Representing Uncertainty

MacEachren Chpt. 4 & Slocum Chpt

• One of the many maps designed by Cindy Brewer and Alan MacEachren for the National Center for Health Statistics’ Atlas of United States Mortality. These choropleth maps use both a diverging color scheme (to focus attention of the highest- and lowest-values) and an innovative cross-hatch pattern (using parallel black-and-white lines) to distinguish between reliable and unreliable data values.

The Elements of Uncertainty

• “Original Sin” – error in the raw data
• Errors in processing data
  – Interpolation
  • Different methods
• Error in symbolization
• Read MacEachren pp.74-76.
Quality and reliability

• FIPS 173 (NIST 1994) 5 categories for assessing data quality:
  – Lineage
  – Positional Accuracy
  – Attribute Accuracy
  – Logical Consistency
  – Completeness

Assessing Data Quality

• Lineage – historical development of digital data, analogue source, metadata
• Positional Accuracy - locational horizontal and vertical
• Attribute Accuracy – classification, identification
• Logical Consistency – fidelity of relationships encoded in data structure
• Completeness – information about selection criteria (what is included / excluded.)

Representing Data Quality

• Intrinsic visual variables
  – Saturation
• Extrinsic visual variables
  – Added clues – dials, thermometers, arrows, bars
    • Tissot’s Indicatrix
      – Ellipse of distortion
  • Clarity
    – Crispness – fuzzy boundaries
    – Resolution – level of detail
    – Transparency – “Fog” Reliable data clearly seen; unreliable not
Uncertainty Metaphors

Table 1: Spatial data metaphors used for uncertain data

<table>
<thead>
<tr>
<th>Technique</th>
<th>Fog – Detail</th>
<th>Etos – Form</th>
<th>Hitting Probs – Stability</th>
<th>Colour Mix – Clarity</th>
<th>Print Mix – Fuzziness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certain Dose</td>
<td>Clear</td>
<td>Sharp Focus</td>
<td>No Blending</td>
<td>High Saturation</td>
<td>Single Mix – Low Fuzziness</td>
</tr>
<tr>
<td>Metaphor of</td>
<td>– Receding</td>
<td>– Focused</td>
<td>– No Movement, consistent</td>
<td>– High clarity, less receptive</td>
<td></td>
</tr>
<tr>
<td>Uncertain Dose</td>
<td>Fuzzy</td>
<td>Hazy</td>
<td>Hitting over areas</td>
<td>Low Saturation</td>
<td>Multiple Mixes – High Fuzziness</td>
</tr>
<tr>
<td>Metaphor of</td>
<td>– Hiding</td>
<td>– Unfocused</td>
<td>– Less visible, more</td>
<td>– Low clarity, more reactive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Detail</td>
<td>and emerging</td>
<td>unsettled</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Methods to show uncertainty

• Separate maps
• Attribute and uncertainty portrayed on same map
• Date exploration tools – animation etc.

Uncertainty – as disagreement

• Figures 1 and 2 use color to represent regions of uncertainty in the output from meteorological computer models. These dynamic maps highlight regions where the models disagree the most.
Animating Uncertainty

- By animating these maps, it is also possible to see how the meteorological models disagree over time (i.e. the further out they try to forecast).
- Click here to learn more about the animation aspect of our research or here to learn more about this particular project.

RVis

- Reliability Visualization (RVis) system developed by David Howard and Alan MacEachren was designed to show uncertainty in data of the Chesapeake Bay.
- Their research focused on two important questions:
  - what kind of user interface is most effective for Geovisualization and,
  - what graphic variables are appropriate for showing different kinds of uncertainty?

RVis

- The map on the left displays dissolved inorganic nitrogen with darker representing more nitrogen.
- The map on the right represents reliability (measured as the variance estimate from kriging).
- An interactive reliability threshold allows users to hone-in on areas of greatest uncertainty.
Visualization of Uncertainty Techniques

- Which techniques do users prefer?
- NZ Survey assessed visual appeal, speed of comprehension, and overall effectiveness of uncertainty visualization techniques. (41 respondents)
- Visualising Uncertainty in Spatially-Referenced Attribute Data using Hierarchical Spatial Data Structures, J.D. Kardos, A. Moore and G.L. Benwell Department of Information Science, University of Otago, NZ

The techniques involved in the survey included:

- **Adjacent maps** - where two maps can be used to show the uncertainty, one to show the actual information and another to show the uncertainty (MacEachren et al. 1999).
- **Overlay** - where a single choropleth map can be used to show the attribute information with an overlay of the uncertain information shown as textures on top.
- **Blurring** - where the clarity of an area boundary is used to show the uncertainty of the data. A blurred area indicates certain information, whilst a more approximate pattern definition would indicate uncertain information (MacEachren 1992).
- **Fog** - where uncertain parts on a map become partially hidden making the areas unclear to see. The thicker the fog the more uncertainty is in that part of the map. The fact that fog obscures data from view is not an issue here since fog is part of the method (MacEachren 1992).
- **Pixel mixture** - where pixels are divided into sub-pixels and an appropriate class value is given to each sub-pixel proportional to the membership function calculation (De Gruijter et al. 1997).
- **Saturation of color** - where saturation of color is used to visualize uncertainty. The more saturated (richer) a color representing a particular class, the more certain the information is in that part of the map. This creates a multi-dimensional uncertainty chart, with brightness between classes and uncertainty in the data represented by hue and saturation (Hengl et al. 2002).
- **Sound** - provides a level of uncertainty at a particular location on a map through a variable pitch. A low pitch sound depicts low uncertainty, and a high pitch sound for large uncertainty (Fisher 1994; Krygier 1994).
- **Blinking pixels** - where information in the spatial display is manipulated causing it to blink, hence highlighting those uncertain areas to the user (Fisher 1995; Monmonier and Gluck 1994; Evans 1997).
- **Animation** - a movie of map realizations (generated from a Monte Carlo simulation) highlighting areas where data is considered to be uncertain. Ehlschlaeger et al. (1997) note that if there is little change between realizations then one can be fairly convinced about the extent of the uncertainty.

**Summary**

- Many types of data quality
- Many means to represent uncertainty
  - Intrinsic graphic variables
  - Extrinsic graphic variables
- Need for more research on perception and use of uncertainty