

Distributed Quantitative Precipitation Forecasts Combining Information from Radar and Numerical Weather Prediction Model Outputs

by

Auroop R. Ganguly

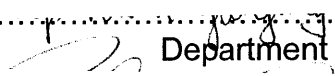
B. Tech. (Hons.), Civil Engineering
Indian Institute of Technology, Kharagpur-India, 1993


M. S., Civil Engineering
University of Toledo, Toledo, OH, 1997


Submitted to the Department of Civil and Environmental Engineering in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

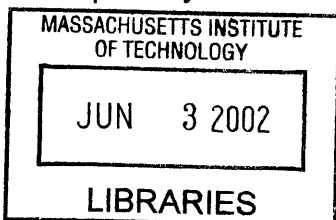
at the
Massachusetts Institute of Technology
June, 2002

© 2002 Massachusetts Institute of Technology
All rights reserved

Signature of Author.....
 Department of Civil and Environmental Engineering
February, 2002

Certified by.....
 Rafael L. Bras
Bacardi and Stockholm Water Foundations Professor
Thesis Supervisor

Accepted by.....
 Oral Buyukozturk
Chairman, Departmental Committee on Graduate Studies



BARKER

Distributed Quantitative Precipitation Forecasts Combining Information from Radar and Numerical Weather Prediction Model Outputs

Auroop R. Ganguly

Submitted to the Department of Civil and Environmental Engineering in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ABSTRACT

Applications of distributed Quantitative Precipitation Forecasts (QPF) range from flood forecasting to transportation. Obtaining QPF is acknowledged to be one of the most challenging areas in hydrology and meteorology. Recent advances in precipitation physics, Numerical Weather Prediction (NWP) models, availability of high quality remotely sensed measurements, and data dictated forecasting tools, offer the opportunity of improvements in this area. Investigative studies were performed to quantify the value of available tools and data, which indicated the promise and the pitfalls of emerging ideas. Our studies suggested that an intelligent combination of NWP model outputs and remotely sensed radar measurements, that uses process physics and data dictated tools, could improve distributed QPF.

Radar measurements have distributed structure, while NWP-QPF incorporate large scale physics. Localized precipitation processes are not well handled by NWP models, and grid average NWP-QPF are not too useful for distributed QPF owing to the spatial variability of rainfall. However, forecasts for atmospheric variables from NWP have information relevant for modeling localized processes and improving distributed QPF, especially in the Summer. Certain precipitation processes like advection and large scale processes could be modeled using physically based algorithms. The physics for other processes like localized convection or residual structures are not too well understood, and data dictated tools like traditional statistical models or Artificial Neural Networks (ANN) are often more applicable.

A new strategy for distributed QPF has been proposed that utilizes information from radar and NWP. This strategy decomposes the QPF problem into component processes, and models these processes using precipitation physics and data dictated tools, as appropriate and applicable. The proposed strategy improves distributed QPF over existing techniques like radar extrapolation alone, NWP-QPF with or without statistical error correction, hybrid models that combine radar extrapolation with NWP-QPF, parameterized physically based methods, and data dictated tools alone. New insights are obtained on the component processes of distributed precipitation, the information content in radar and NWP, and the achievable precipitation predictability.

Thesis Supervisor: Rafael Bras

Title: Bacardi and Stockholm Water Foundations Professor, CEE, MIT

Acknowledgments

This work was partially supported through the National Weather Service (NOAA, US Department of Commerce) Award NA07WH0437 as part of the Cooperative Agreement on the Comparison of Distributed and Lumped Hydrologic Models, and the Consiglio Nazionale delle Ricerche of Italy as part of the Cooperative Agreement in Climate Change and Hydrogeological Disasters. I would like to acknowledge the support I received at Oracle Corporation, where I have been employed full time for the last three and a half years.

I am grateful for the suggestions and help I received from my adviser Prof. Rafael Bras, and all my committee members: Profs. Dennis Mclaughlin and Elfatih Eltahir of the MIT Parsons Lab, Dr. Amar Gupta of the MIT Sloan School of Management, and Prof. Shafiqul Islam of the University of Cincinnati. I am grateful for the suggestions from Prof. Dara Entekhabi, to whom I owe many of the starting assumptions of this research. I would also like to thank Dr. Earle Williams of the Parsons Lab at MIT for his initial comments, and the helpful suggestions at various stages from Prof. Ana Barros at Harvard University. I gratefully acknowledge the help from the National Weather Service for making the radar data (in Topeka, KS) available and the help from Prof. Ana Barros and her collaborator Dr. Robert Kuligowski on information regarding this data and the necessary transformations. The NWP-Eta FOUS model outputs were purchased from the NCDC (thanks to Ms. Diane Dicus), and the gage corrected NEXRAD data in Oklahoma were downloaded from the ABRFC web page.

I cannot begin to acknowledge the constant support of my parents, Shreemati Deepali Ganguly and Shree Nirmal Kumar Ganguly, my “kid” brother Swaroop, and my wife Debashree. Without their constant support and encouragement, I could never have finished this work. Thanks also to all my friends, relatives, and coworkers at Oracle Corporation, MIT, IIT-KGP, and elsewhere, who supported and encouraged me while I was a full time graduate student at MIT, and especially during the final three and a half years, as I juggled between my doctoral research in rainfall forecasting at MIT, and full time work as the Product Manager for Demand Planning (a forecasting application) at Oracle. Special thanks to Ms. Elaine Healy and Ms. Cynthia Stewart of MIT for their continual help during the last three and a half years when I was a non-resident student at MIT, and during the preparation of the thesis as well as for other coordination issues.

Contents

Pages

1.0 Introduction.....	12-16
1.1 Benefits of Improved Distributed QPF	12-12
1.2 Challenges and Issues.....	13-14
1.3 Improving Distributed QPF.....	14-16
2.0 Executive Summary and Problem Setting.....	17-19
2.1 Problem Setting	17-17
2.2 Executive Summary	18-19
3.0 Recent Advances and Distributed QPF Literature.....	20-33
3.1 Numerical Weather Prediction Models and Precipitation Physics.....	20-23
3.2 Remote Sensing Methods and Ground Measurements	23-25
3.3 Traditional and Emerging Statistical and Data Dictated Tools.....	25-29
3.4 Radar Extrapolation, NWP Model Outputs and Precipitation Physics	30-31
3.5 Meteorological and Hydrological Approaches to Distributed QPF	32-33
4.0 Data Analyses and Investigation of QPF Strategies.....	34-55
4.1 Analysis of Information Content in the Available Data	35-42
4.2 Analysis of Distributed Precipitation Processes and Strategies.....	43-50
4.3 Exploring the Value of Conceptual Approaches to Distributed QPF	51-52
4.4 Consolidation of Insights for Improved Distributed QPF	53-55

5.0 Decomposition of Precipitation Processes and a Hybrid Model.....	56-78
5.1 Decomposition Rationale	56-56
5.2 Overview of the Hybrid Modeling Strategy.....	58-59
5.3 Details of the Radar Extrapolation Component.....	60-62
5.4 Details of the Large Scale Physics Component.....	63-68
5.5 Details of the Localized Evolution Component.....	69-74
5.6 Details of the Residual Structures Component	75-78
6.0 Reference Models and Skill Scores.....	79-87
6.1 Reference Models for Distributed QPF	79-82
6.2 Measures of Skill for Distributed QPF	83-87
7.0 Description and Availability of Data.....	88-96
7.1 Numerical Weather Prediction Model Outputs	88-92
7.2 Radar Rainfall	93-94
7.3 Precipitation Events	95-96
8.0 Results.....	97-137
8.1 Contribution of the Component Processes of Precipitation	97-102
8.2 Improvements to Distributed QPF.....	103-137

9.0 Discussions and Performance Analysis.....	138-148
9.1 Summer Precipitation Events.....	140-142
9.2 Winter Precipitation Events.....	143-147
9.3 All Seasons Combined.....	148-148
10.0 Summary	149-151
11.0 Future Research	152-158
11.1 Event Based and Continuous QPF Formulations	154-157
11.2 Distributed Precipitation Maps as Geometrical Objects	158-158
12.0 Conclusions	159-160
Appendix A: Use of Artificial Neural Networks (ANN).....	161-185
A.1 Hydrologic Scope of the QPF Problem.....	161-161
A.2 ANN Fundamentals and Forecasting Applications	162-176
A.3 ANN Formulation Details	177-185
Appendix B: Directional Structure in Precipitation and QPF Errors.....	186-204
Bibliography.....	205-218

List of Figures

Pages

Figure 4-1: Spatial Mean and Variance of Hourly Radar Rainfall	37
Figure 4-2: Spatial Mean of Radar Rainfall, Persistence, and Advection	38
Figure 4-3: Distributed Skills in Persistence and Advection.....	39
Figure 5-1: Flowchart for the Proposed Hybrid Modeling Strategy	59
Figure 5-2: Radar Extrapolation Strategy	62
Figure 5-3: Large Scale Physics Strategy: Temporal Disaggregation	66-67
Figure 5-4: Large Scale Physics Strategy: Spatial Disaggregation.....	68
Figure 5-5: Localized Evolution Strategy: Function Approximation	73
Figure 5-6: Localized Evolution Strategy: ANN.....	74
Figure 5-7: Residual Structures Strategy.....	78
Figure 6-1: The NIMROD QPF Strategy.....	82
Figure 7-1: NWP FOUS locations in the US	91
Figure 8-1: Contribution of Component Processes - Distributed Skills	101
Figure 8-2: Contribution of Component Processes - Aggregate Errors	102
Figure 8-3: Average 1/NMSE for all Seasons Combined	112
Figure 8-4: Average Threat Scores for all Seasons Combined	113
Figure 8-5: Average 1/NMSE for the Winter	114
Figure 8-6: Average Threat Score for the Winter	115
Figure 8-7: Average 1/NMSE for the Summer.....	116
Figure 8-8: Average Threat Score for the Summer	117
Figure 8-9: Precipitation Error Contour Plots for Summer Storm "A"	118
Figure 8-10: Precipitation Error Surface Plots for Summer Storm "A"	119

Figure 8-11: Precipitation Error Contour Plots for Summer Storm “B” 120

Figure 8-12: Precipitation Error Surface Plots for Summer Storm “B” 121

Figure 8-13: Precipitation Error Contour Plots for Summer Storm “C” 122

Figure 8-14: Precipitation Error Surface Plots for Summer Storm “C” 123

Figure 8-15: Precipitation Error Contour Plots for Winter Storm “D” 124

Figure 8-16: Precipitation Error Surface Maps for Winter Storm “D” 125

Figure 8-17: Precipitation Error Contour Plots for Winter Storm “E” 126

Figure 8-18: Precipitation Error Surface Plots for Winter Storm “E” 127

Figure 8-19: Precipitation Error Contour Plots for Winter Storm “F” 128

Figure 8-20: Precipitation Error Surface Plots for Winter Storm “F” 129

Figure 8-21: Contour Plots of Radar Rainfall, Proposed QPF, and pQPF 131

Figure 8-22: Contour Plots of Radar Rainfall, QPF and pQPF Errors for Storm A 132

Figure 8-23: Contour Plots of Radar Rainfall, QPF and pQPF Errors for Storm B 133

Figure 8-24: Contour Plots of Radar Rainfall, QPF and pQPF Errors for Storm C 134

Figure 8-25: Contour Plots of Radar Rainfall, QPF and pQPF Errors for Storm D 135

Figure 8-26: Contour Plots of Radar Rainfall, QPF and pQPF Errors for Storm E 136

Figure 8-27: Contour Plots of Radar Rainfall, QPF and pQPF Errors for Storm F 137

Figure 11-1: “Continuous” QPF Performance using Simulated Data 157

Figure B-1: Measured Precipitation and Persistence Contours: Storm A 187

Figure B-2: Measured Precipitation and Persistence Contours: Storm A 188

Figure B-3: Persistence Errors: Storm A 189

Figure B-4: Measured Precipitation and Persistence Contours: Storm B 190

Figure B-5: Measured Precipitation and Persistence Contours: Storm B 191

Figure B-6: Persistence Errors: Storm B 192

Figure B-7: Measured Precipitation and Persistence Contours: Storm C..... 193

Figure B-8: Measured Precipitation and Persistence Contours: Storm C..... 194

Figure B-9: Persistence Errors: Storm C 195

Figure B-10: Measured Precipitation and Persistence Contours: Storm D..... 196

Figure B-11: Measured Precipitation and Persistence Contours: Storm D..... 197

Figure B-12: Persistence Errors: Storm D 198

Figure B-13: Measured Precipitation and Persistence Contours: Storm E 199

Figure B-14: Measured Precipitation and Persistence Contours: Storm E 200

Figure B-15: Persistence Errors: Storm E 201

Figure B-16: Measured Precipitation and Persistence Contours: Storm F 202

Figure B-17: Measured Precipitation and Persistence Contours: Storm F 203

Figure B-18: Persistence Errors: Storm F..... 204

List of Tables

Pages

Table 4-1: Aggregate Radar Rainfall and NWP Outputs: Variability Explained	42
Table 4-2: Performance of Preliminary QPF for Advection and Evolution	47
Table 4-3: Comparison of Simple Disaggregation and Localized Evolution	50
Table 7-1: Sample FOUS data for the NWP-Eta model	88
Table 7-2: Significant Precipitation Events in Oklahoma	96
Table 9-1: Performance of QPF Component Processes in the Summer	142
Table 9-2: Performance of QPF Component Processes in the Winter	147

1.0 Introduction

1.1 Benefits of Improved Distributed QPF

Distributed Quantitative Precipitation Forecasts (QPF) are the primary forcing to lumped and distributed models of hydrology (Pessoa et al., 1993; Garrote and Bras, 1995; Ibbitt et al., 2000), and critical for accurate predictions of floods and flash floods. Of all natural disasters, floods are the leading cause of damage to human lives and property in the United States (NRC, 1996; Kim and Barros, 2001; Droegemeier et al., 2000) and other countries. Between 1989 and 1999, floods resulted in 988 human lives lost in the United States and \$4.5 billion property damage (US Army Corps of Engineers, 1999; Kim and Barros, 2001). Approximately 75% of all presidential declarations of natural disasters have involved floods (NRC 1996).

As indicated by Smith and Austin (2000), applications of short-term, distributed QPF range from urban hydrology (Cluckie et al, 1999) to space shuttle landings (Bauman and Businger, 1996). Browning and Collier (1989) tabulated the benefits of nowcasting (e.g., 1-3 hour QPF). Severe weather warnings for heavy rain and gust fronts, and water management operations like flood forecasting, sewer operations and spillway operations could save lives and property. Agriculture, irrigation and civil engineering operations like concrete pours could benefit from advance warning systems that use accurate forecasts of rainfall and floods. QPF is useful for the transportation industry, including air, road, sea, and rail transport. Diverse activities like military operations and leisure are influenced by QPF, as are industry sectors like power and communication.

1.2 Challenges and Issues

Distributed QPF is acknowledged to be among the most challenging areas in hydrology and meteorology (Fritsch et al., 1998; Antolik, 2000; Collier and Kryzysztowicz, 2000), and improvements to QPF over the decades have been somewhat limited (Smith and Austin, 2000). The degree of belief in distributed QPF is so low that until recently, operational models of hydrology were run (and often continue to be run) without these inputs. However, not using QPF corresponds to using forecasts of no rainfall (Collier and Kryzysztowicz, 2000), which could significantly underestimate the likelihood of floods.

Several factors contribute to make precipitation "notorious for being difficult to quantitatively predict accurately" (Gaudet and Cotton, 1998). The atmospheric processes that cause precipitation could occur at several scales, ranging from large air mass movements to extremely localized convective events (Rogers and Yau, 1989). The relations of precipitation quantity to the basic meteorological state variables are indirect. The meteorology of precipitation is influenced by air motion and turbulent eddies leading to cloud formation, variation of cloud and aerosol properties, and microphysical processes that dictate droplet growth and evaporation. The net effect of all of the above is that precipitation is extremely variable in space and time, and hence very difficult to forecast.

The requirements and achievable predictability for the temporal and spatial resolution at which QPF needs to be produced vary by forecast lead times. At very short lead times (0-1 hour), high forecast resolutions are needed. Forecasts based on extrapolation of observations like radar rainfall are often found to be most skillful in this

context (Golding, 2000; Collier, 1991). At lead times greater than about 6 hours, Numerical Weather Prediction (NWP) models by themselves provide the best available forecasts of precipitation, at lower resolutions in space and time. Researchers like Krishnamurti et al. (1999) have demonstrated that an optimal combination of the outputs from several NWP models could improve forecasts at these resolutions. This research focuses on 1-6 hour QPF, for which forecasts are required at high resolutions for hydrologic applications (NOAA, 1997). For these forecast scenarios, an optimal combination of information from NWP model outputs and distributed radar measurements could lead to improved distributed QPF.

1.3 Improving Distributed QPF

Better understanding of the physics of the weather system and improved computer speeds have enabled Numerical Weather Prediction (NWP) models to forecast precipitation and atmospheric quantities in real time at higher resolutions, and to consider effects like topography and convective parameterizations. The best available and archived NWP model results currently available in the United States are those from the 48-km Eta model, in the form of FOUS (Forecast Outputs for the United States). The forecasts from NWP-Eta are 6 hourly and 48 km grid average.

Availability of high quality measurements from remote sensing tools like radar and satellites are continually improving in accuracy, coverage and resolutions. NEXRAD (NEXt generation RADar) covers the continental United States extensively, and precipitation measurements in the HRAP (Hydrologic Rainfall Analysis Project) grid are available at 4x4 km resolutions. Uncalibrated Stage I and calibrated Stage III precipitation estimates from radar are made available by the River Forecast Centers

(RFC). The highest quality radar data are available in the ABRFC (Arkansas-red Basin River Forecast Center), the best calibration with ground based measurements is in Oklahoma due to the presence of the Oklahoma MesoNet. Antolik (2000) of the National Weather Service (NWS) mentions that future research needs to explore methods that combine the 48-km NWP-Eta model outputs with Stage III radar data.

Evaluation of available and emerging tools and methods for distributed QPF indicate that the best strategy to combine radar and NWP information could be through a hybrid modeling strategy that decomposes the problem into component processes, and handles each process using precipitation physics or data dictated tools, whichever is most applicable. The hybrid modeling strategy must be designed to extract the maximum information from NWP and radar at different space-time resolutions.

This work decomposes the distributed QPF problem into four component processes. These are "Radar Extrapolation", "Large Scale Physics", "Localized Evolution", and "Residual Structures". A hybrid modeling strategy is proposed to account for these processes, each component successively builds upon the results of the previous components. Radar extrapolation is handled through advection. QPF produced by Numerical Weather Prediction models (NWP-QPF) is assumed to represent large scale physics at lower resolutions, this is disaggregated in time and space using the results of advection. Localized evolution is assumed to be conditioned on the results of advection and large scale physics, and uses NWP forecasts of atmospheric variables in a data dictated Artificial Neural Network (ANN) formulation. Error analysis of distributed and aggregate (in space and time) predictions indicate that improvements can be obtained by combining various outputs in an "optimal" way.

The proposed hybrid model is shown to appropriately combine the individual component processes, and succeeds in improving distributed QPF over the "state of the art". The use of data dictated tools alone, as well as the use of parameterized physically based models, are shown to be sub-optimal to a hybrid modeling strategy. We consider persistence, advection, and a hybrid model that combines advection and NWP-QPF, as the commonly used reference models. These are compared with the proposed hybrid model in terms of skill measures. The goodness of a forecast is quantified by an average distributed measure of skill based on the Root Mean Squared Errors (RMSE), and a surrogate cost measure that is based on a redefined threat score, which considers the number of times the forecasts and observations lie within given thresholds. The improvement in distributed QPF is also depicted by contour and surface plots of the QPF errors. Detailed analyses of performance of the various QPF components offer interesting insights on the precipitation processes and on the information content in radar measurements and NWP model outputs.

2.0 Problem Setting and Executive Summary

2.1 Problem Setting

In the United States, high resolution precipitation estimates are obtained from the NEXRAD (NEXt generation RADar) system. Uncalibrated radar measurements are referred to as “Stage I”. The River Forecast Centers (RFC) calibrate radar measurements with ground based information, to obtain “Stage III” precipitation estimates. The available resolution of radar precipitation is 4x4 km and 1 hour, in the HRAP (Hydrologic Rainfall Analysis Project) grid. The best radar calibrations are available in the ABRFC (Arkansas-red Basin River Forecast Center) in Oklahoma, owing to the presence of the Oklahoma Mesonet.

Numerical Weather Prediction (NWP) models take into account large scale physics of the atmosphere, and generate low resolution forecasts for precipitation and atmospheric variables. The best NWP model for which archived results are readily available is the 48-km Eta. Outputs from the NWP-Eta are available at 48 km grid spacing and for 6, 12, 18, 24, 30, 36, 42, and 48 hours.

The objective of this work is to obtain distributed Quantitative Precipitation Forecasts (QPF) at 1-6 hour lead times, which are useful for hydrologic applications like flood forecasting. The achievable resolutions are 4x4 km and 1 hour (i.e., radar resolution). The first part of this work performs an analysis of the information, and evaluates traditional and emerging tools and strategies for distributed QPF. The second part attempts to use these insights to develop a new strategy for distributed QPF that improves over the “state of the art”.

2.2 Executive Summary

Distributed QPF is useful for a variety of applications ranging from the prediction of floods and flash floods, to transportation and space shuttle landings. This research evaluates existing methods for distributed QPF like radar extrapolation (i.e., persistence and advection), NWP-QPF (raw, or after statistical post-processing), hybrid models that combine radar extrapolation with NWP-QPF, data dictated tools alone (which include traditional statistical techniques and nonlinear methods like ANN), and parameterized physically based methods. A new strategy is proposed that improves distributed QPF over existing methods.

Improvement in distributed QPF is achieved by decomposing the problem into component processes. The decomposition strategy attempts to utilize available precipitation physics and data dictated tools, through the use of relevant information from NWP model outputs (including QPF and atmospheric variables) and remotely sensed radar measurements. The proposed strategy has four component processes, which we shall call “Radar Extrapolation”, “Large Scale Physics”, “Localized Evolution” and “Residual Structures”.

Radar Extrapolation provides distributed structure, and is best modeled using advection. The proposed advection formulation uses two-dimensional spatial correlation to estimate the velocity scales at lagged time steps, and single exponential smoothing to forecast the velocity scales. These are in turn used to translate the current precipitation map. Large Scale Physics is best represented by NWP-QPF, which is available at lower resolution. This is disaggregated in time using linear interpolation, and then corrected for errors by comparing with

spatially aggregated radar measurements, through an Auto-Regressive statistical time series technique. The time disaggregated and error corrected QPF from NWP is then disaggregated in space using the results of Extrapolation. Localized Evolution is assumed to occur on top of Extrapolation and Large Scale Physics. This is assumed to be a function of the precipitation intensity at neighboring pixels, and the large scale atmospheric state, which in turn is indicated by NWP forecasts of atmospheric variables. The physics at these scales is not well known, so localized evolution is modeled using an ANN based nonlinear regression framework. The Residual Structures strategy attempts to reduce the distributed and aggregate errors at all lead times by utilizing the complementary skills of the other components processes.

The proposed strategy produces spatially distributed QPF and pQPF (i.e., upper and lower confidence bounds) for 1-6 hour lead times. Improvements to distributed QPF are demonstrated through quantitative skill scores, and visually through plots of the errors.

3.0 Recent Advances and Distributed QPF Literature

3.1 Numerical Weather Prediction Models and Precipitation Physics

Numerical models of the weather have been improving continually over the years (see Hartmann, 1994), currently atmospheric centers around the globe routinely run their own versions of these models. NWP models attempt to describe atmospheric state evolution with three dimensional partial differential equations, and solve these by numerical integration. The boundary conditions are derived from data assimilation techniques, which use measurements and results of lower resolution models. Initial conditions are obtained from measurements and previous model runs.

The greatest source of uncertainty in NWP model outputs are subgrid scale processes like convection, clouds, and Planetary Boundary Layer (PBL) eddies. The coverage of the NWP models and the constraints imposed by the numerical schemes preclude the possibility of grid resolutions that are high enough to capture these subgrid processes, however these processes could significantly influence the weather system. Convection and clouds, for example, have large impacts on rainfall generation and cause the greatest uncertainty in NWP model outputs. NWP models generate grid average forecasts of various quantities including QPF, the latter is often thought to be relatively less accurate and is typically used as a measure of model performance (Mesinger, 1996; Antolik, 2000).

There is a belief that gradual improvements to NWP models (both the physics as well as the resolutions) and computer speeds would be the best way to improve distributed QPF in the future. Given the nature of the distributed precipitation processes

(Islam et al., 1993a; Rodriguez-Iturbe et al., 1989; Schertzer and Lovejoy, 1987), and the limited understanding of the precipitation physics at those scales (Greco and Krajewski, 2000; Rogers and Yau, 1989), this assumption remains to be proven (see also, Georgakakos 2000; Nakakita et al., 1996). Scientists are beginning to generally acknowledge that the ultimate limit to predictability from NWP models is dictated by the chaotic nature of the weather system (Lorenz, 1982; Islam et al., 1993b; Li et al., 1995), which causes the numerical model results to be very sensitive to the prescribed initial conditions. This is reflected on the widespread use of ensemble techniques that attempt to run NWP models with a number of possible initial conditions to generate a set of probable outcomes (e.g., Du et al., 1997). Another consideration is that the currently achievable resolutions from NWP models are not adequate for generating QPF at scales useful for hydrologic purposes (Antolik, 2000; Georgakakos, 2000). However, NWP models are rapidly evolving, and it is conceivable that the forecast resolutions would approach hydrologically useful scales in the future (Droegemeir et al., 2000).

Some researchers have suggested using radar measurements to initialize NWP model runs through data assimilation, or “nudging” NWP models with radar estimates (Jones and Macpherson, 1997). However, these nudging methods do not address several key issues, as pointed out by Golding (2000) and others. First, these methods do not alleviate the problems caused by the lower NWP model resolutions. Initializations and calibrations still need to be performed at NWP resolutions. Second, the models themselves are not a good approximation of reality owing to the relatively poor understanding of the precipitation physics at distributed scales. Finally,

precipitation is not a state variable of the weather equation but needs to be derived, which causes the accuracy to suffer as mentioned by Antolik (2000) and others.

While NWP models cannot account for distributed precipitation physics too well, parameterized physically based models do exist for distributed QPF (Georgakakos and Bras, 1984 a,b; Lee and Georgakakos, 1990, 1996; French et al., 1994 a,b; Dolcine et al., 2000). The parameters of these models could be potentially related to (and estimated from) NWP model outputs. For example, Nakakita et al. (1996) obtained improvement in QPF at 1-3 hour lead times by using forecasts for atmospheric variables (but not QPF) from a NWP model, through a physically based scheme developed for distributed scales (see also, Sugimoto et al., 2001).

Previous researchers have attempted to make use of statistical and fractal models of rainfall, and "downscale" NWP-QPF in a manner that preserves the statistical properties of distributed rainfall. For example, the downscaling algorithm of Perica and Fofoula-Georgiou (1996) uses the CAPE (Convective Available Potential Energy). Bindlish and Barros (2000) suggest a way to disaggregate rainfall in regions of complex terrain. While these disaggregation models could be useful to understand and forecast the statistical properties of rainfall at distributed scales, they might not be optimal in forecasting applications that attempt to generate distributed QPF for hydrologic or other applications.

In the United States, operational rainfall forecasts at selected locations are made using MOS (Model Output Statistics) techniques, which use linear regression equations to relate NWP model outputs to point rainfall (Antolik, 2000). These seem to work better in the "cooler seasons", possibly because localized convection dominates in the

"warmer seasons". An interesting observation from these results is that the forecasts for precipitation (QPF) from the NWP models are not found to be the most important predictor for rainfall at a point. This underscores the spatial variability of precipitation at distributed scales, and the relevance of NWP forecasts of atmospheric variables to precipitation generation. In the US, the best NWP model for which outputs are currently readily available is the 48 km NWP-Eta (Black, 1994). The model outputs are forecasts of precipitation (QPF), and atmospheric variables like wind, pressure, temperature, humidity and instability indices. The forecasts are for 48 km grid average and every 6 hours, till 48 hours (6, 12, 18, 24, 30, 36, 42, 48 hours).

3.2 Remote Sensing Methods and Ground Measurements

Advanced remote sensing systems for precipitation estimation, like the NEXt generation RADar (NEXRAD), and geostationary satellites, are in place in countries like the US.

Satellite observations have been used for rain rate estimation (Sorooshian et al., 2000) and for forecasting precipitation (Browning, 1979) and floods (Kim and Barros, 2001). Satellite observations are not available at resolutions directly useful for hydrologic purposes, and could have significant errors, especially when related to quantities relevant to precipitation generation (Smith and Austin, 2000).

The information from weather radar is believed to be more directly relevant to distributed QPF. Weather radar is used to measure reflectivity, and atmospheric variables like Vertically Integrated Liquid (VIL) water content at near surface layers (Rogers and Yau, 1989). The intensity of the back-scattered radiation is quantified by the radar reflectivity, Z . The quantity Z is related to the size and distribution of the water

droplets (or hydrometeors) in the atmosphere. These hydrometeors cause precipitation, hence empirical relations have often been used to relate Z to the actual rainfall rate R . Empirical power laws of the type $Z=aR^b$ (a , b are constants) have often been used, as in the Marshall-Palmer relation, which assumes an exponential form for the Drop Size Distribution (DSD) of the hydrometeors. These power law assumptions could cause significant errors in QPF and hydrographs. Some researchers have tried the Probability Matching Method (PMM) or variations thereof (see Sun et al., 2000, and the references therein) to obtain rainfall from radar.

Remotely sensed measurements typically need to be compared and calibrated with ground based measurements. Precipitation estimated from radar measurements often show systematic bias, and at times fail to adequately represent the true rainfall observed at the ground (Browning and Collier, 1989). Rain gages are used to measure rainfall at specific locations, and represent the "ground truth" at these points. However, even fairly dense rain gage networks often fail to provide significant distributed information owing to the extreme variability of rainfall in space (Golding, 2000). One compromise that is often attempted is to use radar measured quantities to estimate rainfall with a correction factor from rain gage measurements (e.g., for removal of any systematic bias). Some researchers tend to believe that an optimal combination of radar and rain gages produces the best results, although there is some debate on the value of calibration for distributed rainfall estimation (Kitchen and Blackall, 1992).

In the United States, the ABRFC (Arkansas-red Basin River Forecast Center) has the most dense NEXRAD radar coverage. In the Hydrologic Rainfall Analysis Project (HRAP) grid, the NEXRAD resolution is 4x4 km and hourly. Uncalibrated Stage I

radar data, as well as calibrated Stage III data, are made available through the RFC (River Forecast Centers). The best quality calibrated radar data are in the state of Oklahoma, where high quality ground based measurements are available from the Oklahoma MesoNet.

3.3 Traditional and Emerging Statistical and Data Dictated Tools

Statistical techniques for QPF have been evolving continuously over the years. One approach is to model, analyze and forecast time series and space-time phenomena from observations. Another could be to estimate and update the parameters of physically based formulations in real-time with new observations, for example through the use of discrete Kalman filter formulations. The tools available for statistical analysis of data have improved over the years. Among these complex tools are Artificial Neural Networks (ANN, see Bishop, 1996), which have been shown to efficiently model nonlinear relationships among data sets (Vemuri and Rogers, 1994; Weigend and Gershenfeld, 1994).

As pointed out by Grecu and Krajewski (2000), the physics of rainfall at distributed scales in space and time is not well understood. Previous researchers have attempted to obtain mathematical or stochastic formulations for the description and forecasting of distributed rainfall. An example is the cluster model of Rodriguez-Iturbe and Eagleson (1987), where statistical distributions were used to describe rainfall cells in space and time. McLaughlin et al. (1990) demonstrated how the parameters of these models could be determined from data through the use of a discrete Kalman filter formulation. Other researchers suggested parameterized, physically based models, and estimated the parameters thereof using a Kalman filter (for example, Georgakakos and

Bras, 1984 a,b). These are conceptually appealing, but the inherent physics or statistics is typically a crude parameterization, and the applicability is often limited to specific situations for which these models were designed (Smith and Austin, 2000).

Forecasts of rainfall using statistical time series methods have been attempted before. Burlando et al. (1996) used a multivariate ARIMA (AutoRegressive Integrated Moving Average) formulation. Toth et al. (2000) forecast rainfall time series using a variety of methods like ARMA, ANN, and K-NN (K Nearest Neighbor), and verified the QPF by routing these through models of hydrology. The method based on ANN performed the best. To combine information from various models or measurement sources, Kryzysztowicz (1999) suggested using Bayesian statistics. Similar methods based on human computer interaction were alluded to by Browning and Collier (1989). These methods might not fully address the need to generate distributed forecasts, or to combine information at disparate space-time resolutions.

Estimation and forecasting based on Artificial Neural Networks (ANN) have shown promise in geophysical applications (Elsner and Tsonis, 1992; Grieger and Latif, 1994; Navone and Ceccato, 1994; Tsonis et al., 1994), including Quantitative Precipitation Estimation (QPE) and Forecasting (QPF). ANN are complex data dictated tools that have been shown to act as "universal function approximators", and converge faster than other traditional approximators (Bishop, 1996; MacKay 1991, 1992). The Santa Fe time series competition saw some of the best forecasts from ANN based methods (Weigend and Gershenfeld, 1994). However, there is a need for caution when these tools are used. The applicability of ANN for a specific situation needs to be determined from an understanding of the statistics, domain knowledge and data

analysis. The discussions by Chatfield (1993) illustrate how it could be easy to misuse these tools, and the Sante Fe competition showed that misleading results could be obtained from not so good ANN techniques.

ANN based techniques have been shown to perform better than existing methods for fusion of radar and rain gage estimates in most situations (Matsoukas et al., 1999). Multiple information from satellite remote sensing (IR and/or VIS) were optimally combined for QPE using an ANN based technique by Hsu et al. (1999) and Sorooshian et al. (2000). QPE using satellite brightness temperature was also attempted by Hsu et al. (1997). Radar measurements were related to rainfall rate using an ANN based methodology by Xiao and Chandrasekar (1997), which improved over existing methods. Kuligowski and Barros (1998 a,b) obtained 0-6 hr QPF at a point using an ANN based time series forecasting method and information from multiple sources. Barros et al. (1996) forecast 1-hour QPF time series at a point using wind information and ANN. Hall (1996) used NWP-Eta model outputs and ANN for 24-hour gage mean and probabilistic rainfall forecasts. Single station QPF using pattern detection was attempted by Dumais and Young (1995). Basin average and basin maximum QPF were produced from NWP outputs through an ANN based approach by Gillispie (1993 a,b,c), results improved over the operational weather service forecasts. Lindner and Krien (1993) used ANN to determine the maximum rain in Mississippi state from radiosonde. Kim and Barros (2001) improved flood forecasts by the use of ANN and multisensor information. ANN based cloud classification have been attempted by Lee et al. (1990), Visa et al. (1995) and in the Gandolf system of Golding (2000). Analysis and tracking of radar images using ANN were attempted by Ding et al. (1993)

and Denoeux and Rizand (1995). Marzban and Stumpf (1996) modeled a tornado using ANN. McCann (1992) modeled and forecast thunderstorms using ANN, and attempted to interpret the weights to obtain insights on the process physics. Lightning was modeled by Frankel et al. (1993) using ANN. An ANN based approach was used to forecast the Indian monsoon by Navone and Ceccato (1994).

The promise and pitfalls of ANN usage in the context of QPF are illustrated by Toth et al. (2000) and Grecu and Krajewski (2000). Toth et al. (2000) forecast precipitation time series, a method based on ANN performed the best. The works of Kuligowski, Barros, Gillispie, McCann and others described earlier also seem to indicate that ANN have the potential to significantly improve QPF. However, Grecu and Krajewski (2000) showed that ANN methods are not always appropriate for precipitation forecasting, specifically for modeling certain processes like advection. This suggests that ANN methods could be useful to handle certain components of the QPF problem, while other techniques might be more appropriate for the remaining components. This leads to the possibility of a hybrid modeling strategy for QPF.

Hybrid models for QPF that statistically combine results from NWP models and extrapolation techniques have been used by the meteorological community (see Smith and Austin, 2000, for a review). Golding (2000) mentions that improving QPF at 1-6 hour lead time is a key requirement, and this could be achieved by statistically combining radar and NWP information. As discussed earlier, at 0-1 hour, NWP model outputs do not improve over radar extrapolation alone (Collier, 1991), while for longer leads (greater than 6 hours), forecasts are skillful only at aggregate space-time scales, at which NWP model outputs are available. The currently operational NIMROD system

in the U. K. Meteorological Office (Golding, 2000) combines radar advection with NWP-QPF using relative weights, and was designed for 1-6 hour QPF. This hybrid strategy improves distributed QPF at short lead times (1-6 hours) over radar extrapolation alone and NWP-QPF alone. Smith and Austin (2000) hail this NIMROD strategy as a great step forward in improving QPF.

Our studies indicate that complex phenomena like precipitation could be best modeled and forecast by decomposing the problem into component processes, and using process physics where these are most applicable, and data dictated tools like ANN or traditional statistics where these are best suited. In other words, this suggests a hybrid modeling approach, that uses the most applicable tool for each component process. The "state of the art" models for distributed QPF are in essence simple hybrid models, that combine the results of radar extrapolation achieved through advection or pattern recognition, with the results of large scale physics as represented by grid average QPF from NWP models (e.g., Golding, 2000). Hybrid models have been used in other areas related to forecasting and optimization. For example, in an inventory optimization problem undertaken at the MIT Sloan School of Management by the author along with others, inventory was reduced and customer satisfaction levels were maintained by using a hybrid model that combined domain specific knowledge with data dictated ANN tools (Reyes-Aldasoro et al., 1999).

3.4 Radar Extrapolation, NWP Model Outputs, & Precipitation Physics

Greco and Krajewski (2000) conclude that advection (or translation based on estimates of a velocity scale) is the only physical process that contributes to QPF improvement from radar extrapolation. They tried two methods to estimate the velocity scales for advection, one based on spatial correlation, and another based on ANN. Improvement from the ANN was minimal to zero, which led them to conclude that complex statistical models do not improve forecasts based on extrapolation. Smith and Austin (2000) and Browning and Collier (1989) cite references to conclude that complex or higher order methods do not improve radar extrapolation results over simpler and/or linear methods. Browning and Collier (1989) mention that the decorrelation time for precipitation structure is of the order of an hour. Mecklenburg et al. (2000) demonstrated that local structures persist for about 40 min for convective situations and for slightly higher lead times for stratiform. The relative skills obtained by Golding (2000) show that advection could have two distinct skill behaviors (and modeling strategies). The 15-minute (or sub-hourly) Advection, which uses 15-minute precipitation maps, shows significant relative skills at lead times of 0-1 hour, but decays sharply thereafter. The skills from the 1-hour Advection are relatively low initially, but the decay rate is lower for 1-6 hour leads. Golding (2000) uses either spatial correlation or NWP wind profile information for estimating the velocity scales for advection, whichever improves the forecast at analysis time. Browning and Collier (1989) mention that the use of measured or NWP wind has not been conclusively proven to improve advection forecasts.

We mentioned that NWP-MOS studies, which linearly regress NWP outputs with point rainfall, have discovered that QPF from NWP is not the most important predictor (Antolik, 2000). Precipitation exhibits high spatial variability, and NWP models do not predict QPF as well as atmospheric variables (Mesinger, 1996). Regression relations between atmospheric variables (or, instability indices) and convective rainfall have been reported (Zawadzki et al., 1981). Pepler and Lamb (1989) reported regression relations for “growing season” rainfall, where convective and stratiform regimes might not be easily separable. These results, when viewed together, appear to suggest that while a simple disaggregation of NWP-QPF in space might have limited predictive power for distributed QPF, forecasts of atmospheric state variables and instability indices from NWP could have additional information content, especially in convective situations. This possibility is also suggested from the results of Perica and Fofoula-Georgiou (1996), and Nakakita et al. (1996). However, previous researchers (e.g., Grecu and Krajewski, 2000) acknowledge that precipitation physics at distributed scales is not too well understood, and existing models for convective processes often rely on simplifying assumptions and parameterizations (Smith and Austin, 2000). In a nutshell, previous research seems to strongly indicate that there is information in NWP model outputs other than QPF (e.g., forecasts of atmospheric variables) that could potentially be exploited to improve distributed QPF. However, how best to exploit this information remains an open research area.

3.5 Meteorological and Hydrological Approaches to Distributed QPF

The meteorological community has been improving NWP models continuously. There is a belief that better NWP models, faster computer speeds, and better data assimilation techniques that use (for example) radar measurements to initialize or “nudge” NWP models, could be sufficient to improve distributed QPF. We agree with Golding (2000) that data assimilation or initialization with distributed measurements would not make up for the inability of the NWP models to handle localized subgrid scale processes like convection. One line of argument is that NWP model resolutions and convective (and orographic) parameterizations would improve in the future, thus the best approach would be to estimate the parameters of the NWP models themselves (e.g., using discrete filter formulations) from radar measurements. We believe that given the spectrum of spatial and temporal scales at which rainfall processes could occur, the disparity between the resolutions of NWP models and measurement devices like radar would continue to exist. To circumvent this, the meteorological community has often developed hybrid models that combine extrapolation techniques with NWP-QPF (Golding, 2000; Kryzysztowicz, 1999; Browning, 1979).

The value of NWP model outputs for distributed QPF is yet to be generally acknowledged or proven, and the physics of precipitation is not too well understood at distributed scales. Thus a line of argument could be that data dictated models alone (e.g., based on traditional statistics, fractal schemes, or ANN) would be best in these situations. The hydrologic community has often used different mathematical or stochastic approaches to describe and forecast aspects of rainfall structure and evolution. For example, Johnson and Bras (1980) proposed a new technique to

forecast rainfall in space and time using rain gage information, while Rodriguez-Iturbe and Eagleson (1987) developed stochastic models to describe rainfall structures in space and time. However, previous studies of advection (Grecu and Krajewski, 2000) and nowcasting (Smith and Austin, 2000; Tsonis et al, 1981) demonstrate that complex data dictated tools might not be best to model certain component processes of precipitation. For example, the lack of improvement in distributed QPF from direct application of these tools (Browning and Collier, 1989; Grecu and Krajewski, 2000) indicate that complex data dictated methods by themselves, either fail to significantly improve estimates for velocity scales in advection, and/or fails to model the dynamics of the distributed precipitation evolution.

Finally, given the fact that the value of NWP model outputs for distributed QPF is an open research area, the use of NWP model outputs as surrogate measurements could be questioned. Krishnamurti et al. (1999) have shown that combining the results of different NWP models based on relative skills (a “multimodel superensemble”) could improve forecasts at NWP resolutions. The work of researchers like Kuligowski and Barros (1998a) suggest that post-processing of NWP model outputs using statistical methods could have value in distributed QPF applications.

4.0 Data Analyses and Investigation of QPF Strategies

In this section, we analyze the data from NWP and radar and outline the results of our investigative studies to distributed QPF processes and strategies. These confirm insights of past researchers described in the previous chapter, and suggests new ones.

The studies carried out here could be broken down into three broad categories:

- Analysis of Information Content in the Available Data
- Analysis of Distributed Precipitation Processes and Strategies
- Applicability of Parameterized, Physically Based Models

These studies were designed to explore the potential value of the available strategies in their present form for distributed QPF, given the nature of the available data. The insights gained from these analyses were used to formulate a hybrid strategy for distributed QPF, which is described in later sections.

4.1 Analysis of Information Content in the Available Data

Two sources of information were considered in this research, distributed precipitation estimates from radar, and space-time averaged NWP model outputs for precipitation and atmospheric variables. The behavior of radar measurements with time, both in terms of their aggregate statistics and distributed structure, was explored. Correlation studies were carried out to explore the relationship of NWP model outputs to aggregated radar rainfall.

4.1.1 Analysis of Hourly Radar Rainfall

Spatially aggregate statistics of the hourly radar rainfall, as a function of time, were investigated. Uncalibrated Stage I radar rainfall was used for events in Topeka, Kansas, while calibrated Stage III radar rainfall from the ABRFC region in Oklahoma was used. We looked at the spatial mean of the radar map and the variance around this mean. Data analyses on hourly radar rainfall demonstrate the non-stationary nature, both in terms of the spatial mean and the variance around this mean. These were calculated at each hourly time step using the precipitation intensities of individual pixels. The hourly distributed precipitation maps show minimal persistence, but exhibit relatively higher skills upon advection, especially in the Winter. The skills are more evident in terms of the spatially aggregate statistics of the distributed rainfall.

Figure 4-1 uses a Summer and a Winter precipitation event, and demonstrates the non-stationarity in the rainfall process in terms of both the spatial mean, and the variance around this mean, as a function of the lead time in hours. Figure 4-2 show the spatially aggregate mean from radar, and the corresponding forecasts obtained from persistence and advection. These indicate the extent to which the aggregate

information is preserved in hourly rainfall maps, with or without translation. While significant information is not preserved in either case, advection does show more skill than simple persistence in terms of the aggregate statistics. The skills are worse in terms of the distributed measures. Figure 4-3 shows the lack of distributed skills from either persistence or translation for hourly QPF. The distributed skill was quantified in terms of the Root Mean Squared Errors (RMSE) normalized by the standard deviation of the verification data.

Both the mean and the variance showed significant non-stationarity in the 1-6 hour range. The non-stationarity was higher in the Summer than in the Winter. The hourly radar maps showed limited persistence, both in an Eulerian or in a Lagrangian sense (i.e., after advection), especially in the Summer. The large changes in aggregate statistics could indicate the influence of large scale precipitation processes. The lack of persistence, whether in an Eulerian or in a Lagrangian sense, could indicate the possibility of localized processes.

In general, we observed that radar advection performed better in the Winter than in the Summer, especially with respect to the overall distributed error measure (i.e., the NMSE when all pixels were considered) and in terms of the aggregate error measure (i.e., the error in the spatial mean). However there were a few situations, more predominant in the Winter, when the advection forecasts missed the rainfall areas almost completely. An overall distributed error measure, given raining conditions, was obtained by considering pixels with non-zero rainfall only (for the NMSE calculations). In terms of this NMSE for “wet” pixels only, skills from advection and persistence were comparable in the Winter, but advection performed slightly worse in the Summer.

Figure 4-1: Spatial Mean & Variance of Rain as a function of Time in Hours

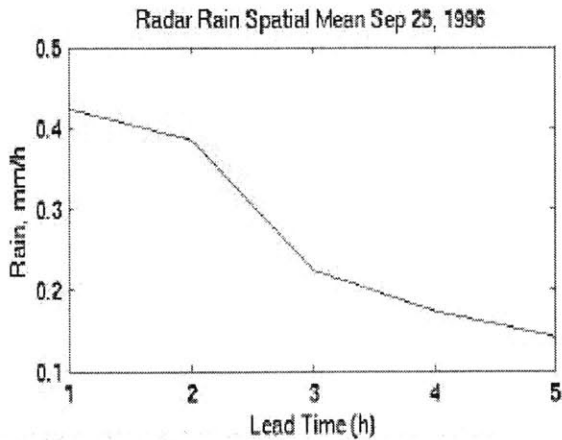


Figure a: Spatial Mean of Radar Rainfall (mm/h) for an event in the Winter as a function of lead time in hours.

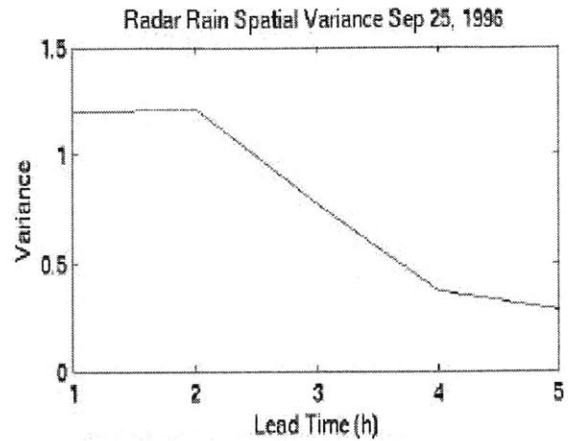


Figure b: Spatial Variance of Radar Rainfall (around the mean) for an event in the Winter as a function of lead time in hours.

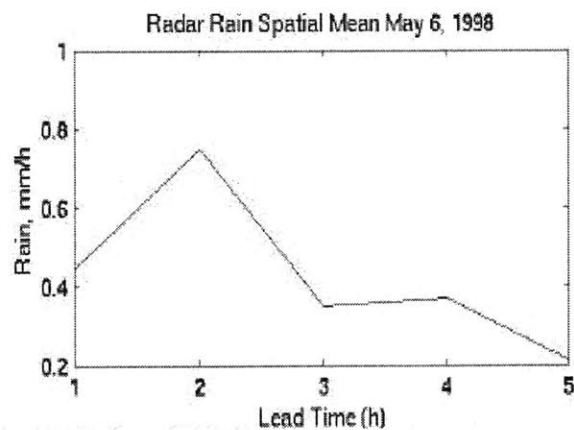


Figure c: Spatial Mean of Radar Rainfall (mm/h) for an event in the Summer as a function of lead time in hours.

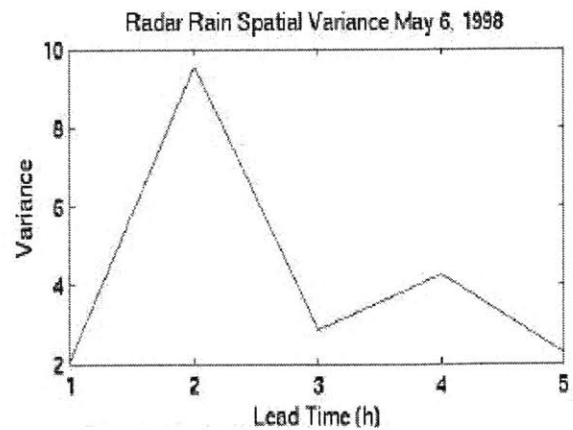


Figure d: Spatial Variance of Radar Rainfall (around the mean) for an event in the Summer as a function of lead time in hours.

Figure 4-2: Spatial Mean of Radar Rainfall, Persistence and Advection

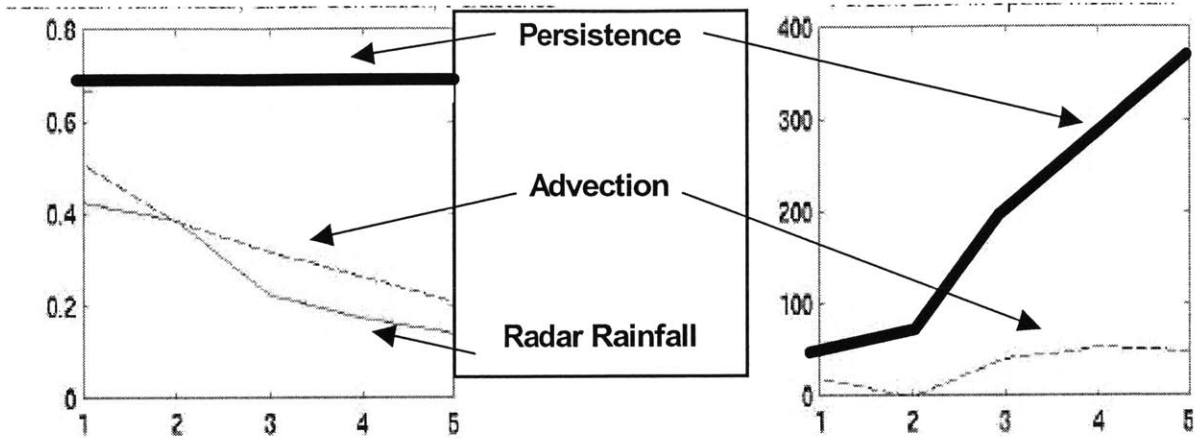


Figure a: Spatial Mean of Radar Rainfall Rainfall (mm/h), Persistence, and Advection, for a precipitation event as a function of hourly lead. A Winter event.

Figure b: Percent error in the spatial mean from Advection and Persistence, for a precipitation event, as a function of lead time in hours. A Winter event.

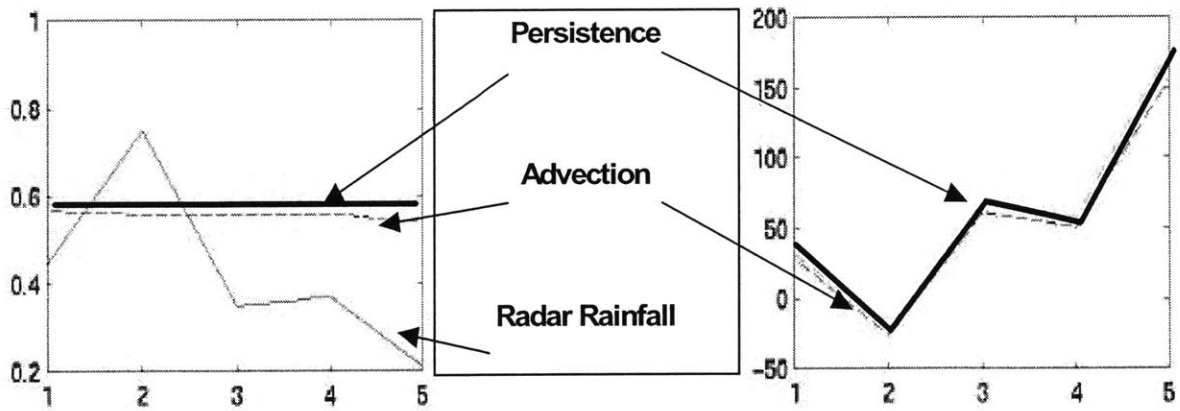


Figure c: Spatial Mean of Radar Rainfall Rainfall (mm/h), Persistence, and Advection, for a precipitation event as a function of hourly lead. A Summer event.

Figure d: Percent error in the spatial mean from Advection and Persistence, for a precipitation event, as a function of lead time in hours. A Summer event.

Figure 4-3: Distributed Skills in Persistence and Advection for Hourly Rainfall (NMSE and Wet-NMSE)

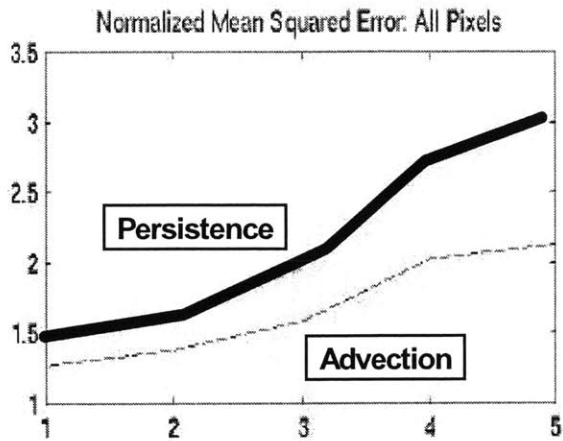


Figure a: A Winter event. RMSE normalized by the variance of the verification data (NMSE) for all radar pixels, as a function of lead in hours.

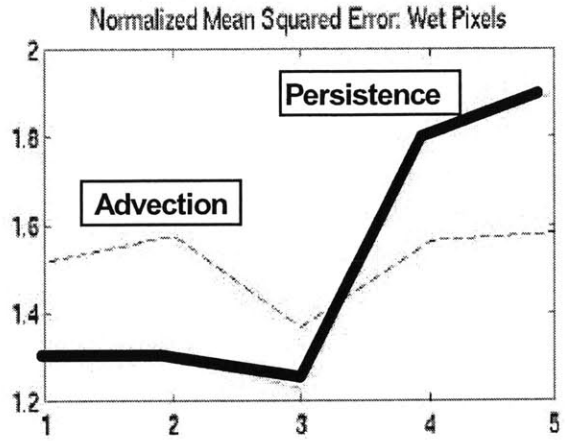


Figure b: A Winter event. NMSE for raining pixels, as a function of the lead time in hours. Provides a measure of the skills, given rain.

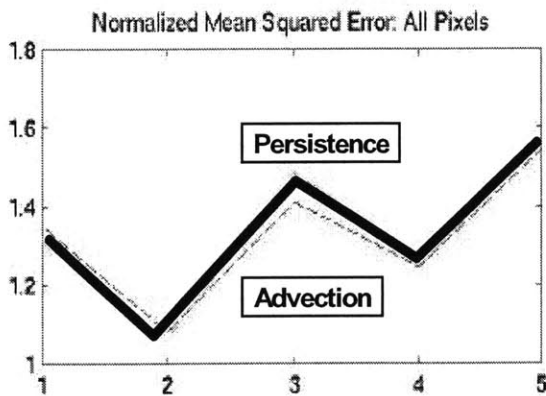


Figure c: RMSE normalized by the variance of the verification data (NMSE) for all radar pixels, as a function of the lead time in hours. A Summer event.

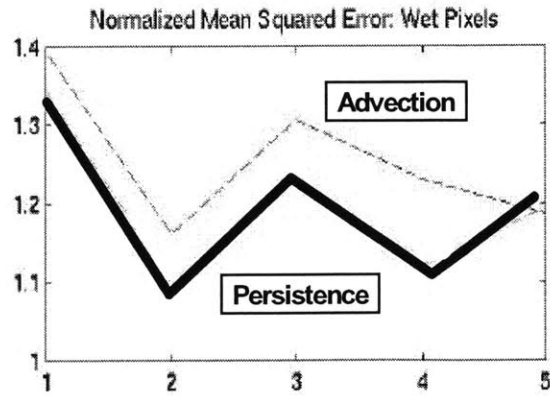


Figure d: NMSE for raining pixels, as a function of the lead time in hours. Provides a measure of the distributed skills, given rain. A Summer event.

4.1.2 Analysis of Information Content in NWP Model Outputs

Radar rainfall was aggregated up both in space (e.g., 4 km to 48 km) and in time (e.g., 1 hour to 6 hours). We used high quality Stage III radar rainfall in the ABRFC in Oklahoma and the 48-km Eta model outputs. Table 4-1 shows how well the variability in aggregate radar rainfall is explained by (a) NWP-QPF alone and (b) by all NWP model outputs, which include QPF and atmospheric variables. The atmospheric variables used were LI (Lifted Index), T (Temperature at 3 Layers), P (Sea Level Pressure), RH (Average Relative Humidity), u and v (Wind Velocity), ω (Vertical Velocity). The results are shown for the Summer and for the Winter, using both linear and nonlinear (ANN) techniques. We analyzed the residuals from ANN based regression, and compared with the analysis of residuals from linear regression, to understand the usefulness of nonlinear strategies. The ANN used were standard Multi-Layer Perceptrons (MLP), the parameters and configuration of which were determined from data analysis (see Bishop, 1996; Jordan & Bishop, 1996). Classification of storm events into “Summer” and “Winter” improved the linear correlation estimates. The information content in all NWP model outputs was found to be higher than that in NWP-QPF alone. The information content in NWP-QPF alone was found to be higher in the Winter than in the Summer. On the whole, these results agreed with NWP-MOS studies (Antolik, 2000). The results of our data analyses also showed that while the information in NWP-QPF was insignificant in the Summer, NWP forecasts of atmospheric variables had significant information content. The degree of non-linearity (as quantified by the performance of linear regression versus ANN algorithms, see Appendix A) was seen to be higher in the Summer. Nonlinear ANN techniques (see Appendix A) did not improve

over linear methods in the Winter, and “optimally trained” ANN were obtained at a very small number of “epochs” (Bishop, 1996), indicating limited nonlinearity in the Winter. In the Summer, the degree of nonlinearity was higher. This is to be expected, as convective and other nonlinear processes are expected to dominate in the Summer. One other interesting observation from these analyses was that when all NWP model outputs (QPF and atmospheric variables) were used, the variability in aggregated radar rainfall could be better explained in the Summer than in the Winter. In contrast during the Winter NWP model outputs beyond QPF did not improve predictability. We also observed that the degree of nonlinearity (as indicated by improvements resulting from ANN) was higher when all NWP model outputs were used, compared to when only NWP-QPF was used (except for Winter).

We performed a sensitivity analysis using NWP outputs. When all seasons were combined, NWP-QPF, Top and Middle Layer Temperature, Lifted Index, and Vertical Velocity, in that order, were most useful. In the Summer, the most influential were Top, Middle, and Bottom Layer Temperature, and Vertical Velocity, in that order. This indicates the dominance of convective processes. The other influential variables, with less influence than the previous ones, were QPF and Relative Humidity, in that order. This suggests that the QPF produced from NWP in the Summer retains minimal information. In the Winter, QPF was most influential. This indicates that at the lower NWP resolutions, the QPF from NWP is fairly useful in the Winter. The other variables with some information were Bottom Layer Temperature, Pressure, and Vertical Velocity, in that order.

**Table 4-1: Aggregate Radar Rainfall and NWP Outputs:
Percent Variability Explained by Regression**

		<i>NWP-QPF</i>	<i>All NWP</i>
<u>ALL SEASONS</u>	<i>LINEAR</i>	18%	40%
	<i>NON-LINEAR</i>	19%	91%
<u>SUMMER</u>	<i>LINEAR</i>	02%	75%
	<i>NON-LINEAR</i>	12%	97%
<u>WINTER</u>	<i>LINEAR</i>	37%	65%
	<i>NON-LINEAR</i>	60%	63%

Table 4-1 shows the percent variability in the aggregated radar measurements that are explained by the linear and the nonlinear regression strategies, using either NWP-QPF alone or all NWP model outputs (QPF and atmospheric variables) as the dependent variables. Results are shown for the Summer, Winter, and for all seasons combined. These are based on 3 events in the Summer, and 3 in the Winter, in S. W. Oklahoma in the Arkansas-red Basin River Forecast Center (ABRFC).

4.2 Analysis of Distributed Precipitation Processes and Strategies

Component processes of distributed QPF could be decomposed into what is called advection, or translation of entire sections of the precipitation map, and what could be termed evolution, or a change in the distributed structure of rainfall intensities as they are advected (see also, Smith and Austin, 2000). Both of these processes need to be initialized using distributed precipitation information. The evolution component could potentially exploit information on the large scale atmospheric state, as indicated by NWP forecasts of atmospheric variables. We developed an ANN based evolution strategy that modeled the change in distributed rainfall intensity as a nonlinear (ANN) function of the NWP atmospheric forecasts and the current intensity of the neighboring pixels.

Distributed QPF could be obtained by space-time disaggregation of low resolution QPF from NWP. The physics of precipitation at lower resolutions is relatively better understood, and NWP models are designed to handle these large scale physical processes. The low resolution QPF obtained from NWP could be disaggregated in space and time using statistical strategies. We developed a simple disaggregation strategy to achieve this. Linear interpolation was used to disaggregate NWP-QPF from 48 km and 6 hr to 48 km and 1 hr. The errors were cast in the form of a time series, and forecast using Auto-Regressive time series models. Disaggregation in space, from 48 km and 1 hr to 4 km and 1 hr, was obtained by preserving the hourly spatial mean and using the distributed structure from advection.

The evolution strategy could also be viewed as a type of disaggregation, where NWP outputs of atmospheric variables could be used to forecast the change in the

precipitation structure. This could be especially useful in situations where QPF from NWP has lower information content compared to other NWP model outputs. The following sections expand on the above concepts of evolution and advection.

4.2.1 Advection and Evolution

As described earlier, previous research indicates that statistical and/or complex methods do not improve advection results over translation based on spatial correlation. To test these hypotheses, we modeled advection using an ANN based method, a linear regression method, and a method based on velocity scale estimation through spatial correlation. The spatial correlation method used radar (1-hour NEXRAD) data alone, and estimated the velocity scales by two-dimensional correlation in space using lagged maps. The regression strategies used NWP (48 km Eta) model outputs in addition to radar data. The dependent variable for regression was the rainfall evolution (change of rainfall rate from previous time step) at each pixel, the independent variables were the radar precipitation (“state”) at the pixel and its neighbors, and the NWP forecasts of atmospheric variables. The test site where the storm events occurred was Topeka, Kansas. The measure of skill was the Root Mean Squared Errors (RMSE) of the raining pixels for the verification data. Table 4-2 shows an example of these analyses.

Both the ANN based method and the method based on estimation of velocity scales from spatial correlation improved over linear regression by nearly 50% on the average. In fact, the linear regression strategy performed worse than simple persistence (i.e., where the rainfall rates at future times are assumed to remain identical to the previously observed time steps). This indicates that the rainfall evolution process at distributed scales is not linear, and that the linear regression method fails to model

the advection and evolution components. The ANN based model performed worse than the one based on spatial correlation by about 3-6% on the average. This shows that the ANN is barely able to model the advection component. Further, the use of additional information from NWP and the nonlinear modeling does not seem to add value to extrapolation, at least in an Eulerian frame of reference.

Based on the above results and on those of previous researchers, we conclude that ANN and other statistical methods might not add any extra value over advection (which uses velocity scale estimates calculated from spatial correlation) for the extrapolation component. This holds when the ANN and linear regression techniques are conditioned on an Eulerian frame of reference, in which case advection often becomes the primary component to be modeled. We hypothesized that if these regression methods were conditioned on a Lagrangian frame of reference, i.e., on the results of advection, these could model localized processes like convection. This hypothesis implicitly separates the two processes, advection of the larger scale rainfall map, and evolution at higher resolutions that occur on top of advection. The separability of the advection and the evolution components has been corroborated by previous research in this field. For example, Smith and Austin (2000) mention that “It is generally acknowledged that the primary reason why the advection forecast fails to achieve universal success does not originate from errors in the forecast of advection velocity.” They go on to add that the reasons for the failure are the extreme variability of precipitation in space and time, and the fact that “precipitation structures of interest undergo internal development as they are advected.”

When we modeled the evolution of rainfall conditioned on the advected radar map (i.e., in a Lagrangian frame of reference) for Summer storms, the results were interesting. The linear regression strategy marginally improved (about 2-5% on the average) over advection alone, while the ANN based method improved by as much as 10-25%. This indicates that modeling localized evolution is useful only if it is conditioned on advection. This also indicates that the governing dynamics for this high resolution evolution is nonlinear, and that NWP forecasts for atmospheric variables have information relevant for modeling this component.

Previous researchers have indicated that convective processes are nonlinear (Rogers and Yau, 1989), and that these are dictated by the large scale atmospheric state (Peppler and Lamb, 1989), which in turn are indicated by atmospheric instability indices and other variables (Zawadzki et al., 1981). Our results support these suggestions, and further demonstrate the usefulness of both atmospheric forecasts from NWP and nonlinear strategies like ANN for distributed QPF.

Table 4-2: Typical Performance (1 hour Lead time) in the Summer of Preliminary QPF Algorithms for Advection and Evolution

QPF Methods <u>(Advection and Evolution Strategies)</u>	Performance <u>(RMSE: Raining Cells)</u>
Linear Regression based Eulerian Evolution	4.4771
Persistence	3.3932
ANN based Eulerian Evolution	2.8131
Spatial Correlation based Advection	2.7530
(Lagrangian) Linear Evolution Conditioned on Advection	2.7180
(Lagrangian) ANN Evolution Conditioned on Advection	2.5261

Table 4-2 shows the performance of the various preliminary QPF strategies in terms of the RMSE for pixels with non-zero rainfall, for 1 hour lead time. The ANN and the linear regression based evolution methods used NWP model outputs in addition to radar data. Persistence and Advection used radar data alone. Performance of regression (linear or ANN) strategies in an Eulerian frame of reference results in poor performance. ANN does better than linear regression, and manages to beat persistence. Spatial correlation based Advection beats both these methods. However, when the regression is conditioned upon Advection, both linear and ANN methods improve performance. The improvement from ANN is more significant. No advantage was apparent if linear components were isolated prior to using ANN. The RMSE from a (Lagrangian) model, that performed linear regression before ANN was 2.6900, i.e., the performance was better than linear regression, but worse than ANN alone.

4.2.2 ANN based Localized Evolution versus NWP-QPF Disaggregation

We have described earlier that localized evolution conditioned on advection could potentially be handled by data dictated tools like ANN (see Appendix A), as well as by parameterized physically based models. Preliminary comparisons of the two strategies indicated that the data dictated model is more generally applicable. Previous researchers (Perica and Foufoula-Georgiou, 1996; Golding, 2000) have obtained distributed precipitation forecasts by disaggregating NWP-QPF in space and time. In this section, we briefly compare and contrast a “simple disaggregation” approach with a “localized evolution” strategy.

As a rough analogy, the process of distributing the grid average NWP-QPF in space could be compared to preserving the forecast “mean” (in this case, the spatially aggregate value, as well as a “mean” distributed structure), while the localized evolution model could correspond to a forecast of the perturbations around the mean. We devised an algorithm to disaggregate NWP-QPF in time and space, by using linear interpolation in time and by scaling the results of advection in space. This was compared to a data dictated (ANN based) localized evolution model. This localized evolution strategy was conditioned on the results of the disaggregation strategy. Thus if evolution were to result in improvement, the overall skills would improve. A lower value of skills would indicate that the evolution component caused the performance to be worse than before.

The improvement in average distributed skills, as indicated by the RMSE, was significantly higher for 1-3 hour lead times when the evolution component was used. However, the aggregate statistics (e.g., the space-time aggregate forecast) was better

preserved by the simple disaggregation algorithm alone. Also, the distributed error statistics (RMSE) indicated that the localized evolution model performed significantly worse after around 3-4 hour lead times. Together this suggested that the ANN based evolution improves distributed forecasts at short leads, but results in loss of information at longer leads and in terms of aggregate statistics.

After around 3 hours, the skill in the ANN based localized evolution model decayed rather sharply, and the low skills at 4-6 hour lead times caused the aggregate errors to go up. This phenomena was especially true in the Summer, while in the Winter the decay in skills was more gradual. This seems to suggest that the functional form of the localized evolution component modeled by the ANN remains valid for 1-3 hour leads, especially in the Summer. Preliminary experiments with a parameterized, physics based model also performed well until at most 2-3 hour lead times in the Summer (not shown). These results suggest the possibility of a fundamental time scale of about 3 hour lead times, over which localized evolution models (whether these are parameterized physics or ANN based) remain valid. This could correspond to an average life time for summer time convective activity, at least in the Oklahoma and Kansas regions, which is modeled and forecast by both the ANN and the physics based method. After around 3 hour lead times, it appears that the behavior that governs precipitation generation changes, and complex data dictated tools like ANN begin to lose information and perform worse than simpler techniques. An example result is presented in Table 4-3, which shows the average performance with 3 storms (2 Summer, 1 Winter) in Topeka, KS.

Table 4-3: Average Performance (1/NMSE) of ANN based Localized Evolution and Simple Disaggregation for 3 Storms (2 Summer, 1 Winter) in Topeka, KS

QPF Strategies	Lead 1	Lead 2	Lead 3	Lead 4	Lead 5
ANN based Evolution	1.05	0.98	0.92	0.72	0.54
Simple Disaggregation	0.86	0.85	0.68	0.70	0.69
Persistence	0.82	0.83	0.70	0.68	0.61

Table 4-3 shows the performance of the various preliminary QPF strategies in terms of 1/NMSE (or, the inverse of the Normalized Root Mean Squared Errors) for 1 to 5 hour lead times. The ANN based Localized Evolution method used the results of Simple Disaggregation. Improvement over simple disaggregation implies the ANN based evolution adds value to the forecasting process, while lower 1/NMSE corresponds to a loss of information. Persistence (where the forecasts are assumed to be identical to the current time step) is shown as a baseline for comparison.

4.3 Parameterized, Physically Based Models

A controversial area is the use of parameterized, physically based models for distributed QPF. Some researchers have suggested that this physics of distributed QPF is not too well understood (Grecu and Krajewski, 2000), and that existing models are not generally applicable and represent crude parameterizations (Smith and Austin, 2000). However, others have used these models for research purposes (Georgakakos, 2000; Lee and Georgakakos, 1990, 1996; Georgakakos and Bras, 1984 a,b), and have suggested estimating the parameters thereof from NWP model outputs (Nakakita et al., 1996; Sugimoto et al., 2001). We conducted a study to determine how well such a model would perform on real data, and also how the performance and ease of application would compare to other data dictated or hybrid models.

We used the model proposed by Lee and Georgakakos (1990), keeping in mind that our grids are square (i.e., $\delta x = \delta y$), and with the assumption that the gradient of the cloud bottom in space could be considered negligible. The parameters of this “physically based” model were estimated from the available NWP model outputs if possible, or else estimated from data (previous time steps).

We discovered that these physically based and parameterized models do show some promise, and in certain situations (e.g., in terms of distributed skills in the Summer, for pixels with non-zero rainfall) perform as well or at times even marginally better than techniques based on extrapolation, NWP-QPF, and statistical methods. However we also found that these models, in their current state of evolution, are not consistent and could perform worse than even simple persistence in certain other situations (e.g., in terms of aggregate skills, especially in the Winter). This is not

surprising for models explicitly developed to handle convective situations and localized rainfall evolution. We also found that the parameter estimation problem could be rather involved, primarily because not all the parameters can be easily related to NWP model outputs. This seems to agree with some of the comments in Smith and Austin (2000) and others like Collier (2001).

Based on these preliminary conclusions (see also the similar conclusions in Sugimoto et al., 2001), we did not use these techniques further in the model building exercise described in later sections. However, we feel that that this research area needs to be explored further before firm conclusions can be reached.

4.4 Consolidation of Insights for Improved Distributed QPF

Existing models for distributed QPF include radar extrapolation alone (e.g., persistence and advection), NWP-QPF alone (raw or after statistical error correction, as in NWP-MOS described by Antolik, 2000), hybrid models that combine radar extrapolation with NWP-QPF (e.g., the NIMROD model described in Golding, 2000), pure data dictated models (e.g., based on statistical or ANN based techniques alone, like French et al., 1992), and parameterized physically based models (e.g., Lee and Georgakakos, 1990 and Georgakakos and Bras, 1984 a,b). This research seems to suggest that a decomposition of the QPF problem into component processes, followed by a modeling strategy that optimally combines precipitation physics and data dictated tools for each process, could improve distributed QPF over existing techniques.

Extrapolation methods like persistence and advection using radar information alone cannot account for large scale physics, cannot model high resolution localized evolution that occurs in rainfall “as they are advected” (Smith and Austin, 2000), and cannot handle residual structures. However extrapolation is a useful component in the QPF process, and is best modeled through advection based on estimates of velocity scales. These conclusions are supported from the results of our studies (see Table 4-2), and from the results of previous researchers like Golding (2000).

Hybrid models like NIMROD combine radar advection with NWP-QPF. However, the NIMROD system (or similar hybrid methods) uses information from NWP-QPF alone, and not other NWP model outputs that indicate the state of the atmosphere. Results from our data analysis (see Table 4-1) and previous research (e.g., Nakakita et al., 1996) indicate that there is additional information content in NWP model outputs

other than QPF that could be utilized for distributed rainfall forecasts. Further, NIMROD does not model localized evolution at higher resolutions, and does not handle residual structures. These could be a significant drawbacks, as NWP models cannot account for localized, high resolution processes too well (see Mesinger, 1996 and Antolik, 2000; and also note the lack of significant improvement obtained over persistence by disaggregating NWP-QPF, as shown in Table 4-3). Finally, NIMROD uses grid average NWP-QPF directly. NWP-QPF does provide information on the spatially aggregate and cumulative precipitation, and does consider large scale physics. However, a statistical “post-processing” of the NWP-QPF and other model outputs using high quality measurements often add value (see Kuligowski and Barros, 1998a; Krishnamurti et al., 1999).

Pure data dictated tools are not optimal in handling all the component processes of distributed precipitation. For example, previous researchers (Grecu and Krajewski, 2000; Browning and Collier, 1989 and the references therein) and QPF studies carried out as part of this research (see Table 4-2) shows that ANN and complex statistical tools are not optimal to handle advection and Eulerian evolution.

Our preliminary experiments that compare the performance of a parameterized, physically based algorithm with a hybrid modeling strategy appear to indicate that the latter could be more consistent in terms of forecast skills. Further, some of the parameters of the “physical” model might not easily relate to NWP model outputs in all situations, resulting in numerically expensive computational methods.

Our studies appear to indicate that the resolutions of NWP models and radar measurements, as well as the relevant precipitation physics, could lead to an

identification and separation of the dominant component processes of precipitation. For example, our studies have demonstrated that localized evolution improves distributed QPF only if conditioned on the results of advection. We have also shown that disaggregation of NWP-QPF in space could use the results of advection. Finally, we have shown that a simple disaggregation of NWP-QPF, and an ANN based localized evolution have complementary skills. This suggests that an optimal combination of the two strategies, that retains the information in both, could be an improvement over either.

We have attempted to use and build upon these insights while developing a new hybrid modeling strategy for distributed QPF, which is described in detail in the next section.

5.0 Decomposition of Precipitation Processes and a Hybrid Model

5.1 Decomposition Rationale

The distributed QPF problem has been decomposed into four component processes. The motivation for the decomposition was to optimize the use of available information from radar and NWP, and to make the best use of the available process physics, hydrologic insights, and data dictated tools.

5.2 Overview of the Hybrid Modeling Strategy

The component processes of precipitation are handled through a hybrid modeling strategy. The processes are modeled in succession, and each is superimposed on the results of the previous components. A schematic flowchart for the proposed hybrid modeling strategy is shown in Figure 5-1. We first present an overview of the hybrid method, and then discuss the details of each component.

The first component is radar extrapolation, which is handled through advection of the hourly radar maps. Velocity scales at each time in the past are estimated using two-dimensional spatial correlation by comparing with the map at the previous time step. An exponential smoothing based time series formulation is used for estimating the velocity scales in the future. This formulation applies relative weights to the past estimates, with the weights decaying as one goes further into the past.

The second component is large scale physics, as represented by NWP-QPF. NWP-QPF is grid averaged and cumulative (48 km and 6 hour for NWP-Eta), this is disaggregated in time and space. Disaggregation in time is performed by linear interpolation for QPF and other NWP model outputs. Temporal disaggregation is

followed by an AR (AutoRegressive) error correction model for the grid average QPF, by comparing with aggregated radar rainfall. The temporally disaggregated, and error corrected, grid averaged QPF is “superimposed” on the results of advection. The distributed structure obtained from advection is retained, but the values of the pixels with non-zero rainfall are scaled to reflect the large scale mean from NWP-QPF.

Localized evolution is modeled at radar pixel resolutions as a function of NWP atmospheric forecasts and the distributed precipitation “state”, which in turn is the result of radar extrapolation followed by NWP-QPF disaggregation. As a rough analogy, we could compare the disaggregation of the NWP-QPF as an attempt to retain the grid average precipitation forecast based on the distributed structure from advection, and the localized evolution as an attempt to model the perturbations from the mean. The evolution attempts to model localized precipitation processes that cannot be handled well by the lower resolution NWP model physics, for example convective processes. The assumption is that this evolution would be a function of the distributed state (rainfall) of the pixel and its neighbors, as well as the large scale NWP atmospheric forecasts which are assumed to be spatially uniform over the NWP grid. The function to be approximated could be nonlinear, and the physics is not too well understood. Data dictated Artificial Neural Networks (ANN) based tools are used for this function approximation.

The final component is the residual structures (errors) at both distributed and aggregate scales. Distributed errors are handled at higher (radar pixel) resolutions by an ANN based strategy that combines the results of the component processes described earlier (simple disaggregation and localized evolution). The errors at

aggregate scales (lower resolutions) are handled by combining the spatial means of the rainfall intensities from the individual component processes.

**Figure 5-1:
Flowchart for the Proposed Hybrid Modeling Strategy**

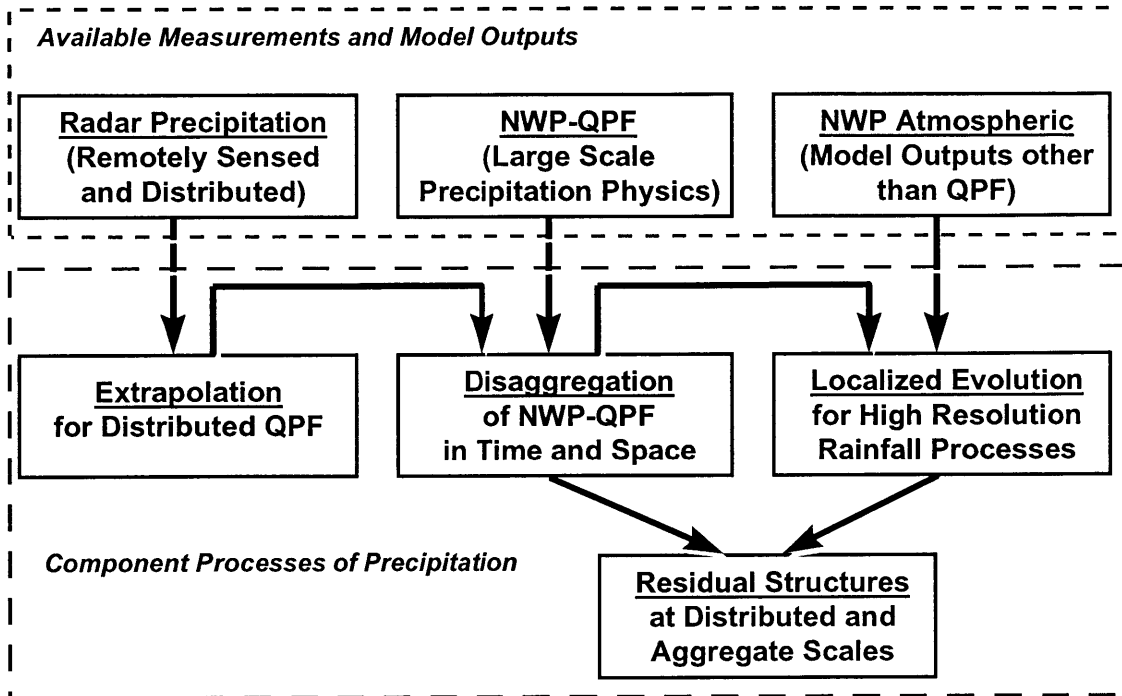


Figure 5-1 shows a flow diagram for the proposed QPF strategy. The available information are radar rainfall at 4km and 1 hour, and NWP model outputs for precipitation and atmospheric variables at 48 km and 6 hours. The QPF problem has been decomposed into four components: Radar Extrapolation, Large Scale Physics, Localized Evolution, and Residual Structures.

5.3 Details of the Radar Extrapolation Component

A radar map at a given time t is compared to the lagged map at time $(t-1)$. Two dimensional spatial correlation is calculated, by “moving” the current map over the lagged map. The maximum correlation obtained in this fashion provides an estimate for the velocity scale at any time t (see Figure 5-2). The metric (for 2-D covariance) to be maximized takes the form:

$$S = \sum \{ [X(i-i_g, j-j_g, t) - \langle X(t) \rangle] \bullet [X(i,j,t-1) - \langle X(t-1) \rangle] \} \quad 5.1$$

X represents the radar state (rainfall) at pixel (i,j) in time t . The two-dimensional correlation in space is calculated in terms of the metric S . The values of (i_g, j_g) that maximize the metric S are denoted by (i_{gM}, j_{gM}) . These provide estimates of the velocity scales, for time t .

Velocity scales could similarly be estimated for all times t in the past, in the context of the precipitation event. The time series of the velocity scales is used for forecasting this quantity into the future, using a single exponential smoothing formulation:

$$i_{gMF} = \alpha i_{gM}(t) + \alpha (1 - \alpha) i_{gM}(t-1) + \dots + \alpha (1 - \alpha)^n i_{gM}(t-n) \quad 5.2$$

$$j_{gMF} = \alpha j_{gM}(t) + \alpha (1 - \alpha) j_{gM}(t-1) + \dots + \alpha (1 - \alpha)^n j_{gM}(t-n) \quad 5.3$$

The parameter α is the exponential smoothing parameter estimated from data, this determines the decay applied to the weights for estimates further in the past. The parameter is less than unity (for our QPF studies, the range was 0.3-0.7), thus the most recent velocity estimate gets the highest relative weight. Note that in the single exponential smoothing strategy (Mills, 1990), the forecasts (in this case, for the velocity scale) remains identical for all lead times.

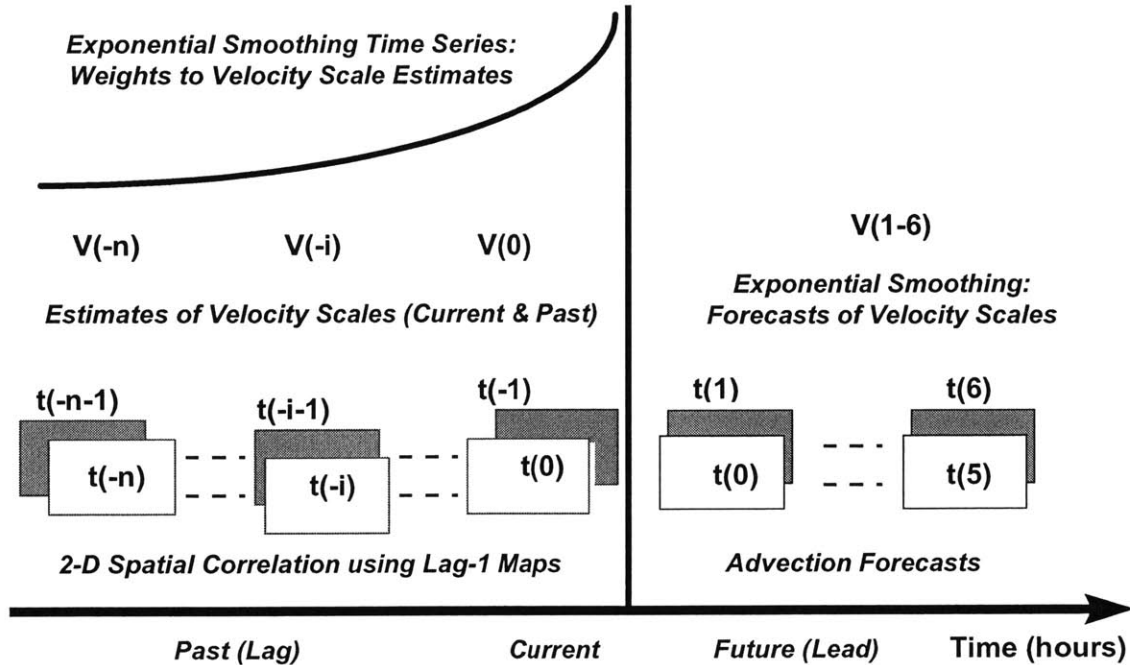
The radar forecasts are initiated with the current map, which is translated in the future for each radar pixel using the forecasts for the velocity scales:

$$X^g(i, j, t + t_L) = X(i - i_{gMF}, j - j_{gMF}, t) \text{ for all } (i, j) \quad 5.4$$

The extrapolation strategy does not use NWP wind forecasts. Our studies indicated that the wind available from FOUS at 6 hour resolutions, have no significant correlation with the calculate advection velocities, and do not improve forecasts for the velocity scales. Advection strategies that rely on rain cell identification and tracking are usually not relevant for 1-hour rainfall maps, owing to the hourly decorrelation times for distributed rain structure (Browning and Collier, 1989; Mecklenburg et al., 2000).

In a nutshell, the extrapolation strategy calculates the velocity scales at each previous time step using 2-D correlation in space, and uses these velocity estimates in a single exponential smoothing formulation to forecast the velocity scale in the future. The forecast of the velocity scale is used to advect the current precipitation map for all lead times.

Figure 5-2: Radar Extrapolation Strategy - Advection



The first component of the proposed QPF strategy is Radar Extrapolation, which is based on advection. The velocity scales for advection are obtained using single exponential smoothing, a time series forecasting technique. Estimates for velocity scales in previous time steps, for the precipitation event, are performed using 2-D correlation in space. These are used to calculate the advection velocity.

5.4 Details of the Large Scale Physics Component

As shown in Figure 5-3a, we first disaggregate NWP-QPF in time from 6 hour cumulative to hourly resolution through linear interpolation. We apply an error correction (Figure 5-3b) based on an auto-regressive (AR) time series model for the errors. The model selection and parameter estimation were performed using MATLAB at the forecast lead times. The errors are obtained by comparing 1-hour QPF (after temporal disaggregation) with spatially aggregated hourly radar rainfall. The third step (Figure 5-4) is to achieve a spatial disaggregation of the hourly QPF from 48 km (Eta) to 4 km (radar) resolution. NWP-QPF represent the result of large scale physics at low resolutions in space and time, radar extrapolation does not take physics into account. However the grid average NWP-QPF does not have the distributed structure of radar extrapolation. We combine these two processes by scaling the result of radar advection at all forecast lead times of relevance using the spatially average estimate of rainfall from 1-hour NWP-QPF (i.e., QPF from NWP after temporal disaggregation and error correction). This scaling is applied to all radar pixels in a given NWP grid.

Temporal Disaggregation

The first step is to disaggregate the NWP-QPF from 48 km and 6 hour resolution to 48 km and 1 hour resolution.

$$r(t+k*\Delta t) = [\{ (T+\Delta T)-(t+k*\Delta t) \} \{R(T)/6\} + [\{ (t+k*\Delta t) - T \} \{R(T+\Delta T)/6\}] / \Delta T \quad 5.5$$

In Equation 5.5, r represents the rainfall at 1-hour resolution, t the current time, Δt the radar time step (1-hour), and k an integer. The left hand side of the equation represent the aggregate forecasts for time steps at lead- k . On the right hand side, T is

the current (or last available) NWP time, $(T+\Delta T)$ is the next available NWP time, and R the average (6 hour) NWP-QPF. The equation as shown assumes that the forecast lead times lie within the intervals for which NWP outputs are available.

The hourly QPF obtained from Equation 5.5 are compared with spatially aggregate radar rainfall. The errors are modeled and forecast as AR (autoregressive) time series. Error correction is used only if the forecasts are improved at analysis time, which is checked by forecasting the errors at verification time steps.

The time series of the aggregate errors is $Z(\tau) = [\langle X(t-i\Delta t) \rangle - r(t-i\Delta t)]$, where i is the index for times in the past, $\tau=(t-i\Delta t)$, and $\langle X \rangle$ is the spatial mean of the observed radar rainfall. The AR model for error correction (at lead k) is $Z_{t+k} \cong \alpha_1 Z_{t+k-1} + \alpha_2 Z_{t+k-2} + \dots + \alpha_p Z_{t+k-p}$. Note that $\Delta t = 1$ hour, and forecasts generated at each lead time could be used in the AR formulation for forecasts at successive lead times. The forecasts for the spatial average hourly forecast following error correction (F) becomes:

$$F(t+k) = r(t+k) + \underline{Z}_{t+k} = r(t+k) + \sum \alpha_i Z_{t+k-i} \quad 5.6$$

Spatial Disaggregation

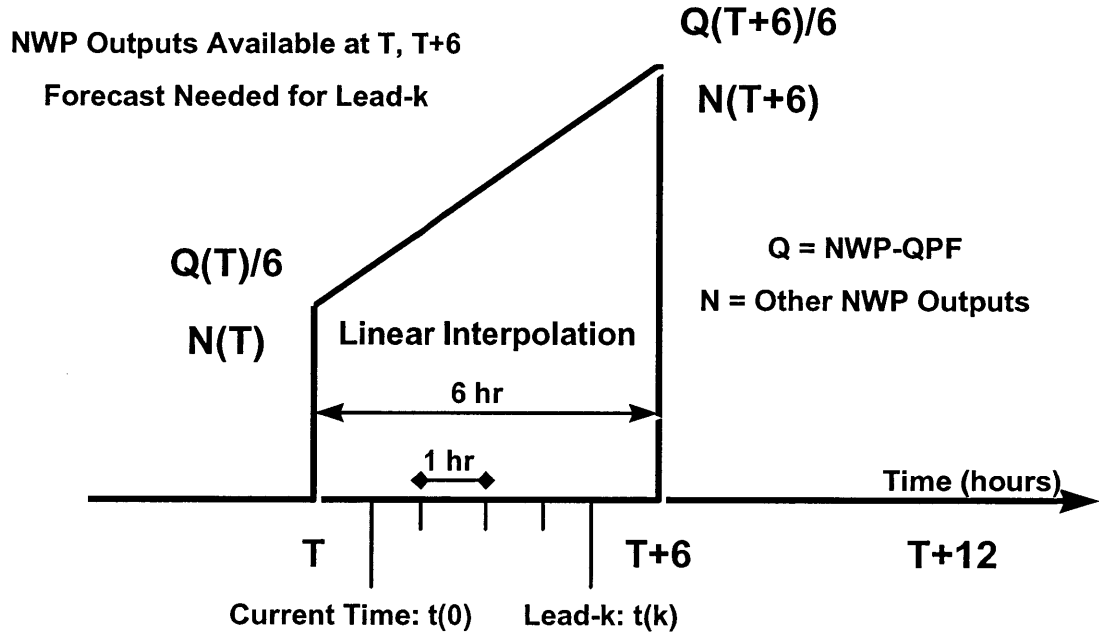
The temporally aggregated QPF obtained from Equation 5.6 is used to scale the distributed map obtained from radar advection. The pixels with zero (or below a certain minimum threshold) rainfall are left unaltered. Pixels with non-zero rainfall are changed to reflect the grid average QPF. NWP-QPF is best interpreted as a grid average quantity. To preserve the nature of this information, the non-zero pixels are modified equally using the average deficit or excess grid average rainfall indicated by Equation 5.6.

If $X^g(i,j,t+k*\Delta t) > 0$ (or, a small threshold ε),

$$X^D(i,j,t+k*\Delta t) = X^g(i,j,t+k*\Delta t) + N_{\text{pixel}}\{F(t+k*\Delta t) - \langle X^g(t+k*\Delta t) \rangle\} / N_{\text{wet-pixel}} \quad 5.7$$

Quantities within $\langle \rangle$ denote spatial average, N_{pixel} the total number of radar pixels, $N_{\text{wet-pixel}}$ the total number of radar pixels with non-zero rainfall, and X^D is the distributed rainfall forecast from this disaggregation strategy.

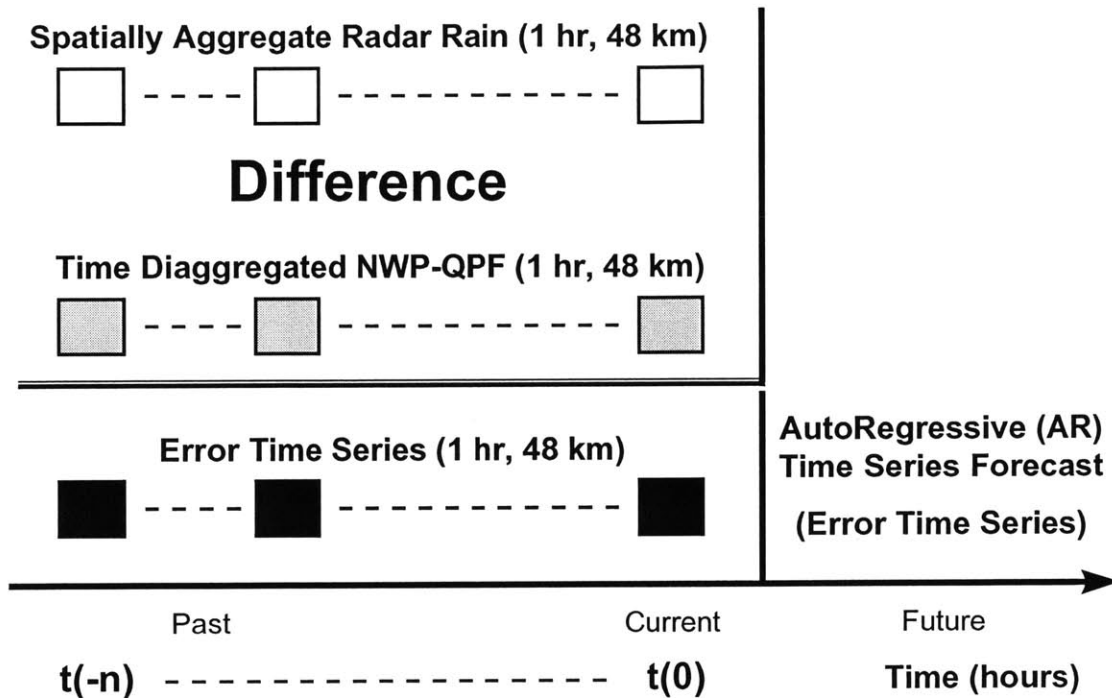
**Figure 5-3a: Large Scale Physics Strategy -
1. Temporal Disaggregation of NWP-QPF and Other Outputs**



The second component of the proposed QPF strategy is Large Scale Physics. The physics itself is assumed to be incorporated in the NWP-QPF, which are at 6 km and 48 km resolutions. These are disaggregated in time and space.

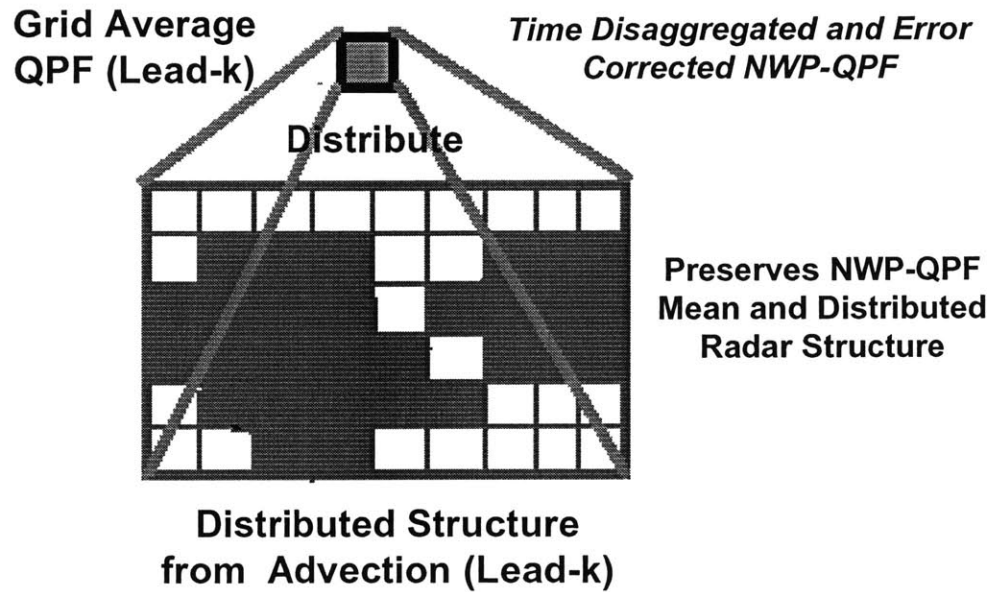
The first step in this process is diaggregation in time, from 6 hr, 48 km to 1 hr, 48 km, using linear interpolation.

**Figure 5-3b: Large Scale Physics Strategy -
2. Error Correction for Time Disaggregated NWP-QPF**



The second step in the disaggregation strategy is error correction of the hourly time series of the spatial mean rain forecast. All available previous time steps for the entire precipitation could be used for the correction. Errors are obtained by comparing with observed values, and forecast using the best AR model. These are used to post-process the NWP-QPF, after disaggregation in time.

**Figure 5-4: Large Scale Physics Strategy -
3. Spatial Disaggregation of NWP-QPF**



The third step is the spatial disaggregation of the time disaggregated and error corrected NWP-QPF . Advection forecasts are used as the basis for the spatial disaggregation. The rainfall map is scaled to retain the distributed structure from advection and the spatial mean from NWP-QPF.

5.5 Details of the Localized Evolution Component

The evolution of the distributed rainfall “state”, after advection and scaling, is assumed to be modeled by a functional form (see Figure 5-5).

$$d\underline{X}/dt = \Phi^{(E)}(\underline{X}, \underline{X}_N, \underline{\Theta}^{(E)}, \underline{\Omega}) + \eta \quad 5.8$$

In Equation 5.8, \underline{X} represents the vector of the distributed state in a Lagrangian frame of reference conditioned on and scaled by NWP-QPF, and $d\underline{X}/dt$ denotes the evolution of this state. This evolution is assumed to be a function of the state \underline{X} at the pixel under consideration, and the state at neighboring pixels \underline{X}_N . We assume that the nature of this evolution (the functional form) remains the same over the area covered by each NWP grid for the given precipitation event, this uniformity is represented by the large scale atmospheric state E . The NWP model outputs for the atmospheric variables are assumed to serve as indicators of this large scale atmospheric state, these outputs are denoted by $\underline{\Omega}$. Finally, we assume a set of parameter vectors $\underline{\Theta}^{(E)}$, which could themselves be a function of the large scale atmospheric state. The functional form for the state evolution is denoted by $\Phi^{(E)}$. The term η denotes model errors and “noise”, which could result (for example) from the processes that occur at resolutions greater than radar observations.

To model the evolution hypothesized in Equation 5.8, we use data dictated ANN techniques that do not need to make *a priori* assumptions about the functional form or the parameters relating the input vectors to the outputs. The ANN model (see Figure 5-6) could be expressed as:

$${}^E X\text{-}^D X = f_{ANN}({}^D X, {}^D X_N; LI, P, T, RH, u, v, \omega) + \eta \quad 5.9$$

In Equation 5.9, E^X denotes the effects of evolution, while D^X denotes the pixel state in a Lagrangian frame of reference conditioned on large scale NWP-QPF. The term D^X_N denotes the state of the neighboring pixels of the scaled advection map.

NWP model outputs other than QPF were also disaggregated in time, in the same way as illustrated for precipitation in Equation 5.5. The $[R(T)/6]$ term in Equation 5.5 was replaced with $[W(T)]$, where W denotes the NWP output. For these outputs, we did not need to divide by 6. Spatial disaggregation was not attempted for these outputs as (a) there were no high resolution forecasts or measurements to serve as bases for disaggregation, and (b) the atmospheric outputs are more uniform in space than QPF (Antolik, 2000; Mesinger, 1996).

The NWP model outputs used are the relevant ones available in the FOUS (Forecast Outputs for the United States) format. These are Lifted Index (LI), P (Sea Level Pressure), T (Average Air Temperature), RH (Average Relative Humidity), u , v (Wind Vector) and ω or updraft velocity. We know from process physics (Rogers and Yau, 1989; Lee and Georgakakos, 1990) that evolution is sensitive to these quantities, we have confirmed this from data analysis at aggregate scales (as described earlier). We include all of these as ANN inputs as we cannot reject *a priori* the hypothesis that these do not have information pertinent for distributed rainfall. The strength and nature of this information was assumed to be optimally derived from the available data by the ANN.

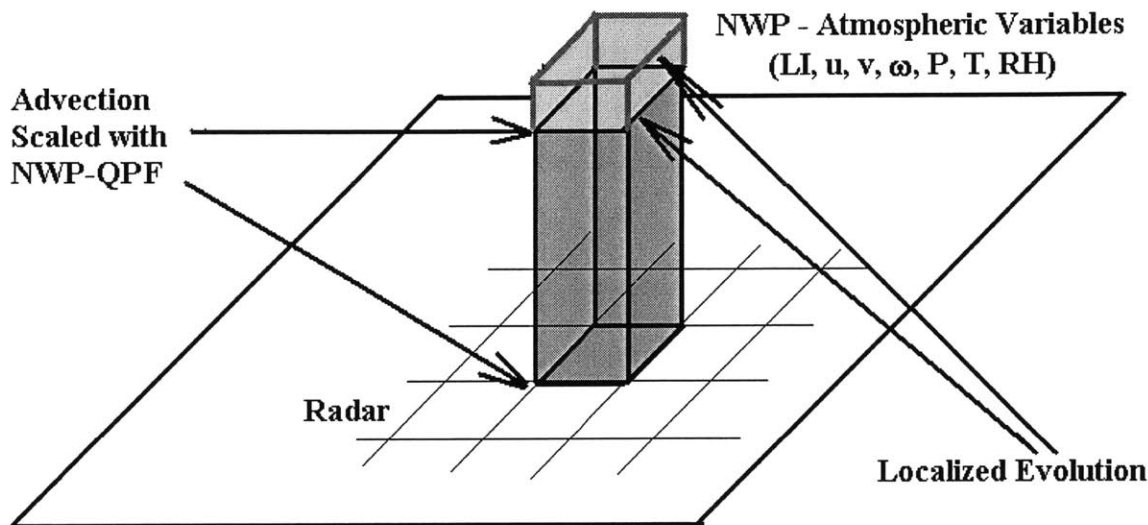
The ANN used was a modified form of the NARMA (Nonlinear AutoRegressive Integrated Moving Average) model (Connor et al., 1994). NARMA models handle the autoregressive and moving average components of a traditional time series model

(Mills, 1990) using nonlinear ANN functions instead of linear formulations. We extended the NARMA implementation such that the NAR and NMA components were separated out and modeled using individual MLP (MultiLayer Perceptron; see Bishop, 1996). The “Bayesian ANN” methods of MacKay (1994) were used to generate error bars and most probable forecasts. This method first trains an ANN, then calculates the error statistics on the cross-validation data and uses these to generate an ensemble of identically distributed random error simulations. Realizations from the random simulations are then summed with the observed output variables to generate “pseudo-outputs”. An ensemble of ANN models are obtained by re-training the originally trained ANN with each of the pseudo-outputs. This ensemble of models provide ensemble forecasts at each verification time step, which could be used to calculate the expected error statistics. The training iterations for the ensemble are not too high, as these are initialized with previously calibrated ANN. However, to reduce forecast generation time, only the most likely forecasts (obtained from the original outputs) and upper and lower error bounds were used for forecast generation at successive time steps. The number of training time steps, the neighboring window size, and the optimal MLP architecture were chosen from data analysis. Note that the functional form to be modeled by the ANN was assumed to remain invariant within the context of the “localized evolution event” for the training and forecast generation times in the domain of interest. The pixels in space at each training time step were divided randomly into “training” and “cross-validation” data, roughly in the ratio of 2:1.

To summarize, the localized evolution model calculates the rainfall intensities at individual pixels through an ANN based nonlinear function approximation. This function

approximation is conditioned on the advected radar map, scaled with spatially averaged QPF obtained from NWP. The inputs to this ANN model are the NWP forecasts of atmospheric variables, and the scaled advected rainfall intensities.

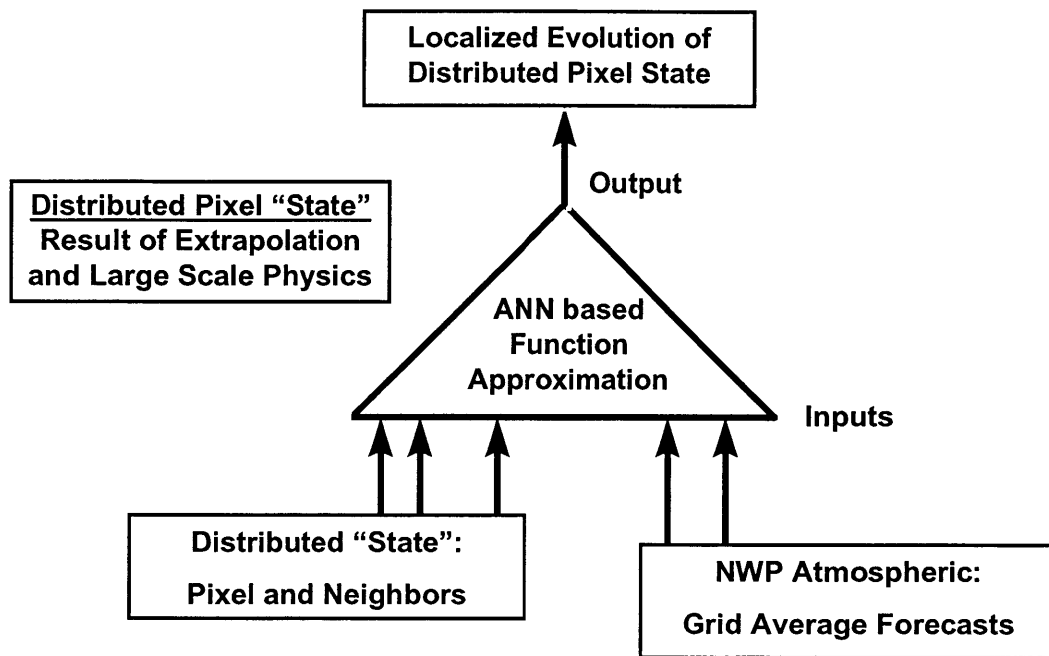
**Figure 5-5: Localized Evolution Strategy -
1. Conditioned on Extrapolation and QPF Disaggregation**



**Accounts for “Perturbations around Mean” NWP-QPF
Disaggregation, caused by Localized High Resolution Processes**

The third component of the proposed QPF strategy is Localized Evolution. The scaled, advected map is used as a starting point. At each radar pixel, there could be localized perturbations, which could be functions of NWP forecasts of atmospheric variables, and the existing forecasts at neighboring pixels. This could be especially true when convective situations dominate, like in the Summer.

**Figure 5-6: Localized Evolution Strategy -
2. Function Approximation Strategy**



The Localized Evolution component uses an ANN based strategy to approximate the functional form that dictates the change in rainfall intensity at a pixel, as a function of the intensity at neighboring pixels, and NWP forecasts of atmospheric variables.

5.6 Details of the Residual Structures Component

We obtain distributed precipitation forecasts from the previously discussed components. The first could be called a “Disaggregation Model”, where extrapolation results are used for disaggregating NWP-QPF. The other could be referred to as an “Evolution Model”, this is ANN based and handles localized evolution, and is assumed to be conditioned on the results of the Disaggregation Model. Both these approaches consider advection as a baseline. The first strategy attempts to preserve the distributed radar structure, as well as the aggregate information from NWP. The second approach tries to capture the localized evolution at smaller scales in a Lagrangian frame of reference. It is conceivable that there are residual structures in the distributed and aggregate errors from both these strategies, which could be modeled as a function of the individual forecast errors.

Results of the individual modeling strategies indicate somewhat complementary skills, as shown later. The strategy that combines extrapolation and NWP-QPF does relatively better in the Winter, in terms of aggregate statistics in space and time, and for longer lead times (~4-6 hours), as shown later. The approach that models high resolution evolution using NWP atmospheric variables and ANN, exhibits better skills in the Summer, at distributed scales, and for shorter lead times (~1-3 hours). We hypothesize that an approach that could capture the distributed errors and preserve the aggregate information by combining the results of both these individual strategies could be an improvement over either. The strategy for modeling residual structures is depicted in Figure 5-7.

To achieve a distributed (i.e., on a pixel by pixel basis) combination, we use a data dictated ANN approach that regresses the observed rainfall on the results of what we could call the “Disaggregation” (transforming NWP-QPF to higher resolutions) and the “Evolution” (function approximation using NWP atmospheric forecasts and scaled advected rainfall) models. The use of ANN precludes the necessity to assign a functional form *a priori* to this relation.

$${}^{DE}X(i, j, t) - {}^AX(i, j, t) = g_{ANN}\{[{}^DX(i, j, t) - {}^AX(i, j, t)], [{}^EX(i, j, t) - {}^AX(i, j, t)]\} + \eta \quad 5.10$$

In Equation 5.10, AX is the result of advection alone, ${}^{DE}X$ is the distributed map obtained by combining the results of disaggregation and evolution on a pixel by pixel basis, DX the scaled advection, EX the localized evolution, and $X(i, j, t)$ the observed rainfall at a radar pixel (i, j) at time t . The result of pure Advection is removed prior to the ANN based function approximation, and added back to the outputs later. The details of the ANN formulation were similar to the one used for localized evolution.

To retain the optimal spatial aggregate information, we combine the spatial mean from the Disaggregation and the Evolution models by assigning relative weights.

$$\langle {}^OX \rangle = \{ (1/\sigma_D) \cdot \langle {}^DX \rangle + (1/\sigma_{DE}) \cdot \langle {}^{DE}X \rangle \} / \{ (1/\sigma_D) + (1/\sigma_{DE}) \} \quad 5.11$$

In Equation 5.11, the terms within $\langle \rangle$ represent the spatial aggregate. $\langle {}^OX \rangle$ is the overall spatial aggregate, DX represents scaled advection, ${}^{DE}X$ represents the combined map that accounts for distributed errors, σ_D is the aggregate error (RMSE) for the scaled advection, and σ_{DE} is the aggregate error (standard deviation of the forecast ensemble) for ${}^{DE}X$, obtained from the “Bayesian ANN” technique described earlier.

The QPF map obtained by combining the results of the Disaggregation and the Evolution model (Equation 5.10) is assumed to best account for the errors at distributed

scales. The weighted aggregate forecast (Equation 5.11) is assumed to best preserve the aggregate information. We combine these by scaling the former distributed map (Equation 5.10), with the latter estimate for the aggregate spatial mean. The details of the scaling algorithm are similar to the spatial disaggregation strategy for NWP-QPF described earlier (Equation 5.7).

Thus, the scaling takes the form:

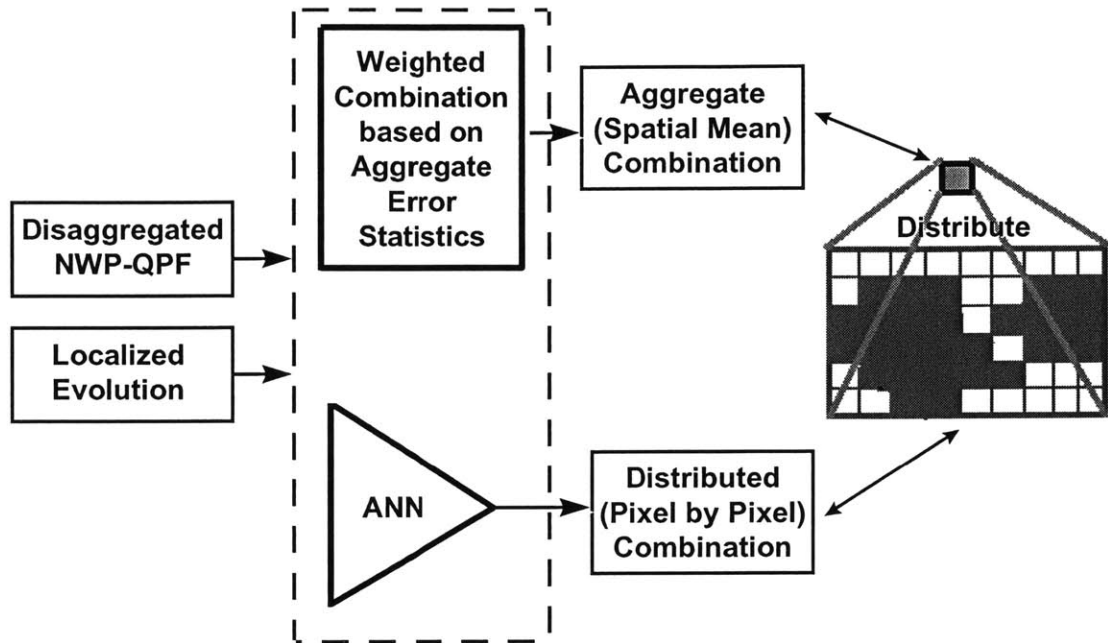
If ${}^{DE}X(i,j,t+k*\Delta t) > 0$ (or, a small threshold ε),

$$X^C(i,j,t+k*\Delta t) = {}^{DE}X(i,j,t+k*\Delta t) + N_{\text{pixel}}\{\langle {}^OX(t+k*\Delta t) \rangle - \langle {}^{DE}X(t+k*\Delta t) \rangle\} / N_{\text{wet-pixel}} \quad 5.12$$

In Equation 5.12, X^C represents the final combined map. The notations are similar to Equation 5.7 and/or Equation 5.11.

The process of disaggregating NWP-QPF from lower to higher resolution, and the use of NWP forecasts of atmospheric variables to evolve the scaled advected radar maps, could introduce errors both at aggregate and at distributed scales. Our results (described later) not only confirm this, but also show that the nature of the behavior of the two approaches at aggregate and distributed scales are not identical. We have hypothesized that the error structures at different scales are separable, and could be modeled separately. This is the rationale for modeling the distributed and aggregate errors separately, and then combining these to obtain the final forecasts. The improvements in skills obtained from this approach, shown later, vindicate our hypothesis.

Figure 5-7: Residual Structures Strategy - Distributed and Aggregate Scales



The fourth and final component of the proposed QPF strategy is Residual Structures. The complementary skills of the Disaggregation (Large Scale Physics) and Evolution strategies are combined. Distributed combination is achieved at each pixel level using ANN. A weighted combination in terms of the aggregate spatial means is achieved by using the aggregate error statistics.

6.0 Reference Models and Measures of Skill

6.1 Reference Models for Distributed QPF

We compare the results from our proposed approach for distributed QPF, with three most commonly used “baseline” methods. The baseline methods are persistence, advection, and a hybrid model that combine NWP-QPF with advection. These are described briefly in this section.

6.1 Persistence

Persistence implies that the forecast at any point in space at all lead times would be exactly equal to the current value. Expressed mathematically,

$$X(i, j, t+\Delta t) = X(i, j, t); \text{ for all } i, j, \Delta t \quad 6.1$$

In Equation 6.1, (i, j) represent a distributed (radar) pixel, t the current time (or the time when the last available rainfall map is available to initialize the forecast generation process), and Δt the forecast lead time (multiples of an hour for the purposes of this research).

6.2 Advection

Advection refers to translation of radar images in forecast lead times, using estimates of velocity scales. This could be expressed mathematically:

$$X(i, j, t+\Delta t) = X(i-i_g, j-j_g, t); \text{ for all } i, j, \Delta t \quad 6.2$$

Equation 6.2 is identical to the equation that describes persistence (6.1), with the exception of the (i_g, j_g) terms, which denote the movements in the x- and y- coordinates. This equation assumes that the highest resolution that can be estimated would be in

pixel units, this holds for our data scenario as there is no measurement higher than radar resolution (i.e., smaller than radar grid size). The estimate of a velocity scale is implicit in Equation 6.2, this is needed to determine the optimal translation vector (i_g, j_g) . Previous researchers have tried different methods to estimate these velocity scales, as described by Browning and Collier (1989) and the references therein, Charnoboy et al. (1994), Einfalt et al. (1990), Grecu and Krajewski (2000), Mecklenberg et al. (2000) and Golding (2000), among others. The maximum temporal resolution of our data (and hence the forecast lead times) is hourly, which precludes the use of local and echo identification or tracking methods, as well as mathematical or stochastic descriptions of precipitation fields. Previous research seems to indicate that the decorrelation times for distributed rainfall structure is of the order of an hour. Based on these previous research and investigative data analyses (see previous sections and the references therein), we used two-dimensional global correlation in space to estimate the velocity scales.

4.3 Hybrid Models Combining Extrapolation and NWP-QPF

Combining information from radar and NWP model outputs is a recent area of research, the NIMROD system of Golding (2000) in the U. K. Meteorological Office represents the latest attempt in this area. This NIMROD system essentially assigns relative weights to the results of advection and NWP-QPF.

NIMROD considers two types of Advection, one using 15 minute rainfall, another using hourly rainfall. The forecast lead time is 15 minutes to 6 hours. The 15-minute Advection component shows high initial skills, but no skills beyond about an hour. This component is not considered in this research. The relative skills from 1-hour Advection

and NWP-QPF were found to be comparable, with 1-hour Advection performing slightly better at 1-3 hour lead times, and NWP-QPF at 4-6 hours (Golding, 2000).

For comparison with the proposed approach, we used an enhanced scheme to combine radar extrapolation and NWP-QPF. Radar extrapolation was handled through advection, and NWP-QPF was distributed in time using linear interpolation followed by error correction, and in space by using the results of advection. This method preserved the distributed structure of the advection results, as well as the large scale mean from NWP-QPF. This strategy, which we shall compare with the proposed approach, could be viewed as a distributed version of NIMROD (see Figure 6-1).

Figure 6-1: The Nimrod QPF strategy (after Golding, 2000)

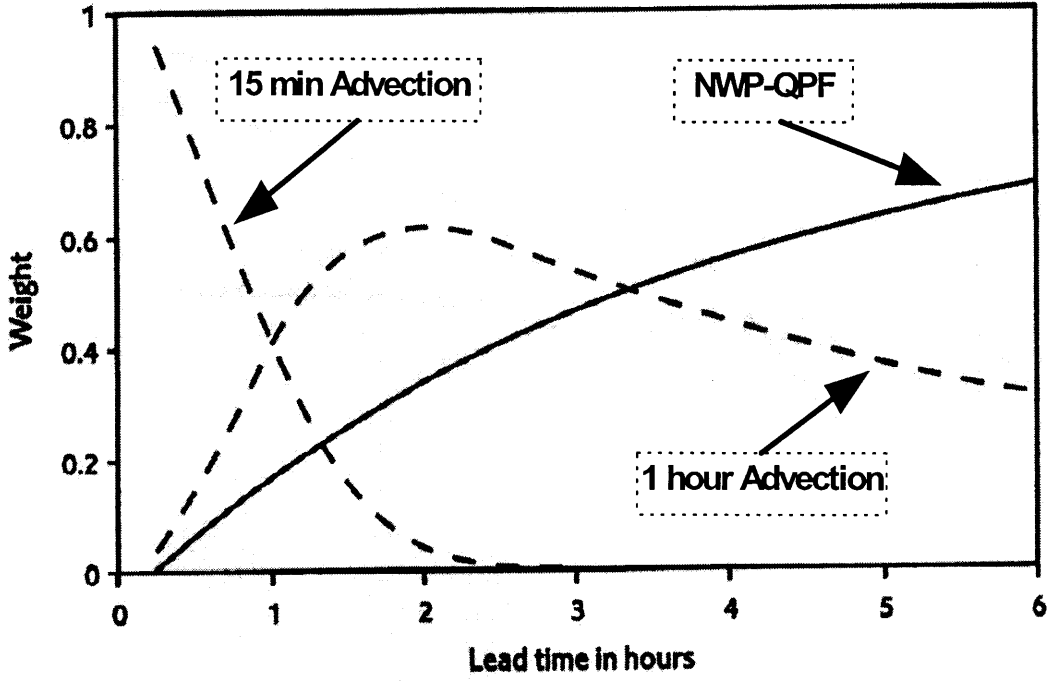


Figure 6-1 describes the relative weights attached to the results of 15-minute Advection, 1-hour Advection, and NWP-QPF as a function of forecast lead times in hours. The 15-minute Advection component uses 15-minute radar rainfall, while the 1-hour Advection component uses 1-hour radar rainfall. The relative weights are obtained from data analysis at the U. K. Meteorological Office, and could be indicative of the relative skills. The 15-minute Advection component appears to have no skills after about an hour. Skills from the 1-hour Advection component appear comparable to that from NWP-QPF. NWP-QPF does not appear to show any skill decay for 1-6 hour leads.

6.2 Measures of Skill for Distributed QPF

6.2.1 Choice of an Optimal Measure of Skill for Distributed QPF

Murphy (1993) indicates that three types of skills could be used to measure forecast performance: Type I skills are subjective and refers to whether a forecasters' best available knowledge and information have been utilized. Type II skills measure the performance of a forecast on verification data using error statistics. Type III skills measure the cost associated with the gain in forecasts in the context of their eventual use. Past QPF researchers have used all of these to various degrees (Ebert and McBride, 2000; Mesinger, 1996; Toth et al., 2000).

6.2.2 Average Distributed Skill: Inverse Normalized Root Mean Squared Errors

To measure the average "closeness" of the distributed forecast and the observed values (i.e., a Type II skill measure), we used the Root Mean Squared Errors (RMSE) of the verification data normalized by the standard deviation (Normalized root Mean Squared Errors, NMSE). As in Ebert and McBride (2000) and Golding (2000), we considered the pixels with non-zero rainfall only. We used the inverse of the NMSE as a measure of the skill. The average $1/\text{NMSE}$ was used across multiple precipitation events to get a generic estimate.

If a forecast were to exhibit "perfect skill", the root mean squared errors, and hence the NMSE, would be zero. For a stationary process, an NMSE of unity indicates that the forecast is no better than the mean of the verification data, i.e., the forecast has "no skill". For non-stationary processes like hourly rainfall, the NMSE provides an indication of the relative performance.

6.2.3 Surrogate Cost Measure: Redefined Threat Score for Conditional Bias

The eventual use of distributed, short term QPF is typically flood and flash flood prediction. A logical estimate of cost could be obtained by routing the precipitation forecasts through lumped or distributed models of hydrology, and estimating the accuracy of the discharge at outlets, distributed stream flow, or soil moisture. However, to isolate the effect of improved QPF alone in that context would be rather difficult (Golding, 2000). The variability associated with infiltration and runoff processes, including parameters like antecedent soil moisture conditions, would tend to dominate. These considerations led us to explore statistical measures that could give an indication of the possible cost or “threat” associated with a forecast which is either too high or too low. If the distributed short term rainfall forecast is significantly higher than the observed rainfall, there might be a dollar value associated with it in terms of the use of early warning systems and unnecessary evacuation or other preparatory measures, but possibly more importantly the longer term erosion of the degree of belief on the forecasts. The cost of a low forecast could be damage to property and maybe even human life, as floods and flash floods might be under-predicted. A true estimate of this cost would also need to factor in the effectiveness of the early warning systems and the expected response.

Given that a true estimate of the dollar value or other costs could be difficult to obtain for distributed QPF, and that both high and low predictions could have associated costs, we choose to use a redefined version of the “threat score” (Mesinger, 1996) as a surrogate cost measure. The threat is defined as the ratio of the number of times both the forecast and the observed values match a predefined criteria (a “hit”), to

the number of times either of these meet that same criteria. Thus, Threat = $N_{hit}/(N_{obs}+N_{fct}-N_{hit})$; where N_{hit} is the number of hits, and N_{fct} and N_{obs} are the number of forecasts and observations satisfying the predefined criteria. The threat is often used in the context of QPF from NWP, where it is defined as the exceedence ratio, or the number of times the forecast or observed values exceed a given threshold. For distributed QPF, we would want to penalize the number of times the values lie outside (above or below) a few sets of prescribed threshold ranges, and take an average for all these ranges. In situations where the models produce probabilistic QPF, the quantification would be in terms of the number of times the forecast confidence bounds intersect the prescribed thresholds.

The redefined Threat Score measures whether the forecasts and observed values happen to lie within the same predefined ranges, with the possible implication that there could be an associated cost if these were not to do so. We selected several arbitrary threshold ranges. The minimum and maximum values of the ranges were, in mm/hr, 0.05-1.00, 0.10-1.00, 0.15-1.00, 0.20-2.00, 0.25-2.00, 0.30-2.00, 0.40-3.00, 0.50-3.00, 1.00-5.00, 2.00-6.00, and 5.00-10.00. The cost measure was calculated as the average, based on all of these arbitrary choices for the threshold ranges, in an attempt to make the measure independent of the choice of a specific threshold range.

The threshold ranges are overlapping. However, the definition of a “hit” and the fact that we take an average over all thresholds, avoid double counting. To understand what this quantifies, let us first consider the usual definition of a threat score, which measures exceedence ratios (i.e., the upper threshold range is infinity). This threat score measures exceedence over several arbitrary thresholds, for example in the range

0.1 mm/hr to 1 mm/hr. For each such threshold, the number of “hits” and hence the calculated threats (ratio of hits to the ratio of either the forecast or the observed exceeding the threshold), are different. These individual threat values are then plotted as a function of the corresponding thresholds. We explored ways to modify this measure to obtain an aggregate measure of skill (or cost). One such measure could be a simple average of the individual threats obtained for each threshold. This measure could be used to compare the performance of different forecast strategies at various lead times, provided the thresholds are identical for each comparison point. The interpretation of such a measure would be the skill or cost that applies on an average over the entire range of thresholds. Thus, if the threats are 40%, 35%, and 30% for threshold ranges of 0.1, 0.2, and 0.3 mm/hr, the average threat over 0.1-0.3 mm/hr would be 35%. Note that these are all overlapping thresholds, since the upper bound is always infinity. Our next step was to extend the definition of threat to consider not just the exceedence ratios (i.e., an upper threshold of infinity), but to quantify the ability to forecast within precipitation ranges (i.e., finite upper thresholds). Once again, we chose several overlapping thresholds. The aggregate cost measure was defined as before, through a simple average. As before, this represents an aggregate measure for the average skill or cost over the entire range of threshold bounds.

For a perfect forecast, $N_{hit} = N_{obs} = N_{fct}$, leading to a threat score of $N_{hit} / (N_{hit} + N_{hit} - N_{hit}) = 1$. The threat score for a random forecast could be approximately derived. Let us start with the assumption that for a perfectly random forecast, the probability of either the forecast or the observed value falling within the prescribed thresholds is equal to that of their falling above or below the thresholds. This assumption could

approximately hold on the average for a set of threshold ranges, depending on the choice of the thresholds. There could be three possible “states” for the observed value, that of exceeding the threshold, falling within the threshold, and falling below the threshold. Similarly, there could be three states for the forecast, leading to a total of 9 combined states. Of these 9 states, the number of states in which a “hit” occurs (both observed and forecast lying within threshold) is 1. The number of times the forecast could fall within the threshold is 3, which is the same as the number of times the observed could lie within the thresholds. Based on our assumption that the probability of the “states” are equal, the threat for a random forecast is approximately $\{1/9\}/(\{3/9\}+\{3/9\}-\{1/9\}) = 1/(3+3-1) = 1/5 = 0.2$. A threat score of 0.2 would therefore represent no skill, according to the definition of threat used here.

6.2.3 Qualitative Visual Measure: Surface Plots of the Errors

Besides statistical measures of skill, it might be of interest to get a feel for the improvement in distributed QPF visually from the precipitation maps or from the plots of their errors. The challenge there is more in the depiction, especially given that for hourly rainfall, the observed rainfall structures rarely persist, if at all. A careful study of the contour plots of the rainfall and the forecast errors, as well as surface plots of the errors themselves, often help identify the characteristics of the forecasts and the improvements obtained. Surface plots of the probabilistic QPF bounds and the measured QPF could provide an indication of the overall performance of the algorithms in terms of forecasting both the mean and the error bounds.

7.0 Description and Availability of Data

This section describes the availability and quality of radar data and the NWP model outputs. We also describe the precipitation events used for model building, calibration, and verification.

7.1 Numerical Weather Prediction Model Outputs

We have used outputs from the 48-km NWP-Eta model. Model outputs from the NGM and the 48-km Eta are archived for the continental United States by the National Weather Service (NWS), the Eta results are thought to be more accurate. Black (1994) provides a detailed description of the NWP-Eta model. As explained earlier, among the NWP model outputs, QPF is usually the one with the highest variability in space and time, and thought to be the least accurate (Mesinger, 1996; Antolik, 2000). The NWP model outputs which correspond to forecasts of atmospheric variables (e.g., Lifted Index - LI, Temperature - T), are not only more uniform in space and time, but also thought to be more accurate. The NWP models are run twice daily, at 00z and 12z UTC (Universal Time Coordinate). Archived model outputs are available at selected stations in the US as FOUS (Forecast Outputs for the United States). Table 7-1 shows a sample FOUS data set for the NWP-Eta.

Table 7-1: Sample FOUS data for the NWP-Eta model

```

FOUS 69 KWBC 200000
OUTPUT FROM ETA 00Z FEB 20 97
TTPTR1R2R3 VVLI PSDDFF HHT1T2T3
OKC//817450 03802 170910 60141209
06023918955 11300 130914 60131310
12035967854 01601 131610 59131310
18058968955 11800 071103 60151309
24066958255 05000 040108 59151108
30031979148 06004 083627 52100604
36013867328 00411 143628 45050201
42000754716 -1010 213624 39050100
48000644023 -1407 243618 37070200

```

Table 7-1 shows sample outputs from FOUS for the 48-km NWP-Eta. The first row is a header line. The second row states that these outputs are from the Eta, and was generated for 20th February, 1997 for the model run at 00z UTC. The third row shows the fields allocated for the lead times and the model forecasts. Thus TT represents lead time in hours, and the number of fields allocated are two (06, 12, 18, 24, 30, 36, 42, or 48). Detailed explanation for each of the output fields are given later. The fourth row states that the FOUS station identifier is OKC (Oklahoma City), and indicates the model initializations. The fifth row onwards gives the model outputs (forecasts) from 6 hour to 48 hour lead times, at 6-hourly increments. In the fifth row, the first 2 fields (06) correspond to TT (lead time in hours), and the following 3 fields (023) correspond to PTT (6-hour cumulative precipitation in hundredths of inches) .

Table 7-1 shows a typical FOUS output, which are the NWP-Eta model outputs at selected stations. The 14 columns (TT, PTT, R1, R2, R3, VVV, LI, PS, DD, FF, HH, T1, T2, T3) represent a different forecast variable, while the last 8 rows represent the forecast lead times of 6, 12, 18, 24, 30, 36, 42, 48 hours. Across a row, one sees the values of the forecast variables at a given lead time, down a column one reads the value of one forecast variable at different lead times. The first header line indicates that this belongs to “FOUS 69”, individual stations are grouped by different FOUS numbers. The second header line indicates that this is an output of the Eta model produced on the 20th of February, 1997, at 00z. Note that “Z” in the context of time is the Greenwich Mean Time, GMT, identical to the Universal Time Coordinate, UTC, which is five hours ahead of the Eastern Standard Time, EST. The third header line symbolizes the forecast lead time and the model outputs, as summarized later in this section. The fourth line starting with “OKC” indicates the station identifier, in this case Oklahoma City. The rest of the third line provides the model initializations, note that there is no forecast lead time (TT) or cumulative precipitation (PTT) for the initialization.

The last 8 rows can be interpreted as below (adapted from the FOUS guidelines):

1. TT: 2 Fields: Forecast projection (hours); 06, 12, 18, 24, 30, 36, 42, or 48.
2. PTT: 3 Fields: 6-hour accumulated precipitation (hundredths of inches); 0.23” for first 6 hours, 0.35” for 6-12 hours, 0.58” for 12-18 hours and so on.
3. R1: 2 Fields: Average Relative Humidity (%) from the surface layer to a depth of approximately 35 mb; 91% for the 6 hour projection.
4. R2: 2 Fields: Average Relative Humidity (%) from 35 mb above the surface layer to 500 mb; 89% for the 6 hour projection.

5. R3: 2 Fields: Average Relative Humidity (%) from 500 mb to the tropopause; 55% for the 6 hour projection.
6. VVV: 3 Fields: Vertical Velocity (tenths of microbars per second) at 700 mb, positive numbers denote rising air while negative numbers denote subsidence, -99 is the lowest coded value; 11.3 $\mu\text{b/s}$ for 6 hour projection.
7. LI: 2 Fields: 6 layer Lifted Index in $^{\circ}\text{C}$, negative values are added to 100; 0 $^{\circ}\text{C}$ for 6 hour projection.
8. PS: 2 Fields: Sea level pressure in mb, leading 9 or 10 is not encoded; 1013 mb for the 6 hour projection.
9. DDFF: 4 Fields: Wind direction (tens of degrees) and average speed from 0-35 mb; 14 knots and 90 degrees at 6 hour.
10. HH: 2 Fields: 1000-500 mb thickness (decameters), leading 4 or 5 is not encoded; 560 dm for the 6 hour projection.
11. T1: 2 Fields: Temperature ($^{\circ}\text{C}$) of 0-35 mb layer, negative values are added to 100; 13 $^{\circ}\text{C}$ for the 6 hour projection.
12. T2: 2 Fields: Temperature ($^{\circ}\text{C}$) of layer centered at 900 mb, negative values are added to 100; 13 $^{\circ}\text{C}$ for the 6 hour projection.
13. T3: 2 Fields: Temperature ($^{\circ}\text{C}$) of layer centered at 800 mb, negative values are added to 100; 10 $^{\circ}\text{C}$ for the 6 hour projection.

Note that precipitation forecasts are usually the least certain of all the forecast numbers from NWP, and are often biased low for large events.

Figure 7-1: FOUS locations in the United States (after NWS Internet site)

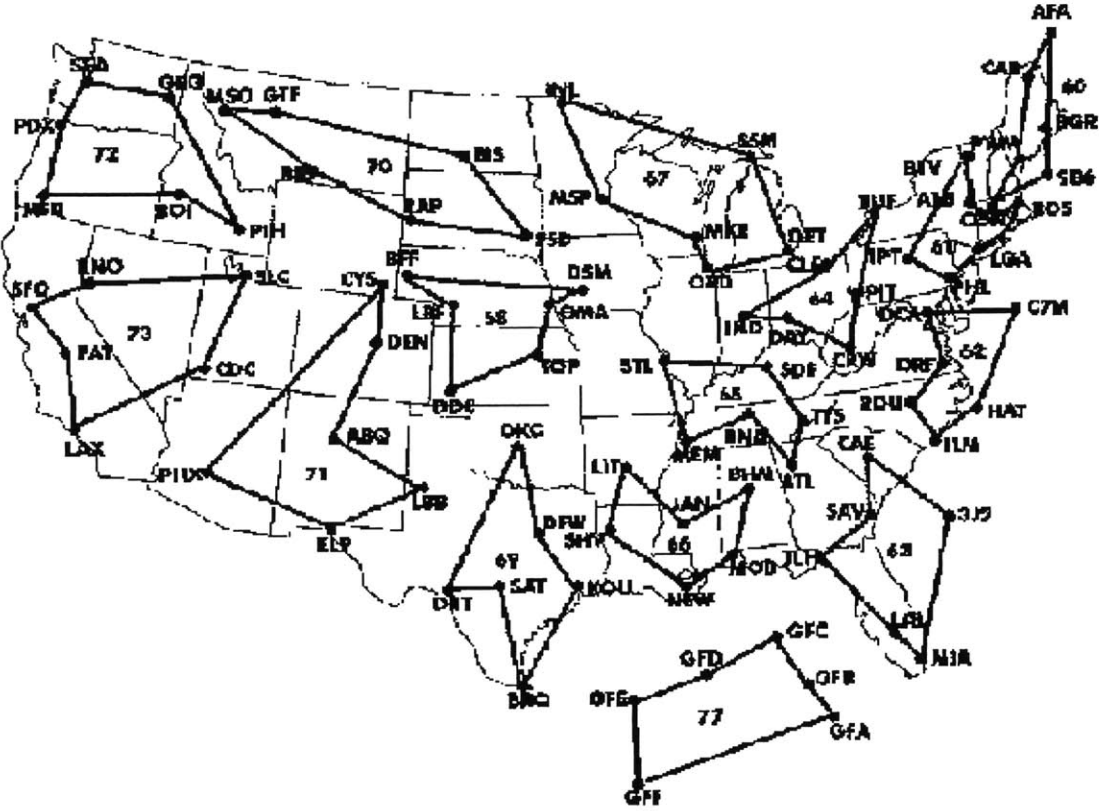


Figure 7-1 shows the locations in the United States for which FOUS (Forecast Outputs for the United States) values are available from the NWP-Eta model.

7.2 Radar Rainfall

We have used radar rainfall from the NEXRAD (NEXT generation RADar) system, which consists of S-band WSR-88D (Weather Surveillance Radar - 1998 Doppler) radar. The geographical coverage includes the continental US and some selected overseas stations. There is continuous spatial and temporal coverage. The NEXRAD radar output is the Equivalent Reflectivity Factor (Z_e). These raw outputs are processed through the Precipitation Processing System (PPS).

Radar measurements are calibrated with automated, real-time surface weather data, the final processing is carried out at the 13 River Forecast Centers (RFCs) of the National Weather Service (NWS). Detailed analysis is performed for the Hydrologic Rainfall Analysis Project (HRAP) region. The grid size in the HRAP region is approximately 4x4 km, the temporal resolution is 1 hour.

In the United States, the highest quality gage corrected radar rainfall estimates are archived in the ABRFC (where the coverage of radar and rain gages is dense), for the regions in and around Oklahoma (owing to the presence of high quality ground measurements from the Oklahoma MesoNet).

The Stage III calibrated radar was developed in the late 1980s by the NWS Office of Hydrology (OH). The calibration methodology is summarized in an Internet site of the NWS (“Operational Precipitation Processing Methodologies”, by J. Schmidt, B. Lawrence and B. Olsen of the Arkansas-red Basin River Forecast Center, in <http://www.srh.noaa.gov/abrhc/p1vol.html>). This methodology “computes an average bias across a radar’s umbrella by comparing some observed gage reports to the corresponding WSR-88D precipitation estimate. The average bias of these sites is then

applied to every grid cell within that radar's coverage. Individual radars are then mosaicked together. Areas that are covered by more than one radar are resolved by either accepting the largest value or the average value of all multisensor precipitation estimates, depending on the user's preference. This is a rather simplistic overview of Stage III, but it captures the key assumption that the ABRFC believes accounts for the underestimation of rainfall by Stage III, the use of a single average bias per radar”.

There is some debate in the hydrological and meteorological community on the value of calibration for radar rainfall. Gage measurements are often not indicative of even small surrounding areas, and are often deemed unsuitable for QPF use (Kitchen and Blackall, 1992). However, uncalibrated radar measurements could have significant bias and other errors (Browning and Collier, 1989).

This research uses both uncalibrated Stage I radar rainfall and calibrated Stage III radar rainfall. The Stage I radar was obtained in Topeka, Kansas, from the NHDS (National Hydrologic Data Server). The Stage III radar information is best calibrated in the ABRFC (Arkansas-red Basin River Forecast Center) in South West Oklahoma. We obtained calibrated radar rainfall archived by the ABRFC, from their Internet repository.

7.3 Precipitation Events

Three precipitation events were used from uncalibrated Stage I radar rainfall in Topeka, Kansas (one in the Fall, one in early Summer, and another in mid Summer). Each of these were of fairly long duration (about 12 hours or more).

The first 7 hours were used for calibration, the 7th was used to initialize, and the 8th-12th hours for 1-5 hour forecast generation and verification. The dates for these storm events were for September 25, 1996; June 29-30, 1997; and May 5-6, 1998.

For calibrated Stage III radar data, a total of 6 storms were used in the Arkansas-red Basin River Forecast Center (ABRFC) in South West Oklahoma. An Internet site of the NCDC (National Climatic Data Center) mentions that eight major precipitation events occurred in Oklahoma from 1993 onwards (see Table 7-2), six of these have been used for this research (two of these were measurements corresponding to an identical precipitation event, another was too localized). Of the six events, three were in the Summer or the “warmer seasons”, and three in the Winter or the “cooler seasons”. Precipitation events have often been classified into warmer and cooler seasons by previous researchers (e.g., Antolik, 2000).

The 3 “Winter” events were for February 20th 1997; October 5th 1998; and October 17th 1998. The 3 “Summer” events were in April 27th 1998; May 4th 1999; and June 16th 1999. A total of 16 hour maps were used, 10 for calibration, the 10th for initialization, and the rest for 1-6 hour forecast generation and verification.

Table 7-2: Significant Precipitation Events in Oklahoma
(Courtesy: NCDC Storm Events Database)

8 PRECIPITATION event(s) were reported in **Oklahoma** between **01/01/1993** and **09/30/2000**.

Mag: Magnitude
Dth: Deaths
Inj: Injuries
PrD: Property Damage
CrD: Crop Damage

Click on Location or County to display Details.

Location or County	Date	Time	Type	Mag	Dth	Inj	PrD	CrD
Oklahoma								
1 Idabel Arpt	02/20/1997	07:00 PM	Heavy Rain	N/A	0	0	0	0
2 Valliant	02/20/1997	07:00 PM	Heavy Rain	N/A	0	0	0	0
3 (tul)tulsa Intl Arpt	03/01/1998	12:01 AM	Heavy Rain	N/A	0	0	0	0
4 (tul)tulsa Intl Arpt	04/27/1998	12:00 AM	Heavy Rain	N/A	0	0	0	0
5 (tul)tulsa Intl Arpt	10/05/1998	12:00 AM	Heavy Rain	N/A	0	0	0	0
6 (tul)tulsa Intl Arpt	10/17/1998	12:00 AM	Heavy Rain	N/A	0	0	0	0
7 Tulsa Intl Arpt	05/04/1999	12:00 AM	Heavy Rain	N/A	0	0	0	0
8 (mlc)mc Alester Muni	06/16/1999	12:00 AM	Heavy Rain	N/A	0	0	0	0
TOTALS:					0	0	0	0

Table 7-2 shows the significant precipitation events in Oklahoma from 1993 onwards, as archived by the NCDC. The event for 02/20/1997 were repeated for two locations. For the event on 03/01/1998, there was not much distributed information (possibly this was a localized event). Thus we were able to work with six precipitation events in Oklahoma.

8.0 Results

This chapter presents the results obtained from the proposed strategy for distributed QPF (as well as the components thereof), and the results from the reference QPF methods.

Section 8.1 provides a summary of how the individual component processes of the proposed QPF strategy contribute to the overall forecast skills. The behavior of skills at distributed and aggregate scales are often different, these are therefore shown separately. (Note that detailed discussions on the skills of the individual components of the proposed QPF strategy, as well as how these perform at different spatial and temporal scales in the Winter and in the Summer, are presented in Chapter 9).

Section 8.2 compares and contrasts the results obtained from the proposed strategy for distributed QPF and the “state of the art” models. Results are compared through statistical skill and surrogate cost measures, as well as through visual plots of the forecasts and uncertainty bounds. Improvements over the “state of the art” models are highlighted.

8.1 Contribution of the Component Processes

The proposed strategy for distributed QPF, as defined and described in detail earlier, comprises four components, which are applied in sequence. The first component is “Extrapolation”, followed by “Large Scale Physics” and “Localized Evolution”, and finally by the strategy for handling “Residual Structures”. This section summarizes the performance of the individual component processes in terms of the

distributed and aggregate error statistics, and/or measures of skills. For comparison, results from a baseline method (persistence) is also presented.

In general, we found limited improvement over persistence when we used Extrapolation alone for the hourly precipitation maps. This is not surprising, as the Extrapolation strategy relies on advection based on estimates of a velocity scale, which in turn are obtained from spatial correlation. Neither large scale precipitation physics, nor localized processes, are taken into account.

The Large Scale Physics strategy uses disaggregated QPF from NWP to scale the advected radar map. The QPF from NWP is a result of the large scale physics considered by the NWP models. This strategy resulted in minimal improvement in distributed skills (especially in the Summer), but significant improvement in aggregate skills and the relative skills at higher (4-6 hours) lead times. This was expected, because while neither advection nor NWP-QPF are capable of accounting for localized effects like convection or cell dynamics, NWP-QPF accounts for large scale physics.

The Localized Evolution component, based on a high resolution ANN model, was designed to handle localized processes. When this component was applied, the results were much improved at lower lead times (1-3 hours), especially in the Summer and in terms of distributed skills. Thus the performance of the NWP-QPF disaggregation (i.e., the Large Scale Physics strategy) and the ANN based localized evolution (i.e., the Localized Evolution component) appear to be complementary. Since the latter is conditioned on the former, this implies that there are domains where the Localized Evolution component significantly improves the results (e.g., distributed scales and shorter leads), and others where they actually cause the skills to go down (e.g.,

aggregate scales and for longer leads). This is not surprising, as the validity of the complex data dictated models are not expected to last to longer lead times. As long as the models remain valid, they are expected to result in significant improvement, but are known to perform poorly thereafter (for example, see discussions in Weigend and Gershfeld, 1994, and also comments in a QPF context by Browning and Collier, 1989). Also, since our formulations are designed to explicitly deal with localized processes at higher resolutions, the distributed skills typically improve, but the aggregate skills usually tend to suffer, especially for higher lead times.

The strategy for “Residual Structures” is an attempt to use the error statistics from the results of the simple disaggregation strategy alone (i.e., Large Scale Physics), and after the application of the ANN based high resolution model (i.e., Localized Evolution), to improve the overall forecasts at all domains (for example, at both aggregate and distributed scales and for 1-6 hour lead times). As expected, this combination consistently performed better than the individual components.

We show two examples in Figures 8-1 and 8-2. The measures of skill used were $1/NMSE$ for quantifying distributed skills (Figure 8-1), and the average error at NWP resolutions (i.e., 6 hour and 48 km) for quantifying aggregate skills (Figure 8-2). In these figures, the asterisk shows the skills in the baseline persistence method. The circles represent the advection (i.e., Extrapolation) component, the dashed line is for advection scaled with error corrected NWP-QPF (i.e., Large Scale Physics). The dotted-dashed line shows the performance of the ANN based localized (high resolution) evolution in a Lagrangian frame of reference, conditioned on scaled advection (i.e., Localized Evolution). The solid line represents the proposed model that combines the results of

the scaled advection with that of the localized evolution by modeling residual structures (i.e., the proposed strategy for Residual Structures) at distributed and aggregate scales. Note that higher values of the ordinate in Figure 8-1 represent greater skills, but in Figure 8-2 higher values represent greater errors, or lower skills. Figure 8-1 represents an average of three precipitation events in Topeka, Kansas, for which Stage I data were available. Figure 8-2 represents an average of six precipitation events for which we used Stage III data in the ABRFC in Oklahoma.

Figure 8-1 shows the typical distributed skill ($1/\text{NMSE}$) behavior for the proposed model and the individual components thereof. The proposed model performs better for all lead times and for all seasons combined. Figure 8-2 shows that the proposed model is also the best in terms of the aggregate skills. We note that the relative performance of the individual component processes depend on the forecast lead time, and the resolution at which the skills are quantified. The proposed hybrid model is able to optimally combine these components and yield the most skillful forecasts. This validates our overall QPF strategy of decomposing the problem into component processes, and the corresponding hybrid modeling strategy.

Figure 8-1: Contribution of Component Processes - 1/NMSE

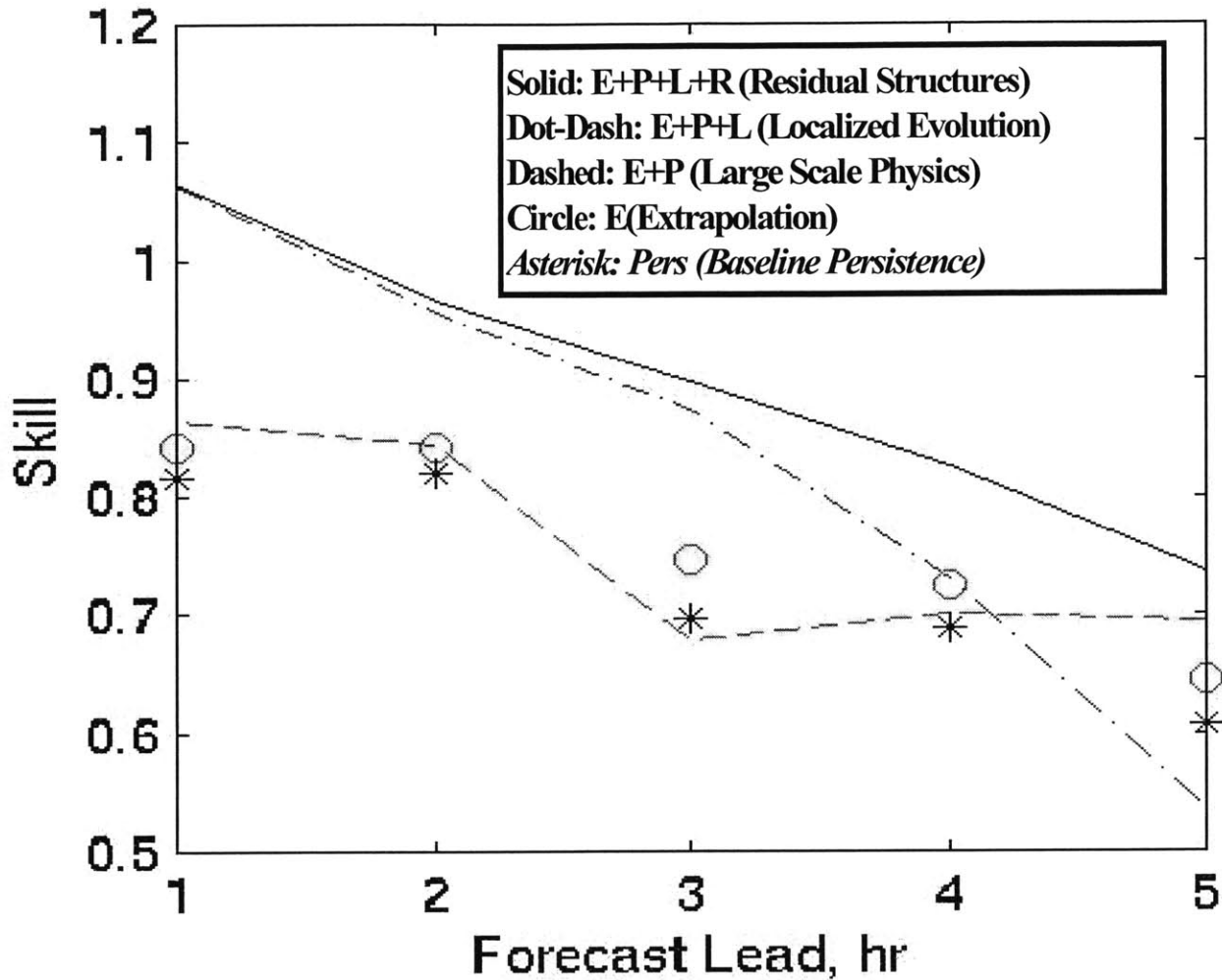


Figure 8-1 shows the contribution of the component processes in terms of distributed skills (1/NMSE using all pixels), as a function of lead time in hours. For non-stationary processes, the NMSE provides an indication of relative skills. Persistence is used as a baseline for reference. 1-hour Advection marginally improves over Persistence, the improvement from distributed NWP-QPF is minimal to zero. ANN based Localized evolution improves skills at 1-3 hour leads, but decays thereafter. The rate of decay is lower for distributed NWP-QPF. The combined method that models residual structures performs the best for all lead times (1-6 hours).

**Figure 8-2: Contribution of Component Processes
Aggregate Precipitation Errors (mm) at NWP scales (6 hr, 48 km)**

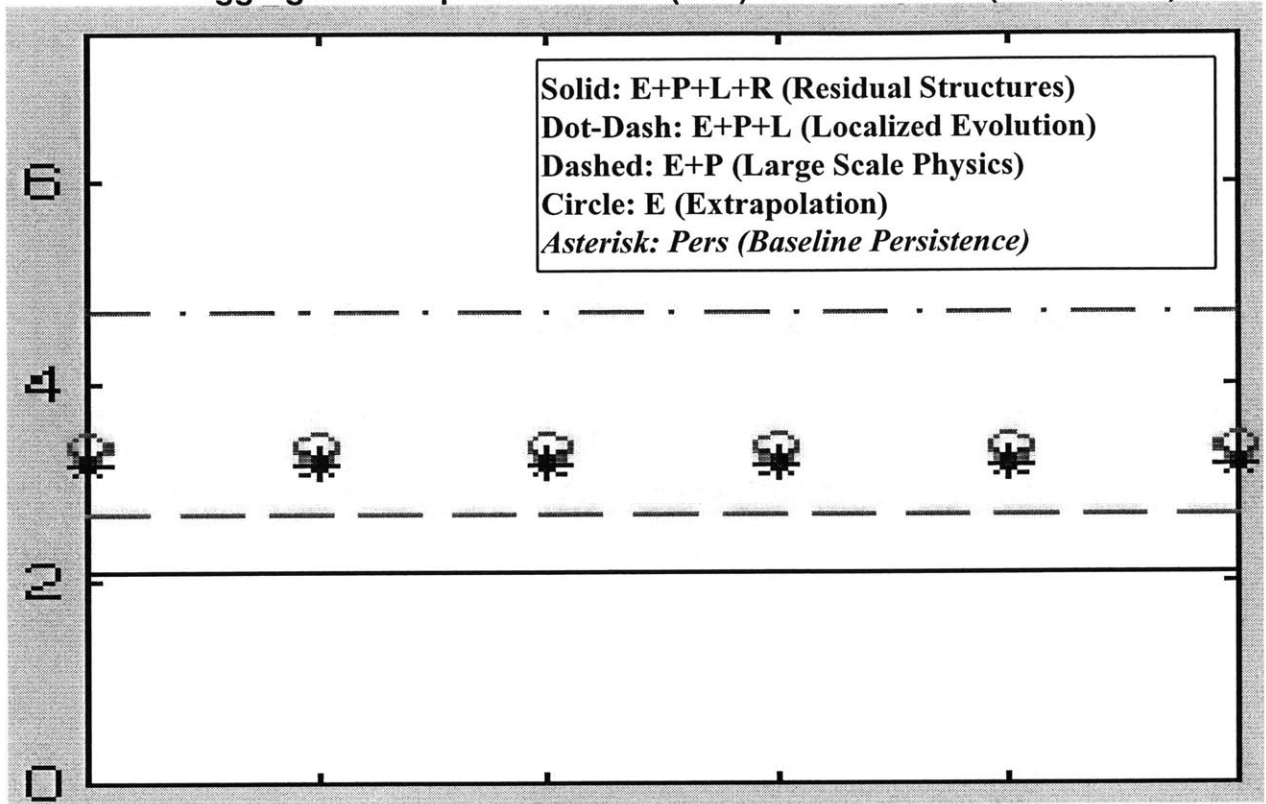


Figure 8-2 shows the contribution of the component processes in terms of aggregate errors at NWP scales (48 km and 6 hour). The x-axis in the figure is for visual clarity only, the y-axis shows the errors at 6 hours and 48 km. Persistence is used as a baseline for reference. 1-hour Advection marginally improves over Persistence. The improvement from distributed NWP-QPF appears significant. High resolution evolution actually performs worse in terms of the aggregate skills, this is caused by the sharp decay in skills at higher (4-6 hour) leads, especially for the spatial mean. The combined method that models residual structures performs the best.

8.2 Improvements to Distributed QPF

As discussed earlier, reference models for distributed QPF include Persistence, Advection, and Hybrid models that combine Advection with NWP-QPF. This section presents the results obtained from these reference QPF strategies, and the corresponding improvements obtained by the proposed strategy for distributed QPF. For these results, we focus on Stage III radar data available from the ABRFC.

A few notations have been used interchangeably in the figures and discussions. “PROP” refers to the proposed QPF strategy. “GA00” refers to Advection (after Grecu and Krajewski, 2000, who suggested that advection is the primary component that contributes to improving distributed QPF). “DN00” refers to the combination of Extrapolation (i.e., advection) and NWP-QPF. The combination methodology used here, as described earlier, could also be viewed as a distributed (and, enhanced) version of the Nimrod strategy of Golding (2000). “PERS” refers to Persistence.

To understand the skills from the various strategies, we have used two quantitative skill measures and two ways of visual depiction (i.e., qualitative “measures”). A normalized version of the RMSE has been used as a measure of the mean distributed skills. In terms of this measure, the proposed strategy performs significantly better than all the other models in the Summer at all lead times, but in the Winter the performance is mixed. A surrogate cost measure has also been defined, by extending the definition of the commonly used “threat score”, in an attempt to quantify the probability of the forecasts and the observed values falling within arbitrary thresholds. The skills of the proposed strategy in terms of this measure is significantly higher than that from the reference methods, both in the Summer and in the Winter.

None of the QPF methods, with the exception of the proposed strategy, appears to perform any better than a “random forecast” in terms of this measure. Contour and surface plots of QPF errors have been shown for the proposed and the reference strategies. Lack of structures in the errors indicate better forecasts. The errors from the proposed strategy almost always has less structure. This is especially true in the Summer, and for several occasions in the Winter as well. DN00 appears to occasionally perform as well or better in terms of this indicator in the Winter. Contour plots are also shown for the forecast errors from the proposed strategy, as well as the errors in the corresponding confidence bounds. These show the extent of the measured precipitation that is not contained within the confidence bounds, thus indicating (visually) the ability of the proposed strategy to produce pQPF (probabilistic QPF) estimates.

On the whole, the proposed method improves over the existing methods in almost all situations and in terms of most of the quantitative and qualitative (visual) measures of skill. While there is some variability on a storm by storm basis depending on the type of skill measure, the season, the resolution, and the lead time, the proposed strategy seems to consistently and significantly improve both in terms of the average distributed skills and in terms of the surrogate cost, as well as through the qualitative visual indicators. This demonstrates the value of the proposed strategy for distributed QPF.

Results are presented in this section through Figures 8-3 to 8-27. First we present the distributed skills (Figure 8-3) and the surrogate cost (Figure 8-4) for all seasons combined. Then we present the corresponding figures for the Winter (Figures 8-5 and 8-6), and the Summer (Figures 8-7 and 8-8). Finally, we depict the QPF errors

for each storm, first through contour plots and then through surface plots. For example, Figure 8-10 show the surface plot and Figure 8-11 the contour plot of the QPF errors for the first storm. Figures 8-12 through 8-20 show the errors for the other five storms. This section discusses the overall performance, while Section 9 discusses the performance in each season (Summer or Winter) in detail.

8.2.1 Improvements to Distributed QPF in terms of Distributed Skills

Figures 8-3, 8-5 and 8-7 present the average distributed skills in the proposed and reference QPF strategies, in terms of the inverse Normalized root Mean Squared Errors ($1/NMSE$). Figure 8-3 shows the skills averaged across all six storms (three Summer and three Winter) in the ABRFC. Figure 8-5 shows the skills averaged for the three Winter storms, while Figure 8-7 shows the skills averaged across the three Summer storms.

The proposed method significantly improves over the “state of the art” techniques in the Summer in terms of $1/NMSE$. This is to be expected for two primary reasons. First, data analysis at aggregate scales and preliminary studies indicate that the information content in NWP forecasts of atmospheric variables is higher in the Summer. The proposed method uses this additional information from NWP, unlike the existing methods. Second, data analysis and preliminary QPF studies also indicate that nonlinear relations dominate in the Summer, which are likely to be captured by the ANN based methods.

The existing methods show some skills, especially in the Summer at short lead times. The improvement of 1-hour Advection over Persistence is minimal, primary due to occasional missed cells and the structures not being preserved for hourly rainfall.

The method that combines Advection and NWP-QPF, usually improves over pure advection or persistence in terms of aggregate skills, but the gain is marginal if any in terms of distributed skills. For 1-hour resolutions and 1-6 hour leads, skills from 1-hour Advection and NWP-QPF have been found to be comparable (Golding, 2000). Finally, the distribution of NWP-QPF in space and time is error prone, as discussed earlier.

8.2.2 Improvements to Distributed QPF in terms of Surrogate Costs

Figures 8-4, 8-6 and 8-8 present the skills in the proposed and reference QPF strategies, in terms of the surrogate cost measure. As described earlier, the field site is the ABRFC in Oklahoma, and the data used were calibrated Stage III NEXRAD (radar) measurements and 48-km NWP-Eta model outputs from FOUS. Figure 8-4 shows the skills averaged across all six storms (three Summer and three Winter) in the ABRFC. Figure 8-6 shows the skills averaged for the three Winter storms, while Figure 8-8 shows the skills averaged across the three Summer storms.

None of the “state of the art” methods demonstrate any skills (i.e., no improvement over a perfectly random forecast) in terms of the cost measure, while the proposed method shows significant skills and improvement over the existing methods. This implies that the proposed method would be able to predict the possibility of the rainfall exceeding or lying within given thresholds better than the existing methods. This could result in significant cost benefits, both in terms of activating advance systems that could warn residents of potential threats, and in terms of reducing the number of false alarms. Neither Advection nor the distributed version of the Nimrod succeeds in consistently improving the cost measure over simple Persistence.

The proposed method significantly improves over all the existing methods, for two primary reasons. The first is the ability of the proposed strategy to better capture the structure of precipitation, as seen from the error plots and the NMSE, especially in the Summer. The second is the ability of the proposed model to generate good estimates of forecast error bounds, both in the Summer and in the Winter.

There is a slight decaying trend in the skills obtained from the proposed model with forecast lead times. The decay rate is rather flat compared to that of the distributed skill measure. This is an indicator of the skills in the probabilistic QPF estimates, and the ability of the QPF strategy to predict the precipitation range better. The rate of decay of the skills from the proposed model, in terms of this surrogate cost measure, appears to be more pronounced in the Summer than in the Winter. Conditions in the Winter are expected to be relatively more stable, hence the predictability in terms of the QPF lying within prescribed bounds is higher.

8.2.3 Visual Improvements to Distributed QPF

8.2.3.1 Comparison of the Proposed and Reference QPF Strategies

We have used hourly rainfall maps for model development and calibration, and hourly maps for verification and forecast generation. Previous research and data analysis indicates that the rainfall structure is not too well preserved at spatially distributed hourly scales, as we have discussed earlier. To visually demonstrate QPF improvement, we looked at the rainfall maps and QPF error plots for each precipitation event (storm). Contour plots showed the structure in the residual errors. Even though the contour values were indicated in the figures, these often failed to effectively depict

the error magnitudes. Surface plots of the hourly QPF errors (forecast less observed) summarized the improvements in distributed QPF. The surface plots from the proposed strategy showed marked improvement over the reference methods in the Summer. For the Winter, the visual behavior of the distributed errors appeared to be more erratic. In general, the distributed errors in the Winter were higher than that in the Summer.

To facilitate a storm by storm analysis of the QPF errors, we first present the contour plots and then the surface plots for each storm in sequence. These are Figures 8-9 and 8-10 for Storm 1, Figures 8-11 and 8-12 for Storm 2, Figures 8-13 and 8-14 for Storm 3, Figures 8-14 and 8-15 for Storm 4, Figures 8-16 and 8-17 for Storm 5, and Figures 8-19 and 8-20 for Storm 6. The exact dates and locations are indicated in the figures. The first three are Summer storms, while the other three are Winter storms.

Overall, the proposed hybrid model appears to perform significantly better than the existing methods for all Summer storms, and occasionally (but not always) improves the forecasts for Winter storms. The improvements from Advection over Persistence appears to be marginal if any. The use of error corrected NWP-QPF usually improves the skills in the Winter and occasionally in the Summer. The localized evolution and residual modeling components appear to result in consistent improvements in the Summer, and occasionally improves in the Winter as well.

8.2.3.2 QPF and pQPF Errors from the Proposed QPF Strategy

The performance of the proposed strategy, both in terms of producing the most likely forecasts (i.e., QPF) and the confidence bounds (i.e., pQPF), are visually depicted through contour plots in Figures 8-21 through 8-27. Each of these figures represents one storm. Figure 8-21 is for a storm in Topeka, Kansas (where Stage I radar data was

available), while Figures 8-22 through 8-27 show each of the six storms in the ABRFC (where Stage III radar data was available). The first four storms are for the Summer, while the last three are for the Winter.

Each of the figures (8-22 through 8-27) have four component sub-figures. The first sub-figure shows the measured precipitation at the forecast lead time of 1-hour, while the second sub-figure shows the QPF errors from the most likely forecasts produced by the proposed strategy. The third sub-figure shows the contour plots for the errors that are above the upper confidence bounds, i.e., the measured precipitation less the upper confidence bounds for situations where the former exceeds the latter. The fourth sub-figure shows the contour plots for the errors that are below the lower confidence bounds, i.e., the magnitude of the difference of the measured precipitation and the lower confidence bounds, for situations where the former is less than the latter. Thus, the contour maps depict the errors in both the QPF and the pQPF.

The results indicate that the proposed strategy for distributed QPF performs consistently well in predicting the confidence bounds. In the Summer, the predictions appear to be relatively more accurate, and the pQPF errors often do not indicate any structure. In the Winter, the performance is not as good as in Summer. The structures in the pQPF errors are typically similar to the structures in the QPF errors themselves.

8.2.3.3 Apparent Directional Structure of Precipitation and QPF Errors

We note that visually, the plots of the QPF errors often appear to indicate a directional pattern (i.e. more stretched in the one direction than the other). While one may be tempted to conclude that this is an artifact of the proposed QPF strategy or one of its components, detailed inspection indicates otherwise.

First, the errors from the proposed strategy, or any of the components thereof, do not appear to be any more directional than the errors from hourly persistence (which merely compares the observed precipitation maps at successive hourly time steps). This can be verified by the contour and surface plots of the errors shown in this section (Figures 8-9 through 8-20). Second, even the contour plots of the observed precipitation exhibit this visual behavior to a similar degree (qualitatively) as the errors. This can be verified from Figures 8-22 through 8-27, which show contour plots for the measured precipitation. Thus the issue, if any, seems to be with the directional behavior of the measured precipitation itself, rather than that for the QPF errors. Further evidence is presented in Appendix B, where contours are plotted for measured precipitation as well as for errors from hourly persistence. An interesting side note is that the directional behavior was only observed when the Stage III data were used in the ABRFC, but not for the Stage I data in Topeka, Kansas (e.g., see Figure 8-21).

We investigated two potential causes for the directional behavior in the measured precipitation. First, there could be an issue with the software or the contouring algorithm (note that we used standard MATLAB functions). Second, this behavior could be inherent in the data, either due to some artifact of the measurement or the calibration process, or even as an intrinsic characteristic of precipitation structures in the ABRFC.

We discovered that the directional structures were more apparent for lower contour values. In fact, if we plotted the contour values with higher magnitudes only, these directional structures were either not observed or greatly reduced. This was true both for the measured precipitation, and for the errors from any QPF strategy, including

persistence. Evidence for this is presented in Appendix B. This led us to conclude that while an appropriate selection of the contour values might make this behavior less visible, the behavior did indeed exist in the data, and was more prominent for the lower intensity values. The issue therefore, was not with the contouring algorithm or software.

Assuming our conclusions above were correct, the directional behavior were a characteristic of the measured Stage III precipitation in the ABRFC. There could be two potential reasons for this observed behavior. First, this could be an artifact of the measurement or the calibration process. Second, this could be caused by an inherent property of the storms in the ABRFC region. The second hypothesis could be further qualified. One could argue that since we considered storms in the NCDC database that were known to have generated significant amounts of precipitation, this could be a property of only the more intense storms in the ABRFC. While we cannot totally reject this second hypothesis, it seems less likely. If the structures were caused by some physical process, one would expect them to persist for higher contour values as well. The other possibility (i.e., the directional structures being the possible effect of the calibration or the measurement process) appears more likely, especially as the Stage I precipitation data in Topeka did not appear to exhibit this property. However, we do not have sufficient evidence to accept or reject either of these hypotheses.

Figure 8-3: Average 1/NMSE for All Summer and Winter Storms Combined as a function of the Forecast Lead Time in Hours

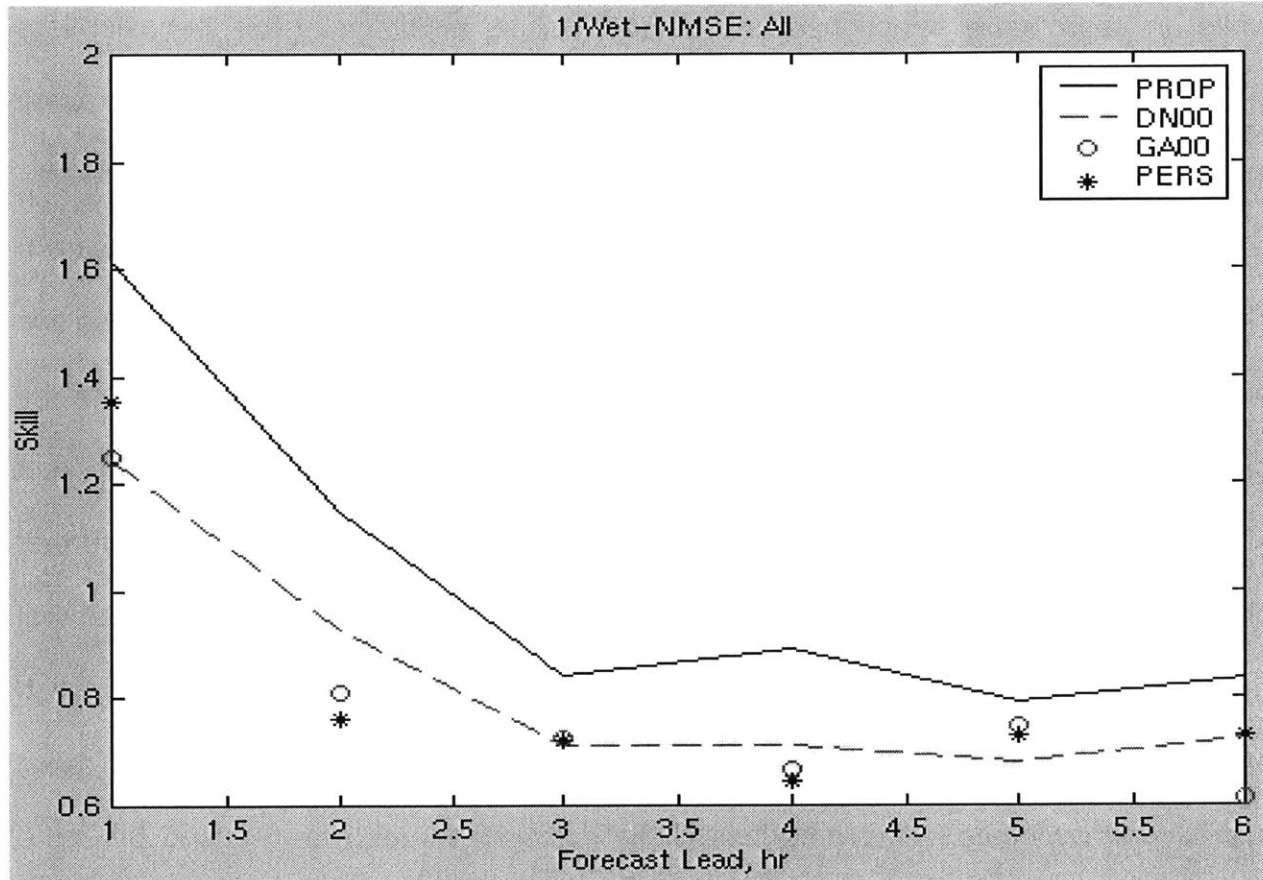


Figure 8-3 shows the performance of the Proposed (PROP) and “state of the art” strategies for distributed QPF, using calibrated radar rainfall and NWP-Eta model outputs, for all seasons combined. The measure of skill used is the average 1/NMSE, which quantifies distributed skills. The skill is shown as a function of the forecast lead time in hours, and is averaged for 6 precipitation events. Persistence (PERS, asterisk) is used as a baseline for reference. Improvement from 1-hour Advection (circle, GA00) is minimal to zero. The distributed version of Nimrod (DN00, dashed line) that combines 1-hour Advection and NWP-QPF, shows minimal improvement. The Proposed strategy (PROP, solid line) improves over all the existing methods for all lead times.

Figure 8-4: Average Threat Score for All Summer and Winter Storms Combined as a function of the Forecast Lead Time in Hours

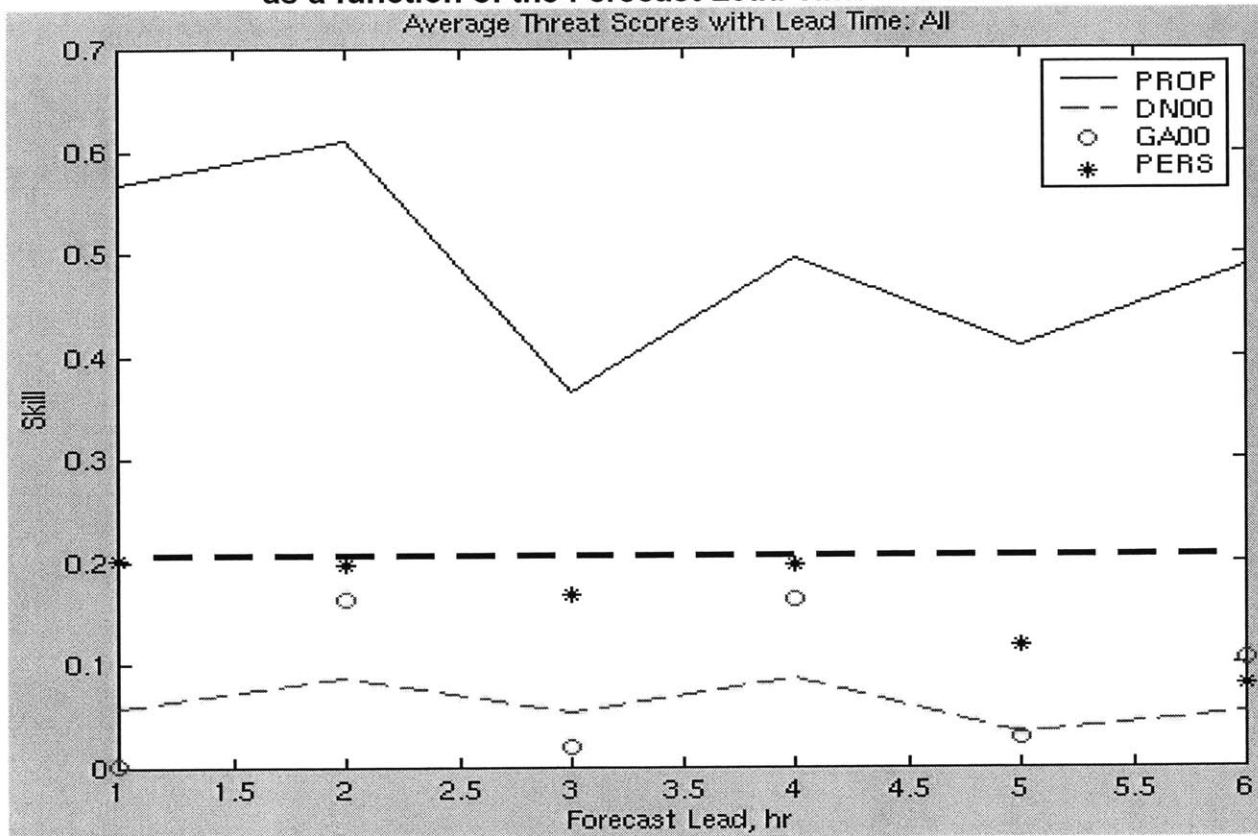


Figure 8-4 shows the performance of the Proposed (PROP) and “state of the art” QPF, using calibrated radar rainfall and NWP-Eta outputs. The average redefined “threat” score quantifies a surrogate “cost” by penalizing values that lie outside prescribed threshold ranges (see text for the definition of the cost measure). The skill is shown as a function of lead time in hours, and is averaged for 6 precipitation events. Persistence (PERS, asterisk) is used as a baseline for reference. Improvement from 1-hour Advection (circle, GA00) is minimal to zero. The distributed version of Nimrod (DN00, dashed line) that combines 1-hour Advection and NWP-QPF, shows no improvement. None of the existing methods show any skill over a random forecast. The Proposed strategy (PROP, solid line) improves over all the existing methods for all lead times, and shows significant skills. The decay of skill with lead time is much less sharp compared to the distributed skill (1/NMSE) in Figure 8-3.

Figure 8-5: Average 1/NMSE for Winter as a function of Lead Time in Hours

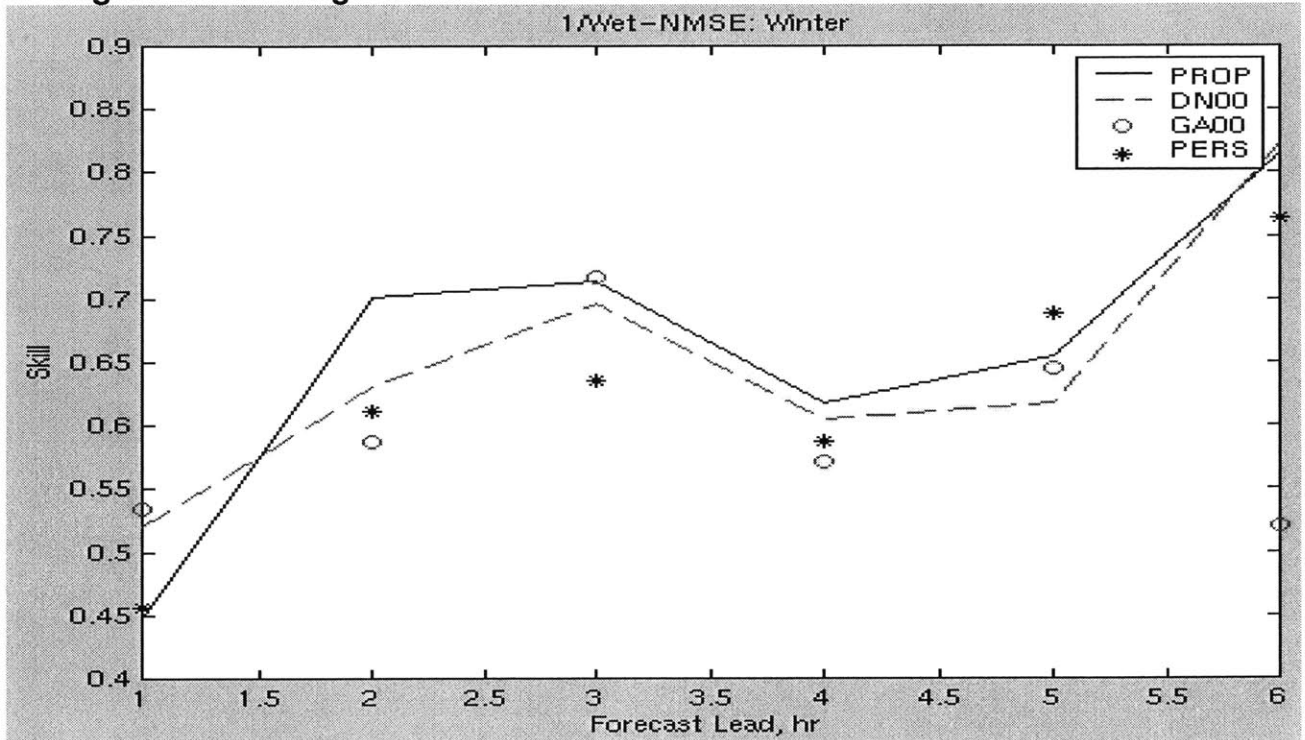


Figure 8-5 shows the performance of QPF strategies, using calibrated radar rainfall and NWP-Eta model outputs, for the Winter. The average 1/NMSE for 3 events is shown as a function of the forecast lead time in hours. Neither the existing methods (Persistence, PERS; 1-hour Advection, GA00; distributed Nimrod, DN00), nor the Proposed method (PROP) show any significant distributed skill. Previous researchers indicate that conditions in the Winter are more stable, and precipitation predictability is higher. We found this to be true for the aggregate forecasts, and for a skill score that measures the number of times the forecasts and observations lie within prescribed thresholds (Figure 8-6). For distributed skills, this was not found to be true. The skills from 1-hour Advection showed a lot of variation in the Winter, with the translation missing the observed values more often. The information content in NWP-QPF was found to be higher for the aggregate, but did not translate to significant distributed skills. The NWP atmospheric outputs were found to have limited information content in the Winter. These contributed to a lack of significant average distributed skills in the Winter.

Figure 8-6: Average Threat Score for Winter as a function of Lead Time in Hours

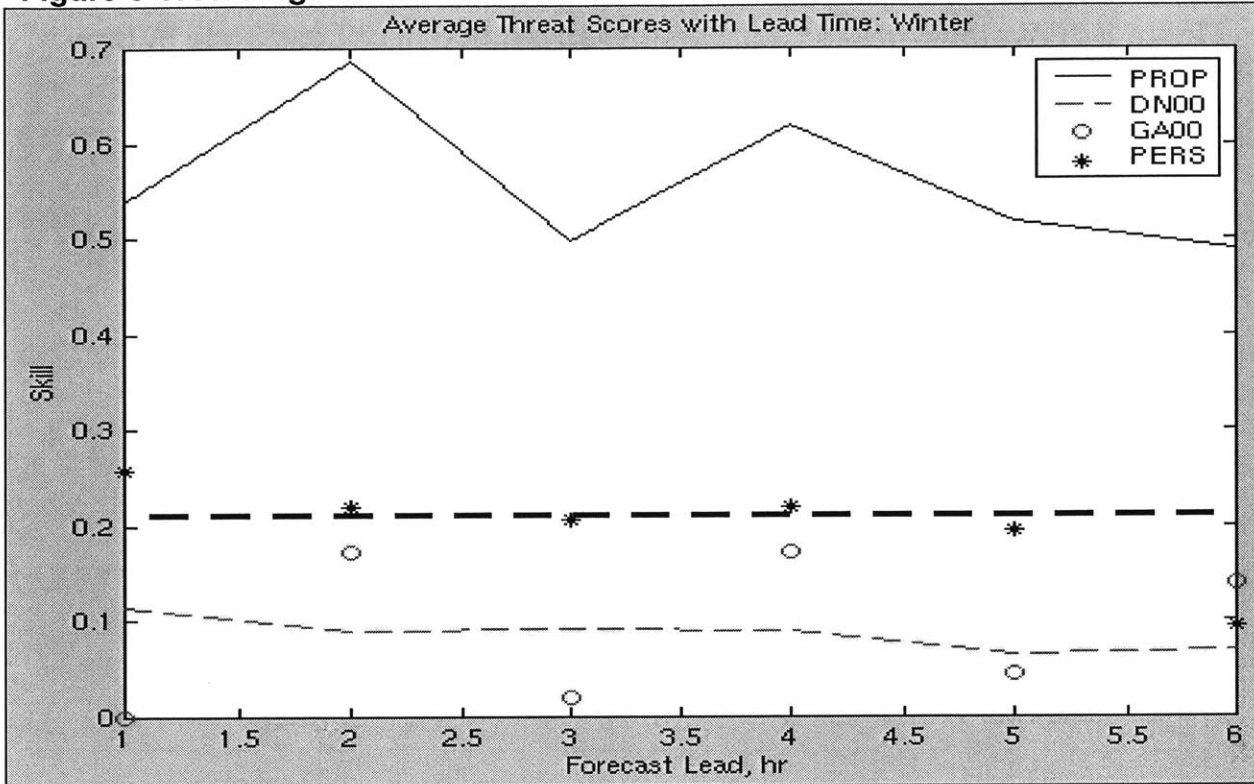


Figure 8-6 shows the performance of the Proposed (PROP) and “state of the art” QPF, using calibrated radar rainfall and NWP-Eta outputs, for 3 Winter events. The average redefined “threat” score quantifies a surrogate “cost” by penalizing values that lie outside prescribed thresholds (see text). Persistence (PERS, asterisk) is used as a baseline for reference. There is no improvement from 1-hour Advection (circle, GA00), or the distributed version of Nimrod (DN00, dashed line) that combines 1-hour Advection and NWP-QPF. None of the existing methods show any skill over a random forecast. The Proposed strategy (PROP, solid line) improves over all the existing methods, and shows significant skills. The skills for the Winter are higher than that in the Summer (Figure 8-8), possibly owing to more stable conditions. The skill decay with lead time is less pronounced for PROP, indicating good estimates for the confidence bounds. This contrasts with the distributed skills from $1/NMSE$ (Figure 8-5). Together, these show that the average distributed forecasts might not be skillful in the Winter in terms of the MSE, but the possibility of lying within prescribed bounds is well predicted.

Figure 8-7: Average 1/NMSE for Summer as a function of Lead Time in Hours

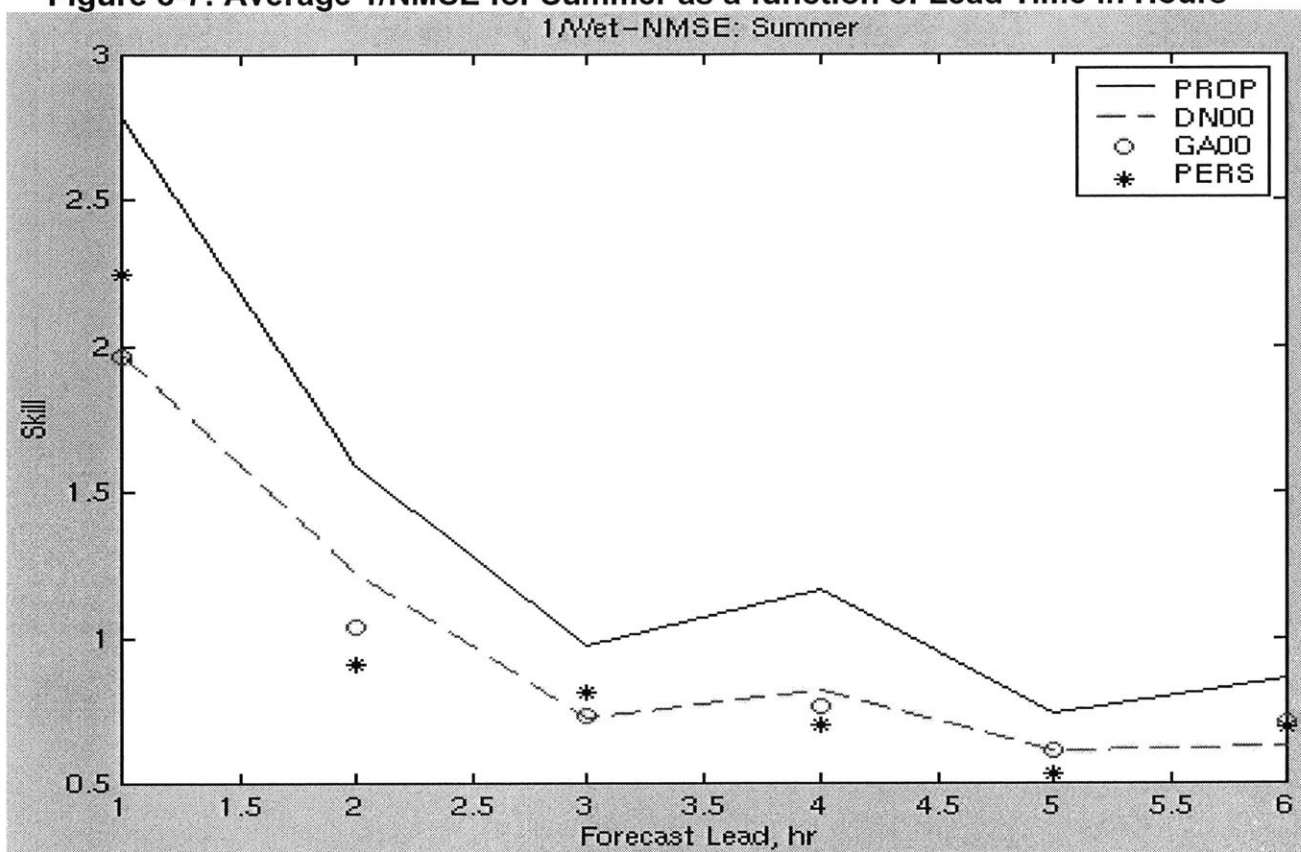


Figure 8-7 shows the performance of the Proposed (PROP) and “state of the art” strategies for distributed QPF, using calibrated radar rainfall and NWP-Eta model outputs, for the Summer. The measure of skill used is the average 1/NMSE, which quantifies distributed skills. The skill is shown as a function of the forecast lead time in hours, and is averaged for 3 precipitation events. Persistence (PERS, asterisk) is used as a baseline for reference. Improvement from 1-hour Advection (circle, GA00) is minimal to zero. The distributed version of Nimrod (DN00, dashed line) that combines 1-hour Advection and NWP-QPF, shows minimal improvement. The Proposed strategy (PROP, solid line) improves over all the existing methods for all lead times. This contrasts with the lack of significant distributed skills in the Winter (Figure 8-5).

Figure 8-8: Average Threat Score for Summer as a function of Lead Time in Hours

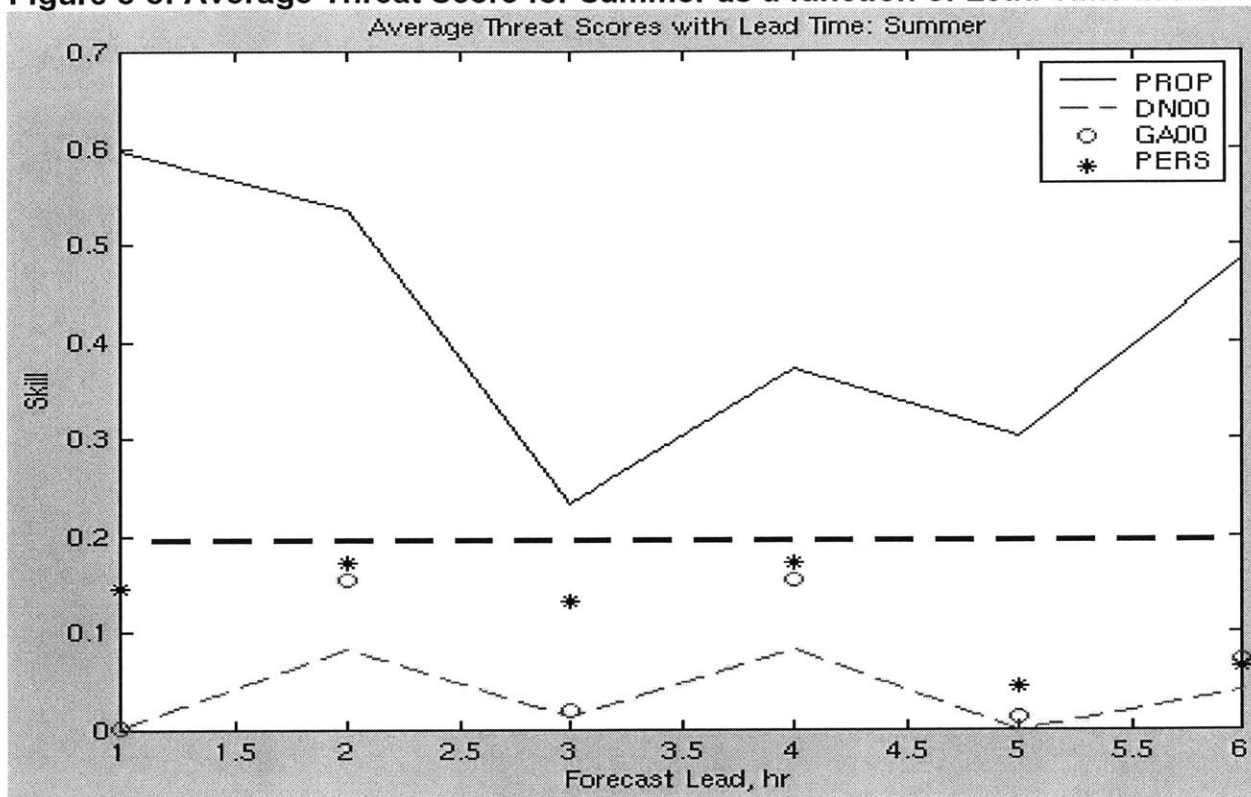


Figure 8-8 shows the performance of the Proposed (PROP) and “state of the art” QPF, using calibrated radar rainfall and NWP-Eta outputs, for 3 Summer events. The average redefined “threat” score quantifies a surrogate “cost” by penalizing values that lie outside prescribed thresholds (see text for the exact values). The skill is shown as a function of lead time in hours. Persistence (PERS, asterisk) is used as a baseline for reference. Improvement from 1-hour Advection (circle, GA00) is minimal to zero. The distributed version of Nimrod (DN00, dashed line) that combines 1-hour Advection and NWP-QPF, shows no improvement. None of the existing methods show any skill over a random forecast. The Proposed strategy (PROP, solid line) improves over all the existing methods for all lead times, and shows significant skills. The decay of skill with lead time is not as sharp as the distributed skill ($1/NMSE$), in Figure 8-7. However, the decay with lead time is sharper than that in the Winter (Figure 8-6), possibly owing to less stable conditions.

**Figure 8-9: Precipitation Error (mm/hr) Contour Maps - Summer Storm "A"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)**

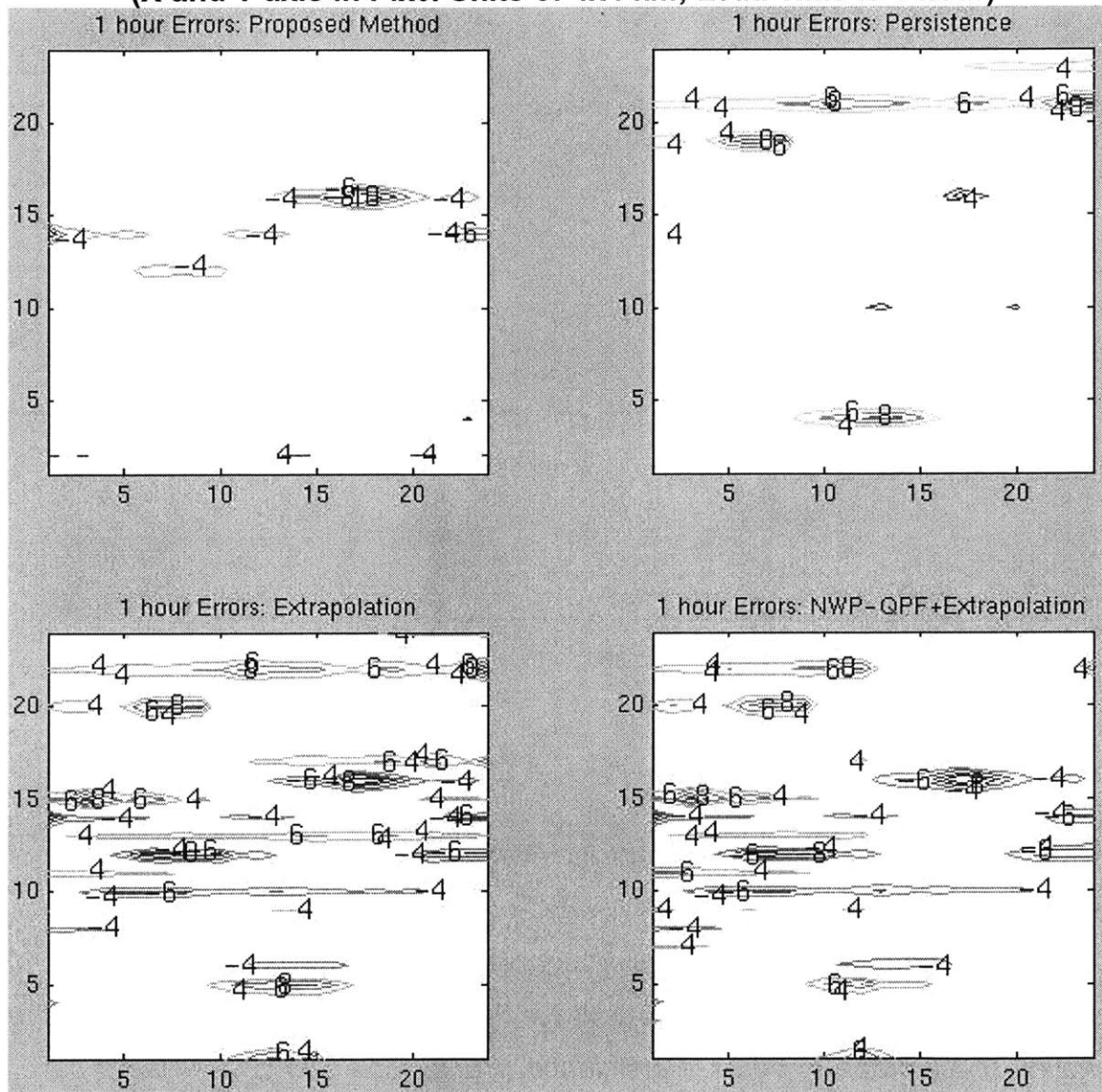


Figure 8-9 shows contour plots for the distributed errors (April 27, 1998 in the ABRFC). Visually, neither Extrapolation (GA00) nor the combination of Extrapolation with NWP-QPF (DN00) seems to have improved over Persistence (PERS) in terms of the residual error structures. The Proposed strategy (PROP) seems to have reduced the error structures compared to the existing methods.

Figure 8-10: Precipitation Error (mm/hr) Surface Plots - Summer Storm "A"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)

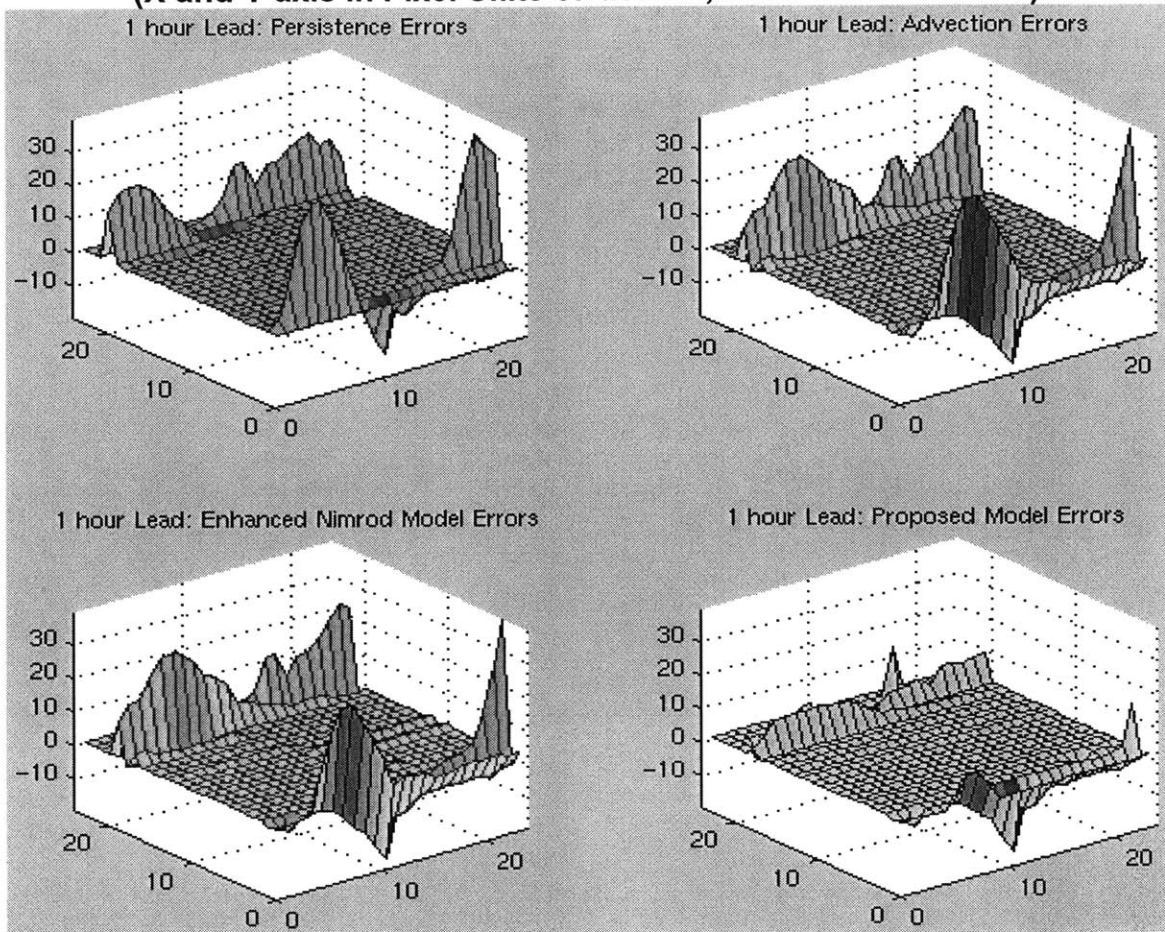


Figure 8-10 demonstrates the improvement achieved from the proposed method over the "state of the art" (April 27, 1998 in the ABRFC). Errors from Persistence have a lot of structure, this is to be expected for hourly rainfall especially in the Summer. The bias (or inaccuracy in spatial aggregate) in the rainfall at 1-hour lead times is evident from the surface plot for Persistence errors. The errors from 1-hour Advection do not show any improvement over Persistence, the error structure changes due to the translation of radar maps. The improvement from the combination of NWP-QPF and Advection ("Distributed Nimrod") appears to be minimal, when compared to Advection alone. The error plot for the Proposed model show that a significant portion of the structure has been modeled. Improvements over all the "state of the art" techniques, both in terms of the distributed and the aggregate skills, are strongly suggested from this plot.

**Figure 8-11: Precipitation Error (mm/hr) Contour Maps - Summer Storm "B"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)**

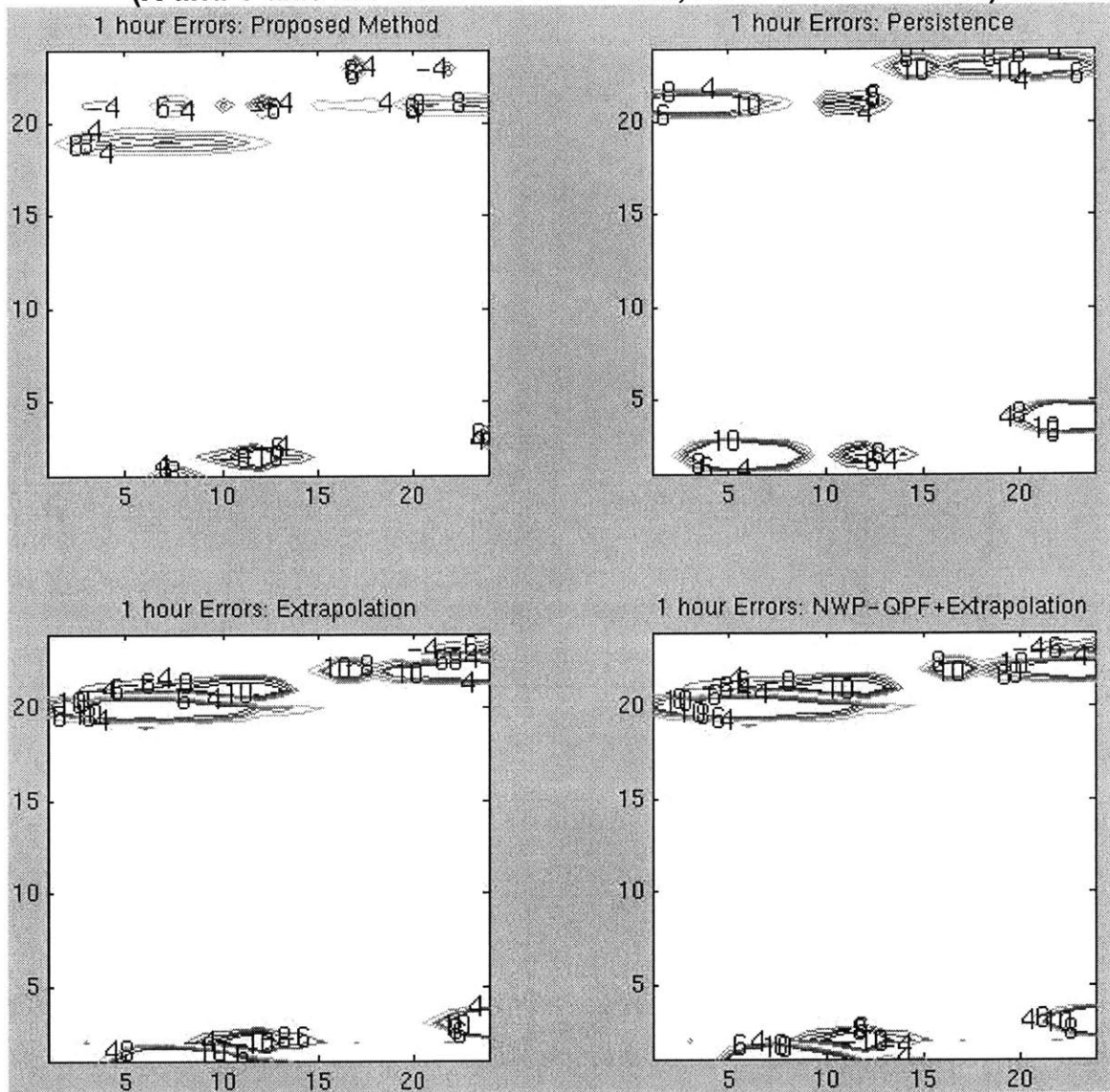


Figure 8-11 shows contour plots for the distributed errors (May 4-5, 1999 in the ABRFC). Visually, neither Extrapolation (GA00) nor the combination of Extrapolation with NWP-QPF (DN00) seems to have improved over Persistence (PERS) in terms of the residual error structures. The Proposed strategy (PROP) seems to have reduced the error structures compared to the existing methods.

**Figure 8-12: Precipitation Error (mm/hr) Surface Plots - Summer Storm "B"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)**

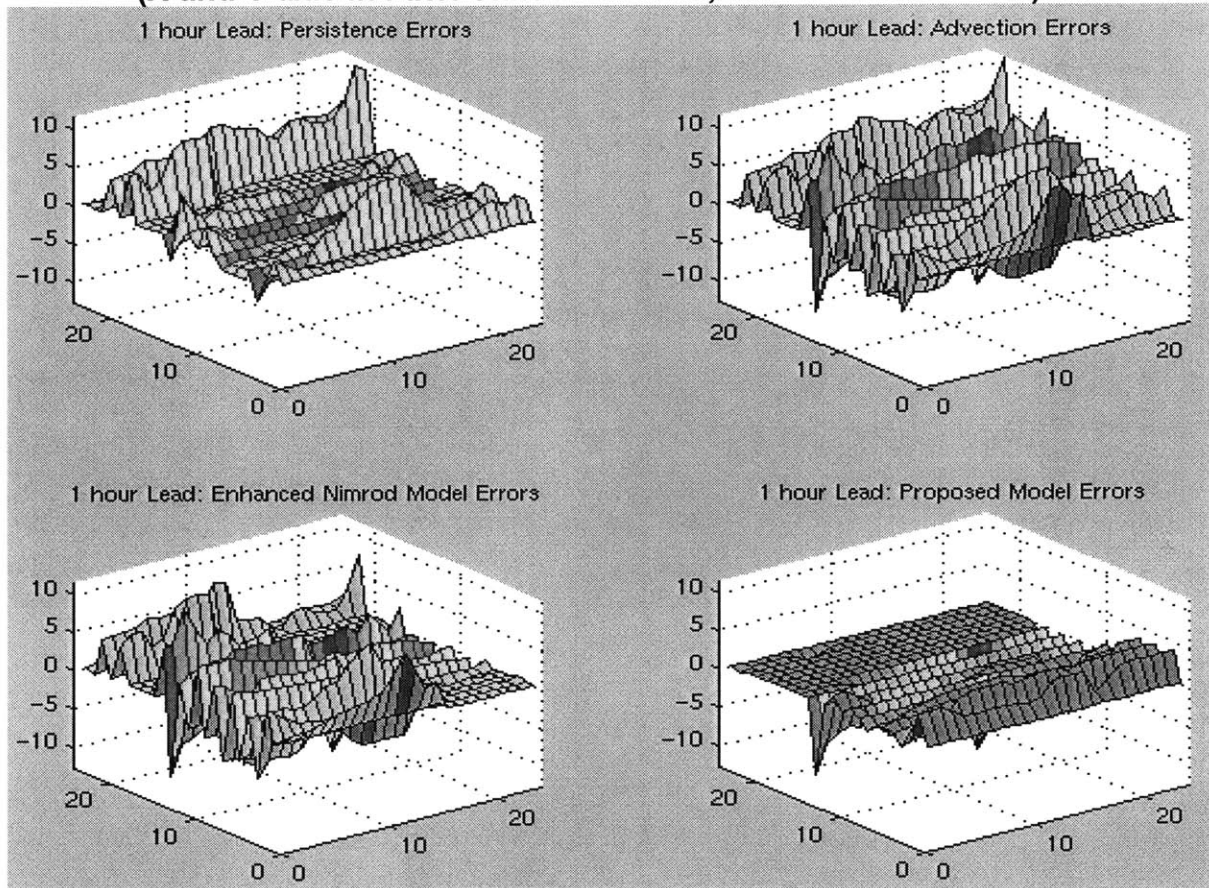


Figure 8-12 shows the surface plots for the errors from a summer storm (May 4-5, 1999 in the ABRFC). Results are similar to that obtained for the previous storm (Figure 8-10). Persistence retains significant error structures. 1-hour Advection does not improve visually over Persistence, but the error structure changes. The use of error corrected NWP-QPF by distributed Nimrod does seem to improve the aggregate skills marginally. The proposed strategy significantly improves over the existing methods, both in terms of the distributed and the aggregate errors structures.

**Figure 8-13: Precipitation Error (mm/hr) Contour Maps - Summer Storm "C"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)**

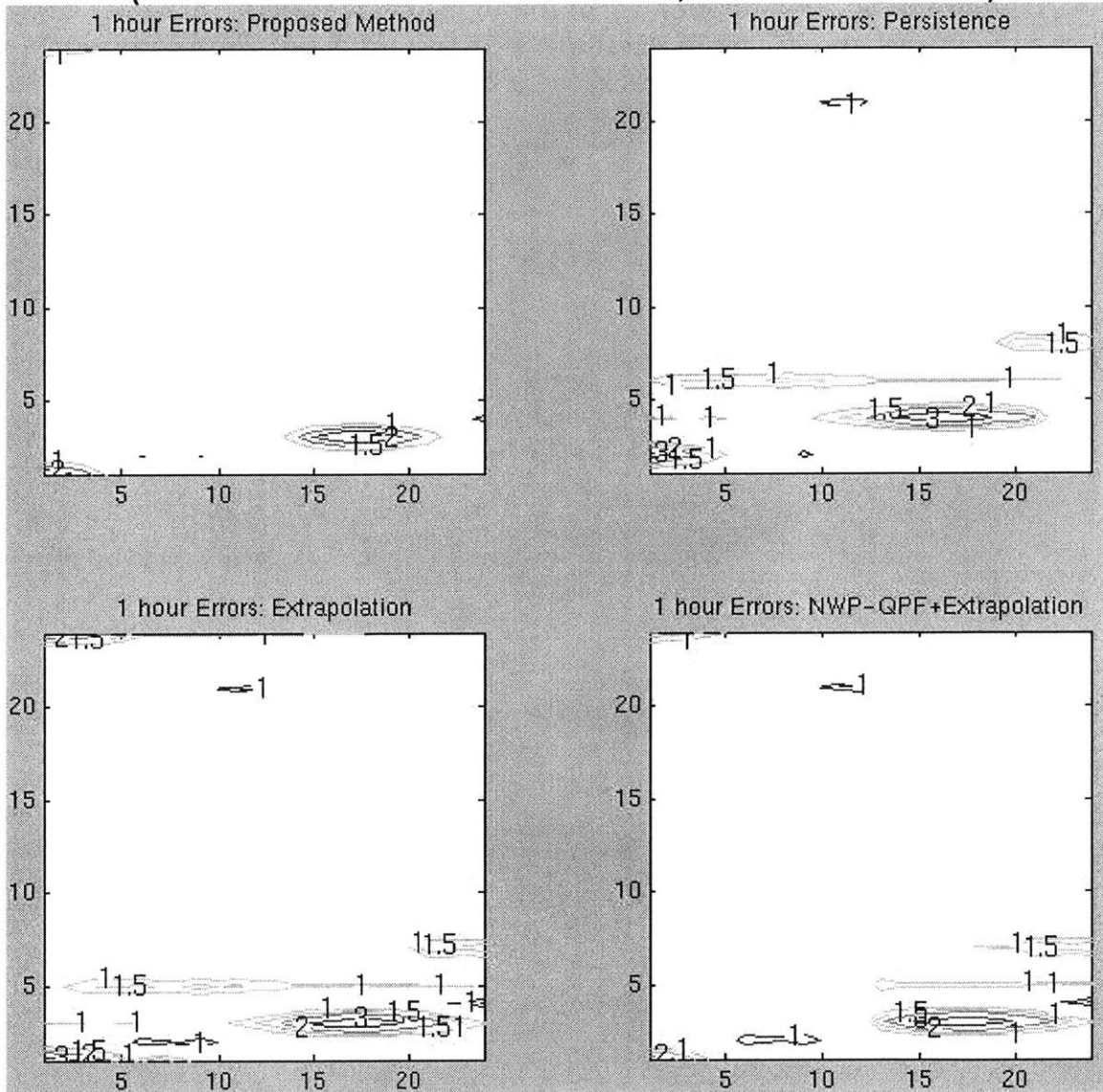


Figure 8-13 shows contour plots for the distributed errors (June 16, 1999 in the ABRFC). Visually, neither Extrapolation (GA00) nor the combination of Extrapolation with NWP-QPF (DN00) seems to have improved over Persistence (PERS) in terms of the residual error structures. The Proposed strategy (PROP) seems to have reduced the error structures compared to the existing methods.

**Figure 8-14: Precipitation Error (mm/hr) Surface Plots - Summer Storm "C"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)**

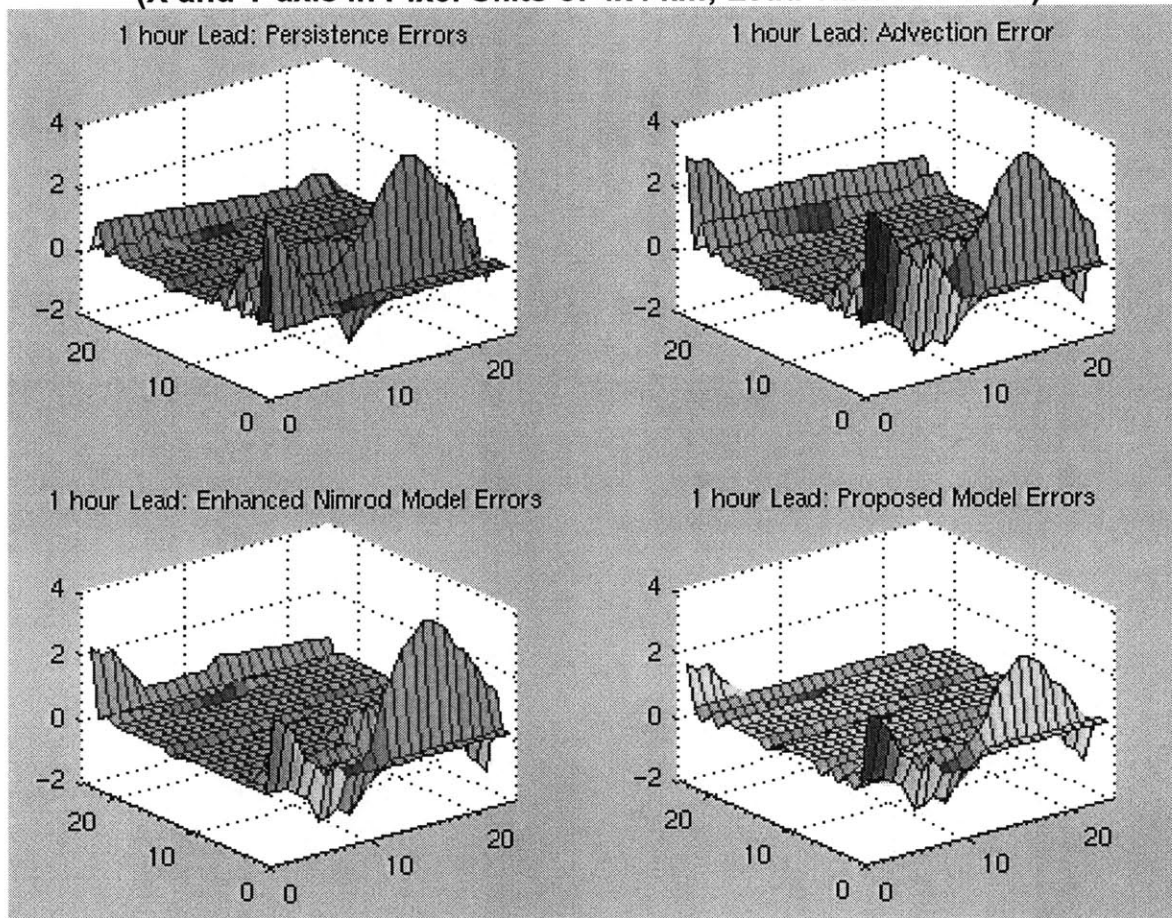


Figure 8-14 shows the surface plots for the errors from a summer storm (June 16, 1999 in the ABRFC). Results are similar to that obtained for the previous storms (Figures 8-10 and 8-12). Persistence retains significant error structures. 1-hour Advection does not improve visually over Persistence, but the error structure changes. The use of error corrected NWP-QPF by distributed Nimrod does seem to improve the aggregate skills. The proposed strategy significantly improves over the existing methods, both in terms of the distributed and the aggregate errors structures.

Figure 8-15: Precipitation Error (mm/hr) Contour Maps - Winter Storm "D"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)

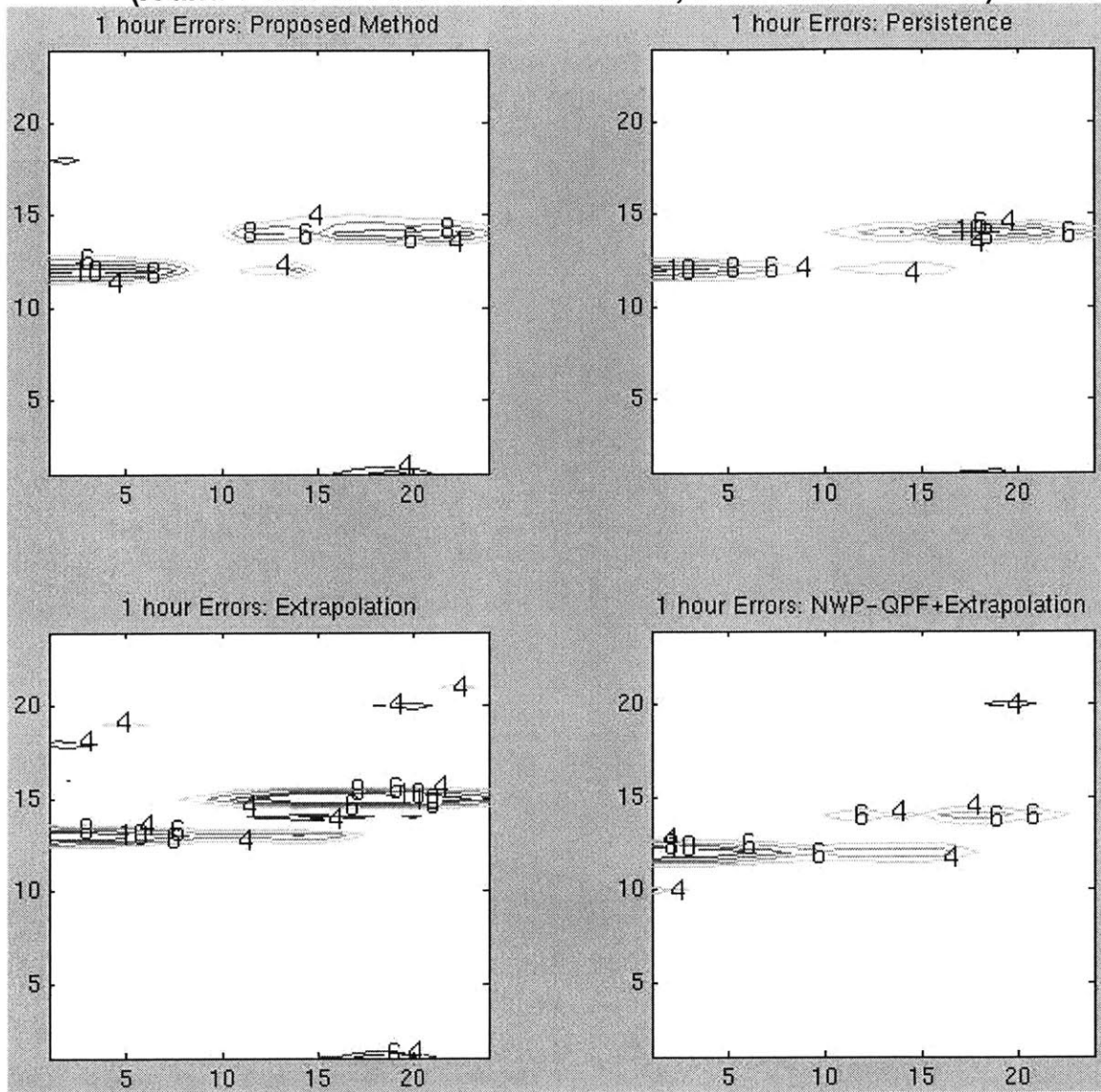


Figure 8-15 shows contour plots for the distributed errors (February 20, 1997 in the ABRFC). Visually, Extrapolation (GA00) does not appear to improve over Persistence (PERS) in terms of the residual error structures. The combination of Extrapolation with NWP-QPF (DN00) and the Proposed strategy (PROP) seem to have improved over both Persistence and Extrapolation. The performance of DN00 and PROP appear to be comparable, but PROP seems to be marginally better.

Figure 8-16: Precipitation Error (mm/hr) Surface Plots - Winter Storm "D"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)

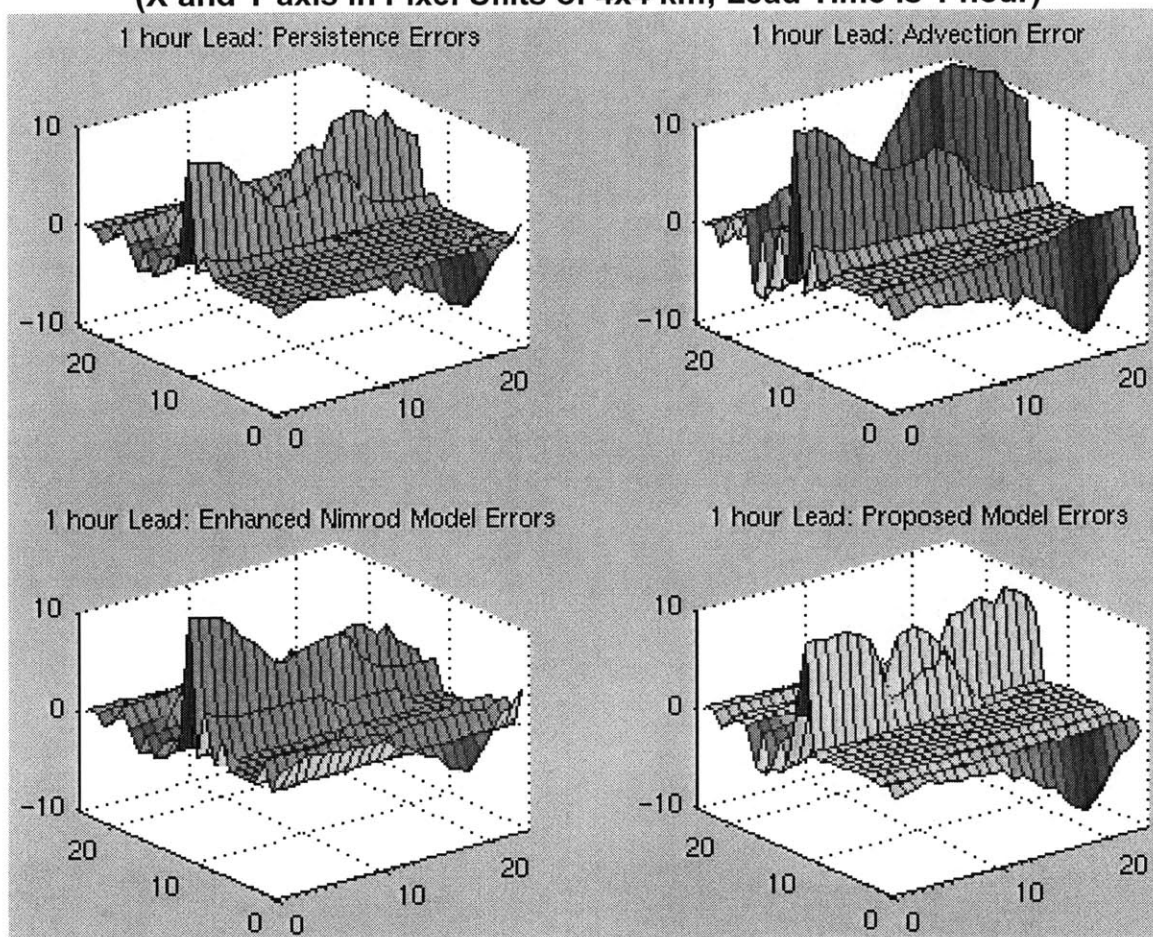


Figure 8-16 shows the surface plots for the errors from a winter storm (February 20, 1997 in the ABRFC). Results are different from the summer storms (Figure 8-10, 8-12 and 8-14). Persistence retains significant error structures. 1-hour Advection does not improve visually over Persistence, but seems to cause more error structures. The use of error corrected NWP-QPF by distributed Nimrod does seem to improve the aggregate skills over advection. The proposed strategy marginally improves over advection and distributed Nimrod, both in terms of the distributed and the aggregate errors structures. The overall improvement over persistence appears to be marginal, if any.

Figure 8-17: Precipitation Error (mm/hr) Contour Maps - Winter Storm "E"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)

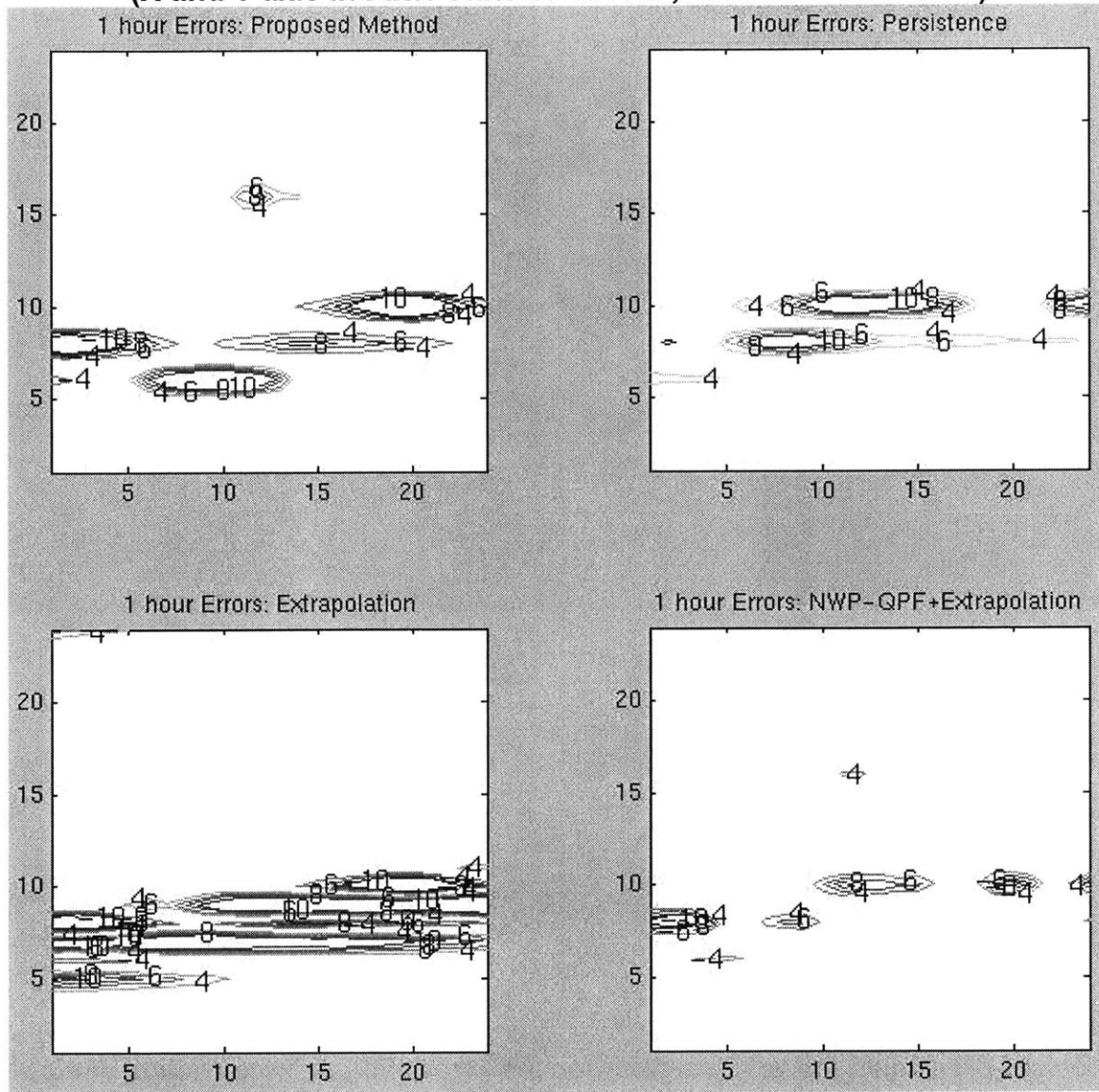


Figure 8-17 shows contour plots for the distributed errors (October 05, 1998 in the ABRFC). Visually, Extrapolation (GA00) does not appear to improve over Persistence (PERS) in terms of the residual error structures. The combination of Extrapolation with NWP-QPF (DN00) seems to have improved over both Persistence and Extrapolation. The Proposed strategy (PROP) appears to beat GA00, but does not appear to perform as well as DN00.

Figure 8-18: Precipitation Error (mm/hr) Surface Plots - Winter Storm "E"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)

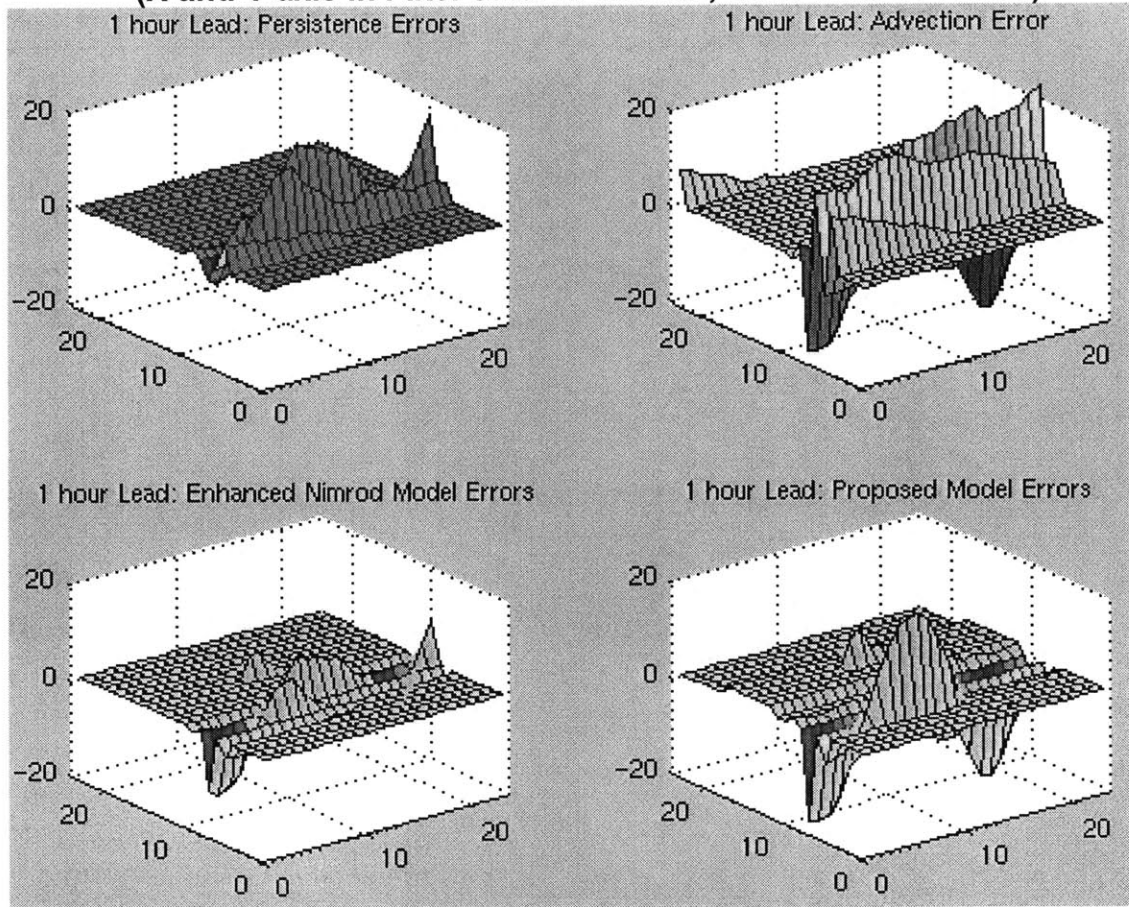


Figure 8-18 shows the surface plots for the errors from a winter storm (October 05, 1998 in the ABRFC). Persistence retains significant error structures. 1-hour Advection does not improve visually over Persistence, but seems to cause more error structures. The use of error corrected NWP-QPF by distributed Nimrod does seem to improve the aggregate skills over advection. The proposed strategy marginally improves over advection, but does not appear to improve over the distributed Nimrod, both in terms of the distributed and the aggregate errors structures. There seems to be some improvement over persistence overall.

Figure 8-19: Precipitation Error (mm/hr) Contour Maps - Winter Storm "F"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)

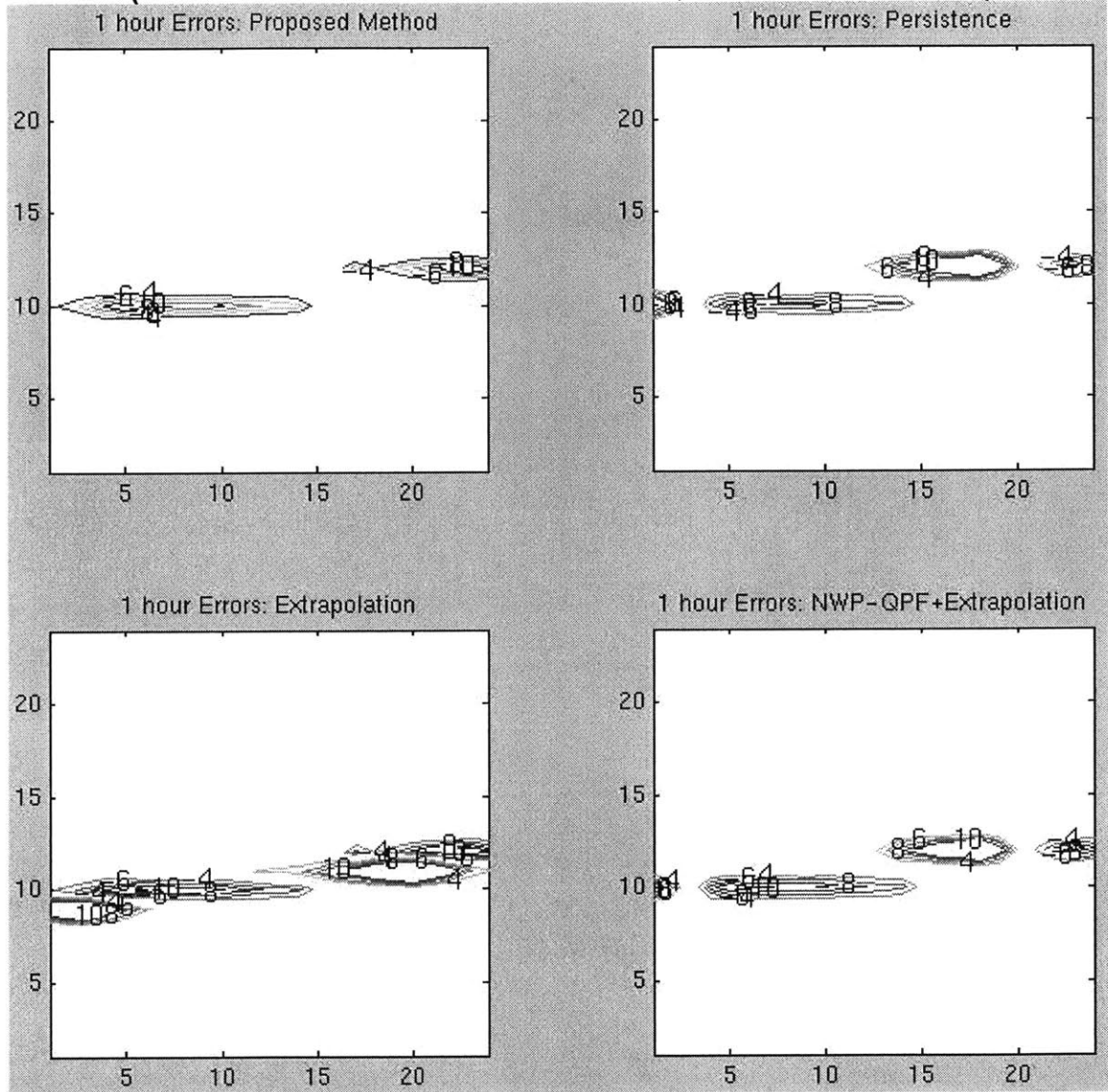


Figure 8-19 shows contour plots for the distributed errors (October 17-18, 1998 in the ABRFC). Visually, neither Extrapolation (GA00) nor the combination of Extrapolation with NWP-QPF (DN00) appear to improve over Persistence (PERS) in terms of the residual error structures. The Proposed strategy (PROP) seems to have improved over the existing methods in terms of the magnitude of the errors (see the reduced contour values), although not much is obvious from the shape of the contours in this figure.

Figure 8-20: Precipitation Error (mm/hr) Surface Plots - Winter Storm "F"
(X and Y axis in Pixel Units of 4x4 km; Lead Time is 1 hour)

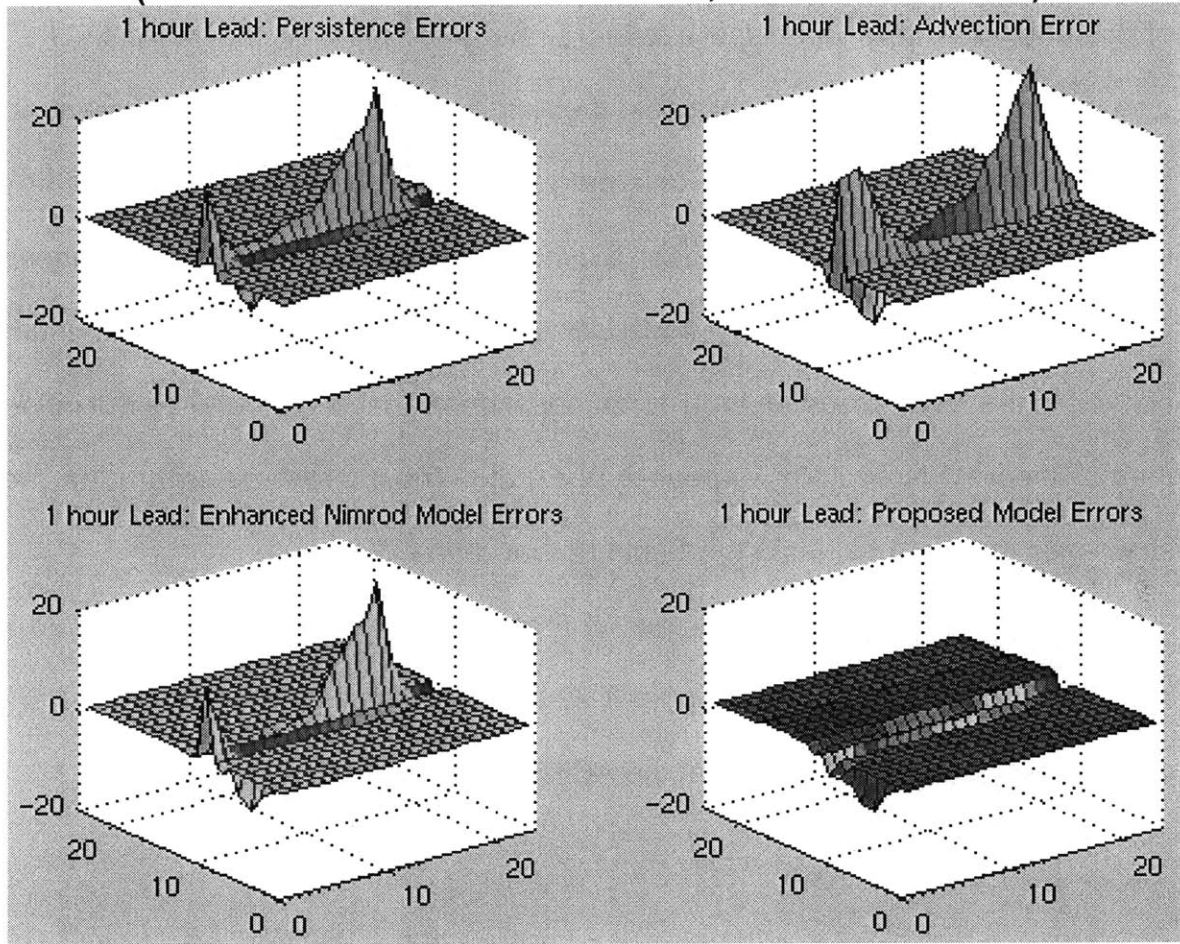


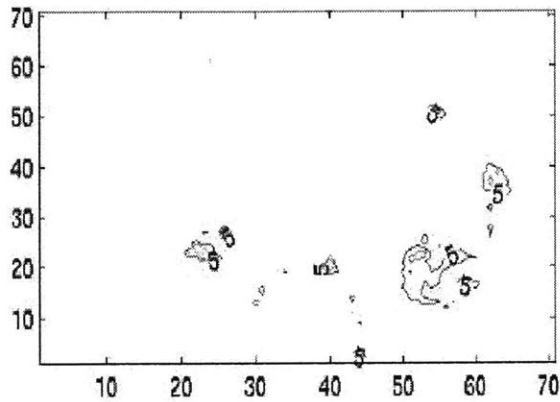
Figure 8-20 shows the surface plots for the errors from a winter storm (October 17-18, 1998 in the ABRFC). Persistence retains significant error structures. 1-hour Advection does not improve visually over Persistence. The use of error corrected NWP-QPF by distributed Nimrod does seem to marginally improve the aggregate skills over advection. The proposed strategy seems to improve significantly over all the existing methods, both in terms of the distributed and the aggregate errors structures.

Figures 8-9 through 8-20 summarize the QPF improvements from our research. In general, these clearly depict that the proposed strategy has been able to reduce the error structures more than any existing strategy for most of the storms. The simple disaggregation strategy which uses the large scale information inherent in error corrected NWP-QPF appears to have the significant value in the Winter. In the Winter, modeling the localized and residual structures using NWP forecasts of atmospheric variables through ANN is only occasionally useful. In the Summer, the information content in the NWP model outputs for atmospheric variables, as well as the influence of the nonlinear ANN strategy, appear to dominate. The improvement from the proposed strategy is therefore higher in the Summer than in the Winter.

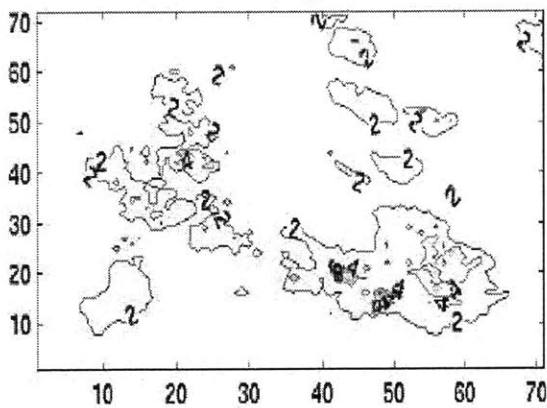
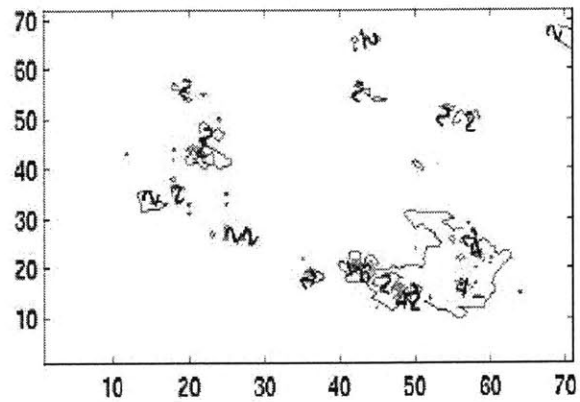
Figure 8-21 shows an example QPF and pQPF result using uncalibrated Stage I radar data at Topeka, Kansas. Figure 8-22 and 8-27 show QPF and pQPF errors using calibrated Stage III radar rainfall at the ABRFC in Oklahoma.

Figure 8-21: Contour plots of Measured Radar Rainfall, Proposed QPF, and pQPF (X and Y axes are in pixel units: 4x4 km; Rainfall is in mm/hr; Lead time is 1 hour)

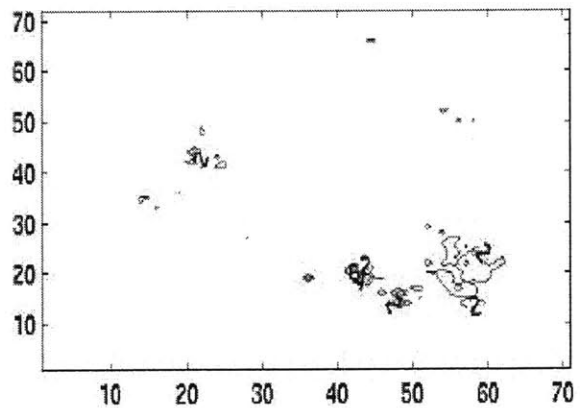
Measured Radar Rainfall (mm/hr)



Mean of Forecast Rainfall (mm/hr)



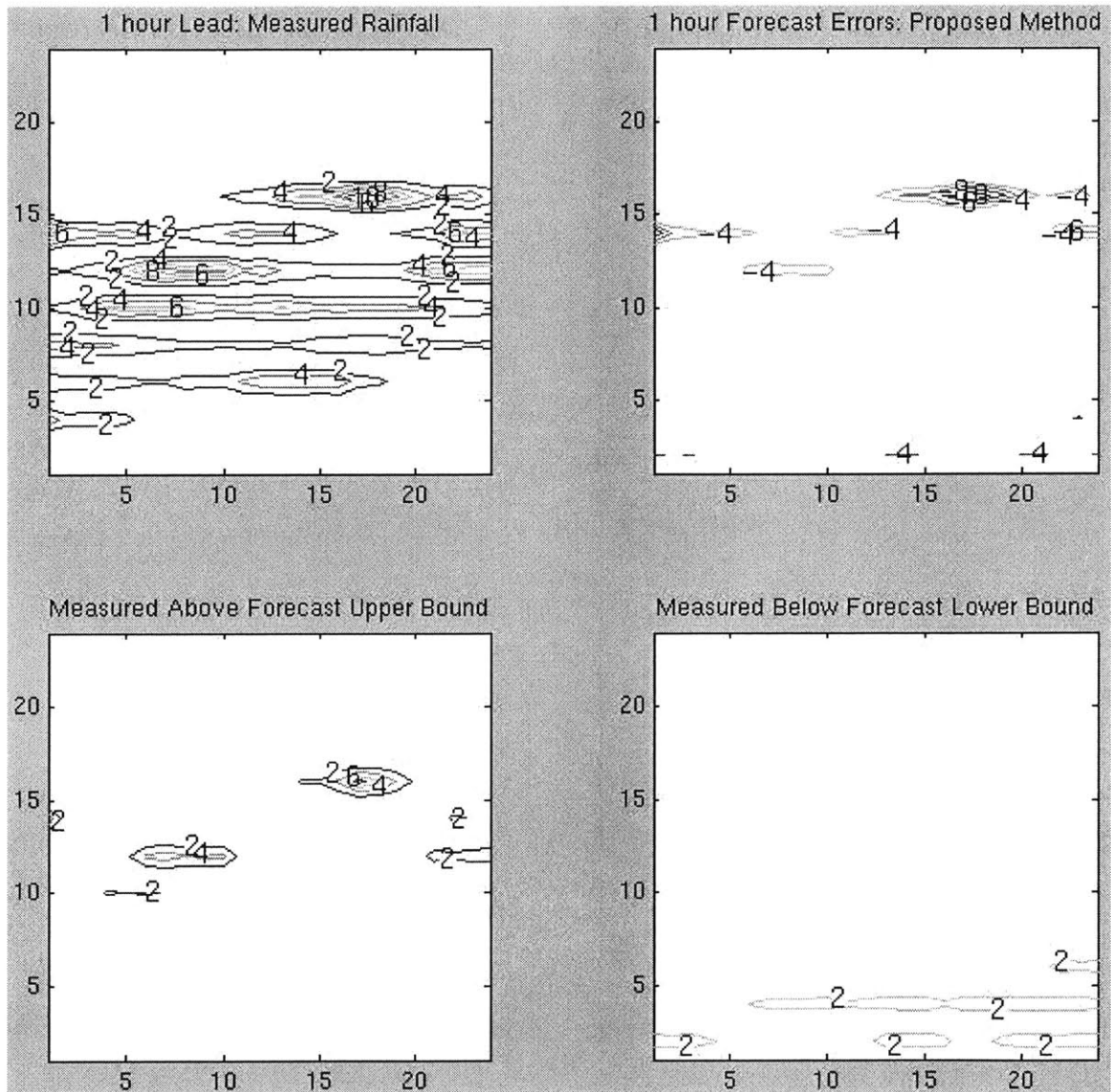
Upper Bound of Forecast Rainfall (mm/hr)



Lower Bound of Forecast Rainfall (mm/hr)

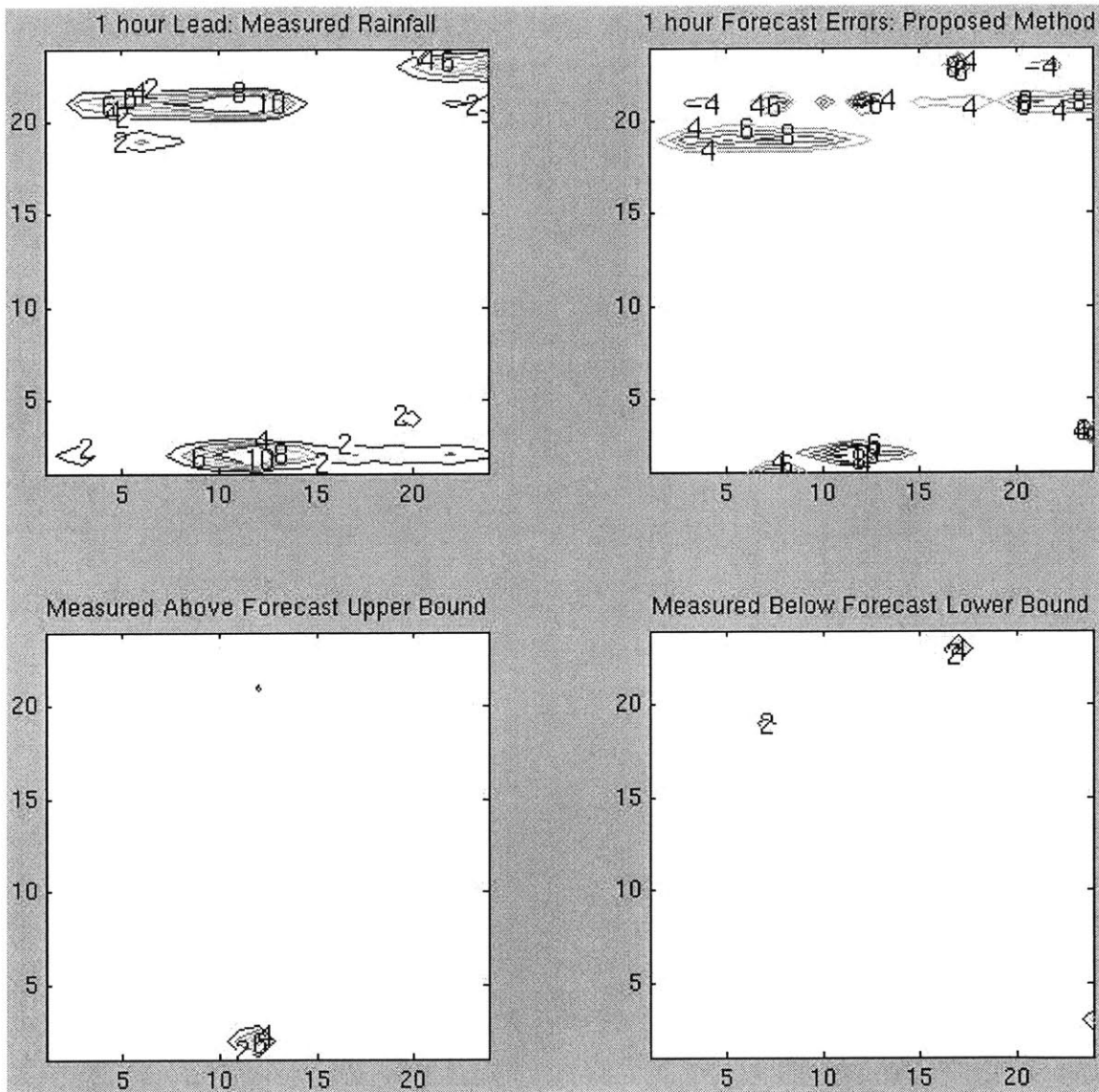
The 1-hour lead contour plots, for a precipitation event (May 5-6, 1998) in Topeka-Kansas. Figure 8-21 shows that the forecast map from the Proposed model agrees with the observed, but appears to be more dispersed in space. This dispersion is more for the upper confidence bound, as expected. However, the lower confidence bound shows that the areas of rainfall concentration matches the forecast values.

Figure 8-22: Contour plots of Stage III Radar Rain, Proposed QPF & pQPF Errors (X and Y axes are in pixel units: 4x4 km; Rainfall is in mm/hr; Lead time is 1 hour)



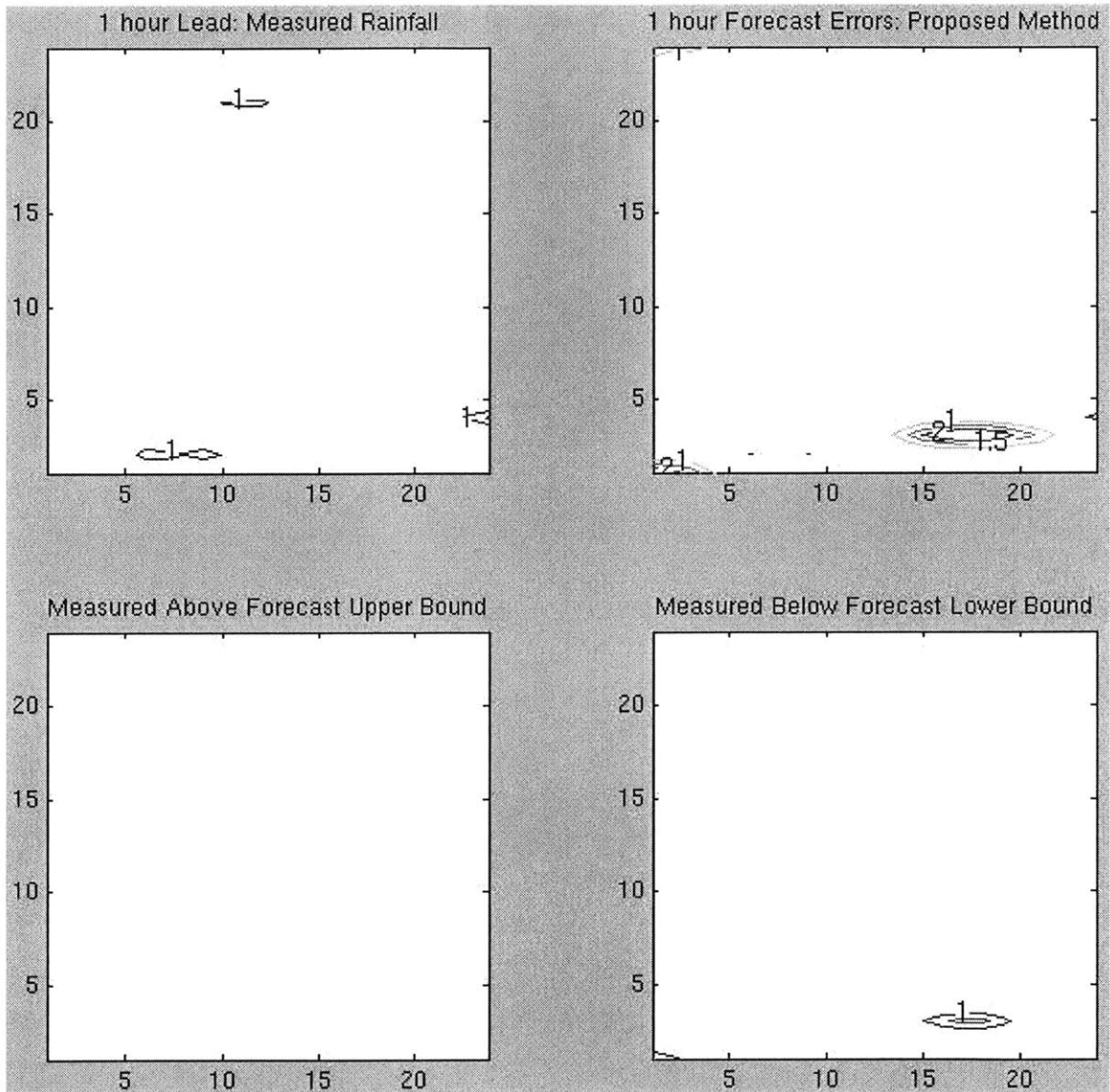
The 1-hour lead contour plots, for "Summer Storm A" (April 27, 1998) in ABRFC, Oklahoma. Figure 8-22 shows the structure in the measured Stage III precipitation. This is compared with the errors in QPF from the proposed strategy, as well as the errors in pQPF (i.e., the upper and lower confidence bounds). Lack of structures in the errors indicate good forecasts.

Figure 8-23: Contour plots of Stage III Radar Rain, Proposed QPF & pQPF Errors (X and Y axes are in pixel units: 4x4 km; Rainfall is in mm/hr; Lead time is 1 hour)



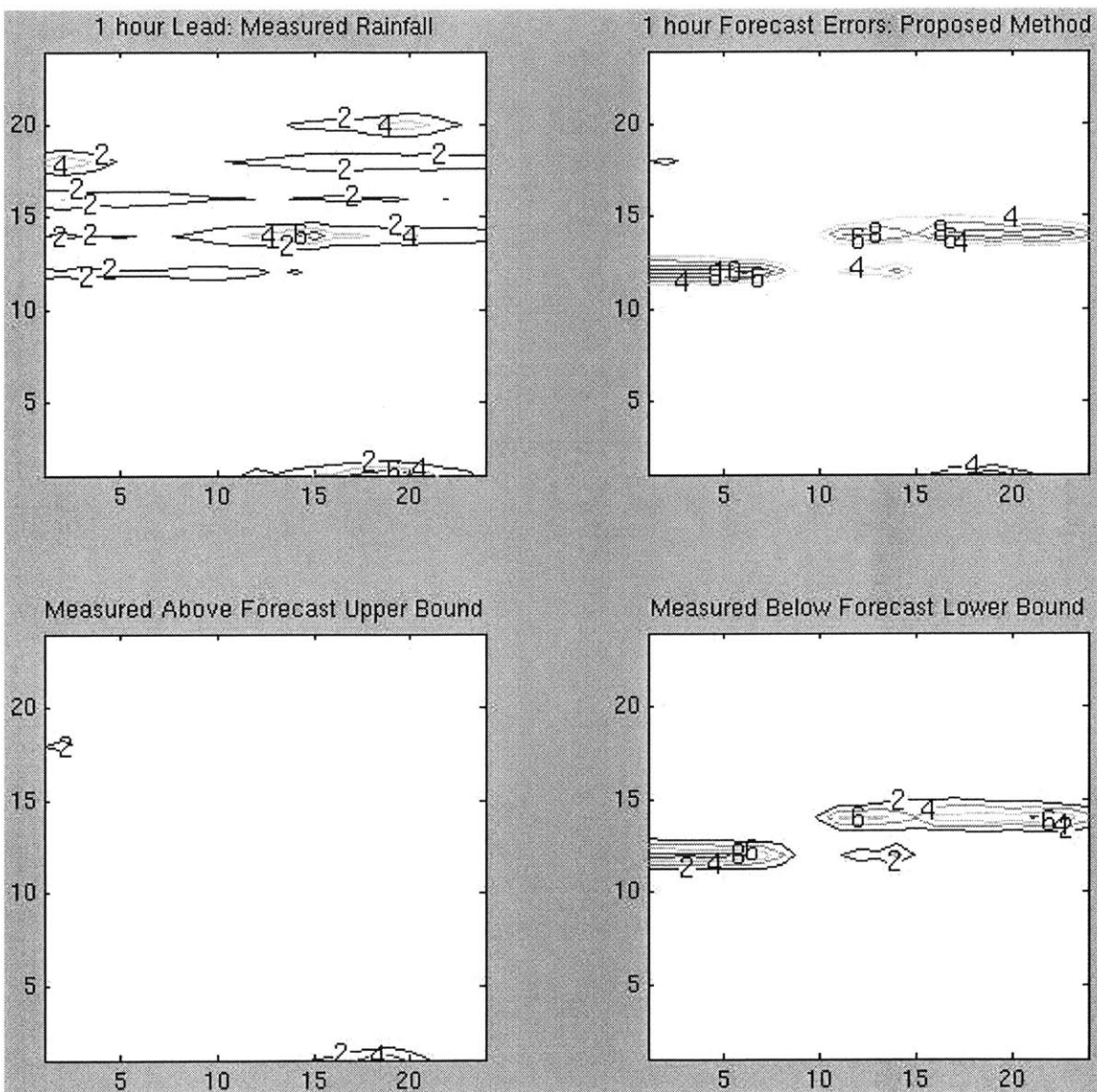
The 1-hour lead contour plots, for “Summer Storm B” (May 4-5, 1999) in ABRFC, Oklahoma. Figure 8-23 shows the structure in the measured Stage III precipitation. This is compared with the errors in QPF from the proposed strategy, as well as the errors in pQPF (i.e., the upper and lower confidence bounds). Lack of structures in the errors indicate good forecasts.

Figure 8-24: Contour plots of Stage III Radar Rain, Proposed QPF & pQPF Errors (X and Y axes are in pixel units: 4x4 km; Rainfall is in mm/hr; Lead time is 1 hour)



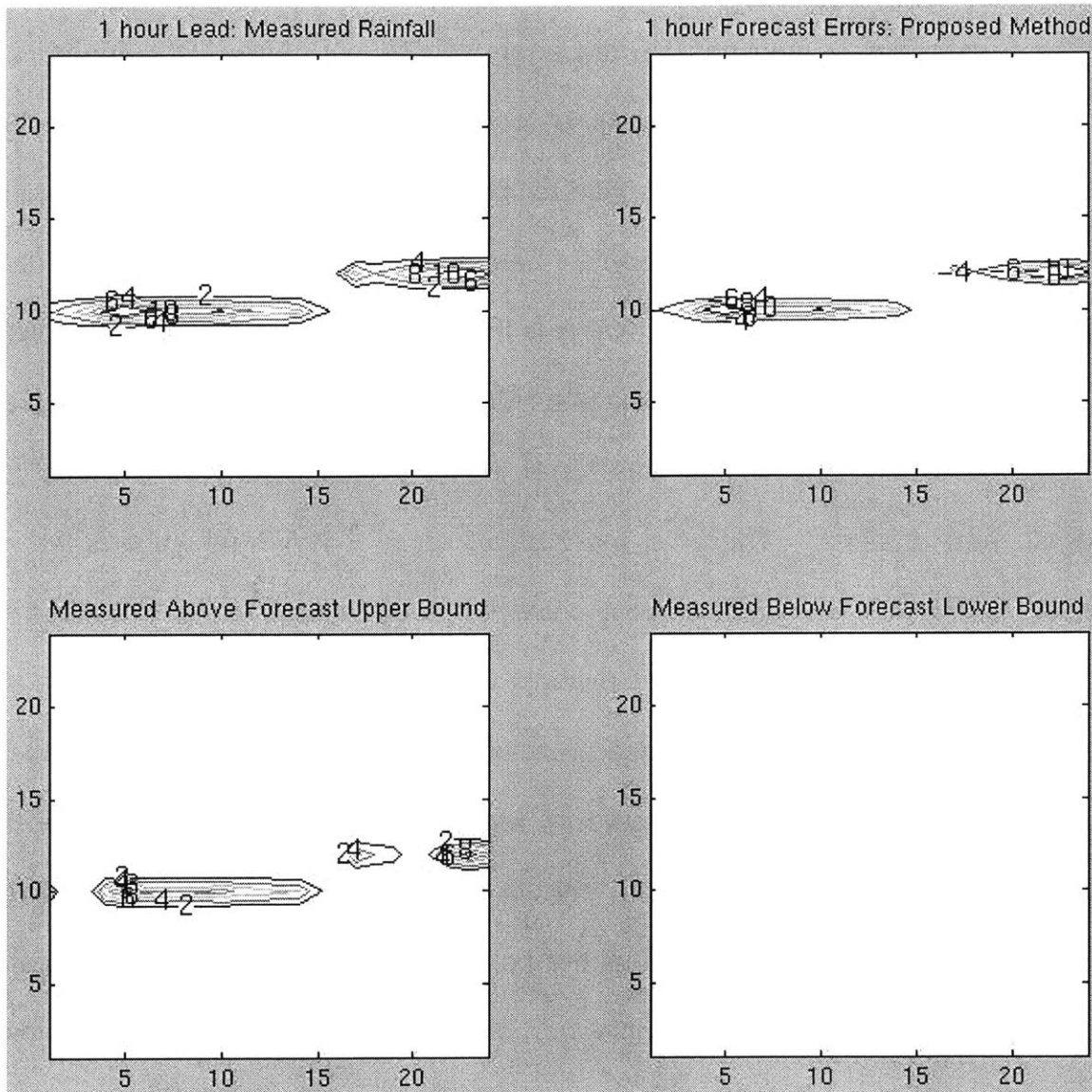
The 1-hour lead contour plots, for "Summer Storm C" (June 16, 1999) in ABRFC, Oklahoma. Figure 8-24 shows the structure in the measured Stage III precipitation. This is compared with the errors in QPF from the proposed strategy, as well as the errors in pQPF (i.e., the upper and lower confidence bounds). Lack of structures in the errors indicate good forecasts.

Figure 8-25: Contour plots of Stage III Radar Rain, Proposed QPF & pQPF Errors (X and Y axes are in pixel units: 4x4 km; Rainfall is in mm/hr; Lead time is 1 hour)



The 1-hour lead contour plots, for "Winter Storm D" (February 20, 1997) in ABRFC, Oklahoma. Figure 8-25 shows the structure in the measured Stage III precipitation. This is compared with the errors in QPF from the proposed strategy, as well as the errors in pQPF (i.e., the upper and lower confidence bounds). Lack of structures in the errors indicate good forecasts.

Figure 8-27: Contour plots of Stage III Radar Rain, Proposed QPF & pQPF Errors (X and Y axes are in pixel units: 4x4 km; Rainfall is in mm/hr; Lead time is 1 hour)



The 1-hour lead contour plots, for "Winter Storm F" (October 17-18, 1998) in ABRFC, Oklahoma. Figure 8-27 shows the structure in the measured Stage III precipitation. This is compared with the errors in QPF from the proposed strategy, as well as the errors in pQPF (i.e., the upper and lower confidence bounds). Lack of structures in the errors indicate good forecasts.

9.0 Discussions and Performance Analyses

Ebert and McBride (2000) mention that a way to understand the behavior of QPF could be to decompose and analyze the errors at different scales. We analyzed the errors from the individual components of our proposed QPF strategy. The error statistics were determined at three spatio-temporal scales, (a) NWP resolutions, or 6 hour and 48 km, (b) Hourly spatial means, 1 hour and 48 km, and (c) Radar resolutions, 1 hour and 4 km. We used the three events in the Summer, and the three in the Winter in Oklahoma, where Stage III radar data and NWP-Eta model outputs were available.

We considered conditional bias (error in the overall mean) as a measure of aggregate error statistics. Thus the measure of error at NWP resolutions was the magnitude of the difference between the observed precipitation and that predicted from our proposed approach for a total of 6 hours over the 48-km NWP grids. The error statistic used for the hourly spatial mean resolution was identical, except that this was calculated for each hour over the same 48 km grids. At radar resolutions, the error statistic was the mean squared errors calculated from the errors at individual radar pixels (i.e., 1 hour and 4 km). Note that the behavior of the error statistics at spatially aggregate scales might not be a simple sum of the behavior at distributed scales. In other words, the errors in QPF could be scale dependent, this is suggested but not proven in Figures 4-2 and 4-3. As a simple example of scale dependent errors in this context, it is conceivable that NWP models predict the 6 hour, 48-km precipitation in the Winter relatively well, but the error introduced during the disaggregation process causes the distributed error statistics to be significant.

We have tabulated the results for the Summer (Table 9-1) and the Winter (Table 9-2), as the percent improvement over baseline persistence (with negative values indicating that the results were worse than persistence). We noted the errors as each additional component of our QPF strategy was added. Given the inherent variability of precipitation between and within events, we do expect and obtain variability in the results. However, a closer inspection does reveal some interesting patterns.

We used results from calibrated Stage III radar data and NWP-Eta model outputs for the analyses in this section (i.e., Tables 9-1 and 9-2). The average distributed skills have been quantified through $1/\text{NMSE}$, which considers all 4x4 radar pixels in the hourly verification time steps. These have shown earlier in Figure 8-3 (all seasons combined), Figure 8-5 (Winter), and Figure 8-7 (Summer). The spatially aggregate errors for the 48 km NWP grids at 1-hour forecast lead times, as well as the overall aggregate errors at NWP scales (48 km grid average and 6 hour cumulative) are shown in this section.

9.1 Summer Precipitation Events

Figures 8-7 and Table 9-1 show the average errors for the 3 precipitation events in the Summer, in the ABRFC in Oklahoma, where Stage III radar data was available.

The average distributed skills (4 km, 1-hour) show the poor performance from Persistence, 1-hour Advection, as well as NWP-QPF combined with 1-hour Advection. The overall strategy, which models high resolution Lagrangian evolution and the residual structures in addition to the other components, perform better at all lead times. The distributed skills of the various QPF strategies in the Summer are quantified in Figure 8-7.

The aggregate errors at NWP scales (48 km, 6 hour), from the component models for distributed QPF, are shown for the Summer in Table 9-1. 1-hour Advection improves over Persistence in terms of the aggregate errors. When the results from 1-hour Advection are scaled with NWP-QPF, the aggregate errors do not show any improvement. This is expected, as previous researchers and our data analyses indicate that NWP-QPF does not have significant information in the Summer. The aggregate (i.e., spatial mean) performance of Advection scaled with NWP-QPF actually worsens relative to pure Advection, however both improve over Persistence. The ANN based localized evolution method performs worse than all other methods, including Persistence and Advection, in terms of the spatial mean. This is caused by the relatively sharp decay of skills at higher lead times, which is expected. The ANN model learns the distributed and localized evolution through complex function approximation, the validity of which could be expected to decay rather sharply with lead time. The proposed

method, which models residual structures in addition to the other components, performs the best overall in terms of distributed and aggregate skill measures.

As we have discussed, the performance of the ANN component, which attempts to capture localized evolution in a Lagrangian frame of reference, is good for lower lead times (e.g., 1-3 hour) and in terms of distributed skill measures. This improvement is more evident in the Summer, as shown earlier. This is to be expected. First, localized convective effects are expected to dominate in the Summer. Second, data analysis at aggregate scales demonstrate that the information content in NWP atmospheric outputs are significant in the Summer. Third, we have shown that nonlinear relations predominate in the Summer, especially in the relation of precipitation and NWP atmospheric outputs. However, at higher lead times the skills decay sharply, especially in terms of the spatial aggregate. This is also expected, as the conditions in the Summer are expected to be more dynamic and the validity of the evolution modeled by the ANN is not expected to prolong to extended lead times. This is shown in Table 9-1, which shows the errors in the estimate of the spatial aggregate as a function of lead time in hours. The average errors in the ANN based evolution model increases sharply after about 4 hour lead time. The 1-hour Advection component performs relatively better, and always beats Persistence. Scaling Advection with NWP-QPF rarely improves the estimate of the hourly spatial aggregate over Advection in the Summer. The overall proposed method performs better than all of these baseline techniques and/or component processes in terms of the hourly, spatially aggregate errors from the hourly, distributed forecasts.

Table 9-1: Performance of QPF Component Processes in the SUMMER

	Radar Extrapolation	Large Scale Physics	Local ANN Evolution	Residual Structures
	(E)	(E+P)	(E+P+L)	(E+P+L+R)
(Conditional "Bias") 6 hr, 48 km				
	29	7	-48	46
(Conditional "Bias") 1 hr, 48 km				
Lead: 1 hr	0	50	20	40
Lead: 2 hr	50	-50	9	50
Lead: 3 hr	40	40	0	40
Lead: 4 hr	50	-43	5	55
Lead: 5 hr	33	30	-78	67
Lead: 6 hr	28	11	-155	70
(Mean Squared Errors) 1 hr, 4 km				
Lead: 1 hr	-50	-25	60	70
Lead: 2 hr	30	-26	55	65
Lead: 3 hr	30	32	69	77
Lead: 4 hr	30	-27	39	70
Lead: 5 hr	45	33	8	89
Lead: 6 hr	46	40	-61	92

Table 9-1 shows the percent improvement over persistence by a successive addition of the QPF component processes. The results are averaged over 3 Summer Events in Oklahoma, using data from Stage III NEXRAD and NWP-Eta. The "conditional bias" measures the error in the spatial mean. The bias and MSE are averaged over the three precipitation events. Negative quantities indicate that the performance was worse than persistence.

9.2 Winter Precipitation Events

Figures 8-5 and Table 9-2 show the average errors for the 3 precipitation events in the Winter, in the ABRFC in Oklahoma where Stage III radar data was available.

Figure 8-6 demonstrates that there is skill in the proposed model, slightly more than that in the Summer (Figure 8-8), in terms of the number of times the forecast and the observed values lie within prescribed thresholds. The relatively small decay of this skill in the Winter as evident from Figure 8-6 also indicate that the confidence bounds are able to predict the variation of precipitation with lead time better than that in the Summer. This is to be expected as the conditions are expected to be more stable in the Winter. The relatively higher skill compared to Summer and the lesser rate of decay agrees with the observations from previous researchers.

The proposed method shows minimal improvement over existing methods in terms of the average distributed skills (4 km, 1-hour), as quantified in Figure 8-5 and Table 9-2. Previous researchers often indicate that precipitation is easier to predict in the Winter (Antolik, 2000). Thus, Figure 8-5 represents an apparent contradiction, which is explored further in this section.

There are several causes that contribute to the lack of distributed skills in the Winter, as apparent from Figure 8-5. First, the performance of baseline Persistence and 1-hour Advection in the Winter is no better than that in the Summer. This is to be expected, as the typical decorrelation times for rainfall structure is of the order of an hour both in the Winter and in the Summer. Mecklenburg et al. (2000) found the structures persist in a Lagrangian sense for relatively small rainfall areas only till 40 minutes for convective storms, and slightly higher for stratiform. The former is expected

to dominate in the Summer, and the latter in the Winter. In either case, the persistence of rainfall structures, in an Eulerian or a Lagrangian frame of reference, typically does not extend to more than an hour.

Second, the 1-hour Advection component performs rather poorly on the average, significantly worse than even baseline Persistence. This is clearly seen in the overall MSE (mean squared errors), calculated by considering all the radar pixels (Table 9-2) in a verification time step. This effect is significant enough to cause the performance from 1-hour Advection to be worse than Persistence in terms of the hourly spatial aggregate (Table 9-2), and even in terms of the 6-hour spatial aggregate (Table 9-2). An investigation of the Advection maps revealed that this is primarily caused by the forecast rainfall areas missing the observed rainfall areas. We call this the “problem of the missed cells”, this is seen to be more predominant for the Winter events than for the Summer. This led us to believe that the velocity scales applied to the hourly rainfall maps, estimated from global correlation in space using lagged rain maps and an exponential smoothing formulation, were more error prone in the Winter.

We discovered that while the 1-hour Advection often performed better in the Winter than in the Summer (see Figures 5-1 and 5-2), the average performance was overwhelmed by the occurrence of “missed cells”. Based on this, we hypothesize that for 1-hour Advection in the Winter, spatial correlation alone is not a good way to estimate the velocity scales, and there might be a need to look for larger scale rain area movements and possibly wind profile information. However, this hypothesis needs to be verified by detailed analysis of translation patterns for Winter events, and the possible causes need to be explored further.

The couple of causes discussed earlier indicate that radar extrapolation alone, whether from Persistence or from Advection, would not contain much information for distributed, hourly QPF in the Winter. Next, we explore the information content in NWP model outputs relevant to the prediction of distributed precipitation in the Winter.

Previous researchers indicate that the QPF from NWP is better predicted in the Winter than in the Summer (Mesinger, 1996), which could be one cause for better QPF in the “cooler seasons” (Antolik, 2000). This was independently reinforced by the results of our data analyses at aggregate NWP scales, as described earlier. We showed that the information content in NWP-QPF, relevant to aggregated precipitation (from calibrated radar), was significantly higher in the Winter than in the Summer. When the distributed errors from the results of the various QPF strategies are aggregated in space and time to 48 km and 6 hours, the error corrected NWP-QPF also performs the best (Figure 9-5). However, the process of distributing this rainfall forecast in time and space introduces significant errors, as quantified by Figures 9-3 and 9-4. Antolik (2000) mentions that among the lessons learnt from the operational Model Output Statistics (MOS) is that QPF from NWP is not the most important predictor for rainfall at a point. Thus the information content in NWP-QPF in the Winter is relatively high for the aggregate, but low for distributed skills. This is the third cause that contributes to the lack of distributed QPF skills in the Winter.

The fourth and final cause for the low distributed skills in the Winter is the lack of any significant information content in the NWP atmospheric outputs. This is indicated by the results of our data analyses described earlier, and contrasts with the situation in the Summer. For the Winter, there is neither significant information in the NWP

atmospheric outputs, nor any significant nonlinearity that could be easily extracted by the use of nonlinear function approximation tools like ANN.

Overall, it appears that the low average distributed skill in the Winter, from the 3 storms that we analyzed, is caused by the lack of skills in hourly radar extrapolation, the advection forecast often missing the observed rain areas, the errors introduced during the space-time distribution of the grid-averaged cumulative QPF from NWP, and the lack of significant information content in the NWP model outputs other than QPF.

Table 9-2: Performance of QPF Component Processes in the WINTER

	Radar Extrapolation	Large Scale Physics	Local ANN Evolution	Residual Structures
	(E)	(E+P)	(E+P+L)	(E+P+L+R)
(Conditional "Bias") 6 hr, 48 km				
	-58	26	-45	9
(Conditional "Bias") 1 hr, 48 km				
Lead: 1 hr	0	37	38	50
Lead: 2 hr	-3	9	-50	10
Lead: 3 hr	-130	30	-130	0
Lead: 4 hr	0	15	-67	5
Lead: 5 hr	-300	-12	-160	-60
Lead: 6 hr	-108	55	-45	0
(Mean Squared Errors) 1 hr, 4 km				
Lead: 1 hr	-120	1	-25	-30
Lead: 2 hr	0	0	5	35
Lead: 3 hr	-115	23	-2	3
Lead: 4 hr	-2	-10	6	25
Lead: 5 hr	-105	-28	9	2
Lead: 6 hr	-111	26	5	42

Table 9-2 shows the percent improvement over persistence by a successive addition of the QPF component processes. The results are averaged over 3 Winter Events in Oklahoma, using data from Stage III NEXRAD and NWP-Eta. The "conditional bias" measures the error in the spatial mean. The bias and MSE are averaged over the three precipitation events. Negative quantities indicate that the performance was worse than persistence.

9.3 All Seasons Combined

When the skills or errors are averaged for all 6 precipitation events (3 in the Summer, and 3 in the Winter), the overall combined model (i.e., the proposed QPF strategy) performs significantly better than both the existing methods and the individual components of the combined model itself. This holds true in terms of both the aggregate and the distributed skills.

Figure 8-3 shows that the proposed QPF strategy beats all existing models in terms of the distributed skills for rainfall areas. Figure 8-4 shows similar improvements in terms of the forecasts and observed values lying within prescribed thresholds. These are reinforced by the results from Summer and Winter events. Figures 8-1 and 8-2 demonstrate that the overall strategy optimally combines the skills from the individual components. The overall strategy improves over the individual components in terms of both the distributed skills (Figure 8-1) and the aggregate skills (Figure 8-2). Figures 8-10 and 8-11 demonstrate the improvement achieved over existing techniques in terms of the distributed error structures, for two example precipitation events.

The improvement in QPF, demonstrated through quantitative measures and visual indications, is a contribution of this research. We demonstrate that the information content in the NWP model outputs, both QPF and atmospheric variables, can be used for improving distributed QPF at 1-6 hour leads. We also demonstrate that an intelligent decomposition of the problem to take best advantage of the process physics and data dictated ANN tools result in significant QPF improvement, especially when the overall hybrid modeling strategy is optimized to make the best use of the available information.

10.0 Summary

For the radar rainfall, both the spatial mean and the variance around this mean were seen to be non-stationary. Limited information was retained by persisting the structures for 1 hour lead time or more. NWP-QPF was seen to have more information content in the Winter, but the forecasts of the atmospheric variables were more pertinent in the Summer. Nonlinear relations were found to dominate in the Summer, but not in the Winter.

Advection, or translation based on estimation of velocity scales, was the only component found to have value for extrapolation using radar data alone, as per Grecu and Krajewski (2000). Previous researchers have found limited value in complex or higher order strategies for this component (Browning and Collier, 1989). The estimates of velocity scales from complex ANN strategies did not improve estimates for Grecu and Krajewski (2000). In our investigative studies, linear regression or nonlinear strategies performed worse than advection based on spatial correlation. This 1-hour Advection component typically performed marginally better in the Winter than in the Summer. However, the problem of missed rainfall cells occurred more frequently in the Winter, this overwhelmed the average distributed skill measure. On the whole, the gain in QPF skills from 1-hour Advection was marginal.

NWP-QPF was found to have significant information content for aggregate forecasts, especially in the Winter. For distributed forecasts, the information content in NWP-QPF was found to be rather poor, both in the Winter and in the Summer. This agrees with the lessons learnt from MOS, which show that QPF from NWP is not the

most important predictor for rainfall at a point (Antolik, 2000). The relative skills from this component was typically higher for lead times of 4-6 hours.

Modeling localized and high resolution precipitation evolution was found to have value only if these were performed in a Lagrangian frame of reference, and when NWP forecasts of atmospheric variables were used. For this component, nonlinear ANN based models were found to be more skillful than linear modeling strategies. This component performed the best in terms of distributed skills and at shorter (1-3 hour) lead times, especially in the Summer.

The final component of the proposed hybrid QPF strategy, which modeled the residual structures at both aggregate and distributed scales by combining the results of the other components, performed the best overall in terms of aggregate and distributed skills, and for 1-6 hour lead time in the Winter and in the Summer.

Our investigative studies indicate that a hybrid modeling strategy would be optimal for distributed QPF, when compared with linear or nonlinear data dictated tools alone, and/or parameterized physically based model alone.

The predictability of distributed, quantitative precipitation as a function of lead time would not depend just on the component process and the modeling strategy. The other relevant factors would be the resolution in space and time (both for the forecast generation and/or for the measure of skill), the season and event types, the type of measure used to quantify the skill scores, the specific modeling strategy (e.g., linear or nonlinear), as well as the types of outputs used from the NWP models and the information content in these outputs.

Greco and Krajewski (2000) conclude that the predictability of precipitation increases with lower resolution, and for more intense precipitation events. This research shows that the predictability could be a complex function not only of the resolution, but other criteria as well. Predictability is higher in the Winter in terms of aggregate QPF. In the Winter, QPF from NWP has more information content than other NWP model outputs at aggregate scales, and linear modeling strategies seem to perform relatively better. For distributed QPF, the predictability could be higher in the Summer, but only if nonlinear strategies are used that exploit the information content in the NWP forecasts of atmospheric variables. Further, the behavior of skills in terms of the sum of the mean squared errors could be different from the skills in terms of the ability to correctly predict a precipitation range, and these could be significantly different from the skills in terms of the aggregate error statistics.

11.0 Future Research

Several areas of future research have been identified by our studies. While we have performed or initiated preliminary research in a few of these areas, we believe these need to be explored further.

The proposed hybrid modeling scheme for QPF implicitly or explicitly assume that certain conditions remain identical within the scope of a “precipitation event”. The classification of events into Summer and Winter was seen to improve correlation and forecasts. In a continuous QPF scenario, the nature and type of the conditions that dictate precipitation generation could depend on the event type. These event types could broadly be classified into convective and stratiform, with a spectrum of processes that fall within these broad categories (e.g., see Browning and Collier, 1989).

We have used a parameterized, physically based, model for comparing with the proposed QPF modeling strategy. However, we believe there is significant room for further investigations. For example, it is conceivable that an optimal combination of a parameterized physically based model and an ANN based evolution model could improve the localized evolution component, especially in the Summer.

As NWP model resolutions improve, the QPF from these models could become more useful for flood forecasting applications. One area of investigation could be to evaluate the QPF obtained from a higher resolution NWP model which has better physical parameterizations for higher resolution precipitation processes, and which is initialized with radar observations. Another potential direction could be to enhance the understanding of the effects of topography on precipitation processes. One other research direction could be a more detailed investigation of the value and accuracy of

the NWP model outputs, including but not limited to QPF. The error estimates of the NWP model outputs could relate to the error bounds of the distributed QPF.

Finally, one could try to explain the distributed error structure and the performance of the various QPF strategies by a better understanding of the physical processes. Interesting questions would be whether and how the distributed error structures are related to topography, diurnal cycle, and/or to the skills in NWP model outputs, the relation of the skill of the precipitation components (e.g., the decay of skill of the localized evolution component after around 3 hour leads) to process physics (e.g., life time of convective storms), and similar insights.

11.1 Event Based and Continuous QPF Formulations

The scope of our problem is event based QPF. One logical future step could be to extend this to a continuous formulation, however this has both hydrologic and implementation implications.

Let us consider the relation of atmospheric instability indices to the precipitation rate. For convective situations, Zawadzki et al. (1981) reported good regression relations. However, mixed regression results were obtained by Peppler and Lamb (1989) when relating atmospheric instability indices to “growing season rainfall” (when both convective and stratiform regimes could be present, and regime separation is difficult to obtain). Issues like these indicate that going from an event based QPF model to a continuous one might not be a simple extension.

To go towards a continuous formulation, one option could be to have a classification scheme for handling multiple events and another QPF strategy (ANN based or otherwise) for handling event based rainfall evolution. We tested a similar strategy by creating a combined ANN based classification and regression scheme. Note the similarity with the ANN based approach suggested by Kim and Barros (2001) in the context of flood forecasting.

In our experiment, regression ANN were optimized to handle specific event types or regimes. Browning and Collier (1989) indicates that there could be a broad spectrum of processes within the two broad regimes (convective or stratiform) that dictates rainfall evolution. In a continuous application, ANN could be optimized for each of these regimes. For a given event type, these ANN would be used to model the evolution of the radar state on the current radar state and NWP model outputs.

The classification ANN used atmospheric indicators (NWP model outputs or some functions thereof) and aggregate radar statistics (centered around the current pixel) to assign probabilistic weights to the results of each of these regime specific regression ANN, for each pixel.

The final result was a Bayesian combination that reflected (a) the probabilities of the pixel at some “unseen” time step belonging to known event types or regimes (from the classification ANN), and (b) quantitative rainfall estimates given that the pixel belongs to these event types (from each of the regression ANN).

We generated simulated data using simplified physically based models to test this strategy for a continuous QPF formulation. We used simplified formulations from Lee and Georgakakos (1990) and as well as a simple translation and propagation (growth or decay) algorithm to generate simulated data. We generated simulated data for three regimes: “convective”, “stratiform”, and “mixed”. The parameters of the simulation models were adjusted to make the convective regime the most dynamic, the stratiform relatively stable, and the mixed somewhere in between.

Regression ANN were optimized to handle convective and stratiform situations, and a classification ANN was developed to determine the probability of belonging to one of these two regimes. The approach proposed here corresponds to a Bayesian formulation, which takes the form $Q_T = \sum \{ [Q|E_i] \times \text{Prob.}[E_i] \}$. Here Q_T is the forecast quantitative precipitation, Q is the quantitative precipitation forecast, E_i the event type (e.g., convective, stratiform, or the result of a more granular classification), $[Q|E_i]$ is the Q given the event type E_i , and $\text{Prob.}[E_i]$ is the probability of the current weather

regime belonging to one of the “known” event types. The probabilistic approach is suitable in this context, as individual regimes could be mixed in a typical weather event.

We verified our algorithms using “unseen” data, for convective, stratiform, as well as “mixed” situations. Note that no regime specific ANN existed for “mixed”. We were able to demonstrate skills in terms of the RMSE for each of the three situations, although the skills for the “mixed” regime was lower than that for the other two (Figure 11-1).

From the lessons learnt from previous QPF research, and the results of our analysis based on simulated data, we feel that a transition from the proposed event based QPF to a continuous formulation could be feasible. One approach could be an extension of the Bayesian formulation suggested in the example in this section. Several QPF “experts” (i.e., linear or nonlinear regression models, or physically based formulations if appropriate) could be optimized or calibrated for “representative” storms that occurred in the past. Certain characteristics of the storms useful for classification of the storm type (relevant NWP outputs, aggregated radar measurements, and possibly satellite measurements) could be used to identify a new storm (as a probabilistic combination of the existing storm types). Prediction could involve a weighted combination of the individual forecasts from the “experts”. A second approach could be the “mixture of experts” ANN architecture (see Bishop, 1996), which attempts to use an overall ANN architecture to perform the classification and regression tasks at once, rather than separating them out. Each method could have its own advantages and disadvantages. However, the implications in terms of data analysis, hydrologic process modeling, and implementation of these modeling strategies, could be fairly significant.

Figure 11-1: "Continuous" QPF Performance using Simulated Data
 (X- and Y- axes in pixel units [L]; Contour plots are for simulated VIL [ML^{-2}])
Simulated Verification Data QPF-ANN (Multiple Events Strategy)

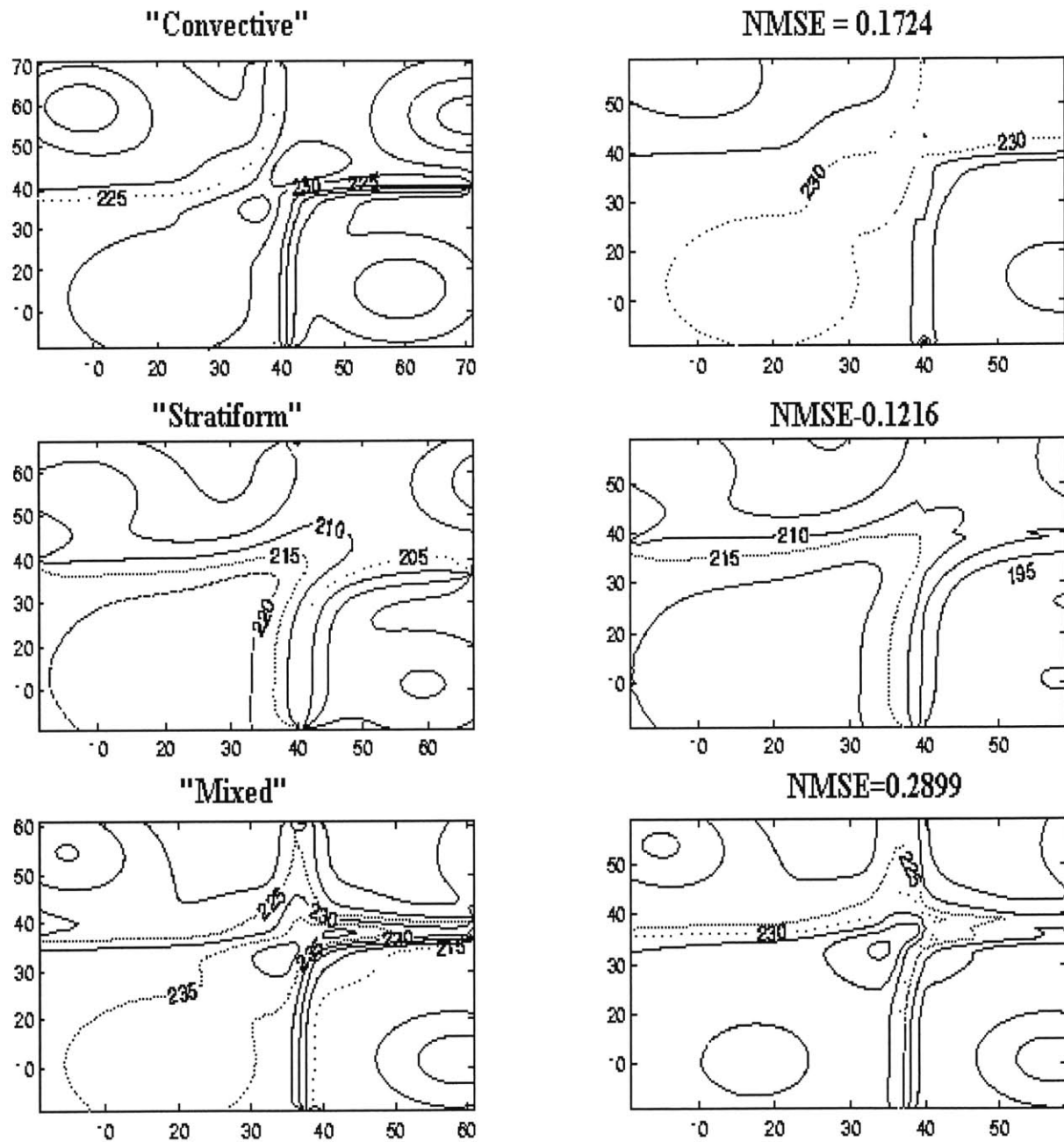


Figure 11-1 shows the performance of the suggested ANN based strategy for continuous QPF. Regime specific ANN were trained for "convective" (dynamic) and "stratiform" (stable), but not for "mixed". Classification ANN was trained to assign relative weights for convective or stratiform to unseen data on a pixel by pixel basis. This ANN based strategy was used on verification data, skillful QPF was obtained for all situations. The performance was the best for "stratiform", followed by "convective" and finally "mixed".

11.2 Distributed Precipitation Maps as Geometrical Objects

While the focus of a distributed QPF strategy is on both the aggregate measures and the distributed structure, it is the latter that is usually the most difficult to preserve for hourly rainfall. We have discussed earlier some of the approaches that have been tried and reported in the QPF literature (see Browning and Collier, 1989, and Smith and Austin, 2000, Duda and Blackmer, 1972, and the relevant references therein). One of the major problems with such approaches has been the decorrelation of observed precipitation structure after about an hour (Browning and Collier, 1989; Mecklenburg et al., 2000). However, it is conceivable that recent advances and emerging techniques in the areas of image recognition, pattern matching, or generic spatio-temporal evolution could offer new possibilities for QPF. Research in the evolution of spatio-temporal patterns might offer better ways to model and forecast the distributed structure of precipitation. As Swinney (1994) points out however, this area is still at a nascent stage, although interesting developments are already occurring (Swinney, 1997).

Researchers in QPF have already begun to explore areas like “Gaussian mixtures” to describe and forecast rainfall structure (Dorffner et al. 2001; Schellner et al., 2000). These methods are not new to the pattern matching community (Bishop, 1996), and are still being developed and explored further (Ma et al., 2000). However, similar models for rainfall cells (composed of Gaussian like entities with mean and covariance structures) have been explored by several researchers in the past (see for example, Rodriguez-Iturbe and Eagleson, 1987 and French, 1990 and the references therein). The added benefit from “Gaussian mixture” formulations, if any, need to be carefully evaluated.

12.0 Conclusions

Distributed QPF has diverse and important applications. In the area of flood forecasting, this has the potential of saving human lives and property. Improving distributed QPF also happens to be among the most challenging problems in hydrology and meteorology. This paper evaluates existing QPF strategies, and discusses the promise and pitfalls of recent advances that could be relevant for this area. Availability of better NWP models and high quality radar observations, as well as emerging data dictated tools and the insights gained from the combined wisdom of previous researchers, appear to indicate that a window of opportunity might have emerged for improving distributed QPF. Investigative studies indicate that decomposing the distributed QPF problem into component processes, and making the best use of available process physics and data dictated tools, could be a way forward.

A new hybrid modeling strategy has been proposed for distributed QPF at 1-6 hour lead times, that optimizes the use of information from radar measurements and NWP model outputs, and makes the best use of available process physics and data dictated tools. The strategy is to decompose the distributed QPF problem into four component processes, which are Radar Extrapolation, Large Scale Physics, Localized Evolution, and Residual Structures. The rationale for this decomposition rests on data analyses, investigative QPF studies, and hydrologic insights from previous research.

High quality NEXRAD data have been used in conjunction with the 48-km NWP-Eta model outputs to demonstrate improvements in distributed QPF. Improvements are quantified by a measure of average distributed skill based on the RMSE, a surrogate cost measure that indicates the predictability of precipitation thresholds, and depicted

by surface plots of distributed QPF errors. The proposed hybrid model improves distributed QPF over “state of the art” techniques like radar extrapolation alone, NWP-QPF alone, and hybrid models that combine radar extrapolation with NWP-QPF.

Investigative QPF studies described earlier also indicate that the hybrid model would perform better than pure data dictated tools, as the latter cannot model all the components of precipitation adequately. These studies also indicate that the hybrid model would perform consistently better than parameterized physically based models. The results of the investigative studies, as well as an analysis of the distributed QPF skills, offer new insights on precipitation processes, the information content in radar and NWP model outputs, as well as on the achievable precipitation predictability.

Appendix A: Use of Artificial Neural Networks (ANN)

For a thorough review of ANN concepts and their use in forecasting and function approximation, the reader is referred to Bishop (1996), Jordan and Bishop (1996), Weigend and Gershenfeld (1994), and Vemuri and Rogers (1994). We describe how we have made best use of these tools to handle aspects of the event based, distributed, quantitative precipitation forecasting problem.

A.1 Hydrologic Scope of the QPF Problem

A.1.1 Hydrologic Motivations, Forecast Requirements, Data Availability

Hydrologic considerations have guided the evolution of our ANN strategies, along with an understanding of ANN basics. Thus, the hydrologic requirement is to generate spatially distributed QPF at 1-6 hour leads, the motivations are similar to Golding (2000) and Antolik (2000). Forecasting horizon, lead time, resolution, and similar requirements were guided by these requirements and by the availability of data.

The ANN formulations were tested for scalability using real data, and for generic applicability by testing with a total of nine precipitation events. As described earlier, previous researchers have used ANN to solve other problems in hydrology, including but not limited to rainfall forecasting at different resolutions and lead times. For example, there have been successful use of ANN for forecasting rainfall time series (Toth et al., 2000), and the Indian monsoon (Navone and Ceccato, 1994). However, these are beyond the scope of this research.

A.2 ANN Fundamentals and Forecasting Applications

A.2.1 Taxonomy of ANN used for Function Approximation and Temporal Processing

Artificial Neural Networks (ANN) refer to a class of models, which could be broadly classified into “feedforward” and “recurrent” networks. ANN consist of one or more layers, each with a series of nodes. The first layer receives the inputs, and is called the input layer. The final layer produces the output, and is the output layer. An ANN could have zero, one or more hidden layers. In feedforward network the outputs of a given layer are always used as inputs to a successive layer. In a recurrent neural network, the outputs (or a transformation) of a layer could be used as inputs to one or more previous layers.

Feedforward ANN have been shown to act as universal functions approximators (Hornik et al., 1989; Cybenko, 1989), and to converge faster than other approximators that are linear sum of fixed functions. A special class of feedforward ANN are called MLP (MultiLayer Perceptron), which have no explicit time varying component. An MLP could be used to “learn” the Nonlinear Auto-Regressive (NAR) component of a NARMA (Nonlinear Auto-Regressive Moving Average) formulation, if the output is a time series variable, and the inputs are the lagged variables. If the MLP were to consist of one node with a linear transfer function, this could be identical to a ARMA model with the appropriate choice of inputs and output.

Another class of feedforward ANN are called TDNN (Time Delayed Neural Networks), which are often functionally similar to recurrent networks (Wan, 1994). Similar models include the “FIR Network” of Wan (1994), that uses FIR (Finite Impulse Response) linear filters for the ANN nodes. These have demonstrated value in forecast

applications, the model by Wan was judged to be the best in the Santa Fe time series competition. However, these models require a vast number of forecast parameters, for example the “FIR Network” of Wan (1994) used 1105 “parameters” to fit 1000 data points. The reason these models still work is due to “early stopping”, which keeps the number of effective parameters low. However, even the model of Wan was not the best for long lead time forecasts, when the complex function approximated by the network was possibly no longer as valid.

Mozer (1994) provides a comprehensive taxonomy of recurrent neural network architectures, which includes TDNN (and variations thereof) as special cases. According to Mozer (page 244), “conventional neural net architectures and algorithms are not well suited for patterns that vary over time”. According to him, prediction is a two step process, where the first step “constructs a short-term memory that retains aspects of the input sequence relevant to making predictions”, while the second step “makes a prediction based on the short-term memory”. Mozer classifies recurrent ANN based on what he calls the form, the contents, and the adaptability of the short-term memory. The forms could be “tapped delay-line memory” (identical to delay space embedding and corresponds to a nonlinear variant of the auto-regressive or AR model), the “exponential trace memory” (similar to a tapped delay line but assumes an exponential form for the relative strength of the lagged inputs, and is a special case of the Moving Average of order 1 or MA(1) model), the “gamma memory” (which generalizes the depth of the memory or the number of lagged variables and the resolution or the extent of information retained on each variable), and other forms of memory like a Gaussian memory. The contents of the short-term memory could be I (the inputs), TI (transformed

inputs, for example through a logistic transfer function), TIS (transformed input and state, which is a nonlinear function of the current input and the state), O (outputs), TO (transformed outputs), and TOS (transformed outputs and state). The adaptability of the memory refers to whether the parameters of the form and contents are fixed (“static memory”, leading to less flexibility but more parsimony) or whether these can adapt with data (“adaptive memory”, leading to more flexibility but increased parameters). These could lead to several types of ANN architectures, representations, and algorithms, of which not all have been explored fully. Mozer states that the types of recurrent ANN that have been explored do not suggest that these could significantly improve forecasts. For example, he mentions that (page 263) “a nonlinear extension of the auto-regressive models - the basic I-delay architecture - fared no worse than a more complex architecture”, and quotes another study that supports his conclusions.

Connor et al. (1994) proposed and used a special type of recurrent neural networks for modeling the NARMA(p,q) model, where the lagged input variables and the lagged errors (which represents a transformation of the neural network outputs) were used as inputs. This form of ANN showed improved skills in predicting time series.

While the potential for ANN for prediction tasks have been demonstrated by previous researchers (Vemuri and Rogers, 1994; Weigend and Gershenfeld, 1994), numerous unresolved issues remain. Among these considerations are long-term memory, short-term memory, and model parsimony. However, the value of ANN (specifically, MLP and similar feedforward networks) to handle pattern recognition problems (e.g., classification and regression, where the functional form is not time varying) have been demonstrated. As discussed later, Bishop (1996) suggests that

ANN methods might be of limited use for prediction unless longer term trends are removed and the functional form is not time-varying.

There have been numerous discussions on “connectionist” (ANN) versus statistical techniques. On the one hand, statisticians have been weary of the “media hype” and have tried to put ANN methods in place, by stating that “in one sense neural networks are little more than nonlinear regression and allied optimization methods” (Ripley, 1993). Others have gone to the extent of stating that connectionist researchers “routinely reinvent methods that have been known in the statistical or mathematical literature for decades or even centuries” (Sarle, 1994). On the other hand, recent trend among the ANN community (primarily those involved in classification, data analysis and prediction tasks) to move closer to traditional statistics (Bishop, 1996; Jordan and Bishop, 1996; MacKay 1992, 1994) appear to lend support to the view that when ANN are used for data analysis and prediction, these need to be looked at as extensions of traditional statistical tools, and interpreted in that light as far as possible. Another “connectionist researcher”, Flexer (1995), mention that “neural network models can be described properly within a statistical framework”, and hopes that for future research in ANN, “references and links to statistics will become obligatory, and hopefully, a must for acknowledgment of scientific soundness”. Flexer (1995) adds that “hopefully ... the steady growth in the number of different models and algorithms in connectionism will stop.”

However, it is important to note that well designed ANN often perform better than traditional methods (Wan, 1994; Weigend and Gershenfeld, 1994; Vemuri and Rogers, 1994; McCann, 1992). In particular, the statistical community has recognized the

MultiLayer Perceptron (MLP) as “especially valuable because ... the complexity of the model [can be changed] from a simple parametric model to a highly flexible, nonparametric model” (Sarle, 1994). However, other ANN models like “recurrent networks are already being examined for their commonality with NARIMA models”, and “the underlying statistical principles of such networks are still unexplored” (Flexer, 1995).

A.2.2 ANN Basics: Definitions and Key Concepts

A.2.2.1 The Interconnection “Architecture”

The interconnection “architecture” consists of the inputs, the outputs, and how the mapping between these are described via the nodes (which apply a transfer function like linear, logistic, or hyperbolic tangent), and connection “weights” (which are model parameters that assigns relative weights to the inputs of each node). An MLP could have an input layer (which receives the input vector), an output layer (which yields the outputs), and hidden layers. Each layer consists of one or more nodes, each with a transfer function. The outputs from each layer are fed to each node of the successive layer, after applying the connection weights.

A.2.2.2 Activation Functions

An activation [or, “squashing”] function is associated with each node. Some common forms of the activation functions include logistic or sigmoid, i.e.,

$$g(\underline{X}) = 1/\{1+\exp(-w\underline{X})\},$$

where w is the vector of weights or parameters, and \underline{X} the vector of inputs, as well as linear and hyperbolic tangent. These nonlinear activation functions cause the

mapping between the inputs and outputs to be nonlinear, the overall mapping is often viewed as a “superposition of nonlinearities”. The training or calibration process starts with values of the weights or parameters chosen from Gaussian distributions with zero mean and small variance. Since the activation functions behave in a linear fashion when the weights are near zero, the overall network initially behaves in a linear fashion, getting more nonlinear as training proceeds and the weights get strengthened.

A.2.2.3 The Cost Function

A cost function is the error measure which needs to be minimized. The choice of the cost function is based on the kind of function that the ANN is trying to “learn” or model. For example the mean squared error (MSE) is often used in regression problems, and the cross-entropy in certain classification and recognition problems. The iterative scheme that modifies the weights based on the training data is guided by the specific combination of the cost function and the activation function that the ANN uses.

A.2.2.4 Training Algorithms

A “training algorithm” is used for parameter optimization, which determines the most likely values of the connection weights (parameters), given the data. The available data is typically sub-divided into two segments, the first for training and cross-validation, and the second for verification. For forecasting applications, the first set would comprise data in the past, while the second would be data for forecast lead times (“unseen” data). Typically, the ratio is about 2:1. The first set is then further divided into the training and the cross-validation data, roughly in the ratio of 2:1. For forecasting, the data could be randomly divided into training and cross-validation. The training data is

used for adjusting the model parameters, while the cross-validation data is used to test the performance of the model on “held out” data. The training algorithm is usually an iterative scheme (but could also work in a “batch” mode) that modifies the values of the weights at successive iterations. Exit criteria for the iterations include a minimum specified value for the cost function on the training or cross-validation data, or if convergence cannot be achieved, when the cost functions fails to improve significantly after successive iterations.

We discussed earlier that the ANN based mapping gets more and more nonlinear with more training iterations, leading to an increase in what Weigend and Gershenfeld (1994) calls the number of “effective parameters”. This often has the effect that while the cost function keeps on decreasing with respect to the training data as the number of iterations (or training “epochs”) are increased, the cost function on the “held out” or cross-validation data remains stagnant or even tends to increase. The performance on the cross-validation data is therefore typically used as the exit criteria for ANN training.

The iterative schemes used for modifying the ANN weights are based on standard parameter optimization techniques like gradient descent or conjugate descent. The algorithm used to modify the weights use the chain rule of calculus, where the errors at each node are propagated backwards to modify the weights (this leads to the much used backpropagation algorithm and variations thereof, see Bishop, 1996). Various heuristics are then applied to make the parameter optimization efficient and to prevent overfitting. For example, a momentum term could be used to control the weight modification at each iteration as a function of the current value of the weight, a learning

rate could be specified to determine how fast the weights should be modified to account for the error at the current iteration, a regularization term could be added to the cost function that penalizes model complexity so that the ANN remains parsimonious.

A.2.3 Single Node MLP with Linear Activation for the ARMA(p,q) model

Let us first consider the simplest ANN (an MLP with just one layer) , and how that could be used for time series forecasting. Let the vector of inputs be \underline{X} , with the elements $\{x_1, x_2, \dots, x_i, \dots, x_n\}^{(k)}$, where k represents a particular data set (or pattern, which is an input-output pair) and x_i represents the data elements. The output (we assume a single output for the purposes of this illustration) could be denoted by $y^{(k)}$. Let us assume that the ANN has a single node, with a linear activation function. For the k^{th} input-output pair (or “pattern”), the output is given by

$$y^{(k)} = \sum w_i x_i^{(k)}$$

where “i” represents the number of interconnection weights. Given that the activation function is linear, the output is a linear combination of the inputs. Note that a constant term could be included in this framework by allowing for an extra set of inputs x_0 , the value of which always remains 1. This is referred to as the “bias” (not to be confused with the statistical nomenclature) for the ANN node.

The simple ANN described above could include as special cases the traditional statistical formulations for modeling the random component of a time series, for example the AR (Auto-Regressive), MA (Moving Average), or ARMA (Auto-Regressive Moving Average) components by an appropriate choice of the input and output pairs, and also the single exponential smoothing formulation by an appropriate choice of the weights. This could be accomplished by using successive values of a time series as the

outputs, and the corresponding lagged (“time delayed”) values as the inputs (the AR model), or lagged errors as inputs (the MA model), or a combination of both (the ARMA model).

Once an ANN has been set up in the manner described, the parameters or the weights need to be estimated, just as in any ARMA formulation. For the ANN, this will involve the choice of an error function (e.g., $E^{(k)} = (y^{(k)} - t^{(k)})^2$ for least squares, where t represents the observed data used for training), and a method for parameter optimization (for example, gradient descent). For the least squares cost function and the gradient descent optimization method, one could show that the iterative weight update scheme takes the following form:

$$w_i^{(new)} = w_i - \eta (\delta E / \delta w_i)$$

While we have shown that ANN could theoretically model the ARMA(p,q) model once p and q are known, we have not discussed how we shall estimate these lag windows. For the traditional ARMA model, the Box-Jenkins methodology allows us to do that, by selecting initial estimates of p and q from data analysis (e.g., by looking at autocorrelation functions and partial autocorrelation functions), and then successively refining these estimates. Automated techniques often rely on quantifiable measures like Information Criteria (e.g., AIC, the Akaike Information Criteria, or BIC, the Bayesian Information Criteria), but these might not always yield the best model choice. In the traditional ANN literature, the way to deal with this kind of situation would be through trial and error, and verifying the performance of each trial on the cross-validation or held out data. Recent trends often attempt to combine statistical insights with ANN, and could offer several possibilities. First, model identification could be done through

statistical techniques prior to estimation, or linear statistical models could be used to pre-process the data before passing them on to the ANN. These sequential combinations of linear and nonlinear methods could be very useful in certain application scenarios. However, these might not always be applicable in situations where the motivation of using ANN is to discover and account for nonlinear behavior. Second, one could start with fairly large ANN with a large number of inputs, and selectively “prune” these based on the relative strengths of the weights. The third approach could be to start with small network, and successively add additional input layers based on performance on cross-validation data. Bishop (1996) discusses these methods. Finally, Bayesian techniques could be used with ANN (Bishop, 1996; MacKay, 1992, 1994) to ensure model parsimony.

A.2.4 Single Node MLP with Nonlinear Activation

Let us now extend the single node, linear activation MLP discussed previously to a single node MLP with a nonlinear activation function, say the most commonly used logistic activation function, $g(\alpha) = \{1 + \exp(-\alpha)\}^{-1}$, where $g(\alpha)$ is the transfer function applied to the input α . Using the same notations for the inputs and outputs as in the previous section, the output could be represented as a nonlinear function of the inputs, for example:

$$y^{(k)} = g\{\sum w_i x_i^{(k)}\} = \{1 + \exp(-\sum w_i x_i^{(k)})\}^{-1}$$

Jordan and Bishop (1996) proved that this seemingly simplistic formulation could model a wide range of functional forms and statistical distributions. For example, in classification problems, this could model the “exponential family” of distributions, which refer to a generic class of distributions that include the Gaussian, Poisson, Bernoulli,

Exponential, Gamma and several others as special cases. In regression problems, this would correspond to “logistic regression”, one of the most commonly used forms of nonlinear regression methodologies. The output from this simplistic model would always lie between (0,1), hence the training data must be scaled to lie between these ranges. In regression problems, the output from the ANN could be scaled back to the original units. The parameter estimation, or training algorithm, is a straightforward extension of the special case of the linear activation function described earlier.

A.2.5 The Generic MultiLayer Perceptron

The generic form of the MLP could consist of a selected number of hidden layers, each with a set of nodes with potentially nonlinear activation functions and a set of (adjustable) interconnection weights linking one layer to the next (starting from the input layer to the output layer, with hidden layers in between).

The generic MLP could have n layers, with the n^{th} layer yielding the outputs, the 1st layer being the input layer, and the rest being the hidden layers. The number of nodes in the output or input layers would be equal to the number of outputs or inputs respectively. The optimal number of hidden layers, and the number of nodes in each of these layers would depend on the complexity of the underlying mapping between the outputs and the inputs. These are determined from preliminary data analysis, where one would start with a simple network and progressively add layers and nodes depending on the performance of the cost function for the held out data.

The relation between the inputs (x) and the output (y) takes the form (for a given output unit k):

$$y_k = g_n[\sum w_{kp}^{(n)} \dots \sum w_{mj}^{(2)} \{g_1(\sum w_{ji}^{(1)} x_i)\}]$$

where i represents an input unit, j the units in the first hidden layer, and so on, p the units in the last hidden layer, and k the units in the output layer. The w 's represent the interconnection weights (from one layer to the next starting with the input layer), these are the MLP parameters that need to be adjusted. The g 's represent the activation function of each layer (assumed to be the same for each node in a given layer. This could take the form of the logistic function: $g(\alpha) = 1/\{1+\exp(-\alpha)\}$. The initial value of the weights are selected using random numbers with $\sim N(0,\varepsilon)$ where $\varepsilon \rightarrow 0$, which initially keeps the g 's linear.

The training (weight adjustment) step is just the chain rule of calculus, and the adjustment at each node depends on the rate of change of the cost function $-\delta E/\delta w$. Derivation of the rule for what is known as the "error backpropagation" becomes a straightforward application of the chain rule (see Bishop, 1996, page 140-146 for a full derivation and a simple example). For the output layer, the error is obtained from the difference of the network output and the "observed" or target value. This error is then propagated backwards, this involves using the derivatives of the activation function. For the commonly used activation functions like the logistic or the hyperbolic tangent, this is relatively easy to obtain. Note that the choice of the cost function to be minimized and the nature of the activation function determines the final form of the derivative.

A.2.6 ANN for Forecasting

While the value of ANN techniques for pattern recognition (classification and regression) problems are well established (Bishop, 1996; Jordan and Bishop, 1996), their applicability in forecasting applications have always been the subject of considerable debate. For example, Chatfield (1993) argued that while ANN could indeed improve forecasts in certain situations, these might not even be applicable in all situations. In fact, Chatfield used an example to demonstrate how easily one could obtain misleading results with ANN. This is not surprising, as ANN often cannot deal well with extrapolation situations (i.e., when the training or calibration data do not contain the domain of possible behavior). The other problem with ANN is that these are not parsimonious models, and could “overfit” the training data leading to poor generalization. While one could argue that the number of “effective parameters” in ANN could be controlled by reducing the number of “training epochs” (these terms are defined earlier), achieving the correct balance could prove to be difficult.

On the other hand, when used appropriately, ANN could result in significant improvement in forecasting applications, for example the best forecast in the Sante Fe time series competition was obtained by an ANN based technique (Weigend and Gershenfeld, 1994; Wan 1994). Vemuri and Rogers (1994) provide several examples of the use of ANN in forecasting applications, and Kim and Barros (2001) provide an example of ANN usage in the context of hydrologic forecasting. Connor et al. (1994) demonstrated how an ANN could model the nonlinear form of the traditional ARMA model (or, the NARMA). One could show that the nonlinear statistical models like the TAR (Threshold Auto-Regressive) and Volterra are special cases of the NARMA model.

These considerations suggest that while the application of ANN to forecasting problems could add significant value, they must be used with care. For example, one primary consideration is model parsimony, another is the presence of time-varying trends. Bishop (1996; page 302-304) goes so far as to suggest that the use of ANN in time series applications might be inappropriate unless the time dependent components are removed by de-trending the data, and if the “generator of the data itself evolves with time”. However, linear models like ARIMA and exponential smoothing have severe limitations in certain situations, Weigend and Gershenfeld (1994; page 16-17) mention that “if and only if the power spectrum is a useful characterization of the relevant features of a time series, an ARIMA model will be a good choice for describing it”, and demonstrate through examples situations where linear models break down. Weigend and Gershenfeld further mention [page 34] that while state-space embedding techniques might not work too well for high dimensional, stochastic, and nonlinear data, in the case of ANN “much of their [ANN] promise comes from the hope that they can learn to emulate unanticipated regularities in a complex signal”, which in turn could lead to “the best as well many of the worst forecasts” with ANN.

The review of ANN in the context of forecasting applications draws one closer to the conclusions reached by Chatfield (1993) in the early days of such applications. While demonstrated evidence now exists on the potential of ANN techniques, these might not be applicable for all forecasting scenarios. Further, the ANN community has often been overwhelmed with “optimal architectures” and corresponding representations and algorithms, at times at the cost of the insights gained over years of research in the statistical and forecasting community. One example is the need to

consider model parsimony, ANN architectures like complex recurrent nets often do not take that into consideration. Recent research in ANN is moving the field closer to traditional statistics, where ANN are treated as tools for nonlinear regression and classification. Bishop (1996) and Jordan and Bishop (1996) provide examples of interpreting ANN function approximation in the light of traditional statistics, while the Bayesian approaches of MacKay (1991, 1992, 1994) and others deal with issues like model parsimony and uncertainty propagation. The NARMA formulation of Connor et al. (1994) is a similar effort. We go a step further and assume that the NAR and NMA are linearly separable. This not only allows us to understand and model the two components separately, but also provides the flexibility to use the NMA component only if it is useful. Together, these capabilities make the overall model parsimonious.

Previous researchers have discussed other ways in which traditional time series and forecasting concepts could be used in conjunction with ANN for better results. For example, longer term trends and certain short-term or time varying components could be isolated and modeled using standard time series techniques, while nonlinear time-invariant components could be modeled using ANN. These topics are discussed in some detail in the literature (Flexer, 1995; Bishop, 1996), as well as in the more informal media like the FAQ area for the Internet newsgroups on neural networks.

The QPF formulations described earlier restricted the use of ANN to functional forms that are assumed to be invariant in time, with lower resolution NWP physics accounting for large scale effects, and extrapolation based on advection accounting for shorter trends. The optimal use of ANN in the context of distributed QPF was one of several considerations that led to the development of the proposed approach.

A.3 ANN Formulation Details

We have described earlier our rationale for using ANN based methods to model certain components of the overall precipitation process, and how we selected the inputs and outputs. This section gives a very brief overview of the capabilities of ANN, and the specific techniques used in this research.

There could be different strategies for developing ANN formulations. Both the ANN basis functions (neurons), and the connections between them, could take different forms leading to a variety of possible ANN “architectures”. One approach is a “brute force method”, where different architectures are tested with a large number of data points till the “correct” architecture is identified. It could be argued that these are somewhat ad-hoc. However, implicit in these methods are often the search for entities like the optimal choice of the basis function in certain kinds of regression problems, or the optimal size of the lag window in a time series forecasting problem, which are difficult choices in traditional statistics.

Recent research efforts attempt to relate ANN to traditional statistics (Jordan and Bishop, 1996; Bishop, 1996; Weigend and Gershenfeld, 1994; Connor et al., 1994; MacKay, 1992). These studies suggest that the best results from ANN could be obtained not by treating these as mere computational tools, but by using these as nonlinear statistical tools. In this work, we have attempted to steer away from methods that could be perceived as “ad hoc”, and have attempted to relate our formulations to statistical concepts and theoretically proven ANN capabilities.

A.3.1 ANN as Function Approximators

Hornik et al. (1989) and Cybenko (1989) proved that MLP (MultiLayer Perceptrons) could act as “universal function approximators”. MLP have an input layer, an output layer, and one or more hidden layers. The number of nodes in the input and output layers are dictated by the number of inputs and outputs between which a functional mapping is sought. The basis functions used for each node, the number of hidden layers and the nodes in each hidden layer depend on the nature of the functional form to be modeled.

A variety of universal approximators exist in the literature. As Barron (1993) and Bishop (1996) indicate, the residual sum-of-squares of the errors falls as $\sim O(1/M)$ for ANN and as $\sim O(1/M^{2/d})$ for linear combinations of fixed functions, where M is the number of hidden nodes or functions and d the number of input dimensions. Algorithms exist (Bishop, 1996; Hertz et al., 1991) to search for the best ANN parameters using cost functions like mean squared errors and optimization methods like gradient descent.

A.3.2 MLP for Nonlinear Regression

The ability of MLP to model nonlinear regression problems is well known (Bishop, 1996; Jordan and Bishop, 1996):

$$X_{t+1} = f_{\text{MLP}}(Y_{t+1}^1, Y_{t+1}^2, \dots, Y_{t+1}^n) + \varepsilon \quad \text{A.1}$$

In Equation A.1, “X” is the vector of outputs (dependent variables), and “Y” the vector of inputs (independent variables) in the nonlinear regression problem, at time $(t+1)$. The lagged values of “Y” could be modeled separately to obtain Y_{t+1} , or they could be included in a generic ANN formulation.

A.3.3 Nonlinear ARMA and MLP

Connor et al. (1994) showed that an MLP could model the Nonlinear ARMA (AutoRegressive Moving Average), or NARMA(p,q) model as:

$$X_{t+1} = f_{MLP}(X_t, X_{t-1}, \dots, X_{t-p+1}; \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) + \varepsilon_{t+1} \quad A.2$$

We assume that Nonlinear AR (NAR) and Nonlinear MA (NMA) are linearly separable:

$${}^{NAR}\underline{X}_{t+1} \cong f_{MLP}(X_t, X_{t-1}, \dots, X_{t-p+1}) \quad A.3a$$

$${}^{NMA}\underline{X}_{t+1} = X_{t+1} - {}^{NAR}\underline{X}_{t+1} = f_{MLP}(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) \quad A.3b$$

The assumption of linear separability reduces some generality, but decomposes the problem into an autoregressive component and another that models the errors. The MLP for the NMA component is only used if this improves the forecasts at analysis time, this reduces the computation time. Note that this formulation includes the ARMA, ARIMA (AutoRegressive Integrated Moving Average), and the (Nonlinear) TAR (Threshold AutoRegressive) models in traditional time series as special cases.

We used the formulation described in this section for the ANN based functions approximation in our study.

A.3.4 Proposed Extension to Multivariate NARMA

The NARMA formulation could be extended to the multivariate case, where independent variables other than the past values of the quantity to be predicted need to be used in the prediction equation. This leads to the multivariate NAR model:

$${}^{(M)NAR}X_{t+1} = f_{MLP}(X_t, X_{t-1}, \dots, X_{t-p+1}; {}^1Y_{t+1}, \dots, {}^nY_{t+1}) + \varepsilon_{t+1} \quad A.4a$$

$${}^{NMA}\underline{X}_{t+1} = X_{t+1} - {}^{(M)NAR}\underline{X}_{t+1} = f_{MLP}(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) \quad A.4b$$

In Equations A.4, the “Y” terms denote the independent variables at time (t+1). The form of the NMA model remains identical to Equation A.3b. Note that both the regression formulation of Equation A.2 and the NARMA formulation of Equation A.3 are special cases of the generic formulations of Equations A.4. Lagged values of the independent variable “Y” could be used directly in an ANN formulation, or could be used to obtain forecasts for Y_{t+1} (the latter is assumed in Equations A.4).

A.3.5 Proposed QPF Formulation from the Multivariate NARMA

In the context of the QPF problem, ANN are used to model (a) the localized evolution of rainfall in a Lagrangian frame of reference, and (b) the distributed residual structure as a function of the results of two individual components, scaled advection and ANN evolution. For these processes, we assume that the functional form in space and time for a given rainfall event is identical for (a) all radar pixels contained in one NWP grid, and (b) ANN training and cross-validation time steps (e.g., 3 hours into the past) and ANN forecast lead times (e.g., 6 hours into the future). This is a fairly strong assumption, and implies that all possible functional forms in the forecast generation phase, for the given NWP grid and precipitation event, would be learnt by the ANN from the training data.

A.3.5.1 ANN for the Localized Evolution Problem

With the assumptions described earlier, the multivariate NARMA (with linearly separable multivariate NAR and NMA components) could be extended for the function approximation problems in space using all radar pixels in one NWP grid. For localized

evolution, we do not consider lagged values of the state (X), or the independent variables (Y), for reasons explained later.

$${}^{(M)NAR} \underline{X}_{t+1,x,y} \cong f_{MLP}(X_{t,x,y}, \dots, X_{t,x-i,y-j}, \dots, X_{t,x-m,y-n}, {}^1Y_{t+1,s}, \dots, {}^nY_{t+1,s}) \quad A.5a$$

$${}^{NMA} \underline{X}_{t+1,x,y} \cong f_{MLP}(\varepsilon_{t,x,y}, \dots, \varepsilon_{t,x-i,y-j}, \dots, \varepsilon_{t,x-m,y-n}) \quad A.5b$$

In Equations A.5a and A.5b, which correspond to the localized Lagrangian evolution problem, $X_{t,x-i,y-j}$ is the radar observation at neighboring grid (x-i, y-j), while “Y” represent NWP outputs which are uniform (in space, denoted by s) over the NWP grid.

In a generic time series problem, lagged values of the state (X) would need to be included. In the distributed QPF problem, this would amount to solving the advection problem with ANN, which could result in worse performance, as indicated earlier. The hourly decorrelation time of distributed rainfall precludes the inclusion of the lagged state even in a Lagrangian frame of reference, especially since the resolution is hourly.

In a generic space time problem, the lagged values of the independent variables (Y) could also be considered. In the context of the QPF problem, these are NWP model outputs. NWP models integrate the state based on best available physical models, starting with the current state as initial conditions. The use of lagged NWP model outputs in an ANN formulation could reduce the value of this physics based integration. This does not contradict earlier observations that there could be value in error correction of the NWP outputs based on comparisons with observations.

A.3.5.2 ANN for Distributed Residual Structures

The ANN formulation used for handling residual structures on a pixel by pixel basis could be expressed as:

$${}^{(M)NAR}\underline{X}_{t+1,x,y} \cong f_{MLP}({}^1Y_{t+1,x,y}; {}^nY_{t+1,x,y}) \quad A.6a$$

$${}^{NMA}\underline{X}_{t+1,x,y} \cong f_{MLP}(\varepsilon_{t,x,y}, \dots, \varepsilon_{t,x-i,y-j}, \dots, \varepsilon_{t,x-m,y-n}) \quad A.6b$$

The (M)NAR component in Equation A.6a takes as inputs the results of the two component strategies (disaggregation of NWP-QPF in space and time using advection results, and localized Lagrangian evolution).

In Equation A.6a, no lagged values of “Y” are used, as these components are modeled separately. No lagged values of “X” are used as advection or evolution is not being modeled here.

A.3.6 Bayesian ANN for Error Bars

Mackay (1992, 1994) proposed “Bayesian ANN” techniques, and suggested a method to generate error bars using ANN as part of this proposal.

The first step is to model the data using the ANN (e.g., MLP), and obtain a converged MLP, ${}^{Conv}f_{MLP}$. The next step is to determine the error statistics of the cross-validation data. If zero-mean Gaussian errors are assumed, the variance is the only parameter. Random numbers (η) are generated with the same statistical properties as obtained from the error statistics, these are added to the outputs of the cross-validation data to generate a set of pseudo-outputs ($\underline{Y}' = \underline{Y}_{test} + \eta$). The converged MLP obtained earlier is then re-trained using the pseudo outputs with unchanged input vectors, this gives an ensemble of ANN models. This completes the ANN calibration process.

During the verification or forecast generation phase, the converged MLP is used to generate the most likely forecast. The ensemble of ANN models are then used to generate an ensemble of forecasts at each forecast point. The statistics of these ensemble forecasts are calculated. If Gaussian distributions are assumed, only the

variance is relevant. Confidence bounds at each forecast or verification point are calculated from the statistics of the ensemble.

A.3.7 Proposed Technique for Generation of Error Bars at Multiple Lead Times

For single step forecasts, we use the Bayesian ANN concept described earlier to generate error bars. We extend this concept for multiple lead times. At lead time (T+1), we could have “n” number of ensemble forecasts for each forecast point. We could use each of these to generate “n” ensembles at forecast lead time (T+2), yielding to total of n^2 forecast ensembles. The confidence bound is expected to widen with higher lead times. We implicitly assume that the error statistics at a lead time (T+2) are a function not only of the input variables relevant for the (T+2), but also of the error statistics at the previous lead time (T+1).

We recognize that if a large number of ensembles are needed to model the error statistics or if there are several forecast time steps, the computation cost could get large. We could reduce this cost by a simplifying implementation assumption that the ensembles would be only generated from the upper and lower bounds. This is not a strong assumption. This merely implies that the lower/upper confidence bound at a forecast lead of (T+1) would be the starting point for the lower/upper confidence bound at a forecast lead of (T+2). This simplifying assumption will keep the number of ensembles to $2xn$ at each forecast lead, n being the number of ensembles. We have typically used this approach in this research.

A.3.8 ANN Architectures and Training Parameters

We used MLP (MultiLayer Perceptrons), which are a commonly used ANN type, to perform a variety of tasks. The specific inputs and outputs for the MLP used in our research were decided based on hydrologic considerations and statistical data analysis. The hidden node configurations, and the basis functions, were selected from preliminary ANN analysis. The results were typically not found to be too sensitive to the exact choice of the basis function or the MLP configuration. The MLP and the procedures used for selecting these were based on standard techniques. For detailed discussions, refer Bishop (1996), Jordan and Bishop (1996), and Hertz et al. (1991). For implementing the MLP, we used the MATLAB toolbox. We used the modeling language of MATLAB to use the MLP for specific tasks.

The MLP used to model the localized high resolution evolution process had 16 inputs and 1 output. Radar extrapolation provided 9 inputs, for the pixel under consideration and the 8 neighbors. For localized evolution, this represents a neighboring area of 12x12 km, larger areas did not improve forecasts. Relevant NWP model outputs yielded 7 inputs, these were Wind vectors u and v , updraft velocity ω , the lifted index LI, average temperature T , sea level pressure P , and average relative humidity H . The output was the rainfall evolution forecast at the corresponding pixel. For QPF using calibrated radar, a 16-9-5-1 MLP, was selected from prior data analysis. The basis functions were the hyperbolic tangent sigmoid transfer function for the input layer, and linear transfer function for the output layer. The hidden layers used the log sigmoid transfer function. Gradient descent with momentum weight/bias learning function was used for training. The momentum term was 0.9 and the learning rate was

0.01. The error criteria was MSE. The training algorithm was Levenberg-Marquardt backpropagation. The maximum number of training epochs for the most likely estimates was restricted to 200 based on preliminary observations and analysis. The number of training epochs for the Bayesian ensemble was restricted to a maximum of 50.

The MLP used to model the distributed residual structures had 2 inputs and 1 output. The inputs were the QPF results for the pixel under consideration from the two components of the proposed model: Advection scaled with NWP-QPF, and ANN based localized, high resolution evolution in a Lagrangian frame of reference. The output was the combined radar rainfall at the corresponding pixel. The ANN based function approximation was conditioned on the results of advection. For QPF using calibrated radar, a 2-3-2-1 MLP, was selected based on data analysis. The basis functions were the hyperbolic tangent sigmoid transfer function for the input layer, and linear transfer function for the output layer. The hidden layers used the log sigmoid transfer function. Gradient descent with momentum weight/bias learning function was used for training. The momentum term was 0.9 and the learning rate was 0.1. The error criteria was MSE. The training algorithm was Levenberg-Marquardt backpropagation. The maximum number of training epochs for the most likely estimates was typically restricted to 150 based on preliminary observations and data analysis. The number of training epochs for the Bayesian ensemble was restricted to a maximum of 20.

In this section, we have described the ANN techniques and models that we have used for our QPF work. We have attempted to highlight how we built upon existing research, and our proposed innovations. A detailed exposition of ANN is beyond the scope of this thesis, the interested reader is referred to the existing literature.

Appendix B: Directional Structure in Measured Precipitation and QPF Errors - Evidence for Assertions

This appendix provides evidence for some of the assertions in Section 8.2.3.3 (see pp. 109-111), which in turn deal with the directional structure in the QPF errors and measured precipitation. First, the figures here demonstrate that the directional behavior are present in the measured precipitation and the hourly persistence errors. This suggests that these are not artifacts of the proposed strategy for QPF, or any of the components thereof. These figures also demonstrate the effect of the lower and higher contour values on the directional structures. While the lower values appear more scattered and directional, this behavior is less prominent when only higher contour values are plotted. This behavior is observed in varying degrees for all the storms in the ABRFC that use Stage III data.

For each storm (precipitation event), we show a set of three figures. The first and second sets show the measured precipitation at the 1-hour ahead verification time step, and the corresponding precipitation produced from Persistence. Persistence uses the current precipitation value at each pixel as the forecast at the lead times. The measured precipitation and the corresponding persistence map (i.e., two measured precipitation maps at successive hourly time steps) are shown adjacent to one another. Successive maps use only the higher values of the contours. The third set of figures show the errors from persistence, and the successive maps once again use higher contour values only.

Besides presenting evidence for the directional structure of the observed precipitation maps, the objective of this section is to visually depict how these maps and the errors from hourly persistence appear when plotted in the form of contours.

**Figure B-1: Measured Precipitation and Persistence (mm/hr)
 Summer Storm "A" (April 27, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

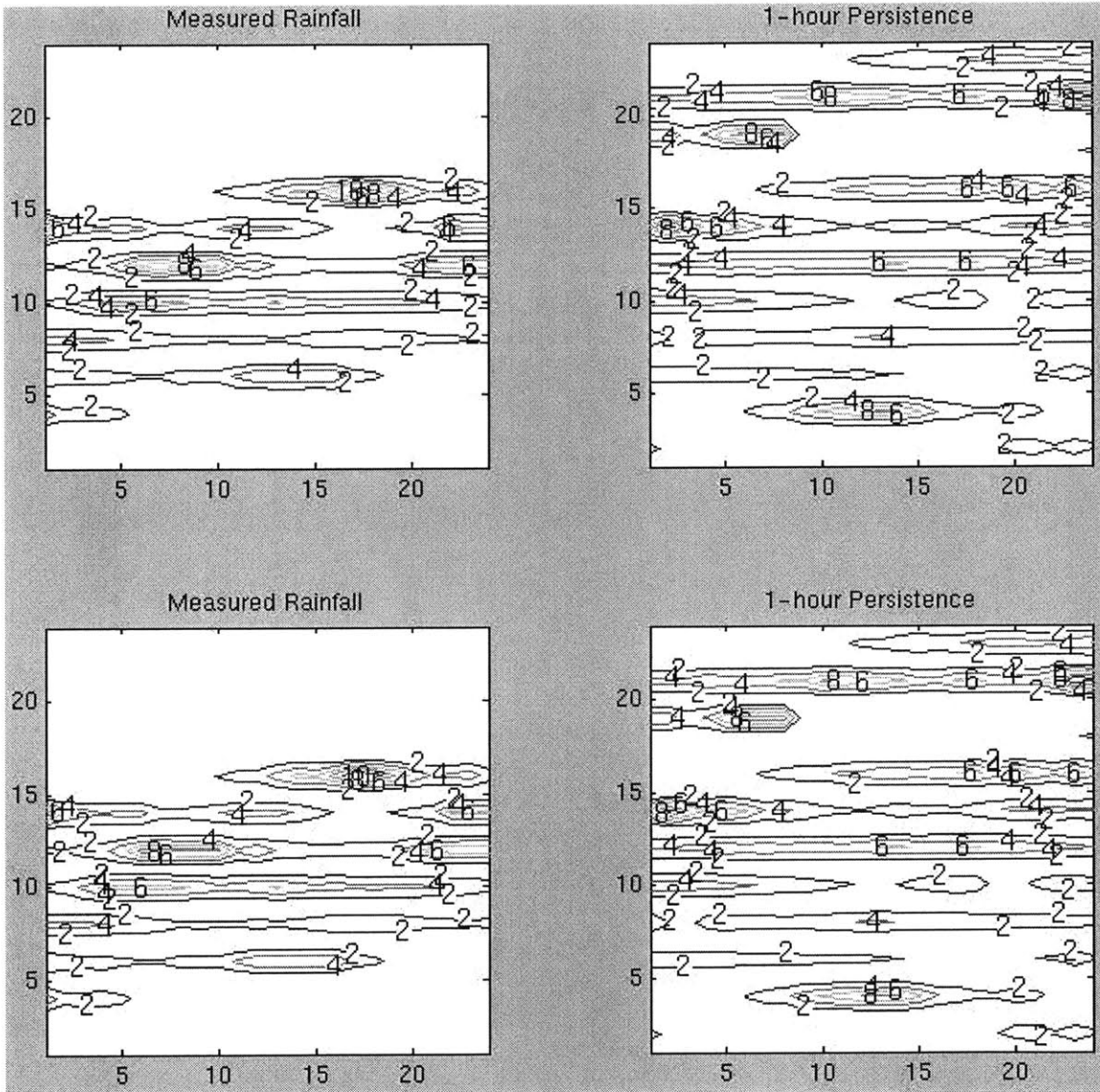


Figure B-1 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (and then onto the next figure), we see the same contour plots, but with the lower contour values successively removed. For this particular storm, the top and the bottom set of figures are identical, even though the latter (unlike the former) does not use contour values below 2 mm/hr. However, this behavior is not typical, as seen later.

**Figure B-2: Measured Precipitation and Persistence (mm/hr)
 Summer Storm "A" (April 27, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

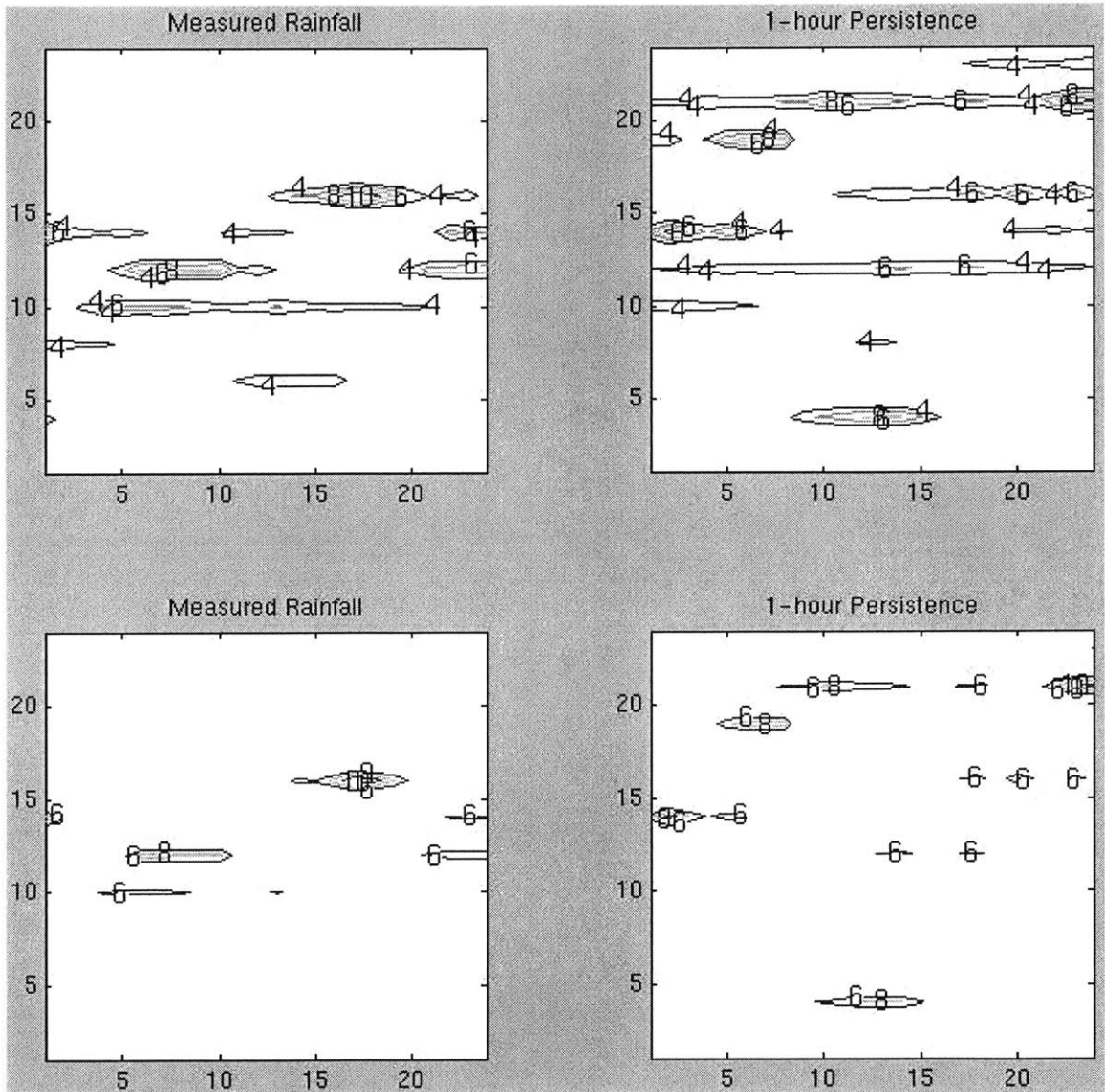


Figure B-2 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (starting from the preceding figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-3: Persistence Errors (mm/hr)
 Summer Storm "A" (April 27, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

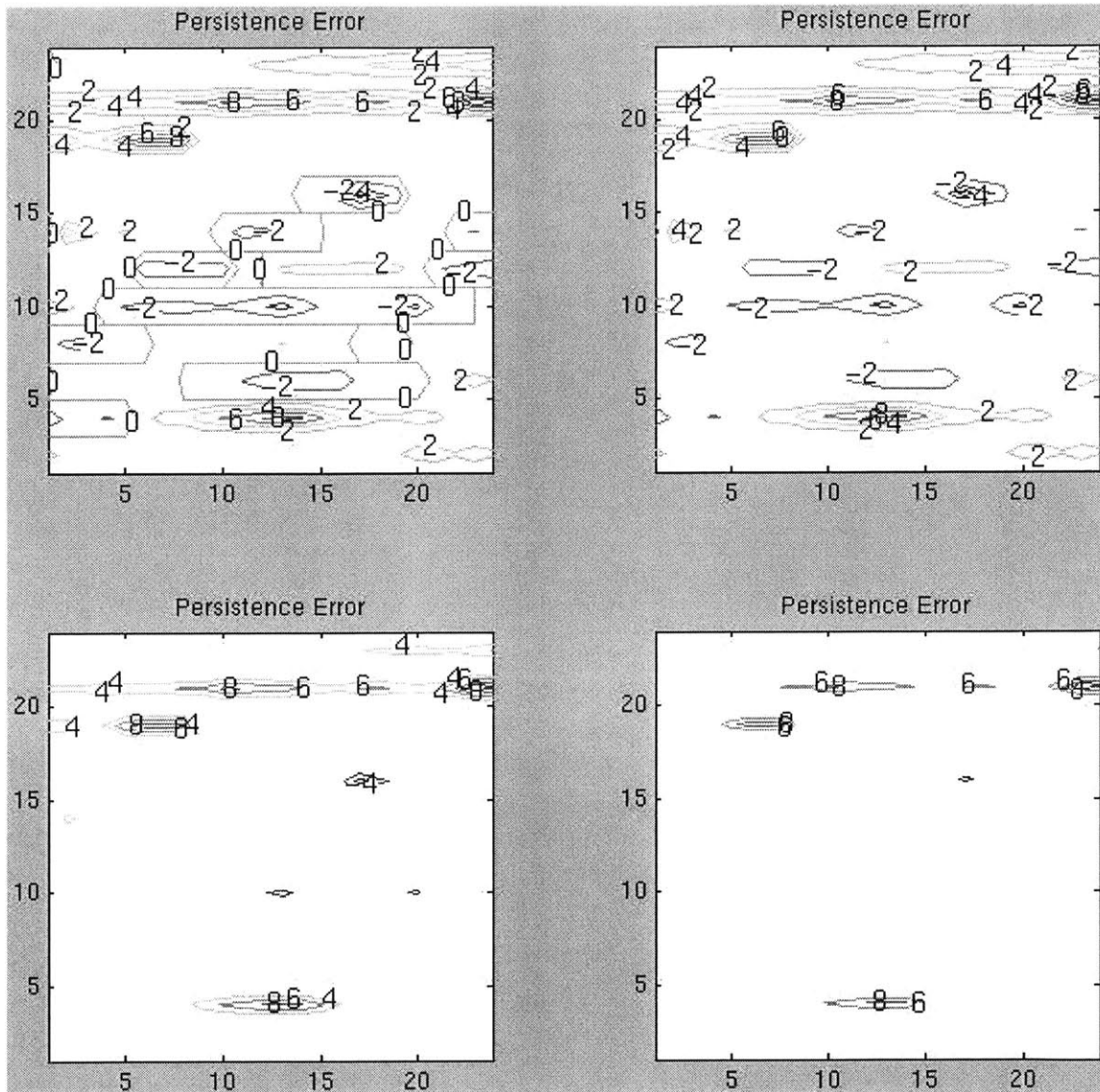


Figure B-3 shows the errors from persistence at 1-hour lead times, which is the same as the difference between successive hourly measured precipitation maps. As we move from top left to bottom right, we see the same contour plots, but with the lower contour values successively removed.

**Figure B-4: Measured Precipitation and Persistence (mm/hr)
 Summer Storm "B" (May 4-5, 1999)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

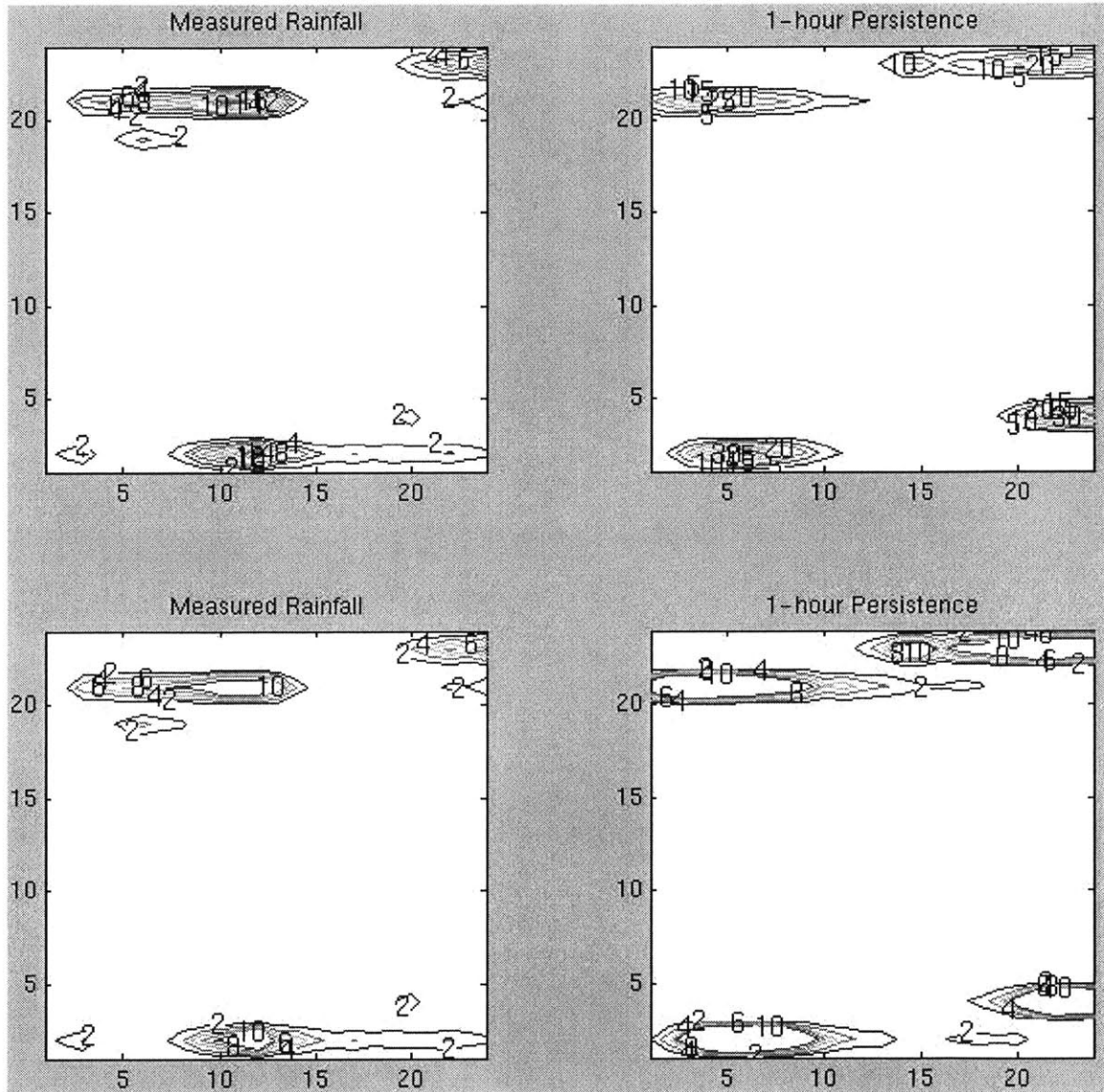


Figure B-4 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (and then onto the next figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-5: Measured Precipitation and Persistence (mm/hr)
 Summer Storm "B" (May 4-5, 1999)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

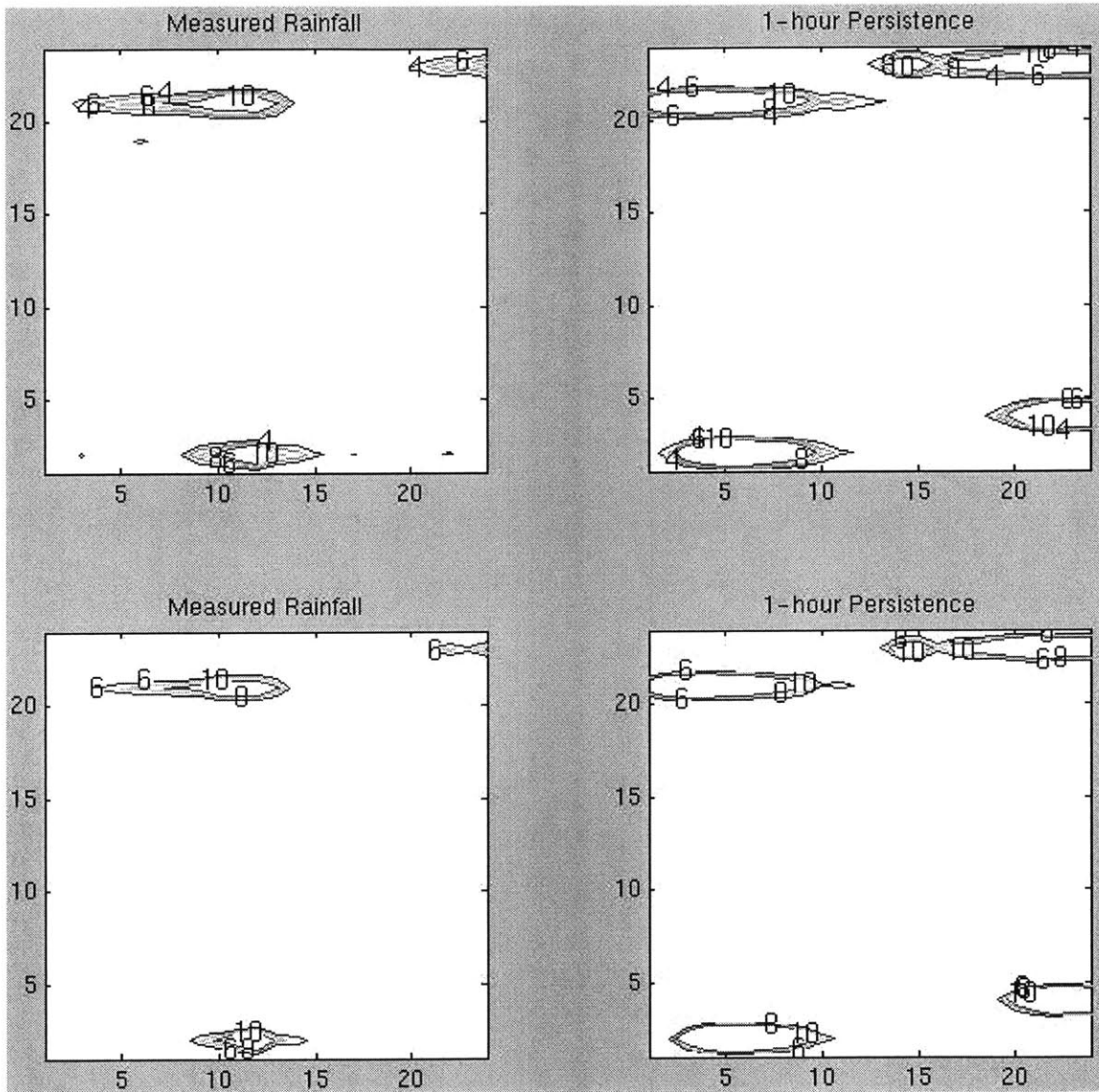


Figure B-5 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (starting from the preceding figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-6: Persistence Errors (mm/hr)
 Summer Storm "A" (May 4-5, 1999)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

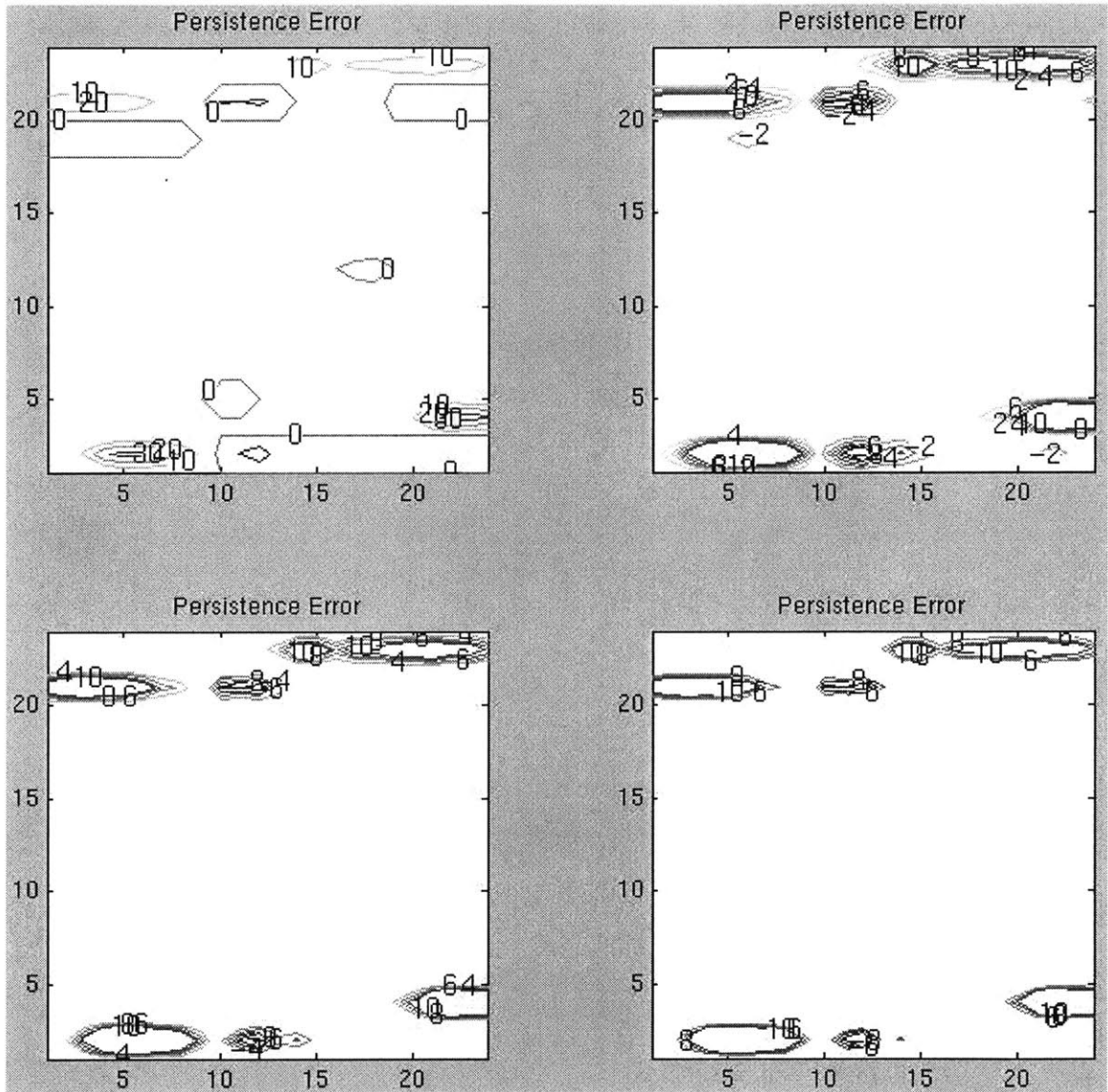


Figure B-6 shows the errors from persistence at 1-hour lead times, which is the same as the difference between successive hourly measured precipitation maps. As we move from top left to bottom right, we see the same contour plots, but with the lower contour values successively removed.

**Figure B-7: Measured Precipitation and Persistence (mm/hr)
 Summer Storm "C" (June 16, 1999)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

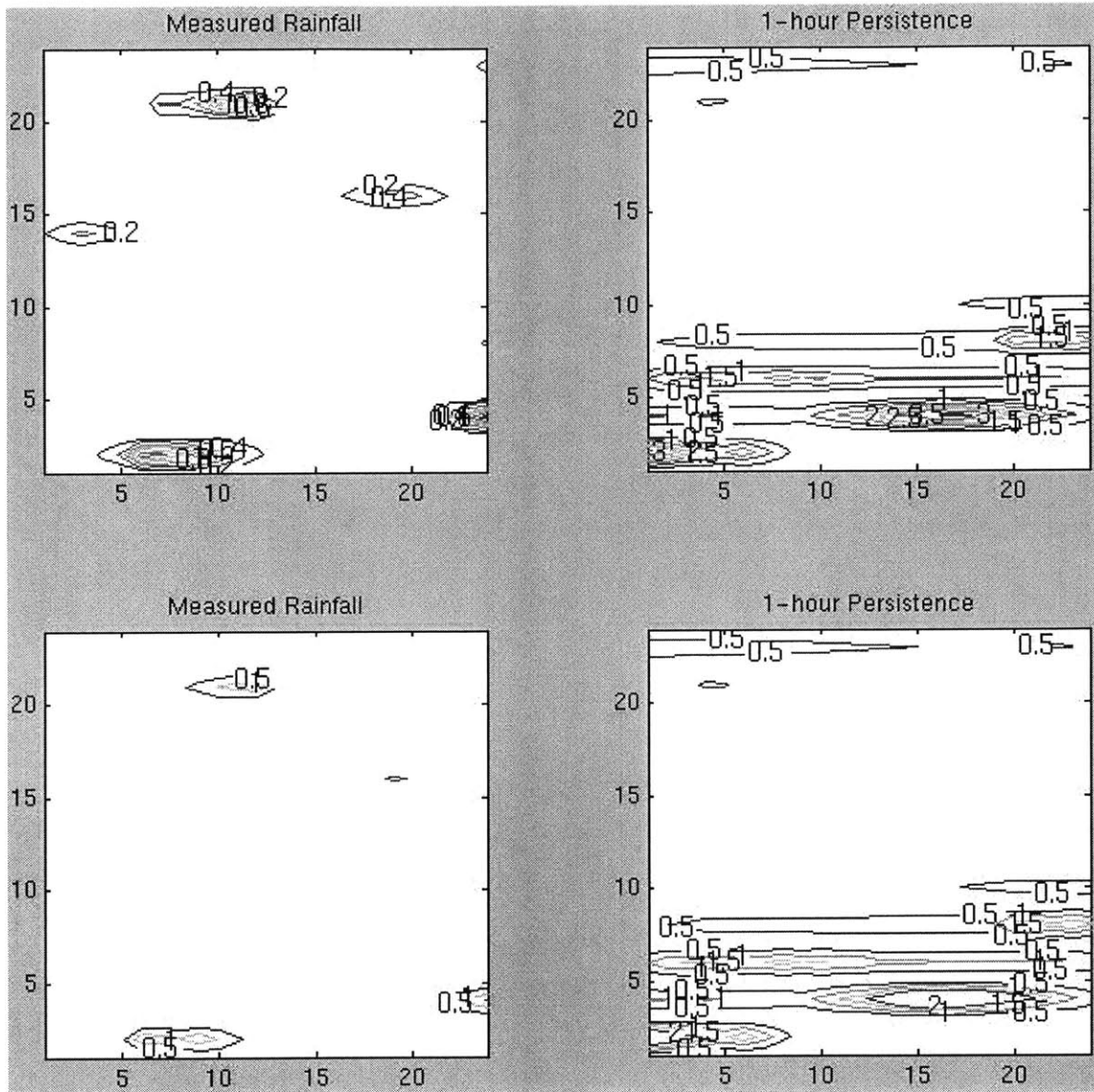


Figure B-7 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (and then onto the next figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-8: Measured Precipitation and Persistence (mm/hr)
 Summer Storm "C" (June 16, 1999)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

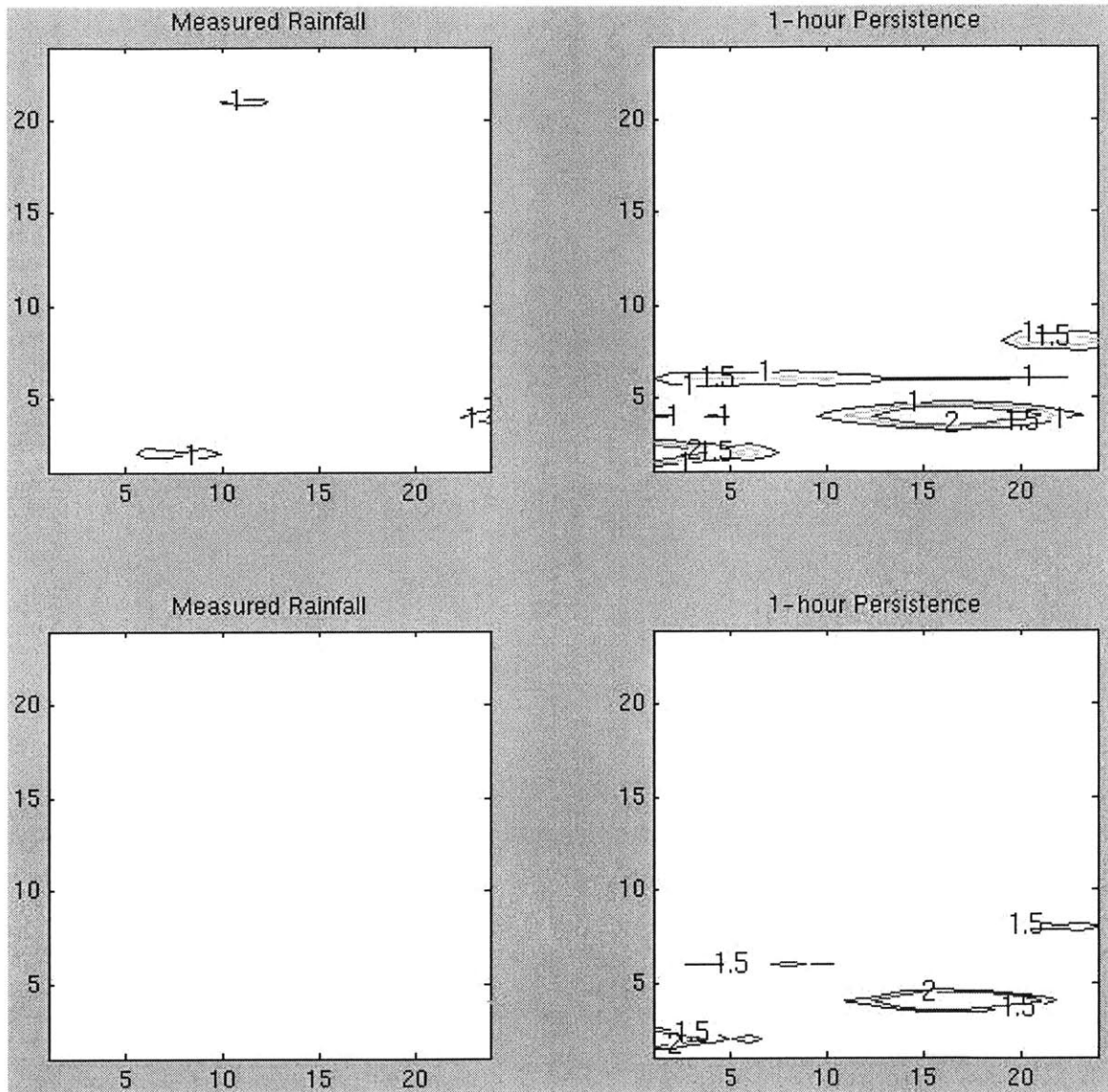


Figure B-8 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (starting from the preceding figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-9: Persistence Errors (mm/hr)
 Summer Storm "C" (June 16, 1999)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

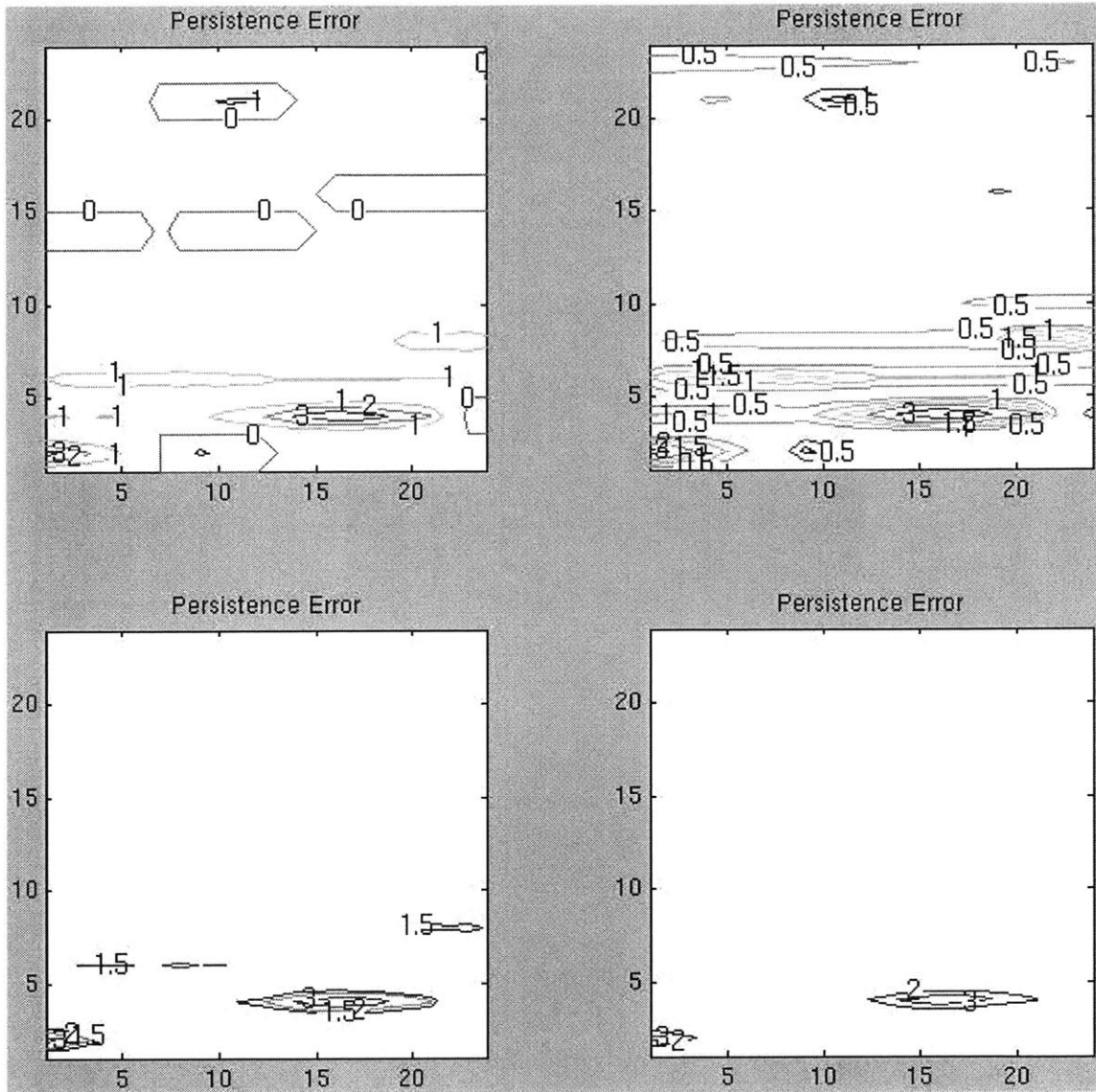


Figure B-9 shows the errors from persistence at 1-hour lead times, which is the same as the difference between successive hourly measured precipitation maps. As we move from top left to bottom right, we see the same contour plots, but with the lower contour values successively removed.

**Figure B-10: Measured Precipitation and Persistence (mm/hr)
 Winter Storm "D" (February 20, 1997)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

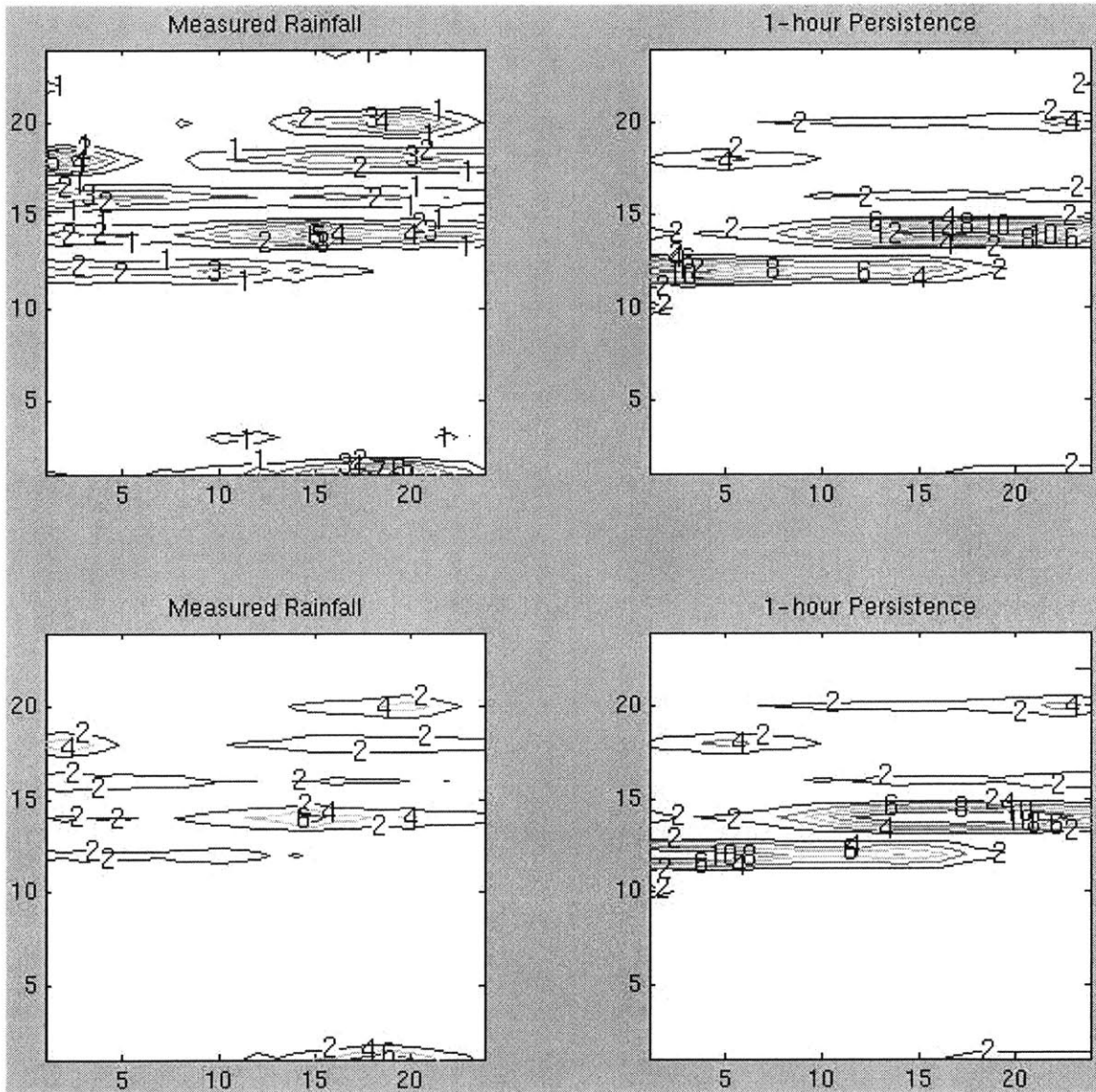


Figure B-10 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (and then onto the next figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-11: Measured Precipitation and Persistence (mm/hr)
 Winter Storm "D" (February 20, 1997)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

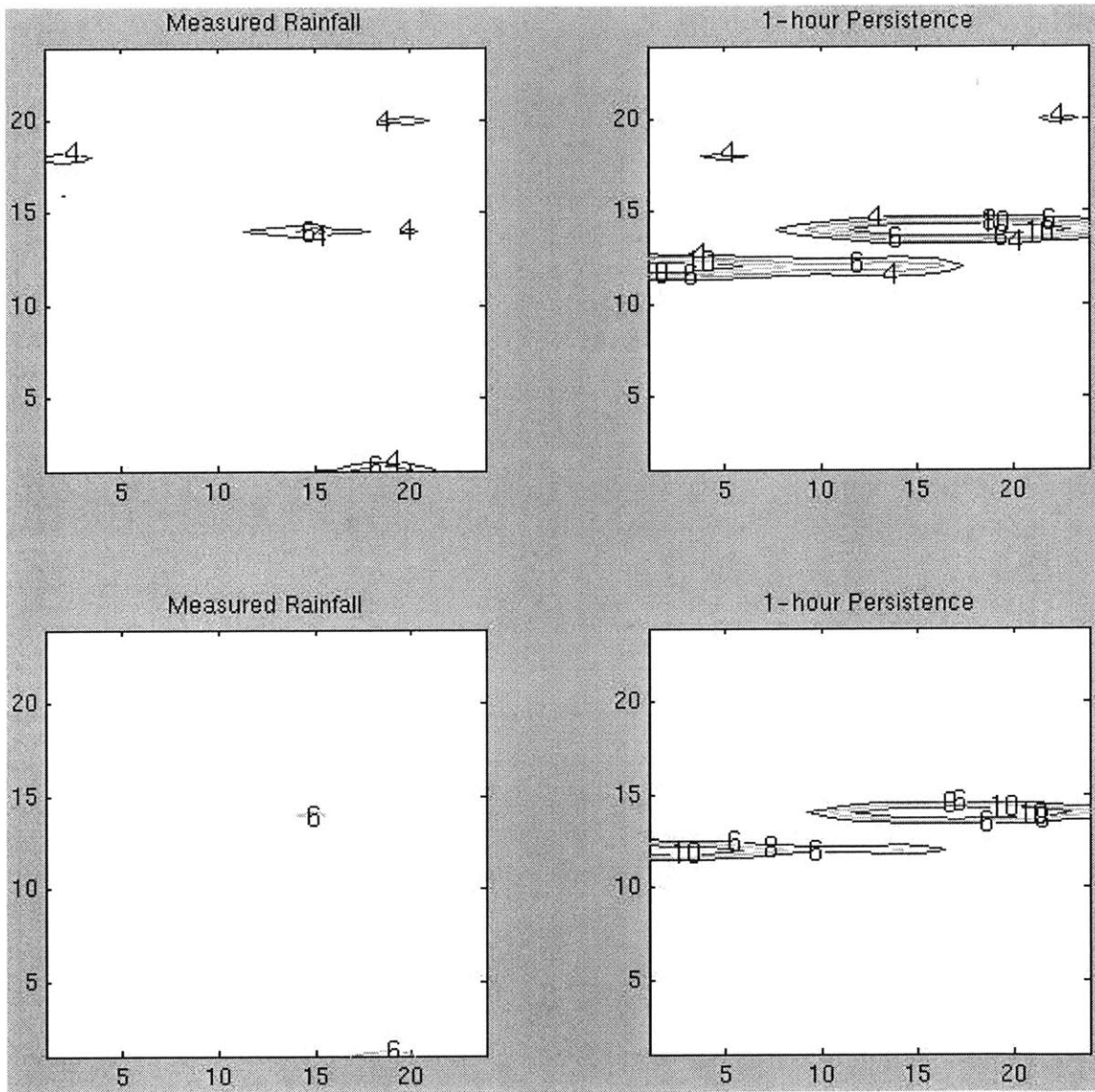


Figure B-11 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (starting from the preceding figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-12: Persistence Errors (mm/hr)
Winter Storm "D" (February 20, 1997)
(X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

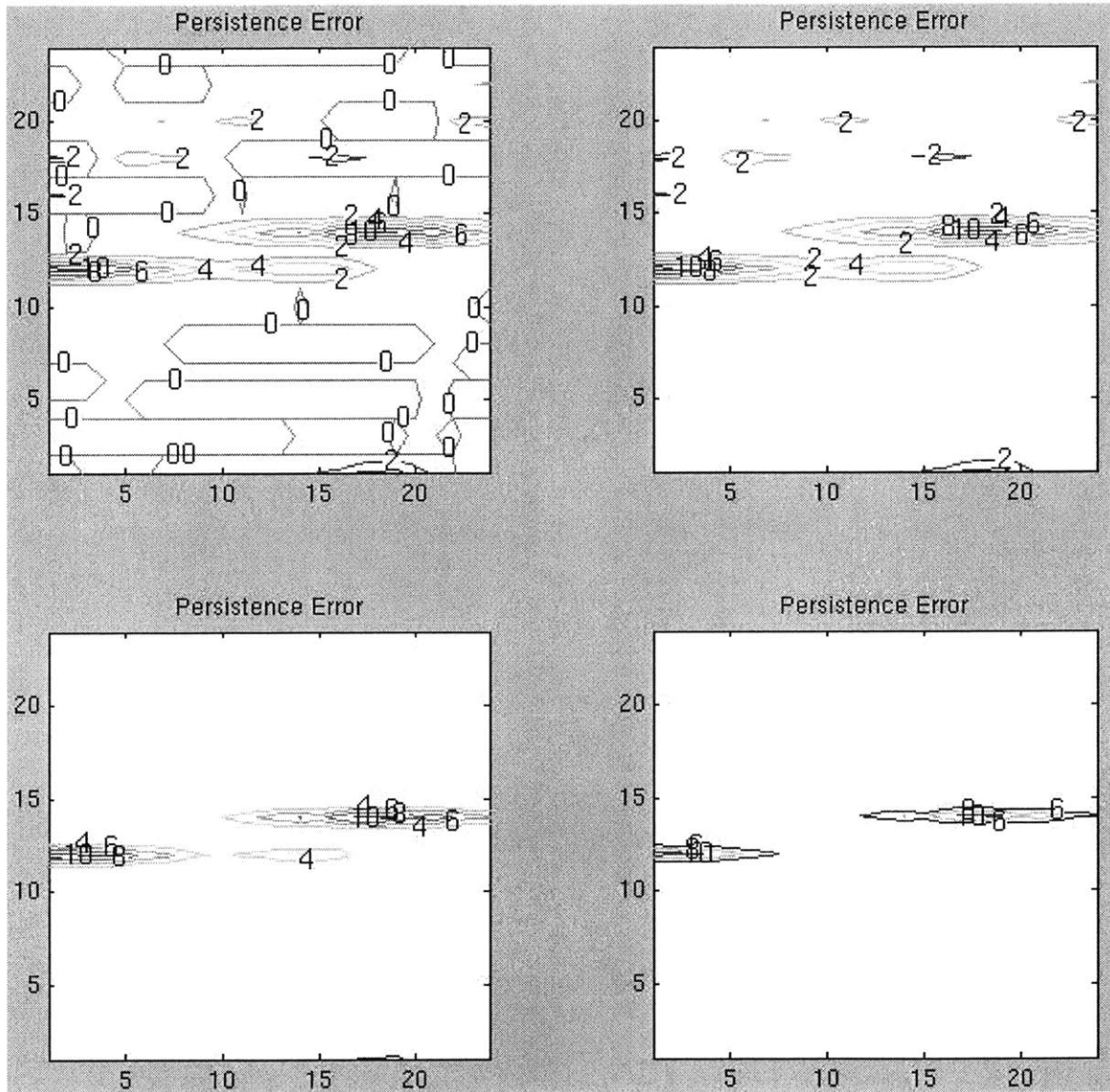


Figure B-12 shows the errors from persistence at 1-hour lead times, which is the same as the difference between successive hourly measured precipitation maps. As we move from top left to bottom right, we see the same contour plots, but with the lower contour values successively removed.

**Figure B-13: Measured Precipitation and Persistence (mm/hr)
 Winter Storm "E" (October 5, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

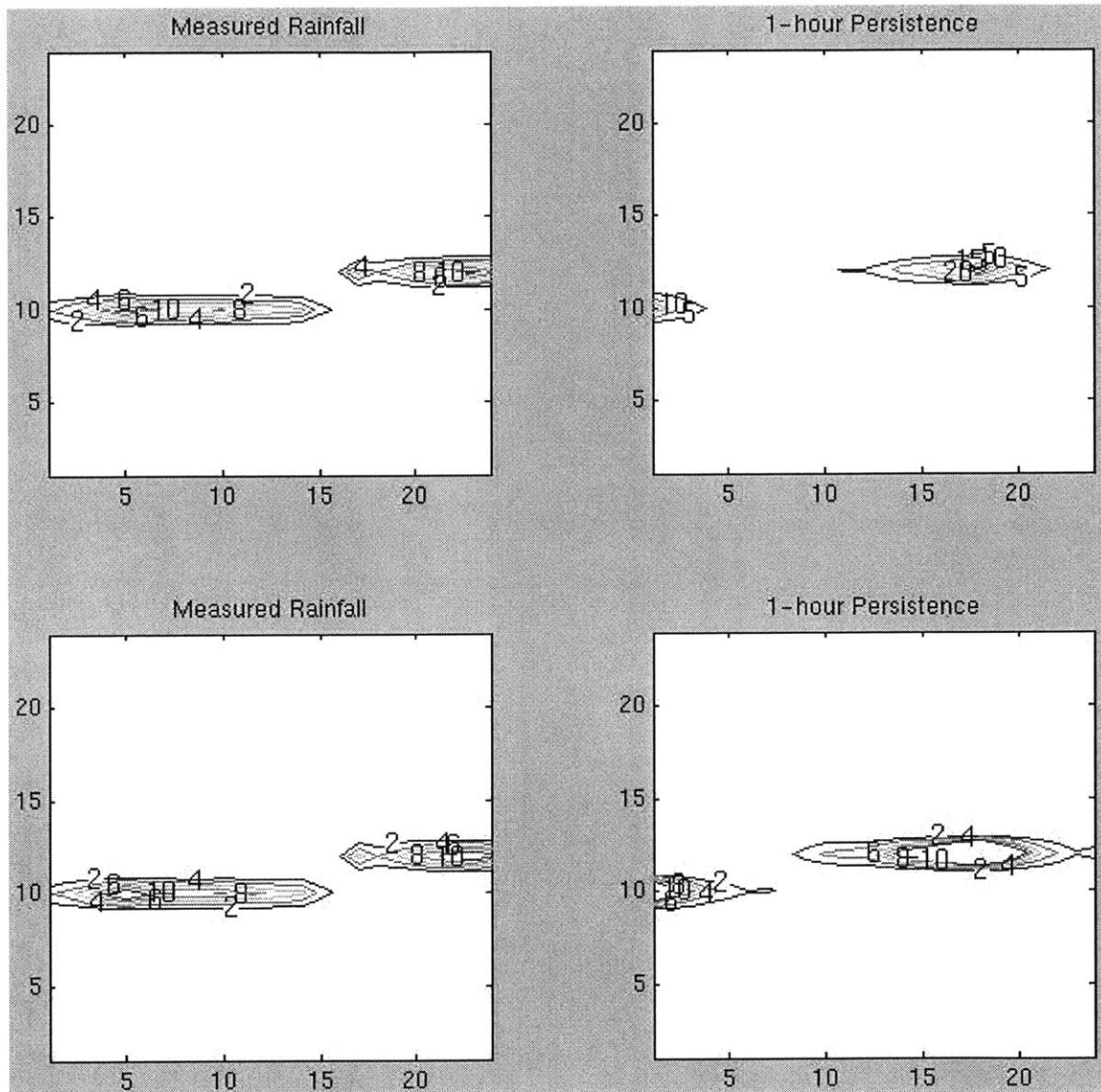


Figure B-13 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (and then onto the next figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-14: Measured Precipitation and Persistence (mm/hr)
 Winter Storm "E" (October 5, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

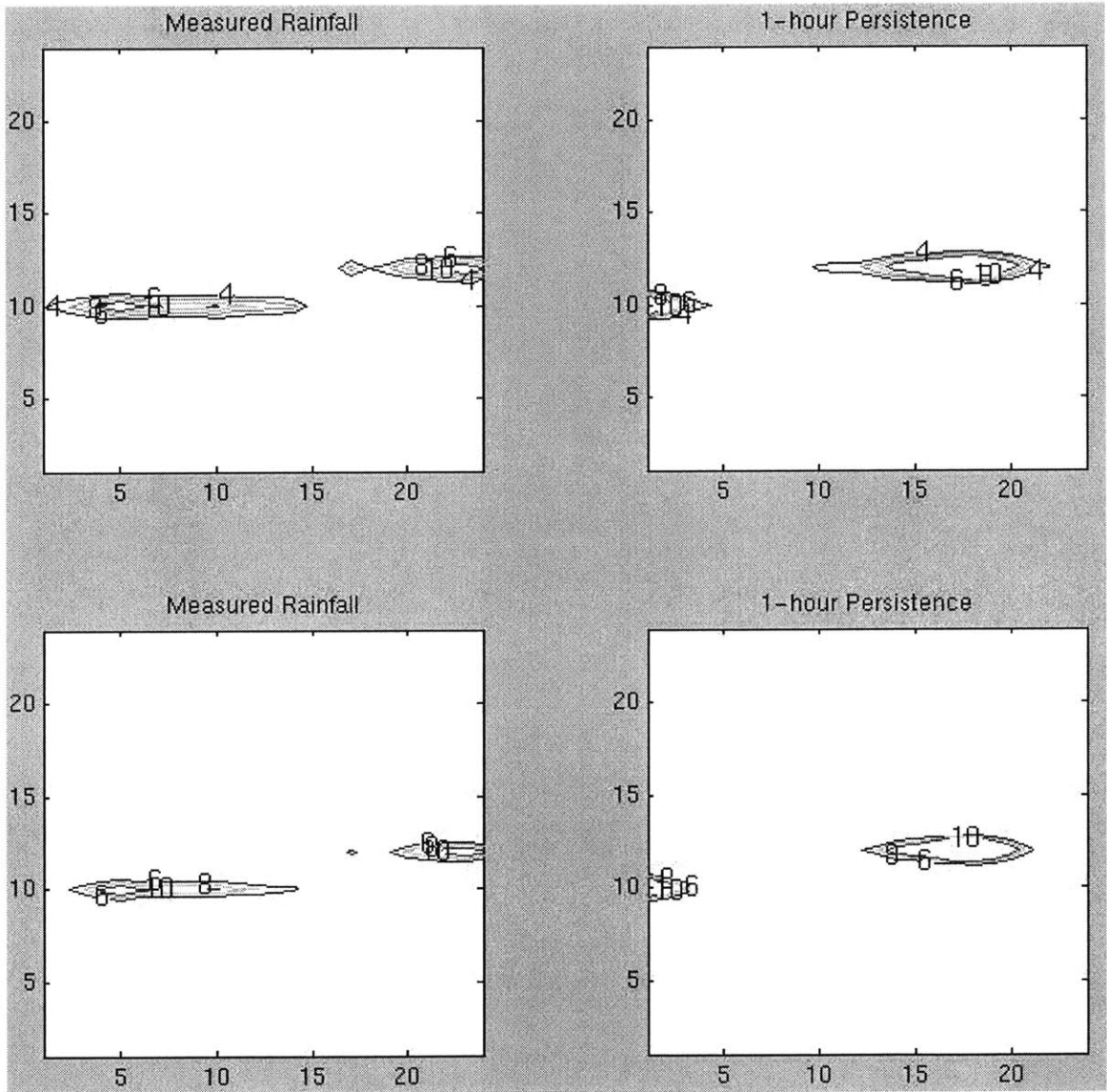


Figure B-14 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (starting from the preceding figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-15: Persistence Errors (mm/hr)
 Winter Storm "E" (October 5, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

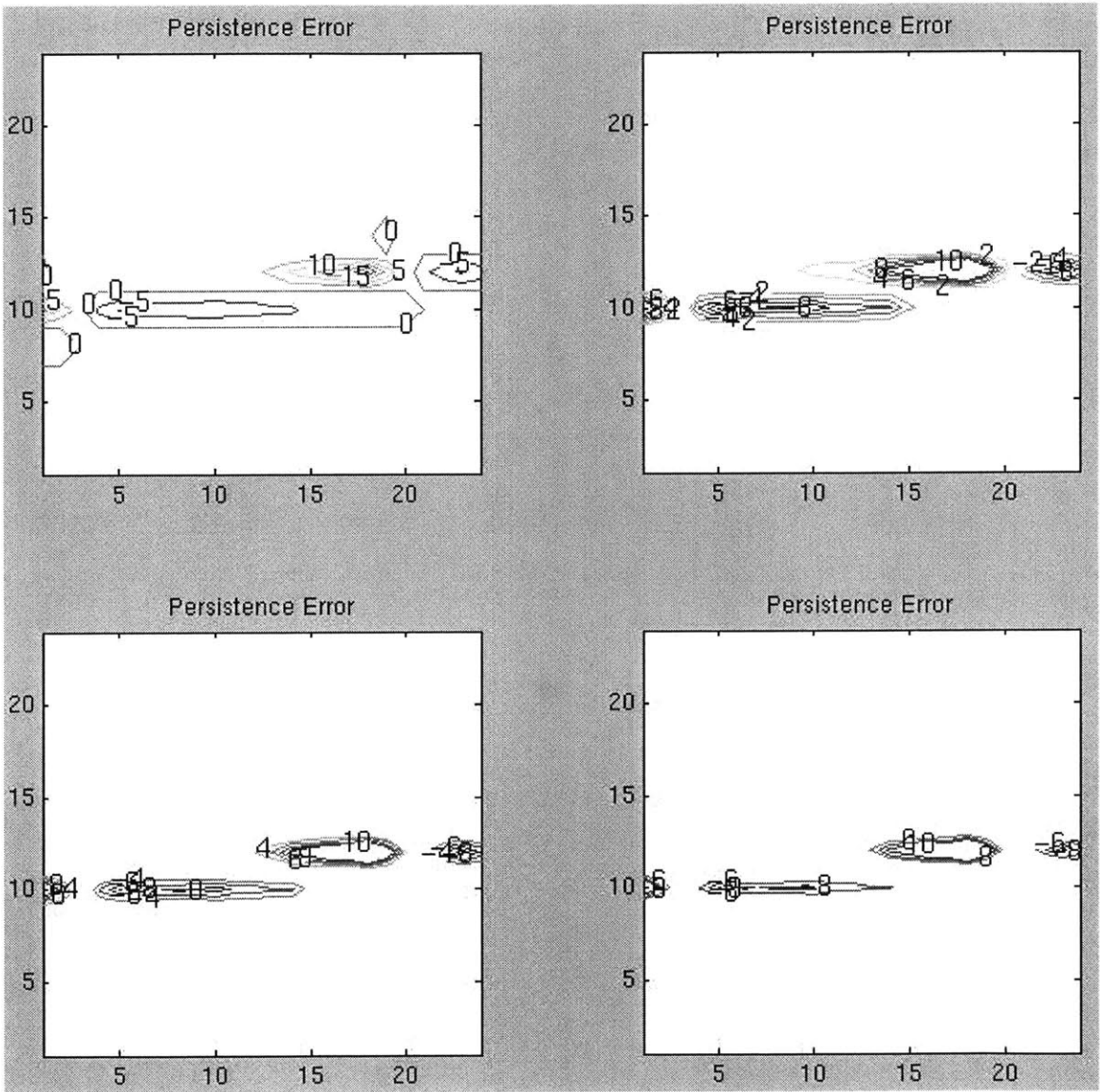


Figure B-15 shows the errors from persistence at 1-hour lead times, which is the same as the difference between successive hourly measured precipitation maps. As we move from top left to bottom right, we see the same contour plots, but with the lower contour values successively removed.

**Figure B-16: Measured Precipitation and Persistence (mm/hr)
 Winter Storm "F" (October 17-18, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

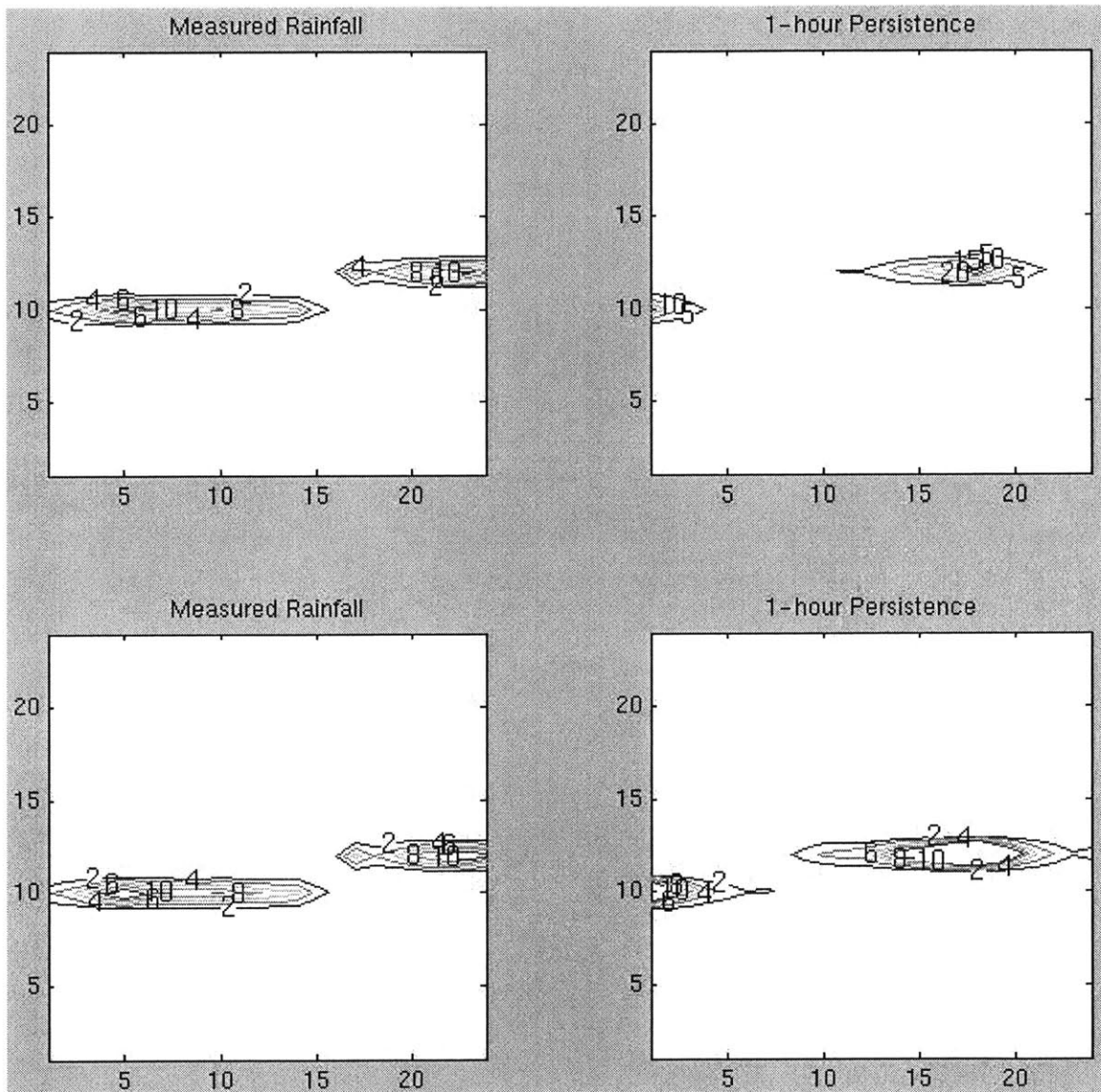


Figure B-16 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (and then onto the next figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-17: Measured Precipitation and Persistence (mm/hr)
 Winter Storm "F" (October 17-18, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

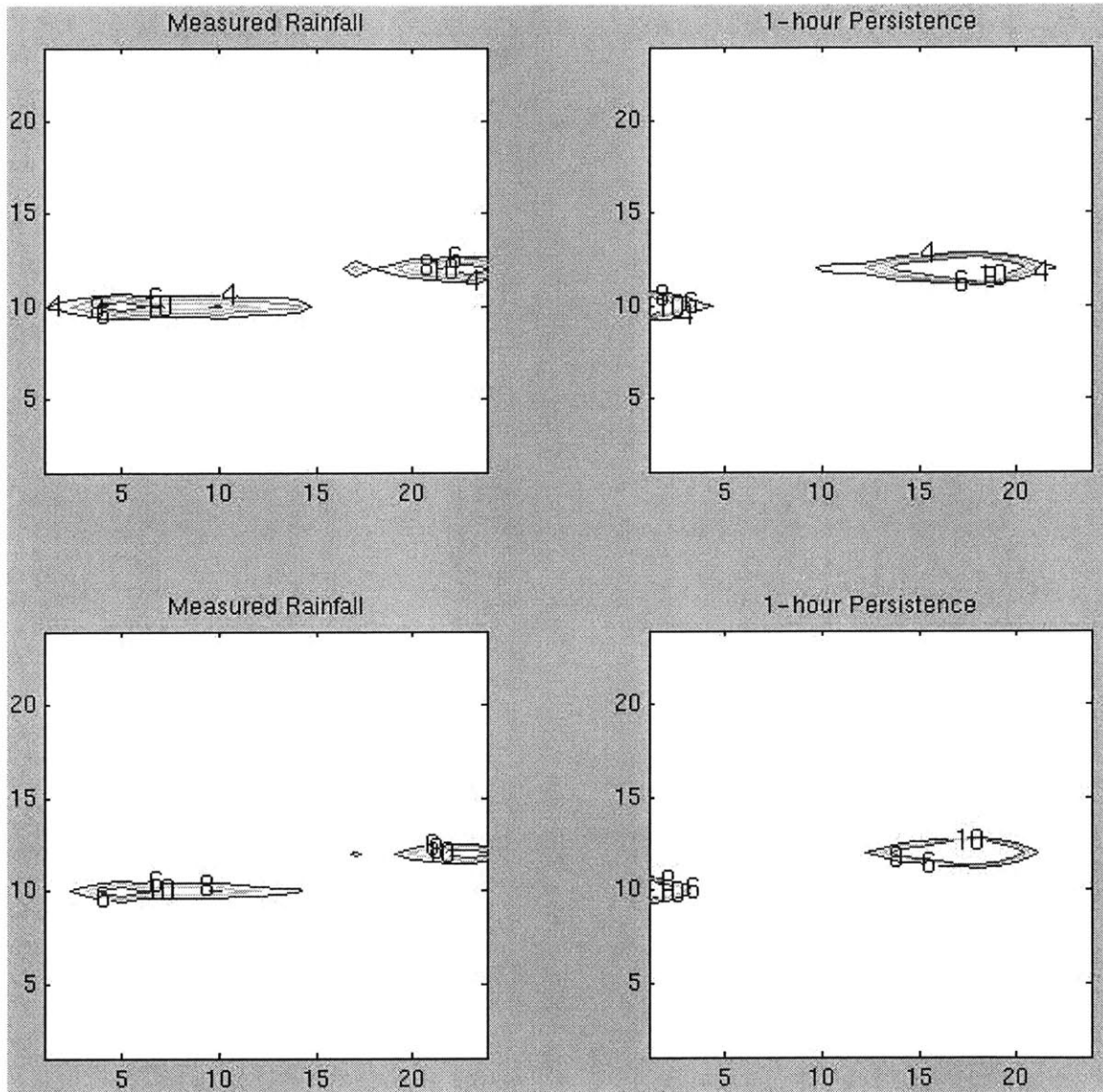


Figure B-17 shows the measured precipitation (plots on the left) at 1-hour lead time and the corresponding 1-hour persistence (plots on the right). As we move from top to bottom (starting from the preceding figure), we see the same contour plots, but with the lower contour values successively removed.

**Figure B-18: Persistence Errors (mm/hr)
 Winter Storm "F" (October 17-18, 1998)
 (X and Y axis Units are 4x4 km; Lead Time is 1 hour)**

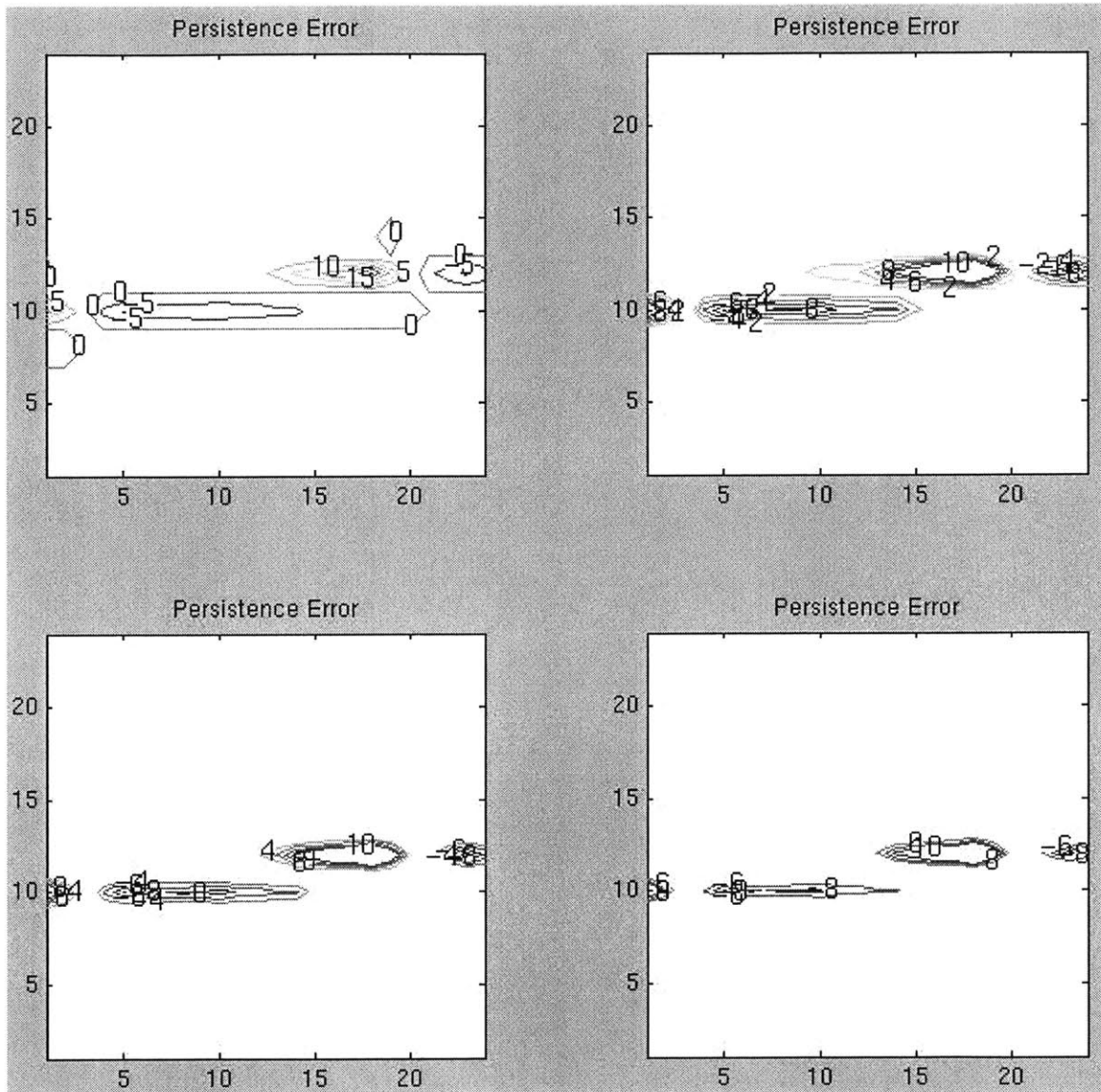


Figure B-18 shows the errors from persistence at 1-hour lead times, which is the same as the difference between successive hourly measured precipitation maps. As we move from top left to bottom right, we see the same contour plots, but with the lower contour values successively removed.

Bibliography

Andrieu, H., M. N. French, V. Thauvin, and W. F. Krajewski [1996]. "Adaptation and application of a quantitative rainfall forecasting model in a mountainous region." Journal of Hydrology. 184: 243-259.

Antolik, M. S. [2000]. "An overview of the National Weather Service's centralized statistical quantitative precipitation forecasts." Journal of Hydrology. 239: 306-337.

Arakawa, A. and V. R. Lamb [1977]. "Computational design of the basic dynamical processes of the UCLA general circulation model." Methods in Computational Physics. 17: 173-265.

Austin, G. L. and A. Bellon [1974]. "The use of digital weather records for short-term precipitation forecasting." Quarterly Journal of the Royal Meteorological Society. 100: 658-664.

Barros, A. P., R. J. Kuligowski and J. T. Ostrowski [1996]. "Quantitative precipitation forecasting using artificial networks." UCAR Final Outreach Report. S95-59721.

Bauman, W. H. and S. Businger [1996]. "Nowcasting for space shuttle landings at Kennedy Space Center, Florida." Bulletin of the American Meteorological Society. 77: 2295-2305.

Benediktsson, J. A., P. H. Swain and O. K. Ersoy [1990]. "Neural Network Approaches versus Statistical Methods in Classification of Multisource Remote Sensing Data." IEEE Transactions on Geoscience and Remote Sensing. 28(4): 540-552.

Beven, K. J. and A. M. Binley [1992]. "The future of distributed models: Model calibration and uncertainty prediction." Hydrologic Processes. 6: 279-298.

Bindlish R. and A. P. Barros [2000]. "Disaggregation of Rainfall for one-way Coupling of Atmospheric and Hydrological Models in Regions of Complex Terrain." Global and Planetary Change. 25.

Bishop, C. M. [1996]. Neural Networks for Pattern Recognition. Oxford University Press.

Black, T. L. [1994]. "The use of NMC Mesoscale Eta Model: Description and Forecast Examples." Weather and Forecasting.

Bond, J. E., K. A. Browning, and C. G. Collier [1981]. "Estimates of surface gust speeds using radar observation of showers". Meteorological Magazine. 110: 29-40.

Bras, R. L. and I. Rodriguez-Iturbe [1985]. Random functions and hydrology. Addison-Wesley.

Browning, K. A. [1964]. "Airflow and precipitation trajectories within severe local storms which travel to the right of the winds." Journal of the Atmospheric Sciences. 21: 634-639.

Browning, K. A. [1979] "The FRONTIERS plan: A strategy for using radar and satellite imagery for very-short-range precipitation forecasting." The Meteorological Magazine. 108: 161-184.

Browning, K. A. and C. G. Collier [1989]. "Nowcasting of precipitation systems." Reviews of Geophysics. 27(3).

Brubaker, K. L. and D. Entekhabi [1996]. "Analysis of feedback mechanisms in land-atmosphere interaction." Water Resources Research. 32(5): 1343-1357.

Burlando, P., A. Montanari and R. Ranzi [1996]. "Forecasting of storm rainfall by combined use of radar, rain gages and linear models." Atmospheric Research. 42: 199-216.

Burlando, P., R. Rosso, L. Cadavid and J. D. Salas [1993]. "Forecasting of short-term rainfall using ARMA models." Journal of Hydrology. 144: 193-211.

Charnoboy, E. S., A. M. Matlin and J. P. Morgan [1994]. "Automated storm tracking for terminal air traffic control." The Lincoln Laboratory Journal. 7: 427-448.

Chatfield, C. [1993]. "Neural Networks: Forecasting breakthrough or passing fad?" International Journal of Forecasting. 9: 1-3.

Cluckie, I. D., A. Lane, and J., Yuan [1999]. "Modeling large urban drainage systems in real time." Water Science and Technology. 39: 21-28.

Cobourn, W. G., L. Dolcine, M. French and M. C. Hubbard [2000]. "A comparison of nonlinear regression and neural network models for ground-level ozone forecasting." Journal of the Air and Waste Management Association. Accepted for publication 05/2000.

Collier, C. G. [1981]. "Objective rainfall forecasting using data from the United Kingdom weather radar network." Proceedings of the International Association of Meteorology and Atmospheric Physics Symposium, Hamburg, 25-28 August. Report ESA SP-165: 201-206.

Collier, C. G. [1991]. "The combined use of weather radar and mesoscale numerical model data for short-period rainfall forecasting." Hydrological Application of Weather Radar, edited by Ellis Horwood, 644 pp., Chichester, England.

Collier, C. G. [2001]. "Quantitative Precipitation Forecasting for hydrological applications in the United Kingdom." Presentation at the 7th International Precipitation Conference, Rockport, Maine.

Collier, C. G. and R. Kzyzysztowicz [2000]. "Preface: Quantitative precipitation forecasting". Journal of Hydrology. 239: 1-2.

Connor, J. T., R. D. Martin and L. E. Atlas [1994]. "Recurrent Neural Networks and Robust Time Series Prediction." IEEE Transactions on Neural Networks. 5(2).

Cotton, W. R. and R. A. Anthes [1989]. Storm and Cloud Dynamics. Academic Press.

Damrath, U., G. Doms, D. Fruhwald, E. Heise, B. Richter and J. Steppeler [2000]. "Operational quantitative precipitation forecasting at the German Weather Service." Journal of Hydrology. 239: 260-285.

Denoeux, T. and P. Rizand [1995]. "Analysis of radar images for rainfall forecasting using neural networks." Neural Computing and Applications. 3

Ding, X., T. Denoeux, and F. Helloco [1993]. "Tracking rain cells in radar images using multilayer neural networks." Proceedings of ICANN'93, S. Gielen and B. Kappen, eds., Springer-Verlag, London. 962-967.

Dolciné, L., H. Andrieu and M. N. French [1997]. "Rainfall forecasting in mountainous region using a weather radar and ground meteorological observations." Physics and Chemistry in the Earth. 22(3-4): 247-252.

Dolciné, L., H. Andrieu and M. N. French [1998]. "Evolution of a conceptual rainfall forecasting model from observed and simulated rain events." Hydrology and Earth System Science. 2(2-3): 173-182.

Dolcine, L., H. Andrieu, M. N. French and J. D. Creutin [2000]. "Implementation considerations of a conceptual precipitation model." Journal of Geophysical Research. 105(D2): 2291-2297.

Dorffner G., Schellner K., and Prem E. [2001]. "Regularized Gaussian Mixture Models for Effective Short-Term Forecasting of Rainfall Pattern", Elektrotechnik und Informationstechnik, 118(7/8)371-378.

Droegemeier, K. and Co-Authors [2000]. "Hydrological Aspects of Weather Prediction & Flood Warnings". Bulletin of the American Meteorological Society. 81(11).

Du, J., S. J. Mullen and F. Sanders [1997]. "Short Range Ensemble Forecasting of Quantitative Precipitation." Monthly Weather Review. October.

Duda, R. O. and R. H. Blackmer [1972]. "Application of pattern recognition techniques to digitised weather radar." Stanford Research Institute. Menlo Park, California.

Dumais, R. E., and K. C. Young [1995]. "Using a self learning algorithm for single-station Quantitative Precipitation Forecasting in Germany." Weather and Forecasting. 10: 105-113.

Ebert, E. E. and J. L. McBride [2000]. "Verification of precipitation in weather systems: determination of systematic errors." Journal of Hydrology. 239: 179-202.

Einfalt, T., T. Denoeux and G. Jacquet [1990]. "A radar rainfall forecasting method designed for hydrological purposes." Journal of Hydrology. 114: 229-244.

Elsner, J. B. and A. A. Tsonis [1992]. "Nonlinear prediction, chaos, and noise." Bulletin of the American Meteorological Society. 73: 49-60.

Eltahir, E. A. B. and J. S. Pal [1996]. "Relationship between surface conditions and subsequent rainfall in convective storms." Journal of Geophysical Research. 101: 26237-26245.

Entekhabi, D. and P. S. Eagleson [1989]. "Land surface hydrology parameterization for atmospheric General Circulation models including subgrid scale spatial variability." Journal of Climate. 2(8): 816-831.

Entekhabi, D. E., R. L. Bras, E. A. Williams and C. E. Forest [1997]. "Improving Hydrometeorologic Hazards Prediction using Quantitative Precipitation Forecasts." Unpublished Notes, Massachusetts Institute of Technology, Cambridge, MA..

Flexer A. [1995]: "Connectionists and Statisticians, Friends or Foes?", in Mira J. and Sandoval F.(eds.), From Natural to Artificial Neural Computation, Proc. of International Workshop on Artificial Neural Networks, Malaga-Torremolinos, Spain, Springer, LNCS 930, pp. 454-461.

Frankel, D. S., R. J. Olsen, S. L. Frankel and S. W. Chisholm [1989]. "Use of a Neural Net Computer System for Analysis of Flow Cytometric Data of Phytoplankton Population." Cytometry. 10: 540-550.

Frankel, D. S., J. S. Draper, J. E. Peak and J. C. McLeod [1993]. "Artificial Intelligence Needs Workshop: 4-5 November 1993, Boston, Massachusetts." Bulletin of the American Meteorological Society. 76(5): 728-738.

French, M. N. [1990]. "Real Time Forecasting of Rainfall in Space and Time." Master's Thesis. Ralph M. Parsons Laboratory, Department of Civil and Environmental Engineering. Massachusetts Institute of Technology, Cambridge, MA 02139.

French, M. N., W. F. Krajewski and R. R. Cuykendall [1992]. "Rainfall forecasting in space and time using a neural network." Journal of Hydrology. 137: 1-31.

French, M. N. and W. F. Krajewski [1994a]. "A model for real-time quantitative precipitation forecasting using remote sensing 1. Formulation." Water Resources Research. 30(4): 1075-1083.

French, M. N., H. Andrieu and W. F. Krajewski [1994b]. "A model for real-time quantitative precipitation forecasting using remote sensing 2. Case Studies." Water Resources Research. 30(4): 1084-1097.

French, M. N., H. Andrieu and W. F. Krajewski [1995]. "Uncertainty in vertically integrated liquid water content due to radar reflectivity observation error." Journal of Atmospheric and Ocean Technology. 12(2): 404-409.

Fritsch, J. M., and Co-Authors [1998]. "Quantitative precipitation forecasting: Report of the Eighth Prospectus Development Team, U.S. Weather Research Program." Bulletin of the American Meteorological Society. 79: 285-299.

Garrote, L. and R. L. Bras [1995]. "A distributed model for real-time flood forecasting using digital elevation models." Journal of Hydrology. 167.

Gaudet, B. and W. R. Cotton [1998]. "Statistical characteristics of a real-time precipitation forecasting model." Weather and Forecasting. 13: 966-982.

Gelhar, L. [1994]. Stochastic Subsurface Hydrology. Prentice Hall.

Georgakakos, K. P. [2000]. "Covariance propagation and updating in the context of real-time radar data assimilation by quantitative precipitation forecast models." Journal of Hydrology. 239: 115-129.

Georgakakos, K. P. and R. L. Bras [1984a]. "A Hydrologically Useful Station Precipitation Model 1. Formulation." Water Resources Research. 20(11): 1585-1596.

Georgakakos, K. P. and R. L. Bras [1984b]. "A Hydrologically Useful Station Precipitation Model 2. Case Studies." Water Resources Research. 20(11): 1597-1610.

Georgakakos, K. P. and M. D. Hudlow [1984]. "Quantitative rainfall forecast techniques for use in hydrologic forecasting." Bulletin of the American Meteorological Society. 65(11): 1180-1200.

Gershenfeld, N. A. and A. S. Weigend [1994]. "The Future of Time Series: Learning and Understanding." Time Series Prediction: Forecasting the Future and Understanding the Past. Chapter 1. Edited by A. S. Weigend and N. A. Gershenfeld. Addison-Wesley. Reading, MA.

Gillispie, M. [1993a]. "The Use of Neural Networks for making Quantitative Precipitation Forecasts." Western Region Technical Attachment SR/SSD 93-42. National Weather Service. Fort Worth, Texas.

Gillispie, M. [1993b]. "Neural Networks: A New approach to Quantitative Precipitation Forecasting in Texas." Weather Service Forecast Office. Sioux Falls.

Gillispie, M. [1993c]. "Creating Operational Neural Networks." National Weather Service Forecast Office. Sioux Falls, SD.

Giorgi, F. and R. Avissar [1997]. "Representation of heterogeneity effects in earth system modeling: Experience from land surface modeling." Reviews of Geophysics. 35: 413-438.

Glahn, H. R. and D. L. Lowry [1972]. "The use of Model Output Statistics (MOS) in objective weather forecasting." Journal of Applied Meteorology. 11: 1203-1211.

Golding, B. W. [2000]. "Quantitative precipitation forecasting in the UK." Journal of Hydrology. 239: 286-305.

Greco, M. and W. F. Krajewski [2000a]. "A large-sample investigation of statistical procedures for short-term quantitative precipitation forecasting." Journal of Hydrology. 239: 69-84.

Greco, M. and W. F. Krajewski [2000b]. "Simulation study of the effects of model uncertainty in variational assimilation of radar data on rainfall forecasting." Journal of Hydrology. 239: 85-96.

Grieger, B. and M. Latif [1994]. "Reconstruction of the El Nino attractor with neural networks." Climate Dynamics. 10: 267-276.

Hall, T. [1996]. "BRAINMAKER: A new approach to quantitative and probability of precipitation forecasting." NWS Southern Region Technical Attachment SR/HSD 96-2, NOAA, U.S. Department of Commerce. 7 pp.

Hartmann, D. L. [1994]. Global Physical Climatology. Academic Press.

Hetz, J., A. Krogh and R. G. Palmer [1991]. Introduction to the Theory of Neural Computation. Addison-Wesley.

Hsu, K.-L., H. V. Gupta, X. Gao and S. Sorooshian [1999]. "Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation." Water Resources Research. 35(5): 1605-1618.

Hsu, K.-L., X. Gao, S. Sorooshian and H. V. Gupta [1997a]. "Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks." Journal of Applied Meteorology. 36: 1176-1190.

Hsu, K.-L., X. Gao, S. Sorooshian and H. V. Gupta [1997b]. "Rainfall Estimation From Satellite Infrared Imagery Using Artificial Neural Networks." Department of Hydrology and Water Resources. University of Arizona. HWR No. 97-010.

Ibbitt, R. P., R. D. Henderson, J. Copeland and D. S. Wratt [2000]. "Simulating mountain runoff with meso-scale weather model rainfall estimates: a New Zealand experience." Journal of Hydrology. 239: 19-32.

Islam, S., R. L. Bras and I. Rodriguez-Iturbe [1993a]. "A Possible Explanation for Low Correlation Dimension Estimates for the Atmosphere." Journal of Applied Meteorology. 32(2).

Islam, S., R. L. Bras and K. Emanuel [1993b]. "Predictability of mesoscale rainfall in the tropics." Journal of Applied Meteorology. 32(2): 297-310.

Jinno, K., A. Kawamura, R. Berndtsson, M. Larson, J. Niemczynowicz [1993]. "Real-time rainfall prediction at small space-time scales using a two-dimensional stochastic advection-diffusion model." Water Resources Research. 29(5): 1489-1504.

Johnson, E. R. and R. L. Bras [1980]. "Multivariate short-term rainfall prediction." Water Resources Research. 16: 173-185.

Jones, C. D. and B. Macpherson [1997]. "A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model." Meteorological Applications. 4: 269-277.

Jordan, M. I. and R. A. Jacobs [1994]. "Hierarchical Mixture of Experts and the EM Algorithm." Neural Computation. 6: 181-214.

Jordan, M. I. and C. M. Bishop [1996]. "Neural Networks". CRC Handbook of Computer Science. Edited by A. Tucker.

Kim, G. and A. P. Barros [2001]. "Quantitative flood forecasting using multisensor data and neural networks." Journal of Hydrology. 246(2001): 45-62.

Kitchen, M. and R. M. Blackall [1992]. "Representativeness errors in comparisons between radar and gauge measurements of rainfall." Journal of Hydrology. 134: 13-33.

Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil and S. Surendran [1999]. "Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble." Science. 285.

- Krzystofowicz, R., W. J. Drzal, T. R. Drake, J. C. Weyman and L. A. Giordano [1993]. "Probabilistic Quantitative Precipitation Forecasts." Weather and Forecasting. 8(4).
- Krzystofowicz, R. [1999]. "Bayesian theory of probabilistic forecasting via deterministic hydrologic model." Water Resources Research. 35(9): 2739-2750.
- Kuligowski, R. J. and A. P. Barros [1999]. "High-resolution short-term quantitative precipitation forecasting in mountainous regions using a nested model." Journal of Geophysical Research. 194(D24): 31533-31564.
- Kuligowski, R. J. and A. P. Barros [1998a]. "Localized Precipitation Forecasts from a Numerical Weather Prediction Model Using Artificial Neural Networks." Weather and Forecasting. 13(4): 1194-1204.
- Kuligowski, R. J. and A. P. Barros [1998b]. "Experiments in short-term precipitation forecasting using artificial neural networks." Monthly Weather Review. 126: 470-482.
- Lee, T. H. and K. P. Georgakakos [1990]. "A two-dimensional stochastic-dynamical quantitative precipitation forecasting model." Journal of Geophysical Research. 95(D3): 2113-2126.
- Lee, H. T. and K. P. Georgakakos [1996]. "Operational rainfall prediction on meso- γ scales for hydrologic applications." Water Resources Research. 32(9): 987-1003.
- Lee, J., R. C. Weger, S. K. Sengupta and R. M. Welch [1990]. "A neural network approach to cloud classification." IEEE Transactions on Geoscience and Remote Sensing. 28(5): 846-855.
- Li, Q., R. L. Bras, and S. Islam [1995]. "Growth and decay of errors in numerical cloud models due to small initial perturbations and parameter changes." Journal of Applied Meteorology. 34(7): 1622-1632.
- Ligda, M. G. H., and W. A. Mayhew [1954]. "On the relationship between the velocities of small precipitation areas and geostrophic winds." Journal of Meteorology. 11: 421-423.
- Lindner, A. J. and A. S. Krien [1993]. "A neural network for forecasting heavy precipitation." 13th Conference on Weather Analysis and Forecasting, Vienna, VA, American Meteorological Society. 612-615.
- Lorenz, E. [1982]. "Atmospheric predictability experiments with a large numerical model." Tellus. 34: 505-513.
- Lowry, D. A. and H. R. Glahn [1976]. "An Operational Model for Forecasting Probability of Precipitation." Monthly Weather Review. 104(3): 221-232.

Ma, J., L. Xu, and M. I. Jordan [2000]. "Asymptotic convergence rate of the EM algorithm for gaussian mixtures." Neural Computation. 12: 288-290.

Mackay, D. J. C. [1991]. "Bayesian Methods for Adaptive Models." Ph.D. Dissertation. California Institute of Technology.

Mackay, D. J. C. [1992]. "Probable networks and plausible predictions - a review of practical Bayesian methods for supervised neural networks." Neural Computation. 4: 448-472.

Mackay, D. J. C. [1994]. "Bayesian non-linear modelling for the prediction competition." ASHRAE Transactions. 100(2): 1053-1062.

Marzban, C. and G. J. Stumpf [1996]. "A neural-network for tornado prediction based on Doppler radar-derived attributes." Journal of Applied Meteorology. 35: 617-626.

Matsoukas, C., S. Islam and R. Kothari [1999]. "Fusion of Radar and Raingauge Measurements for an Accurate Estimation of Rainfall." Journal of Geophysical Research. 104(D24): 31437-31450.

McCann, D. W. [1992]. "A neural-network short term forecast of significant thunderstorms." Weather and Forecasting. 7: 525-532.

McLaughlin, D., S. Islam, D. Entekhabi, M. French and R. Bras [1990]. "A distributed filtering approach to real-time rainfall forecasting." Preprint: Eighth Conference on Hydrometeorology. October. American Meteorological Society.

McLaughlin, D [1995]. "Recent developments in hydrologic data assimilation." Reviews of Geophysics, Supplement. 977-984.

Mecklenburg, S., J. Joss and W. Schmid [2000]. "Improving the nowcasting of precipitation in an Alpine region with an enhanced radar echo tracking algorithm." Journal of Hydrology. 239: 46-68.

Mesinger, F. [1996]. "Improvements in Quantitative Precipitation Forecasts with the Eta Regional Model at the National Centers for Environmental Prediction: The 48-km Upgrade." Bulletin of the American Meteorological Society. 77: 2637-2649.

Mills, T. C. [1990]. Time Series Techniques for Economists. Cambridge University Press.

Mozer, M. C. [1994]. "Neural Net Architectures for Temporal Sequence Processing", in A. S. Weigend and N. A. Gershenfeld (Eds.), Time Series Prediction: Forecasting the Future and Understanding the Past. 243-264.

Murphy, A.H. [1993]. "What is a good forecast? An essay on the nature of goodness in weather forecasting." Weather and Forecasting. 8: 281-293.

Nakakita, E., S. Ikebuchi, T. Kanamuri, M. Okuda, A. Yamaji and T. Takasao [1996]. "Short-term rainfall prediction model using a volume scanning radar and grid point value data from numerical weather prediction." Journal of Geophysical Research. 101: 26181-26197.

Navone, H. D. and H. A. Ceccato [1994]. "Predicting Indian monsoon rainfall: a neural network approach." Climate Dynamics. 10: 305-312.

Neal, R. M. [1994]. "Bayesian Learning for Neural Networks." Ph.D. Dissertation. Department of Computer Science. University of Toronto.

NOAA [1997]. "The End-to-End Forecast Process for Quantitative Precipitation Information: Hydrometeorological Requirements, Scientific Issues, and Service Concepts." National Weather Service - Office of Meteorology. US Department of Commerce.

NRC, National Research Council [1996]. "Towards a New National Weather Service: Assessment of Hydrologic and Hydrometeorological Operations and Services." A report of the National Research Council, Commission on Engineering and Technical Systems, National Weather Service Modernization Committee, December.

Olson, D. A., N. W. Junker and B. Korty [1995]. "Evaluation of 33 years of Quantitative Precipitation Forecasting at the NMC." Weather and Forecasting. 10: 498-511.

Park, S. K. [1999]. "Nonlinearity and predictability of convective rainfall associated with water vapor perturbations in a numerically simulated storm." Journal of Geophysical Research. 104(D24): 31575-31587.

Park, S. K. and K. K. Droegemeier [1999]. "Sensitivity Analysis of a Moist 1D Eulerian Cloud Model Using Automatic Differentiation." Monthly Weather Review. 127: 2180-2196.

Parsons, D. B. and P. V. Hobbs [1983]. "The mesoscale and microscale structure of cloud and precipitation in midlatitude cyclones, XI, Comparison between observational and theoretical aspects of rainbands." Journal of the Atmospheric Sciences. 40: 2377-2397.

Perica, S. and E. Foufoula-Georgiou [1996]. "A model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions." Journal of Geophysical Research. 101(D21): 26347-26361.

Peppler, R. A. and P. J. Lamb [1988]. "Tropospheric static stability and the central North American growing season rainfall." Monthly Weather Review. 117: 1156-1180.

Pessoa, M. L., R. L. Bras and E. R. Williams [1993]. "Use of weather radar for flood forecasting in the Sieve river basin: a sensitivity analysis." Journal of Applied Meteorology. 32: 462-475.

Reyes-Aldasoro, C. C., A. R. Ganguly, G. Lemus, A. Gupta [1999]. "A hybrid model based on dynamic programming, neural networks, and surrogate value for inventory optimisation applications." Journal of the Operational Research Society. 50(1).

Ripley, B. D. [1993]. "Statistical Aspects of Neural Networks", in Barndorff-Nielsen, O. E., Jensen, K. L., Kendall, W. S. (Eds.), Networks and Chaos: Statistical and Probabilistic Aspects, London: Chapman and Hall.

Rodriguez-Iturbe, I. and P. S. Eagleson [1987]. "Mathematical Models of Rainstorm Events in Space and Time." Water Resources Research. 23(1): 181-190.

Rodriguez-Iturbe, I., B. F. Power, M. B. Sharifi and K. P. Georgakakos [1989]. "Chaos in Rainfall." Water Resources Research. 25(7): 1667-1675.

Rogers, R. R. and M. K. Yau [1989]. A Short Course in Cloud Physics. Pergamon.

Rogers, E., D. Deaven and G. J. DiMego [1995]. "The regional analysis system for the operational "early" Eta model: original 80-km configuration and recent changes." Weather and Forecasting. 10: 810-825.

Rogers, E., T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. Baldwin, N. W. Junker and Y. Lin [1996]. "Changes to the operational "Early" Eta analysis/forecast system at the National Centers for Environmental Prediction." Weather and Forecasting. 11: 391-413.

Rosenfeld, D., E. Amitai and D. B. Wolff [1995]. "Improved accuracy of radar WPMM estimated rainfall upon application of objective classification criteria." Journal of Applied Meteorology. 34: 212-223.

Rosenfeld, D., D. B. Wolff and E. Amitai [1994]. "The window probability matching method for rainfall measurements with radar." Journal of Applied Meteorology. 33: 683-693.

Salvucci, G. D. and D. Entekhabi [1995]. "Hillslope and climatic controls on hydrologic fluxes." Water Resources Research. 31(7): 1725-1740.

Sarle, W. S. [1994]. "Neural Networks and Statistical Models." Proceedings of the 19th Annual SAS Users Group International Conference, Cary, NC: SAS Institute. 1538-1550.

Schellner K., Dorffner G., and Prem E. [2000]. "Predicting Rainfall Patterns Using Regularized Gaussian Mixture Models", in Trapp R. (Ed.), Cybernetics and Systems, Oesterreichische Studiengesellschaft fuer Kybernetik, Wien, Vols. 1+2, pp.564-569.

Schertzer, D. and S. Lovejoy [1987]. "Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes." Journal of Geophysical Research. 92: 9692-9714.

Schmidt, J., B. Lawrence and B. Olsen. "Operational Precipitation Processing Methodologies." Arkansas-red Basin River Forecast Center. <http://www.srh.noaa.gov/abr/c/p1vol.html>

Scofield, R. A. [1987]. "The NESDIS operational convective precipitation estimation scheme." Monthly Weather Review. 115: 1773-1792.

Seo, D.-J., S. Perica, E. Welles and J. C. Schaake [2000]. "Simulation of precipitation fields from probabilistic quantitative precipitation forecast." Journal of Hydrology. 239: 203-229.

Seo, D. J. and J. A. Smith [1992]. "Radar-based short term rainfall prediction." Journal of Hydrology. 131: 341-367.

Serafin, R. J. and J. W. Wilson [2000]. "Operational Weather Radar in the United States: Progress and Opportunity." Bulletin of the American Meteorological Society. 81(3): 501-518.

Smith, D. L. [1975]. "The application of manually digitised radar data to short-range precipitation forecasting." Preprint Volume of the Sixteenth Radar Meteorology Conference, Houston Texas. 347-352. American Meteorological Society, Boston, MA.

Smith, K. T. and G. L. Austin [2000]. "Nowcasting precipitation - a proposal for a way forward." Journal of Hydrology. 239: 34-45.

Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite [2000]. "Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall." Bulletin of the American Meteorological Society. 81(9): 2035-2046.

Staudenheimer, M, Jr. [1996]. "A description of the Meso-Eta model." Western Region Technical Attachment 96-06. National Weather Service.

Sugimoto, S., E. Nakakita, and S. Ikebuchi [2001]. "A stochastic approach to short-term rainfall prediction using a physically based conceptual rainfall model." Journal of Hydrology. 242: 137-155.

Sun, X., R. G. Mein, T. D. Keenan and J. F. Elliott [2000]. "Flood estimation using radar and raingage data." Journal of Hydrology. 239: 4-18.

Swinney, H. L. [1997]. "Emergence and evolution of patterns", in V. L. Fitch, D. R. Marlow and M. A. E. Dementi (Eds.), Critical Problems in Physics. Princeton University Press. 51-74.

Swinney, H. L. [1994]. "Spatiotemporal patterns: observation and analysis", in A. S. Weigend and N. A. Gershenfeld (Eds.), Time Series Prediction: Forecasting the Future and Understanding the Past. 557-567.

Tatehira, R., H. Sato, and Y. Makino [1976]. "Short-term forecasting of digitised echo pattern." Kisho-cho Kenkyu Jiho. 28: 61-70.

Thielen, J., B. Boudevillain and H. Andrieu [2000]. "A radar data based short-term rainfall prediction model for urban areas - a simulation using meso-scale meteorological modeling." Journal of Hydrology. 239: 97-114.

Toth, E., A. Brath and A. Montanari [2000]. "Comparison of short-term rainfall prediction models for real-time flood forecasting." Journal of Hydrology. 239: 132-147.

Tsonis, A. A., and G. L. Austin [1981]. "An evaluation of extrapolation techniques for short-term prediction of rain amounts." Atmospheres-Ocean. 19: 54-65.

Tsonis, A. A., G. N. Triantafyllou, and J. B. Elsner [1994]. "Searching for determinism in observed data: a review of the issue involved." Nonlinear Processes in Geophysics. 1: 12-25.

US Army Corps of Engineers annual flood damage report to the congress for fiscal year 1999, <http://www.usace.army.mil/inet/functions/cw/cecwe/flood99/>.

Vemuri, V. R. and R. D. Rogers [1994]. Artificial Neural Networks: Forecasting Time Series. IEEE Computer society Press.

Veneziano, D., R. L. Bras and J. D. Niemann [1996]. "Nonlinearity and self-similarity of rainfall in time and a stochastic model." Journal of Geophysical Research. 101(D21): 26371-26392.

Visa, A., J. Iivarinen, K. Valkealahti, and O. Simula [1995] "Neural Network Based Cloud Classifier." Proceedings of ICANN'95.

Vislocky, R. L. and J. M. Fritsch [1995a]. "Generalized Additive Models versus Linear Regression in Generating Probabilistic MOS Forecasts of Aviation Weather Parameters." Weather and Forecasting. 10(4).

Vislocky, R. L. and J. M. Fritsch [1995b]. "Improved Model Output Statistics Forecasts through Model Consensus." Bulletin of the American Meteorological Society. 76(7).

Wan, E. A. [1994]. "Time Series Prediction by Using a Connectionist Network with Internal Delay Lines", in A. S. Weigend and N. A. Gershenfeld (Eds.), Time Series Prediction: Forecasting the Future and Understanding the Past. 195-217.

Warner, T. T., D. N. Yates and G. H. Leavesley [2000]. "A community hydrometeorology laboratory for fostering collaborative research by the atmospheric and hydrologic sciences." Bulletin of the American Meteorological Society. 81(7).

Weigend, A. and N. Gershenfeld [1994]. Time Series Prediction: Forecasting the Future and Understanding the Past. Addison-Wesley.

Wetzel, S. W. and J. E. Martin [2001]. "An Operational Ingredients-Based Methodology for Forecasting Midlatitude Winter Season Precipitation." Weather and Forecasting. 16(1).

Wiggert, V., S. S. Ostlund, G. J. Lockett, and J. V. Stewart [1976]. "Computer software for the growth histories of weather radar echoes." NOAA Technical Memorandum EDRL WMPO-35, 86 pages.

Wilson, J. W. [1966]. "Movement and predictability of radar echoes." U.S. Weather Bureau contract CWB-11093. Travelers Weather Research Center, Hartford, Connecticut.

Xiao R. and V. Chandrasekar [1997]. "Development of a Neural Network Based Algorithm for Rainfall Estimation from Radar Observations." IEEE Transactions on Geoscience and Remote Sensing. 35(1).

Zawadzki, I., E. Torlaschi and R. Sauvegeau [1981]. "The relationship between mesoscale thermodynamic variables and convective precipitation." Journal of the Atmospheric Sciences. 38: 1535-1540.

Zipser, E. [1990]. "Rainfall Predictability: When will extrapolation based algorithms fail?" Preprint Volume of the Eighth Conference on Hydrometeorology, Alberta, Canada. 138-142. American Meteorological Society, Boston, MA.

Zittel, W. D. [1976]. "Computer applications and techniques for storm tracking and warning." Preprint Volume of the Seventeenth Radar Meteorology Conference, Seattle. 514-521. American Meteorological Society, Boston, MA.