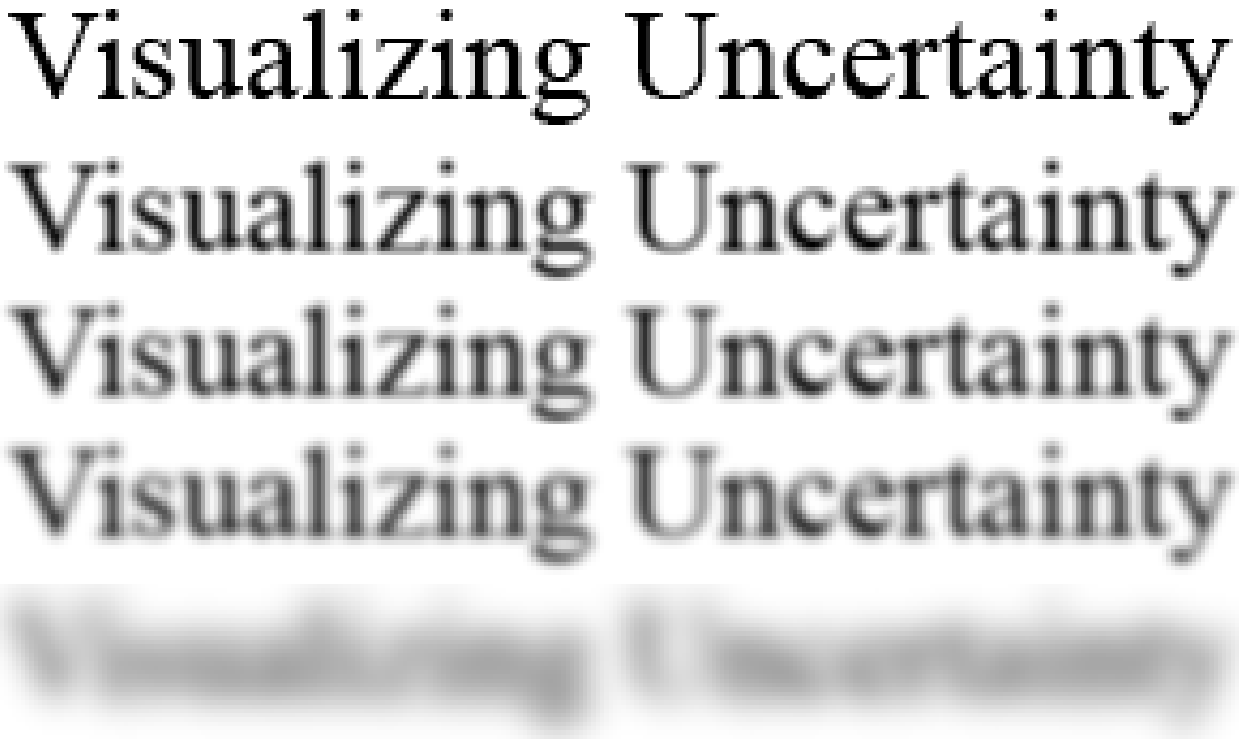
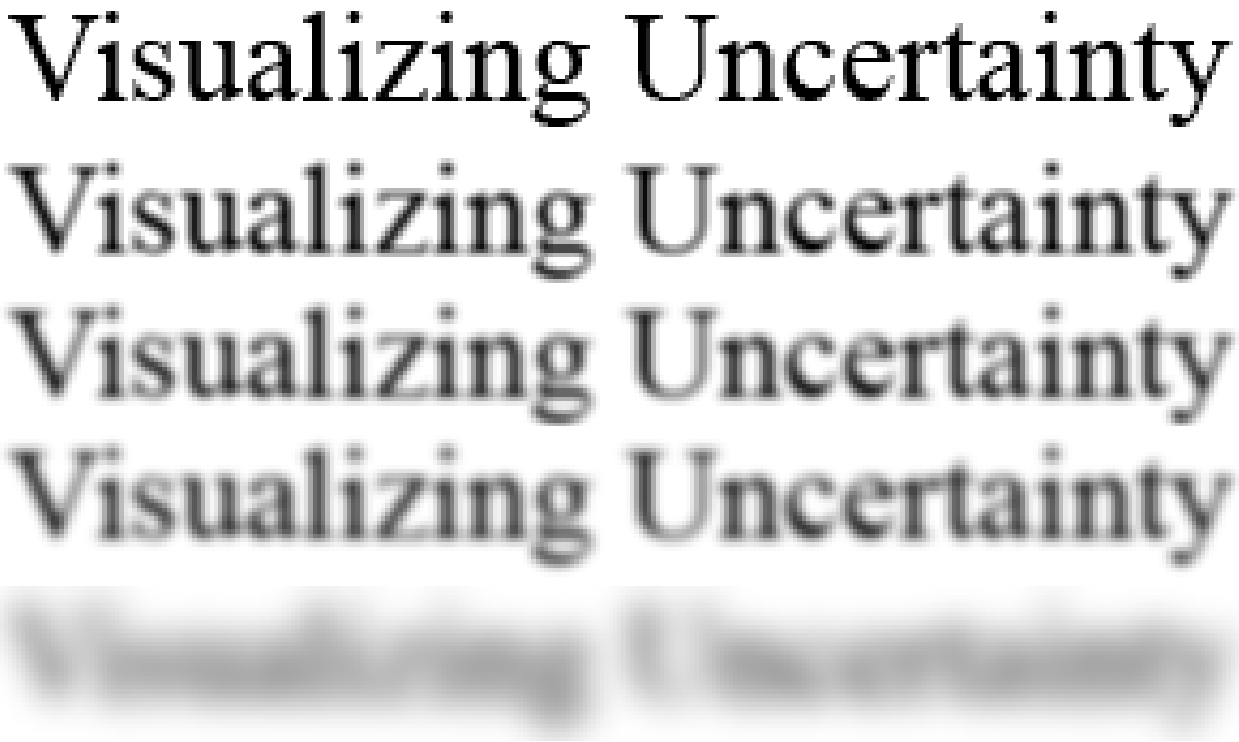
Visualizing Uncertainty: Computer Science Perspective

Ben Shneiderman, Univ of Maryland, College Park
Alex Pang, Univ of California, Santa Cruz

National Academy of Sciences Workshop
March 3-4, 2005, Washington, DC


Visualizing Uncertainty
Visualizing Uncertainty
Visualizing Uncertainty
Visualizing Uncertainty



What do we mean by uncertainty?

Why is this an issue now?

Sources of uncertainty

- Measurement problems
- Ranges/Summaries
- Missing data
- Human ratings
- Potential deceptions
- Privacy protection
- Risk assessments
- Forecasts

- Scientific data
- Intelligence sources
- Statistical analyses
- Medical images
- Gene expression
- Simulations
- Financial models
- Weather
- Consumer ratings

Visualizations

- Text, statistical measures…

SciViz

- 1D: Lists, documents, numeric ranges…
- 2D: Geographic Info
- 3D: Scientific Visualization

InfoViz

- Multi-Variate: Information Visualization
- Temporal: Patient histories, web logs …
- Tree: Taxonomies, org charts, directories …
- Network: Social, communication…

Text: Statistical uncertainty

- Text

- 1D
- 2D
- 3D

- Multi-V
- Temporal
- Tree
- Network

Poll	Dems	Reps
Bush	14%	81%
Kerry	79%	7%
Margin of error +/- 3%		

	Frustration (N=372)
Time Variables	
Time Lost (Incident)	**.293
Time to Fix (Incident)	**.233
Computer Years	-.041
Hours per Week	*-.124
* = p<.05 ** = p<.01	

	Rain inches	Reliability
SW	1.0	low
SE	1.5	high
NW	1.2	low
NE	0.8	med

Risk/Danger vs. Trust/Validity/Security

- Text

- 1D
- 2D
- 3D

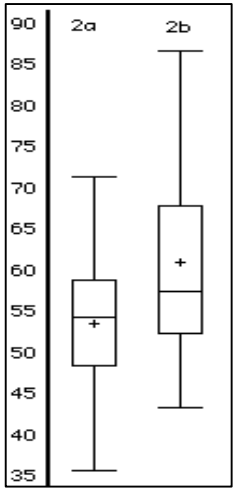
- Multi-V
- Temporal
- Tree
- Network

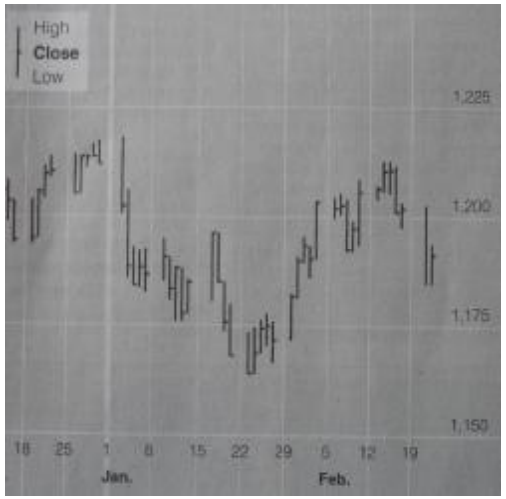
	Risk
Municipal Bonds	**low**
Blue Chip Stocks	**med**
Tech Stocks	**high**
Real Estate	**med**

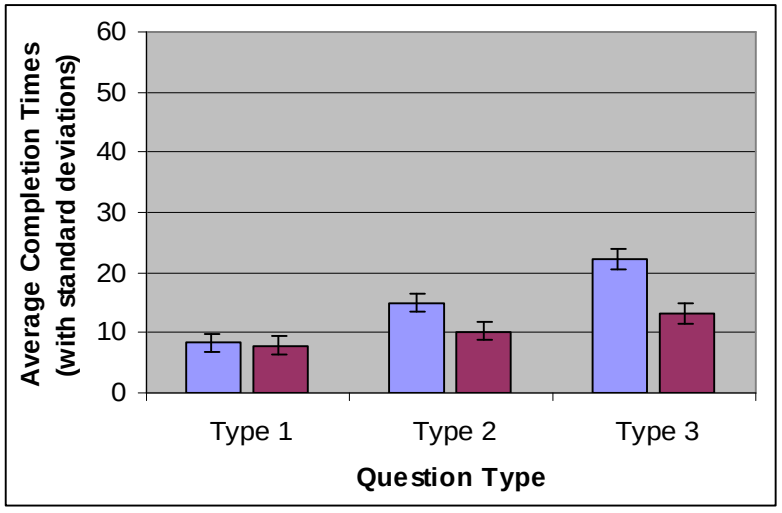
Highly Unlikely	Unlikely	About Even	Likely	Highly Likely

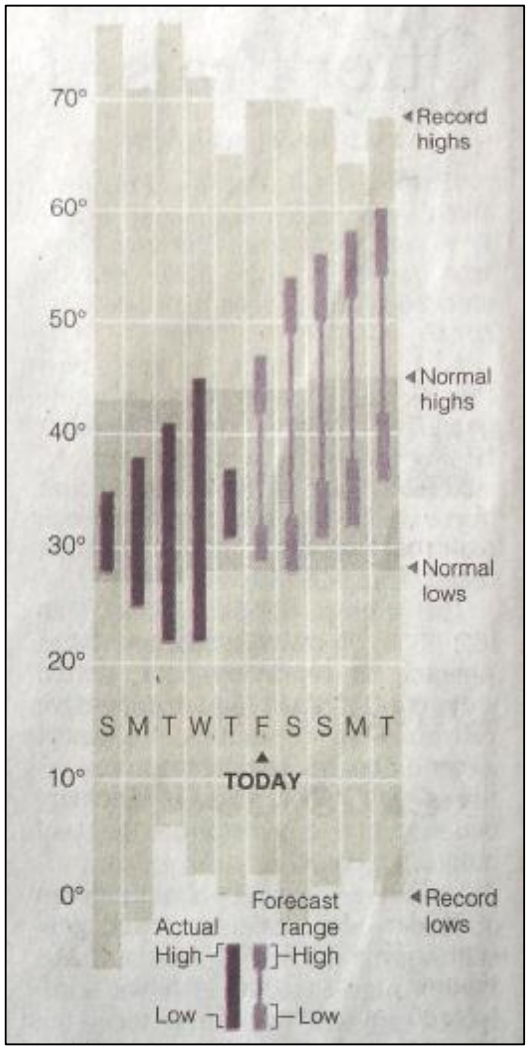
1D: Ranges, variations & forecasts

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network





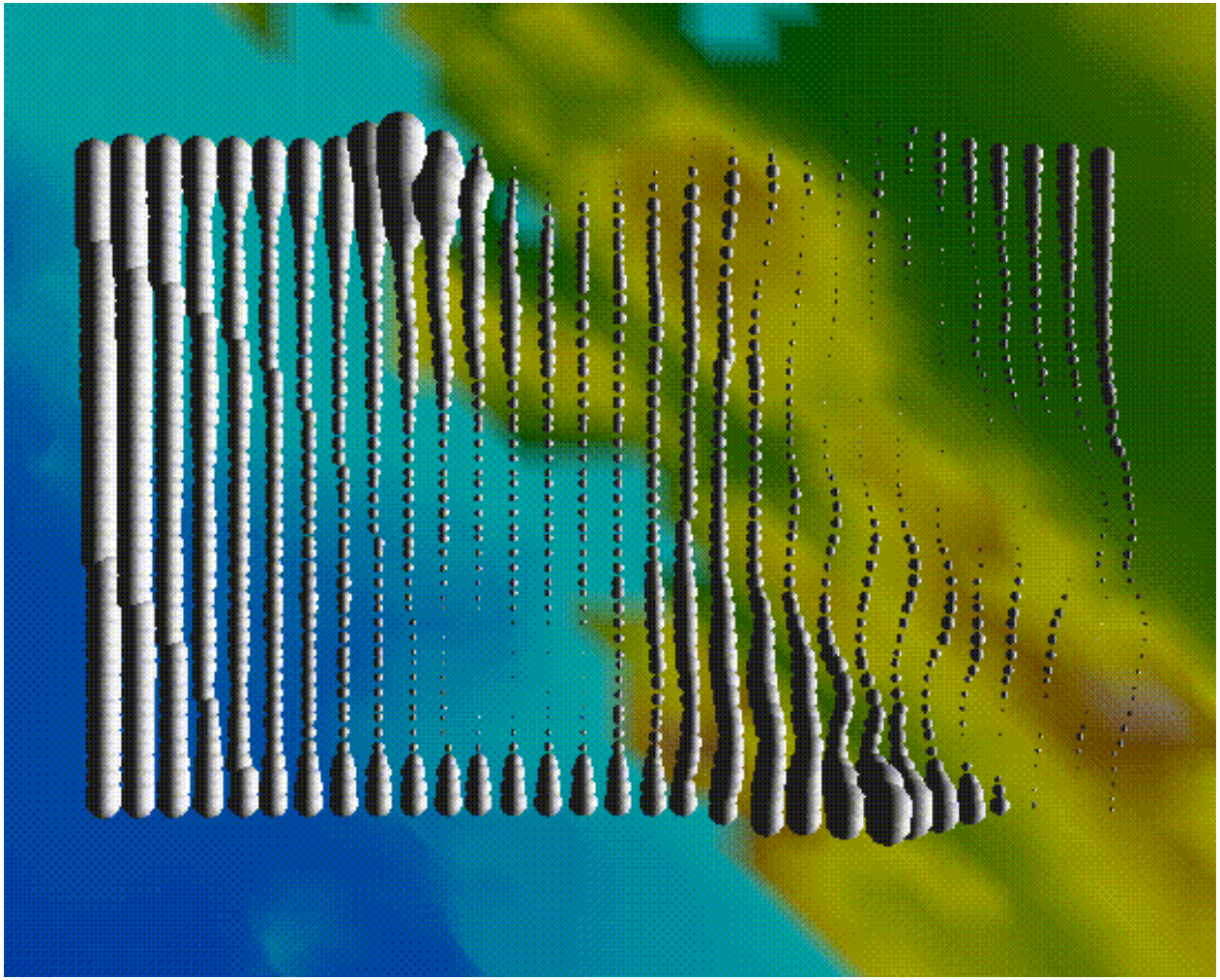




2D: Ball glyphs

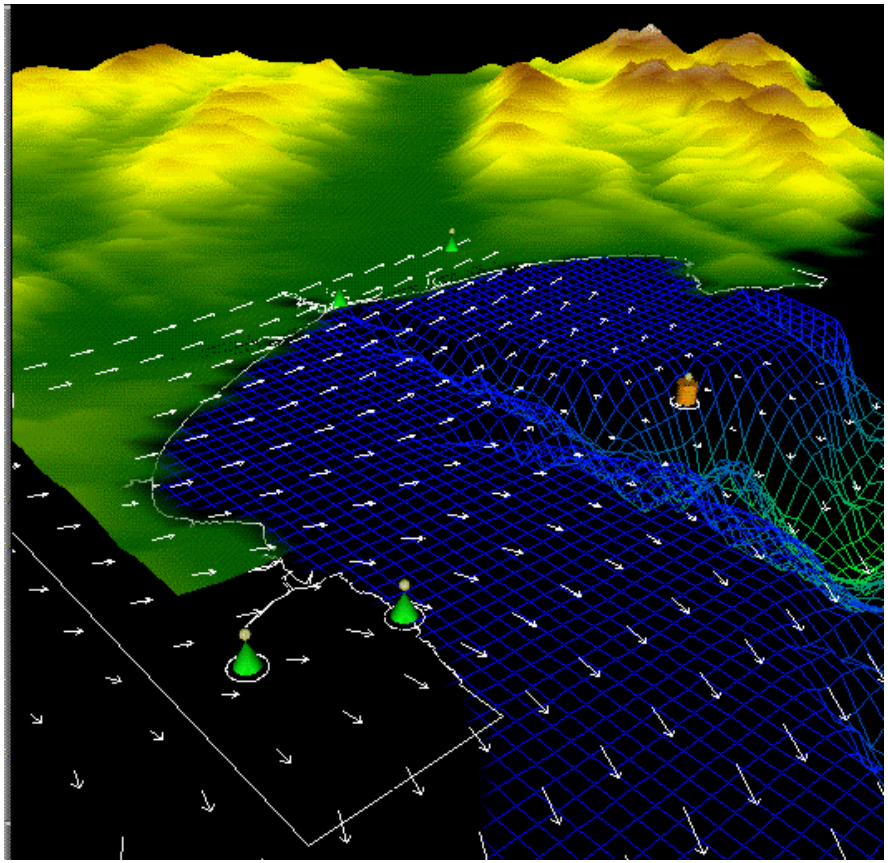
Delta (observation - forecast)

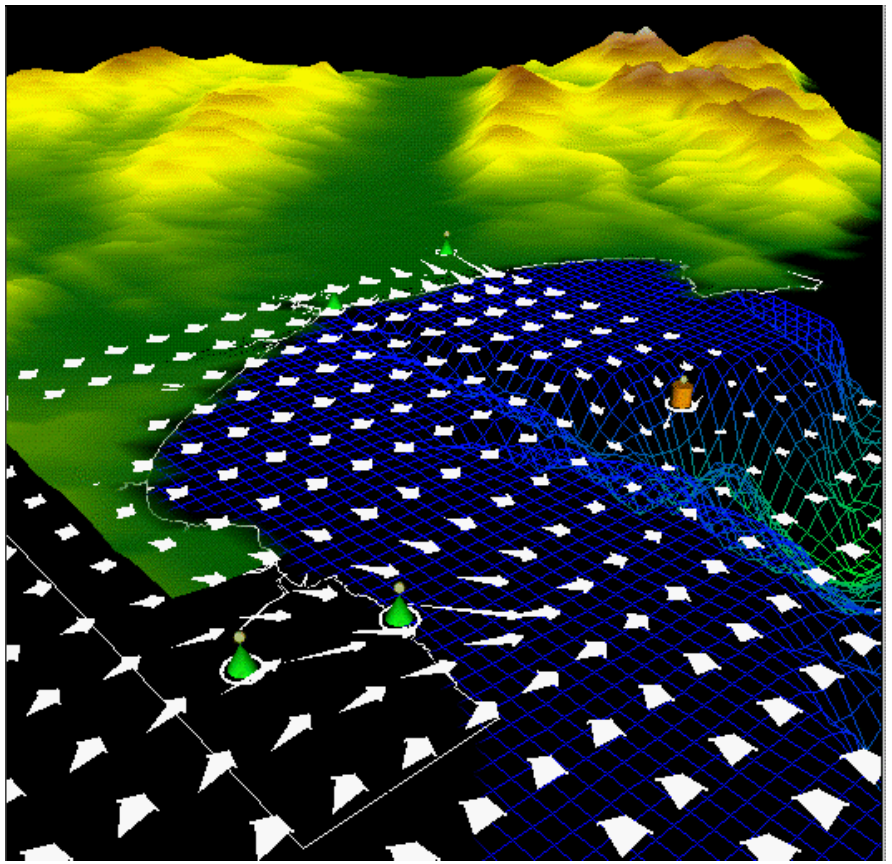
- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network



http://www.cse.ucsc.edu/research/slvg/assim.html

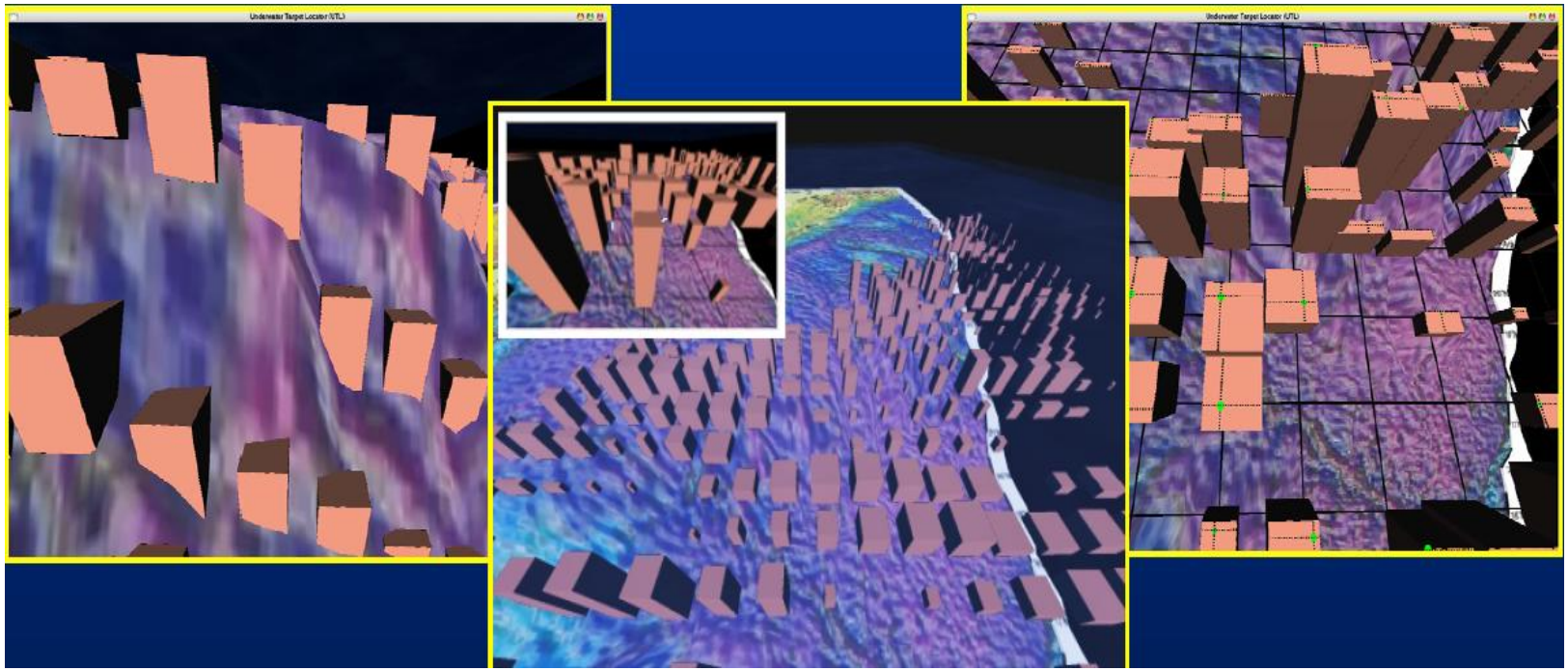
2D: Arrow glyphs (Direction & velocity)





http://www.cse.ucsc.edu/research/avis/unvis.html

2D: Box glyphs



Schmidt et al., 2004, Underwater Environmental Uncertainty, IEEE CG&A
http://csdl.computer.org/comp/mags/cg/2004/05/g5056abs.htm

2D: Grid with transparency/shading

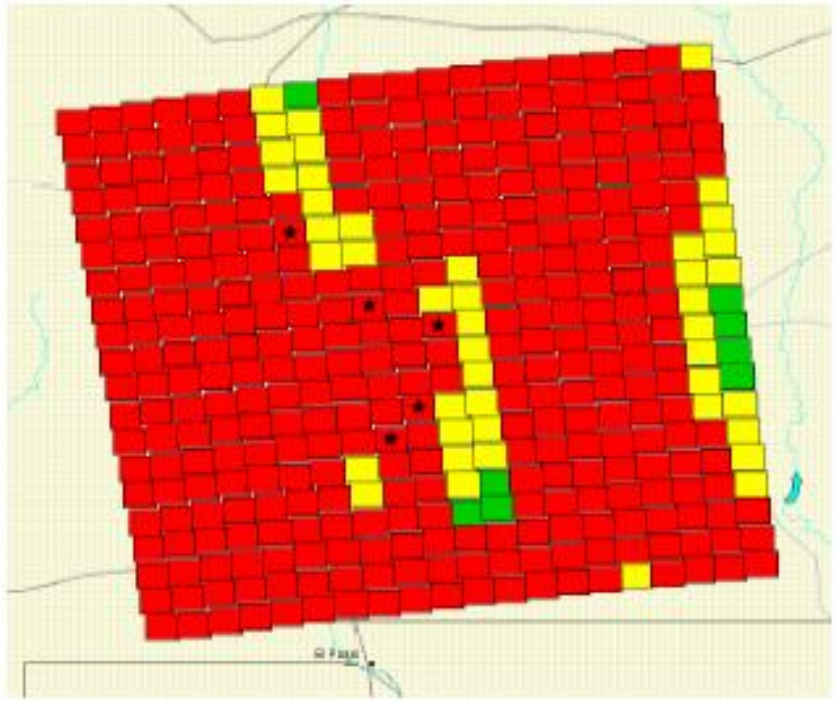


Figure 5: Current IWEDA-like spatial display showing the cloud ceiling at one particular time over the PUMAR domain. The black stars were points of interest during the T-UAV demonstration scenario. The red, yellow, and green "pixels" were assigned based on operational T-UAV cloud ceiling sensitivities.

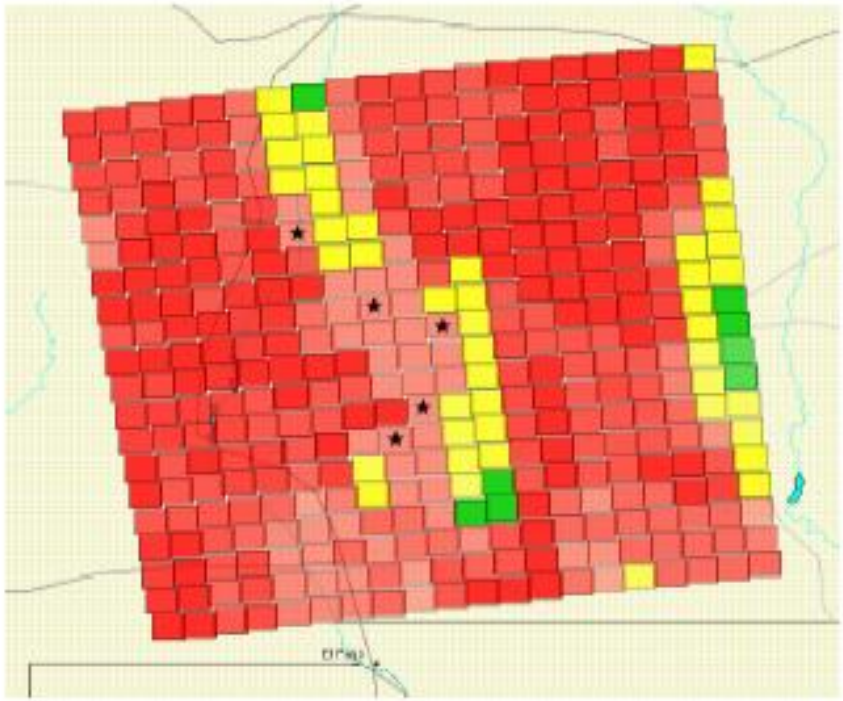
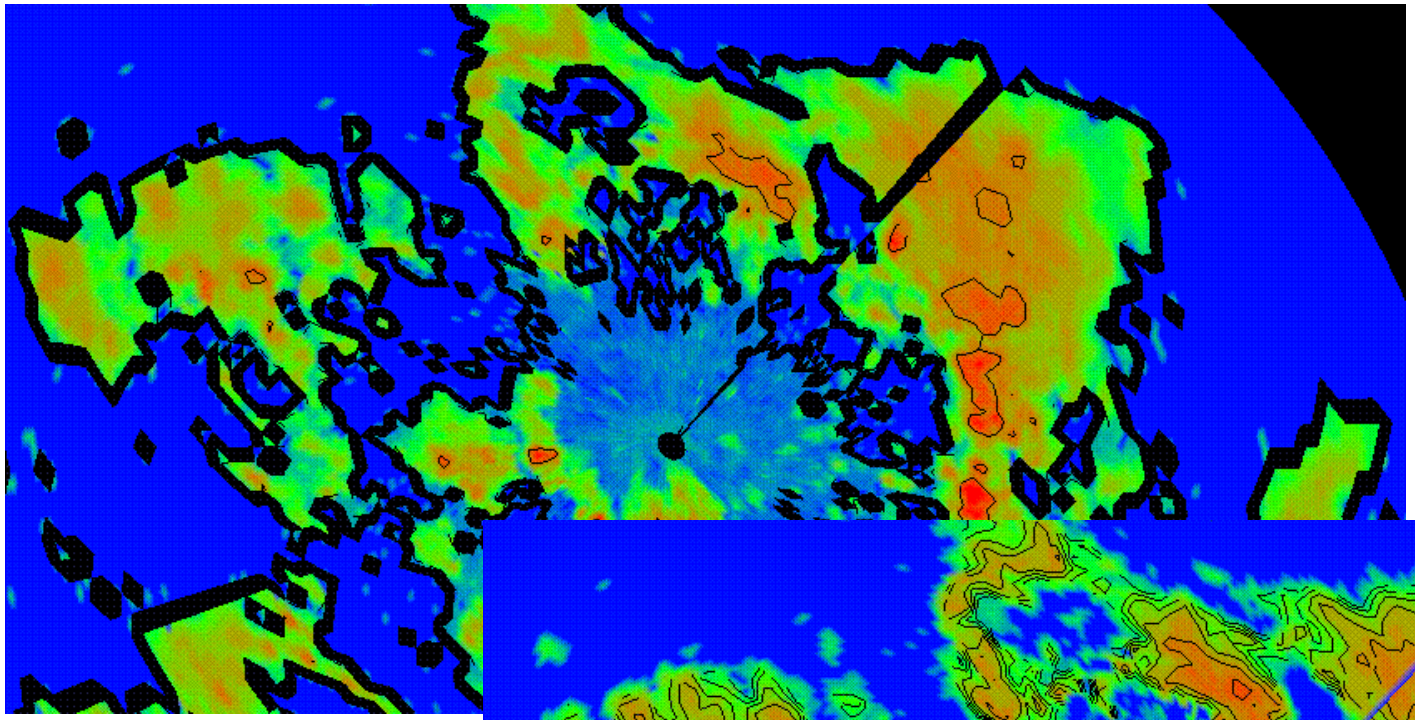


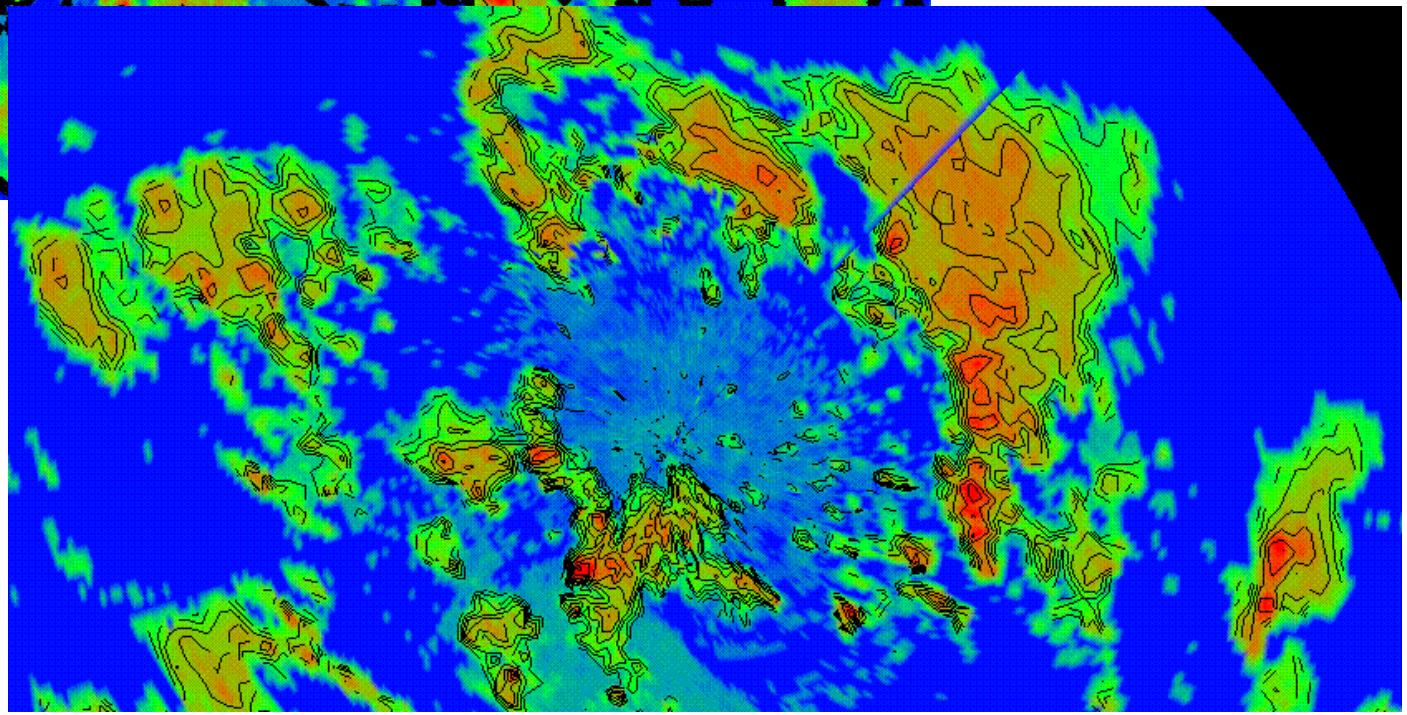
Figure 6: The same display as in Fig. 4, except for the application of cognitive engineering techniques to change the hue (color) and transparency of the cloud ceiling impacts based on the uncertainty of the predicted cloud ceiling.

Lefevre, Pfautz & Jones
http://ams.confex.com/ams/pdfpapers/82400.pdf

2D: Isolines with missing values

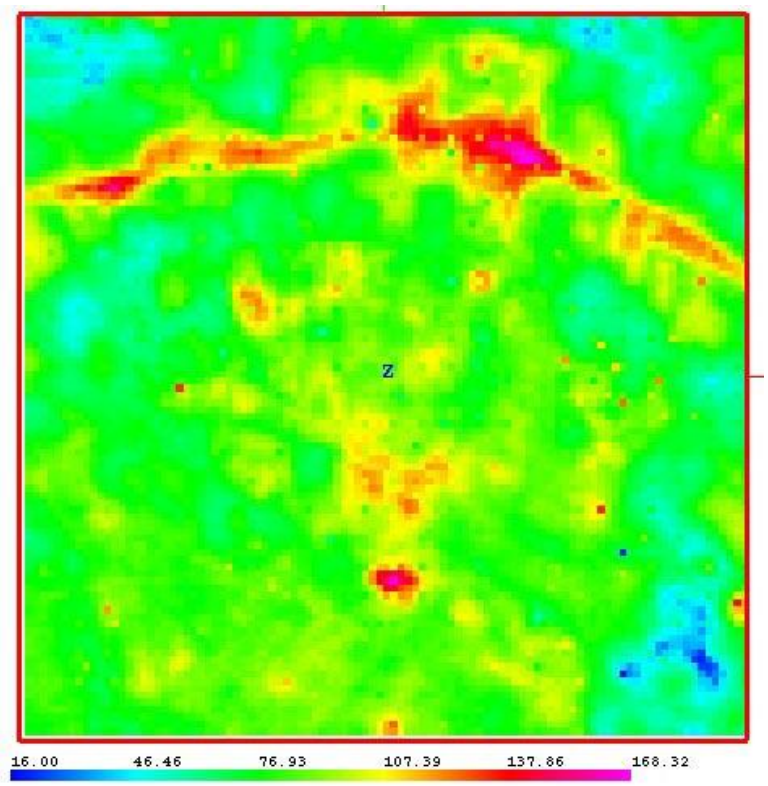


Off-the-shelf software can give incorrect contours on data with lots of missing values.



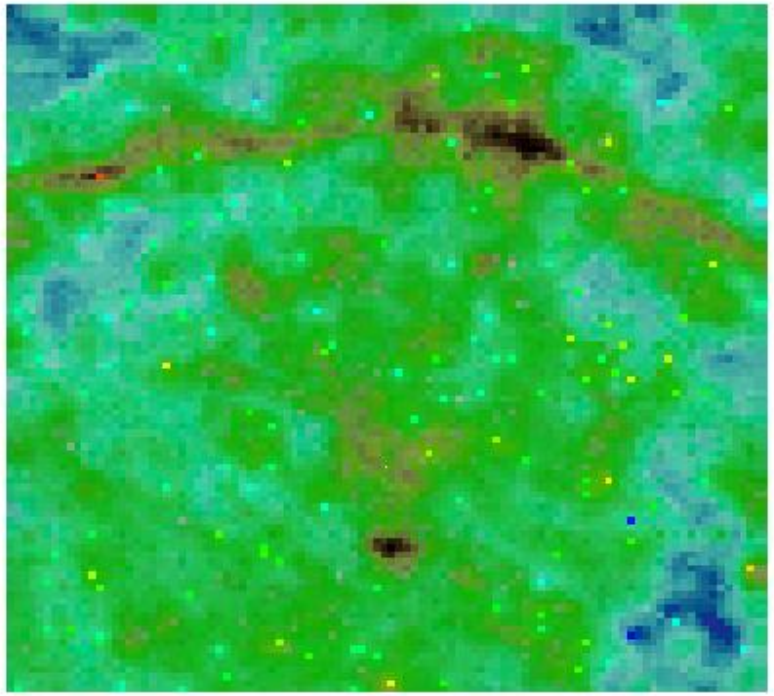
Modifications to contouring algorithm to account for large number of missing values.

2D: Pseudocolor shows mean values

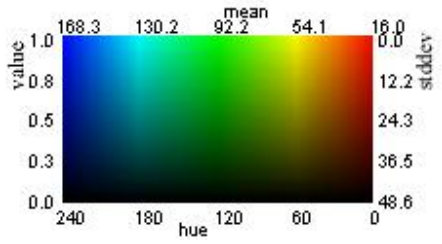


Luo, Kao & Pang, 2003, EuroVis
http://www.soe.ucsc.edu/~pang/op.pdf

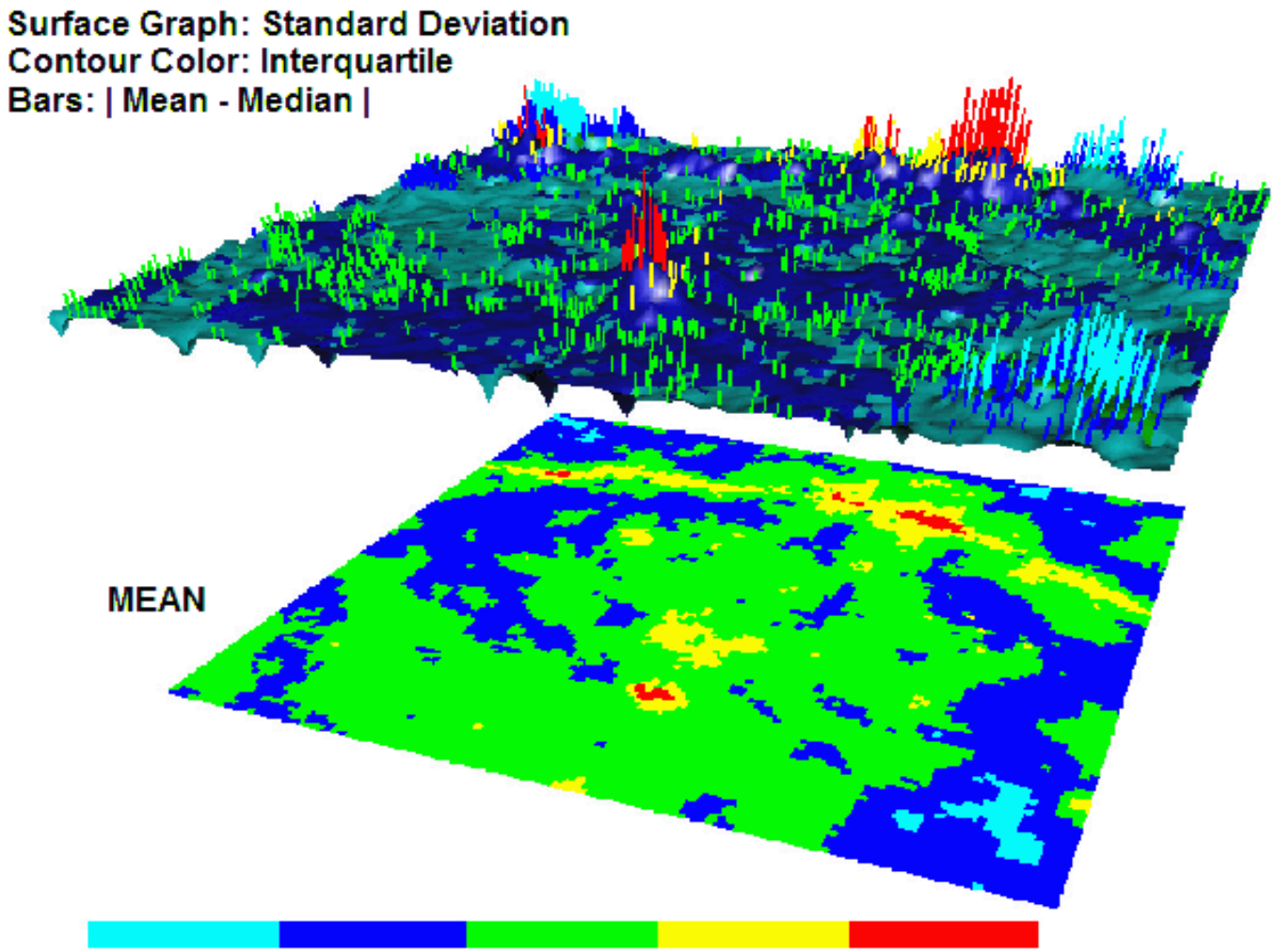
2D: Darkness = uncertainty (high stddev)



Mean = hue
Skew = 1/saturation
Stddev= 1/value

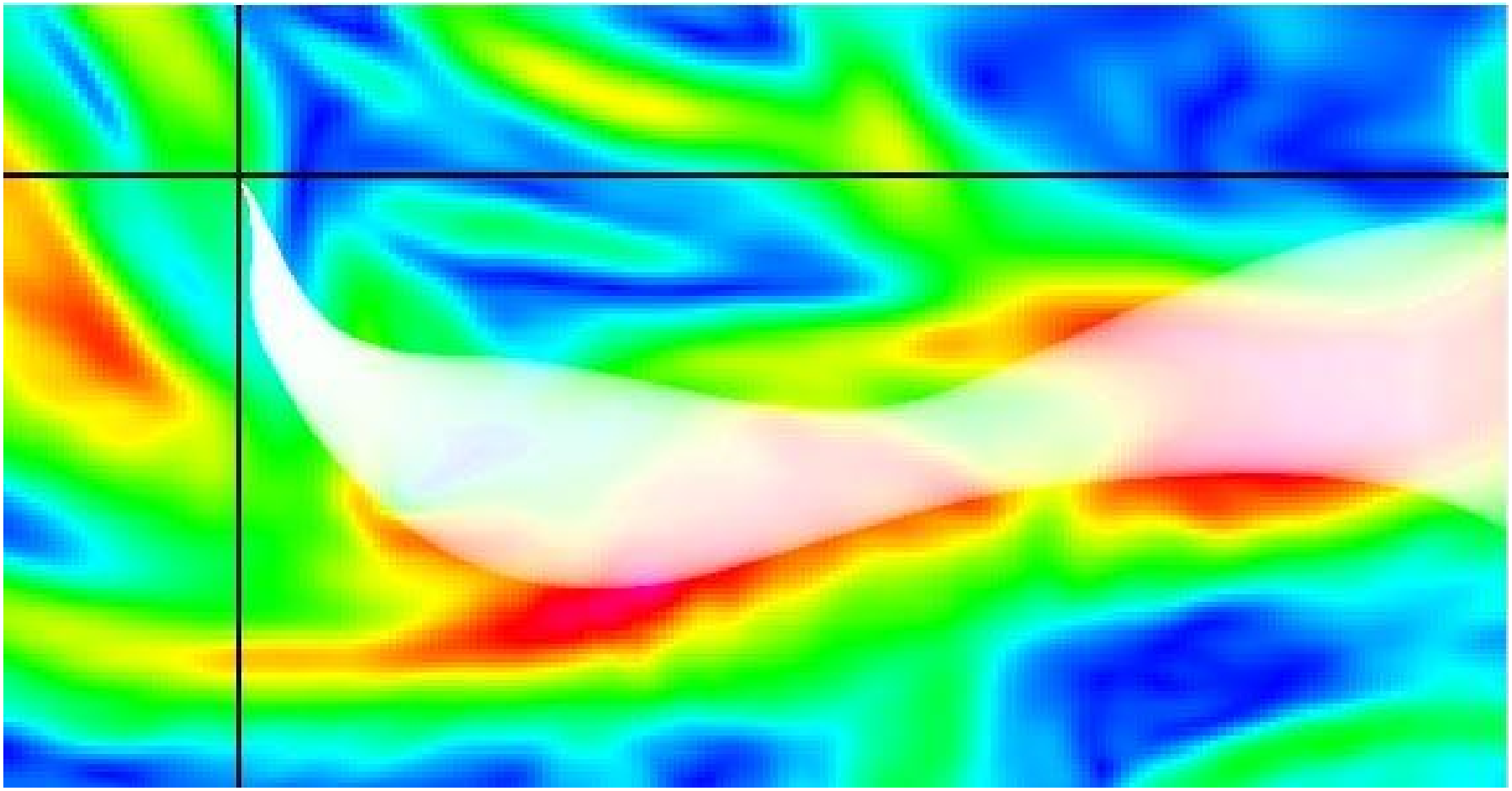


2D: Separate layer



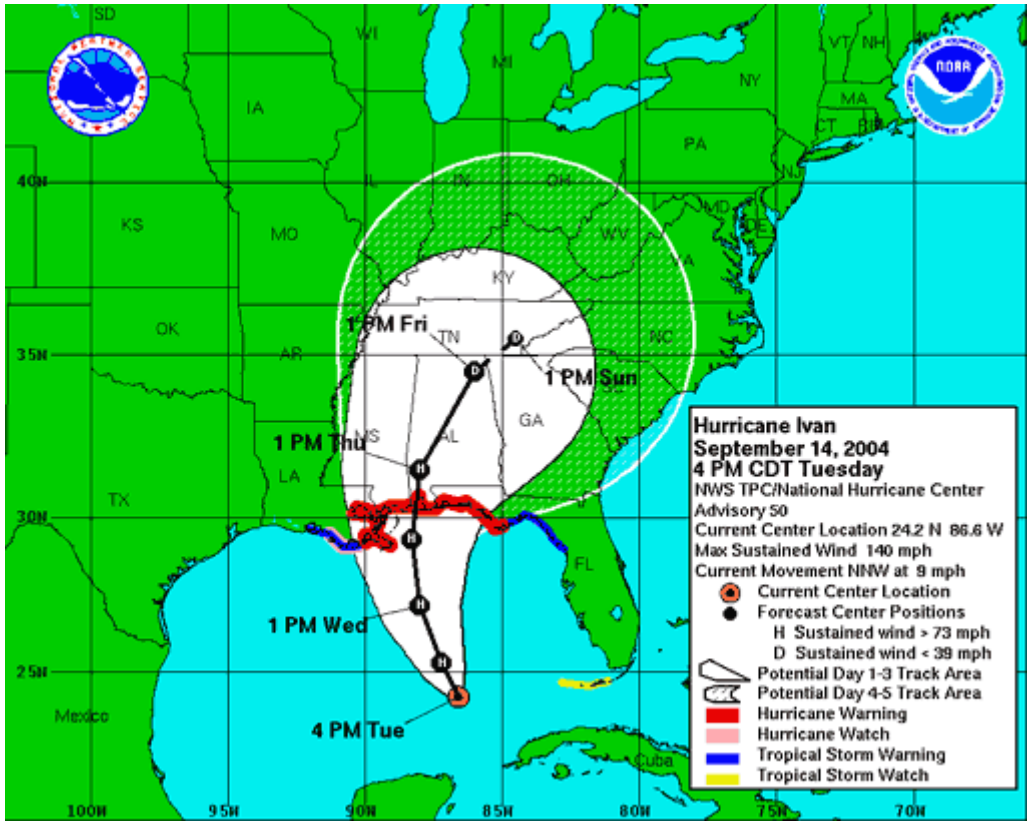
http://www.cse.ohio-state.edu/~bordoloi/Pubs/pdfCluster.pdf

2D: Streamlines with binwise +

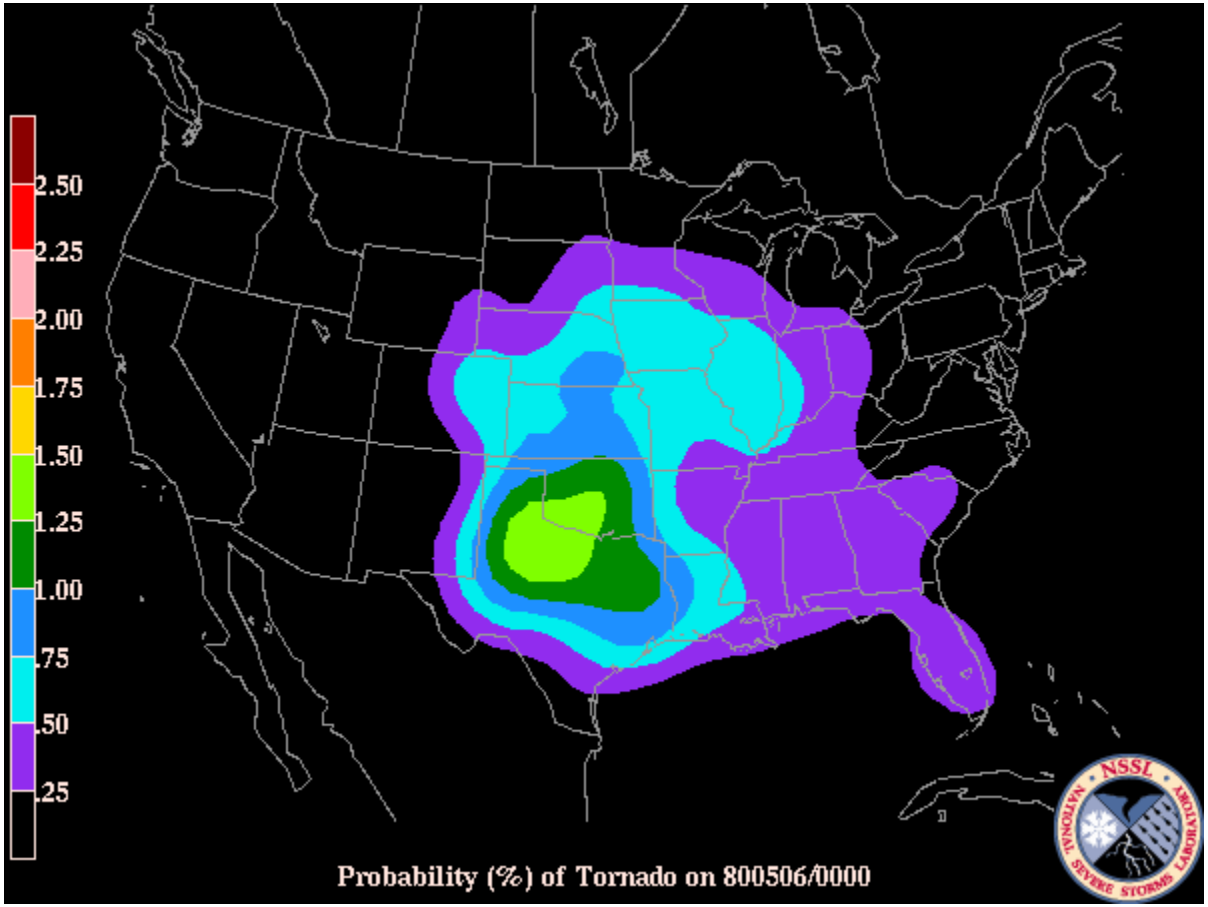


Luo, Kao & Pang, 2003
http://www.soe.ucsc.edu/~pang/op.pdf

2D: Weather forecast

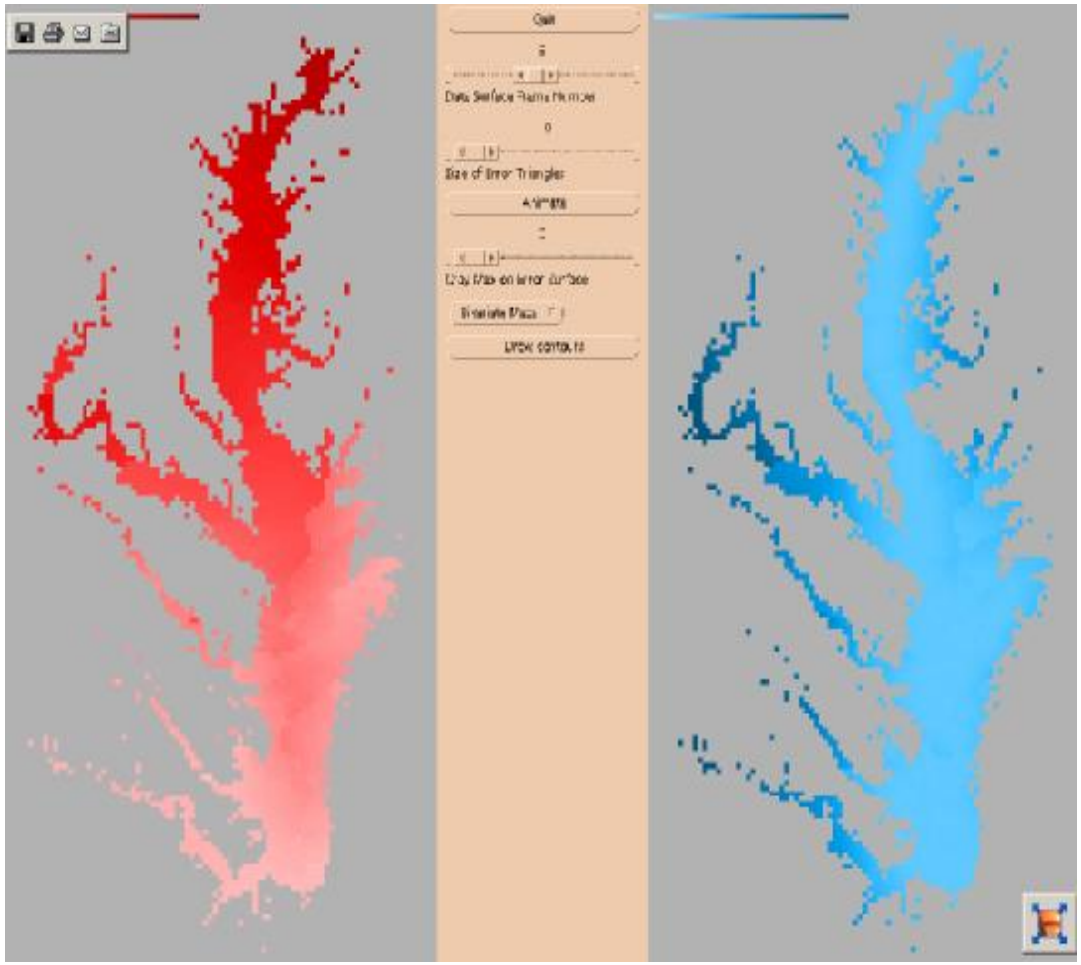


2D: NOAA Storm Prediction Center



http://www.spc.noaa.gov/products/

2D: Dual views & grid lines



Dark shows pollution Dark shows Certainty

(Howard & MacEachren, 1996)

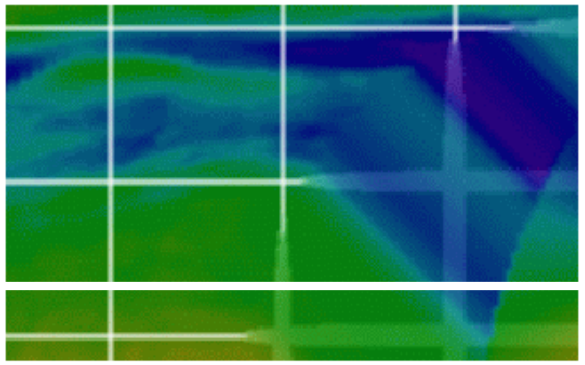


Figure 8: Closeup on the Width variation technique.

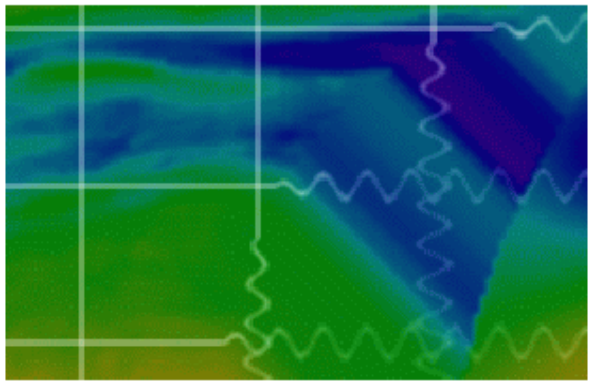
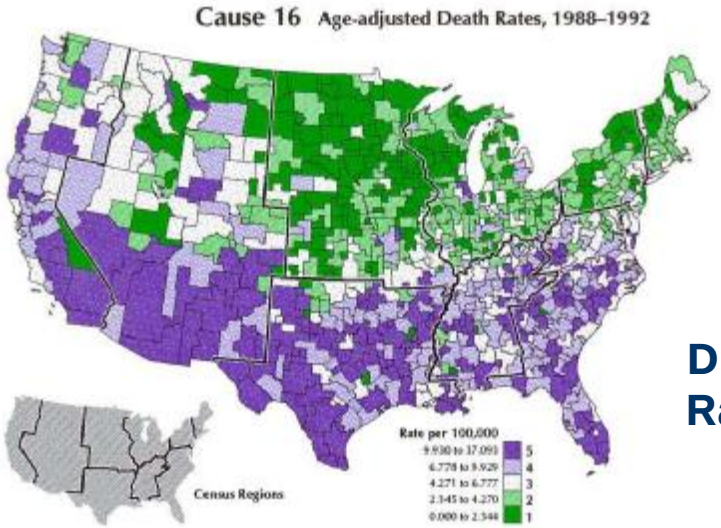


Figure 12: Closeup on the amplitude modulated distortion.

(Cedlink & Rhenigas, 2000)

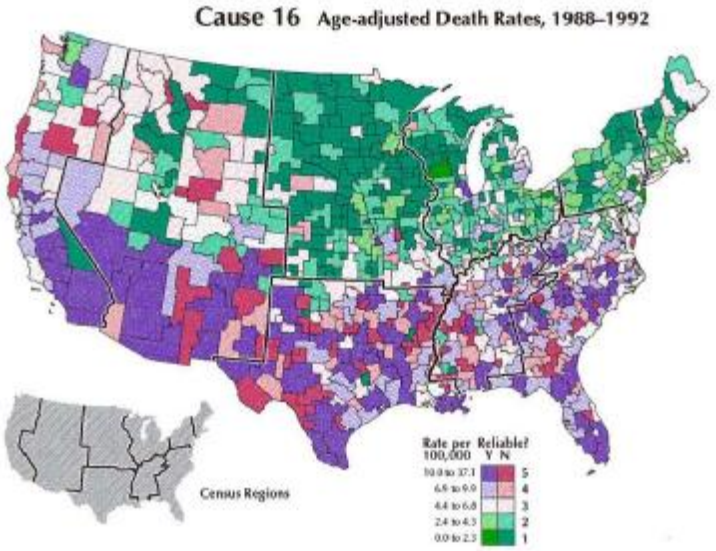


Dual maps for
Rate and reliability

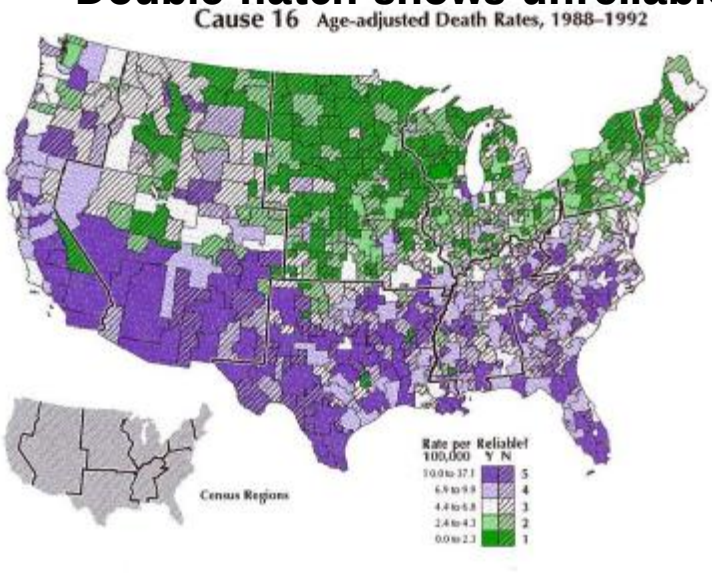
Sample Test Maps:

- cause: 16
- scale: quarter
- color scheme: purple-green
- reliability schemes:
  - map pairs
  - color change
  - texture overlay

Bivariate color scheme

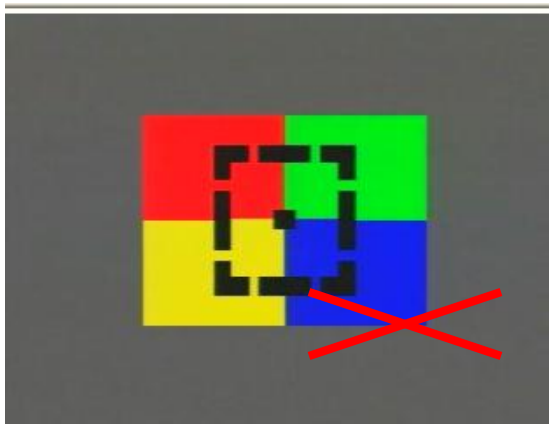


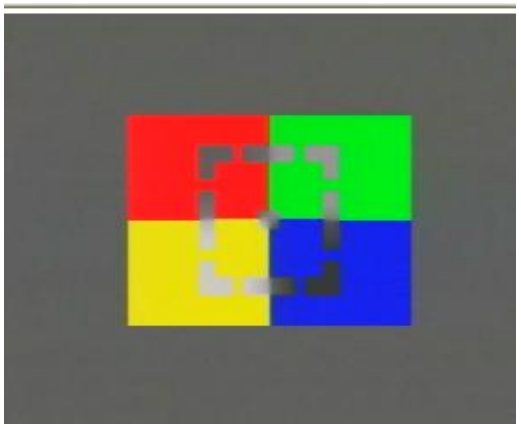
Double hatch shows unreliable

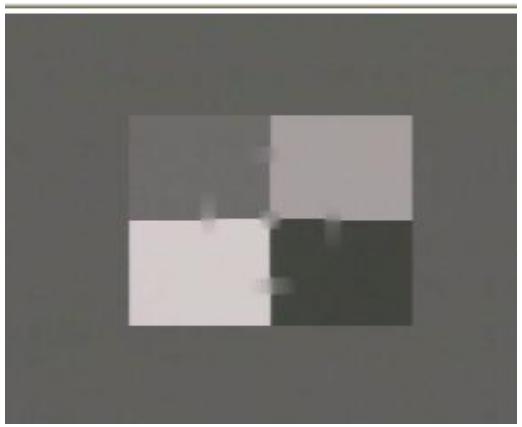


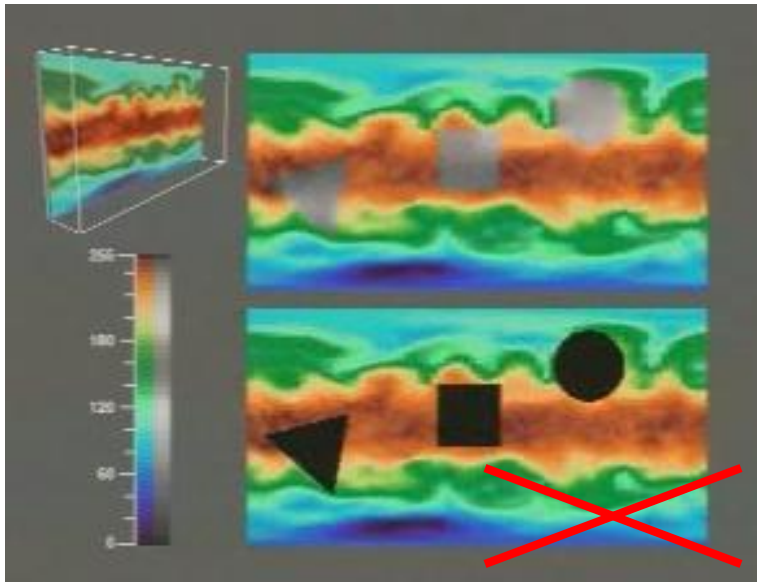
(MacEachren et al., 1998)

2D: Gray for missing & interpolation









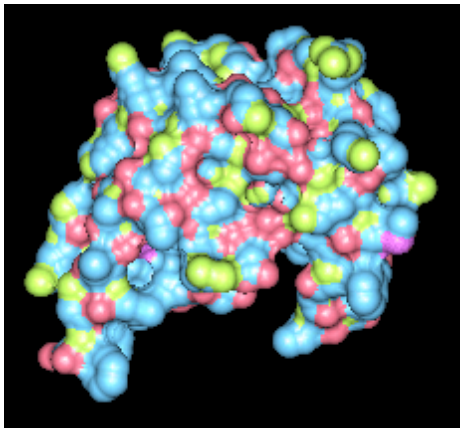
Gray shows missing & interpolated value, superior to using black only

Twiddy, R., Cavallo, J., and Shiri, S. 1994. Restorer*:* A visualization technique for handling missing data. *IEEE Visualization 94*, 212-216

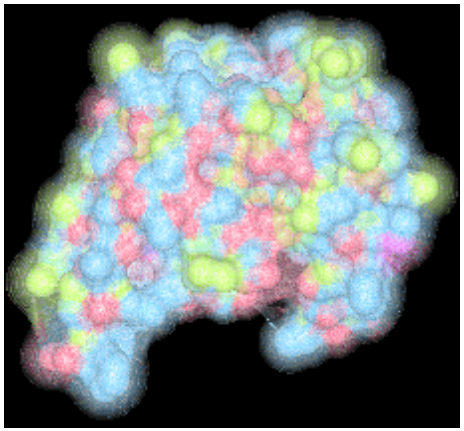
http://svs.gsfc.nasa.gov/vis/a000000/a000000/a000010/index.html

3D: Fuzzy molecular surface: HIV protease

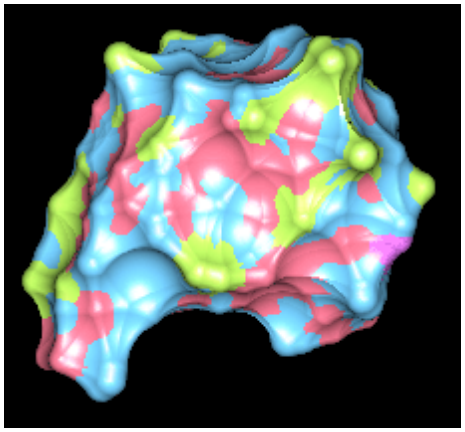
- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network



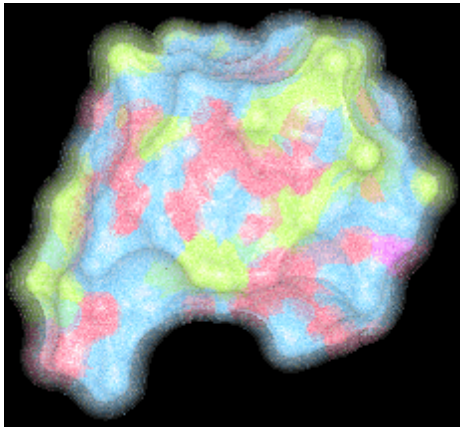
Crisp molecular surface
Probe radius 1.4



Fuzzy molecular surface
Probe radius 1.4



Crisp molecular surface
Probe radius 5.0

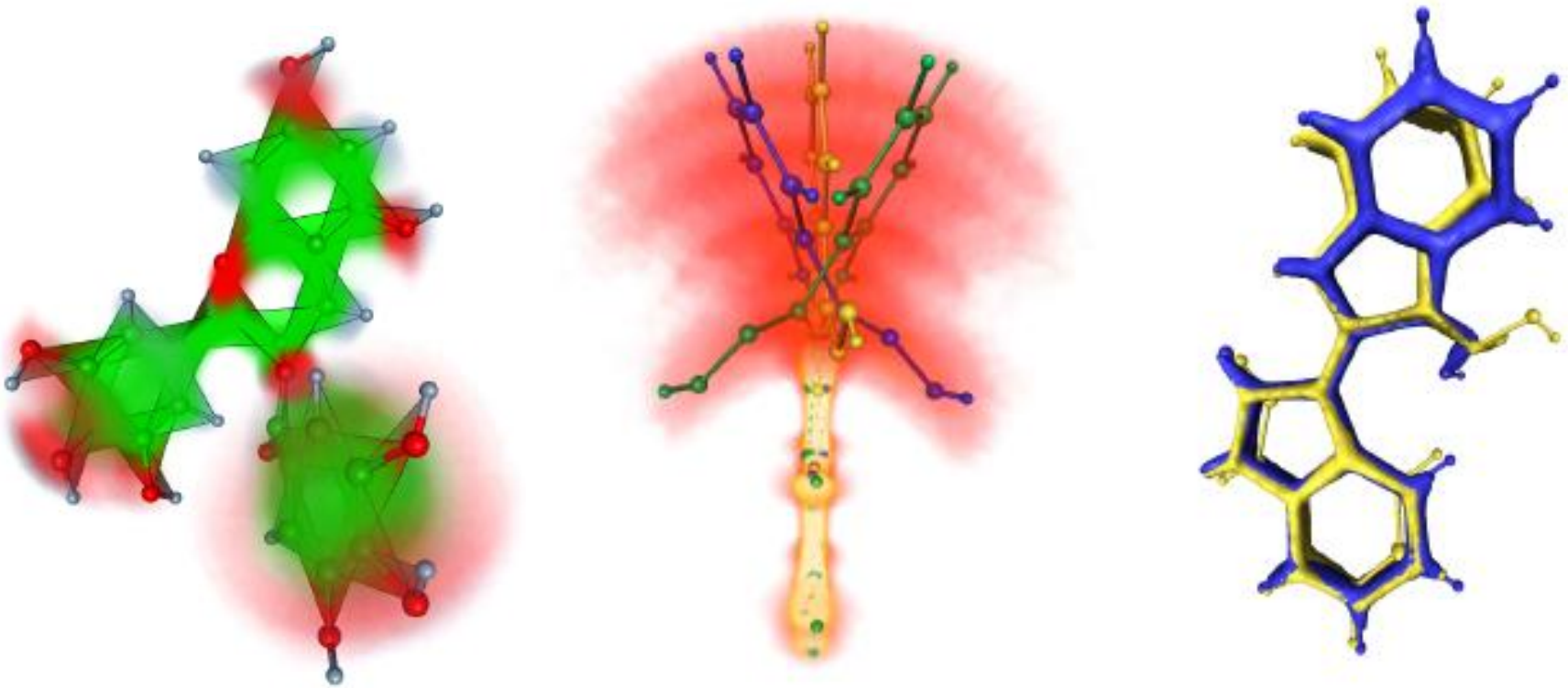


Fuzzy molecular surface
Probe radius 5.0

Lee & Varshney (2002), UM Graphics and Visual Informatics Lab

http://www.cs.umd.edu/gvil

3D: Fuzzy molecular densities



3D: Uncertainty ‘dust’

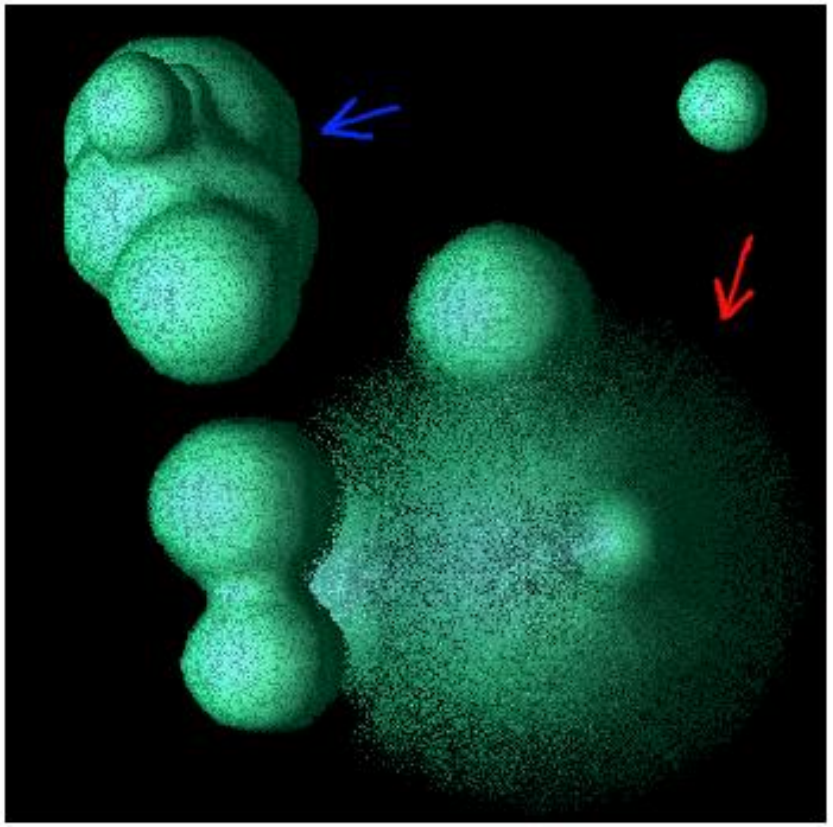


Figure 5: **Hybrid of point-based and polygonal models with transparency. Representation of uncertainty is more intuitive since uncertain regions are simply the ones which look uncertain the eye.**

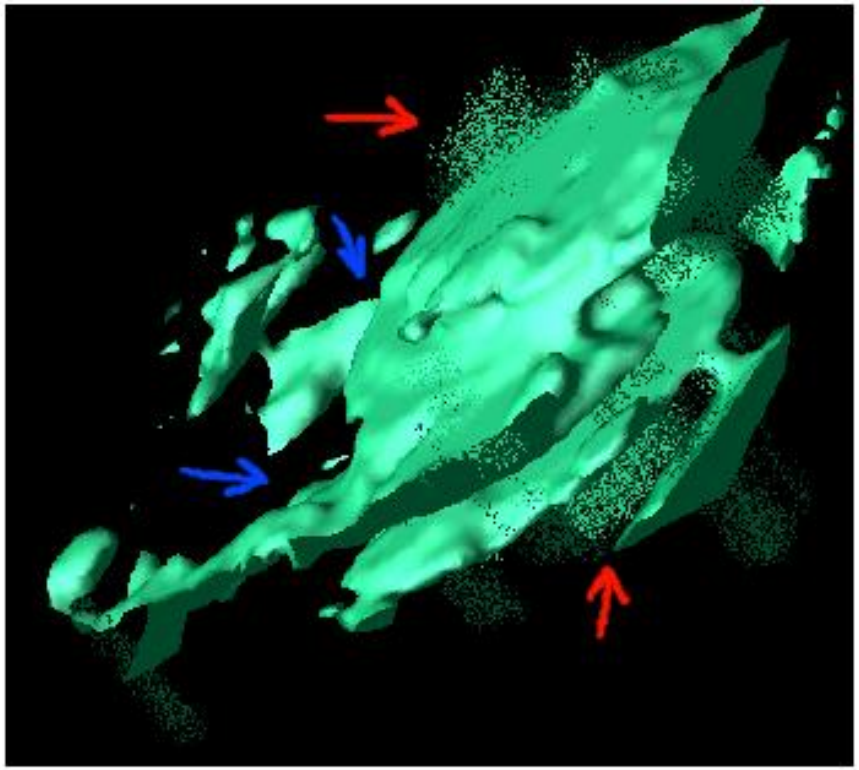
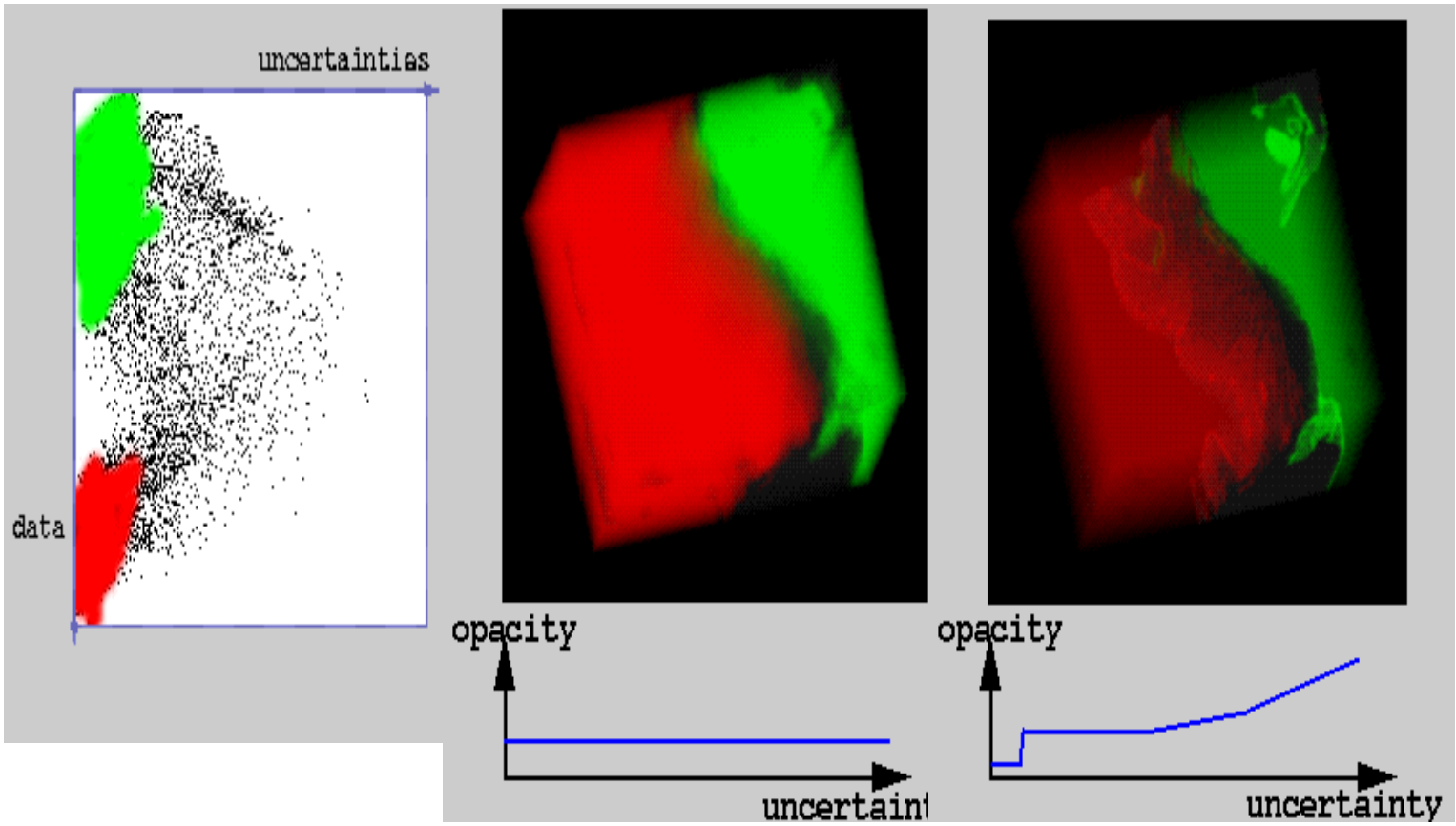


Figure 7: **Real tumor data.**

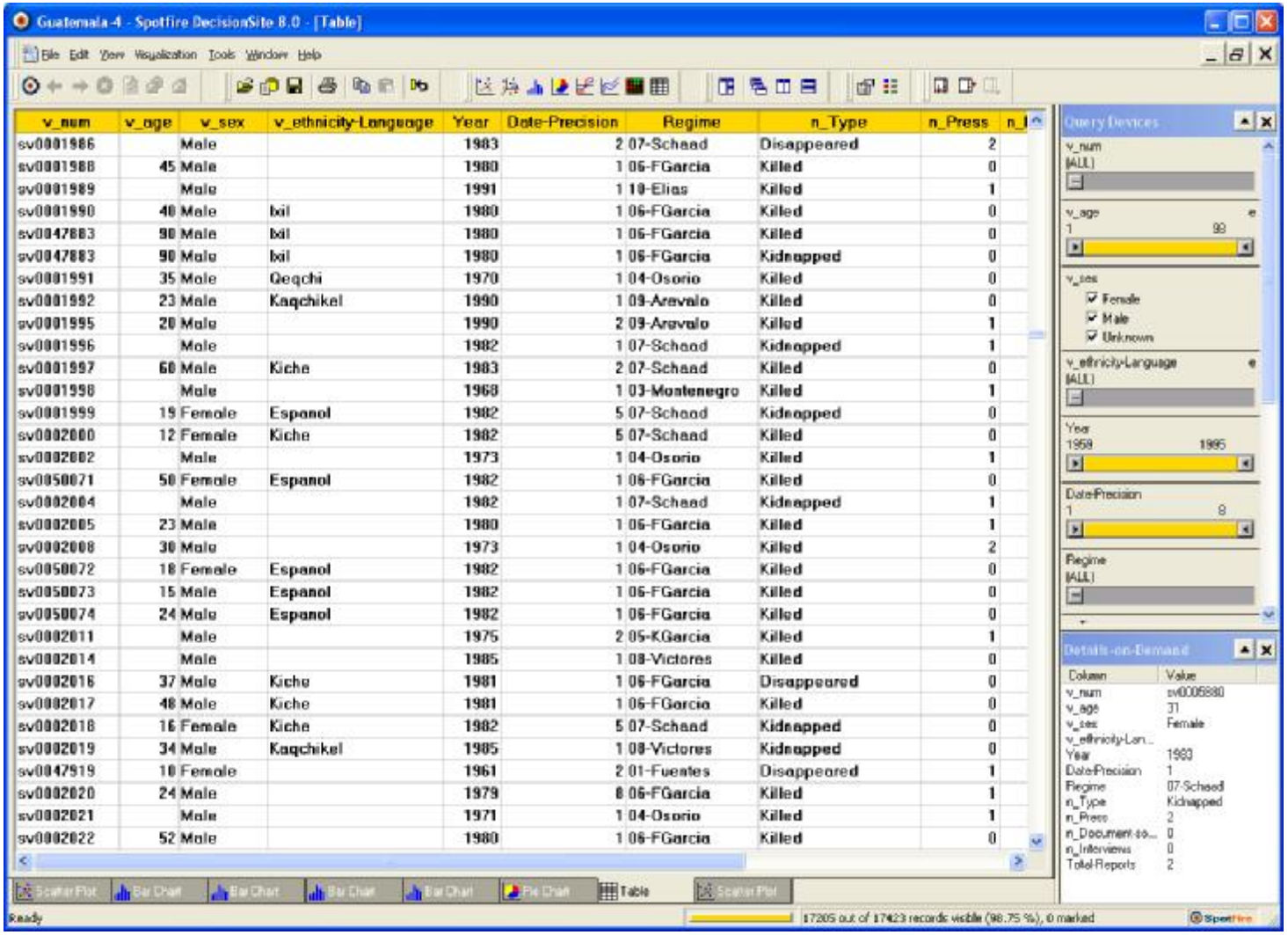
clearly see that areas pointed to by red arrows are of particularly high uncertainty while those pointed to by blue arrows are rather certain and strictly shaped. Figure 8 visualizes a larger subvolume of the same data set. Uncertainty distribution is just as intuitive as it is in Figure 3.

3D: Color & opacity



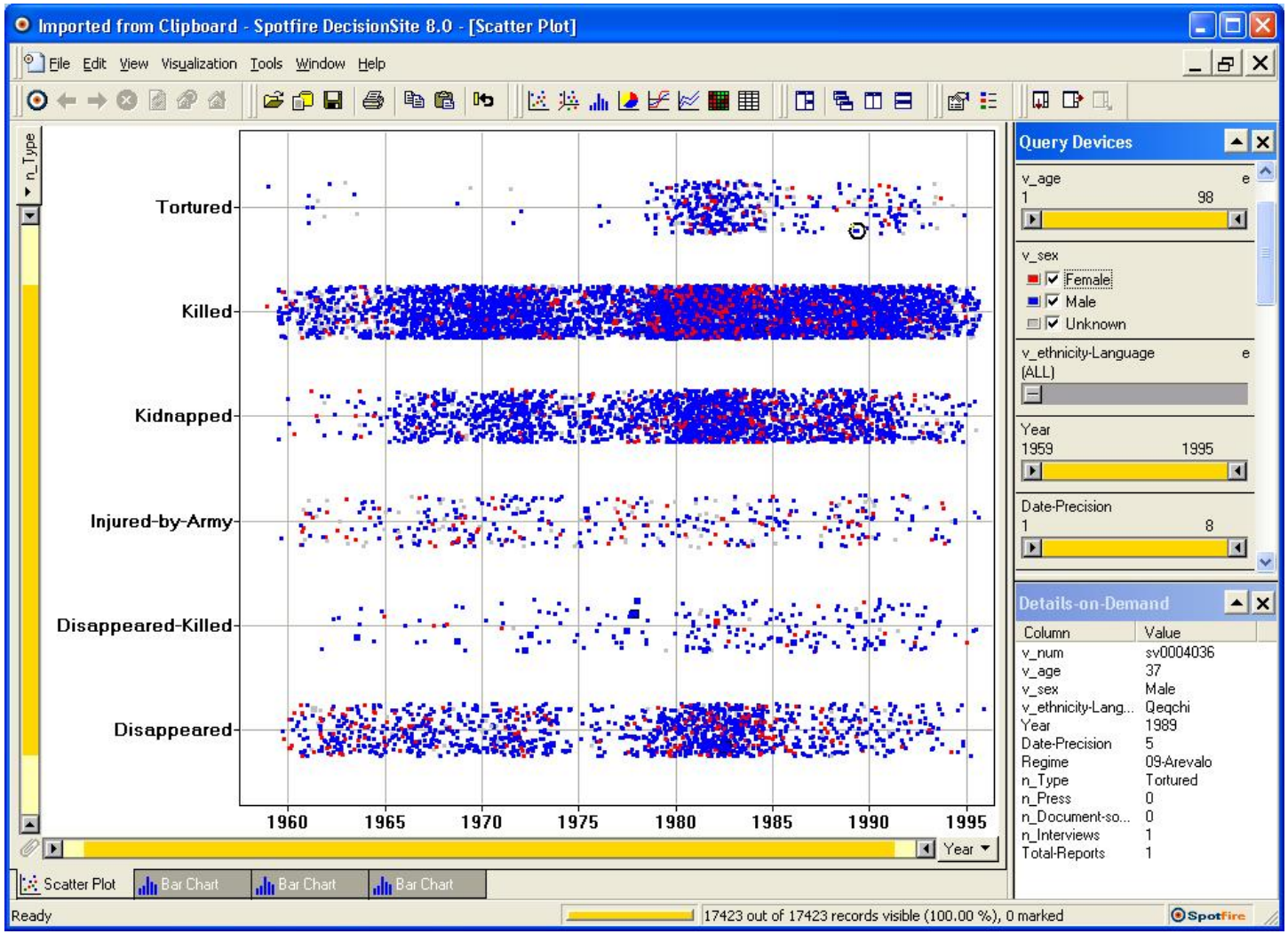
Multi-V: Database/spreadsheet tables

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network



Multi-V: Database/spreadsheet tables

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network



Temporal: Granularity of time

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network

April 14, 1865 - The Stars and Stripes is ceremoniously raised over Fort Sumter. That night, Lincoln and his wife Mary see the play Cousin" at Ford's Theater. At 10:13 p.m., during the third act of the play, John Wilkes Booth shoots the president in the head. Docto president in the theater then move him to a house across the street. He never regains consciousness.

April 15, 1865 - President Abraham Lincoln dies at 7:22 in the morning. Vice President **Andrew Johnson** assumes the presidency.

April 18, 1865 - Confederate Gen. Joseph E. Johnston surrenders to Sherman near Durham in North Carolina.


Funeral Procession on Pennsylvania Ave. - April 19, 1865

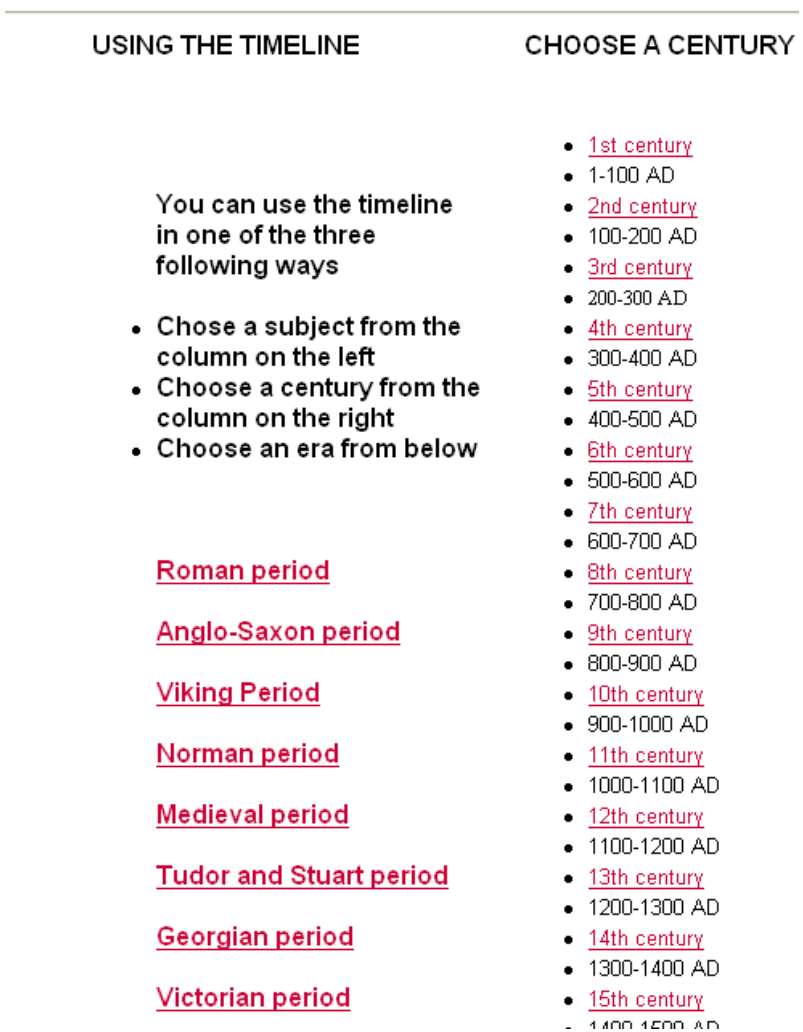
April 26, 1865 - John Wilkes Booth is shot and killed in a tobacco barn in Virginia.

May 4, 1865 - Abraham Lincoln is laid to rest in Oak Ridge Cemetery, outside Springfield, Illinois.

In May - Remaining Confederate forces surrender. The Nation is reunited as the Civil War ends. Over 620,000 Americans died in disease killing twice as many as those lost in battle. 50,000 survivors return home as amputees.

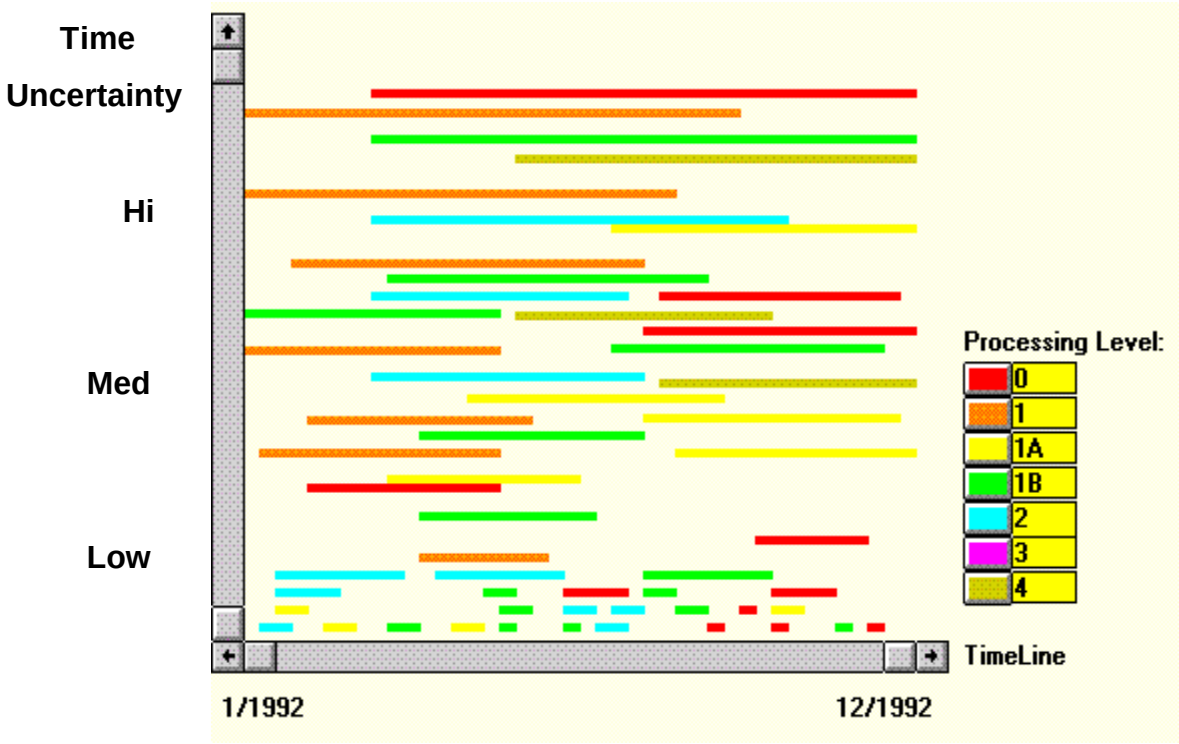
Temporal: Granularity of time

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network



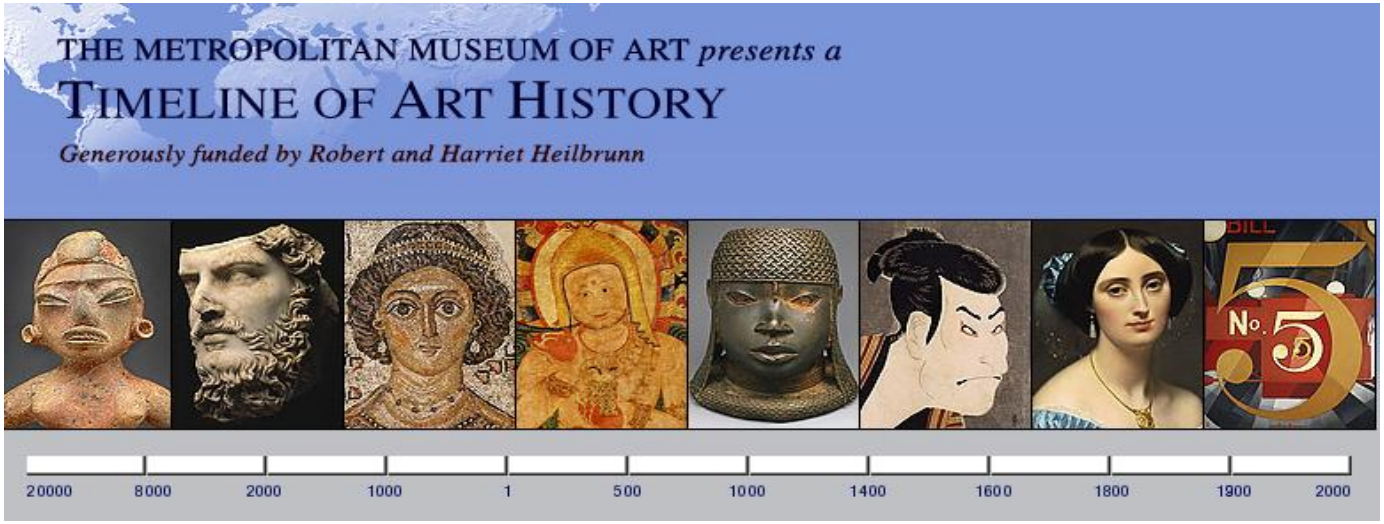
Temporal: Granularity of time

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network



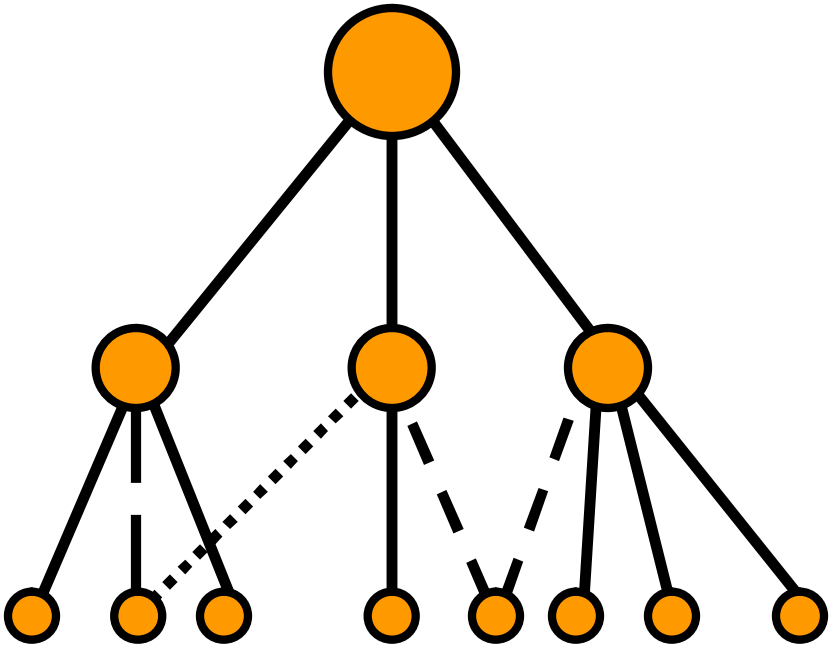
Temporal: Granularity of time

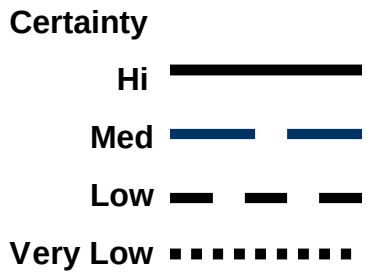
- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network



Tree: Topology, values & names

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network

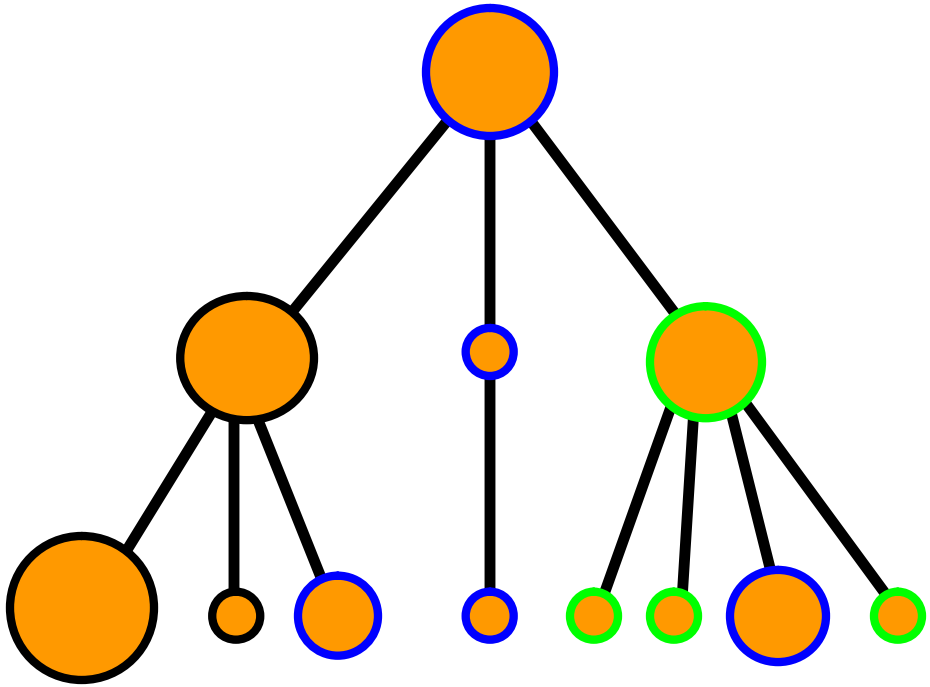


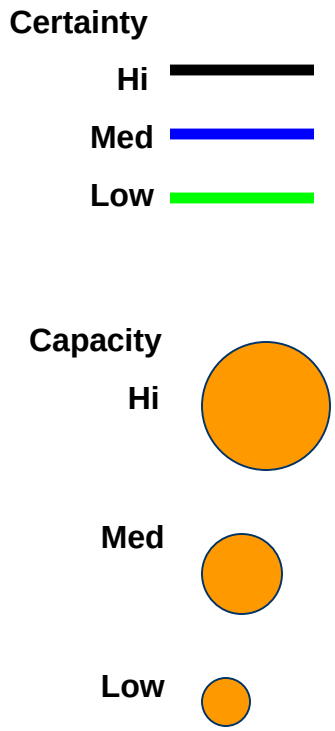


- Gene Ontology
- Tree of Life
- Medical Subject Heading (MeSH)
- Chain of command/Org chart

Tree: Topology, values & names

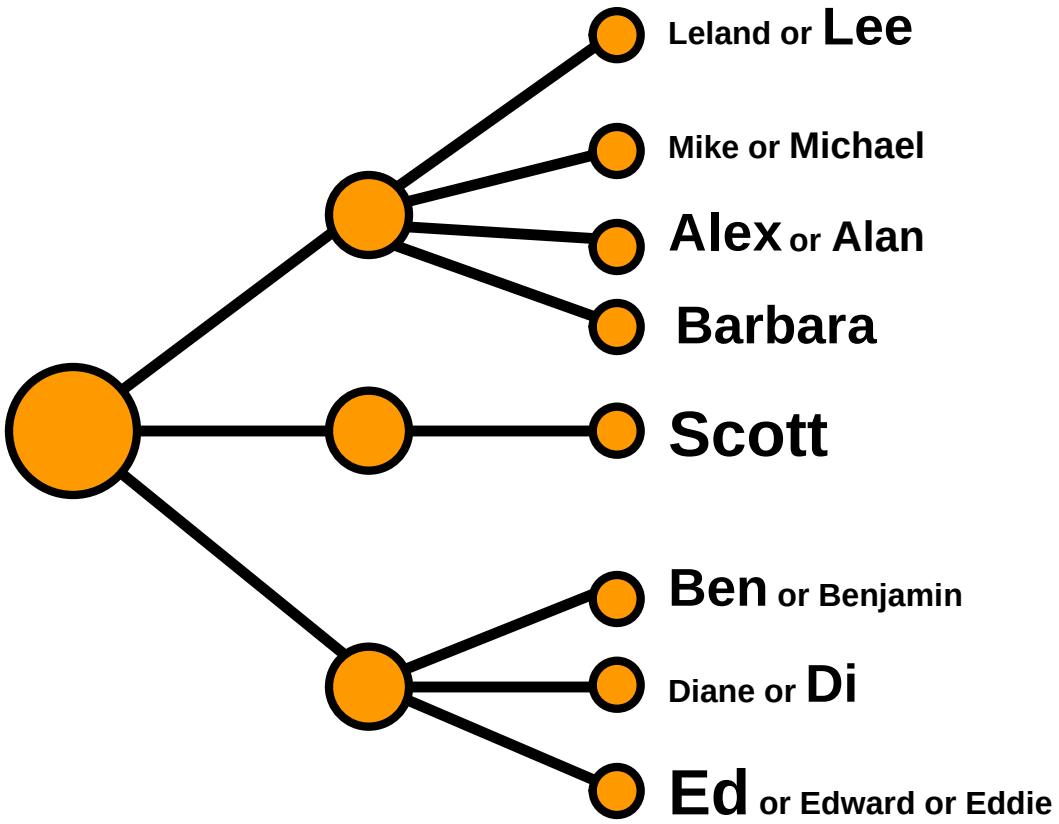
- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network

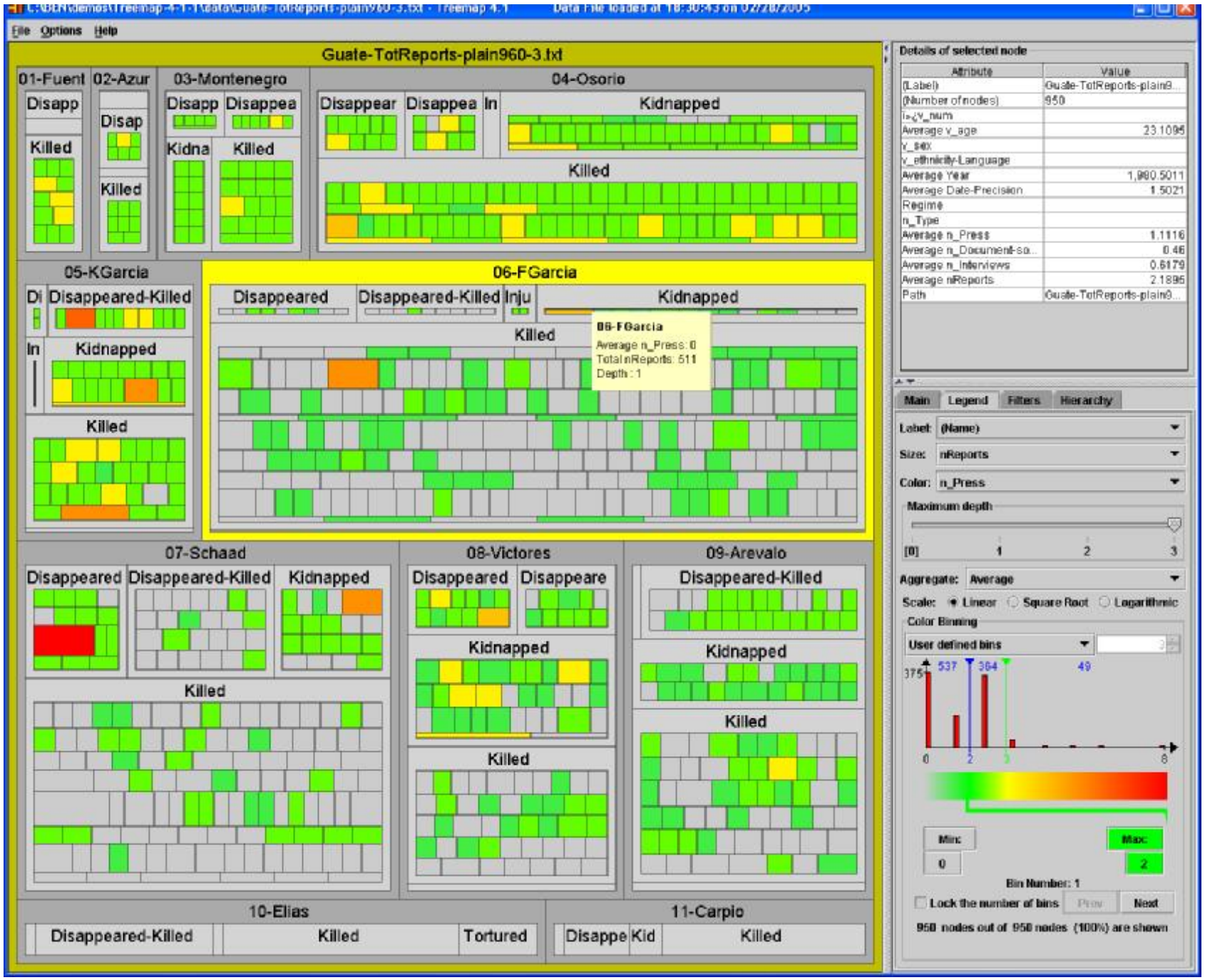




Tree: Topology, values & names

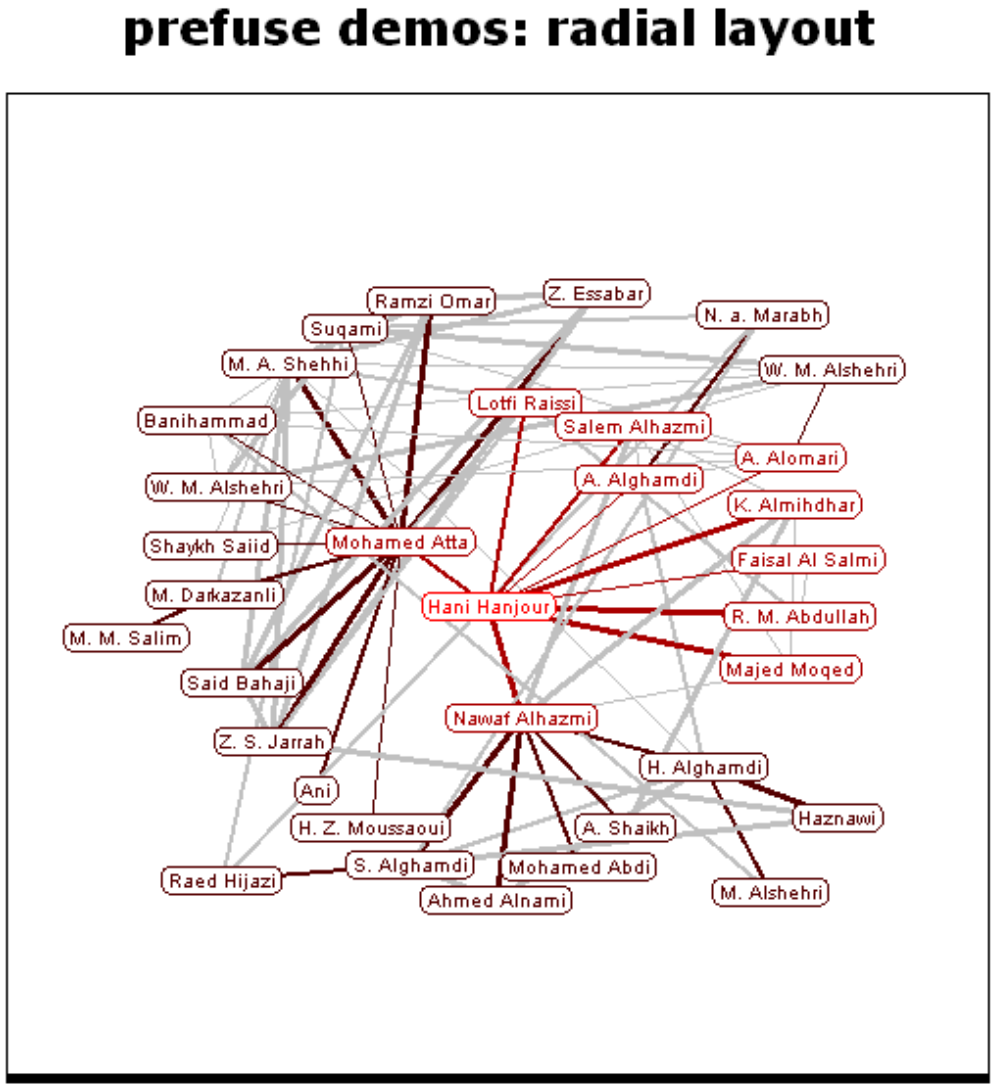
- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network




File Options Help
Guate-TotReports-plain960-3.txt
01-Fuent
Disapp
Killed
02-Azur
Disap
Killed
03-Montenegro
Disapp
Disappea
Kidna
Killed
04-Osorio
Disappear
Disappea
In
Kidnapped
Killed
05-KGarcia
Di
Disappeared-Killed
In
Kidnapped
Killed
06-FGarcia
Disappeared
Disappeared-Killed
Inju
Kidnapped
Killed
06-FGarcia
Average n_Press: 0
Total nReports: 511
Depth : 1
07-Schaad
Disappeared
Disappeared-Killed
Kidnapped
Killed
08-Victores
Disappeared
Disappeare
Kidnapped
Killed
09-Arevalo
Disappeared-Killed
Kidnapped
Killed
10-Elias
Disappeared-Killed
Killed
Tortured
11-Carpio
Disappe
Kid
Killed
Details of selected node
Attribute
Value
(Label)
Guate-TotReports-plain9...
(Number of nodes)
950
Average v_age
23.1095
Average Year
1,980.5011
Average Date-Precision
1.5021
Average n_Press
1.1116
Average n_Document-so...
0.46
Average n_Interviews
0.6179
Average nReports
2.1895
Path
Guate-TotReports-plain9...
Main
Legend
Filters
Hierarchy
Label: (Name)
Size: nReports
Color: n_Press
Maximum depth
Aggregate: Average
Scale: Linear Square Root Logarithmic
Color Binning
User defined bins
375
537
364
49
Minc
0
Max
2
Bin Number: 1
Lock the number of bins
Prev
Next
950 nodes out of 950 nodes (100%) are shown

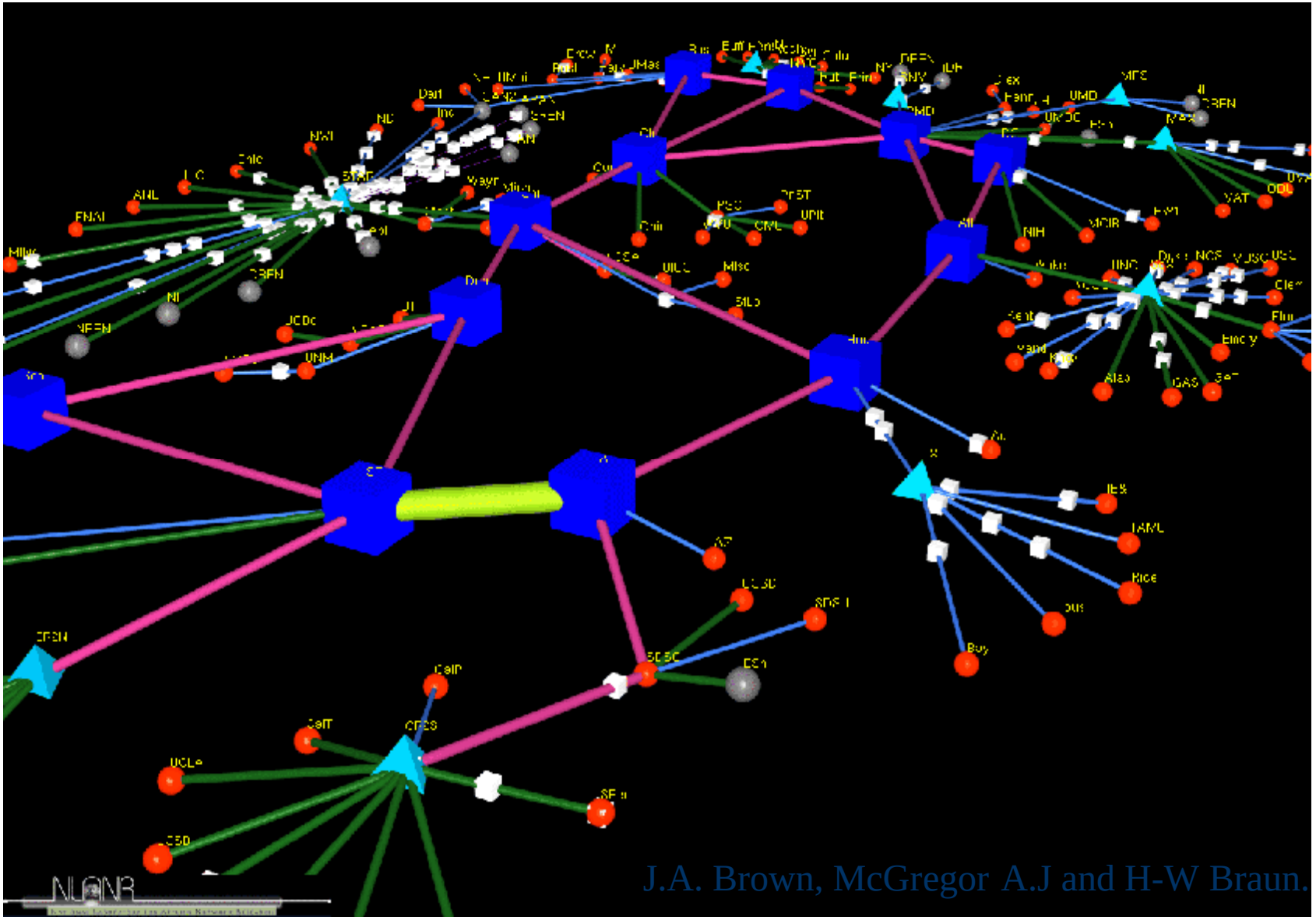
Network: Social relationship

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network



http://prefuse.sourceforge.net/demos-radial.html

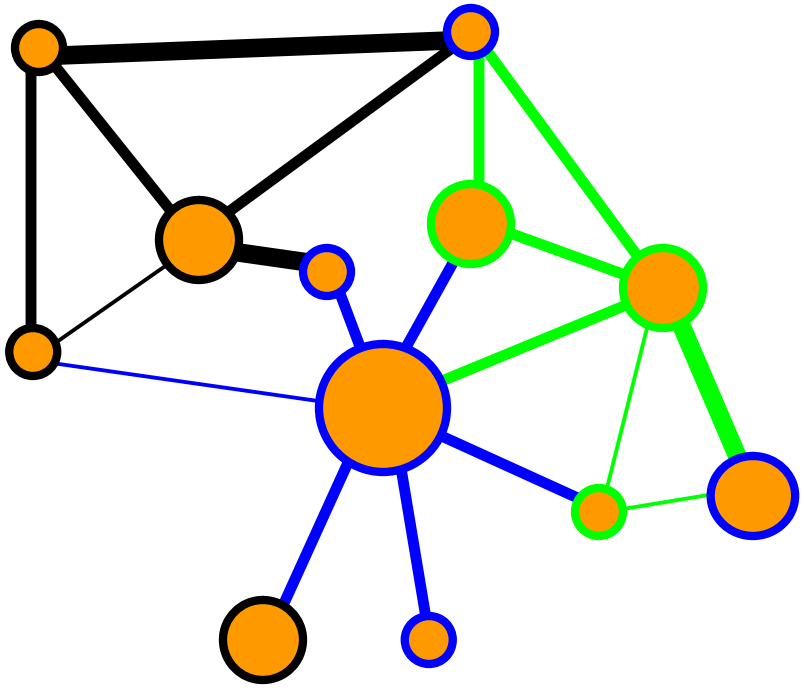
Network: Communication capacity

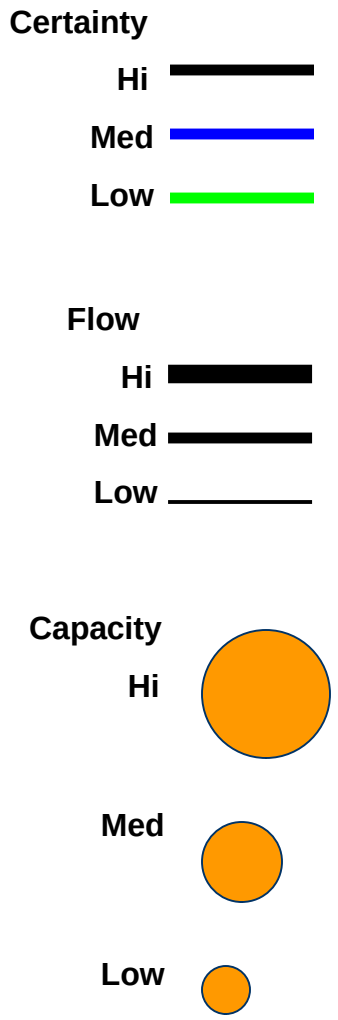


J.A. Brown, McGregor A.J and H-W Braun.

Network: Node & edge uncertainty

- Text
- 1D
- 2D
- 3D
- Multi-V
- Temporal
- Tree
- Network





Next steps

Explore novel approaches to:

- Text: standard terms, percent, probabilities
- Box plots, whiskers, ranges
- Uncertainty glyphs, isoclines,…
- Contours, surfaces, volume rendering
- Hue, saturation, value, focus, haze, dust,…
- Dual views, probes
- Animation, blinking, shaking, flipping,…
- Sound, haptics,…

Next steps

- Heighten awareness of the problem among public, professionals, researchers & developers
- Support multi-valued data representation standards
- Explore techniques for each data type
- Develop guidelines for implementers:
  - Data formats
  - Interactive interfaces
  - Visual presentations
- Develop human performance evaluation methods
- Publish benchmark datasets & evaluation metrics
- Form guidelines for how to propagate/integrate uncertainty markers