

**COMPUTATIONAL MODELS FOR THE  
PREDICTION OF SEVERE THUNDERSTORMS  
OVER EAST INDIAN REGION**

*Thesis submitted by*

**LITTA A J**

*in partial fulfilment of the  
requirements for the award of the degree of*

**DOCTOR OF PHILOSOPHY**

*Under the Faculty of Technology*

**DEPARTMENT OF COMPUTER SCIENCE  
Cochin University of Science and Technology  
Cochin - 682 022, Kerala, India**

*November 2013*

*COMPUTATIONAL MODELS FOR THE PREDICTION OF  
SEVERE THUNDERSTORMS OVER EAST INDIAN REGION*

*Ph.D Thesis*

*Author:*

*Litta A J*

*Department of Computer Science  
Cochin University of Science and Technology  
Cochin - 682 022, Kerala, India  
littaaj@gmail.com*

*Supervisor:*

*Dr. Sumam Mary Idicula*

*Professor and Head  
Department of Computer Science  
Cochin University of Science and Technology  
Cochin - 682 022, Kerala, India  
sumam@cusat.ac.in*

*November 2013*

## CERTIFICATE

This is to certify that the work presented in this thesis entitled **“Computational Models for the Prediction of Severe Thunderstorms over East Indian Region”** submitted to Cochin University of Science and Technology, in partial fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy in Computer Science** is a bonafide record of research work done by **Litta A. J.** in the Department of Computer Science, Cochin University of Science and Technology, under my supervision and guidance and the work has not been included in any other thesis submitted previously for the award of any degree.

Kochi  
November 2013

**Dr. Sumam Mary Idicula**  
(Supervising Guide)



## *Declaration*

I hereby declare that the work presented in this thesis entitled **“Computational Models for the Prediction of Severe Thunderstorms over East Indian Region”** submitted to Cochin University of Science and Technology, in partial fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy in Computer Science** is a record of original and independent research work done by me under the supervision and guidance of Dr. Sumam Mary Idicula, Professor and Head, Department of Computer Science, Cochin University of Science and Technology. The results presented in this thesis have not been included in any other thesis submitted previously for the award of any degree.

Kochi  
November 2013

**Litta A. J.**



## *Acknowledgement*

---

*The research work leading to PhD is a very intense process, which produced challenging, and interesting moments and experiences in my life. On those moments it was great to have the help of many people, who have strongly supported me in each phase of the thesis work. It is now my turn to express my gratitude to all of them.*

*First and foremost thanks and supreme glory to the God Almighty for providing me the health and wisdom towards the completion of this research work,*

*I would like to express my immense gratitude to Dr. Sumam Mary Idicula, Professor and Head, Department of Computer Science, Cochin University of Science and Technology for her guidance throughout the thesis work. I thank her for providing me an opportunity to do research under her supervision. Her knowledge, invaluable comments, caring and supportive attitude etc. were the main driving forces of my work,*

*I am grateful to Dr. K. Poulose Jacob, Professor, Department of Computer Science and Pro-Vice Chancellor, Cochin University of Science and Technology for providing constant encouragement, support and scholarly advice throughout my research period.*

*I would like to thank Dr. U. C. Mohanty, Professor, Centre for Atmospheric Sciences, Indian Institute of Technology (IIT), Delhi for providing guidance and necessary computer facilities to carry out the Numerical Weather Prediction (NWP)*

*modeling part of the thesis. His constant encouragement and valuable suggestions could help me in undertaking this observational and NWP modeling work,*

*Special thanks must also go to Dr. Someshwar Das, Scientist, National Centre for Medium Range Weather Forecasting, Noida, Dr. Ajith Abraham, Machine Intelligence Research Labs, Washington, Dr. K. Mohankumar, Dr. P. V. Joseph and Mr. B. Chakrapani, Department of Atmospheric Sciences, Cochin University of Science and Technology for their valuable suggestions and comments which have indeed helped me to improve my works. Thanks are due to the data providers specially India Meteorological Department (IMD).*

*I extend my sincere thanks to Dr. David Peter S, Ms. Sonia Sunny, all teaching and non-teaching staff and research scholars of my department for their cordial relation and help. I am thankful to all my teachers throughout my education for making me what I am today.*

*I owe heartfelt thanks to my parents, Mr. A. U. John and Mrs. Rosy A. A, whose encouragement and support always kept me overcoming hard times during this work, I extend my gratitude to my husband Mr. C. Naveen Francis, who has always been a constant source of energy, support and encouragement during the difficult situations. Thanks to all my relatives, friends and well-wishers for their good wishes, love and support at various stages of my study.*

***Litta A. J.***



## *Preface*

Thunderstorm is one of the most spectacular weather phenomena in the atmosphere. Many parts over the Indian region experience thunderstorms at higher frequency during pre-monsoon months (March-May), when the atmosphere is highly unstable because of high temperatures prevailing at lower levels. Most dominant feature of the weather during the pre-monsoon season over the eastern Indo-Gangetic plain and northeast India is the outburst of severe local convective storms, commonly known as 'Nor'wester' or 'Kalbaishakhi'. The severe thunderstorms associated with thunder, squall line, lightning and hail cause extensive losses in agriculture, damage to structure and also loss of life. The casualty due to lightning associated with thunderstorms in this region is the highest in the world. The highest numbers of aviation hazards are reported during occurrence of these thunderstorms. In India, 72% of tornadoes are associated with this thunderstorm.

The severe thunderstorms have significant socio-economic impact over eastern and northeastern parts of India. An accurate location specific and timely prediction is required to avoid loss of lives and property due to strong winds and heavy precipitation associated with

these storms. Forecasting thunderstorms is one of the most difficult tasks, due to their rather small spatial and temporal extension and the inherent non-linearity of their dynamics and physics. The improvement in prediction of these important weather phenomena is highly handicapped due to lack of observations and insufficient understanding. Realizing the importance of improved understanding and prediction of this weather event, an attempt is made to study severe thunderstorms during the pre-monsoon season of 2006, 2007 and 2009. The improvement in the prediction of this severe weather phenomenon has been done in this work using empirical and dynamical approaches. The most widely used empirical approach for weather prediction is artificial neural network (ANN). ANN based approach can be used to model complex relationships between inputs and outputs or to find patterns in data. The recent advances in neural network methodology for modeling nonlinear, dynamical phenomena along with the impressive successes in a wide range of applications, motivated to investigate the application of ANNs for the prediction of thunderstorms.

The second approach is based upon equations and forward simulations of the atmosphere, and is often referred to as computer modeling (Numerical Weather Prediction (NWP)). These models are computer programs that take the analysis as the starting point and evolve

the state of the atmosphere forward in time using the understanding of physics and fluid dynamics. The complicated equations which govern how the state of a fluid changes with time require high performance computers to solve them. The output from the model provides the basis of the thunderstorm forecast. Accurate prediction requires knowledge about “where” and “when” storms will develop and how they will evolve. NWP models can allow forecasters to anticipate not only, whether or not thunderstorms will develop in an environment, but also such things as thunderstorm movement, type, severity and longevity. In India, studies related to modeling of clouds are very scarce, particularly in intense thunderstorm events. Understanding the importance of these weather events and their socio-economic impact, this research has been initiated for analyzing and predicting severe thunderstorm events over east Indian region with most commonly known NWP models namely Non-hydrostatic Mesoscale Model (NMM) and Advanced Research WRF (ARW) model core of Weather Research and Forecasting (WRF) system.

The thesis, presented in seven chapters deals with the work carried out in designing and developing computational models for the prediction of severe thunderstorms over east Indian region.

Chapter 1 introduces the severity of thunderstorms over Indian region and its social impact and prediction challenges.

Chapter 2 provides a brief description about computational models used for the prediction of thunderstorms over Indian region. It gives the introduction to ANN and NWP modeling. In this chapter, the definition of neural network, a brief history, the architecture of neural networks, the various activation functions used, the different learning processes, and the various learning algorithms are dealt with. This chapter also gives a brief introduction of numerical modeling and its governing equations, grid structure, boundary conditions and parameterization. The details of WRF modeling system and its dynamic cores like ARW and NMM models are also introduced.

Chapter 3 describes the design and development of neural network model for the prediction of thunderstorms over Kolkata. In this work, the capabilities of six different learning algorithms in predicting thunderstorms were studied and their performances were compared. The results indicate that multilayer perceptron network (MLPN) model with Levenberg-Marquardt (LM) algorithm well predicted thunderstorm affected surface parameters as compared to other learning algorithms. This model was tried to find its usefulness for the advanced prediction of thunderstorms with 1, 3, 6, 12 and 24 h gaps. The results show that 1 h, and 3 h MLPN models are able to predict hourly temperature and relative humidity adequately with sudden fall and rise.

Chapter 4 discusses the prediction of thunderstorms using the NMM model. In this study, an attempt has been made to understand the relative role of initial conditions, convective parameterization schemes and microphysics schemes for thunderstorm predictability. Three sets of initial conditions are experimented using NMM model for a thunderstorm event on 20 May 2006. The trends shown by various meteorological fields of 24 h simulation were in good agreement with each other and very much consistent with dynamic and thermodynamic properties of the atmosphere for the occurrence of a severe thunderstorm. The sensitivity experiments are conducted with NMM model by changing the convective parameterization schemes for two severe thunderstorm cases (20 May 2006 and 21 May 2007) at Kolkata and validated the model results with observation. This study shows that the prediction of thunderstorm affected parameters is sensitive to convective schemes. The Grell and Devenyi scheme is well predicted the thunderstorm activities, in terms of time, intensity and the region of occurrence of the events, as compared to other convective schemes and also explicit scheme. Another sensitivity experiments have been conducted with three microphysics schemes for a severe thunderstorm event on 15 May 2009. The results show that the NMM model with Ferrier microphysics scheme appears to reproduce the cloud and

precipitation processes more realistically than other schemes for the prediction of this severe thunderstorm event.

Chapter 5 gives a comparative study of two numerical models namely NMM and ARW in the prediction of severe thunderstorms. In this study, an attempt has been made to compare the predicted results of severe thunderstorm events during May 2009 and validated the model results with the observations. Both models are able to broadly reproduce several features of the thunderstorm events, such as spatial pattern and temporal variability over east region of India. Comparison of model simulated thunderstorm affected parameters with that of the observed show that NMM has performed better than ARW in capturing the sharp rise in humidity and drop in temperature. The genesis, intensification and propagation of thunderstorms are well captured by NMM model than ARW.

Chapter 6 gives the performance evaluation of computational models for the prediction of thunderstorms over Kolkata. For this, the performance of thunderstorm affected parameters for the prediction of thunderstorm events using ANN and NWP models (NMM and ARW) were considered. The 24 h forecast data of surface temperature and relative humidity at Kolkata during severe thunderstorm days of May 2009 were used to test these models. Performance and reliabilities of the

models were then evaluated by a number of statistical measures. Comparison of observed and simulated results from 3 models indicates the superiority of NMM model in simulating thunderstorm over Kolkata on these severe thunderstorm cases. The results suggest that NMM model holds promise for prediction of surface weather parameters with reasonable accuracy in severe thunderstorm cases over east Indian region.

Chapter 7 gives a brief summary and conclusions of the work and the future directions of ANN and NWP model studies.

.....କରକର.....





# Contents

List of figures .....	vii
List of Tables.....	xv
List of Abbreviations .....	xix
<b>1. INTRODUCTION.....</b>	<b>1</b>
1.1 Structure and Formation of Thunderstorms.....	4
1.2 Severe Thunderstorms in India.....	10
1.2.1 Tornadoes in India.....	18
1.2.2 Hailstorms in India.....	21
1.3 Objectives.....	24
1.4 Layout of the Thesis.....	26
<b>2. COMPUTATIONAL MODELING .....</b>	<b>29</b>
2.1 Neural Network Modeling .....	31
2.1.1 Architecture of neural networks.....	35
2.1.2 Types of learning in neural networks .....	38
2.1.3 Activation functions .....	40
2.1.4 Multilayer perceptron network.....	41
2.1.5 The learning algorithms.....	47
2.1.6 Applications of neural networks.....	52
2.2 Numerical Weather Prediction Modeling.....	54
2.2.1 Governing equations .....	55
2.2.2 Grid structure .....	58

2.2.3 Vertical levels.....	59
2.2.4 Spatial boundary conditions .....	60
2.2.5 Parameterizations .....	61
2.2.6 Space scale.....	63
2.3 WRF Modeling System .....	65
2.3.1 ARW model.....	67
2.3.2 NMM model.....	69
2.3.3 WRF system requirements.....	70
2.3.4 WRF software framework.....	72
2.3.5 WRF workflow .....	75
<b>3. ARTIFICIAL NEURAL NETWORK MODEL FOR THUNDERSTORM PREDICTION .....</b>	<b>81</b>
3.1 Data and Methodology.....	85
3.1.1 ANN experimental setup.....	85
3.1.2 Statistical analysis.....	90
3.2 Learning Algorithms for the Present Study.....	93
3.2.1 STP algorithm .....	95
3.2.2 MOM algorithm.....	96
3.2.3 CG algorithm.....	97
3.2.4 LM algorithm.....	97
3.2.5 QKP algorithm .....	98
3.2.6 DBD algorithm.....	98
3.3 Case Description.....	100

3.4 Results and Discussion.....	102
3.4.1 Comparison of learning algorithms .....	103
3.4.2 Comparison of different advanced predictions .....	111
3.5 Chapter Summary .....	120
<b>4. WRF-NMM MODEL FOR THUNDERSTORM PREDICTION .....</b>	<b>123</b>
4.1 Data and Methodology.....	126
4.1.1 Initial and boundary conditions.....	127
4.1.2 Experiment 1- Study with different initial conditions .....	130
4.1.3 Experiment 2 - Study with different CPSs .....	132
4.1.4 Experiment 3 - Study with different microphysics schemes .....	133
4.1.5 Observational data.....	134
4.2 Case Description .....	136
4.3 Results and Discussion .....	137
4.3.1 Sensitivity study with different initial conditions.....	137
4.3.1.1 Stability indices.....	138
4.3.1.2 Surface parameters .....	143
4.3.1.3 Composite radar reflectivity .....	147
4.3.2 Sensitivity study with different CPSs.....	154
4.3.2.1 Stability indices .....	154
4.3.2.2 Surface parameters .....	157
4.3.2.3 Composite radar reflectivity.....	165
4.3.3 Sensitivity study with different microphysics schemes .....	170

4.3.3.1 Stability indices .....	171
4.3.3.2 Surface parameters .....	172
4.3.3.3 Composite radar reflectivity .....	178
4.4 Chapter Summary .....	182
<b>5. COMPARISON OF NUMERICAL MODELS FOR THUNDERSTORM PREDICTION .....</b>	<b>187</b>
5.1 Data and Methodology .....	191
5.2 Results and Discussion .....	194
5.2.1 Analysis of stability indices .....	195
5.2.2 Analysis of surface relative humidity and temperature .....	197
5.2.3 Analysis of precipitation .....	203
5.2.4 Analysis of composite radar reflectivity .....	207
5.2.5 Analysis of cloud top temperature .....	212
5.3 Chapter Summary .....	224
<b>6. EVALUATION OF COMPUTATIONAL MODELS FOR THUNDERSTORM PREDICTION .....</b>	<b>227</b>
6.1 Data and Methodology .....	231
6.1.1 Numerical model .....	231
6.1.2 ANN model .....	232
6.2 Results and Discussion .....	233
6.2.1 Analysis of surface relative humidity .....	233
6.2.2 Analysis of surface temperature .....	236
6.3 Chapter Summary .....	242

<b>7. SUMMARY AND CONCLUSIONS</b> .....	<b>245</b>
7.1 Future Directions.....	249
<b>REFERENCES</b> .....	<b>253</b>
<b>LIST OF PUBLICATIONS</b> .....	<b>281</b>
<b>APPENDIX</b> .....	<b>285</b>



## List of Figures

<i>Table No</i>	<i>Title</i>	<i>Page No</i>
1.1	Structure of a severe thunderstorm.....	6
1.2	The life cycle of an ordinary single-cell thunderstorm: (a) towering cumulus stage (b) mature stage (c) dissipating stage.....	8
1.3	Annual number of thunderstorm days.....	13
1.4	Pie diagram showing the percentage of occurrence of thunderstorms for six different zones.....	15
1.5	Climatological annual thunderstorm days over India.....	17
1.6	Climatological frequency (1981-2009) of thunderstorm occurrences over Dum Dum (Kolkata) during April and May.....	17
1.7	The distribution of tornadoes in the Indian subcontinent.....	20
1.8	The monthly frequency of tornadoes in Indian subcontinent between 1839 and 1999.....	20
1.9	Photographs of the tornado over Orissa of 31 March 2009 and a typical damage photograph due to the tornado .....	21
1.10	Mean annual frequency distribution of hail days.....	22
1.11	Hailstorm occurrences over India for a 100 year period.....	23
1.12	(a) Monthly distribution of moderate and severe hail for India (b) diurnal variation of hailstorms.....	25
2.1	Graphical representation of single $n$ -input artificial neuron.....	34

2.2	A single layer feed-forward network .....	36
2.3	An architecture of feed-back network .....	37
2.4	Common non-linear function used for synaptic inhibition. Soft non-linearity: (a) sigmoid and (b) tanh; Hard non-linearity: (c) signum and (d) step.....	42
2.5	Activity diagram for the learning problem in the multilayer perceptron.....	45
2.6	The MLPN with three layers.....	46
2.7	Arrangement of the variables in a staggered grid cell.....	58
2.8	Horizontal cross section of a nested grid structure. Density fields are placed at the center and velocity fields are on the edges of each grid square .....	61
2.9	WRF modeling system infrastructures .....	67
2.10	WRF software framework .....	73
2.11	WRF workflow chart .....	79
3.1	The geographical location of Kolkata in West Bengal.....	84
3.2	Basic flow for designing ANN model.....	87
3.3	Architecture of MLPN for the prediction of (a) temperature (b) relative humidity.....	89
3.4	Comparison of ANN predicted hourly surface temperature ( $^{\circ}$ C) using different learning algorithms with observation on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.....	106
3.5	Comparison of ANN predicted hourly relative humidity (%) using different learning algorithms with observation on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.....	107



3.6	Performance accuracy of learning algorithms for the prediction of (a) temperature and (b) relative humidity during thunderstorm days.....	111
3.7	Comparison of ANN predicted hourly temperature ( $^{\circ}$ C) using different advanced prediction models with observation on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.....	114
3.8	Comparison of ANN predicted hourly relative humidity (%) using different advanced prediction models with observation on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009 .....	117
3.9	Performance accuracy of different advanced prediction models for the prediction of (a) temperature and (b) relative humidity during thunderstorm days.....	119
4.1	The geographical location of study area.....	127
4.2	Domain of NMM model.....	130
4.3	The inter-comparison of observed and model simulated (a) surface temperature ( $^{\circ}$ C) and (b) relative humidity (%) with different initial conditions over Kolkata valid for 20 May 2006 at 0000 UTC to 21 May 2006 at 0000 UTC .....	144
4.4	The inter-comparison of observed and model simulated diurnal variation of 24 h accumulated rainfall (mm) with different initial conditions over Kolkata valid for 20 May 2006 .....	147
4.5	The 3 h accumulated rainfall (mm) with different initial conditions over Kolkata valid for 20 May 2006 at 0900 UTC to 1200 UTC.....	148

4.6	Kolkata DWR composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 20 May 2006 .....	149
4.7	NMM simulated composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 20 May 2006 with Ex-1.....	151
4.8	NMM simulated composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 20 May 2006 with Ex-2.....	152
4.9	NMM simulated composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 20 May 2006 with Ex-3.....	153
4.10	The inter-comparison of observed and model simulated relative humidity (%) using different CPSs over Kolkata valid for (a) 20 May 2006 (b) 21 May 2007 .....	158
4.11	The inter-comparison of observed and model simulated temperature ( $^{\circ}$ C) using different CPSs over Kolkata valid for (a) 20 May 2006 (b) 21 May 2007 .....	162
4.12	The inter-comparison of observed and model simulated accumulated rainfall (mm) with different CPS over Kolkata valid for (a) 20 May 2006 (b) 21 May 2007 .....	164
4.13	The spatial distribution of 3 h accumulated rainfall (mm) between 0900 and 1200 UTC with different CPSs on 20 May 2006.....	166
4.14	The spatial distribution of 3 h accumulated rainfall (mm) between 0900 and 1200 UTC with different CPSs on 21 May 2007 .....	167

4.15	Kolkata DWR composite radar reflectivity (dBZ) imageries from 0800 to 1100 UTC on 21 May 2007 .....	168
4.16	NMM simulated composite radar reflectivity (dBZ) pictures from 0800 to 1100 UTC on 21 May 2007 with GD scheme .....	169
4.17	The inter-comparison of observed (AWS) and NMM model simulated diurnal variation of (a) surface temperature ( $^{\circ}$ C) (b) relative humidity (%) with different microphysics schemes over Kolkata valid on 15 May 2009 .....	174
4.18	The spatial distribution of 3 h accumulated rainfall (mm) between 1200 and 1500 UTC with different microphysics schemes on 15 May 2009 .....	176
4.19	Kolkata DWR composite radar reflectivity (dBZ) imageries from 1000 to 1300 UTC on 15 May 2009 .....	179
4.20	NMM simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009 using FERR microphysics scheme.....	180
4.21	NMM simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009 using WSM6 microphysics scheme.....	181
4.22	NMM simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009 using THOM microphysics scheme.....	182
5.1	Domain of NMM and ARW model.....	192
5.2	Inter-comparison of NMM and ARW model simulated and observed diurnal variation of surface relative	

	humidity (%) over Kolkata on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009 .....	199
5.3	Inter-comparison of NMM and ARW model simulated and observed diurnal variation of surface temperature ( $^{\circ}$ C) over Kolkata on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009 .....	202
5.4	Comparison of NMM and ARW simulated 24 h accumulated rainfall during 3 thunderstorm events (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009 .....	204
5.5	Kolkata DWR composite radar reflectivity (dBZ) imageries from 1000 to 1300 UTC on 3 May 2009 .....	209
5.6	NMM model simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 3 May 2009 .....	210
5.7	ARW model simulated composite radar reflectivity (dBZ) from 1000 to 1300 UTC on 3 May 2009.....	213
5.8	Kolkata DWR composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 11 May 2009 .....	214
5.9	NMM model simulated composite radar reflectivity (dBZ) pictures from 0900 to 1200 UTC on 11 May 2009 .....	215
5.10	ARW model simulated composite radar reflectivity (dBZ) pictures from 0900 to 1200 UTC on 11 May 2009.....	216
5.11	Kolkata DWR composite radar reflectivity (dBZ) imageries from 1000 to 1300 UTC on 15 May 2009 .....	217
5.12	NMM model simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009 .....	218

5.13	ARW model simulated composite radar reflectivity (dBZ) from 1000 to 1300 UTC on 15 May 2009.....	219
5.14	Kalpana satellite derived CTT ( $^{\circ}$ C) imageries from 1000 to 1300 UTC on 3 May 2009 .....	221
5.15	NMM model simulated CTT ( $^{\circ}$ C) from 1000 to 1300 UTC on 3 May 2009 .....	222
5.16	ARW model simulated CTT ( $^{\circ}$ C) from 1000 to 1300 UTC on 3 May 2009 .....	223
6.1	Inter-comparison of NMM, ARW and ANN models simulated and observed diurnal variation of surface relative humidity (%) over Kolkata on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009 .....	235
6.2	Inter-comparison of NMM, ARW and ANN models simulated and observed diurnal variation of surface temperature ( $^{\circ}$ C) over Kolkata on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.....	238
6.3	Performance accuracy of NMM, ARW and ANN models for the prediction of temperature (TMP) and relative humidity (RH) during 3 thunderstorm days .....	239



## List of Tables

<i>Table No</i>	<i>Title</i>	<i>Page No</i>
2.1	Scale definitions.....	64
2.2	A list of the supported combinations of hardware and software for WRF.....	71
3.1	Performance comparison of different learning algorithms in hourly temperature prediction.....	108
3.2	Performance comparison of different learning algorithms in hourly relative humidity prediction.....	109
3.3	Performance comparison of different advanced predictions for hourly temperature during thunderstorm days.....	115
3.4	Performance comparison of different advanced predictions for hourly relative humidity during thunderstorm days.....	118
4.1	The input meteorological parameters for NMM model.....	129
4.2	NMM model configuration.....	131
4.3	The different stability indices and their critical values for severe thunderstorm.....	141
4.4	NMM model simulated stability indices over Kolkata at 1200 UTC using different initial conditions.....	141
4.5	Statistical analysis of simulated and observed temperature and relative humidity over Kolkata based on MAE, RMSE and CC.....	146

4.6	The inter-comparison model simulated stability indices with different CPSs over Kolkata valid for 20 May 2006 at 1200 UTC (Case 1) and 21 May 2007 at 1100 UTC (Case 2).....	157
4.7	Statistical analysis of relative humidity with different CPSs over Kolkata valid for 20 May 2006 (Case 1) and 21 May 2007 (Case 2) .....	160
4.8	Statistical analysis of temperature with different CPSs over Kolkata valid for 20 May 2006 (Case 1) and 21 May 2007 (Case 2).....	163
4.9	NMM model simulated stability indices over Kolkata at 1300 UTC using different microphysics schemes.....	172
4.10	Statistical analysis of simulated and observed temperature and relative humidity over Kolkata based on MAE, RMSE and CC .....	175
4.11	The comparison of model simulated 24 h accumulated precipitation using different microphysics schemes of 6 meteorological stations with rain gauge observations.....	177
5.1	NMM and ARW model configuration .....	193
5.2	Comparison of NMM and ARW model simulated stability indices for three thunderstorm events during May 2009.....	196
5.3	Statistical analysis of simulated and observed relative humidity over Kolkata based on MAE, RMSE and CC.....	200
5.4	Statistical analysis of simulated and observed temperature over Kolkata based on MAE, RMSE and CC.....	203



5.5	Comparison of modeled precipitation of three thunderstorm cases with rain gauge observations.....	206
5.6	Statistical analysis of simulated and observed precipitation for three thunderstorm cases .....	207
6.1	Statistical analysis of simulated and observed relative humidity over Kolkata based on MAE, RMSE and CC.....	237
6.2	Statistical analysis of simulated and observed temperature over Kolkata based on MAE, RMSE and CC.....	240



## *List of Abbreviations*

ACPW	-	Asymmetric Coplanar Waveguide
AFWA	-	Air Force Weather Agency
AiWS	-	Air Weather Service
ANN	-	Artificial Neural Network
ARG	-	Automatic Rain Gauge
ARPS	-	Advanced Regional Prediction System
ARW	-	Advanced Research WRF
AS	-	Arakawa-Schubert
AWS	-	Automatic Weather Station
BMJ	-	Betts-Miller-Janjic
BP	-	Back Propagation
CAPE	-	Convective Available Potential Energy
CAPS	-	Center for Analysis and Prediction of Storms
CC	-	Correlation Coefficients
CG	-	Conjugate Gradient
CPS	-	Convective Parameterization Scheme
CRM	-	Cloud Resolving Model
CTT	-	Cloud Top Temperature
DBD	-	Delta Bar Delta
DWR	-	Doppler Weather Radar
EPS	-	Ensemble Prediction Systems
FAA	-	Federal Aviation Administration
FNL	-	NCEP Final analysis
FSL	-	Forecast System Laboratory
GD	-	Grell-Devenyi
GDAS	-	Global Data Assimilation System

GFDL	-	Geophysical Fluid Dynamics Laboratory
GFS	-	Global Forecast System
GrADS	-	Grid Analysis and Display System
GSLV	-	Geo-Stationary Launch Vehicle
GTS	-	Global Telecommunications System
GWB	-	Gangetic West Bengal
HPC	-	High Performance Computing
IMD	-	India Meteorological Department
KF	-	Kain-Fritsch
KI	-	K Index
LI	-	Lifted Index
LM	-	Levenberg Marquardt
MAE	-	Mean Absolute Error
MLPN	-	Multilayer Perceptron Network
MM5	-	Fifth-Generation Penn State/NCAR Mesoscale Model
MMM	-	Mesoscale and Microscale Meteorology
MOM	-	Momentum
MSE	-	Mean Square Error
MYJ	-	Mellor Yamada Janjic
NCAR	-	National Centre for Atmospheric Research
NCEP	-	National Centers for Environmental Prediction
NHAC	-	Northern Hemispherical Analysis Centre
NMM	-	Non-hydrostatic Mesoscale Model
NOAA	-	National Oceanic and Atmospheric Administration
NRL	-	Naval Research Laboratory
NWP	-	Numerical Weather Prediction

NWS	-	National Weather Service
PC	-	Percent Correct
QKP	-	Quick Propagation
QN	-	Quasi Newton
RAMS	-	Regional Atmospheric Modeling System
RMSE	-	Root Mean Square Error
STP	-	Step
TTI	-	Total Total Index
US	-	United States
USGS	-	United States Geological Survey
VHRR	-	Very High Resolution Radiometer
WMO	-	World Meteorological Organization
WPP	-	WRF Postprocessor
WPS	-	WRF Preprocessing System
WRF	-	Weather Research and Forecasting



## Chapter-1

### INTRODUCTION

---

#### Contents

- 1.1 Structure and Formation of Thunderstorms
  - 1.2 Severe Thunderstorms in India
  - 1.3 Objectives
  - 1.4 Layout of the Thesis
- 

A thunderstorm is a high frequency cumulus or cloud scale weather phenomenon characterized by the presence of lightning and its effect: thunder, which develops due to intense convection. It is usually accompanied by heavy rain and sometimes snow, hail, or no precipitation at all. Thunderstorms may line up in a series, and strong or severe thunderstorms may rotate which lead to catastrophe over the particular location. It is the towering cumulus or the cumulonimbus clouds of the convective origin and high vertical extent that are capable of producing lightning and thunder. The surface parameters play a significant role in the genesis whereas the strength of the upper air pull is required to assess the growth of the thunderstorms. Usually, thunderstorms have the spatial extent of a few kilometers and life span less than an hour. However multi-cell thunderstorms developed due to organized intense convection may have a life span of several hours and may travel over a few hundreds of

kilometers. Thunderstorms are one or more convective cells in which electrical discharges are seen as lightning or heard as thunder.

Each year, many people are killed or seriously injured by severe thunderstorms despite the advance warning. While severe thunderstorms are most common in the summer, they can occur just about any time of year. Many thunderstorms are typically short-lived (up to an hour) and limited in size (up to 10 km in diameter) but can traverse large distances during that time and are capable of inflicting significant damage (Kessler 1983). They can produce some hazardous weather conditions. Through lightning strikes, floods and tornadoes, thunderstorms have created massive property damage and death. Thunderstorms have been known to occur in almost every part of the world, although they are rare in the Polar Regions. Nearly 2000 thunderstorm cells are estimated to be present over the planet at any given time. It is estimated that globally there are 16 million thunderstorms each year. In the United States (US) the areas of maximum thunderstorm activity are the Florida peninsula and the coast of the Gulf of Mexico (70 - 80 days per year). The global distribution of thunderstorms is rather complex, but the influence of certain controls is visible. The frequency generally tends to decrease in colder seasons. There are relationships, although not perfect, with topography, land - sea configuration, air mass movements, and airflow on all scales. Thunderstorms are most frequent at low latitudes, where the atmosphere's low layers are heated mostly by contact with warm ground or



water and there by conditioned for an overturning process essential to thunderstorms.

Europe and Australia have few seasons with more than 20 thunderstorm days annually. In Asia, only in the southeastern sector and around Bangladesh does the frequency exceed 60. South America and Africa view for the most thunderstorm-prone continent. The pattern is intricate over central South America where as elsewhere additional data and examination of physical factors should contribute insight into the causes. The tropical oceanic regions around 20<sup>0</sup> north and south, regions of semi-permanent high pressure, are relatively free from thunderstorms. In the northern hemisphere, relatively few thunderstorms occur north of 10<sup>0</sup>N in winter (Dec-Feb). In the southern hemisphere the location of inter-tropical convergence zone dominates the pattern, although warm onshore winds and topography are also important, as in eastern Australia. Central Africa and Indonesia have been considered to have the world's greatest incidence of thunderstorms. Between 1916 and 1919, the city of Bogur in Indonesia averaged 322 thunderstorms per year. However the accepted record is 242 thunderstorm days per year, recorded over a 10 year period at Kampala, Uganda just north of Lake Victoria. In this area, as often elsewhere in the equatorial regions, local influences are very strong. Many regions of the world also have a seasonal preference for strong storms, including spring and summer for the south-central US,

June-August in the Sahel and March-May over the Gangetic Plain and Bangladesh.

Many parts over the Indian region experience thunderstorms at higher frequency during pre-monsoon months (March-May), when the atmosphere is highly unstable because of high temperatures prevailing at lower levels. The main regions of high thunderstorm activity in India are east-northeast India, southwest peninsula (particularly Kerala) and northwest India. There are as much as 30 to 40 days of thunderstorm in parts of east-northeast India and in south Kerala during this season. Thunderstorm activity progressively increases from March to May. Though there is considerable thunderstorm activity in India during the monsoon season, the severity of thunderstorms is marked only in the pre-monsoon season when they are accompanied by violent squalls.

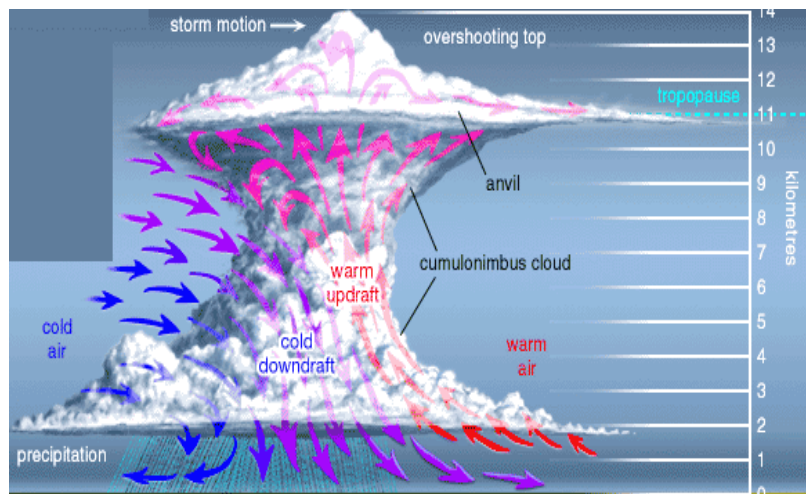
### **1.1 Structure and Formation of Thunderstorms**

Thunderstorms are generated by thermal instability in the atmosphere, and represent a violent example of convection - the vertical circulation produced in a fluid made thermally unstable by the local addition or subtraction of heat and the conversion of potential to kinetic energy. The convective overturning of atmospheric layers that sets up a thunderstorm is dynamically similar to convective circulations observed under laboratory conditions, where distinct patterns are generated in liquids by unequal heating. The orderly circulations produced in a

laboratory are rarely encountered in the atmosphere, where areas corresponding to the rising core of laboratory convective cells are marked by cumulus and cumulonimbus clouds. Clouds are parcels of air that have been lifted high enough to condense the water vapor they contain into very small, visible particles. These particles are too small and light to fall out as rain. As the lifting process continues, these particles grow in size by collision and coalescence until they are large enough to fall against the updrafts associated with any developing convective clouds. Cumulus (for accumulation) clouds begin their towering movement in response to atmospheric instability and convective overturning. Warmer and lighter than the surrounding air, they rise rapidly around a strong, central updraft. These elements grow vertically, appearing as rising mounds, domes, or towers. The atmospheric instability in which thunderstorms begin may develop in several ways. Radiational cooling of cloud tops, heating of the cloud base from the ground, and frontal effects may produce an unstable condition. This is compensated in air, as in most fluids, by the convective overturning of layers to put denser layers below less-dense layers. Figure 1.1 shows the structure of a severe thunderstorm.

Extensive studies indicate that thunderstorms go through a cycle of development from birth to maturity and to decay (Byers and Braham 1949). All thunderstorms, regardless of type, go through three stages: the cumulus stage, the mature stage, and the dissipation stage. Depending on the conditions present in the atmosphere, these three stages can take

anywhere from 20 minutes to several hours to occur. The first stage of a thunderstorm is the cumulus stage, or developing stage. In this stage, masses of moisture are lifted upwards into the atmosphere. The trigger for this lift can be insolation heating the ground producing thermals, areas where two winds converge forcing air upwards, or where winds blow over terrain of increasing elevation.



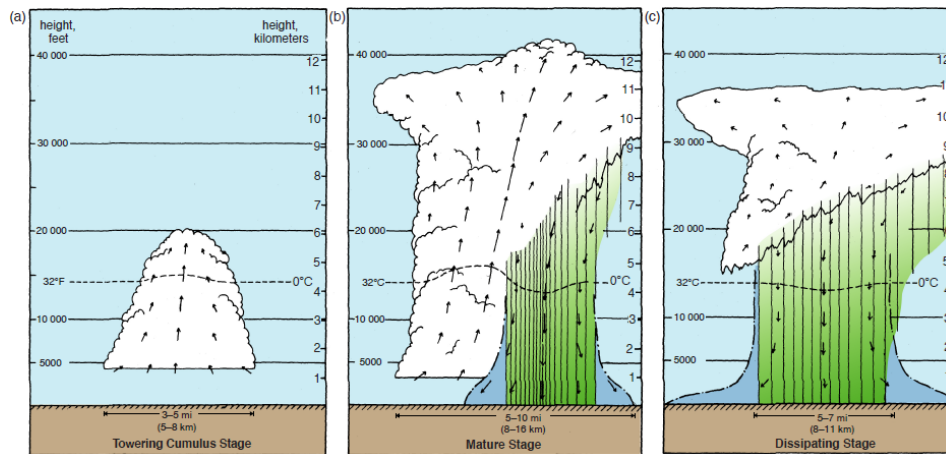
**Figure 1.1:** Structure of a severe thunderstorm (Britannica).

The moisture rapidly cools into liquid drops of water, which appears as cumulus clouds. As the water vapor condenses into liquid, latent heat is released which warms the air, causing it to become less dense than the surrounding dry air. The air tends to rise in an updraft through the process of convection (hence the term convective system). This creates a low-pressure zone beneath the forming thunderstorm. In a typical thunderstorm, approximately  $5 \times 10^8$  kg of water vapor are lifted,

and the amount of energy released when this condenses is about equal to the energy used by a city with population of 100,000 during a month.

In the mature stage of a thunderstorm, the warmed air continues to rise until it reaches existing air which is warmer, and the air can rise no further. Often this cap is the tropopause. The air is instead forced to exist, giving the storm a characteristic anvil shape. The resulting cloud is called cumulonimbus incus. The water droplets coalesce into heavy droplets and freeze to become ice particles. As these fall they melt to become rain. If the updraft is strong enough, the droplets are held aloft long enough to be so large that they do not melt completely and fall as hail. While updrafts are still present, the falling rain creates downdrafts as well. The simultaneous presence of both an updraft and downdrafts marks the mature stage of the storm and during this stage considerable internal turbulence can occur in the storm system, which sometimes manifests as strong winds, severe lightning, and even tornadoes. Typically, if there is little wind shear, the storm will rapidly enter the dissipating stage and rain itself out, but if there is sufficient change in wind speed and/or direction the downdraft will be separated from the updraft, and the storm may become a super-cell, and the mature stage can sustain itself for several hours. In certain cases however, even with little wind shear, if there is enough atmospheric support and instability in place for the thunderstorm to feed on, it may even maintain its mature stage a bit longer than most storms. In the dissipation stage, the thunderstorm is dominated by the

downdraft. If atmospheric conditions do not support super-cellular development, this stage occurs rather quickly, approximately 20-30 minutes into the life of the thunderstorm. The downdraft will push down out of the thunderstorm, hit the ground and spread out. The cool air carried to the ground by the downdraft cuts off the inflow of the thunderstorm, the updraft disappears and the thunderstorm will dissipate. Figure 1.2 shows the airflow during the three stages of thunderstorm.



**Figure 1.2:** The life cycle of an ordinary single-cell thunderstorm: (a) towering cumulus stage (b) mature stage (c) dissipating stage, from Markowski and Richardson (2010).

The types of thunderstorms could be classified as four, they are single-cell (Byers and Braham 1949), multi-cell cluster, multi-cell line (also called as squall line) (Browning 1962) and super-cell (Browning 1964). Which type forms depends on the instability and relative wind conditions at different layers of the atmosphere (wind shear). Single-cell

technically applies to a single thunderstorm with one main updraft. Within a cluster of thunderstorms, the term cell refers to each separate principal updraft. Thunderstorm cells can and do form in isolation to other cells. Such storms are rarely severe and are a result of local atmospheric instability; hence the term air mass thunderstorm. These are the typical summer thunderstorm in many temperate locales. They also occur in the cool unstable air which often follows the passage of a cold front from the sea during winter. While most single cell thunderstorms move, there are some unusual circumstances where they remain stationary. Multi-cell storms form as clusters of storms but may then evolve into one or more squall lines. They often arise from convective updrafts in or near mountain ranges and linear weather boundaries, usually strong cold fronts or troughs of low pressure.

Multi-cell line storms, commonly referred to as squall lines, occur when multi-cellular storms form in a line rather than clusters. They can be hundreds of miles long, move swiftly, and be preceded by a gust front. Heavy rain, hail, lightning, very strong winds and even isolated tornadoes can occur over a large area in a squall line. Bow echoes can form within squall lines, bringing with them even higher winds. An unusually powerful type of squall line called a derecho occurs when an intense squall line travels for several hundred kilometers, often leaving widespread damage over thousands of square kilometers. Occasionally, squall lines also form near the outer rain band of tropical cyclones. The

squall line is propelled by its own outflow, which reinforces continuous development of updrafts along the leading edge. Super-cell storms are large, severe quasi-steady-state storms which feature wind speed and direction that vary with height (wind shear), separate downdrafts and updrafts (i.e., precipitation is not falling through the updraft) and a strong, rotating updraft (a mesocyclone). These storms normally have such powerful updrafts that the top of the cloud (or anvil) can break through the troposphere and reach into the lower levels of the stratosphere and can be 24 km wide. These storms can produce destructive tornadoes, sometimes F3 or higher, extremely large hailstones (10 cm diameter), straight-line winds in excess of 130 kilometer per hour (kmph) and flash floods. In fact, most tornadoes occur from this type of thunderstorm.

## **1.2 Severe Thunderstorms in India**

A common feature of the weather during the pre-monsoon season (March-May) over the Indo-Gangetic plain and northeast India is the outburst of severe local convective storms, commonly known as ‘Nor’westers’ (as they move from northwest to southeast) or ‘Kalbaishakhi’ (which means calamities in the month of Baishakh) (Desai 1950). Nor’westers are mesoscale convective systems, which can develop under large-scale envelope of the seasonal low-level trough over West Bengal – Bihar – Jharkhand belt with a possible embedded low-pressure area. Nearly 28 severe thunderstorms occur in this region during April and May. Strong heating of landmass during mid-day initiates convection over Chhotanagpur



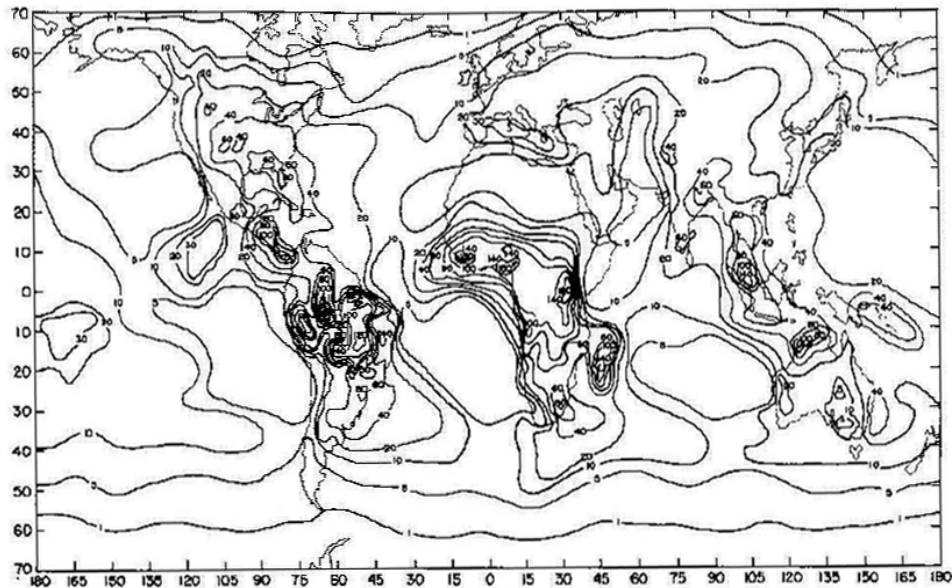
Plateau which moves southeast and gets intensified by mixing with warm moist airmass. The severe thunderstorms associated with thunder, squall lines, lightning and hail cause extensive losses in agriculture, damage to structure and also loss of life. The casualty due to lightning associated with thunderstorms in this region is the highest in the world. The strong wind produced by the thunderstorm downdrafts after coming in contact with the earth surface spreads out laterally and is referred as downbursts. These are real threat to aviation. The highest numbers of aviation hazards are reported during occurrence of these thunderstorms. In India, 72% of tornadoes are associated with Nor'westers (Science plan 2005).

Convective dust-storms occur over northwest India during the pre-monsoon season mid-March to mid-June. They are locally known as 'Andhi'. Over northwest India in the pre-monsoon season the lowest atmospheric layers have very high temperature and relatively low moisture content which makes the thunderstorms to have high bases above the ground of the order of 3 to 4 km. The ground being dry over long periods, there is loose and fine dust available in plenty. The rain falling down from these storms evaporate off before reaching the ground, particularly because of their high bases and the low relative humidity of the air below. These factors enable severe thunderstorms of northwest India generate dust-storms. Joseph et al. (1980) made a study of 40 cases of Andhi that occurred at Delhi airport during the period of 1973 to 1977, using a transmissometer (to measure the variation of horizontal visibility

as the dust wall moved across the airport), a weather radar (to study the movement of the associated thunderstorm cloud) and wind, temperature, humidity and pressure measuring instruments. From the nature of variations of horizontal visibility and wind speed near the ground level associated with these dust-storms, it was found that 4 types of Andhi occur. From the radar study it was found that the distance between the cumulonimbus cloud and the associated Andhi dust-wall on the ground can be as large as 30 km. It is observed that, the horizontal visibility is reduced to less than 100 meters during strong dust-storms at Delhi airport.

Considerable numbers of literatures are available on thunderstorm studies over the Indian region during the last three decades in which many successful investigations have been made to study the climatology on frequency, diurnal variation, month wise and season wise distribution of thunderstorms. The earliest study of thunderstorm frequency in India was by Dallas (1900) who took only 10 stations data of India during the year 1897. The first series of published charts of monthly frequency of days of thunder in India and neighborhood based on data for a short period was published in the climatological atlas for airmen (India Meteorological Department (IMD) 1943). The average monthly and annual frequency of thunderstorm days for all Indian and neighboring stations are given in the climatological tables of observatories in India (IMD 1953). Later on climatological tables have also been prepared by IMD based on data of 1931-60 (IMD 1969) and 1951-80 (IMD 1995). Simultaneously, the

World Meteorological Organization (WMO) published the average frequencies of thunderstorm days in the WMO publication (WMO 1953) and is shown in Figure 1.3. These averages are based on data for a uniform period of 15 years. In this (Figure 1.3), highest annual frequency of thunder in India is given as 60 days over east-northeast India.



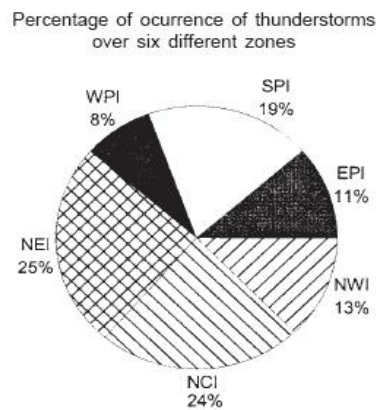
**Figure 1.3:** Annual number of thunderstorm days (WMO 1953).

Rao and Raman (1961) used data of 20 years to present monthly and annual frequency of thunderstorms in India. Their study showed highest thunderstorm activity occurs over east-northeast India including Assam, West Bengal, Jharkhand and Orissa. The annual average of thunderstorm frequency for these areas exceeds 75 days/year. Raman and Raghavan (1961) for the first time systematically studied the diurnal

variation of thunderstorm occurrence over India. Alvi and Punjabi (1966) examined diurnal variation in squalls which usually accompany thunderstorms. They also worked out annual frequency of thunderstorms as 75 days/year over Bangladesh, West Bengal, adjoining Orissa and northeast India. However, the northeast Assam is the most thundery area in India with average exceeding 100 days/year. The annual thunderstorm frequency is about 50 days over western Himalyas, southern parts of Kerala and adjoining Tamil Nadu. However later study by Rao (1981) gives maximum frequency of 60 to 80 days over West Bengal and adjoining Jharkhand and Orissa with relatively lower frequency of 40 to 60 days over Bangladesh and Assam, whereas annual mean number of thunderstorm given by Pant and Rupa Kumar (1997) shows thunderstorm activity of 60 days over northeast India, Bangladesh, West Bengal and adjoining areas with maximum number thunderstorm as 80 over northeast Assam.

Manohar et al. (1999) has been studied the average seasonal thunderstorm activity over India using monthly data from a large number of Indian stations. In this study, the latitudinal inter-month comparison of the thunderstorm activity during the pre-monsoon season showed a significant increase in the number of thunderstorm days, and their activity decreased with increasing latitude. Kandalgaonkar et al. (2005) made a climatological study by analyzing 30 years (1951–1980) of mean monthly thunderstorm days for six different zones North-West India (NWI),

North-Central India (NCI), North-East India (NEI), West-Peninsular India (WPI), South-Peninsular India (SPI), East-Peninsular India (EPI) with 260 Indian observatories spread uniformly over the country. Figure 1.4 shows the pie diagram of percentage of occurrence of thunderstorms for six different zones. From this figure it is seen that the highest (25%) percentage of occurrence of thunderstorm is noticed in NEI and the lowest (8%) in WPI, whereas the percentage of occurrence of TS in the other four zones is 24% in NCI, 13% in NWI, 11% in EPI and 19% in SPI.



**Figure 1.4:** Pie diagram showing the percentage of occurrence of thunderstorms for six different zones (Kandalgaonkar et al. 2004).

Tyagi (2007) studied the thunderstorm climatology over Indian region based on latest representative climatological data including 390 IMD observatories, 50 Indian Air Force (IAF) observatories, six Bangladesh observatories, two Pakistan observatories and one each in

Nepal and Sri Lanka. The study has brought out higher (100-120 days) annual frequency of thunderstorms as compared to those given by earlier studies (80-100 days). The highest annual frequency (100-120 days) is observed over Assam and Sub-Himalayan West Bengal in the east and Jammu region in the north (Figure 1.5). The lowest frequency (less than 5 days) is observed over Ladakh region. Mukherjee and Sen (1983) studied the diurnal variation of thunderstorm for some selected stations to understand the influence of different physical features viz., plain stations, hill stations, coastal stations, island stations etc.

In addition to above there have been several studies based on limited period of data like, Gupta and Chorghade (1961) studied thunderstorm occurrences at Agartala based on period of three years (1957-1959), Viswanathan and Faria (1962) for Bombay, Krishnamurthy (1965) for pune, Awadeshkumar (1992) for Lucknow, Moid (1995) for Mohanbari airport and Santosh et al. (2001) for three aerodrome stations in Kerala. Mukherjee (1964) showed that the frequency of thunderstorm over Guahati was highest in night time during pre-monsoon months. He reported that hills in the region plays profound role in the development of thunderstorm. Figure 1.6 shows the climatological frequency of thunderstorm occurrences over Dum Dum (Kolkata) station during April and May. The maximum number of thunderstorms occurred in the year of 1997 and minimum in 1987. An average of 16 numbers of thunderstorms

has occurred over this station which is very close to the average climatology of thunderstorm occurrence between 1951 and 1980.

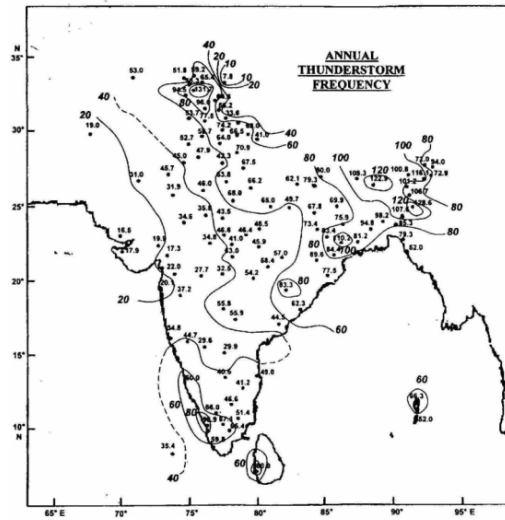


Figure 1.5: Climatological annual thunderstorm days over India (Tyagi 2007).

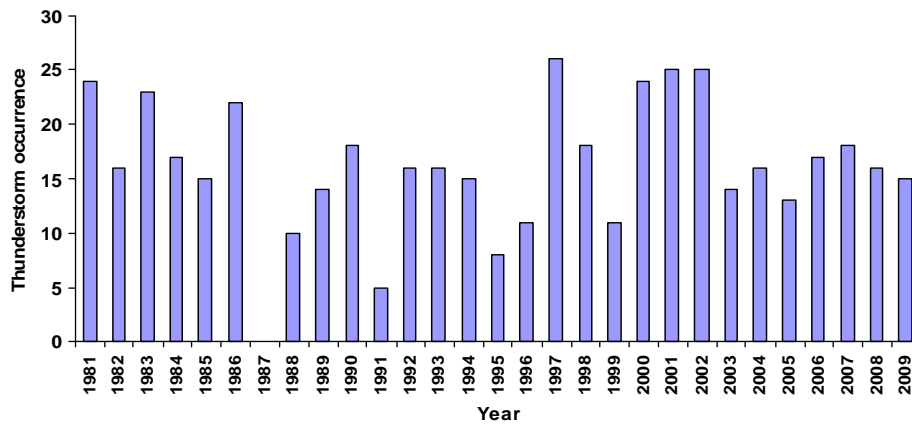


Figure 1.6: Climatological frequency (1981-2009) of thunderstorm occurrences over Dum Dum (Kolkata) during April and May.

### 1.2.1 Tornadoes in India

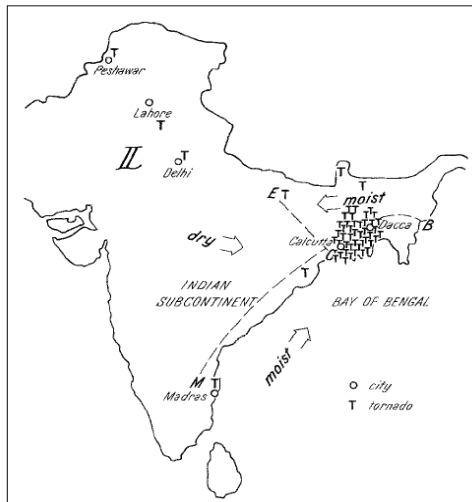
A tornado is a dark rotatory fragment of a severe thunderstorm cloud, generally super-cell type, descending down like a funnel, often swinging like the trunk of an elephant. It can have various other forms also. Tornadoes have been observed to occur in every continent except Antarctica. This dangerous phenomenon occurs mostly in the United States, but occasionally occurs in other parts of the world. India is also not free from occurrences of such tornadoes. Eastern parts of India particularly West Bengal and Orissa are vulnerable to tornadoes during pre-monsoon season (March-May). Several climatological studies of tornadoes for the Indian subcontinent have been conducted. The most comprehensive works were by Petersen and Mehta (1981), which documented 51 possible tornadoes across Bengal, 18 of which killed 10 people or more. Twelve of these occurred from 1838 to 1963 and 24 occurred after 1968. However, there might exist a tendency to report only the relatively significant tornadoes that leave more damage and attract more attention. Between 1972 and 1978, 13 tornado events occurred in the area approximately coinciding with Bangladesh. Figure 1.7 shows the distribution of tornadoes in the Indian subcontinent. Tornadoes concentrate in Bangladesh and east-northeastern India. Considering the entire area of the country, this gives a frequency rate of occurrence of about  $1 \times 10^{-5} \text{yr}^{-1} \text{km}^{-2}$  (Goliger and Milford 1998).



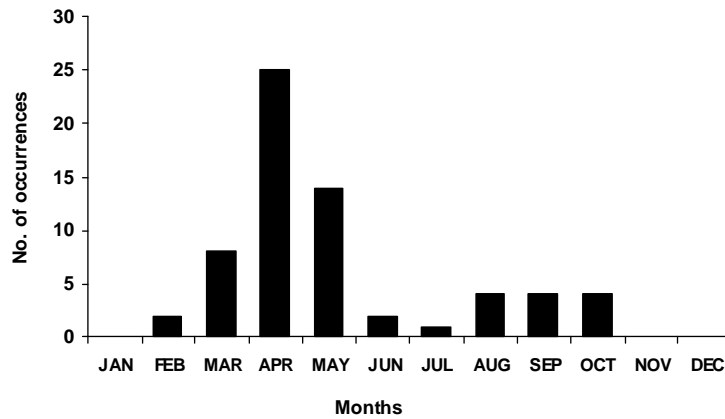
Saha (1967) tabulated relatively more prominent reported tornadoes in India during the period 1838 to 1950. This study shows that northeast India and neighbourhood are prone to tornado genesis more than other parts of India. Saha also pointed out that the region extending from Peshawar to Allahabad including Delhi gets tornadoes but less frequently than east-northeast India. Goldar et al. (2001) documented 36 possible spring tornadoes over West Bengal, 14 of which killed 10 people or more. While some events may not have been tornadic, this study partially fills the gap from the 1890's to early 1900's. Figure 1.8 shows the monthly frequency of tornadoes in the Indian subcontinent between 1839 and 1999. Most tornadoes occur during the pre-monsoon season, peaking in April. Other studies such as Singh (1981) have listed a few tornadoes for India. The associated wind speeds have been estimated to be of the order of 200-400 kmph. Litta et al. (2009) has been studied about a tornado (F3 on the Fujita-Pearson scale) over Rajkanika block of Kendrapara district of Orissa, India (20.7<sup>0</sup>N, 86.68<sup>0</sup>E) in the afternoon of 31 March 2009 (Figure 1.9). The devastation caused by the tornado consumed 15 lives, left several injured with huge loss of property.

Northwest India does not frequently experiences this violent weather phenomenon; but there have been a few cases over the region. In northern Delhi, 28 people were killed and 700 were injured by a tornado that cut a path 5 km long and about 50 m wide on 17 March 1978.

Another tornado is reported to have killed 10 people near Ludhiana (Punjab) on 10 March 1975 (Kumar and Singh 1978; Kumar et al. 1979).



**Figure 1.7:** The distribution of tornadoes in the Indian subcontinent (Petersen and Mehta 1995).



**Figure1.8:** The monthly frequency of tornadoes in Indian subcontinent between 1839 and 1999 (Goldar et al. 2001).

Bhan (2007) has been studied about a tornado (F0 on the Fujita-Pearson scale) close to Ludhiana airport (Punjab), northwest region of India on 15 August 2007. Relatively less damage occurred as it passed through in the open fields, but there were minor injuries to some cattle and damage to property. Although only a few tornadoes occur over this part of the country, they have a great potential of causing damage to property and loss of life.

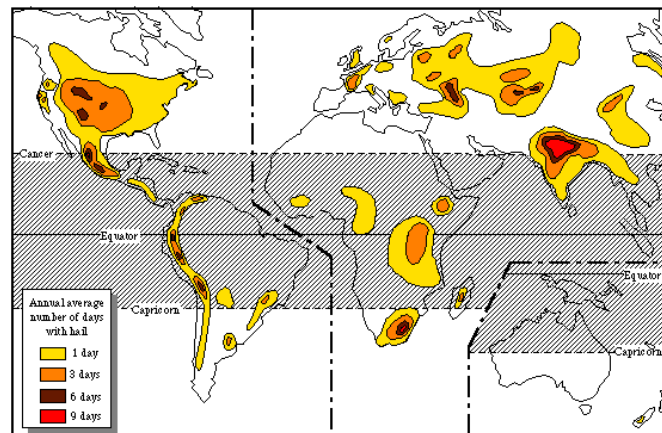


**Figure 1.9:** Photographs of the tornado over Orissa of 31 March 2009 and a typical damage photograph due to the tornado (orissadiary.com).

### **1.2.2 Hailstorms in India**

Severe thunderstorms tend to give precipitation, part of which reaches the ground as hail. Hail is more common along mountain ranges because mountains force horizontal winds upwards (known as orographic lifting), thereby intensifying the updrafts within thunderstorms and making hail more likely. Hailstorms are sufficiently important owing to their economic impact worldwide that some records are kept in most

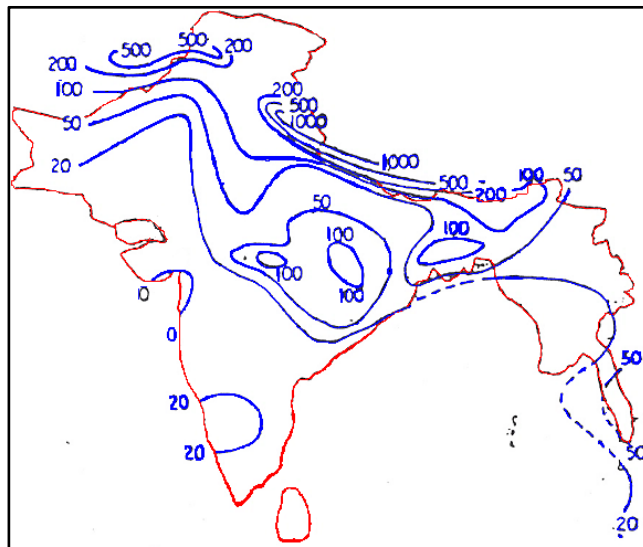
nations that have hail falls at all regularly. Figure 1.10 shows one picture of the global annual hail day (i.e., a day with one or more hail events) frequency distribution, indicating where the frequency is at least one hail day/year (Doswell and Bosart 2001). One of the more common regions for large hail is across the eastern and northeastern region of India, which reported one of the highest hail-related death tolls on record in 1888. China also experiences significant hailstorms. Across Europe, Croatia experiences frequent occurrences of hail. Hailstorms have been the cause of costly and deadly events throughout history. One of the earliest recorded incidents occurred around the 9th century in Roopkund, Uttarakhand, India (Gokhale 1975).



**Figure 1.10:** Mean annual frequency distribution of hail days (Doswell and Bosart 2001).

India is among the countries in the world having large frequency of hail. Figure 1.11 shows hail occurrences over India for a hundred year

period from Ramamurthy (1983). In the northeast Indian subcontinent the maximum average frequency of about one hailstorm annually is found in the foothills. In some areas as many as nine hailstorms have been reported in a year, and some very large hailstones probably occur in this region. The complex topography produces great variations over short distances. In the Irrawaddy delta of Burma the maximum frequency is during autumn (September-November), but in the northern hill stations it is from April to July or August. In the arid areas of southwestern Asia hail is rare, although occasional reports are received from the Yemeni highlands (Frisby and Sansom 1967).



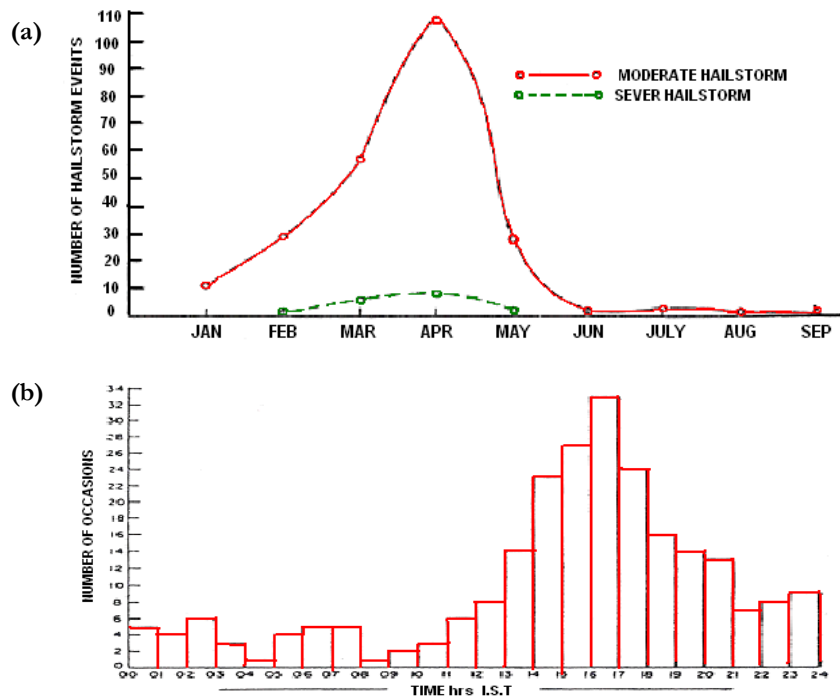
**Figure 1.11:** Hailstorm occurrences over India for a 100 year period (Ramamurthy 1983).

Reviewing the annual reports of IMD from 1982 to 1989, Nizamuddin (1993) finds that there were 228 hail days (about 29 per year) of moderate to severe intensity. Hail size comparable to mangoes, lemons and tennis balls has been observed. Eliot (1899) found that out of 597 hailstorms in India 153 yielded hailstones of diameter 3 cm or greater. These events killed 250 persons and caused extensive damage to winter wheat crops. India and Bangladesh are different from other northern hemisphere tropical stations in that hail is observed in the winter and pre-monsoon seasons with virtually no events after the onset of the southwest monsoon. Chaudhury and Banerjee (1983) show that the percentage of hailstorm days out of thunderstorm days decreases from 5% to less than 2% from March to May for east-northeastern India and Bangladesh. Figure 1.12 taken from their study shows the monthly distribution of moderate and severe hailstorms (Figure 1.12a) and the diurnal variation of hailstorms (Figure 1.12b).

### **1.3 Objectives**

Forecasting thunderstorms is one of the most difficult tasks in weather prediction, due to their rather small spatial and temporal extension and the inherent non-linearity of their dynamics and physics. The improvement in prediction of these important weather phenomena is highly handicapped due to lack of mesoscale observations and insufficient understanding. An accurate location specific and timely prediction is

required to avoid loss of lives and property due to strong winds and heavy precipitation associated with severe local storms.



**Figure 1.12:** (a) Monthly distribution of moderate and severe hail for India (b) diurnal variation of hailstorms (Chaudhury and Banerjee 1983).

This research is expected to improve both understanding and prediction of thunderstorms over Indian region. Brief objectives of this research works are as follows:

- Understand the genesis, development and propagation of severe thunderstorms over India.

- Thunderstorm prediction with Artificial Neural Network (ANN) model.
- Customization of Non-hydrostatic Mesoscale Model (NMM) core of Weather Research and Forecasting system (WRF) with improved forecast skill for the prediction of thunderstorms.
- Compare the skills of different numerical models namely NMM and Advanced Research WRF (ARW) for the prediction of severe thunderstorms.
- Evaluate the performance of ANN, ARW and NMM models for the thunderstorm prediction over Kolkata and find out suitable model.

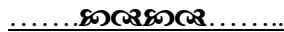
#### **1.4 Layout of the Thesis**

The rest of the thesis is laid out as follows:

- Chapter 2 provides a brief description about computational models used for the prediction of thunderstorms over Indian region.
- Chapter 3 describes the design and development of neural network model for the prediction of thunderstorms over Kolkata.
- Chapter 4 discusses the prediction of thunderstorms using NMM model and the sensitivity study of NMM model with different initial conditions, convective parameterization schemes and microphysics schemes.



- Chapter 5 gives a comparative study of two numerical models namely NMM and ARW in the prediction of severe thunderstorms.
- Performance evaluation of computational models namely ANN, NMM and ARW models for the prediction of severe thunderstorms over Kolkata are given in Chapter 6.
- A brief summary and conclusion of the work and the scope for future work are given in Chapter 7.
- References are listed after Chapter 7 along with the details of publications made by the author.





## Chapter- 2

# COMPUTATIONAL MODELING

---

### Contents

- 2.1 Neural Network Modeling
  - 2.2 Numerical Weather Prediction Modeling
  - 2.3 WRF Modeling System
- 

Accurate forecasting of thunderstorms and severe thunderstorms is critical for a large range of users in the community. The general public can benefit from timely forecasts and warnings of impending severe thunderstorms. The aviation industry in particular is one user group particularly affected by thunderstorms and one that can benefit greatly from enhanced forecasting services. In this case, the value can be expressed in terms of economic efficiency as well as in terms of safe operations of aircraft. Thunderstorm forecasting typically has proved to be one of the most difficult tasks, due to their rather small spatial and temporal extension and the inherent non-linearity of their dynamics and physics (Orlanski 1975). Generally, two methods are used to forecast weather: (a) the empirical approach and (b) the dynamical approach (Lorenz 1969).

The first approach is based upon the occurrence of analogues and is often referred as analogue forecasting. This is using past weather data to predict future events. The most widely used empirical approaches for weather prediction are regression, ANN, stochastic, fuzzy logic and group method of data handling. ANN based approach can be used to model complex relationships between inputs and outputs or to find patterns in data. ANN can be viewed as a mathematical model or computational model that is inspired by the structure or functional aspects of biological neural networks. ANNs are trainable self-adaptive systems that can "learn" to solve complex problems from a set of examples and generalize the "acquired knowledge" to solve unforeseen problems as in environmental prediction (Bishop 1995). The recent advances in neural network methodology for modeling non-linear, dynamical phenomena along with the impressive successes in a wide range of applications, motivated to investigate the application of ANNs for the prediction of thunderstorms.

The second approach is based upon equations and forward simulations of the atmosphere, and is often referred to as computer modeling. Weather forecasting using computer models is known as numerical weather prediction (NWP). NWP in the recent years emerged as an important discipline that requires increased computing power and accurate modeling. It is regarded as a part of the grand challenge application that requires teraflop/petaflop capabilities. The rapid advent in

the techno-scientific research related to high resolution downscaling/ forecasting of the weather in the last few decades has led to the development of complex mathematical models for several spatio-temporal scales of evolution of the atmosphere. These models are computer programs that take the analysis as the starting point and evolve the state of the atmosphere forward in time using the understanding of physics and fluid dynamics. The complicated equations which govern how the state of a fluid changes with time require high performance computers to solve them. Computer models have become an integral part of the forecast process by allowing forecasters to combine conceptual models of meteorological systems with forecast states of the atmosphere to produce accurate forecasts. Thunderstorms are one such phenomenon where computer models have led to real improvement in forecasts. Models can allow forecasters to anticipate not only whether or not thunderstorms will develop in an environment but also such things as thunderstorm movement, type, severity and longevity.

## **2.1 Neural Network Modeling**

Neural networks have seen an explosion of interest over the last few years, and are being successfully applied across an extraordinary range of problem domains, in areas as diverse as finance medicine, engineering, geology and physics. Indeed, anywhere that there are problems of prediction, classification or control, neural networks are being introduced. Neural networks could be define as an interconnected

of simple processing element whose functionality is based on the biological neuron. Biological neuron is a unique piece of equipment that carries information or a bit of knowledge and transfers to other neuron in the chain of networks. Artificial neuron imitates these functions and their unique process of learning (Fausett 1994). The interest in neural networks comes from the networks' ability to mimic human brain as well as its ability to learn and respond. As a result, neural networks have been used in a large number of applications and have proven to be effective in performing complex functions in a variety of fields. These include pattern recognition, classification, vision, control systems and prediction (Haykin 1994).

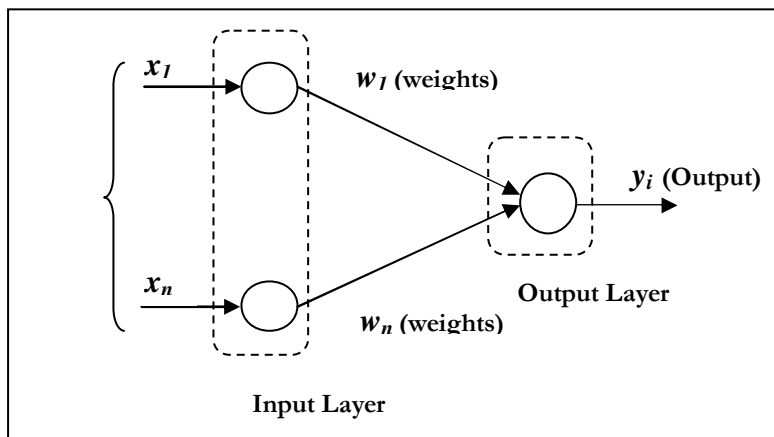
A first wave of interest in neural networks emerged after the introduction of simplified neurons by McCulloch and Pitts (1943) also known as connectionist models. Four years later, the same authors explored network paradigms for pattern recognition using a single layer perceptron (Pitts and McCulloch 1947). Hebb (1949) developed the first learning rule that is if two neurons are active at the same time then the strength between them should be increased. In the 1950s and 1960s, a group of researchers combined these biological and psychological insights to produce the first ANN (Rosenblatt 1962). Initially implemented as electronic circuits, they were later converted into a more flexible medium of computer simulation. However, researchers such as Minsky and Papert (1969) later challenged these works. They showed that perceptron could

not learn those functions which are not linearly separable. They strongly believed that intelligence systems are essentially symbol processing of the kind readily modeled on the von Neumann computer. Parker (1985) and LeCun (1986) discovered a learning algorithm for multi-layer networks called back propagation that could solve problems that were not linearly separable.

ANN is a network of collections of very simple processors ("Neurons") each possibly having a (small amount of) local memory. The units operate only on their local data and on the inputs they receive via the connections or links which are unidirectional (Ajith 2006). A network unit has a rule for summing the signals coming in and a rule for calculating an output signal that is then, sent to other network units. The rule for calculating the output is known as the activation function. ANNs can learn to perform a task through repeated adjustments of weights. During the learning process, a rule is used to decide by how much the value of the weight should be changed (Callen 1999). Negnevitsky (2002) mentioned that, there are more than a hundred different learning algorithms in ANNs. Adaptation or learning is a major focus of neural net research that provides a degree of robustness to the ANN model. In predictive modeling, the goal is to map a set of input patterns onto a set of output patterns. ANN accomplishes this task by learning from a series of input/output data sets presented to the network. The trained network is

then used to apply what it has learned to approximate or predict the corresponding output (Haykin 1994).

ANNs are nonlinear information (signal) processing devices which are built from interconnected elementary processing devices called neurons. An artificial neuron is a  $n$ -input single-output signal-processing element, which can be thought of a simple model of a non-branching biological neuron. Graphically, an artificial neuron is represented in Figure 2.1. In the figure, various inputs to the network are represented by the mathematical symbol,  $x_n$ . Each of these inputs are multiplied by a connection weight. These weights are represented by  $w_n$ . In this case, these products are simply summed ( $z_{in}$ ), fed through an activation function to generate results and then delivered as output  $y_i$ .



**Figure 2.1:** Graphical representation of single  $n$ -input artificial neuron.



Below,  $z_{in}$  refers in all cases to the weighted sum of all the inputs to the neuron, i.e. for  $n$  inputs, where  $w$  is a vector of synaptic weights and  $x$  is a vector of inputs.

$$z_{in} = \sum_{i=0}^n x_i w_i = x_0 w_0 + x_1 w_1 + x_2 w_2 + \dots + x_n w_n \quad (2.1)$$

$$y_i = \Phi(z_{in} + b) \quad (2.2)$$

where  $\Phi(\cdot)$  is the activation function,  $b$  is the bias and  $y_i$  is the output signal of the neuron.

An artificial neuron is characterized by:

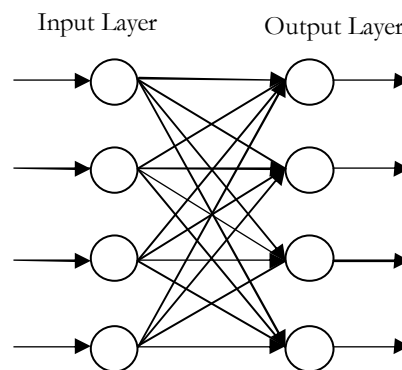
- Architecture (connection between neurons).
- Training or learning (determining weights on the connections).
- Activation function.

### 2.1.1 Architecture of neural networks

The arrangement of neurons into layers and the pattern of connection within and in-between layer are generally called as the architecture of the network. Neural network architectures can be divided usefully into two main categories. They are Feed-forward or Feed-back networks.

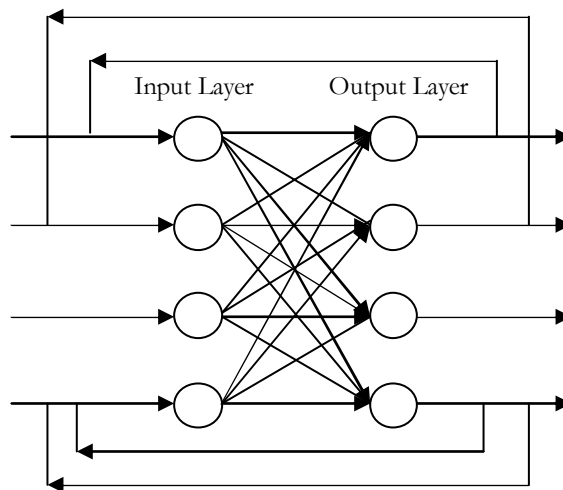
**Feed-forward networks:** Feed-forward ANNs allow signals to travel one way only; from input to output. There is no feedback (loops) i.e. the output of any layer does not affect that same layer. At any particular time,

an input pattern results in an output determined completely by the mapping function of the weights. The power of these networks is in the internal representations formed by the hierarchy of neuron layers. This type of network has the advantage of being unconditionally stable and fast. Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition. Examples of feed-forward networks are Perceptron and Adaline (Haykin 1994). Figure 2.2 shows a single layer feed-forward network. A neural network in which the input layer of source nodes projects into an output layer of neurons but not vice-versa is known as single feed-forward or acyclic network. In single layer network, ‘single layer’ refers to the output layer of computation nodes as shown in Figure 2.2. In multilayer feed-forward network consists of one or more hidden layers, whose computation nodes are called hidden neurons or hidden units.



**Figure 2.2:** A single layer feed-forward network.

**Feed-back networks:** Feed-back networks can have signals traveling in both directions by introducing loops in the network. Feed-back networks are very powerful and can get extremely complicated. Feed-back networks are dynamic; their 'state' is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Information storage capacity is improved in comparison to feed-forward networks. Figure 2.3 shows the architecture of feed-back network. Feed-back architectures (Figure 2.3) are also referred to as interactive or recurrent, although the latter term is often used to denote feed-back connections in single-layer organizations. Recurrent networks are well suited to associative memory, optimization, or retrieval type tasks, as exemplified by the Hopfield network (Hopfield 1982).



**Figure 2.3:** An architecture of feed-back network.

### 2.1.2 Types of learning in neural networks

A neural network has to be configured such that the application of a set of inputs produces (either 'direct' or via a relaxation process) the desired set of outputs. Various methods to set the strengths of the connections exist. One way is to set the weights explicitly, using a priori knowledge. Another way is to 'train' the neural network by feeding it teaching patterns and letting it change its weights according to some learning rule. The learning methods in neural networks can be classified into three categories. These are supervised learning, unsupervised learning and reinforcement learning.

**Supervised learning:** In supervised learning, both inputs and outputs are provided. The network then processes inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights, which control the network. This process occurs over and over as the weights are continually tweaked. The set of data, which enables the training, is called the "training set." During the training of a network, the same set of data is processed many times, as the connection weights are ever refined. The term supervised originates from the fact that the desired signals on individual output nodes are provided by an external teacher. Examples of supervised learning are the delta rule and the perceptron rule.

**Unsupervised learning:** In this type, the network is provided with inputs but not with desired outputs. The system itself must then decide what features it will use to group the input data. This is often referred to as self-organization or adaption. These networks use no external influences to adjust their weights. Instead, they internally monitor their performance. These networks look for regularities or trends in the input signals, and makes adaptations according to the function of the network. Even without being told whether it's right or wrong, the network still must have some information about how to organize itself. This information is built into the network topology and learning rules. An unsupervised learning algorithm might emphasize cooperation among clusters of processing elements. In such a scheme, the clusters would work together. If some external input activated any node in the cluster, the cluster's activity as a whole could be increased. Likewise, if external input to nodes in the cluster was decreased, that could have an inhibitory effect on the entire cluster. In this type of learning external teacher is not present. Typical examples are the Hebbian learning rule and the competitive learning rule (Fausett 1994).

**Reinforcement learning:** This kind of learning is based upon both supervised and unsupervised learning. Here the learning machine does some action on the environment and gets a feedback response from the environment. The learning system grades its action good (rewarding) or bad (punishable) based on the environmental response and accordingly adjusts its parameters. Generally, parameter adjustment is continued until

an equilibrium state occurs, following which there will be no more changes in its parameters. Reinforcement learning is also called learning with a critic as opposed to learning with a teacher (Sutton and Barto 1998).

### 2.1.3 Activation functions

An activation function  $\Phi(v)$  performs a mathematical operation on the outputs of neuron  $Z_{in}$ . It acts as a squashing function, such that the output of a neuron in a neural network is between certain values (usually 0 and 1, or -1 and 1). The activation functions are selected according to the types of problem to be solved by the network. In general, there are three types of activation functions, denoted by  $\Phi(v)$ .

First, there is the Threshold function which takes on a value of 0 if the summed input is less than a certain threshold value ( $v$ ), and the value 1 if the summed input is greater than or equal to the threshold value.

$$\Phi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases} \quad (2.3)$$

Secondly, there is the Piecewise-Linear function. This function again can take on the values of 0 or 1, but can also take on values between that depending on the amplification factor in a certain region of linear operation.

$$\Phi(v) = \begin{cases} 1 & v \geq \frac{1}{2} \\ v & -\frac{1}{2} > v > \frac{1}{2} \\ 0 & v \leq -\frac{1}{2} \end{cases} \quad (2.4)$$

Thirdly, there is the Sigmoid function. This function can range between 0 and 1, but it is also sometimes useful to use the -1 to 1 range. This non-linear function is the most common type of the activation used to construct the neural networks. It is mathematically well behaved, differentiable and strictly increasing function. An example of the sigmoid function is the hyperbolic tangent function.

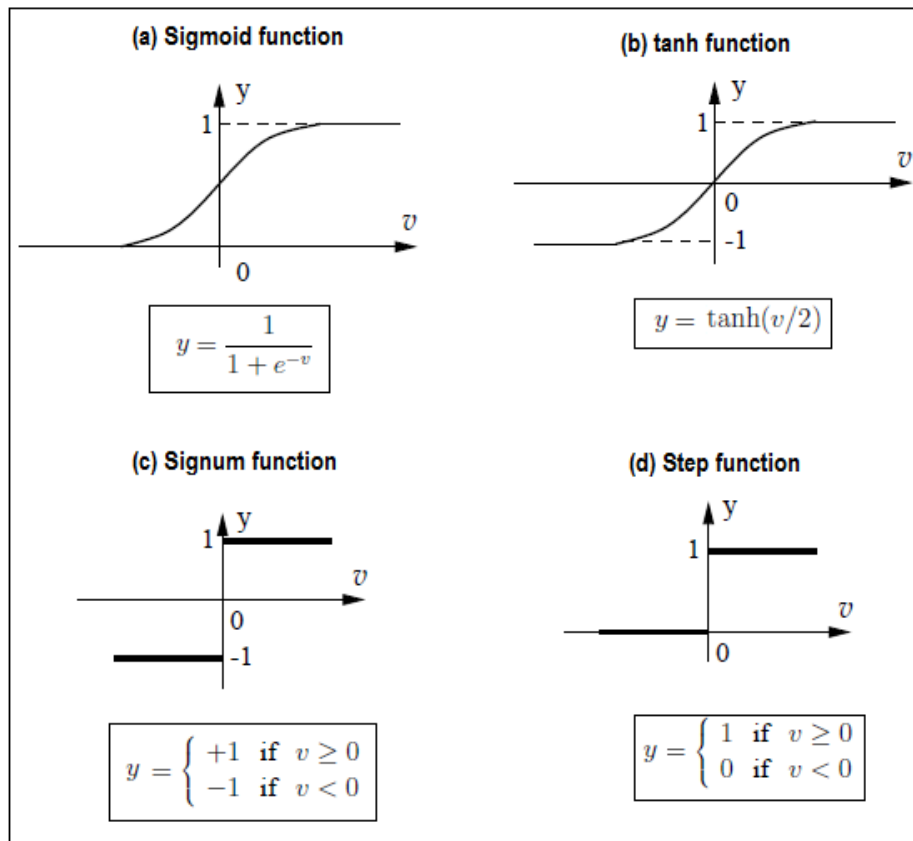
$$\Phi(v) = \tanh\left(\frac{v}{2}\right) = \frac{1 - \exp(-v)}{1 + \exp(-v)} \quad (2.5)$$

Figure 2.4 shows common non-linear function used for synaptic inhibition.

#### **2.1.4 Multilayer perceptron network**

When used without qualification, the ANN usually refers to a Multilayer Perceptron Networks (MLPNs). However, there are many other types of neural networks including Probabilistic Neural Networks, General Regression Neural Networks, Radial Basis Function Networks, Cascade Correlation, Functional Link Networks, Kohonen Networks, Gram-Charlier Networks, Learning Vector Quantization, Hebb Networks, Adaline Networks, Hetero-associative Networks, Recurrent Networks and Hybrid Networks. MLPNs constitute an important class of feed-forward ANNs, developed to replicate learning and generalization abilities of humans with an attempt to model the functions of biological neural networks. Most of the literature in the field is referred to this neural

network. This type of neural network is known as a supervised network because it requires a desired output in order to learn. The goal of this type



**Figure 2.4:** Common non-linear function used for synaptic inhibition. Soft non-linearity: (a) sigmoid and (b) tanh; Hard non-linearity: (c) signum and (d) step.

of network is to create a model that correctly maps the input to the output using historical data so that the model can then be used to produce the output when the desired output is unknown (Zurada 1992). The MLPN is



a nonparametric technique for performing a wide variety of detection and estimation tasks (Haykin 1994).

In this section, the learning problem in the MLPN is formulated. The multilayer perceptron is characterized by a neuron model, network architecture, associated objective functional and learning algorithm. These four concepts are briefly described as follows:

**Neuron model:** A neuron model is a mathematical model of the behavior of a single neuron in a biological nervous system. The characteristic neuron model in the multilayer perceptron is the so called perceptron. The perceptron neuron model receives information in the form of a set of numerical input signals. This information is then integrated with a set of free parameters to produce a message in the form of a single numerical output signal.

**Network architecture:** In the same way a biological nervous system is composed of interconnected biological neurons, an ANN is built up by organizing artificial neurons in network architecture. In this way, the architecture of a network refers to the number of neurons, their arrangement and connectivity. The characteristic network architecture in the multilayer perceptron is the so called feed-forward architecture.

**Objective functional:** The objective functional plays an important role in the use of a neural network. It defines the task the neural network is required to do and provides a measure of the quality of the representation

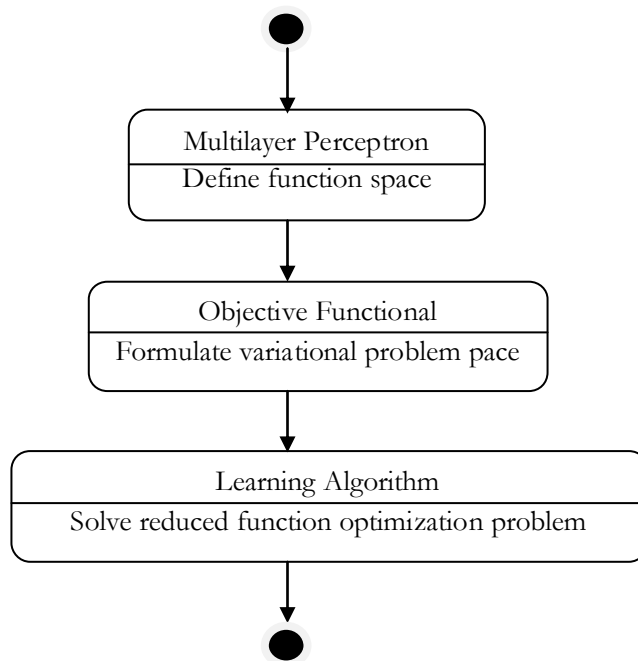
that the network is required to learn. The choice of a suitable objective functional depends on the particular application.

**Learning algorithm:** The procedure used to carry out the learning process is called training algorithm, or learning algorithm. The learning algorithm is applied to the network in order to obtain a desired performance. The type of learning is determined by the way in which the adjustment of the free parameters in the neural network takes place.

Figure 2.5 depicts an activity diagram for the learning problem in the multilayer perceptron. The solving approach here consists of three steps. The first step is to choose a suitable parameterized function space in which the solution to the problem is to be approximated. The elements of this family of functions are those spanned by a multilayer perceptron. In the second step the variational problem is formulated by selecting an appropriate objective functional, defined on the function space chosen before. The third step is to solve the reduced function optimization problem. This is performed with a learning algorithm capable of finding an optimal set of parameters (Lopez 2008).

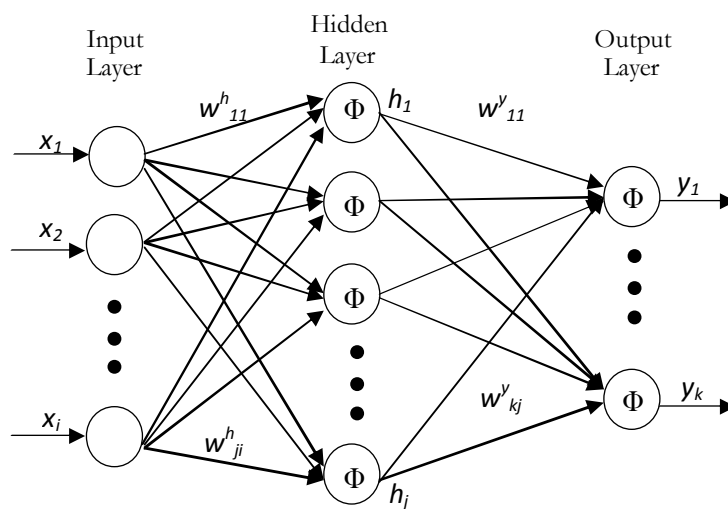
The architecture of a neural network refers to the number of neurons, their arrangement and connectivity. Any network architecture can be symbolized as a directed and labeled graph, where nodes represent neurons and edges represent connectivities among neurons. An edge label represents the free parameter of the neuron for which the flow goes in

(Chen and Haykin 2002). Most neural networks, even biological neural networks, exhibit a layered structure. In this work layers are the basis to determine the architecture of a neural network. Thus, a neural network typically consists on a set of sensorial nodes which constitute the input layer, one or more hidden layers of neurons and a set of neurons which constitute the output layer. As it was said above, the characteristic neuron model of the multilayer perceptron is the perceptron. On the other hand, the multilayer perceptron has feed-forward network architecture (Lopez 2008).



**Figure 2.5:** Activity diagram for the learning problem in the multilayer perceptron.

Figure 2.6 illustrates a MLPN with three layers. This network has an input layer (on the left), one hidden layer (in the middle) and an output layer (on the right).  $x_i$  stands for the input variables.  $y_k$  represents output variables.  $w$  stands for weight and  $\Phi$  stands for activation function.



**Figure 2.6:** The MLPN architecture with three layers.

**Input layer:** A vector of predictor variable values ( $x_1 \dots x_i$ ) is presented to the input layer. The input layer (or processing before the input layer) standardizes these values so that the range of each variable is -1 to 1. The input layer distributes the values to each of the neurons in the hidden layer. In addition to the predictor variables, there is a constant input of 1.0, called the *bias* that is fed to each of the hidden layers; the bias is multiplied by a weight and added to the sum going into the neuron.

**Hidden layer:** Arriving at a neuron in the hidden layer, the value from each input neuron is multiplied by a weight ( $w_{ji}$ ), and the resulting weighted values are added together producing a combined value  $u_j$ . The weighted sum ( $u_j$ ) is fed into an activation function,  $\Phi$ , which outputs a value  $h_j$ . The outputs from the hidden layer are distributed to the output layer.

**Output layer:** Arriving at a neuron in the output layer, the value from each hidden layer neuron is multiplied by a weight ( $w_{kj}$ ), and the resulting weighted values are added together producing a combined value  $z_j$ . The weighted sum ( $z_j$ ) is fed into an activation function,  $\Phi$ , which outputs a value  $y_k$ . The  $y$  values are the outputs of the network.

### 2.1.5 The learning algorithms

Learning algorithms are an integral part of ANN model development. The best-known example of ANN learning algorithm is back propagation (BP) (Patterson 1996; Haykin 1994; Fausett 1994). In BP, the gradient vector of the error surface is calculated. This vector points along the line of steepest descent from the current point, so we know that if we move along it a "short" distance, we will decrease the error. A sequence of such moves (slowing as we near the bottom) will eventually find a minimum of some sort. The difficult part is to decide how large the steps should be. Large steps may converge more quickly, but may also overstep the solution or (if the error surface is very

eccentric) go off in the wrong direction. A classic example of this in neural network training is where the algorithm progresses very slowly along a steep, narrow, valley, bouncing from one side across to the other. In contrast, very small steps may go in the correct direction, but they also require a large number of iterations. In practice, the step size is proportional to the slope (so that the algorithm settles down in a minimum) and to a special constant: the learning rate. The correct setting for the learning rate is application-dependent, and is typically chosen by experiment; it may also be time-varying, getting smaller as the algorithm progresses (Carling 1992).

The algorithm is also usually modified by inclusion of a momentum term: this encourages movement in a fixed direction, so that if several steps are taken in the same direction, the algorithm "picks up speed", which gives it the ability to (sometimes) escape local minimum, and also to move rapidly over flat spots and plateaus. The algorithm therefore progresses iteratively, through a number of epochs. On each epoch, the training cases are each submitted in turn to the network, and target and actual outputs compared and the error calculated. This error, together with the error surface gradient, is used to adjust the weights, and then the process repeats. The initial network configuration is random and training stops when a given number of epochs elapses, or when the error reaches an acceptable level, or when the error stops improving.

More sophisticated techniques for non-linear function optimization have been in use for some time. These methods include Conjugate Gradient (CG), Quasi-Newton (QN), and Levenberg-Marquardt (LM), which are very successful forms of two types of algorithm: line search and model-trust region approaches. They are collectively known as second order learning algorithms. A line search algorithm works as follows: pick a sensible direction to move in the multi-dimensional landscape. Then project a line in that direction, locate the minimum along that line (it is relatively trivial to locate a minimum along a line, by using some form of bisection algorithm), and repeat. In this context, the direction of steepest descent is an obvious choice for a sensible direction (the same direction that would be chosen by back propagation). Actually, this intuitively obvious choice proves to be rather poor. Having minimized along one direction, the next line of steepest descent may spoil the minimization along the initial direction (even on a simple surface like a parabola a large number of line searches may be necessary). A better approach is to select conjugate or non-interfering directions - hence CG (Bishop 1995).

The idea here is that, once the algorithm has minimized along a particular direction, the second derivative along that direction should be kept at zero. Conjugate directions are selected to maintain this zero second derivative on the assumption that the surface is parabolic (speaking roughly, a nice smooth surface). If this condition

holds,  $N$  epochs are sufficient to reach a minimum. In reality, on a complex error surface the conjugacy deteriorates, but the algorithm still typically requires far less epochs than back propagation, and also converges to a better minimum (to settle down thoroughly, back propagation must be run with an extremely low learning rate). QN training is based on the observation that the direction pointing directly towards the minimum on a quadratic surface is the so-called Newton direction. This is very expensive to calculate analytically, but QN iteratively builds up a good approximation to it. QN is usually a little faster than CG, but has substantially larger memory requirements and is occasionally numerically unstable.

A model-trust region approach works as follows: instead of following a search direction, assume that the surface is a simple shape such that the minimum can be located (and jumped to) directly - if the assumption is true. Try the model out and see how good the suggested point is. The model typically assumes that the surface is a nice well-behaved shape (e.g. a parabola), which will be true if sufficiently close to a minima. Elsewhere, the assumption may be grossly violated, and the model could choose wildly inappropriate points to move to. The model can only be trusted within a region of the current point, and the size of this region isn't known. Therefore, choose new points to test as a compromise between that suggested by the model and that suggested by a standard gradient-descent jump. If the new point is good, move to it, and



strengthen the role of the model in selecting a new point; if it is bad, don't move, and strengthen the role of the gradient descent step in selecting a new point (and make the step smaller). LM uses a model that assumes that the underlying function is locally linear (and therefore has a parabolic error surface).

LM is typically the fastest of the learning algorithms, although unfortunately it has some important limitations, specifically: it can only be used on single output networks, can only be used with the sum squared error function, and has memory requirements proportional to  $w_2$  (where  $w$  is the number of weights in the network; this makes it impractical for reasonably big networks). CG is nearly as good, and doesn't suffer from these restrictions. BP can still be useful, not least in providing a quick (if not overwhelmingly accurate) solution. It is also a good choice if the data set is very large, and contains a great deal of redundant data. BP's case-by-case error adjustment means that data redundancy does it no harm (for example, if you double the data set size by replicating every case, each epoch will take twice as long, but have the same effect as two of the old epochs, so there is no loss). In contrast, LM, QN, and CG all perform calculations using the entire data set, so increasing the number of cases can significantly slow each epoch, but does not necessarily improve performance on that epoch (not if data is redundant; if data is sparse, then adding data will make each epoch better). BP can also be equally good if

the data set is very small, for there is then insufficient information to make a highly fine-tuned solution appropriate (Hill and Lewicki 2007).

### **2.1.6 Applications of neural networks**

The utility of ANN models lies in the fact that they can be used to infer a function from observations. This is particularly useful in applications where the complexity of the data or task makes the design of such a function by hand impractical. Neural networks have been successfully applied to broad spectrum of data-intensive applications (Nielsen 2001).

**Language processing:** These applications include text-to-speech conversion, auditory input for machines, automatic language translation, secure voice keyed locks, automatic transcription, aids for the deaf, aids for the physically disabled which respond to voice commands and natural language processing.

**Character recognition:** Neural network based products are available which can recognize hand printed characters through a scanner. It is 98% accurate for numbers, a little less for alphabetical characters. Quantum Neural Network software package (Qnspec) is available for recognizing characters, including cursive.

**Image (data) compression:** Neural networks can do real-time compression and decompression of data. These networks can reduce eight

bits of data to three and then reverse that process upon restructuring to eight bits again.

**Pattern recognition:** Many pattern recognition applications are in use like, a system that can detect bombs in luggage at airports by identifying from small variances and patterns from within specialized sensor's outputs, a BP neural network which can discriminate between a true and a false heart attack, a network which can scan and also read the PAP smears etc. Many automated quality control applications are now in use, which are based on pattern recognition.

**Signal processing:** Neural networks have proven capable of filtering out electronic noise. Another application is a system that can detect engine misfire simply from the engine sound.

**Financial:** Banks, credit card companies and lending institutions deal with many decisions that are not clear-cut. They involve learning and statistical trends. Neural networks are now trained on the data from past decisions and being used in decision-making.

**Servo control:** A neural system known as Martingale's Parametric Avalanche - a spatio-temporal pattern recognition network is being designed to control the shuttle during in-flight maneuvers. Another application is ALVINN, for Autonomous Land Vehicle.

**Weather prediction:** Weather forecasting for the future is one of the most important and ever-challenging areas of investigation for scientists

due to many sectors are largely dependent on the weather conditions. The applicability of ANN technology to improve weather forecasting is the recent development.

## **2.2 Numerical Weather Prediction Modeling**

The atmosphere is a hydro-thermodynamic system where the motions obey the laws of physics (Holton 2004). Numerical models are designed to solve the fundamental equations that govern these motions in the atmosphere. These equations are derived from the Newtonian mechanics and thermodynamic laws, especially the conservation laws for mass, energy and momentum. These sets of equations are known as the primitive (so called since they are derived from conservation principles) or the fundamental equations, and are equations of momentum, mass continuity, thermodynamics and moisture. The equations of motion are highly non-linear partial-differential equations and complex in nature. So far there has been no success in solving the full governing equations analytically. Lewis Fry Richardson in 1922 conducted the first experimental numerical weather forecasting by solving the equations using a mechanical desk calculator. It took him six weeks to do a 6 hour forecast. Later, John Von Neumann and a group of scientists used the first digital computer for weather forecasting (Wallace and Hobbs 2006). However, with the invention of high speed computers today, it is possible to approximate these equations in their non-linear form with an

exceptional degree of accuracy. The equations describing fluid motion are generally known as Navier-Stokes equations.

### **2.2.1 Governing equations**

A complete set of equations that govern the evolution of the atmosphere can be described as (Kalnay 2003):

- Newton's second law of conservation of momentum (three equations for the three components of velocity);
- Conservation of mass or the continuity equation;
- Conservation of energy or 1st law of thermodynamics;
- The equation of state for gas;
- Conservation equation for water mass.

The equations for horizontal and vertical motion are derived from Newton's second law or the law of conservation of momentum (Washington and Parkinson 2005). In the atmosphere, the major forces involved in motion are the force of gravity, the pressure gradient force and the Coriolis force. The pressure gradient force and the Coriolis force account for the major forces in the horizontal direction. Whereas in the vertical, the two main forces responsible are the force of gravity and the pressure gradient force, due to the variation of pressure with height. When the vertical pressure gradient is in balance with the gravitational force the motion is considered to be in hydrostatic equilibrium. In the early days, the models were designed to assume hydrostatic equilibrium due to

limitations in computing power. However, recently developed numerical codes provide the non-hydrostatic option, so that higher resolutions of the order of a few meters can be used to resolve the small scale circulations such as cumulus convection and boundary layer circulations (AMS 2000; Kalnay 2003).

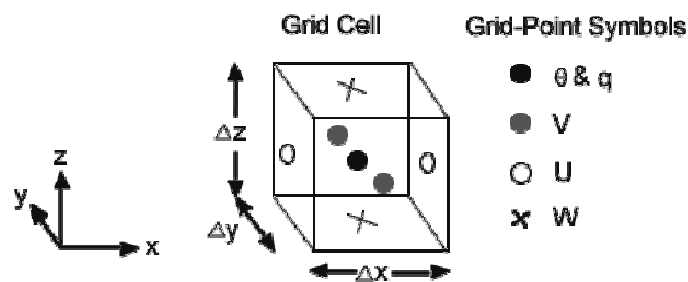
The conservation of mass or the continuity equation ensures that the mass of air parcels remain the same through time. That is, the rate at which mass enters a system is equal to the rate at which mass leaves the system, for constant density/pressure. The temperature of a parcel in the atmosphere could be modified by mixing with warmer or colder air or due to expansion or contraction of the parcel. Conservation of energy or the first law of thermodynamics is used so that if heat is applied to a parcel, this heat can be used to increase the internal energy and/or to produce work of expansion. The atmosphere is assumed to be a perfect gas (Kalnay 2003). The equation of state is used to predict the state of gases or liquids and is another form of thermodynamic equation relating the three thermodynamic variables of pressure, density and temperature. It is applied to the atmosphere, relating the change in temperature of a parcel of air to energy transfer between the parcel and its environment and work done by or on the parcel. The equation for conservation of water mass indicates that the total amount of water vapour in a parcel is conserved as a function of advection, evaporation or condensation.

The governing equations contain unknown variables ( $u$ ,  $v$ ,  $w$ ,  $\rho$ ,  $p$  and  $T$ ), where  $u$  is the zonal,  $v$  is the meridional, and  $w$  is the vertical component of wind,  $\rho$  is the density,  $p$  is the pressure and  $T$  is the temperature, but as a solvable system (Washington and Parkinson 2005). Since the governing equations are higher-order non-linear partial-differential equations, no analytical solution has been obtained. To get a possible solution, alternative techniques are used (Stull 2000). One method is to find an exact analytical solution by highly simplifying the governing equations. Highly simplified forms of these equations can be used to understand many of the atmospheric motions, such as the geostrophic wind, gradient wind and surface winds around high and low pressure centers and atmospheric waves (Jacobson 2005). Another method of solving the equations is by using finite difference approximations and this method is implemented in the modern day NWP models. Numerical models solve the governing equations by discretizing them using various numerical schemes. The most commonly used numerical schemes in mesoscale models are the interpolation schemes and the finite difference schemes (Pielke 2002). An interpolation scheme uses polynomials to approximate the dependent variables in one or more spatial directions. Finite difference schemes approximate the differential terms in the governing equations using a truncated form of the Taylor series expansion and writes the equations as a form of difference equations. The latter is more widely used, due to its conceptual simplicity

and ease of computational programming (Pielke 2002). The variables of the governing equations are defined on a grid and integrated in time using the finite difference schemes to arrive at a forecast.

### 2.2.2 Grid structure

To solve these equations at every point in the atmosphere would incur an extensive amount of computer time. Therefore the equations are solved on a finite number of regularly spaced points known as a grid. Figure 2.7 shows the arrangement of the variables within a grid cell or a grid volume. Variables are arranged in the three Cartesian directions  $\Delta x$ ,  $\Delta y$  and  $\Delta z$ . The resolution of the models is determined by the dimension of these grid cells, which would be set according to the phenomena of interest to be simulated. One common way of arranging the variables in the grid is to represent the thermodynamic variables such as potential temperature, specific humidity, liquid water content, etc. in the center of the grid cell. Velocity components  $u$ ,  $v$  and  $w$  are placed at the boundaries of the grid cells, usually at  $1/2\Delta x$ ,  $1/2\Delta y$  and  $1/2\Delta z$ .



**Figure 2.7:** Arrangement of the variables in a staggered grid cell (Stull 2000).



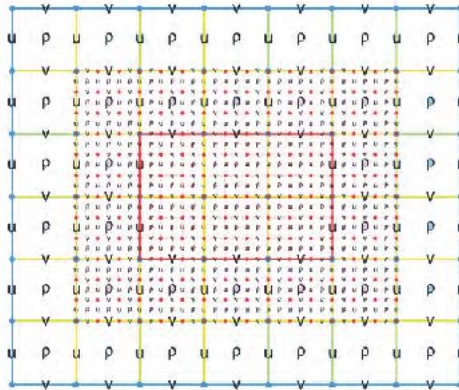
There are several ways in which the variables are arranged on a grid and detailed discussions of the properties of the various arrangements are discussed in Arakawa and Lamb (1977) and Messinger and Arakawa (1976). The arrangement shown in Figure 2.8 is known as a staggered grid and one of the most commonly used of this type (the Arakawa C grid). The type of grid used is related to the computational stability of the numerical schemes used and has its own pros and cons. Moreover, staggering introduces considerable complexity in, for example, diagnostic studies and in post-processing of the graphical outputs.

### **2.2.3 Vertical levels**

Several methods are used to represent the vertical coordinates in a model. Vertical coordinates are converted to pressure coordinates (Eliassen 1949) and are chosen to represent the large scale motions due to their hydrostatic nature. This coordinate system is widely used since it greatly simplifies the governing equations, and due to its easiness in relating the quantities from observation such as radiosonde that provide the altitude of observations in pressure values (Satoh 2004). However, using pressure coordinates comes with its drawbacks as they do not represent the presence of complex orography very well. To overcome this problem, ‘normalized pressure’ or sigma coordinates were introduced by Phillips (1957). This is the most widely used vertical coordinate system in numerical models and is sometimes referred to as the terrain-following coordinate system. The vertical coordinate,  $\sigma$ , is defined as:

$$\sigma = P/P_s \quad (2.6)$$

where  $P$  is the atmospheric pressure at the point in question and  $P_s$  is the surface pressure below the point in question. Usually, more levels are defined near the surface in order to better resolve the processes in the boundary layer.



**Figure 2.8:** Horizontal cross section of a nested grid structure. Density fields are placed at the center and velocity fields are on the edges of each grid square (Wiki 2008).

#### 2.2.4 Spatial boundary conditions

In contrast to global models, limited area or mesoscale models have their grids artificially closed on the sides of the domain. Therefore, it becomes necessary to define the dependent variables at the perimeter of the domain. These defined values are known as the boundary conditions. Each boundary top, bottom and the lateral boundaries are treated differently in mesoscale models (Pielke 2002). The main idea behind

using a boundary condition is to filter out or damp the disturbances such as internal gravity waves from being reflected back into the simulation domain and modifying the solutions of interest. The bottom boundary of a model is where the ‘real’ boundary is defined as it is where the surface conditions are characterized. Transfer of physical properties such as heat and moisture across the bottom boundary plays a fundamental role in the development of meteorological circulations within the model.

At the top of the model, one of the techniques used to damp disturbances is to use the top as a rigid lid. The vertical velocity is set to zero at the top level and pressure is adjusted to account for mesoscale perturbations at that level (Pielke 2002). Another commonly used technique introduced by Klemp and Lilly (1978) is to use a damping or an absorbing layer at the model top, where disturbances are effectively removed. In mesoscale models, it is always a practice to keep the lateral boundaries far away from the region of interest. In general, two types of lateral boundary condition are utilized (Pielke 2002): *open lateral boundary* – where mesoscale perturbations are allowed to pass in and out of the domain, and *closed lateral boundary* – where perturbations are not allowed to enter or exit.

### **2.2.5 Parameterizations**

Parameterization is a method of approximating an unknown term by using one or more known terms or factors (Stull 2000). Some of the

physical processes in the atmosphere are well understood but too complex or too unwieldy to formulate in a model. Other physical processes are not sufficiently well understood to formulate physical laws and some of these important processes are not explicitly resolved by numerical models. These non-explicitly resolved processes are known as sub-grid-scale processes. In order to represent these processes in numerical models, these processes are parameterized. An example of one of the important sub-grid-scale processes is the turbulent mixing in the boundary layer. Surface heating leads to formation of turbulent eddies which have a scale of a few meters to about a hundred meters, which is an important process in heat and moisture transfer, and crucial to the development of thunderstorms. However, due to their small scale, they might not be resolved by the models with a horizontal grid size of the order of a few kilometers. Instead of calculating the motion and heat transport by each and every eddy, the net vertical heat flux transport by the eddies is parameterized to be represented over the large scale domains. This is also known as turbulence parameterization. Moreover, several physical processes such as cloud microphysics, radiation, surface properties and vegetation effects, which are not explicitly resolved, are parameterized in the models. Main parameterization schemes used for NWP models are:

- Microphysics: Bulk schemes ranging from simplified physics suitable for mesoscale modeling to sophisticated mixed-phase physics suitable for cloud-resolving modeling.

- Convective parameterizations: Adjustment and mass-flux schemes for mesoscale models.
- Surface physics: Multi-layer land surface models ranging from a simple thermal model to full vegetation and soil moisture models, including snow cover and sea ice.
- Planetary boundary layer physics: Turbulent kinetic energy prediction or non-local K schemes.
- Atmospheric radiation physics: Longwave and shortwave schemes with multiple spectral bands and a simple shortwave scheme. Cloud effects and surface fluxes are included.

### **2.2.6 Space scale**

Meteorological phenomena occur over a wide range of space and time scales. Phenomena having short time scales also tend to have small spatial scales, and vice versa. Curiously, the ratio of (horizontal) space to time scales, which has units of velocity ( $\text{ms}^{-1}$ ), is roughly the same order of magnitude for all features ( $\sim 10 \text{ ms}^{-1}$ ). Orlanski (1975) proposed a set of scales that include the micro, meso and macro scales. Table 2.1 shows these three definitions, which have gained wide acceptance, despite an even newer proposal by Fujita (1981).

A meteorological numerical model is a simplified abstraction of the real atmosphere, which is valid for a certain length and timescale. The model is given by a set of equations and the corresponding numerical

solvers. Within the model, a scale dependent discretization of the atmosphere in space and time is necessary. Temporal and spatial resolutions of a mesoscale model are better than in a macroscale model but coarser than in a microscale model. Mesoscale models are used for purposes ranging from weather forecasting, to air-quality regulatory applications, and to basic research.

**Table 2.1:** Scale definitions (Orlanski 1975).

Name	Scale Range	Meteorological phenomena
Microscale- $\gamma$	< 20 m	Turbulence, plumes, roughness
Microscale- $\beta$	20-200 m	Dust devils, thermals, wakes
Microscale- $\alpha$	200-2000 m	Tornadoes, short gravity waves
Mesoscale- $\gamma$	2-20 km	Thunderstorm convection, complex terrain flows, urban effects
Mesoscale- $\beta$	20-200 km	Nocturnal low-level jets, cloud clusters, sea breezes
Mesoscale- $\alpha$	200-2000 km	Fronts, low-pressure systems, hurricanes
Macroscale- $\beta$	2000-20000 km	Baroclinic waves
Macroscale- $\alpha$	> 20000 km	Tidal waves

Over the last decade, sophisticated mesoscale models like MM5 (Fifth-Generation Penn State/NCAR Mesoscale Model), RAMS (Regional Atmospheric Modeling System), ARPS (Advanced Regional Prediction System), ARW, NMM, etc. were developed. The physical complexity of these models allows today's most accurate simulations. However, its use needs expensive computational resources as well as years of professional experience. To acquire spatially distributed

information in two or three dimensions, a model is often the only possibility. In the mesoscale, the objects of interest vary on small distances, thus requiring spatially highly resolved information. Mesoscale models also help understanding processes by allowing full control over environmental parameters. Hence it is possible to determine the steering factors of a phenomenon and also to test sensitivity against changes in environmental conditions. Scientific goals of mesoscale modeling include accurate numerical simulations of mesoscale processes to understand the role of synoptic scale parameters for generation and evolution of mesoscale phenomena, to find the limits of predictability by means of sensitivity studies, and to understand interactions of the mesoscale with smaller and larger scales.

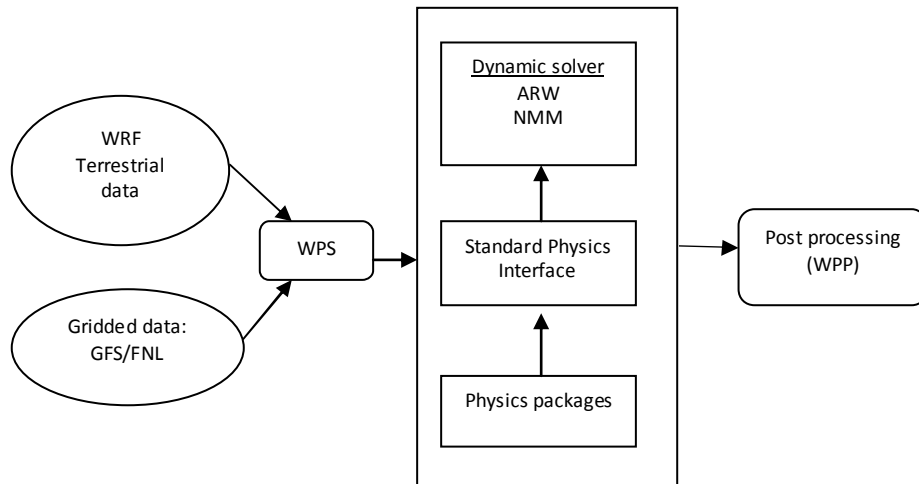
### **2.3 WRF Modeling System**

The development of the WRF modeling system is a multi-agency effort intended to provide a next-generation mesoscale forecast model and data assimilation system to advance both the understanding and prediction of mesoscale weather and accelerate the transfer of research advances into operations. The WRF modeling system was developed as a collaborative effort among the US organizations, the National Centre for Atmospheric Research (NCAR) Mesoscale and Microscale Meteorology (MMM) Division, the National Oceanic and Atmospheric Administration's (NOAA), National Centers for Environmental Prediction (NCEP) and Forecast System Laboratory (FSL), the Department of Defense's Air

Force Weather Agency (AFWA) and Naval Research Laboratory (NRL), the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma, and the Federal Aviation Administration (FAA) along with the participation of a number of university scientists. The model is designed to be a flexible, state-of-the-art, portable code that is efficient in a massively parallel computing environment. A modular single-source code is maintained that can be configured for both research and operations. It offers numerous physics options, thus tapping into the experience of the broad modeling community. Advanced data assimilation systems are being developed and tested in tandem with the model. WRF is maintained and supported as a community model to facilitate wide use, particularly for research and teaching, in the university community. It is suitable for use in a broad spectrum of applications across scales ranging from meters to thousands of kilometers. Such applications include research and operational NWP, data assimilation and parameterized-physics research, downscaling climate simulations, driving air quality models, atmosphere-ocean coupling, and idealized simulations (e.g. boundary-layer eddies, convection, baroclinic waves). The WRF modeling system infrastructure is given in Figure 2.9. The WRF system consists of these major components:

- WRF Preprocessing System (WPS)
- Dynamic solver
- WRF Postprocessor (WPP) and graphics tools





**Figure 2.9:** WRF modeling system infrastructures.

There are two dynamics solvers in the WRF for its computation of the atmospheric governing equations, and the variants of the model are known as: the ARW solver (originally referred to as the Eulerian mass or “em” solver) developed primarily at NCAR and the NMM solver developed at NCEP.

### 2.3.1 ARW model

The ARW system consists of the ARW dynamics solver together with other components of the WRF system needed to produce a simulation. Thus, it also encompasses physics schemes, initialization routines, post processing and a data assimilation package. ARW solver main characteristics are:

- Equations: Fully compressible, Euler non-hydrostatic with a run-time hydrostatic option available. Conservative for scalar variables.
- Prognostic Variables: Velocity components  $u$  and  $v$  in Cartesian coordinate, vertical velocity  $w$ , perturbation potential temperature, perturbation geopotential, and perturbation surface pressure of dry air. Optionally, turbulent kinetic energy and any number of scalars such as water vapour mixing ratio, rain/snow mixing ratio, and cloud water/ice mixing ratio.
- Vertical Coordinate: Terrain-following hydrostatic-pressure, with vertical grid stretching permitted. Top of the model is a constant pressure surface.
- Horizontal Grid: Arakawa C-grid staggering.
- Time Integration: Time-split integration using a 3rd order Runge-Kutta scheme with smaller time step for acoustic and gravity-wave models.
- Spatial Discretization: 2nd to 6th order advection options in horizontal and vertical.
- Turbulent Mixing and Model Filters: Sub-grid scale turbulence formulation in both coordinate and physical space. Divergence damping, external-mode filtering, vertically implicit acoustic step off-centering. Explicit filter option also available.

- Initial Conditions: Three dimensional for real-data, and one-, two- and three-dimensional using idealized data. A number of test cases are provided.
- Lateral Boundary Conditions: Periodic, open, symmetric, and specified options available.
- Top Boundary Conditions: Gravity wave absorbing (diffusion or Rayleigh damping).  $w = 0$  top boundary condition at constant pressure level.
- Bottom Boundary Conditions: Physical or free-slip.
- Earth's Rotation: Full Coriolis terms included.
- Mapping to Sphere: Three map projections are supported for real-data simulation: polar stereographic, Lambert-conformal, and Mercator. Curvature terms included.
- Nesting: One-way, two-way, and moving nests.

### **2.3.2 NMM model**

The NMM is designed to be a flexible, state-of-the-art atmospheric simulation system that is portable and efficient on available parallel computing platforms. The model is suitable for use in a broad range of applications across scales ranging from meters to thousands of kilometers. It consists of the NMM dynamics solver together with other components of the WRF system needed to produce a simulation. Thus, it also encompasses physics schemes, initialization routines and post processing.

Model physics parameterizations are quite similar in both dynamic solvers. The key features of NMM are:

- Fully compressible, non-hydrostatic model with a hydrostatic option.
- Hybrid (sigma-pressure) vertical coordinate.
- Arakawa E-grid.
- Forward-backward scheme for horizontally propagating fast waves, implicit scheme for vertically propagating sound waves, Adams-Bashforth Scheme for horizontal advection, and Crank-Nicholson scheme for vertical advection. The same time step is used for all terms.
- Conservation of a number of first and second order quantities, including energy and enstrophy.

### 2.3.3 WRF system requirements

The WRF modeling system software installation is fairly straightforward on the ported platforms. The model-component portion of the package is mostly self-contained, meaning that WRF model requires no external libraries (such as for various linear algebra solvers). The WPS package, separate from the WRF source code, has additional external libraries that must be built (in support of GRIB2 processing). The one external package that both of the systems require is the netCDF library, which is one of the supported I/O API packages. The WRF code

has been successfully ported to a number of Unix-based machines. Table 2.2 shows a list of the supported combinations of hardware and software for WRF. The WRF model may be built to run on a single processor machine, a shared-memory machine (that use OpenMP API), a distributed memory machine (with the appropriate MPI libraries), or on a distributed cluster (utilizing both OpenMP and MPI).

**Table 2.2:** A list of the supported combinations of hardware and software for WRF.

Vendor	Hardware	Operating System	Compiler
Cray	X1	UniCOS	vendor
Cray	AMD	Linux	PGI/PathScale
IBM	Power Series	AIX	vendor
SGI	IA64 / Opteron	Linux	Intel
COTS*	IA32	Linux	Intel/ PGI/gfortran/ g95/PathScale
COTS*	IA64 / Opteron	Linux	Intel/ PGI/ gfortran/ PathScale
Mac	Power Series	Darwin	xlf/ g95/ PGI/ Intel
Mac	Intel	Darwin	g95/ PGI/ Intel

The WRF model is written in FORTRAN (what many refer to as FORTRAN 90). The software layer, RSL-LITE, which sits between WRF and the MPI interface, are written in C. Ancillary programs that perform file parsing and file construction, both of which are required for default building of the WRF modeling code, are written in C. Thus, FORTRAN 90/95 and C compilers are required. Additionally, the WRF build mechanism uses several scripting languages: including perl (to handle

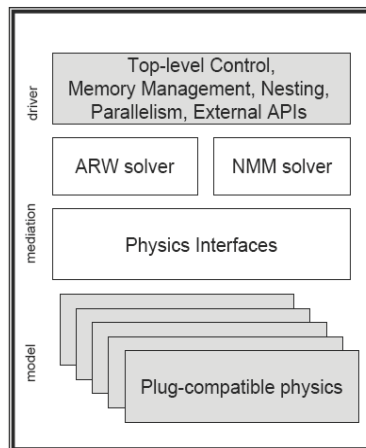
various tasks such as the code browser designed by Brian Fiedler), C-shell and Bourne shell. The traditional UNIX text/file processing utilities are used: make, M4, sed, and awk. If OpenMP compilation is desired, OpenMP libraries are required. The WRF I/O API also supports netCDF, PHD5 and GriB-1 formats, hence one of these libraries needs to be available on the computer used to compile and run WRF.

### 2.3.4 WRF software framework

The WRF Software Framework (WSF) (Figure 2.10) provides the infrastructure that allows efficient use of an array of high performance computing (HPC) systems, architectures which continue to evolve as we move into petascale computing and beyond. The architecture accommodates multiple dynamics solvers, physics packages that plug into the solvers through a standard physics interface and programs for initialization. The WFS is organized functionally as a three-level hierarchy (Figure 2.10) superimposed over the model subroutine call tree. The highest levels of the call tree correspond to the *driver layer* and the lowest levels correspond to the *model layer*. A *mediation layer* provides the interface between the driver and model layers.

The driver is responsible for top-level control of initialization, time-stepping, I/O, instantiating domains, maintaining the nesting hierarchy between domain type instances, and setting up and managing domain decomposition, processor topologies and other aspects of

parallelism. The model layer comprises the subroutines that perform actual model computations. Model subroutines are written to be callable over an arbitrarily shaped piece of the three-dimensional model domain. The mediation layer provides the glue between the model and driver layers. The mediation layer contains information pertaining to both the model layer and the driver layer: model-specific information such as the flow of control to compute a time step on a domain and driver-specific mechanisms such as tiling and communication.



**Figure 2.10:** WRF software framework.

WRF simulations are typically carried out on HPC clusters as they require an effective compute resource that can handle complex and parallel simulations. HPC clusters are scalable performance compute solutions based on industry standard hardware connected by a private system high speed network. The main benefits of clusters are

affordability, flexibility, availability, high performance and scalability. A cluster uses the aggregated power of compute server nodes to form a high performance solution for parallel applications such as the WRF model. When more compute power is needed, it can be simply achieved by adding more server nodes to the cluster. The way HPC clusters are architected (i.e. multi-core, multi-processor based HPC servers with high speed interconnects) has a great influence on the overall application performance and productivity. In order to meet the demand of more powerful HPC servers, more execution cores (e.g. dual, quad-core) are being integrated into each processor and more processors are being tightly connected (e.g. 4, 8, 16 processors connected through HyperTransport™ technology, a packet-based, high-bandwidth, scalable, low latency point-to-point technology that links processors to each other, processors to coprocessors and processors to I/O and peripheral controllers). The cluster interconnect is very critical to deliver efficiency and scalability for the applications as it needs to handle the networking requirements of each CPU core without imposing additional networking overhead. In a multi-core multi-socket HPC server based cluster, the driving factors of performance and scalability for WRF have shifted from the frequency and cache size per core to the memory and interconnect throughput per core. The memory bottleneck can be solved by using interconnects that support Direct Memory Access (DMA), Remote DMA and zero-copy transactions.



### **2.3.5 WRF workflow**

Configuring and running WRF is a very complicated process. The typical workflow entails editing multiple configuration files (namelist files), setting environment variables, and running executables in the correct order. Figure 2.11 illustrates the program components and data flow in WRF model. Running real cases in WRF requires the following steps.

**WPS:** The WPS is a set of three programs (geogrid.exe, ungrib.exe, and metgrid.exe) whose collective role is to prepare input to the real program for real-data simulations. Each of the programs performs one stage of the preparation: geogrid defines model domains and interpolates static geographical data to the grids; ungrib extracts meteorological fields from GRIB formatted files; and metgrid horizontally interpolates the meteorological fields extracted by ungrib to the model grids defined by geogrid. The work of vertically interpolating meteorological fields to WRF eta levels is performed within the real program. The data flow between the programs of the WPS is shown in Figure 2.11. Each of the WPS programs reads parameters from a common namelist file, as shown in the figure. This namelist file has separate namelist records for each of the programs and a shared namelist record, which defines parameters that are used by more than one WPS program. A brief description of each of the three main programs is given below:

**geogrid.exe:** The purpose of geogrid is to define the simulation domains, and interpolate various terrestrial data sets to the model grids. The simulation domains are defined using information specified by the user in the “geogrid” namelist record of the WPS namelist file, namelist.wps. In addition to computing the latitude, longitude, and map scale factors at every grid point, geogrid will interpolate soil categories, land use category, terrain height, annual mean deep soil temperature, monthly vegetation fraction, monthly albedo, maximum snow albedo, and slope category to the model grids by default. Besides interpolating the default terrestrial fields, the geogrid program is general enough to be able to interpolate most continuous and categorical fields to the simulation domains. Output from the geogrid is written in the NetCDF I/O format.

**ungrib.exe:** The ungrib program reads GRIB files, "degrib" the data, and writes the data in a simple format, called the intermediate format. The GRIB files contain time-varying meteorological fields and are typically from another regional or global model, such as NCEP's NAM or GFS models. The ungrib program can read GRIB Edition 1 (GRIB1) and, if compiled with a GRIB Edition 2 option (GRIB2). GRIB files typically contain more fields than are needed to initialize WRF. Both versions of the GRIB format use various codes to identify the variables and levels in the GRIB file. Ungrib uses tables of these codes – called Vtables, for "variable tables" – to define which fields to extract from the GRIB file

and write to the intermediate format. Vtables for common GRIB model output files are provided with the ungrib software.

**metgrid.exe:** The metgrid program horizontally interpolates the intermediate-format meteorological data that are extracted by the ungrib program onto the simulation domains defined by the geogrid program. The interpolated metgrid output can then be ingested by the WRF real program. The range of dates that will be interpolated by metgrid are defined in the “share” namelist record of the WPS namelist file, and date ranges must be specified individually in the namelist for each simulation domain. Since the work of the metgrid program, like that of the ungribprogram, is time-dependent, metgrid is run every time a new simulation is initialized. Output from the metgrid is written in the netCDF I/O.

**WRF:** The WRF code contains an initialization program (real.exe/real\_nmm.exe) and a numerical integration program (wrf.exe). The real.exe/real\_nmm.exe portion of the code generates initial and boundary conditions for the wrf.exe program that are derived from output files provided by the WPS. The real.exe/real\_nmm.exe program performs the following tasks:

- Reads data from the namelist and allocates space.
- Initializes remaining variables.
- Reads input data from the WPS.

- Prepares soil fields for use in the model (usually vertical interpolation to the requested levels).
- Checks to verify soil categories, land use, land mask, soil temperature and sea surface temperature are all consistent with each other.
- Vertically interpolates to the models computational surfaces.
- Generates initial and lateral condition file.

The wrf.exe is the a numerical integration program that reads the initial conditions and the lateral boundary conditions set up for the forecast domain during the preprocessing steps to provide forecast over the specified duration. The settings in the namelist.input file are used to configure WRF model (real.exe/real\_nmm.exe and wrf.exe). The data flow between the programs of the WRF module is shown in Figure 2.11.

**WPP:** The WPP (wrfpost.exe and copygb.exe) was designed to interpolate both ARW and NMM output from their native grids to National Weather Service (NWS) standard levels (*pressure, height, etc.*) and standard output grids (*AWIPS, Lambert Conformal, polar-stereographic, etc.*) in NWS and WMO GRIB1 format. This package also provides an option to output fields on the model's native vertical levels.

**GRADS:** The Grid Analysis and Display System (GrADS) is an interactive tool that is used for access, manipulation, and visualization of earth science data with different file format (GRIB, NetCDF, HDF and

BUFR). Operations are executed interactively by entering FORTRAN-like expressions at the command line and also add their own functions as external routines written in any programming language. Grads can visualize the WPP output in image format, ASCII format and to standard out.

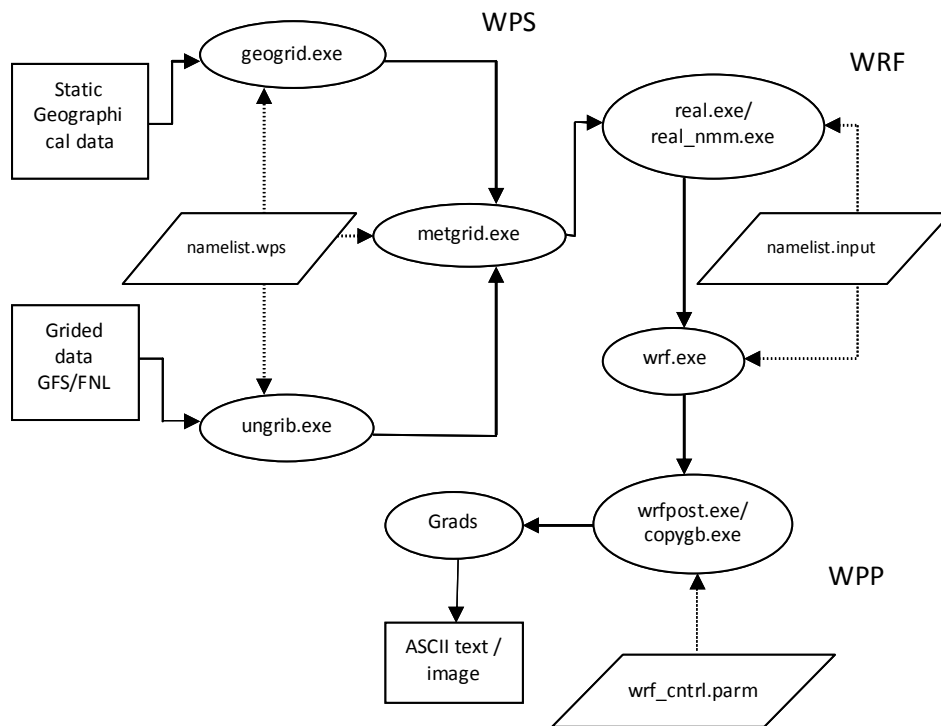


Figure 2.11: WRF workflow chart.





## Chapter-3

# ARTIFICIAL NEURAL NETWORK MODEL FOR THUNDERSTORM PREDICTION

---

### Contents

- 3.1 Data and Methodology
  - 3.2 Learning Algorithms for the Present Study
  - 3.3 Case Description
  - 3.4 Results and Discussion
  - 3.5 Chapter Summary
- 

Accurate forecasting of thunderstorms and severe thunderstorms is critical for a large range of users in the community. The general public can benefit from timely forecasts and warnings of impending severe thunderstorms. Thunderstorm forecasting typically has proved to be one of the most difficult tasks, due to their rather small spatial and temporal extension and the inherent non-linearity of their dynamics and physics (Orlanski 1975). ANN based approach can be used to model complex relationships between inputs and outputs or to find patterns in data. ANN can be viewed as a mathematical model or computational model that is inspired by the structure or functional aspects of biological neural networks. Neural networks are designed to extract existing patterns from noisy data. The procedure involves training a network (training phase) with a large sample of representative data, after which one exposes the network to data not included in the training set (validation or prediction

phase) with the aim of predicting the new outcomes (Bishop 1995). The interest in neural networks comes from the networks' ability to mimic human brain as well as its ability to learn and respond. As a result, neural networks have been used in a large number of applications and have proven to be effective in performing complex functions in a variety of fields (Haykin 1994).

ANN have proven to be powerful and general technique for machine learning (Shavlik et al. 1991). Most successful application of neural networks involved pattern recognition, statistical mapping or modeling (Sentiono et al. 2002). According to Bailey and Thompson (1990), successful application can include signal validation, process monitoring, diagnostics, signal and information processing and control of complex system. James et al. (2000) mentioned that, ANNs have the ability to tackle the problem of complex relationships among variables that cannot be accomplished by more traditional methods. ANNs are excellent tools for complex manufacturing processes that have many variable and complex processes. According to Palade et al. (2001), ANNs represent an excellent tool that has been used to develop a wide range of real world applications, especially in case when traditional solving methods fail. The advantages of ANNs such as ideal learning ability from data, classification capabilities and generalization for situation not contain training data set, computationally fastness once trained due to parallel processing, noise tolerance. There were these advantages that make



ANNs to be successfully applied to various real world problems, including medical diagnosis (Hayashi et al. 2000), image computing, speech recognition, process control and modeling (Palade et al. 2001). However, a little attention was paid to the use of ANNs in weather forecasting (Maqsood et al. 2000 and 2004; Bodri and Cermak 2000; Luk et al. 2000; Wedge et al. 2005; Steidley et al. 2005; Chaudhury and Chattopadhyay 2005).

Bodri and Cermak (2000) developed an ANN using 38 year of rainfall data to predict monthly and yearly precipitation levels for multiple sites in the Czech Republic. Using spatial and temporal data of recent rainfall, Luk et al. (2000) developed an ANN for short-term precipitation prediction focused on predicting flash flood rainfall amounts for 15 min ahead for various areas of western Sydney, Australia. Maqsood et al. (2004) used an ensemble of ANNs to provide 24 hour (h) predictions for air temperature, wind speed, and humidity at the Regina Airport in Canada. Wedge et al. (2005) developed an ANN for prediction of waves spilling over sea walls in using sea conditions and wall properties as inputs. Steidley et al. (2005) used ANNs to predict tidal water levels for periods of 3–48 h ahead for a shallow embayment on the coast of Texas in the United States. Chaudhury and Chattopadhyay (2005) designed a feed-forward multi-layered ANN model to estimate the maximum surface temperature and relative humidity. The recent advances in neural network methodology for modeling non-linear, dynamical phenomena along with the

impressive successes in a wide range of applications, motivated to investigate the application of ANNs for the prediction of hourly temperature and relative humidity needed for the genesis of severe thunderstorms over Kolkata.

In this chapter, experiments are conducted with ANN model to predict severe thunderstorms that occurred over Kolkata ( $22.52^{\circ}\text{N}$ ,  $88.37^{\circ}\text{E}$ ) using thunderstorm affected meteorological parameters. The geographical location of the study area is given in Figure 3.1.



**Figure3.1:** The geographical location of Kolkata in West Bengal.

The performance of six learning algorithms namely Step (STP), Momentum (MOM), Quick Propagation (QKP), Delta-Bar-Delta (DBD), CG and LM are evaluated using predicted hourly surface temperature and relative humidity during thunderstorm days. The accuracy of the predictions was evaluated by the correlation coefficient (CC), the root mean-square error (RMSE), the

mean absolute error (MAE) and Percent Correct (PC) between the measured and predicted values. The developed ANN model with LM algorithm was applied to derive thunderstorm forecast from 1 to 24 h ahead at Kolkata. The goal of this study was to use ANNs to predict hourly temperature and relative humidity during thunderstorm days from 1 to 24 h ahead using prior weather data as inputs.

### **3.1 Data and Methodology**

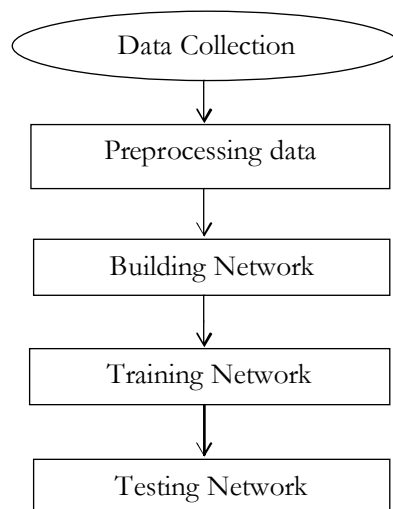
#### **3.1.1 ANN experimental setup**

The developed ANN model is based on one of the neural network architecture named MLPN model (also known as multilayer feed-forward network). This is the most popular network architecture in use today. This is the type of network which the units each perform a biased weighted sum of their inputs and pass this activation level through a transfer function to produce their output, and the units are arranged in a layered feed forward topology. The network thus has a simple interpretation as a form of input-output model, with the weights and thresholds (biases), the free parameters of the model. Such networks can model functions of almost arbitrary complexity with the number of layers and the number of units in each layer, determining the function complexity. Important issues in multilayer perceptron design include specification of the number of hidden layers and the number of units in these layers (Hagan et al. 1996). Once the number of layers and number of units in each layer have been

selected, the network's weights and thresholds must be set so as to minimize the prediction error made by the network. This is the role of the learning algorithms. The best known example of a neural network learning algorithm is BP. Modern second-order algorithm such as CG and LM are substantially faster for many problems (Bishop 1995; Shepherd 1997). There are also heuristic modifications of BP which work well for some problem domains, such as QKP (Fahlman 1988) and DBD (Jacobs 1988).

This study evaluates the utility of MLPN for estimating hourly surface temperature and relative humidity. Designing ANN model follows a number of systemic procedures. In general, there are five basic steps: (1) collecting data, (2) preprocessing data, (3) building the network, (4) train and (5) test performance of model. The basic flow in designing ANN model is given in Figure 3.2. The hourly surface weather parameters namely mean sea level pressure (hPa), relative humidity (%) and wind speed ( $\text{ms}^{-1}$ ) of 3 years (April and May 2007 to 2009) collected from the India meteorological department (IMD) of Kolkata were used as the input data for training and testing the ANN model which will be used for the prediction of hourly temperature. The hourly surface weather parameters namely hourly mean sea level pressure (hPa), temperature ( $^{\circ}\text{C}$ ) and wind speed ( $\text{ms}^{-1}$ ) of 3 years (April and May 2007 to 2009) of Kolkata were used as the input data for training and testing the ANN model which will be used for the prediction of hourly relative humidity. The other

additional input parameters for each model are month, day and hour of the observation. The total length of the data record included for ANN model is more than 4000. Fragment of the data used for this study are attached in the appendix.



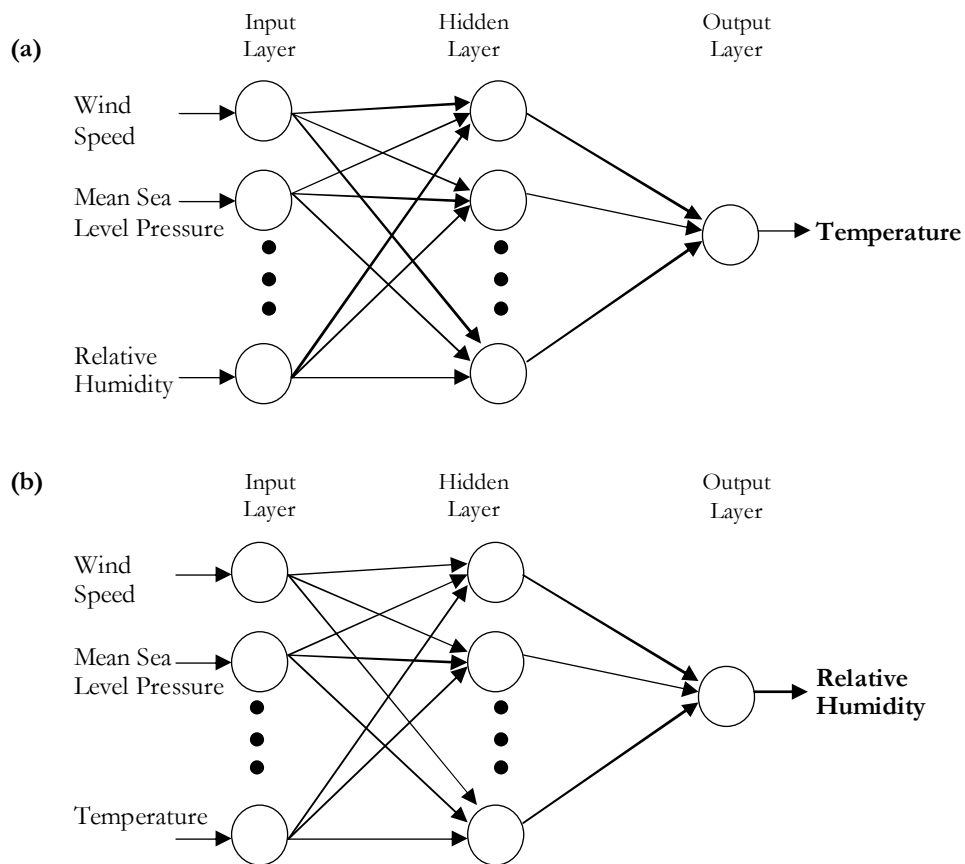
**Figure 3.2:** Basic flow for designing ANN model.

After data collection, two data preprocessing procedures are conducted to train the ANNs more efficiently. These procedures are: (1) solve the problem of missing data and (2) normalize data. The missing data are replaced by the average of neighboring values. Neural networks generally provide improved performance with the normalized data. The use of original data as input to neural network may cause a convergence problem (Khan and Ondrusek 2000). All the weather data sets were therefore, transformed into values between -1 and 1 through dividing the

difference of actual and minimum values by the difference of maximum and minimum values. At the end of each algorithm, outputs were de-normalized into the original data format for achieving the desired result.

The configuration of the neural network depends highly on the problem. In this study, separate models with same configuration have been built to predict both surface parameters namely temperature and relative humidity. A three layer structure (one input layer, one hidden layer and one output layer) was selected with hyperbolic tangent (tanh) transfer function for hidden layer and linear transfer function for output layer. Figure 3.3 provides an overview of the structure of MLPN model for the prediction of temperature and relative humidity. The chosen weather data were divided into two randomly selected groups, the training group and test group. The hourly surface parameters during April and May 2007-2008 and April 2009 correspond to training group and the hourly surface parameters of May 2009 correspond to test group. Major numbers of thunderstorms are occurred over Kolkata in April and May. Thus the hourly data sets of these two months are selected for training and testing. Networks were trained for a fixed number of epochs. The error level was set to a relatively small value ( $10^{-4}$ ). The optimal number of hidden neurons was obtained experimentally by changing the network design and running the training process several times until a good performance was obtained. A random number generator was used to assign the initial values of weights and thresholds with a small bias as a

difference between each weight connecting two neurons together since similar weights for different connections may lead to a network that will never learn. Hourly surface temperature is used as test data for the first model and hourly relative humidity for second model.



**Figure 3.3:** Architecture of MLPN for the prediction of (a) temperature and (b) relative humidity.

The 24 h ANN model outputs of surface temperature and relative humidity at Kolkata (22.52<sup>0</sup>N, 88.37<sup>0</sup>E) during three severe thunderstorm days of May 2009 (3, 11 and 15 May 2009) were used to evaluate these models. The capability of six different learning algorithms in predicting thunderstorms were studied and their performances were compared for the prediction of surface temperature and relative humidity. The learning algorithms used for these studies are STP, MOM, CG, QKP, LM and DBD. Performance and reliabilities of the models were then evaluated by a number of statistical measures like RMSE, MAE, CC and PC. The developed ANN model with LM algorithm was applied to derive thunderstorm forecast from 1 to 24 h ahead at Kolkata from the data of 3 consecutive years (April and May 2007 – 2009). Models were created to predict temperature and relative humidity at hourly intervals with 1, 3, 6, 12 and 24 h ahead. The results are evaluated using MAE, RMSE, CC and PC. The ANN model simulations are carried out using the NeuroSolutions software developed by NeuroDimension Inc. of Florida (NeuroDimension 2005). The details of the NeuroSolutions software are given in appendix.

### **3.1.2 Statistical analysis**

Verifying forecasts of continuous variables measures how the values of the forecasts differ from the values of the observations. The performances of the models developed in this study were assessed using various standard statistical performance evaluation criteria. The following



statistical parameters were calculated for the comparisons between the forecast and observation data.

**MAE:** This is the average over the verification sample of the absolute values of the differences between forecast and the corresponding observation. The MAE is a linear score which means that all the individual differences are weighted equally in the average. The MAE is a common measure of forecast error in time series analysis. It measures accuracy (the level of agreement between the forecast and the observations) for continuous variables. The MAE is given by

$$MAE = \frac{1}{N} \sum_{i=1}^N |F_i - O_i| \quad (3.1)$$

where  $F_i$  is the forecast and  $O_i$  is the observation. The difference between the forecast and the observation is the error. The lower the errors, the greater the accuracy. The range is 0 to infinity. The perfect score is 0.

**RMSE:** It is a frequently used measure of the differences between values predicted by a model and the values actually observed. It measures average error, weighted according to the square of the error. It does not indicate the direction of the deviation. The RMSE puts greater influence on large errors than smaller errors, which may be a good thing if large errors are especially undesirable, but may also conservative forecasting. The RMSE is given by

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_i - O_i)^2} \quad (3.2)$$

**CC:** This indicates the strength and direction of a linear relationship between two random variables. That means it measures the strength of the linear relationship between the forecasts and observations. The  $r$  is given by

$$r = \frac{\sum (F - \bar{F})(O - \bar{O})}{\sqrt{\sum (F - \bar{F})^2} \sqrt{\sum (O - \bar{O})^2}} \quad (3.3)$$

The value of  $r$  is such that  $-1 \leq r \leq +1$ . The + and – signs are used for positive linear correlations and negative linear correlations, respectively. If  $F$  and  $O$  have a strong positive linear correlation,  $r$  is close to +1. An  $r$  value of exactly +1 indicates a perfect positive fit. A correlation greater than 0.8 is generally described as strong, whereas a correlation less than 0.5 is generally described as weak.

**PC:** It is used to represent the number of times out of 100 a task is performed correctly. Percentages represent the portion of 100 outcomes that are successful or meet certain criteria. The temperature forecast is ‘Correct’ when the difference between the forecasted temperature and the actual is  $\pm 2^\circ\text{C}$  or less. The relative humidity forecast is ‘Correct’ when the difference between the forecasted relative humidity and the actual is  $\pm 10\%$  or less.

### **3.2 Learning Algorithms for the Present Study**

In neural network, the learning algorithms plays quite important role in the process. An appropriate topology may still fail to give a better model, unless trained by a suitable learning algorithm. A good learning algorithm will shorten the training time, while achieving a better accuracy. Therefore, training process is an important characteristic of the ANNs, whereby representative examples of the knowledge are iteratively presented to the network, so that it can integrate this knowledge within its structure. There is a number of learning algorithms used to train a MLPN and a frequently used one is called the BP learning algorithm. The BP algorithm, which is based on searching an error surface using gradient descent for points with minimum error, is relatively easy to implement. The pseudo code for BP algorithm (Mitchell 1997) is given as follows:

#### **Pseudo Coding**

```
Assign all network inputs and output
Initialize all weights with small random numbers, typically between -1
and 1
repeat
  for every pattern in the training set
    Present the pattern to the network
//   Propagated the input forward through the network:
    for each layer in the network
```

```
    for every node in the layer
        1. Calculate the weight sum of the inputs to the node
        2. Add the threshold to the sum
        3. Calculate the activation for the node
    end
end
// Propagate the errors backward through the network
    for every node in the output layer
        calculate the error signal
    end
    for all hidden layers
        for every node in the layer
            1. Calculate the node's signal error
            2. Update each node's weight in the network
        end
    end
// Calculate Global Error
    Calculate the Error Function
end
while ((maximum number of iterations < than specified) AND
      (Error Function is > than specified))
```

However, the BP algorithm has some problems for many applications. The algorithm is not guaranteed to find the global minimum of the error

function since gradient descent may get stuck in local minima, where it may remain indefinitely. In addition to this, long training sessions are often required in order to find an acceptable weight solution because of the well-known difficulties inherent in gradient descent optimization. Therefore, a lot of variations to improve the convergence of the BP were proposed such as DBD, QKP (Fahlman 1988; Rumelhart et al. 1986; Jacobs 1988). Optimization methods such as second-order methods (CG, QN and LM) have also been used for ANN learning in recent years. The LM algorithm combines the best features of the Gauss–Newton technique and the steepest-descent algorithm, but avoids many of their limitations. In particular, it generally does not suffer from the problem of slow convergence (Hagan and Menhaj 1994). Six learning algorithms were applied in this study, in order to identify the one which trains a given network more efficiently. All of them are variations of basic BP algorithms. Variations of these learning algorithms are given in the following sections:

### **3.2.1 STP algorithm**

Gradient descent (GD) learning rules provide first order gradient information about the network's performance surface (e.g. BP and real time recurrent learning). The most straightforward way of reaching the bottom (the minima) given which way is up, is to move in the opposite direction. With this scenario, the only variable is the step size (i.e. how far should it move before obtaining another directional estimate). If the steps

are too small, then it will take too long to get there. If the steps are too large, then it may overshoot the bottom, causing it to rattle or even diverge. The STP uses this procedure to adapt the weights of the activation component that it is stacked on (NeuroSolutions Manual 2003).

### 3.2.2 MOM algorithm

Step components try to find the bottom of a performance surface by taking steps in the direction estimated by the attached BP component. Network learning can be very slow, if the step size is small. It can oscillate or diverge if it is chosen too large. For further complicate matters, a step size that works well for one location in weight space may be unstable in another. The momentum provides the gradient descent with some inertia, so that it tends to move along a direction that is the average estimate for down. The amount of inertia (i.e. how much of the past to average over) is imposed by the momentum parameter. The higher the momentum, the more it smoothes the gradient estimate and the less effect a single change in the gradient has on the weight change. The major benefit is the added ability to breakout of local minima that a step component might otherwise get caught in. Note that oscillations may occur if the momentum is set too high. The momentum parameter is the same for all weights of the attached component. An access point has been provided for the step size and momentum allowing access for adaptive and scheduled learning rate procedures (NeuroSolutions Manual 2003).

### **3.2.3 CG algorithm**

The GD algorithms (like “step” and “momentum”) use only the local approximation of the slope of the performance surface (error versus weights) to determine the best direction to move the weights in order to lower the error. Second order methods use or approximate second derivatives (the curvature instead of just the slope) of the performance surface to determine the weight update. This information is very important for determining the optimal update direction. Since this method makes use of the second derivatives of the function to be optimized, it is typically referred to as the second order methods (Jalali-Heravi et al. 2008).

### **3.2.4 LM algorithm**

The LM algorithm is one of the most appropriate higher-order adaptive algorithms known for minimizing the Mean Square Error (MSE) of a neural network. It is a member of a class of learning algorithms called "pseudo second order methods". Standard gradient descent algorithms use only the local approximation of the slope of the performance surface (error versus weights) to determine the best direction to move the weights in order to lower the error. Second order methods use the Hessian or the matrix of second derivatives (the curvature instead of just the slope) of the performance surface to determine the weight update, while pseudo-second order methods approximate the Hessian. In

particular the LM utilizes the so called Gauss-Newton approximation that keeps the Jacobian matrix and discards second order derivatives of the error. If the performance surface is quadratic (which is only true in general for linear systems) then using a second order method can find the exact minimum in one step. A key advantage of the LM approach is that it defaults to the gradient search when the local curvature of the performance surface deviates from a parabola, which may happen often in neural computing (NeuroSolutions Manual 2003).

### **3.2.5 QKP algorithm**

The QKP uses information about curvature of the error surface. This requires the computation of the second order derivatives of the error function during training. The QKP assumes the error surface, a function of connection weights, to be locally quadratic (i.e a parabola) and attempts to jump in one step from the current position directly into the minimum of the parabola. The QKP computes the derivatives in the direction of each weight. After computing the first gradient as in regular back propagation, a direct step to the error is attempted by changing the weight (Chakraborty and Chakraborty2002).

### **3.2.6 DBD algorithm**

The DBD is an adaptive step-size procedure for searching a performance surface. The step size and momentum are adapted according to the previous values of the error at the neurons. If the current and past



weight updates are both of the same sign, it increases the learning rate linearly. The reasoning is that if the weight is being moved in the same direction to decrease the error, then it will get there faster with a larger step size. If the updates have different signs, this is an indication that the weight has been moved too far. When this happens, the learning rate decreases geometrically to avoid divergence (Haciismailoglu et al. 2009).

A number of researchers have carried out comparative studies of MLPN learning algorithms. Kisi and Uncuoglu (2005) compared LM, CG and resilient algorithm for stream-flow forecasting and determination of lateral stress in cohesionless soils. They found that LM algorithm was faster and achieved better performance than the other algorithms in learning. Esugasini et al. (2005) considered the problem of breast cancer diagnosis and compared the classification accuracy of the standard steepest descent against the classification accuracy of the gradient descent with momentum and adaptive learning, resilient BP, QN and LM algorithms. The simulations show that the neural network using the LM algorithm achieved the best classification performance. Raju et al. (2011) demonstrated the application of ANNs in predicting the weekly spring discharge with three different learning algorithms like QKP algorithm, batch BP algorithm and LM algorithm. They conclude that the QKP algorithm had a better performance to the application. Ubeyli and Guler (2004) compared BP, DBD, extended DBD, QKP, and LM algorithms to compute the quasi-static parameters, the characteristic impedance and the

effective dielectric constant, of the asymmetric coplanar waveguides (ACPWs). The results of the LM algorithm for the quasi-static parameters of the ACPWs were in very good agreement with the results available in the literature. The results of above studies have illustrated that the relative performance of algorithms depends on the problem being tackled.

### **3.3 Case Description**

For the ANN model validation, three severe thunderstorm cases of May 2009 (3, 11 and 15 May 2009) have been taken and the description of each case is as follows:

Case 1 was a severe thunderstorm, which was reported on 3 May 2009 over Kolkata with a maximum speed of 61.2 kilometer per hour (kmph) lasting for a few minutes. This intense convective event produced 31.4 mm (millimeter) rainfall over Kolkata. In the synoptic charts at 0000 UTC (Coordinated Universal Time) a low pressure area was found at the surface over north Chattisgarh and adjoining Jharkhand and a trough from this extending southward up to interior Tamilnadu across Andhra Pradesh is found. At 1.5 km above sea level (a.s.l) cyclonic circulation is seen over west Uttar Pradesh, a trough from this extends southeastwards up to south peninsula across east Madhya Pradesh and Andhra Pradesh. No significant trough in mid troposphere. No subtropical westerly jet maxima were seen over the region. A few places recorded moderate rainfall over Gangetic West Bengal (GWB) and isolated rainfall over Orissa,

Chattisgarh and Bihar. Bankura recorded 24.9 mm and Sriniketan 38.2 mm of rainfall.

Case 2 was a severe thunderstorm, which was reported on 11 May 2009 over Kolkata with squally winds of the order of 87 kmph. Rainfall of 33.3 mm was reported over Kolkata. The synoptic charts show a trough at sea level chart from east Uttar Pradesh to north Tamilnadu across east Madhya Pradesh and Andhra Pradesh. Cyclonic circulation in lower levels is found over Bihar and neighborhood. Trough from this extends up to extreme south peninsula across Chattisgarh, Telangana and Rayalaseema. Another cyclonic circulation existed over Arunachal Pradesh and adjoining Assam and Meghalaya. A trough from Arunachal Pradesh to northwest Bay of Bengal was found in middle troposphere. Sub-tropical westerly jet maxima were found over the region. Light to moderate rain occurred at few places over Orissa and GWB with Midnapore and Alipore reporting 17.8 mm and 21.9 mm respectively.

Case 3 was a severe thunderstorm, which was reported on 15 May 2009. A squall passed over Kolkata on 15 May 2009 with a maximum speed of 68.4 kmph. This intense convective event produced 16.9 mm rainfall over Kolkata. The synoptic charts show a trough at sea level from east Madhya Pradesh to south coastal Tamilnadu across Telangana and another trough to northeast Bay of Bengal across Orissa. Cyclonic circulation seen in lower levels over west Uttar Pradesh and a trough from this extends up to coastal Andhra Pradesh across Vidarbha with

embedded cyclonic circulation over Telangana. Trough in mid – troposphere is found from Arunachal Pradesh to north Bay of Bengal. Sub-tropical westerly jet maxima were found over the region. A few places of GWB recorded moderate rainfall and isolated rainfall over Orissa and Bihar. Bankura recorded 34.0 mm and Midnapore 51.6 mm of rainfall (Mohanty et al. 2009).

### **3.4 Results and Discussion**

According to the previous studies of (Doswell 1987; Johns and Doswell 1992; McNulty 1995) the general preconditions for the initiation of thunderstorms are conditional instability, a sufficiently deep humid layer in the lower and mid-troposphere and an uplifting mechanism to initiate convection. The formation of thunderstorms is an interaction between these conditions on different scales. The surface parameters play a significant role in the genesis whereas the strength of the upper air pull is required to assess the growth of the thunderstorm (Asnani 2006). The greater the density differences between air masses (temperature and humidity) the greater the atmospheric instabilities that develop, and the greater the intensity of these thunderstorms (Price 2006). Recent studies show a high positive correlation between surface temperature and lightning activity (William 2005). The temperature and relative humidity on the surface are useful tool in forecasting the likelihood occurrence of a thunderstorm (Lopez et al. 2007). A sudden drop in temperature or

sudden increase in relative humidity during the day indicates for the occurrence of thunderstorm (Asnani 2006).

### **3.4.1 Comparison of learning algorithms**

The ANN model predicted surface temperature and relative humidity with different learning algorithms during severe thunderstorm cases are explored in the following section. Analysis of the results of these experiments are helpful to understand the impact of learning algorithms on the prediction of severe thunderstorm events and assist in the customization of model for future severe thunderstorm predictions over east and northeast Indian region.

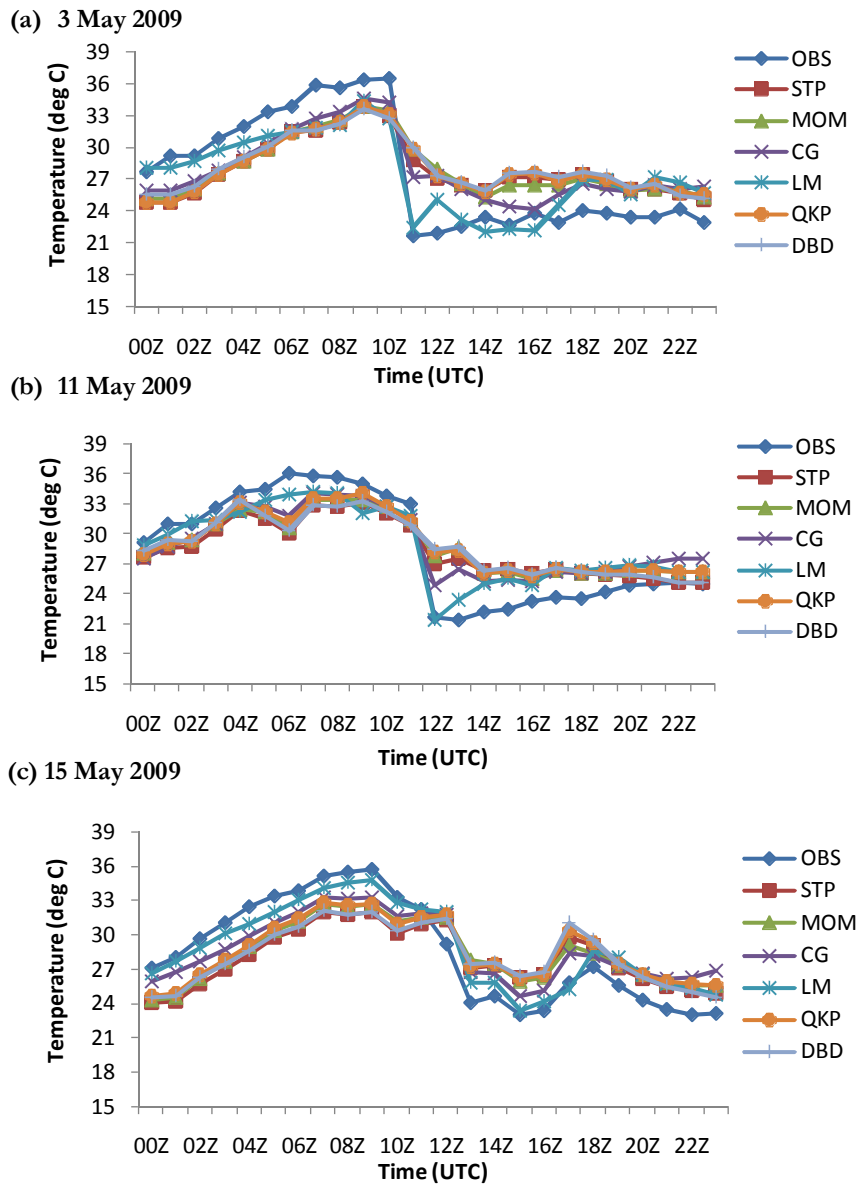
Figure 3.4 shows the inter-comparison of observed and ANN model predicted diurnal variation of surface temperature ( $^{\circ}\text{C}$ ) with different learning algorithms over Kolkata valid for 3, 11 and 15 May 2009. From the figures, it is clearly visible that the observed data (OBS) show a sudden drop in temperature in all three thunderstorm days. The ANN model with different learning algorithms captured the temperature drop during the thunderstorm hour for all the three cases. But the predicted intensity is different for different algorithms. For the first case (Figure 3.4a), the observed temperature showed a sudden drop of  $15^{\circ}\text{C}$  from  $36.7^{\circ}\text{C}$  to  $21.7^{\circ}\text{C}$  at 1000 UTC. The ANN model prediction with LM showed a drop from  $33^{\circ}\text{C}$  to  $22^{\circ}\text{C}$  ( $11^{\circ}\text{C}$ ) at 1000 UTC, whereas CG presented a drop from  $34^{\circ}\text{C}$  to  $27^{\circ}\text{C}$  ( $7^{\circ}\text{C}$ ) at 1000 UTC. All other

algorithms showed a difference less than  $4^{\circ}\text{C}$  during thunderstorm hour. The DBD has least performance than other algorithms. In the second thunderstorm case (Figure 3.4b), observed temperature fall is from  $33.1^{\circ}\text{C}$  to  $21.7^{\circ}\text{C}$  ( $11^{\circ}\text{C}$ ) at 1200 UTC, whereas LM indicated a drop from  $32^{\circ}\text{C}$  to  $21^{\circ}\text{C}$  ( $11^{\circ}\text{C}$ ) at the same thunderstorm hour. CG showed only  $6^{\circ}\text{C}$  difference between predicted and observed values. The other algorithms presented less intensity in difference between predicted and observed values during thunderstorm hour. For the third case (Figure 3.4c), observed temperature showed a drop from  $29^{\circ}\text{C}$  to  $24^{\circ}\text{C}$  ( $6^{\circ}\text{C}$ ) at 1300 UTC, whereas LM showed a drop from  $32^{\circ}\text{C}$  to  $27^{\circ}\text{C}$  ( $5^{\circ}\text{C}$ ). All other algorithms are also captured the sudden fall with almost same intensity of observation and LM algorithm for this thunderstorm case.

Relative humidity at surface level has been taken into account, as it is an essential factor in intense convection. Storm days require a sufficiently humid and deep layer in the lower and middle atmosphere (Johns and Doswell 1992). Figure 3.5 shows the inter-comparison of observed and ANN model predicted relative humidity (%) with different learning algorithms over Kolkata for severe thunderstorm days. For all the thunderstorm cases, ANN model with different algorithms have captured the increase in relative humidity during thunderstorm hour as in the observation. But the predicted intensity is different for different learning algorithms. In the first case (Figure 3.5a), the observed relative humidity showed a rise of 48% from 52% to 100% at 1000 UTC. The

ANN model prediction with LM showed a rise from 53% to 95% (42%) at 1000 UTC. All other algorithms except CG show same change in intensity (32%) at 1000 UTC, whereas CG presented a rise from 55% to 81% (25%) at 1100 UTC. The performance of CG algorithm is poor than all other algorithms during first thunderstorm case. In the second case (Figure 3.5b), observed relative humidity rise is from 66% to 100% (34%) at 1200 UTC, whereas LM indicated a rise from 68% to 100% (32%) at the same time. As in the previous case, the CG shows increase in relative humidity at 1400 UTC with 16% change in intensity. The changes in intensity predicted by other algorithms are also same and the intensity of sudden increase is 23%. For the third case (Figure 3.5c), observed relative humidity showed a rise from 63% to 100% (37%) at 1300 UTC, whereas LM showed a rise from 73% to 95% (22%). The other algorithms showed an intensity rise around 10%. From these analyses of temperature and relative humidity, it can be clearly seen that ANN model with LM algorithm well predicted diurnal variation during thunderstorm days and captured the sudden drop and rise with almost same intensity of observation as compared to other algorithms.

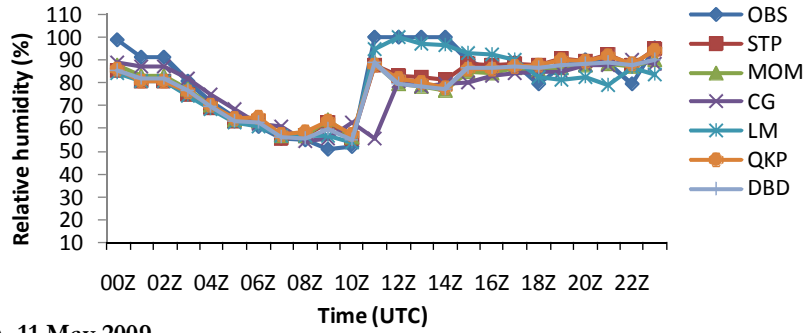
The results of statistical analysis based on MAE, RMSE and CC to evaluate forecasted temperature and relative humidity are shown in Table 3.1 and Table 3.2. The results of Table 3.1 indicated that, LM algorithm has less MAE and RMSE as compared to all other algorithms for these 3 thunderstorm cases. The CG algorithm has also given moderate results.



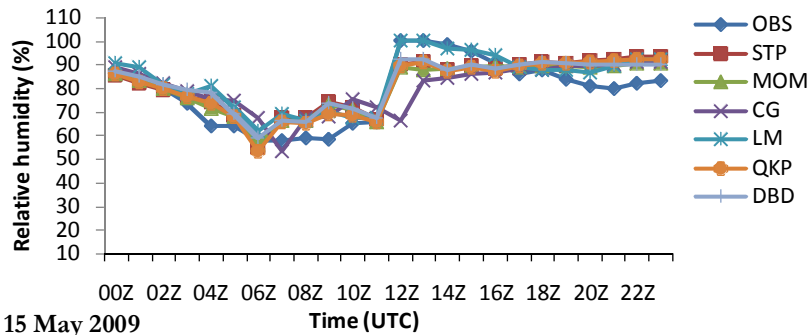
**Figure 3.4:** Comparison of ANN predicted hourly surface temperature ( $^{\circ}$ C) using different learning algorithms with observation on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.



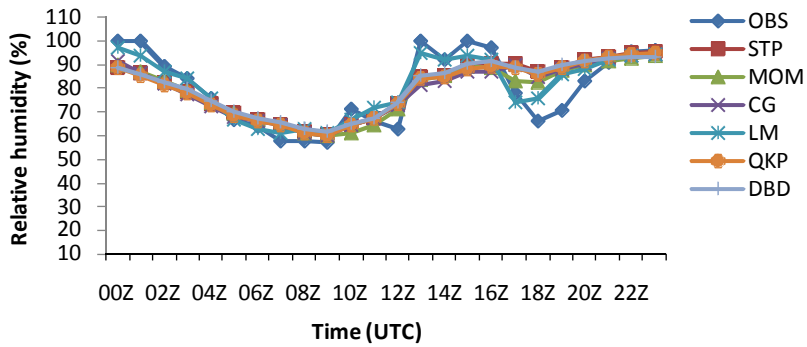
(a) 3 May 2009



(b) 11 May 2009



(c) 15 May 2009



**Figure 3.5:** Comparison of ANN predicted hourly relative humidity (%) using different learning algorithms with observation on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.

All other algorithms displayed more error in all thunderstorm cases as compared to LM and CG algorithms. The average MAE and RMSE from these 3 cases are also less for LM algorithm than other 5 algorithms. Another verification method used for this study is correlation coefficient. From the table, it can be seen that all algorithms are positively correlated. The LM algorithm has the highest CC in all three cases as compared to all other algorithms. The average CC of LM and CG are high and the values are more than 0.9. The CC of other algorithms is less than 0.85. The performance of DBD algorithm is less efficient than other algorithms. The analysis shows that LM algorithm is best for hourly temperature prediction over Kolkata during thunderstorm days.

**Table 3.1:** Performance comparison of different learning algorithms in hourly temperature prediction.

Statistical Analysis	Dates	STP	MOM	CG	LM	QKP	DBD
MAE	3-May-09	3.36	3.24	2.69	2.08	3.48	3.41
	11-May-09	2.69	2.57	2.27	1.72	2.54	2.62
	15-May-09	2.93	2.66	2.08	1.21	2.69	2.90
	<b>MEAN</b>	<b>2.99</b>	<b>2.82</b>	<b>2.35</b>	<b>1.67</b>	<b>2.90</b>	<b>2.98</b>
RMSE	3-May-09	3.54	3.50	2.90	2.35	3.70	3.69
	11-May-09	3.07	3.02	2.44	1.90	2.99	3.19
	15-May-09	3.07	2.76	2.20	1.41	2.78	3.04
	<b>MEAN</b>	<b>3.23</b>	<b>3.09</b>	<b>2.51</b>	<b>1.89</b>	<b>3.16</b>	<b>3.31</b>
CC	3-May-09	0.82	0.82	0.90	0.93	0.79	0.80
	11-May-09	0.89	0.89	0.94	0.97	0.89	0.86
	15-May-09	0.74	0.80	0.91	0.96	0.80	0.74
	<b>MEAN</b>	<b>0.82</b>	<b>0.84</b>	<b>0.92</b>	<b>0.95</b>	<b>0.83</b>	<b>0.80</b>

The results of Table 3.2 indicated that, LM algorithm has less error as compared to all other algorithms for these 3 thunderstorm cases as in temperature study. All other algorithms have also given moderate results except CG algorithm. The CG algorithm displayed more error in all cases. But in temperature prediction (Table 3.1), CG algorithm performed well than other 4 algorithms, namely STP, MOM, DBD and QKP. The average MAE and RMSE of LM algorithm has least value than other 5 algorithms.

**Table 3.2:** Performance comparison of different learning algorithms in hourly relative humidity prediction.

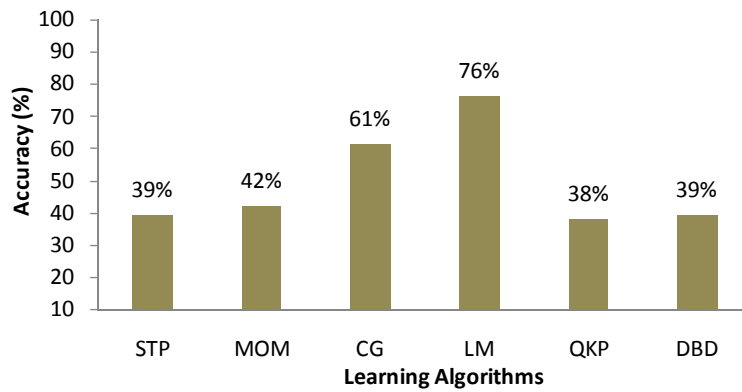
Statistical Analysis	Dates	STP	MOM	CG	LM	QKP	DBD
MAE	3-May-09	6.62	6.78	8.21	5.24	7.06	6.65
	11-May-09	6.80	5.77	8.46	5.20	6.19	6.26
	15-May-09	9.23	9.20	9.86	3.02	9.53	8.98
	<b>MEAN</b>	<b>7.55</b>	<b>7.25</b>	<b>8.84</b>	<b>4.49</b>	<b>7.60</b>	<b>7.30</b>
RMSE	3-May-09	8.77	9.50	12.64	6.55	9.43	9.42
	11-May-09	7.99	7.06	10.71	6.76	7.25	7.30
	15-May-09	10.33	9.94	10.93	3.71	10.56	9.89
	<b>MEAN</b>	<b>9.03</b>	<b>8.83</b>	<b>11.43</b>	<b>5.67</b>	<b>9.08</b>	<b>8.87</b>
CC	3-May-09	0.86	0.86	0.69	0.93	0.84	0.86
	11-May-09	0.84	0.88	0.64	0.95	0.86	0.89
	15-May-09	0.68	0.76	0.67	0.95	0.68	0.69
	<b>MEAN</b>	<b>0.80</b>	<b>0.83</b>	<b>0.67</b>	<b>0.95</b>	<b>0.80</b>	<b>0.82</b>

From the table (Table 3.2), it can be seen that all the algorithms are positively correlated. The LM algorithm has the highest CC in all three cases as compared to all other algorithms. The average CC of 3 thunderstorm cases is more for LM algorithm and which is more than 0.9.

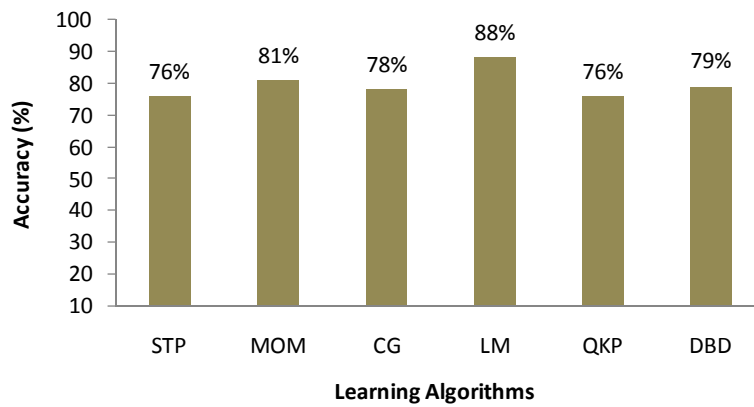
The CC of other algorithms except CG is 0.8 which is strong correlation. The performance of CG algorithm is less efficient than other algorithms for the prediction of hourly surface relative humidity during thunderstorm days. The results show that LM algorithm is best for hourly relative humidity prediction over Kolkata during thunderstorm days.

Figure 3.6 gives the performance accuracy of learning algorithms for hourly temperature and relative humidity prediction. The PC of temperature presented a percentage number of the times when the forecast is accurate to within  $\pm 2^{\circ}\text{C}$ . The results (Figure 3.6a) indicated that overall accuracy of LM algorithm for three events is 76%. CG gave a moderate accuracy of 61%. Other algorithms displayed less accuracy. The PC of relative humidity presented a percentage number of the times when the forecast is accurate to with  $\pm 10\%$  confidence range. The results (Figure 3.6b) indicated that overall accuracy of LM algorithm for three events is 88%. MOM algorithm also shows a good accuracy of 81%. The other algorithms displayed a moderate accuracy, which is more than 75%. The time-series plots and statistical analysis of temperature and relative humidity revealed that LM algorithm is well predicted the occurrence and intensity of all 3 thunderstorm cases as in the observation. The results suggest that the ANN model with LM algorithm holds promise for prediction of surface weather parameters with reasonable accuracy in severe thunderstorm cases.

(a) Temperature



(b) Relative humidity



**Figure 3.6:** Performance accuracy of learning algorithms for the prediction of (a) temperature and (b) relative humidity during thunderstorm days.

### 3.4.2 Comparison of different advanced predictions

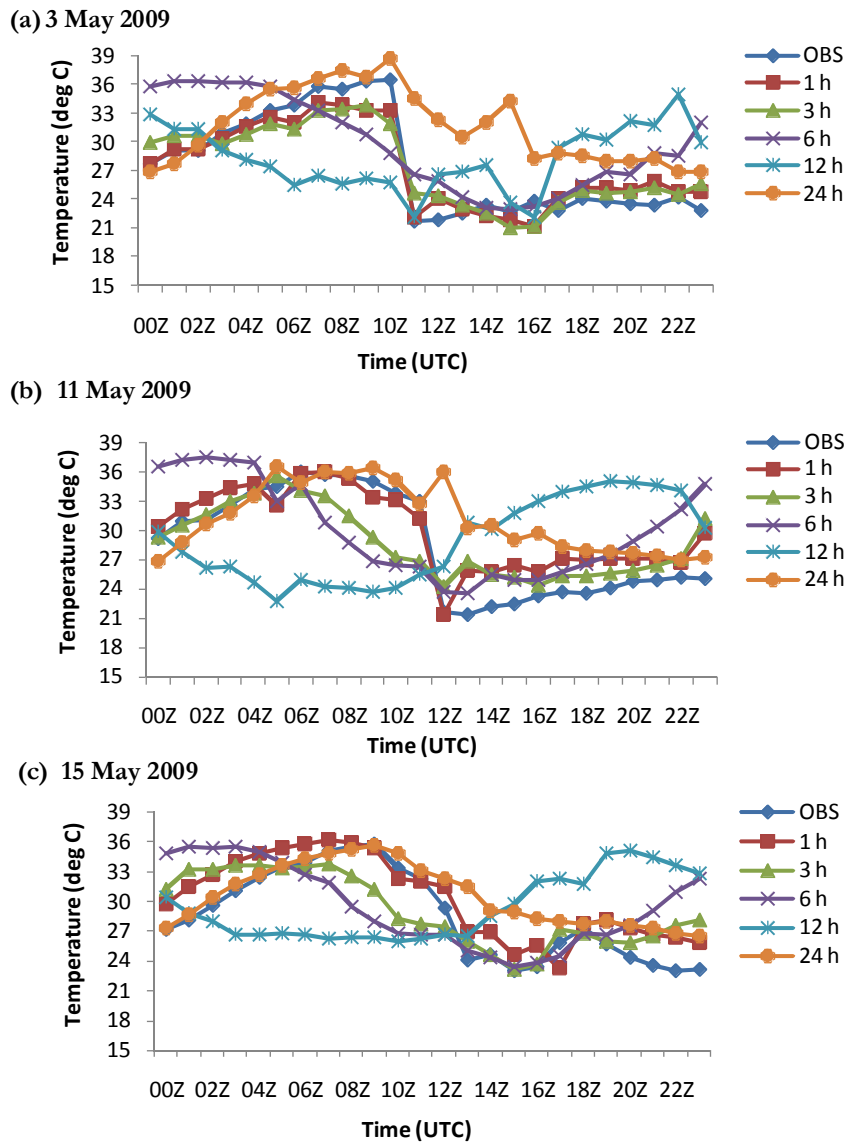
The developed ANN model was applied to derive thunderstorm forecast from 1 to 24 h ahead at Kolkata from the data of 3 consecutive

years (April and May 2007–2009). The ANN models were created to predict surface temperature and relative humidity at hourly intervals with 1, 3, 6, 12 and 24 h ahead during severe thunderstorm cases and the results are evaluated in the following section. Analysis of the results of these experiments are helpful to understand the efficiency of developed ANN model to predict severe thunderstorm events in advance and can apply operationally over east and northeast Indian region.

The comparison between observed and predicted surface temperature for 1 to 24 h advanced forecasting on 3, 11 and 15 May 2009, is shown in Figure 3.7. As seen in the figure, 1 h advanced forecast could forecast quite accurately. It was captured the sudden fall in temperature during thunderstorm hours for all 3 thunderstorm days. The 3 h forecast was in the next position and very close to the observation for the first case. Some deviations are there in the second and third cases. The 6 and 12 h forecast failed to capture the entire pattern. For the first case, (Figure 3.7a), the observed temperature showed a sudden drop of  $15^{\circ}\text{C}$  from  $36.7^{\circ}\text{C}$  to  $21.7^{\circ}\text{C}$  ( $15^{\circ}\text{C}$ ) at 1000 UTC. The 1 h ahead forecast model showed a drop from  $33^{\circ}\text{C}$  to  $22^{\circ}\text{C}$  ( $11^{\circ}\text{C}$ ) at 1000 UTC, whereas 3 h presented a drop from  $32^{\circ}\text{C}$  to  $25^{\circ}\text{C}$  ( $7^{\circ}\text{C}$ ) at 1000 UTC. In the second thunderstorm case (Figure 3.7b), observed temperature fall is from  $33.1^{\circ}\text{C}$  to  $21.7^{\circ}\text{C}$  ( $11^{\circ}\text{C}$ ) at 1200 UTC, whereas 1 h ahead forecast model indicated a drop from  $31^{\circ}\text{C}$  to  $21^{\circ}\text{C}$  ( $10^{\circ}\text{C}$ ) at the same thunderstorm hour. The other models failed to capture a sudden fall during thunderstorm

hour. For the third case (Figure 3.7c), observed temperature showed a drop from 29<sup>0</sup>C to 24<sup>0</sup>C (6<sup>0</sup>C) at 1300 UTC, whereas 1 h advanced prediction model showed a drop from 32<sup>0</sup>C to 27<sup>0</sup>C (5<sup>0</sup>C). The temperature forecast at 24 h ahead model also gave good results as compared to 6 and 12 h ahead forecasts. The hourly temperature variation of 24 h ahead model show a fall with an intensity of 6<sup>0</sup>C at 1600 and 1300 UTC for the first two thunderstorm cases, and 3<sup>0</sup>C at 1400 UTC for the third thunderstorm case. This model captured sudden fall with 6 hour time lag for the first case and 1 hour for second and third cases.

The statistical analyses of the ANN model performance for the advanced prediction of surface temperature during thunderstorm days are given in Table 3.3. Both MAE and RMSE are less for 1 and 3 h advanced prediction and also have a high positive correlation. The errors are high and have low correlation (+ve or -ve correlation) for 6 and 12 h advanced predictions. The 24 h ahead forecast models are better performed than 6 and 12 h ahead prediction and the average correlation is 0.70. The results were highly satisfactory for temperature forecast with 1 to 3 h ahead. The 6 and 12 h ahead forecast accuracy was very poor as compared to 1, 3 and 24 h. The PC of these 5 models are given in Figure 3.9a. The figure clearly indicates that overall accuracy of 1 h ahead forecast for three events is 70%. The 3 and 24 h ahead forecast models are also close to this with 65% and 54%. The other two models (6 and 12 h) displayed less accuracy.



**Figure 3.7:** Comparison of ANN predicted hourly temperature ( $^{\circ}\text{C}$ ) using different advanced prediction models with observation on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.



**Table 3.3:** Performance comparison of different advanced predictions for hourly temperature during thunderstorm days.

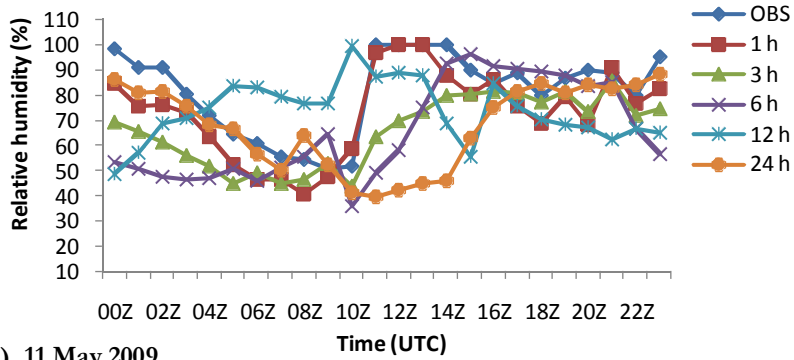
Statistical Analysis	Date	1 h	3 h	6 h	12 h	24 h
MAE	3-May-09	1.34	1.88	4.04	6.06	4.61
	11-May-09	2.10	2.44	4.62	8.47	3.21
	15-May-09	2.07	2.32	3.84	6.54	2.11
	<b>MEAN</b>	<b>1.84</b>	<b>2.22</b>	<b>4.17</b>	<b>7.02</b>	<b>3.31</b>
RMSE	3-May-09	2.36	2.13	4.94	6.81	6.35
	11-May-09	2.42	3.06	5.17	8.97	4.71
	15-May-09	2.29	2.95	4.82	7.19	2.93
	<b>MEAN</b>	<b>2.35</b>	<b>2.71</b>	<b>4.98</b>	<b>7.66</b>	<b>4.66</b>
CC	3-May-09	0.97	0.93	0.63	-0.19	0.53
	11-May-09	0.95	0.82	0.48	-0.80	0.70
	15-May-09	0.94	0.76	0.38	-0.77	0.91
	<b>MEAN</b>	<b>0.95</b>	<b>0.84</b>	<b>0.50</b>	<b>-0.59</b>	<b>0.71</b>

The comparison between observed and predicted relative humidity for 1 to 24 h ahead forecasting on 3, 11 and 15 May 2009 is shown in Figure 3.8. The results show, 1 h ahead forecast captured sudden increase in relative humidity at thunderstorm hours during all three severe thunderstorm cases. The 3 h advanced prediction model was able to predict the rise in relative humidity during thunderstorm hour in the first two cases. The 24 h forecast was also close to the observation for the second thunderstorm case even though one hour time lag exists. The 6 and 12 h forecast failed to capture the entire pattern for all 3 cases as in temperature prediction. In the first case (Figure 3.8a), the observed relative humidity showed a rise of 48% from 52% to 100% at 1000 UTC.

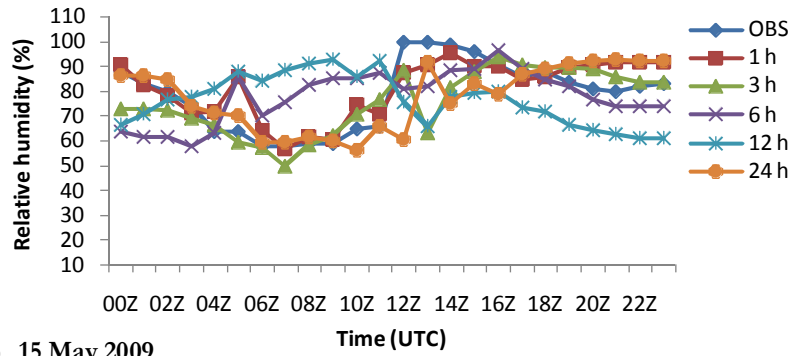
The 1 h advanced ANN model prediction shows a rise from 60% to 97% (37%) at 1000 UTC. The 3 h ahead model showed a rise with an intensity of 20% at 1000 UTC. The other models failed to capture sudden rise as in 1 and 3 h ahead forecast. In the second case (Figure 3.8b), observed relative humidity rise is from 66% to 100% (34%) at 1200 UTC, whereas 1 h ahead model indicated a rise from 70% to 88% (18%) at the same time. The intensity of increase is less (12%) for 3 h ahead model and 6 and 12 h ahead models failed to capture the sudden rise during thunderstorm hour. The 24 h ahead model showed a sudden rise from 61% to 92% (31%) with one hour time lag for this thunderstorm event. For the third case (Figure 3.8c), observed relative humidity showed a rise from 63% to 100% (37%) at 1300 UTC, whereas 1 h ahead model showed a rise from 70% to 92% (22%). The 3 and 24 h ahead models showed an intensity of rise around 3% and 12% respectively. From these analyses of temperature and relative humidity, it can be seen that 1 h advanced prediction model well predicted diurnal variation during thunderstorm days and captured the drop and rise with almost same intensity of observation as compared to other models.

The statistical analyses of the ANN model performance for the advanced prediction of hourly relative humidity during thunderstorm days are given in Table 3.4. Both MAE and RMSE are less for 1, 3 and 24 h advanced prediction. The highest correlation coefficient is for 1 h advanced prediction model (0.80). The 3 and 24 h model have also a + correlation.

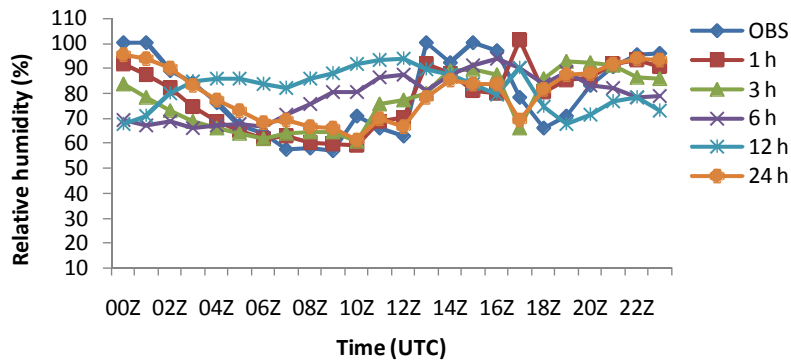
(a) 3 May 2009



(b) 11 May 2009



(c) 15 May 2009



**Figure 3.8:** Comparison of ANN predicted hourly relative humidity (%) using different advanced prediction models with observation on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.

The errors are high and have low correlation for 6 and 12 h advanced prediction models as in the temperature forecast case. The results were satisfactory for relative humidity forecast with 1, 3 and 24 h ahead as in temperature forecast. The PC of these 5 models are given in Figure 3.9b. The figure clearly indicates that overall accuracy of 1 h ahead forecast for three events is 74%. The 3 and 24 h ahead forecast models are also close to this with 61% and 54%. The 6 and 12 h ahead model forecast accuracy was very poor as compared to other three advanced prediction models.

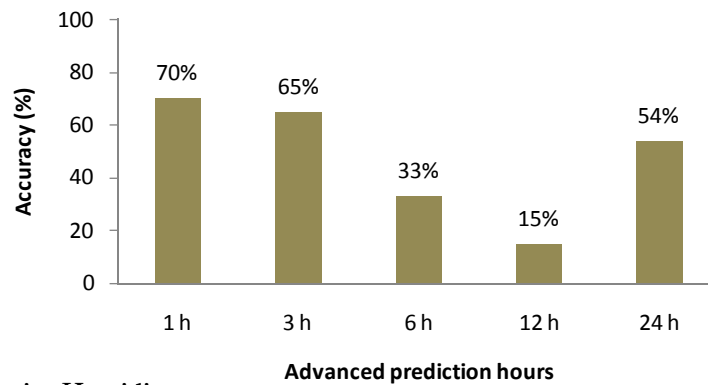
**Table 3.4:** Performance comparison of different advanced predictions for hourly relative humidity during thunderstorm days.

Statistical Analysis	Date	1 h	3 h	6 h	12 h	24 h
MAE	3-May-09	9.08	15.72	19.26	21.39	20.06
	11-May-09	5.96	7.43	13.26	20.04	11.65
	15-May-09	8.03	11.02	14.54	17.40	7.31
	<b>MEAN</b>	<b>7.69</b>	<b>11.39</b>	<b>15.69</b>	<b>19.61</b>	<b>13.00</b>
RMSE	3-May-09	10.63	18.73	25.01	24.59	27.65
	11-May-09	7.73	10.58	15.47	21.63	14.05
	15-May-09	9.98	12.55	16.94	19.67	8.71
	<b>MEAN</b>	<b>9.45</b>	<b>13.95</b>	<b>19.14</b>	<b>21.96</b>	<b>16.80</b>
CC	3-May-09	0.92	0.77	0.40	-0.39	0.45
	11-May-09	0.85	0.70	0.19	-0.65	0.61
	15-May-09	0.76	0.61	0.08	-0.40	0.72
	<b>MEAN</b>	<b>0.84</b>	<b>0.69</b>	<b>0.22</b>	<b>-0.48</b>	<b>0.59</b>

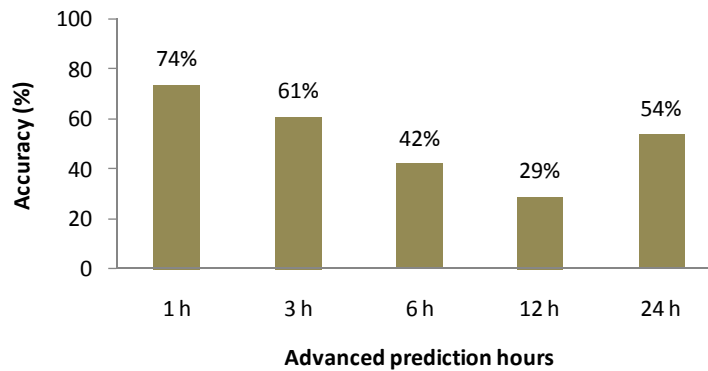
The models developed in this section show how surface temperature and relative humidity can be predicted for 1 to 24 h ahead

with an ANN model. Although the results varied, the 1 and 3 h ahead ANN models were able to predict hourly temperature and relative humidity adequately with sudden fall and rise. Even 24 h advanced prediction model can able to predict features of thunderstorm with reasonable accuracy.

(a) Temperature



(b) Relative Humidity



**Figure 3.9:** Performance accuracy of different advanced prediction models for the prediction of (a) temperature and (b) relative humidity during thunderstorm days.

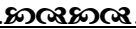
Even though the model performance of 6 and 12 h ahead forecasting was low and the forecasting was not as accurate as expected, the developed model can still be useful in decision making for meteorologists and others who work with real-time thunderstorm forecast.

### **3.5 Chapter Summary**

In this chapter, sensitivity experiments have been conducted with ANN model to test the impact of learning algorithms on severe thunderstorms prediction that occurred over Kolkata on 3, 11 and 15 May 2009 and selected LM algorithm for further studies. The developed ANN model with LM algorithm was applied to derive thunderstorm forecasts from 1 to 24 h ahead at Kolkata. The objective of this study was to use ANNs to predict temperature and relative humidity during thunderstorm days from 1 to 24 h ahead using prior weather data as inputs. A statistical analysis based on MAE, RMSE, CC and PC is also performed for comparison among predicted and observed data with different learning algorithms and advanced predictions.

The model setups were identical except for the use of different learning algorithms for the sensitivity experiments of learning algorithms. Hence the differences in the prediction results attributed to the sensitivity of learning algorithms. It is clearly demonstrated that LM algorithm performance is significantly better than other algorithms. After analyzing the results, it can be concluded that the ANN model with LM algorithm

has well predicted the hourly temperature and relative humidity in terms of sudden fall of temperature and rise of humidity during thunderstorm hours. The ANN models were created to predict surface temperature and relative humidity at hourly intervals with 1, 3, 6, 12 and 24 h ahead during same severe thunderstorm cases. Analysis of the results reveals that, the 1 and 3 h ANN models were able to predict hourly temperature and relative humidity adequately with sudden fall and rise. The efficiency of ANN models were reduced as the forecast lead time increased from 6 to 12 h. The 24 h advanced prediction model can able to predict features of thunderstorm with reasonable accuracy. The results of these analyses demonstrated the capability of ANN model in prediction of severe thunderstorm events over eastern Indian region and will helpful for real time thunderstorm forecast.

..........





## Chapter-4

# WRF-NMM MODEL FOR THUNDERSTORM PREDICTION

---

### Contents

- 4.1 Data and Methodology
  - 4.2 Case Description
  - 4.3 Results and Discussion
  - 4.4 Chapter Summary
- 

Weather forecasting is the application of science and technology to predict the state of the atmosphere for a future time and a given location. Human beings have attempted to predict the weather informally for millennia, and formally since at least the nineteenth century. Weather forecasts are made by collecting quantitative data about the current state of the atmosphere and using scientific understanding of atmospheric processes to project how the atmosphere will evolve. Once an all-human endeavor based mainly upon changes in barometric pressure, current weather conditions and sky conditions, forecast models are now used to determine future conditions. Human input is still required to pick the best possible forecast model to base the forecast upon, which involves pattern recognition skills, teleconnections, knowledge of model performance and knowledge of model biases. The chaotic nature of the atmosphere, the massive computational power required to solve the equations that

describe the atmosphere, error involved in measuring the initial conditions and an incomplete understanding of atmospheric processes mean that forecasts become less accurate as the difference in current time and the time for which the forecast is being made (the range of the forecast) increases (Doswell 2004).

The understanding of the dynamical/physical mechanisms of thunderstorms is essential for improving the forecast of these systems. One of the ways to understand the physics and dynamics of these severe thunderstorms is to simulate these systems with the help of mesoscale models. A number of studies have been carried out (e.g. Brooks and Wilhelmson 1992; Farelly et al. 1992) to simulate thunderstorms for studying various dynamical and physical processes occurring within them. Accurate simulation requires knowledge about “where” and “when” storms will develop and how they will evolve. The high resolution non-hydrostatic mesoscale models with sophisticated parameterization schemes for the important physical processes would be very useful tool for reasonably accurate prediction of these severe thunderstorms (Weiss et al. 2006). However, mesoscale research and forecasting in India could not keep pace with developments of the post-1970 period, especially in respect of mesoscale observational techniques (Doppler Weather Radar (DWR), wind profilers, meso-network), mesoscale analysis and mesoscale NWP (Tyagi 2000). In India, studies related to modeling of clouds are very scarce and in particular intense thunderstorm events (Das

1999). Simulation of severe thunderstorms with high-resolution mesoscale models over Indian region have been attempted recently (Vaidya 2007; Chatterjee et al. 2008; Rajeevan et al. 2010).

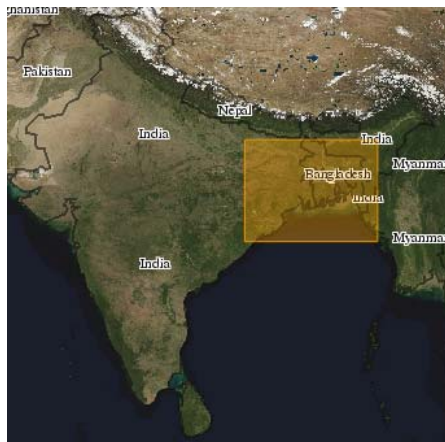
A simulation study was carried out by Vaidya (2007) for a pre-monsoon thunderstorm over east coast of India. In this study, the model performances of ARPS and ARW models are compared by examining the predicted parameters. They found that in case of idealized simulation of thunderstorm, ARPS was able to predict the spatial distribution of rainfall and diverging winds better compared to ARW model. Mesoscale model MM5 with some modifications in the cloud microphysics scheme of Schultz has been used to simulate two hailstorm events over Gangetic Plain of West Bengal by Chatterjee et al. (2008). The authors found that the MM5 model has the ability to simulate hailstorm if the cloud-microphysics scheme of Schultz is modified suitably. Rajeevan et al. (2010) simulated the features associated with a severe thunderstorm event over Gadanki (over southeast India) using WRF model and examined its sensitivity to four different microphysics schemes validated with many observations. This study suggests large sensitivity of the microphysics schemes in the simulations of the thunderstorm. The study also emphasizes the need for a comprehensive observational campaign using multi-observational platforms to improve the parameterization of the cloud microphysics and land surface processes over the Indian region.

Contemporary atmospheric numerical models contain a large number of physical parameterization schemes in order to represent the various atmospheric processes that take place in sub-grid scales. However, modeling systems also reflect inherent errors and uncertainties in specifying the initial state of the atmosphere, and simplifications in physics and parameterization of sub-grid scale processes further contribute to errors in model forecasts. It is believed that physics errors become more important as model resolution increases (e.g. Stensrud et al. 2000; Wandishin et al. 2001). In this chapter, sensitivity experiments have been conducted with the NMM model to examine the impact of different initial conditions, convective parameterization schemes (CPSs) and microphysics schemes in capturing the severe thunderstorm events occurred over Kolkata during 2006, 2007 and 2009. The geographical location of study area is given in the Figure 4.1. The goal of this study is to determine the usefulness of high resolution NMM model when it comes to the severe thunderstorm prediction over east and northeast region of India.

#### **4.1 Data and Methodology**

The NMM core of the WRF system is a next-generation mesoscale forecast model that will be used to advance the understanding and the prediction of mesoscale convective systems. The NMM model has been designed to be an efficient and flexible mesoscale modeling system for use across a broad range of weather forecast and idealized research

applications, with an emphasis on horizontal grid sizes in the range of 1-10 km. This state-of-the-art mesoscale model (WRF-NMMV3.2.1) is used in this study to perform cloud-resolving simulation of the thunderstorm events over Kolkata. The details of NMM model specified in Chapter 2.



**Figure 4.1:** The geographical location of study area.

#### 4.1.1 Initial and boundary conditions

Since NMM is a limited area model, it needs to be fed with initial and lateral boundary conditions for the forecasts. The process of providing initial value data to a model is known as *initialization*. Initial conditions are typically provided by a numerical synthesis of available observations. The initial and boundary conditions for the NMM model are obtained from the NCEP Final Analysis (FNL) datasets. These data are on  $1^\circ \times 1^\circ$  grids (110 x 110 km) (DS-083.2) prepared operationally every six hours (<http://rda.ucar.edu/datasets/ds083.2>). This product is from the

Global Data Assimilation System (GDAS), which continuously collects observational data from the Global Telecommunications System (GTS), and other sources, for many analyses. The FNLs are made with the same model which NCEP uses in the Global Forecast System (GFS), but the FNLs are prepared about an hour or so after the GFS is initialized. The FNLs are delayed so that more observational data can be used. The GFS is run earlier in support of time critical forecast needs, and uses the FNL from the previous 6 hour cycle as part of its initialization. The analyses are available on the surface, at 26 mandatory (and other pressure) levels from 1000 to 10 mb (millibar), in the surface boundary layer and at some sigma layers, the tropopause and a few others. Parameters include surface pressure, sea level pressure, geopotential height, temperature, sea surface temperature, soil values, ice cover, relative humidity, u and v winds, vertical motion, vorticity and ozone. The lists of input meteorological parameters used for NMM model are given in Table 4.1.

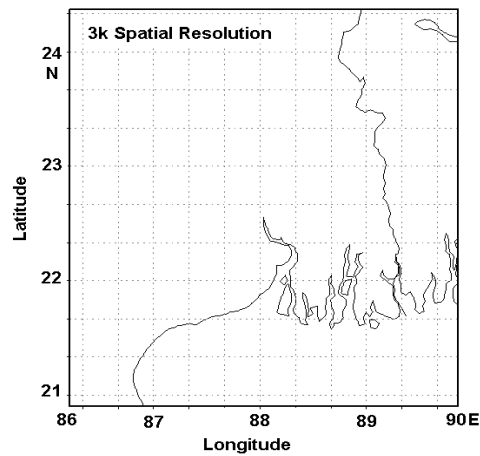
Topography and the surface characteristics for all the grids are driven from the United States Geological Survey (USGS) dataset. This is a global dataset which comprises the land use, vegetation type, vegetation fraction, albedo, topographic heights and other variables that are provided at various resolutions. The dataset with the highest resolution of 30 seconds (approximately 1 km) is used in this research. These datasets are used in the NMM model to define the surface properties.

**Table 4.1:** The input meteorological parameters for NMM model.

Description	Units
Pressure	Pa
Pressure reduced to MSL	Pa
Geopotential height	gpm
Total ozone	Dobson
Temperature	K
Potential temperature	K
Geopotential height anomaly	gpm
u-component of wind	$\text{m s}^{-1}$
v-component of wind	$\text{m s}^{-1}$
Vertical velocity (pressure)	$\text{Pa s}^{-1}$
Absolute vorticity	$\text{s}^{-1}$
Specific humidity	$\text{kg kg}^{-1}$
Relative humidity	%
Precipitable water	$\text{kg m}^{-2}$
Water equivalent of accumulated snow depth	$\text{kg m}^{-2}$
Total cloud cover	%
Land cover	fraction
Ice concentration	fraction
Surface lifted index	K
Best (4 layer) lifted index	K
Vertical speed shear	$\text{s}^{-1}$
Volumetric soil moisture content	fraction
Ozone mixing ratio	$\text{kg kg}^{-1}$
Convective inhibition	$\text{J kg}^{-1}$
Convective available potential energy	$\text{J kg}^{-1}$
Planetary boundary layer height	m
5-wave geopotential height	gpm
5-wave geopotential height anomaly	

#### 4.1.2 Experiment 1- Study with different initial conditions

In this study, the model was integrated with multiple initial conditions starting from 19 May 2006 at 0000 UTC, 19 May 2006 at 1200 UTC and 20 May 2006 at 0000 UTC for a period of 48 h (simulation is designated as Ex-1), 36 h (as Ex-2) and 24 h (as Ex-3) respectively and established the robustness of the results. A single domain with 3 km horizontal spatial resolution was configured as shown in Figure 4.2, which is reasonable in capturing the mesoscale cloud clusters. The dataset used in this research is the  $1^\circ \times 1^\circ$  grid FNL dataset (DS 083.2) and is used to update the boundary conditions every six hours.



**Figure 4.2:** Domain of NMM model.

The domain covers  $86.0^\circ\text{E}$  to  $90.0^\circ\text{E}$  and  $21.0^\circ\text{N}$  to  $24.0^\circ\text{N}$ . The grids are centered at  $88.0^\circ\text{E}$ ,  $22.5^\circ\text{N}$  with 167 X 165 grid points. The domain is configured with vertical structure of 38 unequally spaced sigma



(non-dimensional pressure) levels. The physical parameterizations used in this study are Geophysical Fluid Dynamics Laboratory (GFDL) for longwave and shortwave radiation (Schwarzkopf and Fels 1991; Lacis and Hansen 1974), NMM Land surface scheme (Eket al. 2003) for land surface, Mellor Yamada Janjic (MYJ) scheme (Janjic 2002) for planetary boundary layer, Ferrier scheme (Ferrier 2002) for microphysics, Janjic similarity scheme (Janjic1994) for surface layer and Grell-Devenyi cloud ensemble scheme (Grell and Devenyi 2002) for convective parameterization. All the above schemes are well tested for NMM model and are used operationally at NCEP. Table 4.2 shows the model configuration of the present study. A statistical analysis based on MAE, RMSE and CC is performed for comparisons between the simulated and observed data with different initial conditions.

**Table 4.2:** NMM model configuration.

Description	Parameters
Horizontal-Spatial resolution	3 km
Integration time step	6 second
Map projection	Rotated latitude and longitude
Horizontal grid system	Arakawa E-grid
Vertical co-ordinate	38 sigma levels
Radiation parameterization	GFDL/GFDL
Surface layer parameterization	Janjic similarity scheme
Convective parameterization	Grell-Devenyi ensemble scheme
Land surface parameterization	NMM Land surface scheme
PBL parameterization	Mellor-Yamada-Janjic
Microphysics scheme	Ferrier (new eta) scheme

### 4.1.3 Experiment 2 - Study with different CPSs

The simulations from numerical models are known to be sensitive to the representation of the physical processes. In order to obtain realistic results it is necessary to incorporate appropriate physics into the model. It is believed that physics errors become more important as model resolution increases (Stensrud et al. 2000), such that numerical prediction of precipitation and associated convective processes remain a key challenge. Numerical simulations by means of NMM model with different CPSs have been carried out for the present study. The model was integrated for a period of 24 h, starting at 0000 UTC of 20 May 2006 as initial time for the first case and starting at 0000 UTC of 21 May 2007 as initial time for the second case. A single domain with 3 km horizontal spatial resolution is configured as shown in Figure 4.2, which is reasonable in capturing the mesoscale cloud clusters. Initial conditions for the 3 km domain are derived from 6 h FNL datasets (DS 083.2) at  $1^\circ \times 1^\circ$  grids. The domain is configured with vertical structure of 38 unequally spaced sigma (non-dimensional pressure) levels.

In this study, four simulations have been done for each case by changing the CPSs of the NMM model. The first simulation used the Kain-Fritsch scheme (KF), based on Kain (2004) and Kain and Fritsch (1993). The second simulation used Betts-Miller-Janjic (BMJ) parameterization, which is based on Janjic (1996) and Janjic (2000). The third one used Grell-Devenyi ensemble (GD) parameterization, based on

Grell and Devenyi (2002). Finally, the simulation used simplified Arakawa-Schubert scheme (AS), based on Arakawa and Schubert (1974) as simplified by Grell (1993). In addition, a simulation without a convective scheme (NO) is performed for each case to determine if the model could simulate the convection explicitly. The other physical parameterizations used in this study are same as experiment 1. To compare the differences among the CPSs, simulations are performed for a particular time period utilizing the same initial and boundary conditions and other physical parameterizations for each CPSs and then the model outputs are compared with observation. A statistical analysis based MAE, RMSE and CC is performed for comparisons between the simulated and observed data with different CPSs.

#### **4.1.4 Experiment 3 - Study with different microphysics schemes**

Microphysics schemes are used to handle clouds and precipitation in mesoscale models. The impact of cloud microphysics on cloud resolving simulations is an important issue in NWP and regional climate modeling. Because of the wide variety of cloud microphysics schemes currently being used, a natural concern is the sensitivity of the prediction and simulation of precipitation in high-resolution numerical models to the microphysics parameterization. Another important issue is whether realism is consistently gained with increasingly sophisticated cloud microphysics. The dependency of model results on microphysics parameterizations has traditionally been addressed through short-term

idealized simulations of a single precipitation system (e.g. a thunderstorm or a squall line) (Otkin and Greenwald 2008). Sensitivity experiments have been conducted for a severe thunderstorm event on 15 May 2009 with three microphysics schemes namely Ferrier (FERR) (Ferrier et al. 2002), WRF Single Moment 6 class (WSM6) (Hong et al. 2004) and Thomson scheme (THOM) (Thompson et al. 2004).

In the present study, the model was integrated for a period of 24 h. A single domain with 3 km horizontal spatial resolution was configured. Initial conditions for the 3 km domain are derived from 6 h FNL data (DS 083.2) at  $1^\circ \times 1^\circ$  grids. The domain is configured with vertical structure of 38 unequally spaced sigma (non-dimensional pressure) levels. The physical parameterization schemes other than microphysics scheme used in this study are same as experiment 1. In all experiments, the model setups were identical except for the use of different microphysics schemes. The model results are analyzed and compared to the available surface observations in order to identify the parameterizations that provide the best representation of the spatio-temporal variability of thunderstorm affected parameters.

#### **4.1.5 Observational data**

A brief description of the observational data used for the present study is given in this section. In particular, DWR reflectivity images, surface observations from Automatic Weather Station (AWS) records and Automatic Rain Gauge (ARG) station data were used for validating the

model results. The DWR was installed over Kolkata (22.57<sup>0</sup>N, 88.35<sup>0</sup>E) in April 2002. It has a beam width of 1 degree and nominal range of about 450 km for reflectivity and 250 km for radial velocity as well as spectrum width. Maximum unambiguous radial velocity estimates following unfolding techniques is about 64 ms<sup>-1</sup> (approximately 220 kmph) with a radial resolution of 1 km and 1 degree in azimuth. Reflectivity factors (Z), radial velocity (V) and velocity spectrum width (W) are the three base data directly observed/measured by the radar. It is capable of monitoring clouds, precipitation systems and winds over large areas of more than 400 km from the radar location. DWR has the unique capability to continuously track and predict fast evolving weather systems such as thunderstorms, cyclones and cloudbursts. The detailed description of the radar was given by Bhatnagar et al. (2003). The DWR reflectivity factors Z (dBZ) over Kolkata during pre-monsoon thunderstorms are taken for the present study.

IMD has augmented its AWS network under its modernization programme considering its utility in monitoring and predicting weather events. The network is being expanded with additional 550 AWS and 1350 ARG stations. At present, there are about 524 AWS and 456 ARG operationally working all over India. Each ARG Station is configured to measure hourly and cumulative rainfall for the day (Mohapatra et al. 2010). The numbers of AWS stations are increased to 17 in the West Bengal region by the end of 2010. The surface observations from AWS

(Kolkata) and ARG data obtained during pre-monsoon thunderstorm days are used for validation. The weather charts include Daily Northern Hemispherical Analysis Centre (NHAC) operational charts at surface, 925, 850, 700, 500, 300 and 200 hPa. The other observations which are used for the present studies are past weather reports, particularly of thunderstorms/thundershowers and associated severe weather phenomena such as squalls/hail storm etc.

## 4.2 Case Description

The occurrence of pre and post monsoon thunderstorms over Indian continent is a special feature. Thunderstorms are associated with heavy rainfall during short duration of 2–3 hours. For the present study three severe thunderstorm cases of 20 May 2006, 21 May 2007 and 15 May 2009 have been taken and the description of each case is as follows.

Case 1 was a severe thunderstorm, which was reported on 20 May 2006 at 1200 UTC over Kolkata. This intense convective event produced 52 mm rainfall over Kolkata. The weather situation started with a squall passing Kolkata airport on 20 May 2006 at 1100 UTC with a maximum speed of  $19 \text{ ms}^{-1}$  lasting for a few minutes. A few places recorded moderate rainfall over GWB and isolated rainfall over Orissa, Chattisgarh and Bihar. Dum Dum recorded 50 mm and Alipore 40 mm of rainfall (Mohanty et al. 2006). Case 2 was another severe thunderstorm occurred over Kolkata on 21 May 2007 at 1100 UTC. A squall was reported over

Kolkata at 1100 UTC from northwesterly direction with max speed of  $19 \text{ ms}^{-1}$  lasted for 1 minute. This convective event produced 20 mm rainfall over Kolkata. A few places recorded moderate rainfall over GWB and isolated rainfall over Orissa, Bihar and Jharkhand (Mohanty et al. 2007). The descriptions of third thunderstorm case (15 May 2009) are already specified in Chapter 3.

### **4.3 Results and Discussion**

Today there are a number of parameters available that may be used to characterize pre-convective conditions and predict the beginning of convection. Johns and Doswell (1992) reviewed severe thunderstorm and tornado forecasting in detail. According to him, three of the most important factors to examine in determining occurrence of severe thunderstorm events are intense instability, a sufficiently deep humid layer in the lower and middle troposphere and an updraft to initiate convection. The formation of thunderstorms is an interaction between these conditions on different scales. The occurrence and intensity of severe thunderstorms are examined in the following sections by the analysis of observed and model simulated results.

#### **4.3.1 Sensitivity study with different initial conditions**

Atmospheric flow is sensitively dependent on initial conditions. Initial conditions involve the same model, with the same forcing, run from variety of different start dates. Because the weather system is

chaotic, tiny changes in things such as temperatures, winds, and humidity in one place can lead to very different paths for the system as a whole. In this study, an attempt has been made to understand the relative role of initial conditions for thunderstorm predictability. The model simulated results with different initial conditions are explored in the following section. Analysis of the results of these experiments is helpful to understand the impact of different initial conditions on the simulation of 20 May 2006 severe thunderstorm event and assist in the customization of model for future severe thunderstorm simulations over east Indian region. The storm initiation and development is examined by the analysis of stability indices, surface temperature and relative humidity, 24 h accumulated rainfall and composite radar reflectivity.

#### ***4.3.1.1 Stability indices***

The formation, intensification and propagation of thunderstorms are mostly governed by the synoptic situation and localized dynamic and thermodynamic conditions of the atmosphere. Stability indices have been a corner stone in the forecasting of convection for many decades and often are used in the research literature as well. These indices are very helpful in predicting thunderstorms. To obtain a quick check on thunderstorm possibility, various thunderstorm indices and parameters have been developed. The indices are having critical values and above these critical values, there is possibility of thunderstorm. Studies on the efficiency of different stability indices for the thunderstorm prediction



have been made by several authors (Schultz 1989; Jacovides and Yonetani 1990). Jacovides and Yonetani (1990) found that combining a thermodynamic index with a kinematic parameter improved the skill of forecasting non-frontal thunderstorms in Cyprus. Advection of warm air in the lower levels and cold air in the upper levels (generally associated with deep troughs in upper tropospheric westerlies) increases the conditional instability in the atmosphere and favor outbreak of severe thunderstorms in Kolkata region (Alvi and Punjabi 1966; Rao et al. 1971). Mukhopadhyay et al. (2003) worked on objective forecast of thundery/non-thundery days using conventional indices over three northeast Indian stations. Tyagi et al. (2011) studied some thermodynamic indices for the prediction of thunderstorm occurrence over Kolkata during April and May 2006–2008 and proposed suitable threshold values in forecasting these thunderstorms.

Most of the indices analyzed describe the stability of the atmosphere, as opposed to shear or moisture. They can be determined by mathematical formulae or by plotting on a skew-T/log-p diagram. Often a certain threshold value is defined above (below) which the possibility of thunderstorms is considered. Table 4.3 shows the selected indices for severe convective weather forecasting and their critical values for severe thunderstorms suggested by Air Weather Service (AIWS) Technical Report (1990) and Tyagi et al. (2011) for Kolkata. In this study, an attempt is made to examine different stability indices namely Convective

Available Potential Energy (CAPE), Lifted Index (LI), Total Total Index (TTI) and K Index (KI) obtained from NMM model during these severe local storm days over Kolkata (22.52<sup>0</sup>N, 88.37<sup>0</sup>E). The CAPE represents the amount of buoyant energy available to accelerate a parcel vertically and a CAPE value greater than 1500 Jkg<sup>-1</sup> is suggested by Rasmussen and Wilhelmson (1983) as being necessary for super-cells to form. Brooks et al. (2003) demonstrated that threat of significant severe convective weather increases with increasing CAPE. Table 4.4 shows the NMM model simulated stability indices over Kolkata at 1200 UTC using different initial conditions. It can be seen from table that all 3 experiment results show a high value (2128, 2150 and 1909 Jkg<sup>-1</sup>) during the thunderstorm hour, which is a favorable condition for severe thunderstorms. The CAPE values of three experiments are greater than critical levels (1500 and 1000 Jkg<sup>-1</sup>).

The LI measures the difference between a parcel's temperatures compared with the environmental temperature at 500 hPa, after the parcel has been lifted from the Lifting Condensation Level (AiWS Technical Report 1990). The LI has proved useful for indicating the likelihood of severe thunderstorms. The chances of a severe thunderstorm are best when the LI is less than or equal to -3. This is because air rising in these situations is much warmer than its surroundings and can accelerate rapidly and create deep and violent thunderstorms. Tyagi et al. (2011) suggested the same critical level (-3) for thunderstorms over Kolkata

region. The NMM model simulated LI with different initial conditions show a value of -5 at 1200 UTC for these three thunderstorm cases, which is a favorable environment for thunderstorm occurrence (Table 4.4).

**Table 4.3:** The different stability indices and their critical values for severe thunderstorm.

Stability Indices	Description	Critical values	For Kolkata (Tyagi et al. 2011)
Lifted Index (LI)	$T_{500} - T_{parcel}$	$\leq -3$	$\leq -3$
K Index (KI)	$(T_{850} - T_{500}) + Td_{850} - (T_{700} - DT_{700})$	$> 33$	$> 24$
Total Total Index (TTI)	$(T_{850} + Td_{850}) - 2(T_{500})$	$> 44$	$> 46$
CAPE	$\int_{z_f}^{z_n} g \left( \frac{T_{parcel} - T_{env}}{T_{env}} \right) dz$	$> 1500$	$> 1000$

**Table 4.4:** NMM model simulated stability indices over Kolkata at 1200 UTC using different initial conditions.

Stability Indices	Ex-1	Ex-2	Ex-3
CAPE	2128	2150	1909
LI	-5	-5	-5
KI	36	34	35
TTI	45	42	46

The KI is a combination of the Vertical Totals (VT) and lower tropospheric moisture characteristics. The VT is the temperature difference between 850 and 500 hPa, while the moisture parameters are the dew point depression at 850 and 700 hPa. The KI has proved useful in

indicating the probability of severe thunderstorms. As the KI increases, so does the probability of having a severe thunderstorm. Although KI values can be correlated to a probability of thunderstorm occurrence, these values will vary with seasons, locations, and synoptic settings. The values between 31 to 35 shows 60-80% chance for thunderstorm occurrence and the values more than that will give 80-100% confidence (AiWS Technical Report 1990). Tyagi et al. (2011) suggested a critical level for KI, which is more than 24 over Kolkata. The model simulated KI values of all 3 experiments are more than the critical level suggested by AiWS Technical Report (1990) and Tyagi et al. (2011) (Table 4.4). Miller (1972) introduced the TTI for identifying areas of potential thunderstorm development. It accounts for both static stability and the presence of 850 hPa moisture. A TTI of greater than 44 indicates favorable conditions for development of severe thunderstorms (AiWS Technical Report 1990). Higher values for TTI are associated with a greater probability of thunderstorms. The TTI value of Ex-3 is equal to 46, which is a favorable condition. The TTI value (45) of Ex-1 is also greater than the critical level suggested by AiWS Technical Report (1990). Ex-2 simulated TTI value (42) is less than both critical levels. Examination of the model simulated stability indices with different initial conditions clearly indicated that the NMM model well captured the instability of the atmosphere at 1200 UTC for the occurrence of a severe thunderstorm. Thus the thermodynamic

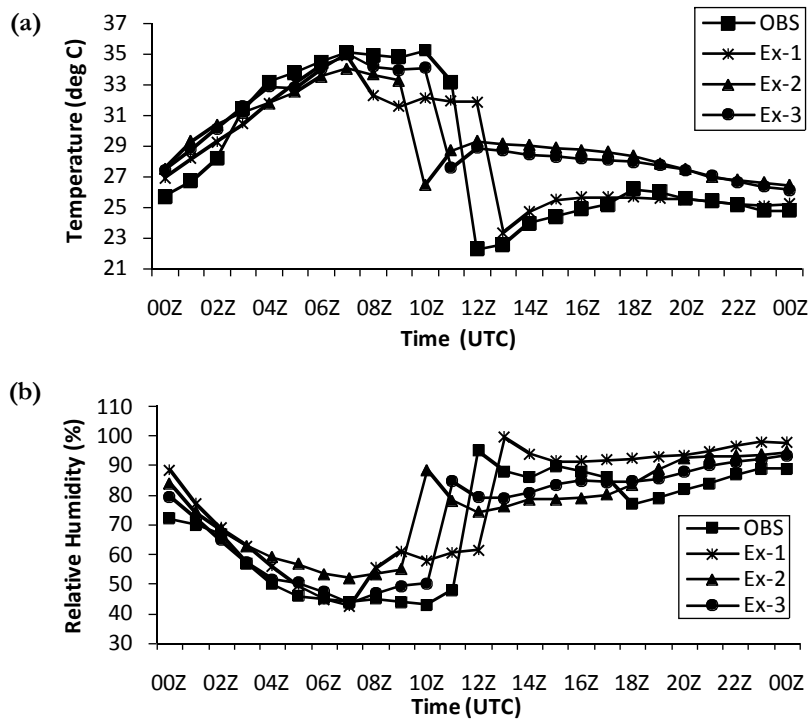
structure over Kolkata becomes conducive for a thunderstorm occurrence on the evening of 20 May 2006.

#### ***4.3.1.2 Surface parameters***

The initiation and intensification of this severe thunderstorm is examined by the analysis of surface parameters namely precipitation, surface temperature and relative humidity. Surface temperature and relative humidity are useful parameters in determining the likelihood occurrence of a thunderstorm. Figure 4.3a shows the hourly variation of model simulated surface temperature at different initial conditions with the observed values over Kolkata (AWS data), from 20 May 2006 at 0000 UTC to 21 May 2006 at 0000 UTC. The observed temperature showed a sudden drop from 33<sup>0</sup>C to 22<sup>0</sup>C at 1200 UTC. Ex-3 captured the variation with drop in temperature at 1100 UTC from 33.5<sup>0</sup>C to 27<sup>0</sup>C, one hour before the time of thunderstorm, which could be attributed to the cooling of the surface temperature due to precipitation by the thunderstorm system. Ex-2 has captured the temperature drop at 1000 UTC from 33<sup>0</sup>C to 27<sup>0</sup>C, which is two hours prior to the actual occurrence. Ex-1 showed a dip at 1300 UTC from 32<sup>0</sup>C to 23<sup>0</sup>C.

Figure 4.3b shows the inter-comparison of observed and model simulated relative humidity using different initial conditions over Kolkata valid for 20 May 2006 at 0000 UTC to 21 May 2006 at 0000 UTC. The observed relative humidity values peaked from 48% to 95% at 1200 UTC

whereas Ex-3 showed a sharp rise from around 50% to 85% at 1100 UTC, which is one hour prior to the thunderstorm occurrence. Ex-2 showed a sudden rise from 55% to 88% at 1000 UTC whereas Ex-1 showed a peak from 61.5% to 99.6% at 1300 UTC. Ex-3 has well captured the time and intensity of temperature and relative humidity as compared to other experiments.



**Figure 4.3:** The inter-comparison of observed and model simulated (a) surface temperature ( $^{\circ}\text{C}$ ) and (b) relative humidity (%) with different initial conditions over Kolkata valid for 20 May 2006 at 0000 UTC to 21 May 2006 at 0000 UTC.

Verifying forecasts with statistical analysis measures how the values of the forecasts differ from the values of the observations. A statistical analysis based on MAE, RMSE and CC was performed for comparisons between the simulated and observed surface parameters namely temperature and relative humidity over Kolkata valid for 20 May 2006 at 0000 UTC to 21 May 2006 at 0000 UTC and are given in Table 4.5. The CC gives us a numerical measure of correlation. From the table, it can be clearly seen that, simulated temperature and relative humidity of Ex-3 have well correlated to the observation as compared to all other two experiments (Ex-1 and Ex-2). The CC of both parameters are 0.87 for Ex-3, which is a strong correlation. When verifying deterministic forecasts, two commonly used error statistics to measure quantitative accuracy are the MAE and RMSE. The MAE and RMSE of each weather parameters have been computed to test the prediction error of the model. The MAE and the RMSE can be used together to diagnose the variation in the errors in a set of forecasts. The MAE and RMSE values of temperature in Ex-3 are 1.62<sup>0</sup>C and 2.28<sup>0</sup>C respectively. The same errors of relative humidity in Ex-3 are 7.55% and 10.70%. The errors are less in Ex-3 as compared Ex-1 and Ex-2.

Precipitation is recognized as one of the most difficult parameters to forecast in NWP despite the fact that the accuracy of numerical models has increased during the past several decades (Wang and Seaman 1997). Accurate estimates of precipitation at both temporal and spatial

resolutions are required for many applications. Figure 4.4 shows the comparison of observed and model simulated accumulated progressive rainfall with different initial conditions at Kolkata.

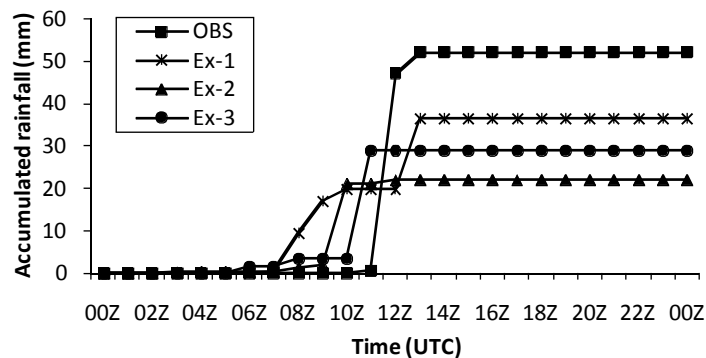
**Table 4.5:** Statistical analysis of simulated and observed surface temperature and relative humidity over Kolkata based on MAE, RMSE and CC.

Statistical Analysis	Parameters	Ex-1	Ex-2	Ex-3
MAE	Temperature	2.19	2.82	1.62
	Relative humidity	10.95	9.66	7.55
RMSE	Temperature	2.83	3.51	2.28
	Relative humidity	14.14	11.86	10.70
CC	Temperature	0.82	0.74	0.87
	Relative humidity	0.72	0.82	0.87

Ex-3 is able to capture 29 mm of rainfall at 1100 UTC, which is less compared to actual observation (52 mm). Ex-2 is only able to simulate 22 mm at 1000 UTC, which is very less. The total accumulated rainfall of Ex-1 is 36.6 mm. But, the rainfall amount during thunderstorm hour (between 0900 and 1200 UTC) is only 16.6 mm. The results show the total rainfall between thunderstorm hours is well captured by Ex-3. The model simulated spatial distribution of 3 h accumulated rainfall between 0900 and 1200 UTC under different initial conditions is shown in Figure 4.5. From the spatial pattern of rainfall, it can be clearly seen that, the rainfall amount and spread are well captured by Ex-3 than other two experiments. Ex-2 is also captured rainfall intensity and distribution



over the study area. Ex-1 has failed to capture the intensity between thunderstorm hours.

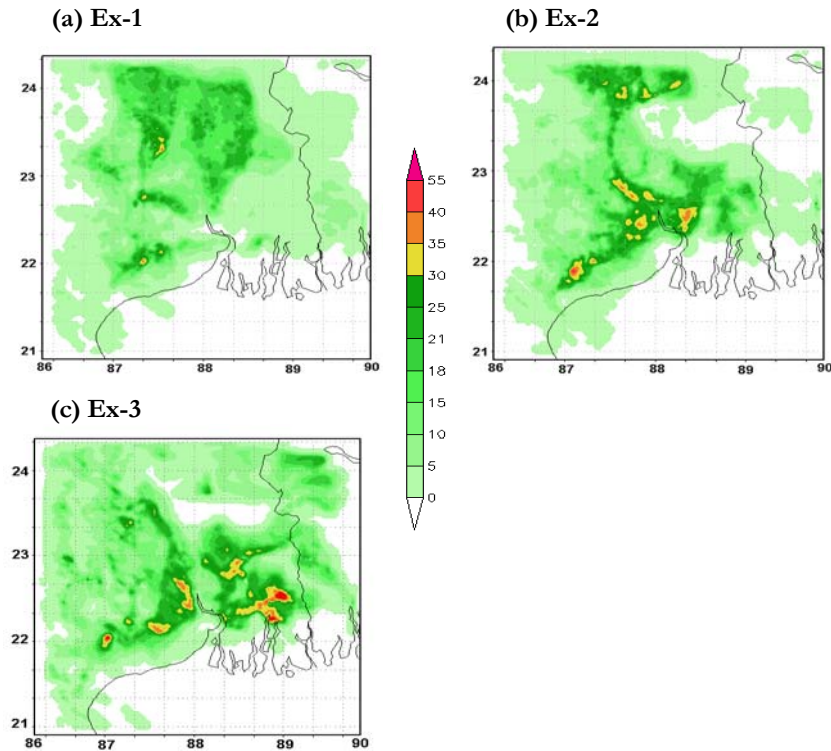


**Figure 4.4:** The inter-comparison of observed and model simulated diurnal variation of 24h accumulated rainfall (mm) with different initial conditions over Kolkata valid on 20 May 2006.

#### 4.3.1.3 Composite radar reflectivity

The use of composite radar reflectivity fields as a model output product has become increasingly popular recently as a means for display of high-resolution numerical model fields, mainly for convective weather scenarios. The reflectivity product offers significant advantages over traditional precipitation forecast displays, including the obvious fact that radar reflectivity is easier to verify in real time by directly comparing with readily available, observed reflectivity products. The chief advantage of the model reflectivity product appears to be that it allows one to more easily see detailed mesoscale and near-storm scale structures capable of being simulated by finer resolution models, such as the structure of deep

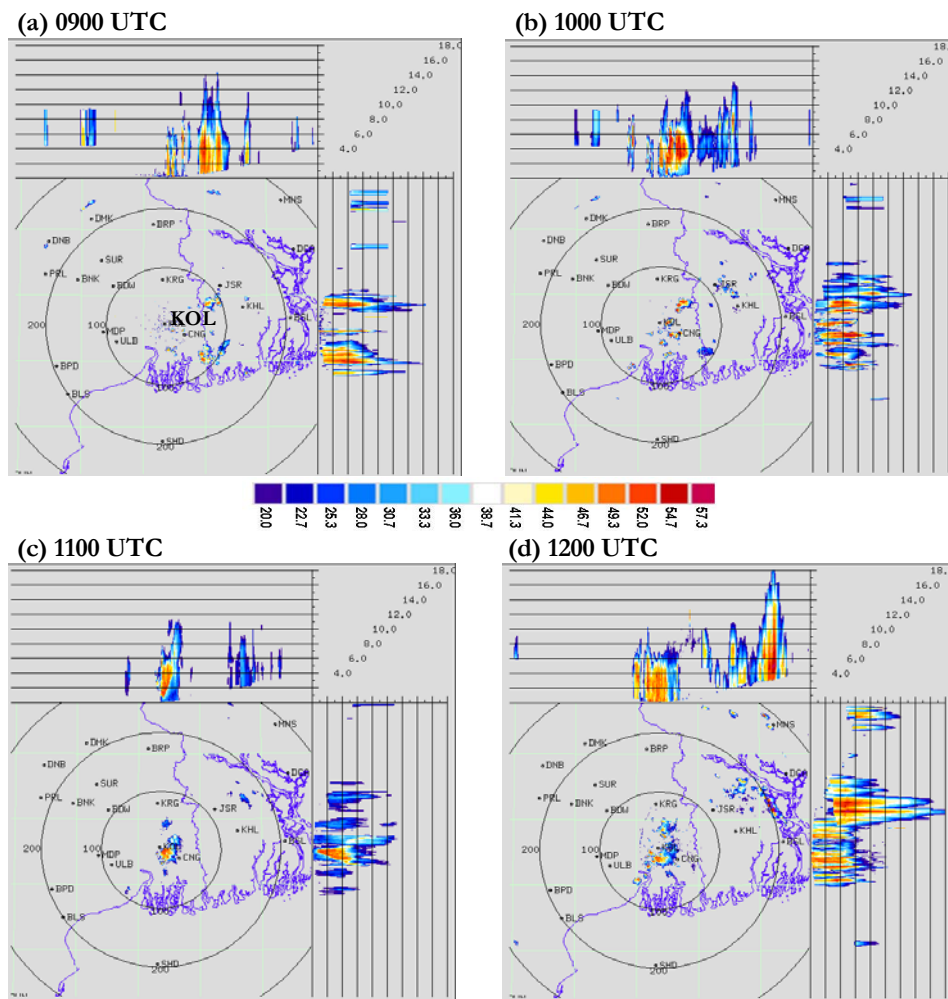
convection, movement of squall line and frontal precipitation bands (Koch et al. 2005).



**Figure 4.5:** The 3 h accumulated rainfall (mm) with different initial conditions over Kolkata valid for 20 May 2006 at 0900 UTC to 1200 UTC.

Recently, installation DWR has highlighted the better prospect of mesoscale prediction in the nowcast to very short-range time scale over Indian region (e.g. Srivastava et al. 2010). Kolkata DWR composite radar reflectivity imageries on 20 May 2006 from 0900 to 1200 UTC is shown in Figure 4.6. From the DWR products it can be seen that scattered echoes

developed northeast of Kolkata (KOL) at 0900 UTC and moved towards Kolkata at 1000 UTC. This echo was intensified and over Kolkata at 1100 UTC. This echo disappeared at 1300 UTC (Mohanty et al. 2006).

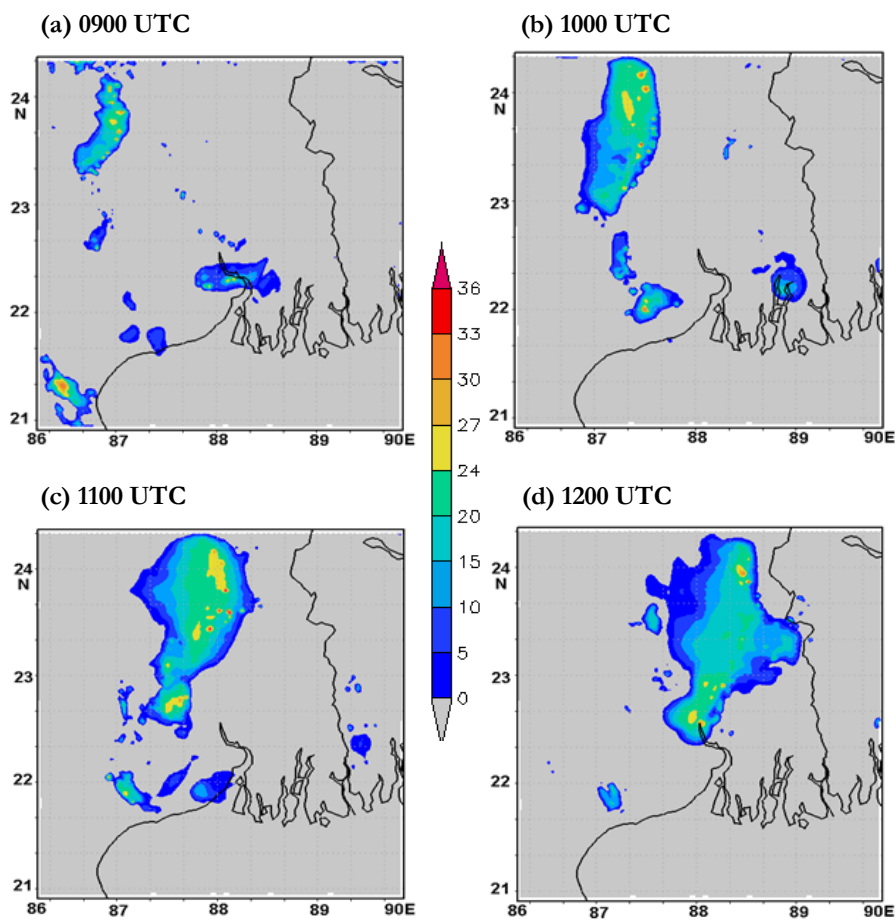


**Figure 4.6:** Kolkata DWR composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 20 May 2006.

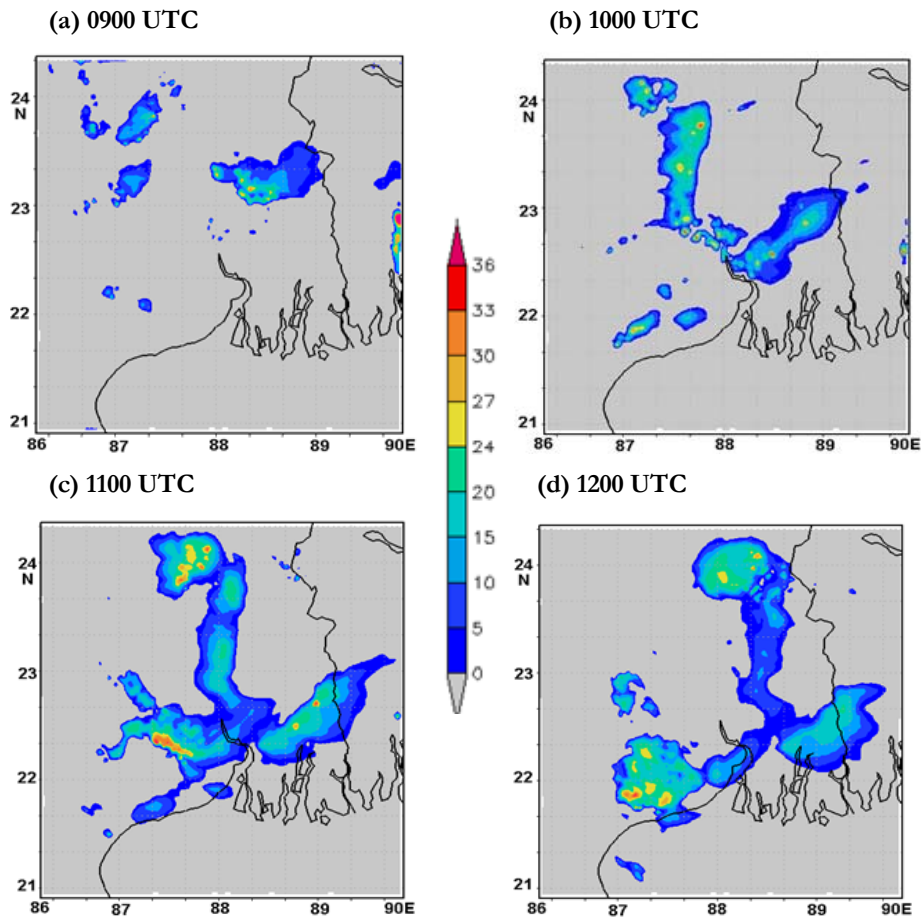
The storm structure and movement can be evaluated by comparing the modeled radar reflectivity to the observed radar reflectivity. In order to achieve that composite radar reflectivity (dBZ) calculated in different initial conditions are compared with observed one. NMM model simulated composite radar reflectivity on 20 May 2006 from 0900 to 1200 UTC with Ex-1 is shown in Figure 4.7. Same for Ex-2 and Ex-3 are shown in Figure 4.8 and 4.9 respectively. By analyzing NMM model simulated composite radar reflectivity pictures of Ex-1 (Figure 4.7), scattered echoes are developed northwest of Kolkata at 0900 UTC. This echo moved towards Kolkata at 1000 and 1100 UTC. This echo was nearby Kolkata by 1200 UTC. The movement of the system (northwest to southeast) in Ex-1 is opposite to that of the observed radar reflectivity (northeast to southwest), which is shown in Figure 4.6. By analyzing NMM model simulated composite radar reflectivity pictures of Ex-2 (Figure 4.8), scattered echoes are developed northeast of Kolkata at 0900 UTC. One more echo developed and intensified at 1000 UTC over northwest of Kolkata. The first echo moved southwest direction and the second one moved eastwards by 1100 UTC. The movement of first echo was faster than the observation (Figure 4.6).

By analyzing NMM model simulated composite radar reflectivity pictures of Ex-3 (Figure 4.9), scattered echoes are developed northeast of Kolkata at 0900 UTC. This echo was moving southwestwards at 1000 UTC and intensified at 1100 UTC. This echo moved further in southwest

direction at 1200 UTC. The NMM model with Ex-3 simulated this squall line movement by simulated composite radar reflectivity fields as observed by DWR imageries. The squall line movement in DWR pictures is well matching with Ex-3 than Ex-2 and Ex-1.



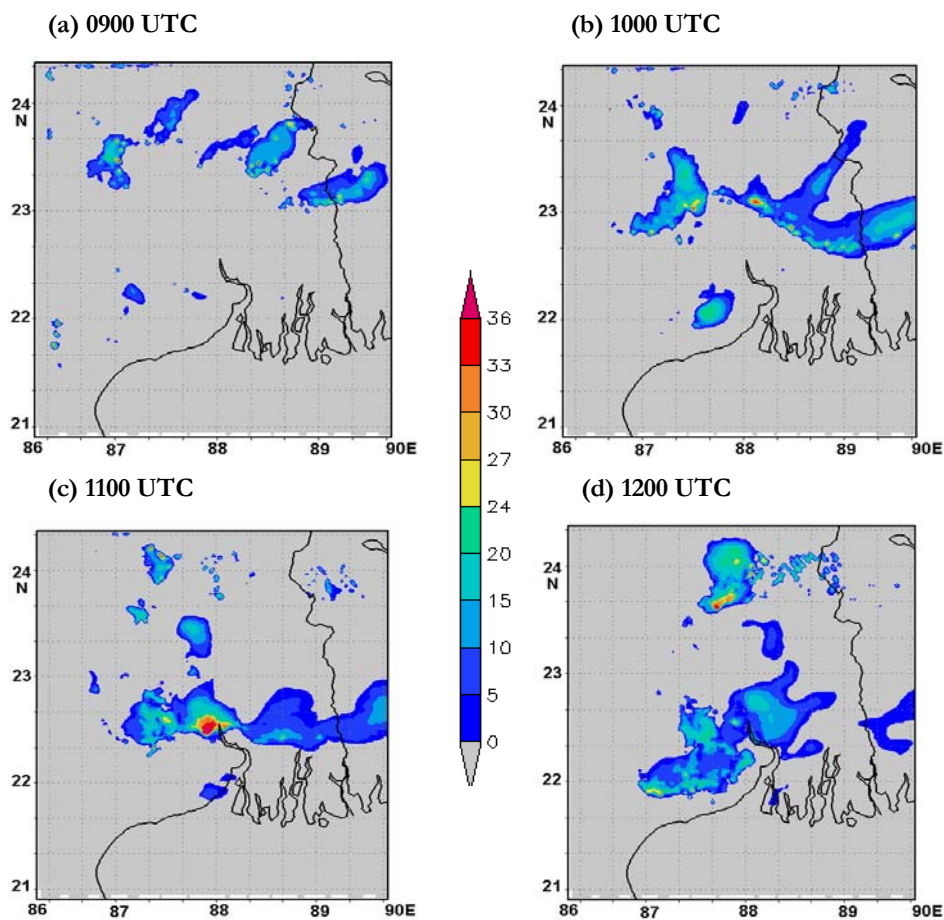
**Figure 4.7:** NMM simulated composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 20 May 2006 with Ex-1.



**Figure 4.8:** NMM simulated composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 20 May 2006 with Ex-2.

The trends shown by various meteorological fields of Ex-3 are in good agreement with each other and very much consistent with dynamic and thermodynamic properties of the atmosphere for the occurrence of a severe thunderstorm. Ex-2 has also well captured all the meteorological

parameters with two-hour time lag. Ex-1 is failed to capture the direction of squall line movement. It may also be noted that all the characteristic properties of the genesis, occurrence and life cycle of severe thunderstorm were well simulated by the Ex-3.



**Figure 4.9:** NMM simulated composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 20 May 2006 with Ex-3.

### 4.3.2 Sensitivity study with different CPSs

The convective processes are implemented in NWP models through parameterizations because they are not resolved in the grid systems of most large scale and mesoscale models. The CPSs are procedures that attempt to account for the collective influence of small-scale convective processes on large-scale model variables. They are representing dynamic and thermodynamic processes of moist convection occurring at sub-grid scales. No universal framework exists for CPSs, which led to the development of numerous different schemes. Properly parameterizing the effects of convection is still a challenging problem for NWP. The model simulated results with different CPSs for two severe thunderstorm cases (20 May 2006 and 21 May 2007) are explored in the following section. Analysis of the results of these experiments are helpful to understand the impact of CPSs on the simulation of severe thunderstorm events and assist in the customization of NMM model for future severe thunderstorm simulations over east Indian region.

#### 4.3.2.1 Stability indices

Table 4.6 shows the inter-comparison of model simulated stability indices with different CPSs over Kolkata valid for 20 May 2006 at 1200 UTC (Case 1) and 21 May 2007 at 1100 UTC (Case 2). The results of first thunderstorm case show that, only GD scheme is able to simulate a high CAPE value ( $1909 \text{ Jkg}^{-1}$ ) during the thunderstorm hour, which is a



favorable condition for severe thunderstorms. The CAPE value of NO scheme ( $1433 \text{ Jkg}^{-1}$ ) is also close to the critical level suggested by AiWS Technical Report (1990) and greater than the value suggested by Tyagi et al. (2011) for Kolkata. But CAPE values of KF and AS schemes are less than  $1500 \text{ Jkg}^{-1}$  and greater than  $1000 \text{ Jkg}^{-1}$  during thunderstorm hour. The BMJ simulated CAPE is very less ( $983 \text{ Jkg}^{-1}$ ) as compared to all other CPSs. The GD scheme captured the lowest LI value (-5) compared to all other CPSs as in CAPE (Table 4.6). The NO scheme simulated LI is -4. The AS and KF simulated LI (-3) is equal to the critical level. The BMJ simulated LI is -2, which is higher than the critical value and not favorable for thunderstorm occurrences. The model simulated KI values with different CPSs are more than the critical level suggested by AiWS Technical Report (1990) and Tyagi et al. (2011), which is favorable for severe thunderstorm. All the CPSs are able to capture a TTI of greater than or equal to 44. The GD and KF simulated TTI is 46 which is equal to the critical value suggested by Tyagi et al. (2011) (Table 4.3). The BMJ and AS are captured the least value (44) compared to other CPSs and equal to the critical level suggested by AiWS Technical Report (1990). By comparing all the stability indices with critical level (Table 4.3), it can be concluded that all the CPSs are well simulated the overall pattern except BMJ scheme. The GD scheme simulated stability indices are well shown the instability of the atmosphere at 1200 UTC for the occurrence of a severe thunderstorm.

The results of second thunderstorm case show that, GD, AS and BMJ schemes are able to capture a high CAPE value (4413, 3742, 2128  $\text{Jkg}^{-1}$  respectively) during the thunderstorm hour, which is a favorable condition for severe thunderstorms. The other schemes namely KF and NO (1071, 1034  $\text{Jkg}^{-1}$  respectively) are not able to capture a value greater than the critical level (1500  $\text{Jkg}^{-1}$ ) suggested by AiWS Technical Report (1990), but the values are more than 1000  $\text{Jkg}^{-1}$ , which is the suggested critical value by Tyagi et al. (2011). The GD scheme simulated LI is -8, which is the lowest value among all other CPSs. The AS and BMJ schemes are also able to capture a low value (-7 and -5) during the thunderstorm hour. The NO scheme simulated LI is -4. The KF simulated LI is equal to the critical level (Table 4.6). The NMM model simulated KI values with different CPSs are very high except BMJ scheme. BMJ scheme simulated KI value (26) is less than the critical level suggested by AiWS Technical Report (1990), but greater than the critical value (24) suggested by Tyagi et al. (2011) over Kolkata (Table 4.3). All the CPSs are able to capture a TTI greater than or equal to 46, which is a favorable condition for severe thunderstorms. The AS scheme simulated TTI is the highest among all other schemes (48). The GD and KF simulated TTI is 47. The BMJ scheme is also captured a high value which is equal to 46 (Table 4.6). The NO scheme is failed to capture a value greater than or equal to 46 as suggested by Tyagi et al. (2011), but greater than the value suggested by AiWS Technical Report (1990). By comparing all the

stability indices of different CPSs, it can be concluded that GD, AS, and BMJ schemes are well simulated the stability indices which is shown the instability of the atmosphere at 1100 UTC for the occurrence of a severe thunderstorm.

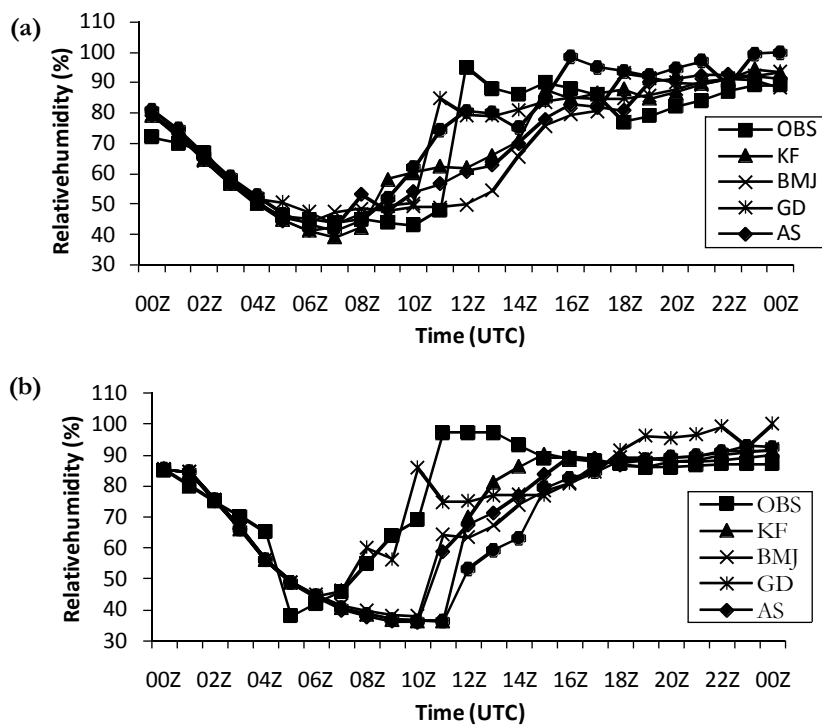
**Table 4.6:** The inter-comparison of model simulated stability indices with different CPSs over Kolkata valid for 20 May 2006 at 1200 UTC (Case 1) and 21 May 2007 at 1100 UTC (Case 2).

Stability Indices	Cases	KF	BMJ	GD	AS	NO
CAPE	Case 1	1215	983	1909	1244	1433
	Case 2	1071	2128	4413	3742	1034
LI	Case 1	-3	-2	-5	-3	-4
	Case 2	-3	-5	-8	-7	-4
KI	Case 1	34	37	35	35	34
	Case 2	40	26	40	43	41
TTI	Case 1	46	44	46	44	45
	Case 2	47	46	47	48	45

#### **4.3.2.2 Surface parameters**

This study presents an inter-comparison of a few CPSs in WRF-NMM model with different thunderstorm affected meteorological parameters like relative humidity, temperature, and precipitation. Figure 4.10a shows the inter-comparison of observed and model simulated relative humidity (%) using different CPSs over Kolkata valid from 20 May 2006 at 0000 UTC to 21 May 2006 at 0000 UTC. The observed relative humidity values are peaked from 48% to 95% (47% increase) at

1200 UTC whereas GD scheme showed a sharp rise from around 50% to 85% (35%) at 1100 UTC, which is one hour prior to the thunderstorm occurrence. But all other CPSs are failed to capture the sudden rise, which is a characteristic feature of thunderstorm.



**Figure 4.10:** The inter-comparison of observed and model simulated relative humidity (%) using different CPSs over Kolkata valid for (a) 20 May 2006 (b) 21 May 2007.

Figure 4.10b shows the inter-comparison of observed and model simulated relative humidity (%) using different CPSs over Kolkata valid from 21 May 2007 at 0000 UTC to 22 May 2007 at 0000 UTC. GD

scheme has well captured the intensity of relative humidity during the model simulated thunderstorm hour as in the observation. The observed relative humidity values peaked from 69% to 97% (28% increase) at 1100 UTC whereas model showed a sharp rise from around 56% to 86% (30%) at 1000 UTC, which is one hour prior to the observed. All other parameterization schemes failed to capture the intensity and time as compared to the observation and GD scheme.

A statistical analysis based on MAE, RMSE and CC is performed for comparison between the simulated and observed relative humidity with different CPSs valid for 20 May 2006 (Case 1) and 21 May 2007 (Case 2) (Table 4.7). From the table, it can be clearly see that, GD scheme has less error as compared to all other schemes for both thunderstorm cases. There is not much variation between MAE values of KF and AS scheme, which is less compared to BMJ and NO schemes. In the case of RMSE, GD scheme has the least error compared to all other CPSs. The next position is for KF and AS schemes for the first thunderstorm case and AS and BMJ schemes for the second thunderstorm case. The NO scheme simulated results have the most error (MAE and RMSE). Another verification method used for this study is CC. From the table it can be clearly seen that, all the CPSs are positively correlated. The GD scheme has the highest correlation coefficient (0.87) for the first thunderstorm case. There is not much variation between the correlation coefficient of NO (0.82), AS (0.82) and KF (0.81) schemes. The BMJ scheme has the

least correlation (0.73) than other CPSs in the first thunderstorm case. All the schemes have strong correlation ( $> 0.8$ ) except BMJ scheme. In the second case also, GD scheme is well correlated (0.78) to the observation than all other CPSs. In this case AS and BMJ schemes are more correlated to the observation than KF scheme. The NO scheme has the least correlation as in the error.

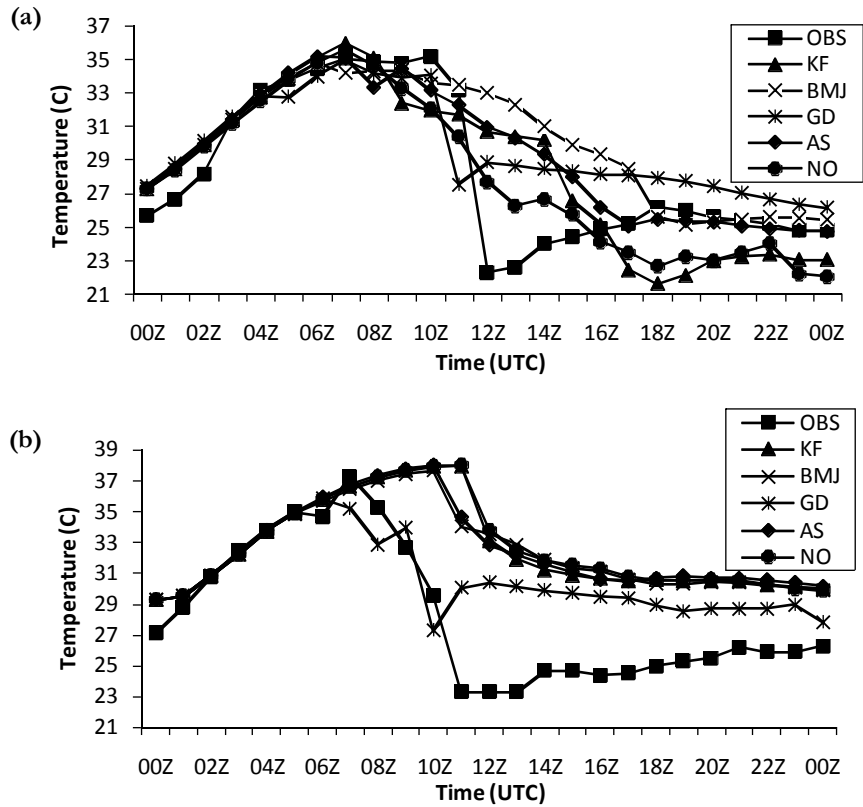
**Table 4.7:** The statistical analysis of relative humidity with different CPSs over Kolkata valid for 20 May 2006 (Case 1) and 21 May 2007 (Case 2).

Statistical Analysis	Cases	KF	BMJ	GD	AS	NO
MAE	Case 1	7.81	8.36	7.55	7.68	8.59
	Case 2	9.85	10.52	9.14	9.56	13.16
RMSE	Case 1	10.79	14.11	10.70	10.87	11.78
	Case 2	16.87	15.25	11.16	15.12	20.49
CC	Case 1	0.81	0.73	0.87	0.82	0.82
	Case 2	0.69	0.72	0.78	0.74	0.55

Figure 4.11a shows the inter-comparison of observed and model simulated surface temperature ( $^{\circ}\text{C}$ ) using different CPSs over Kolkata valid from 20 May 2006 at 0000 UTC to 21 May 2006 at 0000 UTC. The observed temperature values have a sudden fall from  $33^{\circ}\text{C}$  to  $22^{\circ}\text{C}$  ( $11^{\circ}\text{C}$ ) at 1200 UTC whereas GD scheme showed a sharp fall from  $33.5^{\circ}\text{C}$  to  $27^{\circ}\text{C}$  at 1100 UTC, which is one hour prior to the thunderstorm occurrence. But all other CPSs are failed to capture the sudden fall, which is a characteristic feature of thunderstorm. Figure 4.11b shows the inter-comparison of observed and model simulated surface temperature ( $^{\circ}\text{C}$ )

using different CPSs over Kolkata valid from 21 May 2007 at 0000 UTC to 22 May 2007 at 0000 UTC. GD scheme has well captured the variation with drop in temperature during the model simulated thunderstorm hour (1000 UTC) as in the observation. The observed temperature values have a sudden fall from 30<sup>0</sup>C to 23<sup>0</sup>C (7<sup>0</sup>C) at 1100 UTC whereas model showed a sudden drop from around 34<sup>0</sup>C to 27<sup>0</sup>C (7<sup>0</sup>C) at 1000 UTC, which is one hour prior to the observed. The intensity of temperature fall with GD scheme is same as in the observation. All other parameterization schemes are failed to capture the intensity and time as compared to the observation and GD scheme.

The statistical analysis of temperature (<sup>0</sup>C) based on MAE, RMSE and CC is performed for comparisons between the simulated and observed values with different CPSs valid for 20 May 2006 (Case 1) and 21 May 2007 (Case 2). The results are shown in Table 4.8. By taking both error measurement (MAE and RMSE) values together, it can be clearly seen that, GD scheme has less error as compared to all other schemes in both thunderstorm cases. The errors are less for NO scheme in the first thunderstorm case. While AS scheme has less MAE and RMSE, when two thunderstorm cases taken together. GD scheme has an average error of 2.62<sup>0</sup>C for temperature and 10.80% for relative humidity. The next position is for AS scheme. All other schemes have more errors for the simulation of thunderstorm affected parameters. The GD scheme has the highest CC (0.87 and 0.82) in both cases as compared to all other CPSs.



**Figure 4.11:** The inter-comparison of observed and model simulated temperature ( $^{\circ}\text{C}$ ) using different CPSs over Kolkata valid for (a) 20 May 2006 (b) 21 May 2007.

The CC of NO scheme is higher for first thunderstorm case, but least for second case. When average correlation of two cases is taken, GD scheme simulated results are highly correlated to the observation than all other schemes. The next position is for AS scheme as in the MAE and RMSE. All other schemes have less correlation.

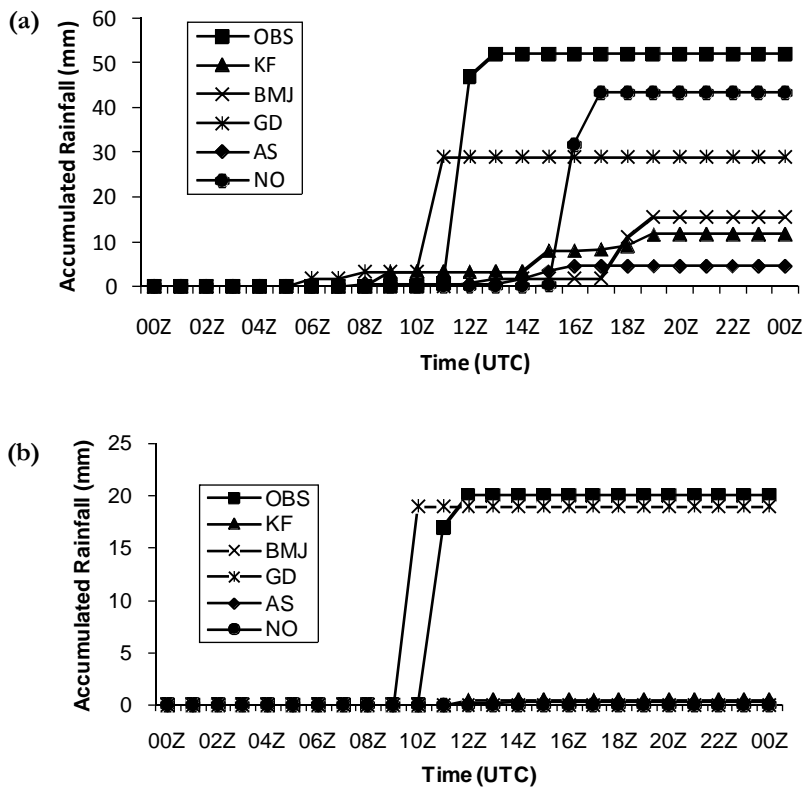


**Table 4.8:** The statistical analysis of temperature with different CPSs over Kolkata valid for 20 May 2006 (Case 1) and 21 May 2007 (Case 2).

Statistical Analysis	Cases	KF	BMJ	GD	AS	NO
MAE	Case 1	2.44	2.17	1.62	2.20	1.89
	Case 2	4.44	4.35	2.96	4.40	4.58
RMSE	Case 1	3.28	3.67	2.28	2.96	2.80
	Case 2	5.66	5.42	3.64	5.43	5.84
CC	Case 1	0.74	0.68	0.87	0.79	0.81
	Case 2	0.56	0.63	0.82	0.66	0.53

The rainfall fields are examined by temporal and spatial pattern. Figure 4.12a shows the inter-comparison of observed and NMM model simulated diurnal variation of 24 h accumulated rainfall (mm) with different CPSs over Kolkata valid from 20 May 2006 at 0000 UTC to 21 May 2006 at 0000 UTC. GD scheme is able to capture 29 mm of rainfall, which is less compared to actual observation (52 mm). GD scheme has predicted the rainfall at 1100 UTC, which is one hour prior to the actual severe thunderstorm occurrence (1200 UTC). The NO scheme is well captured the intensity (43 mm) with five hour time lag. But other schemes are failed to capture the intensity and time of occurrence. Figure 4.12b shows the inter-comparison of observed and NMM model simulated accumulated progressive rainfall with different CPSs at Kolkata valid from 21 May 2007 at 0000 UTC to 22 May 2007 at 0000 UTC. The GD scheme is able to capture 18.5 mm of rainfall at 1000 UTC, which is very close to the actual observation (20 mm). The GD scheme has predicted

the rainfall at 1000 UTC, which is one hour prior to the actual thunderstorm occurrence (1100 UTC). The GD scheme is well simulated the intensity and time of occurrence of precipitation over Kolkata on 21 May 2007. But other schemes are failed to capture the rainfall amount and time of occurrence as compared to GD scheme.

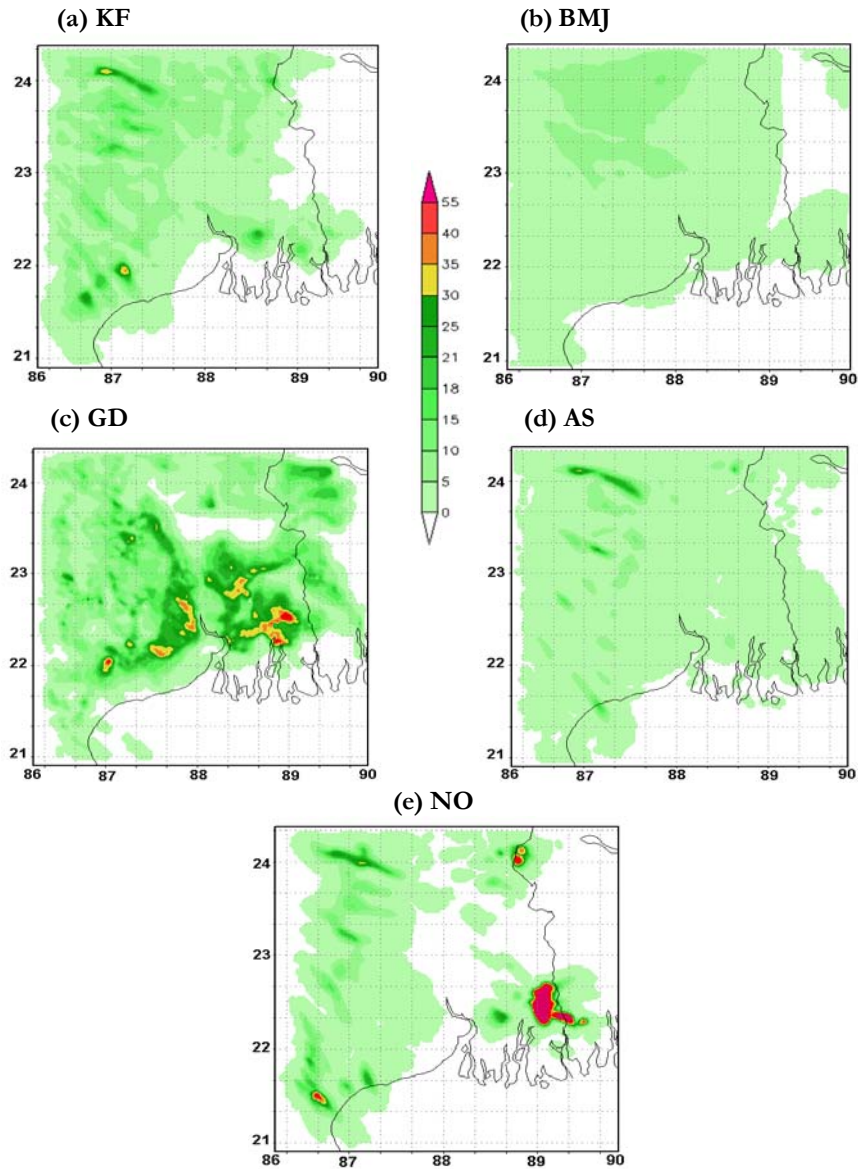


**Figure 4.12:** The inter-comparison of observed and model simulated accumulated rainfall (mm) with different CPS over Kolkata valid for (a) 20 May 2006 (b) 21 May 2007.

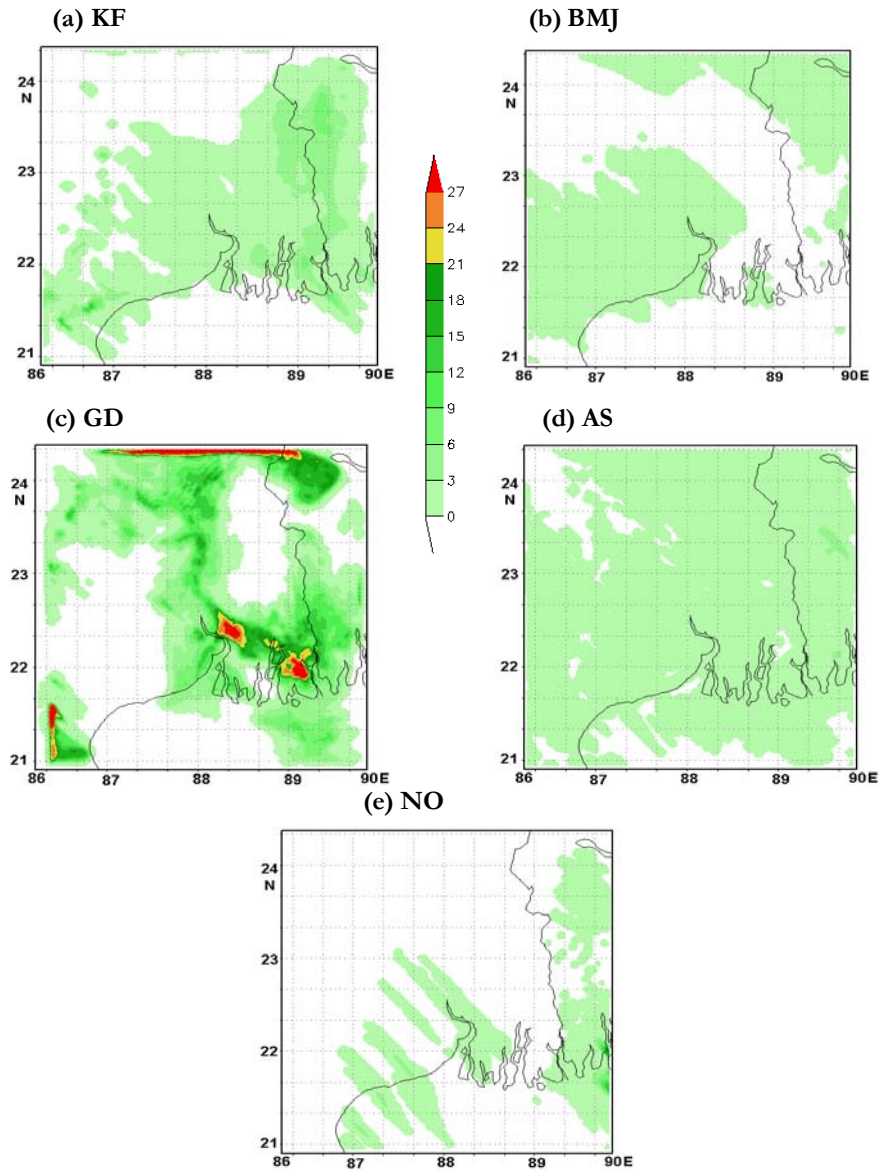
The spatial distribution of 3 h accumulated rainfall (mm) between 0900 and 1200 UTC with different CPSs on 20 May 2006 is shown in Figure 4.13. From the figures, it can be seen that, GD scheme is well simulated the rainfall intensity as compared to other schemes during the thunderstorm hours. NO scheme is also able to simulate the intensity, but the location is shifted to eastwards (near Bangladesh border). All other CPSs are failed to capture the intensity and time of this severe thunderstorm event. The spatial distribution of 3 h accumulated rainfall (mm) between 0900 and 1200 UTC with different CPSs on 21 May 2007 is shown in Figure 4.14. From the figures, it can be clearly seen that GD scheme is well simulated the rainfall intensity as compared to other schemes during the thunderstorm hours. All other CPSs are failed to capture the intensity and time of this severe thunderstorm event.

#### ***4.3.2.3 Composite radar reflectivity***

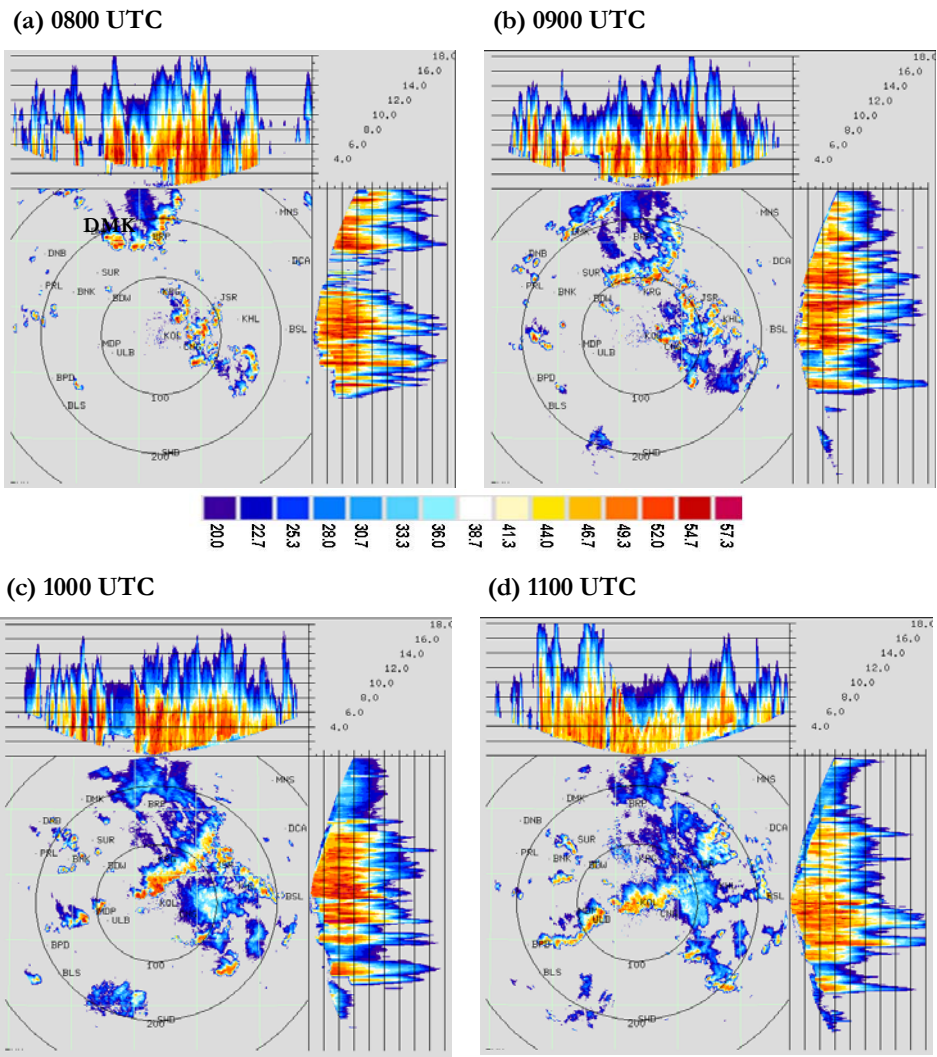
Kolkata DWR composite radar reflectivity imageries on 21 May 2007 from 0800 to 1100 UTC is shown in Figure 4.15. By analyzing DWR imageries, scattered echoes are developed near Dumka (DMK) at 0800 UTC and moving south eastwards at 0900 UTC. This echo is intensified into a squall line (30 km north of Kolkata) at 1000 UTC. This squall line moved further in southeast direction (Mohanty et al. 2007). NMM model with GD scheme simulated composite radar reflectivity on 21 May 2007 from 0800 to 1100 UTC is shown in Figure 4.16. By analyzing NMM model simulated radar reflectivity pictures, scattered



**Figure 4.13:** The spatial distribution of 3 h accumulated rainfall (mm) between 0900 and 1200 UTC with different CPSs on 20 May 2006.



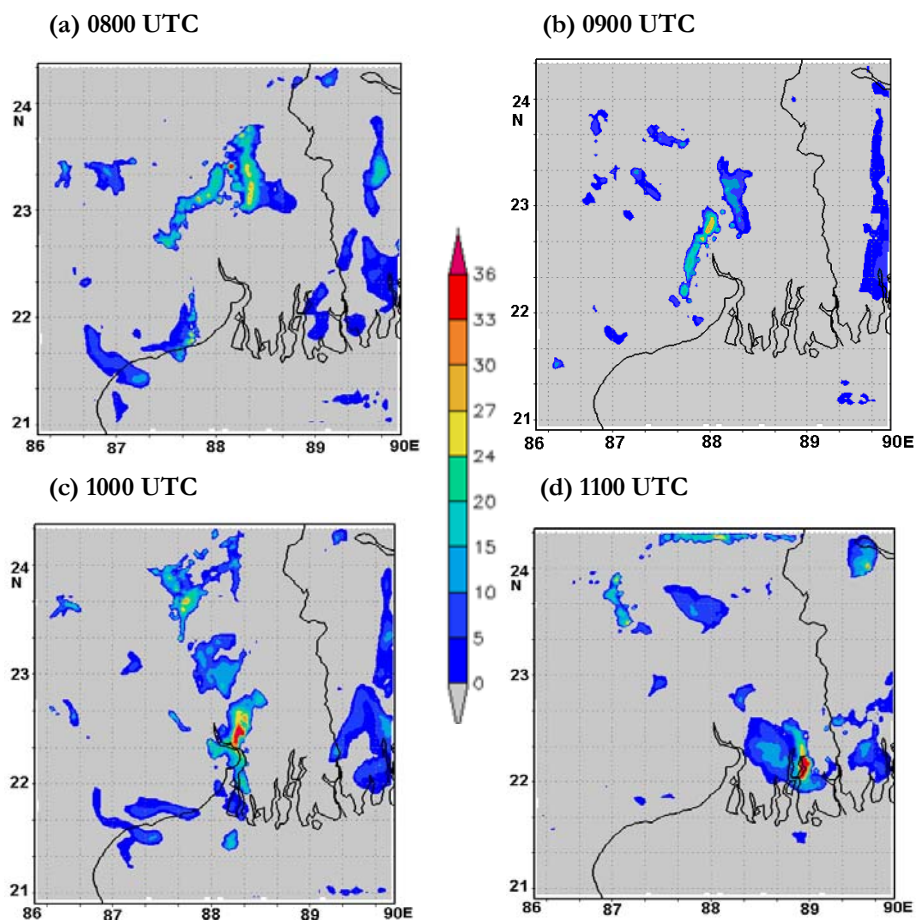
**Figure 4.14:** The spatial distribution of 3 h accumulated rainfall (mm) between 0900 and 1200 UTC with different CPSs on 21 May 2007.



**Figure 4.15:** Kolkata DWR composite radar reflectivity (dBZ) imageries from 0800 to 1100 UTC on 21 May 2007.

echoes developed northwest of Kolkata at 0800 UTC. This echo was moving south eastwards at 0900 UTC and intensified at 1000 UTC. This

echo moved further in southeast direction and disappeared over sea. The GD scheme is well simulated this thunderstorm movement with simulated composite radar reflectivity fields. The other schemes are failed to capture the echoes and movement of this system (Figures are not shown).



**Figure 4.16:** NMM simulated composite radar reflectivity (dBZ) pictures from 0800 to 1100 UTC on 21 May 2007 with GD scheme.

The trends shown by various meteorological fields of NMM model with GD scheme are in good agreement with each other and very much consistent with dynamic and thermodynamic properties of the atmosphere for the occurrence of severe thunderstorms on 20 May 2006 and 21 May 2007 even though one hour time lead exists.

### **4.3.3 Sensitivity study with different microphysics schemes**

NWP models contain a large number of physical parameterization schemes in order to represent the various atmospheric processes that take place in sub-grid scales. Cloud microphysical processes play an important role through direct influences on the cold pool strength (due to rainfall evaporation) and latent heating (due to condensation). Therefore, microphysical parameterizations could be a principal source of uncertainty in convection allowing high resolution NWP models. It is important to quantify the uncertainty associated with the cloud microphysics parameterization – a salient concern in convection permitting models. It is also important to assess whether increasingly sophisticated cloud microphysics gives consistently better results. Therefore, assessing the cloud microphysics schemes is not only of practical significance but also helpful for guiding the future improvement of cloud microphysics parameterizations. The model simulated results with different microphysics schemes are explored in the following section. Analysis of the results of these experiments is helpful to understand the impact of microphysics on the simulation of 15 May 2009



severe thunderstorm event and assist in the customization of model for future severe thunderstorm simulations over east Indian region.

#### ***4.3.3.1 Stability indices***

An attempt is made to examine different stability indices namely CAPE, LI, TTI and KI obtained from NMM model during this severe local storm day over Kolkata. Table 4.9 shows the NMM model simulated stability indices with three microphysics schemes (FERR, WSM6 and THOM) over Kolkata at 1300 UTC on 15 May 2009. The model simulated CAPE values with different microphysics schemes are high and greater than the critical levels (1500 and 1000  $\text{Jkg}^{-1}$ ) during this thunderstorm hour, which is a favorable condition for severe thunderstorm occurrence. The NMM model simulated LI values with different microphysics schemes are very less (-6) at 1300 UTC, which is less than the critical level suggested by AiWS Technical Report (1990) and Tyagi et al. (2011) and favorable for thunderstorm occurrences. The model simulated KI value with FERR scheme (31) is close to the critical level suggested by AiWS Technical Report (1990) during the thunderstorm hour. The KI value of WSM6 and THOM schemes are very less than the critical level suggested by AiWS Technical Report (1990), but more than the critical level suggested by Tyagi et al. (2011) (Table 4.3). The NMM model simulated TTI with FERR schemes are showing high values, which is greater than or equal to both critical levels (46 and 44) during the thunderstorm hour of these severe local storm (Table 4.9).

The TTI value of WSM6 and THOM schemes are less than the critical level suggested by Tyagi et al. (2011). Examination of all the model simulated stability indices at 1300 UTC of 15 May 2009 clearly indicated that NMM model with FERR microphysics scheme has well captured the instability of the atmosphere for the occurrence of a severe thunderstorm. The model simulated thermodynamic structure over Kolkata was conducive for a thunderstorm occurrence as suggested by AiWS Technical Report (1990) and Tyagi et al. (2011). CAPE and LI values are well simulated by the other microphysics schemes, but failed to produce accurate values of KI and TTI indices suggested by AiWS Technical Report (1990) and Tyagi et al. (2011).

**Table 4.9:** NMM model simulated stability indices over Kolkata at 1300 UTC using different microphysics schemes.

Stability Indices	FERR	WSM6	THOM
CAPE	2428	2664	2603
LI	-6	-6	-6
KI	31	27	25
TTI	46	45	45

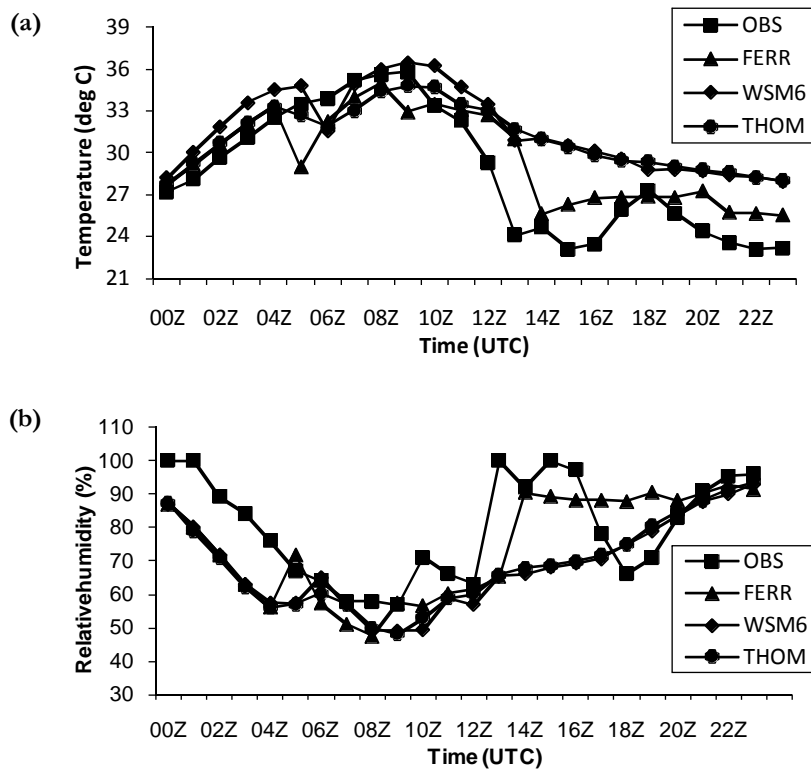
#### 4.3.3.2 Surface parameters

Figure 4.17a shows the inter-comparison of observed (AWS) and NMM model simulated diurnal variation of surface temperature ( $^{\circ}\text{C}$ ) with different microphysics schemes over Kolkata valid from 15 May 2009 at 0000 UTC to 16 May 2009 at 0000 UTC. From the figures, it can be

clearly seen that the NMM model with FERR scheme captured the sudden temperature drop during the model simulated thunderstorm hour (1400 UTC), one hour after the observed. The observed temperature showed a drop from 29<sup>0</sup>C to 24<sup>0</sup>C (5<sup>0</sup>C) at 1300 UTC, whereas NMM simulation shows a drop with same intensity of 5<sup>0</sup>C from 31<sup>0</sup>C to 26<sup>0</sup>C at 1400 UTC. The other microphysics schemes are failed to capture the sudden temperature fall, a feature for thunderstorm occurrence. Figure 4.17b shows the inter-comparison of observed (AWS) and NMM model simulated diurnal variation of surface relative humidity (%) with different microphysics schemes over Kolkata valid from 15 May 2009 at 0000 UTC to 16 May 2009 at 0000 UTC. NMM model with FERR schemes has captured the sudden rise of relative humidity values during the model simulated thunderstorm hour as in the observation. The observed relative humidity showed a rise from 63% to 100% (37%) at 1300 UTC, whereas NMM simulation showed a rise from 65% to 91% (26%) at 1400 UTC. The WSM6 and THOM schemes are failed to capture sudden rise of relative humidity as in the temperature. The NMM model well simulated the sudden rise and fall of surface relative humidity and temperature during the model simulated thunderstorm hour with one hour time lag.

A statistical analysis based on MAE, RMSE and CC was performed for comparison between the simulated and observed surface temperature and relative humidity over Kolkata valid for 15 May 2009 at 0000 UTC to 16 May 2009 at 0000 UTC. The results of the analysis are

given in Table 4.10. The simulated temperature and relative humidity with FERR scheme have more CC than other two schemes. The CC of temperature and relative humidity with FERR scheme are 0.87 and 0.72 respectively. The CC of relative humidity with WSM6 and THOM schemes are very less and which is very below the range of strong correlation (0.8).



**Figure 4.17:** The inter-comparison of observed (AWS) and NMM model simulated diurnal variation of (a) surface temperature ( $^{\circ}\text{C}$ ) (b) relative humidity (%) with different microphysics schemes over Kolkata on 15 May.

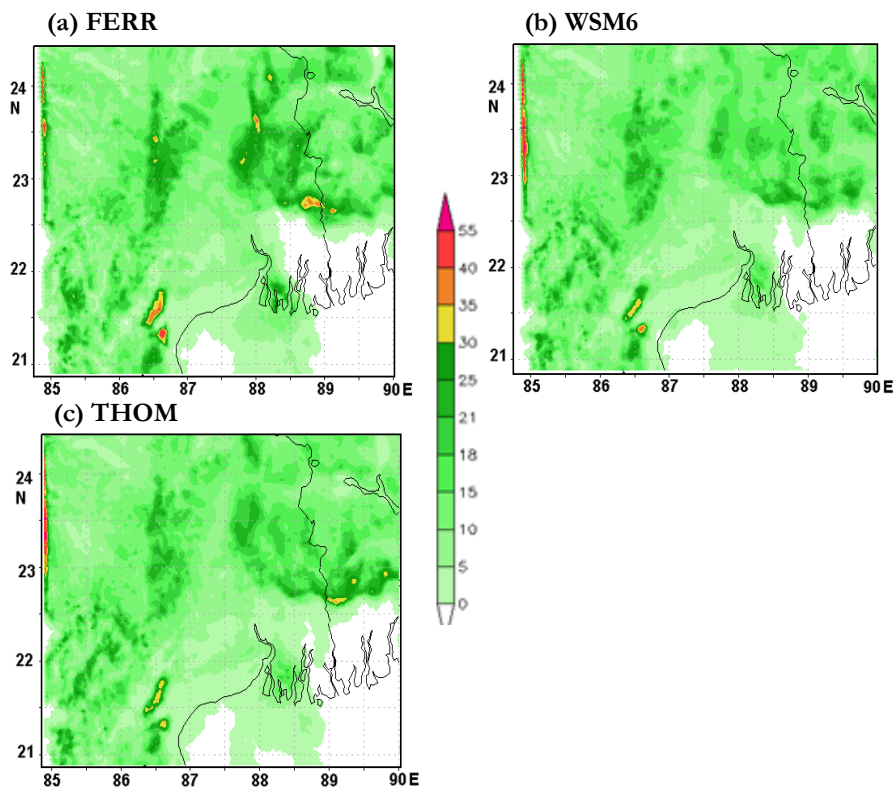
The MAE and RMSE of each weather parameters have been computed. The MAE and RMSE values are less for FERR scheme as compared to other two schemes. The MAE and RMSE of temperature with FERR scheme are 1.93<sup>0</sup>C and 2.47<sup>0</sup>C respectively. The same errors of relative humidity with FERR scheme are 10.89% and 13.80%. The errors are high for other schemes. The results suggest that the NMM model with FERR scheme holds promise for prediction of surface weather parameters with reasonable accuracy in severe thunderstorm cases.

**Table 4.10:** Statistical analysis of simulated and observed surface temperature and relative humidity over Kolkata based on MAE, RMSE and CC.

Statistical Analysis	Parameter	FERR	WSM6	THOM
MAE	Temperature	1.93	3.28	3.07
	Relative Humidity	10.89	12.90	12.75
RMSE	Temperature	2.47	3.88	3.80
	Relative Humidity	13.80	16.28	16.07
CC	Temperature	0.87	0.81	0.74
	Relative Humidity	0.72	0.54	0.55

The comparison of 3 h accumulated NMM model simulated rainfall with three microphysics schemes on 15 May 2009 is plotted in Figure 4.18. Prior studies have shown that a model's microphysical parameterization scheme can strongly influence the magnitude of predicted precipitation (Otkin et al. 2006). The overall rainfall distribution is reasonably well captured by all three microphysics schemes. The

maximum intensity of model simulated rainfall with FERR scheme is 40 to 55 mm. But the other two parameterization schemes are failed to capture this maximum intensity over West Bengal region. FERR scheme show a rainfall spread in the range of 25 to 40 mm over north, northwest and northeast of Kolkata. The other schemes are showing this spread, but failed to capture the intensity as in FERR scheme.



**Figure 4.18:** The spatial distribution of 3h accumulated rainfall (mm) between 1200 and 1500 UTC with different microphysics schemes on 15 May 2009.

Table 4.11 shows a quantitative assessment of model simulated 24 h accumulated rainfall data using different microphysics schemes and surface rain gauge observations of IMD over West Bengal region on 15 May 2009. The results show that FERR scheme was closer to the records than WSM6 and THOM schemes. The observed and simulated rainfalls with FERR scheme are closely agreed at Krishnagar and Dum Dum stations. The other two microphysics schemes are underestimated the rainfall amount over all 5 stations except Krishnagar as in the spatial rainfall distribution. The average rainfall from 6 stations show that FERR scheme simulated average rainfall is close to the observation. The results suggest an overall improvement from NMM model with FERR scheme with respect to ground based measurements.

**Table 4.11:** The comparison of model simulated 24 h accumulated precipitation using different microphysics schemes of 6 meteorological stations with rain gauge observations.

STATION	LAT	LON	OBS	FERR	WSM6	THOM
Dum Dum	22.39	88.27	16.90	17.20	10.78	9.24
Bankura	23.13	87.04	34.00	14.70	15.32	16.41
Krishnagar	23.24	88.31	19.60	18.70	25.81	22.35
Digha	21.50	87.48	21.00	18.40	10.00	11.87
Midnapore	22.25	87.19	51.60	26.50	12.65	13.16
Haldia	22.04	88.04	33.20	30.40	15.71	18.74
MEAN			<b>29.38</b>	<b>20.98</b>	<b>15.04</b>	<b>15.29</b>

#### ***4.3.3.3 Composite radar reflectivity***

The storm structure can be evaluated by comparing the modeled radar reflectivity to the observed radar reflectivity. In order to achieve that composite radar reflectivity (dBZ) calculated in different simulation times are compared with observed one. Figure 4.19 shows the Kolkata DWR imageries for the thunderstorm event, which occurred on 15 May 2009. The analysis of these imageries revealed that a strong echo was developed near Purulia (PRL) at 1000 UTC, which intensified into north-south oriented squall line by 1100 UTC. This echo gradually moved southeastwards at 1200 UTC. This echo was passed over Kolkata by 1300 UTC. The intensity of the squall was reduced thereafter and disappeared at 1500 UTC. The movements of this severe thunderstorm are from northwest to southeast as in the typical Norwester's (Mohanty et al. 2009).

NMM model simulated composite radar reflectivity with FERR scheme on 15 May 2009 from 1000 to 1300 UTC is shown in Figure 4.20. By analyzing simulated reflectivity plots it can be seen that a strong echo developed northwest of Kolkata at 1000 UTC. This echo moved towards Kolkata at 1100 UTC and over Kolkata at 1300 UTC as in the DWR imageries. The movement of squall line was well captured by FERR scheme as in the observation. The other microphysics schemes namely WSM6 and THOM (Figure 4.21 and 4.22) also captured this squall line



movement which was initiated northwest of Kolkata at 1000 UTC and moved further towards Kolkata during the following hours.

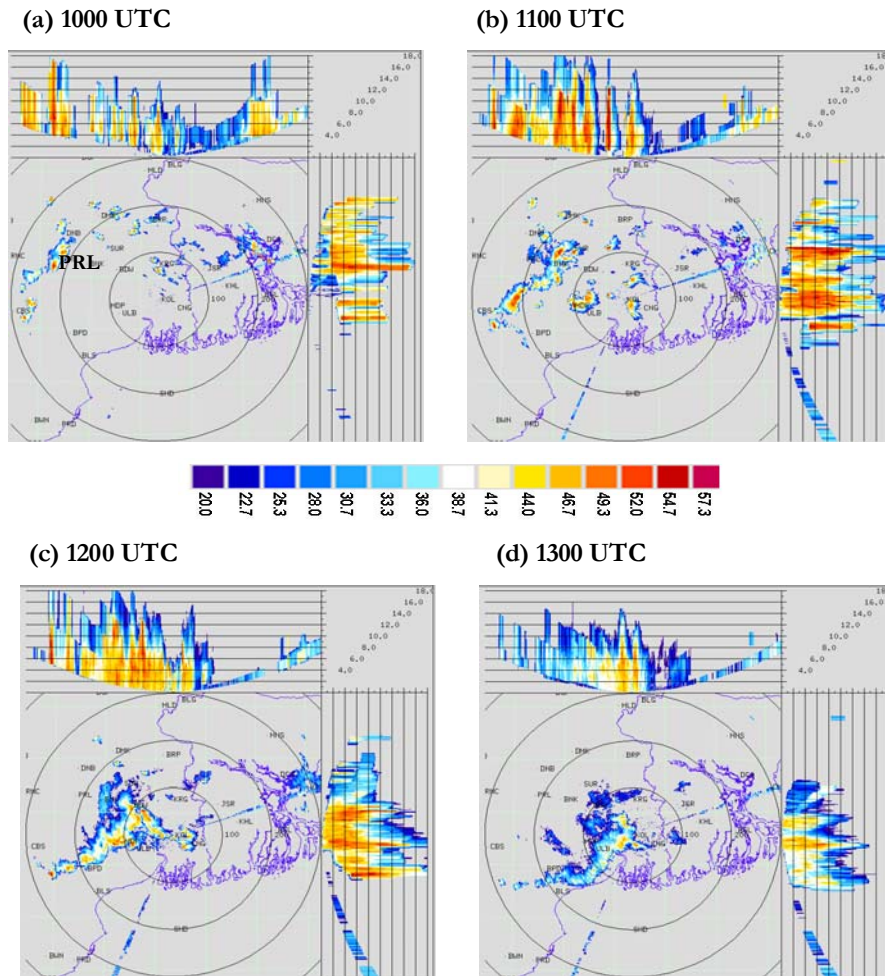
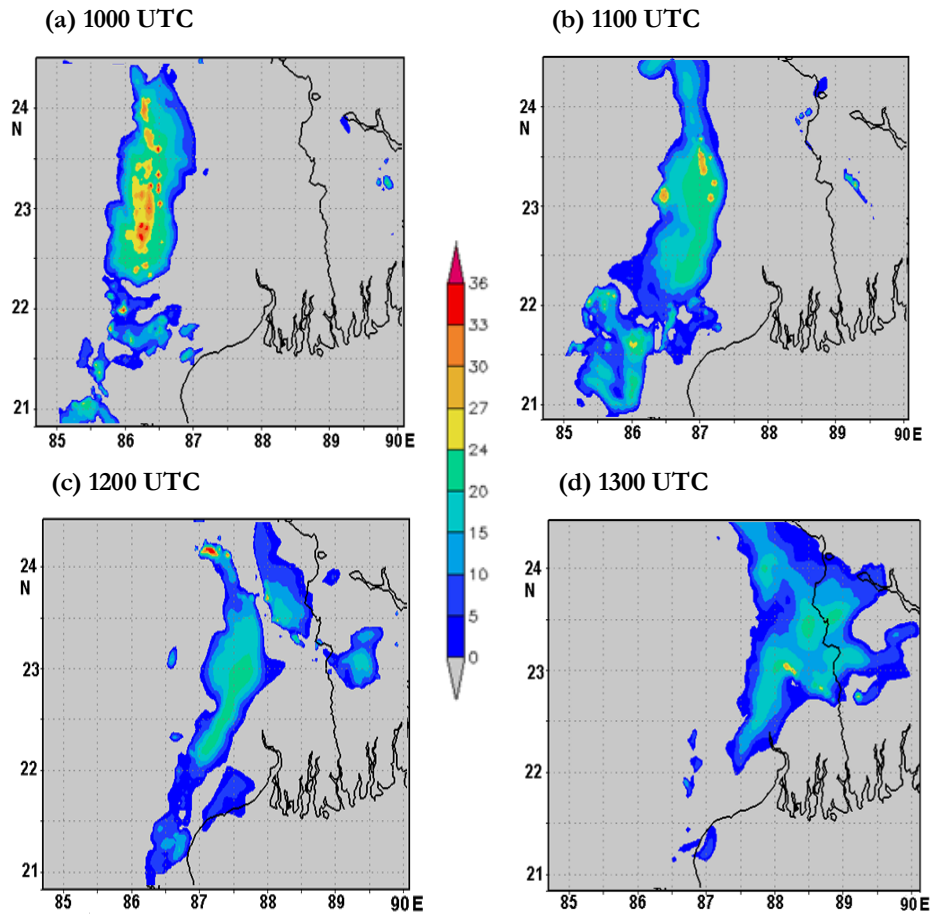


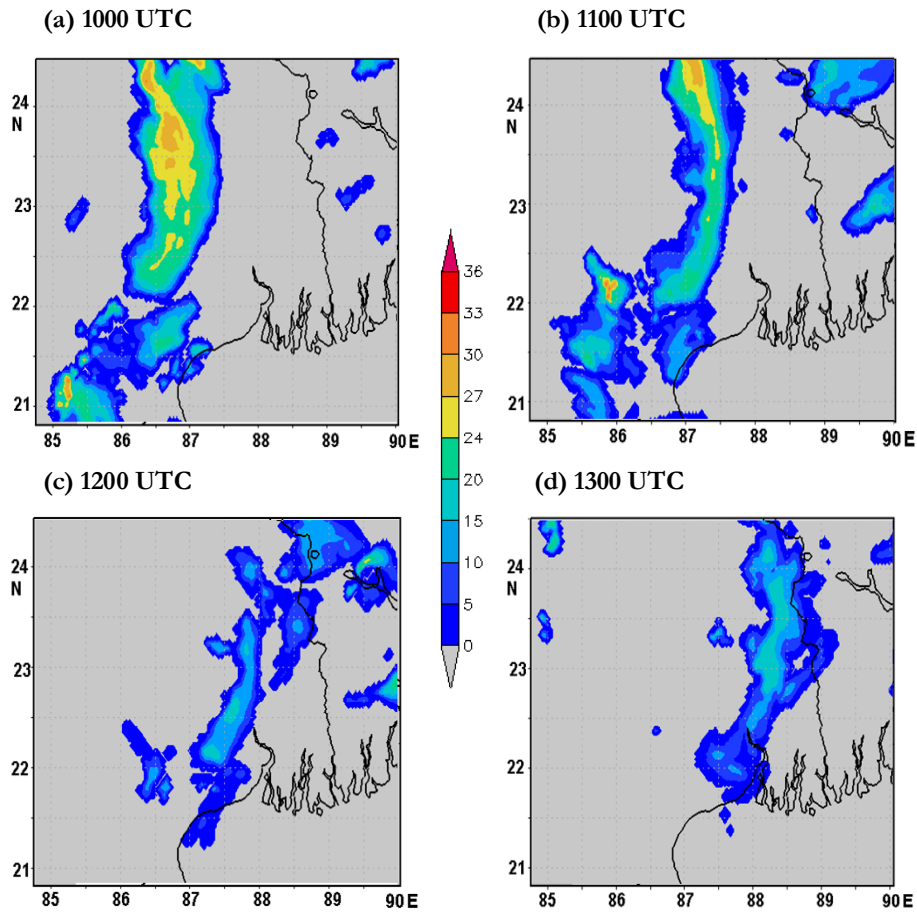
Figure 4.19: Kolkata DWR composite radar reflectivity (dBZ) imageries from 1000 to 1300 UTC on 15 May 2009.



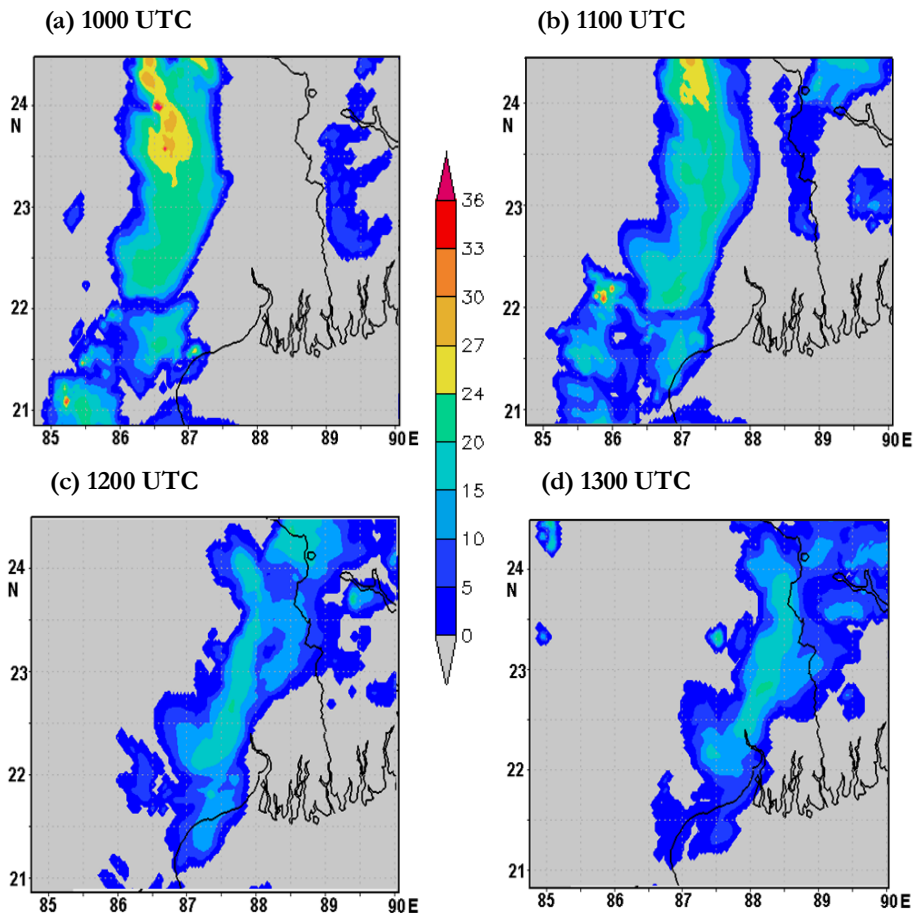
**Figure 4.20:** NMM simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009 using FERR microphysics scheme.

The trends shown by various meteorological fields of NMM model with FERR scheme are in good agreement with each other and very much consistent with dynamic and thermodynamic properties of the atmosphere

for the occurrence of a severe thunderstorm on 15 May 2009 even though one hour time lag exists.



**Figure 4.21:** NMM simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009 using WSM6 microphysics scheme.



**Figure 4.22:** WRF-NMM simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009 using THOM microphysics scheme.

#### 4.4 Chapter Summary

The sensitivity experiments have been conducted with NMM model to examine the impact of different initial conditions, CPSs and

microphysics schemes in capturing the severe thunderstorm events occurred over Kolkata during May 2006, 2007 and 2009.

The thunderstorm of 20 May 2006 over Kolkata is simulated using NMM model with different initial conditions to resolve mesoscale signature of the atmosphere and establish the robustness of the results. The simulation of the stability indices is good enough with the values indicating higher instability for the thunderstorm to occur. Ex-3 (24 h) has well simulated the thunderstorm initiation in terms of stability indices. Ex-3 has performed well in simulating all the thunderstorm affected parameters namely surface temperature, relative humidity and accumulated rainfall which is useful for occurrence and intensity of the severe thunderstorm even though one hour time lead exists. A statistical analysis based on MAE, RMSE and CC is also revealed that, time and intensity of surface parameters are well captured by Ex-3. From the spatial plots of composite radar reflectivity of Ex-3, it can be seen that a squall line is initiated at 0900 UTC, gradually moved towards Kolkata at 1000 UTC and intensified at 1100 UTC as in DWR pictures. Ex-3 well captured the squall line movement than other initial conditions.

The sensitivity experiments have been conducted with the NMM model to test the impact of CPSs on simulating severe thunderstorms that occurred over Kolkata on 20 May 2006 and 21 May 2007 and validated the model results with observation. In all experiments, the setups were identical except for the use of different convective schemes. Hence

differences in the simulation results may be attributed to the sensitivity of the convective schemes. This study showed that the prediction of thunderstorm affected parameters is sensitive to CPSs. It is clearly demonstrated that GD scheme performance is significantly better than other parameterization schemes including explicit scheme. By comparing both the thunderstorm cases, GD scheme is well simulated the instability of the atmosphere in terms of CAPE, LI, KI and TTI for the occurrence of a severe thunderstorm over Kolkata as compared to all other CPSs. The time-series plot and statistical analysis of surface temperature and relative humidity revealed that GD scheme is well captured the sufficient deep humid layer and dip in temperature for the occurrence of a severe thunderstorm on 20 May 2006 and 21 May 2007 as in the observation. The temporal and spatial patterns of precipitation simulated by GD scheme are in good agreement with the observation. But all other schemes are failed to capture the intensity and time of occurrence for both the thunderstorm cases. From the model simulated spatial plots of composite radar reflectivity, it can be cleared that the squall line movements are also well captured by GD scheme for both thunderstorm cases.

The sensitivity experiments have been conducted for a severe thunderstorm on 15 May 2009 with three microphysical schemes to examine the sensitivity of the simulations to different cloud microphysics. Examination of the model simulated stability indices with different microphysics schemes on 15 May 2009 clearly indicated that NMM

model with FERR scheme has well captured the instability of the atmosphere for the occurrence of a severe thunderstorm. The model simulated surface precipitation with three microphysics schemes are analyzed and compared to the available observations in order to identify the microphysics scheme that provide the best representation of the spatio-temporal variability of precipitation in severe thunderstorm conditions. The best results are produced by FERR microphysics scheme. The analysis of surface temperature and relative humidity revealed that FERR scheme is well captured the sudden fall and rise on 15 May 2009 with one hour time lag. The statistical analysis showed that FERR scheme has performed well in simulating all thunderstorm affected surface parameters. From the model simulated spatial plots of composite radar reflectivity with three microphysics schemes, it can be seen that in all three cases, a thunder squall is initiated northwest of Kolkata, gradually moving towards Kolkata as in DWR imageries. The NMM model with all microphysical schemes has reasonably well simulated the movement of the severe thunderstorm of 15 May 2009 as in the DWR imageries.

The results of these analyses demonstrated the capability of high resolution WRF-NMM model in simulation of severe thunderstorm events and determined that the 3 km model improve upon current abilities when it comes to simulating severe thunderstorms over east and northeast Indian region.

..........





## Chapter-5

# COMPARISON OF NUMERICAL MODELS FOR THUNDERSTORM PREDICTION

---

### Contents

- 5.1 Data and Methodology
  - 5.2 Result and Discussion
  - 5.3 Chapter Summary
- 

Forecasting of severe thunderstorms is a daunting challenge due to great complexity of the processes involved and interplay of many factors. Part of the problem is due to small time scales of these disturbances which enable only short lead times for forecasting. Meteorologists involved in severe weather forecasting have to look at a number of ingredients pertaining to the atmospheric flow and the thermodynamic conditions to demarcate the favorable or unfavorable environments for the growth of severe weather. This is an extremely difficult task as the parameters usually taken into consideration are essentially interdependent and vary in relation to each other in different situations.

The use of NWP output to complement the interpretation of conventional observations can add great value to the forecast process. The higher time and space resolution of the model data enables a forecaster to

view the evolution of the weather situation in much greater detail and can provide an insightful framework within which actual observations can be interpreted. Forecasting thunderstorms has been an issue of debate for quite some time. Many NWP models like MM5, ARW, NMM and ARPS have been in operational use for weather forecasting at many places in the world. A basic characteristic of these models is that their governing equations are non-hydrostatic since the vertical and horizontal scales of convection are similar. Such models are also necessary for explicitly resolving gravity waves triggered by clouds. Presently, NWP models having a resolution less than 9 km are also available for the simulation and prediction of regional weather systems. These models can be used for a variety of applications including simulation and prediction of heavy rainfall, severe storms and tropical cyclones (Mohanty et al. 2004). NWP models have been developed with wide variety of flexibilities in terms of altering horizontal and vertical resolutions, nesting domains, and choosing options for different physical parameterization schemes. By setting some important parameters appropriately, these models can be used in a wide range of applications including thunderstorm forecasting (Mohanty et al. 2003).

Numerical modeling of clouds has evolved in the past 4 decades. In 60's, two-dimensional cloud models were developed to study evolution of clouds in idealized conditions. Subsequently, in 70's, three-dimensional cloud models were developed to quantify the effects of wind

shear on deep convection, squall lines and mid-latitude thunderstorms, which are associated with tornado genesis (Ogura and Takahashi 1971; Orville and Kopp 1977; Schlesinger 1978; Klemp and Wilhelmson 1978). Use of cloud resolving models (also known as cumulus ensemble models) started in late 70's and 80's to study collective effects of convection on the large-scale environment (Arakawa and Schubert 1974; Cotton et al. 1982; Tao and Simpson 1984; Tao et al. 1987). The main objective was to improve cumulus (convective) parameterization scheme based on the knowledge gained from the cloud resolving models (CRM). A primary interest was to study the effect of wind shear on mesoscale convective systems, effect of ice processes on cloud formation and evolution, effects of stratiform rainfall and their relationship with deep convective rainfall. In 1990's CRM was further improved to study multi-scale processes, cloud-chemistry interaction and surface processes (Grabowski et al. 1996; Lin and Arakawa 1997; Lynn et al. 1998; Tao et al. 1999).

Although studies have been conducted for pre-monsoon thunderstorms over Indian region, serious attempt to predict the development was a recent activity (e.g. Dasgupta and De 2007; Chaudhary 2008; Ghosh et al. 2008; Mukhopadhyay et al. 2009; Latha and Murthy 2011; Tyagi et al. 2011). Dasgupta and De (2007) developed logistic regression model for prediction of pre-monsoon thunderstorms over Kolkata. Chaudhury (2008) has studied low level clouds during these thunderstorm days using soft computing technique in the form of rough

set theory. Latha and Murthy (2011) have presented pre-monsoon thunderstorm development in terms of turbulence and wind fields using doppler sodar observations. Use of satellite and DWR data in studying pre-monsoon thunderstorms are also available in the literature (Ghosh et al. 2008; Mukhopadhyay et al. 2009).

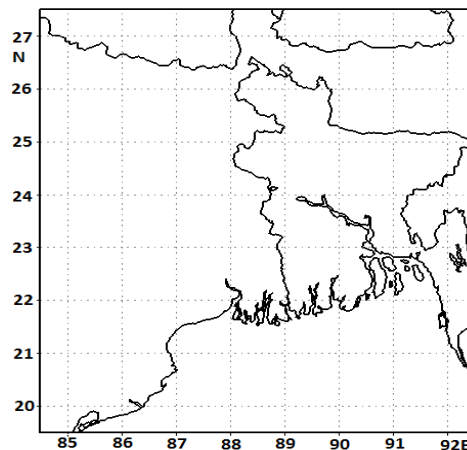
In India, studies related to modeling of clouds and thunderstorms are scarce. The improvement in thunderstorm prediction is also highly handicapped due to lack of mesoscale observations and insufficient understanding (Rajeevan et al. 2012). Motivated by the need to develop NWP interpretation techniques for forecasting of severe weather in east Indian region, an attempt is made to examine the utility of numerical guidance. In this study, the simulated results of three thunderstorm events (3, 11 and 15 May 2009), using NMM and ARW modeling systems are compared. The temporal variations of surface temperature and relative humidity, which are useful for occurrence and intensity of the severe thunderstorms are evaluated and validated the model results with observations. The capacities of the NMM and ARW models in retrieving precipitation fields over east Indian region during three severe thunderstorm events were analyzed, by comparing the outputs of the models with ground observations. A quantitative verification of the results was performed with classical statistics parameters namely MAE, RMSE and CC. The model simulated radar reflectivity and cloud top temperature were compared with the Kolkata DWR and Kalpana satellite

imageries, to verify whether the models were able to simulate the genesis, intensification and propagation of these thunder squalls.

## **5.1 Data and Methodology**

The skill of prediction has been demonstrated in this study through the numerical simulations with NWP models. NMM and ARW modeling systems were used here to perform cloud resolving simulation of thunderstorm events that occurred over east Indian region. The WRF model is a next-generation mesoscale forecast model that will be used to advance the understanding and the prediction of mesoscale convective systems. It features a software architecture allowing for computational parallelism and system extensibility. The WRF model will be used for a wide range of applications, from idealized research to operational forecasting, across scales ranging from meters to thousands of kilometers. The WRF model contains two dynamic cores: NMM (Janjic 2003) core, developed by NOAA/NCEP at USA and the ARW (Skamarock et al. 2005) core, developed by NCAR at USA. NMM runs are initialized through the same basic mechanism as the ARW runs. The WPS reads GRIB data from an initializing model and interpolates it onto the target WRF domain grid. However, the functionality of the WPS had to be expanded to handle the horizontal staggering, map projection, and vertical coordinate used by the NMM, as each is distinct from its ARW counterpart. The details of each model are specified in Chapter 2.

In the present study, an attempt has been made to compare the simulated results of thunderstorm events during 3, 11 and 15 May 2009 using NMM and ARW models. Both ARW and NMM models were integrated for a period of 24 hours starting from 0000 UTC of each day and ending at 0000 UTC of the following day. Boundary and initial conditions for both models are from FNL dataset of NCEP with  $1^{\circ} \times 1^{\circ}$  lat/lon grids. Both models thus have a common starting point, and avoid a potential source of difference. A single domain was configured with 3 km horizontal spatial resolution (Figure 5.1), which is reasonable in capturing the mesoscale cloud clusters. The domain covers  $84.5^{\circ}\text{E}$  to  $92.5^{\circ}\text{E}$  and  $19.5^{\circ}\text{N}$  to  $27.5^{\circ}\text{N}$  and the grids are centered at  $88.5^{\circ}\text{E}$ ,  $23.5^{\circ}\text{N}$ . Both NMM and ARW domains are configured with vertical structure of 38 unequally spaced sigma (non-dimensional pressure) levels.



**Figure 5.1:** Domain of NMM and ARW model.

In this study, same physics options are taken for both the ARW and NMM simulations. The physical parameterizations used in this study are GFDL for longwave and shortwave radiation (Schwarzkopf and Fels 1991; Lacis and Hansen 1974), Noah Land surface scheme (Chen and Dudhia 2001) for land surface, MYJ scheme (Janjic 2002) for planetary boundary layer, Ferrier scheme (Ferrier 2002) for microphysics, Janjic similarity scheme (Janjic 1994) for surface layer and GD cloud ensemble scheme (Grell and Devenyi 2002) for CPS. All the above schemes are well tested for NMM and ARW models. Table 5.1 shows the model configuration for the present study.

**Table 5.1:** NMM and ARW model configuration.

Model	WRF-NMM	WRF-ARW
Dynamics	Non-hydrostatic	Non-hydrostatic
Horizontal resolution	3km	3km
Forecast Length	24 hrs	24 hrs
Map projection	Rotated latitude and longitude	Mercator
Horizontal grid system	Arakawa E-grid	Arakawa C-grid
Vertical co-ordinate	Hybrid sigma to pressure vertical coordinate (38levels)	Terrain following sigma vertical coordinate (38levels)
Radiation	GFDL/GFDL	GFDL/GFDL
Surface layer	Janjic scheme	Janjic scheme
Land surface	Noah land surface	Noah land surface
Convective	Grell-Devenyi	Grell-Devenyi
PBL parameterization	Mellor-Yamada-Janjic	Mellor-Yamada-Janjic
Microphysics	Ferrier (new eta) scheme	Ferrier (new eta) scheme

Output from each model is post-processed to bring them back to a common format that enables direct comparison. The WPP vertically interpolates output from each model onto isobaric surfaces, diagnoses various fields not directly computed by the models, and generates a GRIB file on the model's native projection (rotated latitude longitude for the NMM model and mercator for the ARW model). NCEP's "product generator" horizontally interpolates the data from each model onto a common grid used for visualization and verification. The hourly observations of AWS data, DWR imageries over Kolkata, Kalpana satellite imageries and rain gauge observations from IMD are used in this study for model validation. The details of thunderstorm cases are described in Chapter 3.

## **5.2 Result and Discussion**

The ARW and NMM model have become popular for various applications. Several researches related to comparison of impacts of mesoscale dynamic cores (NMM and ARW) over US have been performed (Gallus 2006). But only single study is available including both ARW and NMM model over Indian region (Pattanayak et al. 2008), which is for cyclone prediction. The main objective of this work is to analyze the skills of NWP models as drivers of thunderstorm modeling system over east Indian region. For that purpose, NMM and ARW models are evaluated during 3 severe thunderstorm cases and the thunderstorm



affected parameters from model outputs are analyzed in the following section.

### **5.2.1 Analysis of stability indices**

Stability indices have been a corner stone in the forecasting of convection for many decades and often are used in the research literature as well. These indices are very helpful in predicting the severe weather events like thunderstorms. In the present study, an attempt is made to examine different stability indices of Kolkata (22.52<sup>0</sup>N, 88.37<sup>0</sup>E) obtained from NMM and ARW model in three thunderstorm days during May 2009. Table 5.2 shows the inter-comparison of NMM and ARW model simulated stability indices over Kolkata at 0000 and 1200 UTC. The NMM and ARW model simulated CAPE values are high and greater than the critical levels (1500 and 1000 Jkg<sup>-1</sup>) suggested by AiWS Technical Report (1990) and Tyagi et al. (2011) at 0000 and 1200 UTC of these three thunderstorm events, which is a favorable condition for severe thunderstorms. The mean of simulated CAPE values of ARW and NMM models at 0000 and 1200 UTC are found to be much larger than the critical level. The average CAPE values at 1200 UTC of both models are very high than 0000 UTC, which indicates that the models tend to simulate large CAPE values at 1200 UTC, when commonly thunderstorms occur. ARW and NMM model simulated LI values are less than critical level during all thunderstorm days. The mean of simulated LI values of both models at 0000 UTC are nearly the same and are less than

the critical level (-3). ARW model resulted mean LI values are lesser than NMM simulated LI value at 1200 UTC.

**Table 5.2:** Comparison of NMM and ARW model simulated stability indices for three thunderstorm events during May 2009.

Stability Indices	Critical Level	Critical Level Tyagi et al. (2011)	Thunder Storm Cases	0000 UTC		1200 UTC	
				NMM	ARW	NMM	ARW
CAPE	> 1500	> 1000	3 May	2947	3338	3361	3583
			11 May	3685	3455	3932	3963
			15 May	3033	3100	2993	3554
			<b>MEAN</b>	<b>3221.7</b>	<b>3297.7</b>	<b>3428.7</b>	<b>3700</b>
LI	$\leq -3$	$\leq -3$	3 May	-7	-8	-7	-8
			11 May	-10	-9	-10	-11
			15 May	-9	-8	-6	-8
			<b>MEAN</b>	<b>-8.7</b>	<b>-8.3</b>	<b>-7.7</b>	<b>-9.0</b>
TTI	> 44	> 46	3 May	50	49	49	50
			11 May	51	51	56	58
			15 May	50	50	43	47
			<b>MEAN</b>	<b>50.3</b>	<b>50.0</b>	<b>49.3</b>	<b>51.9</b>
KI	> 33	> 24	3 May	29	30	29	26
			11 May	28	27	39	36
			15 May	33	34	29	28
			<b>MEAN</b>	<b>30.0</b>	<b>30.3</b>	<b>32.3</b>	<b>30.0</b>

Both models captured a higher TTI values except for the third case. NMM model simulated TTI value (43) is below critical levels (44 and 46) for 15 May 2009 at 1200 UTC. In both time, model results show a high mean value (more than 48) of TTI, which is a favorable for severe thunderstorm occurrence. The mean of simulated KI values of both

models at 0000 UTC are very close to each other as in other stability indices. But ARW simulated value at 1200 UTC is very less than the critical level suggested by AiWS Technical report (1999), but greater than the value suggested by Tyagi et al. (2011).

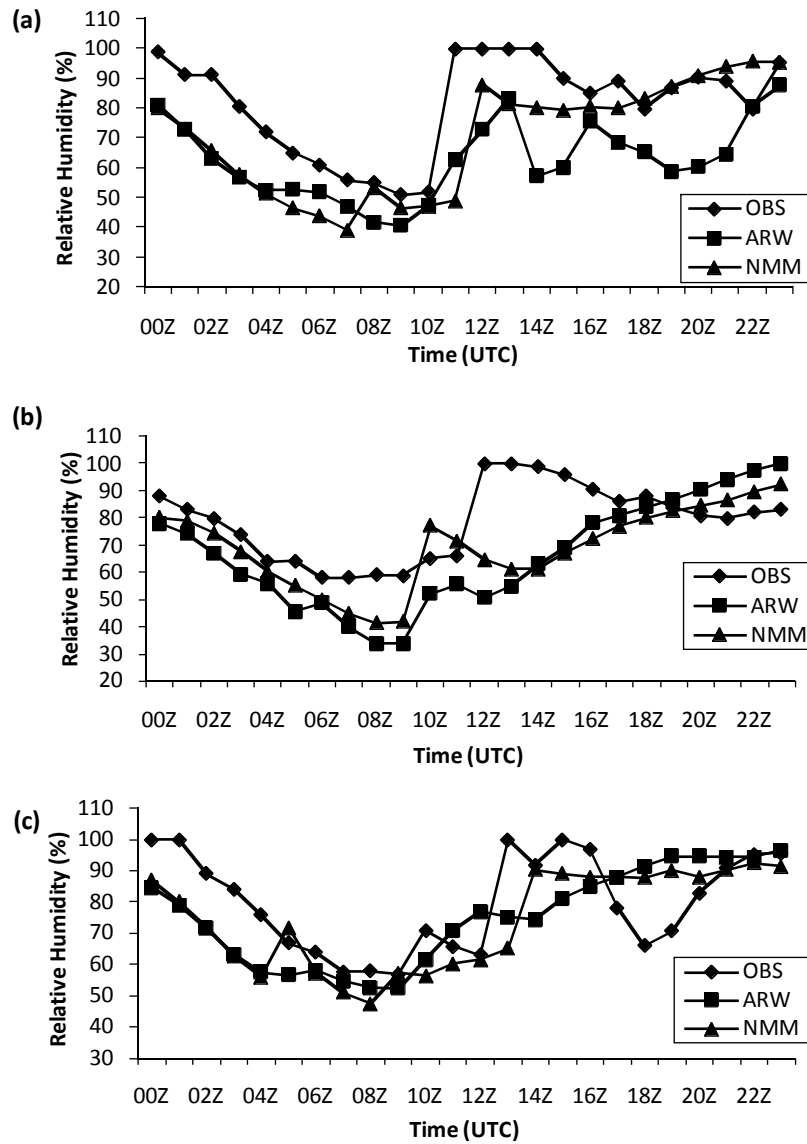
Examination of the model simulated stability indices for each thunderstorm day clearly indicated that both models have done well in capturing the instability of the atmosphere at 0000 and 1200 UTC for the occurrence of a severe thunderstorm. Thus model simulated thermodynamic structure over Kolkata becomes conducive for a thunderstorm occurrence.

### **5.2.2 Analysis of surface relative humidity and temperature**

Relative humidity at surface level has been taken into account, as it is an essential factor in intense convection. Figure 5.2 shows the inter-comparison of observed and model simulated relative humidity (%) using NMM and ARW model over Kolkata valid for 3, 11 and 15 May 2009 at 0000 UTC to next day at 0000 UTC. The observed relative humidity values for 3 May 2009 (Figure 5.2a) peaked from 52% to 100% (48%) at 1000 UTC whereas NMM model showed a sharp rise from around 49% to 88% (39%) at 1200 UTC, which is two hour later than that of the observed. ARW was not able to capture the sharp rise of relative humidity during the thunderstorm hour as in the NMM model. In the second case (Figure 5.2b), observed relative humidity showed a rise from 66% to

100% (34%) at 1200 UTC, whereas NMM simulation showed a rise from 42% to 77% (35%) at 1000 UTC, which is two hour prior than that of the observed. ARW model simulation showed an increase from 34% to 52% (18%) at 1000 UTC. A sharp increase of 35% has been captured by NMM model as in the observed rise of 34%. ARW model is able to capture the rise with less intensity. In the third case (Figure 5.2c), observed relative humidity peaked from 63% to 100% (37%) at 1300 UTC, whereas NMM model showed a sharp rise from 65% to 91% (26%) at 1400 UTC, which is one hour later than that of the observed. In this thunderstorm case also ARW model was not able to capture the sharp rise of relative humidity during the thunderstorm hour as in the NMM model. For all the thunderstorm cases (Figure 5.2), NMM model has captured the sudden rise of relative humidity values during the model simulated thunderstorm hour as in the observations.

A statistical analysis based on MAE, RMSE and CC was performed for comparisons between the simulated and observed surface relative humidity over Kolkata for 3 thunderstorm cases and are given in Table 5.3. From the table, it can be clearly seen that, simulated relative humidity of NMM model has more correlation with observation than ARW model for all 3 thunderstorm cases. The CC of NMM model is more than 0.7 for first and third thunderstorm cases and has an average CC value close to 0.7. The CC value of ARW model is very less for second and third cases and has a less average value.



**Figure 5.2:** Inter-comparison of NMM and ARW model simulated and observed diurnal variation of surface relative humidity (%) over Kolkata on (a) 3 May 2009 (b) 11 May 2011 (c) 15 May 2011.

The MAE and RMSE of each model have been computed to test the prediction error of the model. The MAE and RMSE values of relative humidity are less for NMM model than ARW for all 3 thunderstorm cases. The average errors are high for ARW model and the values are 16.21% and 19.04%.

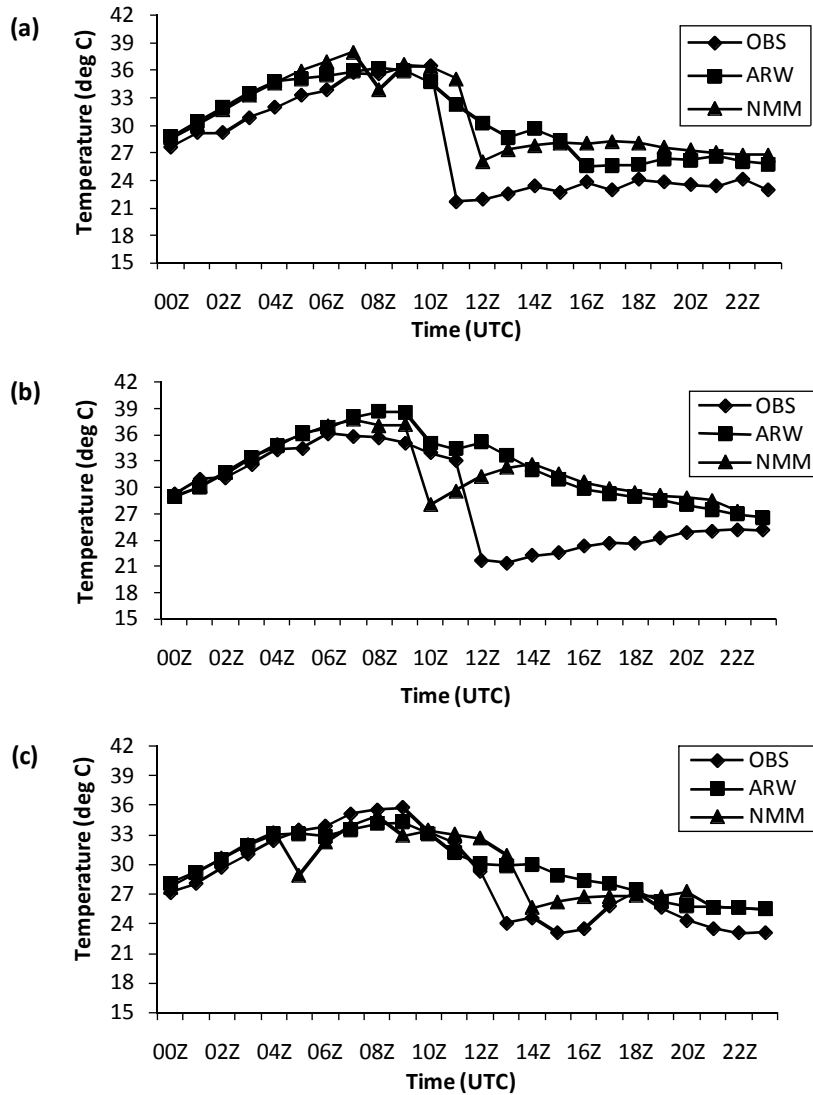
**Table 5.3:** Statistical analysis of simulated and observed surface relative humidity over Kolkata based on MAE, RMSE and CC.

Statistical Analysis	CASES	NMM	ARW
MAE	3-May	13.36	18.97
	11-May	13.11	17.18
	15-May	10.89	12.49
	<b>MEAN</b>	<b>12.45</b>	<b>16.21</b>
RMSE	3-May	17.37	21.62
	11-May	17.00	20.85
	15-May	13.80	14.65
	<b>MEAN</b>	<b>16.05</b>	<b>19.04</b>
CC	3-May	0.76	0.72
	11-May	0.53	0.49
	15-May	0.72	0.57
	<b>MEAN</b>	<b>0.67</b>	<b>0.59</b>

Surface temperature is useful parameter in forecasting the likelihood occurrence of a thunderstorm. Figure 5.3 shows the inter-comparison of observed and model simulated temperature ( $^{\circ}\text{C}$ ) using NMM and ARW model over Kolkata valid for 3, 11 and 15 May 2009 at 0000 UTC to next day at 0000 UTC. The observed temperature (Figure 5.3a) showed a sudden fall from  $36.7^{\circ}\text{C}$  to  $21.7^{\circ}\text{C}$  ( $15^{\circ}\text{C}$ ) at 1000 UTC

whereas NMM model showed a fall from 35.1<sup>0</sup>C to 26.1<sup>0</sup>C (9<sup>0</sup>C) at model predicted hour. For the second case (Figure 5.3b), observed temperature showed a drop from 33.1<sup>0</sup>C to 21.7<sup>0</sup>C (11.4<sup>0</sup>C) at 1200 UTC, whereas NMM simulation shows a drop from 37.1<sup>0</sup>C to 28<sup>0</sup>C (9.1<sup>0</sup>C) at 1000 UTC. In the third case, the observed temperature (Figure 5.3c) showed a sudden fall from 29<sup>0</sup>C to 24<sup>0</sup>C (5<sup>0</sup>C) at 1300 UTC whereas NMM model showed a fall from 31<sup>0</sup>C to 26<sup>0</sup>C (5<sup>0</sup>C) at model predicted hour of 1400 UTC. In all three thunderstorm cases ARW model failed to capture the sudden temperature fall over Kolkata as in NMM model.

A statistical analysis based on MAE, RMSE and CC was performed for comparison between the simulated and observed surface temperature over Kolkata for 3 thunderstorm cases and are given in Table 5.4. NMM model have more correlation with observation than ARW model for all 3 thunderstorm cases. The CC of NMM model is more than 0.8 for first and third thunderstorm cases as in the relative humidity and have an average CC value more than 0.8. The CC value of ARW model is also greater than 0.8 for first and third cases and an average is less than 0.8 due to less correlation in second case. The MAE and RMSE of each model have been computed to test the prediction error of the model. The MAE and RMSE values of temperature are less for NMM model than ARW for all 3 thunderstorm cases. The average errors are less for NMM model than ARW and the values are 2.91<sup>0</sup>C and 3.86<sup>0</sup>C.



**Figure 5.3:** Inter-comparison of NMM and ARW model simulated and observed diurnal variation of surface temperature ( $^{\circ}$ C) over Kolkata on (a) 3 May 2009 (b) 11 May 2011 (c) 15 May 2011.



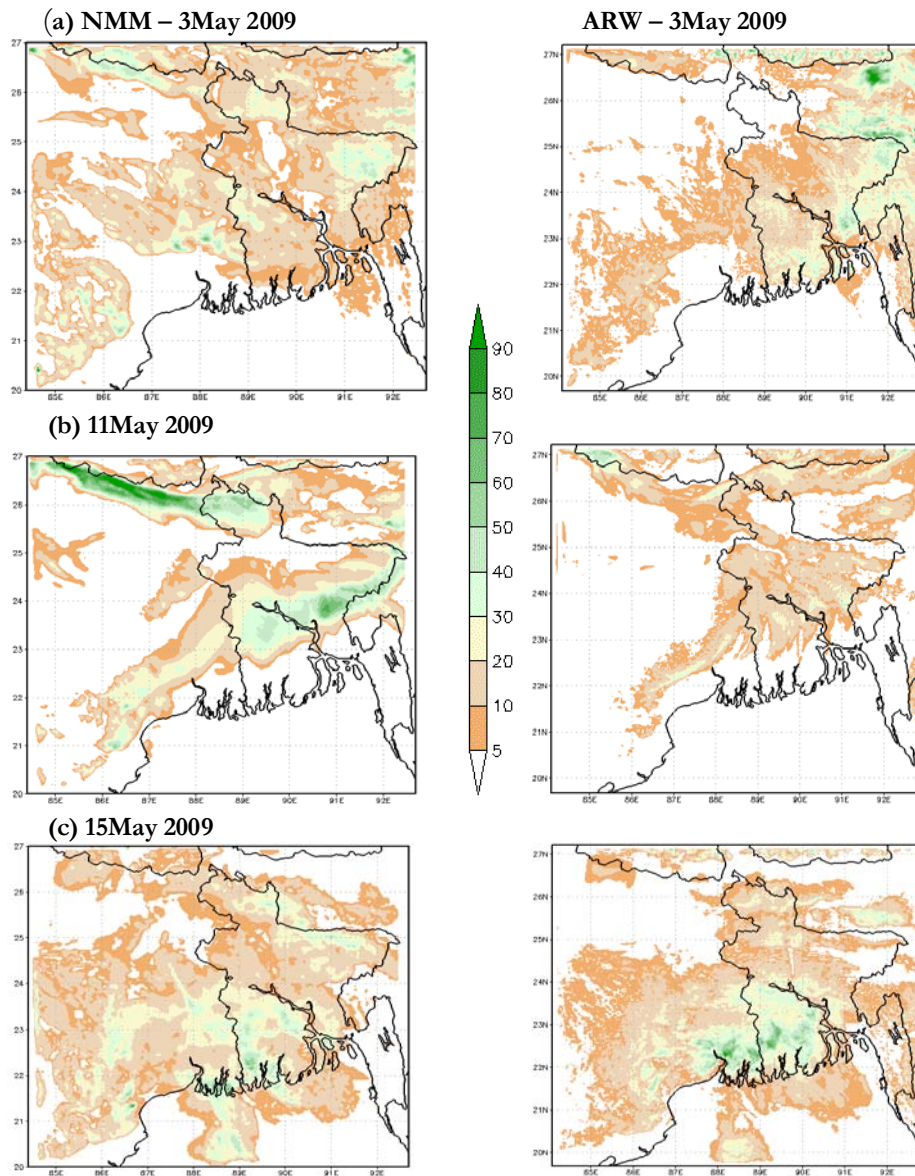
**Table 5.4:** Statistical analysis of simulated and observed surface temperature over Kolkata based on MAE, RMSE and CC.

Temperature	CASES	NMM	ARW
MAE	3-May	3.03	3.45
	11-May	3.78	3.99
	15-May	1.92	1.93
	<b>MEAN</b>	<b>2.91</b>	<b>3.12</b>
RMSE	3-May	3.94	4.28
	11-May	5.19	5.28
	15-May	2.47	2.58
	<b>MEAN</b>	<b>3.86</b>	<b>4.05</b>
CC	3-May	0.86	0.86
	11-May	0.69	0.62
	15-May	0.87	0.85
	<b>MEAN</b>	<b>0.81</b>	<b>0.77</b>

Comparison of the surface parameters simulated by both models indicated the superiority of NMM model in simulating thunderstorms over Kolkata even though one or two hour lead or lag exists.

### 5.2.3 Analysis of precipitation

Precipitation is recognized as one of the most difficult parameters to forecast in NWP. Most of the thunderstorms produce heavy rainfall during their lifecycle of 1-3 hours. The comparison of 24 h accumulated NMM and ARW model simulated rainfall with three thunderstorm cases are plotted in Figure 5.4. The overall rainfall distribution is reasonably well captured by the NMM model for all 3 thunderstorm cases. The rainfall spread simulated by ARW model is less for first two cases.



**Figure 5.4:** Comparison of NMM and ARW simulated 24 h accumulated rainfall during 3 thunderstorm events (a) 3 May 2009 (b) 11 May 2009 and (c) 15 May 2009.

For third case, both model well captured the rainfall spread over West Bengal region. The maximum intensity of NMM model simulated rainfall over West Bengal is 30 to 50 mm for first two cases. But ARW failed to capture this intensity over West Bengal region in first two thunderstorm cases. The ARW model was able to capture the rainfall spread as in NMM model over this region for third thunderstorm cases.

A quantitative assessment of ARW and NMM model simulated rainfall data and surface rain gauge observations of 6 stations over West Bengal region for three thunderstorm cases has been analyzed. The precipitation is accumulated for up to 24 h, starting from 0000 UTC of each day and ending at 0000 UTC of the following day. Comparison of modeled precipitation with rain gauge station observations for all three thunderstorm days is given in Table 5.5. Both models have well simulated the rainfall amount with NMM performing better than ARW as indicated in the table. It can also be seen from the values in Table 5.5 that NMM model had predicted the rainfall amount better than ARW on 3 May 2009 at the stations Dum Dum, Bankura, Basirhat and Balasore while ARW's predictions were better at Sriniketan and Jamshedpur. The average rainfall of NMM model from six rain gauge stations are more than ARW model. On 11 May 2009, NMM model has done better than ARW model in simulating rainfall at all 6 stations. ARW model simulated very less rainfall in 5 stations and over predicted the rainfall amount at Kharagpur. NMM simulated average rainfall on this day is very good as compared to ARW. On 15 May 2009, both

models have well simulated the rainfall amount as in the rainfall distribution plot. The average rainfall value of third case is more for ARW (21.65) model than NMM (20.98) as indicated in Table 5.5. Although ARW has done well occasionally in simulation of rainfall at these stations, the overall performance was better with NMM model.

**Table 5.5:** Comparison of modeled precipitation of three thunderstorm cases with rain gauge observations.

DATE	STATION	LAT	LONG	OBS	NMM	ARW
3 May	Dum Dum	22.39	88.27	31.40	23.26	6.56
	Bankura	23.13	87.04	24.90	14.73	12.62
	Basirhat	22.40	88.53	21.20	14.12	12.14
	Sriniketan	23.39	87.42	38.20	26.06	35.24
	Balasore	21.30	86.56	43.30	31.04	11.80
	Jamshedpur	22.44	86.12	35.80	15.49	32.00
	<b>MEAN</b>			<b>32.47</b>	<b>20.78</b>	<b>18.39</b>
11 May	Dum Dum	22.39	88.27	33.30	23.10	12.48
	Bankura	23.22	87.07	22.00	15.13	3.44
	Canning	22.25	88.67	26.40	21.00	12.87
	Basirhat	22.4	88.53	48.40	24.75	18.74
	Digha	21.83	87.8	24.40	10.08	0.00
	Kharagpur	22.2	87.19	16.80	11.99	19.31
	<b>MEAN</b>			<b>28.55</b>	<b>17.68</b>	<b>11.14</b>
15 May	Dum Dum	22.39	88.27	16.90	17.19	35.30
	Bankura	23.13	87.04	34.00	14.69	20.40
	Krishnagar	23.24	88.31	19.60	18.72	17.17
	Digha	21.5	87.48	21.00	18.39	18.40
	Midnapore	22.25	87.19	51.60	26.54	17.70
	Haldia	22.04	88.04	33.20	30.39	21.20
	<b>MEAN</b>			<b>29.38</b>	<b>20.98</b>	<b>21.69</b>

In order to analyze the modeled precipitation, statistical analysis has been done by calculating the CC, RMSE and MAE which is given in Table 5.6. All statistical parameters are calculated by taking precipitation values of six rain gauge stations for three thunderstorm cases together. NMM model's superior performance is witnessed with high correlation coefficient of 0.565, better than that of ARW. Further it can be seen that RMSE, MAE of NMM are less than that of ARW indicating better efficiency of NMM model in predicting rainfall at different stations. The statistical analysis showed that NMM model's predicted rainfall amounts are closer to that of the observed in comparison with that of ARW. So NMM model has outperformed ARW in rainfall prediction and is superior out of two models.

**Table 5.6:** Statistical analysis of simulated and observed precipitation for three thunderstorm cases.

Statistical Analysis	Description	NMM	ARW
MAE	$MAE = \frac{1}{N} \sum_{i=1}^N  f_i - o_i $	10.905	15.379
RMSE	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (f_i - o_i)^2}$	13.785	18.464
CC	$CC = \frac{\sum(f_i - \bar{f})(o_i - \bar{o})}{\sqrt{\sum(f_i - \bar{f})^2 \sum(o_i - \bar{o})^2}}$	0.565	0.121

#### 5.2.4 Analysis of composite radar reflectivity

DWR is being used worldwide for the study of various severe weather phenomena like thunderstorms, hailstorms, tornadoes and cyclones. In other words, it can measure how fast rain or hail is moving towards or

away from the radar. From a volume scan (a series of  $360^\circ$  sweeps, each tilting a little higher than the last); forecasters can get a detailed look at structures and movements in storms close to the radar (Chatterjee et al. 2008). Kolkata DWR composite radar reflectivity imageries on 3 May 2009 from 1000 to 1300 UTC is shown in Figure 5.5. By analyzing Kolkata DWR composite radar reflectivity (dBZ) imageries on 3 May 2009, a strong echo was developed northwest of Kolkata (Ranchi (RNC)) at 0900 UTC. This echo intensified into north-south oriented squall line by 1000 UTC (Figure 5.5a) and gradually moved towards Kolkata at 1100 UTC (Figure 5.5b). This echo was over Kolkata at 1300 UTC (Figure 5.5d) and disappeared at 1400 UTC (Mohanty et al. 2009).

NMM model simulated composite radar reflectivity (dBZ) on 3 May 2009 from 1000 to 1300 UTC is shown in Figure 5.6. By analyzing NMM model simulated composite radar reflectivity plots, a squall line developed northwest of Kolkata at 1000 UTC. This squall line was moved towards Kolkata at 1100 UTC and was over Kolkata at 1300 UTC as in the DWR imageries. ARW model simulated composite radar reflectivity (dBZ) on 3 May 2009 from 1000 to 1300 UTC is shown in Figure 5.7. ARW model simulated composite radar reflectivity plots also show a squall line, which developed northwest of Kolkata at 1000 UTC as in NMM model. This squall line was moved towards Kolkata at 1100 UTC, but didn't reach Kolkata at 1300 UTC, which indicating the slow movement of the squall line. The squall line movement and intensity was well captured by NMM than ARW.

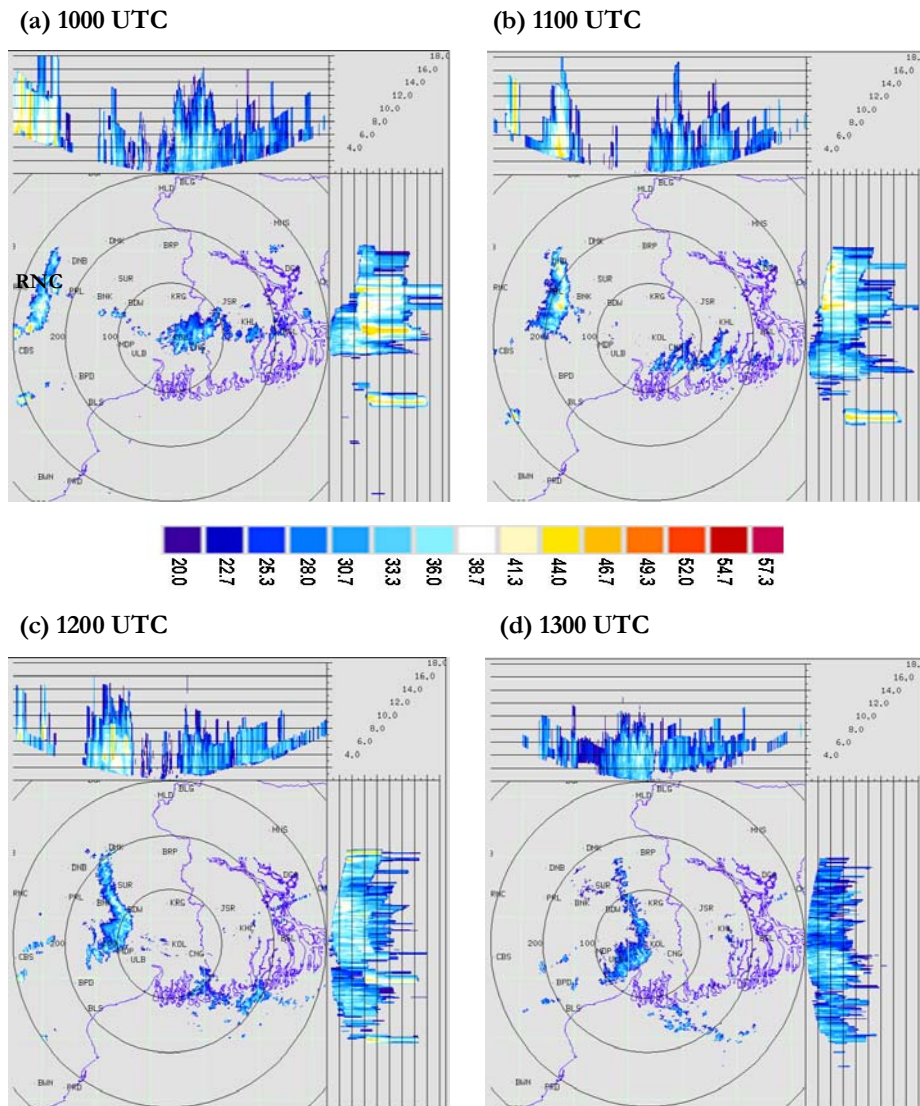
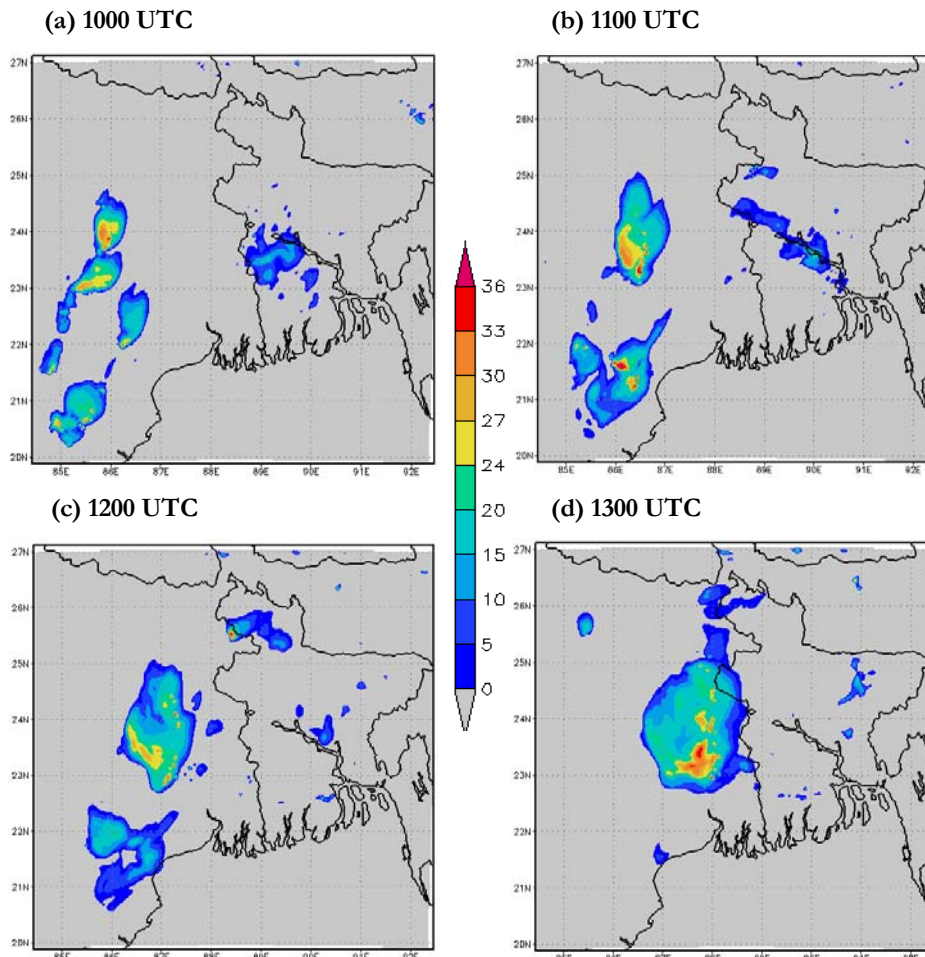


Figure 5.5: Kolkata DWR composite radar reflectivity (dBZ) imageries from 1000 to 1300 UTC on 3 May 2009.



**Figure 5.6:** NMM model simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 3 May 2009.

Kolkata DWR imageries from 1000 to 1300 UTC on 11 May 2009 are given in Figure 5.8. By analyzing Kolkata DWR imageries of 11 May 2009, a strong echo was developed northeast of Kolkata at 1000 UTC, which



intensified into west-east oriented squall line by 1100 UTC. Another strong echo was developed at the northwest of Kolkata at 1100 UTC. These two echoes are merged at 1200 UTC and become intensified. This echo gradually moved towards Kolkata at 1300 UTC (Mohanty et al. 2009). Both NMM (Figure 5.9) and ARW (Figure 5.10) models are failed to capture two strong echoes in their plots. They are able to simulate one echo which was initiated from northeast of Kolkata at 1000 UTC as in observation. It was intensified and moved towards Kolkata at 1100 UTC (Figure 5.9). NMM model well captured this squall line movement as compared to ARW model (Figure 5.10) even though the magnitude of composite radar reflectivity simulated by NMM model is less.

By analyzing Kolkata DWR imageries on 15 May 2009 (Figure 5.11), a strong echo was developed near Purulia (PRL) at 1000 UTC, which intensified into north-south oriented squall line by 1100 UTC. This echo gradually moved towards Kolkata at 1200 UTC. This echo was over Kolkata at 1300 UTC and disappeared at 1500 UTC (Mohanty et al. 2009). NMM model simulated composite radar reflectivity on 15 May 2009 from 1000 to 1300 UTC is shown in Figure 5.12. By analyzing NMM model simulated composite radar reflectivity plots, a squall line developed northwest of Kolkata at 1000 UTC. This squall line was moved towards Kolkata at 1100 UTC and was over Kolkata at 1300 UTC as in the DWR imageries. By analyzing ARW model simulated composite radar reflectivity pictures (Figure 5.13), a squall line developed northwest of Kolkata at 1000 UTC as

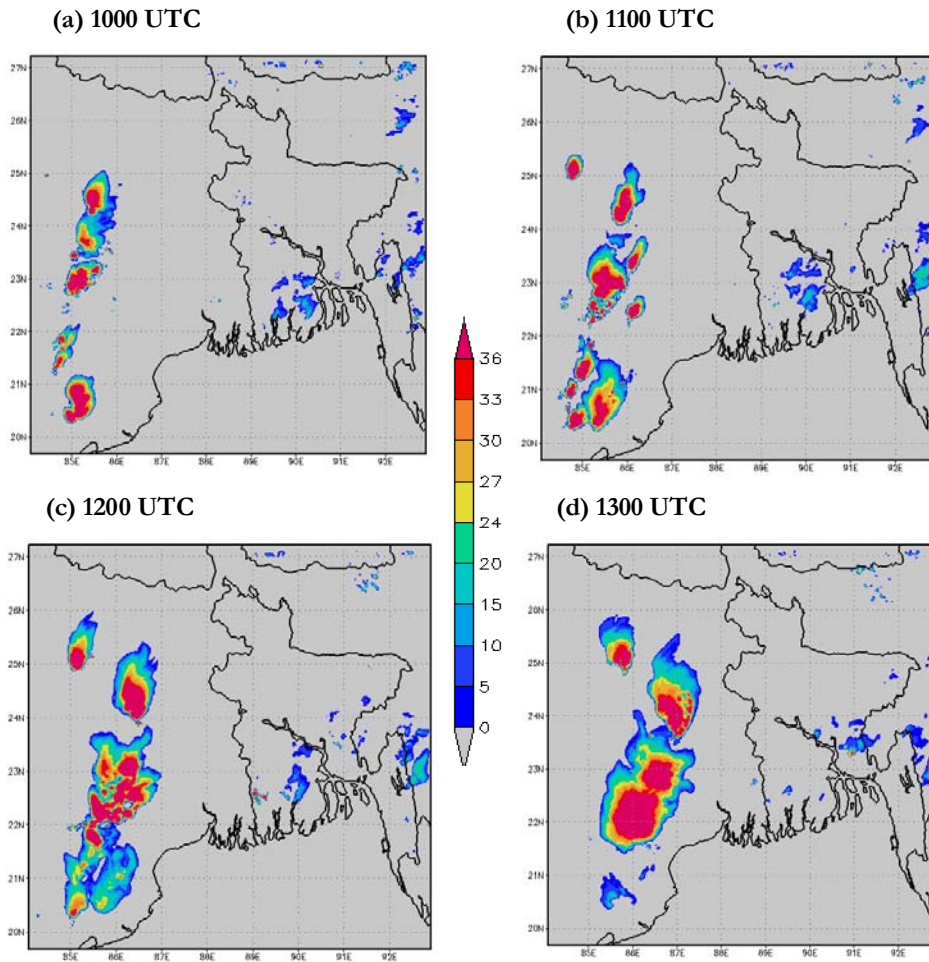
in NMM model. This squall line was moved towards Kolkata at 1100 UTC. This echo was not reached over Kolkata by 1300 UTC as in the DWR imageries and NMM simulated outputs. ARW model simulated more intensity than NMM model as in the previous case. However it is seen that the movement of the squall line was slow in ARW as compared to that of the observed.

From the present analysis of the simulated composite radar reflectivity, it can be concluded that NMM model has reasonably well simulated genesis, intensification and propagation of three severe thunderstorms during 2009 pre-monsoon season over east Indian region as in the DWR radar reflectivity imageries. ARW model well simulated the thunderstorm initiation, while the squall line movement was slow.

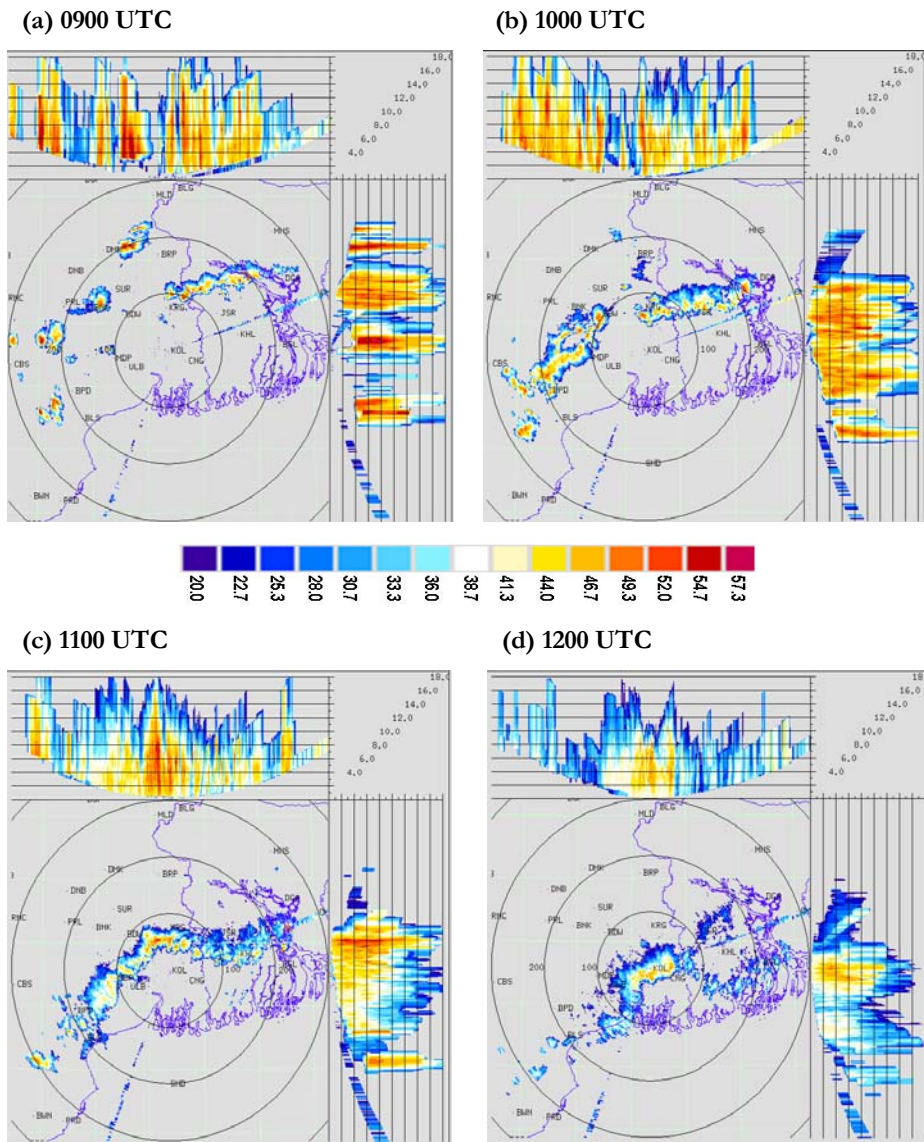
### **5.2.5 Analysis of cloud top temperature**

The ability to accurately forecast cloudiness is necessary in the fields of aviation. In recent years, brightness temperature and cloud top temperature derived from NWP model output have been used to demonstrate the advanced capabilities of these models for severe weather prediction (Otkin and Greenwald 2008). In this section, the ability of NMM and ARW model to realistically simulate the cloud top temperature (CTT) ( $^{\circ}\text{C}$ ) over east Indian region are examined. The comparisons of Kalpana satellite derived and model simulated CTT are presented here. Kalpana satellite is a dedicated, meteorological geostationary Indian

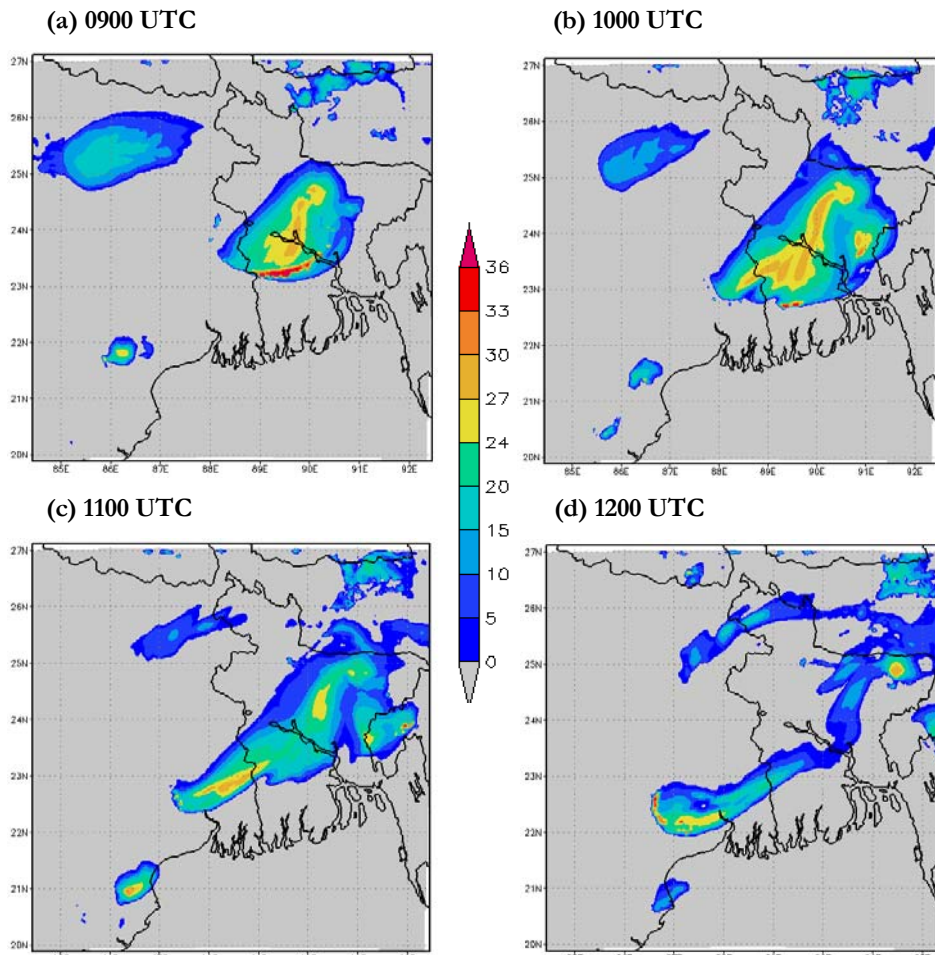
satellite launched by Geo-Stationary Launch Vehicle (GSLV) and operating since 24 September 2002.



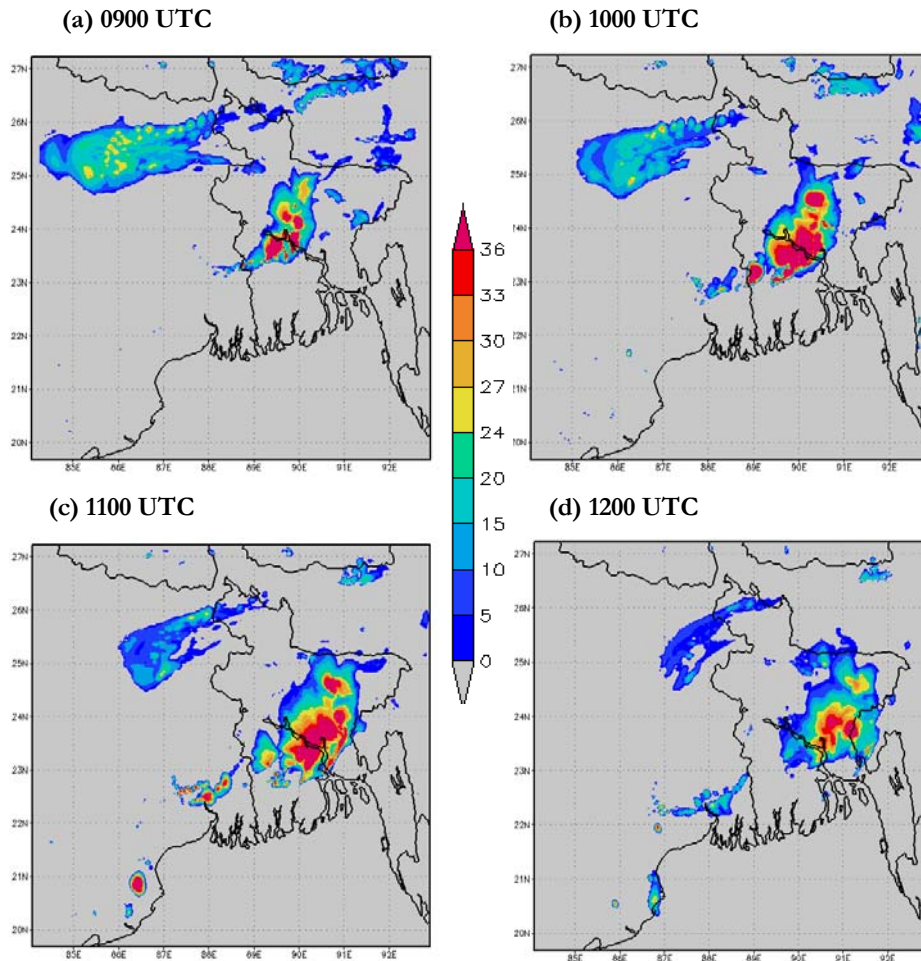
**Figure 5.7:** ARW model simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 3 May 2009.



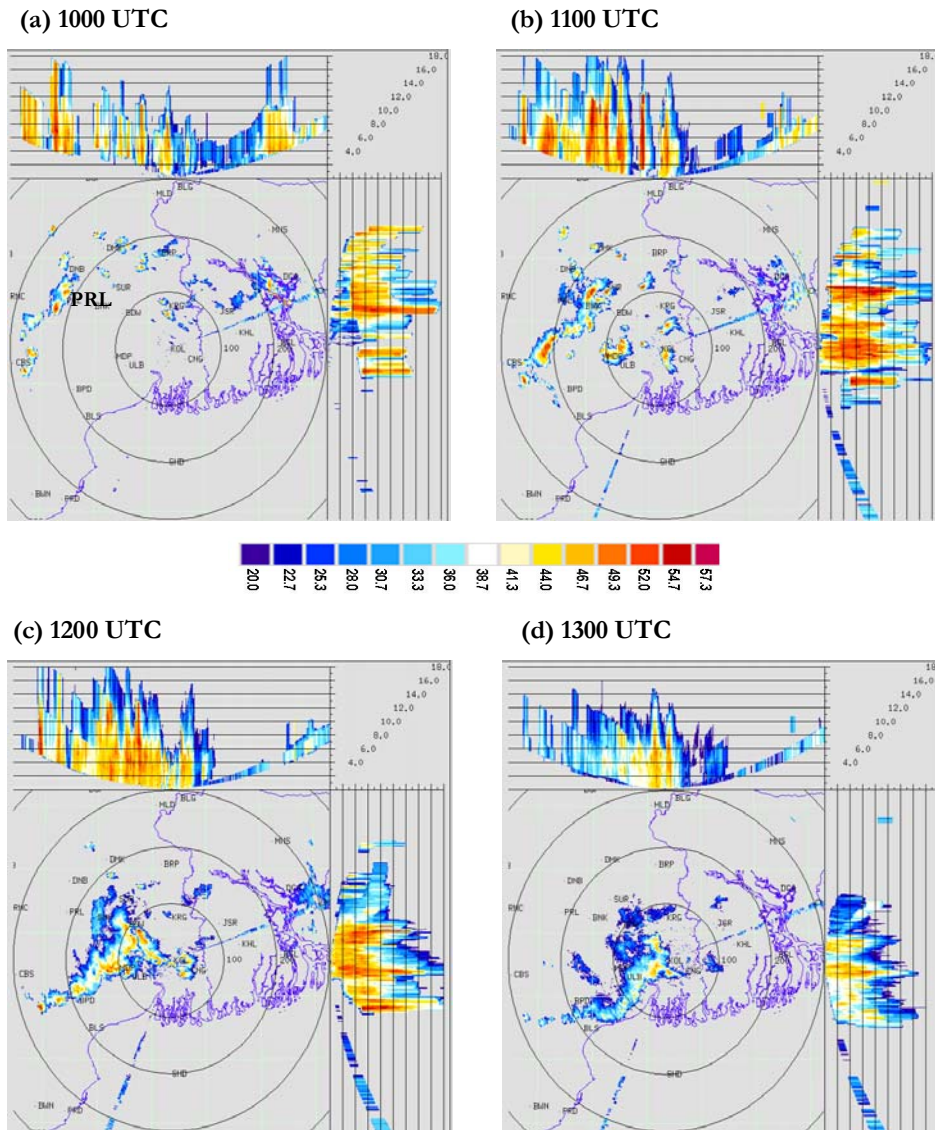
**Figure 5.8:** Kolkata DWR composite radar reflectivity (dBZ) imageries from 0900 to 1200 UTC on 11 May 2009.



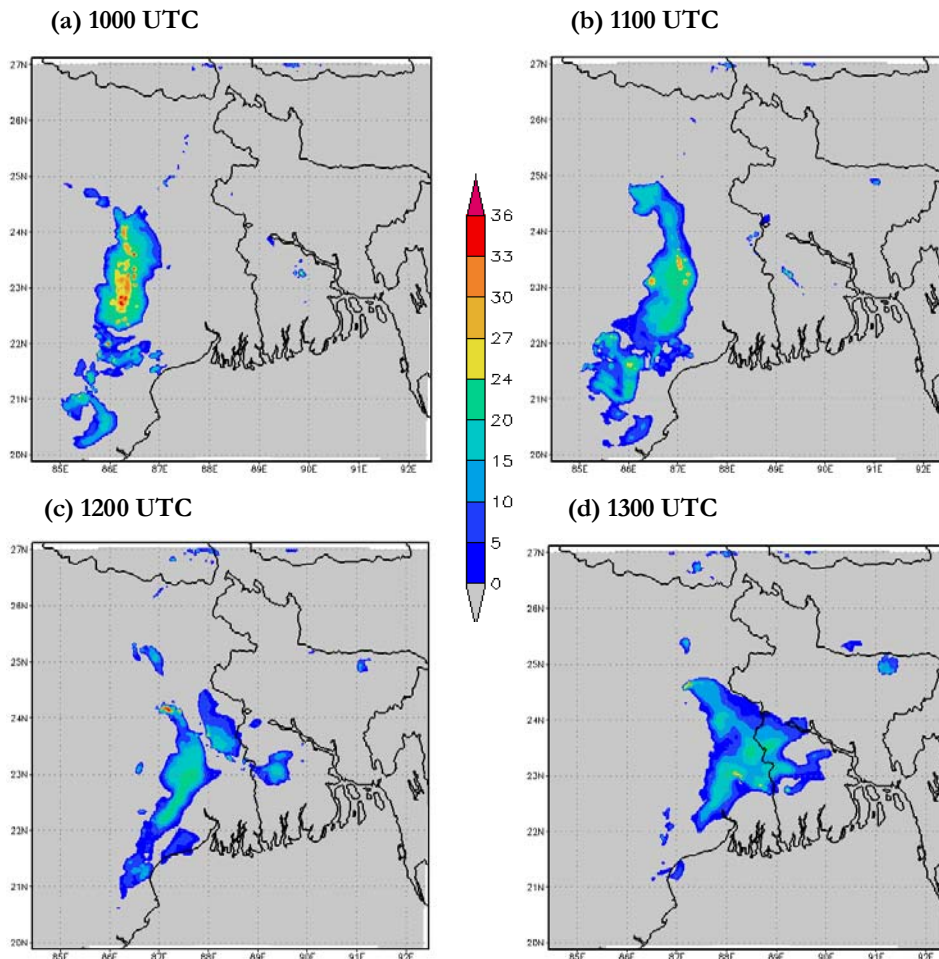
**Figure 5.9:** NMM model simulated composite radar reflectivity (dBZ) pictures from 0900 to 1200 UTC on 11 May 2009.



**Figure 5.10:** ARW model simulated composite radar reflectivity (dBZ) pictures from 0900 to 1200 UTC on 11 May 2009.

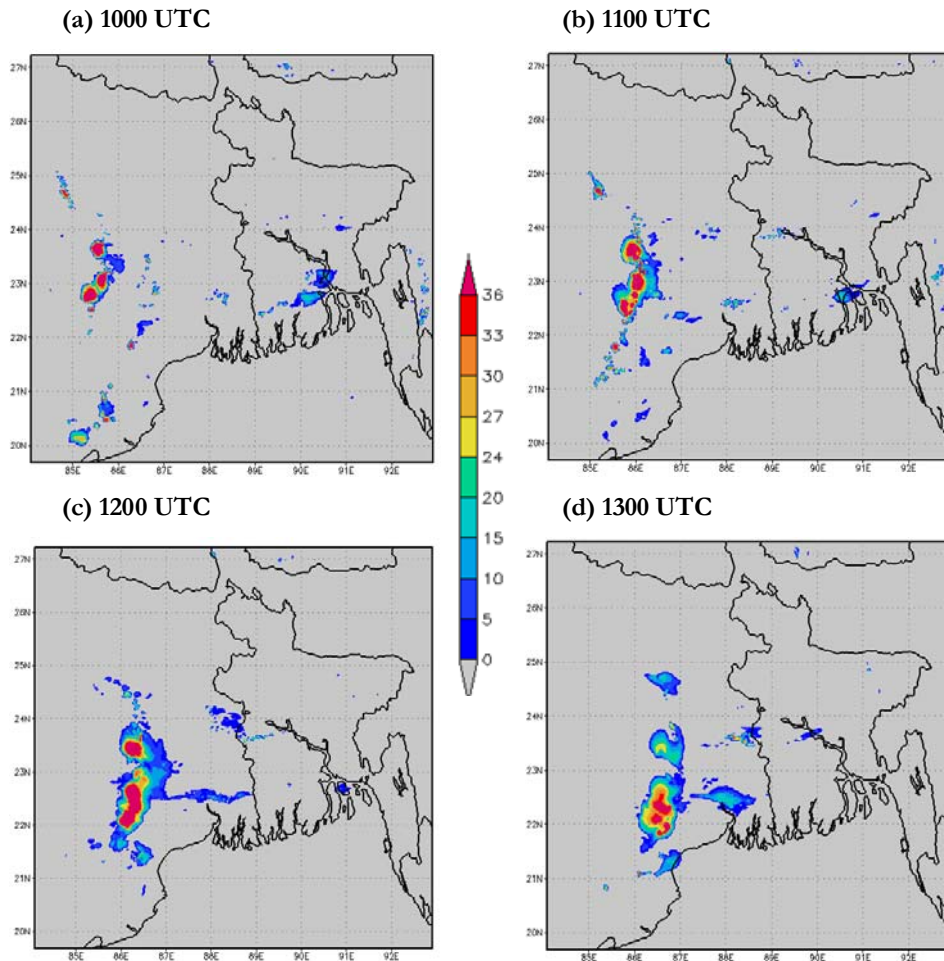


**Figure 5.11:** Kolkata DWR composite radar reflectivity (dBZ) imageries from 1000 to 1300 UTC on 15 May 2009.



**Figure 5.12:** NMM model simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009.



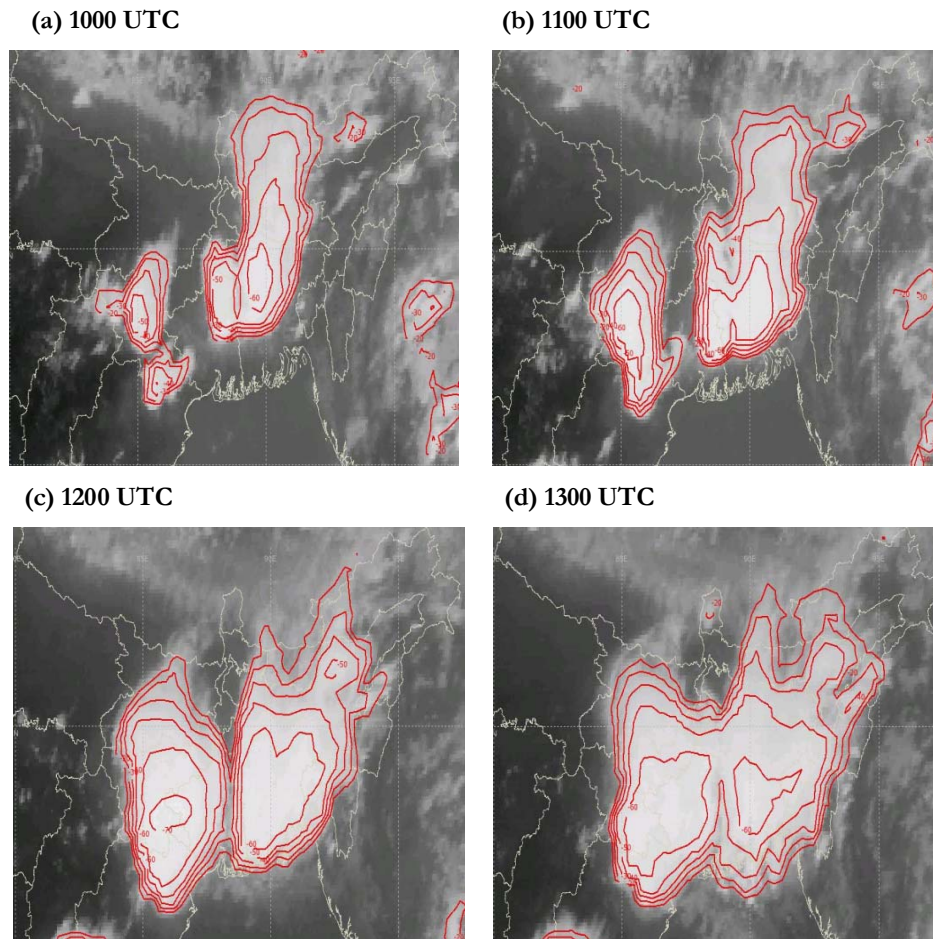


**Figure 5.13:** ARW model simulated composite radar reflectivity (dBZ) pictures from 1000 to 1300 UTC on 15 May 2009.

This geostationary satellite carries onboard a Very High Resolution Radiometer (VHRR) along with other instruments. This sensor operates in three wavelengths bands, viz. (i) thermal IR band (TIR): 10.5–

12.5  $\mu\text{m}$ , (ii) visible band (VIS): 0.55–0.75  $\mu\text{m}$  and (iii) water vapour band (WV): 5.7–7.1  $\mu\text{m}$ . In WV and TIR bands, the spatial resolution is 8 km whereas in VIS band spatial resolution is 2 km. The details of the Kalpana satellite are given by Kaila et al. (2002). This product indicates the contours of the CTT of the IR channel. The range of values for which contours are drawn is  $-20$  to  $-80^{\circ}\text{C}$ .

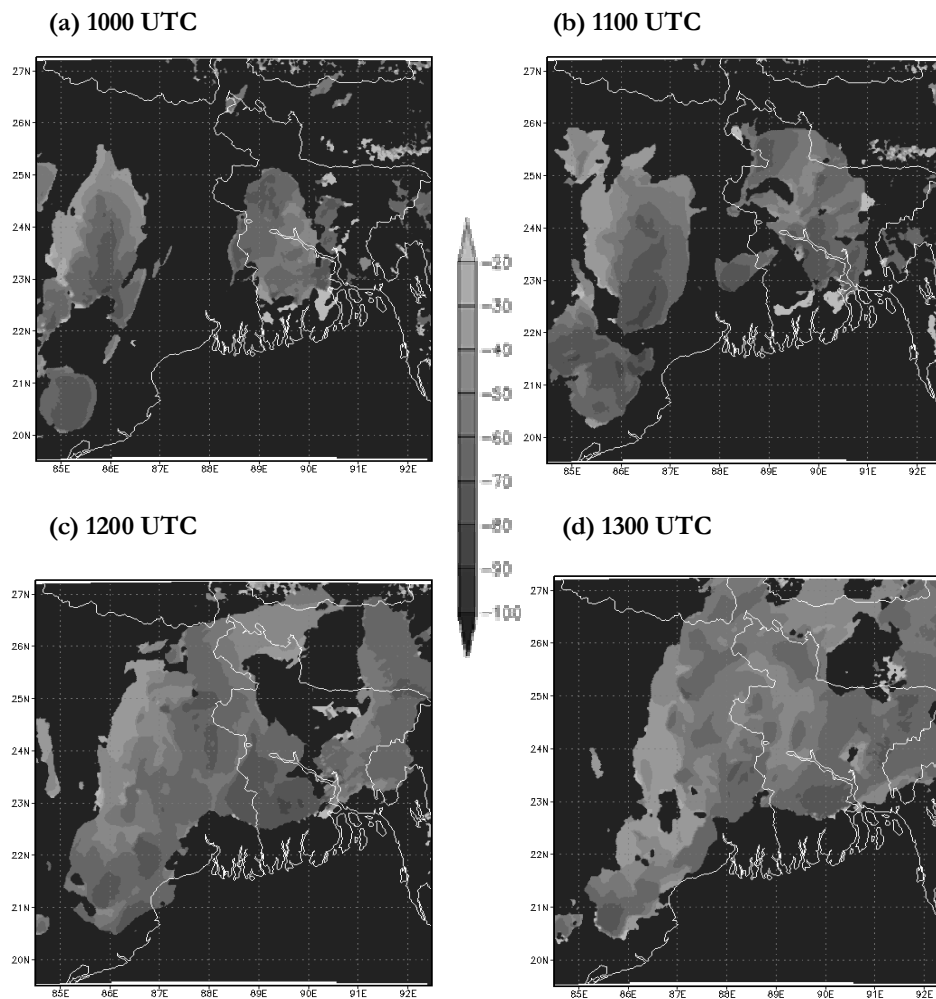
Kalpana satellite derived CTT ( $^{\circ}\text{C}$ ) imageries from 1000 to 1300 UTC on 3 May 2009 are shown in Figure 5.14. The satellite imageries (Figure 5.14) of this thunderstorm case show that two convective cells developed over Bangladesh (northeast of Kolkata) and Jharkhand (northwest of Kolkata) at 1000 UTC. These cells expanded and merged over West Bengal by 1200 UTC and reached a maximum CTT of  $-60^{\circ}\text{C}$ . This cell is more intensified at 1300 UTC and reached upto  $-70^{\circ}\text{C}$  (Mohanty et al. 2009). The NMM model simulated CTT (Figure 5.15) also shows both convective cells over northeast and northwest of Kolkata at 1000 UTC. These cells are merged over West Bengal at 1200 UTC as in the satellite imageries. The model simulated CTT reached upto  $-70^{\circ}\text{C}$  during this cloud formation and movement. The ARW model simulated CTT show the cloud cluster over northwest of Kolkata as in the NMM model (Figure 5.16). The ARW model failed to capture convective cell over northeast of Kolkata as in NMM and observed imageries. The movement of this cloud cluster simulated by ARW model is slow as in DWR imageries.



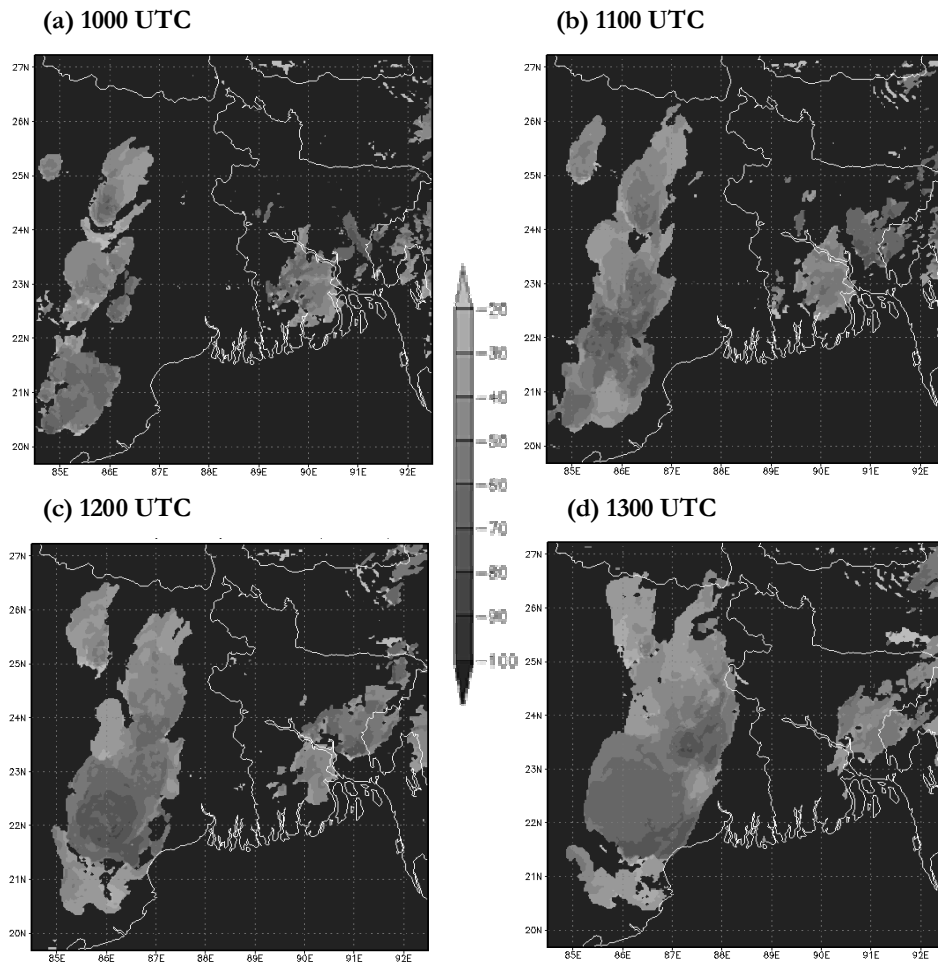
**Figure 5.14:** Kalpana satellite derived CTT ( $^{\circ}\text{C}$ ) imageries from 1000 to 1300 UTC on 3 May 2009.

The NMM model simulated CTT for other two cases also show cloud clusters over West Bengal region as in observations. But ARW model failed to represent the cloud clusters as in observations (Figures are not shown). The convection diagnosed by the CTT from NMM model

appears to be fairly representative of the structure and intensity observed in Kalpana satellite imageries.



**Figure 5.15:** NMM model simulated CTT ( $^{\circ}\text{C}$ ) from 1000 to 1300 UTC on 3 May 2009.



**Figure 5.16:** ARW model simulated CTT ( $^{\circ}$ C) from 1000 to 1300 UTC on 3 May 2009.

### 5.3 Chapter Summary

In this chapter, an attempt has been made to compare the simulated results of three thunderstorm events during May 2009, using NMM and ARW model and validated the model results with observations.

Analysis of the stability indices simulated by both models clearly indicate that ARW and NMM models have performed well in simulating the different thermodynamic indices such as CAPE, LI, TTI and KI at 1200 UTC which is very much favorable for thunderstorm occurrence. Comparison of model simulated thunderstorm affected surface parameters with that of the observed revealed that NMM has performed better than ARW in capturing the sharp rise in humidity and drop in temperature even though one or two hour lag or lead exists. ARW model has failed to capture the rise and drop in humidity and temperature respectively. The statistical analysis of surface temperature and relative humidity indicate the superiority of NMM model in simulating the thunderstorm over Kolkata on these severe thunderstorm cases.

The precipitation forecasts have been analyzed both spatially and temporally. The NMM model performed well for first two thunderstorm cases. Both models are good and the results are very close to each other in the third case. But statistical analysis of precipitation clearly indicate that NMM model has done better with high CC than that of the ARW and also with low RMSE and MAE. Comparison of model simulated radar

reflectivity with that of the observed revealed that both models have done well in simulating the initiation of squall lines. NMM model has simulated well the propagation of the squall lines, which is in good agreement with that of the observed, while the squall line movement was slow in ARW. The NMM model simulated CTT appears to be fairly representative of the structure and intensity observed in satellite imageries than ARW model. From the above results, it can be concluded that NMM model has better capability in prediction of thunderstorms over east Indian region.







## Chapter-6

# EVALUATION OF COMPUTATIONAL MODELS FOR THUNDERSTORM PREDICTION

---

### Contents

- 6.1 Data and Methodology
  - 6.2 Results and Discussion
  - 6.3 Chapter Summary
- 

Accurate and timely thunderstorm forecasts are essential for efficient management of national weather that mitigate associated aviation hazards, including turbulence, icing, hail, and lightning. Forecasting of severe thunderstorms is a challenge due to complexity of the processes involved and interplay of many factors. Thunderstorm prediction approaches are challenged by complex weather phenomena with limited observations and past data. The techniques for predicting thunderstorms can be classified into two groups (Wilson et al. 1998). One method is a historical treatment of thunderstorm extrapolation techniques (knowledge-based expert systems including Fuzzy logic and ANN). The second method is prediction using high resolution NWP models.

Riordan and Hansen (2002) explains this classification as follows: The first method is based upon the happenings of comparable cases (i.e.

similar weather situations). This method is useful for predicting local-scale weather if recorded cases are plentiful (e.g. cloud ceiling and visibility in a few square kilometers around an airport). The second method is based upon equations of the atmosphere and is commonly referred to as computer modeling (NWP). The NWP models start with current weather observations and attempt to predict future weather, describing the physics and dynamics of atmosphere, mathematically. An accurate observation about what the weather is doing now is key to help predict, what it will do in the future. NWP models (i.e. MM5, ARW, NMM etc.) help in deciding, whether the conditions will be favorable for the development of thunderstorms or not. The prediction of thunderstorms is still subjected to forecaster's experience and interpretation of NWP models. The NWP methods have been found to dominate the literature, almost exclusively, for forecasts over 10 h ahead (Hart 2002; Landberg 1995). However, the published ANN methods based on observations appear to be more accurate over shorter periods (minutes to a few hours). They are also much simpler than the NWP methods.

In recent years, a large literature has evolved on the use of ANNs in many weather forecasting applications (Radhika and Shashi 2009; Guhathakurta 2006). Neural networks are particularly appealing because of their ability to model an unspecified non-linear relationship between weather parameters. ANN has capability to extract the relationship between the inputs and outputs of a process, without the physics being

explicitly provided (Zurada 1992). The development of ANN, which perform nonlinear mapping between inputs and outputs, has lately provided alternative approaches to forecast thunderstorms. ANN has several advantages. ANNs are complex and flexible nonlinear systems with properties not found in other modeling systems. These properties include robust performance in dealing with noisy or incomplete input patterns, high fault tolerance, and the ability to generalize from the input data (Patterson 1996). Neural networks offer a number of advantages, including requiring less formal statistical training, ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms. They are not computationally intensive as much as NWP systems. Disadvantages include its "black box" nature, greater computational burden, proneness to over fitting, and the empirical nature of model development.

NWP model forecasts have become an indispensable source of information in virtually every aspect of weather forecasting. Simulation of active mesoscale systems such as tropical cyclones, heavy rainfall and severe thunderstorms with high-resolution mesoscale models has been attempted by many Indian researchers recently (e.g. Mohanty et al. 2004; Vaidya 2007; Chatterjee et al. 2008; Rajeevan et al. 2010). The level of detail in modern models allows for a wide variety of products and forecast

fields to be delivered, for use not only in general meteorology, but also in specialized areas such as aviation and air quality. Advances in data visualization and delivery methods show great promise for the users of meteorological products and the practitioners of the science. However, managing the mass of forecast data created by the models is fast becoming a science of its own. While NWP is the greatest success achieved, its application can still be said to be only partially effective. There are three main reasons for this: One has to face the problem of creating a sufficiently accurate picture of the state of the atmosphere at the outset of the forecast process. Errors introduced at the beginning of the forecast will propagate and amplify at each forecast interval, gradually eroding its accuracy and usefulness. To move from a theoretical understanding of the weather to computer code that can generate a weather forecast, modelers will inevitably have to make some approximations. To be successful, the model must integrate an understanding of many different phenomena and their interactions.

The severe thunderstorms have significant socio-economic impact over eastern and northeastern parts of India and it's very important to improve understanding and prediction of these severe local storms. Realizing the importance of improved understanding and prediction of these weather events, an attempt is made to compare the performance of ANN and NWP model results for the prediction of three severe thunderstorm events (3, 11 and 15 May 2009) over Kolkata using surface

temperature and relative humidity and to see whether these models are reproducible. The hourly surface temperature and relative humidity from ANN, ARW and NMM models, which are useful for occurrence and intensity of the severe thunderstorms, are evaluated and validated the model results with observations. A quantitative verification of the results was performed with classical statistics parameters namely MAE, RMSE, CC and PC.

## **6.1 Data and Methodology**

### **6.1.1 Numerical model**

Both ARW and NMM models were integrated for a period of 24 hours starting from 0000 UTC of each day and ending at 0000 UTC of the following day. Boundary and initial conditions for both models are from FNL dataset of NCEP with  $1^{\circ} \times 1^{\circ}$  lat/lon grids. Both models thus have a common starting point, and avoid a potential source of difference. A single domain was configured with 3 km horizontal spatial resolution, which is reasonable in capturing the mesoscale cloud clusters. The domain covers  $84.5^{\circ}\text{E}$  to  $92.5^{\circ}\text{E}$  and  $19.5^{\circ}\text{N}$  to  $27.5^{\circ}\text{N}$  and the grids are centered at  $88.5^{\circ}\text{E}$ ,  $23.5^{\circ}\text{N}$ . Both NMM and ARW domains are configured with vertical structure of 38 unequally spaced sigma (non-dimensional pressure) levels. In this study, same physics options are taken for both the ARW and NMM simulations. The physical parameterizations used in this study are GFDL for longwave and shortwave radiation (Schwarzkopf and

Fels 1991; Lacis and Hansen 1974), Noah Land surface scheme (Chen and Dudhia 2001) for land surface, MYJ scheme (Janjic 2002) for planetary boundary layer, Ferrier scheme (Ferrier 2002) for microphysics, Janjic similarity scheme (Janjic 1994) for surface layer and GD cloud ensemble scheme (Grell and Devenyi 2002) for CPS. Output from each model is post-processed to bring them back to a common format that enables direct comparison. The details of thunderstorm cases are described in Chapter 3.

### 6.1.2 ANN model

This study evaluates the utility of MLPN for estimating hourly surface temperature and relative humidity with 24 h ahead. The hourly surface weather parameters namely mean sea level pressure (hPa), relative humidity (%), wind speed ( $\text{ms}^{-1}$ ) and temperature ( $^{\circ}\text{C}$ ) of 3 years (April and May 2007 to 2009) collected from IMD of Kolkata were used as the input and desired data for ANN model. The other additional input parameters for each model are month, day and hour of the observation. The total length of the data record included for ANN model is more than 4000. A three layer structure (one input layer, one hidden layer, one output layer) was selected with hyperbolic tangent (tanh) transfer function for hidden layer and linear transfer function for output layer. The chosen weather data were divided into two randomly selected groups, the training group, corresponding to 67% of the patterns, and the test group, corresponding to 33% of patterns; so that the generalization capacity of

network could be checked after training phase. Networks were trained for a fixed number of epochs. Networks were trained for a fixed number of epochs. Hourly surface temperature is used as desired data for the first model and hourly relative humidity for second model. The ANN model with LM algorithm was applied to derive thunderstorm forecast of 24 h ahead at Kolkata. The 24 h observed (AWS) data of surface temperature and relative humidity at Kolkata were used to validate these models during three thunderstorm days of May 2009 (3, 11 and 15 May 2009). The results are evaluated using MAE, RMSE, CC and PC. The ANN model simulations are carried out using the Neuro Solutions software developed by Neuro dimensions Inc. of Florida (Neuro Dimension 2005).

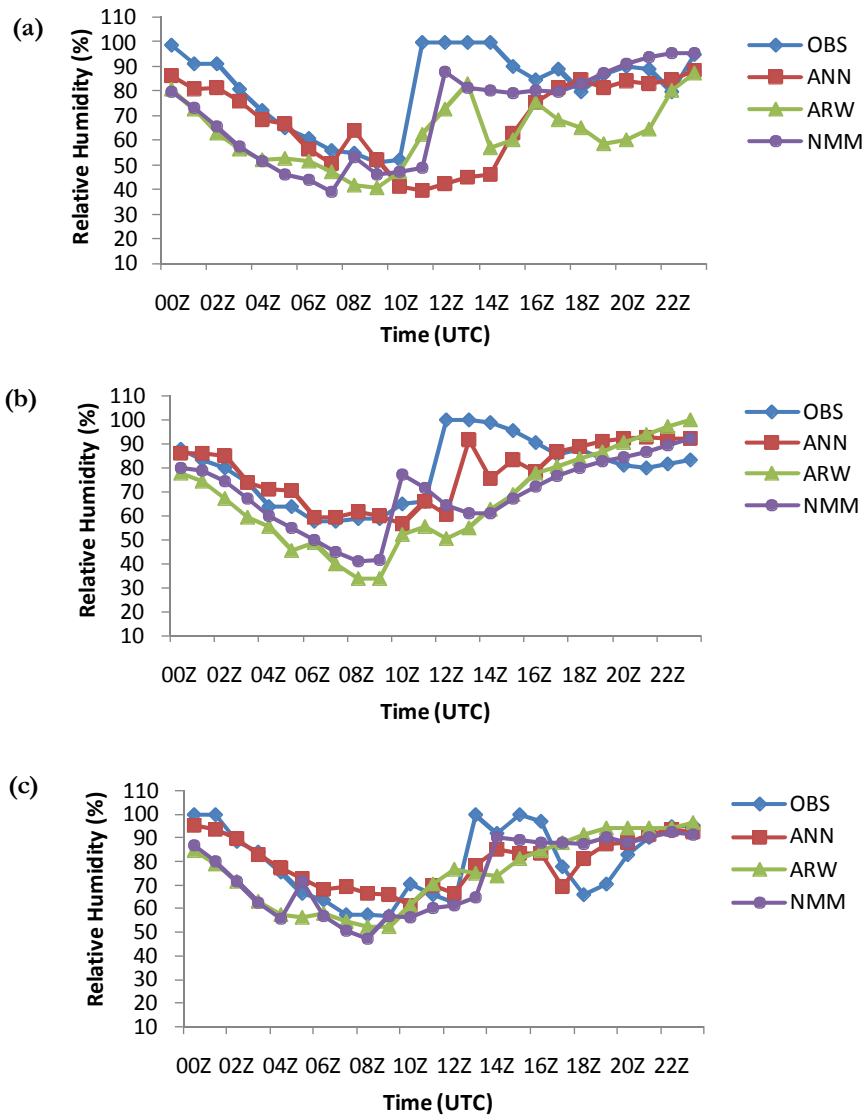
## **6.2 Results and Discussion**

### **6.2.1 Analysis of surface relative humidity**

Thunderstorm development is favored by large values of relative humidity if other conditions are favorable (Asnani 2006). Figure 6.1 shows the inter-comparison of observed and model simulated relative humidity (%) using ANN, NMM and ARW model over Kolkata ( $22.52^{\circ}\text{N}$ ,  $88.37^{\circ}\text{E}$ ) valid for 3, 11 and 15 May 2009 at 0000 UTC to next day at 0000 UTC. The observed relative humidity values for 3 May 2009 (Figure 6.1a) peaked from 52% to 100% (48%) at 1000 UTC whereas NMM model showed a sharp rise from around 49% to 88% (39%) at 1200 UTC, which is two hour later than that of the observed. ARW and ANN

models are failed to capture the sharp rise of relative humidity during the thunderstorm hour as in the NMM model. In the second case (Figure 6.1b), observed relative humidity showed a rise from 66% to 100% (34%) at 1200 UTC, whereas NMM model shows a rise from 42% to 77% (35%) at 1000 UTC, which is two hour prior than that of the observed. ARW model simulation shows an increase from 34% to 52% (18%) at 1000 UTC. ANN model shows a sudden rise from 65% to 92% (27%) at 1300 UTC. A sudden increase of 35% has been captured by NMM model as in the observed rise of 34%. ANN also captured sudden rise of 27% with one hour time lag for this thunderstorm event. ARW model is able to capture the rise with less intensity. In the third case (Figure 6.1c), observed relative humidity peaked from 63% to 100% (37%) at 1300 UTC, whereas NMM model shows a sharp rise from 65% to 91% (26%) at 1400 UTC, which is one hour later than that of the observed. ANN model results show a small increase from 67% to 79% (12%) at 1300 UTC. The intensity of increase is very less as compared to observation and NMM model. In this thunderstorm case also ARW model was not able to capture the sharp rise of relative humidity during the thunderstorm hour. For all the thunderstorm cases (Figure 6.1), NMM model has captured the sudden rise of relative humidity values during the model simulated thunderstorm hour as in the observations. ANN model results are better than ARW model for all three thunderstorm cases.





**Figure 6.1:** Inter-comparison of NMM, ARW and ANN models simulated and observed diurnal variation of surface relative humidity (%) over Kolkata on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.

A statistical analysis based on MAE, RMSE and CC was performed for comparisons between the simulated and observed surface relative humidity over Kolkata for 3 thunderstorm cases and are given in Table 6.1. From the table, it can clearly see that, simulated relative humidity of NMM model have more correlation with observation than ARW and ANN model for first and third thunderstorm cases. The ANN model performed well for second thunderstorm case than NMM and ARW model and it is greater than 0.7 for third case. The CC of NMM model is more than 0.7 for first and third thunderstorm cases and has an average CC value close to 0.7. The CC of ARW model is good for first thunderstorm case and is very less for second and third cases. The average CC of ARW is a less value and equal to ANN model. The MAE and RMSE of each model have been computed to test the prediction error of the model. The MAE and RMSE values of relative humidity are less for NMM model than ARW and ANN model for first thunderstorm case. The error is less for ANN model than ARW and NMM model for second and third thunderstorm cases. The average MAE and RMSE values are high for ARW model than ANN model and the values are 16.21% and 19.04%. The least errors are for NMM model and the values are 12.45% and 16.05%.

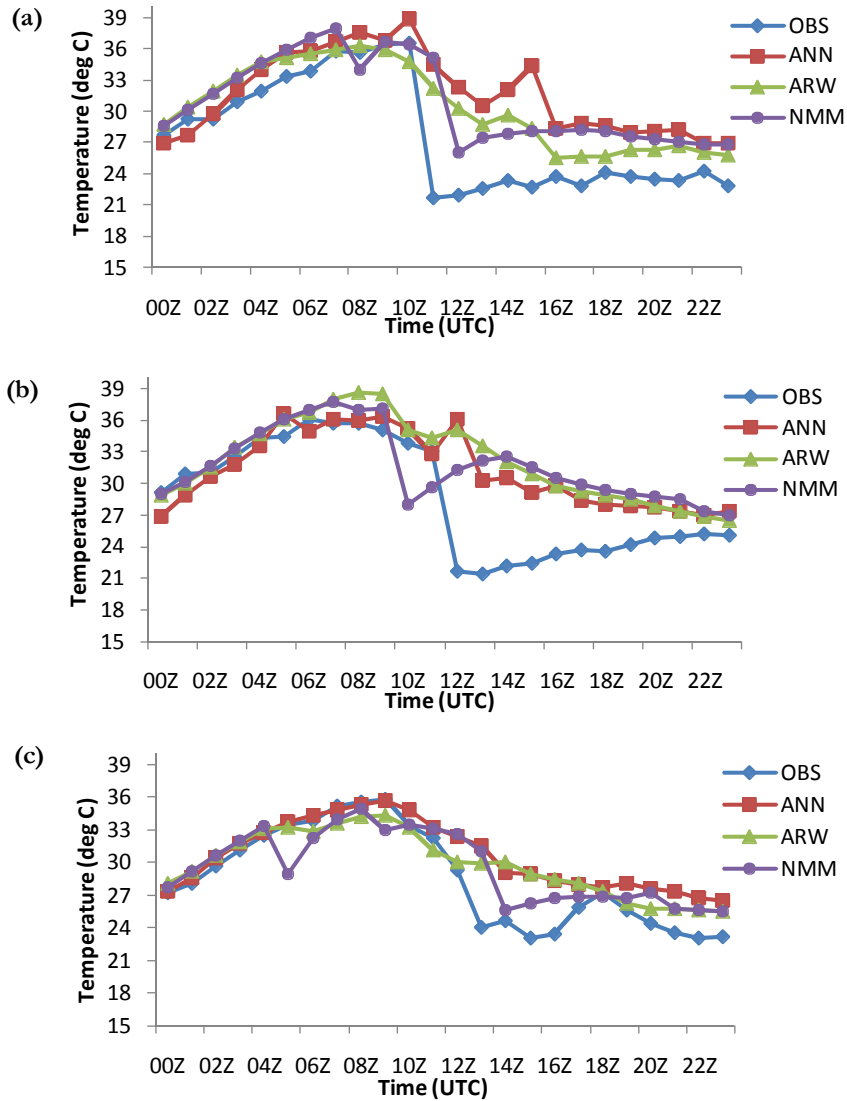
### **6.2.2 Analysis of surface temperature**

Surface temperature is useful parameter in forecasting the likelihood occurrence of a thunderstorm. Figure 6.2 shows the inter-comparison of

observed and model simulated temperature ( $^{\circ}\text{C}$ ) using NMM, ARW and ANN models over Kolkata valid for 3, 11 and 15 May 2009 at 0000 UTC to next day at 0000 UTC. The observed temperature (Figure6.2a) showed a sudden fall from  $36.7^{\circ}\text{C}$  to  $21.7^{\circ}\text{C}$  ( $15^{\circ}\text{C}$ ) at 1000 UTC whereas NMM model showed a fall from  $35.1^{\circ}\text{C}$  to  $26.1^{\circ}\text{C}$  ( $9^{\circ}\text{C}$ ) at model predicted thunderstorm hour (1200 UTC). ANN and ARW model show a small fall at 1600 UTC with an intensity of  $6^{\circ}\text{C}$  and  $4^{\circ}\text{C}$  respectively. For the second case (Figure6.2b), observed temperature showed a drop from  $33.1^{\circ}\text{C}$  to  $21.7^{\circ}\text{C}$  ( $11.4^{\circ}\text{C}$ ) at 1200 UTC, whereas NMM simulation shows a drop from  $37.1^{\circ}\text{C}$  to  $28^{\circ}\text{C}$  ( $9.1^{\circ}\text{C}$ ) at 1000 UTC. ARW model couldn't capture sudden fall for

**Table 6.1:** Statistical analysis of simulated and observed surface relative humidity over Kolkata based on MAE, RMSE and CC.

Statistical Analysis	CASES	ARW	NMM	ANN
MAE	3-May	18.97	13.36	20.06
	11-May	17.18	13.11	11.65
	15-May	12.49	10.89	7.31
	<b>MEAN</b>	<b>16.21</b>	<b>12.45</b>	<b>13.00</b>
RMSE	3-May	21.62	17.37	27.65
	11-May	20.85	17.00	14.05
	15-May	14.65	13.80	8.71
	<b>MEAN</b>	<b>19.04</b>	<b>16.05</b>	<b>16.80</b>
CC	3-May	0.72	0.76	0.45
	11-May	0.49	0.53	0.61
	15-May	0.57	0.72	0.72
	<b>MEAN</b>	<b>0.59</b>	<b>0.67</b>	<b>0.59</b>



**Figure 6.2:** Inter-comparison of NMM, ARW and ANN models simulated and observed diurnal variation of surface temperature ( $^{\circ}\text{C}$ ) over Kolkata on (a) 3 May 2009 (b) 11 May 2009 (c) 15 May 2009.

this case. ANN results show a small fall with an intensity of  $6^{\circ}\text{C}$  at 1300 UTC. In the third case, the observed temperature (Figure 6.3c) showed a sudden fall from  $29^{\circ}\text{C}$  to  $24^{\circ}\text{C}$  ( $5^{\circ}\text{C}$ ) at 1300 UTC whereas NMM model shows a fall from  $31^{\circ}\text{C}$  to  $26^{\circ}\text{C}$  ( $5^{\circ}\text{C}$ ) at model predicted hour of 1400 UTC. ANN model shows a drop from  $32^{\circ}\text{C}$  to  $29^{\circ}\text{C}$  ( $3^{\circ}\text{C}$ ) at 1400 UTC. ARW model failed to capture the sudden temperature fall over Kolkata for second and third case.

A statistical analysis based on MAE, RMSE and CC was performed for comparisons between the simulated and observed surface temperature over Kolkata for 3 thunderstorm cases and are given in Table 6.2. The CC of NMM and ARW models are more than 0.8 for first and third thunderstorm cases. The average CC value of NMM model is more than 0.8 and 0.77 is for ARW model due to less correlation in second case. The CC of ANN model is higher than other two models for second and third case. The average CC of ANN model is less than other two models due to less correlation in the first case. The MAE and RMSE of each model have been computed to test the prediction error of the model. The MAE and RMSE values of temperature are less for NMM model than ARW and ANN models for the first and third thunderstorm cases. The least average errors are for NMM model and the values are  $2.91^{\circ}\text{C}$  and  $3.86^{\circ}\text{C}$ . The MAE and RMSE values are less for ANN model than NMM and ARW models for the second thunderstorm case. The average ANN

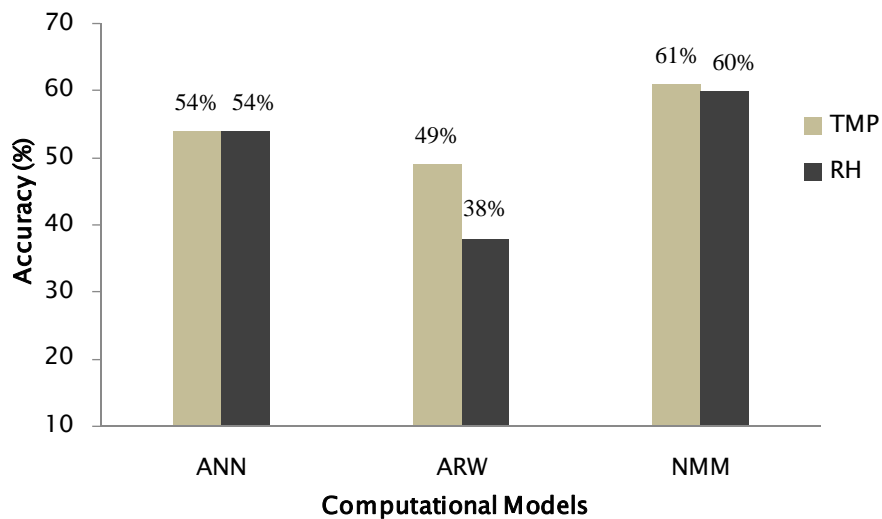
model errors of surface temperature are more than ARW and NMM models.

**Table 6.2:** Statistical analysis of simulated and observed surface temperature over Kolkata based on MAE, RMSE and CC.

Temperature	CASES	ARW	NMM	ANN
MAE	3-May	3.45	3.03	4.61
	11-May	3.99	3.78	3.21
	15-May	1.93	1.92	2.11
	<b>MEAN</b>	<b>3.12</b>	<b>2.91</b>	<b>3.31</b>
RMSE	3-May	4.28	3.94	6.35
	11-May	5.28	5.19	4.71
	15-May	2.58	2.47	2.93
	<b>MEAN</b>	<b>4.05</b>	<b>3.86</b>	<b>4.66</b>
CC	3-May	0.86	0.86	0.53
	11-May	0.62	0.69	0.70
	15-May	0.85	0.87	0.91
	<b>MEAN</b>	<b>0.77</b>	<b>0.81</b>	<b>0.71</b>

Figure 6.3 gives the performance accuracy of computational models for hourly temperature and relative humidity prediction. The accuracy of temperature presented a percentage number of the times when the forecast is accurate to within  $\pm 2^{\circ}\text{C}$ . The result clearly indicated that overall accuracy of NMM model for three events is 61% for the case of temperature prediction. ANN gave a moderate accuracy of 54%. ARW model displayed less accuracy. The accuracy of relative humidity presented a percentage number of the times when the forecast is accurate to with  $\pm 10\%$  confidence range. The result clearly indicated that overall

accuracy of NMM model for three events is 60% for relative humidity prediction. ANN gave a moderate accuracy of 54%. ARW displayed a less accuracy of 38% as in temperature case.



**Figure 6.3:** Performance accuracy of NMM, ARW and ANN models for the prediction of temperature (TMP) and relative humidity (RH) during 3 thunderstorm days.

Comparison of the time-series plots of surface temperature and relative humidity simulated by ARW, NMM and ANN models indicate the superiority of NMM model in simulating thunderstorm over Kolkata on these severe thunderstorm cases even though one or two hour lead or lag exists. The statistical analysis of surface weather parameters revealed that NMM model is well predicted the occurrence and intensity of all 3 thunderstorm cases as in the observation.

### 6.3 Chapter Summary

In this chapter, an attempt is made to compare the performance of ANN and NWP (ARW and NMM) models for the prediction of three severe thunderstorm events (3, 11 and 15 May 2009) over Kolkata using surface temperature and relative humidity and to check the capability of these models for the prediction of severe thunderstorm events over east Indian region. The hourly surface temperature and relative humidity from ANN, ARW and NMM models, which are useful tool for occurrence and intensity of the severe thunderstorms, are evaluated and validated the model results with observations. Performance and reliabilities of the models were then evaluated by a number of statistical measures namely MAE, RMSE, CC and PC.

Comparison of model simulated thunderstorm affected surface parameters with that of the observed revealed that NMM has performed better than ANN and ARW models in capturing the sharp rise in humidity and drop in temperature even though one or two hour lag or lead exists. ANN model reproduced this feature for second and third case with less intensity of sharp changes. ARW model has failed to capture the sharp rise and drop in humidity and temperature respectively. The statistical analysis of surface temperature and relative humidity indicate the superiority of NMM model in simulating the thunderstorms over Kolkata on these severe thunderstorm cases. The results suggest that NMM model holds promise for prediction



of surface weather parameters with reasonable accuracy in severe thunderstorm cases over east Indian region.

.....❧❧❧.....



---

## **SUMMARY AND CONCLUSIONS**

---

The thunderstorms are typical mesoscale systems dominated by intense convection. During April and May, the eastern parts of the country get affected by higher frequency of severe thunderstorms, locally named as ‘Nor’westers’ as it travels from northwest to southeast direction. Severe thunderstorms associated with squalls, lightning, hail, heavy rain cause extensive damage to standing agriculture crops, high rise buildings as well thatched huts, high tension electric poles and wires, and cause flash floods, resulting in loss of life and property. An accurate location specific and timely prediction is required to avoid loss of lives and property due to strong winds and heavy precipitation associated with these storms. Forecasting thunderstorms is one of the most difficult tasks, due to their rather small spatial and temporal extension and the inherent non-linearity of their dynamics and physics. The improvement in prediction of these important weather phenomena is highly handicapped due to lack of mesoscale observations and insufficient understanding. Realizing the importance of improved understanding and prediction of

this weather event, an attempt is made to study these severe thunderstorm events during the pre-monsoon season using empirical and dynamical approaches. The most widely used empirical approach for weather prediction is ANN. The recent advances in neural network methodology for modeling nonlinear, dynamical phenomena along with the impressive successes in a wide range of applications, motivated to investigate the application of ANNs for the prediction of thunderstorms. The second approach is based upon equations and forward simulations of the atmosphere (NWP). The most commonly known NWP models namely NMM and ARW models are used for this study.

The capabilities of six different learning algorithms using MLPN in predicting thunderstorm affected parameters over Kolkata (22.52<sup>0</sup>N, 88.37<sup>0</sup>E) were studied and their performances were compared. The ANN model was found to be efficient in fast computation and capable of handling the noisy and unstable data that are typical in the case of weather data. The results indicated that MLPN model with LM algorithm well predicted thunderstorm affected surface parameters. The developed ANN model with LM algorithm was applied to derive thunderstorm forecast from 1 to 24 h ahead at Kolkata. The studies of advanced prediction of these parameters showed that 1 and 3 h MLPN models were able to predict hourly temperature and relative humidity adequately with sudden fall and rise. The results of these analyses demonstrated the capability of

ANN model in prediction of severe thunderstorm events over eastern Indian region and will be helpful for real time thunderstorm forecast.

The sensitivity experiments have been conducted with NMM model to examine the impact of different initial conditions, CPSs and microphysics schemes in capturing the severe thunderstorm events occurred over Kolkata during May 2006, 2007 and 2009. Three sets of initial conditions (19 May 2006 at 0000 UTC, 19 May 2006 at 1200 UTC and 20 May 2006 at 0000 UTC) were experimented using NMM model for a thunderstorm event on 20 May 2006. The trends shown by various meteorological fields of the third experiment (20 May 2006 at 0000 UTC as initial condition) were in good agreement with each other and very much consistent with dynamic and thermodynamic properties of the atmosphere for the occurrence of a severe thunderstorm. Another sensitivity experiments are conducted with NMM model by changing the CPSs such as KF, BMJ, GD, AS schemes and explicit scheme for two severe thunderstorm cases (20 May 2006 and 21 May 2007) at Kolkata and validated the model results with observation. In all experiments, the setups were identical except for the use of different convective schemes. Hence differences in the simulation results may be attributed to the sensitivity of the convective schemes. This study shows that the prediction of thunderstorm affected parameters is sensitive to convective schemes. The GD scheme has well predicted the thunderstorm activities, in terms of time, intensity and the region of occurrence of the events, as

compared to other convective schemes and also explicit scheme. One more sensitivity experiment has been conducted for a severe thunderstorm event on 15 May 2009 with three microphysics schemes namely FERR, WSM6 and THOM to examine the sensitivity of the simulations to different cloud microphysics. The results show that the NMM model with FERR scheme appears to reproduce the cloud and precipitation processes more realistically than other schemes for the prediction of severe thunderstorm event. The studies found suitable options as 24 h simulation for initial conditions, GD scheme for CPS and FERR for microphysics scheme. These options are used for further studies of thunderstorms with NMM model.

A comparative study with two numerical models namely NMM and ARW were done for the prediction of severe thunderstorm events during May 2009. Both models were able to broadly reproduce several features of the thunderstorm events, such as spatial pattern and temporal variability over east region of India. Comparison of model simulated thunderstorm affected parameters with that of the observed showed that NMM has performed better than ARW in capturing the sharp rise in humidity and drop in temperature. NMM model has predicted well the genesis, intensification and propagation of the squall line, which is in good agreement with that of the observed, while the squall line movement was slow in ARW. The statistical analysis of surface parameters indicates the superiority of NMM model in simulating the thunderstorm over

Kolkata on these severe thunderstorm cases. This suggests that NMM model has the potential to provide unique and valuable information for severe thunderstorm forecasters over east Indian region.

The performance evaluation of computational models namely ANN, ARW and NMM models were done to predict severe thunderstorm events using thunderstorm affected parameters like surface temperature and relative humidity over Kolkata. Thunderstorm prediction is inherently complex process, so it is impossible to wait 100% accurate forecast results since we cannot measure all factors that may be local scales. From the results we can see that, the ANN models are not able to predict the sharp jumps and dips of surface parameters during the thunderstorm hour as it is a complex physical process. NMM model is able to predict these fluctuations adequately with reasonable accuracy. It is concluded from the results that, NMM model is good for thunderstorm prediction temporally and spatially since it is a short-lived mesoscale phenomenon which cannot be much easily predicted from historical data.

### **7.1 Future Directions**

There are strong requirements in India for improved forecasting of rapid and severe thunderstorm over east and northeast Indian region. Even if a reasonable mesoscale analysis can be obtained, imbalances between the dynamic and thermodynamic fields can lead to model spinup which could degrade, e.g. short-term precipitation forecasts, a key requirement

for mesoscale systems. Thus improved mesoscale analyses will rely on the development of assimilation methods which will include alternative methods to replace current analysis approaches in which only data at or near the synoptic times are used. More comprehensive work will be carried out to improve the mesoscale model performance with modified initial condition through sophisticated 3DVAR/4DVAR assimilation system using ample amount of special high quality observations.

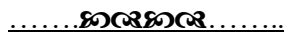
It is clear that high impact weather systems such as thunderstorms, heavy precipitation or tropical cyclones have a high level of uncertainty associated with them. It is therefore more appropriate to provide forecasts of these features in terms of the development of mesoscale ensemble prediction systems (EPS). This will require further research on development of appropriate means of generating initial perturbations and accounting for model uncertainties. Nowcasting and very short range forecasting are concerned with the weather monitoring and forecasting of weather for the shortest time scales, ranging from 0 to 6 hours. These methods are expected to shift from the traditional specialized techniques to an utilization of more general NWP output products. This will be made possible through access to future more powerful computers, through the application of mesoscale NWP models at grid resolutions of the order of 1 km or less and through application of mesoscale assimilation techniques.

The Thunderstorm prediction systems are more likely decision support system than expert systems because they need guidance and



predictions must be evaluated by human interference. The future directions for improving these severe weather events over Indian region can be summarized as follows:

- Comprehensive mesoscale data assimilation utilizing conventional and non-conventional observations from multi-observation platforms.
- Prediction of life cycle of thunderstorm along with associated hazards using very high resolution (1-0.3 km) state-of-the-art mesoscale models.
- Role of physical processes in particular deep convection, cloud microphysics, planetary boundary layer, land surface processes with high resolution mesoscale model and special observational datasets.
- Understanding of cloud microphysics, aerosol concentration and atmospheric electricity in association with severe thunderstorms.
- Systematic error evaluation and bias correction of NWP models in simulation of severe thunderstorms.
- Improve prediction of intensity and time of occurrence of severe thunderstorms by using dynamical-statistical approaches.





## REFERENCES

---

- [1] Air Weather Service Technical Report 79/006, 1990: The use of the skew T, Log P diagram in analysis and forecasting. Air Weather Service, Scott AFB, Illinois.
- [2] Ajith Abraham, 2006: Artificial Neural Networks. Oklahoma State University, Stillwater OK, USA.
- [3] Alvi, S. M. A., and Punjabi, K. G., 1966: Diurnal and seasonal variations of squalls in India. *Indian J. Meteor. Geophys.*, 17, 207-216.
- [4] Alvi, S. M. A., and Punjabi, K. G., 1966: Diurnal and seasonal variations of squalls in India. *Indian J. Meteor. Geophys.*, 7, 207.
- [5] AMS, 2000: Glossary of Meteorology. 2nd edition. Allen Press. New York.
- [6] Arakawa, A., and Lamb, V., 1977: Computational design of the basic dynamical processes in the UCLA general circulation model. *Methods in Computational Physics*, Academic Press, 17, 173–265.
- [7] Arakawa, A., and Schubert, W. H., 1974: Interaction of a cumulus cloud ensemble with the large scale environment. Part I. *J. Atmos. Sci.*, 31, 674-701.

- [8] Asnani, G. C., 2006: Tropical Meteorology (Revised Edition). Vol. II, Sindh Colony: Audh, Pune.
- [9] Awadesh, K., 1992: A climatological study of thunderstorms at Lucknow airport. *Mausam*, 43, 441–444.
- [10] Bailey, D. L., and Thompson, D. M., 1990: Developing Neural Network Applications. *AI Expert*, 5(9), 34-41.
- [11] Bhan, S. C., 2007: Tornado over Ludhiana–Eyewitness account and meteorological aspects. National Symposium on Advances in Meteorology and their Applications, Bhopal, India, 17-19 December.
- [12] Bhatnagar, A. K., Rajesh, R. P., and Kalyanasundaram, S., et al., 2003: Doppler radar- A detecting tool and measuring instrument in meteorology. *Current Science*, 85(3), 256-264.
- [13] Bishop, C. M., 1995: Neural Networks for Pattern Recognition. Oxford University Press, Oxford, UK, 364-369.
- [14] Bodri, L., and V. Cermak, 2000: Prediction of extreme precipitation using a neural network: Application to summer flood occurrence in Moravia. *Adv. Eng. Software*, 31, 311–321.
- [15] Brooks, H. E., and Wilhelmson, R. B., 1992: Numerical simulation of a low-precipitation super-cell thunderstorm. *Meteor. Atmos. Physics*, 49, 3-17.

- [16] Brooks, H. E., Lee, J. W., and Craven, J. P., 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, 67, 73–94.
- [17] Browning, K. A., 1962: Cellular Structure of Convective Storms. *The Meteorological Magazine*, 91(1085), 341-350.
- [18] Browning, K. A., 1964: Airflow and precipitation trajectories within severe local storms which travel to the right of the winds. *J. Atmos. Sci.*, 21, 634-639.
- [19] Byers, H. R., and R. R. Braham, 1949: The Thunderstorm project. U. S. Weather Bureau, U. S. Department of Commerce Tech. Rep., 287 pp.
- [20] Callan, R., 1999: The essence of neural networks, Prentice Hall Europe.
- [21] Carling, A., 1992: *Introducing Neural Networks*. Wilmslow, UK: Sigma Press.
- [22] Chakraborty, B., and Chakraborty, G., 2002: A new feature extraction technique for on-line recognition of handwritten alphanumeric characters. *Inf. Sci.*, 148, 55–70.
- [23] Chatterjee, P., Pradhan, D., and De, U. K., 2008: Simulation of hailstorm event using mesoscale model MM5 with modified cloud microphysics scheme. *Annales Geophysicae*, 26, 3545-3555.

- [24] Chaudhuri, S., 2008: Preferred Type of Cloud in the Genesis of Severe Thunderstorms – A Soft Computing Approach. *Atmospheric Research*, 88, 149-156.
- [25] Chaudhury, A., and Banerjee, A. K., 1983: A study of hailstorms over northeast India. *Vayu Mandal*, 13, 91-95.
- [26] Chaudhury, S., and Chattopadhyay, S., 2005: Neuro-Computing based short range prediction of some meteorological parameters during the pre-monsoon season. *Soft Computing*, 349-354.
- [27] Chen, F., and Dudhia, J., 2001: Coupling an advanced land-surface/hydrology model with Penn State/NCAR MM5 modeling system. Part I: Model description and implementation. *Mon. Wea. Rev.*, 129, 569-585.
- [28] Chen, Z., and S. Haykin, 2002: On different facets of regularization theory. *Neural Computation*, 14, 2790-2845.
- [29] Cotton, W. R., George, R. L., and Knupp, K. R., 1982: An Intense, Quasi-Steady Thunderstorm over Mountainous Terrain. Part I: Evolution of the Storm-Initiating Mesoscale Circulation. *Journal of the Atmospheric Sciences*, 39, 2, 328–342.
- [30] Dallas, W. L., 1900: *Indian Meteorological Memoirs*, 6 Pt V.

- [31] Das Someshwar, 1999: Suggested observational network for simulation of cloud processes during the Indian Ocean Experiment (INDOEX). *Current Science*, 76(7), 912-915.
- [32] Dasgupta, S., and De, U. K., 2007: Binary logistic regression models for short term prediction of pre-monsoon convective developments over Kolkata (India). *International Journal of Climatology*, 27(6), 831-836.
- [33] Desai, B. N., 1950: Mechanism of Nor'westers of Bengal. *Indian J. Meteor. Geophys.*, 1, 74-76.
- [34] Doswell, C. A., 1987: The distinction between large-scale and mesoscale contribution to severe convection: A case study example. *Wea. Forecasting*, 2, 3-16.
- [35] Doswell, C. A. III, and L. F. Bosart, 2001: Extra-tropical synoptic-scale processes and severe convection. *Severe Convective Storms, Meteor. Monogr., No. 50, Amer. Meteor. Soc.*, 27-70.
- [36] Doswell, C. A., 2004: Forecasters's forum: Weather forecasting by humans-heuristics and decision making. *Weather and Forecasting*, 19, 1115 – 1126.
- [37] Ek, M. B., Mitchell, K. E., and Lin, Y., et al., 2003: Implementation of NOAH land surface model advances in the NCEP operational mesoscale Eta model. *J. Geophys. Res.*, 108(22), 8851.

- [38] Eliassen, A., 1949: The quasi-static equations of motion with pressure as independent variable. *Geophys. Publ. (Oslo)*, 17, 44.
- [39] Eliot, J., 1899: Hailstorm in India during the period 1883-1897 with a discussion on their distribution. I. *Met. Memoirs*, 6, 4.
- [40] Esugasini, E., Mashor, M. Y., Isa, N., and Othman, N., 2005: Performance Comparison for MLP Networks Using Various Back Propagation Algorithms for Breast Cancer Diagnosis. *Lecture Notes in Computer Science*, 3682/2005, 123-130.
- [41] Fahlman, S. E., 1988: An empirical study of learning speed in backpropagation networks. *Computer Science Technical Report*, CMU-CS-88-162, Carnegie Mellon University, Pittsburgh.
- [42] Farely, R. D., Wang, S., and Orville, H. D., 1992: A comparison of 3D model results with observations for an isolated CCOPE thunderstorm. *Meteor. Atmos. Physics*, 49, 187-207.
- [43] Fausett, L., 1994: *Fundamental of Neural Networks*, New York, Prentice Hall.
- [44] Ferrier, B. S., Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta model. *Preprints*, 15th Conference on Numerical Weather Prediction, San Antonio, TX, Amer. Meteor. Soc., 280-283.



- [45] Frisby, E. M., and H. W. Sansom, 1967: Hail indices in tropics. *J. Appl. Meteor.*, 6, 339-354.
- [46] Fujita, T. T., 1981: Tornadoes and downbursts in the context of generalized planetary scales. *J. Atmos. Sci.*, 38, 1512-1534.
- [47] Gallus, W. A., and J. F. Bresch, 2006: Comparison of Impacts of WRF Dynamic Core, Physics Package, and Initial Conditions on Warm Season Rainfall Forecasts. *Mon. Wea. Rev.*, 125, 2917–2942.
- [48] Ghosh, A., Lohar, D., and Das, J., 2008: Initiation of Nor'wester in relation to mid-upper and low-level water vapor patterns on METEOSAT-5 images. *Atmospheric Research*, 87, 116-135.
- [49] Gokhale, N. R., 1975: Hailstorms and Hailstone growth. Albany State University of New York Press.
- [50] Goldar, R. N., S. K. Banerjee, and G. C. Debnath, 2001: Tornado in India and Neighborhood and its Predictability Regional Met Centre, Alipore, Kolkata Issued by Office of the Additional Director General of Meteorology (Research), Meteorological Office of India Meteorological Department, Pune, India.
- [51] Goliger, A. M., and R. V. Milford, 1998: A review of worldwide occurrence of tornadoes. *Journal of Wind Engineering and Industrial Aerodynamics*, 74, 111-121.

- [52] Grabowski, W. W., Wu, Xiaoqing, and Moncrieff, M. W., 1996: Cloud-Resolving Modeling of Tropical Cloud Systems during Phase III of GATE. Part I: Two-Dimensional Experiments. *J. Atmos. Sci.*, 53, 24, 3684–3709.
- [53] Grell, G. A., 1993: Prognostic Evaluation of Assumptions Used by Cumulus Parameterizations. *Mon. Wea. Rev.*, 121, 764-787.
- [54] Grell, G. A., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, 29(14), Article 1693.
- [55] Guhathakurta, P., 2006: Long-range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network. *Current Science*, 90, 773-779.
- [56] Gupta, H. N., and Chorghade, S. L., 1961: A statistical study of thunderstorm activity over Agartala airfield, *Indian J. Met. Geophysics*, 12 (1), 109-114.
- [57] Hacıismailoglu, M. C., Kucuk, I., and Derebasi, N., 2009: Prediction of dynamic hysteresis loops of nano-crystalline cores. *Expert Syst. Appl.*, 36, 2225–2227.
- [58] Hagan, M. T., and M. B. Menhaj, 1994: Training feed-forward networks with the Marquardt algorithm. *IEEE Trans. Neural Networks*, 5 (6), 989–993.

- [59] Hagan, M. T., H. B. Demuth, and M. H. Beale, 1996: *Neural Network Design*. PWS Publishing Company, Boston, Massachusetts.
- [60] Hayashi, H. Y., Shi, S. H., and Esteban, J. A., et al., 2000: Driving AMPA receptors into synapses by LTP and CaMKII: requirement for GluR1 and PDZ domain interaction. *Science* 287, 2262–2267.
- [61] Hart, T., 2002: Numerical Prediction Model Performance Summary October to December 2001. *Australian Meteorological Magazine*, 51, 137-141.
- [62] Haykin, S., 1994: *Neural Networks: A Comprehensive Foundation*, Macmillan Publishing.
- [63] Hebb, D. O., 1949: *The Organization of Behavior*. New York: John Wiley & Sons. Introduction and Chapter 4 reprinted in Anderson & Rosenfeld, 1988, pp 45-56
- [64] Hill, T., and Lewicki, P., 2007: *Statistics Methods and Applications*. Tulsa, OK: StatSoft.
- [65] Hobbs, P. V., 1991: Research on Clouds and Precipitation: Past, Present and Future, Part II. *Bull. Amer. Meteor. Soc.*, 72, 2, 184–191.
- [66] Holton, J. R., 2004: *An introduction to dynamic meteorology*. 4th Ed., Elsevier Academic Press, Burlington, Mass.

- [67] Hopfield, J. J., 1982: Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79, 2554-2558.
- [68] IMD, 1943: *Climatological Atlas for Airmen*.
- [69] IMD, 1953: *Climatological Tables of Observatories in India*.
- [70] IMD, 1969: *Climatological Tables of Observatories in India*.
- [71] IMD, 1995: *Climatological Tables of Observatories in India*.
- [72] Jacobs, R. A., 1988: Increased rates of convergence through learning rate adaptation. *Neural Networks*, 1, 295–307.
- [73] Jacobson, M. Z., 2005: *Fundamentals of atmospheric modeling*. 2nd Ed., Cambridge University Press, New York.
- [74] Jacovides, C.P., and Yonetani, T., 1990: An evaluation of stability indices for thunderstorm prediction in greater Cyprus. *Wea. Forecasting*, 5, 559-569.
- [75] Jalali-Heravi, M., Asadollahi-Baboli, M., and Shahbazikhah, P, 2008: QSAR study of heparanase inhibitors activity using artificial networks and Levenberge-Marquardt algorithm. *Eur. J. Med. Chem.*, 43, 548–556.
- [76] James, C., Davis, R., and Meiyer, M., et al., 2000: Aligned micro-contact printing of micrometer-scale poly-l-lysine structures for controlled growth of cultured neurons on planar microelectrode arrays. *IEEE Trans. Biomed. Eng.*, 47, 17–21.

- 
- [77] Janjic, Z. I., 1994: The step–mountain eta coordinate model: further developments of the convection, viscous sub-layer and turbulence closure schemes. *Mon. Wea. Rev.*, 122, 927–945.
- [78] Janjic, Z. I., 1996: The Surface Layer in the NCEP Eta Model. 11th Conf. on NWP, Norfolk, VA, American Meteorological Society, 354–355.
- [79] Janjic, Z. I., 2000: Comments on Development and Evaluation of a Convection Scheme for Use in Climate Models. *J. Atmos. Sci.* 57, 3686.
- [80] Janjic, Z. I., 2002: Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso model, NCEP Office Note, No. 437, 61.
- [81] Janjic, Z. I., 2003: A non-hydrostatic model based on a new approach. *Meteorol. Atmos. Phys.*, 82, 271-285.
- [82] Johns, R. H., and C. A. Doswell, 1992: Severe local storms forecasting. *Wea. Forecasting*, 7(4), 588-612.
- [83] Joseph, P. V., Raipal, D. K., and Deka, S. N., 1980: Andhi, the convective dust-storm of northwest India. *Mausam*, 31, 431- 442.
- [84] Kaila, V. K., Kiran Kumar, A. S., and Sundarmurthy, T. K., et al., 2002: METSAT – a unique mission for weather and climate. *Current Science*, 83, 1081–1088.

- [85] Kain, J. S., 2004: The Kain–Fritsch Convective Parameterization: An Update. *Journal of Applied Meteorology*, 43 (1), 170–181.
- [86] Kain, J. S., and Fritsch, J. M., 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme. The representation of cumulus convection in numerical models, K.A. Emanuel and D.J. Raymond, Eds., Amer. Meteor. Soc., 246.
- [87] Kalnay, E., 2003: Atmospheric modeling, data assimilation, and predictability. 1st Ed., Cambridge University Press, Cambridge.
- [88] Kalnay, E., Kanamitsu, M., and Kistler, R., et al., 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77, 437-471.
- [89] Kandalgaonkar, S. S., M. I. R. Tinmaker, and A. Nath, et al., 2005: Study of thunderstorm and rainfall activity over the Indian region. *Atmosfera*, 18(2), 91-101.
- [90] Kessler, E., 1983: *Thunderstorm Morphology and Dynamics*. University of Oklahoma Press, 79 pp.
- [91] Khan, M. R., and Ondrusek, C., 2000: Short-term electric demand prognosis using artificial neural networks. *Electr. Eng.*, 51, 296-300.

- 
- [92] Kisi, O., and Uncuoglu, E., 2005: Comparison of Three Back-Propagation Training Algorithms for Two Cases Studies. *Indian Journal of Engineering and Materials Sciences*, 12, 434-442.
- [93] Klemp, J. B., and Lilly, D. K., 1978: Numerical simulation of hydrostatic mountain waves. *J. Atmos. Sci.*, 35, 78-107.
- [94] Klemp, J. B., and Wilhelmson, R. B., 1978: The Simulation of Three- Dimensional Convective Storm Dynamics. *J. Atmos. Sci.*, 35, 6, 1070–1096.
- [95] Koch, S. E., B. Ferrier, and M. Stolina, et al., 2005: The use of simulated radar reflectivity fields in the diagnosis of mesoscale phenomena from high-resolution WRF model forecasts. Preprints, 11<sup>th</sup> Conference on Mesoscale Processes, Albuquerque, NM, Amer. Meteor. Soc., CD-ROM, J4J.7.
- [96] Krishnamurthy, V. 1965: A statistical study of thunderstorms over Poona. *Indian J. Met. Geophys.*, 16, 484.
- [97] Lacis, A. A., and Hansen, J. E., 1974: A parameterization for the absorption of solar radiation in the earth's atmosphere. *J. Atmos. Sci.*, 31, 118–133.
- [98] Landberg, L., 1995: Implementing Wind Forecasting at a Utility. Proceedings of the Wind Energy Conversion - 17th British Wind Energy Association Conference, Warwick, U. K.

- [99] Latha, R., and Murthy, B.S., 2011: Boundary layer signatures of consecutive thunderstorms as observed by Doppler sodar over western India. *Atmospheric Research*, 99, 230-240.
- [100] LeCun, Y., 1986: Learning Processes in an Asymmetric Theshold Network. *Disordered Systems and Biological Organization* (Bienenstock, E., Fogelman-Smith, F., Weisbuch, G. (eds)), NATO ASI Series, F20, Berlin: Springer-Verlag.
- [101] Lin, C., and Arakawa, A., 1997: The Macroscopic Entrainment Processes of Simulated Cumulus Ensemble. Part II: Testing the Entraining-Plume Model. *J. Atmos. Sci.*, 54(8), 1044–1053.
- [102] Litta, A. J., Mohanty, U. C., and Sumam Mary Idicula, 2009: Numerical simulation of a Tornado over Orissa (India) on 31 March 2009 using WRF mesoscale model. *International Conference on Modelling and Simulation*, Trivandrum, Kerala, India, 1-3 December.
- [103] Lopez, R., 2008: An Open Source Neural Networks C++ Library. *Universitat Polit`ec- nica de Catalunya*, Barcelona.
- [104] Lopez, L., Garcia-Ortega, E., and Sanchez, J. L., 2007: A short-term forecast model for hail. *Atmospheric Research*, 83(2-4), 176-184.



- 
- [105] Lorenz, E. N., 1969: Three approaches to atmospheric predictability. *Bulletin of the American Meteorological Society*, 50, 345–349.
- [106] Luk, K. C., J. E. Ball, and A. Sharma, 2000: A study of optimal model lag and spatial inputs for artificial neural network for rainfall forecasting. *J. Hydrol.*, 227, 56–65.
- [107] Lynn, B. H., Tao, W. K., and Wetzel, Peter J., 1998: A Study of Landscape-Generated Deep Moist Convection. *Monthly Weather Review*, 126, 4, 928–942.
- [108] Manohar, G. K., S. S. Kandalgaonkar, and M. I. R. Tinmaker, 1999: Thunderstorm activity over India and the Indian southwest monsoon. *J. Geophys. Res.*, 104, D4, 4169–4188.
- [109] Maqsood, I., Khan, M. R., and Abraham, A., 2000: Intelligent weather monitoring systems using connectionist models. *International Journal of Neural, Parallel and Scientific Computations*, 10, 157–178.
- [110] Maqsood, I., M. R. Khan, and A. Abraham, 2004: An ensemble of neural networks for weather forecasting. *Neural Comput. Appl.*, 13, 112–122.
- [111] Markowski, P., and Y. Richardson, 2010: Organization of Isolated Convection. *Mesoscale Meteorology in Midlatitudes*, John Wiley & Sons Ltd., 201-244.

- [112] McCulloch, W. S., and Pitts, W., 1943: A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.
- [113] McNulty, R. P., 1995: Severe and Convective Weather: A Central Region Forecasting Challenge. *Weather and Forecasting*, 10, 187-202.
- [114] Messinger, F., and Arakawa, A., 1976: Numerical methods used in atmospheric models. *GARP Publications Series*, No. 17, 64, pp.115.
- [115] Minsky, M., and Papert, S., 1969: *Perceptron: An Introduction to Computational Geometry*, MIT Press, Cambridge, MA.
- [116] Mitchell, T. M., 1997: *Machine Learning*, Chapter 4, McGraw-Hill.
- [117] Mohanty, U. C., M. Mandal, A. K. Das, and A. P. Dimri, 2003: Mesoscale Modeling of Convective Systems over India: Status and Scope. *Weather and Climate Modeling*, Ed. S.V. Singh, Swati Basu and T.N. Krishnamurti, 63-76.
- [118] Mohanty, U. C., D. R. Sikka, and O. P. Madan, et al., 2006: Weather Summary Pilot Experiment of Severe Thunderstorms-Observational and Regional Modeling (STORM) Programme - 2006, Department of Science and Technology, Government of India, August.

- [119] Mohanty, U. C., D. R. Sikka, and O. P. Madan, et al., 2007: Weather Summary Pilot Experiment of Severe Thunderstorms-Observational and Regional Modeling (STORM) Programme – 2007, Department of Science and Technology, Government of India, November.
- [120] Mohanty, U. C., D. R., Sikka, and Kiran Prasad, S., et al., 2009: Weather Summary, analysis and preliminary evaluation of mesoscale model during Pilot Experiment of Severe Thunderstorms: Observational and Regional Modeling (STORM) Programme – 2009. Department of Science & Technology, Government of India, October.
- [121] Mohanty, U. C., Mandal, M., and Raman, S., 2004: Simulation of Orissa Super Cyclone (1999) using PSU/NCAR Mesoscale model. *Natural Hazards*, 31, 373–390.
- [122] Mohapatra, M., Bandyopadhyay, B. K., Sharma, A. K., and Suman, G., 2010: Utility of Automatic Weather Station (AWS) data for monitoring and prediction of monsoon circulations during 2009, *IMD Met. Monogr. Synop. Met. No. 6*. Published by India Meteorological Department, pp. 109-121.
- [123] Moid, S. A., 1995: A climatological study of thunderstorms at Mohanbari airport. *Mausam*, 46, 202–204.

- [124] Mukherjee, A. K., and Sen, P. N., 1983: Dependence of diurnal variation of thunderstorm on physical features. *Vayu Mandal (Special Issue)*, 13, 105-108.
- [125] Mukhopadhyay, P., Sanjay, J., and Singh, S. S., 2003: Objective forecast of thundery/non thundery days using conventional indices over three northeast Indian stations. *Mausam*, 54(4), 867–880.
- [126] Mukhopadhyay, P., Mahakur, M., and Singh, H. A. K., 2009: The interaction of large scale and mesoscale environment leading to formation of intense thunderstorms over Kolkata Part I: Doppler radar and satellite observations. *Journal of Earth System Science*, 118(5), 441-466.
- [127] Negnevitsky, M., 2002: *Artificial Intelligence. A Guide to Intelligent Systems*. Addison Wesley, Harlow, England; London; New York.
- [128] NeuroDimension, Inc., 2005: Developers of NeuroSolutionsv6.01: Neural Network Simulator. Gainesville, FL. <http://www.nd.com>, 2005.
- [129] NeuroSolutions Manual, 2003: The neural network simulation environment. NeuroDimension Inc., FL.
- [130] Nielsen, F., 2001: *Neural Networks algorithms and applications*. Neil's Brock Business College.

- [131] Nizamuddin, S., 1993: Hail occurrences in India. *Weather*, 48, 90-92.
- [132] Ogura, Y., and Takahashi, T., 1971: Numerical Simulation of the Life Cycle of a Thunderstorm Cell. *Monthly Weather Review*, 99(12), 895-911.
- [133] Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.*, 56, 527-530.
- [134] Orville, H. D., and Kopp, F. J., 1977: Numerical Simulation of the Life History of a Hailstorm. *J. Atmos. Sci.*, 34(10), 1596-1618.
- [135] Otkin, J. A., and Greenwald, T. J., 2008: Comparison of WRF model-simulated and MODIS-derived cloud data. *Mon. Wea. Rev.*, 136, 1957-1970.
- [136] Pant, G. B., and Rupa Kumar, K., 1997: *Climates of south Asia*. John Wiley and Sons, Chichester (UK), 320pp.
- [137] Parker, D., 1985: *Learning Logic*. Technical Report TR-87, Center for Computational Research in Economics and Management Science, MIT, Cambridge, MA.
- [138] Patterson, D. W., 1996: *Artificial Neural Networks: Theory and Applications*. Englewood Cliffs, Prentice Hall.

- [139] Peterson, R. E., and K. C. Mehta, 1981: Climatology of tornadoes of India and Bangladesh, *Arch. Meteor. Geophys. Bioklimat.*, 29B, 345-356.
- [140] Phillips, N. A., 1957: A coordinate system having some special advantages for numerical forecasting. *Journal of Meteorological Society*, 14, 184-185.
- [141] Pielke, R. A., 2002: *Mesoscale meteorological modeling*. 2nd Ed., Academic Press, San Diego.
- [142] Pitts, W., and McCulloch, W. S., 1947: How we Know Universals. *Bull. Math.*, 127-147.
- [143] Price, C., 2006: Global thunderstorm activity, Sprites. In: M. Fullekrug et al. (eds.) *Sprites. Elves and Intense Lightning Discharges*, 85-99.
- [144] Radhika, Y., and Shashi, M., 2009: Atmospheric temperature prediction using support vector machines. *International Journal of Computer Theory and Engineering*, 1, 1793-8201.
- [145] Rajeevan, M., Kesarkar, A., and Thampi, S. B., et al., 2010: Sensitivity of WRF cloud microphysics to simulations of severe thunderstorm event over Southeast India. *Ann. Geophys.*, 28, 603-619.
- [146] Rajeevan, M., Madhulatha, A., and Rajasekhar, M., et al., 2012: Development of a perfect prognosis probabilistic

- model for prediction of lightning over south-east India. *Journal of Earth System Science*, 121 (2), 355-371.
- [147] Raju, M. M., R. K. Srivastava, and Dinesh, et al., 2011: Development of Artificial Neural- Network-Based Models for the Simulation of Spring Discharge. *Advances in Artificial Intelligence*, Article ID 686258.
- [148] Ramamurthy, B. V., 1983: Some cloud physical aspects of local severe storms. *Vayu Mandal*, 13, 3-11.
- [149] Raman, P. K., and Raghavan, K., 1961: Diurnal variations of thunderstorms in India during different seasons. *Indian J. Met. Geophys.*, 12, 115.
- [150] Rao, K. N., and Raman, P. K., 1961: Frequency of days of thunder in India. *Indian J. Met. Geophys.*, 16, 103.
- [151] Rao, K. N., Daniel, C. E. J., and Balasubramanian, L. V., 1971: Thunderstorms over India. IMD Pre-published Scientific Report No.153.
- [152] Rao, Y. P., 1981: The climate of the Indian Subcontinent. *World survey of climatology*, 9, 67-192.
- [153] Rasmussen, R. M., and Wilhelmson, R. B., 1983: Relationships between storm characteristics and 1200 GMT hodographs, low-level shear and stability. Preprints of 13<sup>th</sup> Conf. On Severe Local Storms, Tulsa, American Meteorological Society, 55-58.

- [154] Riordan, D., and Hansen, B. K., 2002: A fuzzy case based system for weather prediction. *Engineering Intelligent Systems*, 3, 139-145.
- [155] Rosenblatt, F., 1962: *Principles of Neurodynamics*, Spartan Press, Washington, DC.
- [156] Rumelhart, D. E., G. E. Hinton, and R. J. Williams, 1986: Learning representations by back-propagating errors. *Nature*, 323, 533–536.
- [157] S. Pattanayak, U. C. Mohanty, and S. R. Rizvi, et al., 2008: A Comparative study on performance of MM5 and WRF (ARW & NMM) models in simulation of tropical cyclone over bay of Bengal, 9th WRF Users' Workshop, Boulder, CO, USA, June 23-27.
- [158] Saha, K. R., 1967: *Tornadoes and their occurrence in India*. PRF-Pub Sci. Rep. No 30, IMD, 14pp.
- [159] Santosh, K., Sarasakumari, R., Gangadharan, V. K., and Sashidharan, N. V., 2001: Some climatological features of thunderstorm at Thiruvananthapuram, Kochi, and Kozhikode. *Mausam*, 52 (2), 357-364.
- [160] Satoh, M., 2004: *Atmospheric Circulation Dynamics and Circulation Models*. 1st Ed., Springer.



- [161] Schlesinger, R. E., 1978: A Three-Dimensional Numerical Model of an Isolated Thunderstorm: Part I. Comparative Experiments for Variable Ambient Wind Shear. *J. Atmos. Sci.*, 35, 4, 690–713.
- [162] Schultz, P., 1989: Relationships of several stability indices to convective weather events in northeast Colorado. *Weather and Forecasting*, 4, 73–80.
- [163] Schwarzkopf, M. D., and Fels, S. B., 1991: The simplified exchange method revisited: An accurate, rapid method for computations of infrared cooling rates and fluxes. *J. Geophys. Res.*, 96, 9075-9096.
- [164] Science Plan, 2005: Severe Thunderstorms – Observations & Regional Modeling (STROM) Programme; Department of Science and Technology, Government of India, New Delhi, December.
- [165] Sen Roy, S., Lakshmanan, V., and Roy Bhowmik, S. K., et al., 2010: Doppler weather radar based nowcasting of cyclone Ogn. *J. Earth Syst. Sci.*, 119, 183–199.
- [166] Sentiono, R., W. K. Leow, and J. M. Zurada, 2002: Extraction of Rules from Artificial Neural Networks for Nonlinear Regression. *IEEE Transactions on Neural Networks*, Vol. 13, No. 3.

- [167] Shavlik, J. W., Mooney, R. J., and Towell, G., 1991: Symbolic and neural learning algorithms - an experimental comparison. *Mach Learning*, 6,111-143.
- [168] Shepherd, A. J., 1997: *Second-Order Methods for Neural Networks: Fast and Reliable Training Methods for Multi-Layer Perceptrons*. Springer, Neural Computing series.
- [169] Singh, R., 1981: On the occurrence of tornadoes in India. *Mausam*, 32(3), 307-314.
- [170] Skamarock, W. C., J. B. Klemp., and J. Dudhia, et al., 2005: A Description of the Advanced Research WRF Version 2. NCAR Tech Note, NCAR/TN-468+STR, 88 pp.
- [171] Srivastava, K., Roy Bhowmik, S. K., Sen Roy, S., Thampi, S. B., and Reddy, Y. K., 2010: Simulation of high impact convective events over Indian region by ARPS model with assimilation of Doppler weather radar radial velocity and reflectivity. *Atmosfera*, 23, 53-73.
- [172] Steidley, C., A. Sadvoski, P. Tissot, and R. Bachnak, 2005: Using an artificial neural network to improve predictions of water level where tide charts fail. *Innovations in Applied Artificial Intelligence*, M. Ali and F. Esposito, Eds., Springer, 599-608.
- [173] Stensrud, D. J., Bao, J. W., and Warner, T. T., 2000: Using initial condition and model physics perturbations in short-

- range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, 128, 2077–2107.
- [174] Stull, R. B., 2000: *Meteorology for Scientists and Engineers*. 2nd Ed., Brooks/Cole, California.
- [175] Sutton, R. S., and Barto, A. G., 1998: *Reinforcement Learning: An Introduction*. Cambridge, MA, MIT Press.
- [176] Tao, W. K., and Simpson, J., 1984: Cloud Interactions and Merging: Numerical Simulations. *J. Atmos. Sci.*, 41(19), 2901-2917.
- [177] Tao, W. K., Simpson, J., and Soong, S. T., 1987: Statistical Properties of a Cloud Ensemble: A Numerical Study. *J. Atmos. Sci.*, 44(21), 3175–3187.
- [178] Tao, W. K., Simpson, J., and Sui, C. H., et al., 1999: Equilibrium States Simulated by Cloud-Resolving Models. *J. Atmos. Sci.*, 56(17), 3128–3139.
- [179] Tyagi, A., 2000: Mesoscale weather prediction. *Current Science*, 79(6), 698-710.
- [180] Tyagi, A., 2007: Thunderstorm climatology over Indian region. *Mausam*, 58(2), 189-212.
- [181] Tyagi, B., Naresh Krishna, V., and Satyanarayana, A. N. V., 2011: Study of thermodynamic indices in forecasting pre-

- monsoon thunderstorms over Kolkata during STORM pilot phase 2006–2008. *Nat. Hazards*, 56, 681-698.
- [182] Ubeyli, E. D., and Guler, I., 2004: Multilayer perceptron neural networks to compute quasi-static parameters of asymmetric coplanar waveguides. *Neuro computing*, 62, 349-365.
- [183] Vaidya, S. S., 2007: Simulation of weather systems over Indian region using mesoscale models. *Meteorology and Atmospheric Physics*, 95, 15-26.
- [184] Vishwanathan, T. R. and J. F. Faria, 1962: A climatological study of thunderstorms at Bombay airport, *Indian J. Met. Geophys.*, 13, 377.
- [185] Wallace, J. M., and Hobbs, P. V., 2006: *Atmospheric Science: An Introductory Survey*. 2<sup>nd</sup> Ed., Academic Press, Burlington, Mass.
- [186] Wandishin, M. S., Mullen, S. L., Stensrud, D. J., and Brooks, H. E., 2001: Evaluation of a short-range multi-model ensemble system. *Mon. Wea. Rev.*, 129, 729–747.
- [187] Wang Wei, and Seaman, N. L., 1997: A Comparison Study of Convective Parameterization Schemes in a Mesoscale Model. *Mon. Wea. Rev.*, 125, 252-278.
- [188] Washington, W. M., and Parkinson, C. L., 2005: *An introduction to three-dimensional climate modeling*. 2nd Ed., University Science Books, Sausalito, Calif.

- [189] Wedge, D., D. Ingram, and D. McLean, et al., 2005: A global-local artificial neural network with application to wave overtopping prediction. *Artificial Neural Networks: Formal Models and Their Applications–ICANN 2005*, Springer, 109–114.
- [190] Weiss, S. J., Bright, D. R., and Kain, J. S., et al., 2006: Complementary Use of Short-range Ensemble and 4.55 KM WRF-NMM Model Guidance for Severe Weather Forecasting at the Storm Prediction Centre. *Preprints, 23<sup>rd</sup> Conf. Severe Local Storms*, ST. Louis MO, Amer. Meteor. Soc.
- [191] Wiki, R., 2008: *Numerical Solution Technique*. September.
- [192] William, E.R., 2005: Lightning and climate: A review. *Atmos. Res.*, 76, 272-287.
- [193] Wilson, J. W., N. A. Crook, C. K. Mueller, J. Sun, and M. Dixon, 1998: Nowcasting Thunderstorms: A Status Report. *Bull. Amer. Meteor. Soc.*, 79, 2079–2099.
- [194] WMO, 1953: World distribution of thunderstorm days. WMO No 21, TP 6 and Supplement (1956), Geneva, Switzerland, 204 and 17 pp.
- [195] Zurada, J. M., 1992: *Introduction to Artificial Neural Systems*. West Publishing Company, Saint Paul, Minnesota.





## **LIST OF PUBLICATIONS**

---

### **International Journal**

- [1] Litta A. J., Mohanty U. C., and Sumam Mary Idicula, 2009: Simulation of Severe Squall Line over Kolkata using WRF-NMM Model. Lectures on Modelling and Simulation (Special Edition -AMSE), 10(1), 73-83.
- [2] Litta A. J., Sumam Mary Idicula, and Mohanty U. C., 2011: A Comparative Study of Convective Parameterization Schemes in WRF-NMM Model. International Journal of Computer Applications, 33(6), 32- 39.
- [3] Litta A. J., Sumam Mary Idicula, Mohanty U. C. and S. Kiran Prasad, 2012: Comparison of thunderstorm simulations from WRF-NMM and WRF-ARW over East Indian Region. The Scientific World Journal, doi:10.1100/2012/951870.
- [4] Litta A. J., Mohanty U. C., Someshwar Das and Sumam Mary Idicula, 2012: Numerical Simulation of Severe Local Storms over East India using WRF-NMM mesoscale model. Atmospheric Research, 116, 161-184.

- [5] Litta A. J., Sumam Mary Idicula, and C. Naveen Francis, 2012: Artificial Neural Network model for the prediction of thunderstorms over Kolkata. *International Journal of Computer Applications*, 50(11), 50-55.

### **National Journal**

---

- [6] Litta A. J., and Mohanty U. C., 2008: Simulation of a severe thunderstorm event during the field experiment of STORM programme 2006 using WRF-NMM model. *Current Science*, 2008, 95(2), 204-215.
- [7] Litta A. J., Mohanty U. C., and Sumam Mary Idicula, 2012: The diagnosis of severe thunderstorms with high-resolution WRF model. *Journal of Earth System Sciences*, 121, 297-316.

### **International Conference**

---

- [8] Litta A. J, Mohanty U. C., and Sumam Mary Idicula, 2009: Simulation of Severe Squall Line over Kolkata using WRF-NMM model. *International Conference on Modelling and Simulation*, Trivandrum, Kerala, India, 1-3 December (Best Paper Award).
- [9] Litta A. J, Mohanty U. C., and Sumam Mary Idicula, 2009: Numerical simulation of a Tornado over Orissa (India) on 31 March 2009 using WRF mesoscale model. *International*



Conference on Modelling and Simulation, Trivandrum, Kerala, India, 1-3 December.

### **National Conference**

---

- [10] Litta A. J., Sumam Mary Idicula, and Mohanty U. C., 2010: WRF Modeling for Effective Severe Thunderstorm Forecasting over East and North-East Indian region. National Symposium on Nowcasting to Forecasting - Needs and Challenges for Naval Operations, Cochin, Kerala, India, 1-2 November.
- [11] Litta A. J., Mohanty U. C., and Sumam Mary Idicula, 2012: Comparison of WRF Model Simulated and Satellite Derived Cloud data during Severe Thunderstorms. National Seminar on Challenges, Developments and Opportunities in Nowcasting, New Delhi, India, 30-31 January.

.....❦.....



## APPENDIX

Appendix I: Fragment of the data between 1 and 15 May 2009 used for testing ANN model is given in the following table. Datasets of severe thunderstorm days are marked as bold and italic.

Month	Day	Hour	Mean Sea Level Pressure (hPa)	Relative Humidity (%)	Wind Speed (ms <sup>-1</sup> )	Temperature (°C)
5	1	0	1005.4	93	2.06	28.5
5	1	1	1005.9	92	1.54	29.0
5	1	2	1006.7	89	4.12	30.4
5	1	3	1008.1	86	2.06	30.9
5	1	4	1007.9	81	2.57	32.5
5	1	5	1006.7	79	2.57	33.2
5	1	6	1005.4	76	2.57	34.6
5	1	7	1005.0	75	2.57	35.0
5	1	8	1004.6	73	4.63	35.8
5	1	9	1003.8	74	4.12	35.5
5	1	10	1003.0	75	3.09	35.0
5	1	11	1002.7	76	3.60	33.9
5	1	12	1002.9	80	3.09	32.2
5	1	13	1003.7	84	4.12	30.8
5	1	14	1004.6	86	3.60	30.3
5	1	15	1005.6	86	3.09	30.2
5	1	16	1006.2	86	3.60	30.1
5	1	17	1006.2	86	2.57	30.0
5	1	18	1005.5	85	2.57	25.0
5	1	19	1005.5	85	2.06	25.0
5	1	20	1005.9	83	1.54	26.4
5	1	21	1005.3	85	2.06	28.8

*Appendix*

5	1	22	1005.2	92	2.57	28.7
5	1	23	1005.9	93	1.54	28.4
5	2	0	1007.4	93	0.51	28.5
5	2	1	1007.7	91	3.09	29.3
5	2	2	1008.1	89	2.06	30.7
5	2	3	1008.2	85	2.06	31.4
5	2	4	1008.4	82	2.57	32.1
5	2	5	1008.2	76	2.57	33.8
5	2	6	1008.6	76	2.57	34.6
5	2	7	1007.4	75	3.09	34.5
5	2	8	1006.9	72	2.06	35.6
5	2	9	1005.6	75	3.09	35.6
5	2	10	1005.3	76	2.06	34.7
5	2	11	1004.8	76	2.06	34.2
5	2	12	1005.4	80	3.09	32.7
5	2	13	1006.0	85	2.57	31.4
5	2	14	1007.4	87	2.57	30.6
5	2	15	1007.8	89	2.57	30.0
5	2	16	1008.6	90	2.06	29.7
5	2	17	1008.6	92	1.54	29.5
5	2	18	1008.3	94	1.54	29.0
5	2	19	1008.0	96	2.06	28.4
5	2	20	1007.5	97	1.03	28.1
5	2	21	1007.3	98	1.03	28.0
5	2	22	1007.2	98	1.03	27.9
5	2	23	1007.4	99	0.51	27.7
<b>5</b>	<b>3</b>	<b>0</b>	<b>1007.8</b>	<b>99</b>	<b>1.03</b>	<b>27.7</b>
<b>5</b>	<b>3</b>	<b>1</b>	<b>1008.6</b>	<b>95</b>	<b>1.03</b>	<b>29.2</b>
<b>5</b>	<b>3</b>	<b>2</b>	<b>1009.7</b>	<b>88</b>	<b>1.54</b>	<b>30.9</b>
<b>5</b>	<b>3</b>	<b>3</b>	<b>1010.6</b>	<b>82</b>	<b>1.54</b>	<b>32.0</b>
<b>5</b>	<b>3</b>	<b>4</b>	<b>1010.7</b>	<b>77</b>	<b>1.54</b>	<b>33.3</b>

5	3	5	1009.8	75	2.57	33.9
5	3	6	1008.7	72	1.54	35.8
5	3	7	1007.9	71	2.06	35.6
5	3	8	1005.6	68	3.60	36.4
5	3	9	1005.0	52	2.06	36.7
5	3	10	1006.7	100	5.17	21.7
5	3	11	1006.3	100	1.54	21.9
5	3	12	1007.5	100	1.03	22.6
5	3	13	1008.6	100	0.51	23.4
5	3	14	1011.1	93	3.09	22.7
5	3	15	1011.6	89	4.12	23.8
5	3	16	1008.0	92	2.06	22.9
5	3	17	1007.1	85	2.57	24.1
5	3	18	1006.1	90	3.09	23.8
5	3	19	1005.6	93	1.54	23.5
5	3	20	1004.7	92	2.57	23.4
5	3	21	1004.5	85	0.00	24.2
5	3	22	1006.0	97	2.06	22.9
5	3	23	1006.8	99	1.03	22.6
5	4	0	1008.1	99	1.03	22.8
5	4	1	1009.1	95	0.51	23.9
5	4	2	1009.4	92	1.03	25.1
5	4	3	1009.6	87	0.51	27.0
5	4	4	1009.5	81	1.03	28.9
5	4	5	1009.2	80	1.03	30.8
5	4	6	1008.1	79	1.03	31.7
5	4	7	1007.2	76	1.54	33.0
5	4	8	1006.2	74	1.54	34.1
5	4	9	1005.0	73	1.03	34.4
5	4	10	1004.3	77	1.54	33.2
5	4	11	1003.8	79	2.06	32.8

*Appendix*

---

5	4	12	1003.7	82	1.03	31.5
5	4	13	1004.0	83	1.54	30.6
5	4	14	1004.2	85	2.06	29.9
5	4	15	1005.0	89	2.57	29.6
5	4	16	1005.3	89	3.09	28.8
5	4	17	1006.3	90	2.06	28.4
5	4	18	1005.7	94	1.54	28.0
5	4	19	1005.4	87	2.06	27.8
5	4	20	1005.1	96	1.54	28.0
5	4	21	1004.7	94	2.06	28.3
5	4	22	1004.6	97	1.03	27.7
5	4	23	1005.3	99	1.03	27.1
5	5	0	1006.6	100	0.51	26.7
5	5	1	1007.5	99	2.06	27.9
5	5	2	1007.9	93	2.57	29.8
5	5	3	1008.2	91	2.06	30.6
5	5	4	1009.0	88	2.57	31.6
5	5	5	1009.1	83	2.06	33.2
5	5	6	1008.3	78	2.57	34.0
5	5	7	1007.4	76	2.57	35.3
5	5	8	1006.4	73	2.57	35.9
5	5	9	1005.6	71	2.06	36.4
5	5	10	1004.9	73	2.57	36.3
5	5	11	1004.4	72	2.06	35.7
5	5	12	1004.5	75	2.06	34.4
5	5	13	1005.2	79	1.54	32.9
5	5	14	1006.3	83	2.06	31.7
5	5	15	1006.5	85	2.06	30.7
5	5	16	1006.8	90	2.57	29.7
5	5	17	1006.2	94	2.57	29.0
5	5	18	1005.8	95	3.60	29.0

5	5	19	1005.2	78	3.09	28.9
5	5	20	1004.6	95	1.03	28.6
5	5	21	1004.2	98	2.06	28.1
5	5	22	1004.3	99	2.06	27.8
5	5	23	1005.0	99	1.54	27.7
5	6	0	1005.7	100	2.06	27.3
5	6	1	1006.6	98	1.03	28.3
5	6	2	1007.7	94	2.06	29.6
5	6	3	1008.1	90	3.09	31.0
5	6	4	1007.9	83	2.57	30.5
5	6	5	1007.9	80	2.06	34.3
5	6	6	1007.0	75	3.60	35.3
5	6	7	1005.9	75	2.57	36.3
5	6	8	1005.0	70	3.09	37.2
5	6	9	1004.1	66	2.57	37.5
5	6	10	1003.2	73	3.09	36.7
5	6	12	1004.4	73	2.57	33.2
5	6	13	1005.2	81	2.06	31.6
5	6	14	1004.9	80	0.00	29.5
5	6	15	1006.2	81	0.51	29.4
5	6	16	1006.6	81	1.54	28.0
5	6	17	1005.9	83	1.54	28.1
5	6	18	1005.1	91	1.03	28.4
5	6	19	1003.4	92	3.60	29.0
5	6	20	1002.8	94	1.54	28.7
5	6	21	1002.7	96	1.54	28.2
5	6	22	1003.0	96	2.06	28.0
5	6	23	1003.3	97	1.54	27.9
5	7	0	1004.1	98	1.54	27.9
5	7	1	1004.9	96	2.57	28.9
5	7	2	1006.0	89	2.57	30.1

*Appendix*

---

5	7	3	1006.8	85	1.54	31.8
5	7	4	1006.7	80	1.03	32.8
5	7	5	1006.3	76	0.51	34.4
5	7	6	1005.6	64	0.51	35.9
5	7	7	1004.8	57	0.00	36.9
5	7	8	1003.8	56	1.03	37.9
5	7	9	1002.7	50	1.03	38.3
5	7	10	1002.1	47	0.00	38.3
5	7	11	1001.8	50	1.03	37.4
5	7	12	1001.9	55	1.54	36.0
5	7	13	1002.3	78	2.06	34.3
5	7	14	1003.3	87	2.06	32.2
5	7	15	1004.0	90	2.57	30.8
5	7	16	1004.5	91	3.60	30.1
5	7	17	1004.4	92	2.57	29.9
5	7	18	1003.8	93	1.54	29.8
5	7	19	1003.5	95	2.06	29.5
5	7	20	1003.3	94	1.03	29.1
5	7	21	1003.2	95	1.54	28.8
5	7	22	1003.2	98	1.03	28.4
5	7	23	1003.5	99	1.54	28.4
5	8	0	1003.9	98	1.54	28.5
5	8	1	1004.9	90	2.06	29.7
5	8	2	1005.6	80	2.57	31.5
5	8	3	1006.1	74	2.57	33.4
5	8	4	1005.7	61	3.09	35.9
5	8	5	1005.2	53	3.09	37.7
5	8	6	1004.6	48	1.03	39.6
5	8	7	1003.5	45	2.06	39.8
5	8	8	1002.6	42	2.06	40.5
5	8	9	1001.8	42	2.06	40.6



5	8	10	1001.1	41	2.57	40.7
5	8	11	1000.7	40	1.54	40.0
5	8	12	1000.8	45	0.51	36.8
5	8	13	1001.4	70	3.09	35.0
5	8	14	1002.1	82	3.09	32.9
5	8	15	1002.9	85	2.57	31.5
5	8	16	1003.5	90	2.06	30.8
5	8	17	1003.4	93	2.06	30.1
5	8	18	1002.9	95	2.06	29.7
5	8	19	1002.6	94	3.09	29.7
5	8	20	1002.1	92	2.06	29.4
5	8	21	1001.9	91	2.57	29.1
5	8	22	1001.8	90	2.06	29.0
5	8	23	1002.0	90	1.03	29.0
5	9	0	1002.3	92	1.03	28.9
5	9	1	1002.3	91	2.06	28.9
5	9	2	1003.9	82	2.06	32.0
5	9	3	1003.9	78	2.06	34.0
5	9	4	1004.0	71	2.06	35.6
5	9	5	1003.4	64	2.06	36.7
5	9	6	1002.8	56	1.54	38.4
5	9	7	1001.5	49	1.54	40.1
5	9	8	1000.3	47	0.51	41.0
5	9	9	999.5	47	1.54	41.3
5	9	10	998.6	46	1.03	41.7
5	9	11	998.3	49	2.06	39.8
5	9	12	998.7	67	3.60	37.8
5	9	13	999.6	78	2.57	34.5
5	9	14	1000.4	82	2.57	32.6
5	9	15	1001.0	88	3.60	31.5
5	9	16	1001.7	90	3.09	30.8

*Appendix*

---

5	9	17	1001.4	91	2.57	30.2
5	9	18	1000.9	92	2.57	29.9
5	9	19	1000.5	91	1.54	29.8
5	9	20	1000.3	93	2.06	29.5
5	9	21	1000.2	96	2.06	29.2
5	9	22	1000.2	98	1.54	28.9
5	9	23	1000.7	98	1.54	28.7
5	10	0	1001.4	98	2.06	28.7
5	10	1	1002.3	96	2.06	29.4
5	10	2	1002.8	91	3.60	30.9
5	10	3	1003.1	82	3.09	33.1
5	10	4	1003.1	80	3.09	34.5
5	10	5	1002.9	77	2.06	36.4
5	10	6	1002.5	75	2.06	36.8
5	10	7	1001.6	69	3.09	37.6
5	10	8	1000.9	63	2.06	38.6
5	10	9	999.9	65	2.06	39.0
5	10	10	999.3	67	3.09	38.3
5	10	11	999.2	76	3.60	35.7
5	10	12	999.2	80	4.12	33.9
5	10	13	999.9	85	3.09	32.4
5	10	14	1000.6	88	3.60	31.1
5	10	15	1001.2	91	4.12	30.4
5	10	16	1001.7	92	2.06	30.1
5	10	17	1001.6	92	2.57	30.0
5	10	18	1001.0	93	2.57	29.9
5	10	19	1000.7	94	3.09	29.8
5	10	20	1001.2	93	1.54	29.8
5	10	21	1001.4	93	1.54	29.7
5	10	22	1001.2	95	1.03	29.2
5	10	23	1001.9	96	2.06	29.0

5	11	0	1002.2	95	2.06	29.2
5	11	1	1003.3	93	2.57	29.8
5	11	2	1004.2	90	3.09	31.0
5	11	3	1004.8	87	2.57	31.1
5	11	4	1004.8	84	3.60	32.6
5	11	5	1004.4	77	5.14	34.3
5	11	6	1005.0	77	3.09	34.5
5	11	7	1004.2	73	0.00	36.1
5	11	8	1003.0	73	3.60	35.8
5	11	9	1002.3	74	3.60	35.7
5	11	10	1000.7	74	4.63	35.1
5	11	11	1000.3	78	3.60	33.9
5	11	12	1002.1	66	2.57	33.1
5	11	13	1007.9	100	5.14	21.7
5	11	14	1003.8	100	4.63	21.4
5	11	15	1004.6	99	1.03	22.2
5	11	16	1005.2	97	1.54	22.5
5	11	17	1005.8	93	0.51	23.3
5	11	18	1005.1	89	1.54	23.7
5	11	19	1004.9	91	1.54	23.6
5	11	20	1004.1	88	1.03	24.2
5	11	21	1004.3	86	1.54	24.9
5	11	22	1004.2	85	1.54	25.0
5	11	23	1004.7	87	1.54	25.2
5	12	0	1004.2	93	1.54	24.6
5	12	1	1005.6	92	1.54	25.4
5	12	2	1006.4	92	1.03	25.7
5	12	3	1006.5	86	1.03	28.0
5	12	4	1006.9	86	1.54	28.7
5	12	5	1008.6	86	3.09	27.8
5	12	6	1007.5	83	2.06	26.8

*Appendix*

---

5	12	7	1007.5	94	2.06	25.1
5	12	8	1005.8	95	1.54	24.8
5	12	9	1004.9	93	1.54	25.4
5	12	10	1003.7	92	1.54	25.5
5	12	11	1001.4	66	2.57	34.9
5	12	12	1001.4	72	0.00	33.7
5	12	13	1001.9	74	1.03	32.5
5	12	14	1003.0	74	1.54	31.6
5	12	15	1004.8	79	2.06	31.0
5	12	16	1005.6	79	2.06	30.6
5	12	18	1005.3	90	2.06	29.5
5	12	19	1004.7	66	1.03	29.1
5	12	20	1004.0	94	1.03	28.6
5	12	21	1003.5	97	0.51	27.7
5	12	22	1003.2	99	0.51	27.1
5	12	23	1003.2	100	0.51	26.8
5	13	4	1005.1	79	2.06	31.0
5	13	5	1004.5	79	2.06	32.3
5	13	6	1003.2	78	3.60	33.3
5	13	7	1002.6	74	3.09	33.8
5	13	8	1001.5	74	3.09	33.7
5	13	9	1000.9	77	3.09	33.3
5	13	10	999.9	78	3.09	33.5
5	13	11	999.2	82	2.57	32.9
5	13	12	999.6	85	2.57	31.6
5	13	13	999.9	86	3.09	30.3
5	13	14	1001.1	91	3.09	29.8
5	13	15	1001.7	93	2.06	29.4
5	13	16	1001.6	94	1.54	29.2
5	13	17	1002.1	95	1.03	29.0
5	13	18	1002.6	98	1.03	28.4

5	13	19	1001.8	99	1.03	28.3
5	13	20	1000.9	100	0.51	28.1
5	13	21	1000.9	99	1.54	27.2
5	13	22	1000.4	100	1.03	24.5
5	13	23	1001.3	100	2.06	24.3
5	14	0	1002.0	100	0.51	23.4
5	14	1	1002.3	99	1.03	24.8
5	14	2	1002.7	94	0.51	27.6
5	14	3	1003.4	91	0.51	28.1
5	14	4	1004.0	87	2.06	30.9
5	14	5	1003.9	74	2.06	31.0
5	14	6	1003.1	72	2.06	31.5
5	14	7	1002.5	75	2.06	31.7
5	14	8	1002.0	77	2.06	32.7
5	14	9	1000.7	77	1.54	33.1
5	14	10	1000.3	76	2.06	33.8
5	14	11	999.6	80	2.57	32.9
5	14	12	1001.1	85	3.60	30.3
5	14	13	1000.7	95	0.51	26.1
5	14	14	1001.2	99	1.03	27.0
5	14	15	1002.2	98	2.06	27.7
5	14	16	1002.9	96	1.54	27.9
5	14	17	1003.1	97	1.03	27.4
5	14	18	1002.7	99	1.03	27.3
5	14	19	1002.9	100	1.03	27.0
5	14	20	1002.8	99	1.03	27.0
5	14	21	1002.7	100	1.03	26.9
5	14	22	1002.8	99	0.51	27.0
5	14	23	1002.9	100	1.03	27.1
<b>5</b>	<b>15</b>	<b>0</b>	<b>1004.2</b>	<b>100</b>	<b>0.51</b>	<b>27.2</b>
<b>5</b>	<b>15</b>	<b>1</b>	<b>1005.0</b>	<b>100</b>	<b>0.51</b>	<b>28.1</b>

---

5	15	2	1005.7	93	1.03	29.7
5	15	3	1005.5	90	1.54	31.1
5	15	4	1005.5	85	1.54	32.5
5	15	5	1005.4	79	1.54	33.5
5	15	6	1004.6	77	1.54	33.9
5	15	7	1003.4	73	2.06	35.2
5	15	8	1002.5	73	1.54	35.6
5	15	9	1001.5	72	1.54	35.8
5	15	10	1001.0	82	1.54	33.4
5	15	11	1001.9	78	2.06	32.3
5	15	12	1001.8	63	2.06	29.3
5	15	13	1003.0	100	2.06	24.1
5	15	14	1005.1	94	2.06	24.7
5	15	15	1007.3	100	2.06	23.1
5	15	16	1006.7	98	2.57	23.5
5	15	17	1001.5	73	6.00	25.9
5	15	18	1002.4	71	3.09	27.3
5	15	19	1004.3	78	1.54	25.7
5	15	20	1004.2	87	1.54	24.4
5	15	21	1003.8	93	1.03	23.6
5	15	22	1003.8	97	1.03	23.1
5	15	23	1004.7	97	1.03	23.2

**Appendix II: Details of the NeuroSolutions software are given in the following section.**

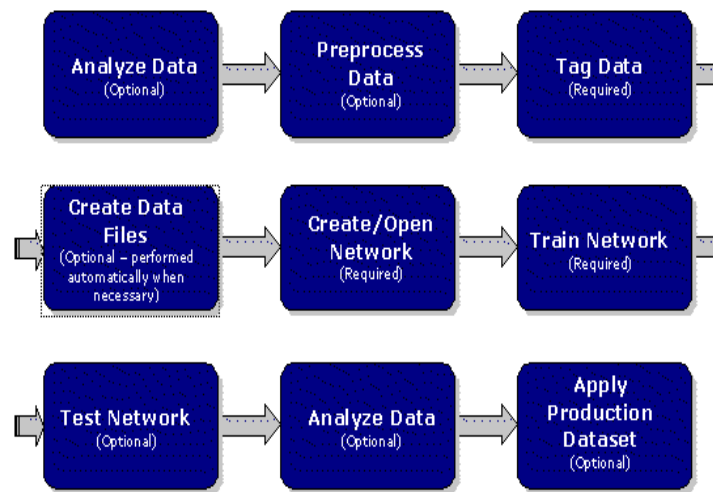
NeuroSolutions is a neural network development environment developed by NeuroDimension Inc., FL. It combines a modular, icon-based (component-based) network design interface with an implementation of advanced learning procedures, such as conjugate gradients, Levenberg -

Marquardt and back propagation through time. The software is used to design, train and deploy neural network (supervised learning and unsupervised learning) models to perform a wide variety of tasks such as data mining, classification, function approximation, multivariate regression and time-series prediction. NeuroSolutions is based on the concept that neural networks can be broken down into a fundamental set of neural components. Individually these components are relatively simplistic, but several components connected together can result in networks capable of solving very complex problems. The network construction wizards will connect these components based on the user's specifications. However, once the network is built the interconnections can be arbitrarily changed and components can be added or removed. The system configuration required for the installation of Neuro Solutions software are given as follows:

- Operating System: Windows XP / Vista / 7
- Memory: 512MB RAM (2GB recommended)
- Hard Drive: 500MB free space
- Video: 800x600 (1024x768 recommended)
- Microsoft Excel 2000/2002(XP)/2003/2007/2010

The ANN model developed in this thesis is used with NeuroSolutions, a Microsoft Excel plug-in that simplifies and enhances the process of getting data into and out of a NeuroSolutions neural network. This tool was designed to develop a complete solution to the problem in one

easy to use package while also giving the flexibility to customize its operation. The following figure presents a flow diagram describing the order in which the NeuroSolutions modules can be used to solve the problem. Notice that, at a minimum, only three operations are required: Tag Data, Create/Open Network, and Train Network. These operations are usually followed up by testing the models performance (Test Network) and applying the model to new input data where the output is unknown (Apply Production Dataset). More information could be found at <http://www.neurosolutions.com/>



.....**SECRET**.....