The authors reply
Os autores respondem

Gilberto Câmara
Antônio Miguel
Monteiro

The authors wish to thank the reviewers for their extensive and careful comments on our paper. We found that most of the remarks complement our work, either further developing certain issues or elucidating differences in our approach to spatial statistical theory. We fully concur with Bailey's remarks that geocomputational techniques should be used to complement more traditional statistical approaches, not as alternatives to them. Indeed, some of the techniques we presented (such as the GAM algorithm) can be used as *a priori* data-mining techniques to investigate data-rich environments. After a significant pattern is found, model-fitting approaches can be applied more effectively. We take note of Bailey's thoughtful comments on the dangers of applying techniques that are not easily associated with statistical measures of sensitivity and robustness, such as neural network or cellular automata.

Although we understand the cautious approach of many of our reviewers, we wish to indicate, as pointed out by both Nobre and Struchiner, that there are circumstances in which exploratory and non-robust geocomputational techniques are useful. Indeed, there are cases where the statistical alternatives are either extremely complex to apply or have yet to be fully developed. Let us consider two types of problems: *multi-dimensional spatial data analysis* and *dynamic spatio-temporal modeling*. As pointed out by Albuquerque, these types of problems arise when we are interested in studying not only the structural elements of space but also the effects of *processes*.

As for multidimensional spatial data analysis, our paper presents a typical situation in which a health researcher searches for areas prone to the incidence of a disease, given a number of possible environmental factors. This problem can be described in general terms as one of prediction, when it is assumed that a causal structure is in place. We proposed to use neural networks as one of several possible solutions to this problem. In his comments, Nobre made the important point that neural networks are "black boxes" in the sense that they do not provide the ability to explain the reasoning used to arrive at a result. Nevertheless, they provide a viable practical alternative to an otherwise difficult problem with traditional techniques, since establishing a spatial correlation structure for such a problem may prove almost intractable from a statistical viewpoint.

In the case of dynamic spatio-temporal modeling, a researcher may be interested in representing geographical space in a detailed way, e.g., a matrix of cells where each cell has unique characteristics. If one is interested in establishing the conditions of disease propagation in such an environment, the sheer size of the problem and the number of variables required to provide a realistic prediction may make the statistical approach unfeasible. In this case, approaches such as cellular automata guided by econometric equations provide a first approximation to an answer.

Assunção makes the important point that *ad hoc* techniques such as neural networks and cellular automata would benefit substantially from the use of statistical techniques for establishing optimality properties and for a better characterization of their variability and the relative impact of each factor. We heartily agree and certainly consider this an important research topic in geocomputation.

A final comment is in order. Geocomputation will not provide any "silver bullets". Indeed, some of its early proponents have in a sense jeopardized its acceptance by the scientific community at large by making rather preposterous claims. For problems where clean and robust statistical techniques are available, these should be used instead of ad hoc approaches like neural networks or cellular automata. However, we hope to have pointed out that there are many situations in which current state-of-the-art statistical methods are not applicable. In such cases, researchers should be encouraged to use the modeling capabilities of computers and allow for different representations of geographical space. It is inevitable and indeed desirable that such a pragmatic approach will bring a new understanding to the analysis of health data sets.