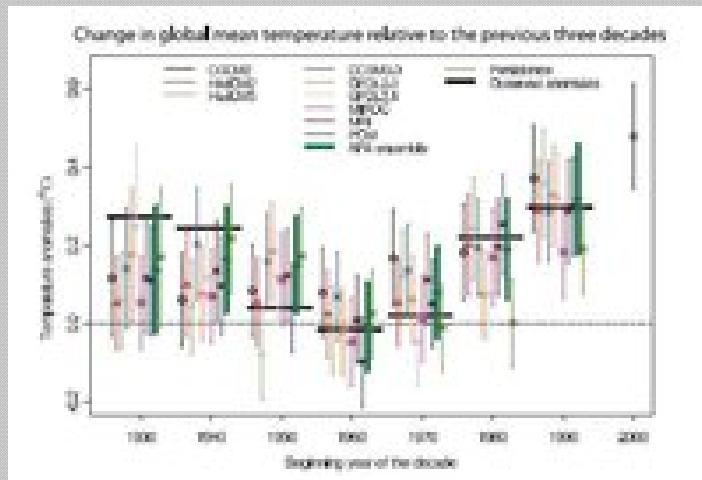


CREST PUBLICATIONS SERIES

Volume 6
No: 05-2008

A ROBUST NEURAL NETWORK SYSTEM ENSEMBLE APPROACH FOR DETECTING AND ESTIMATING SNOWFALL FROM THE ADVANCED MICROWAVE SOUNDING UNIT



By:

Yajaira Mejia, Shayesteh Mahani, Vasil Diyamandoglu and Cezar Kongoli
Civil Engineering Department, The City College of City University of New York



Published By:
NOAA-CREST
National Oceanic and Atmospheric Administration-
Cooperative Remote Sensing Science & Technology Center
City University of New York, NY



© 2008

YAJAIRA MEJIA

All Rights Reserved

ACKNOWLEDGMENT

This study was supported and monitored by the National Oceanic and Atmospheric Administration (NOAA) under Grant Number **NA06OAR4810162**. The views, opinions, and findings contained in this report are those of the author(s) and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. Government position, policy, or decision.

DISCLAIMER

The views, opinions, and findings contained in this report are those of the author(s) and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. Government position, policy, or decision.

A ROBUST NEURAL NETWORK SYSTEM ENSEMBLE APPROACH FOR
DETECTING AND ESTIMATING SNOWFALL FROM THE ADVANCED
MICROWAVE SOUNDING UNIT

by

YAJAIRA MEJIA

A dissertation submitted to the Graduate Faculty in Engineering
in partial fulfillment of the requirements for the degree of Doctor of Philosophy
The City University of New York

2008

© 2008

YAJAIRA MEJIA

All Rights Reserved

This manuscript has been read and accepted for the
Graduate Faculty in Engineering in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Date	Prof. Reza Khanbilvardi Professor of Civil Engineering Chair of Examining Committee
Date	Dean Mumtaz K. Kassir Executive Officer

Supervisory Committee

Prof. Hosni Ghedira

Prof. Shayesteh Mahani

Prof. Vasil Diyamandoglu

Dr. Cezar Kongoli

THE CITY UNIVERSITY OF NEW YORK

Abstract

**A ROBUST NEURAL NETWORK ENSEMBLE APPROACH FOR DETECTING
AND ESTIMATING SNOWFALL FROM THE ADVANCED MICROWAVE
SOUNDING UNIT (AMSU)**

by

Yajaira Mejia

Adviser: Professor Reza Khanbilvardi

The principal intent of this research is to: (a) investigate the potential of passive microwave data from AMSU in detecting snowfall events and in measuring their intensity, and (b) evaluate the effect of both land cover and atmospheric conditions on the retrieval accuracy. A neural-network-based model has been developed and has shown a great potential in detecting and estimating the intensity of snowfall events. This algorithm has been applied for different snow storms occurred in four winter seasons in the North-East of United States. Additional information such as cloud cover and air temperature were added to the process to reduce misidentified snowfall pixels. Only pixels with cloud cover and falling within a specific range of temperature are presented to the snowfall detection model. Surface temperature collected from ground station-based observations and archived by the National Climatic Data Center (NCDC) were used for this test. Different heavy storm events and non-snowfall observations that occurred at the same time as AMSU acquisition were selected. Hourly snow accumulation data collected by the NCDC were used as truth data to train and validate the model. The results indicate that the neural-

network-based model provides a significant improvement in snowfall detection accuracy over existing satellite-based methods. Most importantly, the neural network system product is a map indicating the snowfall area and the respective intensity level for each pixel.

ACKNOWLEDGEMENTS

I would like to thank Dr. Reza Khanbilvardi for providing me the opportunity to continue my studies at the City University of New York with full financial support. His guidance and advice throughout the last four years have been indispensable. I would also like to thank Dr. Shayesteh Mahani for sharing her knowledge and teaching me all the important remote sensing concepts, which I did not have when I joined the program. Her patience and teaching skills made it easier for me. Her dedication in getting all the ideas together for this project is appreciated.

My deepest thanks to Dr. Hosni Ghedira. Without his support, time, knowledge, and advice, this project would have never ended. He guided me and taught me lots of techniques that were applied in this research. His personality as advisor is admirable. I express my immense gratitude to him. I would like to thank Prof. Vasil Diyamandoglu who contributed as member of my PhD. Committee. I am thankful to Dr. Cezar Kongoli for helping me with data and giving me time to discuss and clarify concepts that are presented here.

Most importantly, I want to express my sincere and deep gratitude to the love of my life, Jose Luis for his support. He has helped me in correcting all the papers written in the last three years. His understanding when I did not have time to share with him, especially in weekends is appreciated. Unquestionably, his support and love was key in finishing my doctoral studies.

Last but certainly not least, many thanks to my lovely parents, Loncer and Marina. Without their help I would not be here today. Their immense love and support encouraged me to go forward for a higher degree. I love you so much. Thanks to my brother for his support in the decisions I have made which helped me to continue with my plans.

Table of Contents

1. INTRODUCTION	1
1.1 REMOTE SENSING AND PRECIPITATION.....	1
1.2 THESIS OBJECTIVES	5
1.3 THESIS MOTIVATION.....	5
1.4 THESIS STATEMENT	7
1.5 THESIS OVERVIEW	7
2. LITERATURE REVIEW	9
2.1 PRECIPITATION RAIN/SNOW DISCRIMINATION FROM SATELLITE	9
2.2 SNOWFALL RETRIEVAL.....	12
2.2.1 <i>Using Satellite Information</i>	12
2.2.2 <i>Using Ground-Based Radar Information</i>	14
2.3 NEURAL NETWORK ENSEMBLE APPROACH.....	15
3. RADIO DETECTION AND RANGING (RADAR)	17
3.1 HISTORY	17
3.2 DOPPLER RADAR	19
3.3 DESCRIPTION OF THE ANTENNA.....	20
3.4 HOW RADAR WORKS	20
3.5 RADAR OPERATIONAL MODES.....	22
3.5.1 <i>Clean Air Mode:</i>	22
3.5.2 <i>Precipitation Mode:</i>	23
3.6 REFLECTIVITY	24
3.7 NON-PRECIPITATION ECHOES AND CLUTTER.....	25
3.8 RADAR MAP PROJECTIONS.....	26
3.9 TYPES OF DOPPLER RADAR IMAGES	27
3.9.1 <i>Reflectivity Images:</i>	27
3.9.2 <i>Velocity Images:</i>	29
3.9.3 <i>Precipitation Images:</i>	30
3.9.4 <i>Weather Warnings:</i>	32
3.10 RELATIONSHIP BETWEEN REFLECTIVITY AND PRECIPITATION	33
3.11 PRECIPITATION RADAR LIMITATIONS	34
3.12 THE USE OF DOPPLER RADAR IN SNOWFALL DETECTION	39
4. MICROWAVE INFORMATION	41
4.1 ACTIVE MICROWAVE SENSORS.....	44

4.2	PASSIVE MICROWAVE SENSORS.....	45
4.3	ADVANCED MICROWAVE SOUNDING UNIT (AMSU)	47
4.3.1	<i>Advanced Microwave Sounding Unit-A</i>	48
4.3.2	<i>Advanced Microwave Sounding Unit-B</i>	51
4.4	THE PHYSICS OF PASSIVE MICROWAVE	53
4.4.1	<i>Absorption and Emission by Gases</i>	53
4.4.2	<i>Water Vapor Absorption</i>	56
4.4.3	<i>Oxygen Absorption</i>	56
4.4.4	<i>Extinction and Emission by Clouds and Precipitation</i>	57
4.4.5	<i>Extinction and Backscattering by Snow</i>	57
5.	ARTIFICIAL NEURAL NETWORK	60
5.1	TYPES OF NEURAL NETWORK	62
5.1.1	<i>Multi-Layer Perceptron</i> :.....	62
5.1.2	<i>Radial Basis Function Network</i> :.....	63
5.1.3	<i>Hohonen's Self Organizing Map algorithm</i> :	63
5.2	NEURAL NETWORK ARCHITECTURE	64
5.3	ACTIVATION FUNCTIONS	65
5.4	NEURAL NETWORK TRAINING TECHNIQUES	68
5.5	BACKPROPAGATION ALGORITHM	69
5.6	THE PERCEPTRON LEARNING RULE	70
5.7	TYPES OF BACKPROPAGATION ALGORITHMS.....	72
5.7.1	<i>Variable Learning Rate (GDX)</i> :	72
5.7.2	<i>Resilient Backpropagation (RP)</i> :.....	73
5.7.3	<i>Conjugate Gradient Algorithms</i> :	73
5.7.4	<i>Quasi-Newton Algorithms</i> :	75
5.7.5	<i>Levenberg-Marquardt (LM) algorithm</i> :	76
5.8	NEURAL NETWORK TRAINING	77
5.9	MOST COMMON PROBLEMS IN TRAINING AND/OR TESTING.....	80
5.9.1	<i>Poor training performance</i> :.....	80
5.9.2	<i>Poor generalization performance</i> :	80
6.	DATA ACQUISITION.....	83
6.1	TRAINING AREA AND DATA SET	83
6.2	INPUTS IN THE NEURAL NETWORK MODEL.....	91
6.2.1	<i>Passive microwave</i>	91
6.2.2	<i>Snow cover</i>	95
6.3	DATA FILTERS	95

6.3.1	<i>Cloud cover</i>	96
6.3.2	<i>Surface temperature</i>	98
6.4	VALIDATION STUDY AREA	100
7.	METHODOLOGY, RESULTS AND DISCUSSION	101
7.1	PART A: SNOWFALL DETECTION.....	105
7.1.1	<i>Data noises Filters:</i>	105
7.1.2	<i>Model Inputs</i>	107
7.1.3	<i>Neural Network</i>	107
7.1.4	<i>Neural Network System 1 (NN1)</i>	123
7.1.5	<i>Effect of snow cover on model accuracy</i>	127
7.2	PART B: SNOWFALL ESTIMATION	129
7.2.1	<i>Neural Network system 2 (NN2) using method A</i>	130
7.2.2	<i>Neural Network system 2 (NN2) using method B</i>	147
8.	CONCLUSION	152
	REFERENCES.....	189

List of Tables

Table 1: Satellites Specifications.....	48
Table 2: AMSU-A Instrument Parameters	49
Table 3: AMSU-B Instrument Parameters	52
Table 4: Activation Function Expressions.....	67
Table 5: Local Climatological Data, Page 1.....	84
Table 6: Local Climatological Data, Page 2.....	85
Table 7: Stations's ID and coordinates.....	90
Table 8: Local Climatological Data Unedited Hourly Observations.....	97
Table 9: Cloud Covered Categories	98
Table 10: ANOVA results when using different number of hidden layers and nodes	111
Table 11: ANOVA test results using different combinations of activation functions	112
Table 12: ANOVA test results using different training algorithm	116
Table 13: ANOVA test results for different momentum values	120
Table 14: ANOVA test results for different inputs	123
Table 15: Confusion matrix of neural network output	142
Table 16: ANOVA test for NN accuracy when using different inputs.....	150
Table 17: ANOVA test for the accuracy of 100 nets with 1HL with 11 and 12 nodes	161
Table 18: ANOVA test for the accuracy of 100 nets with 2HL with 6 and 8 nodes	161
Table 19: ANOVA test for the accuracy of 100 nets with 2HL with 9 and 10 nodes	162
Table 20: ANOVA test for the accuracy of 100 nets with 2HL with 8 and 9 nodes	162
Table 21: ANOVA test for the accuracy of 100 nets with 1HL with 5 and 12 nodes	163
Table 22: ANOVA test for the accuracy of 100 trained nets using TTT and TTP	164
Table 23: ANOVA test for the accuracy of 100 trained nets using TTT and LLL.....	164
Table 24: ANOVA test for the accuracy of 100 trained nets using LLP and LLT	165
Table 25: ANOVA test for the accuracy of 100 trained nets using TTL and PPT	165
Table 26: ANOVA test for the accuracy of 100 trained nets using PPL and TTL	166
Table 27: ANOVA test for the accuracy of 100 trained nets using PPT and TTT	166
Table 28: ANOVA test for the accuracy of 100 trained nets using RP and LM	167
Table 29: ANOVA test for the accuracy of 100 trained nets using LM and GDX.....	167
Table 30: ANOVA test for the accuracy of 100 trained nets using BFG and CGB	168
Table 31: ANOVA test for the accuracy of 100 trained nets using SCG and OSS	168
Table 32: ANOVA test for the accuracy of 100 trained nets using OSS and GDX	169
Table 33: ANOVA test for the accuracy of 100 trained nets using CGF and RP	169
Table 34: ANOVA test for the accuracy of 100 trained nets using CGP and LM.....	170
Table 35: ANOVA test for the accuracy of 100 trained nets using LM and SCG.....	170

Table 36: ANOVA test for the accuracy of 100 trained nets using LM and GDX.....	171
Table 37: ANOVA test for the accuracy of 100 nets using a momentum of 0.7 and 0.75	172
Table 38: ANOVA test for the accuracy of 100 nets using a momentum of 0.8 and 0.95	172
Table 39: ANOVA test for the accuracy of 100 nets using a momentum of 0.7 and 1.0	173
Table 40: ANOVA test using all AMSU-B but 89 GHz and all AMSU-B 150 GHz.....	174
Table 41: ANOVA test using all AMSU-B but 183+/-1 GHz and all AMSU-B but 183+/-3 GHz	174
Table 42: ANOVA test using all AMSU-B but 89 and150 GHz and all AMSU-B	175
Table 43: ANOVA test using all AMSU-B but 89 GHz and all AMSU-B	175
Table 44: ANOVA test using all AMSU-B and snow cover information and only all AMSU-B	176
Table 45: ANOVA test using one node in the output layer using different inputs.....	177
Table 46: ANOVA test using one node in the output layer using different inputs.....	177
Table 47: ANOVA test using two nodes in the output layer using different inputs	178
Table 48: ANOVA test using two nodes in the output layer using different inputs	178
Table 49: ANOVA test using all AMSU-B frequencies with one and two nodes.....	179

List of Figures

Figure 1: WSR-57 radar antenna.....	17
Figure 2: Radar covered area.....	19
Figure 3: WSR-88D radar antenna.....	20
Figure 4: Radar beam characteristics	21
Figure 5: Clean air mode	22
Figure 6: Precipitation mode	23
Figure 7: Severe mode.....	24
Figure 8: Radar image	26
Figure 9: Base reflectivity image	27
Figure 10: Composite reflectivity image.....	28
Figure 11: Base velocity image	29
Figure 12: Storm relative motion image.....	30
Figure 13: One hour precipitation image.....	31
Figure 14: Storm total precipitation image.....	32
Figure 15: Weather warnings image.....	33
Figure 16: Clean Air mode scale.....	34
Figure 17: Precipitation mode scale	34
Figure 18: Radar beam diameter	35
Figure 19: Images of precipitation blocked from radar view	36
Figure 20: Shower evaporation below cloud.....	37
Figure 21: Radar beam propagation paths.....	38
Figure 22: Microwave Spectrum	42
Figure 23: Windows and sounder channels.....	43
Figure 24: Forms of passive microwave energy emissions	46
Figure 25: Polar orbiting environmental satellite	47
Figure 26: (a) AMSU-A1 (b) AMSU-A2	49
Figure 27: AMSU A scanning characteristics	50
Figure 28: AMSU B scanning	52
Figure 29: Artificial Neural Network	62
Figure 30: Node process computation	64
Figure 31: Activation functions.....	65
Figure 32: Contribution of each training group to the global training process.....	79
Figure 33: Stations used to collect training data.....	87
Figure 34: Training pixels distribution.....	89
Figure 35: AMSU Temporal Analysis	92

Figure 36: AMSU-89GHz precipitation and non-precipitation pixels comparison	93
Figure 37: AMSU-150 GHz precipitation and non-precipitation pixels comparison.	94
Figure 38: Surface Temperature Threshold.....	99
Figure 39: Validating Study Area.....	100
Figure 40: Overall approach for snowfall detection and estimation.....	104
Figure 41: Snowfall detection.	106
Figure 42: (a) Average Accuracy and (b) Standard Deviation of the Model using One Hidden Layer.....	109
Figure 43: (a) Average Accuracy and (b) Standard Deviation of the Model using Two Hidden Layer	110
Figure 44: (a) Average Accuracy (b) STDEV of the Model using different activation functions.....	113
Figure 45: (a) Average Accuracy (b) STDEV of the Model using different training algorithms.....	115
Figure 46: (a) Average Accuracy (b) STDEV of the Model using different learning rate values.	117
Figure 47: (a) Average Accuracy (b) STDEV of the Model using different momentum values.....	119
Figure 48: Average accuracy of the model output using different combinations of inputs.....	121
Figure 49: (a) Average Accuracy (b) STDEV of the Model using different AMSU-B channels as input.	122
Figure 50: (a) Simulation and (b) true data of the independent area.	125
Figure 51: Simulation output accuracy at different thresholds.....	127
Figure 52: Average accuracy of the model using different inputs in the model.....	128
Figure 53: Snowfall Estimation Approach	129
Figure 54: Snowfall data frequency graph	130
Figure 55: Neural nets accuracy using one hidden layer	133
Figure 56: Neural nets accuracy using two hidden layer.....	134
Figure 57: Neural nets accuracy using different combinations of activation functions.....	135
Figure 58: Neural nets accuracy using training algorithms	136
Figure 59: Neural nets accuracy using different learning rate values.....	138
Figure 60: Neural nets accuracy using different momentum values.....	140
Figure 61: (a)NN output (b)Station-based truth data (c) Radar-based data for January 27, 2004 N15des ..	143
Figure 62: (a) NN output (b) Station-based truth data (c) Radar-based data for January 25, 2004 N16asc ..	144
Figure 63: (a) NN output (b) Station-based truth data (c) Radar-based data for January 14, 2003 N15des ..	145
Figure 64: (a) NN output (b) Station-based truth data (c) Radar-based data for January 16, 2003 N16des ..	146
Figure 65: (a) Average Accuracy (b) STDEV using one node in the output layer.....	148
Figure 66: (a) Average Accuracy (b) STDEV using two nodes in the output layer.	149
Figure 67: (a) NN output (b) Corresponding station-based truth data using method B.....	151

1. Introduction

1.1 Remote Sensing and Precipitation

Precipitation, including rainfall and snowfall, is a key factor generally required in hydrological, and meteorological modeling as well as for other weather related applications such as flood and weather forecasting, climate change assessment, and water resource management. At present, there is no operational global snowfall product that could be used in weather related modeling. For instance, having an accurate detection and estimation of snowfall is indispensable for efficient meteorological nowcasting and modeling. Also, snowfall directly affects the amount of accumulated snow depth. Afterwards, the accumulated snow is used as source of water in some locations.

In the last three decades, remote sensing has been increasingly used in various fields. Satellite-based observations have become a significant source of information for different meteorological applications due to their improved temporal and spatial coverage. In fact, passive microwave remote sensing techniques have been investigated by numerous researchers using various sensors and were found to be potentially effective for detecting and estimating snowfall. Snowfall coverage and intensity are used to update existing Snow Water Equivalent (SWE) measurements. These measurements are required by different scientific communities such as River Forecast Centers (RFCs) and Weather Forecast Offices (WFOs). In the United States, there are traditional ground-based techniques which

provide snowfall depth and its related SWE observations over specific stations but there is no gridded measuring system of snowfall.

In the past four decades, several researches (Sekhon and Srivastava, 1970; Fujiyoshi et al. 1990; Carlson and Marshall, 1972; Boucher and Wieler, 1985; Schmid and Mathis, 2004) have tried to develop different algorithms for detecting snowfall using ground-based radar sensors. They have found that radar information can be useful when there is a strong event. But, it is not as useful when a weak event is present. They also pointed out the limitations of radar information for detecting and estimating snowfall compared to rainfall. For instance, Schmid and Mathis (2004) mentioned in their paper, “Validation of methods to detect winter precipitation and retrieve precipitation type”, the deficiencies of radar information in detecting snowfall. One of the points mentioned by them was the difficulty to detect winter precipitation, especially in mountainous areas, due to shadowing and clutter effects as well as to the small height above ground of precipitating clouds. Twenty years earlier, Boucher and Wieler (1985) have mentioned the difficulty of obtaining a single relationship for all storms. They noticed that the correlation between radar reflectivity and snowfall rate vary from storm to storm due to the variation of snow crystal types and fall speeds between storms and even during the same storm. On the other hand, it is important to mention that not all countries are as well covered by radar as United States. Therefore, the necessity of a northern global product to detect snowfall is obvious. One of the challenges will be how satellite remote sensing can be successfully applied for estimating global precipitation (rainfall/snowfall) particularly over regions where traditional observation techniques such as Doppler radar do not exist or cannot provide any reliable information.

Throughout the years, passive microwave information has been used in a wide range of applications. For instance, several passive microwave sensors have been successfully used in the past decades for snow depth retrievals (Markus et al. 2006), soil moisture and vegetation properties (Tien et al. 2004), and flood forecasting (Bindlish et al. 2004). More recently, the potential of passive microwave in detecting and estimating snowfall events have been investigated (Skofronick-Jackson et al. 2004; Kongoli et al. 2003; Di Michele et al. 2005).

This study aims to develop an algorithm based on an Artificial Neural Network (ANN) model for detecting and estimating snowfall from satellite-based observations. Currently, most of the research groups working on retrieving snow-related information from space are focusing on developing remote sensing based on models that identify snow covered areas and estimate snow cover depth and SWE.

In this study an attempt was made to estimate snowfall from multi-source remote sensing observations and ground based meteorological measurements using a neural network model. The ANN technique has been chosen for this project because of the complexity of the physical relationship between the satellite measurement and the observed environment as well as its easy adaptation to different types of data, multi-sensor satellite imagery, input configurations, and multi-sources data for calibration. Compared to statistical-based techniques, neural networks use their complex configuration to find the best nonlinear

function between the presented inputs and their corresponding outputs (Foody and Arora 1997).

In order to improve the output of the developed product, the contribution of additional information from different sources such as ground observations and meteorological model outputs have been tested. Ground surface information such as air temperature, cloud coverage and type, and snow cover are believed to have an effect on the occurrence and intensity of snow precipitation events.

Brightness temperature channels from the Advance Microwave Sounding Unit (AMSU) from NOAA Polar Orbiting Satellites have been used as the main input to the model. Passive microwave sensors operate at sun-synchronous orbits providing temperature and atmospheric humidity profiles at spatial resolution varying from 25 to 65 km. The passive microwave radiation penetrates cirrus clouds and responds directly to water droplets and ice particles in the precipitation layer (Qui et al. 2005). Also, information from heavy precipitating clouds such as cloud thickness and water/ice content can be inferred with microwave radiances (Scofield and Kuligowsk, 2003). Most importantly, AMSU includes high frequency channels that are shown to be more sensitive to snowfall scattering. For instance, numerous recent studies have demonstrated that high frequency channels (≥ 89 GHz) have less contribution from land surface emissivity but still respond to precipitation scattering signatures (Skofronick-Jackson et al. 2004).

1.2 Thesis Objectives

The primary intent of this thesis is to develop an algorithm for detecting snowfall events and estimating their intensity by using remote sensing information. In order to achieve this goal, five sub-objectives were formulated.

1. Investigate the potential of passive microwave data from the Advance Microwave Sounding Unit in detecting snowfall events.
2. Examine the capability of remote sensing information in classifying the snowfall events into different categories according to the snowfall accumulation in one hour.
3. Evaluate the performance of Neural Network ensemble approach for detecting and estimating snowfall.
4. Evaluate the effect of both land cover such as snow cover and atmospheric conditions on the retrieval accuracy.
5. Compare the developed product with existing passive microwave based techniques.

1.3 Thesis Motivation

Throughout the years, snowfall timing and intensity have had an important effect on human lives and in weather related issues. A small change in temporal, spatial and

intensity distribution of snowfall events can produce big changes in global climate change and variation. Therefore, the necessity of a global snowfall product is evident. Most importantly, having a clear picture of the extent of any specific storm is undoubtedly vital information. Snowfall affects everybody directly or indirectly. For instance, snowfall events affect state and local economies, energy industry, tourism industry, business and transportation. Having an accurate estimate of snowfall can allow governments to prepare a precise cost of snow removal and reduce post-storm damage. Also, in the tourism industry, ski resorts rely on accurate snowfall and reliable information for snowmaking. When analyzing how businesses can be affected by a snow storm, three problems emerge: risk of increased traffic accidents; decline in visits to clients for meetings/sales; and increase in time to deliver merchandise (Yamamoto et al. 2004).

Another motivation in having satellite based snowfall product is the lack of a good gridded snowfall product and coverage, especially in Eastern Europe, Middle Assia and Former Soviet Union countries. Addditionally, as mentioned above, ground based radars have numerous limitations for detecting and estimating snowfall. Therefore, creating a new gridded snowfall product based on remote sensing information can improve and complement observations from radars.

In the future, the increasing number of satellites having passive microwave sensors, as well as merging information provided by existing satellites (European, American, Chinese, Japanese and Indian) will help to reduce the time gap interval and produce a continuous passive microwave data. Real time snowfall estimation is also important to improve prediction of runoff and floods. Snowfall estimation will allow us to improve runoff prediction and flash flood forecasting.

1.4 Thesis Statement

Passive microwave allows us to obtain temperature profiles and atmosphere humidity profiles. Its radiation penetrates cirrus clouds and responds directly to water droplets and ice particles in the precipitation layer (Qiu et al. 2005). Also, information from heavy precipitating clouds such as cloud thickness and water/ice content can be inferred from microwave radiances (Scofield and Kuligowski 2003). Therefore, microwave information could be somehow related to the presence of snowfall events.

The relationship between snowfall occurrences and measured brightness temperatures from passive microwave is complex and difficult to analyze. The Artificial Neural Network (ANN) model is a technique able to be adapted to such complex problems. ANN uses its complex configuration to find the best nonlinear function between input and output data. One of the major advantages of this technique is its easy adaptation to different types of data and input configuration. Also, ANN can easily incorporate ancillary data which would be difficult or impossible with conventional techniques (Ghedira et al. 2000).

1.5 Thesis Overview

This thesis is organized as follows: after this introduction chapter, chapter two provides information about snowfall related researches conducted in the last decades. Chapter three gives an overview of the Radio Detection and Ranging (RADAR) system and their potential and limitations in snowfall detection. Chapter four provides a review of the physics of passive microwave and the different types of sensors and frequencies present in

the current operational satellites. Chapter five describes detailed information of the Artificial Neural Network technique and its application in remote sensing. Chapter six describes the different types of data used. Also, it shows the selected study areas for training and validation. Chapter seven gives a deep detailed description of the steps followed in this study. Chapter eight presents and discusses the results obtained in this study. Finally, Chapter nine concludes this work.

2. Literature Review

2.1 Precipitation rain/snow discrimination from satellite

In the last decade, distinguishing between snow and rain precipitation has become an important issue. No matter which technique is employed, the need of a tool able to discriminate between rain and snow is evident. Specifically, the ability to retrieve rain over ground snow and ice where the signature cannot be separated from surface conditions.

A serious problem occurs in trying to separate rain from snow-covered ground. These situations occur when light rain falls over cold ground, during the spring melt season when the snow pack melts during the day but refreezes during the night, and when strong convective rain occurs (Ferraro et al. 1998). In 1996, Ferraro et al. proposed different techniques that use passive microwave data to identify the extent of rainfall, snowfall cover, deserts, and semi-arid conditions over land; and rain, sea ice, strong surface winds, and clear, calm conditions over ocean. Their research focus on presenting an analysis of various screening procedures that have been developed for use with SSM/I measurements. Four screening techniques were developed in order to generate accurate SSM/I-based rain retrievals. These techniques are: (1) quality control which deals with removal of unrealistic TBs, calibration errors and mislocated orbits, (2) data correction which deals with specials situations where certain corrections need to be applied, (3) geographic screens which deal with the proper identification of non precipitating coastline scenes, and (4) surface screens which deal with the proper removal of aircrafts due to specific surface types.

Over land, cloud Brightness Temperature (TB) is a product of emissivity (ϵ) and surface temperature. In addition, surface type and vegetation cover provided their own emissivity. In the absence of precipitation the atmosphere contribution to the satellite measured TB is 10%, and sometimes its contribution is negligible (e.g. snow cover where atmosphere is generally dry). On the other hand, over oceans, the satellite measured TB is a product of surface and atmosphere conditions. Some factors that affect emissivity over oceans are contribution due to clouds, water vapor, and precipitation (Ferraro et al. 1998). Although these techniques are used for rain retrievals, such techniques have not been applied for snowfall retrievals. In addition, these screening procedures use SSM/I measurements and this study is based on AMSU observations.

An important issue for this research is to discriminate snow from rainy pixels. A robust study made by Kongoli et al. (2003) presents an interesting work on this issue. In their research, the authors have developed an algorithm for detecting snowfall over land. The technique utilizes a combination of channels in the microwave window (23, 31, 89 and 150 GHz), water vapor (183 ± 1 , ± 3 , and ± 7 GHz) and oxygen absorption (50-60 GHz) areas. The snowfall detection algorithm applies the AMSU rain rate algorithm, which embeds surface temperature from forecast to identify rain. Then, when rain is not present, potential snowfall is identified when TB176 and TB180 drops below 255°K and limb corrected temperature at 53.6 GHz is above 245°K . To remove false alarms from potential snowfall, two filters are applied: $\text{TB176-TB180} \leq -20^{\circ}\text{K}$ and $\text{TB150-TB180} \leq -40^{\circ}\text{K}$. The retrievals are compared with NEXRAD radar precipitation and local meteorological observations.

In another study, Di Michele et al. (2005) have evaluated the information content of microwave frequencies between 5 and 200 GHz for rain, snowfall and cloud water retrievals over ocean and land surfaces using optimal estimation theory. The objective of their study was to define the optimized usage of channels in the microwave portion of the spectrum. The results suggested that the most suited frequencies for retrieving meteorological parameters are:

- (i) Rain over oceans: 15-18, 35-40, 80, 145, 118.75 ± 10 -14 GHz;
rain over land: 85-100, 135-140 GHz.
- (ii) Snow over land and oceans: 95-100, 140-150, 187 GHz.
- (iii) Clouds over oceans: 40, 80-85 GHz; clouds over land: 90-100, 135-140 GH.

Since the focus of this study is to estimate snowfall rate, the need of snow detection is primary. Hence, Di Michele et al, (2005) provide crucial information on which channels to use in detecting snow.

Another complex task is to distinguish low clouds with high precipitation potential from snow cover. Mostly, thick middle and high clouds can often be distinguished from snow cover due to the thermal contrast detectable at infrared wavelengths. Allen et al. (1989) developed and evaluated an algorithm for discriminating between clouds, snow-covered land and snow-free land in satellite image data. This technique uses multispectral information from daytime images of NOAA AVHRR channels 1, 3 and 4 at 0.63, 3.7 and 11.0 μm respectively. As result, separation of clouds from snow over land is based primarily on the derived channel 3 reflectance. Reflectance is derived from channel 3 by using the channel 4 emission temperatures to estimate and remove the channel 3 thermal emission. Snow is observed if reflectance in channel 3 is 0.02 to 0.04, land 0.03 to 0.10,

ice clouds 0.02 to 0.27 and liquid water 0.08 to 0.36. A major remaining problem is discrimination between ice clouds and snow cover due to the overlap of thin cirrus and snow signatures. A deeper study needs to be done in order to distinguish snow cover from cirrus clouds when optical data is used; a false signature can reduce the accuracy of the detection results.

2.2 Snowfall Retrieval

2.2.1 *Using Satellite Information*

Snowfall can be detected by using high frequency channels because of their low sensitivity to the variation in land surface emissivity. In addition, light precipitation can also be retrieved using these channels given the fact that they are susceptible to smaller-grained ice particles (Skofronick-Jackson et al. 2004).

Moreover, Skofronick-Jackson et al. (2004) described the methodology and results of physically based retrievals of snowfall over land surfaces. This study was developed for heavy storms occurring during spring season. Brightness temperatures vary depending on time and space. Therefore, the need of a snowfall algorithm for winter time is obvious. Their research utilized vertical distributions of snow, temperature and relative humidity profiles derived from the Mesoscale Model (MM5) cloud model. This data was applied and modified in a radiative transfer model by deriving brightness temperatures consistent with the AMSU-B observations. The retrieval of snowfall distribution was validated with radar reflectivity measurements obtained from a ground-based radar network. The

objective of their study was to derive characteristics of snow whose electromagnetic properties are consistent with microwave brightness temperatures at several frequencies provided by the AMSU-B sensors. The radiative transfer model used in their study employs information generated from a six-parameter model of the atmosphere associated with snow storms. Three of the six parameters are allowed to vary in order to generate different snow cloud and surface conditions. The other three parameters are set to fixed values based on statistics from a cloud resolving model and external measurements. The first and second of the three variable parameters are used to adjust distribution of a vertical structure of snow mass and relative humidity generated by PSU-NCAR MM5. The third adjustable parameter defines the fraction of snow in the ground cover and generates a composite surface emissivity using previously reported emissivities of snow and soil covered surfaces. The three fixed parameters include height levels, temperature profiles, and snow size. The fixed and variable parameters produce snow characteristics used to generate a database of brightness temperatures that would be observed at the AMSU-B frequencies using forward radiative transfer calculations. The optimal estimate of the snow parameters were derived from the best match between computed and measured brightness temperature at all AMSU-B channels. Finally, the retrieved snowfall is compared to the radar reflectivity measurements provided by the National Weather Service operational radars to estimate the validity of the retrieval. The development of an artificial network algorithm for snowfall may improve the snowfall estimates in comparison with a physical model.

Different other techniques have been applied to estimate snowfall. Noh et al. (2005) developed a statistically-based algorithm for snowfall retrieval based on Bayes' theorem using high frequency microwave information. Data from airborne and surface-based radars were used to construct an a-priori database of snowfall profiles. These profiles were used as input to a forward radioactive model to obtain brightness temperatures at high microwave frequencies. The developed algorithm was first verified by airborne microwave and radar observations and then is applied to the AMSU B. As expected, the results showed that the algorithm performs better for dry and heavy snow cases, but it is less accurate for wet and weak snow cases.

In conclusion, the expectation of new instruments and satellites with higher spatial and temporal resolutions may help scientists to improve the results they have acquired until today.

2.2.2 Using Ground-Based Radar Information

As satellite information has been used for retrieving snowfall, several researchers have tried to estimate snowfall using radar-based information. For instance, Fujyoshi et al. (1990), Carlson and Marshall (1972), and Boucher and Wieler (1985) have found different power law relationships between radar reflectivity (Z) and snowfall rate (R). Even though they were able to find accurate estimates of snowfall, one of the disadvantages of these relationships is that they are based and fitted for specific storms. Therefore, these relationships may not work in other storms. Also, as Boucher and Wieler (1985) mentioned in their paper, one single regression line cannot be applied over the entire range

of snowfall rates with the same confidence. However, Z-R relationship may be tailored to fit a limited range of snowfall rates better than others.

Moreover, some other researchers found advantages and disadvantages when using radar information. For instance, Schmid and Mathis (2004) pointed out radars as a key instrument in detecting precipitation. But, they also found some deficiencies when using radar information such as the inability to identify precipitation type (e.g. rain or snow), and the difficulty to detect winter precipitation, especially in mountainous areas. Additional information about the lack of radars for detecting and estimating snowfall is described in chapter 3.

In this study, remote sensing information is used to detect and estimate snowfall to avoid some of the radar problems mentioned in this thesis. At the same time, a product developed from satellites can complement radar-based observations.

2.3 Neural Network Ensemble Approach

Neural Networks have become a technique of interest over the last decades. This technique has been successfully applied across an extraordinary range of problem domains, in areas as diverse as finance, medicine, engineering, geology and physics. Actually, neural networks are used when there are problems of prediction or classification.

A Neural Networks uses its complex configuration to find the best nonlinear function between the input and the output data without the constraint of linearity or pre-specified

non-linearity, which is required in regression analysis (Ghedira et al. 2000). One of the advantages of this technique is its speed in obtaining results.

For a good generalization of Neural Networks, two aspects must be taken into account: the care in collecting and processing data and the architecture. These models are very sensitive to data used to predict the target. Small changes in the data set can produce a high variability in the predicted target. Therefore, the data to use should be carefully processed. In some cases, as in this study, even being careful in the data collection and processing, the data can still be very noisy. Therefore, Neural Network ensemble approach is introduced.

Like the individual neural networks, the neural network ensemble approach is used in a wide range of fields. Ensemble of Neural Networks consists of a set of individually trained neural networks (Opitz and Maclin 1999). They are combined when applied to a new input data to improve the generalization ability. There are different ways to apply the ensemble approach. Training each neural network (a) with different data sets, (b) with different topology (different number of hidden neurons), or (c) with the same architecture but with different initial conditions for the neural network weights (Fox-Rabinovitz et al. 2006).

3. Radio Detection and Ranging (RADAR)

3.1 History

The word RADAR is derived from RAdio Detection And Ranging. This active remote sensing technique was initially developed and used during World War II for detecting aircrafts. But, in the meantime it was discovered that radar is also sensitive to small particles in the atmosphere such as rain drops were it was difficult to detect an aircraft from radar when rain was present. Therefore, in 1941 the radar was officially used to detect precipitation (McNoldy 2003). Afterwards, multiple researches using this technique in different meteorological fields were and numerous advancements have been accomplished in wide array of applications. Actually, there are over 150 public access radars operated by the National Weather Service and also there are several radars that are operated and owned by private companies.



Figure 1: WSR-57 radar antenna

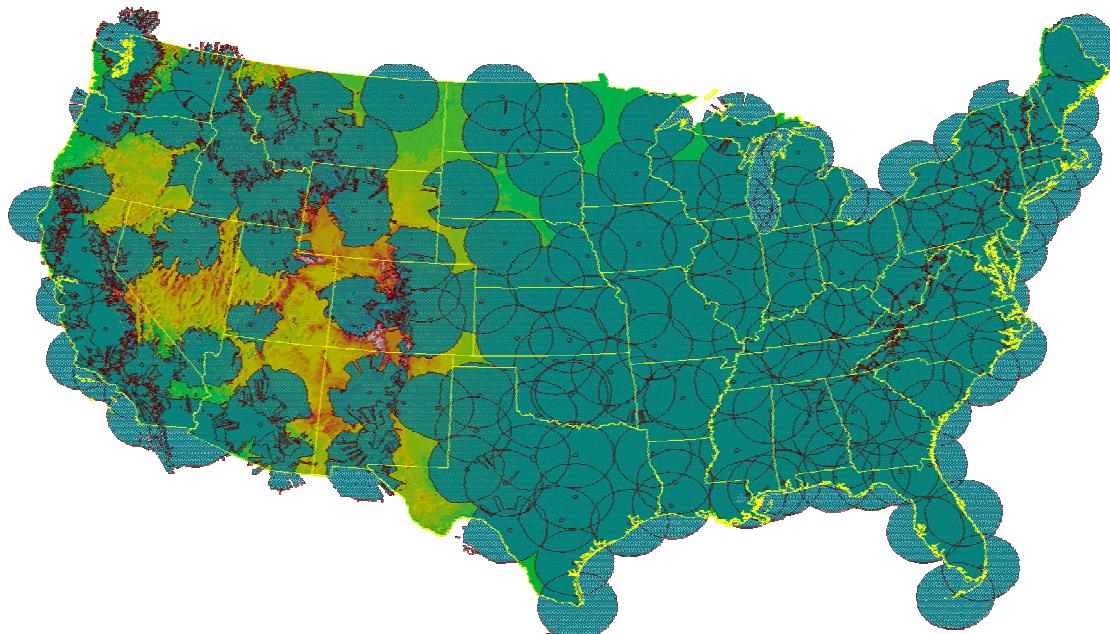
The first generation of radars was the WSR-57 (Weather Surveillance Radar -1957). This generation was built using World War II technology and designed for the national warning network (Figure 1). These radars were able to provide only coarse reflectivity data and no velocity data, which made it extremely difficult to predict tornadoes. The new generation of radars is the WSR-88D (Weather Surveillance Radar -1988 Doppler). This generation has some improvement over the WSR-57 as following:

- **Improved Sensitivity** - This is basically a result of an increased amount of power transmitted and a greater ability to distinguish between smaller returns. The WSR-88D's ability to detect lighter amounts of precipitation has allowed the detection of very light precipitation and even subtle clear air boundaries.
- **Improved Resolution** - This is primarily a function of angular beam width. The narrower the beam, the smaller the width at a given distance. This improved characteristic allowed the WSR-88D to differentiate between objects, thereby increasing the resolution.
- **Volume Scanning** - Rather than scanning along varying azimuth angles (PPI) then stopping to scan vertically (RHI), the radar automatically scan various elevation angels while spinning around 360° of azimuth. Computers generate products based on this volume scan.
- **Enhanced Capabilities and Algorithms** - Sophisticated computer programs have assisted radar operator in detecting various phenomena such as mesocyclones and tornadoes (Tornado Vortex Signature - TVS) and the like. The radar also has a greater range of reflectivities operating in severe and non-precipitation modes.

3.2 Doppler Radar

This special type of radar allows to measure the distance and the velocity of the particle. The way it works is by comparing the frequency of the transmitted signal versus the frequency of the reflected signal. The speed is relative to the radar location. For instance, if any particle flies in a perfect circle around a radar site at the center, the radar will show a velocity of zero because the relative distance between the particle and the radar site would not change.

The WSR-88D network is composed by 158 Doppler radars distributed all over the United States, including the U.S. Territory of Guam and the Commonwealth of Puerto Rico, leaving few gaps (Figure 2). The most important areas are better covered by these radars than those areas that are less important which are covered for one or no one are all.



Source: Maddox, et al. Weather and Forecasting, 2002

Figure 2: Radar covered area

3.3 Description of the Antenna

The radar antenna has a diameter of 28'. It is positioned at the top of a tall tower to elevate it off the ground to minimize ground clutter. This antenna is enclosed in a fiberglass spherical protective dome of 39' in diameter (called radome) (Figure 3).



Figure 3: WSR-88D radar antenna

3.4 How Radar Works

As the radar turns, it transmits a high power beam of radiation, called pulses and receives a small amount of radiation back from whatever the beam hit into. The emitted beam has three fundamental properties:

- Pulse Repetition Frequency (PRF): is the number of pulses of radiation are transmitted per second. The pulses are around 1300 for typical WSR-88 D radar.
- Transmission time: is the duration of each pulse. The transmitted beam travels through the atmosphere at the speed of light. Therefore, the pulse length can be calculated from the transmission time. Usually this length is on average 1 Km.
- Beam width: describes the angular width of the emitted beam, this width is typically about 1° .

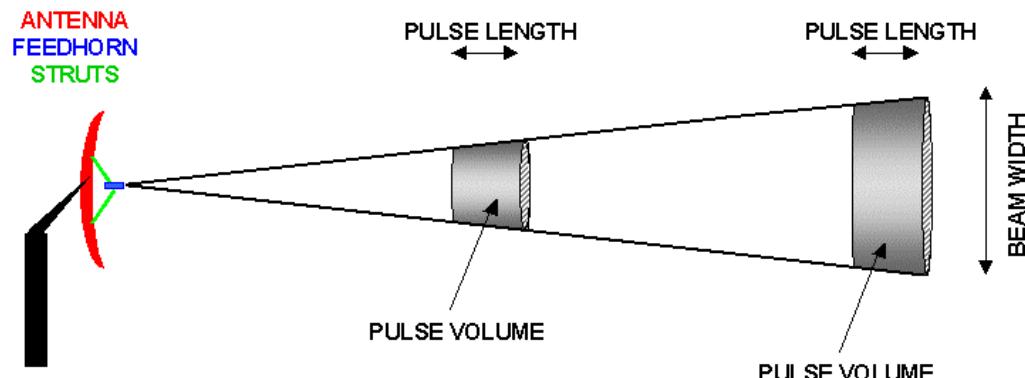


Figure 4: Radar beam characteristics

The pulse volume can be calculated by using the pulse length and the beam width (Figure 4). Sometimes this volume can be enormous, which means that consecutive pulses will receive backscattered radiation from a large number of targets. The radial and the angular resolutions are defined by the pulse length and the beam width respectively.

The radar spends more time listening or receiving (99.9% of time) than transmitting (0.1% of time). More specifically, it might emit a pulse for 0.0000016 seconds then measure the backscattering signal for 0.00019 seconds.

The antenna whirl around making a full 360° scan in just 10 seconds, then it adjusts itself to the next higher elevation scan and makes another 360° full scan. This step is repeated several times using different elevations and when it finish the cycle is repeated. There are up to 20 degrees in elevation allowing the radar to cover a huge volume of the atmosphere.

3.5 Radar Operational Modes

3.5.1 *Clean Air Mode:*

This is the normal mode of operation and is used when there is no significant precipitation in the area. This mode includes five scan angles which range from 0.5° to 4.5° (Figure 5). A full Volume Coverage Pattern (VCP) is complete in ten minutes. In this mode the radar rotates slower and performs fewer scan angles which allows a high resolution.

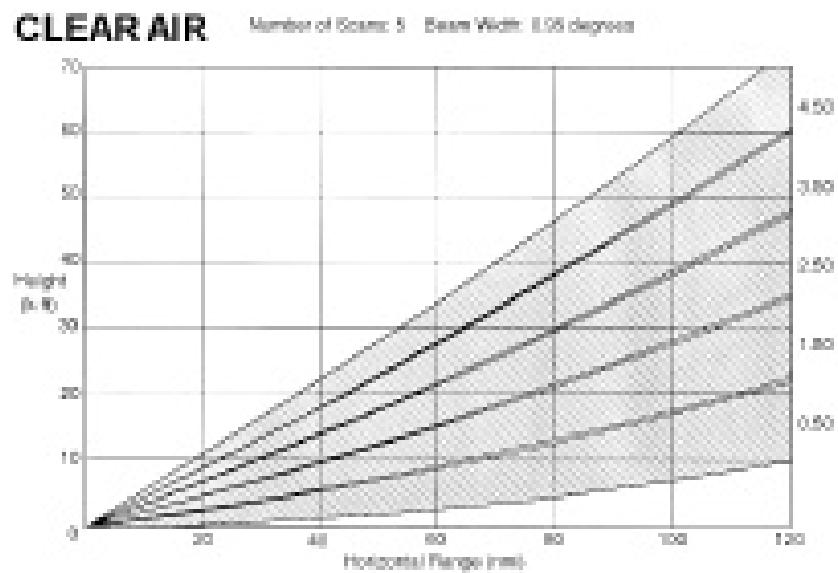


Figure 5: Clean air mode

Therefore, this higher resolution makes this mode more sensitive and allows detecting smoke plumes and large aerosol particles in the atmosphere such as insects, birds and light precipitation drops (rain or snow). The high sensitivity allows the detection of snow showers since snow generally reflects much less energy than other forms of precipitation.

3.5.2 *Precipitation Mode:*

When the radar detects significant precipitation in the area, it automatically switches to precipitation mode. In this mode the resolution is lower due to the radar rotating faster than the clean air mode. Nine angles are performed by this mode. They range from 0.5° to 19.5° (Figure 6). This mode spends shorter time listening for returned signals at the different angles. Therefore, this mode completes a full VCP scan in six minutes. This means it spends less time than the clean air mode making this mode less sensitive.

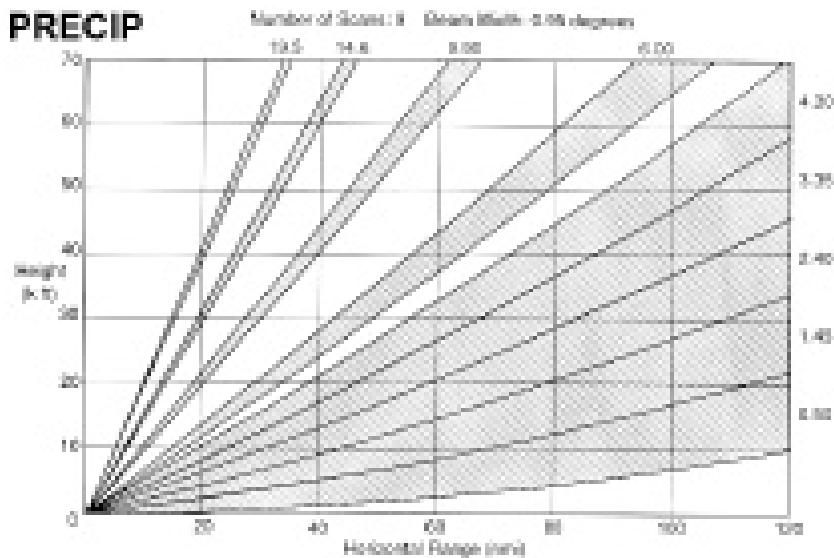


Figure 6: Precipitation mode

There is a special mode for precipitation called *Severe Mode*. In this mode more scans are made at higher angles to capture the full structure of the towering thunderstorms and the rotation rate is increased. The fourteen angles are composed by the same precipitation mode angles and the rest fill the gaps between them (Figure 7). A full VCP scan is complete in five minutes. The final objective is no to get a high resolution, it is to quickly update and complete coverage of the entire storm.

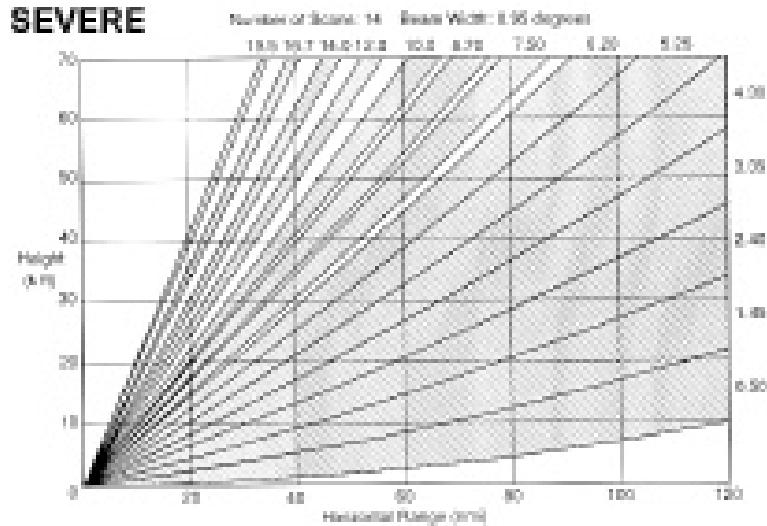


Figure 7: Severe mode

3.6 Reflectivity

Valuable information is taken from the radar as the radar transmits the pulse and analyzes the energy reflected back. The amount of power received by the radar is called *Reflectivity* (z). The reflectivity is computed by using the following empirical equation:

$$dBz = 10 * \log(z/z_0)$$

3-1

where z is the reflectivity factor and z_0 is a constant defined as $1\text{mm}^6/\text{m}^3$. dBZ denotes reflectivity in decibels.

3.7 Non-Precipitation Echoes and Clutter

In some cases weak energy or echo is reflected back to the radar, even when precipitation is not present. It is normal to catch these weak signals when the radar operates in clean air mode which is very sensitive. In general, three types of non-precipitation echoes exist: atmospheric effects, ground clutter and false echoes.

- **Atmospheric Effects:** this refers to small amount of energy the radar receives back from different things in the atmosphere such as cloud, smoke and fog in a clear day. Also, some atmospheric effects can produce small amount of energy to reflect back to the radar such as inversion layers and the variation in air density produced by the variation of temperature.
- **Ground Clutter:** this is another class of non-precipitation echoes. Ground clutter refers to obstacles located on the ground such as buildings, antenna towers and mountains. Also, it includes some objects that are not located on ground such as insects, aircrafts and birds.
- **False Echoes:** this class of non-precipitation echoes refers to variation on the air density that causes the radar signal to refract into the earth.

3.8 Radar Map Projections

The radar images provided from the National Weather Service (NWS) are in an unprojected latitude/longitude except for the state of Alaska making them look oval in shape (or squash) (Figure 8). This projection allows ingesting NWS radar data for display as additional layer with other information such as population density.

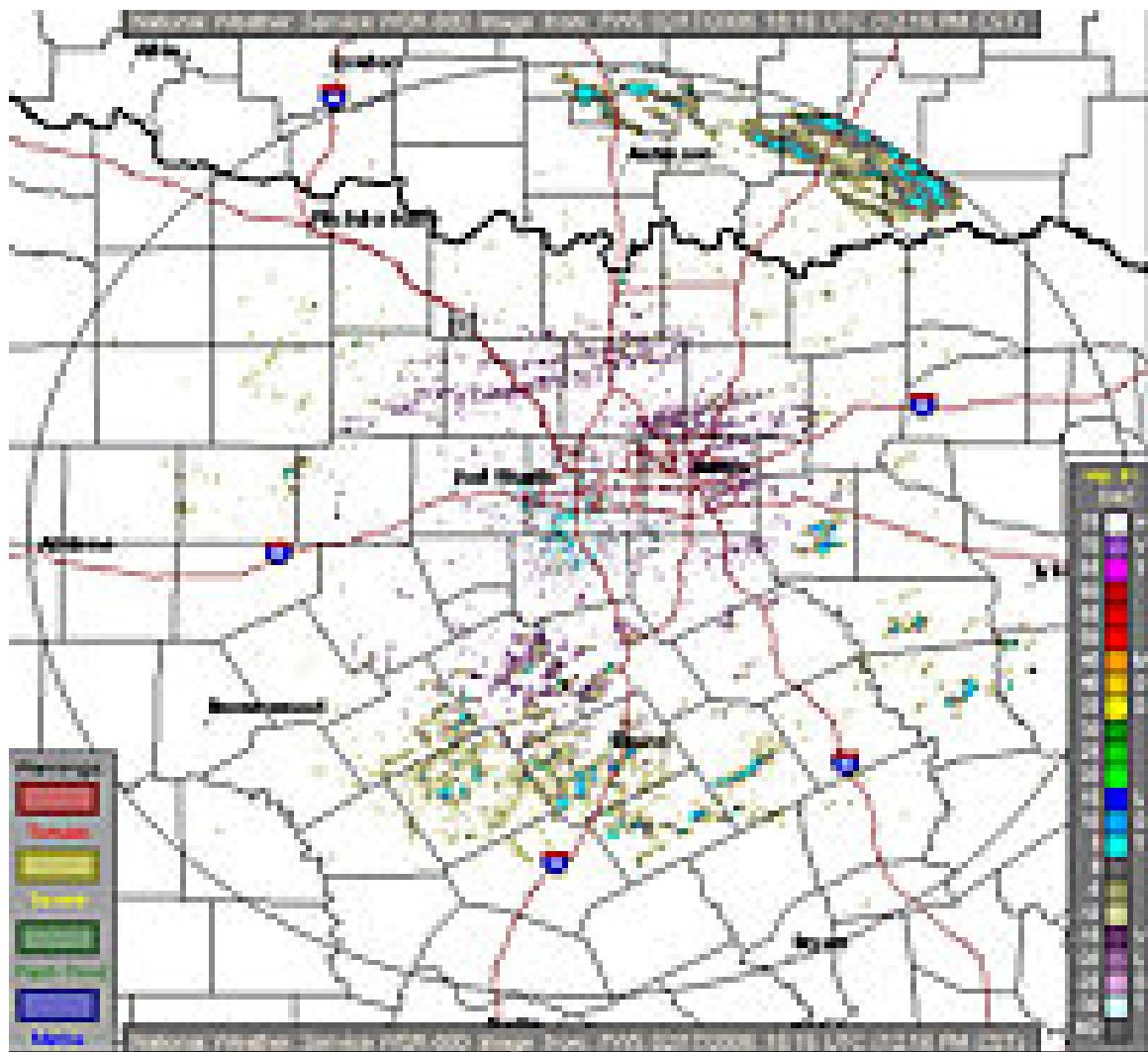


Figure 8: Radar image

3.9 Types of Doppler Radar Images

The radar images use different colors to display the information collected by radars. There are several types of images such as reflectivity, velocity and rainfall information.

3.9.1 Reflectivity Images:

These images are painted as the weather from the energy reflected back from the radar.

There are two types of reflectivity images: Base reflectivity and Composite reflectivity.

- **Base Reflectivity:** these images are taken from the lowest elevation (0.5°). Therefore, a base reflectivity image becomes available after the completion of the 0.5° elevation scan angle during each VCP scan (Figure 9). This is the default image used to see the first sign of what is out there.

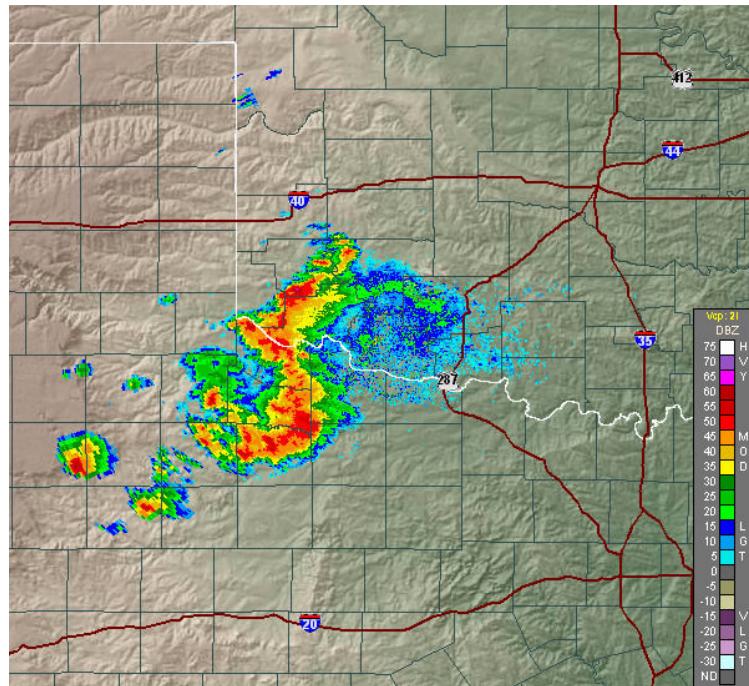


Figure 9: Base reflectivity image

There are two classes of base reflectivity images: the *short range version* which extends out to 124 nm (about 143 miles) and the *long range version* which extends out to 248nm (about 248 miles).

- **Composite Reflectivity:** to produce this image all the scan angles or elevation are used (Figure 10). This image is composed of the greatest reflectivity seen from the radar at any angle or elevation. This image gives a huge advantage compared to the base reflectivity images especially in the presence of mountains in the covered area. Since this type of image utilizes all the scan angles, it is easier to see the weather at the peak of the mountains using this image while it is not possible to see it with the base reflectivity image.

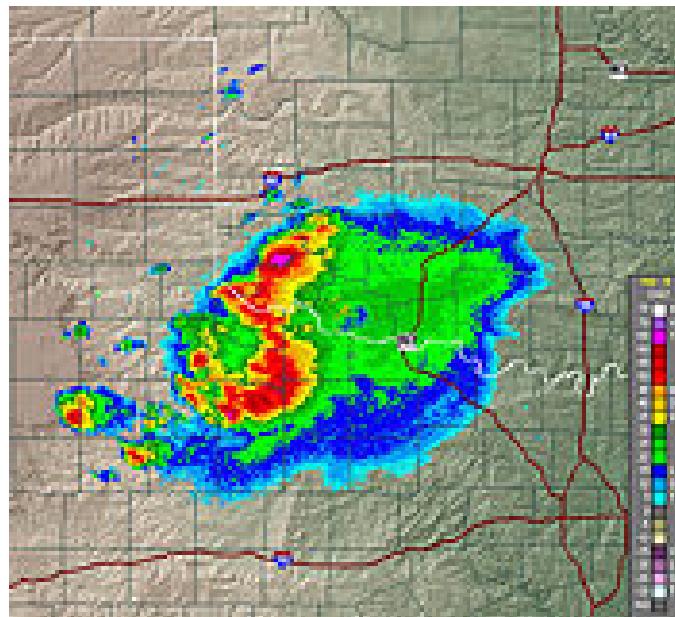


Figure 10: Composite reflectivity image

3.9.2 Velocity Images:

The biggest advantage of the WSR-88D is the ability to detect motion. Although the limitation is that it can see motion only toward or away from the radar, it is really important to know its location. This motion is called radial velocity. The images transmit their information using different colors. In this type of images the red color indicates wind moving away from the radar and the green color represent wind moving toward the radar.

- **Base Velocity Images:** these images provide the basic wind field from the lowest scan angle (0.5°) (Figure 11). These images are key for detecting the velocity of cold fronts and areas of strong winds. The limitation of using this radial velocity is that the strength of the wind will be underestimated unless the wind is moving directly toward or away from the radar.

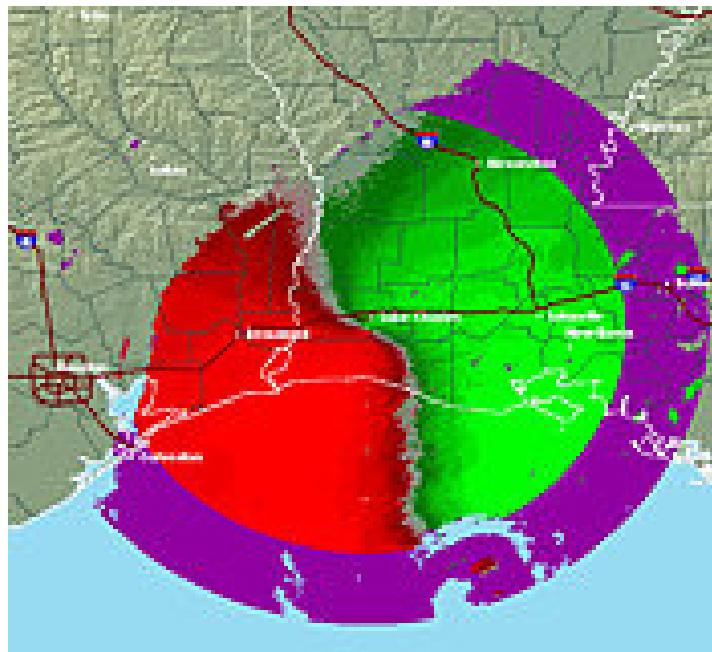


Figure 11: Base velocity image

The wind velocity over surface areas is only provided for areas near the radar. As the distance from the radar increases, the height from surface also increases.

- **Storm Relative Motion Images:** these images are very functional when looking for small circulations in thunderstorms (Figure 12). Sometimes, tornadoes might form in these small circulations. The difference between storm relative motion and base velocity is that the motion of storms is subtracted from the overall flow of the wind. As storms move, small circulation can be hidden by their motion. The motion is removed to see the view of the wind relative to the storm.

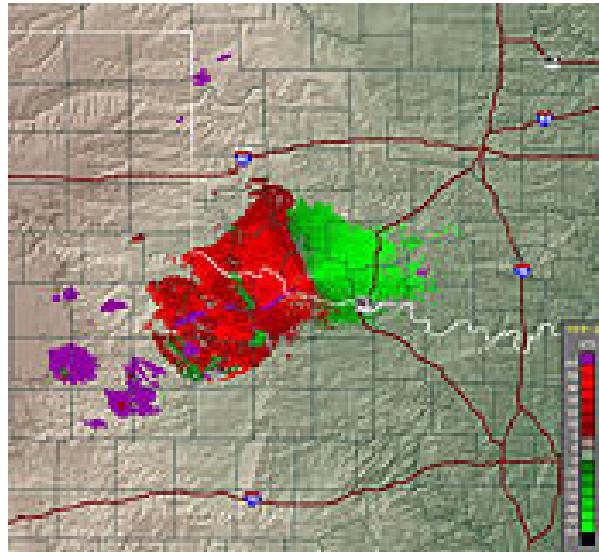


Figure 12: Storm relative motion image

3.9.3 *Precipitation Images:*

Two types of precipitation images are provided: *One Hour Precipitation* and *storm Total Precipitation*. Both types of images give information up to a distance of about 143 miles

from the radar. If precipitation information is needed at a greater distance, adjacent radar should be used to retrieve this data.

- **One-Hour Precipitation:** this type of images illustrates the estimated one hour precipitation accumulation (Figure 13). These images are used to assess rainfall intensities for flash flood warnings, urban flood statements and special weather statements.

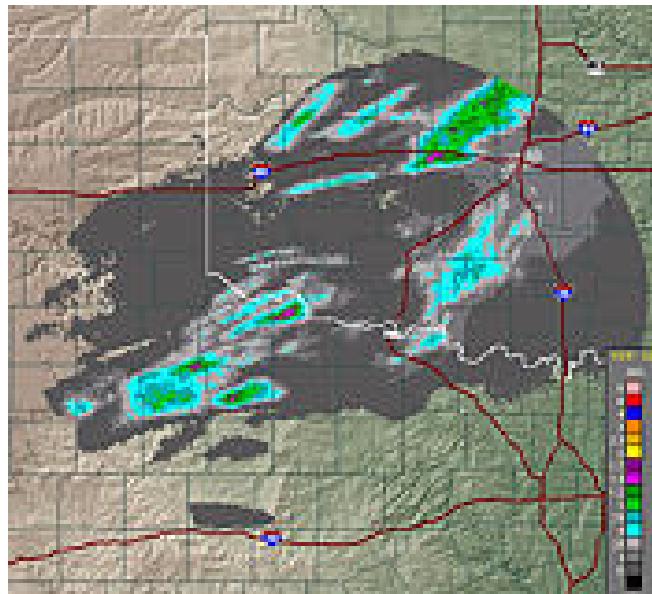


Figure 13: One hour precipitation image

- **Storm Total Precipitation:** this type of images illustrates the estimated accumulated rainfall continuously updated since the last one hour break precipitation (Figure 14). These images are very useful for detecting flood potential over urban or rural areas, to

estimate the total basin runoff and to provide rainfall accumulation for the duration of the event.

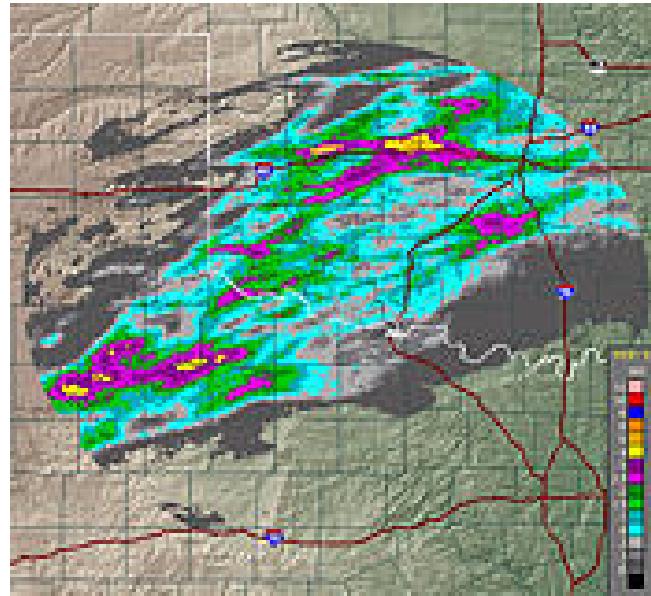


Figure 14: Storm total precipitation image

3.9.4 Weather Warnings:

These images highlight the areas that might be affected by severe weather (Figure 15).

Usually these areas are shown in these images inside a polygon. There are four types of warnings and each one is represented by a color.

- **Red:** Tornado warning. Issued when a tornado is imminent or occurring.
- **Yellow:** Severe thunderstorm warning. Issued when a severe thunderstorm is imminent or occurring. A thunderstorm is defined as severe when hail greater than $\frac{3}{4}$ " are present and/or when wind speed exceeds 58 mph.

- **Green:** Flash flood warning. Issued when a flash flooding is imminent or occurring.
- **Marine:** Special marine warning. Issued for hazardous weather condition (thunderstorms over water, cold air funnels over water, or waterspouts) usually of short duration (2 hours or less) and producing sustained winds or frequent gust of 34 knots or more that is not covered by existing marine warnings.

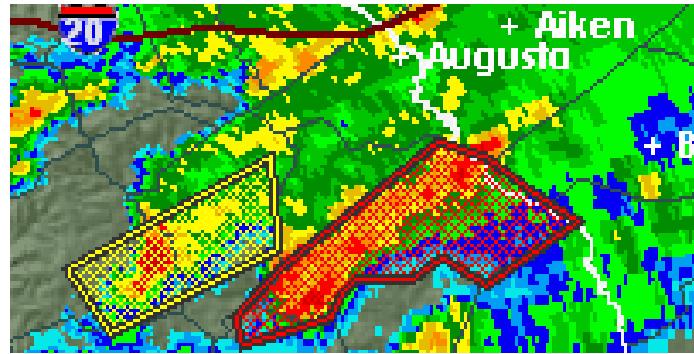


Figure 15: Weather warnings image

3.10 Relationship between Reflectivity and Precipitation

The WSR-88D radar is very sensitive and it is able to detect low reflections such as -28 dBz. Strong precipitation events usually reflects higher energy back to the radar and its reflection is about 15 dBz or more. Precipitation and reflection follow a linear relationship. As stronger is the storm event higher is the energy reflected back to the radar. Reflectivity is represented by a color scale. The clean air mode (Figure 16) is more sensitive than the precipitation mode (Figure 17), therefore two scales are used.

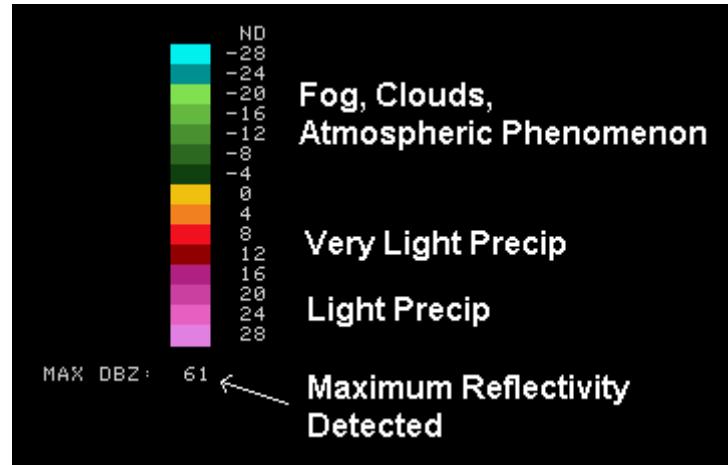


Figure 16: Clean Air mode scale

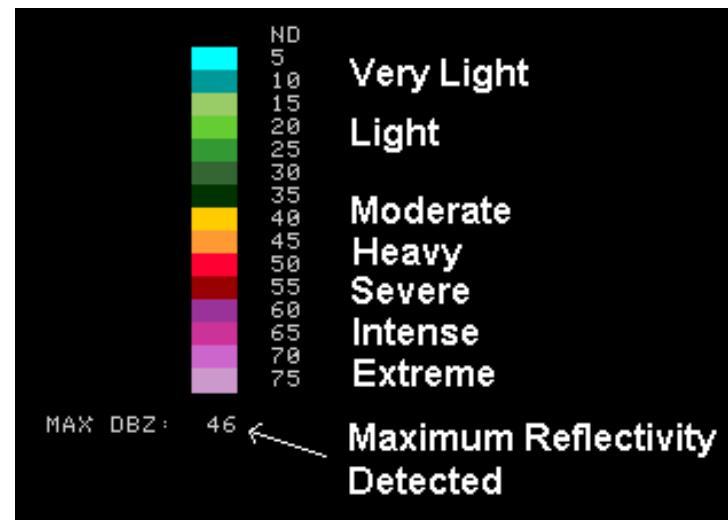


Figure 17: Precipitation mode scale

3.11 Precipitation Radar Limitations

There are several conditions for what conventional and Doppler radars might produce an inaccurate precipitation estimates:

- **Precipitation falls too far from the radar:** there are two reasons why the precipitation estimates can be inaccurate when the event occurs far away from the radar. The first reason is the beam broadening. As precipitation is far from radar, the area coverage by beam increases and the range also increases (Figure 18). Therefore, the beams became too large compared to the precipitation drop. The computations to estimate precipitation are base on targets that fill the beam completely, not allowing the radar to detect these particles. The second reason is overshooting. As the radar beam travels away from radar, the height of the radar beam increases as the range increases. The lowest angle (0.5°) can overshoot.

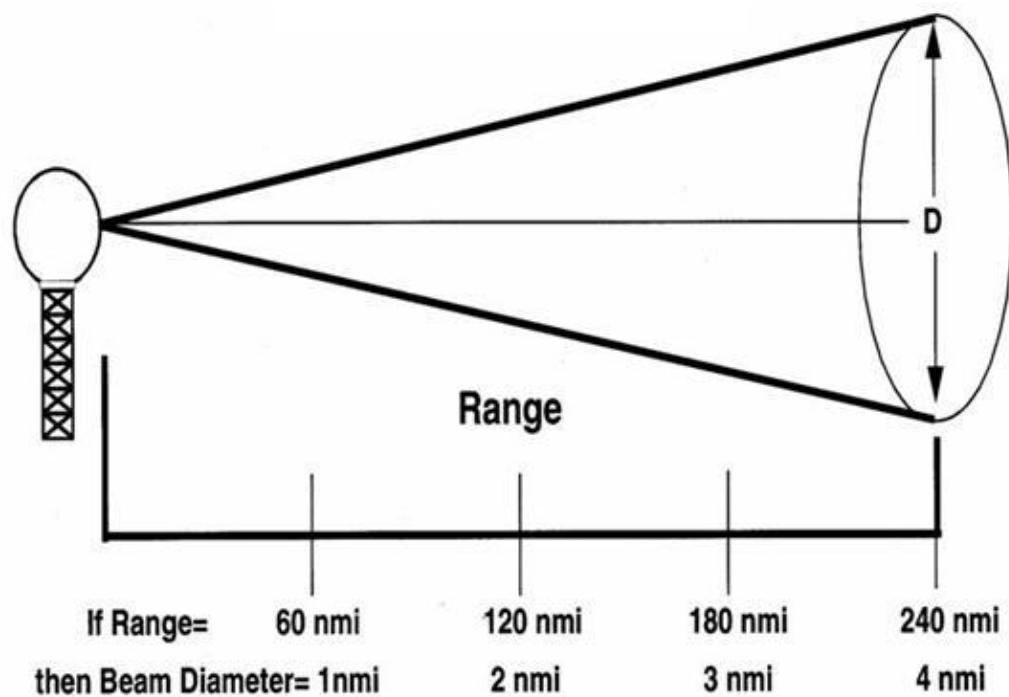


Figure 18: Radar beam diameter

- **Precipitation Blocked from Radar View:** this limitation can occur when a storm event is located in mountainous area (Figure 19). The radar beam would not be able to reach the precipitation particles at the other side of the mountain causing underestimation of the storm event. Also, at low scan angles the radar beam easily intercepts the terrain. This factor reduces the radar coverage, not allowing the radar to reach or detect precipitation.

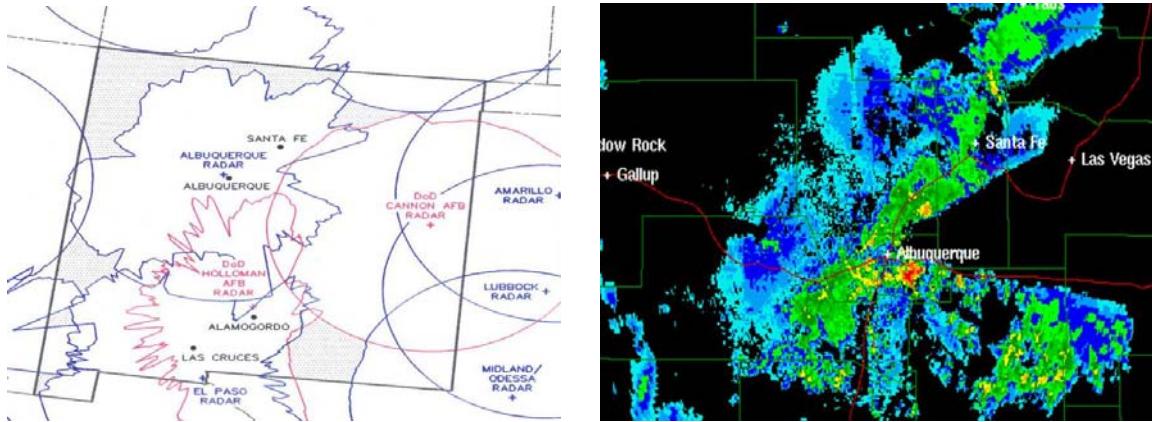


Figure 19: Images of precipitation blocked from radar view

- **Precipitation and Reflectivity Conversion Errors:** some equations are used to retrieve precipitation. In some cases these equation may overestimate or underestimate precipitation. The relationship between precipitation and reflectivity depends on drop-size distribution. The size of the precipitation drop varies for each storm and also during each storm. Mathematically, reflectivity is equal to the 6th power of droplet diameter. The drop diameter looks different to the radar depending on the drop composition. For instance, liquid water drops coat hail, which is less dense than water,

causing it look to the radar with a bigger diameter. Therefore, the equations tend to overestimate precipitation. Also, wet melting snow and sleet drops, both less dense than water stick together near the freezing level looking to the radar with a bigger diameter. Then, precipitation will be also overestimated.

- **Below Beam Effects:** below the beam several factors can produce overestimate or underestimate precipitation (Figure 20). For instance, sometimes it is possible to overestimate precipitation in one location and underestimate precipitation in another one. This can be caused by strong horizontal winds below the lowest beam that blow precipitation from one range bin to the next. As well, overestimated precipitation can be caused by the evaporation of precipitation below the cloud base. On the other hand, 65% of all precipitation events seen on the ground remain undetected by the radar. Also, coalescence of rain drops below the lowest radar beam causes precipitation underestimates.



Figure 20: Shower evaporation below cloud

- **Ground Clutter and Superrefraction:** sometimes power returned from non-weather targets causes precipitation overestimation. An advantage of WSR-88D algorithms is they tend to identify anomalously high reflectivities and remove them before calculating the precipitation rate. Sometimes the radar beam bends downward more than the normal, this is called superrefraction. Superrefraction causes precipitation overestimation when the beam bends far enough to hit the ground (Figure 21).

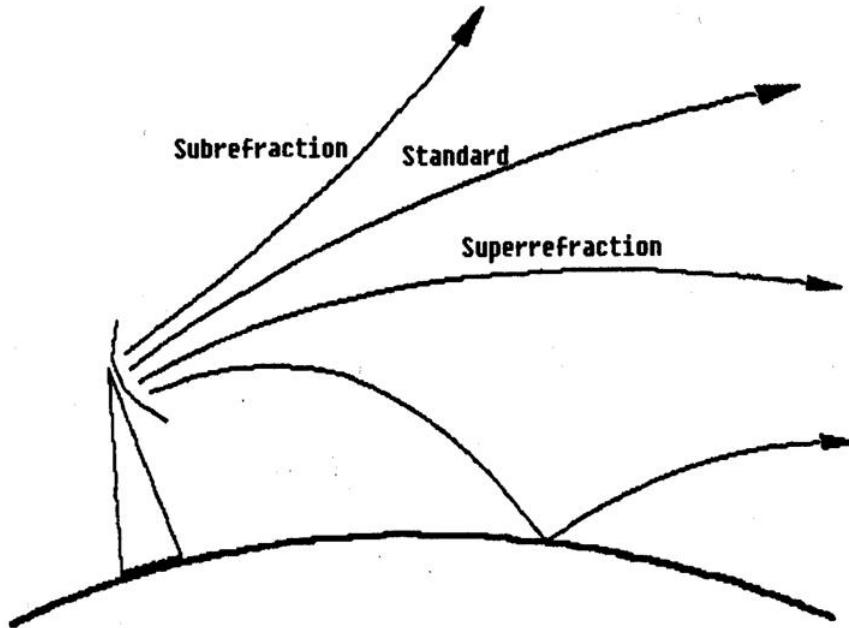


Figure 21: Radar beam propagation paths

- **Hardware Problems:** several internal or external system noises can cause reflectivity to differ from the optimal calibration. These noises can have an effect on the precipitation estimates. For instance, liquid moisture on the dome around the antenna can have two effects. One is that it can attenuate the signal when it is transmitted, and

the second effect is the power of the beam transmitted can be lower therefore the energy reflected back is also lower. These two effects can cause underestimated precipitation.

3.12 The Use of Doppler Radar in Snowfall Detection

Throughout the time several researches had used radar information for solving snowfall aspects related issues such as detecting snowfall, determining snowfall rate and accumulation, and snow size spectra. Some of the research conducted using radar information for snowfall issues are: Sekhon and Srivastava 1970, Fujiyoshi et al. 1990, Carlson and Marshall 1972, Boucher and Wieler 1985, and Schmid and Mathis 2004.

Even though several researches have been conducted in snowfall detection, the results obtained from them are not totally positives. According to the research conducted by Schmid and Mathis on the “Validation of methods to detect winter precipitation and retrieve precipitation type”, the detection rate of a radar for weak snowfall is low (12%). Also, in their paper they established the deficiencies of the radar in detecting snowfall:

- The inability to identify precipitation type (e.g. rain or snow)
- The difficulty to detect winter precipitation, especially in mountainous areas, due to shadowing and clutter effects and due to the small height above ground of precipitating clouds.

In addition, Boucher and Wieler in their research on “Radar Determination of Snowfall Rate and accumulation” mention the difficulty of obtaining a single relationship for all

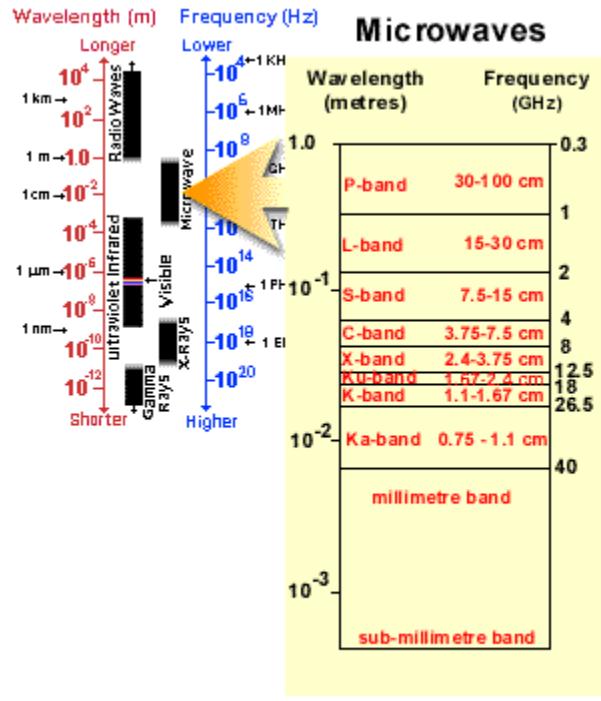
storms. They noticed that the correlation between radar reflectivity and snowfall rate vary from storm to storm due to the differences in snow crystal types and fall speeds between storms and even during the same storm.

In conclusion, snowfall detection is a difficult subject to retrieve from radar. As mentioned above, radar itself has a lot of limitations for snowfall detection retrieval such as the precipitation block from radar view (mountainous areas). In this study, snowfall detection is retrieved using satellite-based data information. Since, they take measurements from space one of the radar limitations (obstacles blocking radar view) is already solved. Also, Satellites-based global information is obtained from space. Therefore, satellites are able to get information of radar gap areas.

4. Microwave Information

Throughout the years remote sensing has become an important source of information in several research fields such as atmosphere, oceanic and land. For instance, in the land field, remote sensing information have help researches to obtain maps of land cover, to estimate geophysical and biophysical characteristics of terrain features, and to monitor changes in land cover. Actually, there are different types of satellites, sensors and frequencies. Therefore, different information from satellite is available. The information to use in any project is selected according to the purpose of the project. Since snowfall is strongly related to microwave wavelengths, this chapter focuses on the description of microwave information.

Passive and active microwave data information is collected worldwide to help in the retrieval of an accurate weather and environmental forecast. Microwave covers approximately from 1mm to 1m in the spectrum making their wavelengths longer compared to the visible and the infrared region of the spectrum (figure 22). The short microwave wavelengths have similar properties to the thermal Infrared region. The long wavelengths permit the penetration of clouds, haze and dust. Also, microwave energy emitted from the Earth surface can be detected from space in almost all weather and environmental conditions. Therefore the measurements can be collected at any time. These microwave sensors are available in space and on the ground and both of them utilize common bands to have more measurements of the Earth surface and the atmosphere.



© CCRS / CCT

Figure 22: Microwave Spectrum

Microwave sensors are able to detect variation in Earth parameters that optical wavelengths are not able to do. Basically, microwave sensors can measure with different levels of accuracy the following parameters:

- Temperature
- Humidity
- Cloud
- Trace gas profiles
- Surface soil moisture

- Sea surface temperature
- Land surface roughness
- Ocean surface wave height
- Moisture content and melt character of ice and snow
- Areas covered by snow and ice

The microwave spectrum is composed by windows and sounder channels (figure 23). The windows channels do not have any contribution from the atmosphere. On the other hand, the sounder channels take measurement of specific layers in the atmosphere.

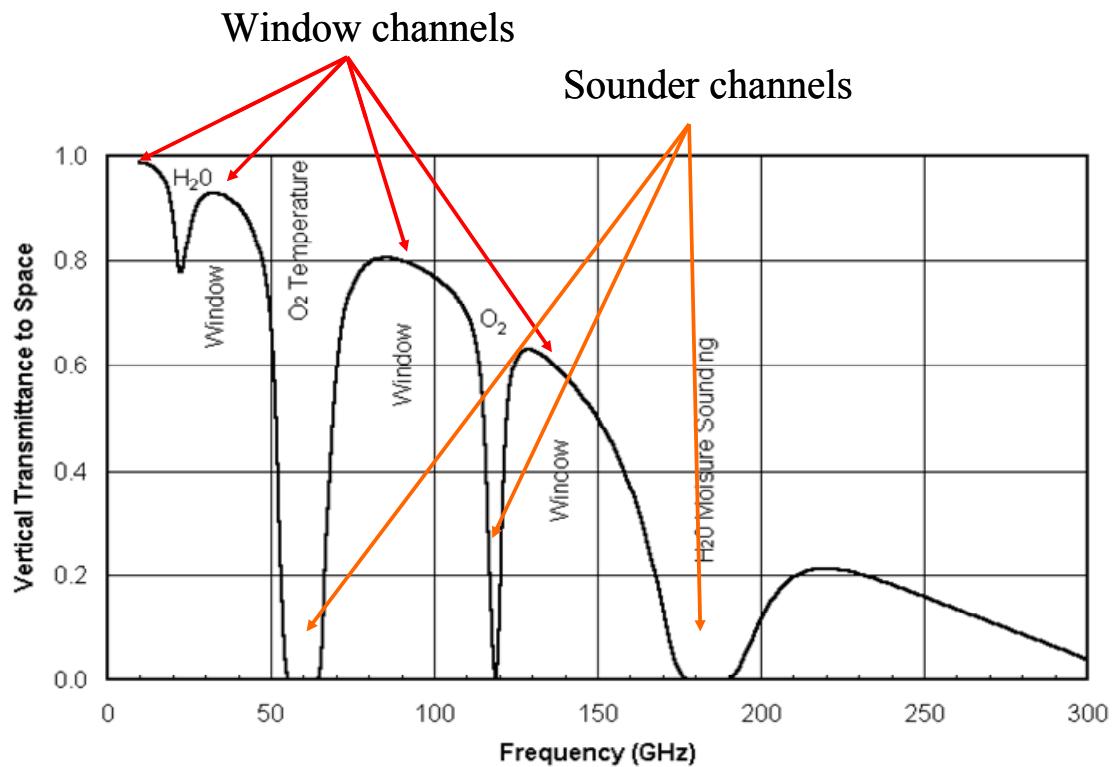


Figure 23: Windows and sounder channels

For instance, any emission captured from the three channels in the 183.1 GHz line, is measurement of the water vapor profile.

Actually, there are several operational microwave-based sensors such as the Advanced Microwave Sounding Unit (AMSU) and instruments on board of the Defense Meteorological Satellite Program's passive microwave weather satellites such as the Special Sensor Microwave/Imager (SSM/I) and Special Sensor Microwave/Temperature (SSM/T).

This chapter describes the characteristics of the passive and the active microwave remote sensing. It will focus on one of the passive microwave instruments, AMSU, due to data sets provided by this instrument are used in this project.

4.1 Active Microwave Sensors

These sensors do not use the sun as source of energy. On the other hand, they provide their own source of energy for illumination. Therefore, one of the advantages of these types of sensors is that they can obtain measurements at any time and with much better spatial resolution than passive microwave sensors. These sensors emit radiation toward the target to be analyzed and then, they measure the energy backscattered back from the obscured scene.

There are two classes of active microwave sensors: *Imaging and Non-imaging*. One of the most common *imaging* active remote sensing sensors is the RADAR (RAdio Detection And Ranging). This sensor takes a two dimensional measurements. One dimension is to

discriminate between different targets. The second dimension is to determine the distance to the target based on the time delay between the transmitted and the reflected signals. Some of the *non-imaging* active microwave remote sensing sensors are altimeters and scatterometers. These sensors take profiling devices measurements in one dimension. Radar altimeters usually are used to measure the distance from the sensor and the target. These radars transmit short microwave pulses and measure the round trip time delay to targets. For instance, radar altimeters are used on aircraft to measure altitude and also on aircraft and satellites to measure sea surface height and for topography mapping. On the other hand, scatterometers are used to measure the amount of energy reflected back from the target. This amount of energy depends on the surface properties and the angle at which the microwave energy strikes the target. For instance, this type of sensor is used to measure winds over ocean based on the sea surface roughness.

4.2 Passive Microwave Sensors

These sensors use the sun as the source of energy or radiation. They are able to measure the electromagnetic energy emitted and scattered by particles in the Earth or atmosphere. The microwave energy emitted is not strongly related to the temperature of the object. On the other hand, it is related to the moisture and physical properties of the object or surface. For instance, soil moisture can be retrieved from passive microwave due to the fact that its emission is influenced by its moisture content. Also, water emits strong energy in the microwave portion of the electromagnetic spectrum.

The energy emitted to passive microwave sensors is very small. Therefore, the measurements of radiation should be taken over large areas. As result, most passive microwave sensors have low spatial resolution.

The passive microwave energy can be collected in different ways (figure 24). (1) Emitted by the atmosphere, (2) Reflected from the surface, (3) Emitted from the surface, or (4) Transmitted from the subsurface.

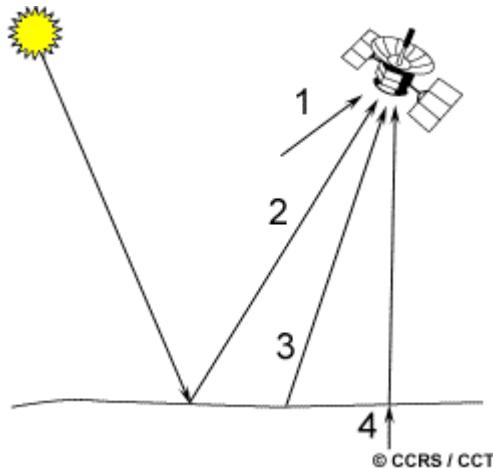


Figure 24: Forms of passive microwave energy emissions

Passive microwave have a wide range of applications in meteorology, hydrology and oceanography. From these sensors several subjects can be detected and measured. Some of these subjects are atmospheric profiles, water and ozone content in the atmosphere, soil moisture, detection of pollutants and mapping sea ice.

The following section focus on one of the passive microwave remote sensing instruments, Advanced Microwave Sounding unit (AMSU). Data set provided by this instrument is used in this project.

4.3 Advanced Microwave Sounding Unit (AMSU)

The Advanced Microwave Sounding Unit (AMSU) instrument is part of the National Oceanic and Atmospheric Administration (NOAA) satellites (NOAA 15, 16 and 17), which are part of the National Polar orbiting Operational environmental satellite system (NPOESS). These polar satellites (figure 25) are able to monitor the entire Earth and collect data on a daily basis, tracking atmospheric variables and providing atmospheric data and Earth's surface images for different applications.

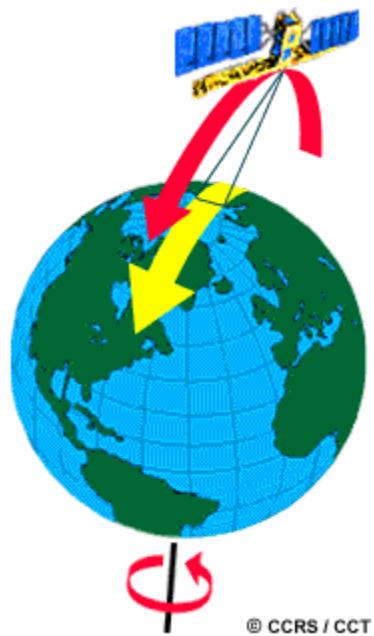


Figure 25: Polar orbiting environmental satellite

The first spacecraft to fly AMSU was NOAA 15 which was launched on May 13, 1998. NOAA 15 flies approximately at 807 Km altitude with an inclination angle of 98.5°. NOAA 15 has a sun synchronous orbit with an orbital period of 101.2 minutes. Following

NOAA 15, NOAA 16, 17 and 18 were launched with similar specifications; they are described in table 1.

Table 1: Satellites Specifications

	Launch date	Approx. Altitude (Km)	Inclination Angle
NOAA 15	May 13, 1998	807	98.5°
NOAA 16	September 21, 2000	849	99.0°
NOAA 17	June 24, 2002	810	98.7°
NOAA 18	May 20, 2005	854	98.74°

Polar Operational Environmental Satellite (POES) consists of two instruments AMSU-A and AMSU-B. A deep description of these two instruments is given in the following section.

4.3.1 Advanced Microwave Sounding Unit-A

AMSU-A is a fifteen-channel microwave radiometer which provides atmospheric information in presence of clouds. These measurements can be used to enhance and/or to correct atmospheric measurements taken from infrared sensor when clouds are present. This instrument permits the calculation of the vertical temperature and moisture profiles for up to 45 Km from the earth's surface. AMSU-A instrument parameters and characteristics are summarized in table 2 (NOAA 2004).

Table 2: AMSU-A Instrument Parameters

Chan. #	Frequency (GHz)	# of Bands	Nominal Bandwidth (MHz)	Nominal Beamwidth (degrees)	NEΔT (K)	Polarization at nadir
1	23.800	1	270	3.3	0.30	V
2	31.400	1	180	3.3	0.30	V
3	50.300	1	180	3.3	0.40	V
4	52.800	1	400	3.3	0.25	V
5	53.596±0.115	2	170	3.3	0.25	H
6	54.400	1	400	3.3	0.25	H
7	54.940	1	400	3.3	0.25	V
8	55.500	1	330	3.3	0.25	H
9	$f_0=57.290.344$	1	330	3.3	0.25	H
10	$F_0\pm 0.217$	2	78	3.3	0.40	H
11	$f_0\pm 0.3222\pm 0.048$	4	36	3.3	0.40	H
12	$F_0\pm 0.3222\pm 0.0022$	4	16	3.3	0.60	H
13	$F_0\pm 0.3222\pm 0.0010$	4	8	3.3	0.80	H
14	$F_0\pm 0.3222\pm 0.0045$	4	3	3.3	1.20	H
15	89.000	1	<6,000	3.3	0.50	V

The Advanced Microwave Sounding Unit-A (AMSU-A) contains modules A1 and A2 (figure 26). Each module has separate spacecraft interfaces.

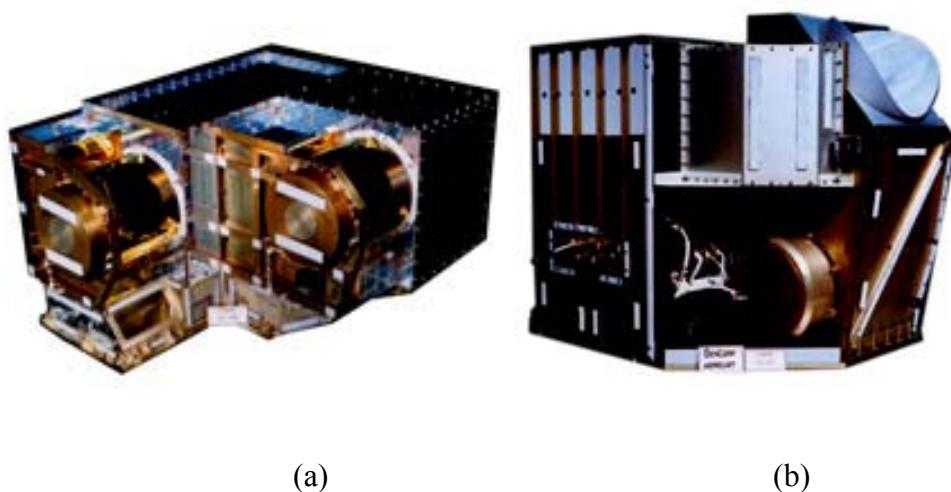


Figure 26: (a) AMSU-A1 (b) AMSU-A2

Module A1 contains channels 3 to 14 in the range of 50.3 to 57.29 GHz and channel 15 at 89.0 GHz, while module A2 contains channels 1 and 2 at 23.8 and 31.4 GHz respectively.

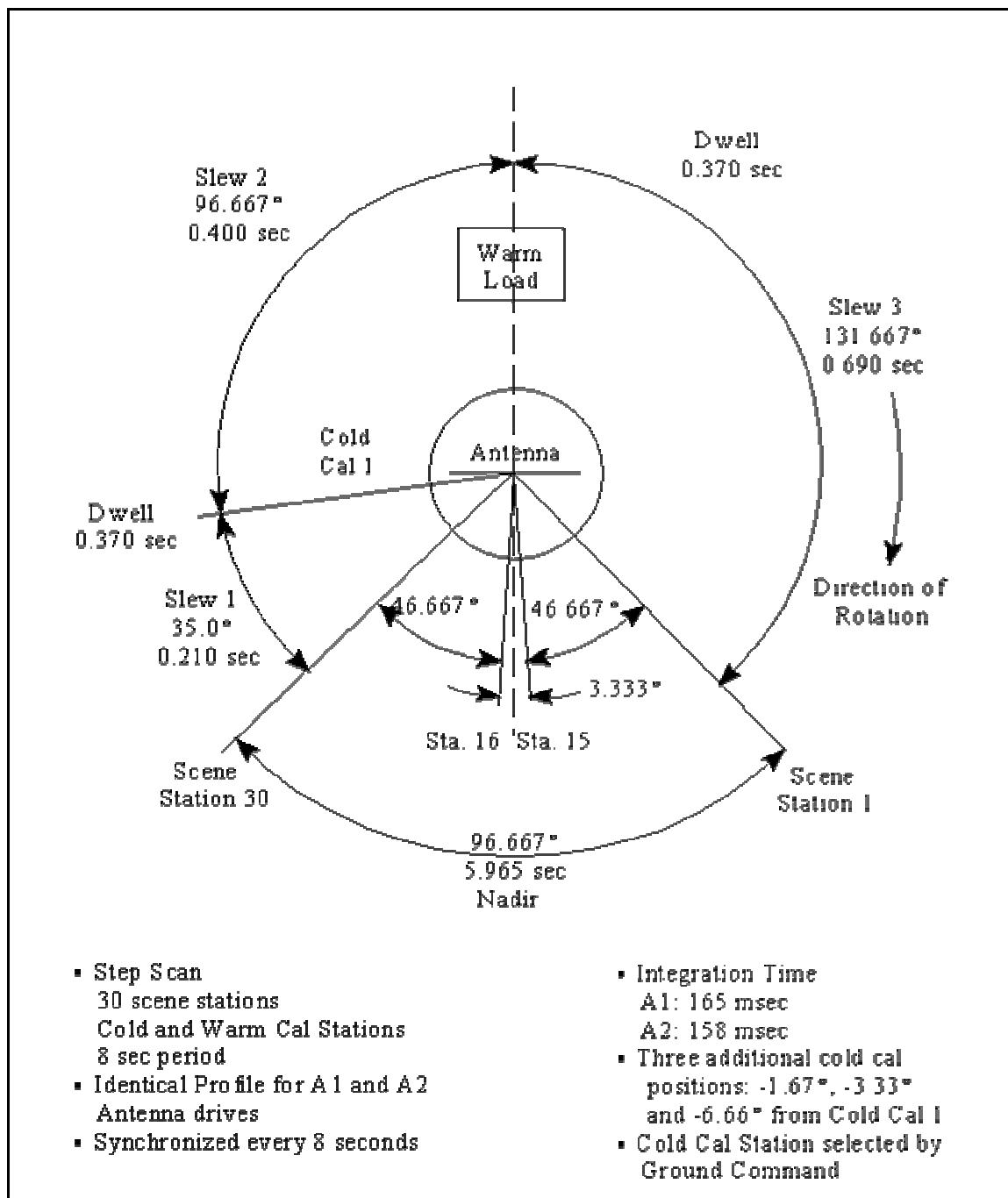


Figure 27: AMSU A scanning characteristics

AMSU A instrument is composed by three receiving antennas. AMSU-A1 has two of them and AMSU-A2 one. These parabolic antennas complete one scan in eight seconds. A single scan is divided into three segments. In the first segment 30 Earth views with a resolution of 50 Km at nadir are measured.

In the second segment a quick scan is performed covering a cold space view and an internal blackbody calibration target. In the third segment each antenna returns to the staring position to start a new scan cycle. A complete description of the AMSU A scanning characteristics is illustrated in figure 27.

4.3.2 Advanced Microwave Sounding Unit-B

The Advanced Microwave Sounding Unit-B (AMSU-B) is a five-channel microwave radiometer at 89, 150, 183.31±1, 183.31±3, and 183.31±7 GHz. The mission of this instrument is to capture observations from different layers of the atmosphere in order to obtain global data on humidity profiles.

AMSU-B has a constant antenna beamwidth of 1.1° (at the halfpower points). Its spatial resolution at nadir is nominally 16 Km (9.94 mi). The receiving antenna makes a crosstrack scan. Each scan takes 2.66 seconds to complete. The antenna scans ±48.95° from the nadir taking a total of 90 Earth views per scan line (figure 28).

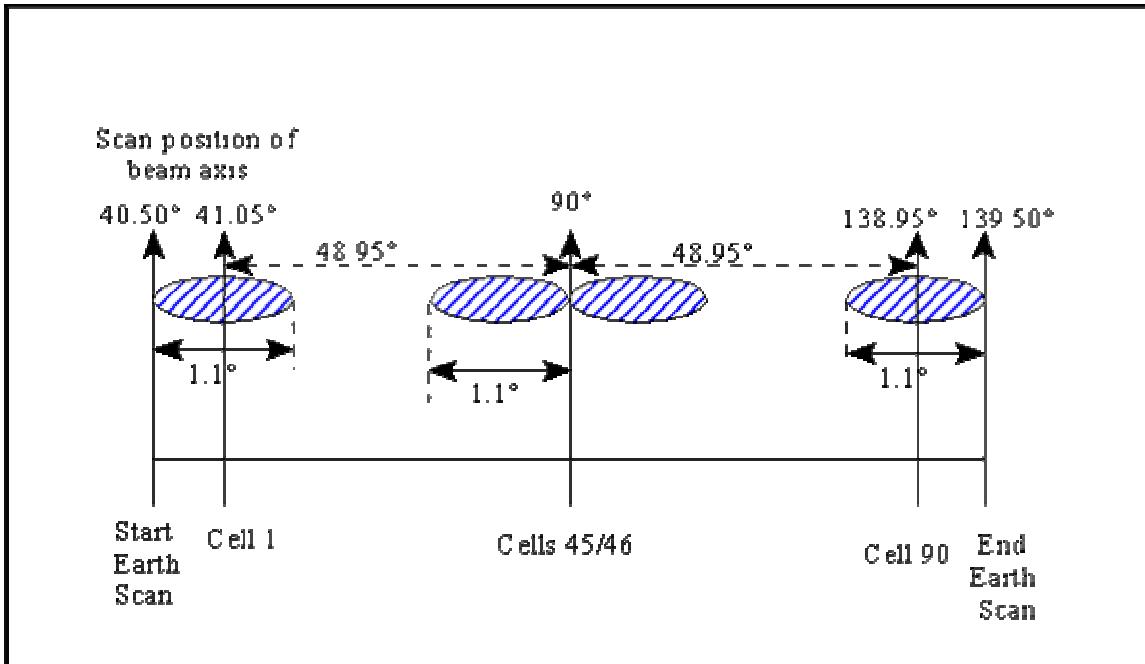


Figure 28: AMSU B scanning

The 183 GHz channels provide the humidity level in the atmosphere. On the other hand, channels 89 and 150 GHz take measurements deeper in the atmosphere to the Earth's surface. AMSU-B instrument parameters and characteristics are summarized in table 3 (NOAA 2004).

Table 3: AMSU-B Instrument Parameters

Channel number	Center freq. of channel (GHz)	No. of pass bands	Bandwidth per passband (MHz)	NEAT (K)	Polarization angle
16	89.0	2	1000	0.37	90-θ
17	150.0	2	1000	0.84	90-θ
18	183.31±1.00	2	500	1.06	90-θ
19	183.31±3.00	2	1000	0.70	90-θ
20	183.31±7.00	2	2000	0.60	90-θ

4.4 The physics of passive microwave

Microwave spectral bands and especially those with higher frequency allow us to get information from atmosphere and weather conditions. Specifically, microwave data is used to measure the spatial distribution of atmospheric variables such as temperature, water vapor, cloud liquid water and wind. Absorption and emission resonances due to water vapor (22.2 and 183.3 GHz) and oxygen (50-70 GHz and 118.7 GHz) can be used to determine the height profiles of the atmospheric water vapor and temperature through radiometric measurements at and near the absorption maxima. On the other hand, Atmospheric windows (low attenuations) are appropriate for surface observations. The material presented in this section is completed from Ulaby et al. (1981), the most cited reference in passive microwave applications.

4.4.1 *Absorption and Emission by Gases*

A molecule consists of three types of energy, electronic, vibrational and rotational. In a transition from lower (or higher) to higher (or lower) energy state in a molecule, a radiation is absorbed (or emitted). The frequency f_{lm} of the absorbed (or emitted) quantum is given by the Bohr formula:

$$f_{lm} = \frac{\xi_m - \xi_l}{h} \quad 4-1$$

Where h is Planck's constant and ξ_m and ξ_l are the internal energies of the higher and lower molecules states, respectively. Absorption line is the absorption spectrum due to a single transition. In the earth's atmosphere only two gases oxygen and water vapor, are constituents that exhibit significant absorption bands in the microwave spectrum. There are other atmospheric gases and pollutants that have absorption lines in the microwave spectrum. Some of them are O₃ SO₂ NO₂ and N₂O. But, their contribution to the microwave gaseous absorption spectrum is negligible in comparison to the oxygen and water vapor contributions. In particular, pressure broadening, which is the collision between molecules, is the most important for atmospheric absorption in the microwave region of the spectrum. The following equation, defines the absorption spectrum for transitions between low and high energy states:

$$k_a(f, f_{lm}) = \frac{4\pi f}{c} S_{lm} F(f, f_{lm}) \quad 4-2$$

where:

k_a = power absorption coefficient, Npm⁻¹

f =frequency, Hz

f_{lm} = molecular resonance frequency for transitions between energy states..

c = velocity of light, 3×10^8 ms⁻¹

S_{lm} =line strength of the lm line, Hz

F = line-shape function, Hz.

The line strength of the lm line of a specific gas is controlled by the number of absorbing molecules of that gas per unit volume, the temperature of the gas and the molecular parameters associated with that transition. The line shape function describes the shape of the absorption spectrum with respect to the resonance frequency f_{lm} . There are several methods to find out the line-shape function but the simplest is the Lorentzian function:

$$F_L(f, f_{lm}) = \frac{1}{\pi} \frac{\gamma}{(f_{lm} - f)^2 + \gamma^2} \quad 4-3$$

γ is defined as half the frequency width at half peak intensity, and is called the linewidth parameter. Generally, this method is used when γ values are much smaller than the transition frequencies.

Van Wleck and Weisskopf (1945) developed a method to find the line-shape function for γ values comparable in magnitude to f_{lm} . The function can be written as:

$$F_{vw}(f, f_{lm}) = \frac{1}{\pi} \left(\frac{f}{f_{lm}} \right) \left[\frac{\gamma}{(f_{lm} - f)^2 + \gamma^2} + \frac{\gamma}{(f_{lm} + f)^2 + \gamma^2} \right] \quad 4-4$$

And finally, Gross (1995) derived another method to find the line-shape function. This can be written as:

$$F_G(f, f_{lm}) = \frac{1}{\pi} \frac{4ff_{lm}\gamma}{(f_{lm}^2 - f^2)^2 + 4f^2\gamma^2} \quad 4-5$$

4.4.2 Water Vapor Absorption

As discussed above, the absorption coefficient of a spectral band is given by equation 4-2

To retrieve water vapor absorption coefficient, $F=FG$, where the Gross line-shape is used because it provides better agreement with experimental observations than other line-shape functions. Also, S_{lm} is defined as:

$$S_{lm} = S_{lm0} f_{lm} \rho_v T^{-5/2} e^{-\xi_l / kT} \quad 4-6$$

Where S_{lm0} is a constant characteristic of the lm transition, ρ_v is the water vapor density and k is the Boltzmann's constant.

4.4.3 Oxygen Absorption

The oxygen microwave absorption region consists of absorption lines between 50-70 GHz and 118.75 GHz. The traditional approach used for computing the absorption spectrum of oxygen has been to add the absorption coefficients due to the individual lines using the Van Vleck-Weisskopf (1945) line shape factor. Also, similar computations have been made using Gross (1955) line-shape factor. The linewidth parameter of an isolated oxygen line should be proportional to the partial pressure oxygen:

$$\gamma = \gamma_0 P_{O_2} \quad 4-7$$

Where γ_0 is the linewidth at $P_{O_2}=1$ mbar and P_{O_2} is the partial pressure of oxygen.

4.4.4 Extinction and Emission by Clouds and Precipitation

The interaction of electromagnetic radiation with particles, such as cloud, fog, snow or rain is important. This process may involve absorption and scattering. The main objective is to examine the case of a large number of particles within the volume of interest. This volume extinction coefficient is controlled by the density, shape, size distribution, and dielectric properties of the particles contained in the volume. Two assumptions are made. One is that the particles are randomly distributed within the volume and the second is that the particles are spherical. The scattering and absorption characteristics of a spherical particle in air are controlled by the electromagnetic wavelength λ_0 , the complex index of refraction n and the radius r of the particle. For purpose of this research this section will focus on snow particles.

4.4.5 Extinction and Backscattering by Snow

For snow particles per unit volume, the extinction coefficient is given by:

$$k_{es} = 4.34 \times 10^3 \left[\sum_{i=1}^{N_V} Q_s(r_i) + \sum_{i=1}^{N_V} Q_a(r_i) \right] \quad 4-8$$

Where N_V is the number of snow particles, k_{es} is in dBkm^{-1} , and $Q_s(r_i)$ and $Q_a(r_i)$ are, respectively the scattering and absorption cross-sections in m^2 of the i th particle with equivalent radius r_i , defined as the radius of a spherical particle with the same mass as the

actual snowflake. For frequencies up to 20 GHz the Rayleigh region, Q_s and Q_a are given by:

$$Q_s = \frac{2\lambda^2}{3\pi} X^6 |K|^2 \quad 4-9$$

$$Q_a = \frac{\lambda^2}{\pi} X^3 \operatorname{Im}\{-K\} \quad 4-10$$

Where K is a complex quantity defined in terms of the complex index of refraction (of the droplet to the background medium). Im signifies “the imaginary part of”. At higher frequencies, Mie formulation should be used for calculating k_{es} . Then:

$$Q_s = \xi_s \times A \quad 4-11$$

$$Q_a = \xi_a \times A \quad 4-12$$

where A is the area (spherical particle = πr^2) and ξ_s and ξ_a are the scattering and absorption efficiency factor, respectively. Using Mie approach:

$$\xi_s(n, x) = \frac{2}{x^2} \sum_{l=1}^{\infty} (2l+1)(|a_l|^2 + |b_l|^2) \quad 4-13$$

$$\xi_e(n, x) = \frac{2}{x^2} \sum_{l=1}^{\infty} (2l+1) \operatorname{Re}\{a_l + b_l\} \quad 4-14$$

$$\xi_a = \xi_e - \xi_s \quad 4-15$$

$$x = k_b r \quad 4-16$$

Where ξe is the extinction efficiency factor, Re means “real part of”, al and bl are Mie coefficients and kb is the wave number in the background medium.

The volume backscattering coefficient of dry snow (σ_{VS}) is given by:

$$\sigma_{vs} = 10^{-10} \frac{\pi^5}{\lambda_0^4} |K_{ds}|^2 Z_s \quad 4-17$$

Where d_s is the diameter of the snowflake and Z_s is given by:

$$Z_s = \sum_{i=1}^{N_p} d_s^6 \quad 4-18$$

In sum, an important issue when dealing with microwave data is the good understanding of the scattering, absorption, and emission behavior of atmospheric constituents in order to monitor atmospheric parameter and weather conditions. Several researches have been conducted in finding a relationship between satellite-based information and snowfall (Skofronick et al. 2004; Kongoli et al. 2003; Noh et al. 2005; Di Michele and Bauer 2005). Successfully, they have found a strong relationship between them. Therefore, based on the scientific research conducted by the authors mentioned above high passive microwave frequencies have been selected for detecting and estimating snowfall.

5. Artificial Neural Network

Artificial Neural Network (ANN) Models have been successfully used in numerous applications. They have been applied in areas such as recognizing images, signal processing problems, pattern recognition, identification, classification, medical prediction and diagnosis and control system design. A neural network model generally is used to perform complex functions that are difficult for conventional computers or human beings. Its capability to be trained rapidly and to deal with data acquired at different levels of measurement precision are the main advantages of neural networks compared to other conventional techniques (Foody and Arora 1997). ANN models have advanced rapidly in the past two decades because of use of scientific method in experimenting with and comparing approaches (Russell and Norvig 2003). Actually, some researches have been developing new neural network approaches for remote sensing problems. In this study, the neural network approach has been selected. The major advantages of ANN models can be summarized in the following points (Ghedira et al. 2000):

- Easy adaptation to different types of data and input configurations (decimal or binary). Moreover, neural networks can easily incorporate ancillary data sources which would be difficult or impossible to do with conventional techniques.
- The traditional parametric classification methods, such as, the Maximum Likelihood Classifier make unreasonable assumptions about the statistical properties of the data. These assumptions are not always satisfied specially when we are dealing with heterogeneous natural land covers.

- A neural network uses its complex configuration to find the best nonlinear function between the input and the output data without the constraint of linearity or pre-specified non-linearity, which is required in regression analysis.
- Unlike to the most statistical classification methods, neural networks have the capacity to weight differently and automatically each data source according to its contribution to ground cover identification (Benediktsson and Sveinsson 1997).

The neural network technique can give positive results in a short time. Two aspects must be taken into account: the care in collecting and processing the training and the validation data and the optimization of its internal parameters. The success of this technique depends on the smart way to apply the knowledge and the scientific background of the presented problem into the design and training of the neural network.

Currently, some existing techniques such as regression are based on assumption of linearity between inputs and outputs linear relationships. On the other hand, Artificial Neural Network (ANN) (figure 29) model can be trained to define a non-linear relationship between the independent variables (input) and dependent variable (output). Sometimes this relationship can be very complex and difficult to analyze by other techniques (e.g. statistical techniques). The relationship between input and output when the ANN is trained cannot be known.

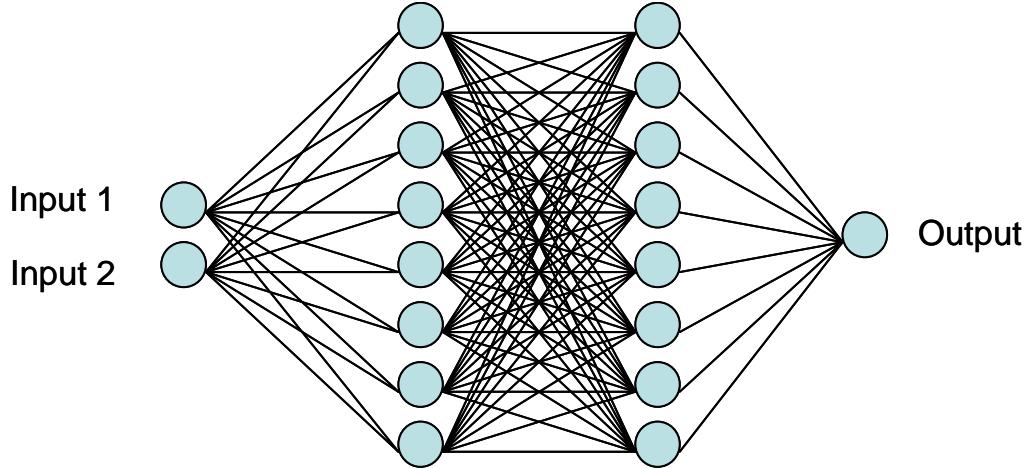


Figure 29: Artificial Neural Network.

5.1 Types of Neural Network

There are several types of neural networks such as Multi-Layer Perceptrons (MLPs), Radial-Basis Function (RBF) and Networks and Kohonen's Feature Map also called Kohonen's Self-Organising Map. A brief description of these types of neural networks is given below and a deeper description of the Multi-layer perceptron network will be presented in one of the following sections (Tarassenko 1998).

5.1.1 *Multi-Layer Perceptron:*

In the 1950's, in a single layer was successfully used in pattern recognition. But, in the 1960's it was demonstrated that a single layer of neurons was not able to learn how to compute a set of binary input. Then, it was discovered that the addition of a second layer

could help in solving the problem. Therefore, another type of neural network was recognized: *Multi-Layer perceptron*.

5.1.2 Radial Basis Function Network:

This network is composed by two layers. One of the main advantages of this type of neural network is the possibility of choosing the parameters for the hidden units without having to perform a full non-linear optimization of the network. The radial basis function as the multi-layer perceptron provides techniques for approximating arbitrary non-linear functional mappings between the inputs and the outputs. The activation used in hidden units is determined by the Euclidean distance between the test pattern and the set of prototype vectors. The radial basis function network uses two different data sets, training and test. In the first stage of the learning the training data set is used without labels, then the same training set or part of it, is used to set the weights in the second layer.

5.1.3 Hohonen's Self Organizing Map algorithm:

This algorithm is mainly used to transform input patterns of arbitrary dimension into two dimensional feature map with topological ordering. This algorithm is usually given in two dimensions but preserving the structure of the data. In the network leaning phase each pattern in the data set is presented to the network, one at the time, in random order. In the initial phase of the training, each pattern gives rise to a localized region of activity in the

feature map against a background of lower activity. After the feature map is trained, the presentation of an input pattern should cause a localized group of units to be active.

5.2 Neural Network Architecture

The neural network model consists of nodes and layers. The nodes are interconnected between them and each one of them operates as a simple processing element (Figure 30). The number of layers and the number of nodes by layer represent the network architecture. The input (I_j) to a node (j) is the weighted sum of the outputs (O_i) from the nodes of the preceding layer (i).

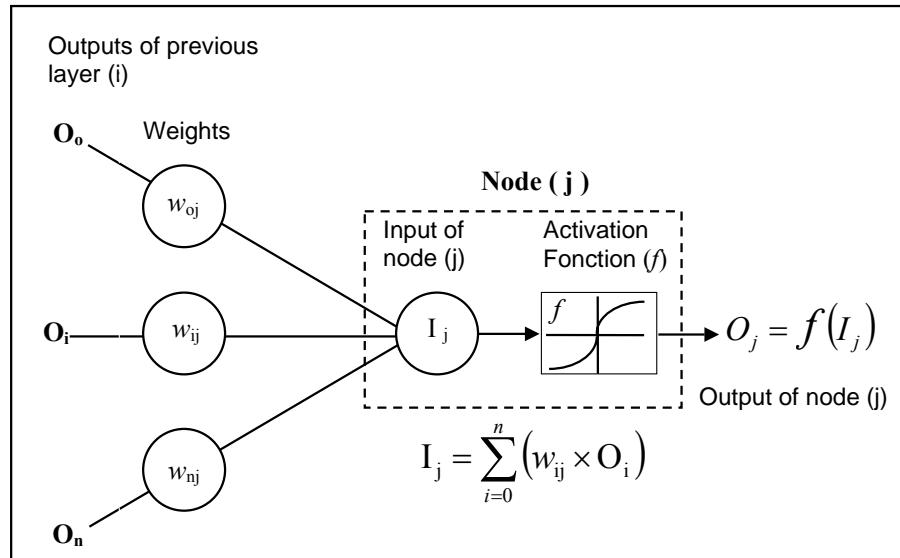


Figure 30: Node process computation

The sum is then passed through an activation function (f) to produce the node's output (O_j) within the range of the activation function. Depending on the activation function

applied, the relationship between input and output varies; see activation functions that will be used in this project in figure 31. The interconnected weight values are adjusted and updated during the training phase to achieve minimal overall training error between the desire and the calculated output vectors (Ghedira et al. 2000).

5.3 Activation Functions

Activation functions are used in the neural network to introduce non-linearity into the system and to face the node output into a specific range.

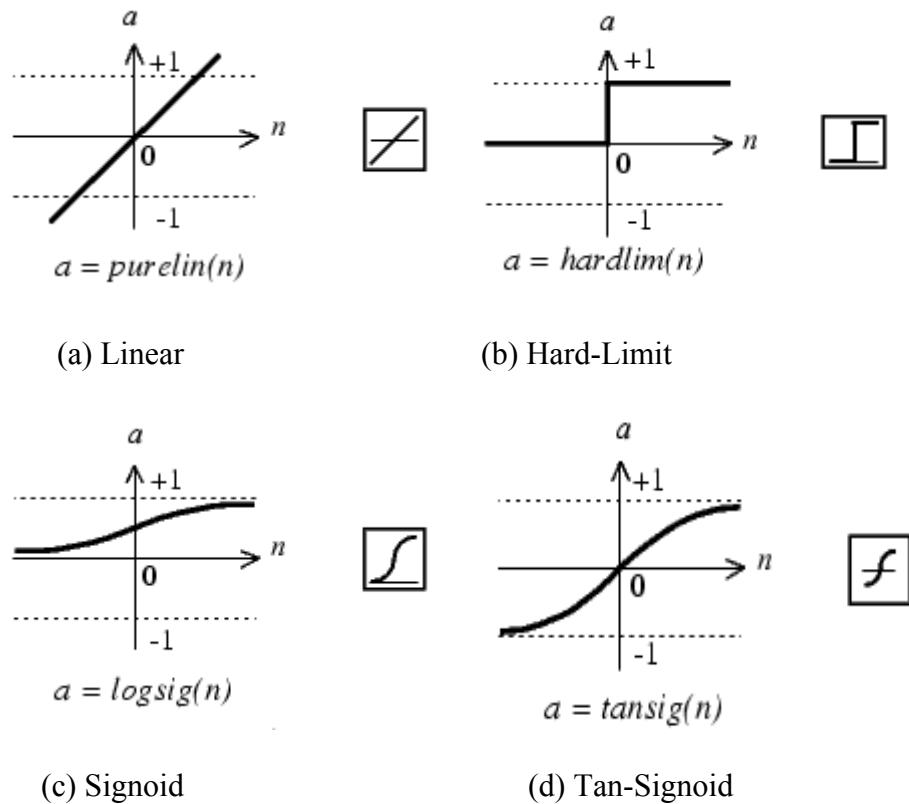


Figure 31: Activation functions

As discussed above, one of the advantages of the neural network is the capability to develop a non-linear relationship between inputs and outputs. This non-linearity relationship is what makes the multilayer neural network powerful.

There are several types of activation functions. Sigmoidal functions such as logistic, the hyperbolic tangent (\tanh) and Gaussian functions are the most common. Some functions produce positive and negative values, others produce only positive values. \tanh and \arctan functions produce both positive and negative values and tend to yield faster training than functions that produce only positive values such as logistic. The mostly used activation functions are described below:

- **Identity:** this activation function is used in the trained of different neural networks such as linear and radial basis function networks.
- **Logistic:** this sigmoid activation function has an output value ranging 0 to 1.
- **Hyperbolic:** this hyperbolic tangent function (\tanh) is *also* a sigmoid function such as the logistic function. The output value ranges from -1 to 1. This function is mostly used in multi-layers, especially in hidden layers.
- **Exponential:** this is a negative exponential activation function. This activation function is ideal for use with radial units.
- **Softmax:** this is an exponential activation function. The results are normalized so the sum of activation functions across the layer is 1. This activation function is mostly

used in the output layer of a multi-layer neural network and especially for classification problems.

- **Unit Sum:** this activation function normalized the output to sum to 1.0.
- **Square Root:** this activation function is used to transform the squared distance activation in a Self Organizing Feature Map (SOFM) network or Cluster network to the actual distance as an output.
- **Sine:** this activation function is useful if recognizing radially distributed data. This activation function is not used as default.
- Ramp: this activation function is a piece-wise linear version of the sigmoid function. This function yield faster training, but with poor training performance.
- **Step:** the output of the neural network can be either 1 or 0. It depends on weather the synaptic value is positive or negative.

The mathematical expressions for each activation function are presented in table 4.

Table 4: Activation Function Expressions

Function	Definition	Range
Identity	X	(-inf,+inf)
Logistic	$\frac{1}{1 - e^{-x}}$	(0,+1)
Hyperbolic	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	(-1,+1)

-Exponential	e^{-x}	(0, +inf)
Softmax	$\frac{e^x}{\sum_i e^{x_i}}$	(0,+1)
Unit sum	$\frac{x}{\sum_i x_i}$	(0,+1)
Square root	\sqrt{x}	(0, +inf)
Sine	$\text{Sin}(x)$	[0,+1]
Ramp	$\begin{cases} -1 & x \leq -1 \\ x & -1 < x < +1 \\ +1 & x \geq +1 \end{cases}$	[-1,+1]
Step	$\begin{cases} 0 & x < 0 \\ +1 & x \geq 0 \end{cases}$	[0,+1]

5.4 Neural Network Training Techniques

Training the neural network could be *supervised* or *unsupervised*. Supervised training uses inputs and their corresponding outputs (usually from historical data), and learns to infer the relationship between them during an iterative training process. Once the neural network is trained a new set of data called test data is used to evaluate the trained neural network performance. Unsupervised training uses only input data. In many real world problems, it is expensive and time consuming to collect the correspondent outputs of a large set of data. Usually, for a small part of the data set the correspondent output might be collected, but for the other part might not be that easy to collect. In this case, a form of

unsupervised learning called *cluster analysis* is used. The clustering technique makes up a number of assumptions about the data. These assumptions usually depend on specific characteristics of the problem (Tarassenko 1998).

5.5 Backpropagation Algorithm

Backpropagation algorithm is the most popular supervised learning algorithm. This algorithm is often because of its easy adaptation to multiple layer networks and to nonlinear differentiable transfer functions. Also, for classification purpose, the multi-layer neural network trained by backpropagation algorithm is the most commonly used. The standard backpropagation is a gradient descent algorithm.

The neural network is trained by adjusting its weights. To adjust the weights a feedback error is used. This is done by computing the error (E) at the output of the network and then this error is minimized by gradient descent.

The initial weights may be randomly chosen what makes a non-unique solution. The solution will always depend on the values of the initial weights. Then, the gradient of the error with respect to each weight is computed. For instance, for weight w_i is $\delta E / \delta w_i$. The following vector w^1 is obtained by moving a small distance in the direction of steepest descent. For example for a single-layer perceptron:

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i} \quad 5-1$$

Where Δw_i is the difference of the current weight with the new weight and η also called learning rate, is the value that defines the step size. The learning rate value should not be too large or too low. If it is too large the error-correction process can overshoot and divergent oscillation may occur. If it is too low, the weights take a very long time to converge.

5.6 The Perceptron Learning Rule

For two classes A and B, the target values are +1 and -1 respectively. If the perceptron classifies correctly one of the patterns from class A, the expression will be as following:

$$y = f_n(\sum w_i x_i) = t = +1 \quad 5-2$$

Where $\sum w_i x_i > 0$. But, if the perceptron classifies correctly one of the patterns from class B, the expression will be as following:

$$y = f_n(\sum w_i x_i) = t = -1 \quad 5-3$$

Where $\sum w_i x_i < 0$. Using the gradient descent concept , the perceptron criterion becomes:

$$\Delta w_i = \eta x_i^p t^p \quad 5-4$$

This equation is called perceptron learning rule, where $\delta E / \delta w_i = -x_i t$. p represents the pth pattern in the training set. p belongs to the set of input patterns.

Sometimes the classes can not be linearly separable, therefore the perceptron rule does not converge and the decision boundary oscillates between a number of positions.

Sum of Square Function

A new technique implementation is needed to discriminate perfectly between classes and the perceptron criterion is then replaced for the sum of square function:

$$E = \frac{1}{2} \sum_{p=1}^P (y^p - t^p)^2 \quad 5-5$$

Where y^p is the (single) output of the multi-layer network for pattern p and t^p is the correspondent target value.

In a multi-layer neural network having k -class problem, the number of output unit will be k . The error function then becomes:

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^K (y_k^p - t_k^p)^2 = \frac{1}{2} \sum_p \sum_k (g(\sum_j w_{jk} y_j^p) - t_k^p)^2 \quad 5-6$$

Where $y_k = g(a_k) = g(\sum_j w_{jk} y_j)$. Since, $y_j = g(a_j) = g(\sum_i w_{ij} x_i)$, then the equation becomes:

$$E = \frac{1}{2} \sum_p \sum_k (g(\sum_j w_{jk} g(\sum_i w_{ij} x_i^p)) - t_k^p)^2 \quad 5-7$$

From this equation the error backpropagation algorithm for updating the weights in a multi-layer perceptron is derived. For the hidden to output layer weights:

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \partial_k y_j \quad 5-8$$

$$\text{Where } \partial_k = \frac{\partial E}{\partial a_k} = (y_k - t_k)y_k(1 - y_k) \quad 5-9$$

For the input to hidden layer weights:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = -\eta \partial_j y_i \quad 5-10$$

$$\text{Where } \partial_j = \frac{\partial E}{\partial a_j} = \sum_k \partial_k w_{jk} y_j (1 - y_j) \quad 5-11$$

Details of the derivation of these equations can be found in Tarassenko 1998.

5.7 Types of Backpropagation Algorithms

There are several variations of backpropagation training algorithms. Some of these variations are named as follow:

5.7.1 Variable Learning Rate (GDX):

This algorithm converges fast. Also, it uses batch mode which updates the weights after all the inputs are applied to the network. The performance of this algorithm is very sensitive to the learning rate value. If the learning value is too high the network can oscillate and become unstable. If the learning rate value is too low the algorithm takes long time to converge.

5.7.2 Resilient Backpropagation (RP):

The purpose of this algorithm is to eliminate one of the problems presented in some backpropagation algorithms. Backpropagation algorithm uses sigmoid transfer functions in the hidden layers. These functions compress an infinite input range into a finite output range. Also, their slope must approach zero as the input gets large. This causes a problem because the gradient can have a very small magnitude and, therefore cause small changes in the weights and the biases, being these values far away from the optimal values. This problem can be solved by using the resilient backpropagation algorithm.

5.7.3 Conjugate Gradient Algorithms:

In this type of algorithms a search is performed along several conjugate directions to adjust the weights. Usually the standard backpropagation algorithms adjust the weights in the steepest descent direction. In this direction the performance function is decreasing most rapidly but it does not necessarily produce the faster convergence. There are four types of conjugate gradient algorithms which are described below (MATLAB user's guide):

- **Fletcher-Reeves Update (cgf):** the basic conjugate gradient algorithms have a default search to adjust the weights. This is usually in the steepest descent direction as follows:

$$P_0 = -g_0$$

5-12

Where P_0 is the initial gradient.

To determine the optimal distance to move along the current search direction, a line search is then performed:

$$x_{K+1} = x_K + \alpha_K p_K \quad 5-13$$

To determine the new search direction, the new steepest descent direction is combined with the previous search direction.

$$p_K = -g_K + \beta_K p_{K-1} \quad 5-14$$

The types of the conjugate gradient algorithm differ in the way the β_K is computed. In this case:

$$\beta_K = \frac{g_K^T g_K}{g_{K-1}^T g_{K-1}} \quad 5-15$$

- **Polak-Ribiere Update (cgp):** a new version for computing β_K was proposed by Polak and Ribiere, as follows:

$$\beta_K = \frac{\Delta g_{K-1}^T g_K}{g_{K-1}^T g_{K-1}} \quad 5-16$$

- **Powell-Beale Restarts (cgb):** the standard reset point occurs when the number of iterations is equal to the number of the network parameters. There are other reset methods that can improve the efficiency of the training. One reset method was proposed by Powell, as follow:

$$g_{k-1}^T g_k \geq 0.2 \|g_k\|^2$$

5-17

If this condition is satisfied, the search direction is reset to the negative of the gradient.

- **Scaled Conjugate Gradient (scg):** the other types of conjugate gradient algorithm spend long time in the search to adjust the weights. The scaled conjugate gradient algorithm avoids the problem of time-consuming. This algorithm utilizes an approach by combining the model trust region from Levenberg-Maquardt algorithm (explained in one of the following sections) with the conjugate gradient approach.

5.7.4 Quasi-Newton Algorithms:

- **BFGS Algorithm (BFG):** this technique is an option to the conjugate gradient method for fast optimization. The disadvantage of this method is that it uses Hessian matrices. These matrices are complex and expensive to compute for feedforward neural networks. The Newton's method basic step is as follows:

$$X_{k+1} = x_k - A_k^{-1} g_k$$

5-18

where the Hessian matrix of the performance index at the current values of the weights and bias is A_k^{-1} . In this algorithm the approximate Hessian matrix is updated at each iteration. The update is computed as a function of the gradient.

- **One Step Secant Algorithm (OSS):** this algorithm is a balance of the conjugate gradient algorithms and the quasi-Newton algorithms. Since the BFGS requires more storage and computation in each iteration, the One Step Secant (OSS) offers a secant approximation with smaller storage and computation requirements. This algorithm does not store the complete Hessian matrix. It assumes that at each iteration, the previous Hessian matrix was the identity matrix.

5.7.5 Levenberg-Marquardt (LM) algorithm:

This algorithm has the capability to train rapidly networks of moderate sizes. Like quasi-Newton algorithms, the Levenberg-Marquardt algorithm was designed to approach second order training speed without having to compute the Hessian matrix. The Hessian matrix can be approximated as:

$$H = J^T J \quad 5-19$$

and the gradient can be computed as:

$$g = J^T e \quad 5-20$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases and e is a vector of network errors. The computation of the Jacobian matrix is done through a standard backpropagation technique which is much less complex than the computation of the Hessian matrix.

The approximation to the Hessian matrix used by the Levenberg-Marquardt algorithm is:

$$X_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad 5-21$$

This equation becomes Newton's method when μ is zero. When μ is large, this equation becomes gradient descent with a small step size. Since when using the Newton's method the optimization is faster and more accurate near an error minimum, the objective is to shift toward the Newton's method as soon as possible.

5.8 Neural Network Training

There are many factors than can affect the neural network training speed, such as the size of the data set, the goal error and the difficulty of the problem. One of the major problems using neural network is the overtraining. Overtraining occurs when the neural network's generalization ability will be compromised and the classification space becomes narrowly defined around the training pixels (Augustejin et al. 1995). An over-trained neural network will learn the details of the training data instead than overlying input and output mapping. Therefore, an overtrained neural network will give poor performance when simulated new data.

Four steps should be applied in training the neural network: assemble the data, train the network, selecting the optimal network parameters and simulate the network response to new inputs (MATLAB, User's guide, 2006).

- **Assemble the data:** in the case of backpropagation algorithms both the inputs and their correspondent targets are needed to train the network, which is time consuming. In order to prevent overtraining the data set should be divided into three sets. This technique is shown in figure 32. The first is the training set. This set is used for computing and updating the network weights. The second is the validation set. This set ensures the training to be away from overtraining by using the validation error. Therefore, this set is used to stop the training without using this data to update the weights. And the third, the test set. This is used to benchmark the neural network and to compare different models. This data set is presented to the trained neural network to assess its performance.
- **Train the Network:** the training step is an iteration process. All the inputs are presented to the network several times. At each time the output error is measured and the weights are updated using this error. For each training value presented at each epoch, the network output is computed using the current weight set and then is compared to the correspondent target. Then, the square error (E) is known. Normally, as it is the case for the learning set error, the error computed on the validation set decreases during the initial phase of training. However, when the network begins to overfit the learning data, the error on the validation set will begin to increase slowly for the next iterations. At that time, the training process will be stopped, and the neural network weights corresponding to the minimum validation error will be maintained for the next steps (Ghedira et al. 2004).

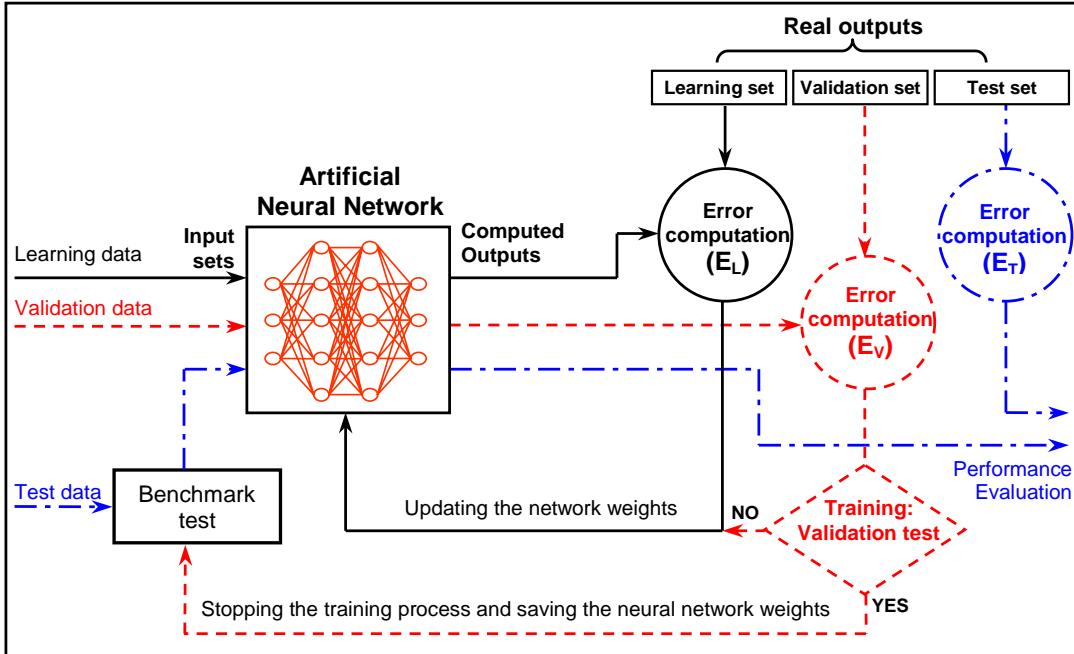


Figure 32: Contribution of each training group to the global training process

- **Selecting the Optimal Neural Network:** the neural network is performed several times and the optimal network or the m best neural networks are selected by looking at the minimum error in the validation set.
- **Simulation of the Neural Network:** the test data set plays an important role in the neural network training process. The performance of the neural network is assessed by using the test data set. This data set is presented to the optimal neural network or to the m best neural networks. Afterwards, the overall performance may be computed by the average accuracy of each of the m best networks.

5.9 Most Common Problems in Training and/or Testing

5.9.1 Poor training performance:

This usually occurs when the training error remains high as the neural network learns from the inputs and their correspondent outputs. This problem may occur for several reasons:

- **Incorrect way to attack the problem:** the inputs may not have any connection with their correspondent targets and especially with the problem. This type of problem may be one of the problems that a neural network could not learn.
- **Wrong set features:** this case is presented when a wrong set of features is chosen to characterize the former even though there is a relationship between inputs and outputs.
- **Stuck units:** this problem is presented when the initialized weights values are too large or when there is a mistake in normalizing the input variables.

5.9.2 Poor generalization performance:

In this case the error on the training set could have reached its minimum value. But, when simulating the trained neural network the error in the test set may be high and not acceptable. This may occur for many reasons, these are describe below:

- **Insufficient number of training patterns:** this problem occurs when the information to determinate the neural network weights properly is not enough. Then, the neural network will not be able to establish any relationship between x and y .

- **Over-fitting:** this problem is caused by the high number of hidden units or high number of weights. This number may be too high for the problem to solve. If the number of hidden units is too high the neural network learns all the details of the training patterns. Also, it will have sufficient degrees of freedom to fit the noise then performance of the trained neural network on the test data will be very poor.
- **Over-training:** this problem occurs when the network is under training for long time without any stop. This could be possible when the validation set is not used or when the training set error is low then the validation error would not begin rise.
- **Unbalance data set:** this problem occurs when the problem to solve is composed by unbalanced data sets. Unbalance data sets means that for one of the classes, especially if it is the class the problem relies on, there is no enough representative data. For instance, this class may represent only 10% of the total data set which is not representative at all.
- **Extrapolation rather than interpolation:** this problem occurs when the neural network is trained at certain conditions and then is simulated under different conditions. Sometimes the conditions under which the neural network was trained change during time. In this case the neural network should be trained again.

In conclusion, neural network models have been increasingly used to solve remote sensing applications problems. In this study, neural network was selected because its capability to solve non-linearity problems that other conventional techniques can not solve. As

discussed in this chapter, there are several types of neural network models that can be used for different problems. The key when selecting and training a neural network model is to understand the characteristics and physics of the problem.

6. Data Acquisition

An important issue in the design, calibration and validation of any prediction or estimate model reside in the rigorous selection of a reliable data to be used in the training and the validation of the model. The selection of the data to use in any study should be based on the knowledge of scientific and physical background of the faced subject. Moreover, the selection of the study area for validation should be independent from the area used in the training of the model in both space and time. In this section, a description of each data set used in this study is given. Also, the areas of training and validation will be described.

6.1 Training Area and Data Set

Hourly snowfall accumulations are used as truth data in this project. An hourly temporal resolution was applied to reduce the potential source of noise in the data. Ground based hourly snowfall precipitation data was collected from the Local Climatological Data (LCD) section of the National Climate Data Center (NCDC). A sample of one day data is illustrated in tables 5 and 6. From table 5 daily observations such as pressure average at the station, pressure average at sea level, maximum, average and minimum temperature, snowfall, and wind information are available. In table 6 hourly precipitations can be found. To determine if the positive amount is related to snowfall or rainfall, hourly surface temperature derived from AMSU-A brightness temperatures is used and it will be described later in this chapter.



JANUARY 2004
LOCAL CLIMATOLOGICAL DATA
NOAA, National Climatic Data Center

YOUNGSTOWN/WARREN, OH

YOUNGSTOWN-WARREN RGNL APT (YNG)

Lat: 41°15' N Long: 80°40' W Elev (Ground): 1190 Feet

Time Zone: EASTERN WBAN: 14852 ISSN #:0198-4047

DATE	TEMPERATURE F							DEG DAYS BASE 65°	WEATHER	SNOW/ICE ON GND(IN)		PRECIPITATION (INCHES)		PRESSURE (INCHES OF HG)		WIND SPEED = MPH DIR = TENS OF DEGREES						DATE		
	MAXIMUM	MINIMUM	AVERAGE	DEP FROM NORMAL	AVERAGE DEW PT	AVERAGE WET BULB	DNLTG			0700 LST	1300 LST	2400 LST	2400 LST	AVERAGE STATION	AVERAGE SEA LEVEL	RESULTANT SPEED	AVG DIR	5-SEC SPEED	AVG DIR	2-MIN SPEED	AVG DIR			
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
01	42	25	34	8	19	31	0	RA		0.0	0.04				4.5	14	17	14	17	17	14	17	01	
02	52	39	46	20	19	47	0	RA DZ BCPG BR		0.0	0.47				8.2	20	17	15	18	20	17	18	02	
03	60*	40	50*	24	50	52	0	RA DZ BR		0.0	0.25	28.61	29.87	5.7	24	10.0	26	33	23	23	23	23	03	
04	40	33	37	11	33	34	0	RA PL BR		T	1.04	28.67	29.95	9.3	03	10.0	22	09	19	06	19	06	04	
05	36	26	31	5	29	31	0	RA SN FG BR		0.5	0.04	28.61	29.91	8.1	30	9.5	23	28	20	27	20	27	05	
06	26	10	18	-7	7	14	47	0	SN BLSN	1	1.3	0.06	28.92	30.24	16.6	26	16.9	32	25	26	25	26	06	
07	18	10	14	-11	6	12	51	0	SN	1	0.7	0.02	28.98	30.33	13.9	25	14.1	32	24	26	25	26	07	
08	25	13	19	-6	13	19	46	0	SN	1	T	28.96	30.29	6.5	26	6.9	15	23	13	26	13	26	08	
09	22	3	13	-12	6	12	52	0	SN	1	T	29.02	30.36	10.7	01	11.0	23	01	18	01	19	01	09	
10	14	1	8	-17	-2	5	57	0		0.0	0.00	29.12	30.48	0.6	01	3.5	13	01	10	33	10			
11	34	10	22	-3	10	19	43	0	SN	1	T	28.83	30.16	11.7	20	11.9	25	23	18	23	23	11		
12	36	31	34	9	29	32	31	0	RA SN BR	1	0.4	0.07	28.70	29.99	11.9	25	12.7	25	23	20	23	12		
13	32	16	24	-1	20	24	41	0	FZDZ SN BR HZ	1	1.7	0.12	28.82	30.12	11.8	30	15.1	29	33	25	23	13		
14	23	12	18	-7	16	18	47	0	SN	1	4.9	0.30	28.61	29.93	4.8	11	8.5	16	33	14	34	14		
15	20	5	13	-12	9	12	52	0	SN BR	6	0.2	3.1	0.12	28.73	30.05	10.7	32	11.1	23	33	18	22	15	
16	15	-5	5	-20	2	6	60	0	SN	6	0.3	0.2	T	28.97	30.32	3.8	32	4.6	15	32	13	33	16	
17	32	-1	16	-8	12	15	49	0	RA FZRA FZDZ SN BLSN	6	0.3	2.0	0.08	28.72	30.06	5.9	16	6.9	16	15	14	14	17	
18	33	16	25	1	20	23	40	0	FZDZ SN BR BLSN	6	1.1	1.4	0.05	28.40	29.69	10.6	29	11.7	28	31	23	22	18	
19	17	14	16	-8	11	14	49	0	SN	7	0.5	3.1	0.13	28.64	29.96	10.8	29	11.1	24	27	18	28	19	
20	17	10	14	-10	6	12	51	0	SN	9	0.7	0.1	T	28.89	30.22	9.1	28	9.5	20	27	15	27	20	
21	27	2	15	-9	6	12	50	0	SN IC	7	0.7	T	T	28.78	30.11	9.0	21	9.2	23	22	17	21	21	
22	32	5	19	-5	8	15	46	0	SN BLSN	6	0.5	0.9	0.03	28.57	29.89	14.4	27	16.2	37*	29	29*	28	22	
23	11	1	6	-18	2	6	59	0	SN BR	6	0.5	2.3	0.05	28.70	30.04	8.9	24	9.8	24	26	20	26	23	
24	16	2	9	-15	4	9	56	0	SN BR	7	0.7	0.3	0.01	28.72	30.06	5.3	33	6.0	21	32	17	33	24	
25	16	-8*	4*	-21	0	6	61	0	SN BLSN	7	0.8	0.4	0.01	28.88	30.23	8.7	09	9.5	25	09	20	10	25	
26	27	14	21	-4	17	19	44	0	RA FZRA FZDZ SN PL BR	8	1.0	2.0	0.22	28.64	29.96	14.9	10	15.2	29	11	24	12	26	
27	33	20	27	2	24	26	38	0	TS RA FZRA SN PL BR	7	1.6	0.8	0.36	28.42	29.71	9.8	19	15.5	32	24	28	25	27	
28	20	15	18	-7	13	16	47	0	SN BR	7	1.5	1.6	0.06	28.61	29.93	16.9	25	17.2	33	24	29	25	28	
29	15	7	11	-14	6	10	54	0	SN BR	7	1.8	2.0	0.05	28.68	30.00	10.1	27	10.5	28	25	21	25	29	
30	11	3	7	-18	3	7	58	0	SN BR BLSN	8	1.9	0.5	0.02	28.56	29.89	11.6	26	11.7	29	25	23	26	30	
31	15	2	9	-16	4	7	56	0	SN	8	1.8	0.1	T	28.81	30.15	9.0	24	9.5	22	23	17	24	31	
	26.4	12.0	19.2			45.5	0.0	< MONTHLY AVERAGES	TOTALS ->		30.3	3.60					10.6	<- MONTHLY AVERAGES						

JANUARY 2004
YOUNGSTOWN/WARREN, OH

Table 6: Local Climatological Data, Page 2.

HOURLY PRECIPITATION

(WATER EQUIVALENT IN INCHES)

YOUNGSTOWN/WARREN, OH

JANUARY 2004 YNG WBAN # 14852

DATE	FOR HOUR (LST) ENDING AT												DATE	FOR HOUR (LST) ENDING AT												DATE	Sum if Different (See Note)	2400 LST
	1	2	3	4	5	6	7	8	9	10	11	12		13	14	15	16	17	18	19	20	21	22	23	24			
01													01												01	0.00	0.04	
02													02												02	T	0.47	
03	T	T	0.01	T	0.01	0.10	0.08	T	0.01	T	0.03	0.04	03	0.01	T	T	T	T	T	T	0.02	0.01	T	T	T	02	0.25	
04	T	0.01	0.01	T	T	T	T	T	T	T	0.05	0.04	04	0.03	0.03	0.01	0.03	0.05	0.06	0.10	0.06	0.11	0.09	0.01	0.03	03	1.04	
05	0.01	0.01	T	T	T	T	T	T	T	T	0.05	0.05	05	T	T	T	T	T	T	T	T	T	T	T	T	04	0.04	
06	0.01	T	0.01	T	T	T	0.01	0.01	0.01	T	0.01	T	06	T	T	0.01	T	T	T	T	T	T	T	T	T	05	0.83	
07	T	0.01	T	T	T	T	T	T	T	T	T	T	07	T	T	T	T	T	T	T	T	T	T	T	T	06	0.06	
08													08														07	0.02
09	T	T	T	T	T	T	T	T	T	T	T	T	09	T	T	T	T	T	T	T	T	T	T	T	T	08	T	
10													10														09	0.00
11																											10	T
12																											11	0.07
13																											12	0.12
14	T	T	T	T	T	T	0.01	T	T	T	0.02	0.02	14	T	T	0.02	0.01	0.01	0.01	0.04	0.06	0.02	0.04	0.01	T	13	0.30	
15	T	T	T	T	T	T	T	T	T	T	0.02	0.01	01	0.01	0.01	0.02	0.01	0.01	0.01	0.04	0.06	0.02	0.01	T	14	0.12		
16	T	T	T	T	T	T	T	T	T	T	T	T	16	T	T	T	T	T	T	T	T	T	T	T	T	15	0.08	
17													17	0.01	T	0.02	0.02	T	0.01	T	T	T	T	T	T	16	T	
18	T	T	0.02	T	T	T	0.01	T	T	T	T	T	18	T	T	T	T	T	T	T	T	T	T	T	T	17	0.08	
19	0.01	0.02	T	T	T	T	T	T	T	T	T	T	19	0.02	0.01	0.01	0.02	0.01	0.01	0.04	0.06	0.02	0.01	0.01	T	18	0.05	
20	T	T	T	T	T	T	T	T	T	T	T	T	20	T	T	T	T	T	T	T	T	T	T	T	T	19	0.13	
21	T	T	T	T	T	T	T	T	T	T	T	T	21	T	T	T	T	T	T	T	T	T	T	T	T	20	T	
22	T	T	T	T	T	T	T	T	T	T	T	T	22	T	T	T	T	T	T	T	T	T	T	T	T	21	0.03	
23	T	T	T	T	T	T	T	T	T	T	T	T	23	T	T	T	T	T	T	T	T	T	T	T	T	22	0.05	
24	0.01	T	T	T	T	T	T	T	T	T	T	T	24	T	T	T	T	T	T	T	T	T	T	T	T	23	0.01	
25													25														24	0.01
26	0.01	0.01	0.01	0.02	T	T	0.01	T	T	T	0.01	26	0.01	T	T	T	T	0.02	0.03	0.03	0.01	0.02	0.02	0.02	25	0.21		
27	0.01	0.05	0.08	0.03	0.02	T	0.01	T	T	T	0.07	0.02	27	0.01	T	T	T	T	0.01	T	T	T	T	T	T	26	0.36	
28	T	T	T	T	T	T	0.01	T	T	T	0.01	T	28	T	T	T	T	T	T	T	T	T	T	T	T	27	0.04	
29	T	T	T	T	T	T	T	T	T	T	T	T	29	0.01	0.01	T	T	T	T	T	T	T	T	T	T	28	0.06	
30	T	T	T	T	T	T	T	T	T	T	T	T	30	T	T	T	T	T	T	T	T	T	T	T	T	29	0.05	
31	T	T	T	T	T	T	T	T	T	T	T	T	31	T	T	T	T	T	T	T	T	T	T	T	T	30	0.02	

The data set used in neural network training was built on a pixel-based approach. Hourly snowfall information was collected during four winter seasons (2002, 2003, 2004 and 2005) from seventeen stations in the North-East of the United States (Figure 33).

After collecting all the data from the seventeen stations during four winter seasons, 108720 snowfall and non-snowfall pixels were collected. From these pixels only 24857 matched the acquisition time with the satellite. Therefore, two filters (cloud cover and surface temperature) were applied to this data set in order to reduce data noises.

The details of these filters are explained later in this chapter. After all, a total of 6056 snowy and non-snowy pixels were collected. To guarantee the collection of the corrected (snowfall) data, only pixels with storm lasting at least three hours were selected. After applying this condition 5690 pixels were left. From these 5690 pixels, 448 were precipitating pixels and the other 5242 were non-precipitating pixels. Therefore, from the 5242 pixels 760 non-precipitating pixels in conjunction with the 448 precipitating pixels were used to train, validate and test the model.

Moreover, the number of pixels is limited because (a) of the difficulty of matching snowfall events with AMSU acquisition time, (b) only pixels with storms lasting more than three hours were selected to reduce the risk of erroneous identification of snowfall pixels and (c) each pixel has to meet the preset filters conditions. Such criteria will undoubtedly increase the level of confidence that snowfall coincides with AMSU acquisition time.

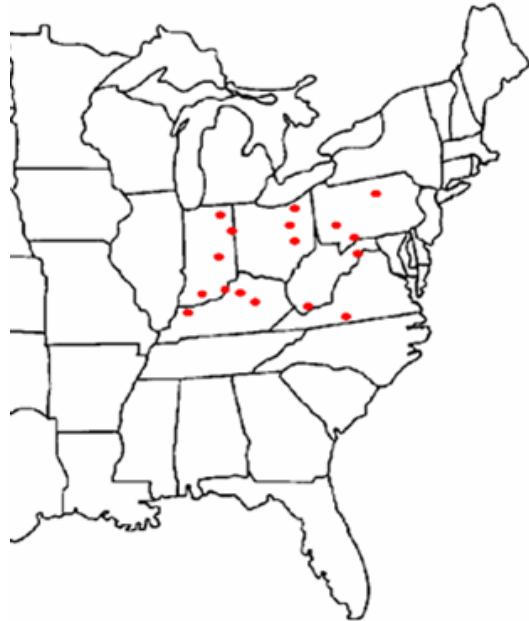
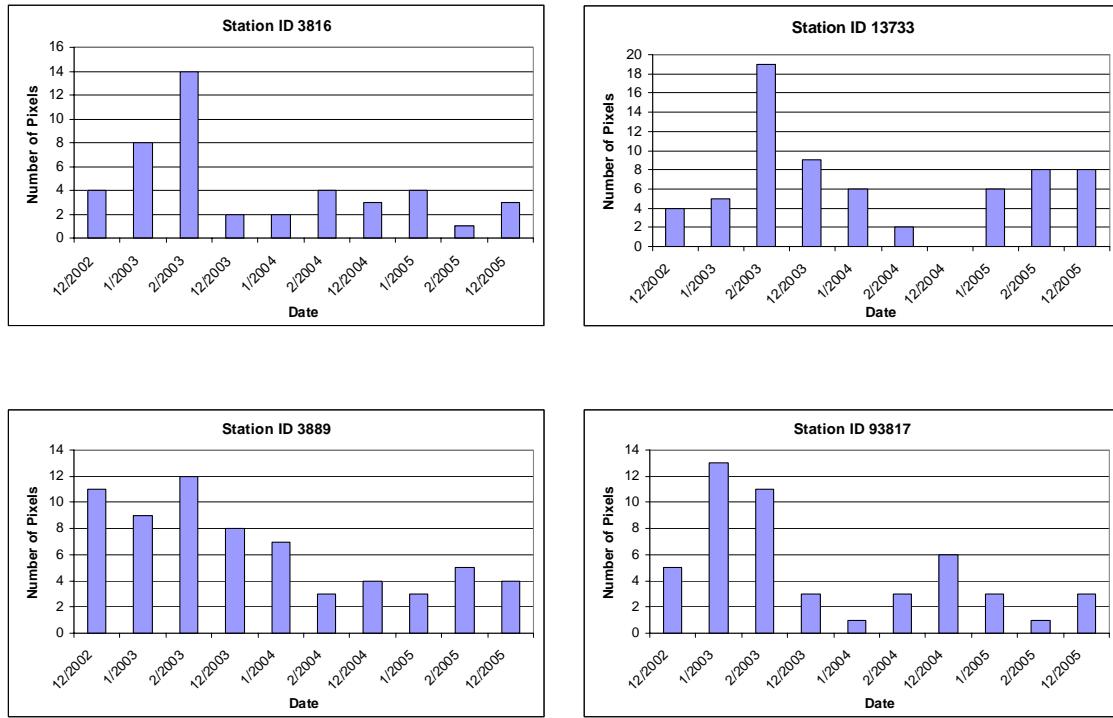
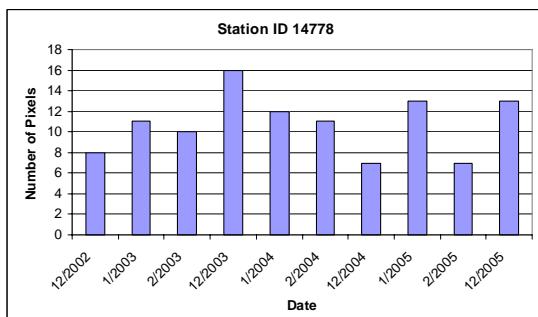
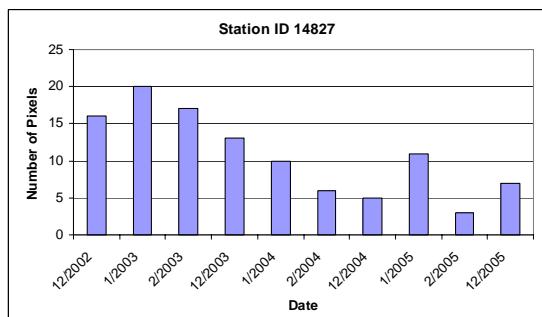
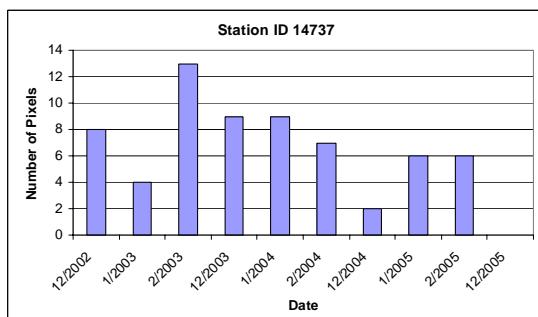
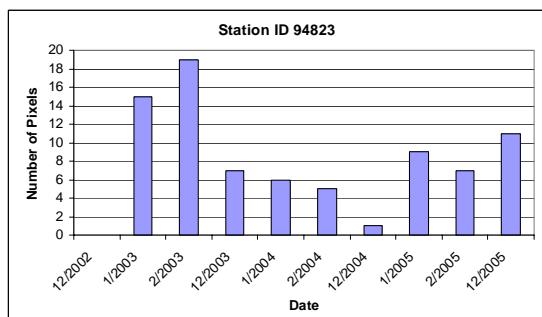
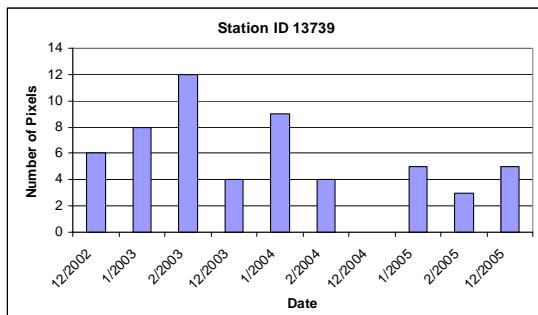
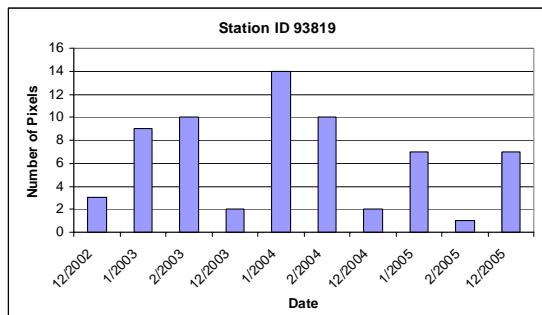
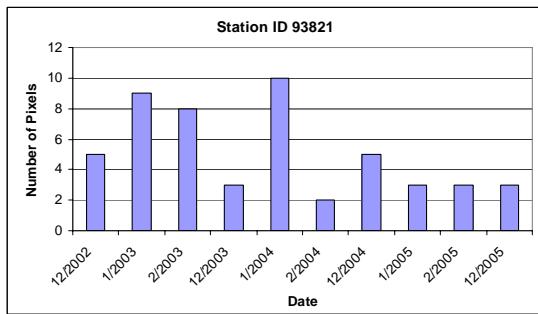
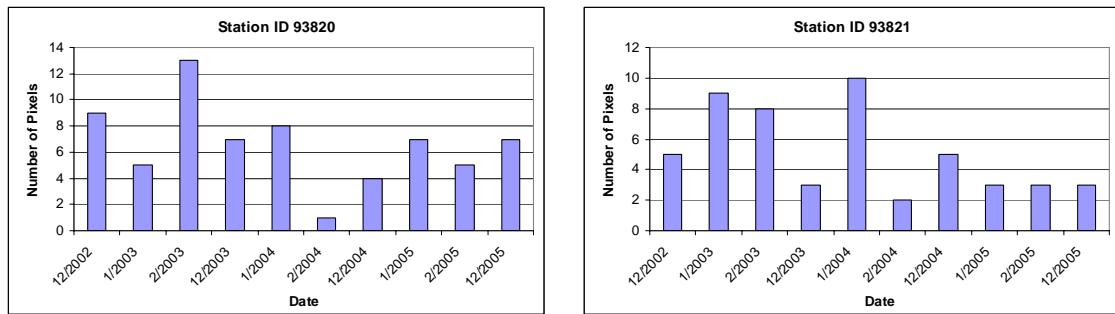


Figure 33: Stations used to collect training data

The distribution in time of the training pixels for each station is shown in the following figure.





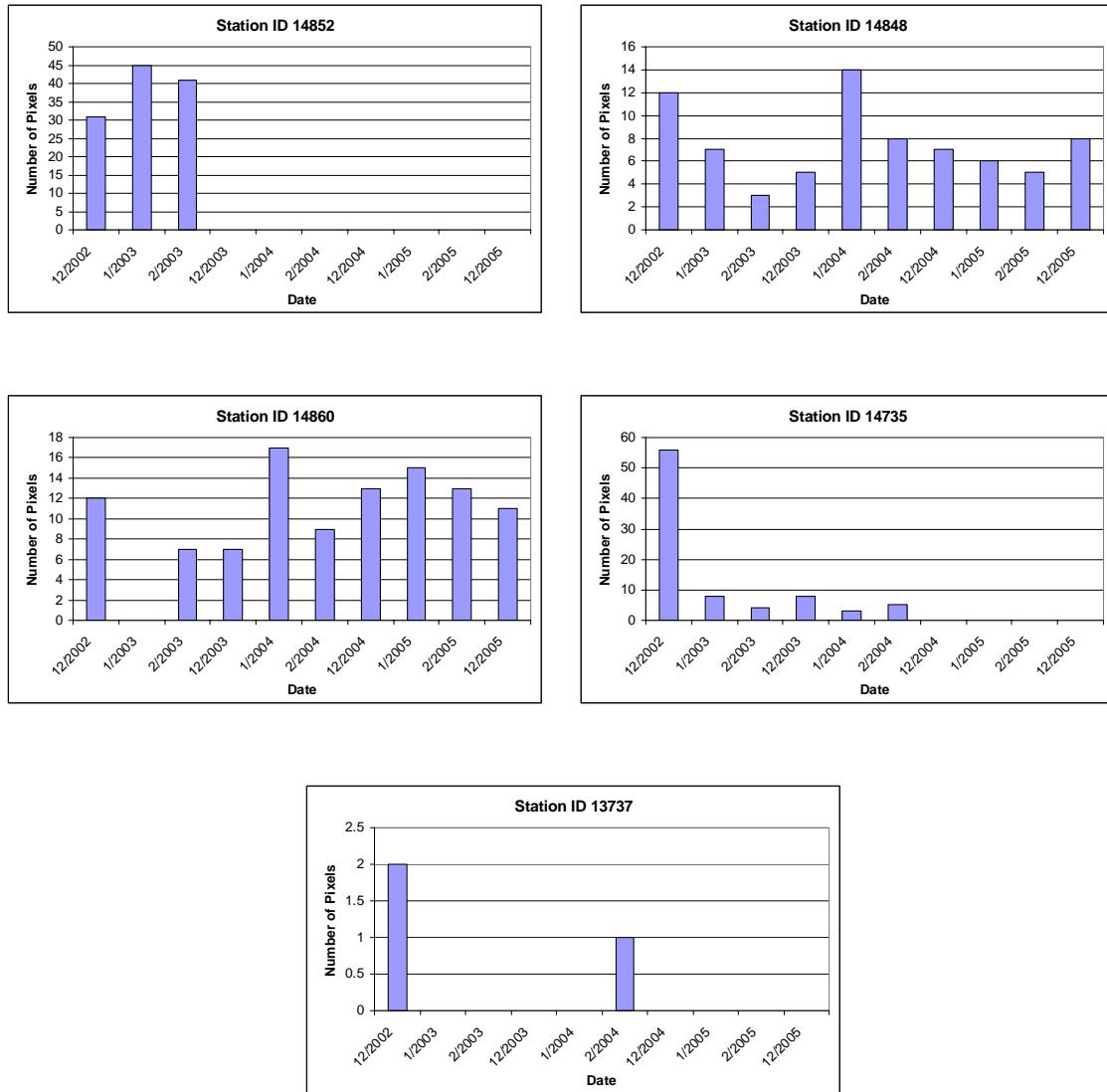


Figure 34: Training pixels distribution

Several stations were used to generate a data set with wide range of topography and climate conditions. The purpose of diversifying the source and characteristics of the data set, is to expose the neural network at its training stage to a maximum number of scenarios that it might be faced later during simulation stage. Topographic information is a relevant factor in snowfall estimation. Snowfall amount varies in time and space. In addition,

altitude is one of the factors that affect climate conditions. In fact, at high elevations the probability to have snowfall is higher. Also, there are no identical winter seasons. Therefore different winter seasons were used to collect this data. The coordinates of the seventeen stations are shown in table 7.

Table 7: Stations's ID and coordinates

	Station ID	Latitude	Longitude
1	13737	36.88	-76.2
2	3816	37.07	-88.77
3	13733	37.33	-79.2
4	3889	37.6	-83.32
5	93817	37.97	-87.55
6	93820	38.03	-84.55
7	93821	38.18	-85.73
8	93819	39.73	-86.27
9	13739	39.88	-75.23
10	94823	40.5	-80.22
11	14737	40.65	-75.43
12	14827	41	-85.22
13	14778	41.25	-76.92
14	14852	41.27	-80.67
15	14848	41.75	-86.17
16	14860	42.08	-80.18
17	14735	42.75	-73.8

6.2 Inputs in the Neural Network Model

As mentioned above, the selection of the neural network inputs should rely on the physics of the problem. Also, it should exist a relationship between the inputs and the subject to investigate. The data used as input to the model is described in the following section.

6.2.1 *Passive microwave*

Passive microwave information from the Advance Microwave Sounding Unit (AMSU) instrument is the main source of information used to retrieve snowfall. Passive microwave allows us to obtain temperature profiles and atmosphere humidity profiles. Its radiation penetrates cirrus clouds and responds directly to water droplets and ice particles in the precipitation layer (Qiu et al. 2005). Also, information from heavy precipitating clouds such as cloud thickness and water/ice content can be inferred with microwave radiances (Scofield and Kuligowski 2003). Also, higher passive microwave frequencies information has been used in other studies related to satellite-based precipitation estimation (Skofronick et al. 2004; Kongoli et al. 2003; Noh et al. 2005; Di Michele and Bauer 2005).

An hourly based temporal analysis was conducted for a selected station in Ohio. The objective was to analyze AMSU measurement behavior during precipitation and non-precipitation. Figure 35 shows the behavior of frequencies 150 and 89 GHz for different days and different amounts of precipitation. Noticeably, brightness temperature drops during precipitation.

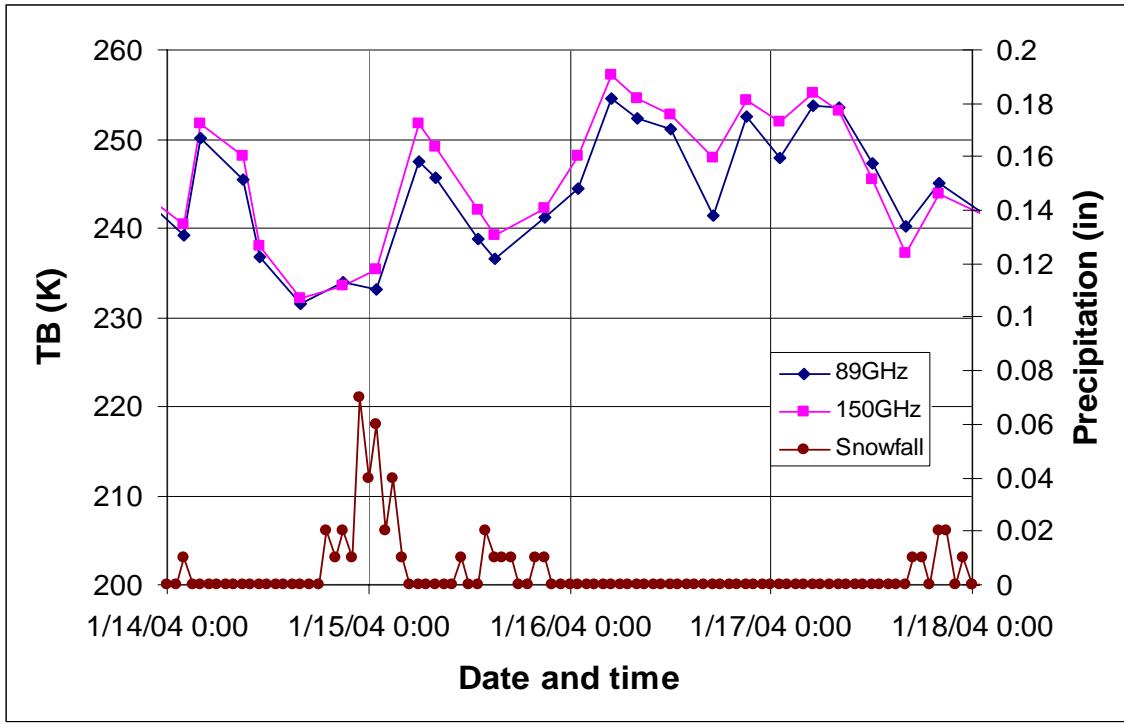


Figure 35: AMSU Temporal Analysis

Two storms were analyzed for the same station in Ohio in 2003 and 2004. AMSU measurements captured at precipitation and non-precipitation times were compared. Figures 36 and 37, illustrate the AMSU measurement behavior at frequencies 89 and 150GHz respectively, for both cases, precipitation and non-precipitation pixels.

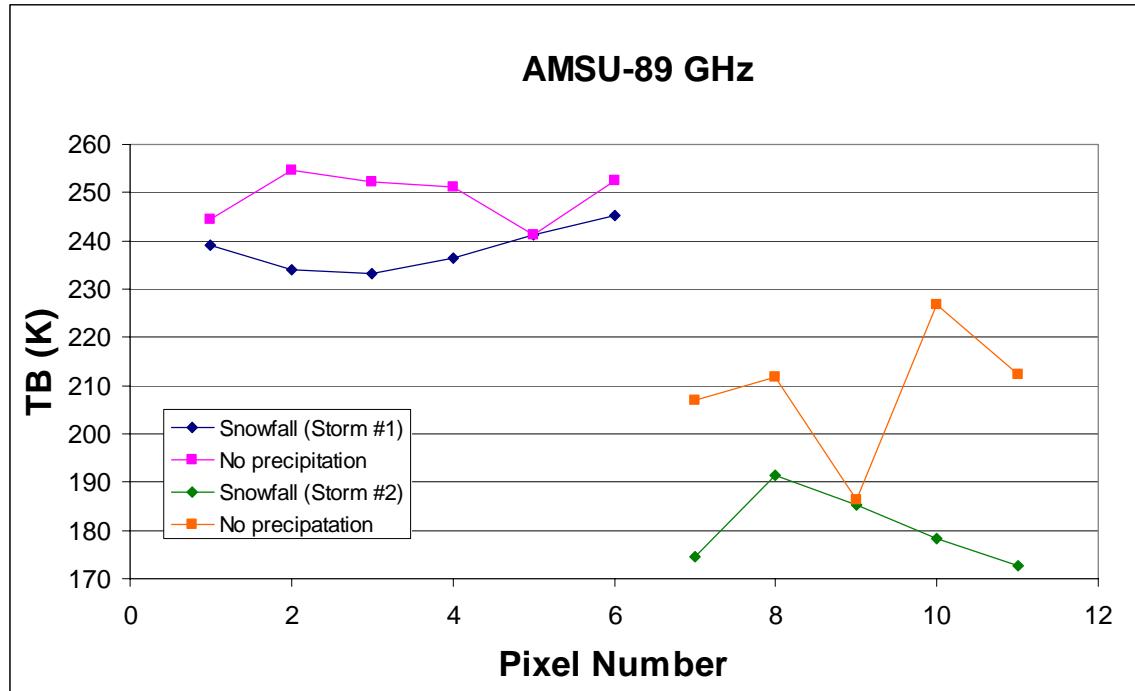


Figure 36: AMSU-89GHz precipitation and non-precipitation pixels comparison

Figure 36 demonstrate that brightness temperatures at 89 GHz are colder in presence of precipitation. Also, there is evidence that brightness temperature at pixel number 5 (storm #1) was not able to distinguish between snowfall and non-snowfall. This may be given due to the fact that this analysis was based on an hourly basis. The storm could have occurred at any time during the hour and AMSU took the measurement at a different time. Using an hourly approach allows us to narrow data noises, but still does not guaranty to have AMSU measurements at the same time of precipitation.

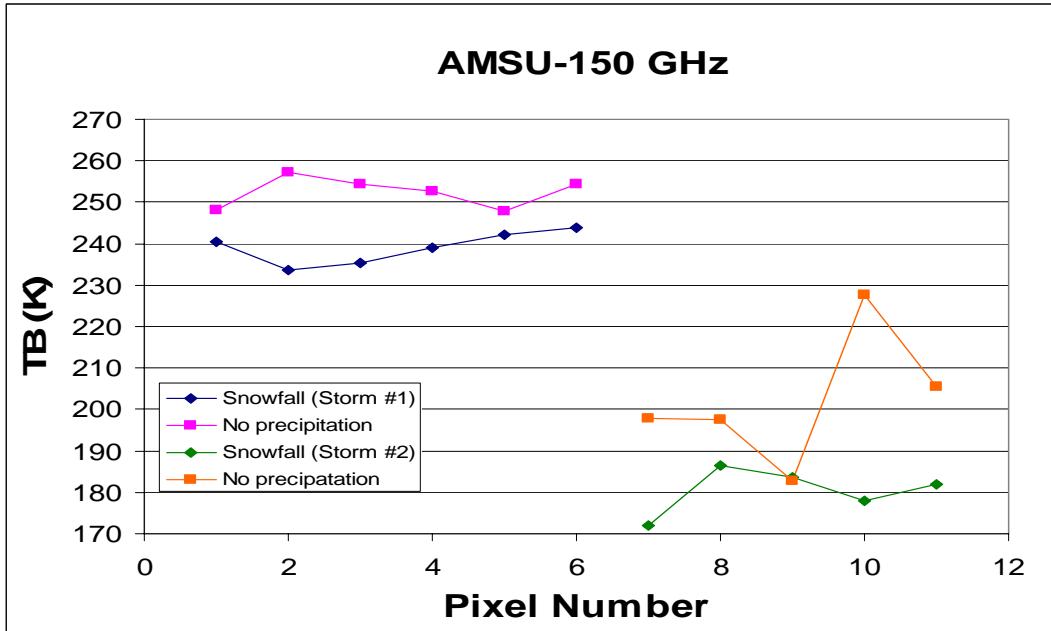


Figure 37: AMSU-150 GHz precipitation and non-precipitation pixels comparison.

Figure 37 also demonstrate that brightness temperatures at 150 GHz are colder in presence of precipitation. Also, on the second storm data set in both 89 and 150 GHz brightness temperatures, pixel number 9 (storm #2) was not distinguished between snowfall and non-snowfall.

In sum, high frequencies from passive microwave have previously shown potential in detecting snowfall events and promise to have an accurate model output. Therefore, AMSU-B high frequencies were selected to be used as model input. In this study a spatial resolution of 25 km was chosen.

6.2.2 Snow cover

Throughout the time several researches have demonstrated the difficulty in detecting snowfall from satellites when snow is present on the ground (Kongoli et al. 2003). Therefore, snow cover was tested as additional input to the model to analyze if it has any effect on the model retrieval accuracy. A product from Microwave Surface and Precipitation Products System (MSPPS) developed by NOAA National Climatic Data Center was used to retrieve snow cover information. The two following equations were used:

$$\Omega_{31.4} = TB_1 - TB_2 - 2.0 \quad 6-1$$

$$\Omega_{89.0} = TB_1 - TB_{16} - 3.0 \quad 6-2$$

Where TB_1 , TB_2 and TB_{16} are the brightness temperatures at 23, 31 and 89 GHz respectively. $\Omega_{31.4}$ and $\Omega_{89.0}$ are scattering indices used to represent the differences between the lowest frequency (AMSU-A, 23.8 GHz) and the higher frequency channels at 31.4 GHz (AMSU-A) and 89.0 GHz (AMSU-B). If $\Omega_{31.4} < 3$ and $TB_1 \leq 215$ K, then the snow type glacial snow is designated. $\Omega_{89.0}$ is used to identify normal snow cover on land and coast. Snow cover is present if the scattering index $\Omega_{89.0} \geq 1$.

6.3 Data Filters

It is important to make sure the data used to train the neural network is most reliable as possible. Therefore, to reduce the noises in the data set, some conditions should be applied in the process of the training data set collection. Any pixel presented to the neural network

as snowfall should meet two conditions. The first one, the pixel should be covered by clouds and the second one, the pixel has to fall into a range of surface temperature. These two conditions are explained in detail below.

6.3.1 Cloud cover

If the pixel is a snowfall pixel, then the pixel is covered by cloud. Therefore to narrow the data noises cloud cover information was used to filter out all cloud free pixels. Cloud cover information from the Local Climatological Data Unedited hourly observations section of the National Climate Data Center (NCDC) is used to retrieve this data. An example of these files is shown in table 8. This table represents hourly information collected for a single station. Column 5 provides hourly cloud cover information for that specific station. For each hour it may appear information of one or more layers. The information of every layer is represented by three letters and three numbers. The first three letters represent the portion of area covered by cloud and the following three numbers represent the cloud height above station level in hundreds of feet. The description of the three letters is described in table 9.

Table 8: Local Climatological Data Unedited Hourly Observations

Date	Time	Station Type	Maint Indic	Sky Conditions	Visibility	Weather Type	Dry Bulb Temp		Wet Bulb Temp		Dew Point Temp		Rel Humd %	Wind Speed (KT)	Wind Dir	Wind Char. Gusts (KT)	Val. for Wind Char.	Station Pressure	Press Tend	Sea Level Pressure	Report Type	Precip. Total
							(F)	(C)	(F)	(C)	(F)	(C)										
01	0012	A020	-	OVC013	10SM	-	-	-	-	-	-	-	4	010	-	0	-	-	-	-	SP	-
01	0054	A020	-	OVC013	10SM	-	64	17.8	62	17	60	16	87	0	000	-	0	28.03	8	249	AA	-
01	0154	A020	-	OVC013	10SM	-	64	17.8	62	17	60	16	87	4	020	-	0	28.02	-	245	AA	-
01	0254	A020	-	OVC013	6SM	BR	63	17.2	62	17	61	16	93	0	000	-	0	28.01	-	242	AA	T
01	0332	A020	-	FEW005 OVC011	6SM	BR	-	-	-	-	-	-	-	0	000	-	0	-	-	-	SP	-
01	0354	A020	-	FEW005 OVC013	10SM	-	63	17.2	62	17	61	16	93	0	000	-	0	28.00	8	241	AA	-
01	0416	A020	-	OVC015	10SM	-	-	-	-	-	-	-	4	010	-	0	-	-	-	-	SP	-
01	0454	A020	-	OVC015	10SM	-	63	17.2	62	17	61	16	93	0	000	-	0	28.00	-	241	AA	-
01	0501	A020	-	OVC013	10SM	-	-	-	-	-	-	-	3	VRB	-	-	-	-	-	-	SP	-
01	0554	A020	-	OVC013	10SM	-	64	17.8	62	17	61	16	90	0	000	-	0	28.01	-	242	AA	-
01	0854	A020	-	OVC011	10SM	-	64	17.8	63	17	62	17	93	4	340	-	0	28.01	3	248	AA	-
01	0754	A020	-	OVC011	10SM	-	65	18.3	63	17	62	17	90	0	000	-	0	28.02	-	248	AA	-
01	0836	A020	-	OVC015	10SM	-	-	-	-	-	-	-	5	350	-	-	-	-	-	-	SP	-
01	0854	A020	-	OVC015	10SM	-	67	19.4	64	18	62	17	84	7	330	-	0	28.02	-	248	AA	-
01	0854	A020	-	OVC019	10SM	-	68	20	64	18	62	17	81	6	340	-	0	28.02	2	248	AA	-
01	1054	A021	-	OVC021	10SM	-	70	21.1	65	18	62	17	78	4	310	-	0	28.00	-	242	AA	-
01	1154	A021	-	OVC023	10SM	-	72	22.2	66	19	63	17	73	4	010	-	0	27.98	-	234	AA	-
01	1254	A021	-	BKN026 OVC031	10SM	-	73	22.8	67	19	63	17	71	5	140	-	0	27.97	8	228	AA	-
01	1345	A021	-	SCT026 BKN035	10SM	-	-	-	-	-	-	-	4	VRB	-	-	-	-	-	-	SP	-
01	1354	A021	-	SCT026 BKN037	10SM	-	76	24.4	69	20	65	18	89	8	140	-	0	27.94	-	218	AA	-
01	1423	A021	-	BKN028 OVC037	10SM	-	-	-	-	-	-	-	8	170	-	-	-	-	-	-	SP	-
01	1430	A021	-	SCT028 BKN035	10SM	-	-	-	-	-	-	-	5	150	-	-	-	-	-	-	SP	-
01	1454	A021	-	SCT028 BKN035	10SM	-	76	24.4	69	20	65	18	89	7	180	-	0	27.92	-	205	AA	-
01	1554	A021	-	SCT026 BKN037	10SM	-	78	25.6	69	21	65	18	84	7	180	-	0	27.90	8	198	AA	-
01	1654	A021	-	SCT029 SCT034	10SM	-	77	25	69	21	65	18	88	8	160	-	0	27.89	-	195	AA	-
01	1754	A021	-	BKN031	10SM	-	75	23.9	68	20	65	18	71	5	120	-	0	27.90	-	199	AA	-
01	1923	A021	-	BKN029 OVC034	10SM	-	-	-	-	-	-	-	6	100	-	-	-	-	-	-	SP	-
01	1844	A021	-	BKN031	10SM	-	-	-	-	-	-	-	5	130	-	-	-	-	-	-	SP	-
01	1854	A021	-	BKN031	10SM	-	72	22.2	67	19	64	18	76	4	140	-	0	27.91	3	205	AA	-
01	1954	A021	-	OVC037	10SM	-	71	21.7	67	19	64	18	79	3	150	-	0	27.93	-	210	AA	-
01	2054	A022	-	FEW034 OVC047	10SM	-	69	20.6	66	19	65	18	87	4	150	-	0	27.94	-	215	AA	-
01	2154	A022	-	FEW027 OVC036	10SM	-	69	20.6	66	19	65	18	87	6	160	-	0	27.94	1	217	AA	-
01	2254	A022	-	FEW012 BKN040 OVC048	10SM	-	68	20	66	19	65	18	90	5	180	-	0	27.94	-	216	AA	-
01	2354	A022	-	BKN013 BKN019 BKN046	10SM	-	67	19.4	66	19	65	18	93	4	200	-	0	27.94	-	215	AA	-
02	0054	A020	-	BKN015	9SM	-	66	18.9	65	18	64	18	93	3	210	-	0	27.93	8	211	AA	-
02	0118	A020	-	SCT015	7SM	-	-	-	-	-	-	-	0	000	-	-	-	-	-	-	SP	-
02	0154	A020	-	SCT014	6SM	BR	65	18.3	64	18	64	18	97	0	000	-	0	27.92	-	204	AA	-
02	0205	A020	-	BKN014	6SM	BR	-	-	-	-	-	-	0	000	-	-	-	-	-	-	SP	-
02	0254	A020	-	OVC012	7SM	-	66	18.9	65	18	64	18	93	0	000	-	0	27.90	-	200	AA	-

Table 9: Cloud Covered Categories

	Fraction	Description
SKC	0/8	Clear Sky
FEW	0/8-2/8	Few
SCT	3/8-4/8	Scattered
BKN	5/8-7/8	Broken
OVC	8/8	Overcast
TCU		Towering Cumulus
VV		Vertical Visibility
CLR		Clear for below 12000 feet

6.3.2 *Surface temperature*

Surface temperature plays an important role in snowfall. Precipitation takes the form of snowfall when the temperature falls below zero. In this project a threshold was assumed to discriminate between rain and snowfall (figure 38). This threshold was selected to ensure the neural network is trained with the correct data.



Figure 38: Surface Temperature Threshold

The ground surface temperature was derived from AMSU A brightness temperatures. The algorithm used to derive surface temperatures is part of the Microwave Surface and Precipitation Products System (MSPPS) developed by NOAA National Climatic Data Center, by using the following linear equation:

$$Ts = 2.9079 * 10^2 - (8.5059 * 10^{-1} - 1.9821 * 10^{-3} TB_1)TB_1 + (6.1433 * 10^{-1} - 2.3579 * 10^{-3} TB_2)TB_2 - (1.1493 - 5.4709 * 10^{-3} TB_3)TB_3 - 1.50 * 10(m - 5.40 * 10^{-1}) \quad 6-3$$

Where TB1, TB2, and TB3 are AMSU brightness temperatures at 23, 31 and 50 GHz respectively and m is the cosine of AMSU zenith angle.

6.4 Validation Study Area

Two independent storms of 4-5 days with intense snowfall occurrences were selected to simulate and evaluate the model efficiency. Figure 39 shows the location of the storms.

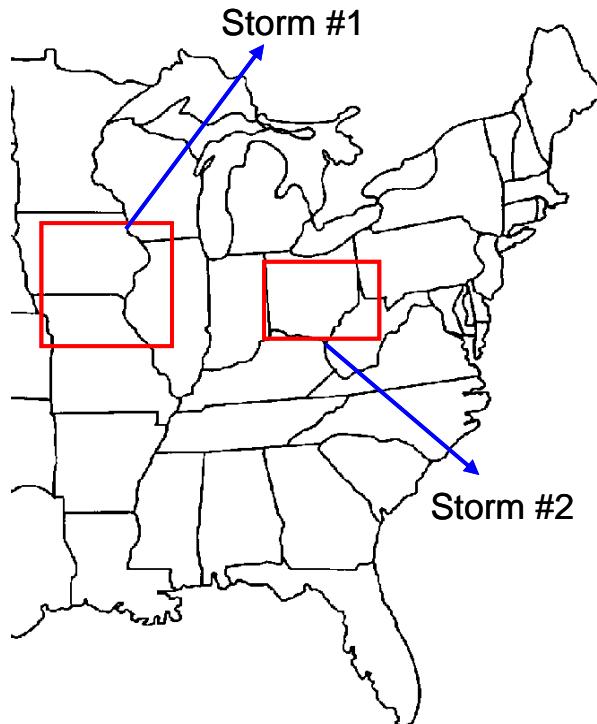


Figure 39: Validating Study Area

Storm #1 occurred in January 14-17, 2003. Its exact location is longitude 89W to 96W and latitude 37N to 43N. Storm #2 occurred in January 23-28, 2004. Its exact location is longitude 78W to 85W and latitude 37N to 41N.

7. Methodology, Results and Discussion

Usually, researchers use a range of methods based on their best understanding of how to achieve successful results. However, in any weather related study, the methodology used has to be selected based on scientific knowledge and understanding of the specific characteristics of the problem. Moreover, it is important to clearly identify the objectives, limitations, goals and priorities of the project. It is also important to ensure that the developed technique can be easily adapted in the future under a variety of topographic and climatic conditions. In this chapter, an innovative approach has been developed and evaluated to detect and estimate snowfall from passive microwave satellites.

Precipitation retrieval algorithms over land using microwave measurements have almost exclusively been limited to rainfall (Kongoli et al. 2003). Several factors influence a good retrieval of atmospheric related parameters including the detection and estimation of snowfall. Some of them are biased or contain errors in the initial conditions that could be caused by spatial and temporal limitations from the truth data. Actually, there is no operational product which provides hourly and gridded snowfall information. Therefore, the use of station-based information is necessary. The stations are located far apart which makes the information sparse. Consequently, one station represents a full area of 25 Km by 25 Km. This is one of the spatial limitations mentioned above. Information from one station is not representative for the whole area. Also, these stations report snowfall information hourly. And, information from stations is compared with information from satellite measurements taken in one shot lasting a couple of seconds. Again, here is the

temporal limitation mentioned above. The station might report snowfall accumulation for that specific period but at the time in which the satellite took the measurement, precipitation might not be present. All of these factors demonstrate the weakness of the truth data.

In this study, a Neural Network based technique was developed for detecting and estimating snowfall. As discussed in chapter 5, this technique was selected for its capability to solve non-linear and complex problems, compared to traditional modeling tools. Neural networks are able to perform complex functions that are difficult to approach with conventional techniques. Due to the multiple potential sources of error, a single neural network is not enough to predict snowfall because it might be trained with a large number of inaccurate data. Therefore, the use ensemble Neural Networks is introduced to this methodology.

An ensemble neural network consists of a set of individually trained Neural Networks (Optiz and Maclin 1999). The Neural Network ensemble approach provides a better approximation and interpolation than an individual neural net and, most importantly, a better accuracy of climate simulation (Fox-Rabinovitz, et al. 2006). In the ensemble Neural Network approach each net should be trained differently from each other. Obviously, combining several identical nets will not be advantageous. Therefore, the nets can differ from each other by training them with (a) different subsets of data, (b) different network architectures, or (c) different initial weights. In Marclin and Shavlik's article

(1995), the idea to initialize the parameters of neural networks over a much larger range of values than normal, was explored. One of the advantages of the Neural Network ensemble (NNE) approach is the additional or complementary information obtained about the target mapping. Therefore, a robust Neural Network system was designed, calibrated and validated to detect and estimate snowfall.

The neural network toolbox from MATLAB was used in this research. The internal parameters used in the systems were changed to find the best generalization of the neural network for this specific problem.

The approach selected in this study is divided into two parts: (a) snowfall detection and (b) snowfall classification. In the first part (snowfall detection), a system of Neural Networks (NN1) was designed to detect snowy pixels. In the subsequent part (snowfall classification), a second Neural Network system (NN2) was developed to categorize the intensity of pixels detected as snowy in the first step into light, moderate or heavy. Figure 40 shows a schematic presentation of the two-steps approach.

In the following sections a detailed description of each part is given. This includes the steps followed in each part, the data set used, the neural network architecture optimization, the internal parameters of the neural network chosen and the assumptions made.

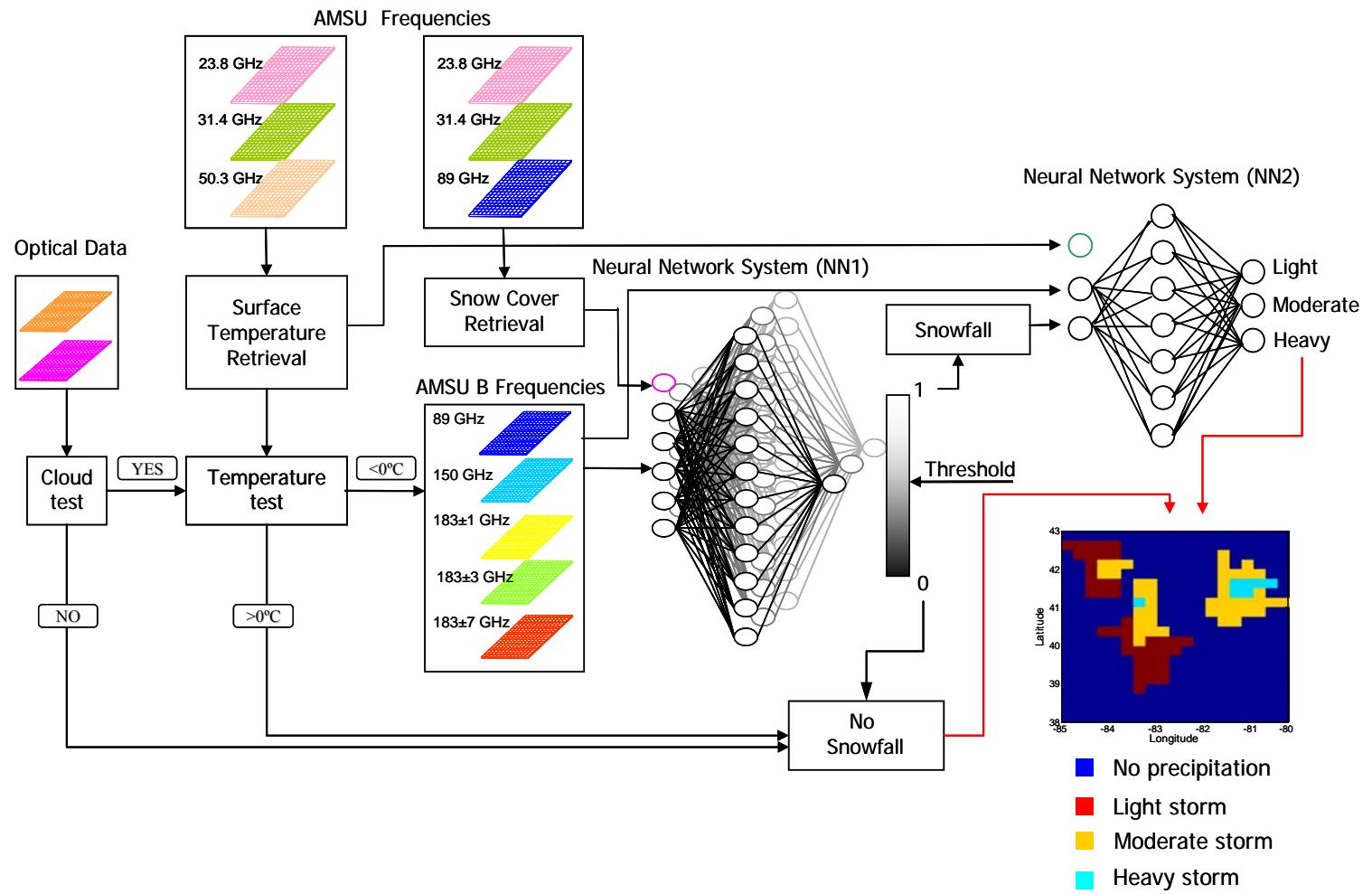


Figure 40: Overall approach for snowfall detection and estimation

7.1 Part A: Snowfall Detection

Obviously, any precipitating (rainfall or snowfall) pixel should be covered by cloud. And, since this study focus only on snowfall, the surface temperature has to be considered in discriminating between rainfall and snowfall. To avoid any mislabeled observation in the data set, cloud and surface temperature filters have been applied to the collected data. The application of these two filters has reduced dramatically the data set size. Indeed, it is difficult to find precipitation pixels matching the same acquisition time with AMSU and passing the two filters conditions. Therefore, these is the reason for extending the temporal and spatial coverage of these study to 17 different stations and four winter seasons as mentioned above in chapter 6.

7.1.1 Data noises Filters:

7.1.1.1 Cloud Cover Information:

First, cloud cover information is used to filter out pixels that are not covered by clouds. The pixels covered by clouds enter to the second filter, surface temperature test. As mention in the data acquisition chapter, cloud cover information is collected from ground stations. In the future work, if this product becomes operational a product from IR/VIS (GOES) will used for retrieving cloud cover.

7.1.1.2 Surface Temperature:

Snowfall precipitation type occurs when surface temperature is below 0°C (Figure 41).

Therefore, ground surface temperature is used to filter out all cloudy pixels with positive surface temperature.

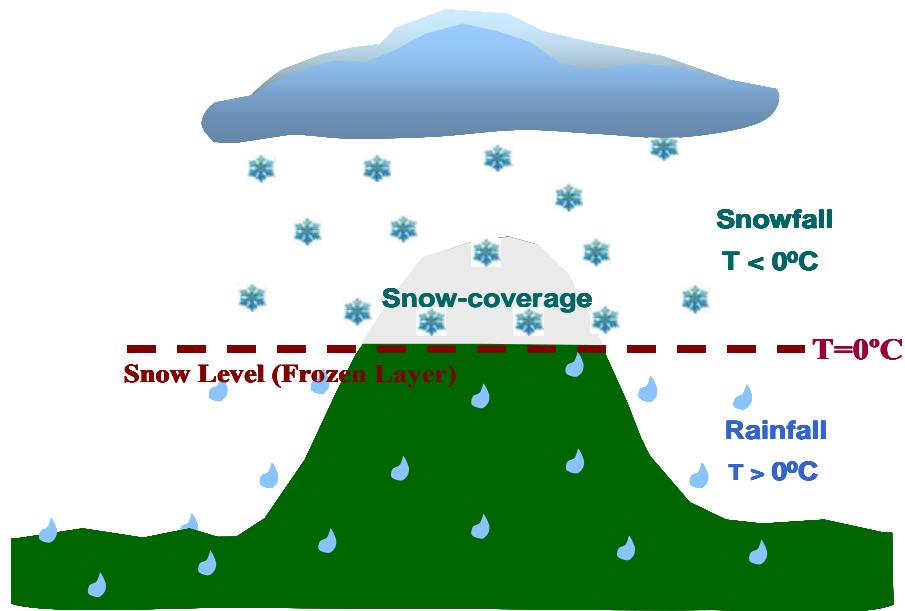


Figure 41: Snowfall detection.

Ground surface temperature information is derived from AMSU-A brightness temperatures. A detail description on the retrieval of this data set is given in chapter 6.

7.1.2 Model Inputs

After filtering out all pixels with no cloud cover and temperature over zero degrees Celsius; the selection of the model was performed. The selection of the variables to use as model inputs, is based on the physics the problem and the relationship of the variable with the corresponding model output. In this study, AMSU-B, surface temperature, snow cover and the station ID number have been tested as inputs. The results obtained in these tests are showed in the neural network optimization parameters section.

7.1.3 Neural Network

Once the snowfall area is detected and the variables to use as model inputs are selected, the pixels are presented to the Neural Network System 1 (NN1).

The neural network structural and training parameters such as activation functions between layers, number of hidden layers and number of nodes in each layer, gradient, and momentum, were optimized by performing one hundred runs for each possible configuration.

A total of 229 snowfall and no-snowfall pixels were used in this optimization process. Each run used a randomly selection of 50% of the data set for training, 25% for validation and the remaining 25% for testing. This process adjusts the connection weights between nodes in order to decrease the error (difference between model outputs and their respective targets). At each training epoch, the error is computed and fed backwards to adjust the network connections.

In the following subsections, a detail description of the optimization of the different parameters is given. The selection of the best parameters is made by looking at the configuration that reaches the highest average accuracy of the hundred runs. Also, in this study, ANOVA (Analysis of Variance) model is used to test for significant differences of results obtained among groups (different inputs, different number of hidden layers and node, activation functions, etc). This technique depends on the number of groups to compared, number of samples in each group (n) and the significant level selected. In this study, the significant level selected was 0.05. Since the variability was measured for two groups of 100 nets, then $n_1=n_2=100$ and the numerator degrees of freedom (df_1) is 1. Therefore, by applying the following equation the denominator degrees of freedom (df_2) can be retrieved.

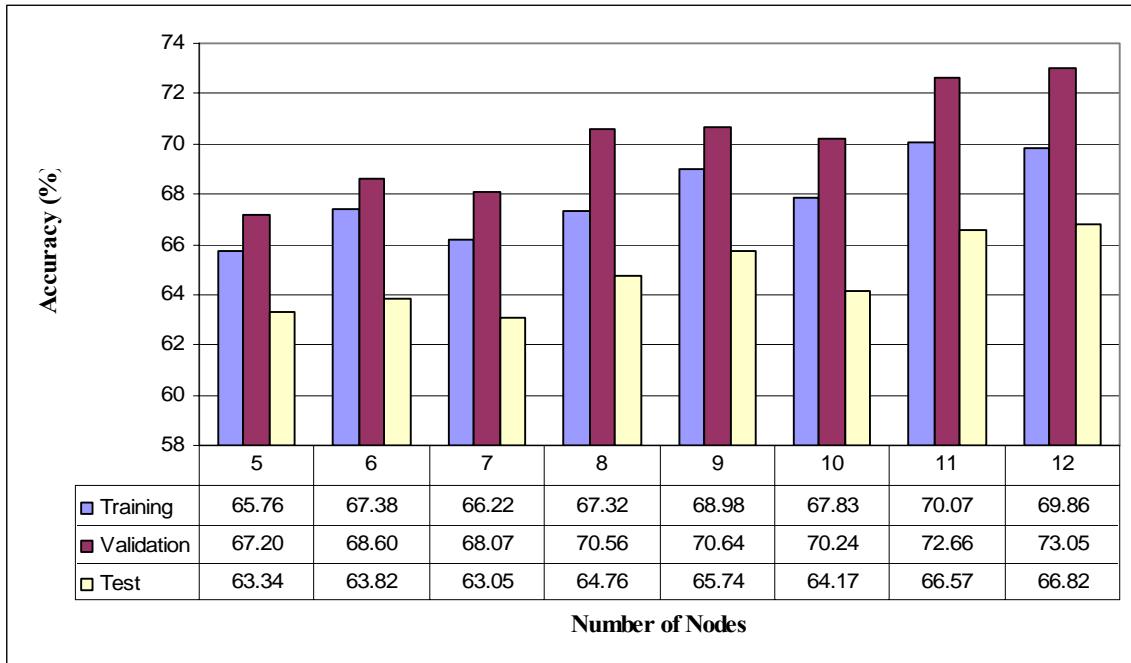
$$df_2 = n_1 + n_2 - 2 = 100 + 100 = 198$$

7-1

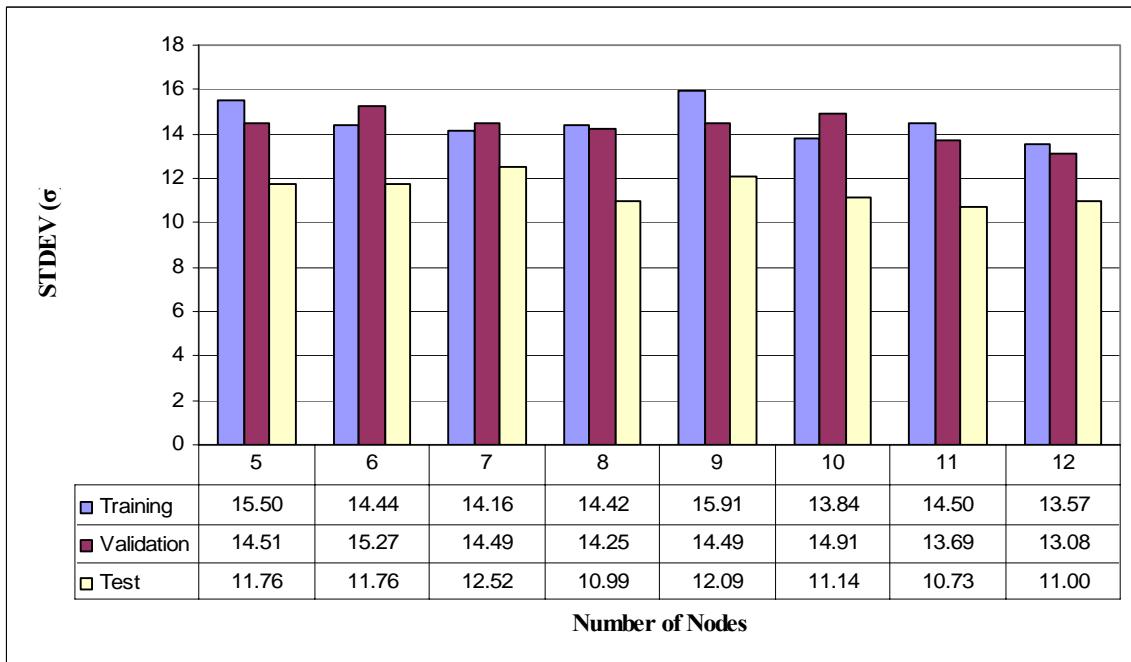
The critical value of F is retrieved using the numerator and the denominator degrees of freedom in the table found in appendix H. F_c is 3.89. An F value greater than F_c , represents a significant difference between the groups.

7.1.3.1 Number of Hidden Layers

The layers between input and output layers are called hidden layers. The number of hidden layers and their respective nodes are assigned experimentally. The number of hidden nodes should be large enough to ensure a sufficient number of degrees of freedom for the network function and small enough to minimize the risk of loss of generalization ability of the network (Ghedira et al. 2000).

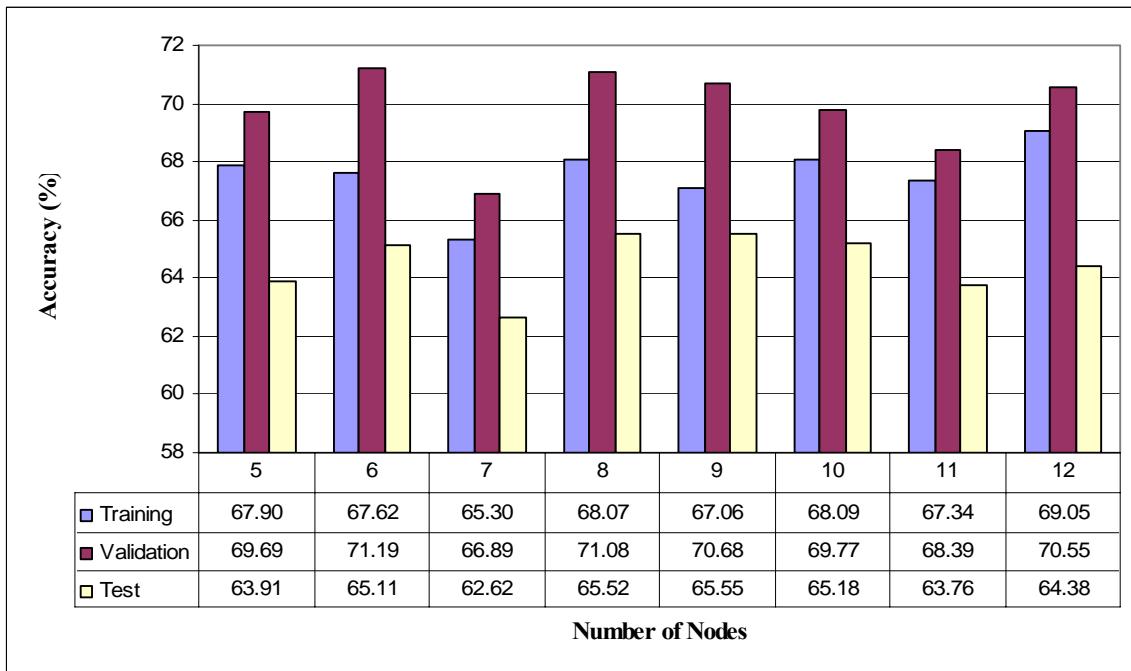


(a)

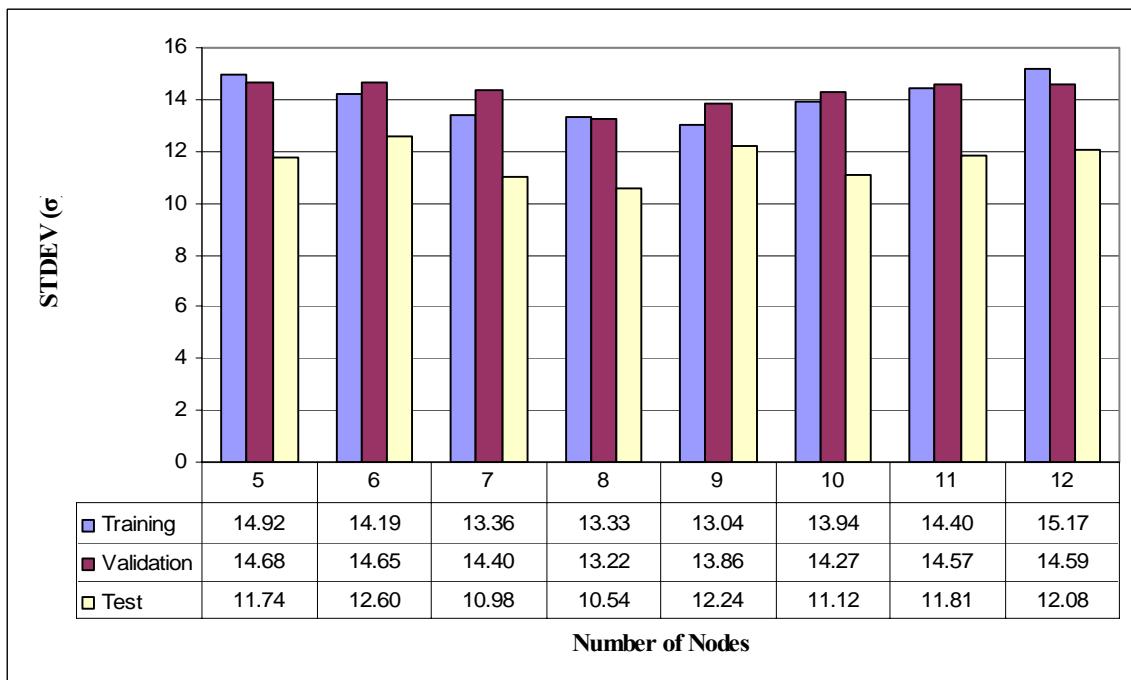


(b)

Figure 42: (a) Average Accuracy and (b) Standard Deviation of the Model using One Hidden Layer



(a)



(b)

Figure 43: (a) Average Accuracy and (b) Standard Deviation of the Model using Two Hidden Layer

In this study one and two hidden layers with different number of nodes have been tested.

Figure 42 and 43 show the model average accuracy and the standard deviation output using one and two hidden layers respectively.

According to figures 42 and 43 the configuration that gives the best average accuracy in the test data set (66.82 %), is one hidden layer with 12 nodes. But, also other configurations show similar accuracy. Therefore, ANOVA model was used to test if there is any significant difference between the different architectures. The following table shows the results of the ANOVA tests. For detailed information of these tests see appendix A.

Table 10: ANOVA results when using different number of hidden layers and nodes

Difference between	F	F>Fc
1HL 12 nodes and 1HL 11 nodes	0.027391	No
2HL 6 nodes and 2HL 8 nodes	0.062985	No
2HL 9 nodes and 2HL 10 nodes	0.049963	No
2HL 8 nodes and 2HL 9 nodes	0.000343	No
1HL 5 nodes and 1HL 12 nodes	4.671749	Yes

According to the ANOVA tests, there is no significant difference when using the configurations that have similar accuracy in the data set. The only configuration that shows significant difference is the one hidden layer with 5 and 12 nodes. Therefore, an architecture one hidden layer with 12 nodes, which has the highest accuracy, is selected and used for the following tests.

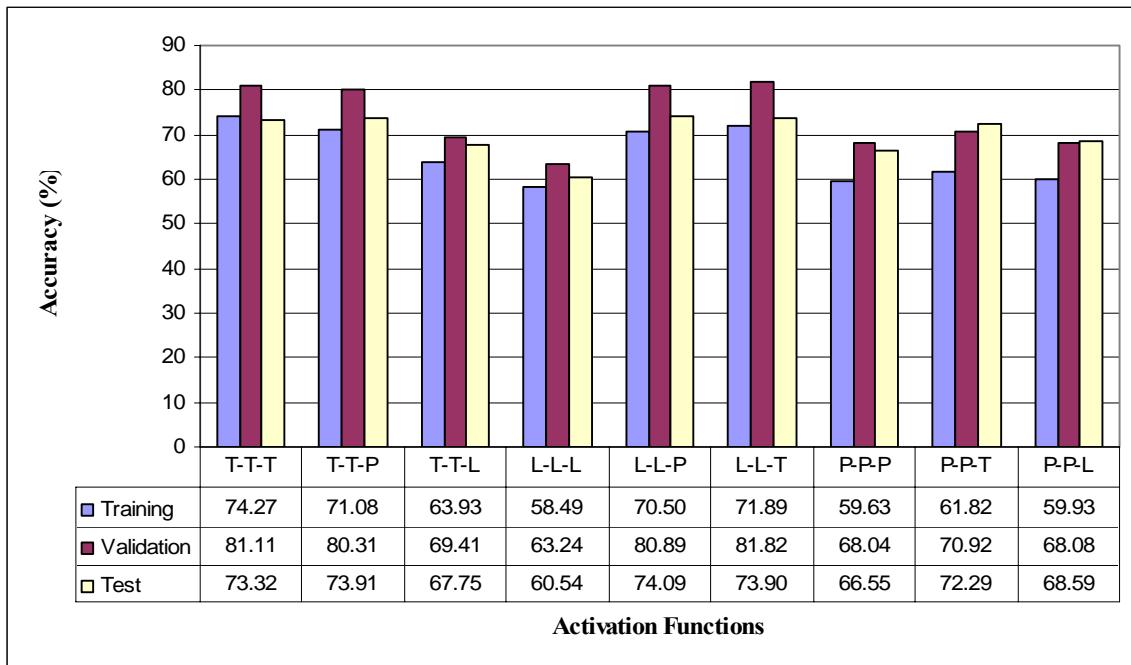
7.1.3.2 Activation Functions

As reminder, activation functions are used in the neural network structure to add non-linearity behavior into the model. The effect of the activation functions used in the training on the model output is illustrated in figure 44. More details about these activation functions are presented in chapter 5. The graphs presented in this figure show the average accuracy and the standard deviation (STDEV) obtained using combination of different activation functions (tansig (T), purelin (P) and logsig (L)) and one hidden layer with twelve nodes.

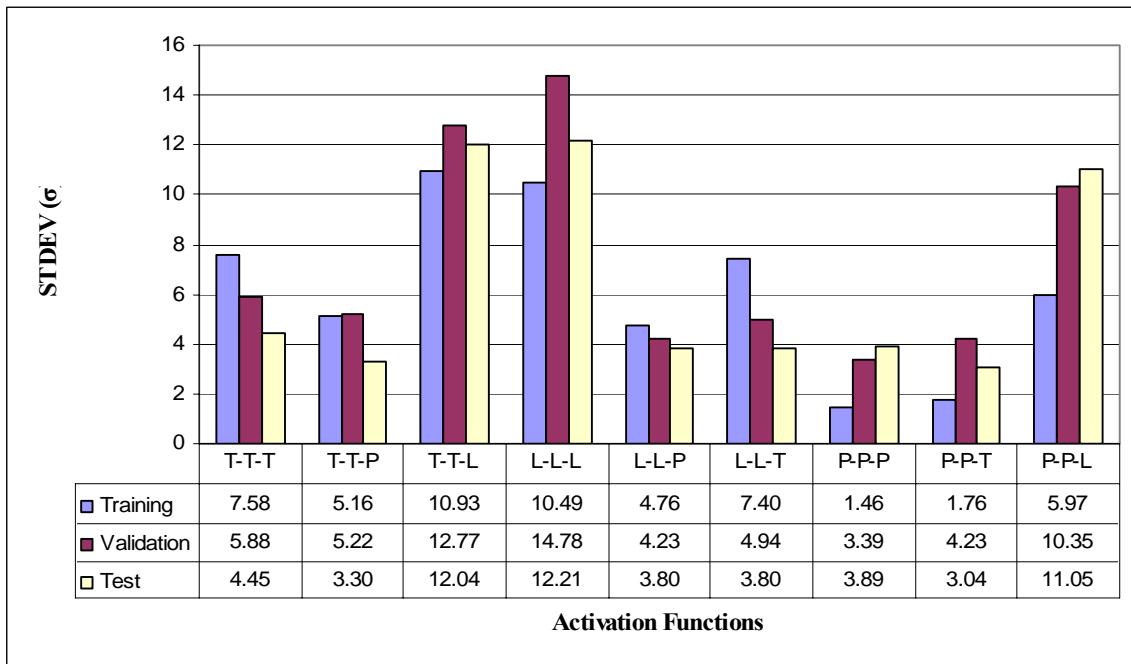
These graphs show that there is not a high variation on the overall accuracy of the model when using different combination of the activation functions. But, they also show that the standard deviation varies significantly. Therefore, ANOVA test is used to compare the difference in the accuracy at the different combinations of activation functions. The results are shown in table 11 (for detailed information of the ANOVA tests results see appendix B).

Table 11: ANOVA test results using different combinations of activation functions

Difference between	F	F>Fc
TTT and TTP	0.056831	No
TTT and LLL	111.1737	Yes
LLP and LLT	0.000259	No
TTL and PPT	13.39296	Yes
PPL and TTL	0.261281	No
PPT and TTT	12.63042	Yes



(a)



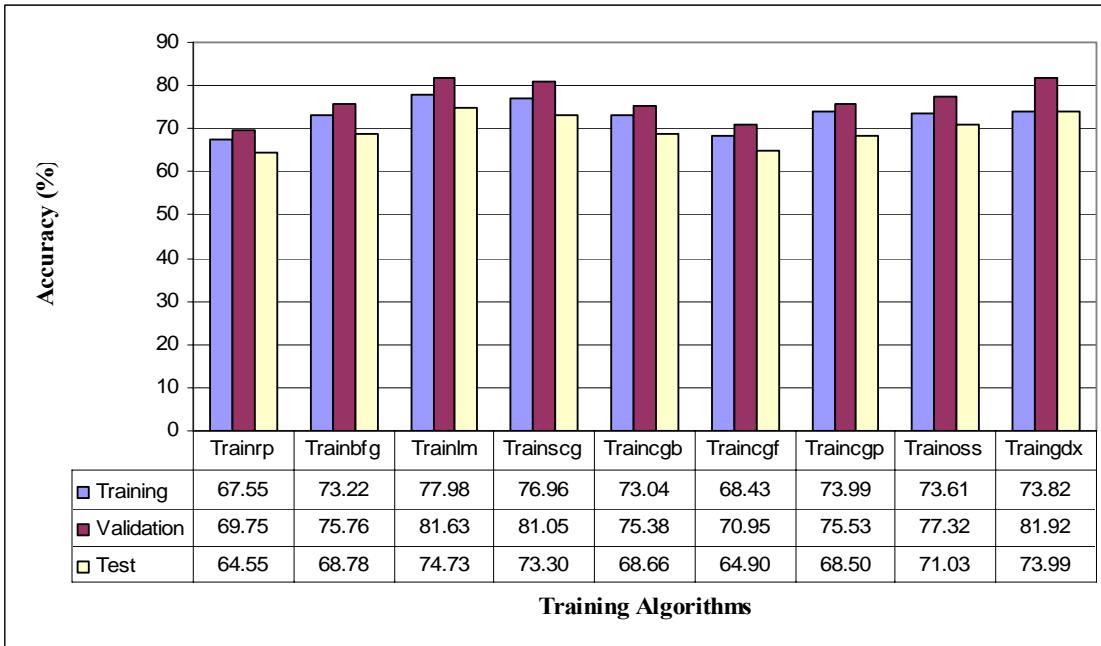
(b)

Figure 44: (a) Average Accuracy (b) STDEV of the Model using different activation functions.

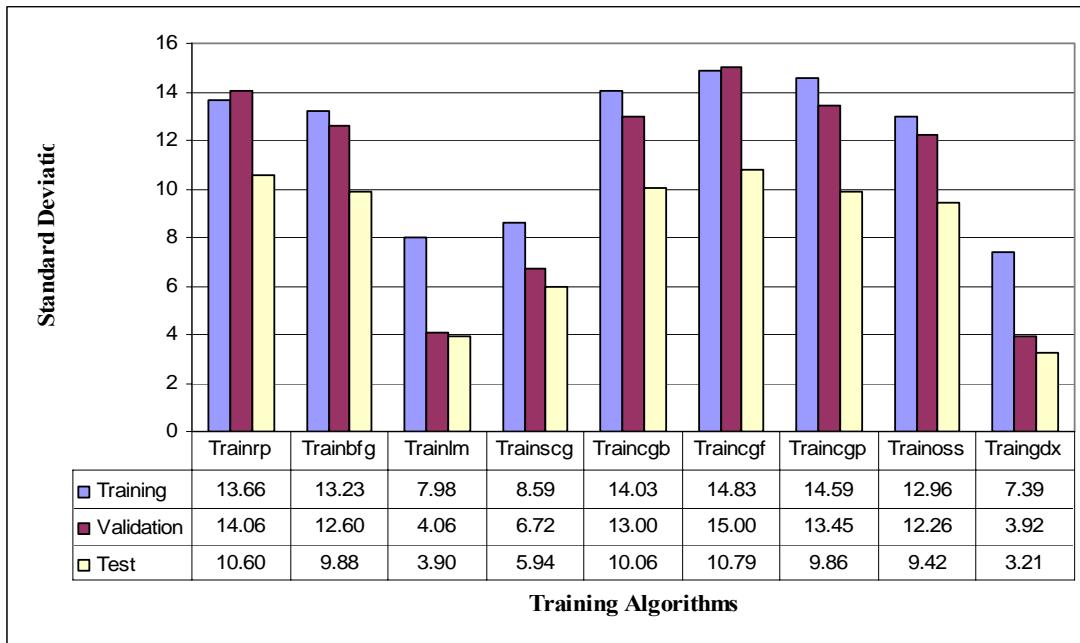
These ANOVA tests show a significant difference between TTT and LLL were TTT shows to have higher accuracy. Also, there is a significant difference between TTL and PPT, having PPT higher accuracy. And finally, ANOVA test also shows a significant difference when comparing PPT and TTT, having TTT higher accuracy. According to the graphs the use of T-T-T as activation functions in the training give one of the highest average accuracy (73.32%) in the model output. Also, the standard deviation is small (4.45). In addition, the activation function tansig allows the output to fall in a wider range [-1 1]. Therefore, TTT configuration is selected for NN1.

7.1.3.3 Training Algorithms

As the activation functions, the selection of the training algorithm to be used is critical. The time consumed to train the model depends on the training algorithm used. Figure 45 illustrates the average accuracy and standard deviation using different training algorithms. This figure shows about 10% of variation when using different training algorithms. Also, it shows variation in the standard deviation. ANOVA model is used to find if the variation in the accuracy is significant. Table 12 shows the ANOVA tests results. Detailed information for these ANOVA tests can be found in appendix C.



(a)



Training Algorithms

(b)

Figure 45: (a) Average Accuracy (b) STDEV of the Model using different training algorithms.

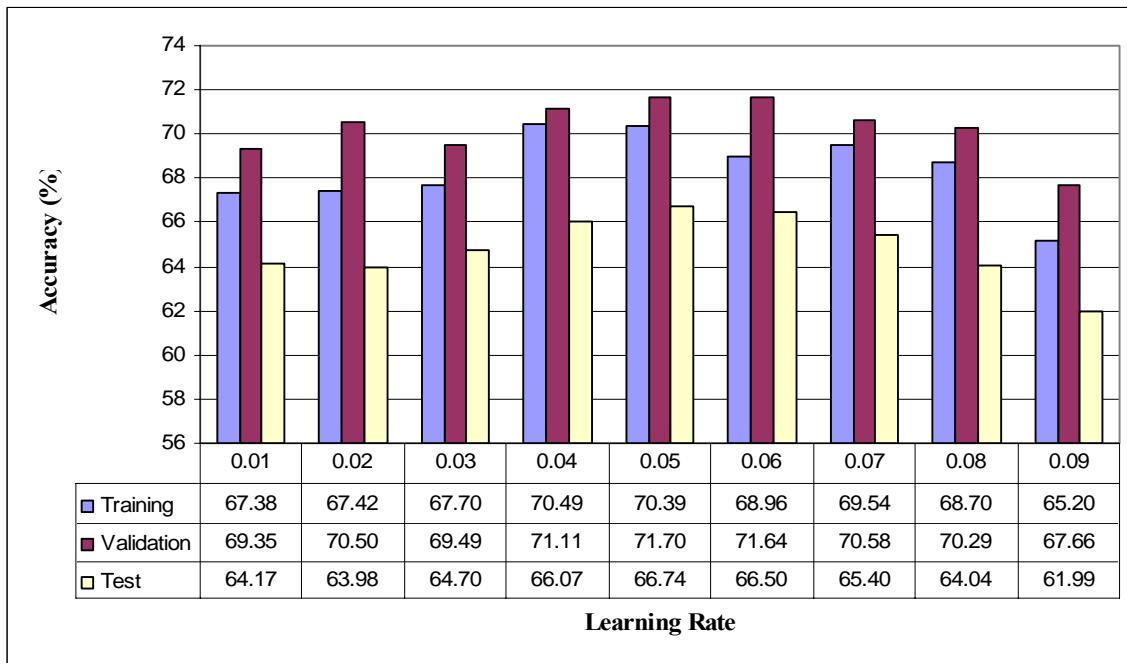
Table 12: ANOVA test results using different training algorithm

Difference between	F	F>Fc
RP and LM	80.90253	Yes
LM and GDX	2.164398	No
BFG and CGB	0.007463	No
SCG and OSS	4.170782	Yes
OSS and GDX	8.85737	Yes
CGF and RP	0.046723	No
CGP and LM	34.63038	Yes
LM and SCG	4.062031	Yes
LM and GDX	2.164398	No

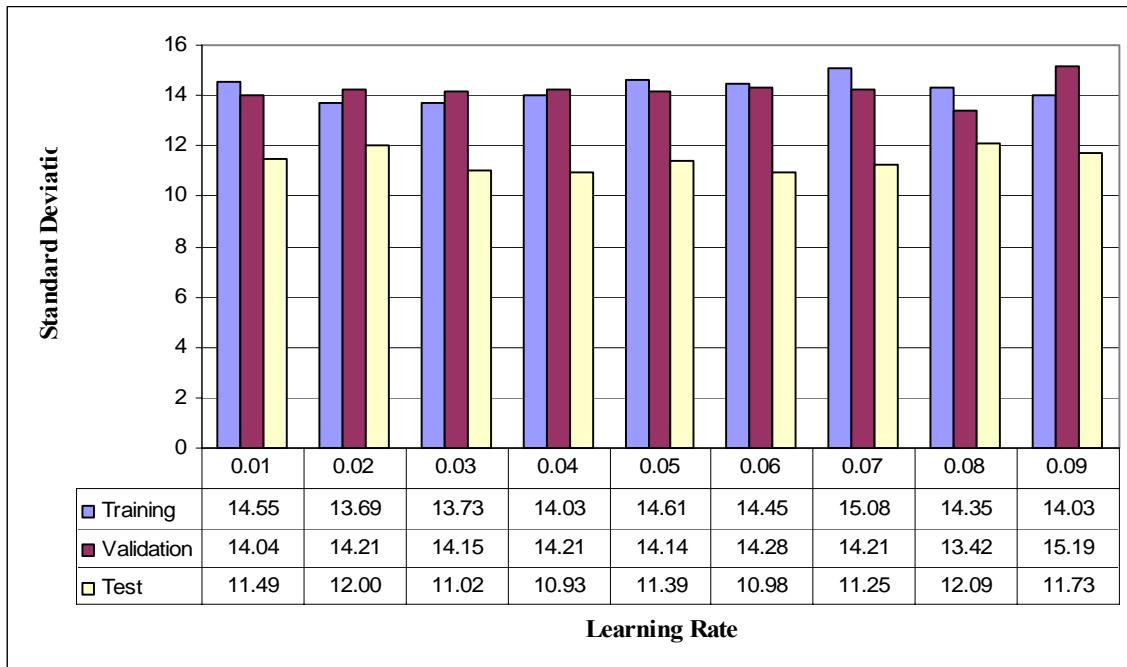
According to figure 45, Lenverg Maquardt algorithm gives the highest average accuracy (74.73 %) model output and the standard deviation falls into the small values, 3.90. Also, table 12 shows a significant difference between some of the training algorithms (RP and LM, SCG and OSS, OSS and GDX, CGP and LM, and LM SCG). By looking at figure 45 and analyzing the ANOVA results, LM is the algorithm that makes a high difference from the others. Therefore, this training algorithm is selected to train the nets.

7.1.3.4 Learning rate

As explained in chapter five, the model behavior during the training depends on the learning rate selected. As mentioned above, sometimes the network can oscillate and become unstable if the learning value is too high. On the other hand, if the learning rate value is too low the algorithm takes much longer to converge.



(a)



(b)

Figure 46: (a) Average Accuracy (b) STDEV of the Model using different learning rate values.

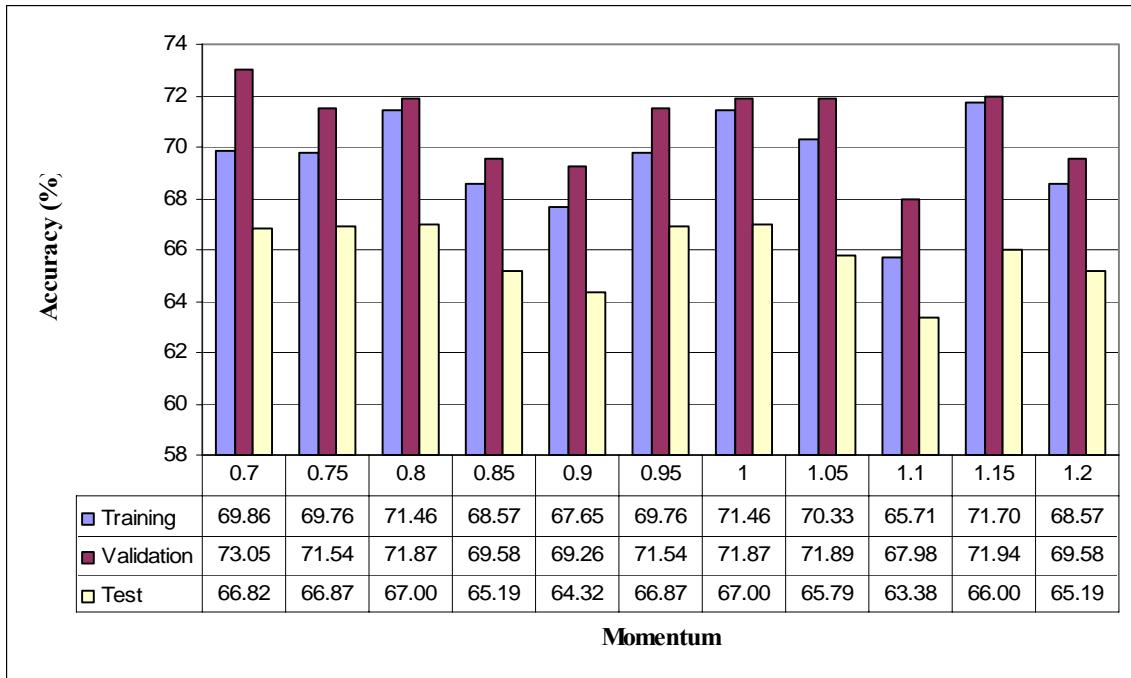
Figure 46 shows the variability of the neural network behavior at different learning rates in terms of average accuracy and standard deviation. As discussed above, the model was tested using one hidden layer with twelve nodes, activation functions T-T-T and Lenverg Maquardt (LM) algorithm.

These graphs show a peak in the accuracy at learning rate of 0.05. Also, the standard deviation does not vary much. Therefore, a learning rate of 0.05 is selected.

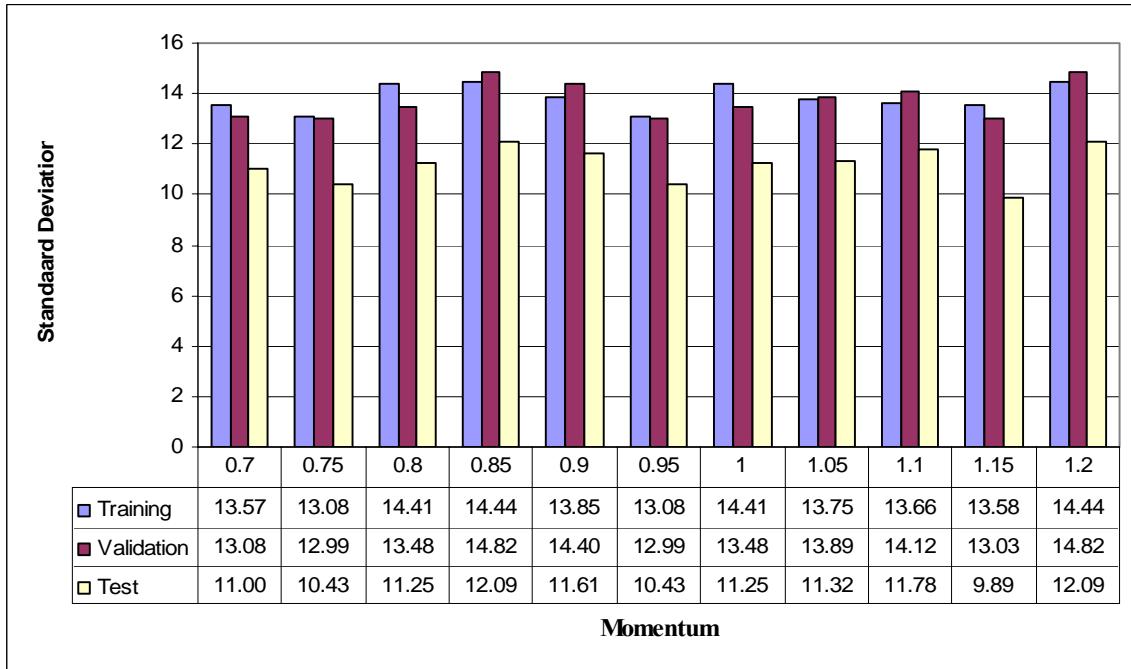
7.1.3.5 Momentum

The use of momentum in the training of a backpropagation algorithm can made the convergence of the trained algorithm faster and can avoid that the training stops at local minima on the error space. In this study, different momentum values were tested. Figure 47 illustrates the average accuracy and standard deviation of the trained neural network at different momentum values.

Figure 47 shows that the variation of the neural network accuracy does not vary much at the different momentum values (63.38 % - 67.00 %). The values that show high accuracies are 0.7, 0.75, 0.8, 0.95 and 1.0. Therefore, ANOVA test is used to find if there is any significant difference between them. The results are shown in table 13. Detailed description of these results can be found in appendix D.



(a)



(b)

Figure 47: (a) Average Accuracy (b) STDEV of the Model using different momentum values.

Table 13: ANOVA test results for different momentum values

Difference between	F	F>Fc
M=0.7 and M=0.75	0.001145	No
M=0.8 and M=0.95	0.006987	No
M=0.7 and M=1.0	0.013022	No

The ANOVA tests do not show any difference by using different momentum values. Therefore, a momentum of 0.8 is selected (default value proposed by Matlab) due to its high accuracy in test data set.

7.1.3.6 Inputs Selection

Different inputs were tested in the training of the neural network. These inputs were chosen according to the physics of the problem and their relationship with the correspondent target. First, the five AMSU-B frequencies were tested as input to the model. Second, station ID was added to the AMSU-B frequencies and they were tested as input to the model. Third, surface temperature and AMSU-B frequencies were tested as model input. And finally, different combinations of AMSU-B frequency brightness temperature were tested.

A number was assign to each station. The objective is that the neural network take indirectly into consideration specific characteristics of the stations such as topography, vegetation, and climate. Figure 48 illustrates the average accuracy of the model at these different inputs.

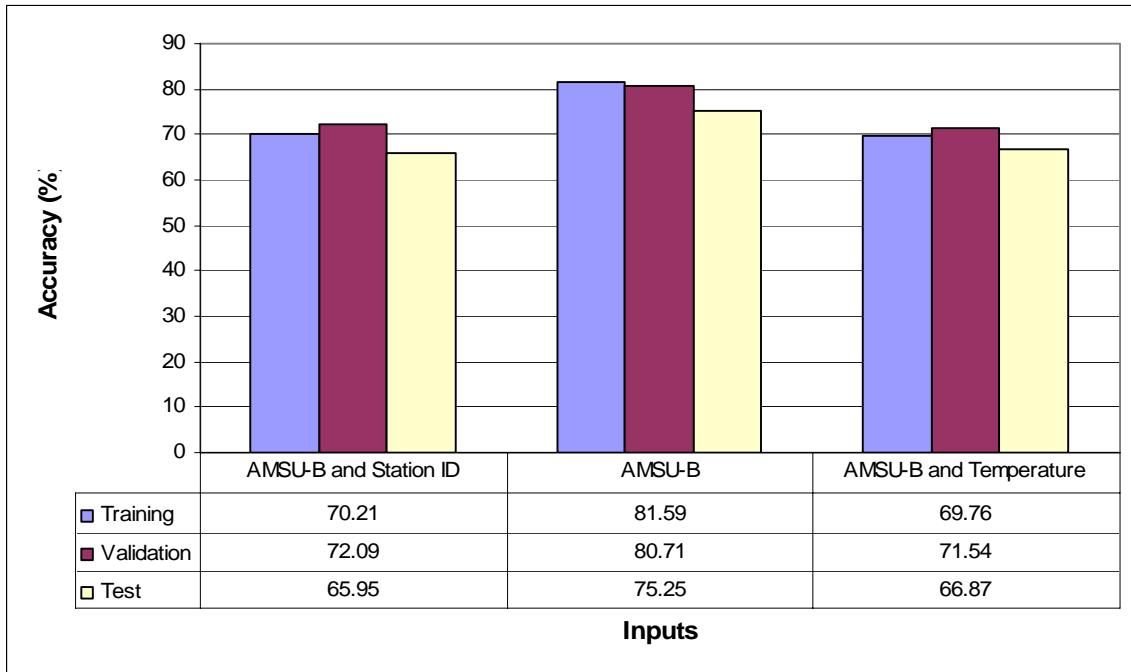
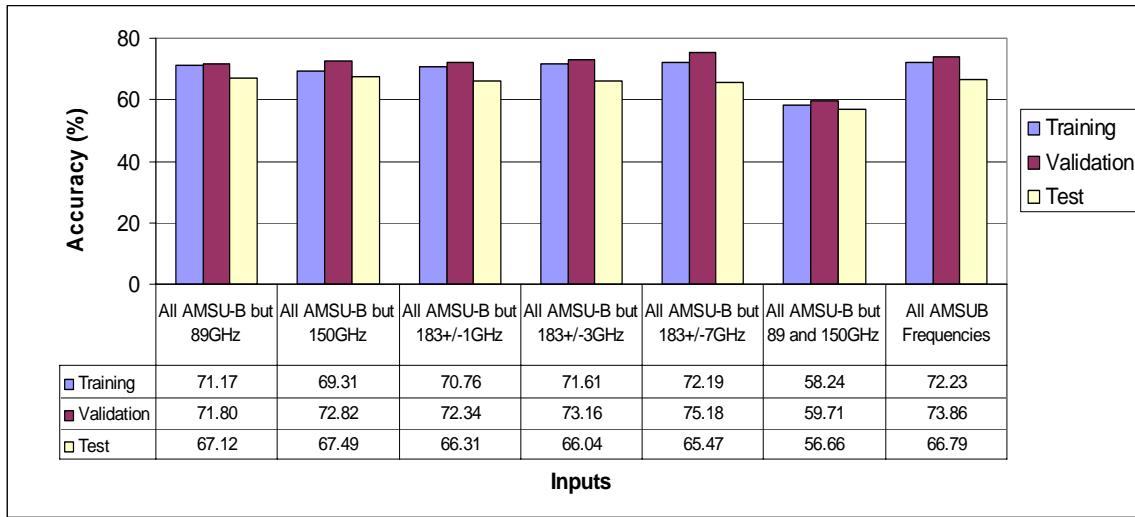


Figure 48: Average accuracy of the model output using different combinations of inputs

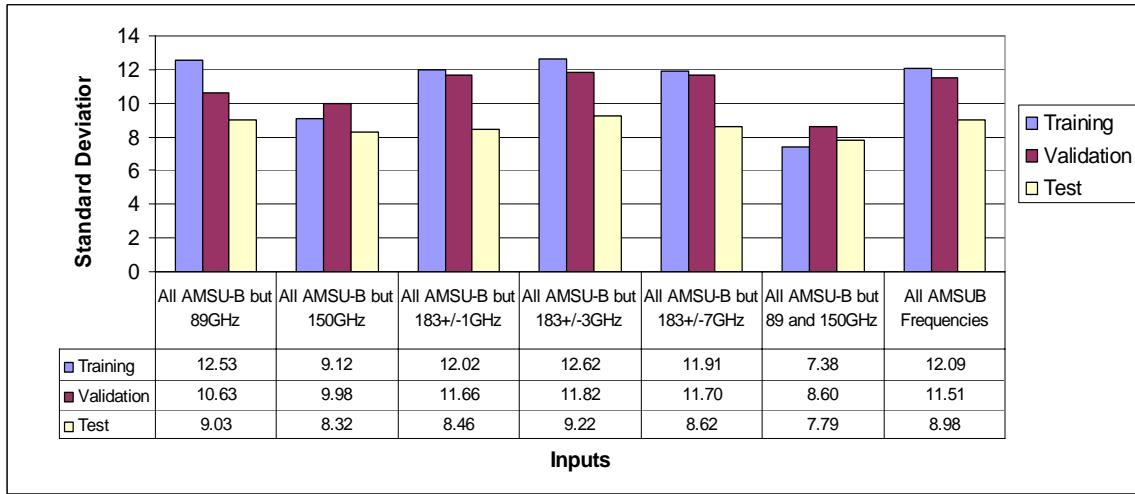
This graph demonstrates that the addition of surface temperature or station ID information to AMSU-B frequencies has no significant contribution in the accuracy of the neural network.

In order to find which of AMSU-B frequencies are more sensitive to snowfall detection, different combinations of frequencies were tested as input to the model. Figure 49 illustrates the results obtained in terms of average accuracy and standard deviation.

The ANOVA test shows a significant difference when using all AMSU-B frequencies but 89 and 150GHz and when using all AMSU-B frequencies. The graphs also show a significant drop in the accuracy when the model is run without 89 and 150 information. Therefore, all AMSU-B frequencies are used as input to the model.



(a)



(b)

Figure 49: (a) Average Accuracy (b) STDEV of the Model using different AMSU-B channels as input.

The ANOVA test was applied to the data in order to find if the difference in the neural network output is significant when using different inputs. The results are shown in the following table (detailed information of these tests can be found in appendix E).

Table 14: ANOVA test results for different inputs

Difference between	F	F>Fc
All AMSU-B but 89GHz and All AMSU-B but 150GHz	0.093799	No
All AMSU-B but 183+/-1 GHz and All AMSU-B but 183+/-3 GHz	0.044868	No
All AMSU-B but 89 and 150GHz and All AMSU-B	72.66782	Yes
All AMSU-B but 89GHz and All AMSU-B	0.065035	No

7.1.4 Neural Network System 1 (NN1)

A total of 1000 pixels (660 non precipitating pixels and 340 snowy pixels) were selected. A second group of 200 pixels (100 snowy and 100 non-snowy) was kept apart to measure the performance of the trained nets. Each net was trained using 600 pixels. These pixels were randomly selected from a pool of 1000 pixels.

The neural network system consists of 100 nets rigorously selected from a group of 500 trained nets. The nets selection was made based on the second group, by looking at the nets having the highest accuracy in detecting snowfall pixels. Appendix G shows the performance of the 200 independent pixels for each of the 500 nets for detecting snowfall and non-snowfall pixels. For instance, for net #1 (second row in the table) the numbers represents the number of pixels classified as the head column denotes. Some nets were excluded because these nets classified almost all the pixels as snowfall. Even though, it may rarely happen, it is better to avoid any additional source of error.

In order to generalize the training process, each of the 500 nets was trained with random initial weights using 600 pixels randomly selected from the training pool of 1000 pixels. To increase the neural network generalization ability, the training data were selected over different ranges of temperatures and land cover types.

As mentioned above, AMSU-B frequencies have been selected as the main input components in network training. AMSU-B data have also been used in other studies related to satellite-based precipitation estimation (Kongoli et al. 2003; Di Michele et al. 2005; Skofronick-Jackson et al. 2002; Noh et al. 2005). The input layer consists of five nodes where each node represents one AMSU-B brightness temperature channel: 89, 150 183 ± 1 , 183 ± 3 and 183 ± 7 GHz.

The pixels are then classified into either snowfall or non-snowfall. Therefore, the output layer consists of one node with values ranging between 0 and 1. A threshold was set at 0.5 where any pixel having an output greater or equal to 0.5 is classified as a snowfall pixel. Any pixel having an output below 0.5 is classified as a non-snowfall pixel.

Figure 50 (a) shows a simulation of the selected 100 nets over an independent area (also called validation area) during a snow storm. Figure 50(b) represents the respective true data for this independent area. No data from this storm have been used during the optimization and the training of the nets. The 100 nets used in this simulation are called Neural Networks System 1 (NN1).

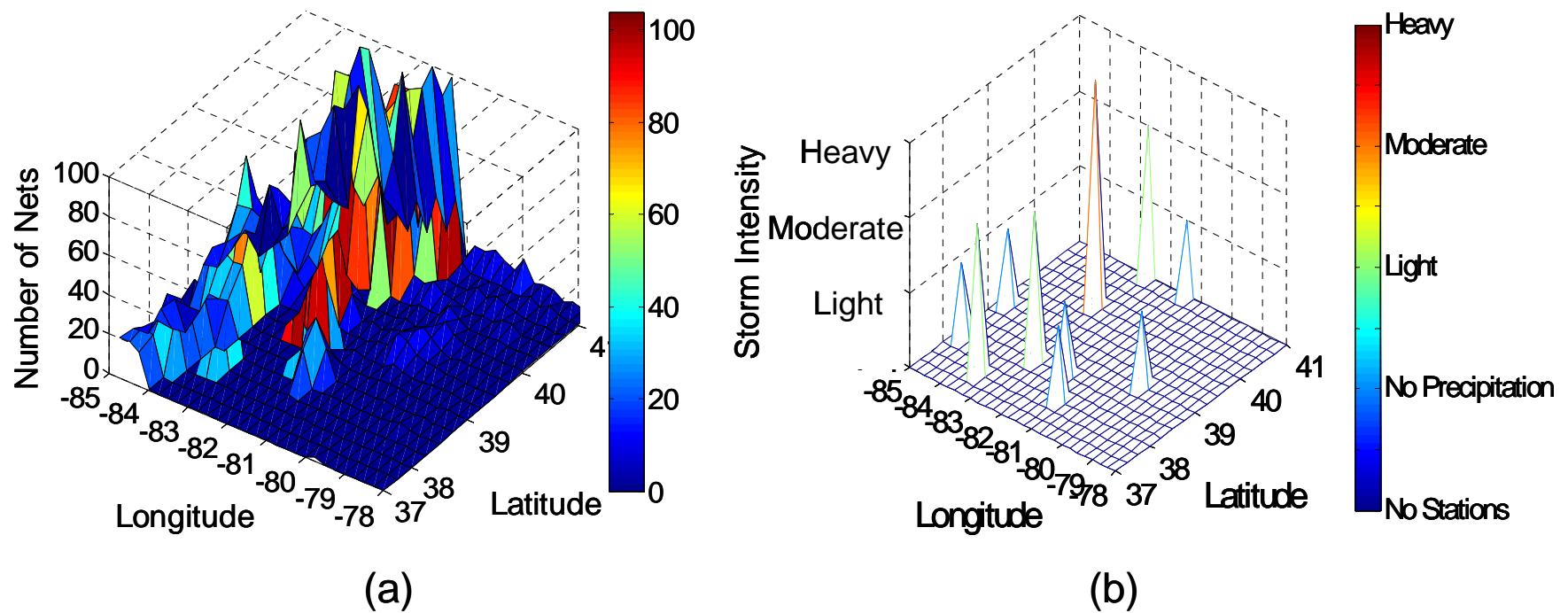


Figure 50: (a) Simulation and (b) true data of the independent area.

These graphs represent the behavior of the best 100 nets over each pixel. Each pixel represents 25 Km by 25 Km. The axes x and y are the latitude and longitude of the area. The z axis in the left graph represents the number of nets classifying the pixel as snowfall. On the other hand, the z axis in the right graph gives information of the truth data for each pixel. For some of the pixels there is no source of information to retrieve truth data. Therefore, they are labeled as “no stations”, meaning that there is no reporting station in that specific area.

Therefore, a threshold in the number of nets is set to select snowfall from non-snowfall pixels. This threshold should allow to obtain the highest accuracy in detecting snowfall and non-snowfall pixels. Consequently, a simulation using a group of 200 independent pixels was performed using different thresholds. The output of this simulation was then compared with the truth data for computing the model accuracy. Figure 51 illustrates the results of this simulation at different thresholds.

According to figure 51 as the threshold increases the accuracy in detecting non-snowfall pixels increases also but the accuracy in detecting snowfall pixels decreases. The optimum threshold to use is the one with 5 nets showing an accuracy of 50% in detecting snowfall pixels and 80% in detecting non-snowfall pixels. Therefore, any simulated pixel is classified as snowfall when more than 5 nets classify it as snowfall.

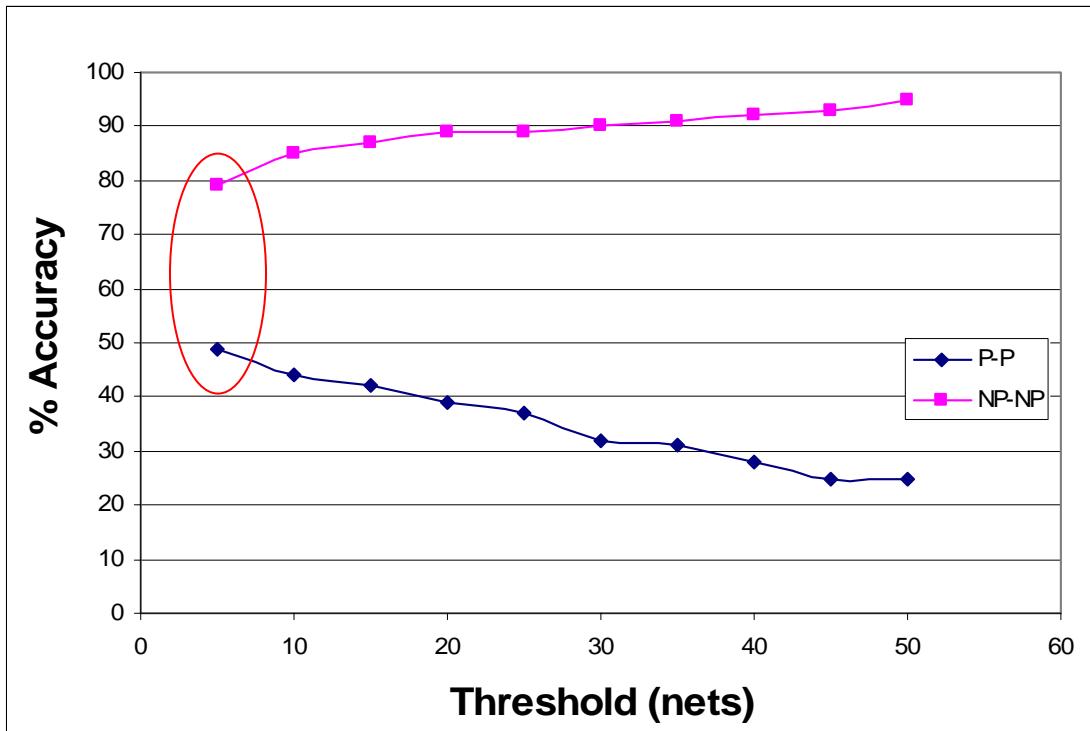


Figure 51: Simulation output accuracy at different thresholds

7.1.5 *Effect of snow cover on model accuracy*

According to previous researches, the detection of snowfall from passive microwave is affected when snow cover is present (Kongoli et al. 2003; Ferraro et al. 1998). Therefore, in this study snow cover have been tested as additional input to the neural network to test if this information could improve the overall accuracy of the model.

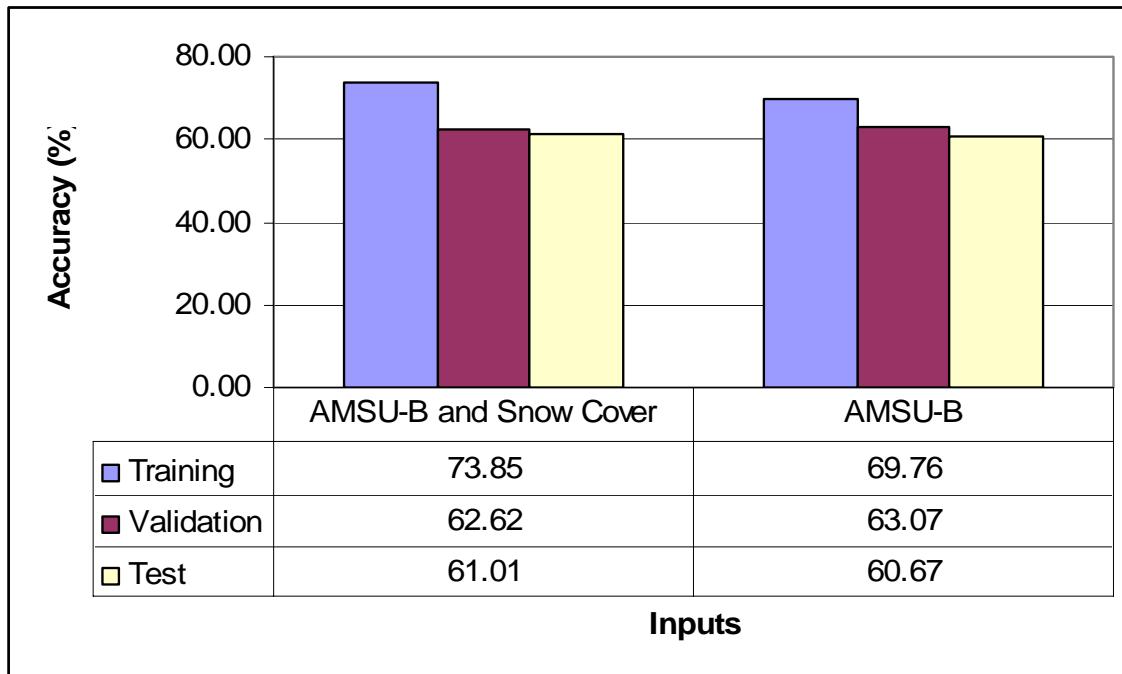


Figure 52: Average accuracy of the model using different inputs in the model

The optimal one hundred net were trained for both cases: 1) using AMSU-B frequencies and snow cover as input, and 2) using only AMSU-B frequencies as input. Figure 52 shows the average accuracy obtained from these two tests.

According to this graph, the addition of snow cover information to the model does not have a significant contribution to the model output accuracy. Also, the ANOVA test was applied. It showed an F value of 0.069063 ($0.069 < F_c = 3.89$), which means that snow cover information do not improve the model accuracy.

7.2 Part B: Snowfall Estimation

After the completion of the snowfall detection part, each pixel detected as snowfall by the first neural networks system (NN1) was presented to a second system (NN2) to perform a qualitative estimation of its intensity rate (figure 53).

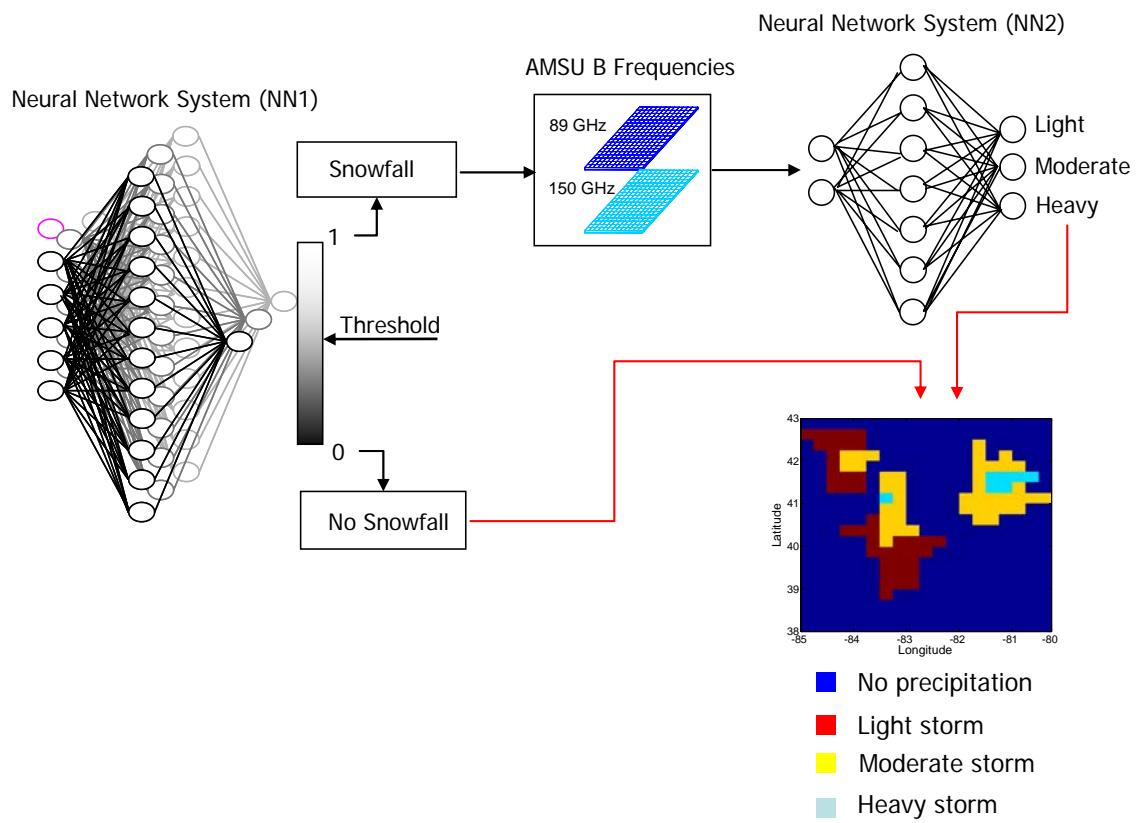


Figure 53: Snowfall Estimation Approach

All the pixels are assigned to one of the categories. To select the categories and their corresponding limits, a data frequency analysis was performed (figure 54).

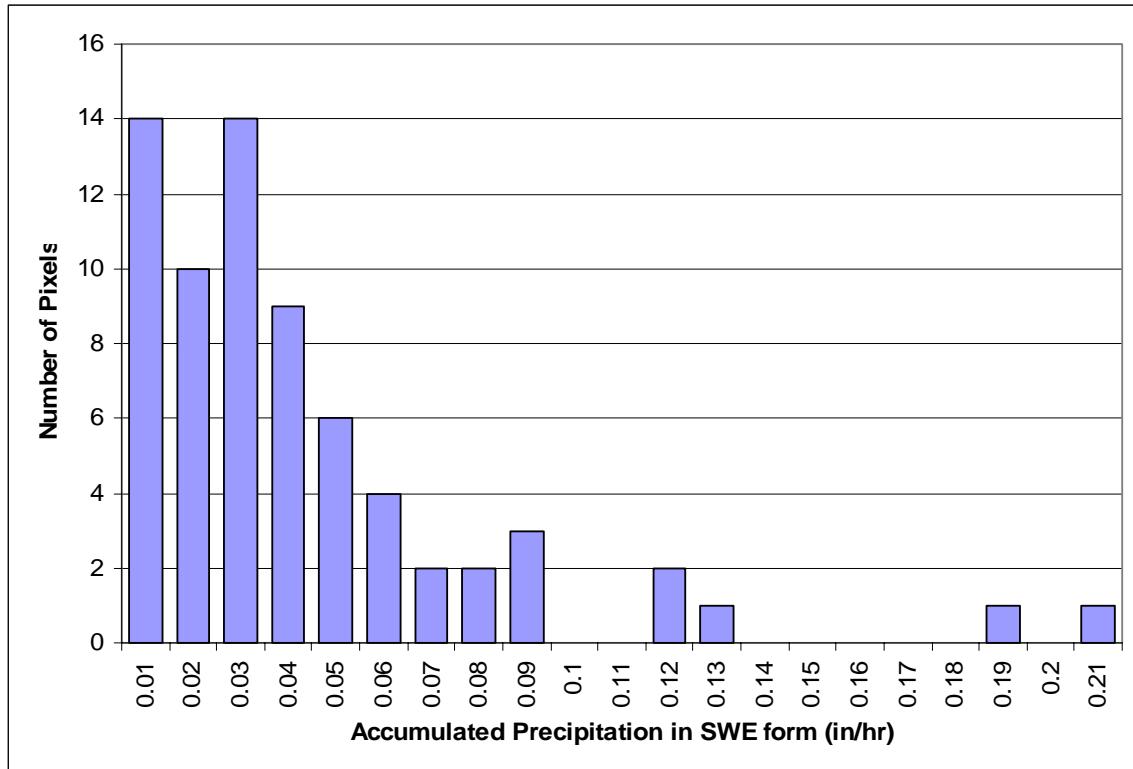


Figure 54: Snowfall data frequency graph

The categories were selected in two different manners: (a) by dividing the highest snowfall accumulation over one hour into three classes and (b) by dividing the total number of precipitating pixels into two classes.

7.2.1 Neural Network system 2 (NN2) using method A

The categories selected were three. The highest snowfall accumulation over one hour (in SWE form) was divided into three. Therefore, all snowy pixels will be classified to one of the following intensity levels: light, moderate or heavy depending on the accumulation

depth over one-hour-period. Snowfall intensities have been categorized as follow: (1) Light when the snow accumulation is between 0.25 and 1 cm; (2) Moderate for accumulations between 1 and 2 cm; and Heavy for accumulations greater than 2 cm.

As for NN1, the NN2 system consists also of one hundred trained and tested nets. A total of sixty-nine snowfall pixels were used to train and test the model. Two AMSU-B frequencies, 89 and 150 GHz, have been used as inputs due to their sensibility to snowfall scattering as showed in chapter 6. The nets used in the NN2 system were configured with two nodes at the input layer for AMSU-B frequencies and three nodes at the output layer for the three storm categories. The presented pixel will be assigned the category of the node having the highest output value.

All the nets have the same internal parameters, which were calibrated and optimized experimentally. One hundred runs for each possible configuration were performed. The best configuration was selected based on the quartile technique. The accuracy percent is divided in quartiles (0%-25%, 25%-50%, 50%-75% and 75%-100%). Therefore, the configuration selected is the one having the most number of nets in the upper quartile.

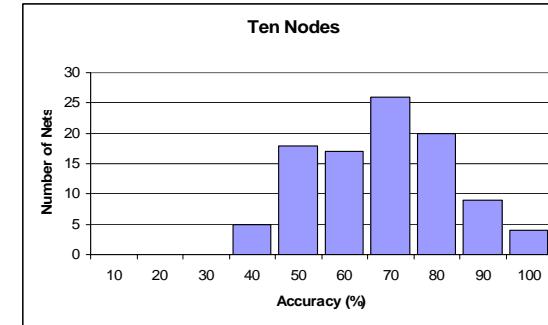
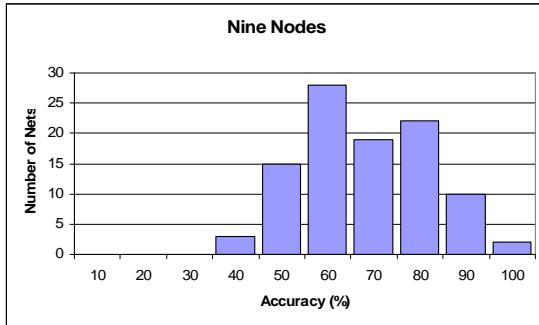
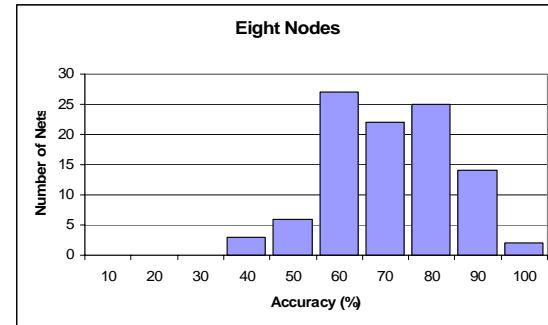
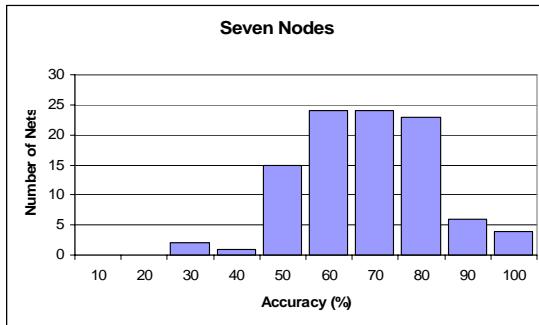
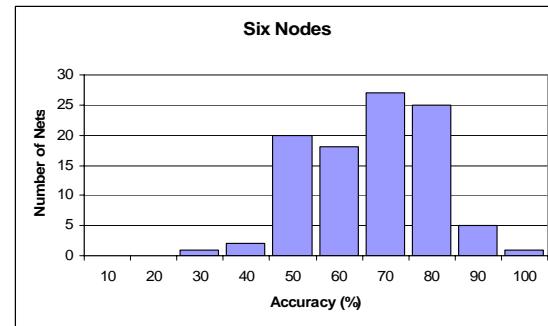
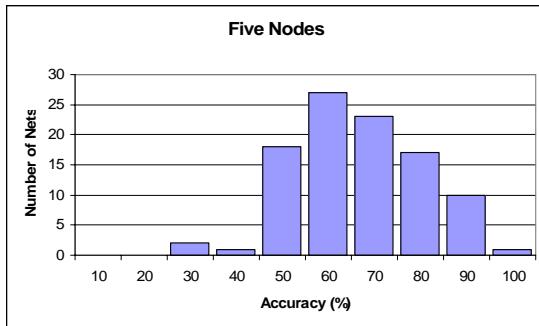
In the following subsections, the details of the optimization parameters such as number of nodes, number of hidden layers, training algorithm, activation functions, learning rate and momentum, are given.

7.2.1.1 Number of Hidden Layers and Nodes

One hundred nets were trained using different number of hidden layers and nodes.

Frequency graphs showing the behavior of the one hundred nets using different number of nodes with one and two hidden layers are illustrated in figures 55 and 56, respectively.

According to these figures, the configuration with the most nets in the upper quartiles is one hidden layer with eleven nodes. Therefore, one hidden layer with eleven nodes was selected to train the NN2.



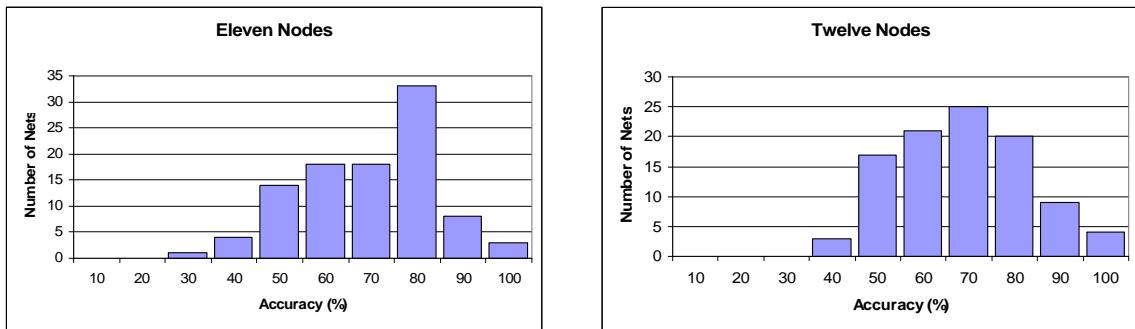
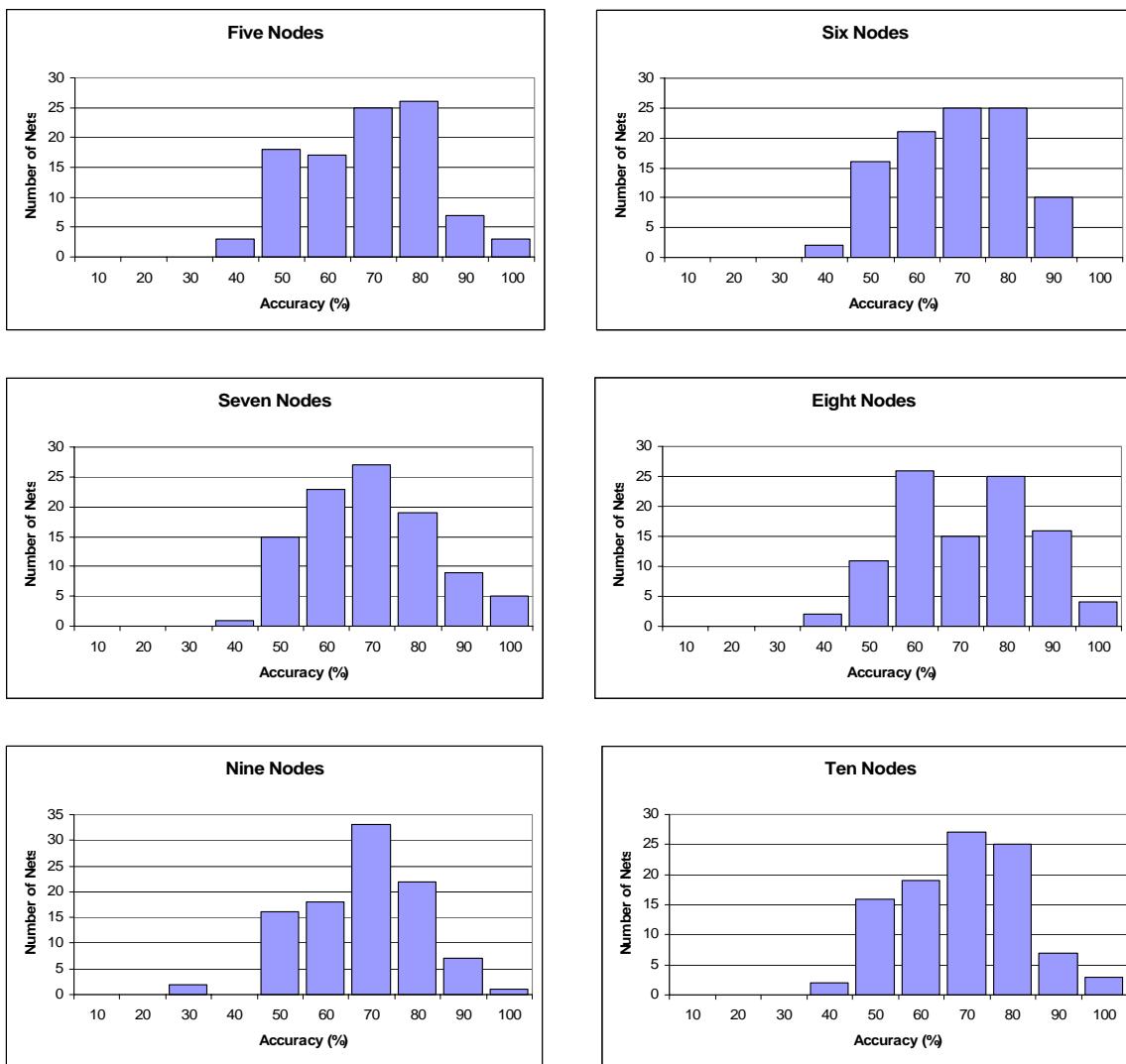


Figure 55: Neural nets accuracy using one hidden layer



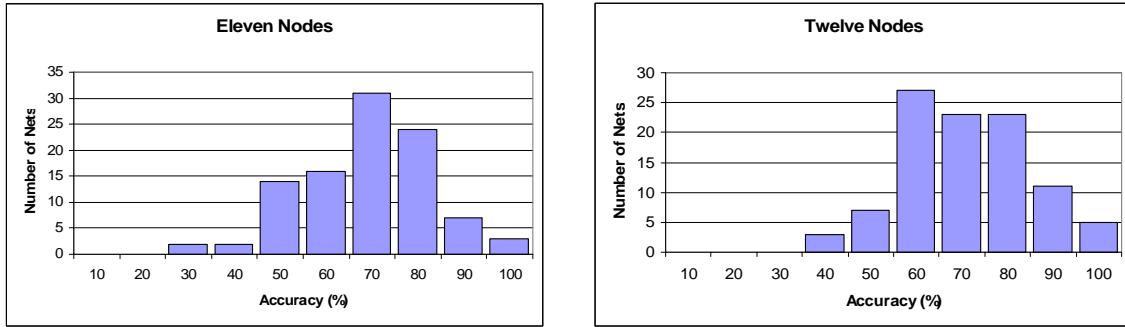
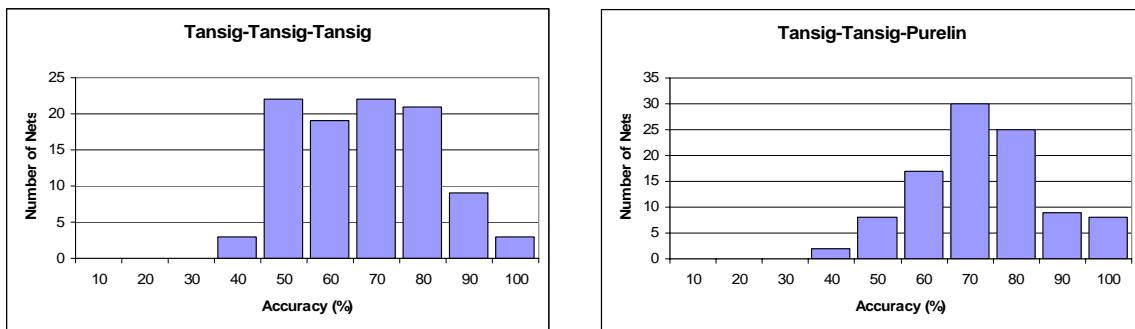


Figure 56: Neural nets accuracy using two hidden layer

7.2.1.2 Activation Functions

As discussed in the section above, the performance of the model output depends on the selected activation functions used to train the model. For selecting the best configuration of activation functions different combinations were used to train the model, using one hidden layer with eleven nodes. Figure 57 shows the frequency graphs of the model output accuracy for the one hundred nets.



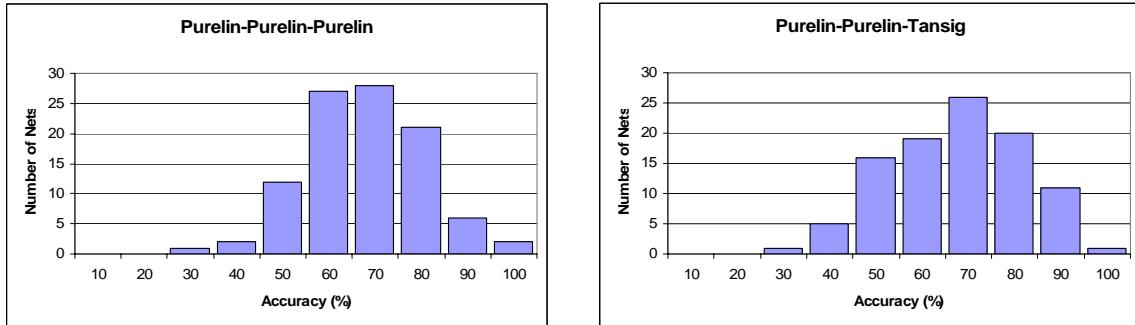
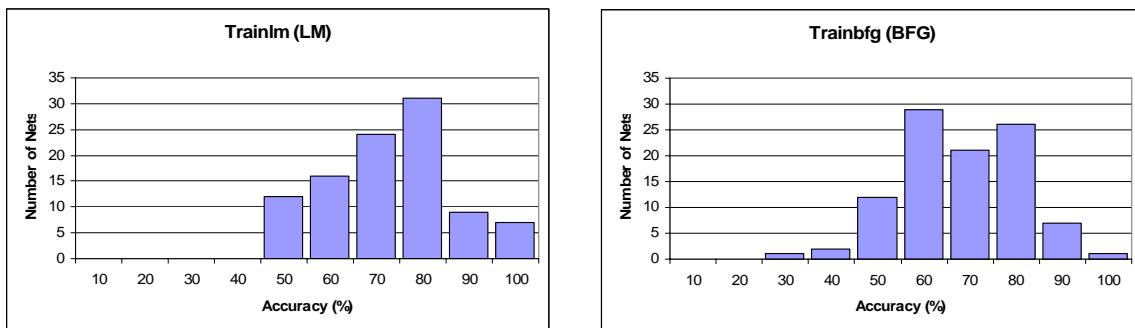


Figure 57: Neural nets accuracy using different combinations of activation functions.

According to the figure above, the best configuration was the one with Tansig-Tansig-Purelin, where 42 nets rated in the upper quartile.

7.2.1.3 Training Algorithms

Different training algorithms were tested as well. As reminder, the model output accuracy and the time taken for training the model depend on the training algorithm selected. Figure 58 illustrates the frequency graphs of the neural net output accuracy using different training algorithms.



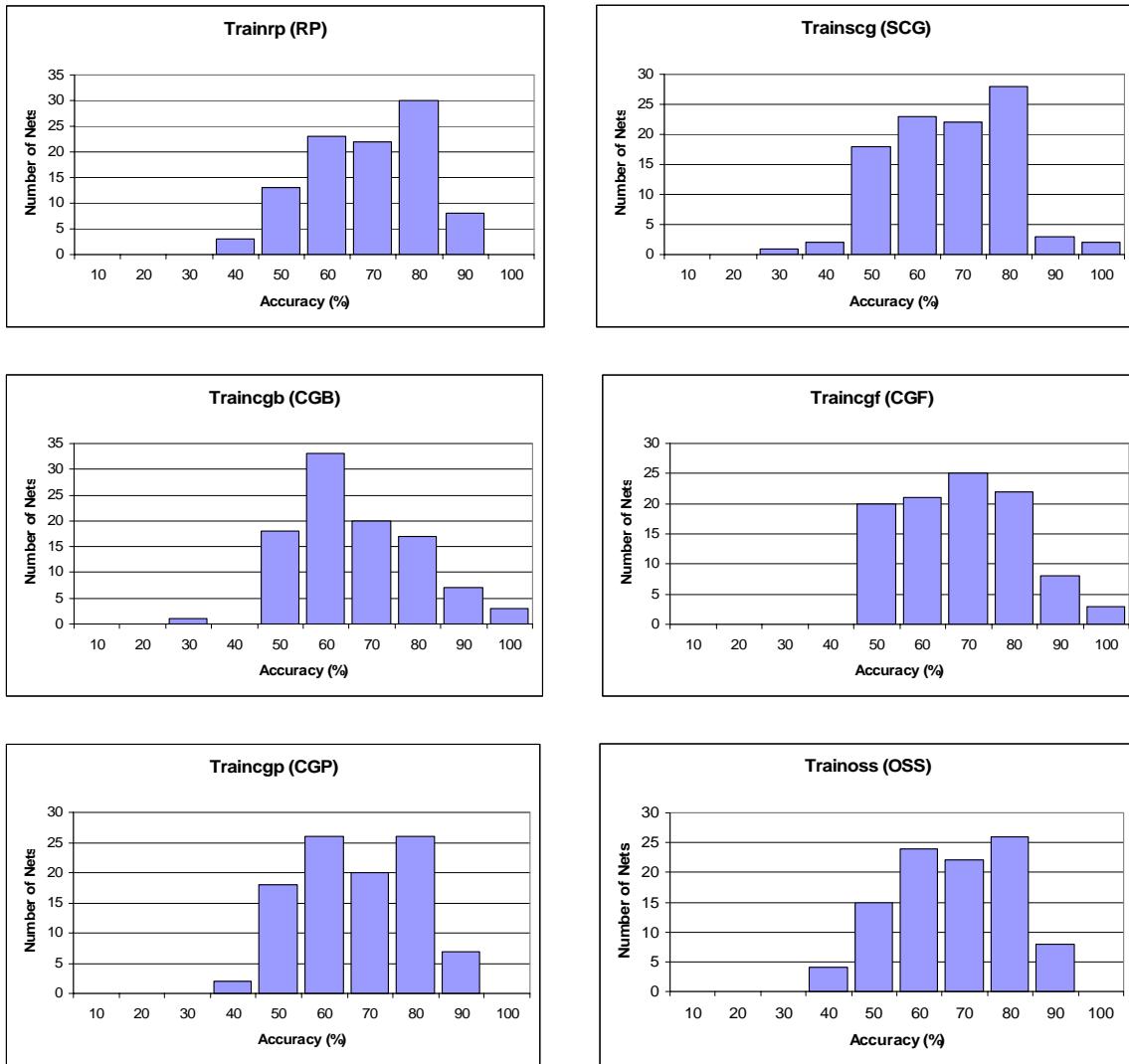
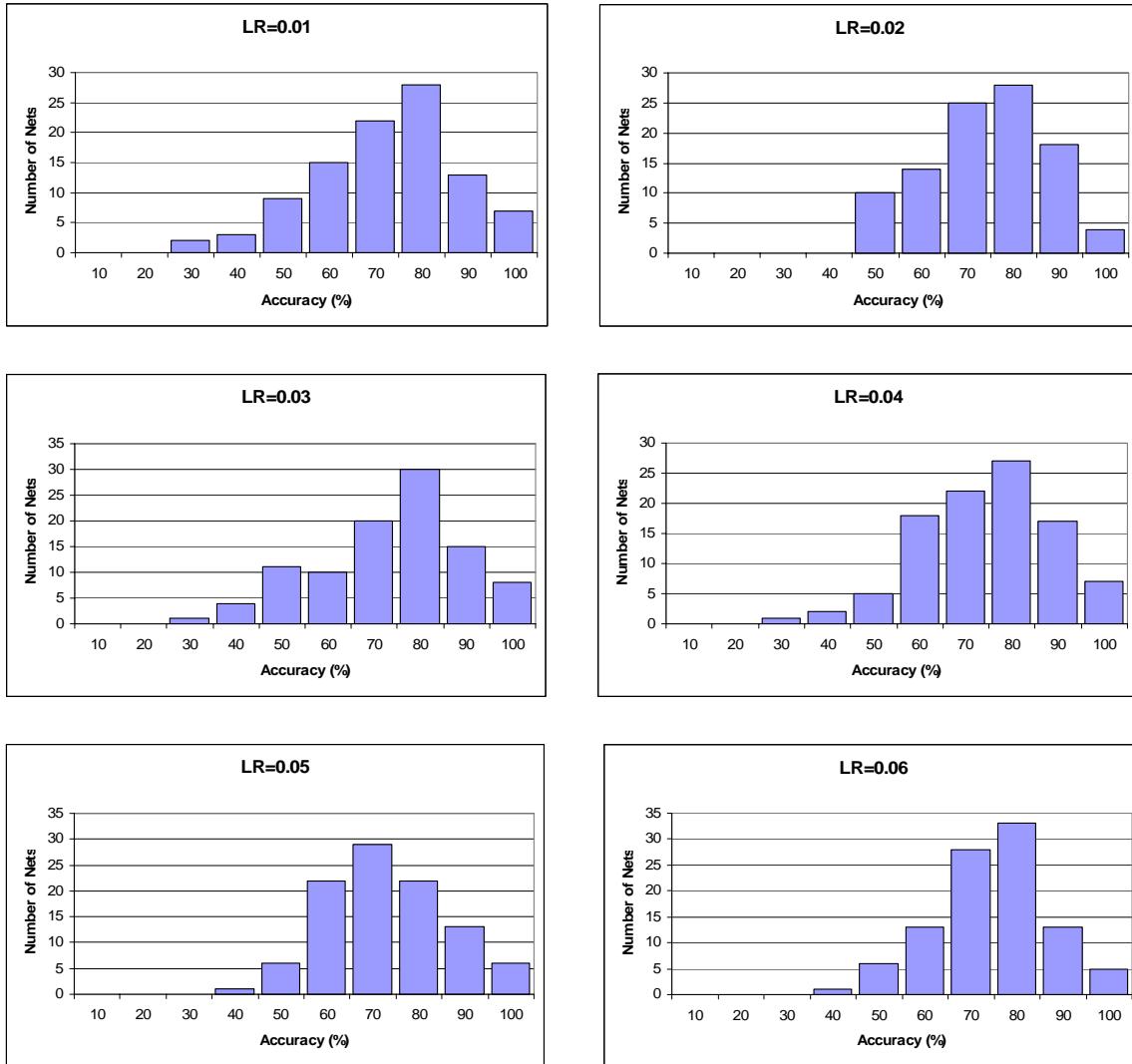


Figure 58: Neural nets accuracy using training algorithms

The best training algorithm showed in the figure above, is Levenberg-Marquardt with 47 nets in the upper quartile.

7.2.1.4 Learning Rate (LR)

As in the snowfall detection part, different learning rate values were tested. Figure 59 shows the frequency graphs of the neural nets output accuracy at different learning rate values.



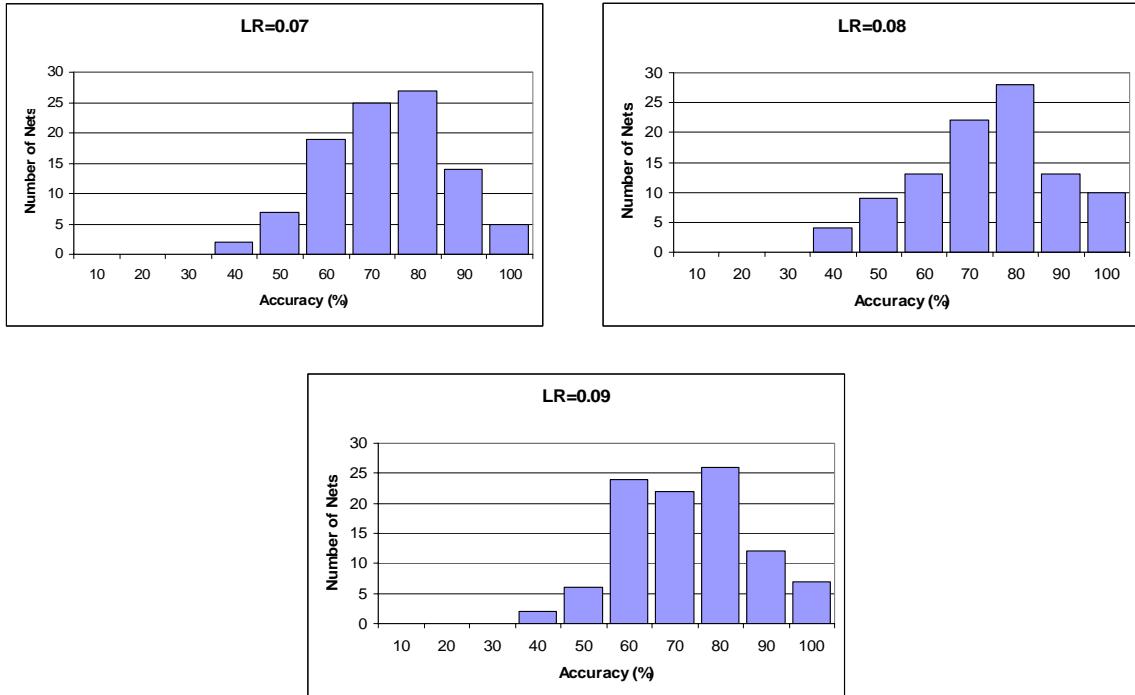
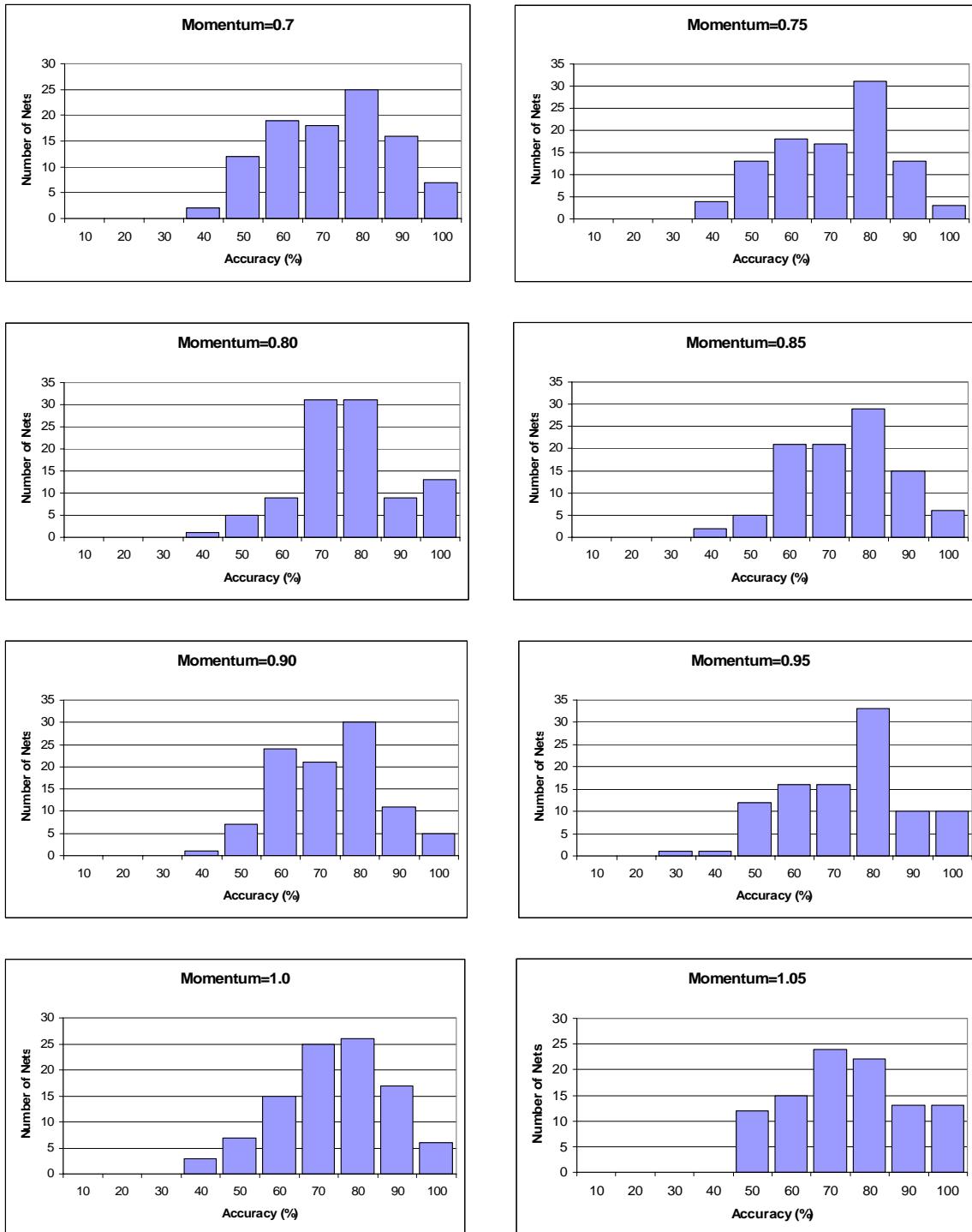


Figure 59: Neural nets accuracy using different learning rate values

The best accuracy obtained in figure 59 is using a learning rate of 0.03. This configuration has 53 nets in the upper accuracy percentage quartile.

7.2.1.5 Momentum

Also, the neural nets were tested at different momentum values to find the best accuracy in the one hundred nets. Figure 60 illustrates the frequency graphs of the neural nets accuracy at different momentum values.



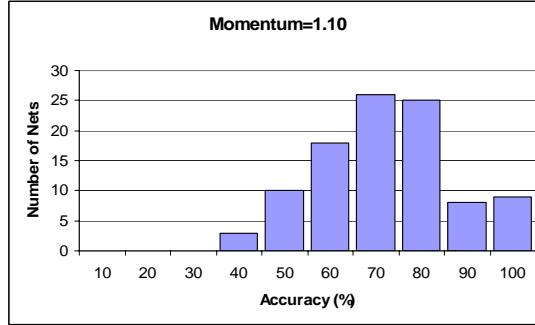


Figure 60: Neural nets accuracy using different momentum values

According to the graphs shown in figure 60, the model accuracy does not varies much at different momentum values. But, still the selected momentum value is 0.95 having 53 nets in the upper accuracy percentage quartile.

After the completion of the detection part, each pixel detected as snowfall by the first neural networks system (NN1) was presented to a second system (NN2) to perform a qualitative estimation of its intensity category level. Therefore, all snowy pixels will be classified to one of the following intensity levels: light, moderate or heavy depending on the accumulation depth over one-hour-period. Snowfall intensities have been categorized as follow: (1) Light when the snow accumulation is between 0.25 and 1 cm; (2) Moderate for accumulations between 1 and 2 cm; and Heavy for accumulations greater than 2 cm.

The NN2 system consists also of one hundred trained and tested nets. All the nets have the same internal parameters, which were calibrated and optimized experimentally. A total of sixty-nine snowfall pixels were used to train and test the model. Two AMSU-B frequencies, 89 and 150 GHz, have been used as inputs due to their sensibility to snowfall scattering. One hundred runs for each possible configuration were performed and the

configurations that obtained the best accuracy were retained. The nets used in the NN2 system were configured with two nodes at the input layer for AMSU-B frequencies and three nodes at the output layer for the three storm categories. The presented pixel will be assigned the category of the node having the highest output value. The simulation results showed that the neural network system (NN2) tends to classify most of the pixels as a light snow storm. This bias can be explained by the fact that the snow accumulation over one hour might not reflect the real intensity during AMSU acquisition time. To reduce this bias, new threshold has been set for both moderate and heavy categories. Thus, a storm will be categorized as heavy when more than five out of the one hundred nets classify it as heavy. The same threshold was applied for the moderate category. Otherwise the pixel is classified as a light storm. This approach has shown to have a significant effect on the estimation accuracy.

A group of one hundred independent snowfall pixels was used to simulate NN2. This group is composed by the three snowfall storm intensities to measure the behavior of NN2 in each one of them. The one hundred pixels consist of 58 Light, 28 moderate and 14 Heavy pixels. The results have shown that 71.9% of light storm, 28.6% of moderate storm and 21.4% of heavy storm pixels were correctly classified (table 15).

Table 15: Confusion matrix of neural network output

		Classified as		
		<i>Light</i>	<i>Moderate</i>	<i>Heavy</i>
Truth	<i>Light</i>	71.9%	22.8%	5.3%
	<i>Moderate</i>	60.7%	28.6%	10.7%
	<i>Heavy</i>	57.2%	21.4%	21.4%

The same storms used to validate NN1 system were also used to validate NN2 system. One purpose of validating these storms is to verify a consistent spatial distribution of the snowfall intensity areas. Only the pixels detected as snowfall were presented to the NN2 system. The results of both detection (NN1) and estimation (NN2) systems snowfall classification are shown in figures 61, 62, 63 and 64. The illustrated in these figures show the spatial consistency of the snowfall classification intensity.

According to these figures, the neural network output show spatial consistency of the snowfall detected areas and its corresponding classification intensity. Also, by comparing the output of the neural network and the radar-based information, it is possible to see a consistency of the storm areas. For instance, figure 63 shows a neural network output accuracy of 100 % according to the truth data, as well as radar-based information.

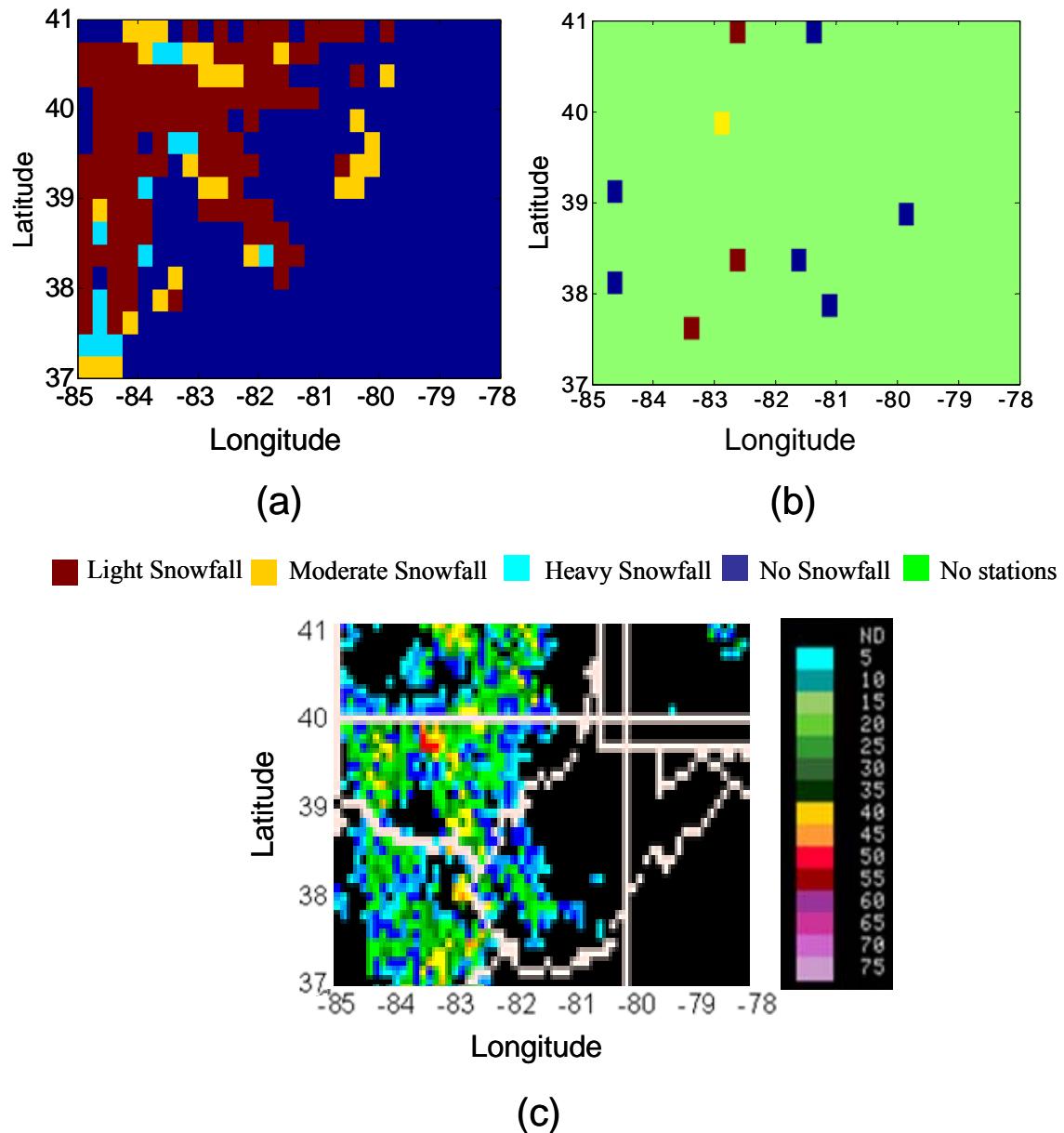


Figure 61: (a)NN output (b)Station-based truth data (c) Radar-based data for January 27, 2004 N15des

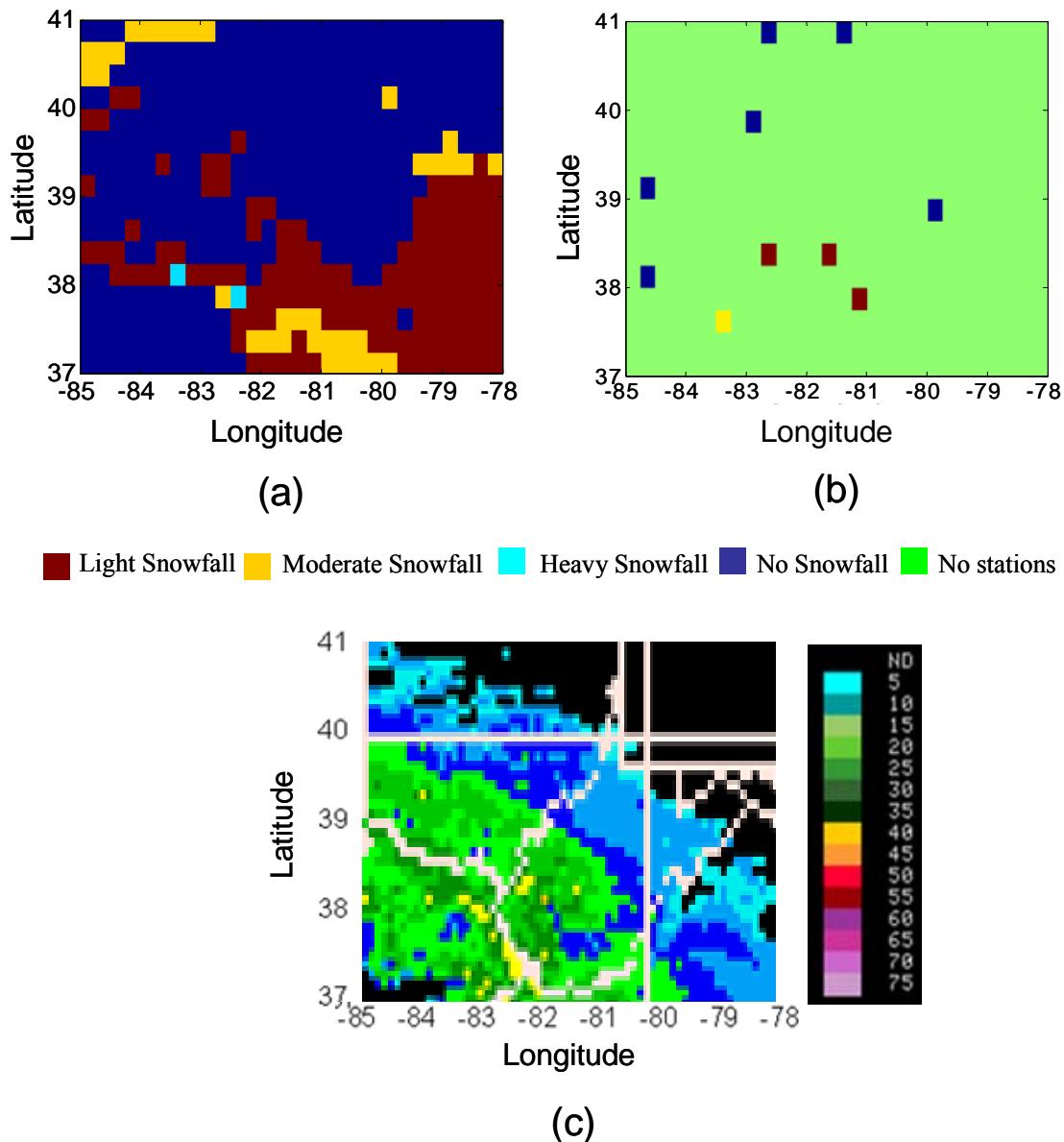
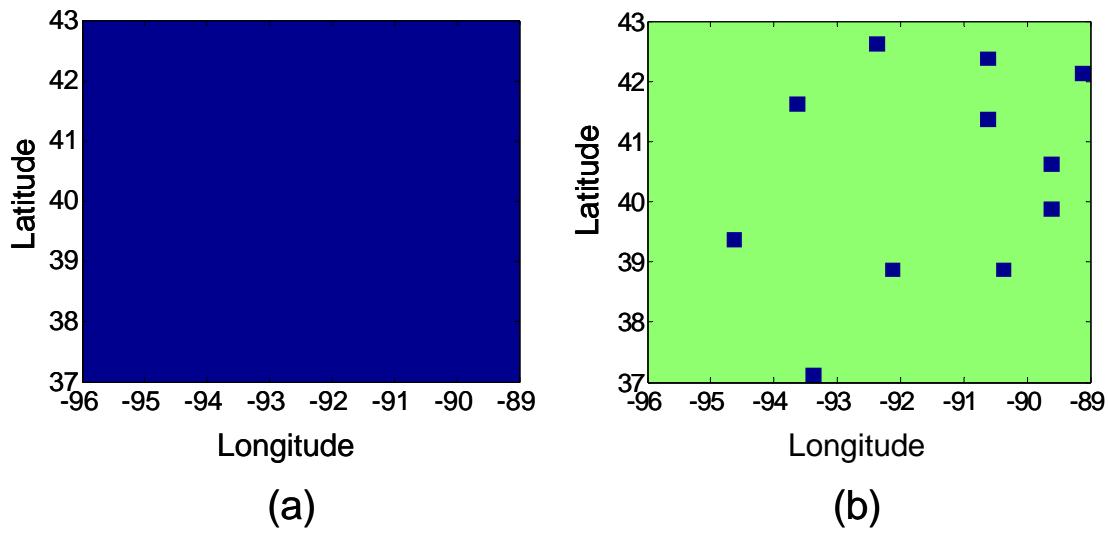


Figure 62: (a) NN output (b) Station-based truth data (c) Radar-based data for January 25, 2004 N16asc



■ Light Snowfall ■ Moderate Snowfall ■ Heavy Snowfall ■ No Snowfall ■ No stations

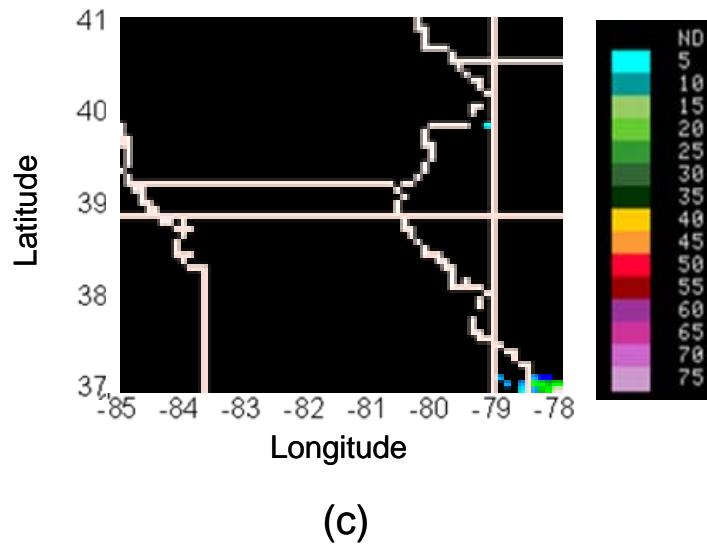


Figure 63: (a) NN output (b) Station-based truth data (c) Radar-based data for January 14, 2003 N15des

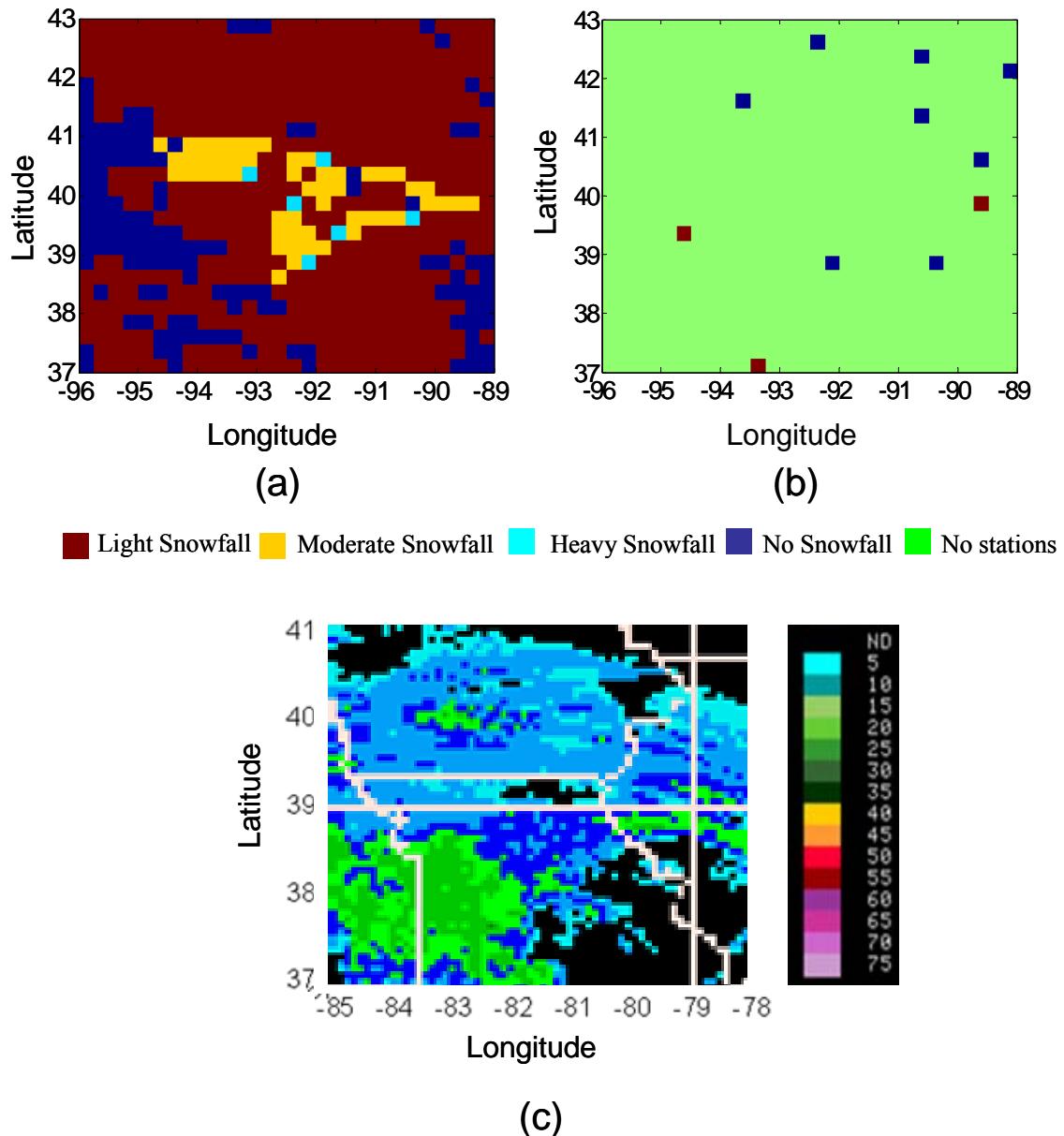


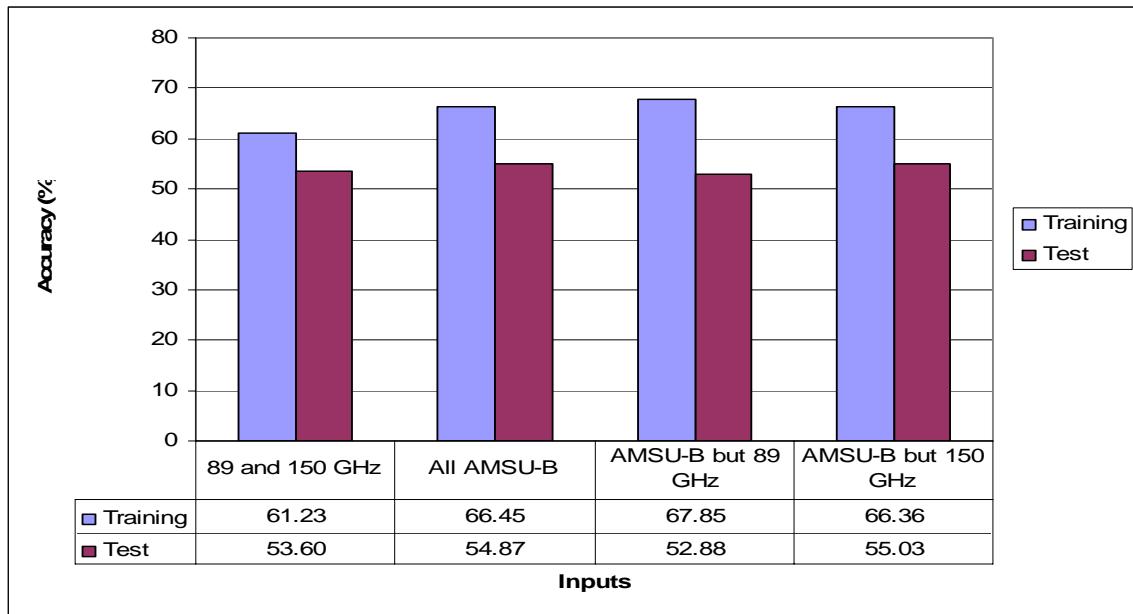
Figure 64: (a) NN output (b) Station-based truth data (c) Radar-based data for January 16, 2003 N16des

7.2.2 Neural Network system 2 (NN2) using method B

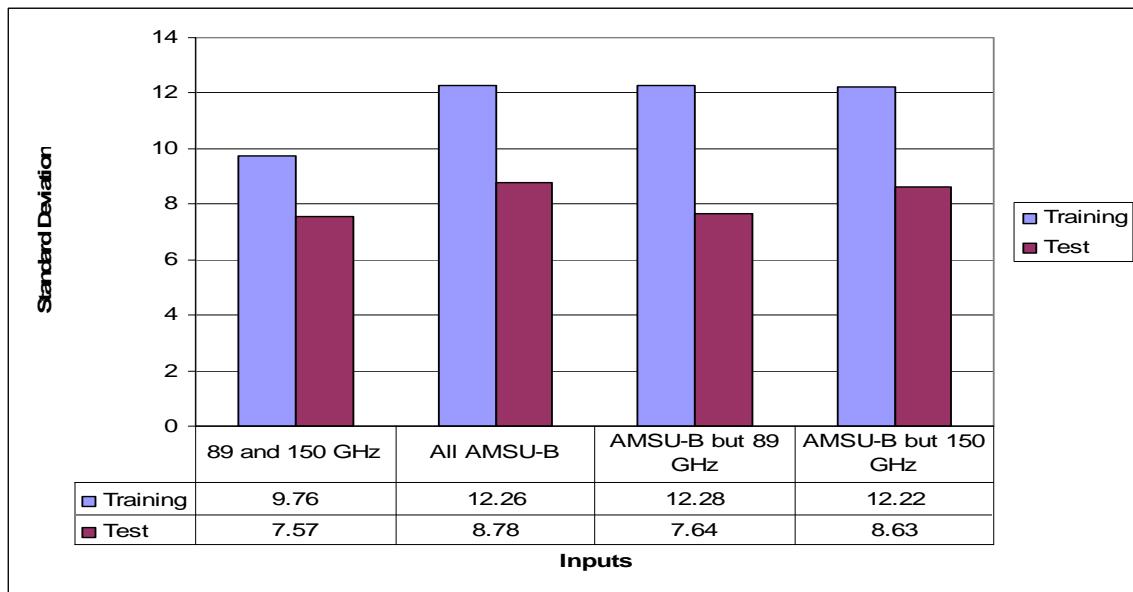
In an attempt to reduce the confusion between intensity classes, another estimation approach with only two categories was tested. The total number of snowfall pixels is divided into two equal quantities. Therefore, all snowy pixels are classified to one of the following intensity levels: light or heavy depending on the accumulation depth over one-hour-period. Snowfall intensities have been categorized as follow: (1) Light when the snow accumulation is between 0.25 and 1 cm; (2) Heavy when the snow accumulation is greater than 1 cm.

7.2.2.1 Selection of Inputs

One hundred runs were performed using the configuration selected in method A. Different inputs were tested to find the more sensitive to snowfall. Also, one and two nodes in the output layer were tested. Since the output ranges from -1 to 1, when using one node a threshold of 0 is defined for selecting between light and heavy. Any pixel with a value greater than 0 is defined to be heavy otherwise is light. Figures 65 and 66 show the results obtained using different inputs and one or two nodes in the output layer.

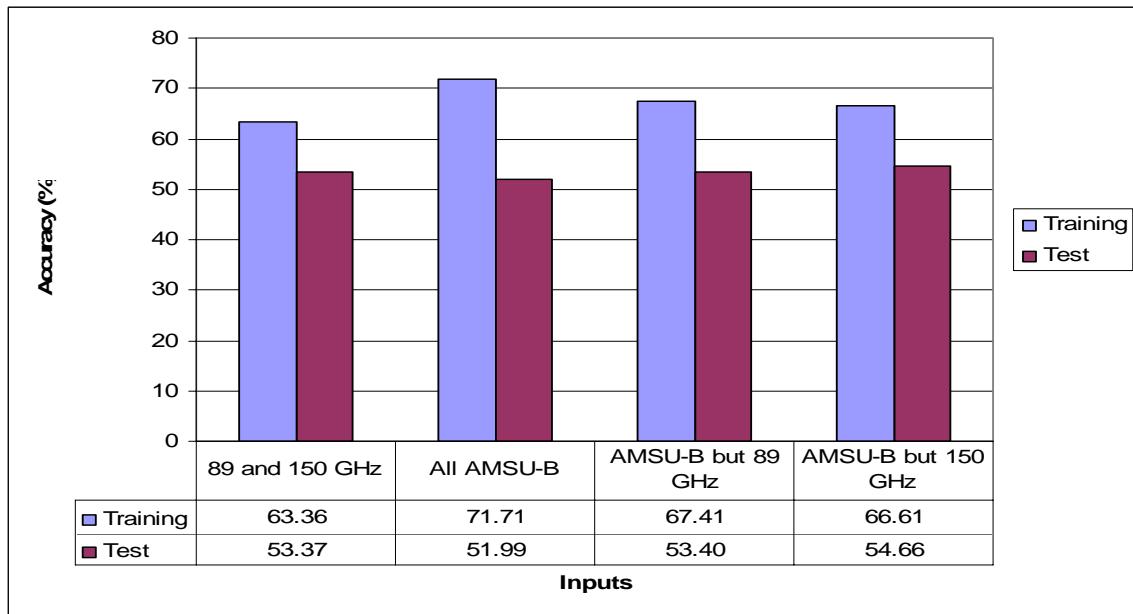


(a)

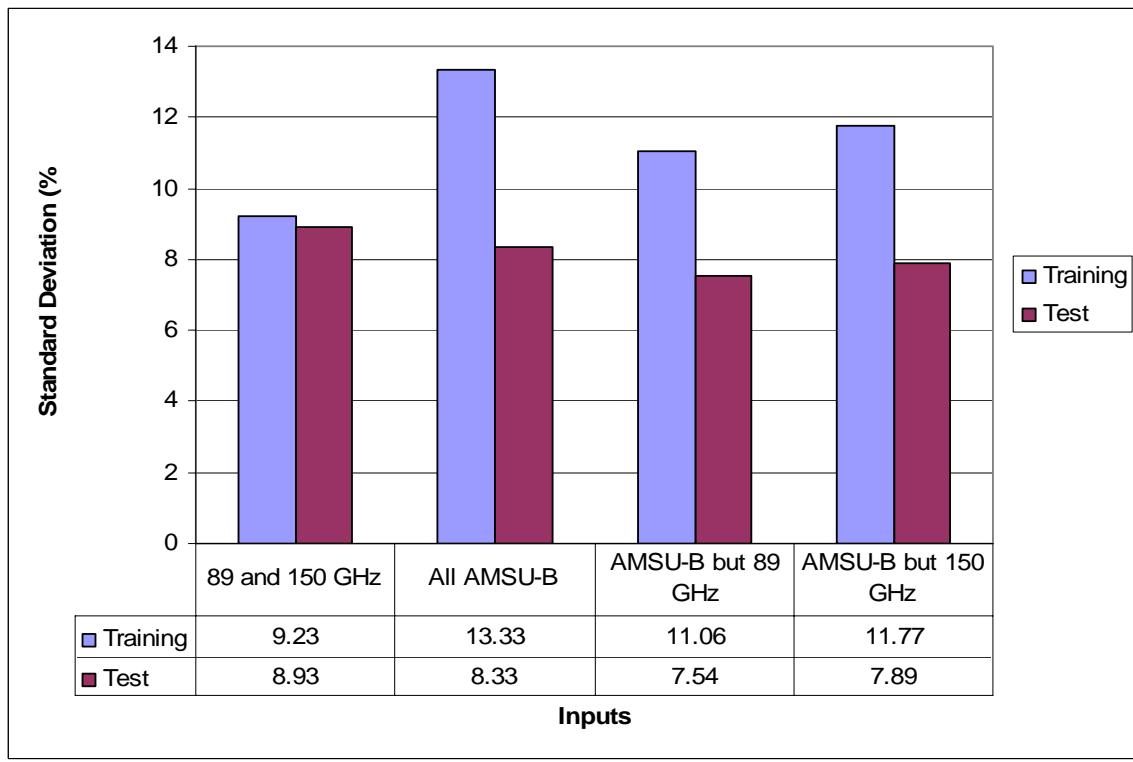


(b)

Figure 65: (a) Average Accuracy (b) STDEV using one node in the output layer.



(a)



(b)

Figure 66: (a) Average Accuracy (b) STDEV using two nodes in the output layer.

According to the graphs, there is no high variation in the accuracy of the model when only one node was used in the output layer (about 2 %). But when two nodes were used, the variation gets higher (about 7%). The ANOVA technique was performed to find if the difference between these groups of inputs is significant. Table 16 shows the ANOVA tests results. Detailed information of these tests can be found in appendix F.

Table 16: ANOVA test for NN accuracy when using different inputs

Difference between	F	F>Fc
89 and 150 GHz and All AMSU-B but 150GHz (One node)	1.550903	No
All AMSU-B and All AMSU-B but 150GHz (One node)	0.016177	No
89 and 150 GHz and All AMSU-B but 150GHz (Two nodes)	1.168115	No
All AMSU-B but 89GHz and All AMSU-B but 150GHz (Two nodes)	1.326379	No
All AMSU-B but 150GHz with one and two nodes	0.099414	No

According to table 16, there is no significant difference by using either combination of frequency as input to the model. Therefore, AMSU-B but 150 GHz with two nodes in the output layer is selected due to its high accuracy in the test data set.

Once both the detection and estimation parts are completed, the final product obtained is a map indicating the snowfall area and the respective intensity level for each pixel. Figure 67 shows a spatial consistency of snowfall classification intensity using method B.

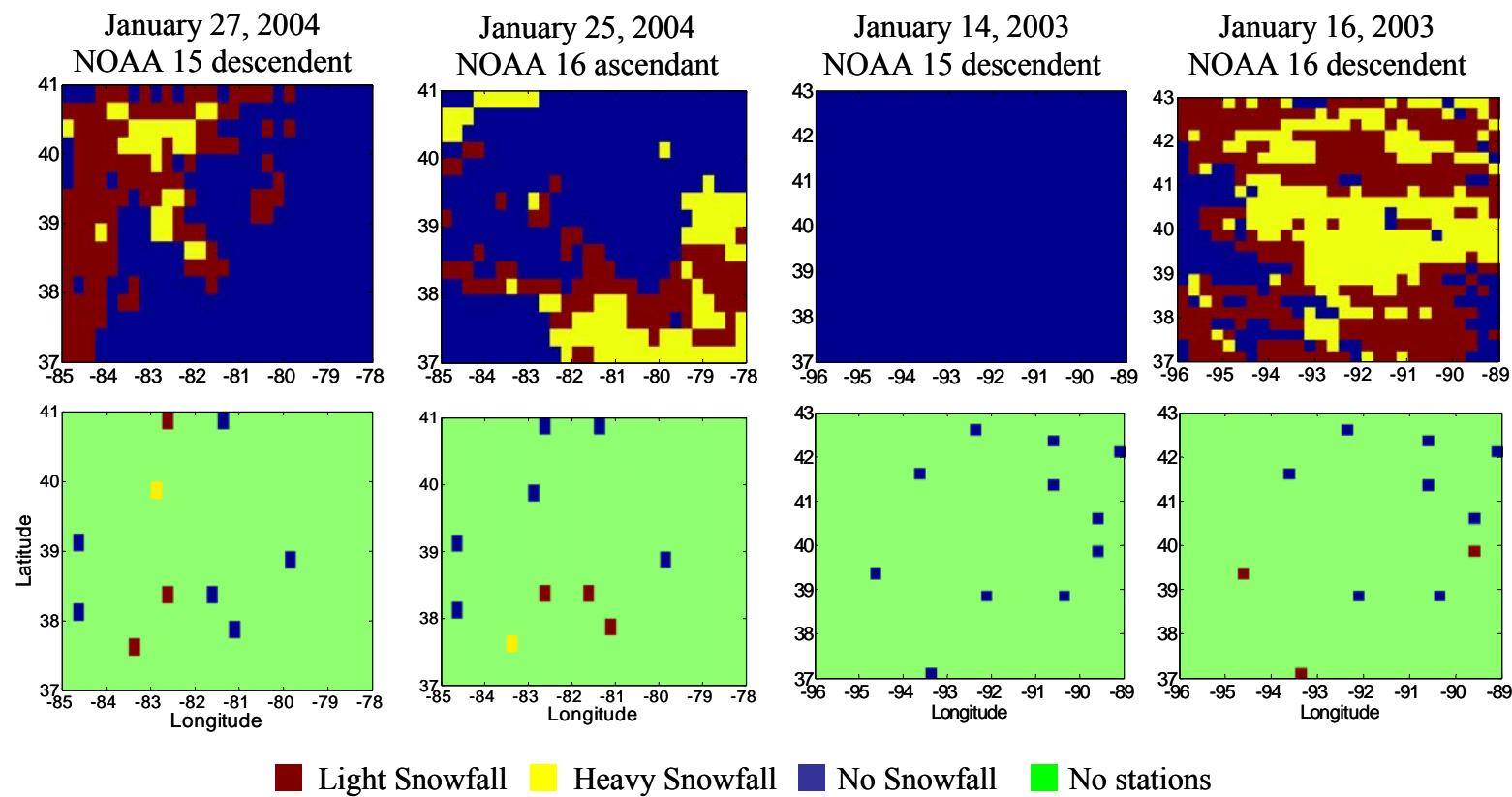


Figure 67: (a) NN output (b) Corresponding station-based truth data using method

8. CONCLUSION

Precipitation, specifically snowfall, is an important factor that affects human lives. The knowledge of the timing and intensity of any precipitating event is important for many weather related applications. Throughout the years several researches have conducted studies for detecting and estimating snowfall from space, but they have not been successful. Still, the necessity of a global snowfall product is evident. Thus, in this study snowfall is retrieved using satellite information and a robust neural network ensemble approach.

Satellite information from the Advanced Microwave Sounding Unit (AMSU) was used for detecting and estimating snowfall. This information is measured from space every four hours by using three NOAA satellites (15, 16 and 17). AMSU instrument is a passive microwave sensor. This sensor is able to capture and measure energy emitted and scattered from particles in both Earth and atmosphere. One of the physical limitations of these types of microwave sensors is that the energy emitted is very small. Therefore, the information has to be taken over large areas making the spatial resolution very low, compared to other satellites-based sensors. Another limitation is the factors that may interfere in the final measured signal. For instance, if the signal emitted from the surface is affected by atmosphere particles, then the signal is not purely taken from surface. The same thing may happen when retrieving snowfall. The signal measured from a snowfall particle might be affected by other atmospheric particles. It might be also affected by the

ground emissivity. Therefore, there is not a one-hundred percent confidence about what exactly contribute to the measured signal.

In this study, hourly ground-based information was used as truth data to train, calibrate and validate the model. Hourly information from seventeen stations was collected during four winter seasons. This information is provided by the National Climate Data Center (NCDC). A total of 108720 snowfall and non-snowfall pixels were collected. This data has been presented to several filters several filters in order to make sure that the information used to train the model was reliable. First, pixels having the same AMSU acquisition time were selected. By applying this condition the data set was reduced by 77 %. Second, pixels covered by clouds and having temperatures below zero degrees Celsius were selected. By applying these filters the data set was reduced by an additional 17 %. Third, pixels with three hour continuous storm were selected. This condition was applied only to snowfall pixels. As a result, the snowfall data set was also reduced. By applying all these filters, the data set became small, and this is one of the limitations for training the model.

In addition, another physical limitation for training the model is the spatial and temporal resolution of the truth data. The stations used to retrieve the truth data are sparse all over United States. Therefore, most of the pixels are represented by only one station. Since snowfall varies in space and depends on topography conditions, one station is not representative of an area of 25 Km by 25Km. Also, this study is based on one hour

temporal resolution. For instance, the snowfall accumulated over one hour is compared to the one shot measurement taken from AMSU. A pixel might reflect snowfall accumulation over one hour, but within that hour, there might be some laps in which there is no precipitation. And, AMSU might take the measurement at that lap. Therefore, a snowfall pixel is categorized by AMSU as a non-snowfall pixel, might be mislabeled.

In this study, a neural network ensemble approach is introduced for reducing the effect of mislabeled pixels in the training data. According to the neural network ensemble approach, each net should be trained differently from each other. In this research, each net is trained using different data sets and different initial weights. By training the nets with different data sets, the chance to train the nets with wrong or mislabel data is reduced. On the other hand, some of the nets might be trained having a higher number of correct label pixels. This approach helps in the training. But, it will not help in the validation part since the pixels used to evaluate the trained nets might be mislabeled and introduce an artificial error to the validation accuracy.

Another factor that affects the neural network generalization is the selection of the neural network internal parameters. The ANOVA technique was used to find out if there is a significant difference in the accuracy of the model when testing different internal parameters. For instance, when using one hidden layer and five nodes, the accuracy of the model obtained is 65.76 %. And, when using one hidden layer with twelve nodes the accuracy of the model improves about 4 %. Therefore, the ANOVA test is used to find if the 4 % improvement is significant. According to the test, there is a significant

improvement in the accuracy when using one hidden layer and twelve nodes, since the F value obtained was greater than the critical F value (F_c) value. The same approach was applied in testing other internal parameters such as training algorithms, activation functions, learning rates, and momentum.

Moreover, the selection of the inputs is one of the most important steps in the calibration and design of the neural network approach. The selection of the variables to use as model inputs is based on the physics of the problem to solve and the relationship of the selected variables with the corresponding outputs. In this research several variables were tested to find the best accuracy of the neural network. Different combinations of AMSU-B frequencies were tested as inputs to the model. Since the snow accumulated in the ground will contribute to the total emissivity measured by the satellite, snow cover information was tested in combination with the AMSU-B frequencies as model input, to find if snow cover information could improve the neural network generalization. Additionally, station ID information was also tested as model input. By adding station ID information, the model gets more information about the specifics characteristics of the station such as topography and climate conditions. By making all these tests, it was found that snow cover and station ID did not give any significant contribution to the accuracy of the model. Also, it was found that by using all AMSU-B frequencies, the model accuracy obtained is the highest.

Once the neural network internal parameters and inputs are selected, the neural network system is trained. The product obtained is a map indicating the snowfall and non-snowfall

areas. Due to the limited number of stations in the validation cases, the comparison between the neural network output and the truth data was made separately for each station. Each station represents one 25 km by 25 km pixel. Therefore, one of the purposes of the validation of the developed algorithm is to check its ability to produce consistent spatial snowfall coverage. As shown in the methodology and results chapter, the Neural Network mapping of the snowfall coverage was consistent and homogenous. By comparing the model outputs to measured snowfall, more than 70 % of the measured events have been correctly detected by the neural-networks-based model. Specifically, one of the validation sites (January 14, 2003 NOAA 15 descending) shows an accuracy of 100 % in detecting the non-snowfall area.

The same approaches followed for selecting the inputs and internal parameters of the neural network system 1 (NN1), were performed for the snowfall estimation part. The truth data used for training the neural network system 2 (NN2) is the same used for NN1. As stated before, an hourly based approach is used in this study. The limitation of using this temporal resolution, is that having an hourly snowfall accumulation compared with the acquisition of brightness temperature from AMSU, taken in one shot, is not as accurate. In one hour, the snowfall accumulation might be high but the snowfall rate is not steady. Therefore, the information acquired during one pass of AMSU is not representative of snow accumulation over one-hour period. This restriction might be reflected in the output of the neural network system. Also, the quality of the truth data used to train these neural networks is not well controlled.

In short, the mislabeled pixels may generate a multiplicative error in the neural network systems. For example, mislabeled pixels presented to NN1 will be wrongly classified. Then, the pixels wrongly classified as snowfall are presented to NN2. Therefore, these pixels are forced to be classified to one of the intensity categories. Consequently, it affects the overall accuracy of the snowfall estimation results.

In the past decades, some researchers have conducted studies on snowfall using satellite-based information and have not been as successful. Some of these studies are physical-based models, others are statistical-based. To our knowledge, this is the first snowfall study based on a neural network model. Even though the overall accuracy of the neural network systems is not too high, at least there is a product what can be improved in the future. This product can be improved by building up truth data with higher quality and by keeping the same approach. For example, a number of meteorological stations can be selected and programmed to report the occurrence of snowfall events and their rates exactly at the same time of AMSU acquisition.

FUTURE WORK

There are several limitations mentioned in this thesis that constrain the results to a lower accuracy. Some of the factors that affect the final accuracy of the model and can be taken into account in a future work are:

- Improvement of quality of truth data: a number of meteorological stations can be selected and programmed to report the occurrence of snowfall events and their rates exactly at the same time of AMSU acquisition for the area covered by the station.
- Adding microwave information from other sensors: this factor can help to reduce the gap between measurements and provide the opportunity to have bigger data set for training the model.
- Wind speed information: this information can give information to the model about the accuracy of the ground station measurements. By adding wind speed information to the model, the overall accuracy of it might improve.

CONFERENCE PROCEEDING AND PRESENTATION

- MEJIA, Y., GHEDIRA, H., SHAYESTEH, M., and R. KHANBILVARDI . "A Robust Neural Network System Design for Detecting and Estimating Snowfall from the Advanced Microwave Sounding Unit (AMSU)". Submitted to *Journal of Applied Remote Sensing*.
- MEJIA, Y., GHEDIRA, H., SHAYESTEH, M., and R. KHANBILVARDI . "A Neural Network Based Approach for Multi-Spectral Snowfall Detection and Estimation". *IEEE International Geoscience and Remote Sensing Symposium*. July 23-27, 2007.
- MEJIA, Y., SHAYESTEH, M., and R. KHANBILVARDI. "Multi-Spectral Remotely Sensed Snowfall Rate Estimation". *Junior Scientist Conference, Vienna University of Technology*, April 19-21, 2006.
- MEJIA, Y., GHEDIRA, H., SHAYESTEH, M., and R. KHANBILVARDI. "Multi-Spectral Remotely Sensed Snowfall Rate Estimation". *21st Conference on Hydrology, of the 87th American Meteorological Society (AMS) Annual Meeting, San Antonio, Texas*, January 14–18, 2007.
- MEJIA, Y., GHEDIRA, H., SHAYESTEH, M., and R. KHANBILVARDI. "Multi-Spectral Remotely Sensed Snowfall Rate Estimation". *NOAA Educational Partnership Program, Education & Science Forum, Florida Agricultural and Mechanical University, Tallahassee, Florida*, October 30-November 1, 2006.
- MEJIA, Y., SHAYESTEH, M., and R. KHANBILVARDI. "Multi-Spectral Remotely Sensed Snowfall Rate Estimation". *American Geophysical Union, Baltimore, Maryland*, May 23-26, 2006.
- MEJIA, Y., SHAYESTEH, M., and R. KHANBILVARDI. "Detecting and Estimating Snowfall using Multi Spectral IR and Microwave Data". *Fourth Annual NOAA-CREST Symposium, University of Puerto Rico at Mayaguez*, February 23-25, 2006.
- MEJIA, Y., SHAYESTEH, M., and R. KHANBILVARDI. "Snowfall Rate Estimation from Multi-Spectrum Remotely Sensed Data". *Microrad06 (9th Specialist Meeting on Microwave Radiometry and Remote Sensing Applications)*, San Juan, Puerto Rico, February 28th-March 3rd, 2006.
- MEJIA, Y., SHAYESTEH, M., and R. KHANBILVARDI. "Detecting and Estimating Snowfall using Multi Spectral Infrared and Microwave Data". *14th Conference on Satellite Meteorology and Oceanography, of the 86th American Meteorological Society (AMS) Annual Meeting, Atlanta, Georgia*, January 29th – February 2nd, 2006.

- *MEJIA, Y., SHAYESTEH, M., and R. KHANBILVARDI.* “*Snowfall Estimation Using Remotely Sensed Information*”. *Einstein’s in the City, New York, April 11-12, 2005.*
- *MEJIA, Y., SHAYESTEH, M., and R. KHANBILVARDI.* “*Snowfall Estimation Using Remotely Sensed Information*”. *NOAA Educational Partnership Program, Education & Science Forum 2004, City College of New York, October 21-23, 2004, pp. 528-532.*

Appendix A: ANOVA Test

Table 17: ANOVA test for the accuracy of 100 nets with 1HL with 11 and 12 nodes

Anova: Single Factor

SUMMARY						
Groups	Count	Sum	Average	Variance		
1HL 12 nodes	100	6682.265	66.82265	120.9692		
1HL 11 nodes	100	6656.838	66.56838	115.0572		

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	3.232484	1	3.232484	0.027391	0.868718	3.888853
Within Groups	23366.62	198	118.0132			
Total	23369.85	199				

Table 18: ANOVA test for the accuracy of 100 nets with 2HL with 6 and 8 nodes

Anova: Single Factor

SUMMARY						
Groups	Count	Sum	Average	Variance		
2HL 6 nodes	100	6510.897	65.10897	158.8548		
2HL 8 nodes	100	6552.137	65.52137	111.1614		

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	8.503523	1	8.503523	0.062985	0.802099	3.888853
Within Groups	26731.6	198	135.0081			
Total	26740.1	199				

Table 19: ANOVA test for the accuracy of 100 nets with 2HL with 9 and 10 nodes

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
2HL 9 nodes	100	6555.128	65.55128	149.8139
2HL 10 nodes	100	6518.162	65.18162	123.6895

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	6.8325	1	6.8325	0.049963	0.823358	3.888853
Within Groups	27076.84	198	136.7517			
Total	27083.68	199				

Table 20: ANOVA test for the accuracy of 100 nets with 2HL with 8 and 9 nodes

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
2HL 8 nodes	100	6552.137	65.52137	111.1614
2HL 9 nodes	100	6555.128	65.55128	149.8139

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	0.044742	1	0.044742	0.000343	0.985245	3.888853
Within Groups	25836.55	198	130.4876			
Total	25836.6	199				

Table 21: ANOVA test for the accuracy of 100 nets with 1HL with 5 and 12 nodes

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
1HL 5 nodes	100	6334.188	63.34188	138.3708
1HL 12 nodes	100	6682.265	66.82265	120.9692

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	605.7856	1	605.7856	4.671749	0.031863	3.888853
Within Groups	25674.66	198	129.67			
Total	26280.44	199				

Appendix B: ANOVA Test

Table 22: ANOVA test for the accuracy of 100 trained nets using TTT and TTP

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
TTT	100	7403.299	74.03299	14.67578
TTp	100	7391.239	73.91239	10.91539

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.727188	1	0.727188	0.056831	0.811823	3.888853
Within Groups	2533.525	198	12.79558			
Total	2534.253	199				

Table 23: ANOVA test for the accuracy of 100 trained nets using TTT and LLL

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
TTT	100	7403.299	74.03299	14.67578
LLL	100	6054.273	60.54273	149.0204

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	9099.351	1	9099.351	111.1737	6.47E-21	3.888853
Within Groups	16205.92	198	81.84808			
Total	25305.27	199				

Table 24: ANOVA test for the accuracy of 100 trained nets using LLP and LLT

Anova: Single Factor

SUMMARY						
Groups	Count	Sum	Average	Variance		
LLP	100	7390.82	73.9082	14.27368		
LLT	100	7389.957	73.89957	14.4234		

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.003722	1	0.003722	0.000259	0.987167	3.888853
Within Groups	2841.01	198	14.34854			
Total	2841.014	199				

Table 25: ANOVA test for the accuracy of 100 trained nets using TTL and PPT

Anova: Single Factor

SUMMARY						
Groups	Count	Sum	Average	Variance		
TTL	100	6775	67.75	144.986		
PPT	100	7229.487	72.29487	9.243244		

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1032.793	1	1032.793	13.39296	0.000324	3.888853
Within Groups	15268.69	198	77.11462			
Total	16301.49	199				

Table 26: ANOVA test for the accuracy of 100 trained nets using PPL and TTL

Anova: Single Factor

SUMMARY						
Groups	Count	Sum	Average	Variance		
PPL	100	6858.548	68.58548	122.1693		
TTL	100	6775	67.75	144.986		

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	34.90126	1	34.90126	0.261281	0.609811	3.888853
Within Groups	26448.37	198	133.5776			
Total	26483.27	199				

Table 27: ANOVA test for the accuracy of 100 trained nets using PPT and TTT

Anova: Single Factor

SUMMARY						
Groups	Count	Sum	Average	Variance		
PPT	100	7229.487	72.29487	9.243244		
TTT	100	7403.299	74.03299	14.67578		

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	151.0537	1	151.0537	12.63042	0.000475	3.888853
Within Groups	2367.983	198	11.95951			
Total	2519.037	199				

Appendix C: ANOVA Test

Table 28: ANOVA test for the accuracy of 100 trained nets using RP and LM

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
RP	100	6457.479	64.57479	112.3306
LM	100	7473.291	74.73291	15.21479

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	5159.371	1	5159.371	80.90253	1.94E-16	3.888853
Within Groups	12626.99	198	63.77268			
Total	17786.36	199				

Table 29: ANOVA test for the accuracy of 100 trained nets using LM and GDX

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
LM	100	7473.291	74.73291	15.21479
GDX	100	7398.931	73.98931	10.33224

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	27.64697	1	27.64697	2.164398	0.142827	3.888853
Within Groups	2529.156	198	12.77352			
Total	2556.803	199				

Table 30: ANOVA test for the accuracy of 100 trained nets using BFG and CGB

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
BFG	100	6878.205	68.78205	97.61909	
CGB	100	6866.025	68.66025	101.1679	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.741738	1	0.741738	0.007463	0.931246	3.888853
Within Groups	19679.91	198	99.39351			
Total	19680.66	199				

Table 31: ANOVA test for the accuracy of 100 trained nets using SCG and OSS

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
SCG	100	7330.129	73.30129	35.24121	
OSS	100	7102.777	71.02777	88.68927	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	258.4435	1	258.4435	4.170782	0.042453	3.888853
Within Groups	12269.12	198	61.96524			
Total	12527.56	199				

Table 32: ANOVA test for the accuracy of 100 trained nets using OSS and GDX

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
OSS	100	7102.777	71.02777	88.68927	
GDX	100	7398.931	73.98931	10.33224	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	438.5351	1	438.5351	8.85737	0.003283	3.888853
Within Groups	9803.129	198	49.51075			
Total	10241.66	199				

Table 33: ANOVA test for the accuracy of 100 trained nets using CGF and RP

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
CGF	100	6490.171	64.90171	116.4179	
RP	100	6457.479	64.57479	112.3306	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	5.343867	1	5.343867	0.046723	0.82909	3.888853
Within Groups	22646.1	198	114.3742			
Total	22651.44	199				

Table 34: ANOVA test for the accuracy of 100 trained nets using CGP and LM

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
CGP	100	6849.572	68.49572	97.12171	
LM	100	7473.291	74.73291	15.21479	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1945.128	1	1945.128	34.63038	1.68E-08	3.888853
Within Groups	11121.31	198	56.16825			
Total	13066.44	199				

Table 35: ANOVA test for the accuracy of 100 trained nets using LM and SCG

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
LM	100	7473.291	74.73291	15.21479	
SCG	100	7330.129	73.30129	35.24121	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	102.4769	1	102.4769	4.062031	0.04521	3.888853
Within Groups	4995.145	198	25.228			
Total	5097.622	199				

Table 36: ANOVA test for the accuracy of 100 trained nets using LM and GDX

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
LM	100	7473.291	74.73291	15.21479	
GDX	100	7398.931	73.98931	10.33224	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	27.64697	1	27.64697	2.164398	0.142827	3.888853
Within Groups	2529.156	198	12.77352			
Total	2556.803	199				

Appendix D: ANOVA test

Table 37: ANOVA test for the accuracy of 100 nets using a momentum of 0.7 and 0.75

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
M=0.7	100	6682.265	66.82265	120.9692	
M=0.75	100	6687.393	66.87393	108.7934	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.131513	1	0.131513	0.001145	0.973043	3.888853
Within Groups	22746.5	198	114.8813			
Total	22746.63	199				

Table 38: ANOVA test for the accuracy of 100 nets using a momentum of 0.8 and 0.95

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
M=0.8	100	6700.214	67.00214	126.452	
M=0.95	100	6687.393	66.87393	108.7934	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.821865	1	0.821865	0.006987	0.933467	3.888853
Within Groups	23289.3	198	117.6227			
Total	23290.12	199				

Table 39: ANOVA test for the accuracy of 100 nets using a momentum of 0.7 and 1.0

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
M=0.7	100	6682.265	66.82265	120.9692
M=1.0	100	6700.214	67.00214	126.452

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	1.610905	1	1.610905	0.013022	0.909265	3.888853
Within Groups	24494.7	198	123.7106			
Total	24496.31	199				

Appendix E: ANOVA Test

Table 40: ANOVA test using all AMSU-B but 89 GHz and all AMSU-B 150 GHz

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
All AMSU-B but 89GHz	100	6711.752	67.11752	81.5416
All AMSU-B but 150GHz	100	6749.359	67.49359	69.23608

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	7.071395	1	7.071395	0.093799	0.759724	3.888853
Within Groups	14926.99	198	75.38884			
Total	14934.06	199				

Table 41: ANOVA test using all AMSU-B but 183+/-1 GHz and all AMSU-B but 183+/-3 GHz

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
All AMSU-B but 183+/-1GHz	100	6630.769	66.30769	71.52471
All AMSU-B but 183+/-3GHz	100	6604.273	66.04273	84.94129

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	3.51019	1	3.51019	0.044868	0.832464	3.888853
Within Groups	15490.13	198	78.233			
Total	15493.64	199				

Table 42: ANOVA test using all AMSU-B but 89 and 150 GHz and all AMSU-B

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
All AMSU-B but 89 and 150GHz	100	5665.812	56.65812	60.68591
All AMSU-B	100	6679.273	66.79273	80.65633

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	5135.516	1	5135.516	72.66782	3.9E-15	3.888853
Within Groups	13992.88	198	70.67112			
Total	19128.4	199				

Table 43: ANOVA test using all AMSU-B but 89 GHz and all AMSU-B

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
All AMSU-B but 89 GHz	100	6711.752	67.11752	81.5416
All AMSU-B	100	6679.273	66.79273	80.65633

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	5.274265	1	5.274265	0.065035	0.798973	3.888853
Within Groups	16057.6	198	81.09897			
Total	16062.87	199				

Table 44: ANOVA test using all AMSU-B and snow cover information and only all

AMSU-B

Anova: Single Factor

SUMMARY

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
AMSU-B and snow cover info	100	6101.098	61.01098	84.0026
AMSU-B frequencies	100	6066.907	60.66907	85.26484

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	5.845088	1	5.845088	0.069063	0.792979	3.888853
Within Groups	16757.48	198	84.63372			
Total	16763.32	199				

Appendix F: ANOVA Test

Table 45: ANOVA test using one node in the output layer using different inputs

Anova: Single Factor

SUMMARY						
Groups	Count	Sum	Average	Variance		
89 and 150 GHz	100	5359.597	53.59597	57.31893		
All AMSU-B but 150 GHz	100	5502.526	55.02526	74.4018		

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	102.1431	1	102.1431	1.550903	0.214474	3.888853
Within Groups	13040.35	198	65.86037			
Total	13142.5	199				

Table 46: ANOVA test using one node in the output layer using different inputs

Anova: Single Factor

SUMMARY						
Groups	Count	Sum	Average	Variance		
All AMSU-B	100	5486.87	54.8687	77.11825		
All AMSU-B but 150 GHz	100	5502.526	55.02526	74.4018		

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1.225567	1	1.225567	0.016177	0.89892	3.888853
Within Groups	15000.49	198	75.76003			
Total	15001.71	199				

Table 47: ANOVA test using two nodes in the output layer using different inputs

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
89 and 150 GHz	100	5336.87	53.3687	79.66149	
All AMSU-B but 150 GHz	100	5465.657	54.65657	62.32921	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	82.93071	1	82.93071	1.168115	0.281104	3.888853
Within Groups	14057.08	198	70.99535			
Total	14140.01	199				

Table 48: ANOVA test using two nodes in the output layer using different inputs

Anova: Single Factor

SUMMARY					
Groups	Count	Sum	Average	Variance	
All AMSU-B but 89 GHz	100	5339.9	53.399	56.90446	
All AMSU-B but 150 GHz	100	5465.657	54.65657	62.32921	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	79.07449	1	79.07449	1.326379	0.250839	3.888853
Within Groups	11804.13	198	59.61683			
Total	11883.21	199				

Table 49: ANOVA test using all AMSU-B frequencies with one and two nodes

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
One Node	100	5502.526	55.02526	74.4018
Two Nodes	100	5465.657	54.65657	62.32921

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	6.796505	1	6.796505	0.099414	0.752866	3.888853
Within Groups	13536.37	198	68.36551			
Total	13543.17	199				

Appendix G: 500 Trained Nets

S-S: Snowfall pixels right classified.

S-NP: Snowfall pixels classified as non-precipitating pixels.

NP-S: Non-precipitating pixels classified as Snowfall pixels.

NP-NP: Non-precipitating pixels right classified.

Selected nets

Eliminated nets

Net #	PP	P-NP	NP-P	NP-NP
1	80	268	40	620
2	28	320	11	649
3	45	303	14	646
4	81	267	49	611
5	0	348	0	660
6	43	305	18	642
7	7	341	17	643
8	7	341	17	643
9	7	341	17	643
10	75	273	34	626
11	44	304	23	637
12	32	316	18	642
13	86	262	47	613
14	0	348	0	660
15	101	247	60	600
16	76	272	33	627
17	7	341	17	643
18	80	268	45	615
19	0	348	0	660
20	109	239	65	595
21	79	269	47	613
22	0	348	0	660
23	75	273	63	597
24	83	265	34	626
25	92	256	46	614
26	93	255	54	606
27	30	318	19	641
28	0	348	0	660
29	77	271	30	630
30	78	270	41	619
31	90	258	43	617
32	93	255	67	593
33	66	282	34	626
34	35	313	9	651
35	39	309	20	640
36	93	255	57	603
37	78	270	38	622
38	83	265	43	617
39	91	257	66	594

40	71	277	36	624
41	89	259	73	587
42	101	247	68	592
43	1	347	2	658
44	35	313	17	643
45	138	210	89	571
46	0	348	0	660
47	90	258	54	606
48	78	270	40	620
49	65	283	23	637
50	7	341	17	643
51	0	348	0	660
52	77	271	42	618
53	0	348	0	660
54	76	272	36	624
55	85	263	63	597
56	85	263	32	628
57	121	227	75	585
58	81	267	41	619
59	93	255	52	608
60	7	341	17	643
61	95	253	50	610
62	8	340	3	657
63	7	341	17	643
64	97	251	56	604
65	84	264	53	607
66	44	304	19	641
67	73	275	48	612
68	8	340	3	657
69	341	7	643	17
70	7	341	17	643
71	26	322	12	648
72	89	259	36	624
73	12	336	3	657
74	0	348	0	660
75	34	314	14	646
76	84	264	47	613
77	0	348	0	660
78	0	348	0	660
79	13	335	3	657
80	94	254	70	590
81	65	283	35	625
82	84	264	45	615
83	92	256	51	609
84	84	264	66	594
85	80	268	49	611
86	0	348	0	660
87	8	340	2	658
88	59	289	30	630
89	86	262	45	615
90	50	298	17	643
91	76	272	37	623
92	13	335	6	654
93	66	282	31	629
94	136	212	105	555
95	81	267	39	621
96	76	272	38	622
97	47	301	31	629
98	19	329	12	648
99	87	261	41	619
100	71	277	36	624
101	65	283	24	636
102	27	321	25	635
103	79	269	45	615
104	76	272	30	630
105	101	247	60	600
106	72	276	32	628
107	80	268	37	623
108	0	348	0	660
109	80	268	51	609
110	0	348	0	660
111	84	264	45	615
112	62	286	21	639
113	101	247	69	591

114	0	348	0	660
115	75	273	40	620
116	70	278	29	631
117	7	341	17	643
118	76	272	37	623
119	83	265	43	617
120	84	264	42	618
121	2	346	2	658
122	88	260	43	617
123	0	348	0	660
124	100	248	53	607
125	68	280	28	632
126	0	348	0	660
127	46	302	22	638
128	51	297	41	619
129	0	348	0	660
130	0	348	0	660
131	0	348	0	660
132	76	272	37	623
133	70	278	43	617
134	76	272	34	626
135	75	273	38	622
136	89	259	54	606
137	34	314	20	640
138	93	255	76	584
139	62	286	26	634
140	79	269	41	619
141	90	258	43	617
142	93	255	70	590
143	83	265	43	617
144	46	302	31	629
145	89	259	63	597
146	54	294	26	634
147	67	281	27	633
148	79	269	39	621
149	67	281	32	628
150	82	266	38	622
151	49	299	39	621
152	66	282	43	617
153	0	348	0	660
154	33	315	11	649
155	2	346	2	658
156	46	302	25	635
157	95	253	51	609
158	66	282	24	636
159	58	290	30	630
160	83	265	45	615
161	7	341	17	643
162	73	275	32	628
163	82	266	52	608
164	75	273	60	600
165	79	269	36	624
166	67	281	32	628
167	7	341	17	643
168	84	264	43	617
169	81	267	40	620
170	87	261	64	596
171	83	265	49	611
172	0	348	0	660
173	87	261	57	603
174	74	274	38	622
175	78	270	30	630
176	75	273	52	608
177	90	258	54	606
178	97	251	67	593
179	93	255	58	602
180	7	341	17	643
181	90	258	59	601
182	0	348	0	660
183	94	254	48	612
184	72	276	28	632
185	67	281	24	636
186	88	260	53	607
187	53	295	19	641

188	83	265	43	617
189	104	244	71	589
190	89	259	49	611
191	56	292	34	626
192	85	263	47	613
193	65	283	28	632
194	71	277	33	627
195	75	273	30	630
196	69	279	33	627
197	35	313	25	635
198	86	262	51	609
199	7	341	17	643
200	100	248	73	587
201	91	257	49	611
202	85	263	64	596
203	102	246	64	596
204	90	258	48	612
205	86	262	50	610
206	45	303	18	642
207	14	334	6	654
208	81	267	40	620
209	84	264	53	607
210	54	294	23	637
211	84	264	35	625
212	79	269	43	617
213	126	222	109	551
214	91	257	52	608
215	75	273	41	619
216	0	348	0	660
217	7	341	17	643
218	80	268	37	623
219	61	287	25	635
220	73	275	34	626
221	81	267	56	604
222	77	271	43	617
223	81	267	47	613
224	7	341	17	643
225	0	348	0	660
226	81	267	45	615
227	76	272	41	619
228	43	305	20	640
229	76	272	34	626
230	51	297	28	632
231	97	251	73	587
232	14	334	4	656
233	75	273	27	633
234	7	341	17	643
235	69	279	39	621
236	81	267	54	606
237	61	287	27	633
238	42	306	14	646
239	0	348	0	660
240	78	270	46	614
241	91	257	56	604
242	96	252	66	594
243	92	256	60	600
244	83	265	36	624
245	79	269	50	610
246	82	266	40	620
247	7	341	17	643
248	85	263	51	609
249	73	275	37	623
250	70	278	32	628
251	87	261	56	604
252	83	265	47	613
253	82	266	60	600
254	97	251	76	584
255	30	318	10	650
256	88	260	60	600
257	89	259	50	610
258	0	348	0	660
259	59	289	26	634
260	0	348	0	660
261	85	263	43	617

262	73	275	36	624
263	80	268	43	617
264	7	341	17	643
265	85	263	42	618
266	53	295	26	634
267	0	348	0	660
268	96	252	47	613
269	75	273	37	623
270	7	341	17	643
271	73	275	33	627
272	6	342	1	659
273	84	264	47	613
274	96	252	52	608
275	48	300	27	633
276	47	301	28	632
277	12	336	3	657
278	0	348	0	660
279	82	266	37	623
280	78	270	39	621
281	92	256	45	615
282	53	295	21	639
283	39	309	30	630
284	82	266	49	611
285	107	241	64	596
286	97	251	74	586
287	38	310	16	644
288	97	251	75	585
289	82	266	45	615
290	36	312	14	646
291	61	287	23	637
292	109	239	69	591
293	107	241	57	603
294	26	322	10	650
295	348	0	660	0
296	7	341	17	643
297	74	274	39	621
298	123	225	90	570
299	0	348	0	660
300	50	298	35	625
301	80	268	47	613
302	7	341	17	643
303	0	348	0	660
304	94	254	43	617
305	66	282	32	628
306	0	348	0	660
307	63	285	22	638
308	3	345	2	658
309	96	252	65	595
310	52	296	15	645
311	7	341	17	643
312	57	291	25	635
313	28	320	25	635
314	68	280	29	631
315	72	276	35	625
316	59	289	26	634
317	107	241	62	598
318	141	207	86	574
319	78	270	51	609
320	80	268	36	624
321	60	288	22	638
322	73	275	33	627
323	0	348	0	660
324	87	261	39	621
325	85	263	73	587
326	36	312	15	645
327	65	283	22	638
328	0	348	0	660
329	0	348	0	660
330	117	231	84	576
331	21	327	6	654
332	0	348	0	660
333	49	299	22	638
334	70	278	31	629
335	28	320	18	642

336	0	348	0	660
337	81	267	41	619
338	75	273	38	622
339	92	256	64	596
340	85	263	53	607
341	71	277	34	626
342	0	348	0	660
343	0	348	0	660
344	92	256	59	601
345	76	272	36	624
346	67	281	35	625
347	82	266	67	593
348	72	276	36	624
349	0	348	0	660
350	76	272	32	628
351	81	267	45	615
352	7	341	17	643
353	75	273	35	625
354	99	249	66	594
355	36	312	19	641
356	85	263	62	598
357	7	341	17	643
358	71	277	37	623
359	78	270	27	633
360	0	348	0	660
361	90	258	61	599
362	77	271	45	615
363	0	348	0	660
364	7	341	17	643
365	66	282	32	628
366	47	301	16	644
367	75	273	29	631
368	81	267	50	610
369	85	263	47	613
370	84	264	54	606
371	28	320	24	636
372	0	348	0	660
373	81	267	54	606
374	74	274	41	619
375	7	341	17	643
376	103	245	71	589
377	82	266	40	620
378	3	345	2	658
379	29	319	12	648
380	56	292	30	630
381	67	281	32	628
382	7	341	17	643
383	81	267	40	620
384	72	276	36	624
385	42	306	18	642
386	39	309	27	633
387	348	0	660	0
388	78	270	41	619
389	72	276	34	626
390	90	258	44	616
391	0	348	0	660
392	81	267	41	619
393	113	235	65	595
394	88	260	50	610
395	55	293	25	635
396	93	255	68	592
397	50	298	18	642
398	77	271	47	613
399	63	285	35	625
400	57	291	18	642
401	56	292	19	641
402	38	310	21	639
403	59	289	25	635
404	88	260	52	608
405	59	289	25	635
406	81	267	42	618
407	89	259	56	604
408	76	272	35	625
409	59	289	29	631

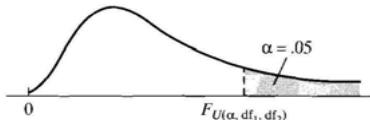
410	7	341	17	643
411	78	270	43	617
412	120	228	84	576
413	74	274	36	624
414	125	223	66	594
415	94	254	51	609
416	55	293	26	634
417	94	254	68	592
418	87	261	50	610
419	7	341	17	643
420	91	257	57	603
421	0	348	0	660
422	62	286	33	627
423	341	7	643	17
424	67	281	31	629
425	77	271	43	617
426	7	341	17	643
427	77	271	47	613
428	0	348	1	659
429	54	294	17	643
430	0	348	0	660
431	92	256	87	573
432	73	275	37	623
433	61	287	44	616
434	95	253	68	592
435	74	274	39	621
436	89	259	61	599
437	80	268	46	614
438	75	273	33	627
439	56	292	18	642
440	0	348	1	659
441	86	262	38	622
442	76	272	38	622
443	66	282	45	615
444	0	348	0	660
445	75	273	36	624
446	89	259	56	604
447	7	341	17	643
448	0	348	2	658
449	88	260	56	604
450	98	250	66	594
451	7	341	17	643
452	54	294	11	649
453	0	348	1	659
454	41	307	17	643
455	100	248	62	598
456	67	281	27	633
457	80	268	46	614
458	106	242	73	587
459	9	339	5	655
460	12	336	8	652
461	85	263	47	613
462	81	267	41	619
463	68	280	34	626
464	87	261	45	615
465	78	270	34	626
466	81	267	46	614
467	91	257	64	596
468	94	254	50	610
469	0	348	0	660
470	58	290	25	635
471	0	348	0	660
472	18	330	16	644
473	7	341	18	642
474	68	280	34	626
475	7	341	17	643
476	140	208	93	567
477	0	348	0	660
478	7	341	17	643
479	0	348	0	660
480	0	348	0	660
481	78	270	36	624
482	74	274	40	620
483	0	348	0	660

484	98	250	47	613
485	107	241	53	607
486	85	263	36	624
487	87	261	63	597
488	61	287	30	630
489	64	284	32	628
490	106	242	60	600
491	61	287	27	633
492	72	276	36	624
493	87	261	49	611
494	86	262	55	605
495	84	264	55	605
496	112	236	68	592
497	8	340	3	657
498	94	254	73	587
499	93	255	45	615
500	7	341	17	643
501	7	341	17	643
502	75	273	37	623
503	49	299	20	640
504	108	240	52	608
505	69	279	32	628
506	121	227	78	582
507	79	269	45	615
508	33	315	19	641
509	42	306	22	638
510	75	273	50	610
511	110	238	70	590
512	93	255	57	603
513	70	278	38	622
514	15	333	26	634
515	82	266	52	608
516	14	334	7	653
517	82	266	37	623
518	0	348	0	660
519	42	306	25	635
520	117	231	66	594
521	92	256	47	613
522	0	348	0	660
523	7	341	17	643
524	90	258	44	616
525	0	348	0	660
526	74	274	36	624
527	88	260	60	600
528	72	276	43	617
529	63	285	30	630
530	86	262	62	598

Appendix H: F Table

CRITICAL VALUES OF F

For a particular combination of numerator and denominator degrees of freedom, entry represents the critical values of F corresponding to a specified upper tail area (α)



Denominator d.f. ₂	Numerator, d.f. ₁																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Source: Reprinted from Pearson, E. S. and H. O. Hartley, eds.; Biometrika Tables for Statisticians, 3d ed.; 1966; by permission of the Biometrika Trustees, London

References

- Allen, R., Durkee, P., and Wash, C. (1989). “Snow/cloud discrimination with multispectral satellite measurements”. *Journal of Applied Meteorology*, vol. 29, pp. 994-1004.
- Augустејин, Е, Клеменс, Е., and Шоу, А. (1995). “Performance Evaluation of Texture Measures for Ground Cover Identification in Satellite Images by Means of a Neural Network Classifier”. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33-3, pp. 616-626.
- Bindlish, R., Crow, W., Jackson, T. (2004). “Potential role of passive microwave remote sensing in improving flood forecasts”. *IEEE International Geoscience and Remote Sensing Symposium*, vol. 3, pp. 1866-1869.
- Boucher, R., and Wieler, J. (1985). “Radar Determination of snowfall rate and accumulation”. *Journal of Applied Meteorology*, vol. 24-1, pp. 68-73.
- Benediktsson, A., and Sveinsson, R. (1997). “Feature extraction for multisource data classification with artificial neural networks”. *International Journal of Remote Sensing*, vol. 18-4, pp. 727-740.
- Carlson, P., and Marshall, J. (1972). “Measurement of snowfall by radar”. *Journal of Applied Meteorology*, vol. 11-3, pp. 494-500.
- Di Michele, S., and Bauer, P. (2005). “Passive microwave radiometer channel selection based on cloud and precipitation information content”. *Royal Meteorological Society*, vol.132, pp. 1299-1323.
- Ferraro, R., Smith, E., Berg, W., and Huffman, G (1998). “A screening methodology for passive microwave precipitation retrieval algorithm”. *American Meteorological Society*.

- Ferraro, R., Weng, F., Grody, N., Zhao, L., Meng, H., Kongoli, C., Pellegrino, P., Qiu, S., and dean, C. (2005). “NOAA operational hydrological products derived from the Advanced Microwave Sounding Unit”. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43-5, pp. 1036-1049.
- Foody, M., and Arora. K. (1997). “An evaluation of some factors affecting the accuracy of classification by an artificial neural network”. *International Journal of Remote Sensing*, vol. 18-4, pp. 799-810.
- Fox-Rabinovitz, M., Krasnopolksky, V., and Belochitski, A. (2006). “Ensemble of neural network emulations for climate model physics: the impact on climate simulations”. *International Joint Conference on Neural Network, Vancouver, Canada*.
- Fujiyoshi, Y., Endoh, T., Yamada, T., Tsuboki, K., Tachibana, Y., and Wakahama, G. (1990). “Determination of a Z-R relationship for snowfall using a radar and high sensitivity snow gauges”. *Journal of Applied Meteorology*, vol. 29-2, pp. 147-152.
- Ghedira, H., Bernier, M., and Ouarda, T. (2000). ” Application of neural networks for wetland classification in Radarsat SAR imagery”. *IEEE International Geosciences and Remote Sensing Symposium, IGARSS'2000*, vol. 2, pp. 675-677.
- Ghedira, H., and Bernier, M. (2004). "The effect of Some Internal Neural Network Parameters on SAR Texture Classification Performance." *In proc. IEEE International Geosciences and Remote Sensing Symposium*, vol-6, pp. 3845-3848.
- Kongoli, C., Pellegrino, P., Ferraro, R., Grody, N., and Meng, H.. (2003). “A new snowfall detection algorithm over land using measurements from the Advanced Microwave Sounding Unit (AMSU)”. *Geophysical Research Letters*, vol.30-14, CLM 6 1-4.

- Maclin, R., and Shavlik, J. (1995). “Combining the predictions of multiple classifiers: using competitive learning to initialize neural networks”. *14th International Joint Conference on Artificial Intelligence*.
- Markus, T., Powell, D., and Wang, J. (2006). “Sensitivity of passive microwave snow depth retrievals to weather effects and snow evolution”. *IEEE transactions on Geoscience and Remote Sensing*. vol. 44-1, pp. 68-77.
- MATLAB. Neural Network Tool box. User’s guide, version five (2006).
- Noh, Y., Liu, G., Seo, E., Wang, J., and Aonashi, K. (2005). “Development of a snowfall retrieval algorithm at high microwave frequencies”. *Submitted to Journal of Geophysical*.
- NOAA. (2004). NOAA-KLM Users Guide [Online]. Available: <http://www2.ncdc.noaa.gov/docs/klm/>.
- Opitz, D., and Maclin, R. (1999). “Popular ensemble methods: an empirical study”. *Journal of Artificial Intelligence Research*. vol. 11, pp. 169-198.
- Qiu, S., Pellegrino, P., Ferraro, R., and Zhao, L. (2005). “The improved AMSU ran-rate algorithm and its evaluation for a cool season event in the western United States”. *American Meteorological Society*, pp.761-774.
- Russell, S., and Norving, P. (2003). “Artificial Intellegence. A modern approach”. *Second Edition*.
- Scofield, R., and Kuligowski, R. (2003). “Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events”.
- Sekhon, R., and Srivastava, R. (1970). “Snow Size Spectra and Radar Reflectivity”. *Journal of the Atmospheric Sciences*, vol. 27-2, pp. 299-307.

- Schmid, W., and Mathis, A. (2004). “Validation of methods to detect winter precipitation and retrieve precipitation type”. 12th SIRWEC conference, Bingen, Germany.
- Skofronick-Jackson, G., Kim, M., Weinman, J., and Chang, D. (2004). “A physical model to determine snowfall over land by microwave radiometer”. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42-5, pp. 1047-1058.
- Tarassenko, L. (1998). “A guide to neural computing applications”.
- Tien, K., Judge, J., and Jacobs, J. (2004). “Passive microwave remote sensing of soil moisture, evaporatranspiration and vegetation properties during a growing season of cotton”. *IEEE International Geoscience and Remote Sensing Symposium*. vol.4, pp. 2795-2798.
- Ulaby, F., Moore, R., and Fung, A. (1981). “Microwave Remote Sensing, Active and Passive”. Vol-1, pp. 256-339.
- Yamamoto, C., Kishi, K., Sato, K., and Hara, F. (2004). “Importance of winter urban traffic issues and performance indicators as rated by businesses”. *Transportation Research E-Circular*, E-C063, pp. 219-236.



For Further Information Contact:

NOAA-CREST

The City College of The City University of New York
Steinman Hall, 140th St. @ Convent Ave.
New York, NY 10031

212-650-5465

Email: noaa-crest@ccny.cuny.edu

www.ccny.cuny.edu/noaa-crest