Renato Martins Assunção

Departamento de Estatística, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

Câmara and Monteiro have the merit of drawing the attention of epidemiologists to new spatial analysis techniques. They have done a good job of summing up the main methodologies recently developed and presenting examples of their use, along with recent bibliographical references. As a statistician, I wish to focus my comments on the relationship between the area referred to as geocomputation by the authors and the usual statistical methods.

Before emphasizing differences, it is necessary to identify commonality. A discipline is merely a label for a set of knowledge and practices exercised by people who use that label to refer to themselves. Such practices and knowledge change dynamically, and statistics is no exception. Currently, many of the methods described by the authors are found in the best and most traditional statistics journals. In particular, the first two topics, GAM (Geographical Analysis Machine) and local spatial statistics are topics of articles and books by statisticians interested in spatial analysis. The other two, neural networks and cellular automata, are less common but not totally absent. As the authors point out, the presence of these topics is due to the current combined availability of data and computational power. I still do not feel comfortable in identifying geocomputation as a defined field of work, since most of the techniques presented emerged in traditional contexts (as in the case of the first two topics) or non-geographical ones (the last two). But this is not relevant for using and learning the techniques presented by the authors, which are useful regardless of labels.

Although the latter two topics, neural networks and cellular automata, are not absent from the statistical literature, they are less present than might be expected nowadays. Thus, one might ask, what can statisticians learn from researchers in neural networks and cellular automata? I believe that we should be less concerned with asymptotic results and optimality and seek methods that function well with large databases. We should deal with large and difficult problems involving a large number of parameters and depending little on hypotheses that cannot be verified. We should be alert to algorithms like the steepest descent with learning rates that can be highly useful in order to avoid over-adjustment in models with many parameters. This could be useful mainly for Bayesian models, which have become increasingly important (Assunção et al., in press).

On the other hand, what can researchers of neural networks and cellular automata learn from statisticians? I believe they should be a little more concerned with the statistical properties of their methods and perhaps slightly more with their optimality. They should make greater effort to compare their methods with others, including simpler traditional statistical methods. Often a linear regression can have a performance similar to that of a multi-layer perceptron. Contrary to what the authors state at the end of their article, the results of techniques like those presented are comparable. A clear example of this is the book by Alexander & Boyle (1997), which presents various techniques, including GAM, used by their respective creators in a set of simulated maps which might or might not present disease foci. The process of generating maps was described to the researchers, but not what each particular map contained. This simulation exercise, although displaying limitations, served to clearly demonstrate that some methods should be abandoned once and for all, and the GAM was not among them.

The techniques and examples presented in the article are a good sample of what new computer-intensive methods can offer. Research on these methods is increasing continuously and will no doubt continue over this decade. The authors are to be congratulated for having raised the topic and for having motivated health researches to take interest in these new methods.

ALEXANDER, F. E. & BOYLE, P., 1997. *Methods for Investigating Localized Clustering of Disease*. Oxford: Oxford University Press.

ASSUNÇÃO, R. M.; REIS, I. A. & OLIVEIRA, C. L., in press. Diffusion and prediction of leishmaniosis in Belo Horizonte, a large Brazilian metropolitan area: A Bayesian spatial-temporal model. *Statistics in Medicine*.