

studies, allowing for better knowledge of the target phenomenon, prediction of new occurrences, and the organization of interventions aimed at prevention.

- BAILEY, T. C. & GATRELL, A., 1995. *Iterative Spatial Data Analysis*. Longman: Scientific & Technical.
- SUSSER, M., 1994a. The logic in ecological I: The logic of analysis. *American Journal of Public Health*, 84:825-829.
- SUSSER, M., 1994b. The logic in ecological II: The logic of design. *American Journal of Public Health*, 84:830-835.

---

*Claudio J.  
Struchiner*

*Programa de  
Computação Científica,  
Fundação Oswaldo  
Cruz, Rio de Janeiro,  
Brasil.*

First I would like to express my appreciation to the authors for this impressively wide-ranging paper. It is a review paper that provides an introduction to geocomputation techniques, i.e., computer-intensive techniques for knowledge discovery in physical and human geography. The authors seem to favor the view that this new interdisciplinary area is to be distinguished from the simple extension of statistical techniques to spatial data. My comments are motivated by questions I have posed to myself after reading their review: How do such methods compare to established techniques? What are their advantages and disadvantages? What are their ranges of applications? Do the new techniques challenge or extend any of the existing paradigms in data analysis?

The computational dimension appears to be the common denominator of the techniques described in this review and goes into the definition of the key concept at stake, geocomputation. Faster and more powerful computers and advances in software engineering have had a profound impact on all areas of statistics. Bootstrap and Monte Carlo Markov Chain (MCMC) methods, for example, allow the estimation of parameters in richer and more realistic model-based representations of natural phenomena, thereby freeing the imagination of the scientific community. In this context, the boundaries of statistical models and statistical theory have been extended, while preserving the current paradigms, i.e., good statistical thinking is based on solid philosophical principles.

Algorithmic thinking also plays an important role in other areas of science. Complex systems can be generated through the use of

very simple building rules, which resemble the functioning of DNA chains. In this context, computer-intensive algorithmic techniques are intimately related to the mechanisms of pattern formation that supposedly occur in nature. In opposition, the procedures under the heading of geocomputation also seek to uncover pattern formation, but their search mechanisms are general in nature and do not bear any relationship to the various possible mechanisms that generate those spatial patterns.

In my view, the geocomputational methods reviewed in this paper do not share the same principles as these extensions. These algorithmic techniques appear to be a computerized version similar, in spirit, to a once very fashionable set of techniques developed by J. Tukey and known as Exploratory Data Analysis. Other statistical techniques put together under different headings such as Data-Driven Procedures and Data Mining attempt to answer similar questions raised here, i.e., "Are there any patterns, what are they, and what do they look like?"

The literature on quantitative methods has acknowledged, at least since the beginning of this century, the existence of two dimensions in research practice, i.e., exploratory versus analytical. For example, R. Ross opposed the concepts of *a priori* versus *a posteriori* pathometry in his Theory of Happenings. Most textbooks make a distinction between descriptive and analytical epidemiology. The debate seems endless and can be naively put by such questions as: "Are there purely descriptive studies? Without knowing what one is looking for, how can one tell when one has found it? If there is some previous knowledge or intuition of a subject, why not make it explicit in a model and see how the available empirical evidence modifies this knowledge or intuition? Do pattern-discovery algorithms carry some sort of built-in intelligence?"

Therefore, by analogy with other computer-intensive techniques mentioned above, one could wonder whether geocomputation, and other modern exploratory data analysis techniques, could benefit from incorporating a causal structure or more specific pattern-formation mechanisms.