

1-1-2011

## Uncertainty visualization of ensemble simulations

Jibonananda Sanyal

Follow this and additional works at: <https://scholarsjunction.msstate.edu/td>

---

### Recommended Citation

Sanyal, Jibonananda, "Uncertainty visualization of ensemble simulations" (2011). *Theses and Dissertations*. 4815.

<https://scholarsjunction.msstate.edu/td/4815>

This Dissertation - Open Access is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact [scholcomm@msstate.libanswers.com](mailto:scholcomm@msstate.libanswers.com).

UNCERTAINTY VISUALIZATION OF  
ENSEMBLE SIMULATIONS

By

Jibonananda Sanyal

A Dissertation  
Submitted to the Faculty of  
Mississippi State University  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy  
in Computer Science  
in the Department of Computer Science and Engineering

Mississippi State, Mississippi

December 2011

Copyright by  
Jibonananda Sanyal  
2011

UNCERTAINTY VISUALIZATION OF  
ENSEMBLE SIMULATIONS

By

Jibonananda Sanyal

Approved:

---

Song Zhang  
Assistant Professor of Computer Science  
and Engineering  
(Major Professor)

---

Philip Amburn  
Research Associate Professor  
(Committee Member)

---

Robert J. Moorhead II  
Professor of Electrical and  
Computer Engineering  
(Committee Member)

---

Edward J. Swan II  
Professor of Computer Science and  
Engineering  
(Committee Member)

---

Edward B. Allen  
Associate Professor of Computer  
Science and Engineering  
(Graduate Coordinator)

---

Sarah A. Rajala  
Dean of the James Worth Bagley College  
of Engineering

Name: Jibonananda Sanyal

Date of Degree: December 9, 2011

Institution: Mississippi State University

Major Field: Computer Science

Major Professor: Dr. Song Zhang

Title of Study: UNCERTAINTY VISUALIZATION OF ENSEMBLE SIMULATIONS

Pages in Study: 152

Candidate for Degree of Doctor of Philosophy

Ensemble simulation is a commonly used technique in operational forecasting of weather and floods. Multi-member ensemble output is usually large, multivariate, and challenging to interpret interactively. Forecast meteorologists and hydrologists are interested in understanding the uncertainties associated with the simulation; specifically variability between the ensemble members. The visualization of ensemble members is currently accomplished through spaghetti plots or hydrographs.

To improve visualization techniques and tools for forecasters, we conducted a user-study to evaluate the effectiveness of existing uncertainty visualization techniques on 1D and 2D synthetic datasets. We designed an uncertainty evaluation framework to enable easier design of such studies for scientific visualization. The techniques evaluated are errorbars, scaled size of glyphs, color-mapping on glyphs, and color-mapping of uncertainty on the data surface. Although we did not find a consistent order among the four techniques for all tasks, we found that the efficiency of techniques used highly depended on

the tasks being performed. Errorbars consistently underperformed throughout the experiment. Scaling the size of glyphs and color-mapping of the surface performed reasonably well.

With results from the user-study, we iteratively developed a tool named ‘Noodles’ to interactively explore the ensemble uncertainty in weather simulations. Uncertainty was quantified using standard deviation, inter-quartile range, width of the 95% confidence interval, and by bootstrapping the data. A coordinated view of ribbon and glyph-based uncertainty visualization, spaghetti plots, and data transect plots was provided to two meteorologists for expert evaluation. They found it useful in assessing uncertainty in the data, especially in finding outliers and avoiding the parametrizations leading to these outliers. Additionally, they could identify spatial regions with high uncertainty thereby determining poorly simulated storm environments and deriving physical interpretation of these model issues.

We also describe uncertainty visualization capabilities developed for a tool named ‘FloodViz’ for visualization and analysis of flood simulation ensembles. Simple member and trend plots and composited inundation maps with uncertainty are described along with different types of glyph based uncertainty representations. We also provide feedback from a hydrologist using various features of the tool from an operational perspective.

Key words: geo-visualization, uncertainty, uncertainty quantification, uncertainty visualization, ensemble simulation, meteorological data, hydrological data, tool design, spaghetti plots, glyph based techniques, visual interaction techniques, exploratory data analysis

## DEDICATION

To pishimuni, Minati Sanyal, who is not around to see this day, to ma, Mira Sanyal, and baba, Jitendra Nath Sanyal, and to my wife Gargi Bhattacharya.

## ACKNOWLEDGMENTS

This dissertation is a product of a lot of hard work, energy, dedication, and CPU cycles. I've spent a lot of time away from my family and parents which I cannot give back in any way. Their support and dedication has been crucial in this endeavour. A big thank you to my baba, Jitendra Nath Sanyal, and ma, Mira Sanyal, for spending so many years away from me while I have been in pursuit of this. There were times when they needed me but they prioritized this over everything else. I salute their sacrifice and dedication and hope that this degree brings them happiness, pride, and a smile on their faces.

Special thanks to my wife Gargi Bhattacharya, who has understood and stood by me for my weird schedules and years spent away from each other. Never has she doubted me, especially when I've been low. She gave everything she could to uplift my spirits and help me through this endeavour. Without her support, this would have never materialized.

Words cannot express my gratitude and love for pishimuni, Minati Sanyal, my aunt, who dreamed of seeing this day. She was the one person whose never ending zeal put me into some of the best schools in India. How I wish she would be around to see me walk the 'walk'. She will always be by my side.

A huge 'thank-you' to kaku, Ranendra Kumar Bhattacharya, who did not stay around long enough to see that transition to baba. In many ways, this was his vision and I hope that wherever he is, I'm making him proud. A special note of thanks to didi, Piyali Chakraborti,



tutuda, Anjan Chakraborti, Anushka Chakraborti, and ma, Bulu Bhattacharya, who provided unquestioned and unparalleled support on all fronts. Anushka especially, who wasn't very sure what a PhD is but prayed all along so that I'd get through!

My advisor, Dr. Song Zhang is a man with great vision. He knows a good opportunity when he sees one. That I have been able to complete this work is a result of his vision, guidance, and steering. His keen sense of what might work and what could lead to a potential wormhole has saved me on numerous occasions and helped me in not just getting papers published but in learning what research is about. There wasn't a single instance when he was too busy to see me. I thank him for his advice and exemplary guidance.

I'd also like to thank Dr. Robert J. Moorhead for his guidance and support. He has funded and transitioned me through multiple projects over the last five years. I must admit I am a fan of his reviews. He can easily spot glaring flaws in a document that everyone else misses. I have learned a great deal about professionalism, scientific scope, and the importance of identifying impact from him.

Dr. Philip Amburn is a person who has left a profound impression on me. He is a great project manager and guide, and an absolutely awesome person. He is particularly good at identifying and answering the "so-what?" question [87]. I deeply cherish our impromptu and absolutely random meetings. I have learned a lot about project and time management from him. In fact, he is one of the people who have been instrumental in getting me to focus and to 'get-out'! Never has he allowed me to deter from my primary goal of getting the degree and I thank him for that.

I have been hugely influenced by Dr. J. Edward Swan too. Although we haven't worked together on a project, he has been immensely supportive of my efforts and his advice and guidance has been immensely valuable. Right from the time when he gave me feedback on my answers to questions in the qualifying examination, to advice on how to schedule my proposal and things to watch out for in my defence. I always made it a point to meet him every few months to update him of my state of things. Since we both love food, it would typically be over lunch. These meetings have always been very helpful. There was a time when I when I went to see him while I was struggling to justify the 'science' in my work. I fondly remember our conversation over the next hour, about research, and about quantifying a person's individual contribution to science. I grateful for his support on that day, and every other day.

Much of my dissertation would not hold together without the expertise of Dr. Jamie Dyer. He has been a domain expert and more for this dissertation. I have used his help to figure out the science, systems, and workings of weather simulation. He taught me how to run the WRF model and produce data, so much data that at one point I nearly brought down the /gri file-system. I cannot thank him enough for teaching me so much about meteorology and hydrology.

Dr. Andrew Mercer's expertise has been critical in this effort. As a domain expert, he helped me immensely in the quantification of uncertainty and in setting up the newest version of WRF. He also helped me in figuring out how to post-process the simulation output. A number of sections in this dissertation is a result of his direct influence and contribution and I thank him for all the help.

I ended up joining Mississippi State University because of Dr. Valentine Anantharaj. I was applying to various universities in the US for a higher degree when he found my profile interesting and contacted me. He invested a lot of time and effort to finally have me here. He was also instrumental in transitioning me to the visualization group which was more aligned with my interests. His contribution to my career is very significant and I'd like to thank him for everything.

I would like to thank Dr. Edward B. Allen, the Graduate Coordinator of the Computer Science and Engineering Department, for his guidance in helping me navigate graduate school.

Special thanks to Keqin Wu for giving the name 'Noodles' to my software prototype.

There are many others who have been of great help. Dr. Patrick J. Fitzpatrick is one such person. It was his evaluation of some of the visualizations in the VERTEX that got me interested uncertainty of weather simulations. I'd also like to thank Dr. T. J. Jankun-Kelly for his feedback on various occasions, Derek Irby, and John van der Zwaag. These guys are awesome!

Special thanks are in order for all my friends with whom I've spend time voicing my frustrations and then forgetting about them, mostly over food and/or beer. In particular, I'd like to thank Andrew Stamps, Adam Jones, Brandon Malone, Teena Garrison, Christianna Taylor, Satish Ganji, Shane Sanders, Shaheen Ahmed, Vijay P. Shah, Meghanand Bharaskar, and Buddy Jones for the wonderful moments of camaraderie. As I graduate, every minute spent with them makes this endeavour worthwhile.

In addition, I'd like to acknowledge the entire open-source community for the tools and libraries used in the dissertation. I'd also like to thank GRI, NGI, and HPC<sup>2</sup> for employing and housing me over the years.

Lastly, I'd like to thank NASA and NOAA for funding the projects that has enabled this research.

Once again, a big THANK YOU to everybody. I've learned a lot from you in these years!

## TABLE OF CONTENTS

DEDICATION . . . . .	ii
ACKNOWLEDGMENTS . . . . .	iii
LIST OF TABLES . . . . .	xii
LIST OF FIGURES . . . . .	xiii
LIST OF SYMBOLS, ABBREVIATIONS, AND NOMENCLATURE . . . . .	xviii
CHAPTER	
1. INTRODUCTION . . . . .	1
1.1 Research Statement . . . . .	3
1.2 Motivation . . . . .	4
1.3 Contributions . . . . .	6
1.4 Organization . . . . .	7
2. BACKGROUND . . . . .	8
2.1 Understanding Uncertainty . . . . .	8
2.1.1 Uncertainty in Data . . . . .	8
2.1.1.1 Uncertainty Visualization Pipeline . . . . .	9
2.1.1.2 Taxonomies and Classification of Proposed Techniques	10
2.1.1.3 Uncertainty Visualization Techniques . . . . .	12
2.1.1.4 Visualization of Geo-Spatial Uncertainty . . . . .	15
2.1.1.5 Perceptual Attributes in Uncertainty Visualization . .	17
2.1.2 Evaluation of Uncertainty Visualization Techniques . . . . .	18
2.2 Application Areas . . . . .	19
2.2.1 Uncertainty in Weather Simulations . . . . .	20
2.2.1.1 Coordinate Systems . . . . .	20
2.2.1.2 Data Formats . . . . .	24
2.2.1.3 User Training . . . . .	25
2.2.1.4 Ensemble Uncertainty Quantification . . . . .	26

2.2.1.5	Technological Gap . . . . .	26
2.2.2	Uncertainty in River-Flow and Inundation Simulations . . . . .	27
2.2.2.1	Accuracy of Topographic Data . . . . .	28
2.2.2.2	Structure Elevations . . . . .	29
2.2.2.3	Inland Flooding . . . . .	29
2.2.2.4	Coastal Flooding . . . . .	30
3.	APPROACH AND TECHNIQUES . . . . .	31
3.1	Uncertainty Evaluation Framework . . . . .	31
3.2	Uncertainty Visualization Techniques . . . . .	33
3.2.1	Glyphs Altered by Size . . . . .	33
3.2.2	Glyphs Altered by Color . . . . .	34
3.2.3	Color-Mapped Lines and Surface . . . . .	34
3.2.4	Gradient and Striped Gradient . . . . .	34
3.2.5	Errorbars . . . . .	37
3.2.6	Boxplots . . . . .	37
3.2.7	Graduated Uncertainty Glyphs . . . . .	40
3.2.8	Uncertainty Ribbon . . . . .	42
3.2.9	Graduated Uncertainty Ribbon . . . . .	43
3.2.10	Animation . . . . .	45
3.2.11	Multivariate Uncertainty Glyphs . . . . .	45
3.2.12	Spaghetti Plots . . . . .	48
4.	QUANTITATIVE EVALUATION OF UNCERTAINTY VISUALIZATION TECHNIQUES . . . . .	49
4.1	Study Design . . . . .	51
4.1.1	Data Generation . . . . .	51
4.1.2	Uncertainty Visualization Techniques Chosen for Evaluation . . . . .	55
4.1.3	Participant Pool . . . . .	58
4.1.4	User-study tasks . . . . .	59
4.1.4.1	Search tasks . . . . .	59
4.1.4.2	Counting tasks . . . . .	61
4.1.5	Interface Design . . . . .	61
4.1.6	Participant Training . . . . .	64
4.1.7	Identifying free parameters . . . . .	65
4.1.8	The Main Study . . . . .	66
4.2	Analysis . . . . .	66
4.3	Results and Discussion . . . . .	68
4.4	Conclusion . . . . .	74
5.	SOFTWARE PROTOTYPE FOR WEATHER ENSEMBLES . . . . .	76

5.1	Ensemble Weather Forecasting . . . . .	76
5.2	Software Engineering Effort . . . . .	78
5.3	Initial Prototype: Noodles . . . . .	79
5.3.1	Uncertainty Metrics . . . . .	81
5.3.1.1	Bootstrapping . . . . .	83
5.3.2	Uncertainty Visualization . . . . .	84
5.3.2.1	Uncertainty Glyphs and Graduated Uncertainty Glyphs	84
5.3.2.2	Uncertainty Ribbon and Graduated Uncertainty Ribbon	87
5.3.2.3	Spaghetti Plots . . . . .	89
5.3.3	Additional Features . . . . .	90
5.3.3.1	Data-Transect Plot . . . . .	90
5.3.3.2	Colormap . . . . .	91
5.3.3.3	Interactive Visual Queries . . . . .	92
5.3.4	Expert Evaluation of the tool . . . . .	93
5.3.4.1	Effectiveness of Uncertainty Visualization . . . . .	93
5.3.4.2	Effectiveness of Interactivity . . . . .	94
5.3.4.3	User feedback . . . . .	95
5.3.5	Case Study: 1993 “Superstorm” . . . . .	96
5.3.5.1	Formulation of the ensemble and running WRF . . . . .	96
5.3.5.2	Expert Evaluation: Model Parametrization Inferences	100
5.4	Iterated Prototype: Noodles 2 . . . . .	105
5.4.1	Software Architecture . . . . .	105
5.4.2	User Interface . . . . .	107
5.4.2.1	Visualization Capabilities . . . . .	109
5.4.2.2	Interaction Capabilities . . . . .	109
5.4.3	Case Studies . . . . .	110
5.4.3.1	Ensemble Formation and Running WRF . . . . .	110
5.4.3.2	Case Study 1: Hurricane Fran’s extra-tropical transition	111
5.4.3.3	Case Study 2: Mississippi-Alabama Tornado Outbreak	114
5.5	Discussion . . . . .	116
6.	<b>SOFTWARE PROTOTYPE FOR RIVER-FLOW AND FLOOD-FORECAST ENSEMBLES . . . . .</b>	<b>117</b>
6.1	Operational Flood Forecasting . . . . .	117
6.1.1	Ensemble Data . . . . .	118
6.1.1.1	Uncertainty Quantification . . . . .	119
6.1.2	Geometry . . . . .	119
6.1.3	Visualization Capabilities in HEC-RAS . . . . .	120
6.2	Software Engineering Effort . . . . .	121
6.3	Visualization Views . . . . .	123
6.3.1	Cross-Section View . . . . .	124

6.3.2	River Profile View . . . . .	124
6.3.3	Plan View . . . . .	124
6.3.4	3D View . . . . .	125
6.4	Uncertainty Visualization Techniques . . . . .	125
6.4.1	Line plots . . . . .	127
6.4.2	Box plots . . . . .	128
6.4.3	Glyph plots . . . . .	129
6.4.4	Star uncertainty glyphs . . . . .	129
6.4.5	Temporal multivariate uncertainty glyphs . . . . .	131
6.4.6	Inundation map with uncertainty . . . . .	132
6.5	Evaluation . . . . .	134
6.6	Discussion and Future Work . . . . .	136
7.	CONCLUSION . . . . .	137
8.	FUTURE WORK . . . . .	140
8.1	Recommendations for Future Research . . . . .	141
8.1.1	Understanding the forecasters' needs . . . . .	141
8.1.2	Visualization Techniques . . . . .	142
8.1.3	Evaluation . . . . .	142
8.1.4	Frameworks . . . . .	143
8.1.5	Engineering prototype software . . . . .	143
	REFERENCES . . . . .	144



## LIST OF TABLES

2.1	Classification of uncertainty visualization techniques by Pang et al. [74]. . . .	11
2.2	Potential heristics and their relevancy to the 8 tested visualizations [109]. . . .	19
4.1	ANOVA results for pairwise comparison on techniques. . . . .	71
4.2	ANOVA results for pairwise comparison on tasks. . . . .	75

## LIST OF FIGURES

1.1	Example of a spaghetti plot used in operational meteorological analysis. . . .	5
1.2	Example of a hydrograph used in operational river forecast analysis. . . . .	5
2.1	Uncertainty visualization pipeline of Pang et al. [74]. . . . .	10
2.2	High-level taxonomy of imperfections in knowledge by Gershon [29]. . . . .	12
2.3	Underwater uncertainty in a CAVE environment [83]. . . . .	13
2.4	Uncertainty in architectural renderings [93]. . . . .	13
2.5	Positional uncertainty of astrophysical data[49]. . . . .	14
2.6	Probabilistic Direct Volume Rendering of medical data [54]. . . . .	14
2.7	Uncertainty in data plots [76]. . . . .	15
2.8	Uncertainty in vector data [108]. . . . .	16
2.9	Ensemble-Vis plugin for CDAT [77]. . . . .	17
2.10	The sigma coordinate system. . . . .	21
2.11	The eta coordinate system. . . . .	22
2.12	The isentropic coordinate system. . . . .	22
2.13	The hybrid coordinate system. . . . .	23
2.14	Example of variables on the Arakawa C grid. . . . .	24
2.15	Flood plain extent with a 30m DEM (left) and 3m Lidar DEM (right) [70]. . .	29
2.16	Differences in heights (NAVD 88 minus NGVD 29) in centimetres [70]. . . .	30

3.1	The Uncertainty Evaluation Framework. . . . .	32
3.2	Uncertainty represented by glyphs altered by size for 1D data. . . . .	33
3.3	Uncertainty represented by glyphs altered by size for 2D data. . . . .	34
3.4	Uncertainty represented by glyphs altered by color for 1D data. . . . .	35
3.5	Uncertainty represented by glyphs altered by size for 2D data. . . . .	35
3.6	Uncertainty represented by altering the line color for 1D data. . . . .	36
3.7	Uncertainty represented by altering the surface color for 2D data. . . . .	36
3.8	Uncertainty represented by the color gradient for 1D data. . . . .	37
3.9	Uncertainty represented by the color gradient on vertical glyphs for 2D data. . . . .	38
3.10	Use of striped gradient for 1D data. . . . .	38
3.11	Uncertainty represented by errorbars for 1D data. . . . .	39
3.12	Uncertainty represented by errorbars for 2D data. . . . .	39
3.13	Construction of graduated uncertainty glyphs. . . . .	40
3.14	Different types of graduated uncertainty glyphs. . . . .	41
3.15	Construction of uncertainty ribbon. . . . .	43
3.16	Construction of graduated uncertainty ribbon. . . . .	44
3.17	Uncertainty represented by animation of glyphs for 1D data. . . . .	45
3.18	Uncertainty represented by animation of glyphs for 2D data. . . . .	46
3.19	Uncertainty represented by animation of surface for 1D data. . . . .	46
3.20	Uncertainty represented by animation of surface for 2D data. . . . .	47
3.21	A multivariate uncertainty glyph. . . . .	47
3.22	Temporal multivariate glyphs illustrating multivariate uncertainty. . . . .	48

4.1	Systemic and random components in the synthetic dataset. . . . .	53
4.2	Generation of data for the user-study. . . . .	53
4.3	1D uncertainty visualizations in the user-study. . . . .	56
4.4	2D uncertainty visualizations in the user-study. . . . .	57
4.5	User interface for questions requiring a count of features. . . . .	62
4.6	User interface for questions requiring a search task. . . . .	63
4.7	Summary of results from the user study. . . . .	67
4.8	User study results for the 1D case. . . . .	69
4.9	User study results for the 2D case. . . . .	69
5.1	Marked up user interface of Noodles. . . . .	80
5.2	Histogram of ensemble means for 8 values. . . . .	82
5.3	Histogram of the bootstrap bootstrap means from Figure 5.2. . . . .	82
5.4	Glyphs illustrating 95% CI of the values. . . . .	85
5.5	Graduated glyphs aligned over the whole grid. . . . .	86
5.6	Graduated glyphs along the mean perturbation pressure. . . . .	86
5.7	Three time-steps illustrating spaghetti plots, graduated glyphs, and colormap. . . . .	87
5.8	Illustration of uncertainty ribbon mapping the IQR. . . . .	88
5.9	Illustration of a graduated uncertainty ribbon. . . . .	88
5.10	The 3×6 color scheme from ColorBrewer [33]. . . . .	89
5.11	The data-transect plot and interactive legend. . . . .	91
5.12	Satellite image showing the extent of the 1993 “Superstorm”. . . . .	97
5.13	Bias in the resolution of the cold-front for perturbation potential temperature. . . . .	101

5.14	Sharp change in data-transect through Figure 5.13. . . . .	102
5.15	Bias in the resolution of the storm center for the pressure perturbation. . . . .	102
5.16	A data transect through Figure 5.15. . . . .	103
5.17	Bias in water-vapor mixing ratio corresponding to Figure 5.13 and Figure 5.15. . . . .	103
5.18	Data transect through Figure 5.15. . . . .	104
5.19	Spaghetti plot of perturbation potential temperature field at upper eta-levels. . . . .	104
5.20	General agreement among schemes corresponding to Figure 5.19. . . . .	105
5.21	Software architecture of Noodles 2. . . . .	106
5.22	Marked up user interface of Noodles 2. . . . .	108
5.23	Satellite image of Hurricane Fran on 5 September 1996. . . . .	112
5.24	Geopotential height and cloud-water mixing ratio for Hurricane Fran. . . . .	113
5.25	Satellite image of the Mississippi-Alabama tornados on 26 April 2011. . . . .	114
5.26	Temperature, water-vapor ratio, and pressure for the tornado outbreak. . . . .	115
6.1	A screen-shot of visualization capability in HEC-RAS. . . . .	120
6.2	High level block diagram of FloodViz. . . . .	121
6.3	A screen-shot of the FloodViz user interface. . . . .	122
6.4	Interactive legend with categorical colors. . . . .	123
6.5	Cross-section view of the river. . . . .	125
6.6	Profile view of the river. . . . .	126
6.7	Plan view illustrating uncertainty in extent of flooding. . . . .	126
6.8	3D View of the extent of flooding. . . . .	127
6.9	Close-up of a box plot. . . . .	128

6.10	Box plot in the profile view illustrating water surface elevation. . . . .	129
6.11	Close-up of the glyphs scaled by size. . . . .	130
6.12	Glyphs illustrating the water surface elevation. . . . .	130
6.13	A star (multivariate) uncertainty glyph. . . . .	131
6.14	Star glyphs illustrating the water surface elevation. . . . .	132
6.15	Temporal glyphs illustrating three variables over time. . . . .	133
6.16	Temporal glyphs in the profile view. . . . .	133

## LIST OF SYMBOLS, ABBREVIATIONS, AND NOMENCLATURE

**NOAA** National Oceanographic and Atmospheric Administration

**NASA** National Aeronautics and Space Administration

**FEMA** Federal Emergency Management Agency

**USGS** United States Geological Survey

**NWS** National Weather Service

**RFC** River Forecast Center

**LMRFC** Lower Mississippi River Forecast Center

**NED** National Elevation Dataset

**NGVD29** National Geodetic Vertical Datum of 1929

**NAVD88** North American Vertical Datum of 1988

**BIPM** Bureau International des Poids et Mesures

**CIPM** Comité International des Poids et Mesures

**NIST** National Institute of Standards and Technology

**SD** Standard Deviation

**IQR** Inter Quartile Range

**CI** Confidence Interval

**ANOVA** Analysis of Variance

**RSS** Root of the sum of squares

**NARR** North American Regional Reanalysis

**ECMWF** European Centre for Medium-Range Weather Forecasts

**RUC** Rapid Update Cycle

**WRF** Weather Research and Forecasting model

**EMS** Environmental Modelling System  
**STRC** Science and Training Resource Center  
**ARW** Advanced Research WRF  
**NMM** Nonhydrostatic Mesoscale Model  
**QPF** Quantitative Precipitation Forecast  
**NCOM** Navy Coastal Ocean Model  
**HEC-RAS** Hydraulic Engineering Center – River Analysis System  
**GRIB** Gridded Binary  
**NETCDF** Network Common Data Form  
**HDF** Hierarchical Data Format  
**DEM** Digital Elevation Model  
**LIDAR** Light Detection And Ranging  
**IDV** Integrated Data Viewer  
**VTK** Visualization Toolkit  
**CDAT** Climate Data Analysis Tools



## CHAPTER 1

### INTRODUCTION

Uncertainty is an inseparable component of almost everything we do. We go about our day-to-day lives with the assumption and expectation that most things will occur per the norm. For example, we depend on the proper functioning of the alarm clock, toaster, car's engine, subway trains, electronic office keys, telephone, and a myriad of other devices for our daily activities. All these devices operate at a certain level of certainty and the odds of something going wrong may be quite low but always existent. We also deal with uncertainty in our dealings with people, in the choices we make, in the outcome of elections, in the resulting policies, in the forces that govern the economy, and in many other manifestations.

Commonly, we strive to understand and minimize uncertainty or use it to our benefit, be it engineering better infrastructure or arriving at policy decisions. Depending on the importance of the phenomenon and magnitude of associated uncertainty, we often try to make efforts to understand the phenomena and identify the sources of uncertainty. Such efforts take time to perfect and we often find ourselves trying to estimate the phenomena, avoid it, or find ways to insure against it. Again, sometimes none of this is possible or feasible.

Weather and flood hazards are two examples of such phenomena. While the general reliability of weather and flood forecasts has improved significantly over the last decade and sufficient warning is often available ahead of time, we are still significantly affected by severe weather events. With the goal of saving life and property, significant resources are invested in the simulation of such events to support the nation's human and economic well being.

These geophysical simulations tend to be large and often prohibitively expensive. The rapid progress in supercomputing has greatly benefited simulation science. Researchers are able to design simulations with ever more sophistication and resolution which produce data-sets that are often magnitudes larger in size than used previously. Scientists have about as much time as they had earlier to analyze and make sense of these datasets. Using conventional approaches and summary descriptors of data are becoming inadequate as large parts of the data tend to get ignored in the analysis.

The modeling itself is far from perfect. Uncertainty abounds in these simulations and its sources are many. Modeling imperfections, accuracy of the input data, resolution of the model, numerical errors, and computational constraints are some of the potential sources of uncertainty. Scientists often use multiple model simulations using different sets of parameters or initial conditions, called 'Ensemble' simulations, to capture some of the modeling uncertainties and to account for various scenarios. These datasets are even larger. To improve forecasts, it is important to understand the associated uncertainties in order to improve the simulation design leading to better approximation of the geophysical processes.

Visualization and visual analytics are emerging as important techniques in enabling scientists in seeing and interacting with huge amounts of data thereby facilitating more in-depth analysis. With the correct set of techniques, valuable information can be derived from an otherwise complex dataset. Visualization is almost becoming indispensable in many disciplines because of the ever increasing size and complexity of datasets. In particular, scientists are beginning to feel a need to be able to quantify and visualize uncertainty in their data.

There is a general lack of techniques and tools for the visualization of uncertainty, particularly in operational setups. Operational personnel at National Oceanographic and Atmospheric Administration's (NOAA) weather forecast offices routinely use ensemble simulations but have limited analysis capabilities. This dissertation is aimed to develop innovative, practical, and usable uncertainty visualization techniques and tools with a focus on improving operational usability. We hope that this effort will bridge some of the gap that exists between advances in simulation and visualization science, and the availability of tools and techniques to analyze the data.

## **1.1 Research Statement**

The objective of this dissertation is the design, development, and evaluation of techniques and tools for uncertainty visualization that are new and improved over the currently used techniques for visualization of uncertainty in ensemble simulations of weather and river flow for operational activities.

## 1.2 Motivation

The motivation for this work stems from initial collaborative research with meteorologists and hydrologists. Both meteorological and hydrological simulations suffer from being imprecise and subject to input and modeling variability. Ensembles are commonly used to account for various possible scenarios in both. An ensemble can be defined as a collection of multiple simulation runs over the same domain or extent and for the same period of time, but with slightly perturbed initial conditions or different parametrizations, and sometime both, with the objective of being able to capture individual model differences and variations. In this dissertation, parameter ensembles of weather simulations and multi-input ensemble of hydrological simulations are discussed as representative examples.

The typical technique used for the analysis of weather ensembles is a spaghetti plot as illustrated in Figure 1.1. The technique available to hydrologists in visualizing ensemble output is by the use of spaghetti plot-like hydrographs as illustrated in Figure 1.2. A spaghetti plot is a collection of iso-valued contours from the different ensemble members over a single spatial domain at fixed pressure levels. A hydrograph is like a spaghetti-plot but constructed by tracing water-level heights at a location over a period of time.

While there are a number of tools to visualize weather simulation data, to the best of our knowledge, there are no tools available to visualize ensemble output. Meteorologists and hydrologists are interested in identifying spatial and temporal regions where the individual model runs agree or disagree. They are also interested in identifying which model runs are poor so as to eliminate them in the final analysis. The use of spaghetti plots alone

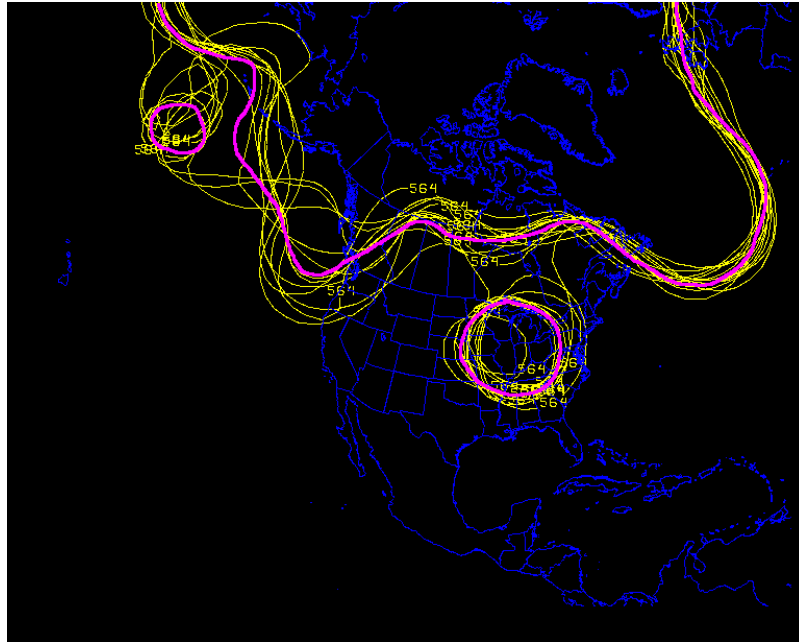


Figure 1.1

Example of a spaghetti plot used in operational meteorological analysis.

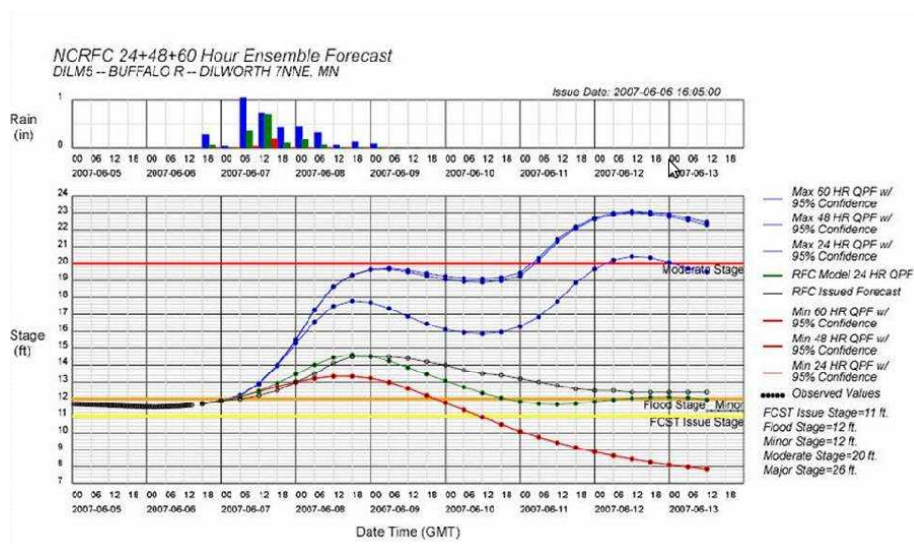


Figure 1.2

Example of a hydrograph used in operational river forecast analysis.

makes it challenging to perform these tasks. Often overlooked, spaghetti plots should not be interpreted without the underlying gradient, nor should they be used to compare spatially separate locations [107].

Meteorological data is inherently 3D and temporal in nature and the unavailability of visualization techniques forces meteorologists to throw away most of this data. Hydrologists are also constrained by the inability to visualize variations of individual ensemble members over a spatial extent or across a cross-section of the river. There is a wide body of research in the field of uncertainty visualization but it has not found its way into operational settings. The tools and techniques available to operational meteorologists and hydrologists for ensemble uncertainty visualization are very limited and this dissertation attempts to address and alleviate some of these limitations.

### 1.3 Contributions

The main contributions of this dissertation are:

- Design of an Uncertainty Evaluation Framework that helps to classify, categorize, and design uncertainty visualizations for scientific data. Note that other existing frameworks can be encompassed in this framework.
- Evaluation of existing uncertainty visualization techniques to find guidelines for the effective visualization of uncertainty in 1D and 2D datasets.
- Design of two new uncertainty visualization techniques, graduated glyphs and graduated ribbons.
- Design of a software prototype that allows interactive multi-variable and multi-technique uncertainty visualization for large ensembles of meteorological simulations.
- Simulation of multiple severe weather events and use of the tool to derive insight into parametrizations used for the modeling.

- Design of uncertainty visualization capabilities for a flood forecasting and analysis tool.

#### **1.4 Organization**

The rest of the document is organized as follows. The next chapter discusses the body of existing literature in the field of uncertainty visualization. It also reflects our initial effort in understanding the state of research and identifying compelling questions that merit investigation. Chapter 3 discusses the Uncertainty Evaluation Framework and uncertainty visualizations techniques developed and used in this dissertation. Chapter 4 presents the results of a user-study designed to understand the effectiveness of commonly used uncertainty visualization techniques. Based on some of the results of the user-study, a prototype of a tool for the visualization of uncertainty in ensemble simulations called ‘Noodles’ is presented in Chapter 5. The chapter also discusses an iterated and improved version of the software prototype called ‘Noodles 2’, along with the studied datasets. Chapter 6 discusses application of some of the uncertainty visualization techniques for enabling ensemble analysis capabilities in a river-flow and flood mapping tool named ‘FloodViz’. Chapter 7 presents concluding remarks and Chapter 8 highlights avenues and ideas for future research.

## CHAPTER 2

### BACKGROUND

This chapter provides an overview of the uncertainty visualization literature followed by a discussion of issues that affect meteorological and hydrological ensembles.

#### 2.1 Understanding Uncertainty

A lot of research has been conducted in understanding, quantifying, and developing techniques to visualize uncertainty. The following is a definition of uncertainty given by Djurcilov et al. [24]: “Uncertainty is a multi-faceted characterization about data, whether from measurements and observations of some phenomenon, and predictions made from them. It may include several concepts including error, accuracy, precision, validity, quality, variability, noise, completeness, confidence, and reliability.”

##### 2.1.1 Uncertainty in Data

For a long time, the term uncertainty was used in a rather loose sense. Researchers recognized the need to clearly define uncertainty as it was a necessary step before trying to solve the visualization problem [15, 38, 74]. The International Bureau of Weights and Measures (BIPM) - International Committee for Weights and Measures (CIPM) recognized the broad scenarios from which uncertainty may arise and suggested two measures



to estimate uncertainty [7, 8]. The first measure, called Type A evaluation of standard uncertainty, is derived by mathematical and statistical approaches such as standard deviations and analysis of variance. The second measure, called Type B evaluation of standard uncertainty, is more subjective and is based on scientific judgement such as experience and specifications. They also defined uncertainty to have ‘random’ and ‘systematic’ components which are conditioned by a mathematical model.

The National Institute for Science and Technology (NIST) [94] suggested a ‘combined standard uncertainty’. A law of the propagation of uncertainty was provided based on a root-of-the-sum-of-the-squares (RSS) method. ‘Expanded uncertainty’ was a term coined to express uncertainty defined by an interval that bounded the measurement. Their definition included correction factors arising from recognized system effects.

Based on the BIPM guide, Mauris et al. [60] proposed a fuzzy approach to measure uncertainty. While these are generic approaches, most scientific disciplines have specialized metrics and methods to quantify uncertainty specific to the domain.

### **2.1.1.1 Uncertainty Visualization Pipeline**

Typically, data undergoes various changes before visualization: it is measured or generated, stored, altered, refined, and eventually visualized. Uncertainty can be introduced to the data at various stages in this process. Pang et al. [74] called this the Uncertainty Visualization Pipeline and divided it into three stages: data acquisition, data refining, and visualization, and showed how uncertainty can be introduced in any of these stages (Figure 2.1).

Uncertainty in acquisition can be introduced because of instrument sensitivities, modeling conditions, and human error. Raw data is rarely used in analysis or visualization. Often various types of transformations such as resampling, rescaling, or quantization are applied to the data which introduce uncertainties. The choice of the visualization technique can also introduce uncertainties in the representation. This can creep in silently. And of course, human perception can play a significant role in the visual representation chosen.

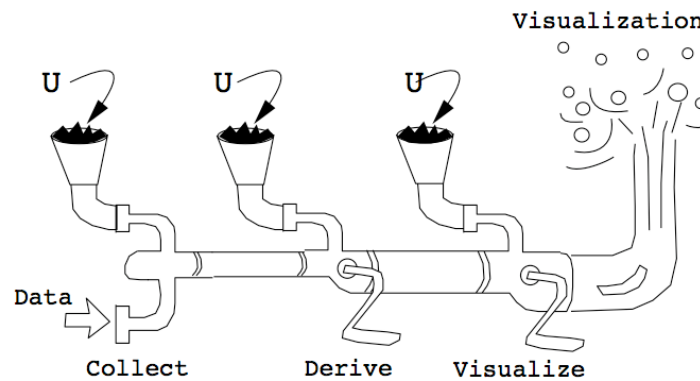


Figure 2.1

Uncertainty visualization pipeline of Pang et al. [74].

### 2.1.1.2 Taxonomies and Classification of Proposed Techniques

A number of techniques have been proposed to study and classify uncertainty visualization techniques. Johnson and Sanderson [41] provided an overview of the current research and identified important goals for further research. MacEachren [55] touched

upon a number of aspects of geospatial uncertainty and discussed conceptual models of spatial uncertainty in a cartographic context.

Pang et al. [74] broadly classified the techniques into adding glyphs, adding geometry, modifying geometry, modifying attributes, animation, sonification, and psycho-visual approaches (Table 2.1). Recently, Thomson et al. [96] presented a typology to visualize uncertain information pertinent to geospatially referenced data. Their typology was developed keeping the tasks of an information analyst in mind. Gershon [29] presented a short discussion on imperfections in information and a taxonomy of the causes of imperfection in knowledge stressing the need to develop better representations (Figure 2.2). Some researchers have also discussed uncertainty cataloguing techniques [19, 53]

Table 2.1

Classification of uncertainty visualization techniques by Pang et al. [74].

Approach	Application			
	Radiosity	Animation	Interpolation	Flow
Add glyphs	spherical	ladders	uncertainty	ellipsoidal
Add geometry		show angles	fat surfaces, bumps	ribbons
Modify geometry	affine transform		IFS, displacement	
Modify Attributes	reflectivity, textures	bump-mapping	pseudo-color	
Animation	magnitude, frequency	oscillate	oscillate	batons-ranking
Sonification			pitch, instrument	duration
Psycho-Visual	left/right		subliminal	

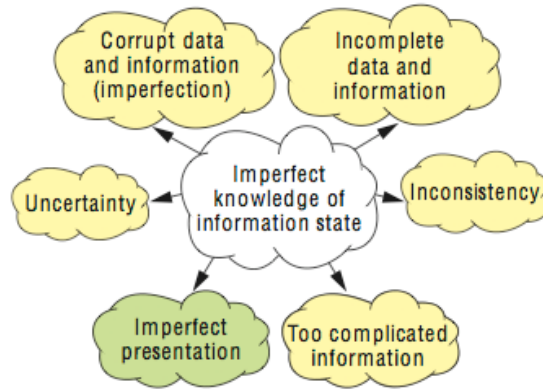


Figure 2.2

High-level taxonomy of imperfections in knowledge by Gershon [29].

### 2.1.1.3 Uncertainty Visualization Techniques

A number of disparate fields of research have successfully researched and applied uncertainty visualization techniques. Schmidt et al. [83] looked at ways of representing the multivariate nature of bathymetric uncertainty (Figure 2.3). Rheingans and Joshi [79] visualized the positional uncertainty of molecules. Strothotte et al. [93] used non-photorealistic techniques to present uncertainty of architectural reconstructions (Figure 2.4). Li et al. [49] visualized uncertainty in astrophysical data (Figure 2.5) and Lundstrom et al. [54] presented a probabilistic animation method to illustrate uncertainty in medical volume renderings (Figure 2.6). Lodha et al. [51] used sound for the depiction of uncertainty and Cedilnik and Rheingans [18] demonstrated procedural annotation techniques.

Olston and Mackinlay [72] argued that visualization methods should be different for statistical uncertainty and bounded uncertainty since statistical uncertainty representations potentially incorporate infinite ranges of values. They proposed ambiguation as the solu-

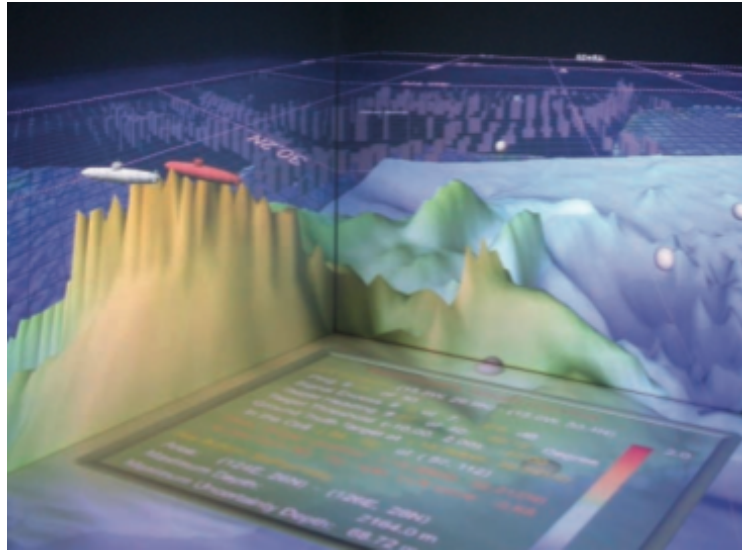


Figure 2.3

Underwater uncertainty in a CAVE environment [83].

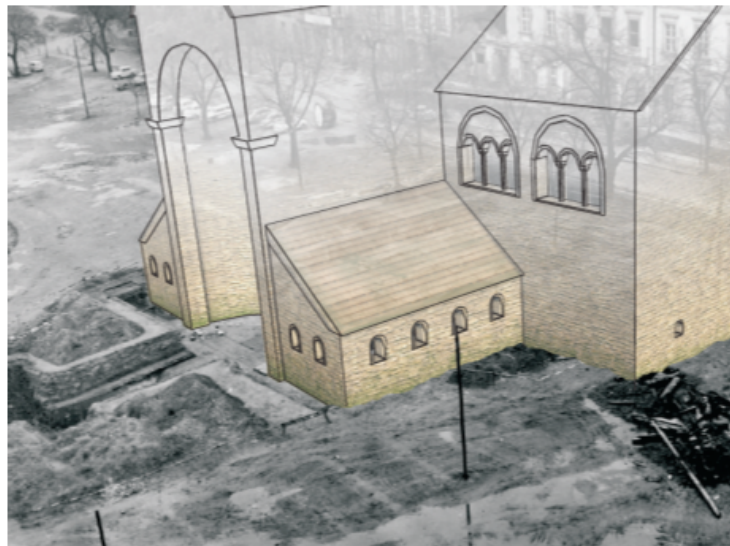


Figure 2.4

Uncertainty in architectural renderings [93].

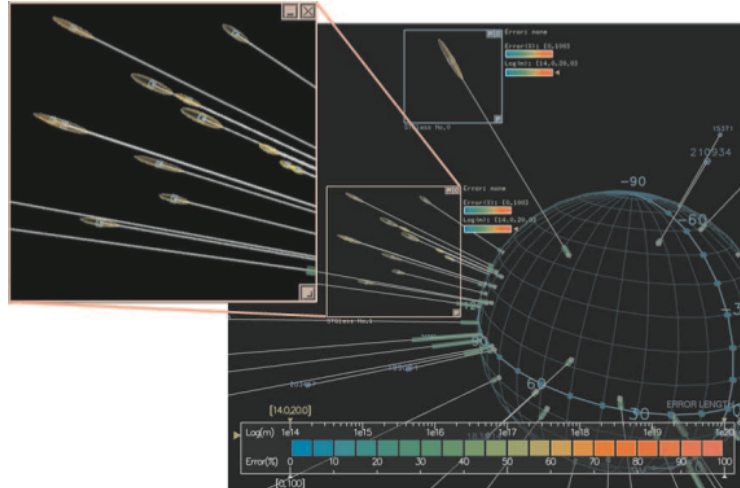


Figure 2.5

Positional uncertainty of astrophysical data[49].

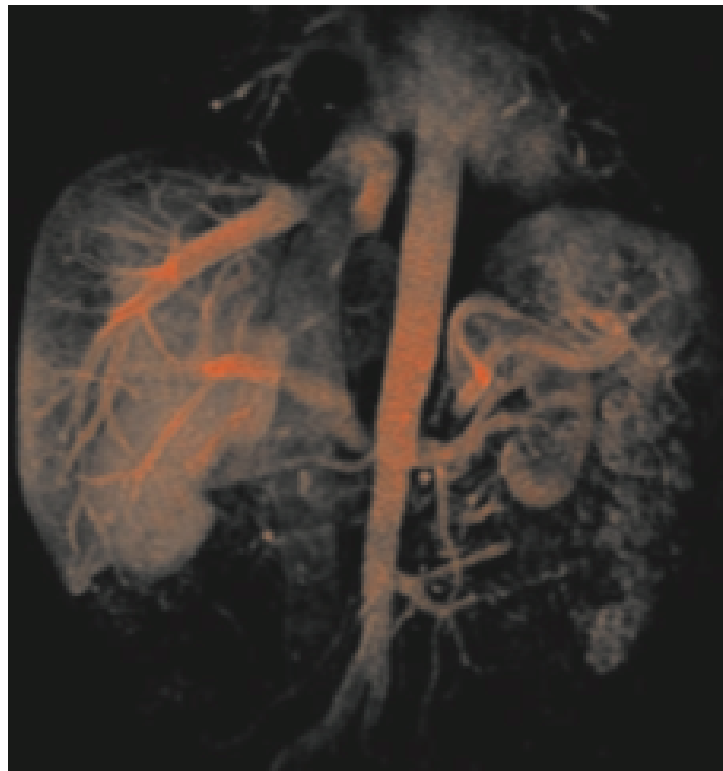


Figure 2.6

Probabilistic Direct Volume Rendering of medical data [54].

tion in which statistical graphics are modified or augmented with visual cues that enhance the notion of unboundedness. Huang [37] showed how a multivariate scatterplot can be created by overloading the visual channels such as color, size, and background color to show the quality of information. Potter et al. [76] illustrated how uncertainty quantified by statistical estimates can be presented in data plots (Figure 2.7).

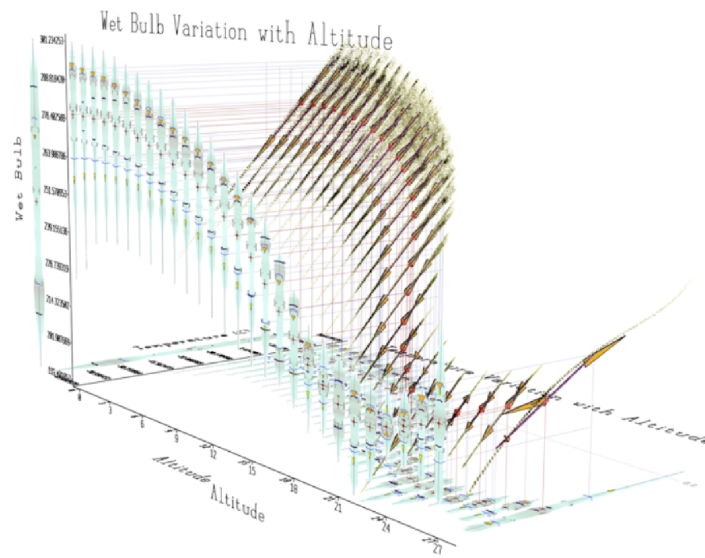


Figure 2.7

Uncertainty in data plots [76].

#### 2.1.1.4 Visualization of Geo-Spatial Uncertainty

Geospatial uncertainty and its representation have been well studied by researchers. MacEachren [55] identified challenges in geospatial uncertainty visualization, underlining the difference between data quality and uncertainty. He addressed the representational

issues and applicability of Bertin's graphic variables [6], proposed conceptual models of spatial uncertainty, and described how they relate to cartography. He also studied visual metaphors and how they affect the interpretation of maps [56] which provides valuable insight into map symbology for cartographic visualization. MacEachren [55] also suggested the use of hue, saturation, and intensity for representing uncertainty on maps. He also stressed the need to understand the objective of a visualization to justify a good design.

Pang et al. identified a family of techniques that are applicable to geospatial uncertainty visualization [73]. Ehlschlaeger et al. [28] showed how animation could be used to depict uncertainty of elevation data. Wittenbrink et al. [108] proposed applying multivariate glyphs for environmental flow visualization (Figure 2.8). Recently, Potter et al. [77] presented an ensemble visualization framework 'Ensemble-Vis' that allows the exploration and generation of visual summaries of a weather ensemble (Figure 2.9).

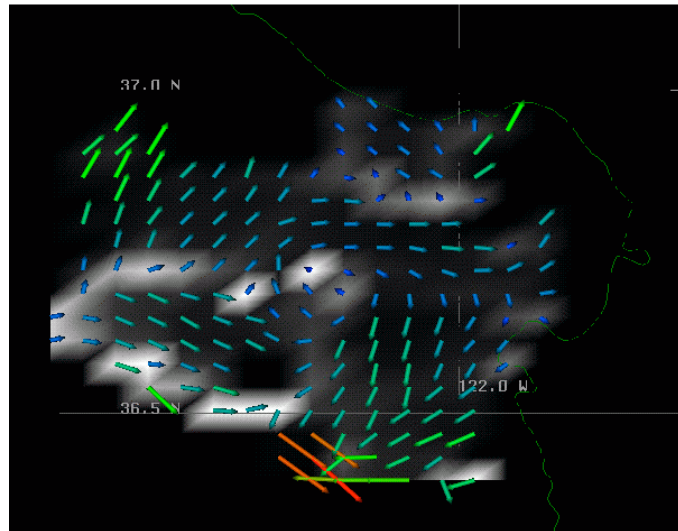


Figure 2.8

Uncertainty in vector data [108].



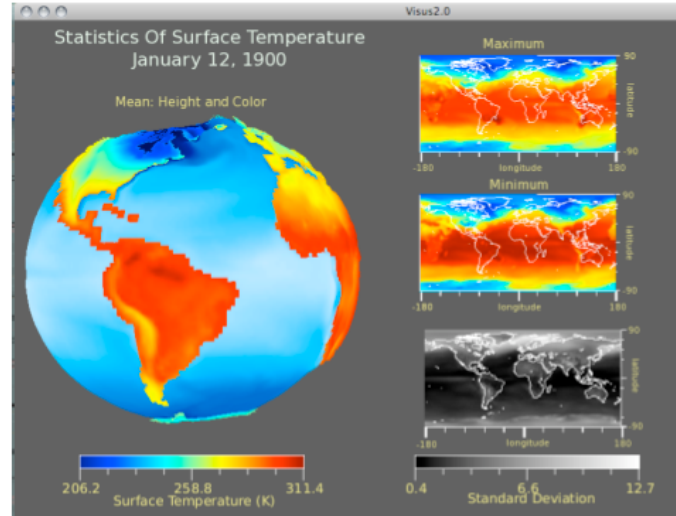


Figure 2.9

Ensemble-Vis plugin for CDAT [77].

### 2.1.1.5 Perceptual Attributes in Uncertainty Visualization

Efforts have been made to identify potential visual attributes that could be used for uncertainty visualization. Hengl and Toomanian [35] illustrated how color mixing and pixel mixing can be used to visualize uncertainty arising from prediction error in spatial prediction models from soil science applications. Jiang et al. [40] used hue and lightness to show fuzzy spatial datasets. Davis and Keller [21] identified value, color, and texture as potentially the best choices for representing uncertainty on static maps. More recently, Hengl [34], like MacEachren [55], made a case for using hue, saturation, and intensity, suggesting the inverse mapping of color saturation to the magnitude of uncertainty.

Generic perception and visualization guidelines by Bertin [6], Tufte [99, 100, 101], and Ware [103] also provide insight that could be of value in representing uncertainty.

### 2.1.2 Evaluation of Uncertainty Visualization Techniques

Some user-studies have also been conducted to bridge the gap between proposed uncertainty visualization techniques and evaluating their effectiveness, but most are domain specific studies [9, 47, 57, 81, 86, 105]. There is some justification to that. Uncertainty representations in many processes are inherent and unique to the task at hand. Blenkinsop and Fisher [9] conducted a user study to evaluate uncertainty visualizations of fuzzy classification of satellite imagery. They found that users were highly successful at determining classification uncertainty using greyscale representations in comparison to random animation and serial animation. They found that serial animation performed the weakest. Zuk and Carpendale [109] presented a theoretical analysis of uncertainty visualization (Table 2.2) in which they evaluated eight uncertainty visualizations from various sources on widely accepted principles from Bertin [6], Tufte [99, 100, 101], and Ware [103]. They presented a set of heuristics and how pertinent each heuristic was with respect to the sampled visualizations. They stressed the need for more research in human factors and perception.

Leitner and Buttenfield [47] conducted an experiment where participants had to make two sitting decisions, one at a park followed by another at an airport based on a set of predetermined planning criteria. The authors found that the addition of certainty information significantly improved the number of correct responses. Additionally, they found that color saturation was not especially effective.

Harrower [32] asked whether the presentation of uncertainty on maps alters the way people solve problems and emphasized the need to conduct longitudinal studies to identify

reasons why a subject makes a correct or an incorrect choice given an uncertainty representation. Couclelis [19] addressed a more fundamental question as to how the uncertain information is processed into knowledge. He identified three different forms of geospatial knowledge production and highlighted the imperfections in each mode. Hunter and Goodchild [38] discussed the issues that need to be addressed to put the theory of uncertainty visualization into practice.

Table 2.2

Potential heuristics and their relevancy to the 8 tested visualizations [109].

Heuristic	Source	Relevance ( <i>n</i> /8)
Ensure visual variable has sufficient length	Bertin and Ware	7
Preserve data to graphic dimensionality	Tufte and Bertin	2
Put the most data in the least space	Tufte	2
Provide multiple levels of detail	Tufte and Ware	2
Remove the extraneous (ink)	Tufte	4
Consider Gestalt Laws	Ware	2
Integrate text wherever relevant	Tufte and Ware	6
Don't expect a reading order from color	Bertin and Ware	1
Color perception varies with size of colored item	Ware and Bertin	2
Local contrast affects color and gray perception	Ware	2
Consider people with color blindness	Ware	2
Preattentive benefits increase with field of view	Bertin and Ware	3
Quantitative assessment requires position or size variation	Bertin	4

## 2.2 Application Areas

In the context of this dissertation, we discuss two application areas which use ensemble simulations that could benefit from improved uncertainty visualization techniques. We

highlight some of the nuances and unique aspects of these fields and discuss some practical issues that one must be aware of.

### **2.2.1 Uncertainty in Weather Simulations**

Weather forecasting is imprecise as the simulation models just attempt to make numerical approximations of the atmospheric processes. The output from these models is sensitive to the initialization conditions, the resolution of the simulation, as well as the choice of physics solvers. In addition, other aspects such as the type of simulation grid, means of data storage, techniques to quantify uncertainty, visualization tools, and finally, experience of the user affect the way uncertainty is perceived. Many of these sources might be difficult if not impossible to eliminate and ensemble simulation offers one way to minimize some of the simulation uncertainties.

While the nature and benefits of using ensemble simulations has been sufficiently highlighted, the following sub-sections discuss some of the other sources of uncertainty and bottlenecks of technology affecting meteorological visualization. Whenever possible, the problem is discussed in the context of ensembles.

#### **2.2.1.1 Coordinate Systems**

When implementing numerical weather prediction simulations, it is common for meteorologists to utilize a sigma, eta, theta, or hybrid vertical coordinate system for finite differentiation [43]. Sigma coordinates (Figure 2.10) consider the ground level as the base. This makes it a terrain following grid with any vertical point calculated as a ratio of the

pressure difference from the point to the top and to the assumed base of the grid. Having a pressure based grid offers computational advantages in solving the governing equations. This is used in models such as the European Centre for Medium-Range Weather Forecasts (ECMWF) [68]. The eta coordinate system (Figure 2.11) is also pressure based; however, the base of the model is taken at mean sea level. The levels in an eta system are horizontal, rendering the numerical formulation of the surface more complex. The theta coordinates (Figure 2.12) uses potential temperature for its vertical coordinate based on the isentropic nature of atmospheric flow. The theta coordinate system performs poorly at the planetary boundary layer where the flow can be strongly non-adiabatic. Thus, hybrid coordinate systems (Figure 2.13) are in use in models such as the Rapid Update Cycle – 2 (RUC–2) [5] which uses the theta coordinate system in combination with a sigma coordinate system to leverage the modeling advantages offered by each system.

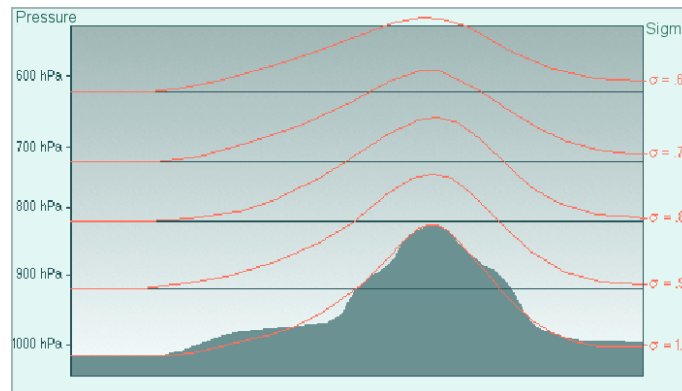


Figure 2.10

The sigma coordinate system.

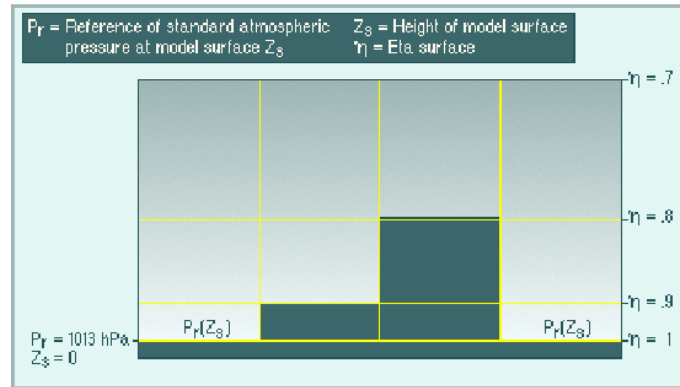


Figure 2.11

The eta coordinate system.

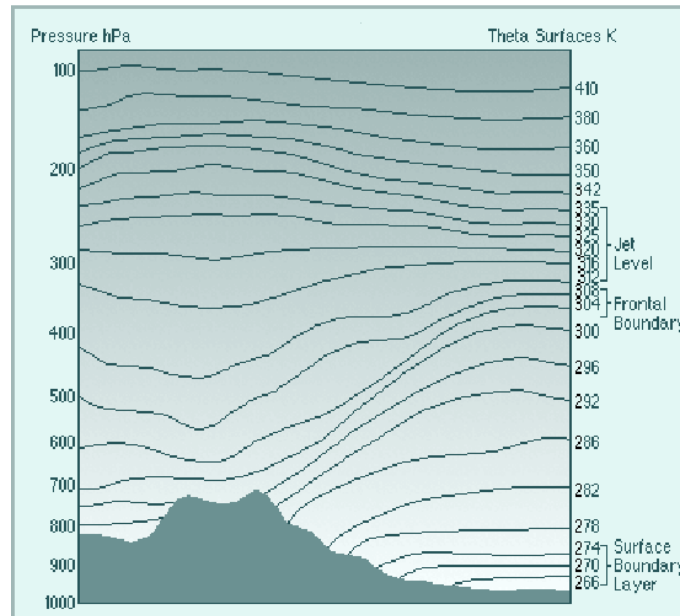


Figure 2.12

The isentropic coordinate system.

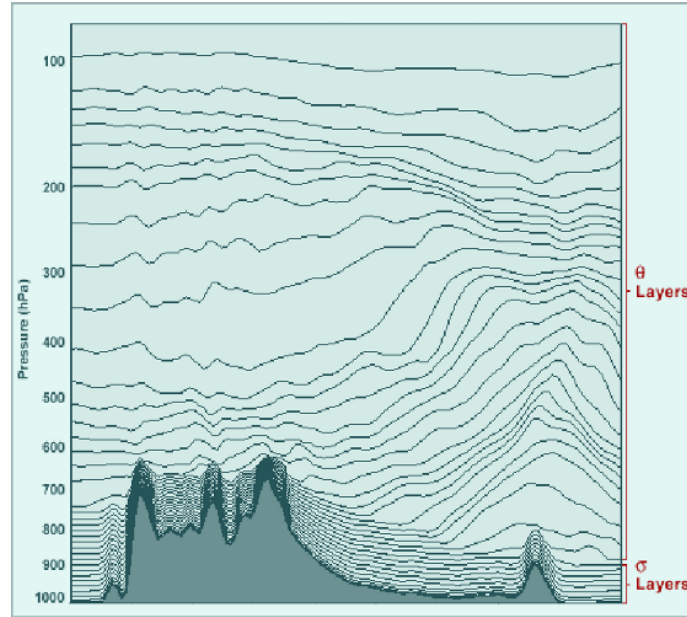


Figure 2.13

The hybrid coordinate system.

Additionally, simulation models often use a staggered simulation grid to minimize aliasing artifacts in the computation. For example, the Weather Research and Forecasting Model (WRF) [67] uses an Arakawa-C grid (Figure 2.14) [1]. This affects the way in which the final output is obtained.

While these coordinate systems offer unique advantages in modeling, it makes the task of visualization much more complex. Meteorologists almost exclusively utilize isobaric or isentropic surfaces when analyzing vertical layers of the atmosphere, often combining several such surfaces to render a mental 3-dimensional image of the physical processes occurring at the given time.

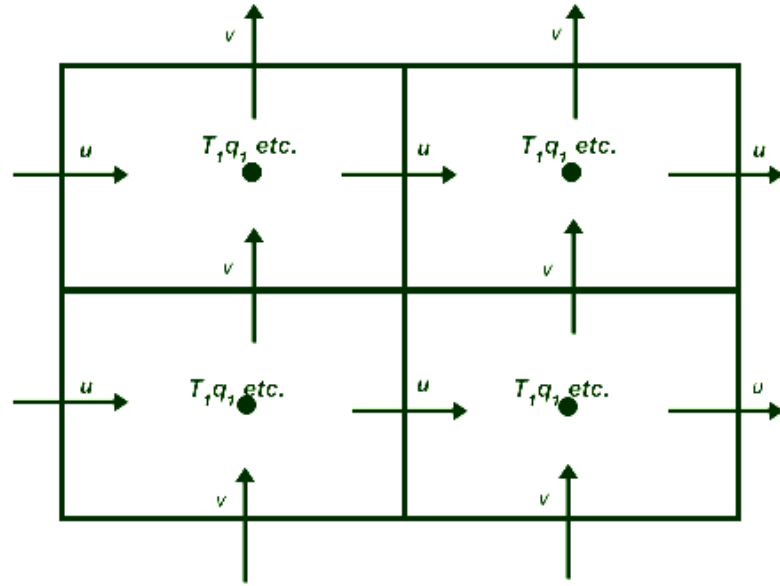


Figure 2.14

Example of variables on the Arakawa C grid.

### 2.2.1.2 Data Formats

There are a large number of geo-scientific data formats such as Grib 1 [91], Grib 2 [22], HDF [71], and NetCDF [78]. The data in a Grib file are stored as messages where each message can be accessed to retrieve the data for a given level. One has to cycle through a set of messages to obtain a 3D block of data. HDF and NetCDF formats have been gaining popularity because of the open nature of the standard which allows one to encode almost any kind of data. This comes at a cost of visualization tools being unable to parse the data structures successfully for some files. Not much driver support is available in terms of parallel and high performance access to data stored in these formats either. Li



et al. [50] have initiated the development of a parallel version of the NetCDF format tuned to High Performance Computing applications which is currently in progress.

Irrespective of the data format, variables are stored on the calculated pressure surfaces. One must access the geopotential height values of the pressure grid to determine the exact physical location for accurate 3D visualization. Often this is a double-redirection which consumes long cpu cycles. In addition, different variables may be on different spatially staggered grids.

Thus, data format issues complicate the process of visualization design. Adding an ensemble dimension complicates the visualization issues even further as the visualization tool must be able to determine the computational grid, associated grid staggering, and appropriately visualize uncertainty. The pressure grids may not agree across model runs and informed choice of a height grid must be made in 3D ensemble visualization.

### **2.2.1.3 User Training**

Primarily due to training and habit, meteorologists typically use 2D slices and spaghetti plots to understand their data. They go through years of training to attain the expertise to mentally visualize the 3D and temporal nature of the atmosphere. Good user training in 2D, 3D, and temporal visualization techniques with interactive analysis could enable expert decision support.

#### **2.2.1.4 Ensemble Uncertainty Quantification**

Metrics such as standard deviation, inter-quartile range, and confidence intervals are often used to quantify the uncertainty in ensemble simulations. While ensemble forecasting consumes significant computing resources by itself, additional CPU time is required for uncertainty calculation. If there are a few ensemble members, bootstrapping [27] can provide a robust approach to estimate uncertainty without being constrained by the requirements of a normal distribution (see section 5.3.1.1). Calculation of the bootstrapped distribution is computationally intensive; however, it is highly parallelizable and scalable.

#### **2.2.1.5 Technological Gap**

The scientific visualization literature provides an abundance of techniques to visualize 2D and 3D data [12, 26, 31, 36, 63]. Techniques such as isosurfacing [52], line-integral convolution [16], and volume visualization [25] have found their way into tools such as Paraview [90], Integrated Data Viewer [69], and VTK [84]. Many other techniques such as multi-resolution [98] and multi-field [80] visualization, information visualization based approaches [2, 92], and illustrative [42] rendering techniques have been developed. In spite of the availability of these tools and techniques, operational meteorologists still use spaghetti plots to visualize ensemble output, partly due to a technology gap between operational requirements and tool capability. While these tools are capable of visualizing weather data, they are not designed for ensemble visualization. These tools can load multiple co-located datasets, but these datasets are not treated as members of an ensemble. As a result, these tools do not allow operational personnel to visualize ensemble mem-

ber uncertainty. There is no complete tool suite available to meteorologists that has been designed specifically for large ensemble data management and analysis. The only such software that exists is a plugin made by Potter et al. [77] for the Climate Data Analysis Tools [106].

### **2.2.2 Uncertainty in River-Flow and Inundation Simulations**

Flooding is the number one cause of disaster and human peril in the United States, contributing to 1,100 disasters which is nearly two-thirds of 1,720 Federal disasters declared from 1953 to 2007 [70]. Thirteen National Weather Service's (NWS) River Forecast Centers (RFCs) provide daily river forecasts, flood warnings, flash flood guidance, and extended forecast information for water resources management with a mission to save lives and decrease property damage and provide for the nation's economic and environmental well being.

There are three types of flooding scenarios, riverine flooding, coastal flooding, and flooding in interconnected ponds. There are various influencing factors affecting the models used for each scenario. Riverine flooding is the most studied and best understood scenario. Uncertainties in simulations of riverine areas stem from inaccurate terrain information (by far the largest contributor), hydrologic uncertainties in flood discharge, and hydraulic uncertainties in converting to a flood-water surface level. Much like meteorology, ensemble simulations of riverine flow attempt to capture some of the modeling uncertainties.

The Federal Emergency Management Agency's (FEMA) Risk Analysis Division and National Oceanographic and Atmospheric Administration's (NOAA) Coastal Services Center conducted a study to understand and evaluate the factors affecting flood map accuracy to improve the mapping, communication, and data used [70]. They also studied the economic impacts of inaccuracies for flooding across riverine, coastal, and ponded landscapes.

Flooding in areas of interconnected ponds (e.g. Florida) is the least understood scenario and recommendations have been made for more research to better model, understand, and predict for such areas. The following sub-sections highlight some of the findings that affect uncertainty in flood modeling and mapping.

#### **2.2.2.1 Accuracy of Topographic Data**

The accuracy of topographic data was found to be the biggest contributor affecting the accuracy of flood-plains. The United States Geological Survey (USGS) National Elevation Dataset (NED) is typically used in floodmap production but the uncertainties in this dataset are over 10 times the acceptable limit determined by FEMA. Lidar (light-detection and ranging) datasets are the most accurate but are not available extensively. The differences are highlighted in Figure 2.15 for for Beaufort County, NC, where the light blue areas represent uncertainty in the extent of inundation at the 95% confidence level for the two types of topographical data.

Additionally, problems also arise from differences in the datums in use: National Geodetic Vertical Datum of 1929 (NGVD29), North American Vertical Datum (NAVD88) of 1988 (Figure 2.16), and numerous tidal datums. An important component is bathymetry

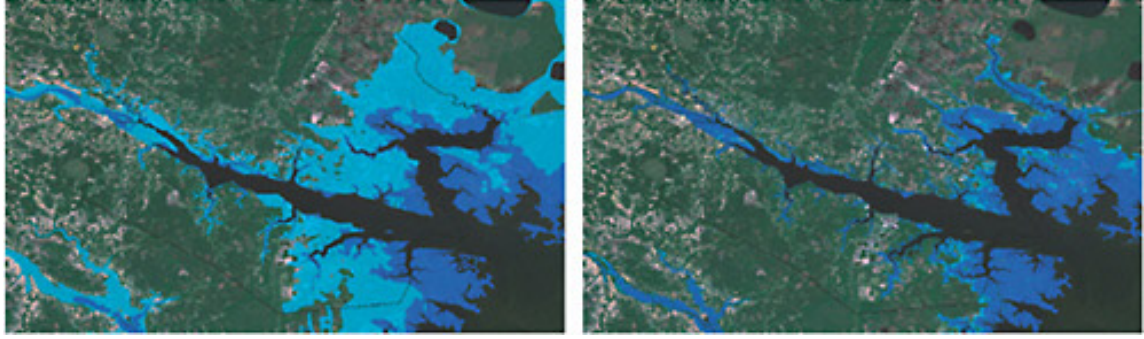


Figure 2.15

Flood plain extent with a 30m DEM (left) and 3m Lidar DEM (right) [70].

data. No technology exists for obtaining detailed and accurate measurements of bottom surfaces of rivers and water bodies.

#### 2.2.2.2 Structure Elevations

The presence of hydraulic structures such as bridges, culverts, and dams affect the base water elevations before and after the structure. It is important to be able to capture this information and model for it.

#### 2.2.2.3 Inland Flooding

Inland flooding is the most studied and best understood type of flooding. There are three main sources of uncertainty and each component introduces unique scientific challenges and scope of study. The three sources are:

- Hydrologic uncertainty in evaluating the base flood discharge
- Hydraulic uncertainty in simulation of the water surface elevation
- Mapping uncertainty of the floodplain boundary

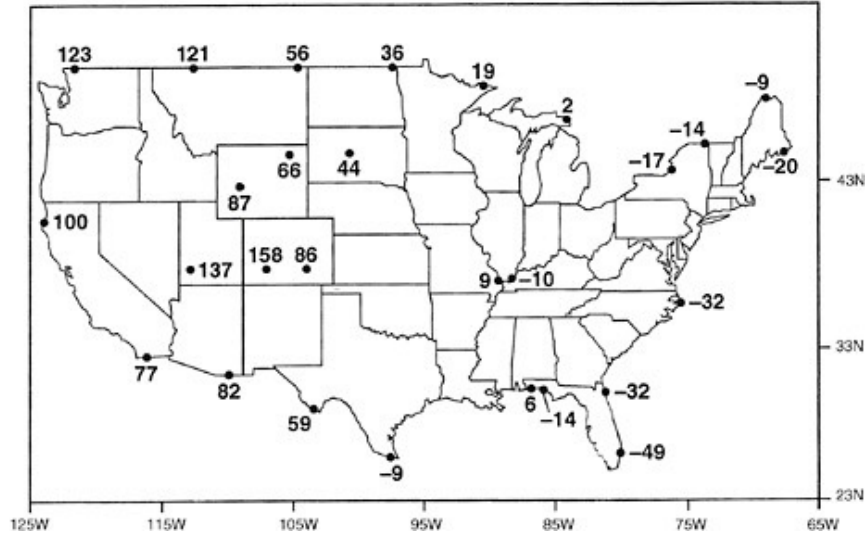


Figure 2.16

Differences in heights (NAVD 88 minus NGVD 29) in centimetres [70].

#### 2.2.2.4 Coastal Flooding

Storm surge, tides, and waves are the greatest causes of coastal flooding. Lately, there have been many improvements in coastal flood modeling. Future recommendations to improve the models used in coastal flood mapping are directed towards using coupled two-dimensional storm surge and wave models. It is worthwhile to note that coastal terrain and bathymetry changes constantly and efforts must be made to keep the data current.

## CHAPTER 3

### APPROACH AND TECHNIQUES

Our initial effort was to design a framework for the evaluation of uncertainty. We first discuss this framework and its merits followed by discussions of various uncertainty visualization techniques derived from the framework and applicable to 1D and 2D data.

#### **3.1 Uncertainty Evaluation Framework**

Visualization techniques are typically designed to operate on data of a certain dimensionality. Sometimes the techniques can be extended to operate on data of higher or lower dimensionality. For example, an isosurface is a 3D version of a 2D contour. We present an Uncertainty Evaluation Framework that provides researchers with a structured classification to evaluate existing uncertainty visualization techniques (Figure 3.1). The framework was designed to consciously think of uncertainty from the perspective of the data being visualized and not by the uncertainty visualization technique employed. This framework also has the potential to provide a basis for development of new techniques and future user-studies.

Spatial data can be thought of as having zero, one, two, or three dimensions. In many applications, it is common to consider time as the outermost dimension. We decouple the temporal dimension and treat it specially because the temporal dimension usually has very

different features and resolution than the spatial dimensions, particularly in simulations. It is worthwhile to note that this may not be true in the field of information visualization, where it is very common to have multidimensional data that is not spatial or temporal [103].

Further, scalars, vectors, and tensors can be thought of as three types of scientific visualization paradigms. Thus, data dimensionality (0D, 1D, 2D, 3D), visualization paradigm (scalar, vector, tensor), and the broad taxonomy of uncertainty visualization techniques (blurring, transparency, noise, etc.) form the three axes that define our classification (Figure 3.1).





<i>Data Dimension</i>	<i>Technique</i>			
	<i>Glyphs-size</i>	<i>Blurring</i>	<i>Transparency</i>	<i>Other Methods</i>
0D 	Point size	Point fadeout	Point visibility	...
1D 	Glyphs on the line	Line fadeout	Line visibility	...
2D 	Glyphs on surface	Surface fadeout	Patch visibility	...
3D 	Volumetric glyphs	Volumetric fadeout	Regional visibility	...

Figure 3.1

The Uncertainty Evaluation Framework.

Our framework allows one to replace the technique axis with other classification schemes such as that of Pang et al. [74]. Additionally, the entire framework has a temporal axis.



This is similar to the taxonomy proposed by Tory and Möller [97] but is more flexible, since this framework allows researchers to structurally extend the technique axis across other classification schemes and data dimensions.

### 3.2 Uncertainty Visualization Techniques

We explored various techniques for the visualization of uncertainty in the context of the types of datasets used in the application domains for this dissertation. The following sections describe these techniques in detail.

#### 3.2.1 Glyphs Altered by Size

This representation alters the size of a round circular glyph depending on the uncertainty in the data. The data itself is represented by a 1D plot or a 2D surface (Figure 3.2 and Figure 3.3).

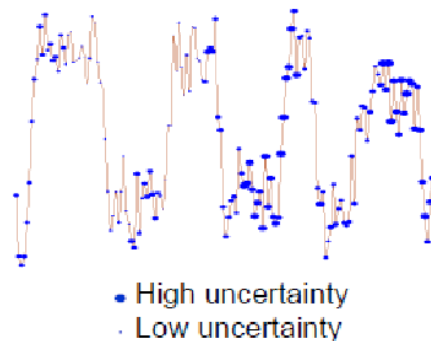


Figure 3.2

Uncertainty represented by glyphs altered by size for 1D data.

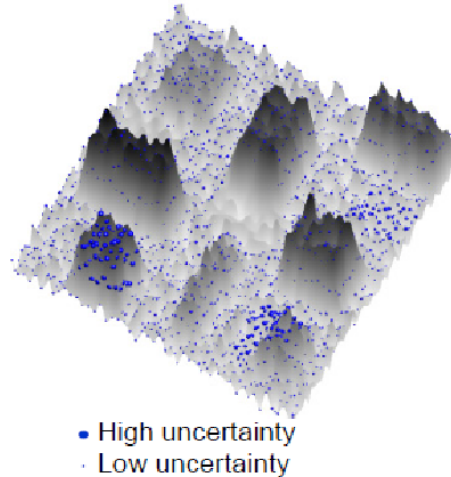


Figure 3.3

Uncertainty represented by glyphs altered by size for 2D data.

### 3.2.2 Glyphs Altered by Color

Altering the color of fixed sized circular glyphs is another method to represent the uncertainty in the data. The data itself is represented by a 1D plot or a 2D surface (Figure 3.4 and Figure 3.5).

### 3.2.3 Color-Mapped Lines and Surface

The uncertainty value itself is mapped to the color of the line in the 1D case and surface in the 2D case (Figure 3.6 and Figure 3.7).

### 3.2.4 Gradient and Striped Gradient

To improve the visual perception of the magnitude of uncertainty, we applied a gradient to the uncertainty range about a data value resulting in a ribbon-like representation

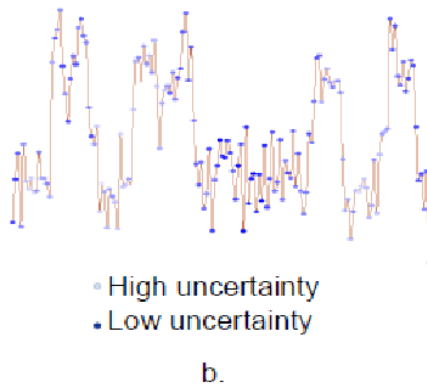


Figure 3.4

Uncertainty represented by glyphs altered by color for 1D data.

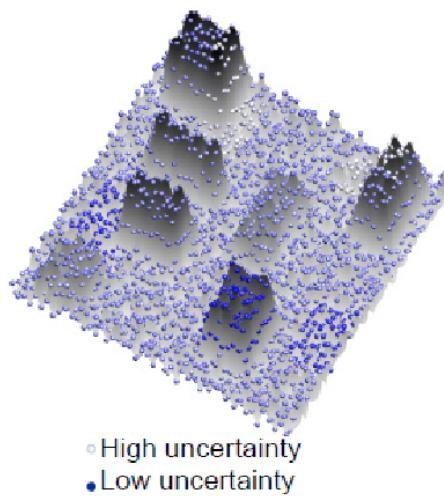


Figure 3.5

Uncertainty represented by glyphs altered by size for 2D data.

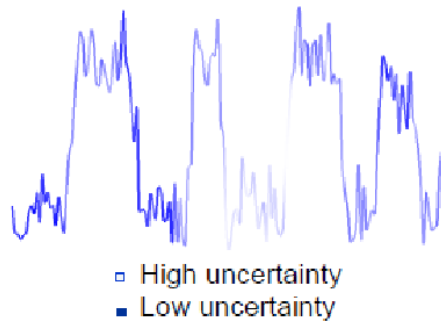


Figure 3.6

Uncertainty represented by altering the line color for 1D data.

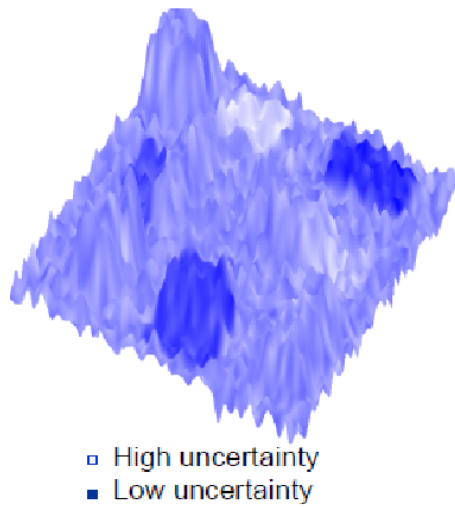


Figure 3.7

Uncertainty represented by altering the surface color for 2D data.

in the 1D case (Figure 3.8) a vertical glyph representation in the 2D case (Figure 3.9). We attempted to improve the representation by binning the uncertainty values and creating a striped gradient representation (Figure 3.10). We found these representations to be difficult to read especially when the gradient of the data surface was steep.

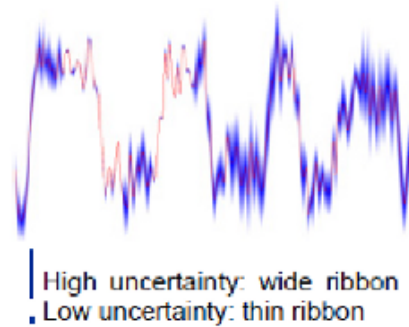


Figure 3.8

Uncertainty represented by the color gradient for 1D data.

### 3.2.5 Errorbars

We experimented with the conventionally used errorbars for both 1D and 2D data resulting in uncertainty representations on a 1D plot as well as a 2D surface (Figure 3.11 and Figure 3.12).

### 3.2.6 Boxplots

The next most logical representation to implement were box-plots. These plots are more descriptive than simple glyph representations as the user can get a sense of the dis-

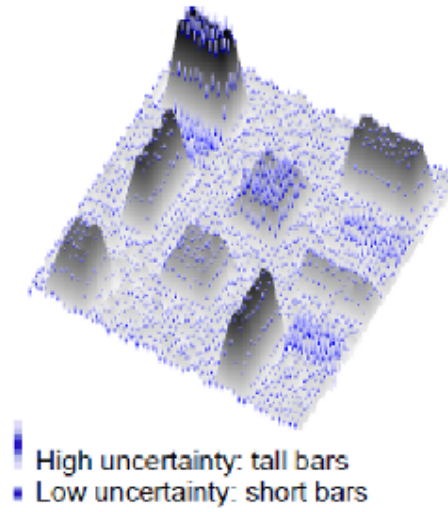


Figure 3.9

Uncertainty represented by the color gradient on vertical glyphs for 2D data.

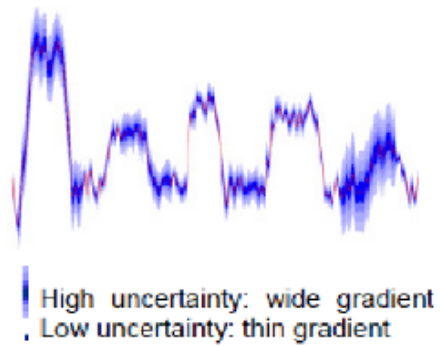


Figure 3.10

Use of striped gradient for 1D data.

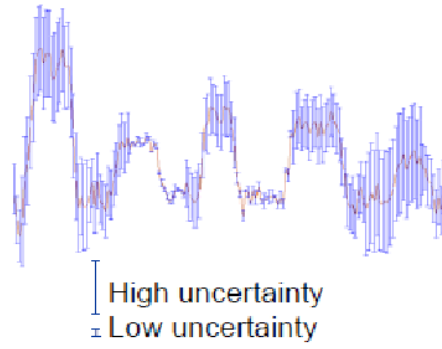


Figure 3.11

Uncertainty represented by errorbars for 1D data.

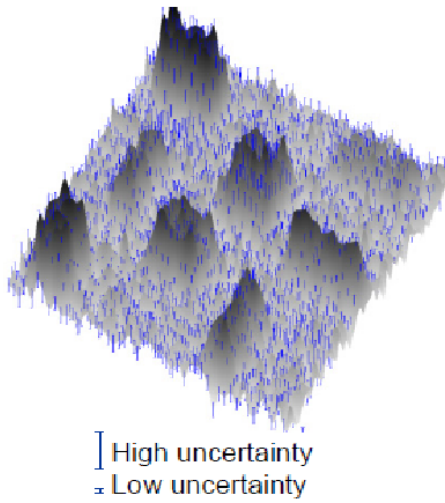


Figure 3.12

Uncertainty represented by errorbars for 2D data.

tribution of the data values as well as statistics such as the minimum, maximum, quartiles, and the median.

### 3.2.7 Graduated Uncertainty Glyphs

It is important to have some sense of the distribution of the individual data values. We experimented with encoding individual data values on circular glyphs as well to create what we call ‘graduated uncertainty glyphs’.

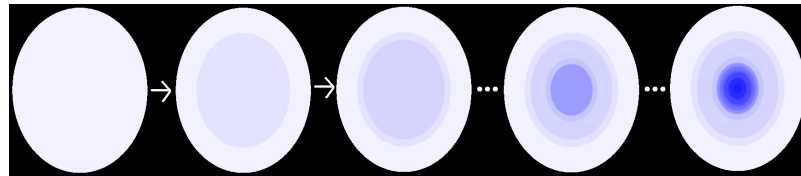


Figure 3.13

Construction of graduated uncertainty glyphs.

Mathematically, let  $u$  be the measure of uncertainty,  $s$  be the maximum desired size of the glyphs, and  $U_m$  be the maximum value of the uncertainty metric. The radius  $r$  of a single glyph is given by

$$r = \frac{su}{U_m} \quad (3.1)$$

Let  $M$  be the mean of  $n$  data values  $v_1, v_2, v_3, \dots, v_n$ . Let the absolute difference of each value  $v_i$  from the mean  $M$  be  $d_i$ . Then, the differences for all data values is given by  $d_1, d_2, d_3, \dots, d_n$ . These difference values are then sorted in increasing order to generate a



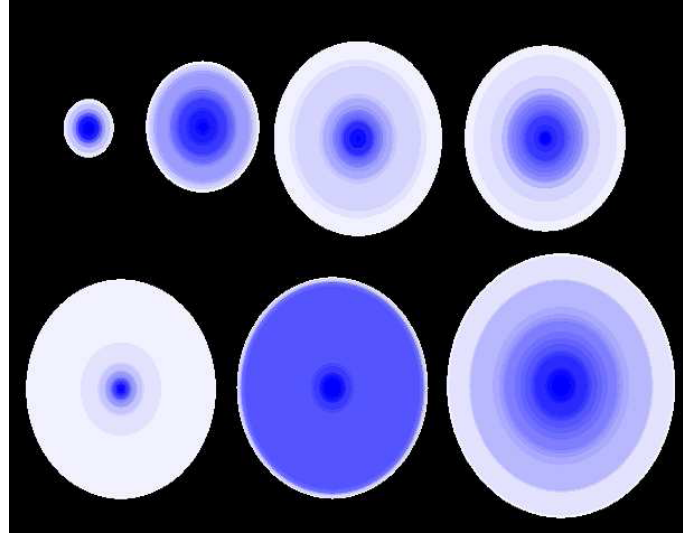


Figure 3.14

Different types of graduated uncertainty glyphs.

new array  $D_1, D_2, D_3, \dots, D_n$ , where  $D_1$  represents the smallest and  $D_n$  the largest of the difference values.

The basic idea is to use these difference values to construct concentric circular glyphs, starting with the largest difference value,  $D_n$  and rendering successively smaller glyphs  $D_{n-1}, D_{n-2}, D_{n-3}, \dots, D_1$ . This ensures that successively rendered glyphs are either smaller-than or equal-to the size of the previously rendered glyphs (Figure 3.13).

Thus, equation (3.1) can be rewritten to give the radius  $r_i$  of the  $i^{th}$  glyph as

$$r_i = \frac{sD_i}{D_m} \quad (3.2)$$

where  $D_m$  is the maximum difference of any data value to the mean for the given variable in the entire data.

Various levels of saturation of a color can be used to color these  $n$  concentric glyphs. The saturation level  $sat_i$  for the  $i^{th}$  glyph is given by

$$sat_i = \frac{i - 1}{n - 1} \quad (3.3)$$

This results in a uniform distribution of the range of saturation values (0 ... 1) over the  $n$  data values. The largest glyph (derived from the largest difference value) ends up having the least-saturated color, and each successive glyph gets a more saturated shade of the color. This creates an overloaded visualization that encodes the distribution and variability between the data values (Figure 3.14).

### 3.2.8 Uncertainty Ribbon

An uncertainty ribbon is generated to quantify the uncertainty along a contour of a value for 2D datasets. The width of the ribbon represents the uncertainty along a contour.

Mathematically, let  $u_i$  be the uncertainty measure at the  $i^{th}$  location along a contour for a given iso-value. The uncertainty measure  $u_i$  can be used to derive the radius  $r_i$  of a hypothetical circle  $c_i$  at that location, which is given by:

$$r_i = w \frac{u_i}{U_m} \quad (3.4)$$

where  $w$  is a user-chosen value that controls the maximum width of the ribbon, and  $U_m$  is the maximum value of the chosen uncertainty metric in the data.

A segment of the uncertainty ribbon can be constructed by first calculating the external tangents,  $t_{ia}$  and  $t_{ib}$ , of circles,  $c_i$  and  $c_{i+k}$ , where  $k$  is a user-specified skip distance along

the contour (Figure 3.15). Connecting the ends of the tangents  $t_{ia}$  and  $t_{ib}$  generates a quadrilateral that forms a segment of the ribbon (Figure 3.15). This is repeated along the contour to generate the complete ribbon. There is a possibility that adjacent circles lie within one another, in which case the algorithm skips to the next non-inscribed circle.

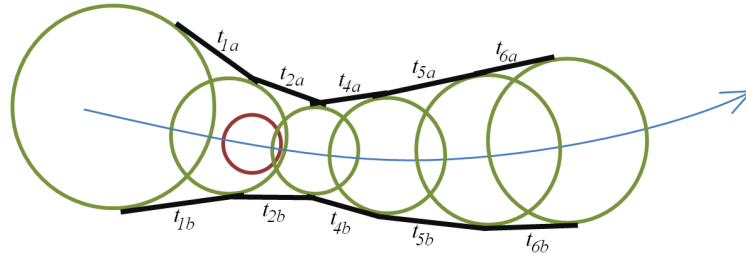


Figure 3.15

Construction of uncertainty ribbon.

### 3.2.9 Graduated Uncertainty Ribbon

Similar in spirit to the graduated uncertainty glyphs, the ‘graduated uncertainty ribbon’ encodes the distribution of the data values. The construction of the graduated uncertainty ribbon is similar to the graduated uncertainty glyphs. While the basic idea stays the same, additional bookkeeping is necessary since consecutive locations are required for calculating the tangents.

Let  $n$  be the number of data values at a given 2D location and let  $k$  be the number of contour segments. Also, let  $d_{i,j}$  be the difference of the  $i^{th}$  data value from the mean  $M$

for the  $j^{th}$  location along the contour. The difference values at each location are sorted in increasing order producing the array  $D_{1,j}, D_{2,j}, D_{3,j}, \dots, D_{n,k}$  for all contour locations.

Once this is complete for all points along the contour, the algorithm must render multiple ribbons along the contour to give the graduated effect (Figure 3.16). This is accomplished by first rendering a wide ribbon for the largest difference values along the contour. Thus,  $D_{n,j}$  for all  $j$  are used to generate the widest uncertainty ribbon. It is also given the least saturated color. Successive ribbons are generated similarly from  $D_{n-1,j}, D_{n-2,j}, D_{n-3,j}, \dots, D_{1,j}$  for all  $j$ . The saturation,  $sat_j$ , of the color for the  $j^{th}$  ribbon is given by

$$sat_j = \frac{j - 1}{n - 1} \quad (3.5)$$

The resulting visualization is a graduated uncertainty ribbon.

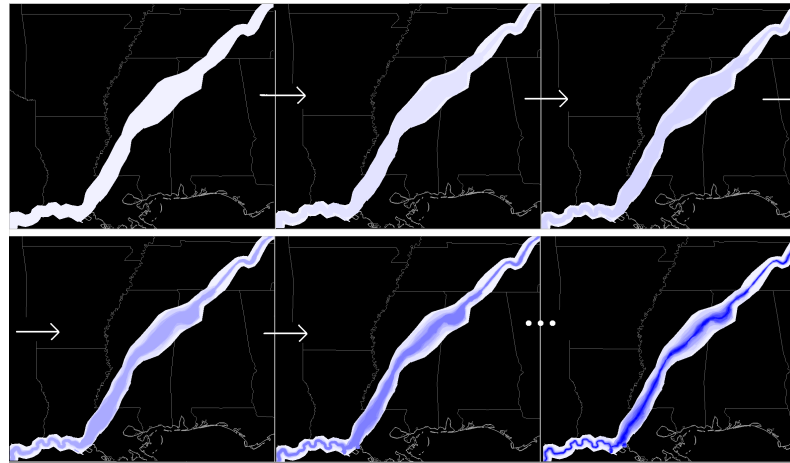


Figure 3.16

Construction of graduated uncertainty ribbon.

### 3.2.10 Animation

We also experimented with animation of glyphs in the uncertainty range about the 1D data plot (Figure 3.17) or 2D surface (Figure 3.17).

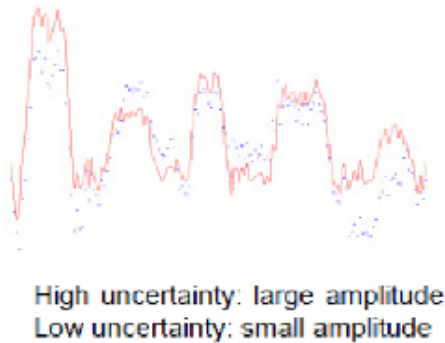


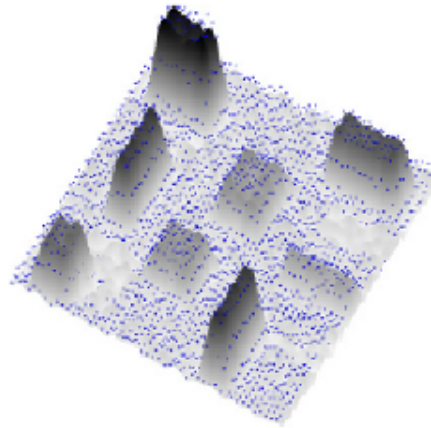
Figure 3.17

Uncertainty represented by animation of glyphs for 1D data.

We also tried to use animation of the plot or surface itself to communicate the uncertainty (Figure 3.19 and Figure 3.20). These representations suffered severely from occlusion and clutter.

### 3.2.11 Multivariate Uncertainty Glyphs

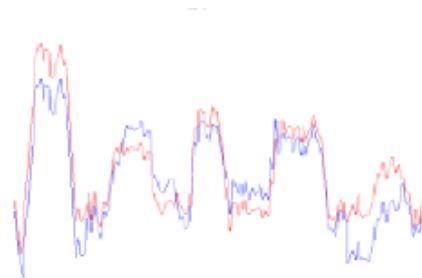
Often multiple variables are essential in data visualization. We used a star representation where each point of the star corresponded to a different variable to create a multivariate visualization. A circle around the glyph demarcated the maximum value possible. The saturation of the point encoded the uncertainty (Figure 3.21).



High uncertainty: large amplitude  
Low uncertainty: small amplitude

Figure 3.18

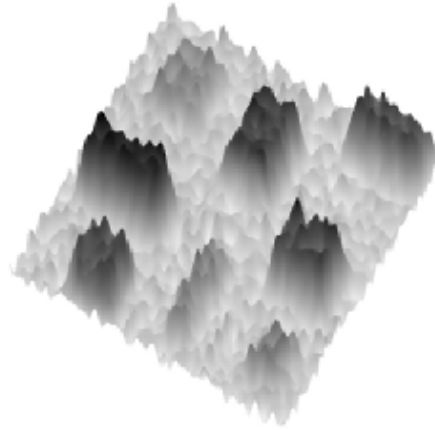
Uncertainty represented by animation of glyphs for 2D data.



High uncertainty: large amplitude  
Low uncertainty: small amplitude

Figure 3.19

Uncertainty represented by animation of surface for 1D data.



High uncertainty: large amplitude  
Low uncertainty: small amplitude

Figure 3.20

Uncertainty represented by animation of surface for 2D data.

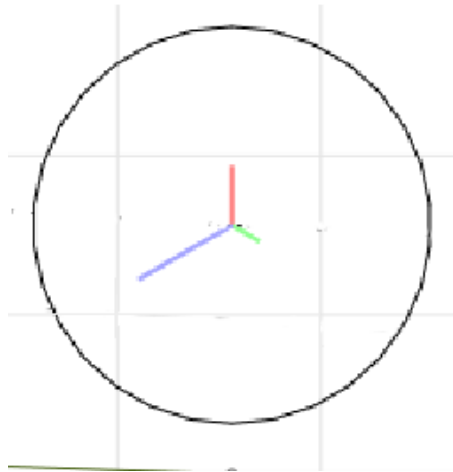


Figure 3.21

A multivariate uncertainty glyph.

We constructed a variation of these plots where each point sweeps out a section of a pie as time progresses illustrating temporal multivariate uncertainty representation (Figure 3.22).

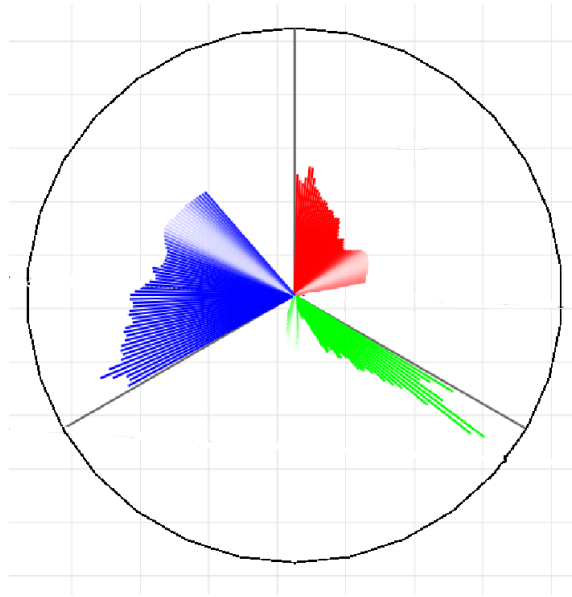


Figure 3.22

Temporal multivariate glyphs illustrating multivariate uncertainty.

### 3.2.12 Spaghetti Plots

We implemented spaghetti plots which are a common and conventional technique in meteorology. These are created as a collection of iso-valued contours from the different members in a 2D dataset.



CHAPTER 4  
QUANTITATIVE EVALUATION OF UNCERTAINTY VISUALIZATION  
TECHNIQUES

Insight and quality assurance can be improved by recording uncertainty along with data. The practical benefits of understanding uncertainty in a scientific process can be manifold. A cursory glance at literature in almost all disciplines will indicate that visual representations to depict the recorded or calculated uncertainties are underdeveloped. The incidence of finding charts of one-dimensional data augmented with errorbars is reasonably high; however, as we move on to data of higher dimensions, visual metaphors to represent the uncertainty are rarely used. Part of the reason is that although a variety of techniques have been suggested, successfully applying them to make insightful visualizations is very challenging and there is a lack of guidance on which uncertainty method will yield the best results. In this chapter, we present our findings from an uncertainty visualization user study which we believe could help to improve future visualization designs.

A literature study indicated that some uncertainty visualization techniques seemed to appear more effective than others, however, little comparison has been done to evaluate the effectiveness of most of these techniques. Keeping this in mind, we constructed a user study to evaluate the effectiveness of four commonly used uncertainty visualization techniques:

- Size of the uncertainty glyphs
- Color of the uncertainty glyphs
- Color of the data surface
- Errorbars

In this study, these techniques were applied to both one dimensional (1D) and two-dimensional (2D) simulated datasets. We define 1D data as samples from a curve, which is a 1D manifold embedded in a 2D Euclidean space. We define 2D data as samples from a surface, which is a 2D manifold embedded in a 3D Euclidean space. Moreover, the definition for our datasets is such that the 1D data is defined by a 1D function  $f(x)$ , and the 2D data defined by a function  $f(x, y)$ . These definitions of 1D and 2D data in this paper were chosen keeping applications of geoscience visualization and analysis in mind.

Our objective was to design a systematic and general user study to evaluate the effectiveness of common uncertainty visualization techniques. We generated synthetic 1D and 2D datasets to avoid being tied to any specific application domain. We also chose common tasks in scientific data analysis, such as searching and counting, to evaluate the techniques chosen for our study. Our user study aims to bridge some of the gap in understanding the circumstances that govern the decision making process in the presence of uncertain information.

We borrowed design ideas from previous uncertainty visualization user studies [9, 32, 86, 105], as well as others, such as the 2D vector field visualization user-study by Laidlaw et al. [45], and the hurricane visualization user study by Martin et al. [58]. The following

section discusses the design of the study, followed by the data analysis and a discussion of the results.

## **4.1 Study Design**

This section discusses the data used, the uncertainty visualization techniques evaluated, the questions asked, the participants, and various aspects of main study.

### **4.1.1 Data Generation**

The data generation process was motivated by geoscience applications of visualization, which typically deal with various types of remotely sensed data, observed data at stations (e.g., buoys), data over a trajectory (e.g., weather balloons), simulated weather data (e.g., output from numerical models) and statistical studies (e.g., temporal correlations). Our objective was to design a controlled synthetic-data generation scheme that would be specific enough to provide immediate insight into geoscience uncertainty representation, as well as be generic enough to potentially have other applications.

We devised a mathematical method to simulate the data acquisition process and hence have complete control over the uncertainties introduced at different stages in a real data collection process. We simulated the process of repeated data collection, where, if any data measurement task is repeated a large number of times, the recorded values end up being normally distributed. If we take a subset of these values, we can derive a mean data value and a corresponding uncertainty value. We also introduced systematic uncertainty components that are an inherent part of any data collection process.

Matlab [95] was used to generate the datasets. We first describe the process for the one-dimensional case and then extend it to two dimensions. We begin with a 1D array, say  $A$ , consisting of 40 zeros and manually implant data features by setting consecutive index locations in  $A$  to a certain value representative of the signal strength, say  $S$ , at that location (dark regions in Figure 4.1). The value  $S$  was generated as a normally distributed random number (using the Matlab *randn* function) with a mean of 0.8 and a standard deviation of 0.2. Thus, we now have an array of zeros with user-defined data features embedded in the array (Figure 4.1). Let this array be  $A'$ . In the next step, we interpolate the array  $A'$  (with the Matlab *interp1* function using cubic spline interpolation) to implant 3 points between every pair of array locations to generate 4 levels between them (Figure 4.1 and Figure 4.2). Let us call this array  $A_{true}$ . For our simulated dataset,  $A_{true}$  will represent the “truth value”, which is analogous to the exact value of a continuous real-life phenomena such as the temperature of a place or water level of a sea surface that no instrument can ever record “exactly”.

We then simulated the act of taking measurements or observations of the data. If a measurement is taken a large number of times, the errors in the observations can also be assumed to be normally distributed around the truth value assuming no systematic error, which constitutes the random uncertainty component. To simulate random uncertainty in our datasets, we generated 50 sets of readings, where every observation is normally distributed about values of the assumed true data  $A_{true}$  (Figure 4.2). Let these sets of values be  $A_0, A_1, \dots, A_{49}$ . To generate these, we first found the mean  $\mu_{true}$  and standard deviation  $\sigma_{true}$  of the true data  $A_{true}$ . We used fractions and multiples of the standard

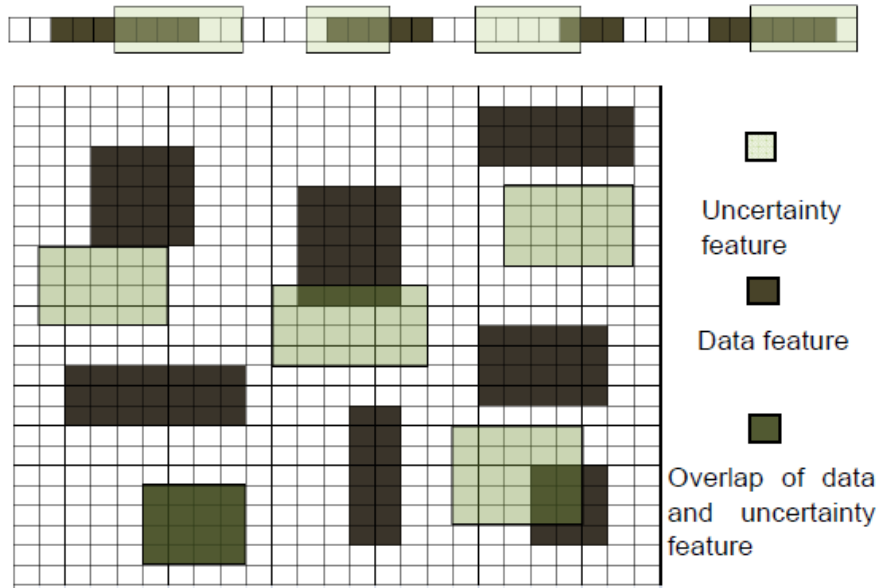


Figure 4.1

Systemic and random components in the synthetic dataset.

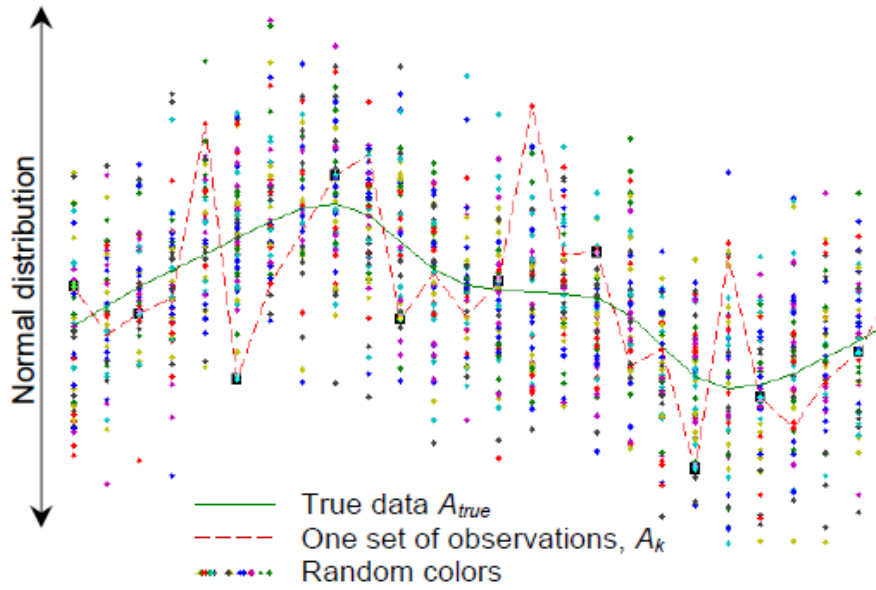


Figure 4.2

Generation of data for the user-study.

deviation  $\sigma_{true}$ , say  $k$ , to generate three types of datasets having three ranges of random uncertainty in the data. Using the standard deviation of the truth data to generate multiple uncertainty levels seemed like a reasonable choice. We used three values for  $k$  which were 0.5, 1.0 and 1.5.

For the dataset  $A_0$ , the  $i^{th}$  element of  $A_0$  corresponds to the  $i^{th}$  element of  $A_{true}$  modified by the generated uncertainty.

Thus,

$$A_{0,i} = A_{true,i} + randn() * k * \sigma_{true} \quad (4.1)$$

We then took the first 10 observation sets,  $A_0, A_1, \dots, A_9$ , and calculated the mean and standard deviation for each index, generating the dataset  $A''$ . This simulated the real-life step of averaging multiple data readings, and the standard deviation represented the uncertainty of the average.

An uncertainty study dataset is incomplete without a systematic uncertainty component. In real-life situations, often the uncertainty in the data exhibits patterns. This can be because of the nature of certain regions of the data, biases in the sensors, and a variety of other reasons. We introduced systematic biases in the generated random uncertainty by manually biasing certain sections of the array (Green regions in Figure 4.1). The bias values, say  $B'$ , were normally distributed random values with a mean of 0.4 and standard deviation of 0.1 and were arbitrarily added or subtracted from the standard deviation values in the chosen sections of the array. This generated our uncertainty features.

The 1D datasets had four different feature layouts (Figure 4.1), each with three values of  $k$ , making a total of twelve 1D datasets. We generated one extra dataset for use in the training of the participants.

The same logic was extended to generate 2D datasets (Figure 4.1). We started with a grid of  $25 \times 25$  zeros and planted rectangular user-defined data features. The grid was interpolated (using Matlab *interp2* function) along both  $x$  and  $y$  axes to create 2 levels between every 2 grid points. 50 sets of pseudo-readings were generated in exactly the same way as that of the 1D case. We took the first 10 observation sets and generated the average signal value, the uncertainty value, and added rectangular uncertainty features. All parameters were kept the same as in the 1D case. Twelve 2D datasets were created for the main study and an additional dataset was created for the training module.

We do acknowledge that using real data from real sources has its merits, most notably being able to establish direct returns from the results of a user-study. We also acknowledge that not all data is normally distributed. We did not perform any tests on any real data or on other data distributions due to constraints of time.

#### **4.1.2 Uncertainty Visualization Techniques Chosen for Evaluation**

Using our uncertainty visualization framework, we chose four visualization techniques that could be applied to both 1D and 2D data. These were scaled sizes of glyphs, altering the color attribute of glyphs keeping the size constant, color-mapping the data surface with the uncertainty, and traditional errorbars (Figure 4.3 and Figure 4.4). The data was displayed in greyscale except where colormapping of the surface was used. The 2D data

surface was rendered with an orthographic projection to minimize 3D perspective effects that may interfere with perception of height. This ensured that the uncertainty representations would be of uniform size regardless of the distance from the eye. A few other techniques such as smooth and striped gradients, animation of glyphs, and animation of the data surface were considered but were not included in the user study.

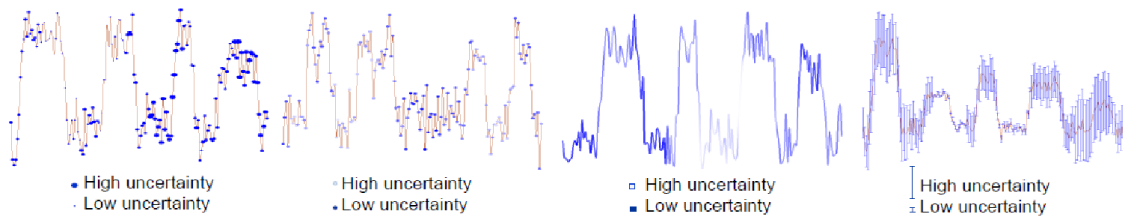


Figure 4.3

1D uncertainty visualizations in the user-study.

There were two considerations in removing some of our suggested techniques from the final study, the first being the inherent merit of the technique and the second being the number of questions it would add to the study. Smooth and striped gradients were the first to be eliminated because they display incorrect uncertainty information across steep slopes. In the 1D case, the uncertainty information in the gradient upon a steep slope aligns itself with the slope, resulting in a thin ribbon. Forcing the ribbon to be always orthogonal is not an elegant solution.

In the 2D case, the surface animation would either hide the data surface, or be itself hidden by the data surface. We eliminated animation of glyphs also because of similar



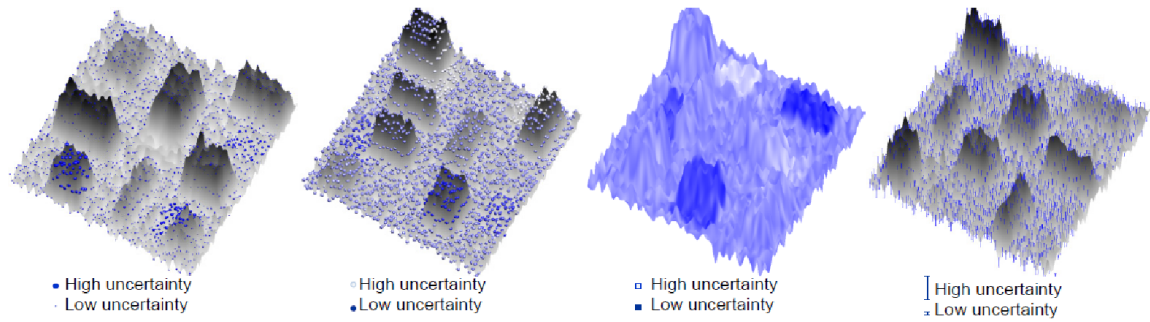


Figure 4.4

2D uncertainty visualizations in the user-study.

occlusion issues in the animation. To be consistent, when we had to eliminate a technique, we eliminated it for both 1D and 2D datasets. We did not want to test a technique that worked for say 1D and did not for the 2D case.

The second concern for us was user fatigue. We did not want to overwhelm the user with too many questions of the same type, which could jeopardise the quality of our results.

The display area had a size of  $800 \times 800$  pixels. The interpretation of visualizations with the errorbars was straightforward [20]. Small bars (smallest being about 8 pixels tall in 1D and about 3 pixels tall in 2D) represented low uncertainty while large bars (largest being about 85 pixels tall in 1D and about 40 pixels tall in 2D) represented high uncertainty. When scaled glyphs were used, large glyphs (largest being about  $10 \times 10$  pixels in both 1D and 2D) represented high uncertainty and small glyphs (smallest being about  $3 \times 3$  pixels in both 1D and 2D) represented low uncertainty. We used flat shading on the glyphs for the 1D dataset, however, we enabled lighting for the 2D dataset to give users a sense of location of the glyphs. For altering the color attribute of glyphs as well as

color-mapping of the surface, we mapped a low uncertainty value to a saturated shade of blue and mapped a high uncertainty value to an unsaturated shade of blue. This is also the scheme suggested by MacEachren [55] and Hengl [34]. The mapping is opposite to the intuitive notion of high and low; however, here we are dealing with negatives, for example “high uncertainty” implies low certainty.

Shades of blue were chosen to convey the uncertainty since blue is a cool color and would cause minimum visual fatigue over the duration of the experiment. Red was used as a preattentive cue to mark regions of interest and highlight user selections [103]. A legend was always provided to aid the user.

#### **4.1.3 Participant Pool**

The participants of our user study were mostly graduates and under-graduates of Mississippi State University. We also had two senior participants who are researchers at the university. We had a total of 36 participants, of which 3 participated in a trial run, 6 participated in a pilot study, and the remaining 27 participated in the main study. Of the 36 participants, 27 were male and 9 were female. None of the participants reported color-blindness while 17 reported 20/20 corrected vision. Most of the participants had some understanding of statistics and used charts and graphs for their day-to-day activities although none of these skills were set as prerequisites to participating in the study. Most users typically spent more than 15 hours weekly using a computer. Each participant was paid \$10 for their time and participation.

#### **4.1.4 User-study tasks**

Not much is known about how domain scientists perceive and use uncertainty. We consulted Dr. Jamie Dyer, a meteorologist, to determine what might be a real world scenario where uncertainty would be a part of his decision making process. He had temperature data in mind and indicated that he would be interested in looking at regions of extreme (high or low) uncertainty. He also wanted to be able to discern features in the data, in the presence of uncertainty. Keeping this in mind, we designed two types of tasks: searching tasks and counting tasks. The searching tasks primarily explored the perception of random uncertainty while the counting tasks explored the perception of systematic uncertainty, along with the cognizance of the underlying data. Dr. Dyer mentioned that he liked to look at the entire data and then focus on a region of interest. The searching and counting tasks were designed to simulate such an exploratory navigation of the data.

##### **4.1.4.1 Search tasks**

The search tasks involved searching for locations of high or low uncertainty from within an area marked in red (Figure 4.5). The entire dataset was always shown to the user. This was done keeping real-life data exploration tasks in mind. Any spot within the marked region could be selected by the user and the corresponding data/uncertainty values would be interpolated whenever necessary. This design decision was made keeping data collection in geosciences in mind, where we take samples at specific locations over a domain and then interpolate if we need values in between. In this document, the searching

for locations of high uncertainty task is labelled Search High Uncertainty (SHU) and the searching for locations of low uncertainty task is labelled Search Low Uncertainty (SLU).

We expected users to perform similarly in both the search tasks, however, our results indicate that there was a significant difference as discussed in our results section. These tasks had more than one correct answer. A user response was considered correct if the chosen location had an uncertainty value within the top 10<sup>th</sup> percentile of the entire range of uncertainty for a task requiring the user to find the location of highest uncertainty, or the bottom 10<sup>th</sup> percentile for a task requiring the user to find the location of lowest uncertainty. We had also tested with the 5<sup>th</sup> percentile but felt that it made the tasks too difficult to perform reasonably. The 10<sup>th</sup> percentile seemed to be a reasonable balance between making the user study impossibly difficult and too easy. Although we did not perform a formal test, we expect the results to remain the same empirically.

Location and proximity of high and low uncertainty areas had an effect on the correctness of the user responses. If a region of interest included both a high and a low uncertainty feature, the range of uncertainty values was much larger than had there been just one uncertainty feature or no uncertainty feature. As a result the number of correct answers changed on a case by case basis. Locating a spot of high or low uncertainty was thereby facilitated, however; a correct answer was not guaranteed by choosing just any location within a feature. There were variations within the feature too, and a user had to make an informed decision as opposed to a blind selection within an approximate region. To control arbitrary effects, we designed the regions of interest to uniformly include high, low, both and none of the uncertainty features.

#### 4.1.4.2 Counting tasks

The counting tasks involved counting either data features or counting uncertainty features within an area marked in red (Figure 4.6). The definition of a data-feature in our study is the presence of any “peak” in the data. Artifacts resulting from the introduction of systematic uncertainty were called uncertainty features in this study, which manifest as regions of extreme glyph-size, glyph-color, errorbar size or surface-color. In this document, the counting of data features task is labelled Count Data Features (CDF) and the counting of uncertainty features task is labelled Count Uncertainty Features (CUF).

One might argue against the merit of having a counting task for data features in an uncertainty visualization experiment. We contend that it is generally important for a user to be always aware of the data and the counting tasks would evaluate the effectiveness of the techniques in retaining a sense of the data.

#### 4.1.5 Interface Design

For the search questions, the interface provided one slider for the 1D data and two sliders for the 2D data to navigate a small red highlight to the chosen answer location (Figure 4.6). Clicking the sliders displayed cross-hair guides to ease the navigation. We were time constrained to implement direct object picking in our interface. We eventually found that users were very comfortable using this interaction metaphor and could reach the desired screen location with at most 2-3 movements of the sliders.

For the counting questions, radio buttons with four static answer choices of 0, 1, 2 and 3 were provided (Figure 4.6). In these questions, the sliders were hidden and four radio

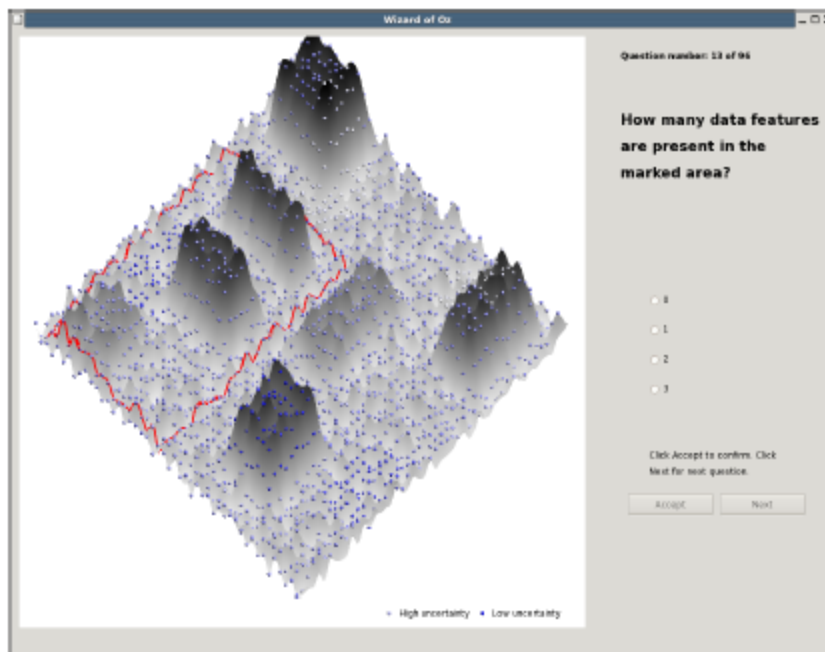


Figure 4.5

User interface for questions requiring a count of features.

button options placed vertically were displayed. Other layouts for the radio buttons were not experimented. Users were expected to make a selection from one of the four radio-button choices. The regions of interest were designed in such a way as to never exceed three data or uncertainty features, and were uniformly designed to include all possibilities. We feel that having a fixed set of choices makes the quality of responses better than having users enter a numeric digit on a prompt. This also ensures a consistent response structure across the methods.

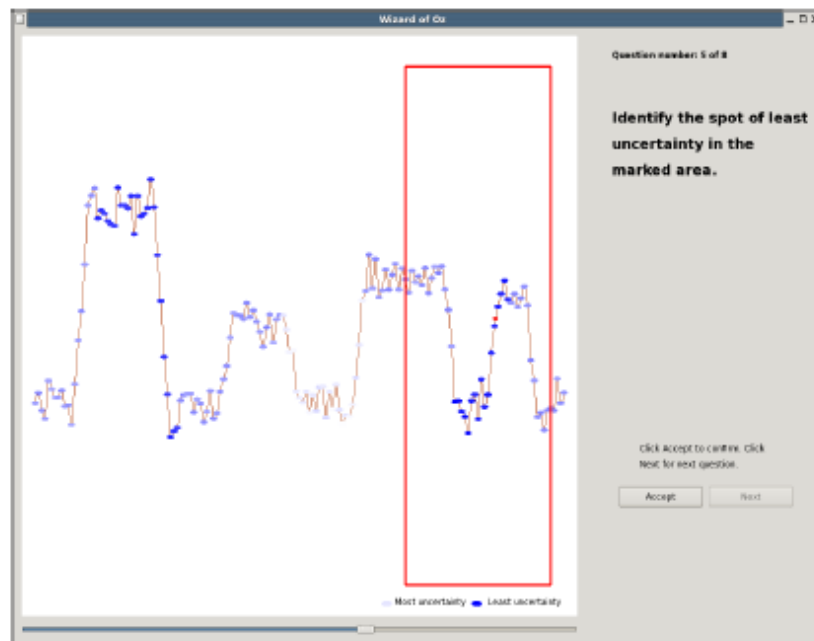


Figure 4.6

User interface for questions requiring a search task.

Users could not skip a question. They clicked an “Accept” button to record their answer and response time, and only then could they click a “Next” button to go to the next question. A break of 5 minutes was given after every 30 questions.

A trial run identified weaknesses with our initial design. Of the three participants in the trial run, two had prior experience designing user studies and their debriefing proved helpful in improving the design. Most notably, the rotation of the 2D surface along the z-axis was fixed to 30 degrees from the original 45 degrees to alleviate some of the artifacts resulting from the view-aligned overlap of errorbars and glyphs. Additionally, the range of the sliders was adjusted to restrict the navigation of the highlight to within the marked region. Users were not allowed to rotate the view or zoom in.

#### **4.1.6 Participant Training**

We typically spent about 15 to 20 minutes to brief the participant about the user study. This involved getting their informed consent and completing a general questionnaire, followed by an explanation of the tasks expected of them. Users were then assigned a computer which ran a training module which was a variation of the software used in the actual study. It familiarized users with the interface and posed 8 questions, one of each type, on the two training datasets. The software highlighted the correct answers to the users to give them feedback on their performance. No person was involved in this process. We felt that users were confident to take on the main user-study after this exercise.



#### 4.1.7 Identifying free parameters

A pilot study was conducted to identify the free parameters in our user-study. These were the size of errorbars and the size of the glyphs. We had 6 participants but we could use data from only 4 of the participants. The quality of the answers from the other 2 was unacceptable because, primarily, they seemed unmotivated and finished too soon. Also, the correctness of their responses was about 50% lower than the others. For this pilot-study, we used three sizes of errorbars and glyphs to compare small, medium, and large representations. The largest of the glyphs was limited to not exceed the size of the grid-cell.

Each participant was asked 144 questions in random order. Although it is difficult to draw meaningful inferences from data from just 4 users, we did find trends that helped us make reasonable assumptions. Users found it easiest to use the smallest errorbars in both 1D and 2D and so we chose to use errorbars of the smallest dimensions in the main study. For glyph colormapping of the 1D data, users found it easiest to use glyphs that had the largest size among the three evaluated sizes. For the 2D data, the three glyph sizes used for glyph color-mapping did not show such a clear trend but had the minimum variance in accuracy of responses for the largest size. So we chose the largest size of the glyphs for use in the main study.

Unlike color-mapping, we did not observe any trends in the responses for the small, medium, and large size ranges used with the glyph-size technique. We attribute it to the inherent nature of the mapping of uncertainty values to size of glyphs making it difficult

to find a separation between the three maximum sizes. So, for this technique, we resorted to using the size we chose for glyph color-mapping.

#### 4.1.8 The Main Study

We had 27 participants in the main study, each answering 96 questions of which 48 questions were on the 1D datasets and the remaining 48 on the 2D datasets. Each set of 48 questions consisted of three sets of 16 questions, each based on data generated using one of the three  $k$  values (0.5, 1.0 and 1.5). The 16 questions asked formed a complete  $4 \times 4$  design of the four visualization techniques explored and the four user-tasks chosen. The response time in milliseconds was recorded for each question. Each user was presented a different shuffled order of questions. The four questions asked were:

- How many data features are present in the marked area?
- How many uncertainty features are present in the marked area?
- Identify the spot of least uncertainty in the marked area.
- Identify the spot of most uncertainty in the marked area.

## 4.2 Analysis

Every correct answer was given a score of 1 and every incorrect answer was given a score of 0. Since there were three questions for a given visualization technique per user task, a participant could achieve a maximum score of 3 for the task, given a visualization technique. Score summaries were created separately for the 1D and 2D datasets.

For each dataset, a  $4 \times 4$  full factorial ANOVA was computed to assess the differences in performances for different questions and different techniques [61]. The summary

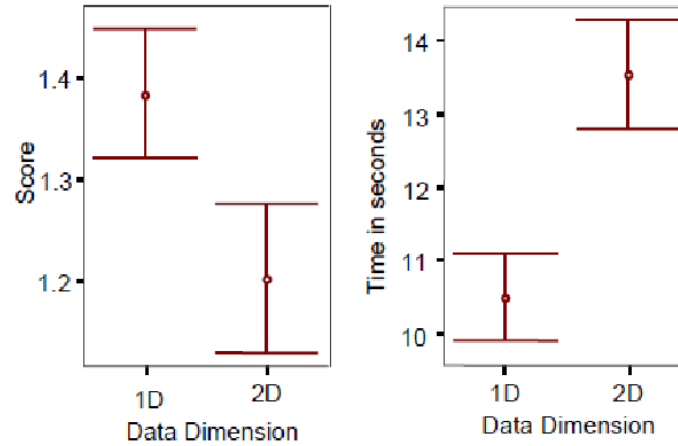


Figure 4.7

Summary of results from the user study.

ANOVA table indicated a significant interaction between type of tasks and techniques used ( $F(9, 416) = 9.968, p < .0001$ ) for the 1D dataset. For the 2D dataset, the summary ANOVA table also indicated a significant interaction between type of tasks and techniques used ( $F(9, 416) = 7.818, p < .0001$ ). This implied that whether there was a significant difference between techniques or not, depended on the type of tasks assigned to the subjects, for both the datasets. Thus, to further explore the results, 8 one-way ANOVAs were computed to capture the Simple Main Effects for each dataset.

The first 4 one-way ANOVAs were intended to see if there were any statistically significant differences between the 4 techniques used with respect to the user tasks. All possible pairwise comparisons were made (6 pairwise comparisons) between the techniques to see if any technique was visibly superior to the rest by creating contrast coefficients to test for the significance of each comparison. The alpha level was set at 0.0083 as compared to widely accepted 0.05 after using Bonferroni's correction ( $\alpha/c$ ; where  $c$  is the number

of comparisons) to control for Type I error. Also, the *t*-test value which does not assume equality of variances was reported for each comparison since the data indicated slight violation of homogeneity of variances. This was computed for the 4 task types, viz. Search Low Uncertainty Locations (SLU), Search High Uncertainty Locations (SHU), Count Uncertainty Features (CUF) and Count Data Features (CDF). The specific findings are listed in Table 4.1. We only report the statistically significant results.

The next 4 ANOVAs were intended to see if there were any statistically significant differences between the 4 user tasks with respect to the visualization techniques used, viz, Glyph-size, Glyph-color, Surface-color and Errorbars. The specific findings are listed in Table 4.2.

We ran our core statistical methods ( $4 \times 4$  full factorial ANOVA) on the obtained scores, from which we identified significantly better performing techniques and tasks. Similar ANOVAs could be based on the analysis of time, however, that would have inordinately complicated the reporting in the time and space available to us which is why we chose graphical techniques to illustrate the time performance (Figure 4.7, Figure 4.8, and Figure 4.9). Also, we were more interested in the accuracy assessment than the time performance in our research goal.

### 4.3 Results and Discussion

We found a consistent trend in the accuracy of responses and the response time for questions on the 1D and 2D datasets (Figure 4.7, Figure 4.8, and Figure 4.9). Since we found a statistically significant interaction between the techniques used and the user-tasks,

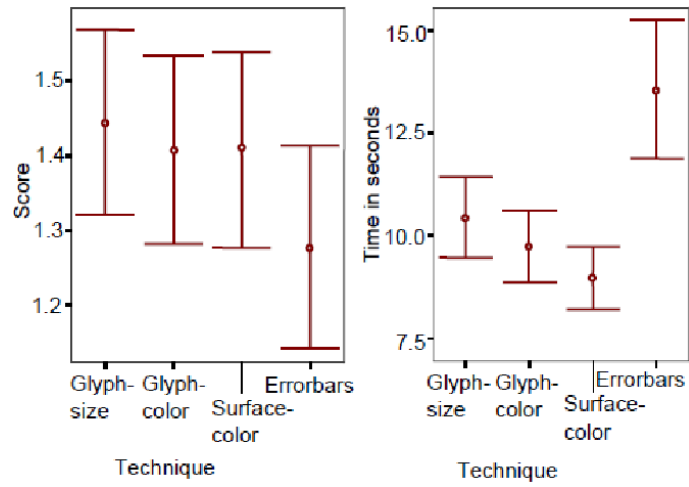


Figure 4.8

User study results for the 1D case.

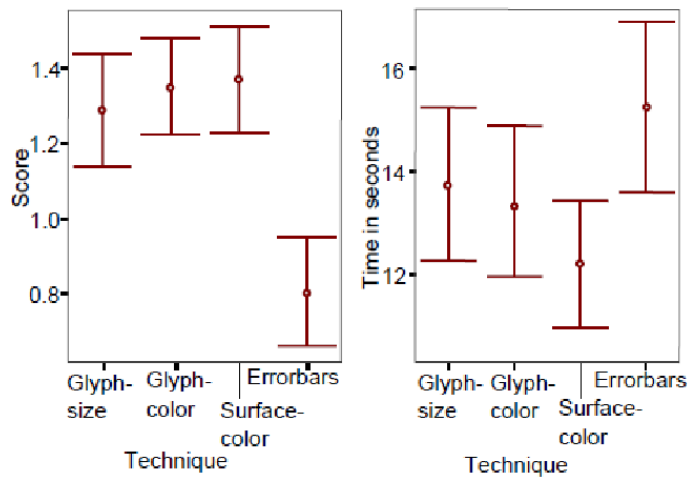


Figure 4.9

User study results for the 2D case.

inferences about a general order of performance of the techniques could not be found, however, we did find several interesting discoveries that we think are useful for uncertainty visualization design. In both cases, Errorbars performed significantly worse than the other techniques studied even though it took substantially more time to answer these questions. One possible explanation of the difference between the errorbars and glyphs is the difference in area between the two representations. We had tested with different errorbar sizes and glyph sizes only in our pilot study from which we determined the most effective size to use in the main study. It would be interesting to test with other shapes of glyph as well.

Visualization researchers agree that the choice of a visualization technique is heavily context dependent. All the visualizations in the study have the same data-density. So, they are fair in the sense that all the techniques were being compared vis-à-vis the same conditions. It is also possible that the data density played a role in perception leading to the poor performance of errorbars. This may be taken as a valuable lesson in designing visualizations for both 1D and 2D cases, which have data-densities comparable to our datasets. However, we do not have sufficient anecdotal evidence that might help us understand this.

The first sets of pairwise comparisons were between the different uncertainty visualization techniques for different user tasks. For the 1D tasks (Table 4.1), users performed significantly better using Glyph-size when the task was to search for locations of least uncertainty. However, both Glyph-color and Surface-color performed better than Glyph-size when the task was to search for locations of high uncertainty. We did not expect to find a significant difference between the two search tasks since both the tasks were designed

Table 4.1

ANOVA results for pairwise comparison on techniques.

Data Dim	Task $t$ -values ( $df$ )	Significantly better	Significantly worse
1D	Search for location of low uncertainty (SLU) $F(3, 104) = 16.176, p < .0001$	Glyph size** Glyph size** Glyph size* Error bars*	Glyph color Surface color Error bars Surface color
	Search for location of high uncertainty (SHU) $F(3, 104) = 13.874, p < .0001$	Glyph color** Glyph color** Surface color* Surface color*	Glyph size Error bars Glyph size Error bars
	Count uncertainty features (CUF)	No significant difference	
	Count data features (CDF)	No significant difference	
	Search for location of low uncertainty (SLU) $F(3, 104) = 6.775, p < .0001$	Surface color* Surface color* Surface color*	Glyph size Glyph color Error bars
2D	Search for location of high uncertainty (SHU) $F(3, 104) = 48.144, p < .0001$	Surface color* Surface color** Glyph size** Glyph color**	Glyph size Error bars Error bars Error bars
	Count uncertainty features (CUF) $F(3, 104) = 7.534, p < .0001$	Surface color* Surface color* Surface color*	Glyph size Glyph color Error bars
	Count data features (CDF) $F(3, 104) = 25.910, p < .0001$	Glyph size* Glyph size** Glyph size* Glyph color** Error bars**	Glyph color Surface color Error bars Surface color Surface color
	* $p < .0083$	Count data features (CDF)	
	** $p < .0001$	Count uncertainty features (CUF) Search for locations of high uncertainty (SHU) Search for locations of low uncertainty (SLU)	

to find extremes in the data. This leads us to believe that human perception of uncertainty ranges using Glyph-size, Glyph-color and Surface-color may not be uniform. Our mapping between visual features and uncertainty was linear. One explanation is that the uncertainty was not translated linearly to the visual features and hence the difference in performance between the two search tasks.

For the 2D tasks (Table 4.1), Surface-coloring performed reasonably well for all questions except counting of uncertainty features. Since shape of the surface was the primary visual cue for data features, we feel that color-mapping of the data surface with uncertainty reduces some of the strength of the shape information. Hence, we see that all other techniques outperform Surface-color for the counting of data features task. On the whole it might sound encouraging to use Surface-coloring to represent uncertainty. While this may work well, one must be aware that it reduces a user's awareness of the actual data. Glyph-size and Glyph-color performed somewhat better than Error-bars although they were worse than Surface-color which bolsters the argument for using one of them as a reasonable trade-off.

The second sets of comparisons were between the tasks for different uncertainty visualization techniques (Table 4.2). For the 1D techniques, Searching for Low Uncertainty (SLU) was clearly the easiest task to perform when the technique was Glyph-size. However, Searching for High Uncertainty (SHU) task was significantly easier when the technique was Glyph-color. Interestingly, for the 2D data, Searching for High Uncertainty (SHU) was consistently easier than Searching for Low Uncertainty (SLU) for all tech-



niques except for Errorbars, where Searching for Low Uncertainty (SLU) was significantly easier.

We found that Counting Data Features (CDF) was generally more accurate except for the 2D case when surface-coloring was used.

On the whole, we found that it took consistently longer for users to respond to questions on the 2D datasets than to questions on the 1D datasets. The accuracy of responses was also higher for the 1D dataset (Figure 4.7). This is not surprising because 2D tasks are generally more difficult than 1D tasks.

Our uncertainty visualization user study brings to light several interesting observations. One such result is that user efficiency in the two search tasks that are opposite of one another (locations of high uncertainty vs. locations of low uncertainty) are significantly different. This is contrary to common understanding and may be attributed to a non-linear perception of the mapping between uncertainty and the visual metaphor. This may drive us to find techniques that compensate for our perceptive biases, or design techniques that are unbiased. One such technique is the Linearized Optimal Color Scale introduced by Levkowitz and Herman [48].

Another aspect that merits discussion is the cognitive associability of high uncertainty with faint colors. The term high uncertainty also associates well with large glyph-sizes. We experimented by reframing our questions and the legend with terms like “high certainty” and “least uncertainty”, and eventually stuck with “high uncertainty” and “low uncertainty” as it seemed to facilitate the cognitive mapping.

We feel that some of our results should force us to think again about the techniques we use on a daily basis. Our study questions the effectiveness of the almost universally used errorbars in data visualization. Although the density of errorbars was higher than for “standard” data graphics, where the x-axis is divided into a relatively small number of categories; there are applications where errorbars are used with the density as evaluated in the user-study. Alternative methods may be suitable in many cases.

#### **4.4 Conclusion**

While we were unable to find clear winners, we were still able to identify scenarios where the effectiveness of certain types of uncertainty visualization results in better visualizations. It is worthwhile to note that errorbars were consistently poor performers. 1D tasks were generally easier than 2D tasks although the 2D tasks took longer to finish. The effectiveness of glyph-sizes and glyph-colors was found to be reasonable. Perhaps the most significant realization is that the effectiveness of an uncertainty visualization technique depended on the task being performed. These results could help scientists in visualizing their data.

The design of the Uncertainty Evaluation Framework and the technique to generate synthetic data with uncertainty are also notable contributions.

Table 4.2

ANOVA results for pairwise comparison on tasks.

Data Dim	Technique $t$ -values ( $df$ )	Significantly better	Significantly worse
1D	Glyph size $F(3, 104) = 13.499, p < .0001$	SLU**	SHU
		SLU**	CUF
		CDF**	SHU
	Glyph color $F(3, 104) = 6.880, p < .0001$	SHU*	SLU
		CDF**	SLU
		CDF*	CUF
	Surface color $F(3, 104) = 17.295, p < .0001$	SHU**	SLU
		CDF**	SLU
CDF*		CUF	
Error bars $F(3, 104) = 11.587, p < .0001$	CDF**	SHU	
	CDF**	CUF	
	CDF*	CUF	
2D	Glyph size $F(3, 104) = 16.721, p < .0001$	SHU**	SLU
		SLU**	CUF
		CDF*	SHU
		CDF**	SLU
		CDF**	CUF
	Glyph color $F(3, 104) = 11.780, p < .0001$	SHU**	SLU
		SLU*	CUF
		CDF**	SLU
	Surface color $F(3, 104) = 36.356, p < .0001$	SHU*	SLU
		SHU**	CDF
		CUF*	CDF
		CUF*	SLU
Error bars $F(3, 104) = 14.067, p < .0001$	SLU**	SHU	
	CUF**	SHU	
	CDF**	SHU	

\* $p < .0083$ \*\* $p < .0001$ 

Count data features (CDF)

Count uncertainty features (CUF)

Search for locations of high uncertainty (SHU)

Search for locations of low uncertainty (SLU)

## CHAPTER 5

### SOFTWARE PROTOTYPE FOR WEATHER ENSEMBLES

In this chapter, we discuss the design, development, and subsequent use of a software prototype named Noodles that has been developed collaboratively between meteorologists and visualization researchers. We discuss the iterative software development paradigm employed that has led to two iterations of the tool. Noodles, and subsequently Noodles 2 has been designed to employ uncertainty visualization techniques to meteorological ensemble simulation output to aid experts in their analysis. Meteorologists Dr. Andrew Mercer and Dr. Jamie Dyer provided expert advice and feedback during the development and subsequent analysis of multiple ensemble datasets using the tool. The technical features presented here are an improvement over what is used by operational meteorologists and could help in improving weather forecasts.

#### 5.1 Ensemble Weather Forecasting

Predicting the weather is inexact and computationally expensive. The most common method for weather prediction is through dynamic modeling, in which simulations recreate or predict the conditions of the atmosphere for a period of time. In an effort to reduce errors from individual model simulations, multiple runs of the dynamic model with different initial conditions are used to create an ensemble. Scientists use the average ensemble

output as a forecast and utilize spaghetti plots to analyze the spread of the ensemble members and describe their uncertainty. Individual ensemble members are usually initialized with slightly perturbed initial conditions or with different parametrizations, or sometimes both.

Generating a forecast ensemble is computationally intensive. The advent of large-scale supercomputers has helped the ensemble generation process as the combined power of multiple processors can be used to address some of the computational issues. Additionally, supercomputing has allowed for the manipulation of larger datasets, leading to increased ensemble sizes. In most cases, entire datasets cannot be loaded into memory. As a result, common analysis tools often lose their interactivity. Augmenting these tools to produce uncertainty visualization slows them down further. While analysis tools are often used in research settings, operational meteorologists almost always use spaghetti plots [23], which are generated by combining single mid-tropospheric contours from various ensemble members for typical 500 mb pressure surface height values (i.e. 5400 m, 5700 m). Such an approach restricts the amount of data available to a forecaster while making a prediction. Effective manipulation and uncertainty visualization techniques could help to emphasize the uniqueness of a situation by conveying the contribution of uncertainty, allowing the forecaster to make a more informed prediction.

It is important to note the difference between forecast accuracy and precision. Numerous forecast accuracy studies have been conducted, e.g. [46, 59], and no unifying forecast verification statistic has been devised. Instead, ensemble forecasts were introduced as a means of removing forecast errors resulting from initial conditions or model parametriza-

tions. Biases in the model forecast will remain, so an ensemble forecast with little member spread may still have poor forecast accuracy. The scope of this study is analysis of ensemble member uncertainty, not forecast accuracy. We did not undertake a comparison of observed and predicted weather conditions.

We used the Weather Research and Forecasting Model (WRF) [67] to run our meteorological simulation ensembles. WRF simulations require three steps: pre-processing, running the model with the chosen parametrizations, and then post-processing the output. For the purpose of this study, we used different parametrizations of cumulus and microphysics schemes to create our ensembles. Specific simulations are described in detail as case studies in the following sections.

## **5.2 Software Engineering Effort**

We conducted an initial meeting between two meteorologists and two visualization designers to identify scenarios where uncertainty visualization would be useful. The visualization designers were made aware of some of the common approaches that meteorologists use in data analysis. We agreed that understanding uncertainty in weather ensembles was challenging and that there were many aspects in 2D visualization that could benefit from a better representation. We decided that the initial focus would be on 2D visualization since forecast meteorologists typically rely on 2D slices of the atmosphere when making weather predictions. Additionally, the 2D framework would provide a baseline for a more complex 3D analysis in the future.

With a focus on these goals, we initially developed a prototype of a tool named *Noodles* [82]. Feedback from the meteorologists was used to redesign the tool to incorporate many of the features deemed essential by them to create the current iteration of the tool named *Noodles 2*. We followed a spiral development model where an initial prototype is refined through user evaluation and feedback.

### 5.3 Initial Prototype: Noodles

*Noodles* was implemented in the open source version of Qt 4.5. The user interface for the tool has four sections (Figure 5.1). The first section is the display widget which uses OpenGL rendering for geographic visualization (boxed in red and labeled ‘1’ in Figure 5.1). We used the shapelib library [104] to read in shapefiles of continents and the countries of North America which were then rendered with thick outlines to indicate international borders. Shapefiles of individual states within the countries are also displayed with thinner outlines for better geographic context. An Equirectangular projection (which is a simple mapping of the latitude and longitude to a Cartesian grid) was used for the visualization. The ensemble data were on a Lambert Conformal Projection and were rendered in the Equirectangular Projection.

The second section (boxed in blue and labeled ‘2’ in Figure 5.1) contains the data control widgets placed under the primary display area. The variable selection combo-box, height selection combo-box, the timeline slider, and the isovalue selection box allow the user control over the data being displayed. An ‘animate’ checkbox allows a time-

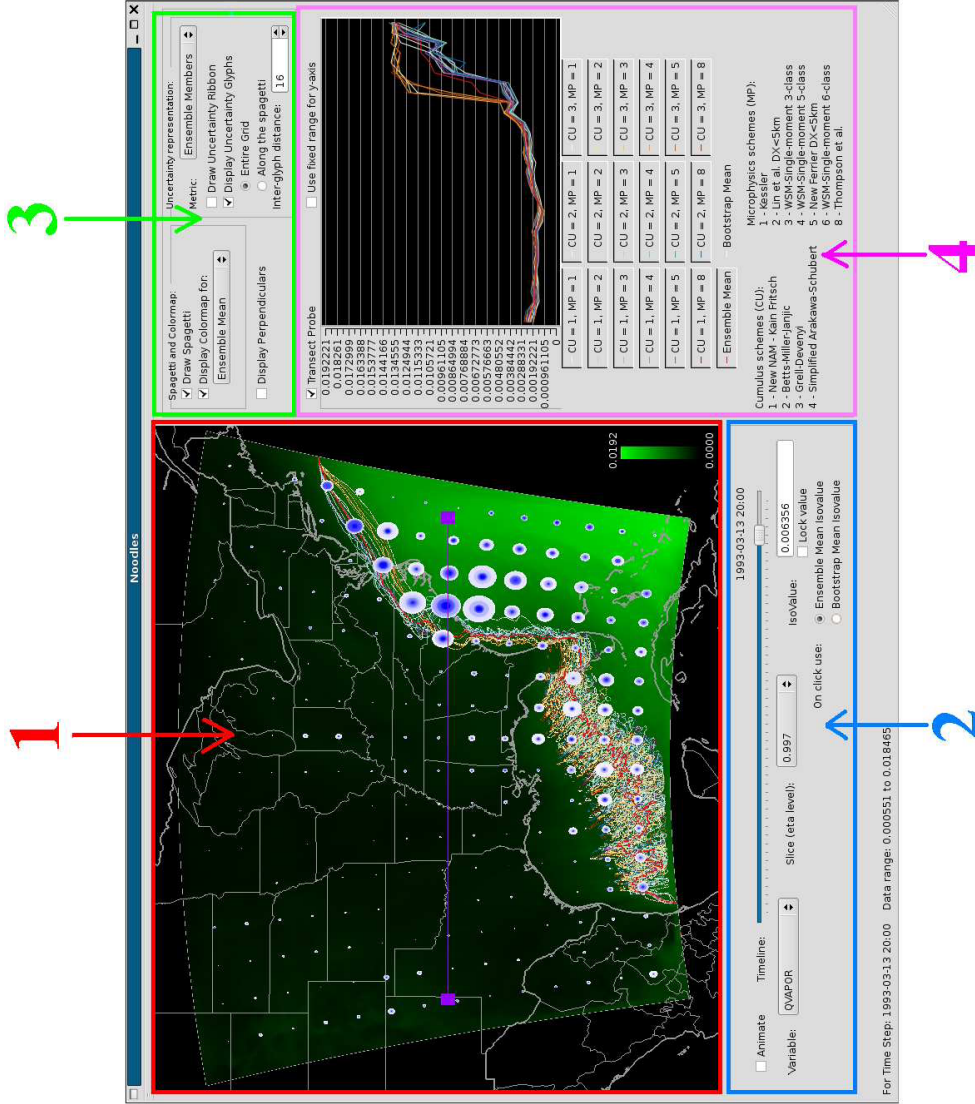


Figure 5.1

Marked up user interface of Noodles.



loop animation of the data. A status bar provides additional information about the current dataset.

The third section (boxed in green and labeled '3' in Figure 5.1) is to the right of the first display area. This section has widgets that enable or disable the various uncertainty visualizations built into the tool.

The transect plot comprises the fourth display area (boxed in pink and labeled '4' in Figure 5.1). The user may optionally enable a data-transect that displays the distribution of the values of the current variable along a straight transect line through the data. The tool has been designed to ensure coordinated views and updates. Changes to one section updates dependant members in other sections to maintain visualization context.

### **5.3.1 Uncertainty Metrics**

One of the most important stages of the uncertainty visualization pipeline [74] is the meaningful quantification of uncertainty. We calculated the ensemble mean, standard deviation (SD), inter-quartile range (IQR), and the 95% confidence intervals (CI) for the ensemble at each grid point. However, these statistics require data that follow a normal distribution to be interpreted properly. It is evident from a sample grid point of perturbation pressure (Figure 5.2) that the ensemble members do not follow a normal distribution, but instead have an unknown distribution that is likely non-normal. To remove the normality constraint, we employed an ensemble mean bootstrap.

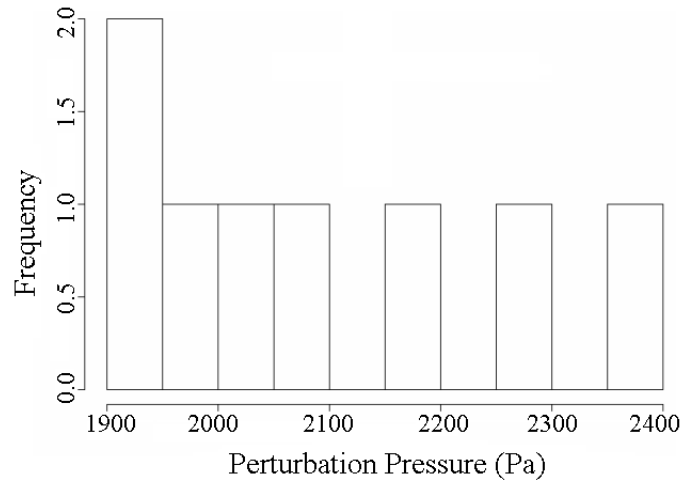


Figure 5.2

Histogram of ensemble means for 8 values.

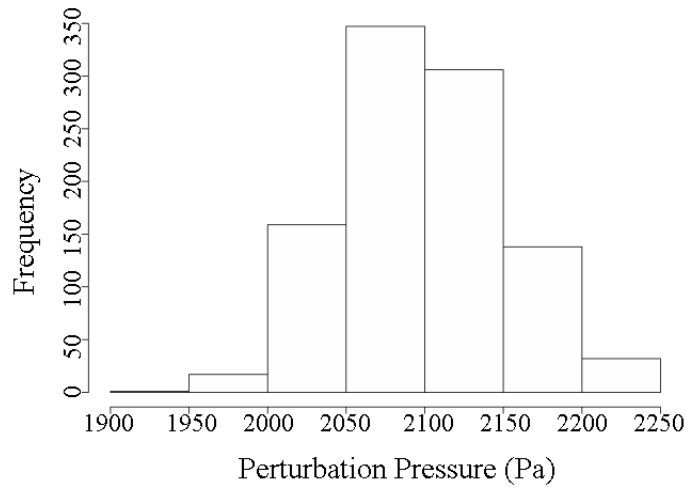


Figure 5.3

Histogram of the bootstrap bootstrap means from Figure 5.2.

### 5.3.1.1 Bootstrapping

Bootstrapping [27] is a statistical method that falls in the general class of resampling statistics typically used to generate an approximating distribution with no assumptions on the type of the source distribution. The ensemble members constitute an independent and identical distribution since they were formulated using the same numerical model with independent initial conditions, thereby satisfying the prerequisites for applying the bootstrap procedure. We formulated the bootstrap means using 1000 random samples of the ensemble members which, by the central limit theorem, are a normally distributed approximation of the true distribution of the ensemble mean. Both the mean and median of these bootstrap replicates (Figure 5.3) are a better representation of the center of the unknown distribution than the sample mean from the ensemble members (Figure 5.2) due to the skewed distribution of the original members. Often this is a small difference (e.g. for the example in Figure 5.2 and Figure 5.3, the sample mean was 2098.06 Pa while the bootstrap mean was 2099.66 Pa), but the issue of distribution is especially important with certain non-normally distributed meteorological parameters (e.g., water vapor mixing ratio, precipitation, etc.) As a result, it is essential that both methods are included for completeness. Likewise, the width of the bootstrap 95% confidence interval and the interquartile range of the bootstrap means yielded uncertainty measures not constrained by a requirement of normality and are easily derived from the bootstrap mean distribution.

### 5.3.2 Uncertainty Visualization

Uncertainty glyphs, graduated uncertainty glyphs, uncertainty ribbon, graduated uncertainty ribbon, and spaghetti plots were available to users of the tool for uncertainty visualization. The visualization itself was done on a Sun Ultra 27 workstation that ran Suse Linux Enterprise Server 10. It had four 3.2 GHz dual-core Intel Xeon processors with 6 gigabytes of memory and an NVIDIA Quadro FX3800 graphics card.

#### 5.3.2.1 Uncertainty Glyphs and Graduated Uncertainty Glyphs

We explored the applicability of circular glyphs scaled in size as one of the uncertainty visualization techniques in the tool. We [81] identified that glyphs altered by size were an effective method to depict uncertainty in 2D datasets. Uncertainty measures of standard-deviation, IQR, and the width of the 95% CI were mapped to the radii of circular glyphs to convey the uncertainty. The user could choose to display these glyphs over the entire grid (Figure 5.4 and Figure 5.5), or along the contour of a value of the ensemble mean or the bootstrap mean (Figure 5.6 and Figure 5.7). The maximum possible size of the glyphs was based on the coarseness of the glyph-spacing such that overlap between adjacent glyphs is minimized.

We also explored the applicability of graduated uncertainty glyphs in the tool. Recall that the basic idea is to use the difference values of ensemble members to the mean to construct concentric circular glyphs (section 3.2.7), starting with the largest difference value,  $D_n$  and rendering successively smaller glyphs  $D_{n-1}, D_{n-2}, D_{n-3}, \dots, D_1$ . This ensures that successively rendered glyphs are either smaller-than or equal-to the size of

the previously rendered glyphs (Figure 3.13). Using a range of saturation values (0 ... 1) over the  $n$  ensemble members, the largest glyph (derived from the largest difference value) ends up having the least-saturated blue, and each successive glyph gets a more saturated shade of blue.

The above process was used to generate glyphs for the whole grid, or along the mean contour. A glyph having a dense and saturated core with a faint periphery indicates that ensemble members mostly agree and have a few outliers. On the other hand, a mostly blue glyph indicates that there are large differences among the members. The overall size of the glyphs across locations indicates the variability of each location with respect to other locations on the grid (Figure 3.14).

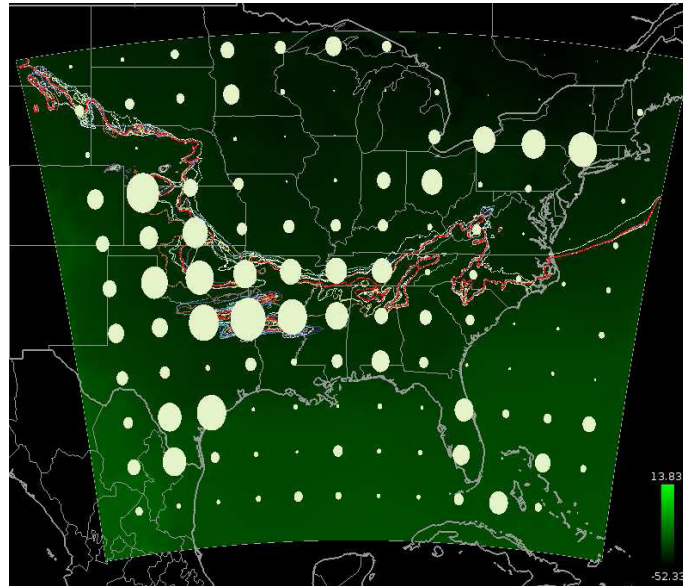


Figure 5.4

Glyphs illustrating 95% CI of the values.

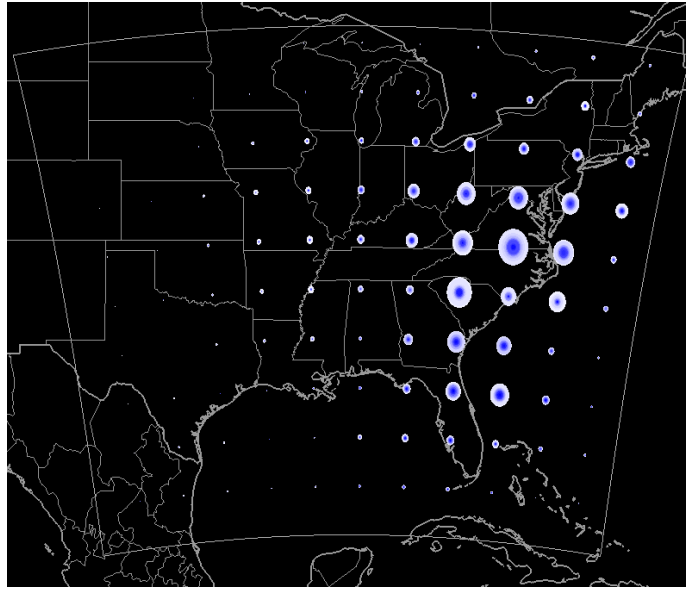


Figure 5.5

Graduated glyphs aligned over the whole grid.

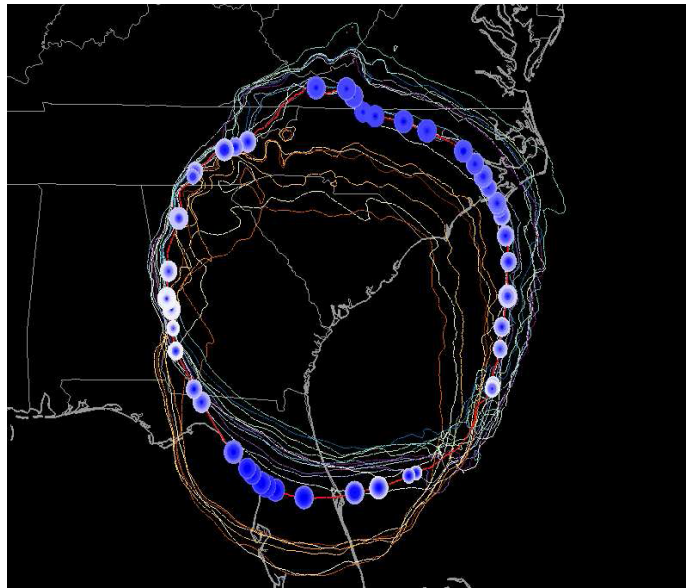


Figure 5.6

Graduated glyphs along the mean perturbation pressure.

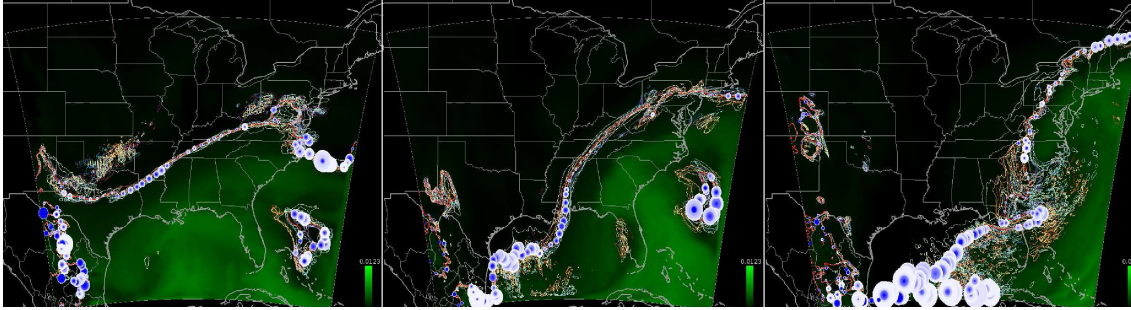


Figure 5.7

Three time-steps illustrating spaghetti plots, graduated glyphs, and colormap.

### 5.3.2.2 Uncertainty Ribbon and Graduated Uncertainty Ribbon

We generated an uncertainty ribbon to quantify the uncertainty along a contour of a value from the ensemble mean or the bootstrap mean (Figure 5.8 and Figure 5.9). The width of the ribbon represented the uncertainty along the contour. Recall that an uncertainty ribbon is generated by connecting external tangents of hypothetical circles along a contour (section 3.2.8). In Noodles, the calculation and rendering of the contours was done using the popular ‘conrec’ routine [11] which outputs contour fragments for a given iso-value. These fragments were converted into connected segments by routines derived from a COCOA implementation of conrec [30]. Uncertainty values could be easily extracted along the generated contour by using the contour fragment location to lookup uncertainty values from the underlying 2D grid.

The tool also supports the rendering of a graduated uncertainty ribbon by overlaying multiple uncertainty ribbons for the number of ensemble members (Refer section 3.2.9).

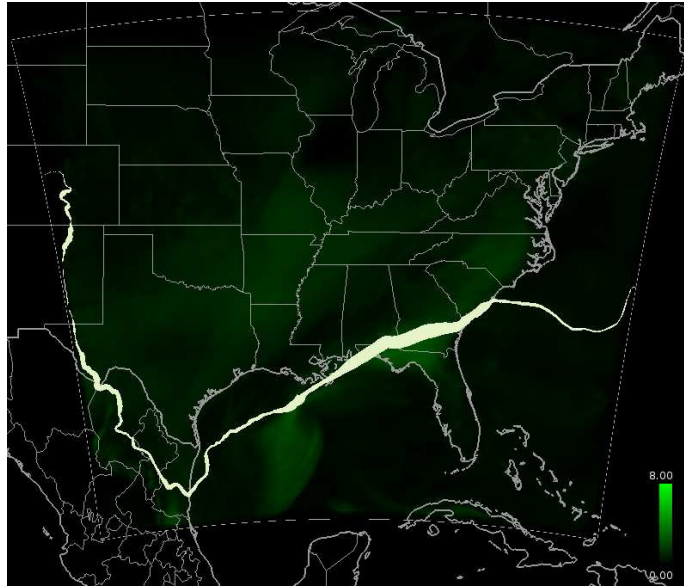


Figure 5.8

Illustration of uncertainty ribbon mapping the IQR.

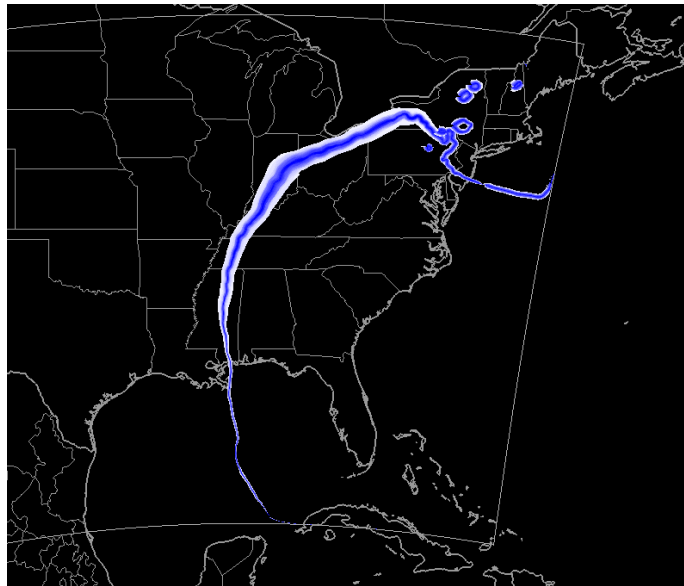


Figure 5.9

Illustration of a graduated uncertainty ribbon.



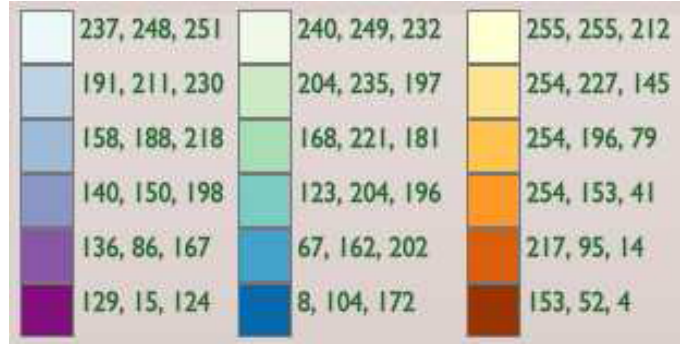


Figure 5.10

The  $3 \times 6$  color scheme from ColorBrewer [33].

### 5.3.2.3 Spaghetti Plots

Spaghetti plots are not a new technique, however we still implemented them in Noodles for the sake of completeness. They also served as a reference visualization that meteorologists were accustomed to. However, the color encoding for the spaghetti plots employed here is a new application in meteorology.

ColorBrewer [33] is an online tool used to generate colormaps for cartographic visualization. Categorical colors for individual contours in the spaghetti plot (and corresponding members in the data-transect discussed later) could not be obtained from ColorBrewer because it has an upper limit of 12 classes for qualitative categorical data while the ensemble had 18 members in the particular case study. Sufficient visual separation between the members cannot be guaranteed for such a large number of categories and so an alternative was designed. Since there were three cumulus parameterizations and six microphysics schemes in the ensemble in the case study, we chose three sequential color schemes, BuPu, GnBu, and YlOrBr from ColorBrewer to map to the cumulus schemes, and used a six

level sequentially graduated hue within each color scheme to map to the six microphysics schemes (Figure 5.10). The color legend also had a  $3 \times 6$  layout. We plotted the ensemble mean with a thick red line to take advantage of its preattentive properties [103]. We plotted the bootstrap mean with a thick white line. It was our experience that using black for the background color allowed the most visual contrast amongst the plots.

### 5.3.3 Additional Features

Various additional features in the tool are as follows.

#### 5.3.3.1 Data-Transect Plot

Meteorologists often look at the distribution of values along transects. To facilitate such use, we designed a data transect plot using the chart widget library Qwt (Figure 5.1). Upon enabling the data-transect checkbox, a purple horizontal line with handles at both ends appears on the main display. The plot on the right side of the window displays the values of all ensemble members along the line. The user can click any of the handles and move the transect, which automatically updates the plot. The scale on the plot can be locked to a predetermined range, or may be allowed to vary locally depending on the dynamic range of the data on the plot. This is often useful in amplifying the local variation when the absolute range is too large (Figure 5.11).

The tool is supplied with an interactive member selection legend (Figure 5.11). Enabling or disabling a legend item enables or disables the corresponding data-transect plot as well as the corresponding contour in the spaghetti plot. This can be used to reduce

clutter in both plots, as well as reduce bias in the original mean by removing the poorly performing members.

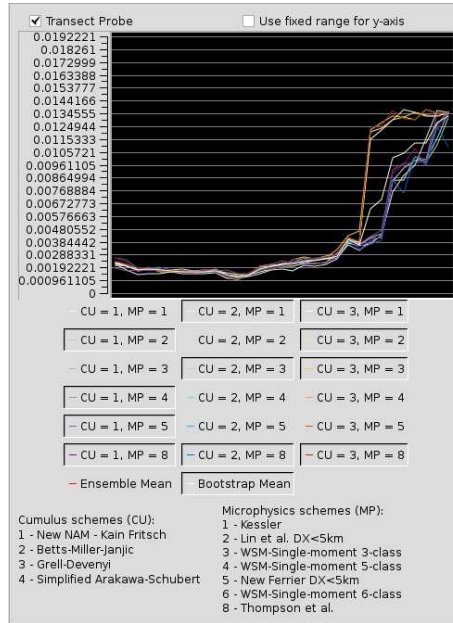


Figure 5.11

The data-transect plot and interactive legend.

### 5.3.3.2 Colormap

Application of a colormap for the chosen uncertainty metric and data values to the geographic extent was particularly useful in exploring the overall distribution of uncertainty in the dataset. It has been our experience that many geo-visualization packages use the rainbow colormap by default in spite of its known perceptual deficiencies [10]. We have made a conscious effort to avoid the use of the rainbow colormap with the objective of

increasing the awareness of its shortcomings in the geo-scientific community. We choose a green saturation colormap as a default for this tool (Figure 5.7).

### 5.3.3.3 Interactive Visual Queries

The spaghetti plots can be generated interactively. Clicking any location on the primary display within the geographic extents of the simulation triggers a visual query that returns either the ensemble mean or the bootstrap mean at that location. The returned value is used to generate a new spaghetti plot. This is extremely useful in situations where the user is curious to see the spaghetti distribution at a specific location. The background colormap can be used to visually identify these regions or observe ‘interesting’ features that may be clicked upon to generate corresponding spaghetti distributions. Most tools that generate spaghetti plots for meteorologists lack features to visually query the data.

Much of the usability of the tool depends on it being reasonably interactive. The ability to change variables, timesteps, or vertical slices independent of other options enables a quick view of different data across the same set of metrics or along the same data-transect. Typical load times for fresh data depended on whether the data was in recent cache, in which case load times were 3 to 4 seconds. These lengthened to about a minute when loading from disk. Once loaded to memory the tool is fully interactive. Apart from selective fetching of data-slices, no other out-of-core memory/data management techniques were employed. The design of Noodles allows such features to be easily added which could significantly enhance the tool. Interactivity of the data-transect plot and its legend

further extend the flexibility of the tool. Typical update time for all elements of the rendering did not produce any perceivable lag.

### **5.3.4 Expert Evaluation of the tool**

We describe the opinions of two meteorologists who used to tool for ensemble analysis. Specific inferences pertaining to the simulation are discussed in the next section.

#### **5.3.4.1 Effectiveness of Uncertainty Visualization**

The uncertainty glyphs and uncertainty ribbon provide an alternative visualization of uncertainty. The glyphs present uncertainty at a point but do not obscure underlying data as much as a colormap of the uncertainty would. An uncertainty ribbon displays the variation of the uncertainty metric along a contour. Both of these techniques compliment the use of spaghetti plots. The spaghetti plots are drawn by generating iso-valued contours which imply geographic separation. When individual ensemble members on a spaghetti plot do not coincide, inferring spatial uncertainty is incorrect. Also, attempting to quantify such uncertainty by using techniques such as finding the shortest distance between contours is not physically meaningful. A simple example of this would be trying to compare the weather at Philadelphia and New York using the distance between them! Additionally, separation within members of a spaghetti plot is highly dependent on the gradient of the 2D field [107] (Figure 5.19 and Figure 5.20). While spaghetti members may be close to one another in regions of steep gradient (Figure 5.13) and be widely separated in regions of low gradient (Figure 5.19), the actual uncertainty might be higher in the region

of the steep gradient where the spaghetti members appear clumped together because even nearby locations may exhibit large value differences (the corresponding data-transects, Figure 5.14 and Figure 5.20). Thus, spaghetti plots need to be interpreted as physically separate contours and not as a measure of spatial uncertainty. On the other hand, uncertainty glyphs and uncertainty ribbons represent the uncertainty at specific grid locations which are geographically co-located and are in essence different from spaghetti plots.

The meteorologists asserted that the uncertainty ribbon and the glyph based visualization were both useful to get an overview of the uncertainties in the data. They still found that the uncertainty ribbon was more effective than the uncertainty glyphs. Meteorologists traditionally work with iso-contours and find a continuous ribbon-like visualization more easily interpretable. They found the glyphs with graduated colors to be difficult to comprehend. Although the glyphs were a good technique to indicate the distribution of the members, they tended to exaggerate the outliers. Small glyphs that have large variances may appear smaller than glyphs with relatively smaller variances but with large outliers. It is possible that the glyphs were overloaded with information, partly because it is easier to interpret other uncertainty metrics such as the 95% confidence interval, which have a single value mapped to the size of the glyphs.

#### **5.3.4.2 Effectiveness of Interactivity**

The meteorologists found the interactivity of the tool extremely useful. Being able to switch variables while keeping the data transect and the visualization techniques consistent was very useful. The coordinated data-transect plots allowed them to easily verify data

distributions across the cold or warm fronts. Being able to easily move in the timeline made it convenient to observe the evolution of the storm and the associated uncertainty (Figure 5.7).

In particular, they found the interactive legend and the color-scheme to be very useful for distinguishing between ensemble members (Figure 5.11). The ability to select which ensemble members to visualize helps to reduce clutter and provides control in the analysis. The color scheme significantly helped to visually differentiate between members. The zoom and pan options were also beneficial, especially because of the lack of such interactive features in operational use.

#### **5.3.4.3 User feedback**

The meteorologists spent a lot of time with the data-transect plot and the spaghetti plot. Part of the reason could be because of their previous training with spaghetti plots. The time-line and the variable swapping widget were also used often. It must be noted that they focused primarily on looking at the deviation of a particular cumulus scheme (Grell-Devenyi scheme, section 5.3.5.2). In general, the experts were able to make sense of the physical processes governing the uncertainty in the ensemble. Previous research (e.g., [65]) has shown that, for WRF simulations, forecast degradation is not significant until at least 72 hours after initialization. As such, forecast errors at 48 hours, while present, were not overwhelming to the meteorologists. They expressed the desire to be able to load multiple datasets and use the tool to look at other simulation runs.

### 5.3.5 Case Study: 1993 “Superstorm”

The 1993 “Superstorm,” also referred to as the “Storm of the Century,” was a unique weather event that occurred between 12 March and 14 March [44] and affected a region stretching from Central America to Canada (Figure 5.12). The storm formed from a weak cut-off low pressure system over the Gulf of Mexico and quickly intensified, undergoing bomb cyclogenesis (a central pressure decrease of 24 mb in 24 hours) as it tracked north-eastward along the East Coast [44, 102]. It caused record low pressure, low temperatures, winds, and snowfall in the Eastern United States and resulted in over 250 fatalities. The storm was also a significant milestone for numerical weather prediction in the United States. For the first time, a number of computer models successfully predicted the severe threat days ahead of its occurrence [102] (note the Weather Research and Forecasting [67] model is not part of this list as it was developed within the last decade). The size, intensity, and time of occurrence of the storm make it unique and important, and contribute to it being a well-studied severe weather event, e.g., [17, 39, 85].

This section describes the design, visualization, and subsequent analysis of the severe weather event by two meteorologists.

#### 5.3.5.1 Formulation of the ensemble and running WRF

The National Weather Service Science Operations Officers / Science & Training Resource Center (STRC) Weather Research and Forecasting (WRF) [67] Environmental Modeling System (EMS) Version 2.0 was used for the simulation. The Advanced Research WRF numerical solver core was used in this simulation experiment since it is intended to





Figure 5.12

Satellite image showing the extent of the 1993 “Superstorm”.

be a research core allowing one to conduct a more thorough simulation. WRF simulations are conducted in three steps involving pre-processing, running of the model, and then post-processing the output. Additionally, for ensemble simulations, one must rerun the three-stage process for the different scenarios. The appendices describe each of these stages in more detail.

Ensemble numerical simulations are conducted through one of two possible methods: a perturbation of initial conditions, or by altering model parametrization schemes. A parametrization ensemble was selected since a fundamental aspect of this work is to identify the uncertainty associated with model physics schemes. Core model physics parametrizations [89] of cumulus and microphysics schemes were identified by a meteorologist to be of research interest.

The cumulus scheme determines the convective behavior of clouds in the atmosphere. A cumulus scheme is recommended if the grid coarseness is greater than 10 kilometers.

The cumulus parametrization schemes tested were the:

1. New NAM Kain-Fritsch scheme
2. Betts-Miller-Janjic scheme
3. Grell-Devenyi scheme

Microphysics schemes determine how the models resolve atmospheric microphysical processes by the use of moments to determine flow of water and ice. The microphysics parameterization schemes tested were the:

1. Kessler scheme
2. Lin et al. scheme
3. WSM Single-Moment scheme
4. WSM Single-Moment 5-class scheme
5. New Ferrier scheme
6. Thompson et al. scheme

The Lin et al. scheme is recommended for fine resolution runs having a grid spacing of less than 2 kilometers. The meteorologists were mildly curious to see how a scheme recommended for fine grids would perform on a relatively coarser grid and so the Lin et al. scheme was included in the study.

Planetary boundary layer physics schemes resolve the near surface atmospheric anomalies which can be very different and more turbulent than the rest of the atmosphere. The

Mellor-Yamada-Janjic planetary boundary layer physics scheme was tested for the simulation, except for the combination of Betts-Miller-Janjic cumulus scheme and Kessler microphysics scheme, where the required planetary boundary layer physics was the Yonsei University scheme. Combinations of the 3 cumulus schemes and 6 microphysics schemes resulted in 18 ensemble members.

The simulation grid had dimensions of 269 (west-east)  $\times$  240 (south-north)  $\times$  30 (vertical). The latitudinal extent was from 21.19 N to 42.26 N while the longitudinal extent was from 103.14 W to 67.36 W, resulting in an average grid-spacing of 12 kilometers. In this case, the North American Regional Reanalysis (NARR) dataset [66] was the source of the input weather data, interpolated and re-gridded to the domain internally by WRF.

The simulation was configured to run on two nodes of a cluster, with each node having two dual-core AMD Opteron 2218 (2.6 GHz processors and 8 gigabytes of memory). The cluster used gigabit ethernet for internode communication. The typical runtime for each simulation was about 10 hours. The model run produced hourly predictions up to 48 hours. The ensemble simulation generated a 133 gigabyte dataset. The variables of interest, water-vapor mixing ratio, perturbation potential temperature, and perturbation pressure, were extracted during post-processing to create a 22 gigabyte binary dataset. No interpolation was necessary because all three variables used an identical Arakawa grid and could be used directly for visualization in Noodles.

### 5.3.5.2 Expert Evaluation: Model Parametrization Inferences

Two meteorologists used the tool to perform analysis of the ensemble. Recall that the scope of the study is ensemble precision, not ensemble accuracy. Verification of the ensemble forecast against ground truth is outside the scope of this work.

One of the parametrizations that the meteorologists were curious about during the design phase was testing the performance of the Lin et al. microphysics scheme, recommended by WRF for fine grids of less than 2 kilometer spacing, on coarser grids. No divergence of the performance of the scheme was found with respect to other ensemble members.

However, the Grell-Devenyi cumulus scheme produced output that was not consistent with the other ensemble members for spatial regions that tracked the cold front (Figure 5.13 – Figure 5.18). A cold front is a boundary between a cold and warm air mass typically associated with a synoptic cyclone. This clashing of air masses was expected to correspond to an area of high uncertainty, which was observed using the tool. The coordinated geographic and data-transect plot revealed a sharp difference in the perturbation potential temperature profile in the Grell-Devenyi scheme (Figure 5.13 and Figure 5.14). Correspondingly, the position of the warm sector was displaced farther west in the Grell-Devenyi ensemble members. A similar bias was noted in the perturbation pressure field, as the location of the low pressure center as presented by the Grell-Devenyi scheme was much farther south than for the other two cumulus parametrizations (Figure 5.15 and Figure 5.16). Additionally, the water-vapor mixing ratio appeared more sensitive to the cumulus scheme than the microphysics scheme in proximity to the cold front, which was

not surprising due to the convective thunderstorm activity associated with the front (Figure 5.17 and Figure 5.18).

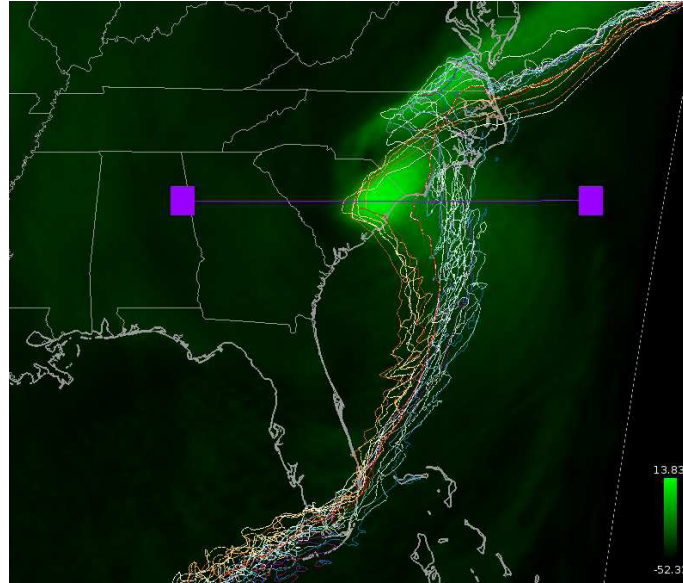


Figure 5.13

Bias in the resolution of the cold-front for perturbation potential temperature.

The meteorologists concluded that the Grell-Devenyi scheme demonstrated previously unknown bias in the perturbation potential temperature and perturbation pressure fields at lower levels of the atmosphere. The meteorologists also confirmed that the scheme was in general agreement with the other ensemble members at higher levels (Figure 5.19 and Figure 5.20). Further investigation into the cause of these biases should be conducted to improve model forecasting.

This uncertainty visualization tool can provide important feedback about a model forecast to operational meteorologists. Likewise, the ability to interactively remove parametriza-

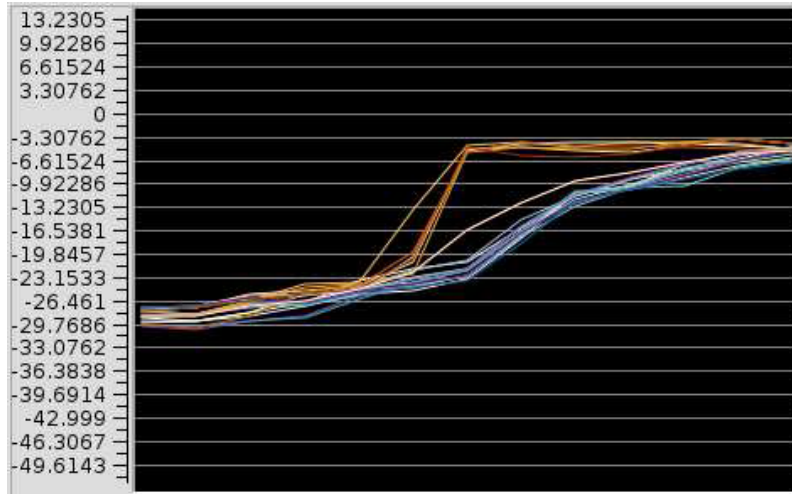


Figure 5.14

Sharp change in data-transect through Figure 5.13.

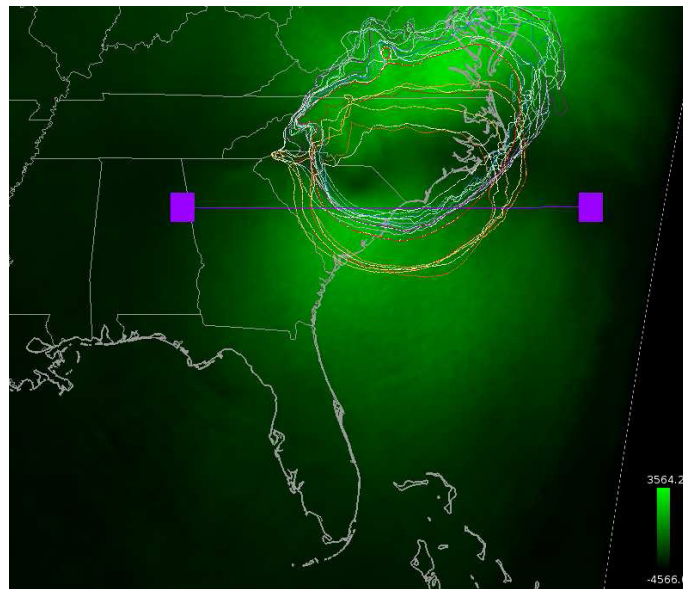


Figure 5.15

Bias in the resolution of the storm center for the pressure perturbation.

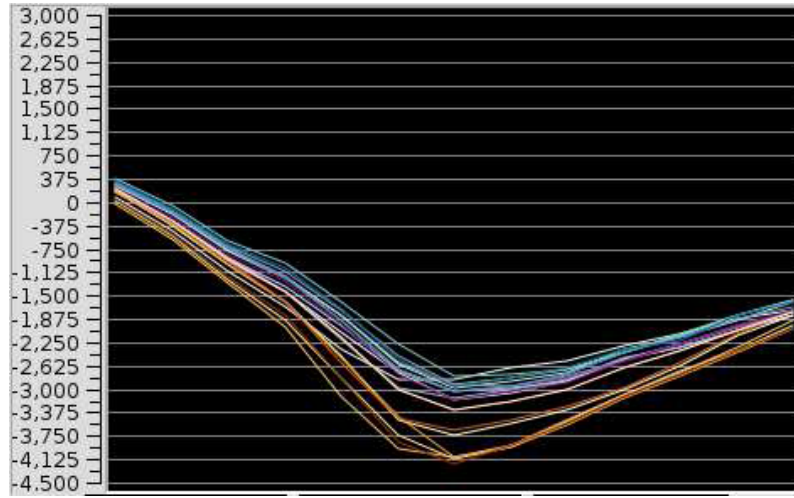


Figure 5.16

A data transect through Figure 5.15.

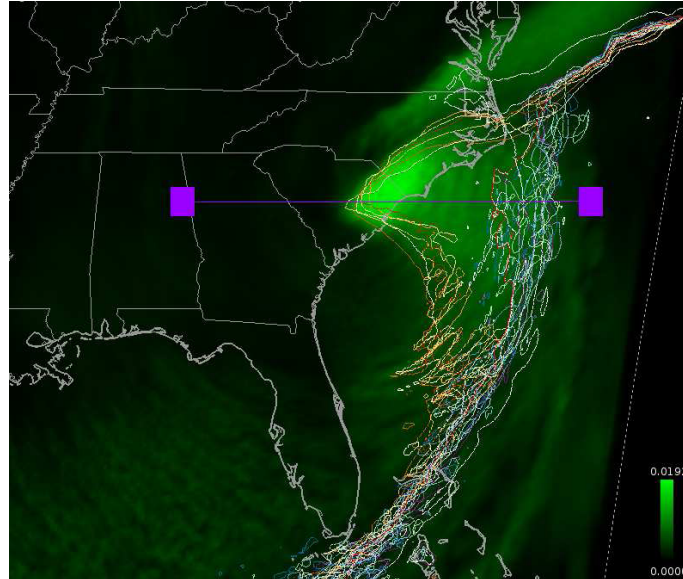


Figure 5.17

Bias in water-vapor mixing ratio corresponding to Figure 5.13 and Figure 5.15.

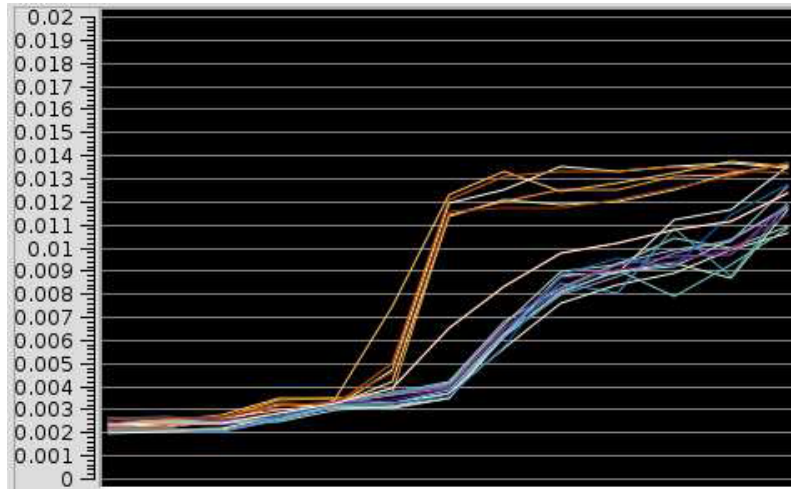


Figure 5.18

Data transect through Figure 5.15.

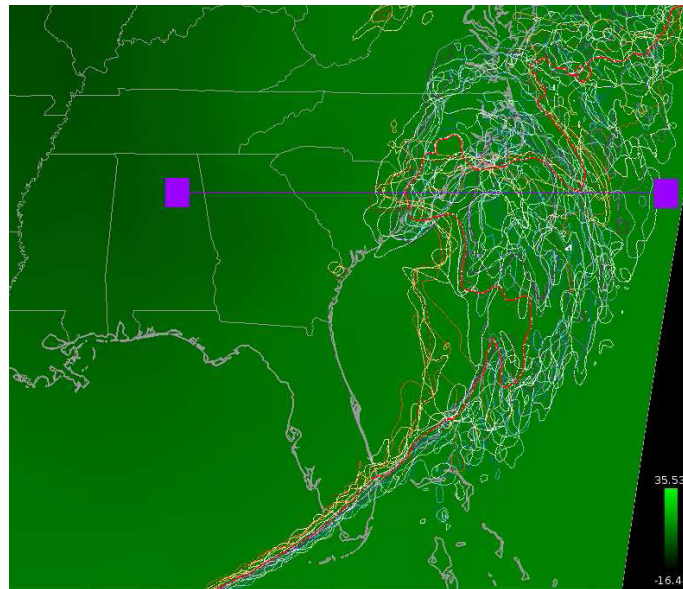


Figure 5.19

Spaghetti plot of perturbation potential temperature field at upper eta-levels.



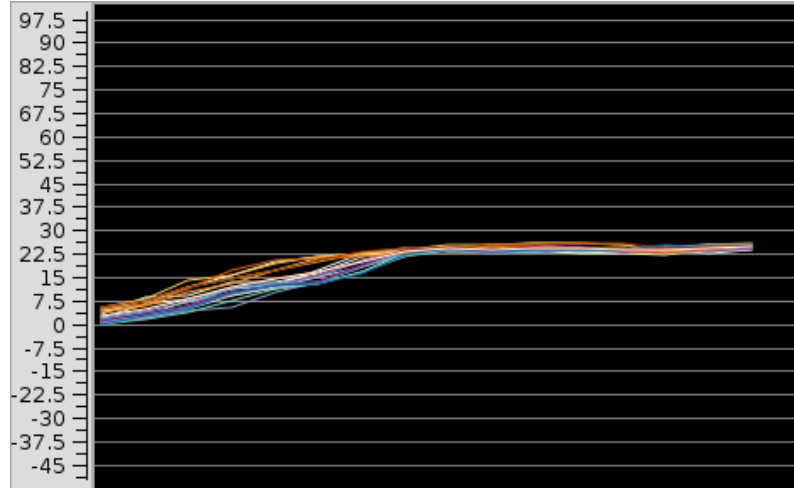


Figure 5.20

General agreement among schemes corresponding to Figure 5.19.

tions from the ensemble will help to improve the accuracy of ensemble forecasts by systematically removing biases in the ensemble members. Currently, operational meteorologists do not have this option and would greatly benefit from such a tool.

#### 5.4 Iterated Prototype: Noodles 2

The results from Noodles were very encouraging. We received valuable feedback from the domain experts based on which we reiterated the software prototype to design ‘Noodles 2’. The following sections describe the software architecture, user interface, visualization capabilities, and case studies of severe weather events.

##### 5.4.1 Software Architecture

*Noodles 2* has been designed following a layered software architecture as illustrated in Figure 5.21. The user interface layer sits at the very top and ‘drives’ the layers below

it. This allows one to easily add or modify interface components without any changes to the visualization subsystem. The user interface triggers signals that the underlying visualization layer intercepts and uses to update itself. This allows one to open multiple independent views with the same or different variables and choose independent visualizations. Additionally, custom visualizations can be easily added to existing visualization, views or completely separate views can be created.

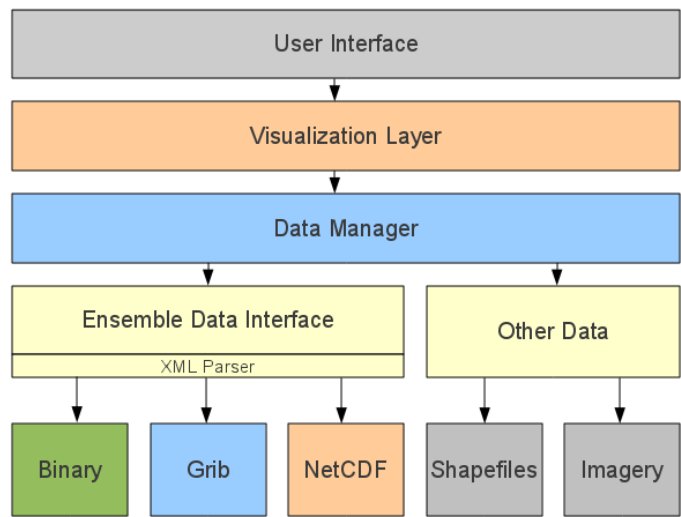


Figure 5.21

Software architecture of Noodles 2.

The visualization layer makes requests to the Ensemble Data Manager for all ensemble data accesses. The software has been designed to manage ensemble data separately from other auxiliary data. Variables and data are loaded by the Ensemble Data Manager layer dynamically and transparently from the visualization layer. Only the current time step and slice of currently loaded variables are kept in memory. With support for the NetCDF [78]

data format in Noodles 2, a large number of variables can be read and we are no longer limited to just the three variables that were studied in the previous prototype. The modular architecture of the system allows for easy extensibility. All other data requests such as the loading and unloading of auxiliary shapefiles or imagery is managed by a generic data manager and shared across visualization views.

### **5.4.2 User Interface**

Noodles 2 presents users with an interface as illustrated in Figure 5.22. The region marked '1' is the primary display area. Region '2' is an area that provides the user with information about the dataset. Region '3' constitutes all the different controls that enable a user to interact with the visualization. These allow a user to choose a variable, select a vertical level, enable spaghetti, enable or disable ensemble members, choose the type of uncertainty visualization, look at the data transect, and change properties of the shapefiles. In addition, a time slider, region '5', is provided to animate through time. The description of the ensemble is inputted to the tool by an XML file.

To allow the analysis of more than one variable, the tool allows users to bring up multiple views with different loaded variables (Figure 5.24 and Figure 5.26). Each view has the associated visualization controls to the right.

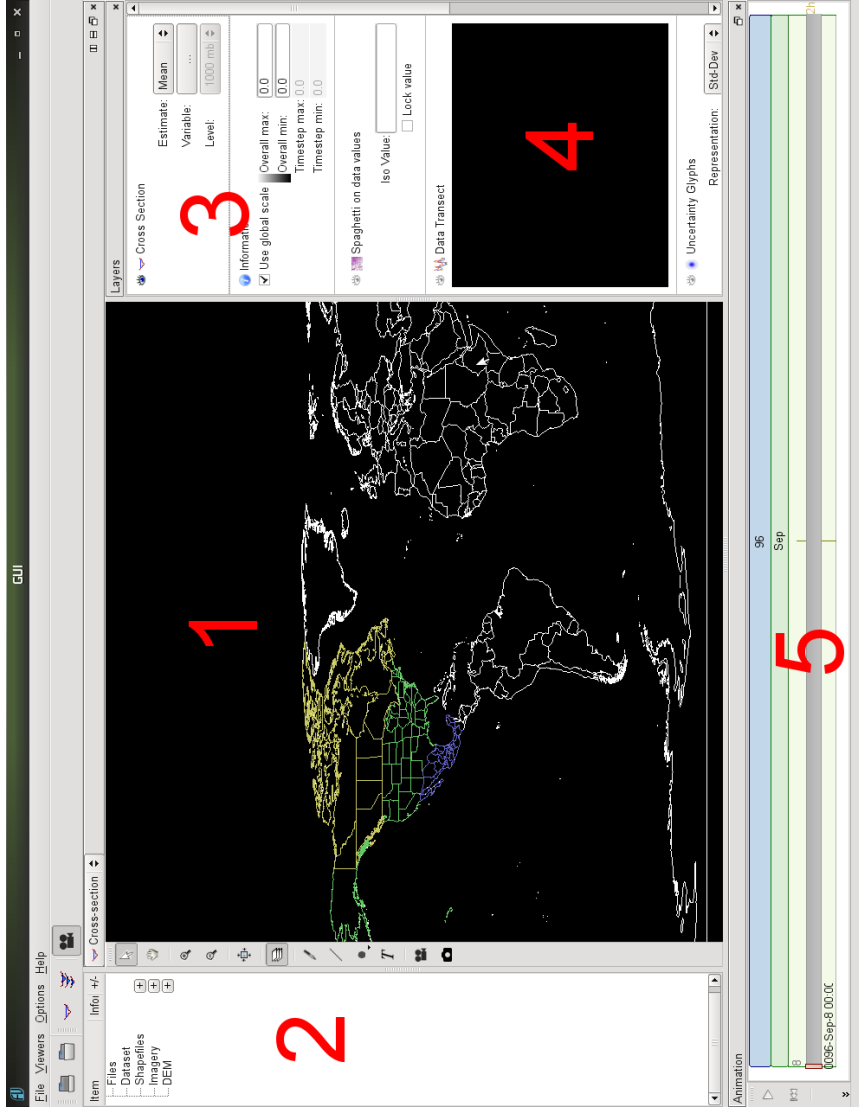


Figure 5.22

Marked up user interface of Noodles 2.

#### **5.4.2.1 Visualization Capabilities**

The visualization subsystem and the Ensemble Data Manager are the most critical components and work closely with one another. The GUI serves as a wrapper for all components.

We ported much of the visualization capabilities of ‘Noodles’ to ‘Noodles 2’. Users can visualize uncertainty by means of glyphs and ribbons, as well as spaghetti plots. In particular, the tool allows the user to input their choice of colors for the spaghetti making it very convenient to design and experiment with different color schemes. The multiple visualization views with independent subset of tools provides users with multiple coordinated views of different variables for analysis.

#### **5.4.2.2 Interaction Capabilities**

Much like its previous prototype, the tool allows one to interactively visualize the data. Although caching or streaming techniques were not utilized, the tool was still usable with tolerable delay (in the order of a few seconds) during the loading of the data. Once loaded, the visualizations could be rendered fairly interactively. One can double click the main view to probe the data value under the mouse cursor and interactively generate the spaghetti plot for the probed value. In addition, the data transect allows users to look at cross-sections of values through the grid (Figure 5.26).

### 5.4.3 Case Studies

We used the Weather Research and Forecasting (WRF) model version 3.3 [88] with the Advanced Research WRF (ARW) numerical solver core to run the simulations in the case studies. This version of the simulation model provides a much larger choice of parametrizations. With the objective of evaluating these parametrizations for severe weather events, we generated two ensemble data sets for evaluation using Noodles 2. We first describe the various cumulus and microphysics parametrizations that we used to create the ensemble and then describe the two severe weather events.

#### 5.4.3.1 Ensemble Formation and Running WRF

Various cumulus parametrizations that were included in this study were:

1. No cumulus scheme
2. Kain-Fritsch scheme
3. Betts-Miller-Janjic scheme
4. Grell-Devenyi scheme
5. Simplified Arakawa-Schubert
6. Grell-3 scheme
7. Tiedtke scheme
8. Zhang-McFarlane scheme
9. New SAS scheme

Various microphysics parametrizations that were included in this study were:

1. No microphysics scheme
2. Kessler scheme
3. Lin (Purdue) scheme

4. WSM Single Moment 3 class scheme
5. WSM Single Moment 5 class scheme
6. Eta (Ferrier)
7. WSM Single Moment 6 class scheme
8. Goddard 6 class scheme
9. Thompson scheme
10. Milbrandt-Yau double moment scheme
11. Morrison double moment scheme
12. SBU-YLin scheme
13. WRF Double Moment 5 class scheme
14. WRF Double Moment 6 class scheme

A total of 126 runs for all combinations of the above options were simulated. The data obtained was on a spatially staggered Arakawa-C grid. We performed a post processing step using the Universal Post Processor (UPP) 1.0 which not only interpolated the data back to an A grid, but also generated many derived meteorological variables such as convective inhibition (CIN), relative humidity (RH), and Convective available potential energy (CAPE).

#### **5.4.3.2 Case Study 1: Hurricane Fran's extra-tropical transition**

Hurricane Fran (Figure 5.23) [62] was a major hurricane to hit the eastern United States in the 1996 Atlantic hurricane season. It emerged from the west coast of Africa on 22 August as a tropical wave and became a tropical storm on 27 August. By 31 August, it had reached hurricane strength and followed a west-northwestward direction in the wake of Hurricane Edouard [75]. It recorded surface wind speeds of upto 15 knots on 5

September as a category 3 hurricane. It weakened to a tropical storm over North Carolina and subsequently became a tropical depression over Virginia. It gradually lost its warm core over the eastern Great Lakes and became extra-tropical at about 00:00 UTC on 9 September.

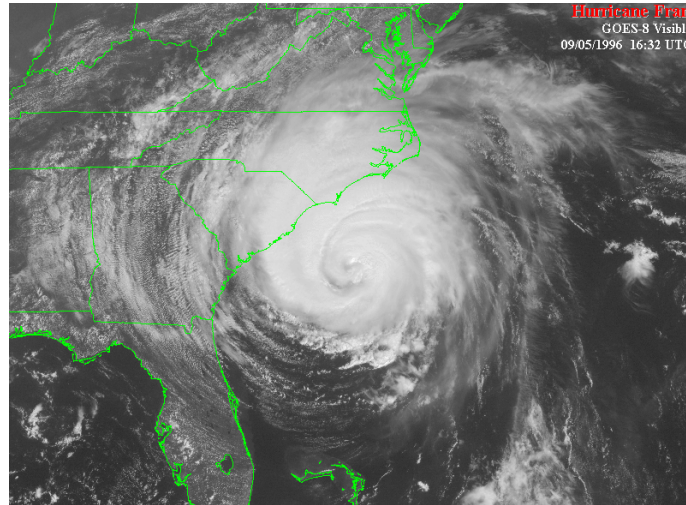


Figure 5.23

Satellite image of Hurricane Fran on 5 September 1996.

In this simulation, we trace the extra-tropical transition of Hurricane Fran. This was a 48 hour simulation starting at 1996-09-08 00:00 UTC. The simulation grid had dimensions of 298 (west-east)  $\times$  298 (south-north)  $\times$  37 (vertical). The latitudinal extent was from 16.95 N to 49.81 N while the longitudinal extent was from 105.41 W to 58.60 W, resulting in an average grid-spacing of 12 kilometers. The North American Regional Reanalysis (NARR) dataset [66] was the source of the input weather data, interpolated and re-gridded to the domain internally by WRF.



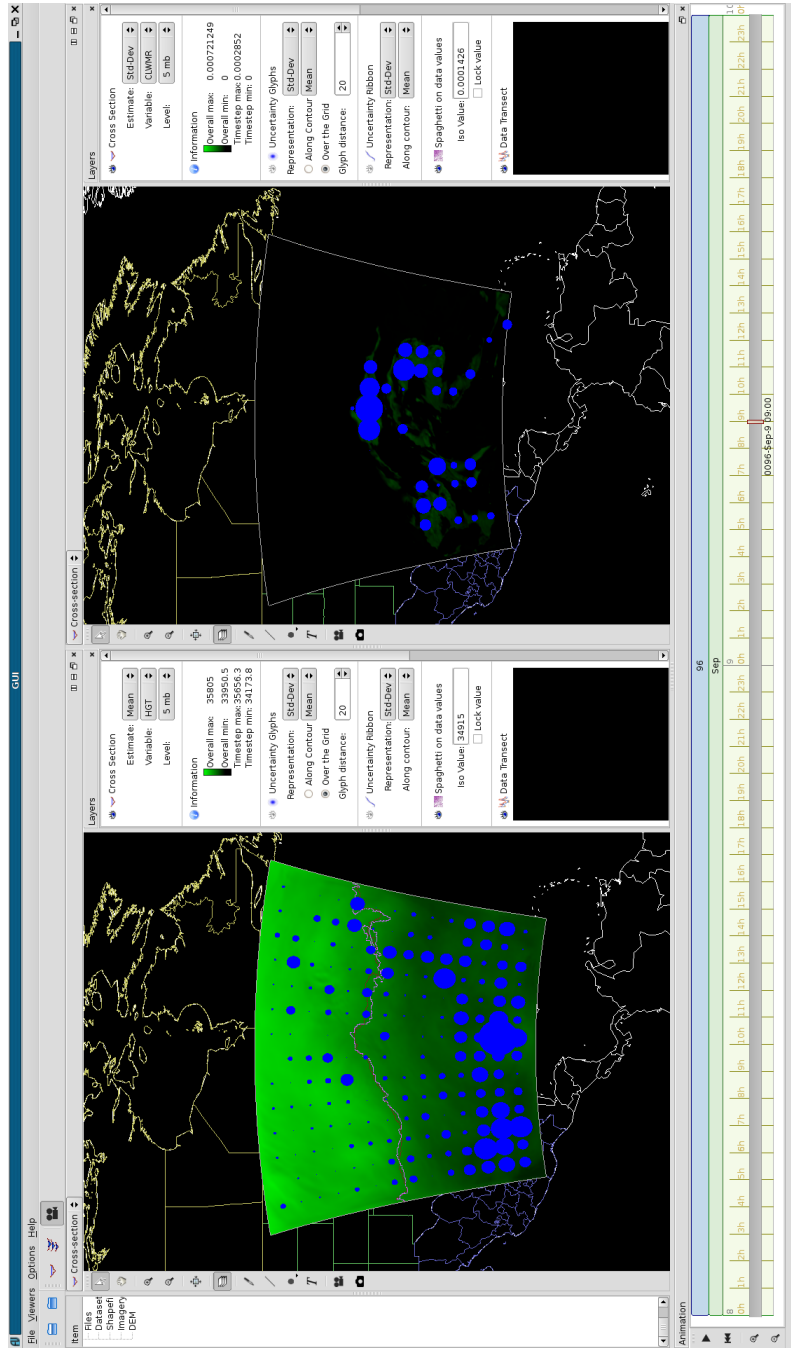


Figure 5.24  
Geopotential height and cloud-water mixing ratio for Hurricane Fran.

Post-processing of the data was a two step process with UPP performing the computations producing Grib files and subsequent conversion to NetCDF files. This resulted in a final dataset that was about 1.6 TB in size.

Various variables were visualized using the tool as illustrated in Figure 5.24.

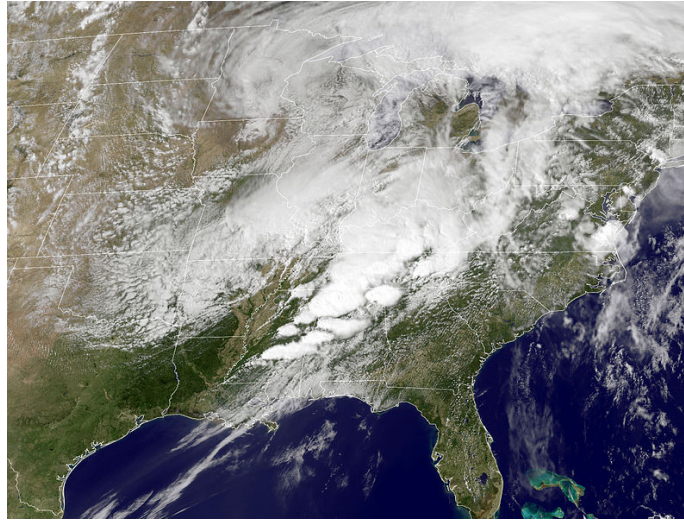


Figure 5.25

Satellite image of the Mississippi-Alabama tornados on 26 April 2011.

#### **5.4.3.3 Case Study 2: Mississippi-Alabama Tornado Outbreak**

The largest tornado outbreak ever recorded happened from 25 April 2011 to 28 April 2011 (Figure 5.25). It affected very large portions of the United States with catastrophic destruction over Alabama. Over 330 tornadoes were recorded during this period. A powerful jet stream, strong wind shear, low pressure center moving north-eastwards, moist warm air, and conducive temperatures resulted in a very active storm system.

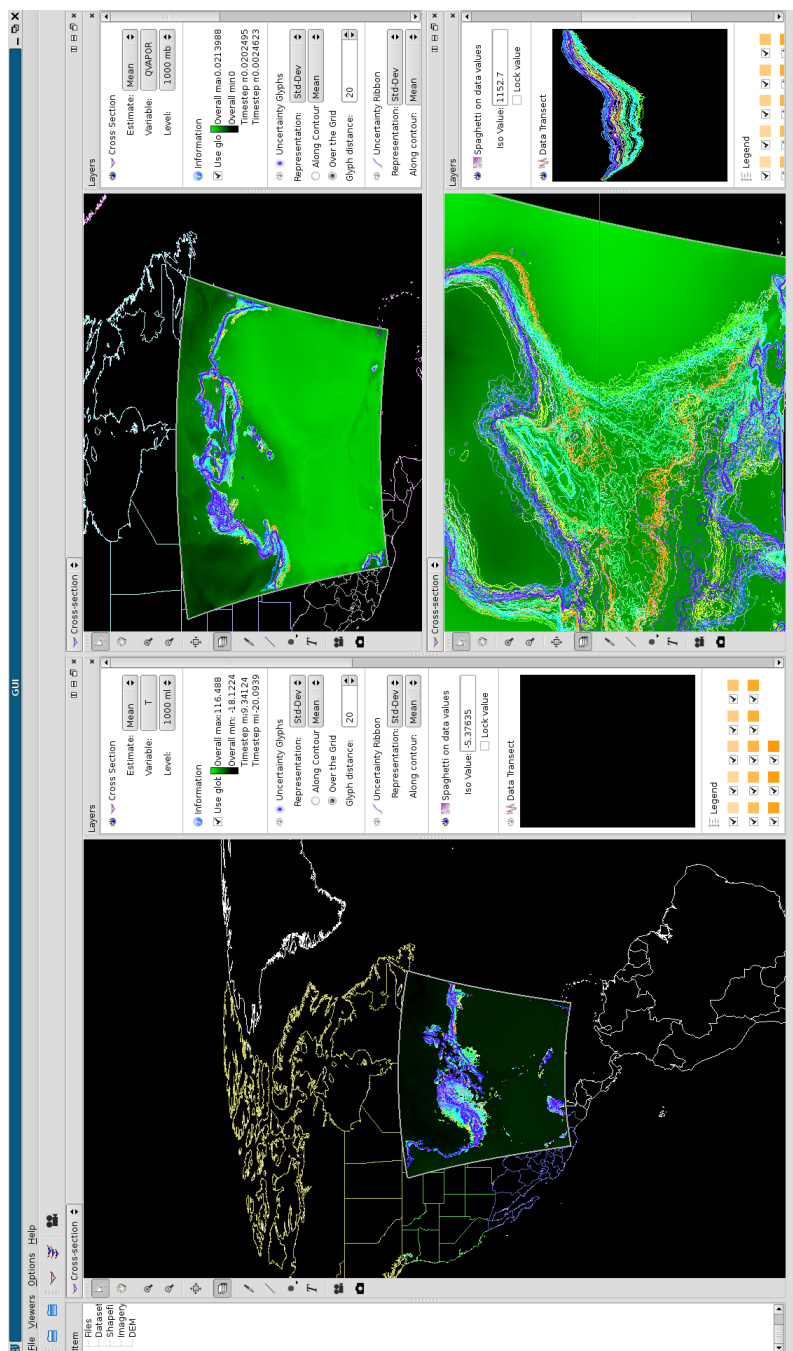


Figure 5.26

Temperature, water-vapor ratio, and pressure for the tornado outbreak.

In this simulation, we attempt to recreate the conditions of the tornado outbreak with a 96 hour simulation starting at 2011-04-24 12:00 UTC. The simulation grid had dimensions of 298 (west-east)  $\times$  298 (south-north)  $\times$  37 (vertical). The latitudinal extent was from 16.5 N to 47.0 N while the longitudinal extent was from 107.0 W to 57.5 W, resulting in an average grid-spacing of 12 kilometers. Here too, the North American Regional Reanalysis (NARR) dataset [66] was the source of the input weather data.

Post-processing of the data involved running UPP to produce Grib files which were converted to NetCDF files resulting in a 3.3 TB dataset.

Various variables visualized in Noodles 2 is illustrated in Figure 5.26.

## 5.5 Discussion

We presented the iterative design of two prototypes for meteorological ensemble visualization. We also reported the opinions of two experts in using the tool to perform analysis on severe weather events. They found many of the features to be useful in performing meteorological analysis and speculated that this could be very useful operationally.

CHAPTER 6  
SOFTWARE PROTOTYPE FOR RIVER-FLOW AND FLOOD-FORECAST  
ENSEMBLES

In this chapter, we describe the development of a tool named *FloodViz* and its visualization capabilities that will allow operational personnel in National Weather Service's (NWS) Lower Mississippi River Forecast Center (LMRFC) to visualize and analyze riverine flood simulation output. In particular, it will have the capability to visualize ensemble simulations and the resulting uncertainty. This is a collaborative effort between us at Mississippi State University and hydrologists at the Lower Mississippi River Forecast Center (LMRFC) who provide stake-holder feedback and expert guidance. We expect *FloodViz* to provide improved visualization and systems capabilities to help hydrologists in determining the extent of flooding, increasing their knowledge and understanding of such effects. The tool will also allow forecasters to relay more information to the emergency management community while issuing forecasts to help protect life and property.

### 6.1 Operational Flood Forecasting

The simulation model in operational use at the Lower Mississippi River Forecast Center (LMRFC) is the Hydrologic Engineering Center's River Analysis System (HEC-RAS) [14]. The software package allows personnel to perform one-dimensional steady flow,

unsteady flow, sediment transport/mobile bed computations, and water temperature modeling.

The HEC-RAS modeling system is comprised of four components for steady flow water surface profile computations, unsteady flow simulation, movable boundary sediment transport computations, and water quality analysis. The system uses various types of input data for the steady and unsteady flow simulation such as precipitation, plan of the simulation extent, geometry and surface properties, and sediment data, grouped into a project space.

### **6.1.1 Ensemble Data**

Simulation of river flow is imprecise and various types of uncertainties contribute to the inaccuracy and imprecision in the simulation. To account for some of the precipitation and modeling uncertainties, the hydrologists at LMRFC use a 13 member river forecast ensemble. The members of the ensemble are constituted from the Qualitative Precipitation Forecast (QPF) ensemble as input to the HEC-RAS model. The 13 QPF inputs are comprised of the average precipitation estimates for 0, 12, 24, 48, and 72 hours, and maximum and minimum precipitation estimates for 12, 24, 48, and 60 hours. The choice of these input quantities have been determined by operational requirements. Generally, the precipitation estimates tend to deviate substantially after 48 hours.

All visualizations represented here are based on a routine operational simulation run for a flood event between 10 July 2010 and 1 August 2010 which was provided to us by the LMRFC personnel.

### 6.1.1.1 Uncertainty Quantification

In many RFCs, ensemble runs are still at an experimental stage. As a result, methods to quantify the resulting uncertainty in the distribution are limited. Discussions with our LMRFC partners has indicated that they would initially be interested in descriptive statistics such as mean and variation of the water level and then figure out which other advanced metrics might be appropriate and applicable.

### 6.1.2 Geometry

To perform a hydraulic study, it is necessary to collect data both upstream and downstream of the selected area in order to model the boundary conditions effectively. A chosen study area with a network of connected streams and rivers along with the floodplain is referred to as a plan (Figure 6.1).

For each river, multiple cross-sections provide information about the ground-surface properties that are used by the modeling system to route the water. Cross-sections are typically 1D surveys from one side of the river to the other and perpendicular to the flow. They extend outwards much beyond the banks of the river and into the floodplain. The size of the cross-sections is very important in determining various flooding scenarios. Information of levees, bridges, embankments, and other channel modifications are important and may affect simulation results. Cross-sections must be carefully chosen to capture the effects of the presence of these features.

A collection of channel data from a selection of river cross-sections is called a river profile (Figure 6.1). A group of related data elements is collectively called as a HEC-RAS project.

### 6.1.3 Visualization Capabilities in HEC-RAS

The HEC-RAS tool suite is capable of generating plots of the river-system plan, cross-sections, profiles, rating curves, hydrographs, and very basic 3D scenes rendered with connected lines. The tool also allows limited pairwise comparisons of model runs. Figure 6.1 illustrates some of the current features available in HEC-RAS.

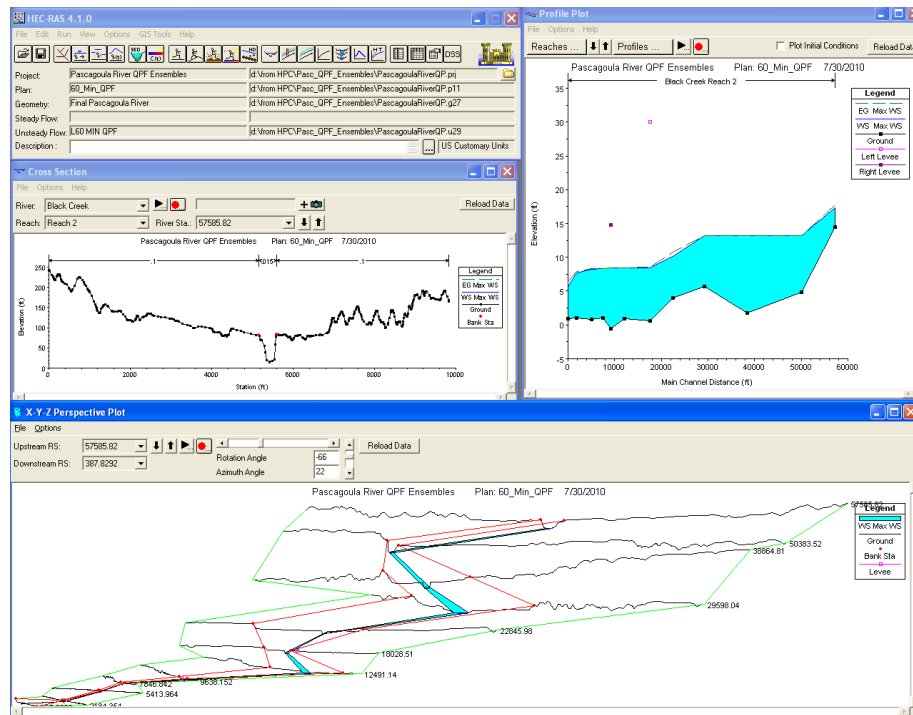


Figure 6.1

A screen-shot of visualization capability in HEC-RAS.



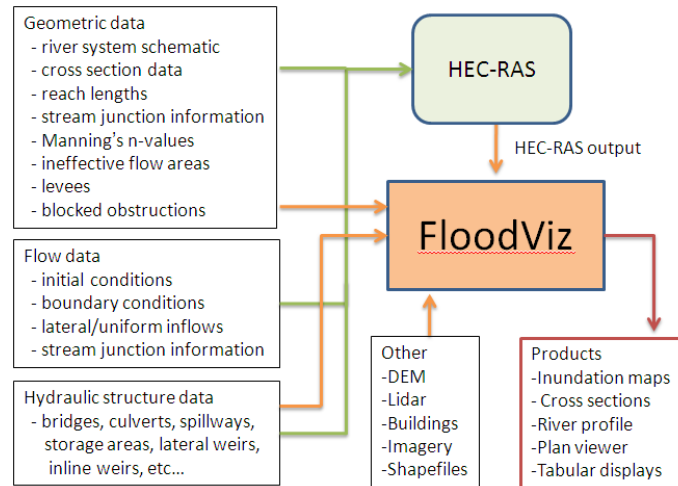


Figure 6.2

High level block diagram of FloodViz.

## 6.2 Software Engineering Effort

The software development objective is to design a ‘future-ready’ tool that will integrate into National Weather Service plans to implement a next generation forecast and modeling system called the Community Hydrologic Prediction System (CHPS) [64]. Figure 6.2 illustrates a high-level block-diagram of the software framework for Floodviz. The tool focuses on providing improved visualization and analysis capabilities for its users.

A spiral software development model has been followed by setting up long-term goals and executing software development cycles of short-term milestones, alpha-testing, and beta-testing. Development has been done in C++ and OpenGL. The GUI framework has been built using the Qt 4.6 software development kit. A number of open-source libraries have been used such as pthreads, boost, and GDAL/OGR. The system is cross-platform compatible and is capable of running on Linux, Mac OS, and Windows.

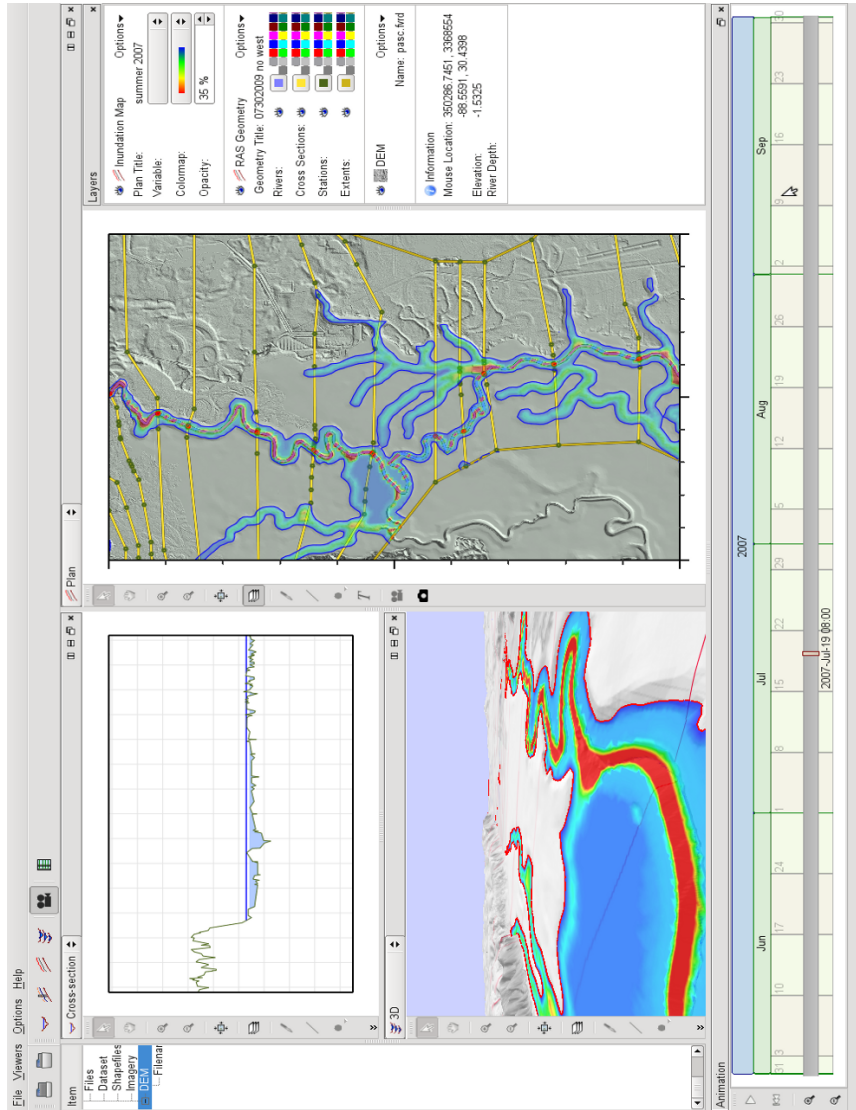


Figure 6.3

A screen-shot of the FloodViz user interface.

An important component in the development efforts has been the design of the HEC-RAS data reader. We were unable to obtain a C/C++ API for reading the data. Only a Java version of the HEC-RAS API was available. We had to resort to develop C++ wrappers around the Java API for FloodViz. This results in a performance downgrade although not as much as to affect its interactive usability.

Digital Elevation Models and tiled imagery for the river-system extent can also be read into Floodviz and are rendered with a level-of-detail algorithm to preserve the river channel while maintaining interactive frame rates.

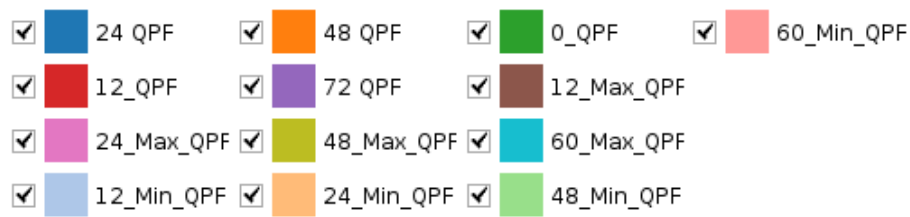


Figure 6.4

Interactive legend with categorical colors.

### 6.3 Visualization Views

FloodViz provides four different views for visualization of simulation output as described in the sections below. Besides providing various controls to adjust the colormap, transparency, and rendering order, it also provides an interactive categorical color legend (Figure 6.4) that allows users to easily enable or disable ensemble members for ease of

analysis. In addition, the multiple linked views provide an indication of the currently active cross-section view in the plan-view.

### **6.3.1 Cross-Section View**

A cross-section view shows the water level at fixed cross-sections along the river. Geometry data and simulation output is illustrated in this view. It is important to note that there are some scenarios in which the cross-sections do not go out far enough into the floodplain. In such situations the model is unable to account for the changes to the flood level as a result of the ‘ineffective areas’ leading to uncertainty. The cross-section view helps in the identification of such areas. Figure 6.5 illustrates one such cross-section on the Red Creek river in the Pascagoula region, MS.

### **6.3.2 River Profile View**

A profile view shows the water level at cross-sections downstream for a reach of a river. Figure 6.6 illustrates the river profile for the Red Creek river.

### **6.3.3 Plan View**

The plan view provides an overview of the inundation extent (Figure 6.7). A Digital Elevation Map (DEM) of the terrain is used to determine the extent of flooding by extending the the output water level to intersect with corresponding heights in the DEM thereby generating the flood boundary. The DEMs tend to be enormous in size and we use a level-of-detail approach to allow interactivity while providing a high resolution inunda-

tion map. The view is also useful in identifying regions where improvements in the survey are essential. In addition, this view also helps in easily identifying poor model behavior.

### 6.3.4 3D View

A river-channel preserving 3D continuous level-of-detail flyover of the terrain allows hydrologists to view the extent of the flooding in 3D (Figure 6.8). This sub-section is provided for completeness since uncertainty visualization techniques for this view are yet to be implemented.

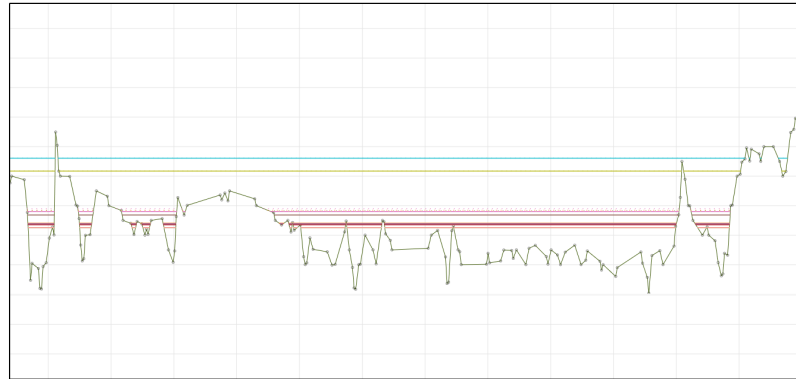


Figure 6.5

Cross-section view of the river.

## 6.4 Uncertainty Visualization Techniques

Various types of uncertainty visualization techniques are provided in the different visualization views that enable hydrologists to compare and contrast between the different model runs.

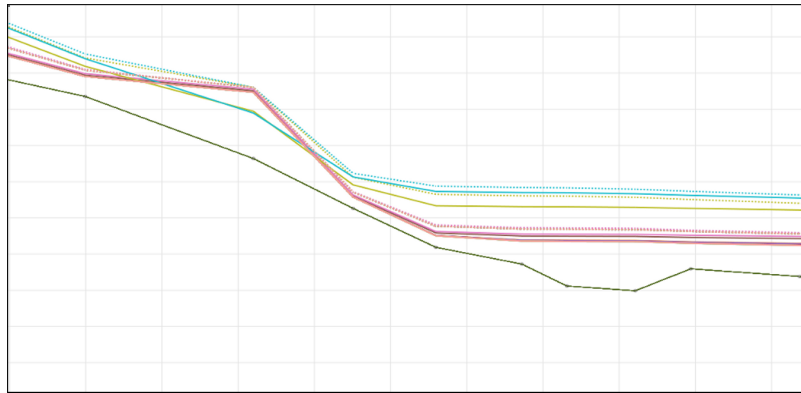


Figure 6.6

Profile view of the river.

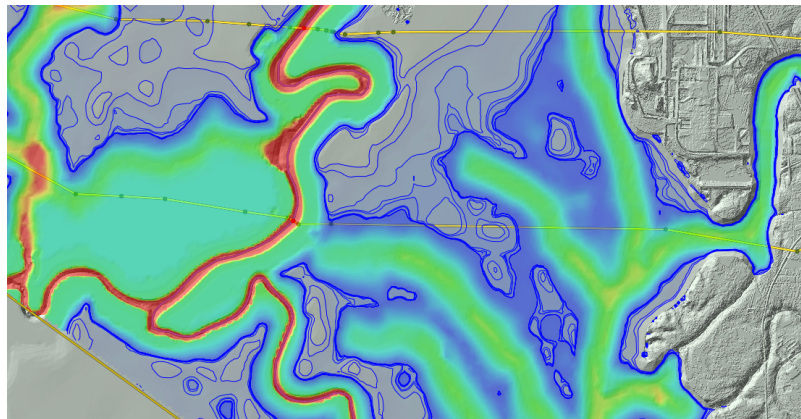


Figure 6.7

Plan view illustrating uncertainty in extent of flooding.

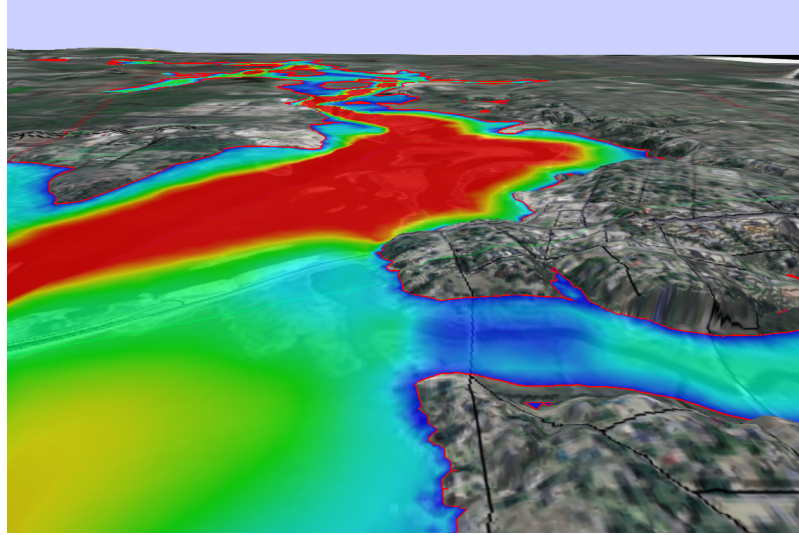


Figure 6.8

3D View of the extent of flooding.

#### 6.4.1 Line plots

Line plots are the simplest form of graphical representation in the cross-section and profile views (Figure 6.5 and Figure 6.6). The value of water surface elevation from each of the model outputs are rendered with the maximum predicted water surface. This simple visualization provides users with a detailed rendering of the different levels expected at the banks and on the flood plains for cross-sections. The individual lines provide the users with a sense of the uncertainty in the output. For the river profile view, it provides a trend for comparison between different simulations, especially with temporal animation. The interactive legend provides users with the capability to add or remove unwanted or poorly performing members of the ensemble.

While line plots are adequate for water surface elevations in perspective of the ground level, other variables in the simulation such as stream velocity and discharge amounts

cannot be represented by a line at a certain distance from the ground. Other overloaded or multivariate visualizations were explored as detailed in later sub-sections.

#### 6.4.2 Box plots

To derive a sense of the distribution of the ensemble members, we constructed box-plots of the data (Figure 6.9 and Figure 6.10). The mean of the water surface for the time-step and the mean of the maximum water surface elevations are displayed along with small notches for individual ensemble members under the box-plots. The resulting visualization provides users with an illustration of the uncertainty of the distribution along with individual member levels. This is particularly useful in the river profile view as trends between cross-sections can be analyzed. Still, this visualization does not illustrate non-elevation type of variables.

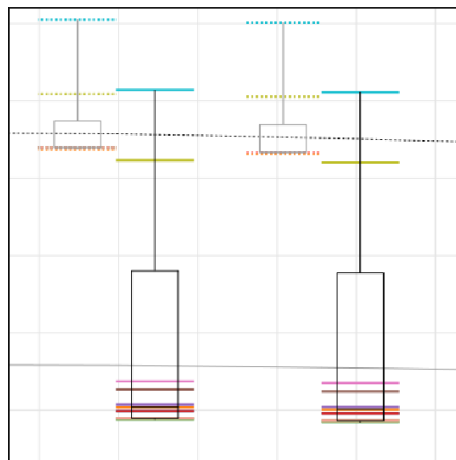


Figure 6.9

Close-up of a box plot.



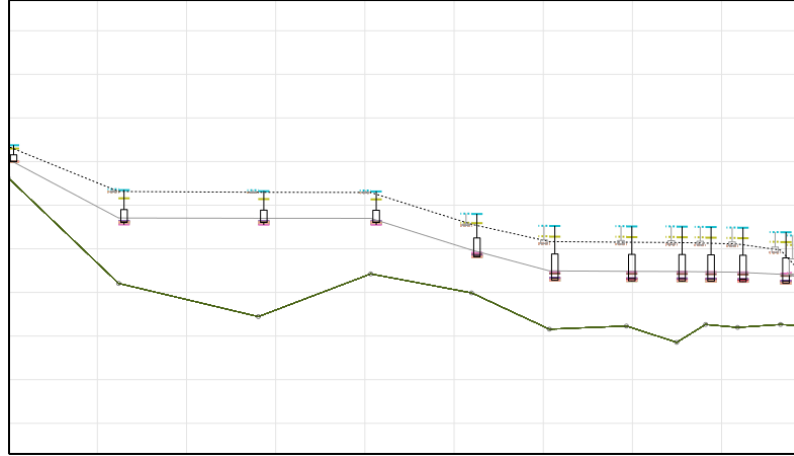


Figure 6.10

Box plot in the profile view illustrating water surface elevation.

### 6.4.3 Glyph plots

We adopted a glyph based approach to encode other non-elevation variables (Figure 6.11 and Figure 6.12). In this view, circular glyphs are rendered at the water surface level for each member of the ensemble. The glyphs are color coded for easy identification. The radius of each glyph is mapped to the value of the variable normalized to a maximum and minimum size. This representation encodes the water-level to the vertical position of the glyph and another chosen variable of interest to map to the size of the glyphs. The differences between glyphs from the model runs illustrates the individual differences highlighting the uncertainty.

### 6.4.4 Star uncertainty glyphs

The glyph plots restricts us to two, maybe three variables: water surface elevation, maximum water surface elevation, and another variable of interest. We used a star plot

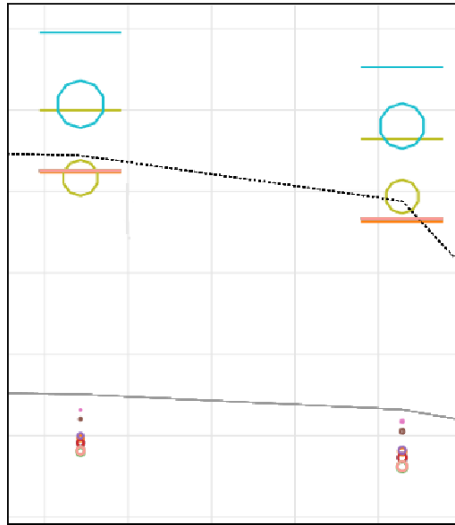


Figure 6.11

Close-up of the glyphs scaled by size.

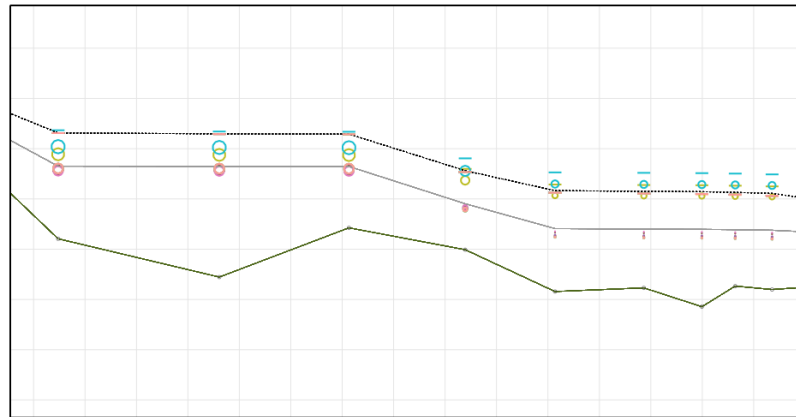


Figure 6.12

Glyphs illustrating the water surface elevation.

to encode more than three variables in a visualization (Figure 6.13 and Figure 6.14). The length of an arm of the star plot represents the derived mean of a variable. The saturation of the color of the arm represents the standard-deviation or uncertainty of the variable. A circle is drawn around the star to illustrate the maximum possible value. Thus, normalized values of the variables are used with the maximum value always corresponding to the radius of the circle. Any number of variables can be represented using this visualization, although the cognitive load will increase tremendously. This approach allows one to explore the uncertainty relationships between different variables.

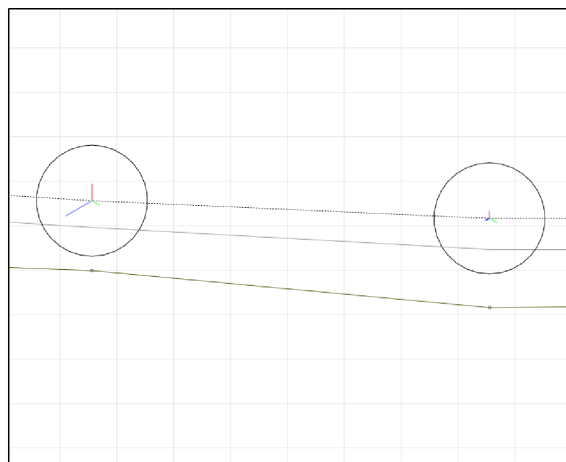


Figure 6.13

A star (multivariate) uncertainty glyph.

#### 6.4.5 Temporal multivariate uncertainty glyphs

To give users a sense of the temporal tendencies of variables, we extended the design of the star uncertainty glyphs to represent temporal trends (Figure 6.15 and Figure 6.16).

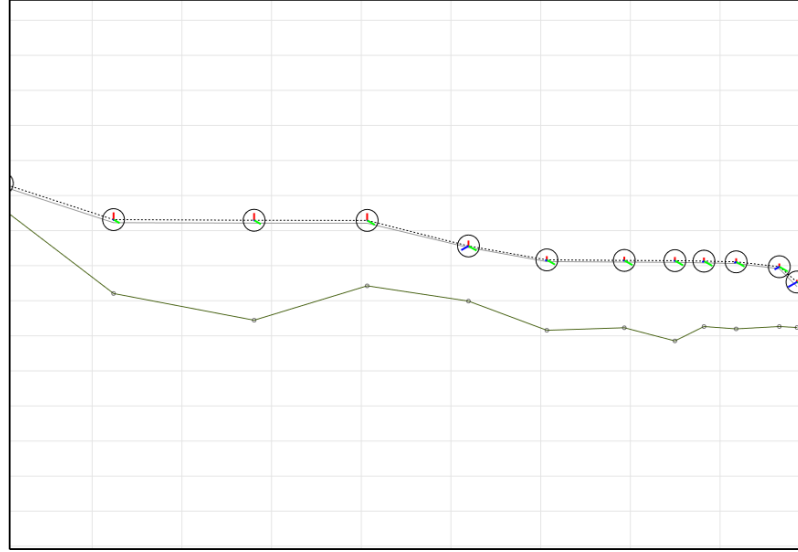


Figure 6.14

Star glyphs illustrating the water surface elevation.

Multiple star plots are rendered for each time-step starting at time-step 0 and progressing upto the current time-step, with each star rendered at a slight angle to the previous such that the entire pie is filled for the final time-step. Thus, for a given time-step, the arm of the star-glyph sweeps out a pie for a variable. The radius used at each time-step is proportional to the mean variable value and the color used is proportional to the standard-deviation of the variable.

This representation can also support any number of variables but will increase the cognitive load significantly as the number of variables increases.

#### 6.4.6 Inundation map with uncertainty

An inundation map depicting the uncertainty of the extent of flooding can be generated from the plan view. We composite multiple inundation maps using a simple trans-

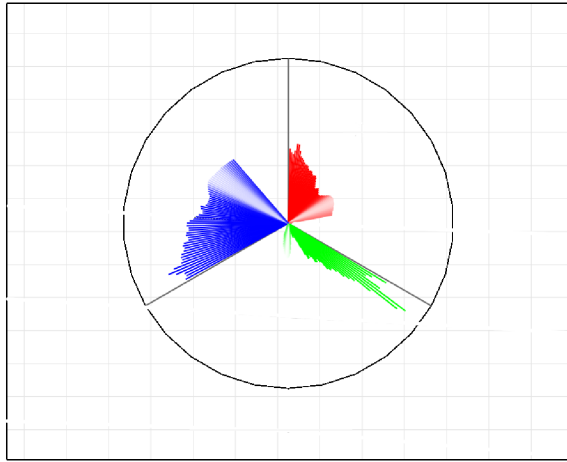


Figure 6.15

Temporal glyphs illustrating three variables over time.

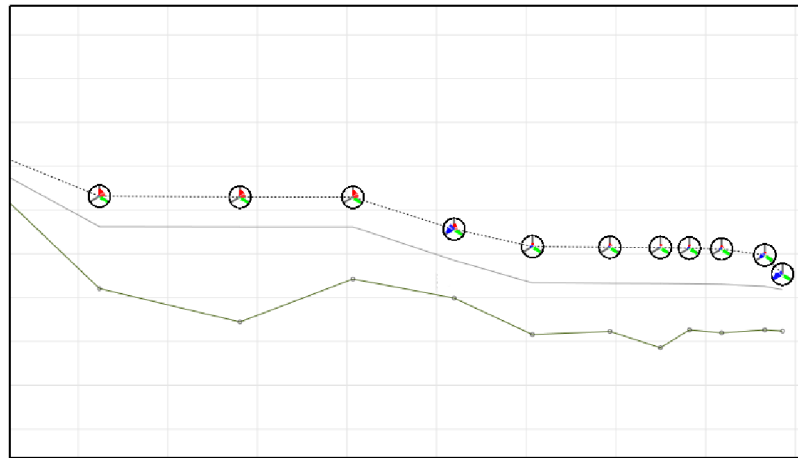


Figure 6.16

Temporal glyphs in the profile view.

parency based blending function as illustrated in Figure 6.7. The transparency, visibility, and color-mapping of variables on individual inundation maps can be controlled leading to a high-dimensional visualization. The representation leads to a visualization where multiple layers of inundation maps provide a sense of the uncertainty in the simulation and extent of flooding. Figure 6.7 illustrates this where the colormap indicates the depth of water. These maps help hydrologists to understand and identify areas that all models agree upon leading to a high probability of inundation. These maps also help them to pick out extreme scenarios which tend to be rare but are usually far more devastating.

## 6.5 Evaluation

The visualizations were presented to a hydrologist who performed evaluations from an operational perspective. He found that the line plots (Figure 6.5 and Figure 6.6) provide a nice and simple representation of the ensemble members and directly describe the uncertainty. This is both easy to read and easy to interpret and are like an extension of currently used hydrographs.

Among the techniques explored, he found the box-plots to have the best appeal (Figure 6.10). Although many of the individual ensemble members tend to remain in the same relative location within each box-plot, it does give a good indication of how the members compare to each other. The box-plots vary with time, which can be changed with the time slider, which is a helpful addition. The uncertainty can also be well estimated using the area of the actual box.

He thought that the star representation (Figure 6.14) definitely allows for more variables to be plotted, but when the actual value of a variable is low, it becomes difficult to assess. In other words, when the length of the arm is small, not much information can be taken from the plots. However, this does provide a relatively efficient way to compare information from two locations along a river reach, despite the fact that the unique ensemble members are not directly represented and only the overall uncertainty is shown.

He thought that the temporal glyphs (Figure 6.16) were a good idea but there appeared to be just too much going on. Also, it is difficult to read the bar graphs associated with each variable since the x-axis is not a straight line. He found himself constantly turning his head to read the graphs. It might be useful to have control widgets that can rotate the glyphs. Also, it is difficult to see information from each individual plot unless zoomed in on the glyph. He suggested that we plot each variable in a different color.

He felt that the glyph plots (Figure 6.12) had a lot of potential. Centering the circular glyphs over the corresponding water surface levels seems to help. More than any other of the approaches presented, the glyphs provide a good indication of how values are changing along the river reach.

He found the uncertainty inundation maps (Figure 6.7) to be extremely useful in determining the extent of the flooding and how the ensemble members deviated from one another.

## 6.6 Discussion and Future Work

In this chapter, we described the architecture and uncertainty visualization capabilities of a tool named FloodViz. In particular, the ensemble uncertainty visualization capabilities were highlighted. We also presented an evaluation by a hydrologist who found some of the techniques such as the box-plots and uncertainty inundation maps to be very helpful while some of the other techniques such as glyph-plots had good potential. He also felt that some of the techniques might be too over-loaded and might require more careful redesign.

Many of the capabilities are constantly being re-evaluated and re-designed. In the coming months, we have planned more evaluations to assess which of the techniques will finally be a part of the tool when deployed operationally. Future plans include further exploration of uncertainty visualization techniques and improved visual analytic capabilities.



## CHAPTER 7

### CONCLUSION

In this dissertation, we presented a user study to compare effectiveness of four uncertainty visualization techniques on 1D and 2D datasets. We had hoped that we would be able to provide guidelines for scientific visualization design with uncertainty from the results of our study but our results are such that we cannot identify clear winners. We did not find a consistent ordering for the four techniques for all the tasks, although the particular findings could be useful for uncertainty visualization design. Errorbars consistently performed poorly compared to the other evaluated techniques. The accuracy of responses for 1D tasks was higher than that of the 2D tasks although 2D tasks consistently took longer to respond to. We also found that effectiveness of uncertainty visualization techniques were highly dependent on the task at hand. User efficiency between the two search tasks was significantly different from one another which raised interesting questions. Surface-coloring worked well except for counting 2D data features. Performance of Glyph-size and Glyph-color seemed reasonable. We feel our results could help researchers in their choice of uncertainty visualization technique for their scientific tasks.

We expect our findings to be useful for researchers who have a need of displaying dense 1D or 2D data with uncertainty. In particular, we see applications from geoscience

such as visualization of severe weather outbreaks, precise terrain modeling, and moving front locations to benefit from this study.

We presented a method to create synthetic data with systemic and random uncertainty components. In addition, we also presented an Uncertainty Evaluation Framework which provides a structured design environment which can be used to create effective uncertainty visualizations across different visualization paradigms.

We also presented a tool named Noodles that was developed for operational meteorologists to visualize ensemble uncertainty. Two new 2D uncertainty visualization techniques were implemented to represent the uncertainty in weather ensembles. Two experts evaluated these techniques alongside the conventionally used spaghetti plots and it appears that there is substantial research and operational benefit of the interactive application of these techniques. For the 1993 ‘Superstorm’ simulation dataset in particular, the Grell-Devenyi cumulus scheme was found to be in disagreement with the other schemes for the cold and warm fronts which raised important questions regarding the choice and effectiveness of the various parameterization schemes for different weather conditions. It also made the meteorologists think that our findings could be of interest to the designers of these schemes.

We then presented the design and software architecture of an iterated software prototype named ‘Noodles 2’. The various improvements and additional features described were based on expert feedback and user interaction. Also, we described the setup and design of a more extensive ensemble simulation to study two additional weather events: Hurricane Fran and the Alabama-Mississippi tornado outbreak. One of the meteorolo-

gists, who served as domain experts who evaluated the tools, remarked that the tool could be the “state-of-the-art” if used in operational meteorology. These are very encouraging comments and results which motivate us to further improve the tool.

Additionally, we described the architecture and visualization capabilities of a tool named FloodViz. In particular, the ensemble uncertainty visualization capabilities were highlighted. We also presented an evaluation of the techniques by a hydrologist who found that some of the techniques such as the box-plots and uncertainty inundation maps to be very helpful while some of the other techniques such as glyph-plots had good potential. He also felt that some of the techniques might be too over-loaded and might require more careful redesign.

This research has illustrated that various uncertainty visualization techniques can be applied to understand uncertainty in ensemble simulations. The tools and techniques developed in this dissertation have the potential to be useful for meteorologists and hydrologists and further work to continue this endeavor appears to be beneficial.

## CHAPTER 8

### FUTURE WORK

Our results on the differences between uncertainty visualization methods motivate future research in this area. With our Uncertainty Evaluation Framework, we plan to use the results from this study to guide our future uncertainty visualization endeavours. Perceptual research to identify the reason why the two search tasks differed so significantly could be potentially beneficial. It may also be very enlightening to research how experts use uncertainty in their decision making process and design experiments around such observations. We plan to extend this research to evaluate uncertainty visualization techniques for 3D data as well as time series data. This may be of benefit to users of data that are inherently 3D and have samples in time. Weather researchers for example may significantly benefit from such knowledge.

The weather simulations open up a host of research questions regarding the sensitivity of models to various parameterizations. These would be research questions that fall in the domain of meteorology. Research questions abound in visualization as well. One such problem is to extend the tool to 3D uncertainty visualization. Implementation of out-of-core and GPU based techniques can significantly improve the performance and enable it to read much larger datasets. We plan to continue working with the subject matter experts

to study ensembles of other events and potentially other datasets such as simulation output from the Navy Coastal Ocean Model (NCOM) [3, 4].

The FloodViz project is still work in progress. Many of the capabilities are constantly being re-evaluated and re-designed. In the coming months, we have planned more comprehensive expert evaluations to assess which of the techniques will finally be part of the tool when deployed operationally. Future plans include further exploration of uncertainty visualization techniques and improved visual analytic capabilities.

## **8.1 Recommendations for Future Research**

The use of visualization techniques for the analysis of ensemble data offers potential benefits to operational forecasters. From our experience, some avenues of research that could lead to significant improvement in the field of weather visualization are:

### **8.1.1 Understanding the forecasters' needs**

The computer scientist is a tool-smith [13]. Close collaboration with domain experts is necessary to understand their unique requirements. Good understanding of meteorology and hydrology can go a long way in the design of effective tools that can help save time and facilitate decision support. Likewise, knowledge of the visualization techniques and their potential drawbacks can greatly help domain experts too. For example, understanding the pseudo-banding effects of rainbow colormaps [10] could improve meteorological data analysis.

### 8.1.2 Visualization Techniques

Research focusing on the development of techniques particularly for weather simulation output and hydrology simulation output is necessary. Visualization research must focus on the analytic requirements (e.g., understanding how meteorologists use multiple variables and the interaction between the variables to come to a conclusion). It is important for the computer scientist to understand decision making from simulation models. Forecasts are often augmented with real-time sensor data. Data fusion for visualization offers many challenges and possible research paths. While improved uncertainty visualization of model simulation output can help the expert, improved visualization for public dissemination is also very important. Geoscientific visualization research could benefit from tools like Colorbrewer [33] and insights from MacEachren [56], Bertin [6], Tufte [100, 99, 101], and Ware [103].

### 8.1.3 Evaluation

Research must be conducted to evaluate the effectiveness of the devised visualization techniques, both for the expert and for the public. It is important to understand that the requirements are somewhat different for each category of users. While experts would want to understand the science behind the simulation, the general public are interested in knowing how the weather would affect them. It is important to convey the notion of ‘possibility’ to the public. All visualizations designed should thus be evaluated for their effectiveness.

#### **8.1.4 Frameworks**

Time invested in the developing visualization frameworks could be of value. As illustrated in *Noodles 2* and *FloodViz*, a robust framework augments the research and development of new visualization techniques and paradigms.

#### **8.1.5 Engineering prototype software**

Investing time and effort to design and engineer prototype software can go a long way in facilitating improvements and new research. Robust platforms for the research and development of new visualization techniques and tools can facilitate technology transfer as well.

## REFERENCES

- [1] A. Arakawa and V. Lamb, “Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model,” *Methods in Computational Physics*, 1977, pp. 173–256.
- [2] M. Baker and C. Bushell, “After the storm: Considerations for information visualization,” *IEEE Computer Graphics and Applications*, vol. 15, no. 3, 1995, pp. 12–15.
- [3] C. Barron, A. Birol Kara, H. Hurlburt, C. Rowley, and L. Smedstad, “Sea Surface Height Predictions from the Global Navy Coastal Ocean Model during 1998–2001,” *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 12, 2004, pp. 1876–1893.
- [4] C. Barron, A. Kara, P. Martin, R. Rhodes, and L. Smedstad, “Formulation, implementation and examination of vertical coordinate choices in the Global Navy Coastal Ocean Model (NCOM),” *Ocean Modelling*, vol. 11, no. 3-4, 2006, pp. 347–375.
- [5] S. Benjamin, J. Brown, K. Brundage, B. Schwartz, T. Smirnova, T. Smith, and L. Morone, “RUC 2–The Rapid Update Cycle Version 2,” *NWS Technical Procedure Bulletin*, vol. 448, 1998.
- [6] J. Bertin, *Semiology of Graphics*, University of Wisconsin Press, 1983.
- [7] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, “Guide to the Expression of Uncertainty in Measurement,” *International Standards Organization, Geneva*, 1st edition 1993.
- [8] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, “Guide to the Expression of Uncertainty in Measurement,” corrected and reprinted 1995.
- [9] S. Blenkinsop, P. Fisher, L. Bastin, and J. Wood, “Evaluating the perception of uncertainty in alternative visualization strategies,” *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 37, no. 1, 2000, pp. 1–14.
- [10] D. Borland and R. Taylor II, “Rainbow color map (still) considered harmful,” *IEEE Computer Graphics and Applications*, 2007, pp. 14–17.



- [11] P. Bourke, "CONREC: A contouring subroutine," *Byte: The Small Systems Journal*, vol. 12, no. 6, 1987, pp. 143–150.
- [12] K. Brodlie, *Scientific visualization: techniques and applications*, Springer, 1992.
- [13] F. Brooks Jr, "The computer scientist as toolsmith II," *Communications of the ACM*, vol. 39, no. 3, 1996, pp. 61–68.
- [14] G. Brunner, "HEC-RAS (River Analysis System)," *North American Water and Environment Congress & Destructive Water*. American Society for Civil Engineers, 2002, pp. 3782–3787.
- [15] B. Battenfield, "Representing data quality," *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 30, no. 2, 1993, pp. 1–7.
- [16] B. Cabral and L. Leedom, "Imaging vector fields using line integral convolution," *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 1993, pp. 263–270.
- [17] P. Caplan, "The 12-14 March 1993 superstorm: Performance of the NMC global medium-range model," *Bulletin of the American Meteorological Society*, vol. 76, no. 2, 1995, pp. 201–212.
- [18] A. Cedilnik and P. Rheingans, "Procedural annotation of uncertain information," *Proceedings of the conference on Visualization '00*, Salt Lake City, Utah, United States, 2000, pp. 77–83.
- [19] H. Couclelis, "The certainty of uncertainty: GIS and the limits of geographic knowledge," *Transactions in GIS*, vol. 7, no. 2, 2003, pp. 165–175.
- [20] G. Cumming, F. Fidler, and D. Vaux, "Error bars in experimental biology," *The Journal of Cell Biology*, vol. 177, no. 1, 2007, pp. 7–11.
- [21] T. Davis and C. Keller, "Modelling and visualizing multiple spatial uncertainties," *Computers & Geosciences*, vol. 23, no. 4, 1997, pp. 397–408.
- [22] C. Dey, *GRIB: The WMO Format for the Storage of Weather Product Information and the Exchange of Weather Product Messages in Gridded Binary Form as Used by NCEP Central Operations*, Tech. Rep. NWS Office Note 388, NOAA/NWS, Washington, DC, 1998.
- [23] P. Diggle, P. Heagerty, K. Liang, and S. Zeger, *Analysis of longitudinal data*, Oxford University Press, United States, 2002.
- [24] S. Djurcilov, K. Kim, P. Lermusiaux, and A. Pang, "Visualizing scalar volumetric data with uncertainty," *Computers & Graphics*, vol. 26, no. 2, 2002, pp. 239–248.

- [25] R. Drebin, L. Carpenter, and P. Hanrahan, "Volume rendering," *ACM Siggraph Computer Graphics*, vol. 22, no. 4, 1988, pp. 65–74.
- [26] R. Earnshaw and N. Wiseman, *An introductory guide to scientific visualization*, Springer-Verlag, New York, United States, 1992.
- [27] B. Efron and R. Tibshirani, *An introduction to the bootstrap*, Chapman & Hall, 1997.
- [28] C. R. Ehlschlaeger, A. M. Shortridge, and M. F. Goodchild, "Visualizing spatial data uncertainty using animation," *Computers & Geosciences*, vol. 23, no. 4, 1997, pp. 387 – 395.
- [29] N. Gershon, "Visualization of an imperfect world," *IEEE Computer Graphics and Applications*, vol. 18, no. 4, 1998, pp. 43–45.
- [30] Graham, "A COCOA implementation of CONREC," <http://www.apptree.net/conrec.htm/>, current March 2010.
- [31] C. Hansen and C. Johnson, *The visualization handbook*, Academic Press, 2005.
- [32] M. Harrower, "Representing uncertainty: Does it help people make better decisions?," *UCGIS Workshop: Geospatial Visualization and Knowledge Discovery Workshop*, Landsdowne, Virginia, 2003.
- [33] M. Harrower and C. Brewer, "Colorbrewer.org: an online tool for selecting colour schemes for maps," *The Cartographic Journal*, vol. 40, no. 1, 2003, pp. 27–37.
- [34] T. Hengl, "Visualisation of uncertainty using the HSI colour model: computations with colours," *Proceedings of the 7th International Conference on GeoComputation*, 2003, pp. 8–17.
- [35] T. Hengl and N. Toomanian, "Maps are not what they seem: representing uncertainty in soil-property maps," *Proceedings 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, 2006, pp. 805–813.
- [36] W. Hibbard and D. Santek, "Visualizing large data sets in the earth sciences," *IEEE Computer*, 1989, pp. 53–57.
- [37] S. Huang, *Exploratory visualization of data with variable quality*, master's thesis, Worcester Polytechnic Institute, 2005.
- [38] G. Hunter and M. Goodchild, "Managing uncertainty in spatial databases: Putting theory into practice," *URISA Journal*, vol. 5, no. 2, 1993, pp. 55–62.
- [39] Z. Huo, D. Zhang, and J. Gyakum, "An application of potential vorticity inversion to improving the numerical prediction of the March 1993 superstorm," *Monthly Weather Review*, vol. 126, no. 2, 1998, pp. 424–436.

- [40] B. Jiang, F. Ormeling, and W. Kainz, "Visualization support for fuzzy spatial analysis," *Technical Papers: American Congress of Surveying and Mapping; American Society for Photogrammetry and Remote Sensing*, 1995, pp. 291–300.
- [41] C. Johnson and A. Sanderson, "A next step: Visualizing errors and uncertainty," *IEEE Computer Graphics and Applications*, vol. 23, no. 5, 2003, pp. 6–10.
- [42] A. Joshi, J. Caban, P. Rheingans, and L. Sparling, "Case Study on Visualizing Hurricanes Using Illustration-Inspired Techniques," *IEEE Transactions on Visualization and Computer Graphics*, 2008, pp. 709–718.
- [43] A. Kasahara, "Various vertical coordinate systems used for numerical weather prediction," *Monthly Weather Review*, vol. 102, no. 7, 1974, pp. 509–522.
- [44] P. Kocin, P. Schumacher, R. Morales Jr, and L. Uccellini, "Overview of the 12-14 March 1993 superstorm," *Bulletin of the American Meteorological Society*, vol. 76, no. 2, 1995, pp. 165–182.
- [45] D. Laidlaw, R. Kirby, C. Jackson, J. Davidson, T. Miller, M. Da Silva, W. Warren, and M. Tarr, "Comparing 2D vector field visualization methods: A user study," *IEEE Transactions on Visualization and Computer Graphics*, 2005, pp. 59–70.
- [46] V. Lakshmanan and J. S. Kain, "A Gaussian Mixture Model Approach to Forecast Verification," *Weather and Forecasting*, vol. 25, no. 3, 2010, pp. 908–920.
- [47] M. Leitner and B. Buttenfield, "Guidelines for the display of attribute certainty," *Cartography and Geographic Information Science*, vol. 27, no. 1, 2000, pp. 3–14.
- [48] H. Levkowitz and G. Herman, "Color scales for image data," *IEEE Computer Graphics and Applications*, vol. 12, no. 1, 1992, pp. 72–80.
- [49] H. Li, C. Fu, Y. Li, and A. Hanson, "Visualizing large-scale uncertainty in astrophysical data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, 2007, pp. 1640–1647.
- [50] J. Li, W. keng Liao, A. Choudhary, R. Ross, R. Thakur, R. Latham, A. Siegel, B. Gallagher, and M. Zingale, "Parallel netCDF: A high-performance scientific I/O interface," *In Proceedings of Supercomputing*, 2003.
- [51] S. K. Lodha, C. M. Wilson, and R. E. Sheehan, "LISTEN: sounding uncertainty visualization," *Proceedings of the 7th conference on Visualization '96*, San Francisco, California, United States, 1996, pp. 189–195.
- [52] W. Lorensen and H. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *ACM Siggraph Computer Graphics*, vol. 21, no. 4, 1987, pp. 163–169.

- [53] K. Lowell and P. Casault, "Outside-in, inside-out: Two methods of generating spatial certainty maps," *Proceedings of GeoComputation 97 and SIRC*, University of Otago, New Zealand, 1997, pp. 26–29.
- [54] C. Lundstrom, P. Ljung, A. Persson, and A. Ynnerman, "Uncertainty visualization in medical volume rendering using probabilistic animation," *IEEE Transactions on Visualization and Computer Graphics*, 2007, pp. 1648–1655.
- [55] A. MacEachren, "Visualizing uncertain information," *Cartographic Perspective*, vol. 13, no. 3, 1992, pp. 10–19.
- [56] A. MacEachren, *How maps work*, Guilford Press, New York, 1995.
- [57] A. MacEachren, C. Brewer, and L. Pickle, "Visualizing georeferenced data: representing reliability of health statistics," *Environment and Planning*, vol. 30, 1998, pp. 1547–1562.
- [58] J. P. Martin, J. E. Swan, R. J. Moorhead, Z. Liu, and S. Cai, "Results of a User Study on 2D Hurricane Visualization," *Computer Graphics Forum*, vol. 27, no. 3, 2008, pp. 991–998.
- [59] C. Marzban, S. Sandgathe, H. Lyons, and N. Lederer, "Three Spatial Verification Techniques: Cluster Analysis, Variogram, and Optical Flow," *Weather and Forecasting*, vol. 24, no. 6, 2009, pp. 1457–1471.
- [60] G. Mauris, V. Lasserre, and L. Foulloy, "A fuzzy approach for the expression of uncertainty in measurement," *Measurement*, vol. 29, no. 3, 2001, pp. 165–177.
- [61] S. Maxwell and H. Delaney, *Designing experiments and analyzing data: A model comparison perspective*, vol. 1, Lawrence Erlbaum, 2004.
- [62] Mayfield, M., "NHC-TPC: Hurricane Fran 23 August - 8 September 1996," <http://www.nhc.noaa.gov/1996fran.html>, current September 2011.
- [63] B. McCormick, "Visualization in scientific computing," *ACM SIGBIO Newsletter*, vol. 10, no. 1, 1988, pp. 15–21.
- [64] J. McEnery, J. Ingram, Q. Duan, T. Adams, and L. Anderson, "NOAA's Advanced Hydrologic Prediction Service," *Bulletin of the American Meteorological Society*, vol. 86, no. 3, 2005, pp. 375–85.
- [65] A. Mercer, C. Shafer, C. Doswell III, L. Leslie, and M. Richman, "Objective Classification of Tornadoic and Nontornadoic Severe Weather Outbreaks," *Monthly Weather Review*, vol. 137, 2009, pp. 4355–4368.
- [66] F. Mesinger, G. DiMego, E. Kalnay, K. Mitchell, P. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E. Berbery, et al., "North American Regional Reanalysis," *Bulletin of the American Meteorological Society*, vol. 87, no. 3, 2006, pp. 343–360.

- [67] J. Michalakes, J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, and W. Wang, "The weather research and forecast model: Software architecture and performance," *Proceedings of the 11th ECMWF Workshop on the Use of High Performance Computing In Meteorology*, 2004, pp. 156–168.
- [68] F. Molteni, R. Buizza, T. Palmer, and T. Petroliaigis, "The ECMWF ensemble prediction system: Methodology and validation," *Quarterly journal of the Royal Meteorological Society*, vol. 122, 1996, pp. 73–119.
- [69] D. Murray, J. McWhirter, S. Wier, and S. Emmerson, "The integrated data viewer-a web-enabled application for scientific analysis and visualization," *Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, 2003.
- [70] National Research Council, Committee on FEMA Flood Maps and United States, Federal Emergency Management Agency, and United States National Oceanic and Atmospheric Administration, *Mapping the Zone: Improving Flood Map Accuracy*, National Academy Press, 2009.
- [71] H. NCSA, "User's Guide HDF 4.2," 2003.
- [72] C. Olston and J. Mackinlay, "Visualizing data with bounded uncertainty," *IEEE Symposium on Information Visualization*, 2002, pp. 37–40.
- [73] A. Pang, "Visualizing uncertainty in geo-spatial data," *Proceedings of the Workshop on the Intersections between Geospatial Information and Information Technology*, 2001.
- [74] A. Pang, C. Wittenbrink, and S. Lodha, "Approaches to uncertainty visualization," *The Visual Computer*, vol. 13, no. 8, 1997, pp. 370–390.
- [75] Pasch, R. J., "NHC-TPC: Hurricane Edouard 19 August - 3 September 1996," <http://www.nhc.noaa.gov/1996edouard.html>, current September 2011.
- [76] K. Potter, J. Kniss, R. Riesenfeld, and C. Johnson, "Visualizing Summary Statistics and Uncertainty," *Computer Graphics Forum*, vol. 29, no. 3, 2010, pp. 823–832.
- [77] K. Potter, A. Wilson, P. Bremer, D. Williams, C. Doutriaux, V. Pascucci, and C. R. Johnson, "Ensemble-Vis: A Framework for the Statistical Visualization of Ensemble Data," *ICDMW '09: Proceedings of the 2009 IEEE International Conference on Data Mining Workshops*, Washington, District of Columbia, United States, 2009, pp. 233–240.
- [78] R. Rew, E. Hartnett, and J. Caron, "NetCDF-4: Software implementing an enhanced data model for the geosciences," *22nd International Conference on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, 2006.

- [79] P. Rheingans and S. Joshi, "Visualization of Molecules with Positional Uncertainty," *Data Visualization '99, Proceedings of the Joint EUROGRAPHICS - IEEE TVCG Symposium on Visualization*, 1999, pp. 299–306.
- [80] K. Riley, D. Ebert, C. Hansen, and J. Levit, "Visually accurate multi-field weather visualization," *IEEE Visualization*, 2003, pp. 279–286.
- [81] J. Sanyal, S. Zhang, G. Bhattacharya, P. Amburn, and R. Moorhead, "A User Study to Compare Four Uncertainty Visualization Methods for 1D and 2D Datasets," *IEEE Transactions on Visualization and Computer Graphics*, 2009, pp. 1209–1218.
- [82] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, and R. Moorhead, "Noodles: A Tool for Visualization of Numerical Weather Model Ensemble Uncertainty," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, 2010, pp. 1421–1430.
- [83] G. Schmidt, S. Chen, A. Bryden, M. Livingston, L. Rosenblum, and B. Osborn, "Multidimensional visual representations for underwater environmental uncertainty," *IEEE Computer Graphics and Applications*, 2004, pp. 56–65.
- [84] W. Schroeder, K. Martin, and B. Lorensen, *The visualization toolkit*, Prentice Hall PTR, 1998.
- [85] D. Schultz, W. Bracken, L. Bosart, G. Hakim, M. Bedrick, M. Dickinson, and K. Tyle, "The 1993 superstorm cold surge: Frontal structure, gap flow, and tropical impact," *Monthly Weather Review*, vol. 125, no. 1, 1997, pp. 5–39.
- [86] D. Schweizer and M. Goodchild, "Data quality and choropleth maps: An experiment with the use of color," *GIS/LIS. American Society for Photogrammetry and Remote Sensing*, 1992, vol. 2, pp. 686–699.
- [87] J. Shapiro, "George H. Heilmeyer," *IEEE Spectrum*, vol. 31, no. 6, 1994, pp. 56–59.
- [88] W. Skamarock, *A description of the Advanced Research WRF version 3*, Tech. Rep., National Center for Atmospheric Research, 2008.
- [89] W. Skamarock, J. Klemp, J. Dudhia, D. Gill, D. Barker, W. Wang, and J. Powers, *A description of the Advanced Research WRF version 2*, Tech. Rep., National Center for Atmospheric Research, 2005.
- [90] A. Squillacote and J. Ahrens, *The paraview guide*, Kitware, 2006.
- [91] J. Stackpole, "GRIB, The WMO Format for the Storage of Weather Product Information and the Exchange of Weather Product Messages in Gridded Binary Form," *National Meteorological, National Weather Service*, 1989.

- [92] C. Steed, P. Fitzpatrick, T. Jankun-Kelly, A. Yancey, and J. Swan, "An interactive parallel coordinates technique applied to a tropical cyclone climate analysis," *Computers & Geosciences*, vol. 35, no. 7, 2009, pp. 1529–1539.
- [93] T. Strothotte, M. Masuch, and T. Isenberg, "Visualizing knowledge about virtual reconstructions of ancient architecture," *Proceedings of CGI '99*, 1999, pp. 1530–1052.
- [94] B. Taylor and C. Kuyatt, *Guidelines for evaluating and expressing the uncertainty of NIST measurement results*, NIST Tech. Note 1297, U.S. Government Printing Office, Washington, District of Columbia, 1994.
- [95] The MathWorks, "Matlab," *Natick Inc.*, 2009.
- [96] J. Thomson, E. Hetzler, A. MacEachren, M. Gahegan, and M. Pavel, "A typology for visualizing uncertainty," *Proceedings of SPIE*, 2005, vol. 5669, pp. 146–157.
- [97] M. Tory and T. Möller, "Rethinking visualization: A high-level taxonomy," *IEEE Symposium on Information Visualization*, 2004, pp. 151–158.
- [98] L. A. Treinish, "Multi-resolution visualization techniques for nested weather models," *Proceedings of the conference on Visualization '00*, Salt Lake City, Utah, United States, 2000, pp. 513–516.
- [99] E. Tufte, *The visual display of quantitative information*, Graphics Press, Cheshire, Connecticut, 1983.
- [100] E. Tufte, *Envisioning Information*, Graphics Press, Cheshire, Connecticut, 1990.
- [101] E. Tufte, *Visual explanations: images and quantities, evidence and narrative*, Graphics Press, Cheshire, Connecticut, 1997.
- [102] L. Uccellini, P. Kocin, R. Schneider, P. Stokols, and R. Dorr, "Forecasting the 12-14 March 1993 superstorm," *Bulletin of the American Meteorological Society*, vol. 76, no. 2, 1995, pp. 183–200.
- [103] C. Ware, *Information Visualization: Perception for Design*, second edition, Morgan Kaufmann Publishers, 2004.
- [104] F. Warmerdam, "Shapefile c library v1. 2," *Web site: <http://shapelib.maptools.org/>*, current October 2011.
- [105] S. Wechsler, "Perceptions of digital elevation model uncertainty by DEM users," *URISA Journal*, vol. 15, no. 2, 2003, pp. 57–64.
- [106] D. Williams, R. Drach, P. Dubois, C. Doutriaux, C. OConnor, K. AchutaRao, and M. Fiorino, "Climate data analysis tool: An open software system approach," *13th Symposium on Global Change and Climate Variations*, 2002.

- [107] L. J. Wilson, W. R. Burrows, and A. Lanzinger, "A Strategy for Verification of Weather Element Forecasts from an Ensemble Prediction System," *Monthly Weather Review*, vol. 127, no. 6, 1999, pp. 956–970.
- [108] C. Wittenbrink, A. Pang, and S. Lodha, "Glyphs for visualizing uncertainty in vector fields," *IEEE Transactions on Visualization and Computer Graphics*, vol. 2, no. 3, 1996, pp. 266–279.
- [109] T. Zuk and S. Carpendale, "Theoretical analysis of uncertainty visualizations," *Proceedings of SPIE*, 2006, vol. 6060, pp. 66–79.