Tessellation Gibbs model for agricultural landscapes : parameter estimation

Master internship project proposal

Hosting institution : INRAE, MaIAGE Research Unit (https://maiage.inra.fr/en)

Internship site : Jouy-en-Josas (Yvelines department)

Supervisors : Katarzyna Adamczyk (INRAE, MaiAGE), Radu Stoica (Université de Lorraine, IECL)

Starting date : march or april 2020, duration : 4 to 6 months

Contact : Katarzyna Adamczyk (Katarzyna.Adamczyk@inra.fr)

Context Agricultural landscape simulators are used to study the impact of landscape elements on spatial processes in agro-ecology, such as spore dispersal, gene flow between species or suface runoff. They allow for studying the processes at a large spatial scale and thus account for the long-range interactions.
The MaIAGE Unit is developing a simulator of agricultural landscapes, approximated by polygonal tessellations. A polygonal tessellation of a two-dimensional domain is a finite subdivision of this domain into polygonal sets with disjoint interiors. The simulator is based on the Gibbs model of tessellations [1], defined by an energy function depending on tessellation summary statistics that we want to control. The model simulation algorithm is implemented in `R` library. Estimation of model parameters from landscape patterns enables simulating the tessellations with summary statistics centered on the values observed in a landscape (see figure 1)

Internship goal The model parameters are currently estimated by Monte Carlo Maximum Likelihood method. This approach is efficient when the number of statistics included in the model remains low. When the parametric dimension of the model increases, the convergence issues of the optimization algorithm appear. Yet, in order to simulate landscape-like tessellations we need to test high-dimensional sets of summary statistics, likely to be correlated. Thus we are looking for the estimation method that would be robust enough in this context.
The goal of the internship is to test alternative estimation approaches. In particular, we would like to compare MCML method with inference based on the ABC Shadow algorithm proposed in [3]. This algorithm allow the approximate sampling of the posterior distribution, hence hypothesis tests, confindence intervals and model choice procedures can be proposed in order to analyze landscape modelling.

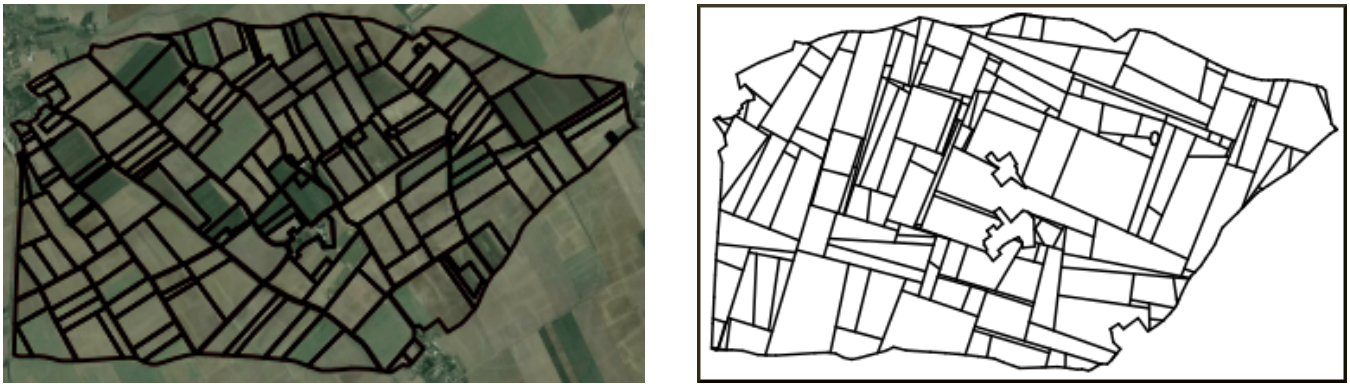
Candidate profile We are looking for a student in Applied Mathematics or Statistics with strong computer skills and knowledge of `R` language, motivated by applications in environmental sciences. First experiences in stochastic modelling and `MCMC` algorithms will be an asset.

References

[1] K. Kiêu, K. Adamczyk-Chauvat, H. Monod, and R. S. Stoica. A completely random T-tessellation model and Gibbsian extensions. *Spatial Statistics*, 6:118–138, 2013.

[2] C. J. Geyer. Markov chain Monte Carlo maximum likelihood. In *Computing Science and Statistics: Proceedins of the 23rd Symposium on the Interface*, pages 156–163, 1991.

[3] R. S. Stoica, A. Philippe, P. Gregori, J. Mateu. An ABC method for posterior sampling of marked point processes. *Statistics and Computing*, 27(5):1225–1238, 2017.



(a) (b)

Figure 1: (a) agricultural landscape, (b) simulation of the model [1] fitted to the landscape data.