Schematized Small Multiples for the Visual Comparison of Geospatial Data

Andreas Reimer, Andrea Unger, Wouter Meulemans, and Doris Dransch

Abstract—The task to compare geospatial variations of data values over other variables appears in many application scenarios. Small multiples is a visualization method to support this task, but it often suffers from too high information density. To address this problem, we propose to apply an automated schematization of the underlying map, which generates a highly abstracted view of the map. The usefulness of our approach is exemplified for the comparison of geoscientific models.

Index Terms—Small Multiples, Information Density, Cartographic Generalization, Schematization.

1 INTRODUCTION
One fundamental task in visualization is the comparison of the variation of data values over other variables. If one of the other variables is geographic location, a map of some sort is the most straightforward way of visualizing the location. The visualization of more than one spatial configuration is known as a cartographic problem with more than two components [3]. For this problem, four major visualization strategies exist:

- **Full-inventory maps**: All data values are shown within one map. Multiple data values need to be represented at the same point in space, leading to the use of diagram-symbols or “glyphs”.
- **Collection of maps**: Show multiple (small) maps. The values are all visualized in separate and comparable maps. This technique is also known as small multiples as popularized by Tufte [15].
- **Animated maps**: A single map is shown that discretely changes over time. Changing data results in perceived movement or changes in the appearance of objects.
- **Synthesized maps**: The essential information is displayed in one layered image, which requires a prior analysis to extract the relevant information.

Full-inventory maps emphasize relations among data values associated with one region. They are less suited to understand variations in geographical space. For this goal, a collection of maps should be used. It allows a direct comparison of the variation of data values and spatial configuration at the same time. Small multiples allow the two dimensions of the plane to be utilized in both directions each. Animated maps, in contrast, allow for only one direction of ordering. They are postulated to work best for showing temporal dynamics, although they have known weaknesses in allowing comparisons between constituent states. Synthesized maps are mainly suitable for the presentation of results.

An adequate choice of the visualization method usually depends on the underlying visualization task. In many applications, the geospatial variations among data values are a crucial aspect for data interpretation and should be emphasized in the visualization. One example is the comparison of multiple geoscientific simulation models, with their inherent geospatial context. In such scenarios, a collection of maps is the preferable visualization strategy. But a collection of maps comes with a disadvantage: a higher density of visual information. This may distract the user from identifying the information relevant to the comparison task.

Acknowledging this shortcoming, Carr and Pierson postulated the need for a spatial “caricature” [4]. Creating such caricatures is the realm of cartographic generalization. An overview of automated generalization methods is given by Li [8] and Mackaness et al. [10]. The successes of these techniques originate from a thorough understanding of topographic maps, leading to design rules, constraints and algorithms. Automated cartographic generalization research offers many solutions for the tasks of line simplification (removal of points) [6, 16] and smoothing (using interpolated points) [9]. Generally, topographic map generalization is concerned with maximizing detail with visibility and clutter as constraining forces.

In contrast, our application demands for a caricature that minimizes detail to enhance perceptual tasks, without losing “identifiability” or “recognizability” of the map. This is covered by the field of schematization, a relatively underdeveloped aspect of generalization. It is concerned with a strong reduction of visual complexity by exaggeration of important features and elimination of non-functional detail. Research in automatic schematization has concentrated on line-network maps such as the London Tube Map [1, 5, 2, 14]. For polygonal schematization, however, automated methods are seldom investigated [11]. The wealth of polygonal schematization examples available in the literature have been created manually [12].

In this poster, we propose to use small multiples in combination with an automatic schematization of polygonal maps to support the task of comparing geospatial variations of data values. Our approach allows us to present only the necessary elements, which are the general geospatial location, the relations among regions where data are available, and the actual values associated with these regions. Information that is irrelevant for the analysis task, such as detailed geographical information, is omitted. This supports an effective comparison of data values with an emphasis on geospatial relations.

After describing our technique to schematize polygonal maps in Section 2, we exemplify our idea with an application example from geoscientific modeling and simulation in Section 3.

2 SCHEMATIZATION
The basic idea for generating schematized maps is based on two characteristics of chorematic diagrams (extremely abstracted maps): they use a very low number of vertices and the edges show a certain amount of parallelism [13]. These are the main characteristics of recognizable manually produced schematizations of outlines. To reproduce these, we detect characteristic points of a given territorial outline and use them as input for a simulated annealing process that attempts to maximize parallelism. The characteristic point detection is done with a modified Imai-Iri algorithm [7]. A binary search is used to generate the specific number of characteristic points as specified by a drawing-space parameter. To address the known problem of starting point dependency for polygons, we cut each polygon at one of the diametric points. To prevent polygons from collapsing onto a line or point, we require that each polygon is represented with at least three vertices. This directly gives a lower bound on the number of vertices in the output. A simulated annealing process follows to optimize parallelism
of the outline by moving the vertices within a given threshold. Some modifications within the simulated annealing algorithm take care of perceptual uncertainty regarding parallelism, and to provide a scoring bonus for facing edges. For full details of the method, including a contextualization regarding related schematization work, please refer to the work of Reimer and Meulemans [13]. A sample input and resulting schematization are shown in Fig. 1. The schematization is used for the application example, presented in the next Section 3.

![Fig. 1. Original outline and a schematization with 30 edges, obtained by applying the method of Reimer and Meulemans [13].](image)

### 3 Application Example

We underline our idea that schematized small multiples support a fast comparison of geospatial data with an application example from the field of geoscientific modeling and simulation. Specifically, we use our technique to simultaneously compare the accuracy of a geoscientific simulation model with varying parameters. This is an important step during the model building phase, as it helps the analyst to determine and refine meaningful configurations of the model. The considered simulation model reproduces the influence of ice masses on the shape of the Earth surface. The output of this model is the elevation of geographical locations on Earth. By comparing simulated elevations with observations, the model’s accuracy can be assessed. More precisely, the accuracy is determined by the percentage of observations that are matched by the simulation data. Usually, the accuracy is not homogeneous in geographical space. In some regions, the model may fit the observations well, while it strongly deviates in others. Hence, it is important to assess accuracy with respect to geographical positions.

In Fig. 2, we show the result of our visualization method. The simulation model has two parameters that both take four different values, resulting in 16 different parameter settings. The spatial distribution of accuracy of each parameter setting is displayed in one map, covering the area of Northern Europe that encompasses the Fennoscandia region in the example. Each colored area is a region whose accuracy is measured. We use a discrete color scheme with four classes, to indicate whether the accuracy in a region is zero, “low”, “medium”, or “high”. Only the information important for the comparison is presented: the general geospatial relations among regions and their corresponding accuracy values. A first feedback from geoscientists indicates that this visualization is useful as an overview on the accuracy of models. Further, thanks to the intuitive representation in geospatial context, the visual representation is a valuable basis to communicate and discuss intermediate modeling results among scientists.

![Fig. 2. An example for small multiples in combination with schematized outlines. The underlying simulation model uses two parameters: the first is varied along the vertical axis, the second along the horizontal axis. The color of a region indicates its accuracy.](image)

### 4 Conclusion

We proposed to combine cartographic schematization techniques with small multiples of a geographical map to obtain a representation that allows to compare different spatial configurations of values. Schematization reduced the information density in each instance of the map. By presenting only the most relevant information, we believe that the usability of small multiples for comparison tasks improves.

### Acknowledgments

The authors owe thanks to Volker Klemann and Sven Schulte for providing data and helpful comments.

### References


