Visual Exploration of the Spatial Distribution of Temporal Behaviours

Gennady Andrienko & Natalia Andrienko
FHG AIS (Fraunhofer Institute for Autonomous Intelligent Systems)
http://www.ais.fraunhofer.de/and

IV conference, Greenwich UK, 8.07.2005

Agenda

1. Background: visual tools for analysing spatial time series data
2. Analytical questions to be supported
3. Visualisation of local behaviours
4. Combining several tools for:
   - getting the general picture for the entire territory;
   - finding spatial patterns of similar behaviours;
   - detecting patterns of similar changes
5. More complex example. Scalability issues.
6. Discussion and conclusion

Recent development of spatial time series analysis (~15 years)

Statistics & InfoVis

1. Access to values by pointing
2. Overlaying lines for comparison, user-controlled line distortion (for facilitating analysis)
3. User-defined selection of lines with particular characteristics
4. Dynamic linking to additional displays (scatter plots, histograms, {maps} etc.)

GeoVis

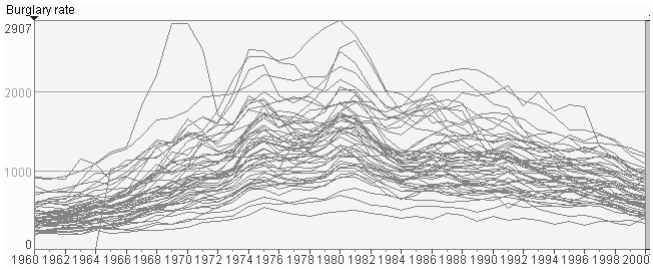
1. Series of maps
2. Animated maps: automatic or user-controlled
3. Temporal focusing and brushing
4. Maps with time diagrams
5. Attention to temporal changes

Unsuitable for large number of time series
Little cooperation of InfoVis and GeoVis methods

Analytical questions for exploring spatial distribution of temporal behaviours

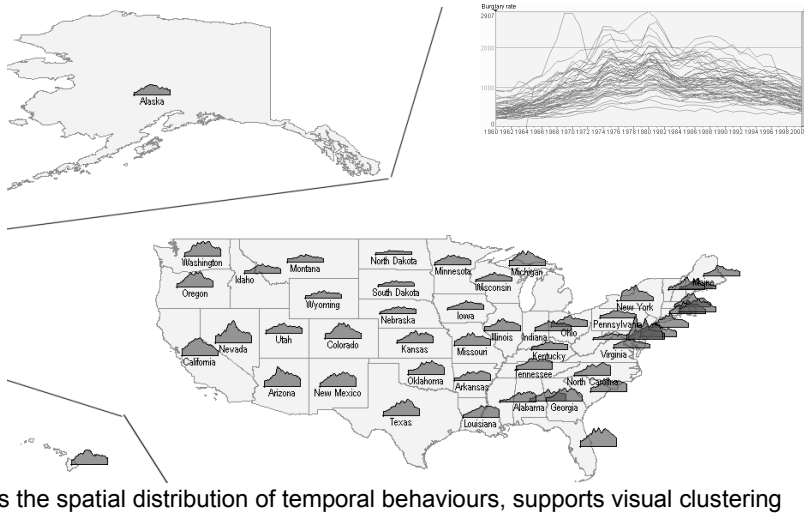
- What is the general dynamics of values over the entire territory?
- What are the general features of the local behaviours in some area and how do they compare to the behaviours on the remaining territory?
- Find locations with the behaviours having specific features and check whether these locations form a spatial cluster.
- Identify spatial clusters of similar behaviors.

Time graph of burglary rate



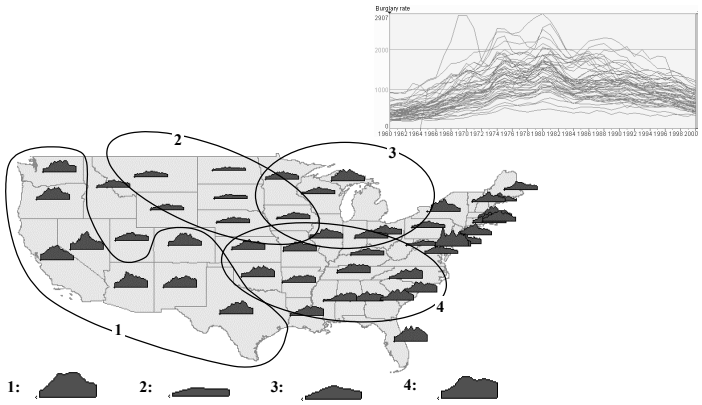
Shows the general pattern of the temporal behaviour

Spatial distribution



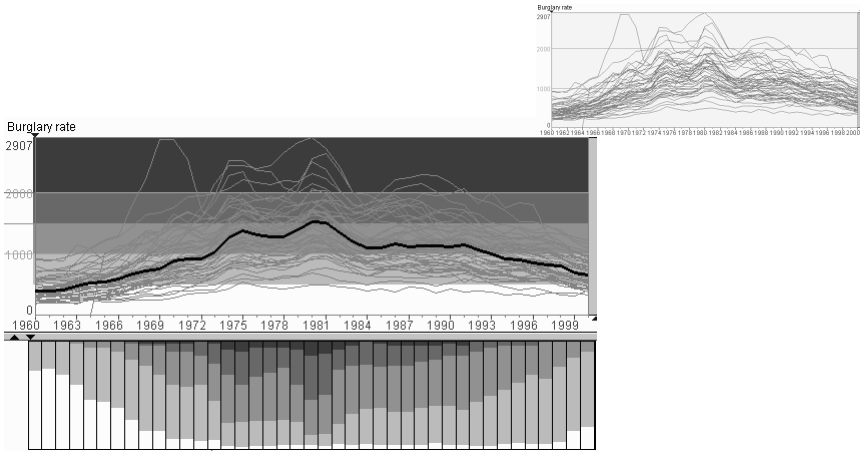
Shows the spatial distribution of temporal behaviours, supports visual clustering

Clusters of similar behaviours



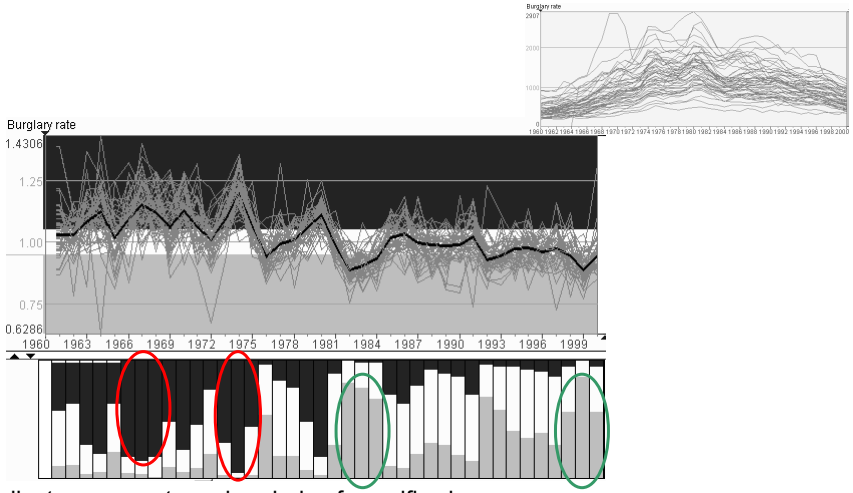
Such visual clustering is applicable to small data sets

Aggregation by values



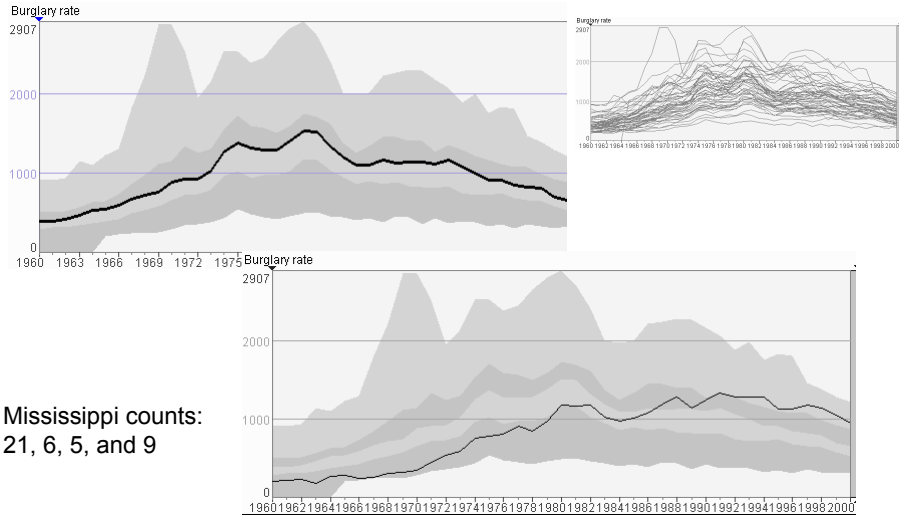
Demonstrates the general pattern of development even for large data sets
Click on any segment highlights lines and spatial objects

Aggregation for transformed data (annual changes)



Indicates moments and periods of specific changes

Aggregation by quantiles



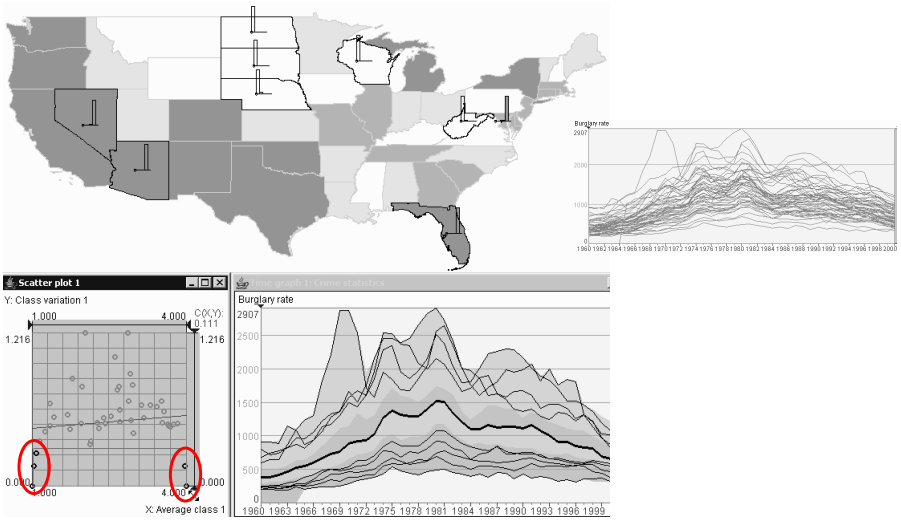
Mississippi counts: 21, 6, 5, and 9

Shows the development of the statistical distribution of values

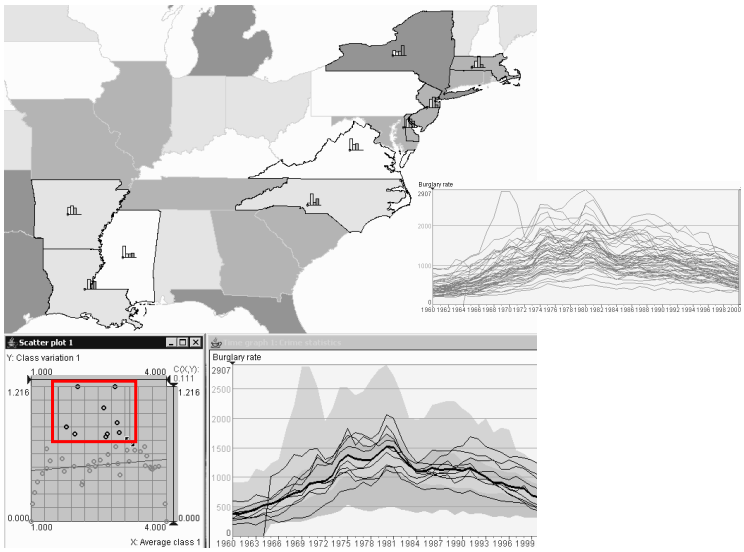
Clusters of similar behaviours

Shows spatial distribution of similar temporal behaviour, applicable to large data sets (without diagrams)

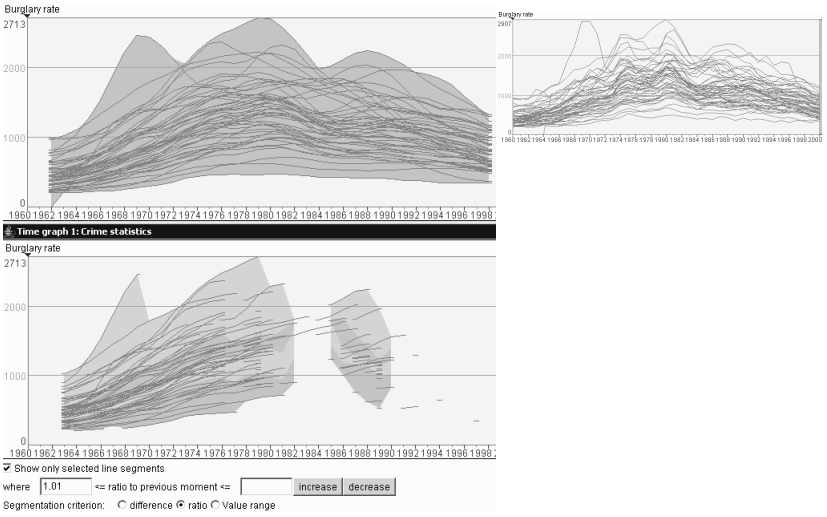
States with persistently low and persistently high values



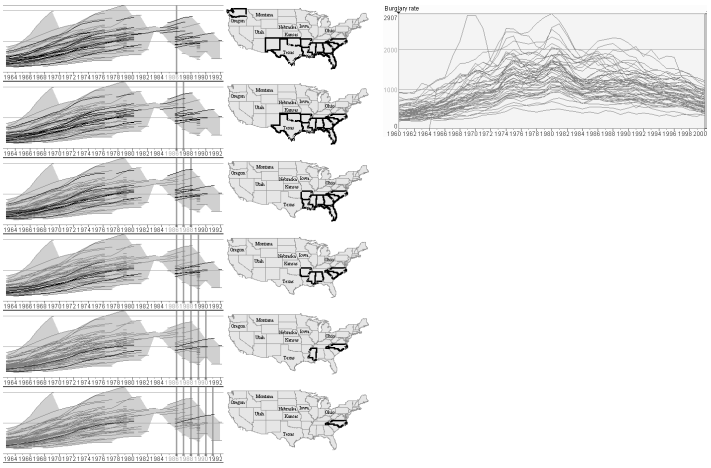
States with high variability of values



Segmentation



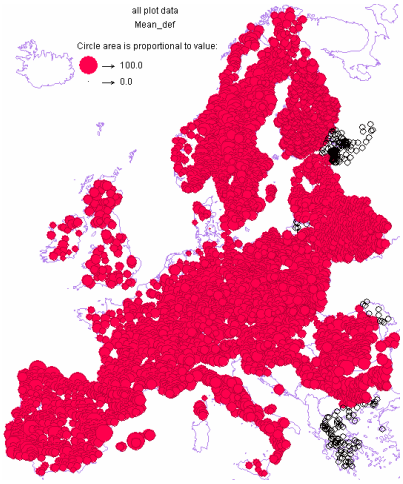
States with stable increase



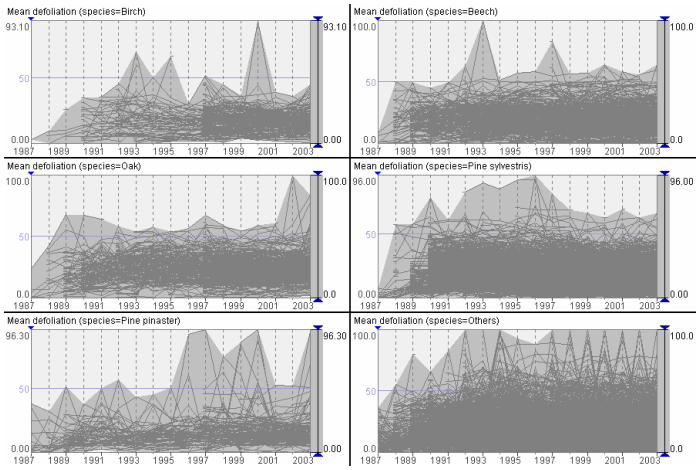
Shows the general pattern of changes, supports sensitivity analysis

More complex example: the defoliation data (NEFIS project)

- Large volume: 6169 spatially-referenced time series (17 years)
- Two dimensions: S&T
- Many missing values
- No full compatibility across countries, species, time etc.

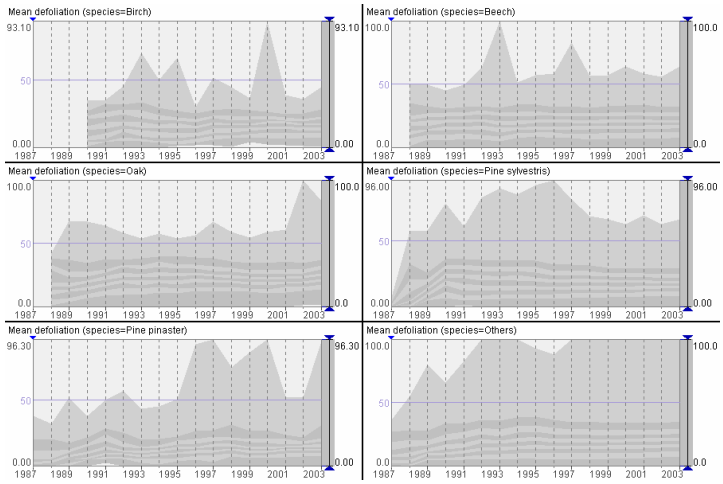


Explore overall temporal trends

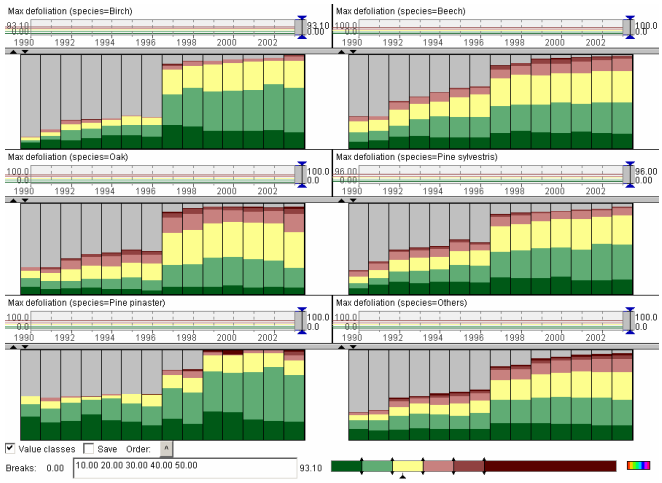


Line overlapping obstructs data analysis
→ apply aggregation

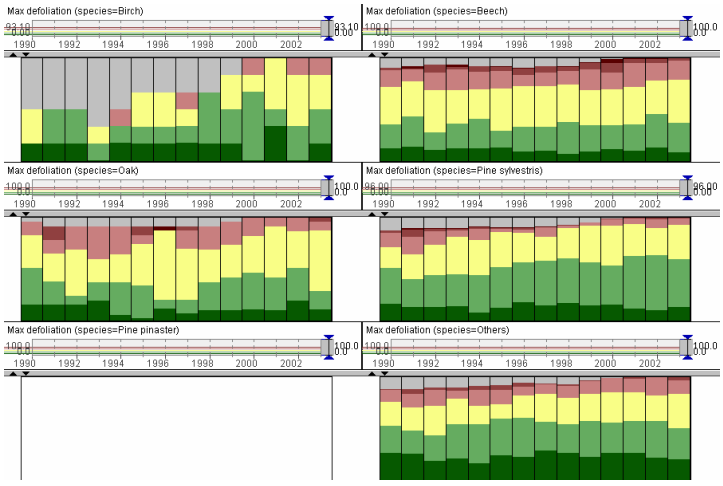
Aggregation method 1: by quantiles



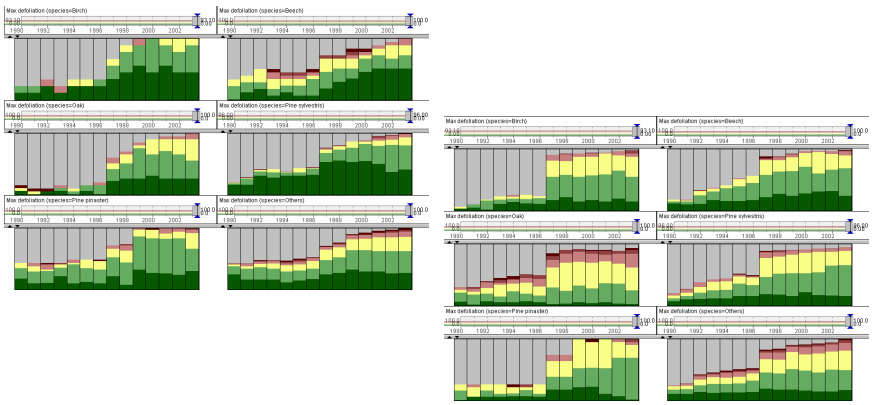
Aggregation method 2: by intervals



Divide and Focus: Germany



Divide and Focus: age groups 1,3



Attend to particulars

Types of particulars (examples):

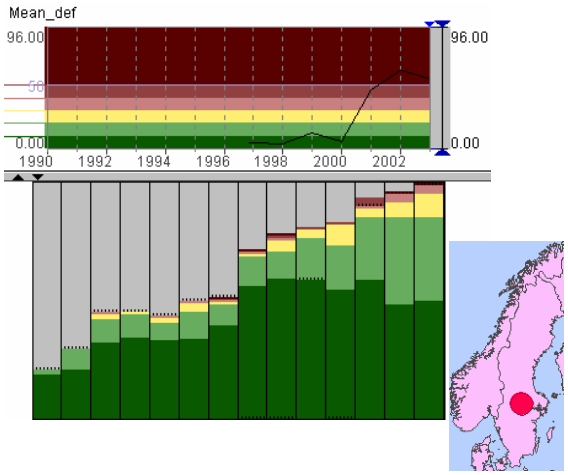
- Extreme values
- Extreme changes
- High variability
- …

Questions:

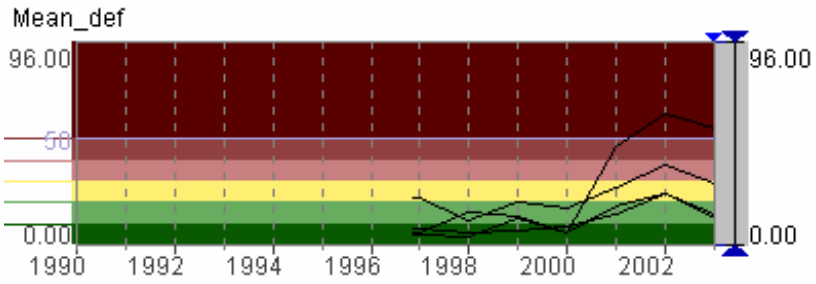
- When?
- Where?
- What is around?
- *Why?* (a question for further, in-depth analysis)
  Domain knowledge is essential

Attend to particulars: extreme values

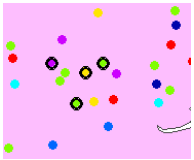
1. Click on a segment corresponding to extreme values
2. The behaviour(s) is(are) highlighted on the time graph
3. The location(s) is(are) highlighted on the map



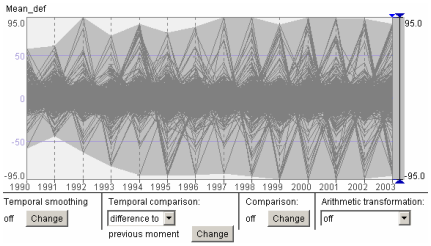
Attend to particulars: what is around?

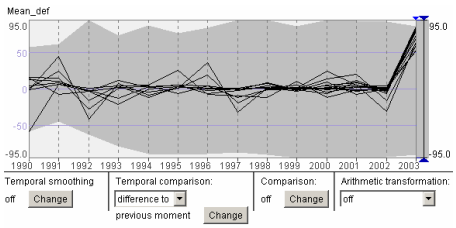


- In some neighbouring places the behaviours during the period 2000 - 2003 are somewhat similar

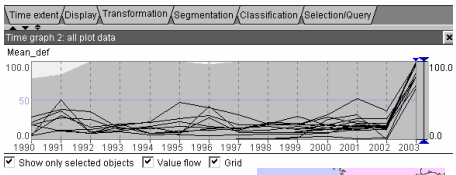


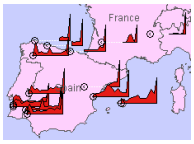
Attend to particulars: extreme changes



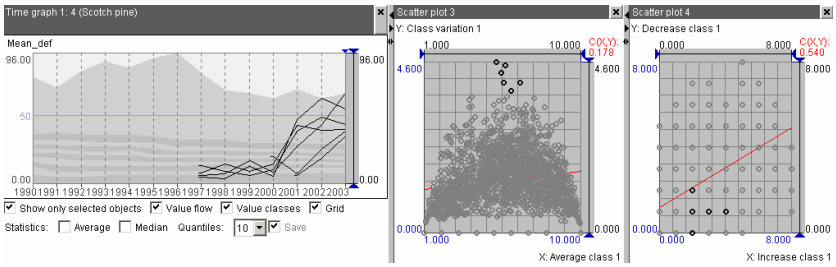


1. Transform the time graph to show changes
2. Select extreme changes in a specific year (here 2003)

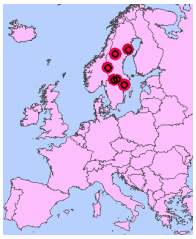




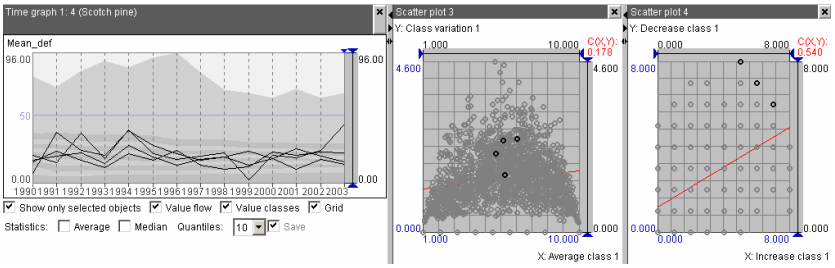
Attend to particulars: high variation



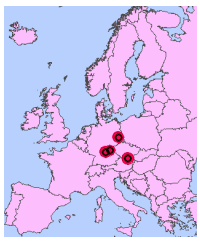
1. Aggregate time graph by quantiles
2. Save counts
3. Visualise e.g. on a scatter plot
4. Select items with high variation



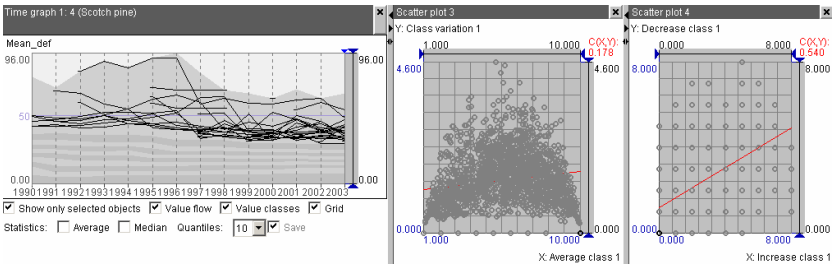
Attend to particulars: high fluctuation



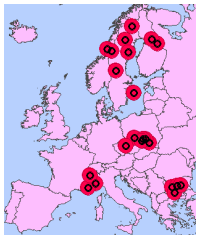
- Select items with maximal number of jumps between quantiles



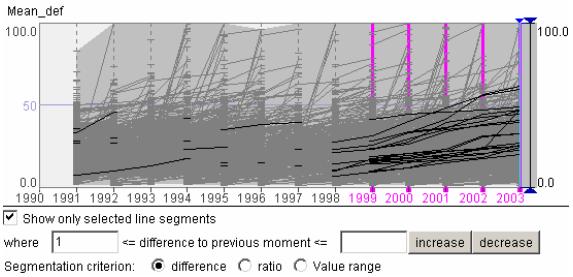
Attend to particulars: stable extremes



- Select items being always in the topmost 10%



Attend to particulars: stable increase



1. Turn the time graph in the segmentation mode
2. Choose “increase” and set minimum difference
3. Select a sequence of years by clicking
4. Check sensitivity to the time period!

Recap: CommonGIS (not a “common GIS”)

A variety of well-integrated tools for EDA

- Time-aware maps + statistical graphics; several mechanisms of display coordination
- Designed to gain synergy of
  - ✓ Visualisation
  - ✓ Display manipulation
  - ✓ Data manipulation
  - ✓ Querying
  - ✓ Computational techniques, including aggregation

Crucial role of tasks in visualisation design!

Scalability issues

- ✓ Most demonstrated methods have the complexity O(n) or O(n*log(n))
  - ✓ They are applicable to large data sets
  - ✓ Interactive manipulation is possible even for large data sets
  - ✓ It is possible to design a client-server system with incremental data loading: first only general statistics for providing graphical overview, then details on demand
- ✓ Long time-series still require new methods

To Learn More:

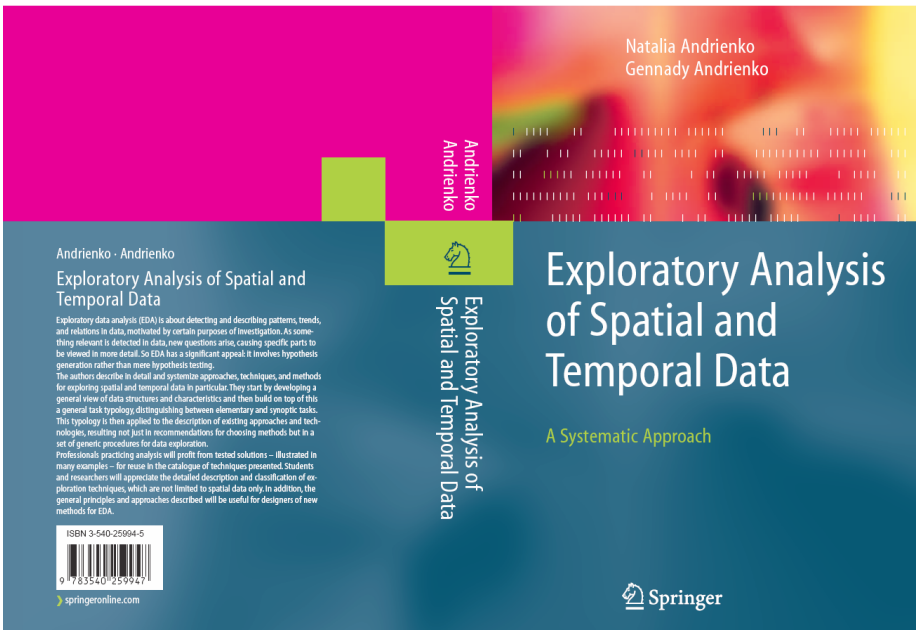
- ➢ Software: http://www.commongis.com
- ➢ Papers, tutorials, on-line demos: http://www.ais.fraunhofer.de/and
- ➢ Book to appear:

N. and G. Andrienko

"Exploratory Analysis of Spatial and Temporal data. A Systematic Approach"

(Springer-Verlag, ≈ end 2005)

A theoretical framework for linking tasks, tools, and principles of data analysis



In press, to appear ≈ end 2005

- ECML/PKDD'05 Workshop on "Mining Spatio-Temporal Data", Porto, Monday, 3.10.2005
- ✓ at 16th European Conference on Machine Learning (ECML'05) and 9th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'05)

  http://www.di.uniba.it/~malerba/activities/mstd/

  Deadline: 25.07.2005 (full paper)
- Special Issue of "Journal of Intelligent Information Systems" after the workshop