimuth Display (VAD). In this method, wind direction and magnitude are constant at a fixed elevation but can vary with elevation.

The radial velocities at an elevation plotted against the azimuth angle should produce a sinusoidal curve if wind is spatially uniform at that elevation. The azimuth corresponding to the maximum positive amplitude of this sinusoid should represent the wind direction (with positive implying movement away from the radar). The radial velocity at this azimuth can be determined by the following expression:

$$V_r = V_h \cos\theta + w\sin\theta \qquad \text{Eq 5.11}$$

where $V_h$ is the resultant horizontal velocity [$=\sqrt{(u^2+v^2)}$] and $w$ is the vertical velocity. The radial velocity at $180^o$ opposite to this azimuth can be expressed as,

$$V_r' = -V_h \cos\theta + w\sin\theta \qquad \text{Eq 5.12}$$

By solving these two equations, one can obtain values of $V_h$ and $w$. Browning and Wexler (1968) later modified the VAD algorithm assuming a linear wind model. As dual or triple doppler radar data is not available here, this study uses single doppler radar data and the VAD algorithm.

Doviak and Zrnic (1993) stated that weather systems covering several hundred kilometres especially generated from low/high pressure system should not cause significant variation in wind speed within the radar covered area. Figure 32 shows that precipitation movement obtained from the correlation tracking algorithm and from VAD at 4 km elevation are consistent.

Environment Canada's online data archive system provides point wind measurements using anemometer and vane at 10 m elevation from the ground at Winnipeg International Airport. It is of interest to compare the point near-surface data with the VAD near-surface data. The lowest elevation is found as 21 m for the $0.5^{o}$ scanning angle. One problem is that airport data is on an hourly scale; hence the VAD is averaged to one hour.

Figure 42 shows a comparison of near-surface hourly averaged wind obtained from the two sensors. The figure shows fairly good agreement between the two sensors except for some outliers.

It is observed that VAD is overestimating speed which may be due to its higher elevation (21 m) compared to the point measurements (10 m). It is noteworthy that hourly averaged point data do not reflect small scale features such as microburst which may last for just 5-30 minutes (Fujita, 1985). Another point is that wind instruments measure the wind speed and direction whereas VAD measures an 'average' speed and direction of wind/ precipitation. The wind speed at a point and the speed of precipitation may not be the same except in the case of snowfall. This may be one reason for the discrepancy between the two sensors.

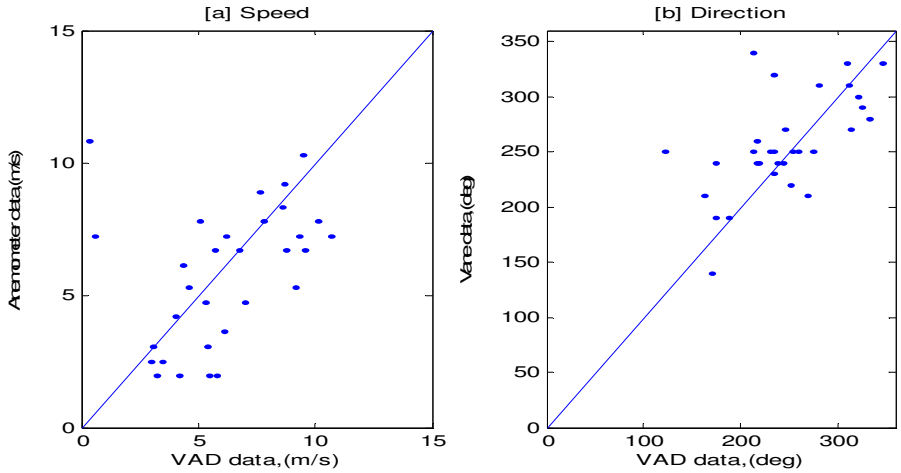


Figure 42: Comparison of hourly [a] wind speed and [b] direction close to the ground using the VAD and anemometer.

### 5.4.4 Model Parameters

Although this study focuses on synoptic scale precipitation systems, the application of VAD data in the trajectory model is still a concern. The reason is that convective precipitation is usually associated with downdraft which may create a radial outflow and divergence at the lower elevations. However, local velocities at a point should be a vector sum of large scale wind (>100 km) motion and local velocity created by downdraft or for some other reasons. This resultant velocity may have greater magnitude at the leading edge of the storm where the two components are co-linear and smaller magnitude (or reverse) at the trailing edge of the storm where the two components are in opposite direction. Hence, a greater portion of the precipitation mass may be advected towards the storm direction and a lesser portion may be advected in other directions. In Figure 41, the higher correlation core is found spread in an area (rather than a point), but skewed towards the VAD wind direction which perhaps implicitly supports this assumption. Admittedly, the trajectory model would be more efficient if velocity data obtained from dual doppler or triple doppler scanning were used.

The vertical velocity is assumed equal to the fall velocity/terminal velocity of the rainfall mass obtained from the reflectivity. Hence, the effect of updraft and downdraft is ignored and *w* is not changing with respect to *x, y, z* and *t* as the precipitation mass of a pixel is traveling towards the earth surface. However, the vertical velocities between two pixels are different, being dependent on their reflectivity. The horizontal velocity components, *u* and *v,* are changing along *z,* but not along *x,* and *y* as the VAD assumes uniform velocity at a horizontal plane. The model also neglects the change of the velocity components with time over the trajectory of rainfall mass, as the total time needed for a rainfall mass to reach the ground should not be more than 6 minutes. With these simplifications, Eq 5.8 becomes

$$\begin{aligned} dx(i) &= u^{t}_{x,y,z-dz(i-1)}dt \\ dy(i) &= v^{t}_{x,y,z-dz(i-1)}dt \\ dz(i) &= w^{t}_{x,y,z}dt \end{aligned} \qquad \text{Eq 5.13}$$

The selection of the time step is a critical issue in the trajectory model. Draxler and Hess (1998) stated that the time step should be such that the advection distance per unit time step is less than 0.75 times the grid spacing. Based on this criterion, a time step of 10 seconds is selected for the trajectory of rainfall mass. The VAD corresponding to the $1.5^{o}$ elevation angle is used as this elevation angle intersects the 1.5 km CAPPI surface at about 50 km radial distance from the radar location. Hence, the velocity provided by the scanning should be representative for Winnipeg which is located approximately 50-70 km from the radar. The assumption behind this segment of VAD scanning is that the wind/precipitation movement is uniform at an elevation up to 1.5 km of the 50 km radius circular area. Due to the geographic location of Winnipeg, one does

not need to assume that wind is uniform over the 200 km radius circular area at a particular elevation.

The reflectivity field at 1.5 km is generated using the tracking algorithm and space-time interpolation with 1-minute resolution. These generated reflectivity fields are passed through the trajectory model. Travel time obtained for each pixel to reach the ground is rounded to the nearest one minute and the displacements in the $x$ and $y$ directions are rounded to the nearest one km. These steps yield reflectivity maps at the earth surface at one-minute resolution. It is then easy to calculate rainfall accumulations with 10-minute or hourly resolution. The space-time interpolation technique thus permits to increase the precision of the trajectory model. As the travel time of the rainfall mass typically varies from 2-6 minutes (less than radar's temporal resolution of 10 minutes), it is indeed necessary to use the interpolated maps when considering space and time shifting criteria.

It is possible that the rainfall mass at two different space-time coordinates may reach the earth surface at the same space-time coordinate. In such cases, the rainfall accumulations corresponding to the two rainfall masses are added together.

### 5.4.5 Results and Discussions

The radar rainfall accumulations at the gauge locations are calculated at the 10-minute scale. Marshall-Palmer's Z-R relationship is used in this context. Figure 43 shows the scatter plot of the gauge and radar accumulations. Figure 44 shows the comparison of the correlations between the two methods, one considers time shifting only (as described in Section 5.3 ) and the other considers both time and space shifting achieved through the

trajectory model. It is observed that the correlation is significantly increased in case of the June 10th and the August 6th rainfall events. For all other cases, the change in correlation is insignificant. The probable reason is that this correction scheme may also be case specific, that is, it should improve only when the wind velocity is strong enough to advect the rainfall mass more than 1 km. Wind direction may also have significant impact on the net advection. For example, if the wind direction is fairly constant within 1.5 km vertical distance, it should cause more advection than a veering or backing wind, even if the magnitude of the wind speed is the same in both cases. Figure 45a shows the mean speed obtained from the VAD profile for the six rainfall events. Figure 45b shows the standard deviation of the direction of wind/precipitation relative to the near-surface direction for the six rainfall events. It is expected that storms with higher variability in direction along the vertical (veering or backing) will have larger standard deviation. It is observed that the June 10th rainfall shows the highest wind/precipitation speed and the August 6th rainfall shows the least standard deviation in direction. These two factors may be held responsible for more advection and thereby improvement in correlation. Admittedly, another important factor is the travel time. If the vertical velocity is significantly high due to downdraft, net advection should be lesser even if the horizontal wind speed is higher.

Another point is that this model should not improve in case of spatially uniform rainfall (e.g. stratiform) even if the wind is strong enough to advect the rainfall mass. For example, the June 13th rainfall appears as stratiform in nature as low intensity rainfall covers the radar umbrella without any convective core and is the result of the movement of a warm front. The July 4th rainfall is also due to the movement of a warm front and

shows widespread stratiform rainfall. However, several unorganized convective cells are found in this rainfall which did not pass over Winnipeg. Due to their widespread spatial uniformity, these two rainfalls may not show significant change due to the wind driven trajectory.

It is interesting to observe that although the VAD represents large scale wind direction (e.g. >100 km) at a particular elevation and does not include small scale movements (e.g. vortex, tornadoes, microburst, divergent flow due to downdraft), it still seems useful in improving radar gauge correlation. This may be due to the fact that precipitation shows a 'net' advection along the large scale wind direction during its 'journey' towards the earth surface.

## 5.5 ATTENUATION AND HAIL

### 5.5.1 Motivation

It is widely believed that the Marshall-Palmer's relationship is useful for stratiform rainfall events but underestimates convective thunderstorms (Sauvageot, 1991, pp 116). Unfortunately, Figure 43 is not supporting this statement. It is observed that the July 7$^{th}$ rainfall is significantly underestimated by radar (bias =1.88) while the July 23$^{rd}$ rainfall is overestimated (bias =0.7) although these two rainfall are both convective thunderstorms. On the other hand, the June 13$^{th}$ and the July 4$^{th}$ rainfall events appear stratiform in nature, but are underestimated by radar (bias 1.69 and 1.28, respectively). The June 10$^{th}$ and the August 6$^{th}$ rainfall events appear convective in nature, but are generally well fitted (bias 0.89 and 1.17, respectively).

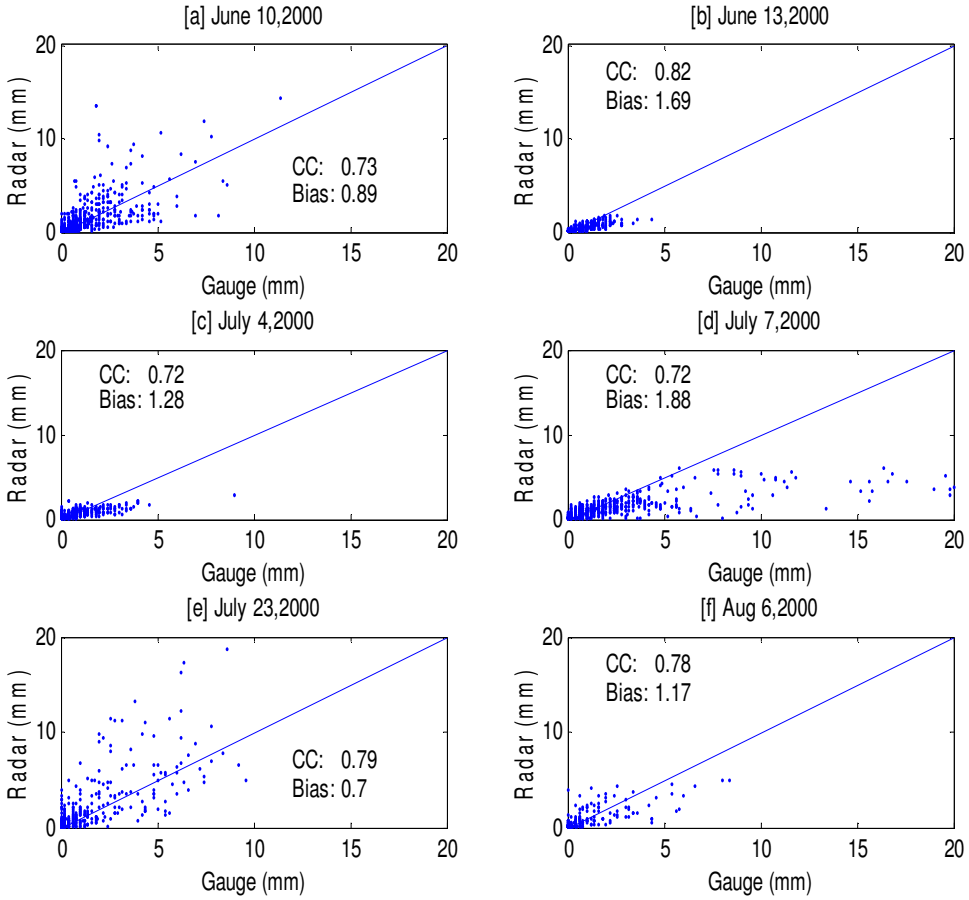


Figure 43: Comparison of 10-minute accumulations of gauge and radar data after space-time interpolation and passing through the trajectory model.

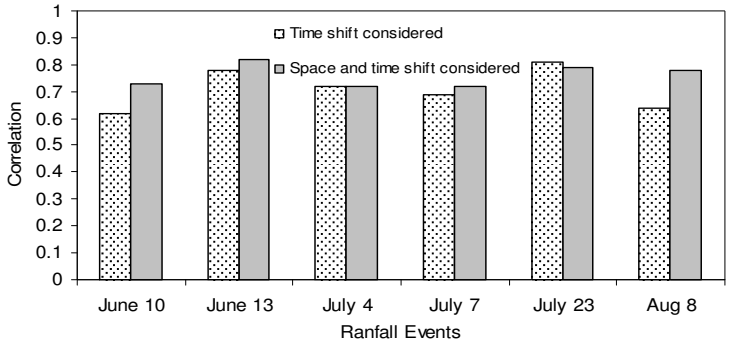


Figure 44: Comparison of correlation obtained with time shifting model and trajectory model. The trajectory model incorporates both time and space shifting.

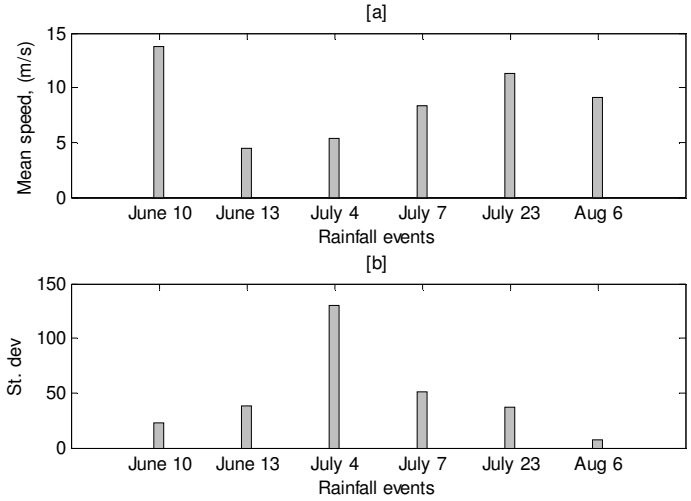


Figure 45: [a] Mean speed of the rainfall events, [b] standard deviation of wind direction (deg) relative to the near-ground direction.

Similar trends are also observed in Figure 14 which shows an 'unprocessed comparison' and in Figure 35 which shows the comparison after space-time interpolation and integration. It is clear that using a different Z-R relationship will not solve this problem. The reason is that if the underestimated rainfall is adjusted by using a suitable Z-R relationship or applying a bias correction, the overestimated rainfall will be further deteriorated. One possible solution that might work well in this case is merging gauge and radar with Co-kriging or other statistical methods. This study attempts to dig out possible meteorological reasons responsible for these chaotic scenarios.

As Environment Canada uses C-band radars, attenuation may be one driving factor for the underestimation. The C-band radars are also being used in Europe and Japan (Liu and Illingworth, 2001), whereas in the USA, S-band radars which are less susceptible to attenuation are being used (Austin, 1987). However, attenuation should not be the reason for overestimation in the case of the July 23$^{rd}$ rainfall event. As the weather observer reported occurrence of hail during the June 10$^{th}$, July 7$^{th}$, and July 23$^{rd}$ rainfall

events (Patrick 2005, personal communication), hail may be another driving factor for the associated discrepancies. As reported by Geotis (1963), reflectivity exceeding 55 dBZ and persistent for several minutes is a definite indication of hail storm. The WSR-88D rainfall algorithm uses 51 dBZ as 'hail cap' for the western mountainous part of the USA (Fulton et al., 1998). Reflectivity above 51 dBZ is replaced with 51 for the aforementioned area considering reflectivity above this threshold to be due to hail. The nationwide default 'hail cap' in the USA is 53 dBZ.

### 5.5.2 Attenuation Correction

This study attempts to apply correction for attenuation. The conventional attenuation correction algorithm is given in Eq 2.10. However different values for the coefficients are reported by different researchers. A list is shown in Table 13 for two-way specific attenuation (attenuation per unit distance).

The equation by Gunn and East (1954) is operationally used in the UK met office (Liu and Illingworth, 2001) and is used in this study. Environment Canada's operational software does not provide any correction for attenuation (Donaldson 2005, personal communication). Figure 47 shows the scatter plot of gauge and radar accumulations after applying the attenuation correction algorithm. It appears that the attenuation correction algorithm makes some improvement in the case of the July 7$^{th}$ rainfall event (correlation increases from 0.72 to 0.83), however it shows no improvement in the June 13$^{th}$ and the

Table 13: Two-way specific attenuation proposed by different researchers.

| | Equation |
|---|---|
| Gunn and East (1954) | $k_s=0.0044R^{1.17}$ |
| Wexler and Atlas (1963) | $k_s=0.006-0.008R$ |
| Liu and Illingworth (2001) | $k_S=0.00124R^{1.31}$ |

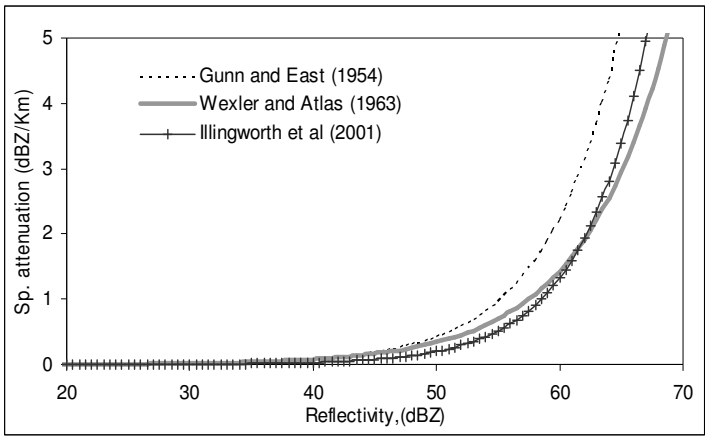


Figure 46: Comparison of three attenuation correction equations

July 4th rainfall events.

As Gunn and East (1954)'s expression provides the highest specific attenuation for a given reflectivity among the three expressions considered (Figure 46), the other expressions should not be able to improve the situation. It is interesting to note that although this expression failed to improve the underestimated rainfall events, it adversely increases the radar estimation of the June 10th and the July 23rd rainfalls. These two rainfall events were already overestimated by radar and the attenuation correction scheme further deteriorated the situation. This study will attempt to explore the possible reasons behind this situation.

It has been reported that the attenuation correction algorithm often causes unusual increase of reflectivity at the far end of the radar scan. To eliminate 'unwanted' enhancement of the reflectivity due to attenuation correction, Lee et al. (2002) proposed to apply a constraint from the gauge data. It appears that it may be better to apply gauge constraints to fix errors due to attenuation rather than errors due to the 'attenuation correction algorithm'. It is interesting to observe that the 'unusual' increase of reflectivity

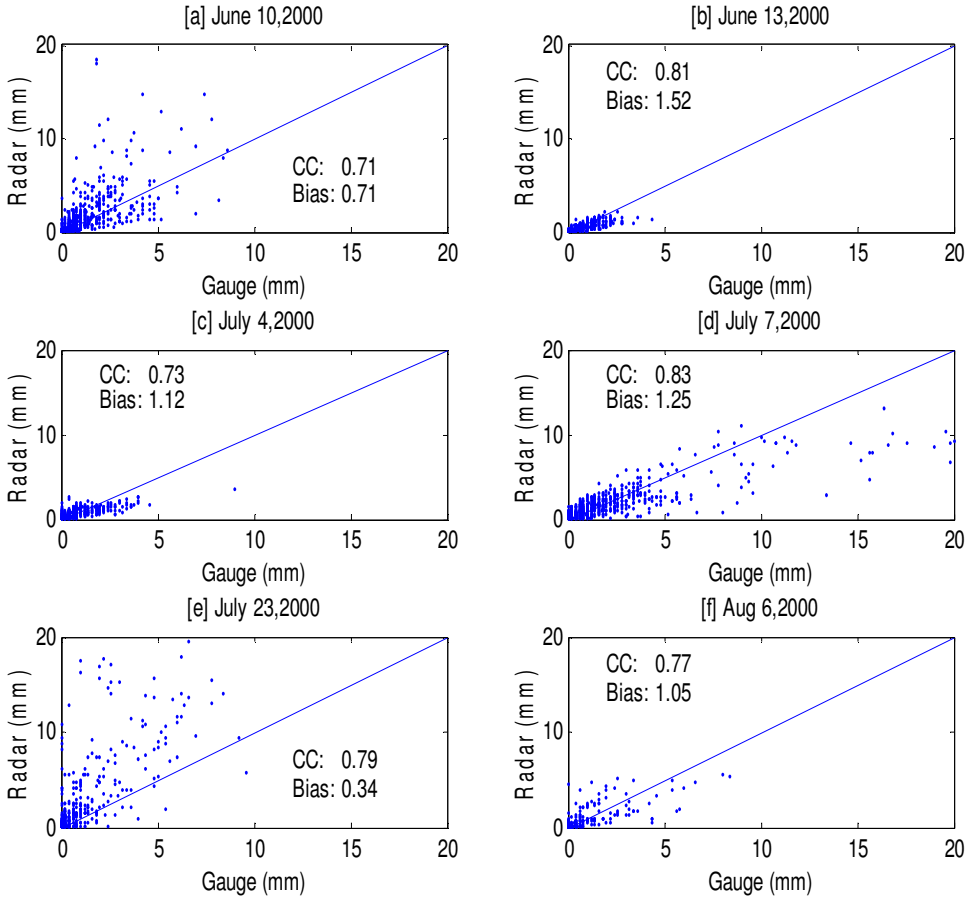


Figure 47: Comparison of gauge and radar accumulations (10-minute) after applying attenuation correction algorithm.

due to attenuation is not observed in case of the August 6$^{th}$, the June 13$^{th}$, and the July 4$^{th}$ rainfall events. Hence, certain weather conditions may be held responsible for such adverse enhancement of reflectivity. This study assumes that hail is one reason for such adverse enhancement of reflectivity as hail was reported during the June 10$^{th}$, the July 7$^{th}$, and the July 23$^{rd}$ rainfall (Patrick 2005, personal communication).

One prime reason for using polarimetric radar is to identify hail covered regions. Hail particles usually tumble and gyrate and therefore produce different differential reflectivity ($Z_{DR}$). The logarithmic ratio of horizontal reflectivity ($Z_{HH}$) and vertical

reflectivity ($Z_{VV}$) is known as the differential reflectivity ($Z_{DR}$). Table 3 shows how the differential reflectivity ($Z_{DR}$) and the horizontal reflectivity ($Z_H$) are simultaneously used to identify hail covered areas.

The polarimetric radar is not available in Winnipeg and improved identification of hail covered area is beyond the scope of this study. However, this study assumes that reflectivity higher than 55 dBZ may result in specific attenuation as high as 1-5 dBZ/km (Figure 46) and an increment of 5 dBZ produces 100% increment in rainfall rate using Marshall-Palmer's Z-R relationship. As the total attenuation of a pixel is the cumulative sum of the specific attenuations, the error generated by the attenuation correction algorithm due to hail will be propagated along the radial till the far edge of the storm and may cause serious overestimation of rainfall. Based on this discussion, this study recommends employing the hail capping algorithm prior to the application of the attenuation correction algorithm.

### 5.5.3 Hail Capping

At this point, one may argue that the actual attenuation caused by the hail storm should compensate for the error in measuring specific attenuation. In light of this opposite argument, it may be better to look at the gauge-radar correlation system when hail capping is applied prior to the attenuation correction algorithm. The ratio of gauge to radar accumulation is computed for all individual scatter points in Figure 43 and plotted against the corresponding radar accumulations in Figure 48. This figure shows that the G/R ratio is less than one for radar accumulations (10-minute) greater than 7 mm which is equivalent to the reflectivity of 49 dBZ. The solid line in this figure shows a 25 point moving average line which falls below one at accumulations greater than 7 mm. Based on

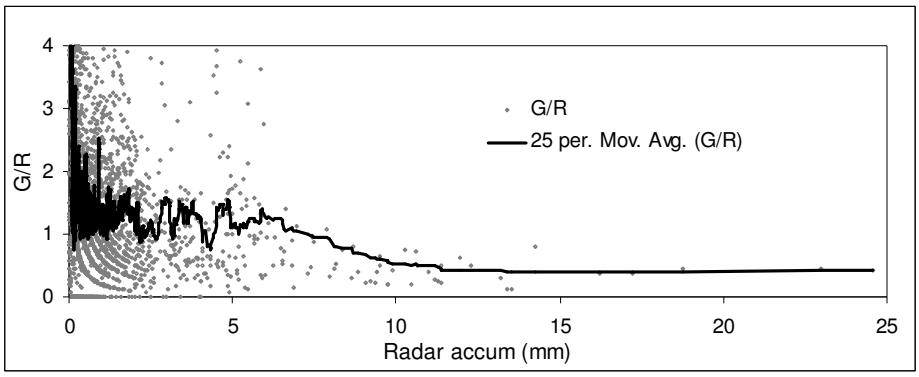


Figure 48: Ratio of gauge to radar accumulation vs corresponding radar accumulation for all scatter pairs of Figure 45.

this observation, this study assumes 50 dBZ as the upper limit of the reflectivity. Figure 49 shows the output from the trajectory model when hail capping is applied prior to the attenuation correction algorithm. This figure shows that the adverse enhancement of reflectivity disappeared and the correlation of the ensemble increases from 0.61 to 0.77. The attenuation correction algorithm alone reduces the correlation from 0.61 to 0.59. Figure 50 shows the comparison of correlations after applying the hail and attenuation correction algorithms. This figure shows that the correlation is improved in the case of the July 7$^{th}$ and the June 10$^{th}$ rainfall events.

### 5.5.4 Radome Attenuation

It is observed that the attenuation correction algorithm has not been able to improve the radar estimates for the June 13$^{th}$ and the July 7$^{th}$ rainfall events (Figure 49b and Figure 49d). The other attenuation equations are not anymore useful as they yield specific attenuations lower than that computed by Gunn and East (1954)'s equation. Another point to observe is that the June 13$^{th}$ rainfall is still underestimated by radar. These issues motivated the consideration of radome attenuation.

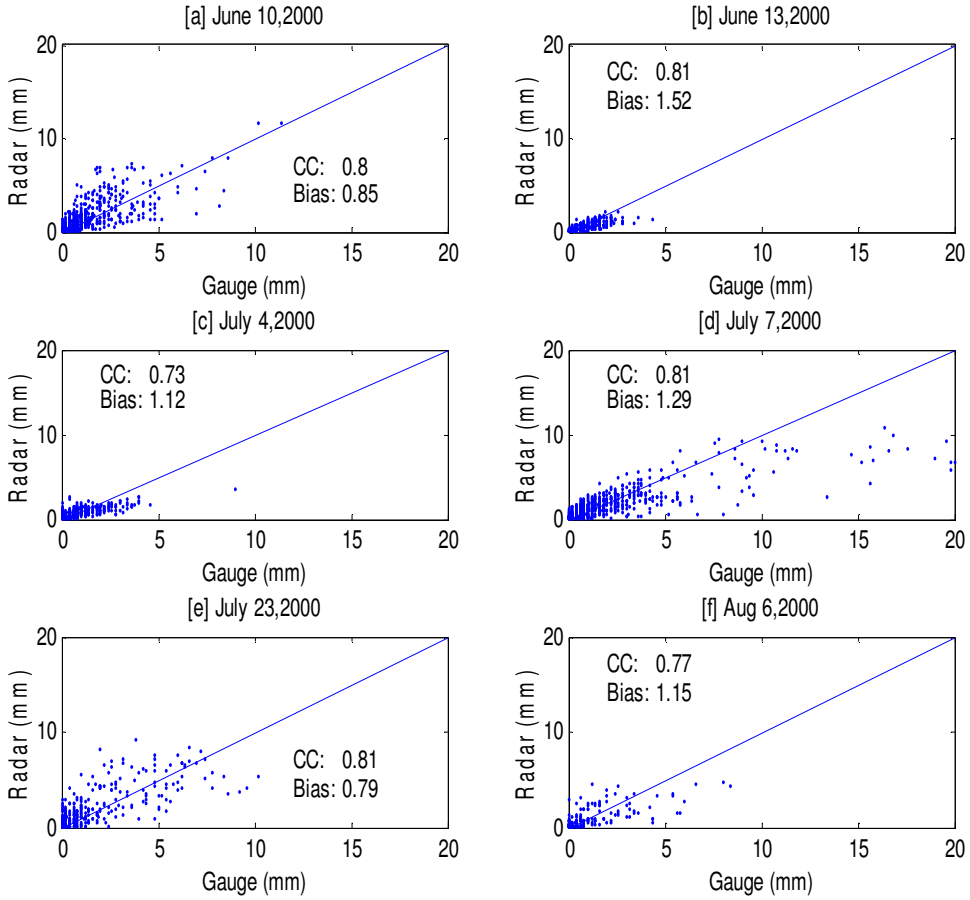


Figure 49: Comparison of gauge and radar accumulation after hail capping is applied prior to the attenuation correction algorithm. The trajectory model is used.

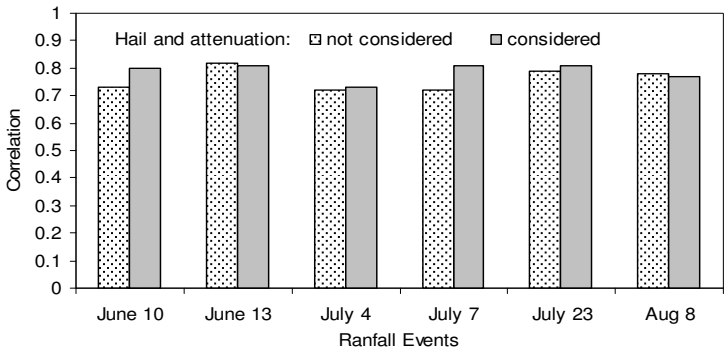


Figure 50: Comparisons of correlation when hail and attenuation correction is considered. The trajectory model is used.

Radome attenuation is caused only when there is rainfall at the radar site, in which case a thin film of water on the radome surface is created (Sauvageot, 1991, pp 32). A wet radome may lead to some attenuation despite the fact that the radome is coated with hydrophobic materials (Patrick 2005, personal communication). Lee et al. (2002) assessed the two-way loss of reflectivity to be around 5 dBZ by using a variational approach.

This study attempts to develop an expression for quantification of wet radome losses to minimize the attenuation in the June 13$^{th}$ and the July 7$^{th}$ rainfall events. The assumption is that the underestimation caused in these two rainfalls is due to the radome attenuation and the parameters of the proposed equation are fitted so that the attenuation in these two rainfalls is minimized. Hence, these two rainfalls are used as a calibration data set. The other rainfalls are used for test purposes. The following expression is found as optimum,

$$k_r = 0.6 dBZ^{0.45} \qquad \text{Eq 5.14}$$

where $k_r$ is the two way radome attenuation for non-zero reflectivity (dBZ). The unit of $k_r$ is dBZ.

Figure 51 shows the comparison of gauge and radar accumulation after applying the proposed radome attenuation scheme. It appears that the June 13$^{th}$ and July 7$^{th}$ rainfall attenuation is alleviated creating no major disturbances in the other rainfall events. The correlation of the ensemble increased from 0.78 to 0.81, indicating a 4% improvement.

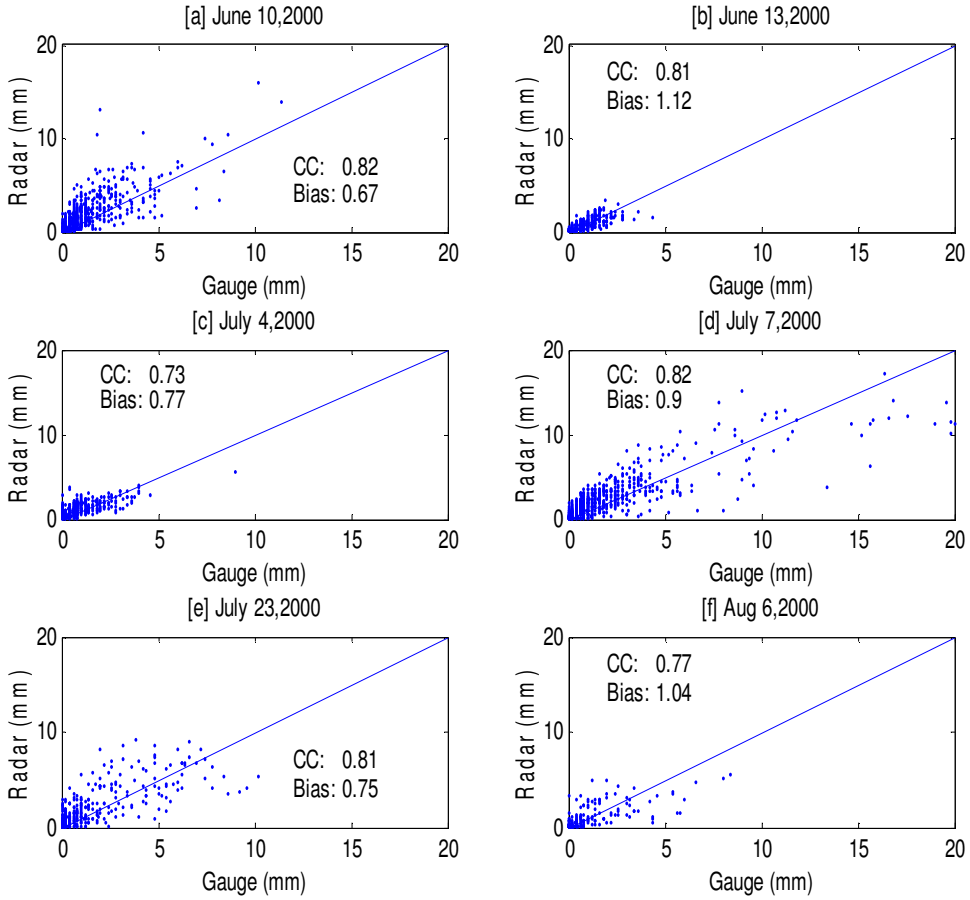


Figure 51 : Comparison of gauge and radar accumulation after applying the proposed radome attenuation correction.

It should be emphasized that radome attenuation occurs only when there is precipitation at the radar site. However, precipitation vertically above the radar is not available at 1.5 km CAPPI as radar scans are limited to a 24.8$^{o}$ tilt angle which produces a 'cone of silence' region (unless a vertically pointing radar is available). Hence, echoes at the second range bin of a radial is used in this scheme.

### 5.6 POST SMOOTHING

Divjak (1999) emphasized the need to carry out a smoothing operation on the radar products as radar products might have various artefacts. These artefacts may be the result

of speckles due to unrecognized clutters (e.g. migratory birds), the synthesis of CAPPI from the discrete scanning, etc. These artefacts may produce discontinuities and sharp gradients. The trajectory model used in the study may cause discontinuities due to incorrect horizontal and vertical velocities.

On the other hand, the smoothing operation may remove fine scale details from the rainfall map. Hence, another assessment of the performance of the smoothing operation is necessary.

This study explores the possibility of a wavelet thresholding approach as a smoothing tool. Wavelet thresholding schemes have been used for eliminating Gaussian white noise from images since the early 1990s (Lin et al., 1998; Starck et al., 2002). The advantage of this technique over linear filtering techniques lies in the fact that it does not distort the edges.

Figure 52 shows a demonstration of the performance of a wavelet-based de-noised scheme. A normally distributed noise (mean 0 and standard deviation 40) is added to an air photo of a city pixel-by-pixel and thereafter it is denoised with a wavelet scheme. If one considers the CAPPI as an image containing random noise, this tool may be helpful in eliminating random noise. This study uses the 'Wavelab' toolbox developed at Stanford University. The description of the algorithm is provided in Appendix C-1.

This study applied several threshold parameters, however improvement in correlation is not observed in any case. Figure 53 shows the scatter plot of gauge and radar accumulations (10-minute) after applying 5 dBZ as threshold parameter.

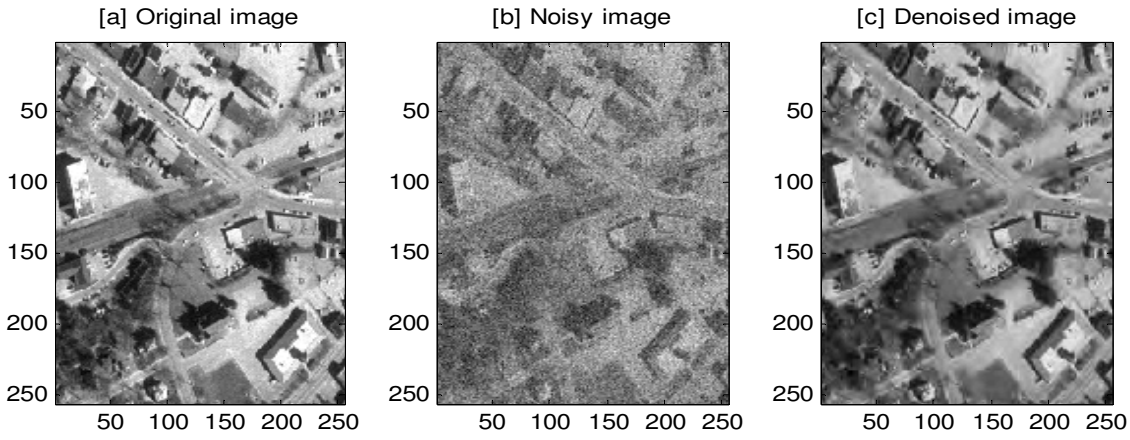


Figure 52: Performance of the wavelet based denoising.

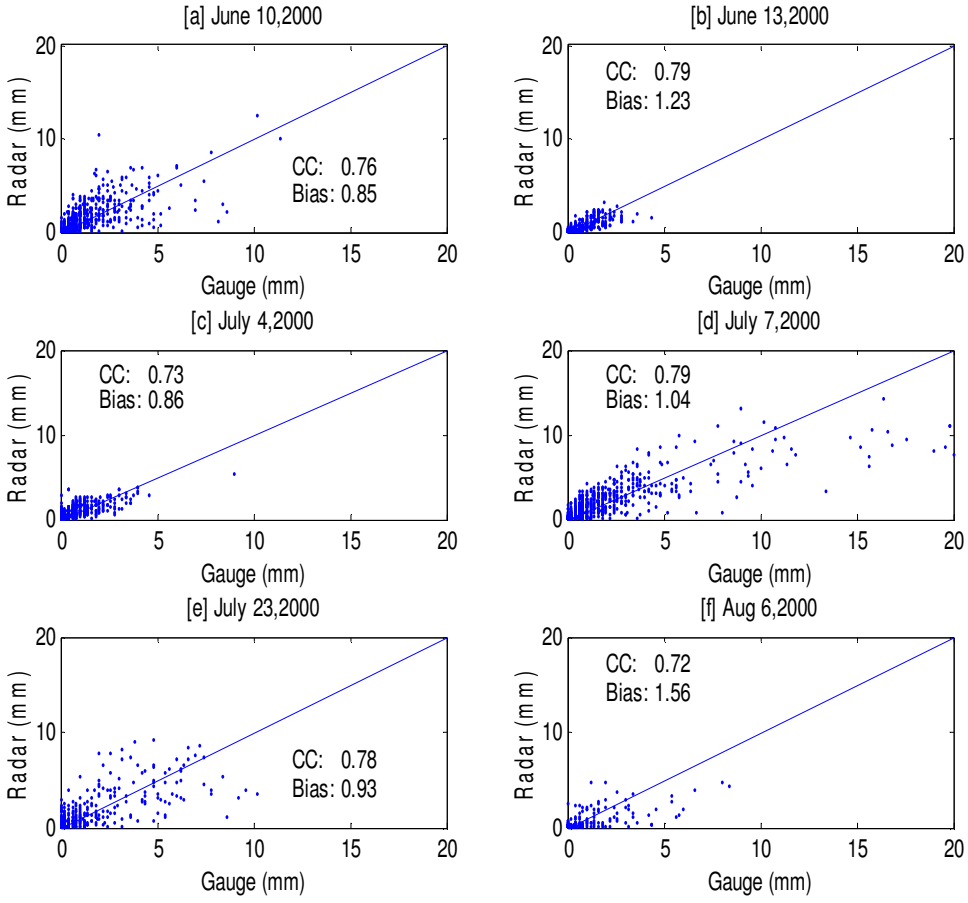


Figure 53: Comparison of gauge and radar accumulations (10-minute) after applying smoothing with wavelet scheme.

The correlation of the ensemble is reduced to 0.78 from 0.81 and bias is reduced from 0.82 to 0.98 for the 5-dBZ threshold parameter. One possible reason may be the smoothing operation reduces peaks which results in reducing bias. Although the bias is improved, the reduction in correlation probably implies that 'real' peaks were eliminated. This is also observed in the case of the July 7$^{th}$ rainfall and the June 10$^{th}$ rainfall events (Figure 53a and Figure 53d). Hence, despite significant improvement in bias, this study is not recommending post smoothing.

## 5.7 OVERALL RESULTS

This study attempts to derive surface rainfall maps from radar observations of the atmosphere. The performance is assessed based on 10-minute accumulations. The very high time resolution allows the model to account for space and time shifting of the rainfall mass as it travels to the earth surface. The model also accounts for the error associated with attenuation and hail.

One drawback of hail capping is that it forces the model not to generate rainfall greater than the 'capped' value, even if the actual rainfall is greater than that. For example, a model is not able to generate rainfall intensities greater than 100 mm/hr if reflectivity is capped at 55 dBZ. This obviously is a problem in flood forecasting models used to estimate peak flows. Although the 'cap' reflectivity is selected based on the long term observation of past weather, there is no guarantee that rainfall will never exceed that limit. This study suggests applying capping prior to the attenuation correction algorithm. In this way, the attenuation correction algorithm may enhance the reflectivity minimizing the adverse effect that may be caused by the capping method. For example, among the six

rainfall events used in the study, the most extreme rainfall is observed during the July 7th rainfall event where 10-minute gauge accumulation reached 25.6 mm (equivalent to 153.6 mm/hr intensity), but the radar reflectivity above the gauge location is only 49 dBZ. The reflectivity is not found higher than 50 dBZ in its upwind direction either. According to the Marshall-Palmer relationship this amount of rainfall should cause reflectivity around 58 dBZ. This suggests that the extreme rainfall may be significantly attenuated based on its spatial distribution and the hail capping method should not cap the real extreme rainfalls. It should cap the 'pseudo' extreme rainfalls caused by hail or by anomalous propagation. An attenuation correction algorithm applied after the hail capping may enhance the reflectivity of the 'real' extreme rainfalls.

The bias of the ensemble is found to be 0.82 which implies that the radar is overestimating to some extent. This is not unexpected as gauge under-catch is not considered. Under-catch may be as high as 40% due to strong winds during thunderstorms (Wilson and Brandes, 1979). No correction is available for the tipping-bucket gauge currently installed in Winnipeg (Hanesiak 2005, personal communication). Underestimation at gauges may also occur in heavy showers due to the loss of precipitation during the tipping process (http://www.webmet.com/). During light rain or drizzle, evaporation is another reason for gauge underestimation. These issues do not affect the accuracy of the radar measurement, but affect the performance evaluation when radar data are compared to gauge data.

Figure 54 shows how correlations improved for each rainfall event as a function of the various correction steps. The correlation of the ensemble of all rainfall events increases from 0.3 to 0.81 after applying all the corrections - a 170% improvement.

Figure 70 shows gauge and radar accumulations (10-minute) for the ensemble of six rainfall events and Figure 71 shows a simplified flow diagram of the proposed model.

## 5.8 LIMITATIONS OF THE MODEL

This model does not consider anomalous propagation (AP) although wake of the thunderstorm is susceptible to be AP echoes. However, it is believed that the precipitation scans generated from the lowest elevation angles are susceptible to anomalous propagation. The WSR-88D precipitation processing algorithm compares the reflectivity map obtained from the lowest elevation angle to the second or higher tilt angle and uses the higher tilt angle if AP is detected (Fulton et al., 1998). This algorithm detects AP if the lowest tilt angle scanning shows precipitation area much higher than suggested by the higher tilt angle scanning.

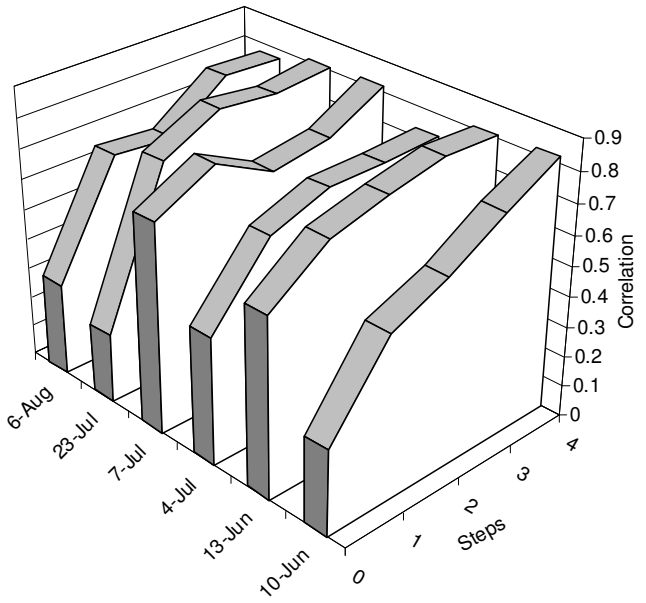


Figure 54: Correlation of each rainfall at each correction step, 0: 'raw' comparison, 1: time integration, 2: time shifting, 3: space-time shifting, 4: attenuation and hail.

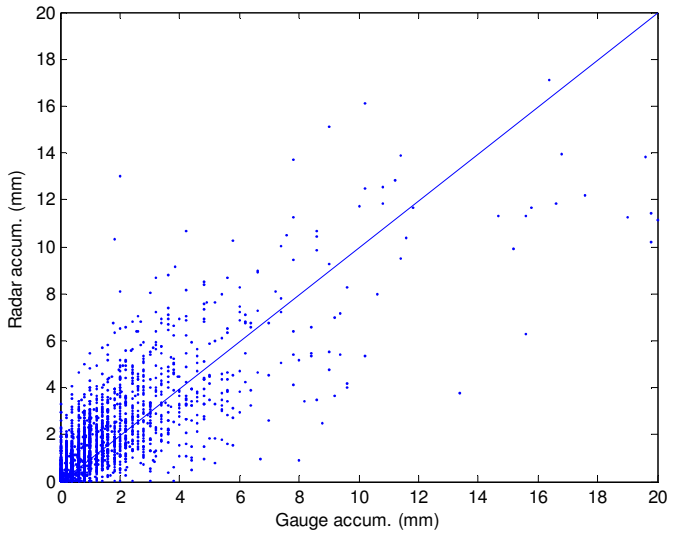


Figure 55: Comparison of 10-minute accumulations of gauge and radar measurement for the ensemble of six rainfall events.

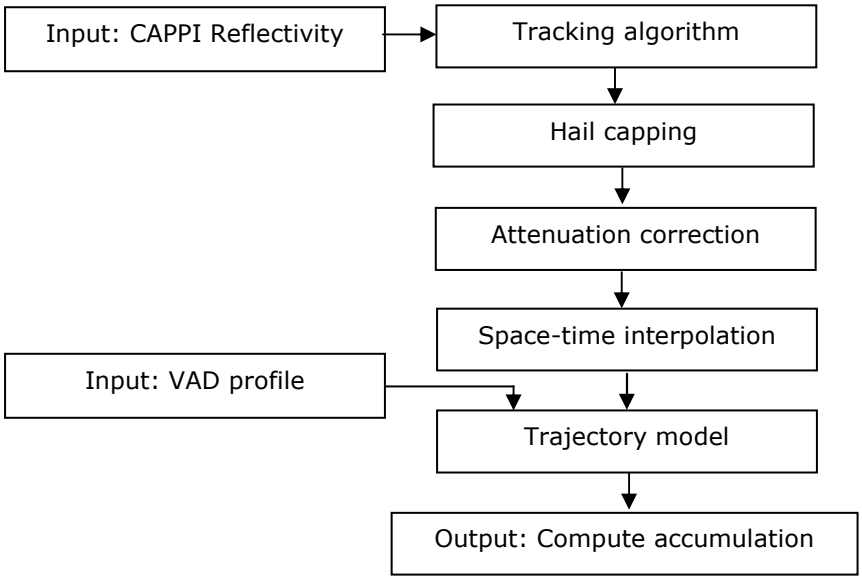


Figure 56: A simplified flow chart for the proposed model.

This study uses the 1.5 km CAPPI instead of the lowest tilt angle scanning. The 1.5 km CAPPI at Winnipeg are obtained from scanning corresponding to the $1.1^{o}$ and $1.4^{o}$ tilt angles. The scanning at these elevation angles should not be affected by AP.

The current study is focused on a range smaller than 100 km. Beyond this range, more errors are generated due to bright band, beam spreading, beam overshooting, etc. which are not considered here. The current study only focused on summer rainfall and the wind drift model developed herein is not applicable to snowfall. This model does not account for atmospheric evaporation of precipitation which may be significant for light rainfall.

## 5.9 CONCLUDING REMARKS

It may be a better choice to use the lowest tilt angle instead of the 1.5 km CAPPI to minimize the error associated with wind shear. The disadvantage of the lowest tilt angle is that it may be affected by the Anomalous Propagation and ground clutter. The WSR-88D algorithm uses the lowest tilt angle and when AP is detected it uses a higher tilt angle and thus a 'hybrid' scan results (Fulton et al., 1998). At higher range (greater than 70 km), the elevation of the lowest tilt angle is approximately 1 km. At such elevations significant errors associated with wind shear may arise. Hence, the lowest tilt angle scanning is also susceptible to errors associated with drift. Given the relative advantage and disadvantage of the CAPPI and the lowest scan, this study arbitrarily selected the CAPPI to work with. At far range (greater than 110 km), the lowest elevation angle is used as the CAPPI and more errors are generated due to bright band, beam spreading, etc. The present study has not focused on these issues.

Although Marshall-Palmer's Z-R relationship is not recommended by NOAA for convective rain showers and thunderstorm, this study found that this expression is suitable for both stratiform and convective rainfall events when attenuation is accounted

for. The attenuation may be the reason for underestimating rainfall by this expression during thunderstorms. Although S-band radars are considered insensitive to attenuation, Ryzhkov and Zrnic (1995) showed that an intense squall line caused significant attenuation in an S-band radar. If Marshall-Palmer's equation is used in the case of S-band radars, it may show underestimation, but this may due to attenuation of the S-band radar rather than the equation itself.

# Chapter 6

# Concluding Remarks

## 6.1 CONCLUSIONS

Efforts have been made over the last forty years to improve the quality of radar data for meteorological applications. The objective of this study is the same. The quality of the radar data is assessed based on gauge data. The basic assumption of this study is that the higher the correlation between the gauge and the radar data, the higher is the quality of the radar data. This study differentiates between the reflectivity map in the atmosphere and at the surface and an attempt is made to get the surface rainfall map from the atmospheric reflectivity map. It is expected that the generated surface rainfall map will be more applicable for hydrologic applications than the atmospheric rainfall map. The overall outcome of this study is summarized in the following sections.

It was initially observed that the performance of the traditional Z-R relationship is below the satisfactory level. The study worked with the six major rainfall events occurring in the year 2000. The correlation coefficient between the gauge and the radar for the ensemble of six rainfall events was found to be 0.3. Besides the Marshall-Palmer's Z-R relationship, alternative Z-R relationships recommended by NOAA for convective ($Z=300R^{1.4}$) and tropical rainfall ($Z=250R^{1.2}$) were explored. These two relationships

gave correlations lower than the Marshall-Palmer's relationship (0.22 and 0.18, respectively). Rainfall was classified into convective and stratiform regions and two different Z-R relationships were applied for the two regions ($Z=200R^{1.6}$ for the stratiform regions and $Z=250R^{1.2}$ for the convective regions), however the correlation was not improved (= 0.18).

It was then assumed that accurate time synchronization plays a vital role in radar-gauge correlation at high time resolution (10-minute accumulations). Before preparing the radar and gauge ensemble used to assess the quality of radar data, one needs to know the exact time when the study area is scanned by the radar. The time interval for gauge accumulations should be prepared in such a way that the time when the radar scans the area falls within the first half of the gauge time interval for a 10-minute radar cycle. In this study, it was found that this time synchronization increased the correlation to 0.5 for the ensemble. Although this is neither error of radar nor error of gauge, it seems important to account for when assessing the quality of radar data or validating any correction scheme with the support of gauge data.

Hourly accumulations show a somewhat higher correlation (0.5). The possible reason for the improvement may be that it smoothens out the error associated with time synchronization. Despite the improvement, hourly accumulations also contain outliers which may be due to hail contamination. Disregarding the outliers, the overall data structure shows an underestimation by radar. The underestimation may be removed by using a different Z-R relationship (e.g. $Z=150R^{1.5}$). Another conclusion may be that the underestimation is a result of attenuation which is not negligible for the C-band radar used in the study.

Development of 'climatological Z-R relationship' was attempted by Miller (1972), Calheiros and Zawadski (1987), and Atlas et al. (1989) by matching the cumulative probability distribution functions. Using this method, this study found the relationship $Z=30R^2$ using the rainfall events from the study. The correlation for hourly accumulations with this relationship was found to be 0.6, which is higher than the Marshall-Palmer's relationship (=0.5). However, correlation measures the linear association and do not account for any systematic overestimation/underestimation. It was observed that the climatological *Z-R* relationship produces a bias of 0.71 and the Marshall-Palmer's relationship produces a bias of 1.05. Hence, in terms of bias, Marshall-Palmer's relationship may be climatologically more accurate.

The performance of various black-box models was explored by considering that the complexity of the reflectivity-rainfall relationship arises due to variability in the weather system. It is expected that a black-box model trained with a group of rainfall events subjected to different weather conditions may work better than the Z-R relationships. The test data was divided into two groups, one containing the reflectivity-rainfall set for the same rainfall events used in the training but at different geographic locations, and the other set containing a new rainfall event. Artificial neural networks were used as a 'black-box' models and three different types of networks were explored: the back propagation network, the radial basis function network, and the generalized regression neural network. Each network was tested with two different types of input data set, one containing a short time series of reflectivity of a pixel and the other containing reflectivity of several pixels pertaining to the same time. It was expected that these two input patterns would help assess the relative impact of temporal /spatial smoothing of the

input data. It was observed that neural networks have some potential to estimate precipitation at new geographic locations in the same rainfall event used for training. The method did not significantly improve the estimation of completely new rainfall events. When applied to the same rainfall event, the correlation was found to be 0.67 which is much higher than the correlation obtained using the Z-R relationship (=0.28). For new rainfall events, its performance is similar to the Z-R relationship (0.32 and 0.28, respectively). One possible reason may be the spatial correlation of rainfall. Frontal precipitation systems are usually synoptic scale phenomena and may extend over the whole radar coverage area. When such precipitation systems pass the area, two geographic points may show similar time series of rainfall, apart from a certain time lag. Zawadski (1973) developed a method to track rainfall using the gauge data based on the space-time autocorrelation of rainfall patterns. This may be what enables the neural network to estimate the rainfall more consistently with gauge data for the same rainfall event used in the training. Hence, the neural network may be recommended as a post-processing tool rather than an operational estimator. The short time series of rainfall as input was found better than the spatial input in all the models tested. Perhaps the trend in the short time series helps to better estimate the rainfall. Among the networks tested, the back-propagation network was found to perform best.

An integrated model has been developed to generate precipitation maps at the surface from the atmospheric precipitation at 10-minute resolution. This model applies correction for attenuation and hail. Using a correlation based tracking algorithm, the speed and direction of movement of precipitation is identified and using this direction nine maps are generated at each 10-minute cycle using the space-time interpolation

technique. The generated scan is then passed through a trajectory model which transfers the rainfall mass to the earth surface using the velocity field obtained from the VAD scanning. The vertical velocity component is assumed equal to the fall velocity and is computed by Rozer's (1964) equation. It was found that the model produced correlations of 0.61.

It was observed that the June 10$^{th}$ and the August 6$^{th}$ rainfall events benefited the most from the model. These two rainfalls were subjected to the least variation in wind direction from the earth surface to the 1.5 km elevation and had significantly higher wind speed than the other events. These two factors may be held responsible for advecting rainfall masses more than the radar's spatial resolution. Although the VAD profile assumes uniform wind at a given elevation, its performance in the trajectory model is encouraging. One possible reason for obtaining a better performance may be that precipitation shows a 'net' advection along the large scale wind direction given by the VAD scanning.

The quality of the VAD profile is checked by comparing it with the velocity measured by an anemometer at the near ground level and by comparing it with the velocity obtained from the tracking algorithm. The direction obtained from the tracking algorithm should reflect the wind direction at the precipitation generation level and the VAD data at 4 km or above shows good agreement with the direction obtained from the tracking algorithm. The maximum variation in direction is about 30$^{o}$. The maximum variation in direction obtained from the anemometer and the VAD is about 50$^{o}$. One possible reason behind this higher variation may be that the anemometer acts as a point estimator of velocity.

This study found that the attenuation correction algorithm makes an enhancement of reflectivity which is not always beneficial. This enhancement improved the highly attenuated July 7$^{th}$ rainfall but deteriorated the radar performance for the June 10$^{th}$ and the July 23$^{rd}$ rainfall events. However, a hail capping method applied prior to the attenuation correction may resolve the situation. Hail produce very high reflectivity (greater than 55 dBZ), and these reflectivities result in high specific attenuation due to the power-law nature of the existing attenuation correction algorithms. The cumulative addition of the specific attenuations along the radial produces overestimated reflectivities at the far edge of the precipitation region and contaminates the whole reflectivity map. Hence, a hail capping method applied prior to the attenuation correction is recommneded.

This study proposes an expression to account for the radome attenuation. The parameters of this expression are chosen by trial and error so that the two underestimated precipitations (June 13$^{th}$ and the July 7$^{th}$) will fit with the gauge measurements. It is observed that the remaining four rainfall events do not deteriorate with this correction algorithm.

It is observed that the computed rainfall accumulations using the generated scans at one-minute resolution perform better (correlation 0.57) than the computed accumulations using the 10-minute resolution scans. In the latter case, one needs to assume that the rainfall intensity is constant over the 10-minute radar cycle which is not necessarily true for fast moving convective cells. Hence, an improved time integration may be achieved by generating precipitation maps at one-minute resolution. It should be emphasized that the generated scans are useful for executing the trajectory model in a more efficient way.

## 6.2 RECOMMENDATIONS

It appears from the study that the quality of both gauge and radar data is influenced by weather conditions. Heavy rainfall extending over large areas may produce severe attenuation as observed in the July 7$^{th}$ rainfall. On the other hand, hail may have caused excessive overestimation by radar in the case of the July 23$^{rd}$ rainfall. A stratiform rainfall may also show underestimation by radar due to the localized heavy rainfall at the radar site. Strong wind shear may advect the rainfall mass horizontally. These issues will not be solved by modifying the *Z-R* relationship. Hence, integration of various correction schemes in a unified model is deemed necessary. The correction schemes may need a wide range of atmospheric data which may not be readily available to an individual. In such cases, a merging with available gauge data is strongly recommended. The reason is that merging with gauge data can implicitly account for attenuation, hail, atmospheric evaporation, bright band issues, etc., by applying constraints from the gauges. In such cases, gauge under-catch due to strong wind and other sources of errors should be accounted for.

The determination of bias or similar correction schemes with the help of gauge data ignoring other possible sources of errors may be misleading. For example, a rainfall may be underestimated due to attenuation or overestimated due to atmospheric evaporation. Hence, bias computed based on one rainfall event may not be applicable to other rainfall events where attenuation is not significant. The attenuation depends on the intensity and the spatial extent of the rainfall and may vary from rainfall to rainfall. If a model is adaptive in nature or is tuned to a particular rainfall, its use for other rainfall events may deteriorate the rainfall estimation rather than improving it. Determining the

bias based on an ensemble of precipitation events may be a solution. This study found the bias for the ensemble of six rainfall events close to unity which implies that a general bias correction will not improve any of the rainfall events. A bias computed for each of a rainfall may be recommended and should improve the precipitation estimate. This process is similar to the merging process.

It is observed that strong winds may advect the precipitation mass up to 3 km. This advection implies that precipitation at the earth surface may differ from atmospheric precipitation above a point. Hence, to feed a hydrologic model, precipitation maps at the earth surface must be derived from the precipitation in the atmosphere. The total accumulation over larger area may not be seriously affected by this advection, but hydrograph generated from a smaller urban watershed may be affected by this advection. This issue is also important when one attempts to measure bias or similar correction schemes or attempt to check the quality of the radar data using the gauge data.

Although the radar provides scans at 10-minute resolution, it is possible to generate scans at higher resolution using a tracking algorithm and space-time interpolation techniques. The generated scans at higher time resolution may be useful for more precise computation of rainfall accumulations which may increase the accuracy of rainfall-runoff models.

In the application of the correlation based tracking algorithm, it is recommended that the size of the window be large enough (e.g. 75km x 75km) to identify the large scale movement. Small scale movements may be associated with the ‘propagation’ effect. The VAD obtained from the single doppler scan may be recommended for obtaining the

direction for the trajectory model in case of frontal precipitation systems. However, the VAD is not recommended for small-scale isolated thunderstorms.

The rainfall selected in this study contains both convective and stratiform components which were identified and separated using Steiner et al. (1995)'s algorithm. It was observed that a single Z-R relationship works fairly well for both convective and stratiform rain in case of hourly accumulation. In the case of 10-minute accumulations, after applying the various correction schemes, classification of rainfall into convective and stratiform regions and the use of two different Z-R relationships did not seem warranted.

Any attenuation correction scheme without prior application of hail-capping may be misleading. The expressions used for computing specific attenuation are exponential in nature with respect to reflectivity and hence generate very high specific attenuation for regions contaminated by hail. When the cumulative addition of specific attenuation is carried out for the whole precipitation covered area, unusually high reflectivity may appear at the far end of the precipitation area from the radar site. Hence, hail capping before attenuation correction is strongly recommended.

Although Marshall-Palmer's relationship is usually recommended for stratiform precipitation, this study found it applicable to both convective and stratiform precipitation when attenuation is properly accounted for. Hence, this relationship is recommended for future studies for both convective and stratiform precipitation events.

## 6.3 FUTURE RESEARCH SCOPE

It has been observed that the proposed correlation-based tracking algorithm shows anomalous directions in a few cases (1%). Hence, the performance of Kalman filtering may be explored to smooth the direction and to remove anomalous directions.

The performance of the trajectory model may be explored with the velocity field obtained from the dual/triple doppler scan. This would provide information on localized updraft and downdraft and may be useful in other meteorological applications.

Effort should be directed to the effect of hail contamination. Reflectivity derived from hail contaminated regions does not follow the conventional Z-R relationship. It might be possible to develop some statistical relationship to estimate the rainfall accumulation using gauge data if a hail contaminated area is identified. Lemon (1978) reported some criteria to identify hail regions using the volume scan which may be utilized if polarimetric data are not available. If reflectivity in the hail covered region and the gauge rainfall in that regions show some trend, it might be possible to estimate rainfall accumulations fairly accurately in those regions. Research should also focus on assessing the effect of hail on tipping-bucket gauge.

Finally, a precipitation processing system similar to NEXRAD PPS (Fulton et al. 1998) should be developed suitable for the Canadian environment including other correction schemes which were not considered in this study. These include bright band contamination and anomalous propagation. The conventional correction scheme of bright band and anomalous propagation requires the volume scans only which are readily available.

# References


Aboufadel, E., and Schlicker, S., 1999, *Discovering Wavelets*, John Wiley and Sons Inc, NY.

Ahrens, C. D., 1994, *Meteorology Today*, 5th Edition, West Publishing Company, MN.

Alberoni, P., Andersson, T., Mezzasalma, P., Michelson, D., and Nanni, S., 2001, Use of the vertical reflectivity profile for identification of anomalous propagation, *Meteorological Applications*, vol 8, no 3, pp 257-266.

Al-Sowayan, S., and Chandrasekar, V., 2001, Convective/Stratiform classification from TRMM precipitation radar, *IEEE International Geoscience and Remote Sensing Proceedings*, vol 6, pp 2507-2509, Sydney, Australia.

Anagnostou, E. N., 2004, A convective/stratiform precipitation classification algorithm for volume scanning weather radar observations, *Meteorological Applications*, vol 11, no 4, pp 291-300.

Anagnostou, E. N., Krajewski, W. F., Seo, D. J., and Johnson, E. R., 1998, Mean-field rainfall bias studies for WSR-88D, *Journal of Hydrological Engineering*, pp 149-159.

Anagnostou, E. N., and Kummarow, C., 1996, Stratiform and convective classification of rainfall using SSM/I 85 GHz brightness temperature observations, *Journal of Atmospheric and Oceanic Technology*, vol 14, pp 570-575.

Atlas, D., Short, D. A., Rosenfeld, D., 1989, Climatologically tuned reflectivity-rain rate relations, *Proceedings of the 24th Conference of radar Meteorology*, American Meteorological Society, pp 666-671.

Atlas, D., and Ulbrich, C. W., 1977, Path and area integrated rainfall measurement by microwave attenuation in the 1-3 cm band, *Journal of Applied Meteorology*, vol 16, pp 1322-1331.

Austin, P. M., 1987, Relationship between measured radar reflectivity and surface rainfall, *Monthly Weather Review*, vol 115, pp 1053-1071,

Baeck, M. L., and Smith, J. A., 1998, Rainfall estimation by the WSR-88D for heavy rainfall events, *Monthly Weather Review,* vol 115, pp 1053-1070.

Battan, L. J., 1976, Vertical air-motions and the Z-R relation, *Journal of Meteorology*, vol 15, pp 1120-1121.

Battan, L. J., 1973, *Radar Observation of the Atmosphere*, The University of Chicago Press, Chicago, IL.

Bellerby, T., Todd, M., Kniveton, D., and Kidd, C., 2000, Rainfall estimation from a combination of TRMM precipitation radar and GOES multispectral satellite imagery through the use of an Artificial Neural Network, *Journal of Applied Meteorology*, vol 39, no 12, pp 2115-2128.

Bech, J., Bebbington, D. H., Codina, B., Sairouni, A., and Lorente, J., 1998, Evaluation of atmospheric anomalous propagation conditions: an application for weather radars. *Proceedings of SPIE Volume: 3499, EUROPTO Conference on Remote Sensing for Agriculture, Ecosystems, and Hydrology*, vol 3449, pp 111-115, Barcelona, Spain.

Bellon A., and Austin, G. L., 1978, The evaluation of two years of real time operation of a short-term precipitation forecasting procedure (SHARP), *Journal of Applied Meteorology*, vol 17, pp 1778-1787.

Bent, A. E., 1946, Radar detection of precipitation, *Journal of Meteorology*, vol 3, pp 78-84.

Bhatnagar, A. K., Rao, P. R., Kalyanasundaram, S. B., Thampi, S. B., Suresh, R., and Gupta, J. P., 2003, Doppler radar - a detecting tool and measuring instrument in meteorology, General article, *Current Science*, vol 85, no 3, India.

Blanchard, D. C., Spencer, A. T., 1975, Experiments of the generation of raindrop size distribution by drop break-up, *Journal of Atmospheric Science*, vol 27, pp 101-108.

Bodiroga, Z., 2004, *Comparison of Gauge and Radar Rainfall Accumulation for Urban Hydrology*, Undergraduate thesis, University of Manitoba.

Borga, M., 2002, Accuracy of radar rainfall estimates for stream flow simulation, *Journal of Hydrology,* vol 267, pp 26-39.

Borga, M., and Tonelli, F., 2002, Long-term assessment of bias adjustment in radar rainfall estimation, *Water Resources Research*, vol 38, no 11, pp 1226-1235.

Browning, K. A., 1979, The FRONTIERS plan, a strategy for using radar and satellite imagery for very short range precipitation forecasting, *The Meteorological Magazine,* vol 10, no 8, pp 161-184.

Browning, K. A., and Wexler, R., 1968, A determination of kinematic properties of a wind field using doppler radar, *Journal of Applied Meteorology*, vol 7, pp 105-113.

Calheiros, R. V., and Zawadski, 1987, Reflectivity-rain rate relationships for radar hydrology in Brazil*, Journal of Climate Applied Meteorology*, vol 26, pp 118-132.

Castro, A. V., 2004, *A Neural Network Approach to Predict Hurricane Intensity in the North Atlantic Basin*, M.Sc. thesis, University of Puerto Rico.

Cheng, M., and Collier, C. G., 1993, An objective method for recognizing and partially correcting bright-band error in radar images, *Journal of Applied Meteorology*, vol 32, American Meteorological Society, pp 1142-1149.

Ciach, C. J., and Krajewski, W. F., 1999, On the estimation of radar rainfall error variance, *Advances in Water Resources*, vol 22, no 6, pp 585-595.

Chandrasekar, H. L., and Gorgucci, E., 2001, Detection of rain/no rain condition on the ground radar based on radar observations, *IEEE Transactions of Geosciences and Remote Sensing*, vol 39, no 3.

Chumchean, S., Sharma, A., and Seed, A., 2004, Correcting of real-time radar rainfall bias using a Kalman filtering approach, *Sixth International Symposium on Hydrological Applications of Weather Radar*, Melbourne, Australia.

Chumchean, S., Sharma, A., and Seed, A., 2003, Radar rainfall error variance and its impact on radar rainfall calibration, *Physics and Chemistry of the Earth*, vol 28, pp 27-39.

Collier, C. G., 1999, The impact of wind drift on the utility of very high spatial resolution radar data over urban areas, *Physics and Chemistry of the Earth*, vol 24, no 8, pp 889-893.

Collier, C. G., 1989, *Application of Weather Radar Systems*, John Wiley and Sons.

Dalezios, N. R., and Kouwen, N., 1990, Radar signal interpretation in warm season rainstorms, *Nordic Hydrology*, vol 21, pp 47-64.

Demuth, H., and Beale, M., 1994, *MATLAB Neural Network Toolbox User's Guide*, version 2, The Math works Inc, MA, USA.

Divjak, M., 1999, *Radar data quality ensuring procedures at European weather radar stations*, Emutnet Opera Programme, Subproject 1c/1, OPERA/WD/9/1999, Final document, http://www.chmi.cz/OPERA/docum/opera_1_fr/WG1_1c1_RadarDataQualityProcedures.pdf (web site accessed on Nov 9, 2005).

Doviak, R. J., and Zrnic, Z. S., 1993, *Doppler Radar and Weather Observations*, 2nd edition, Academic Press, San Diego.

Draxler, R. R., and Hess, G. D., 1998, An overview of the HYSPLIT_4 modelling system for trajectories, dispersion and deposition, Australian Meteorological Magazine, vol 47, pp 295-308.

Draxler, R. R., 1996, Trajectory Optimization for Balloon Flight Planning, *Weather and Forecasting,* vol 11, March, pp 111-114.

Fabry, F., and Zawadzki, I., 1995, Long-term radar observations of the melting layer of precipitation and their interpretation, *Journal of Atmospheric Science.*, vol 52, pp 838-851.

Fabry, F., 1994, *Observations and Uses of High Resolution Radar Data from Precipitation*, PhD Thesis, McGill University.

Fabry, F., Austin, G. L., and Tees, D., 1992, The accuracy of rainfall estimates by radar as a function of range, *Quarterly Journal of Royal Meteorological Society*, vol 118, pp 435-453.

Faisal, H., Anagnostou, E. N., Dinku, T., and Borga, M., 2004, Hydrological model sensitivity to parameter and radar rainfall estimation uncertainty, *Hydrological Processes*, vol 18, pp 3277-3291.

Fujita, T. T., 1985, The downburst: microburst and macroburst, *Satellite and Meteorology Research project*, Research paper no 20, University of Chicago.

Fulton, R. A., Breidenbach, J. P., Seo, D. J., Miller, D. A., 1998, The WSR-88D rainfall algorithm, *Weather and Forecasting*, vol 13, pp 377-394.

Geotis, S. G., 1963, Some radar measurements of hailstorm, *Journal of Applied Meteorology*, vol 2, pp 270-275.

Ghazel, M., 2004, Adaptive Fractal and Wavelet Image Denoising, PhD thesis, University of Waterloo, Waterloo, Canada.

Goddard, J. W. F., Davies, P. G., Thurai, M., and Morgan, K. M., 1991, Radar estimates of attenuation at 30 GHz: comparison with Olympus beacon measurements, *Antennas and Propagation*, *IEEE Conference Publication 333*, York.

Grecu, M., and Krajewski, W. F., 2000, An efficient methodology for detection of anomalous propagation echoes in radar reflectivity data using neural networks, *Journal of Atmospheric and Oceanic Technology*, vol 17, no 2, pp 121-129.

Gunn, K. L. S., and Marshall, J. S., 1955, The effect of wind shear on falling precipitation, *Journal of Applied Meteorology,* vol 12, pp 339-349.

Gunn, K. L. S., and East, T. W. R., 1954, The microwave properties of precipitation particles, *Quarterly Journal of Royal Meteorological Society*, vol 80, pp 522-545.

Hagan, M. T., and Menhaj, M. B., 1994, Training feed forward networks with the Marquardt algorithm, *IEEE Transactions on Neural Networks*, vol 5, no 6. pp 989-993.

Hannesen, R., 2002, *An Enhanced Surface Rainfall Algorithm for Radar Data*, Deliverable 4.4, Multiple-Sensor Precipitation Measurements, Integration, Calibration and Flood Forecasting (MUSIC).

Harouche, I., and Rasmussen, P., 2002, *Preliminary Literature Review of Precipitation Estimation Methods based on Multi-sensor data*. University of Manitoba.

Harrold, T. W., English, E. J., and Nicholass, C. A., 1974, Fall speeds and masses of solid precipitation particles, *Journal of Geophysical Research*, vol 79, pp 2185-2197.

Hessami, M., Anctil, F., and Viau, A. A., 2004, Selection of an Artificial Neural Network Model for the post-calibration of weather radar rainfall estimation, *Journal of Data Science*, vol 2, pp 107-124.

Hessami, M., Anctil, F. and Viau, A. A., 2003, An adaptive neuro-fuzzy inference system for the post-calibration of weather radar rainfall estimation, *Journal of Hydro-informatics*, vol 5, pp 63-70, IWA Publishing.

Hiroshima, K., 1999, Rainfall observation from tropical and rainfall measuring mission (TRMM) satellite*, Journal of Visualization*, vol 2, no 1, pp 93-98.

Hong, Y., Kummerow, C. D., and Olsen, E., 1998, Separation of convective and stratiform precipitation using microwave brightness temperature, *Journal of Applied Meteorology*, vol 38, pp 1195-1213.

Houze, R. A. Jr., 1993, *Cloud Dynamics*, Academic Press.

Hsu, K., Gao, X., Sorooshian, S., and Gupta, H., 1997, Precipitation estimation from remotely sensed information using Artificial Neural Networks, *Journal of Applied Meteorology*, vol 36, no 9, pp 1176-1190.

Hsu, K., Gupta, H. V., and Sorooshian, S., 1995, Artificial neural networking modeling of the rainfall-runoff Process, *Water Resources Research*, vol 3, no 10, pp 2517-2530.

Hunter, S. M., 2005, WSR-88D radar rainfall estimation: Capabilities, potentials and possible improvements, http://www.srh.noaa.gov/mrx/research/precip/PRECIP.htm, (web-site accessed on May 08, 2005).

Hunter, S. M., 1993, A limiting case for the WSR-88D - a severe "gustnado", *Proc. 26th Conference on Radar* Meteorology, American Meteorological Society, Boston, pp 660-663.

Huff, F. A., and Shipp, W. L., 1969, Spatial correlation of storm, monthly and seasonal precipitation, *Journal of Applied Meteorology*, vol 8, pp 543-555.

Issaks, E. H., and Srivasta, R. M., 1989, *An Introduction to Applied Geostatistics*, Oxford University Press, Oxford.

Jayakrsihnan, R., Srinivasan, R., and Arnold, J. G., 2004, Comparison of raingauge and WSR-88D stage III precipitation data over the Texas-Gulf basin, *Journal of Hydrology*, vol 292, pp 135-152.

Johnson, J. T., MacKeen, P. L., Witt, A., Mitchell, E. D., Stumpf, G. J., Eilts, M. D., and Thomas, K. W., 1998, The storm cell identification and tracking (SCIT)

algorithm: an enhanced WSR-88D algorithm, *Weather Forecasting*, vol 13, pp 263–276.

Jordan, P., 2000, Errors in Radar Measurements of Rainfall- Effect on Flood Forecasting, www.smec.com.au/media/papers/hydro200.pdf, website accessed on June 16, 2005.

Kitchen, M., Brown, R., and Davies, A. G., 1994, Real-time correction of weather radar data for the effects of bright band, range and orographic growth in widespread precipitation, *Quarterly Journal Royal Meteorological Society*, vol 120, pp 1231-1254.

Kitchen, M., and Jackson, P. M., 1993, Weather radar performance at long range - simulated and observed, *Journal of Applied Meteorology*, vol 32, pp 975-985.

Kitchen, M., and Blackball, R. M., 1992, Representativeness errors in comparisons between radar and gauge measurements of rainfall, *Journal of Hydrology*, vol 134, pp 13-33.

Klazura, G., and Imy, D., 1993, A description of the initial set of analysis products available from the NEXRAD WSR-88D system, *Bulletin of American Meteorological Society*, vol 74, pp 1293-1311.

Kucera, P. A., Short, D. A., and Thiele, O. W, 2005, An analysis of rainfall intensity and vertical structure from ship borne radars in TOGA COARE, http://www.iihr.uiowa.edu/ products/water_resources/publications.html, (web page accessed on May 20, 2005).

Krajewski, W. F., 2004, Rainfall estimation using weather radar and ground station, http://www.unesco.org.uy/phi/libros/radar/art04.html.

Krajewski, W. F., and Smith, J. A., 2002, Radar hydrology: rainfall estimation, *Advances in Water Resources*, vol 25, pp 1387-1394.

Krajewski, W. F., 1987, Co-kriging radar rainfall and rain-gauge data, *Journal of Geophysical Research*, vol 92, pp 9571-9580.

Krajewski, W. F., and Vignal, B., 2000, Evaluation of anomalous propagation echo detection in WSR 88D data: a large sample case study, *Journal of Oceanic and Atmospheric Technology,* vol 18, no 5, pp 807-814.

Lack, S. A., and Fox, N. I., 2004, Errors in surface rainfall rates retrieved from radar due to wind drift, *Sixth International Symposium on Hydrological Application of Weather Radar Proceedings*, Melbourne, Australia.

Lee, M., 2005, *Combination of Radar and Raingauge Data for Improved Aerial Rainfall Estimation*, B.Sc. Thesis, University of Manitoba.

Lee, G., Berenguer, M., Zawadzki, I., Torres, D. S., 2002, Errors in rain measurement by C-band radar due to attenuation, *QPF Conference 2002 Proceedings*, Royal Meteorological Society, (http://www.rmets.org/event/conference/qpfproc.php) website accessed on Nov 1,2005.

Legates, D. R., 2000, Real-time calibration of radar precipitation estimates, *Professional Geographer*, vol 52, no 2, pp 235-241.

Lemon, L. R., 1978, On the use of storm structure for hail identification, *Proc. 18th Conference of Radar Meteorology*, American Meteorological Society, pp 203-206.

Lhermitte, R. M., 1970, Dual doppler radar observation of convective storm circulation, Proc., 14th Radar Meteorology Conference, *American Meteorology Society,* pp 139-144.

Lhermitte, R. M., and Atlas, D., 1961, Precipitation motion by pulse doppler radar, *Proc. 9th Weather Radar Conference,* pp 218-223.

Lin, S. F., Rohde, G., and Aldroubi, A., 1998, Wavelet denoising of epicardial fluorescence imaging, *Proc. SPIE symposium on to wavelet applications in signal and image processing*, vol 4.

Li, L., Schmid, W., and Joss, J., 1995, Now casting of motion and growth of precipitation with radar data over a complex orography, *Journal of Applied Meteorology*, vol 34, pp 1286-1300.

Liu, C., and Illingworth, A. J., 2001, Polarization algorithm to correct for attenuations at C band, http://www.met.rdg.ac.uk/radar/research/cband_atten/, web site accessed on Oct 28, 2005.

List, R., 1991, A linear radar reflectivity rain rate relationship for steady tropical rain, *Journal of Atmospheric Sciences*, vol 45, pp 3564-3572.

Liu, H., Chandrasekar, V., and Xu, G., 2001, An adaptive neural network scheme for radar rainfall estimation from WSR-88D observations, *Journal of Applied Meteorology*, vol 40, pp 2038-2050, American Meteorological Society.

Lutgens, F. K., and Tarbuck, E. J., 2001, *The Atmosphere,* Eighth Edition, Prentice Hall, New Jersey, pp 238.

Marshall, J. S., 1968, Power-law relations in radar meteorology, *Journal of Applied Meteorology,* vol 8, pp 171-172.

Marshall, J. S., and Gunn, R. E. S., 1954, The effect of wind shear on falling precipitation, *Journal of Atmospheric Sciences,* vol 12, no 4, pp 339-349.

Marshall, J. S., and Palmer, W. M., 1948, The distribution of raindrop with size, *Journal of Meteorology,* vol 5, pp 165-166.

Marshall, J. S., Langille, R. C., and Palmer, W. M., 1947, Measurement of rainfall by radar, *Journal of Meteorology,* vol 4, pp 186-192.

May, P. T., and Keenan, T. D., 2003, Four dimensional micro-physical data from Darwin, *Thirteenth ARM Science Team Meeting Proceedings*, Broomfield, Colorado.

Maynard, R. H., 1945, Radar and Weather, *Journal of Meteorology*, vol 2, no 4, pp 214-226.

Miller, J. R., 1972, A climatological Z-R relationship for convective storms in the northern great plains, *Proceedings 15th Conference of Radar Meteorology,* American Meteorological Society, pp 153-154.

Mittermair, M. P., Hogan, R. J., Illingworth, A. J., 2004, Using meso-scale model winds for correcting wind-drift errors in radar estimates of surface rainfall, *Journal of Royal Meteorological Society,* vol 130, pp 1-20.

Mizzell, H. P., 1999, *Comparison of WSR 88D Derived Rainfall Estimates with Gauge Data in Lexington County*, M.Sc. Thesis, Dept. of Geography, University of South Carolina, USA.

Moszkowicz, M., Ciach, G. J., and Krajewski, W. F., 1993, Statistical detection of anomalous propagation in radar reflectivity patterns, *Journal of Atmospheric and Oceanic Technology*, vol 11, no 4, pp 1026-1034.

MSC, 2004, *Fact sheet, Weather radar*, Meteorological Service of Canada, Environment Canada, http://www.msc-smc.ec.gc.ca/cd/factsheets/weather_radar/index_e.cfm (web page accessed on June 2004).

Mukto, M. A., 2001, *Short Term River Water Level Forecasting using Artificial Neural Networking*, M.Sc. Thesis, BUET, Dahaka.

Pamment, J, A., and Conway, B. J., 1998, Objective identification of echoes due to anomalous propagation in weather radar data. *Journal of Atmospheric and Oceanic Technology*, vol 15, no 1, pp 98-113.

Parker, M. D., and Johnson, R. H., 2000, Organizational modes of mid latitude Mesoscale Convective System, *Monthly Weather Review*, vol 128, pp 3413-3436.

Phil, P., 2000, *Neural Networks*, 2nd Edition, Grass root Series, MacMillan Publisher, England.

Raghavan, S., 2003, *Radar Meteorology,* Kluwer Academic Publishers, The Netherlands.

Rigo, T., and Llsat, M. C., 2004, A methodology for the classification of convective structures using meteorological radar: Application to heavy rainfall events on the Mediterranean coast of the Iberian peninsula, *Natural Hazards and the Earth System Sciences*, vol 4, pp 59-68.

Rinehart, R. E., and Garvey, E. T., 1978, Three-dimensional storm motion detection by conventional weather radar, *Nature*, vol 273, pp 287-289.

Rosenfeld D., Wolff D. B., and Atlas, D., 1993, General probability matched relations between radar reflectivity and rain rate, *Journal of Applied Meteorology*, vol 32, pp 50-72.

Rogers, R. R., 1964, An extension of the Z-R relationship for the Doppler radar, *Proc. 11th Weather Radar Conference,* American Meteorological Society, Boston, pp 170-174.

Ryzkov, A., and Zrnic, D., 1995, Precipitation and attenuation measurements at a 10-cm wavelength, *Journal of Applied Meteorology*, vol 34, pp 2121-2134.

Sauvageot, H., 1991, *Radar Meteorology*, Artech House, Boston, MA.

Sachidananda, M., and Zrnic, D. S., 2000, Clutter filtering and spectral moment estimation for doppler weather radars using staggered pulse repetition time (prt), *Journal of Atmospheric and Oceanic Technology*, vol 17, no 3, pp 323-331.

Seed, A. W., Nicol, J., Austin, G. L., Stow, C. D., and Bradley, S. G., 1996, The impact of radar and raingauge sampling errors when calibrating weather radar, *Meteorological Application*, vol 3, pp 43–52.

Seo, D. J., Breidenbach, J., Fulton, R., Miller, D., Vignal, B., and Krajewski, W., 2000, *Final Report for June 1, 1999 to May 31, 2000,* Interagency Memorandum for Understanding Weather Radar Among the NEXRAD Program, http://www.nws.noaa.gov/oh/hrl/papers/2000mou/Report/Index.html (Web site accessed on May 5, 2005).

Smith, J. A., and Krajewski, W. F., 1991, Estimation of mean field bias of radar rainfall estimates, *Journal of Applied Meteorology*, vol 30, no 4, pp 397-412.

Sokol, Z., 2003, Utilization of regression models for rainfall estimates using radar-derived rainfall data and rain gauge data, *Journal of Hydrology*, vol 278, pp 145-152.

Starck, J. L., Candes, E. J., Donoho, D. L., 2002, The curve-let transform for image denoising, *IEEE Transactions on Image Processing*, vol 11, no 6, pp 670-684.

Steiner, M., and Smith, J. A., 2002, Use of three-dimensional reflectivity structure for automated detection and removal of non-precipitating echoes in radar data, *Journal of Atmospheric and Oceanic Technology*, vol 19, no 5, pp 673-686.

Steiner, M., Houze, Jr. R. A., and Yuter, S., 1995, Climatological characterization of a 3-D storm structure from operational radar and raingauge data, *Journal of Applied Meteorology,* vol 34, pp 1978-2007.

Stull, R. B., 1995, *Meteorology for Scientists and Engineers*, 2nd Edition, Thomson Learning, USA.

Sun, X., Mein, R. G., Keenan, T. D., and Elliott, J. F., 2000, Flood estimation using radar and gauge data, *Journal of Hydrology*, vol 239, pp 4-18.

Todini, E., 2001, A Bayesian approach for conditioning radar precipitation estimates to raingauge measurements, *Hydrology and Earth System Sciences*, vol 5, no 2, pp 187-199.

Torres, S. M., and Zrnic, D. S., 1999, Ground clutter cancelling with a regression filter, *Journal of Atmospheric and Oceanic Technology*, vol 16, no 10, pp 1364-1372.

Uijlenhoet, R. U., Steiner, M., and Smith, J. A., 2003, Variability of raindrop size distributions in a squall line and implication for radar rainfall estimation, *Journal of Hydrometeorology,* vol 4, pp 43-61.

Vieux, B. E., 2002, Radar Rainfall Applications in Hydrology, *Hydrology and Floodplain Analysis,* 3rd edition, Prentice Hall publishers, NJ.

Walker, J. S., 1999, *A Primer on Wavelets and their Scientific Applications*, CRC Press LLC, Florida, USA.

Wasserman, P. D., 1993, *Advanced Method in Neural Computing*, Van Nostrand Reinhold, NY.

Wexler, R., and Atlas, D., 1963, Radar reflectivity and attenuation of rain, *Journal of Applied Meteorology*, vol 2, pp 276-280.

Wilson, J., and Brandes, E., 1979, Radar measurement of rainfall - a summary, *Bulletin of American Meteorological Society,* vol 60, no 9, pp 1048-1058.

Wolfson, M. M., Forman, B. E., Hallowell, R. G., Moore, M. P., 1999, The growth and decay storm tracker, presented at the 79th annual conference of the American Meteorological Society, www.ll.mit.edu/AviationWeather/growth_and_decay_tracker1.pdf (web-site accessed on Sept 18, 2005).

Xiao, R., Chandrasekar, V., and Liu, H., 1998, Development of a neural network based algorithm for radar snowfall estimation, *IEEE Transactions on Geoscience and Remote Sensing*, vol 36, no 3, pp 716-724.

Xiao, R., Chandrasekar, V., and Gorgucci, E., 1998, Detection of rain/no rain condition on ground from radar data using a kohonen neural network, *IEEE Geoscience and Remote Sensing Symposium Proceedings*, pp 159-160.

Xiao, R., and Chandrasekar, V., 1997, Development of a neural network based algorithm for rainfall estimation from radar observations, *IEEE Transactions on Geoscience and Remote Sensing*, vol 35, no 1, pp 160-171.

Zawadzski, I. I., 1973, Statistical properties of precipitation patterns, *Journal of Applied Meteorology,* vol 12, pp 459-472.

Zealand, C. M., 1997, *Short Term Stream Flow Forecasting Using Artificial Neural Networking*, M.Sc. thesis, University of Manitoba, Canada.

Zurada, J. M., 1992, *An Introduction to Artificial Neural Systems,* PWS Publishing Company, USA.

# Appendix A: Weights and Biases

Table A.1: Weight matrix connecting the input layer and the hidden layer and the bias vector for the hidden layer (back-propagation network on temporal pattern).

| | | | | | |
|---|---|---|---|---|---|
| 0.044218 | 0.025613 | -0.033935 | -0.054842 | 0.02226 | -5.056 |
| -0.1742 | 0.53014 | 0.31079 | 0.21091 | 0.28428 | 3.009 |
| 0.026776 | -0.1242 | -0.11801 | -0.0054976 | 0.10656 | 4.5422 |
| 0.91916 | -0.14152 | -2.2581 | -0.77173 | -0.7293 | 3.0881 |
| -0.73799 | -0.70116 | -2.2947 | -2.6188 | 1.4091 | 1.8302 |
| 3.151 | 1.6781 | 0.41823 | 1.3729 | -0.3139 | -2.1704 |
| 0.13259 | -1.58 | -5.2624 | 2.9197 | 3.5311 | -1.0481 |
| 1.4979 | 1.6757 | -3.3774 | -0.31387 | -0.3875 | 3.8112 |
| -0.19077 | 0.31233 | 0.6892 | 0.27341 | -4.1603 | 2.4716 |
| 1.9168 | 1.5506 | 0.70504 | 2.207 | 6.1254 | -4.0941 |
| 0.032303 | 0.019221 | 0.0096734 | -0.055956 | -0.071487 | 5.6381 |
| -0.012003 | 0.011914 | 0.058388 | 0.041171 | 0.047437 | -7.0037 |

Table A.2: Weight matrix connecting the 1st hidden layer to the 2nd hidden layer and the bias vector for the 2nd hidden layer (back-propagation network on temporal pattern).

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.892 | -2.892 | -3.1006 | 1.872 | -0.3834 | 3.5013 | -1.2063 | -0.5771 | 2.2374 | -0.8158 | -1.1938 | -1.0397 | -2.5908 |
| -3.0542 | 3.576 | 0.29897 | 0.06769 | 1.0791 | 3.0549 | -0.1555 | -0.5789 | 1.874 | -3.2722 | -0.1511 | -0.3716 | 2.263 |
| -1.0049 | 0.56898 | 2.4971 | -0.8354 | 1.4688 | 1.7271 | 2.444 | -2.9286 | 2.147 | 2.7374 | -1.6211 | -0.8234 | -2.616 |
| -0.5605 | -0.2672 | 2.7305 | -1.9068 | -1.1792 | 2.9718 | -1.4348 | 0.08024 | -4.7661 | 1.5693 | -3.4314 | 0.45312 | 4.4054 |
| 2.2811 | -2.2097 | -2.9081 | -2.4938 | 0.59494 | -0.6687 | 0.09758 | 0.65929 | -0.8689 | -4.1759 | -1.6559 | -1.771 | 4.8485 |
| 1.5647 | 3.0916 | -2.572 | 2.1636 | 0.38443 | -1.382 | 2.5425 | 1.6245 | -2.7155 | 1.3587 | -1.9579 | -0.1861 | -2.0762 |
| 2.9084 | 0.6955 | -0.3437 | -1.7512 | 3.893 | -3.6291 | 3.3262 | -1.5696 | -6.796 | 4.8026 | 1.2262 | 0.73119 | -2.3278 |
| 2.7605 | -0.6308 | 1.0019 | -0.4660 | -2.4141 | 3.0057 | -1.752 | -2.2825 | 0.24868 | 1.8637 | 1.2405 | 3.1374 | -3.7346 |
| -0.2743 | -1.8784 | 0.9197 | 4.2998 | 3.9784 | -1.3734 | -1.8483 | 4.1265 | -3.5061 | -1.9674 | -2.5731 | -0.6338 | -0.3317 |
| -1.4727 | -1.6779 | -1.8693 | 0.94218 | -4.0499 | 0.75911 | -3.0822 | 0.271 | 0.41549 | 0.68491 | 1.752 | 3.03 | -1.6313 |
| -0.8509 | 3.4219 | 2.1464 | -0.0521 | 0.59249 | 2.3825 | 3.8474 | -0.1830 | -3.6227 | 5.4305 | -0.9155 | -1.3279 | -7.6036 |
| -2.2908 | 1.0293 | 0.21652 | -1.9033 | 3.5653 | 1.4188 | 2.4074 | -0.8396 | -3.0505 | -1.6921 | 2.1109 | 2.8456 | -4.6889 |

Table A.3: Weight matrix connecting the 2nd hidden layer to the output layer and the bias vector for the output layer (back-propagation network on temporal pattern).

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.48143 | -0.5569 | -2.4115 | -3.3946 | 1.5692 | -1.5083 | 5.366 | -1.9091 | -6.3809 | -1.4566 | 3.2471 | -2.6566 | 2.94 |

Table A.4: Weight matrix connecting the input layer and the 1st hidden layer and the bias vector for the hidden layer (back-propagation network on spatial pattern).

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0.082895 | 0.096541 | 0.14608 | 0.06817 | 0.011954 | 0.16102 | 0.090499 | -0.008960 | 0.23129 | 2.6488 |
| 0.67677 | 0.41278 | 0.98111 | 0.42535 | -0.018136 | 0.62271 | 0.66866 | 0.19636 | 1.2006 | -3.0561 |
| 0.96309 | 0.32116 | 1.4556 | 0.48228 | -0.58333 | 0.48762 | 1.1994 | -0.13566 | 1.6598 | -3.918 |
| -0.33836 | 0.48879 | -0.7201 | 0.16151 | 1.2595 | 0.11868 | -0.21632 | 0.94857 | -1.0792 | -0.61681 |
| 0.41233 | 0.09323 | 0.83811 | 0.20029 | -0.42168 | 0.43057 | 0.30854 | -0.18643 | 1.0214 | 0.48154 |
| -0.029896 | 0.29573 | 0.11165 | 0.069264 | 0.29026 | 0.23547 | 0.098294 | 0.23879 | 0.069171 | -1.036 |
| -0.054599 | 0.14602 | -0.41258 | 0.037283 | 0.48505 | -0.24739 | -0.001873 | 0.29293 | -0.72164 | 2.1943 |
| -0.19981 | -0.14387 | -0.30538 | -0.14189 | -0.1285 | -0.25359 | -0.14234 | -0.10131 | -0.3472 | -3.2251 |

Table A.5: Weight matrix connecting the 1st hidden layer to the 2nd hidden layer and the bias vector for the 2nd hidden layer (back-propagation network on spatial pattern).

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.85896 | -1.7965 | -2.2804 | 3.7612 | -3.0486 | 1.361 | 3.4667 | -3.251 | -2.4636 |
| 1.5394 | -1.6105 | -2.9385 | -1.1243 | 2.8556 | -0.51062 | -3.0602 | -4.4964 | 2.7233 |
| -2.4002 | -1.9047 | 3.0175 | -2.0432 | 1.6399 | -2.3223 | -0.32272 | -4.4523 | 6.6583 |
| 2.0928 | -3.7643 | -1.3435 | -3.7882 | -0.32151 | -3.1824 | 0.39959 | 3.0081 | 3.3479 |
| -2.4627 | 2.4702 | 2.8748 | 3.0475 | -2.174 | 2.978 | -1.1395 | 2.8136 | -4.9382 |
| 2.0111 | 2.657 | -1.8154 | 2.5434 | -3.1753 | 1.1355 | -4.6255 | -1.5867 | 0.42237 |
| 0.022257 | -1.9504 | -5.1959 | 3.6499 | -2.0813 | -0.46152 | 0.1273 | 4.3224 | 2.8203 |
| -2.8321 | 3.4397 | 3.1473 | -3.115 | -2.8497 | 1.7423 | -1.484 | -0.17076 | -3.0443 |

Table A.6: Weight matrix connecting the 2nd t hidden layer to the output layer and the bias vector for the output layer (back-propagation network on spatial pattern).

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| -2.4872 | 1.516 | -2.5639 | 1.7678 | 0.79923 | 1.634 | -4.4181 | 2.7866 | 1.1008 |

Table A.7: Weight matrix for the generalized regression network for new validation data.

| | | |
|---|---|---|
| -0.062459 | -0.08825 | 0.0024154 |
| 0.16297 | -0.13862 | 0.012998 |
| 0.062235 | 0.011091 | 0.0039195 |
| 0.5687 | 0.36301 | 0.10163 |
| 0.18426 | 0.23722 | 0.0097669 |
| 0.50477 | 0.013531 | 0.034382 |
| 0.06052 | 0.43213 | 0.0098935 |
| -0.46077 | -0.45976 | 0.0015972 |
| -0.24557 | -0.18007 | 0 |
| -0.45922 | -0.041619 | 0.0018657 |
| 0.48021 | 0.29559 | 0.043518 |
| 0.74994 | 0.71496 | 0.31912 |
| 0.33975 | 0.61713 | 0.083293 |
| -0.0067774 | 0.27261 | 0.0040376 |
| 0.6149 | 0.66064 | 0.2779 |
| 0.15491 | 0.82153 | 0.16452 |
| 0.34799 | 0.10569 | 0.024135 |
| 0.71745 | 0.55218 | 0.30822 |
| 0.59018 | 0.54278 | 0.13837 |
| 0.33148 | -0.034013 | 0.018479 |
| 0.28411 | 0.33157 | 0.013213 |
| 0.11216 | -0.44619 | 0.0073372 |
| 0.45014 | 0.4576 | 0.053268 |
| 0.50518 | 0.79441 | 0.21773 |
| 0.49818 | 0.61833 | 0.13429 |
| 0.69091 | -0.090168 | 0.20923 |
| 0.62929 | 0.25122 | 0.090864 |
| 0.52186 | 0.52841 | 0.080874 |
| 0.37953 | -0.44585 | 0.033914 |
| -0.10503 | 0.11813 | 0.0049966 |
| -0.44432 | 0.34522 | 0.0094774 |
| 0.45426 | 0.37755 | 0.050682 |
| 0.7996 | 0.11088 | 0.2323 |
| 0.16158 | 0.3394 | 0.0094813 |
| 0.076826 | 0.14011 | 0.0038779 |
| 0.28844 | 0.21546 | 0.010361 |
| 0.36588 | 0.4947 | 0.038355 |
| 0.35645 | 0.40482 | 0.043991 |
| 0.4617 | 0.1687 | 0.037675 |
| 0.44724 | 0.54003 | 0.057162 |
| 0.21074 | 0.10692 | 0.011497 |
| 0.82218 | 0.30591 | 0.2487 |
| -0.11434 | -0.44923 | 0.0023614 |
| 0.24242 | 0.47346 | 0.034701 |
| 0.092733 | 0.61392 | 0.08839 |
| -0.10146 | 0.41933 | 0.024963 |
| 0.63256 | 0.45222 | 0.17168 |
| -0.064397 | 0.76371 | 0.12386 |
| 0.52951 | 0.45781 | 0.087891 |
| 0.37077 | 0.28854 | 0.023364 |

# Appendix B: Performance of the ANN

Table B.1: Performance of the back-propagation network for the temporal input pattern

| Net ID | Input Neuron | Hidden layer | Neuron in Hidden Layer | Epoch | RMSE Training | CC Training | RMSE Validation | CC Validation | RMSE Test (i) | CC Test (i) | RMSE Test (ii) | CC Test (ii) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.1 | 1 | 1 | 8 | 1 | 1.8761 | -0.3192 | 2.1994 | -0.3297 | 2.3475 | -0.3049 | 1.0947 | -0.2795 |
| 1.2 | 2 | 1 | 8 | 1 | 1.8765 | -0.3316 | 2.201 | -0.3322 | 2.3492 | -0.3225 | 1.0956 | -0.3710 |
| 1.3 | 3 | 1 | 8 | 30 | 1.3098 | 0.6341 | 1.5697 | 0.63233 | 1.6901 | 0.62792 | 0.90167 | 0.50361 |
| 1.4 | 4 | 1 | 8 | 24 | 1.2519 | 0.67388 | 1.5523 | 0.64626 | 1.6478 | 0.65729 | 0.83897 | 0.58133 |
| 1.5 | 5 | 1 | 8 | 19 | 1.2013 | 0.70693 | 1.461 | 0.69094 | 1.5908 | 0.67679 | 0.91462 | 0.48898 |
| 1.6 | 1 | 1 | 12 | 31 | 1.4935 | 0.47119 | 1.756 | 0.49128 | 1.9314 | 0.4422 | 1.0499 | 0.32321 |
| 1.7 | 2 | 1 | 12 | 71 | 1.3391 | 0.61218 | 1.5783 | 0.62875 | 1.692 | 0.6339 | 0.82054 | 0.60203 |
| 1.8 | 3 | 1 | 12 | 10 | 1.3945 | 0.58197 | 1.5964 | 0.60247 | 1.7078 | 0.61112 | 0.86121 | 0.557 |
| 1.9 | 4 | 1 | 12 | 15 | 1.2888 | 0.6517 | 1.5135 | 0.67079 | 1.6565 | 0.64786 | 0.89859 | 0.51404 |
| 1.10 | 5 | 1 | 12 | 19 | 1.297 | 0.64428 | 1.5132 | 0.6701 | 1.6309 | 0.6716 | 0.9143 | 0.4905 |
| 1.11 | 1 | 1 | 16 | 1 | 1.8761 | -0.2736 | 2.1994 | -0.2793 | 2.3475 | -0.2601 | 1.0947 | -0.2576 |
| 1.12 | 2 | 1 | 16 | 41 | 1.3437 | 0.61072 | 1.5611 | 0.63361 | 1.6727 | 0.63916 | 0.79243 | 0.63789 |
| 1.13 | 3 | 1 | 16 | 32 | 1.3199 | 0.62828 | 1.5609 | 0.63146 | 1.6731 | 0.63448 | 0.93826 | 0.45195 |
| 1.14 | 4 | 1 | 16 | 23 | 1.2573 | 0.67248 | 1.5244 | 0.65829 | 1.6591 | 0.64114 | 0.93558 | 0.43937 |
| 1.15 | 5 | 1 | 16 | 23 | 1.1491 | 0.73519 | 1.4326 | 0.71831 | 1.5738 | 0.693 | 1.0254 | 0.36132 |
| 1.16 | 1 | 1 | 20 | 14 | 1.5046 | 0.45877 | 1.7607 | 0.48488 | 1.9418 | 0.42976 | 1.047 | 0.3097 |
| 1.17 | 2 | 1 | 20 | 26 | 1.3394 | 0.61247 | 1.5749 | 0.62757 | 1.687 | 0.63244 | 0.82024 | 0.60452 |
| 1.18 | 3 | 1 | 20 | 24 | 1.3236 | 0.62528 | 1.5721 | 0.62487 | 1.6782 | 0.63471 | 0.9313 | 0.45573 |
| 1.19 | 4 | 1 | 20 | 29 | 1.2778 | 0.65677 | 1.5357 | 0.673 | 1.681 | 0.64199 | 0.90025 | 0.51399 |
| 1.20 | 5 | 1 | 20 | 19 | 1.2447 | 0.68336 | 1.4623 | 0.6891 | 1.6097 | 0.66781 | 1.2912 | 0.36402 |
| 1.21 | 1 | 1 | 24 | 23 | 1.4927 | 0.47211 | 1.7556 | 0.49155 | 1.931 | 0.44276 | 1.0506 | 0.32304 |
| 1.22 | 2 | 1 | 24 | 35 | 1.3383 | 0.61456 | 1.5647 | 0.63135 | 1.6757 | 0.63573 | 0.80547 | 0.62152 |
| 1.23 | 3 | 1 | 24 | 34 | 1.265 | 0.66551 | 1.5169 | 0.67073 | 1.6593 | 0.64464 | 1.0015 | 0.37094 |
| 1.24 | 4 | 1 | 24 | 22 | 1.2638 | 0.666 | 1.5583 | 0.64512 | 1.6379 | 0.67274 | 0.92269 | 0.45466 |
| 1.25 | 5 | 1 | 24 | 25 | 1.1169 | 0.75327 | 1.4629 | 0.6872 | 1.5409 | 0.70358 | 1.2834 | 0.25219 |
| 1.26 | 1 | 2 | 4 | 35 | 1.4976 | 0.46673 | 1.7604 | 0.48708 | 1.9372 | 0.43602 | 1.0525 | 0.31056 |
| 1.27 | 2 | 2 | 4 | 39 | 1.3347 | 0.61555 | 1.5834 | 0.62617 | 1.6874 | 0.63733 | 0.81488 | 0.61461 |
| 1.28 | 3 | 2 | 4 | 30 | 1.2868 | 0.65034 | 1.5561 | 0.64521 | 1.6617 | 0.64805 | 0.95596 | 0.3996 |
| 1.29 | 4 | 2 | 4 | 16 | 1.3098 | 0.6353 | 1.5743 | 0.6315 | 1.6879 | 0.63384 | 0.86788 | 0.5431 |
| 1.30 | 5 | 2 | 4 | 21 | 1.228 | 0.68925 | 1.4739 | 0.68434 | 1.6446 | 0.64795 | 0.92473 | 0.47801 |
| 1.31 | 1 | 2 | 6 | 35 | 1.5021 | 0.46153 | 1.7615 | 0.48673 | 1.9408 | 0.43244 | 1.044 | 0.31182 |
| 1.32 | 2 | 2 | 6 | 47 | 1.3462 | 0.60698 | 1.5765 | 0.62777 | 1.6901 | 0.63209 | 0.82331 | 0.59948 |
| 1.33 | 3 | 2 | 6 | 25 | 1.2417 | 0.68367 | 1.5564 | 0.62885 | 1.6395 | 0.64586 | 0.9454 | 0.49037 |
| 1.34 | 4 | 2 | 6 | 13 | 1.3374 | 0.61405 | 1.5811 | 0.63232 | 1.6806 | 0.64884 | 0.88229 | 0.51763 |
| 1.35 | 5 | 2 | 6 | 16 | 1.2271 | 0.6899 | 1.4677 | 0.6904 | 1.5914 | 0.68168 | 0.82717 | 0.59897 |
| 1.36 | 1 | 2 | 8 | 36 | 1.4803 | 0.48827 | 1.7435 | 0.48987 | 1.8914 | 0.47317 | 1.1228 | 0.30743 |
| 1.37 | 2 | 2 | 8 | 27 | 1.3327 | 0.61713 | 1.565 | 0.63869 | 1.673 | 0.64475 | 0.81193 | 0.61502 |
| 1.38 | 3 | 2 | 8 | 41 | 1.2366 | 0.68346 | 1.5564 | 0.64404 | 1.685 | 0.62355 | 0.99225 | 0.39534 |
| 1.39 | 4 | 2 | 8 | 43 | 1.136 | 0.74194 | 1.5115 | 0.66844 | 1.6444 | 0.64608 | 1.0424 | 0.29125 |
| 1.40 | 5 | 2 | 8 | 23 | 1.1706 | 0.7231 | 1.4533 | 0.70282 | 1.5898 | 0.68432 | 0.94195 | 0.44627 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.41 | 1 | 2 | 10 | 26 | 1.5027 | 0.46093 | 1.7607 | 0.48603 | 1.9407 | 0.43151 | 1.0467 | 0.31088 |
| 1.42 | 2 | 2 | 10 | 28 | 1.319 | 0.62719 | 1.5729 | 0.63197 | 1.6779 | 0.63998 | 0.77785 | 0.65395 |
| 1.43 | 3 | 2 | 10 | 1 | 1.8768 | -0.3306 | 2.2025 | -0.3256 | 2.3508 | -0.3223 | 1.0966 | -0.2313 |
| 1.44 | 4 | 2 | 10 | 17 | 1.2591 | 0.66909 | 1.5303 | 0.66787 | 1.6485 | 0.65375 | 0.94919 | 0.46461 |
| 1.45 | 5 | 2 | 10 | 24 | 1.138 | 0.74317 | 1.3962 | 0.72808 | 1.5244 | 0.7087 | 1.0124 | 0.37875 |
| 1.46 | 1 | 2 | 12 | 1 | 1.8761 | 0.36636 | 2.1994 | 0.38063 | 2.3475 | 0.34904 | 1.0947 | 0.30018 |
| 1.47 | 2 | 2 | 12 | 33 | 1.3395 | 0.61284 | 1.5654 | 0.63174 | 1.6825 | 0.6327 | 0.82619 | 0.59464 |
| 1.48 | 3 | 2 | 12 | 23 | 1.3375 | 0.61356 | 1.5773 | 0.62803 | 1.6913 | 0.63327 | 0.91206 | 0.48797 |
| 1.49 | 4 | 2 | 12 | 21 | 1.2714 | 0.66089 | 1.5383 | 0.66539 | 1.6485 | 0.66479 | 0.89426 | 0.51495 |
| 1.50 | 5 | 2 | 12 | 34 | 1.1204 | 0.75033 | 1.3748 | 0.75099 | 1.6047 | 0.67347 | 1.0654 | 0.32795 |

Table B.2: Performance of the back-propagation network for the spatial input pattern.

| Net ID | Input Neuron | Hidden layer | Neuron in Hidden Layer | Epoch | RMSE Training | CC Training | RMSE Validation | CC Validation | RMSE Test (i) | CC Test (i) | RMSE Test (ii) | CC Test (ii) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2.1 | 9 | 1 | 8 | 10 | 1.4938 | 0.49423 | 1.6838 | 0.53793 | 1.9202 | 0.44611 | 0.99407 | 0.39725 |
| 2.2 | 9 | 1 | 12 | 13 | 1.4398 | 0.52639 | 1.6868 | 0.55362 | 1.9217 | 0.44589 | 0.94351 | 0.41115 |
| 2.3 | 9 | 1 | 16 | 17 | 1.4278 | 0.54362 | 1.6764 | 0.54391 | 1.9016 | 0.46145 | 0.95906 | 0.39544 |
| 2.4 | 9 | 1 | 20 | 12 | 1.4453 | 0.52253 | 1.6871 | 0.54383 | 1.9222 | 0.44279 | 0.94631 | 0.41006 |
| 2.5 | 9 | 1 | 24 | 12 | 1.4383 | 0.52815 | 1.6957 | 0.54005 | 1.9278 | 0.4388 | 0.94324 | 0.40875 |
| 2.6 | 9 | 2 | 4 | 15 | 1.4253 | 0.54082 | 1.6809 | 0.54566 | 1.9108 | 0.45344 | 0.94533 | 0.40471 |
| 2.7 | 9 | 2 | 6 | 15 | 1.4353 | 0.53539 | 1.6856 | 0.53719 | 1.9186 | 0.44674 | 0.9647 | 0.37153 |
| 2.8 | 9 | 2 | 8 | 15 | 1.433 | 0.53316 | 1.6756 | 0.55625 | 1.9108 | 0.45422 | 0.94441 | 0.40925 |
| 2.9 | 9 | 2 | 20 | 26 | 1.4347 | 0.53133 | 1.6828 | 0.55203 | 1.9254 | 0.44079 | 0.93758 | 0.41473 |
| 2.10 | 9 | 2 | 24 | 10 | 1.451 | 0.52131 | 1.678 | 0.55061 | 1.9204 | 0.44371 | 0.9471 | 0.4125 |

Table B.3: Performance of the radial basis network for the temporal input pattern

| Net ID | No of Cluster | Input Neuron | Width | RMSE Training | CC Training | RMSE Validation | CC Validation | RMSE Test set (i) | CC Test set (i) | RMSE Test set (ii) | CC Test set(ii) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.1 | N/A | 1 | 0.05 | 1.4395 | 0.52651 | 13.812 | 0.13326 | 1.91 | 0.45622 | 1.1381 | 0.29004 |
| 3.2 | 10 | 1 | 0.05 | 1.524 | 0.44574 | 1.8278 | 0.42341 | 1.946 | 0.45185 | 1.0489 | 0.28015 |
| 3.3 | 50 | 1 | 0.05 | 7495.8 | 0.039446 | 9617.2 | 0.020887 | 8840.9 | 0.024593 | 23113 | 0.028706 |
| 3.4 | 100 | 1 | 0.05 | 48091 | 0.036812 | 1.29E+5 | -0.08986 | 27952 | 0.06037 | 67798 | -0.00412 |
| 3.5 | 500 | 1 | 0.05 | 1.4411 | 0.52532 | 171.53 | 0.1062 | 2.0171 | 0.38474 | 1.18 | 0.28235 |
| 3.6 | 1000 | 1 | 0.05 | 1.4684 | 0.51655 | 3.6327 | 0.23916 | 1.9661 | 0.41765 | 1.2696 | 0.26038 |
| 3.7 | N/A | 1 | 0.1 | 1.4569 | 0.50954 | 2.1731 | 0.21098 | 1.8418 | 0.52251 | 1.1306 | 0.30284 |
| 3.8 | 10 | 1 | 0.1 | 1.4837 | 0.48283 | 1.7682 | 0.47939 | 1.9216 | 0.45544 | 1.0452 | 0.3262 |
| 3.9 | 50 | 1 | 0.1 | 9429.2 | -0.01380 | 9756.1 | -0.05700 | 8251.3 | -0.06173 | 22072 | -0.01061 |
| 3.10 | 100 | 1 | 0.1 | 23.446 | 0.054904 | 27.276 | -0.03485 | 23.689 | 0.035002 | 34.303 | 0.022153 |
| 3.11 | 500 | 1 | 0.1 | 1.4572 | 0.50939 | 13.643 | -0.07449 | 1.8751 | 0.49122 | 1.1217 | 0.30445 |
| 3.12 | 1000 | 1 | 0.1 | 1.4703 | 0.5057 | 2.3948 | 0.18444 | 1.8341 | 0.51813 | 1.2313 | 0.29833 |
| 3.13 | N/A | 1 | 0.2 | 1.4624 | 0.50407 | 1.8357 | 0.39853 | 1.8592 | 0.50882 | 1.1185 | 0.30952 |
| 3.14 | 10 | 1 | 0.2 | 1.5236 | 0.44693 | 1.8793 | 0.35634 | 1.9177 | 0.45237 | 1.2416 | 0.25264 |
| 3.15 | 50 | 1 | 0.2 | 57.244 | 0.018183 | 226.85 | 0.11278 | 34.775 | 0.006200 | 1.1567 | 0.2955 |
| 3.16 | 100 | 1 | 0.2 | 5.6933 | 0.078266 | 40.956 | -0.09464 | 2.0733 | 0.35917 | 1.1122 | 0.31076 |
| 3.17 | 500 | 1 | 0.2 | 1.4709 | 0.50132 | 1.854 | 0.39289 | 1.8488 | 0.50762 | 1.1978 | 0.29884 |
| 3.18 | 1000 | 1 | 0.2 | 1.466 | 0.50068 | 1.8202 | 0.41398 | 1.8711 | 0.49546 | 1.1288 | 0.31182 |
| 3.19 | N/A | 1 | 0.5 | 1.4761 | 0.48994 | 1.7592 | 0.48031 | 1.9027 | 0.46907 | 1.086 | 0.32322 |
| 3.20 | 10 | 1 | 0.5 | 5.5077 | 0.13349 | 6.6294 | 0.007500 | 4.8026 | 0.17615 | 8.8416 | 0.043812 |
| 3.21 | 50 | 1 | 0.5 | 1.5926 | 0.40544 | 2.7758 | 0.26056 | 1.9185 | 0.44871 | 1.1081 | 0.31267 |
| 3.22 | 100 | 1 | 0.5 | 1.4741 | 0.49247 | 1.8104 | 0.42366 | 1.903 | 0.46872 | 1.1165 | 0.31496 |
| 3.23 | 500 | 1 | 0.5 | 1.475 | 0.49437 | 1.7488 | 0.48314 | 1.8832 | 0.48255 | 1.1354 | 0.32225 |
| 3.24 | 1000 | 1 | 0.5 | 1.4751 | 0.49437 | 1.7454 | 0.48453 | 1.8866 | 0.47824 | 1.1363 | 0.32332 |
| 3.25 | N/A | 1 | 1 | 1.4855 | 0.47996 | 1.8029 | 0.43221 | 1.9119 | 0.46043 | 1.0652 | 0.32013 |
| 3.26 | 10 | 1 | 1 | 1.5295 | 0.43397 | 1.9807 | 0.26704 | 1.9367 | 0.44033 | 1.0813 | 0.30007 |
| 3.27 | 50 | 1 | 1 | 1.478 | 0.48814 | 1.7467 | 0.4868 | 1.9057 | 0.46418 | 1.0967 | 0.31795 |
| 3.28 | 100 | 1 | 1 | 1.4776 | 0.48951 | 1.7549 | 0.49012 | 1.9069 | 0.46942 | 1.0656 | 0.32055 |
| 3.29 | 500 | 1 | 1 | 1.5083 | 0.46359 | 1.7854 | 0.44393 | 1.925 | 0.43962 | 1.1615 | 0.28219 |
| 3.30 | 1000 | 1 | 1 | 1.4911 | 0.47704 | 1.7725 | 0.46215 | 1.9202 | 0.44841 | 1.1088 | 0.31686 |
| 3.31 | N/A | 2 | 0.05 | 0.68291 | 0.91508 | 23.824 | 0.10521 | 30.963 | 0.13758 | 56.063 | 0.15375 |
| 3.32 | 10 | 2 | 0.05 | 1.9812 | 0.17717 | 2.1934 | 0.16289 | 2.2737 | 0.20539 | 1.8794 | 0.10484 |
| 3.33 | 50 | 2 | 0.05 | 1.4669 | 0.5245 | 1.7872 | 0.50017 | 1.9206 | 0.50449 | 0.92778 | 0.50377 |
| 3.34 | 100 | 2 | 0.05 | 1.3881 | 0.5793 | 1.6814 | 0.55343 | 1.8002 | 0.57051 | 1.0361 | 0.40097 |
| 3.35 | 500 | 2 | 0.05 | 44975 | -0.08355 | 1.02E+5 | -0.02578 | 1.01E+5 | -0.03874 | 2.28E+5 | -0.14485 |
| 3.36 | 1000 | 2 | 0.05 | 90737 | -0.01613 | 4.79E+5 | -0.06662 | 4.29E+5 | -0.00073 | 1.25E+6 | -0.13953 |
| 3.37 | N/A | 2 | 0.1 | 1.0461 | 0.78639 | 28.84 | -0.11028 | 67.327 | -0.00182 | 198.54 | -0.06785 |
| 3.38 | 10 | 2 | 0.1 | 2.0483 | 0.27129 | 2.2269 | 0.24699 | 2.2885 | 0.29497 | 2.4631 | 0.16046 |
| 3.39 | 50 | 2 | 0.1 | 1.4453 | 0.54419 | 1.7182 | 0.5099 | 1.8271 | 0.52494 | 1.6956 | 0.33319 |
| 3.40 | 100 | 2 | 0.1 | 7.5611 | 0.15865 | 7.3681 | 0.11665 | 8.0281 | 0.098464 | 20.953 | 0.19486 |
| 3.41 | 500 | 2 | 0.1 | 13687 | 0.066216 | 43237 | 0.10031 | 42464 | 0.046301 | 1.27E+5 | 0.096457 |
| 3.42 | 1000 | 2 | 0.1 | 822.24 | -0.00723 | 3973.9 | 0.14221 | 8010.3 | 0.017687 | 22878 | 0.06719 |
| 3.43 | N/A | 2 | 0.2 | 1.2172 | 0.69526 | 4.7984 | 0.081358 | 5.7935 | 0.11368 | 11.272 | 0.11168 |
| 3.44 | 10 | 2 | 0.2 | 1.4592 | 0.51712 | 1.738 | 0.50904 | 1.8794 | 0.50176 | 1.0286 | 0.44714 |
| 3.45 | 50 | 2 | 0.2 | 5.1312 | 0.17351 | 3.6797 | 0.20582 | 5.4722 | 0.14587 | 16.055 | 0.16065 |
| 3.46 | 100 | 2 | 0.2 | 488.77 | -0.01827 | 324.67 | -0.07333 | 648.13 | -0.01193 | 1890.8 | -0.06751 |
| 3.47 | 500 | 2 | 0.2 | 254.59 | 0.079415 | 1319.5 | 0.11414 | 602.26 | 0.098195 | 691.37 | 0.043839 |
| 3.48 | 1000 | 2 | 0.2 | 1.8424 | 0.43338 | 10.347 | 0.15733 | 8.6963 | 0.059998 | 18.947 | 0.011094 |
| 3.49 | N/A | 2 | 0.5 | 1.2844 | 0.65173 | 1.8471 | 0.44667 | 1.795 | 0.55137 | 1.1928 | 0.43714 |
| 3.50 | 10 | 2 | 0.5 | 1.3899 | 0.57426 | 1.6404 | 0.58043 | 1.7586 | 0.5865 | 1.0396 | 0.47735 |
| 3.51 | 50 | 2 | 0.5 | 14.095 | -3.43E-5 | 11.387 | 0.071736 | 13.039 | 2.82E-02 | 40.497 | -0.04769 |
| 3.52 | 100 | 2 | 0.5 | 25.823 | 0.032788 | 11.624 | 0.025394 | 40.974 | -0.00816 | 132.03 | 0.012755 |
| 3.53 | 500 | 2 | 0.5 | 1.3448 | 0.61275 | 2.8545 | 0.31081 | 3.8571 | 0.16699 | 8.7236 | 0.050396 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.54 | 1000 | 2 | 0.5 | 1.2973 | 0.64345 | 1.8756 | 0.43302 | 1.758 | 0.57451 | 1.4516 | 0.34345 |
| 3.55 | N/A | 2 | 1 | 1.3198 | 0.62659 | 1.624 | 0.58592 | 1.6746 | 0.64002 | 0.83522 | 0.59759 |
| 3.56 | 10 | 2 | 1 | 5.1738 | 0.30479 | 4.1349 | 0.34061 | 5.2963 | 0.32063 | 8.0296 | 0.27839 |
| 3.57 | 50 | 2 | 1 | 62.662 | 0.028224 | 39.48 | 0.001560 | 38.03 | 0.019386 | 241.13 | 0.003901 |
| 3.58 | 100 | 2 | 1 | 18.408 | 0.052204 | 10.759 | 0.11684 | 49.925 | 0.032502 | 88.993 | 0.029641 |
| 3.59 | 500 | 2 | 1 | 1.3242 | 0.62565 | 1.9707 | 0.36842 | 1.7947 | 0.54875 | 1.1391 | 0.45123 |
| 3.60 | 1000 | 2 | 1 | 1.323 | 0.62425 | 1.6441 | 0.57038 | 1.6767 | 0.63728 | 0.8501 | 0.58759 |
| 3.61 | N/A | 3 | 0.05 | 0.12854 | 0.99712 | 517.31 | 0.003627 | 485.55 | -0.03493 | 1049.6 | -0.22646 |
| 3.62 | 10 | 3 | 0.05 | 1.8676 | 0.10405 | 2.1014 | 0.092443 | 2.2336 | 0.10031 | 1.6081 | 0.064653 |
| 3.63 | 50 | 3 | 0.05 | 1.6909 | 0.27231 | 1.9529 | 0.29687 | 2.106 | 0.26113 | 1.3546 | 0.12459 |
| 3.64 | 100 | 3 | 0.05 | 1.5919 | 0.41033 | 1.9625 | 0.30596 | 2.0924 | 0.3343 | 1.0373 | 0.1239 |
| 3.65 | 500 | 3 | 0.05 | 1.3493 | 0.62355 | 1.8363 | 0.45035 | 2.0138 | 0.3907 | 1.068 | 0.1406 |
| 3.66 | 1000 | 3 | 0.05 | 184.04 | -0.04822 | 343.95 | 0.001796 | 334.84 | -0.02989 | 626.19 | -0.20285 |
| 3.67 | N/A | 3 | 0.1 | 0.32176 | 0.98179 | 379.28 | -0.01820 | 461.63 | 0.14012 | 1353.2 | 0.24409 |
| 3.68 | 10 | 3 | 0.1 | 1.7678 | 0.20563 | 2.024 | 0.18682 | 2.1644 | 0.18189 | 1.4933 | 0.12669 |
| 3.69 | 50 | 3 | 0.1 | 1.5549 | 0.42352 | 1.8312 | 0.44774 | 1.9656 | 0.45373 | 1.0527 | 0.23954 |
| 3.70 | 100 | 3 | 0.1 | 2.9352 | 0.20883 | 3.1413 | 0.16741 | 3.1451 | 0.204 | 5.4629 | 0.21995 |
| 3.71 | 500 | 3 | 0.1 | 23.052 | 0.009405 | 34.28 | 0.015919 | 34.981 | -0.01104 | 86.398 | -0.23554 |
| 3.72 | 1000 | 3 | 0.1 | 90249 | -0.0496 | 2.51E+5 | -0.00891 | 2.45E+5 | -0.04426 | 7.35E+0 | -0.23168 |
| 3.73 | N/A | 3 | 0.2 | 0.67264 | 0.91776 | 39.995 | -0.02386 | 39.643 | 0.093122 | 58.709 | 0.089639 |
| 3.74 | 10 | 3 | 0.2 | 1.6175 | 0.37486 | 1.8886 | 0.36853 | 2.0414 | 0.34994 | 1.3263 | 0.28012 |
| 3.75 | 50 | 3 | 0.2 | 1.3544 | 0.60775 | 1.6086 | 0.61277 | 1.726 | 0.62352 | 1.0892 | 0.43911 |
| 3.76 | 100 | 3 | 0.2 | 4.7289 | 0.17384 | 4.1002 | 0.12017 | 4.2529 | 0.19258 | 13.381 | 0.23616 |
| 3.77 | 500 | 3 | 0.2 | 2871.5 | 0.063831 | 3294 | 0.043893 | 5371.1 | 0.041959 | 14057 | 0.18537 |
| 3.78 | 1000 | 3 | 0.2 | 1792.8 | -0.00164 | 4838 | 0.10926 | 6544.6 | 0.072167 | 17532 | 0.15231 |
| 3.79 | N/A | 3 | 0.5 | 1.0557 | 0.782 | 3.3483 | 0.24047 | 8.2722 | 0.12948 | 6.3645 | 0.039108 |
| 3.80 | 10 | 3 | 0.5 | 1.5482 | 0.40732 | 1.8266 | 0.41795 | 1.9867 | 0.39605 | 0.96913 | 0.3749 |
| 3.81 | 50 | 3 | 0.5 | 1.8224 | 0.40288 | 1.7112 | 0.52907 | 2.0432 | 0.44076 | 3.6012 | -0.13006 |
| 3.82 | 100 | 3 | 0.5 | 13.007 | 0.063318 | 7.7719 | 0.03208 | 15.507 | 0.063717 | 41.726 | 0.29371 |
| 3.83 | 500 | 3 | 0.5 | 91.364 | 0.021626 | 119.14 | -0.04892 | 519.66 | -0.02184 | 533.28 | -0.11112 |
| 3.84 | 1000 | 3 | 0.5 | 2.5571 | 0.39592 | 10.625 | 0.004565 | 3.1745 | 0.31405 | 11.06 | 0.19657 |
| 3.85 | N/A | 3 | 1 | 1.204 | 0.70332 | 1.7566 | 0.50042 | 1.8722 | 0.5225 | 1.6657 | 0.14088 |
| 3.86 | 10 | 3 | 1 | 1.4869 | 0.48108 | 1.7533 | 0.49842 | 1.9029 | 0.48609 | 0.99289 | 0.36403 |
| 3.87 | 50 | 3 | 1 | 17.729 | 0.10587 | 11.687 | 0.15519 | 22.251 | 0.06559 | 43.527 | 0.21858 |
| 3.88 | 100 | 3 | 1 | 135.2 | 0.007150 | 82.458 | 0.009227 | 137.6 | -0.01120 | 410.98 | 0.067896 |
| 3.89 | 500 | 3 | 1 | 3.9427 | 0.22414 | 6.6837 | 0.052283 | 20.264 | 0.047119 | 18.359 | -0.01527 |
| 3.90 | 1000 | 3 | 1 | 1.2341 | 0.69089 | 2.3472 | 0.31427 | 1.6716 | 0.62922 | 3.3966 | 0.050735 |
| 3.91 | N/A | 4 | 0.05 | 0.01695 | 0.99995 | 2613.4 | -0.01850 | 2598.6 | -0.04185 | 4055.3 | -0.14303 |
| 3.92 | 10 | 4 | 0.05 | 2.1208 | -0.01903 | 2.302 | 0.015542 | 2.4246 | -0.04910 | 1.9995 | 0.026362 |
| 3.93 | 50 | 4 | 0.05 | 1.7329 | 0.14216 | 1.9994 | 0.17034 | 2.1485 | 0.13932 | 1.3546 | 0.046486 |
| 3.94 | 100 | 4 | 0.05 | 1.7649 | 0.17821 | 2.0164 | 0.20674 | 2.1695 | 0.1531 | 1.4772 | 0.074994 |
| 3.95 | 500 | 4 | 0.05 | 1.5855 | 0.43019 | 2.026 | 0.24905 | 2.1352 | 0.31716 | 1.0865 | 0.037839 |
| 3.96 | 1000 | 4 | 0.05 | 3.669 | 0.22149 | 5.0631 | 0.060929 | 5.1671 | 0.062803 | 7.0724 | 0.12172 |
| 3.97 | N/A | 4 | 0.1 | 0.050231 | 0.99956 | 465.19 | -0.02462 | 465.61 | -0.08911 | 1231.7 | -0.20434 |
| 3.98 | 10 | 4 | 0.1 | 1.7779 | 0.14703 | 2.0327 | 0.15415 | 2.1768 | 0.13303 | 1.4573 | 0.081286 |
| 3.99 | 50 | 4 | 0.1 | 2.532 | 0.080458 | 2.6673 | 0.081302 | 2.7429 | 0.099492 | 3.0916 | 0.10057 |
| 3.100 | 100 | 4 | 0.1 | 1.5667 | 0.40052 | 1.8978 | 0.35001 | 2.0602 | 0.32508 | 1.0199 | 0.15748 |
| 3.101 | 500 | 4 | 0.1 | 3.0713 | 0.30186 | 3.9771 | 0.12888 | 3.9364 | 0.18651 | 7.3736 | 0.17922 |
| 3.102 | 1000 | 4 | 0.1 | 744.21 | -0.02342 | 1307.1 | -0.00657 | 1315.4 | 0.078004 | 2916.9 | 0.17144 |
| 3.103 | N/A | 4 | 0.2 | 0.17848 | 0.99443 | 12421 | 0.10231 | 22702 | 0.09818 | 1.01E+5 | 0.15542 |
| 3.104 | 10 | 4 | 0.2 | 1.6843 | 0.29129 | 1.9507 | 0.29382 | 2.0983 | 0.28587 | 1.3976 | 0.17909 |
| 3.105 | 50 | 4 | 0.2 | 1.5128 | 0.46329 | 1.7441 | 0.53379 | 1.9306 | 0.48106 | 1.0615 | 0.13942 |
| 3.106 | 100 | 4 | 0.2 | 1.4056 | 0.57679 | 1.6502 | 0.57905 | 1.7881 | 0.57228 | 1.4911 | 0.29735 |
| 3.107 | 500 | 4 | 0.2 | 72.975 | -0.09073 | 78.802 | -0.04600 | 94.417 | -0.11978 | 358.32 | -0.19858 |
| 3.108 | 1000 | 4 | 0.2 | 245.88 | -0.04684 | 421.34 | -0.02159 | 572.35 | -0.09634 | 2404.8 | -0.18881 |
| 3.109 | N/A | 4 | 0.5 | 0.56354 | 0.94305 | 10.456 | 0.23007 | 29.687 | 0.065706 | 68.453 | -0.09749 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.110 | 10 | 4 | 0.5 | 1.5806 | 0.36282 | 1.8535 | 0.38509 | 2.0125 | 0.36805 | 1.0092 | 0.25675 |
| 3.111 | 50 | 4 | 0.5 | 2.7619 | 0.33518 | 2.2967 | 0.38458 | 2.6643 | 0.34407 | 7.69 | 0.19772 |
| 3.112 | 100 | 4 | 0.5 | 1.6389 | 0.52105 | 1.6489 | 0.57653 | 1.9845 | 0.47911 | 3.549 | 0.18581 |
| 3.113 | 500 | 4 | 0.5 | 952.59 | 0.039651 | 537.64 | 0.022508 | 2071.7 | 0.067858 | 7701 | 0.12608 |
| 3.114 | 1000 | 4 | 0.5 | 249.62 | -0.02097 | 254.74 | -0.00278 | 987.08 | -0.04938 | 3473.6 | -0.13804 |
| 3.115 | N/A | 4 | 1 | 0.98347 | 0.81422 | 2.2027 | 0.39268 | 2.6033 | 0.35298 | 9.1772 | 0.026978 |
| 3.116 | 10 | 4 | 1 | 1.5965 | 0.37817 | 1.8413 | 0.40021 | 1.9804 | 0.40452 | 1.3947 | 0.13928 |
| 3.117 | 50 | 4 | 1 | 2.6914 | 0.27331 | 2.1966 | 0.35868 | 2.7606 | 0.29525 | 6.6936 | -0.03119 |
| 3.118 | 100 | 4 | 1 | 9.7207 | 0.051273 | 6.0515 | 0.13794 | 10.678 | 0.093259 | 33.465 | -0.18342 |
| 3.119 | 500 | 4 | 1 | 38.692 | -0.02062 | 25.63 | -0.02857 | 89.629 | -0.03527 | 223.49 | -0.07725 |
| 3.120 | 1000 | 4 | 1 | 8.4897 | 0.14253 | 15.91 | 0.014239 | 20.37 | 0.018338 | 177.45 | 0.021697 |
| 3.121 | N/A | 5 | 0.05 | 0.003876 | 1 | 4117.4 | -0.03256 | 4153.6 | 0.001400 | 5237.8 | -0.07718 |
| 3.122 | 10 | 5 | 0.05 | 1.8137 | 0.063123 | 2.1427 | 0.03254 | 2.2916 | 0.06877 | 1.0558 | 0.029397 |
| 3.123 | 50 | 5 | 0.05 | 1.6919 | 0.097873 | 2.0035 | 0.066635 | 2.1545 | 0.10073 | 1.091 | 0.029681 |
| 3.124 | 100 | 5 | 0.05 | 1.9778 | 0.090326 | 2.196 | 0.07307 | 2.3205 | 0.069013 | 1.8188 | 0.036045 |
| 3.125 | 500 | 5 | 0.05 | 1.6168 | 0.31733 | 1.9838 | 0.16789 | 2.1274 | 0.18909 | 1.0686 | 0.054805 |
| 3.126 | 1000 | 5 | 0.05 | 1.4309 | 0.57444 | 2.0736 | 0.14849 | 2.2018 | 0.19691 | 1.0582 | 0.061468 |
| 3.127 | N/A | 5 | 0.1 | 0.003889 | 1 | 15488 | 0.079984 | 15952 | 0.053525 | 30872 | 0.1298 |
| 3.128 | 10 | 5 | 0.1 | 1.7748 | 0.086648 | 2.0353 | 0.089625 | 2.1809 | 0.07141 | 1.4226 | 0.048732 |
| 3.129 | 50 | 5 | 0.1 | 1.7444 | 0.18145 | 2.0855 | 0.11245 | 2.2152 | 0.18153 | 1.0569 | 0.04107 |
| 3.130 | 100 | 5 | 0.1 | 1.7075 | 0.21542 | 2.0366 | 0.18093 | 2.1643 | 0.22993 | 1.0539 | 0.049389 |
| 3.131 | 500 | 5 | 0.1 | 1.5491 | 0.43588 | 1.9551 | 0.28878 | 2.0766 | 0.32855 | 1.0498 | 0.089021 |
| 3.132 | 1000 | 5 | 0.1 | 5.6798 | 0.15719 | 8.9139 | 0.02156 | 9.1615 | 0.012735 | 14.22 | -0.10517 |
| 3.133 | N/A | 5 | 0.2 | 0.004453 | 1 | 2679.1 | -0.12188 | 3446.9 | -0.09964 | 14235 | -0.11391 |
| 3.134 | 10 | 5 | 0.2 | 1.8252 | 0.18098 | 2.0487 | 0.21263 | 2.2001 | 0.17647 | 1.7244 | 0.11854 |
| 3.135 | 50 | 5 | 0.2 | 1.6453 | 0.29199 | 1.9618 | 0.26491 | 2.1064 | 0.27134 | 1.0611 | 0.17611 |
| 3.136 | 100 | 5 | 0.2 | 1.5222 | 0.4764 | 1.8089 | 0.45423 | 1.94 | 0.46582 | 1.3893 | 0.1839 |
| 3.137 | 500 | 5 | 0.2 | 9.3209 | 0.17057 | 10.932 | 0.17776 | 11.923 | 0.15977 | 36.624 | 0.15078 |
| 3.138 | 1000 | 5 | 0.2 | 160.19 | 0.10669 | 262.09 | 0.10518 | 305.97 | 0.09763 | 1080.6 | 0.14067 |
| 3.139 | N/A | 5 | 0.5 | 0.43255 | 0.96686 | 10.289 | 0.16043 | 17.512 | 0.12143 | 100.11 | 0.069558 |
| 3.140 | 10 | 5 | 0.5 | 1.6303 | 0.28374 | 1.9171 | 0.30338 | 2.0783 | 0.28377 | 1.0324 | 0.17284 |
| 3.141 | 50 | 5 | 0.5 | 1.446 | 0.52346 | 1.6723 | 0.56807 | 1.7943 | 0.57443 | 1.1617 | 0.31636 |
| 3.142 | 100 | 5 | 0.5 | 2.4818 | 0.36933 | 2.1492 | 0.44268 | 2.5282 | 0.36008 | 7.3917 | 0.10093 |
| 3.143 | 500 | 5 | 0.5 | 91.909 | 0.11925 | 78.769 | 0.13287 | 129.84 | 0.046001 | 645.22 | 0.009160 |
| 3.144 | 1000 | 5 | 0.5 | 83.795 | 0.004852 | 90.524 | 0.13662 | 136.28 | 0.01255 | 436.11 | 0.12625 |
| 3.145 | N/A | 5 | 1 | 0.70932 | 0.90815 | 2.3344 | 0.50357 | 3.0276 | 0.32826 | 19.714 | 0.069566 |
| 3.146 | 10 | 5 | 1 | 2.0049 | 0.34193 | 1.9386 | 0.4038 | 2.1643 | 0.36221 | 3.0856 | 0.33032 |
| 3.147 | 50 | 5 | 1 | 2.2167 | 0.22873 | 2.15 | 0.26226 | 2.3486 | 0.27739 | 5.0003 | 0.003617 |
| 3.148 | 100 | 5 | 1 | 18.053 | 0.12727 | 13.059 | 0.15388 | 17.577 | 0.034168 | 53.892 | -0.00642 |
| 3.149 | 500 | 5 | 1 | 915.12 | 0.11069 | 821.51 | 0.18131 | 1283.5 | 0.005026 | 5794.6 | 0.008436 |
| 3.150 | 1000 | 5 | 1 | 42.879 | -0.03867 | 37.792 | -0.04084 | 59.873 | -0.02891 | 326.27 | 0.022812 |

Table B.4: Performance of the radial basis network for the spatial input pattern.

| Net ID | No of Cluster | Input Neuron | Width | RMSE Training | CC Training | RMSE Validation | CC Validation | RMSE Test set (i) | CC Test set (i) | RMSE Test set (ii) | CC Test set(ii) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 4.1 | N/A | 9 | 0.05 | 0.049231 | 0.99958 | 5802.1 | 0.041109 | 5798.4 | 0.028 | 7982 | -0.06452 |
| 4.2 | 10 | 9 | 0.05 | 1.7807 | 0.16879 | 2.0283 | 0.21667 | 2.175 | 0.19941 | 1.4496 | 0.10179 |
| 4.3 | 50 | 9 | 0.05 | 1.6753 | 0.21425 | 1.9553 | 0.234 | 2.1213 | 0.16306 | 1.2224 | 0.10571 |
| 4.4 | 100 | 9 | 0.05 | 2.1888 | 0.14209 | 2.3751 | 0.11803 | 2.4933 | 0.1029 | 2.3168 | 0.095149 |
| 4.5 | 500 | 9 | 0.05 | 1.547 | 0.46082 | 1.966 | 0.26631 | 2.0905 | 0.33154 | 1.0127 | 0.16903 |
| 4.6 | 1000 | 9 | 0.05 | 1.2605 | 0.68994 | 1.8881 | 0.38585 | 2.0336 | 0.3804 | 1.0269 | 0.16105 |
| 4.7 | N/A | 9 | 0.1 | 0.049723 | 0.99957 | 611.58 | -0.15527 | 590.65 | -0.0927 | 1043.5 | 0.050041 |
| 4.8 | 10 | 9 | 0.1 | 1.6031 | 0.44253 | 1.9186 | 0.38828 | 2.0538 | 0.45426 | 1.1212 | 0.15821 |
| 4.9 | 50 | 9 | 0.1 | 1.5553 | 0.47381 | 1.9039 | 0.41238 | 2.0576 | 0.42593 | 1.0605 | 0.15676 |
| 4.10 | 100 | 9 | 0.1 | 1.55 | 0.46238 | 1.8635 | 0.44675 | 2.0403 | 0.41362 | 1.0463 | 0.16584 |
| 4.11 | 500 | 9 | 0.1 | 9.5686 | 0.17368 | 13.034 | 0.18866 | 12.858 | 0.11871 | 20.503 | -0.02897 |
| 4.12 | 1000 | 9 | 0.1 | 1.4578 | 0.64839 | 2.5572 | 0.3879 | 2.7207 | 0.30056 | 3.2184 | 0.012974 |
| 4.13 | N/A | 9 | 0.2 | 0.049697 | 0.99957 | 11373 | -0.13237 | 8645.3 | -0.0763 | 21439 | 0.06004 |
| 4.14 | 10 | 9 | 0.2 | 1.5432 | 0.46394 | 1.8375 | 0.42833 | 1.9636 | 0.4618 | 1.3218 | 0.18442 |
| 4.15 | 50 | 9 | 0.2 | 1.4194 | 0.55946 | 1.7161 | 0.54747 | 1.8498 | 0.56333 | 1.0087 | 0.27987 |
| 4.16 | 100 | 9 | 0.2 | 4.358 | 0.044316 | 5.3109 | -0.03771 | 4.9756 | 0.01292 | 8.9644 | 0.063728 |
| 4.17 | 500 | 9 | 0.2 | 15.945 | -0.03664 | 30.516 | -0.13929 | 25.723 | -0.0533 | 55.662 | 0.039351 |
| 4.18 | 1000 | 9 | 0.2 | 28.75 | -0.03944 | 96.746 | -0.13479 | 76.569 | -0.0666 | 177.99 | 0.055318 |
| 4.19 | N/A | 9 | 0.5 | 0.59834 | 0.93548 | 33.572 | 0.038928 | 27.537 | 0.04696 | 82.661 | -0.00332 |
| 4.20 | 10 | 9 | 0.5 | 1.962 | 0.25478 | 2.3538 | 0.14286 | 2.2941 | 0.24281 | 2.8462 | 0.16897 |
| 4.21 | 50 | 9 | 0.5 | 3.0063 | 0.14632 | 3.8034 | 0.040579 | 3.3736 | 0.10771 | 6.5984 | 0.095444 |
| 4.22 | 100 | 9 | 0.5 | 2.1077 | 0.33025 | 2.6713 | 0.20399 | 2.6662 | 0.20147 | 4.0016 | -0.00438 |
| 4.23 | 500 | 9 | 0.5 | 63.072 | -0.04128 | 138.14 | -0.08421 | 108.57 | -0.0613 | 322.75 | 0.015328 |
| 4.24 | 1000 | 9 | 0.5 | 380.11 | 0.047515 | 1651.4 | 0.089904 | 1082.7 | 0.04764 | 4081.5 | -0.03693 |
| 4.25 | N/A | 9 | 1 | 0.82117 | 0.87453 | 32.195 | 0.01901 | 20.456 | 0.03747 | 82.587 | 0.028076 |
| 4.26 | 10 | 9 | 1 | 3.3596 | 0.11355 | 4.3284 | 0.01335 | 3.6751 | 0.1088 | 6.8009 | 0.17889 |
| 4.27 | 50 | 9 | 1 | 36.127 | 0.1084 | 46.825 | 0.15185 | 37.74 | 0.091 | 96.422 | 0.001884 |
| 4.28 | 100 | 9 | 1 | 17.915 | 0.11654 | 23.364 | 0.1618 | 20.892 | 0.09666 | 53.452 | 0.010023 |
| 4.29 | 500 | 9 | 1 | 24.868 | 0.082403 | 62.692 | 0.042187 | 42.83 | -0.0049 | 113.42 | 0.11829 |
| 4.30 | 1000 | 9 | 1 | 96.513 | -0.01899 | 564.78 | -0.06818 | 334.89 | -0.003 | 1638.4 | 0.028556 |

Table B.5: Performance of the generalized regression network for the temporal input pattern.

| Net ID | No of Cluster | Input Neuron | Width | RMSE Training | CC Training | RMSE Validation | CC Validation | RMSE Test set (i) | CC Test set (i) | RMSE Test set (ii) | CC Test set(ii) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5.1 | N/A | 1 | 0.05 | 1.4724 | 0.49673 | 1.7655 | 0.48446 | 1.895 | 0.48818 | 1.0518 | 0.32271 |
| 5.2 | 10 | 1 | 0.05 | 1.4995 | 0.99986 | 1.7687 | 0.4803 | 1.9446 | 0.43071 | 1.033 | 0.32881 |
| 5.3 | 50 | 1 | 0.05 | 1.4696 | 0.89338 | 1.7578 | 0.48783 | 1.8928 | 0.48738 | 1.063 | 0.32269 |
| 5.4 | 100 | 1 | 0.05 | 1.4715 | 0.94858 | 1.762 | 0.48553 | 1.8891 | 0.49076 | 1.0607 | 0.31995 |
| 5.5 | 500 | 1 | 0.05 | 1.4769 | 0.63079 | 1.7686 | 0.48683 | 1.9063 | 0.48411 | 1.0412 | 0.32147 |
| 5.6 | 1000 | 1 | 0.05 | 1.4678 | 0.57609 | 1.7529 | 0.48622 | 1.8888 | 0.48357 | 1.0858 | 0.32202 |
| 5.7 | N/A | 1 | 0.1 | 1.5085 | 0.46403 | 1.7958 | 0.47737 | 1.9492 | 0.44803 | 1.0001 | 0.32793 |
| 5.8 | 10 | 1 | 0.1 | 1.5049 | 0.99119 | 1.7722 | 0.47952 | 1.9468 | 0.43015 | 1.022 | 0.3258 |
| 5.9 | 50 | 1 | 0.1 | 1.4908 | 0.8679 | 1.7732 | 0.48329 | 1.9345 | 0.45072 | 1.0241 | 0.32861 |
| 5.10 | 100 | 1 | 0.1 | 1.4925 | 0.8503 | 1.7822 | 0.48344 | 1.9392 | 0.45574 | 1.0124 | 0.32926 |
| 5.11 | 500 | 1 | 0.1 | 1.5034 | 0.62079 | 1.7921 | 0.47627 | 1.9418 | 0.45437 | 1.0051 | 0.32699 |
| 5.12 | 1000 | 1 | 0.1 | 1.4906 | 0.5214 | 1.7663 | 0.4839 | 1.9237 | 0.45439 | 1.0352 | 0.32654 |
| 5.13 | N/A | 1 | 0.2 | 1.5572 | 0.41733 | 1.8474 | 0.43608 | 2.0119 | 0.39472 | 0.98348 | 0.31876 |
| 5.14 | 10 | 1 | 0.2 | 1.5283 | 0.97253 | 1.7993 | 0.4567 | 1.9706 | 0.41007 | 1.0086 | 0.32344 |
| 5.15 | 50 | 1 | 0.2 | 1.5449 | 0.68439 | 1.8216 | 0.43218 | 1.9887 | 0.39188 | 1.0073 | 0.31798 |
| 5.16 | 100 | 1 | 0.2 | 1.5411 | 0.79821 | 1.824 | 0.44737 | 1.9901 | 0.40485 | 0.99214 | 0.32167 |
| 5.17 | 500 | 1 | 0.2 | 1.5451 | 0.54888 | 1.8329 | 0.44976 | 1.9968 | 0.40824 | 0.98671 | 0.32172 |
| 5.18 | 1000 | 1 | 0.2 | 1.5414 | 0.46728 | 1.8237 | 0.4502 | 1.9899 | 0.40647 | 0.99237 | 0.32216 |
| 5.19 | N/A | 1 | 0.5 | 1.6224 | 0.35493 | 1.9229 | 0.36779 | 2.0798 | 0.33826 | 1.0103 | 0.2952 |
| 5.20 | 10 | 1 | 0.5 | 1.5976 | 0.83923 | 1.8917 | 0.39878 | 2.0507 | 0.36411 | 1.0057 | 0.30696 |
| 5.21 | 50 | 1 | 0.5 | 1.6136 | 0.75465 | 1.9093 | 0.38632 | 2.0666 | 0.35381 | 1.0202 | 0.30248 |
| 5.22 | 100 | 1 | 0.5 | 1.613 | 0.62492 | 1.9021 | 0.38178 | 2.0589 | 0.34993 | 1.0376 | 0.30085 |
| 5.23 | 500 | 1 | 0.5 | 1.6046 | 0.46762 | 1.8961 | 0.38755 | 2.054 | 0.35485 | 1.0195 | 0.30287 |
| 5.24 | 1000 | 1 | 0.5 | 1.6151 | 0.41721 | 1.9123 | 0.38694 | 2.0696 | 0.35436 | 1.017 | 0.30267 |
| 5.25 | N/A | 1 | 1 | 1.6654 | 0.34957 | 1.968 | 0.36238 | 2.1218 | 0.33342 | 1.0694 | 0.29304 |
| 5.26 | 10 | 1 | 1 | 1.6567 | 0.79811 | 1.9592 | 0.37402 | 2.1135 | 0.34341 | 1.0576 | 0.29768 |
| 5.27 | 50 | 1 | 1 | 1.6639 | 0.7184 | 1.9649 | 0.37114 | 2.1185 | 0.34094 | 1.0766 | 0.29652 |
| 5.28 | 100 | 1 | 1 | 1.6691 | 0.58072 | 1.964 | 0.36529 | 2.1166 | 0.33592 | 1.1139 | 0.29422 |
| 5.29 | 500 | 1 | 1 | 1.6641 | 0.45359 | 1.9588 | 0.36932 | 2.1117 | 0.33938 | 1.1061 | 0.29581 |
| 5.30 | 1000 | 1 | 1 | 1.6665 | 0.39918 | 1.9611 | 0.36858 | 2.1139 | 0.33875 | 1.1109 | 0.29551 |
| 5.31 | N/A | 2 | 0.05 | 1.2566 | 0.6751 | 1.5988 | 0.61739 | 1.7112 | 0.62309 | 0.92262 | 0.54971 |
| 5.32 | 10 | 2 | 0.05 | 1.4677 | 1 | 1.7522 | 0.49083 | 1.8982 | 0.47736 | 0.87629 | 0.53642 |
| 5.33 | 50 | 2 | 0.05 | 1.3262 | 0.99968 | 1.6038 | 0.61523 | 1.7193 | 0.61882 | 0.81898 | 0.61306 |
| 5.34 | 100 | 2 | 0.05 | 1.3049 | 0.99375 | 1.5976 | 0.61511 | 1.7345 | 0.60351 | 0.81657 | 0.61242 |
| 5.35 | 500 | 2 | 0.05 | 1.2697 | 0.89262 | 1.5962 | 0.61624 | 1.6983 | 0.62422 | 0.97361 | 0.53423 |
| 5.36 | 1000 | 2 | 0.05 | 1.2509 | 0.81818 | 1.5974 | 0.60868 | 1.7227 | 0.60797 | 0.94179 | 0.55897 |
| 5.37 | N/A | 2 | 0.1 | 1.3512 | 0.61746 | 1.6474 | 0.61743 | 1.7573 | 0.62881 | 0.82772 | 0.59262 |
| 5.38 | 10 | 2 | 0.1 | 1.4554 | 0.99718 | 1.7429 | 0.51015 | 1.8887 | 0.49878 | 0.87341 | 0.5465 |
| 5.39 | 50 | 2 | 0.1 | 1.3486 | 0.99016 | 1.6213 | 0.61274 | 1.7386 | 0.617 | 0.80091 | 0.62761 |
| 5.40 | 100 | 2 | 0.1 | 1.3423 | 0.97398 | 1.6242 | 0.61424 | 1.7303 | 0.63013 | 0.83611 | 0.58946 |
| 5.41 | 500 | 2 | 0.1 | 1.3521 | 0.83909 | 1.6528 | 0.61068 | 1.7544 | 0.62945 | 0.83825 | 0.58195 |
| 5.42 | 1000 | 2 | 0.1 | 1.3419 | 0.70394 | 1.6348 | 0.61731 | 1.748 | 0.62321 | 0.82335 | 0.60186 |
| 5.43 | N/A | 2 | 0.2 | 1.4552 | 0.55086 | 1.7565 | 0.54827 | 1.8926 | 0.5521 | 0.85119 | 0.61386 |
| 5.44 | 10 | 2 | 0.2 | 1.4587 | 0.95195 | 1.7421 | 0.5224 | 1.8873 | 0.51348 | 0.86748 | 0.5587 |
| 5.45 | 50 | 2 | 0.2 | 1.403 | 0.94876 | 1.6877 | 0.57346 | 1.8167 | 0.57705 | 0.81502 | 0.63493 |
| 5.46 | 100 | 2 | 0.2 | 1.4276 | 0.9086 | 1.7182 | 0.5547 | 1.8487 | 0.56083 | 0.8239 | 0.62086 |
| 5.47 | 500 | 2 | 0.2 | 1.4389 | 0.7747 | 1.744 | 0.54628 | 1.869 | 0.55861 | 0.83446 | 0.58774 |
| 5.48 | 1000 | 2 | 0.2 | 1.4422 | 0.6609 | 1.743 | 0.5528 | 1.8752 | 0.55963 | 0.83887 | 0.61453 |
| 5.49 | N/A | 2 | 0.5 | 1.5691 | 0.43736 | 1.8729 | 0.43587 | 2.0223 | 0.43208 | 0.92282 | 0.48965 |
| 5.50 | 10 | 2 | 0.5 | 1.5876 | 0.85901 | 1.8974 | 0.43147 | 2.0475 | 0.42938 | 0.92845 | 0.49094 |
| 5.51 | 50 | 2 | 0.5 | 1.5508 | 0.79616 | 1.8305 | 0.4486 | 1.974 | 0.44505 | 0.9259 | 0.50294 |
| 5.52 | 100 | 2 | 0.5 | 1.5482 | 0.74915 | 1.8345 | 0.44848 | 1.978 | 0.44765 | 0.92179 | 0.50611 |
| 5.53 | 500 | 2 | 0.5 | 1.5544 | 0.64174 | 1.8527 | 0.44576 | 1.9979 | 0.44769 | 0.91757 | 0.50624 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5.54 | 1000 | 2 | 0.5 | 1.5614 | 0.51445 | 1.8612 | 0.44173 | 2.0092 | 0.43888 | 0.92271 | 0.49784 |
| 5.55 | N/A | 2 | 1 | 1.6368 | 0.40605 | 1.942 | 0.40566 | 2.0934 | 0.40138 | 1.0194 | 0.45321 |
| 5.56 | 10 | 2 | 1 | 1.6527 | 0.79654 | 1.9659 | 0.41194 | 2.1184 | 0.4071 | 0.99741 | 0.45786 |
| 5.57 | 50 | 2 | 1 | 1.6378 | 0.73058 | 1.9263 | 0.4155 | 2.074 | 0.41198 | 1.0843 | 0.46206 |
| 5.58 | 100 | 2 | 1 | 1.6511 | 0.69138 | 1.9356 | 0.41067 | 2.0824 | 0.40789 | 1.1184 | 0.45824 |
| 5.59 | 500 | 2 | 1 | 1.6375 | 0.56758 | 1.9345 | 0.41038 | 2.0841 | 0.40743 | 1.0574 | 0.45826 |
| 5.60 | 1000 | 2 | 1 | 1.639 | 0.49394 | 1.9365 | 0.41027 | 2.0863 | 0.40684 | 1.0586 | 0.45776 |
| 5.61 | N/A | 3 | 0.05 | 0.84372 | 0.87382 | 1.5085 | 0.67627 | 1.7438 | 0.58638 | 0.93886 | 0.47057 |
| 5.62 | 10 | 3 | 0.05 | 1.543 | 1 | 1.8317 | 0.40235 | 1.9907 | 0.38158 | 0.99678 | 0.27546 |
| 5.63 | 50 | 3 | 0.05 | 1.3535 | 0.99999 | 1.5955 | 0.61975 | 1.7615 | 0.58258 | 0.96662 | 0.3905 |
| 5.64 | 100 | 3 | 0.05 | 1.3003 | 0.99979 | 1.5532 | 0.64577 | 1.671 | 0.65442 | 0.92336 | 0.47862 |
| 5.65 | 500 | 3 | 0.05 | 1.1687 | 0.9906 | 1.5646 | 0.63568 | 1.7166 | 0.61494 | 0.95261 | 0.45925 |
| 5.66 | 1000 | 3 | 0.05 | 0.98503 | 0.96776 | 1.4866 | 0.68194 | 1.731 | 0.59372 | 0.92739 | 0.47651 |
| 5.67 | N/A | 3 | 0.1 | 1.1953 | 0.73171 | 1.5977 | 0.66336 | 1.737 | 0.62065 | 0.88216 | 0.52163 |
| 5.68 | 10 | 3 | 0.1 | 1.5488 | 1 | 1.8396 | 0.39616 | 1.9961 | 0.37957 | 0.95928 | 0.37394 |
| 5.69 | 50 | 3 | 0.1 | 1.3485 | 0.99644 | 1.6013 | 0.62651 | 1.7463 | 0.60347 | 0.90116 | 0.48646 |
| 5.70 | 100 | 3 | 0.1 | 1.3252 | 0.99173 | 1.5901 | 0.6353 | 1.7173 | 0.63101 | 0.90588 | 0.48544 |
| 5.71 | 500 | 3 | 0.1 | 1.2136 | 0.91007 | 1.5725 | 0.66888 | 1.7007 | 0.64029 | 0.88422 | 0.51995 |
| 5.72 | 1000 | 3 | 0.1 | 1.2041 | 0.84734 | 1.6199 | 0.63525 | 1.7185 | 0.62862 | 0.9481 | 0.44207 |
| 5.73 | N/A | 3 | 0.2 | 1.4275 | 0.58003 | 1.7541 | 0.55564 | 1.8773 | 0.56542 | 0.88314 | 0.56049 |
| 5.74 | 10 | 3 | 0.2 | 1.5325 | 0.98677 | 1.8249 | 0.43569 | 1.9807 | 0.41875 | 0.94779 | 0.39941 |
| 5.75 | 50 | 3 | 0.2 | 1.3988 | 0.96467 | 1.6748 | 0.5872 | 1.8184 | 0.57595 | 0.87605 | 0.53249 |
| 5.76 | 100 | 3 | 0.2 | 1.3833 | 0.91993 | 1.6764 | 0.5897 | 1.8094 | 0.58491 | 0.859 | 0.56129 |
| 5.77 | 500 | 3 | 0.2 | 1.3927 | 0.78742 | 1.7175 | 0.57826 | 1.838 | 0.58196 | 0.87352 | 0.53947 |
| 5.78 | 1000 | 3 | 0.2 | 1.4006 | 0.67087 | 1.7237 | 0.56891 | 1.8467 | 0.56885 | 0.87795 | 0.52458 |
| 5.79 | N/A | 3 | 0.5 | 1.5657 | 0.43185 | 1.8722 | 0.42295 | 2.0202 | 0.42308 | 0.94477 | 0.42885 |
| 5.80 | 10 | 3 | 0.5 | 1.6049 | 0.8685 | 1.9189 | 0.39899 | 2.0701 | 0.39618 | 0.95275 | 0.43444 |
| 5.81 | 50 | 3 | 0.5 | 1.543 | 0.76646 | 1.8362 | 0.44033 | 1.9805 | 0.44075 | 0.9242 | 0.4697 |
| 5.82 | 100 | 3 | 0.5 | 1.5491 | 0.71689 | 1.8341 | 0.43584 | 1.9767 | 0.43596 | 0.94127 | 0.44454 |
| 5.83 | 500 | 3 | 0.5 | 1.556 | 0.53735 | 1.8573 | 0.43137 | 2.0035 | 0.4324 | 0.93612 | 0.45041 |
| 5.84 | 1000 | 3 | 0.5 | 1.5536 | 0.48741 | 1.854 | 0.43205 | 2.0013 | 0.43036 | 0.93753 | 0.44436 |
| 5.85 | N/A | 3 | 1 | 1.6302 | 0.38907 | 1.9368 | 0.38341 | 2.0876 | 0.38267 | 1.0211 | 0.36878 |
| 5.86 | 10 | 3 | 1 | 1.6596 | 0.86193 | 1.9776 | 0.39254 | 2.13 | 0.39061 | 0.9968 | 0.37929 |
| 5.87 | 50 | 3 | 1 | 1.6249 | 0.69377 | 1.9232 | 0.39596 | 2.0722 | 0.39517 | 1.0438 | 0.39817 |
| 5.88 | 100 | 3 | 1 | 1.6331 | 0.60503 | 1.9213 | 0.3905 | 2.0676 | 0.39047 | 1.082 | 0.39828 |
| 5.89 | 500 | 3 | 1 | 1.6292 | 0.52232 | 1.9333 | 0.39158 | 2.0838 | 0.39002 | 1.0283 | 0.39423 |
| 5.90 | 1000 | 3 | 1 | 1.6296 | 0.42877 | 1.9332 | 0.38971 | 2.0836 | 0.38819 | 1.0325 | 0.38561 |
| 5.91 | N/A | 4 | 0.05 | 0.48277 | 0.95946 | 1.4903 | 0.67336 | 1.7129 | 0.60863 | 1.3578 | 0.21688 |
| 5.92 | 10 | 4 | 0.05 | 1.5523 | 1 | 1.8424 | 0.39029 | 1.9937 | 0.3832 | 0.98672 | 0.30117 |
| 5.93 | 50 | 4 | 0.05 | 1.4149 | 1 | 1.6092 | 0.60752 | 1.8366 | 0.52457 | 1.002 | 0.31351 |
| 5.94 | 100 | 4 | 0.05 | 1.3456 | 1 | 1.6123 | 0.59922 | 1.788 | 0.56563 | 0.95675 | 0.42382 |
| 5.95 | 500 | 4 | 0.05 | 1.1375 | 0.99859 | 1.5413 | 0.65243 | 1.7646 | 0.57142 | 0.99303 | 0.38076 |
| 5.96 | 1000 | 4 | 0.05 | 0.92787 | 0.99268 | 1.5911 | 0.61656 | 1.6756 | 0.63286 | 1.0986 | 0.26613 |
| 5.97 | N/A | 4 | 0.1 | 0.85065 | 0.88051 | 1.5409 | 0.66715 | 1.6546 | 0.66214 | 1.0339 | 0.33198 |
| 5.98 | 10 | 4 | 0.1 | 1.5523 | 1 | 1.8582 | 0.37201 | 2.0009 | 0.37497 | 0.98832 | 0.35558 |
| 5.99 | 50 | 4 | 0.1 | 1.3717 | 0.9984 | 1.6526 | 0.57482 | 1.7714 | 0.59248 | 0.95921 | 0.40481 |
| 5.100 | 100 | 4 | 0.1 | 1.273 | 0.99651 | 1.6078 | 0.60349 | 1.6887 | 0.63596 | 0.98727 | 0.33961 |
| 5.101 | 500 | 4 | 0.1 | 1.1798 | 0.9798 | 1.562 | 0.65177 | 1.6819 | 0.6566 | 0.94205 | 0.43505 |
| 5.102 | 1000 | 4 | 0.1 | 1.0018 | 0.96068 | 1.5859 | 0.64719 | 1.666 | 0.65592 | 1.0026 | 0.36024 |
| 5.103 | N/A | 4 | 0.2 | 1.3261 | 0.6759 | 1.7232 | 0.58787 | 1.8364 | 0.60599 | 0.90315 | 0.5135 |
| 5.104 | 10 | 4 | 0.2 | 1.5602 | 0.99726 | 1.8543 | 0.39542 | 2.0033 | 0.39104 | 0.98311 | 0.32796 |
| 5.105 | 50 | 4 | 0.2 | 1.3988 | 0.97554 | 1.6857 | 0.58034 | 1.8226 | 0.57706 | 0.90179 | 0.48327 |
| 5.106 | 100 | 4 | 0.2 | 1.3714 | 0.95587 | 1.6727 | 0.5994 | 1.8145 | 0.58727 | 0.92767 | 0.43813 |
| 5.107 | 500 | 4 | 0.2 | 1.3312 | 0.87878 | 1.7054 | 0.58365 | 1.8157 | 0.60664 | 0.90626 | 0.48396 |
| 5.108 | 1000 | 4 | 0.2 | 1.3062 | 0.77873 | 1.7114 | 0.58255 | 1.8288 | 0.58642 | 0.89692 | 0.49437 |
| 5.109 | N/A | 4 | 0.5 | 1.5628 | 0.44093 | 1.8751 | 0.42235 | 2.0227 | 0.42425 | 0.94682 | 0.42612 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5.110 | 10 | 4 | 0.5 | 1.5757 | 0.86803 | 1.8785 | 0.40763 | 2.0278 | 0.40658 | 0.9642 | 0.36797 |
| 5.111 | 50 | 4 | 0.5 | 1.5423 | 0.77212 | 1.8399 | 0.44119 | 1.9843 | 0.44236 | 0.93394 | 0.44919 |
| 5.112 | 100 | 4 | 0.5 | 1.5442 | 0.72439 | 1.8437 | 0.43321 | 1.9863 | 0.43751 | 0.93549 | 0.44854 |
| 5.113 | 500 | 4 | 0.5 | 1.5458 | 0.5973 | 1.8555 | 0.43785 | 2.0015 | 0.4397 | 0.93466 | 0.45139 |
| 5.114 | 1000 | 4 | 0.5 | 1.5524 | 0.50392 | 1.863 | 0.42872 | 2.0091 | 0.43114 | 0.93291 | 0.45445 |
| 5.115 | N/A | 4 | 1 | 1.6297 | 0.37731 | 1.9373 | 0.37232 | 2.0882 | 0.37146 | 1.0207 | 0.34497 |
| 5.116 | 10 | 4 | 1 | 1.6425 | 0.77574 | 1.9563 | 0.37974 | 2.1084 | 0.37882 | 1.0076 | 0.34895 |
| 5.117 | 50 | 4 | 1 | 1.6247 | 0.65907 | 1.923 | 0.38808 | 2.0715 | 0.38836 | 1.0492 | 0.37706 |
| 5.118 | 100 | 4 | 1 | 1.6339 | 0.62832 | 1.9277 | 0.38271 | 2.0754 | 0.38284 | 1.0737 | 0.3843 |
| 5.119 | 500 | 4 | 1 | 1.6278 | 0.51554 | 1.928 | 0.37757 | 2.077 | 0.37835 | 1.0475 | 0.35899 |
| 5.120 | 1000 | 4 | 1 | 1.6269 | 0.45987 | 1.9316 | 0.38396 | 2.0817 | 0.38369 | 1.0313 | 0.36597 |
| 5.121 | N/A | 5 | 0.05 | 0.32581 | 0.98152 | 1.4873 | 0.67459 | 1.9024 | 0.5114 | 1.1147 | 0.23352 |
| 5.122 | 10 | 5 | 0.05 | 1.575 | 1 | 1.8322 | 0.41535 | 2.0221 | 0.351 | 1.0877 | 0.15247 |
| 5.123 | 50 | 5 | 0.05 | 1.4578 | 1 | 1.7422 | 0.49875 | 1.8861 | 0.48485 | 0.96292 | 0.42188 |
| 5.124 | 100 | 5 | 0.05 | 1.3832 | 1 | 1.6365 | 0.59156 | 1.8448 | 0.52353 | 1.1207 | 0.33185 |
| 5.125 | 500 | 5 | 0.05 | 1.1762 | 0.99976 | 1.7032 | 0.52853 | 1.8432 | 0.51933 | 1.0147 | 0.36877 |
| 5.126 | 1000 | 5 | 0.05 | 0.9446 | 0.99898 | 1.5281 | 0.64841 | 1.8279 | 0.53828 | 1.0845 | 0.18534 |
| 5.127 | N/A | 5 | 0.1 | 0.60167 | 0.93846 | 1.4416 | 0.72131 | 1.6907 | 0.6215 | 1.014 | 0.32871 |
| 5.128 | 10 | 5 | 0.1 | 1.5717 | 1 | 1.8539 | 0.38391 | 2.0132 | 0.36359 | 0.98754 | 0.34815 |
| 5.129 | 50 | 5 | 0.1 | 1.4568 | 0.99978 | 1.7121 | 0.53818 | 1.887 | 0.49245 | 1.033 | 0.4031 |
| 5.130 | 100 | 5 | 0.1 | 1.3858 | 0.99821 | 1.6755 | 0.56627 | 1.8929 | 0.48072 | 1.107 | 0.32359 |
| 5.131 | 500 | 5 | 0.1 | 1.2037 | 0.98435 | 1.5751 | 0.63308 | 1.6423 | 0.65916 | 0.97994 | 0.36971 |
| 5.132 | 1000 | 5 | 0.1 | 0.95301 | 0.98045 | 1.4708 | 0.68847 | 1.6293 | 0.66504 | 1.0486 | 0.27773 |
| 5.133 | N/A | 5 | 0.2 | 1.2073 | 0.75899 | 1.6547 | 0.65383 | 1.7865 | 0.63858 | 0.90308 | 0.50377 |
| 5.134 | 10 | 5 | 0.2 | 1.5642 | 0.99879 | 1.8444 | 0.40871 | 2.0091 | 0.37818 | 1.0284 | 0.25426 |
| 5.135 | 50 | 5 | 0.2 | 1.484 | 0.97964 | 1.7717 | 0.49775 | 1.9075 | 0.50075 | 0.96537 | 0.38532 |
| 5.136 | 100 | 5 | 0.2 | 1.3924 | 0.97957 | 1.7072 | 0.55271 | 1.812 | 0.59404 | 0.9747 | 0.36922 |
| 5.137 | 500 | 5 | 0.2 | 1.2675 | 0.92489 | 1.6178 | 0.64598 | 1.7802 | 0.61262 | 0.93557 | 0.42602 |
| 5.138 | 1000 | 5 | 0.2 | 1.1948 | 0.86758 | 1.5811 | 0.67952 | 1.7718 | 0.61811 | 0.92988 | 0.4419 |
| 5.139 | N/A | 5 | 0.5 | 1.5544 | 0.46083 | 1.8672 | 0.44461 | 2.0185 | 0.43892 | 0.94629 | 0.43096 |
| 5.140 | 10 | 5 | 0.5 | 1.5866 | 0.92808 | 1.8831 | 0.41633 | 2.0409 | 0.39385 | 0.97968 | 0.33035 |
| 5.141 | 50 | 5 | 0.5 | 1.5583 | 0.78184 | 1.8645 | 0.42914 | 2.0147 | 0.42355 | 0.94562 | 0.43133 |
| 5.142 | 100 | 5 | 0.5 | 1.543 | 0.77663 | 1.844 | 0.45032 | 1.9932 | 0.44426 | 0.94305 | 0.43466 |
| 5.143 | 500 | 5 | 0.5 | 1.5313 | 0.6468 | 1.835 | 0.46745 | 1.9862 | 0.45854 | 0.93803 | 0.44089 |
| 5.144 | 1000 | 5 | 0.5 | 1.5291 | 0.5639 | 1.8332 | 0.46996 | 1.9852 | 0.45967 | 0.93198 | 0.45169 |
| 5.145 | N/A | 5 | 1 | 1.6298 | 0.37667 | 1.9387 | 0.37206 | 2.0898 | 0.37085 | 1.0186 | 0.34587 |
| 5.146 | 10 | 5 | 1 | 1.6405 | 0.86121 | 1.952 | 0.38982 | 2.1054 | 0.37847 | 1.0205 | 0.3196 |
| 5.147 | 50 | 5 | 1 | 1.6307 | 0.6807 | 1.9374 | 0.38421 | 2.0883 | 0.38159 | 1.031 | 0.3611 |
| 5.148 | 100 | 5 | 1 | 1.6308 | 0.62244 | 1.9288 | 0.39135 | 2.0784 | 0.38645 | 1.063 | 0.37372 |
| 5.149 | 500 | 5 | 1 | 1.6258 | 0.54408 | 1.931 | 0.38108 | 2.0811 | 0.38047 | 1.0294 | 0.35737 |
| 5.150 | 1000 | 5 | 1 | 1.6264 | 0.47426 | 1.9311 | 0.3891 | 2.0819 | 0.38532 | 1.0302 | 0.37727 |

Table B.6: Performance of the generalized regression network for the spatial input pattern.

| Net ID | No of Cluster | Width | RMSE Training | CC Training | RMSE Validation | CC Validation | RMSE Test (i) | CC Test (i) | RMSE Test (ii) | CC Test (ii) |
|---|---|---|---|---|---|---|---|---|---|---|
| 6.1 | N/A | 0.05 | 0.77433 | 0.89408 | 1.7628 | 0.4802 | 1.862 | 0.5037 | 1.0488 | 0.34869 |
| 6.2 | 10 | 0.05 | 1.4661 | 1 | 1.7198 | 0.52391 | 1.9079 | 0.46313 | 0.98065 | 0.35931 |
| 6.3 | 50 | 0.05 | 1.4122 | 1 | 1.7322 | 0.50285 | 1.8639 | 0.4984 | 0.99219 | 0.39295 |
| 6.4 | 100 | 0.05 | 1.3972 | 0.99992 | 1.7171 | 0.51505 | 1.8855 | 0.47978 | 0.99869 | 0.39891 |
| 6.5 | 500 | 0.05 | 1.2341 | 0.99363 | 1.6588 | 0.55751 | 1.7985 | 0.54928 | 1.0524 | 0.28006 |
| 6.6 | 1000 | 0.05 | 1.1196 | 0.98337 | 1.6759 | 0.54476 | 1.7898 | 0.55568 | 1.0686 | 0.29763 |
| 6.7 | N/A | 0.1 | 1.205 | 0.70814 | 1.7138 | 0.51391 | 1.8037 | 0.54795 | 1.0035 | 0.3657 |
| 6.8 | 10 | 0.1 | 1.4659 | 1 | 1.7571 | 0.48612 | 1.9122 | 0.46334 | 0.98748 | 0.36473 |
| 6.9 | 50 | 0.1 | 1.4187 | 0.99912 | 1.7203 | 0.51547 | 1.8579 | 0.51241 | 1.0122 | 0.34957 |
| 6.10 | 100 | 0.1 | 1.3854 | 0.99386 | 1.7136 | 0.52331 | 1.8474 | 0.51878 | 0.96885 | 0.40098 |
| 6.11 | 500 | 0.1 | 1.2998 | 0.93929 | 1.6896 | 0.54028 | 1.8181 | 0.54438 | 1.0284 | 0.32488 |
| 6.12 | 1000 | 0.1 | 1.2252 | 0.89916 | 1.6674 | 0.56083 | 1.8185 | 0.54152 | 0.99816 | 0.35022 |
| 6.13 | N/A | 0.2 | 1.4124 | 0.55996 | 1.7271 | 0.53152 | 1.8787 | 0.50519 | 0.98014 | 0.35258 |
| 6.14 | 10 | 0.2 | 1.4564 | 0.99972 | 1.7327 | 0.51311 | 1.9063 | 0.46821 | 0.98292 | 0.36563 |
| 6.15 | 50 | 0.2 | 1.4186 | 0.96653 | 1.7236 | 0.52578 | 1.8718 | 0.51183 | 0.96857 | 0.37729 |
| 6.16 | 100 | 0.2 | 1.431 | 0.93977 | 1.7361 | 0.52375 | 1.8788 | 0.50819 | 0.97057 | 0.36937 |
| 6.17 | 500 | 0.2 | 1.4184 | 0.78889 | 1.7404 | 0.52841 | 1.8912 | 0.50726 | 0.96675 | 0.36962 |
| 6.18 | 1000 | 0.2 | 1.4093 | 0.66528 | 1.7343 | 0.53572 | 1.8921 | 0.50878 | 0.97476 | 0.35575 |
| 6.19 | N/A | 0.5 | 1.531 | 0.4513 | 1.8208 | 0.46926 | 1.9923 | 0.41392 | 0.96461 | 0.34952 |
| 6.20 | 10 | 0.5 | 1.4983 | 0.96174 | 1.77 | 0.48993 | 1.9468 | 0.43517 | 0.98218 | 0.34239 |
| 6.21 | 50 | 0.5 | 1.5283 | 0.85017 | 1.8132 | 0.46954 | 1.985 | 0.41609 | 0.96058 | 0.36487 |
| 6.22 | 100 | 0.5 | 1.5233 | 0.74295 | 1.8013 | 0.46848 | 1.9756 | 0.415 | 0.96915 | 0.35896 |
| 6.23 | 500 | 0.5 | 1.5379 | 0.57179 | 1.8294 | 0.46385 | 1.9999 | 0.40996 | 0.96186 | 0.35537 |
| 6.24 | 1000 | 0.5 | 1.5382 | 0.48926 | 1.8308 | 0.4683 | 2.0016 | 0.4122 | 0.96114 | 0.35527 |
| 6.25 | N/A | 1 | 1.5869 | 0.39199 | 1.8847 | 0.40299 | 2.0478 | 0.36109 | 0.97574 | 0.33176 |
| 6.26 | 10 | 1 | 1.5484 | 0.88541 | 1.8365 | 0.44468 | 2.0052 | 0.39527 | 0.97361 | 0.33952 |
| 6.27 | 50 | 1 | 1.5836 | 0.76783 | 1.8813 | 0.41766 | 2.0447 | 0.37302 | 0.9746 | 0.34023 |
| 6.28 | 100 | 1 | 1.5841 | 0.69151 | 1.8816 | 0.41539 | 2.0452 | 0.37094 | 0.97409 | 0.34088 |
| 6.29 | 500 | 1 | 1.589 | 0.51137 | 1.8845 | 0.39942 | 2.0478 | 0.35726 | 0.97869 | 0.33546 |
| 6.30 | 1000 | 1 | 1.5938 | 0.41743 | 1.8912 | 0.39595 | 2.0536 | 0.3546 | 0.98049 | 0.33276 |

# Appendix C: Wavelet Algorithm

The wavelet based denoising is usually done in three steps, (i) the image is converted to wavelet coefficients using a 2D discrete wavelet transform (DWT), (ii) low coefficients are replaced with zero, which is known as hard thresholding, and (iii) the image is reconstructed from its coefficients (Ghazel, 2004). The magnitude of the lower coefficient threshold can be determined from median filtering techniques or from trial.

The image is considered as a 2D matrix of dimension $m \times n$ where $m$ and $n$ are the total number of rows and columns respectively. One restriction on the size (the number of rows or the number of column) of the matrix is that it should be expressed by $2^a$ where $a$ is an integer. It is not necessary that the number of rows and the number of columns be equal. Each element of the matrix represents brightness of the pixel located in that position.

The 2D discrete wavelet transform mentioned in step 1 is done in two steps-(i) performing 1D wavelet transform on each row of the image and then (ii) performing 1D wavelet transform on each column of the image (Walker, 1999, pp 66). Performing 1D DWT on each row of the image is again done in two steps (i) applying a low pass filter on each element of the row which will produce $n/2$ elements in each row and storing them in a separate matrix of dimension $m \times n/2$ and (ii) applying high pass filter on each element of the row and storing them in a separate matrix having dimension $m \times n/2$ and horizontally concatenate these two matrices together which will again produce an $m \times n$ matrix (Aboufadel and Schlicker, 1999, pp 71). Each elements of this matrix is the wavelet coefficients of the original image matrix.

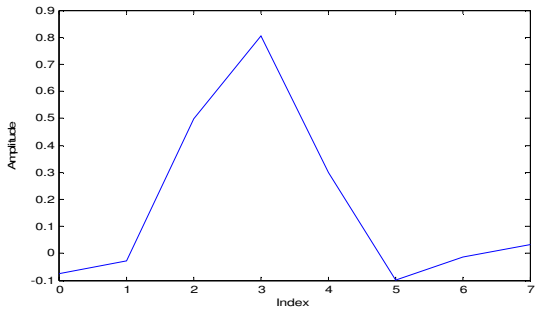


Figure C.1: The Symmlet wavelet.

The filters applied to get the wavelet coefficients are orthogonal wavelets such as Haar wavelets, Coiflet wavelets, Symmlet wavelets, Daubechies wavelets, Vaidanathan wavelets, etc. This study uses symmlet 4 wavelet which is shown in Figure C.1. The 1D DWT is done by the convolution of the signal and the wavelet and the values obtained is termed as wavelet coefficients. In case of image denoising, signal will refer to each row or column vector of the image.

The symmlet wavelet itself is used as a low pass filter. The low pass filtering operation can be mathematically described as follows:

$$L_k = \sum_{j=0}^{J-1} l_j S_{j+2k}, \qquad k = 0,...,(m/2-1) \qquad \text{Eq C.1}$$

where $J$ is the length of the wavelet, $l_j$ refers to each element of the filter, $S$ refers to the signal, $m$ is the length of the signal, and $L_k$ refers to the derived wavelet coefficients after low pass filtering. If the subscript of $S$ exceeds the length of the signal, elements starting from the leading edge of the signal should be used. This will occur in case of $J>2$.

Applying the high pass filter is a little more complicated. It can be derived from the low pass filter used and can be mathematically described as follows:

$$h_j = (-1)^{j+1} l_{J-j-1}, \qquad j = 0,...,(J-1)$$ Eq C.2

where $h_j$ is the mirror filter and $J$ is the number of elements in the filter. A set of low pass and high pass filter is termed as quadratic mirror filter.

The high pass filtering operation can be mathematically described as follows,

$$H_k = \sum_{j=0}^{J-1} h_j S_{j+2k-(J-2)}, \qquad k = 0,...,(m/2-1)$$ Eq C.3

where $H_k$ is the wavelet coefficients obtained after high pass filtering. If the subscript of $S$ gets negative, the elements starting from the trailing edge should be used. This will occur when $J>2$. This way of resolving the boundary problem is known as the periodic method (Aboufadel and Schlicker, 1999, pp 76).

Performing 1D wavelet transform on each column is done in the same way as for the rows. The only difference is that this transformation will produce two matrices each having $m/2$ rows and $n$ columns and these two matrices should be vertically concatenated to obtain an $m \times n$ matrix of wavelet coefficients.

Performing a set of row-wise and column-wise transformations is called 1st level of decomposition. This process may be iterated and the wavelet coefficient matrix obtained after the previous iteration will be used for decomposition. However, the number of rows and columns on which the decomposition is to be performed will be halved. For example, if the size of the image is 256x256, and if a total 3 levels of decomposition are done, the 1st level of decomposition will be carried out over a 256x256 data matrix, the 2nd level of decomposition will be done on 128x128 data matrix and the

rest of the data will remain unaffected. The 3rd level will be done on the 64x64 data matrix.

After completing the decomposition, the lower values of wavelet coefficients are set to zero and this is termed hard thresholding. Then the image is reconstructed from the thresholded wavelet coefficient matrix. The process is termed as re-composition. The re-composition will start column-wise on the dimension on which the final level of decomposition was done. After that, row-wise re-composition will be carried out and a set of column wise and row wise re-composition will complete 1st level of re-composition. Thereafter, the 2nd level re-composition will be carried out. This will continue until the numbers of re-composition is equal to the total number of decompositions employed. If no wavelet coefficient is set to zero or changed, the re-composition process should reproduce exactly the initial image.

During re-composition of a column at a particular level, it is necessary to separate the vector in two halves, the upper half should convolute through a low pass filter and the lower half should convolute through a high pass filter. These filtering will make the length of the vector twice the previous one. For example if the column has eight elements, the first four elements will convolute through a low pass filter and will form eight elements, and the last four elements will convolute through a high pass filter and will form another eight elements. The two generated vectors each having eight elements will then be added together. These process will continue for each column and then for each row and thus the 1st level of re-composition will be completed. The signal re-composition equations are shown below,

$$S_{lk} = \begin{cases} \sum_{j=0}^{J/2-1} l_{2j} L_{k/2-j} & k=0,2,4,...,(n-2) \\ \sum_{j=0}^{J/2-1} l_{2j+1} L_{(k-1)/2-j} & k=1,3,5,...,(n-1) \end{cases} \quad \text{Eq C.4}$$

$$S_{hk} = \begin{cases} \sum_{j=0}^{J/2-1} h_{2j} H_{(k+2)/2-j} & k=0,2,4,...,(n-2) \\ \sum_{j=0}^{J/2-1} h_{2j+1} H_{(k+1)/2-j} & k=1,3,5,...,(n-1) \end{cases} \quad \text{Eq C.5}$$

$$S_k = S_{lk} + S_{hk} \quad \text{Eq C.6}$$

where the $L$ and the $H$ are the wavelet coefficients obtained after low pass and high pass filtering at a particular level and $S$ is the signal reconstructed from the wavelet coefficients. The $n$ is the length of the column which is twice the length of the $L$ or $H$, and $J$ is the length of the filter. The $l$ and $h$ are the low pass and the high pass filter as used previously. If the subscripts of $L$ or $H$ are negative, it will pick elements from the trailing edge of the $L$ or $H$ and if the subscript is positive and greater than the length of the $L$ or $H$, values should be picked from the leading edge of the signal. This will be needed when $J>2$.