Neutral modelling of agricultural landscapes by tessellation methods: the *GenExP-LandSiTes* software - Application to the simulation of gene flow

Le Ber, F. (1,6); Lavigne, C. (2); Adamczyk, K. (3); Angevin, F. (4); Colbach, N. (5); Mari, J.-F. (6), Monod, H. (3)

(1) UMR 7517 LHyGeS, ENGEES, F-67000 Strasbourg, florence.leber@engees.unistra.fr

(2) UR 1115 Plantes et systèmes de culture horticoles, INRA, F-84000 Avignon

(3) Unité MIA, INRA, F-78350 Jouy-en-Josas

(4) UAR 1240 Eco-innov, INRA, F-78850 Thiverval Grignon

(5) UMR 1210 Biologie et Gestion des Adventices, INRA, F-21000 Dijon

(6) UMR 7503 LORIA, F-54500 Vandœuvre-lès-Nancy

Abstract: We present a three steps approach that aimed at simulating neutral agricultural landscape models: (1) we characterized the geometry of three real field patterns; (2) we generated simulated field patterns with two tessellation methods attempting to control the value of some of the observed characteristics and, (3) we evaluated the simulated field patterns. The first two steps were integrated to the *GenExP-LandSiTes* software that thus simulates two-dimensional agricultural landscapes. It is written in Java, and it is freely accessible through a Gnu Public Licence. For the third step, we considered that good simulated field patterns should capture characteristics of real landscapes that are important for the targeted agro-ecological process. Real landscapes and landscapes simulated using either a Voronoi or a rectangular tessellation were thus compared when used as input data within The *MAPOD-maize* gene flow model. The results showed that the Voronoi tessellation performed better than the rectangular tessellation. In our ongoing research we consider random line-based tessellations constrained by a probability distribution penalizing the extreme values of targeted features (for example too large variability of cell areas). We also propose an algorithm for simulating such tessellations. The probability distribution parameters can be fitted from observed landscapes. This should result in generating tessellations similar to real patterns.

Keywords: neutral landscape models; gene flow; tessellation

Introduction

Neutral landscape models (NLM, sensu Gardner et al., 1987) provide random landscape structures as a baseline for comparison with real landscape patterns, or for an evaluation of landscape structure effects on ecological processes. Such models are termed neutral as they model no explicit process giving rise to the landscape pattern, they do model totally random or somewhat constrained covering of the area.

There are several reasons for using neutral models in land-use planning or agronomy. First, real data are not always available or are too specific and thus reduce the scope of application of the model results. Second, in the case of anthropogenic landscapes, it is necessary to prospect new configurations in order to forecast their effects or to find the best configuration with respect to the given agro-ecological process. Finally, neutral landscapes can be used to test the sensitivity of process models to the spatial variability of agricultural landscapes.

We propose to simulate agricultural landscapes, including both their configuration (the field pattern) and their composition (occurrence of categories of land-use) (Li and Reynolds, 1994). As mentioned by Gaucherel et al. (2006b), agricultural landscapes cannot be easily simulated with traditional ecological neutral landscape models for two reasons: (1) they are mainly geometrical, contrarily to less anthropogenic landscapes, and (2) in traditional landscape models, the basic unit is the pixel, and the land-use mosaic (i.e., both the configuration and the composition) only emerges as a result of the simulation, whereas in agricultural landscapes, the field pattern is generally stable and "precedes" the land-use. For handling geometric patches as basic units, a straightforward way is to use tessellation methods, such as proposed by Gaucherel (2008).

However, various tessellation methods are available, that result in somewhat different geometries. In the following, we look for tessellation methods that yield geometrical patterns similar to a real landscape with respect to the features that are of importance for the targeted agro-ecological process. We illustrate the use of such methods in a study of a gene flow between GM and conventional crops.

1. The *GenExP-LandSiTes* software

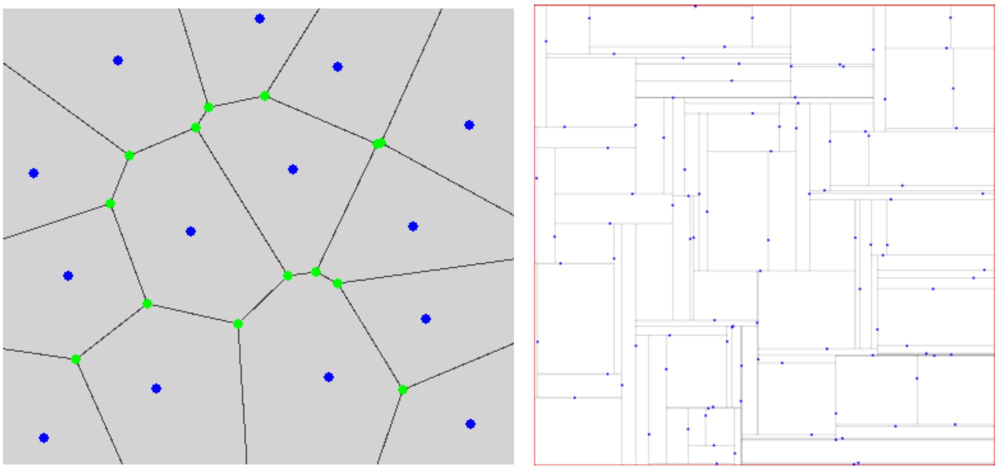
GenExP-LandSiTes is a software simulating two-dimensional agricultural landscapes. It is written in Java, and it is freely accessible through a Gnu Public Licence (http://www.loria.fr/~jfmari/GenExP/). Its main features fall into three categories: (1) the generation of the field patterns based on the simulation of spatial point-processes and on tessellation methods; (2) the allocation of land-uses; and (3) the computation of landscape statistics.

1.1 Generation of the field patterns based on the simulation of spatial point-processes and on tessellation methods

Generation of tessellation seeds. If tessellation seeds are to be generated in a stochastic manner, they can be considered as a realisation of a spatial (2D) point process, within a bounded window. Furthermore, if the seeds correspond to specific points of a landscape pattern, then the point process should fit the spatial distribution of these specific points. Eventually, if examples of such specific points are available, it is possible to estimate the point-process parameters in order to simulate similar configurations of seeds.

Spatial point processes are simulated via the statistical software R, interfaced with GenExp. The library “spatstat” of R (Baddeley and Turner, 2005) provides the complete set of tools for simulation and fit of point process models. The pairwise interaction model enables the generation of suitable patterns for fields centroids: it can yield either an aggregated or a regular centroid pattern and keep the distance between pairs of real centroids above a given threshold, proportional to the minimal field area. The model controls the mean number of points per unit area. The aggregated or regular point patterns are generated according to an interaction function that depends on the distance between the pairs of points.

Figure 1 : The Voronoi and rectangular tessellations



The Voronoi tessellation. Given a set of seed points in the Euclidean plan, a Voronoi tessellation – or Voronoi diagram – is a covering of the Euclidean plan with non-overlapping convex polygons, each surrounding a seed (fig 1, left). The points in a polygon are closer to its seed than to any other seed located in any other polygon. The edges therefore consist in points located at an equal distance from two seeds. The Delaunay triangulation (Okabe et al., 2000) can be used to calculate the polygons. It was implemented using the “3D Hull” algorithm for Delaunay triangulation (O’Rourke, 1998). The major advantage of the Voronoi tessellation is the correspondence between the pair polygon/seed and the pair field/centroid. For example, equally distanced seeds yield regular pattern of cells, aggregated seeds yield clusters of small cells and aligned seeds result in range-structured cells. Similarly, the configuration of field centroids characterises the spatial distribution of the fields: for instance, an aggregated pattern of centroids reveals clusters of small fields. Therefore a simulation of a model fitted from actual fields centroids should result in Voronoi patterns with some characteristics of the corresponding landscape pattern. The drawbacks of this tessellation are related to the shape of the polygons which are convex and often have a higher number of vertices than actual fields.

A rectangular tessellation. In a rectangular tessellation, the area is filled in with non-overlapping rectangles (fig 1, right). We eliminated the simple case when the rectangles are defined by two orthogonal sets of parallel lines and considered only those rectangles sharing T-vertices. In such tessellations, a vertex of a rectangle cannot belong to more than three rectangles (no X-vertex). A method for building such tessellations has been described in (Mackisack and Miles, 1996). The basic principle of the algorithm is to generate the edges of the rectangles by crossing two-directional rays starting from a set of points. In this case, there is no obvious correspondence between the rectangles and the fields, since the basic geometric unit is a segment rather than a polygon and we do not know where a seed is located along an edge. We found no specific points representative of a real landscape to be used as seeds for the rectangular tessellation. Nevertheless, the geometry of the rectangular cells yields a priori more realistic field shapes than the Voronoi cells, and certain overall characteristics of a landscape pattern can be conserved, for example, the number of the polygons.

Pattern replication. Several simulations of a chosen point process can be performed automatically. On these seeds, *GenExP-LandSiTes* performs the same tessellation algorithm, which leads to replicated field patterns with the same basic properties (e.g. number of fields, average area, nearest neighbour distance). This functionality is very useful both for testing the sensitivity of process models to the spatial variability of agricultural landscapes and for assessing the uncertainty associated to the outputs of these models.

Post-processing. After generating a tessellation, the user of *GenExP-LandSiTes* can carry out the following operations: deletion of polygons exceeding landscape limits (clipping); deletion of polygon sides shorter than a given value (fusion of vertices).

1.2 Allocating crops to fields using CarrotAge

Once the field pattern has been simulated, the user can allocate crops to the fields according to various methods. As a starting point, the user chooses a distribution of land-covers, and the land-cover mosaic is built randomly according to the probability of each crop.

GenExP-LandSiTes can also manage stochastic rules of crop successions. The knowledge of the land cover successions is carried out by a data-mining software –CarrotAge– that treats the successions of land covers as high-order Markov chains generated by a second-order Hidden Markov Models (HMM2) (Le Ber *et al.*, 2006). At time 0, *GenExP-LandSiTes* allocates crops to the fields according to a distribution provided by the user (for example: 50% maize, 30% wheat, 20% others); next, at successive time slots, crops are allocated to the fields following the crop successions, e.g. (maize / wheat) or (maize / maize), according to the probabilities of transitions given by the HMM2. The spatial and spatio-temporal correlations between the fields are not yet implemented. We are currently working on two major issues: *(i)* the spatial allocation of land covers based on a Gibbs process, as proposed by Gaucherel *et al.* (2006a), *(ii)* the combination of the temporal and spatial allocations of land covers with hierarchical hidden Markov models (Lazrak *et al.,* 2009).

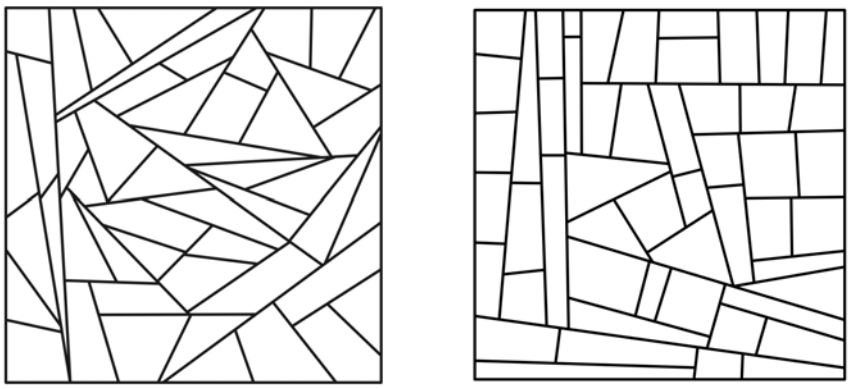
1.3 Computation of landscape statistics

GenExP-LandSiTes provides a library to calculate basic landscape descriptors (area, perimeter, number of vertices, centroid, shape index) in order to characterise the simulated field patterns and to compare them with actual landscapes. The results can be represented as histograms and the user can simultaneously visualise those concerning several landscapes. Furthermore, the coupling with *R* provides high-level tools for statistical analysis of the simulated field patterns.

1.4 Ongoing developments

A main improvement for the short-term concerns the implementation of new tessellation methods. In our ongoing research (Adamczyk et al. 2008) we consider random line-based tessellations constrained by a probability distribution penalizing the extreme values of some targeted features. The model currently implemented penalises a too high variability of cells areas and too small angles between tessellation vertices (see Figure 2). Preliminary results seem promising. We did not simulate yet gene-flow on these resulting field patterns.

Figure 2 :Ttessellation model accounting for (a) variability of cells area (at left), (b) variability of cells area and smallest angle between the adjacent vertices (at right)



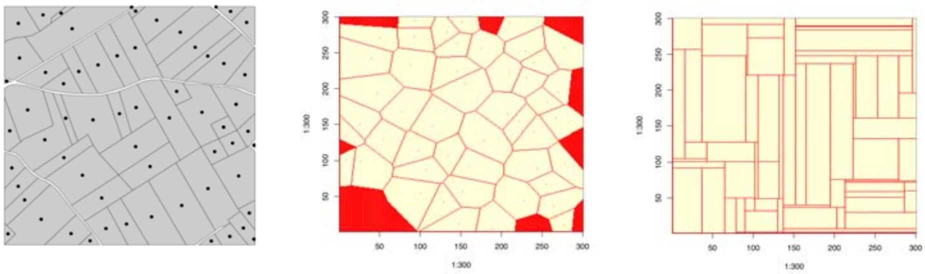
2. Simulating gene flow over neutral agricultural landscapes

2.1 Comparison of original and simulated landscapes

Three French landscapes were chosen for this study. These original landscapes (namely P1, A1, S4), restricted to a 1.5 km×1.5 km window, contained respectively 175, 100 and 63 fields and differed mainly in mean field area (P1:1.09 ha, A1:2.05 ha and S4:3.55 ha respectively) and variability of shapes (shape index[1]: P1:1.50±0.31, A1:1.53±0.28, S4:1.48±0.19). For each region, the pairwise-interaction model was fitted to each pattern of field centroids and five simulated landscapes were generated using each tessellation method (see Figure 3 for the S4 original pattern and the associated simulated field-patterns).

Comparisons of original and simulated landscapes indicate that both tessellation methods produced a correct number of fields of a correct average area. However, the Voronoi tessellation produced fields that were on average somewhat more compact (e.g. shape index: Voronoi: P1:1.32 ± 1.8×10^{-3}, A1:1.34 ± 0.6×10^{-3}, S4:1.32 ± 3.3×10^{-3}) and more similar within one field pattern (e.g. standard deviation of shape index: Voronoi: P1:0.10 ± 1.2×10^{-2}, A1:0.13 ± 1.4×10^{-2}, S4:0.11 ± 2.7×10^{-2}). Contrarily, the rectangular tessellation produced fields that were more elongated (shape index: P1:1.70 ± 0.12, A1:1.75 ± 0.11, S4:1.73 ± 0.19) and more dissimilar within each field pattern (standard deviation of shape index: P1:0.91 ± 0.12, A1:0.97 ± 0.15, S4:0.98 ± 0.18) (Le Ber et al. 2009).

Figure 3: Original landscape (S4), Voronoi and rectangular tessellations



2.2 Comparison of tesselation methods via their impact of cross-pollination

The *GenExP-LandSiTes* software was combined with *Mapod-maize* –a spatially explicit pollen dispersal model (Angevin *et al.*, 2008)– and we assessed the influence of landscape structures on maize gene flow from GM to conventional fields. For each field pattern, we determined the average impurity rate of conventional fields and the proportion of fields with impurity rate above 1% or 0.1%.

Gene flow simulations on Voronoi tessellations generated outputs that were closer to those on the original landscapes, compared to rectangular tessellations (Table 1). When considering a threshold of 1%, however, the two tessellation methods perform similarly (not shown). Comparing the results for the three original patterns (Figure 4) actually shows – at least for the 50% GM maize case where the various patterns give different results – that the rectangular tessellation is better than Voronoi for one pattern (S4) which has rather

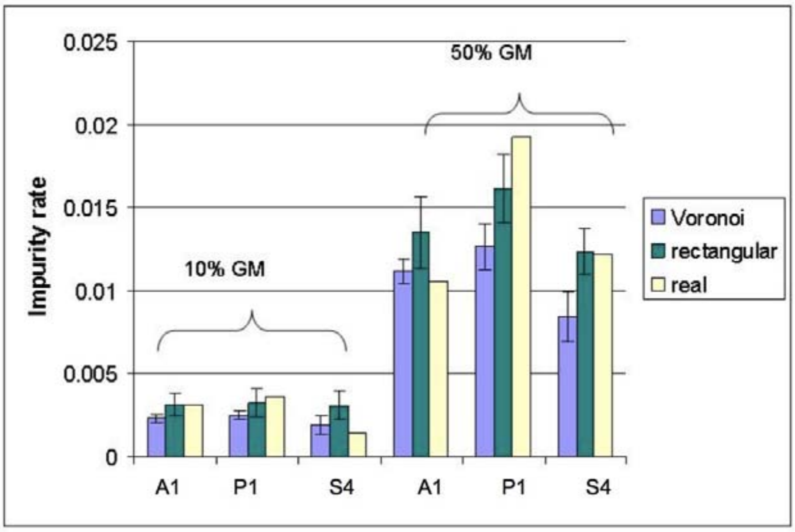
[1]Shape index = perimeter/4√area varies between 1 (for a square) and for example, 2 for a rectangle which length equals 15 times its width

rectangular and homogeneous field shapes. On the contrary the Voronoi tessellation better fits the A1 pattern which has medium sized fields, with variable shapes.

Table 1: Least Square means ± standard deviation of outputs of Mapod-maize on either original or simulated field patterns.

	original	Voronoi	rectangular
log10(impurity rate):	−3.55 ± 0.14	−3.51 ±0.06	−3.18 ± 0.06.
arc sinus(proportion above 0.1%)$^{1/2}$:	0.88 ± 0.04	0.90 ± 0.017	0.98 ± 0.017.
arc sinus (proportion above 1%)$^{1/2}$:	0.52 ± 0.02	0.49 ± 0.01	0.51 ± 0.01.

Figure 4: GM impurity rates as a function of original patterns (A1,P1, S4) and percentage of GM crops over the landscape



Conclusion - perspectives

GenExP-LandSiTes is a first step towards the modelling of field patterns that can be used as a basis for the simulation of agro-ecological processes. It is freely accessible through a Gnu Public Licence (http://www.loria.fr/~jfmari/GenExP/). The tessellation methods studied here account for some key features of an agricultural landscape but still exhibit uncontrolled differences as compared to the original field pattern. Roughly, the Voronoi tessellation underestimates the variability observed in the agricultural landscape whereas the rectangular tessellation overestimates it. This results in lower predicted GM impurity rates with the Voronoi tessellation than with the rectangular tessellation. None of these two tessellations, however, appeared as the most appropriate as the similarity of impurity rates simulated on real and simulated field patterns depended on the characteristics of the real patterns. Differences between original and simulated field pattern characteristics are probably not specific to our field patterns: any other existing tessellation method is likely to inherit only partially real landscape properties. A more direct approach would consist in a tessellation method aiming specifically at landscape representation as the random line-based tessellation in our ongoing research (Adamczyk et al. 2008).

Acknowledgments

Original maps were provided by the Institute for the Protection and Security of the Citizen (Joint Research Centre of the European Union) and AUP (Agence Unique de Paiement / French Payment Agency CAP Support). We acknowledge support from the French national program "ACI OGM et environnement" and the INRA-INRIA project PAYOTE.

References

Adamczyk K., Kieu K., Monod H. and Stoica R., 2008. Tessellation models for agricultural landscapes: parametric inference, *Proceedings of the 7-th French-Danish Workshop on Spatial Statistics and Image Analysis in Biology*, 2008, Toulouse, France.

Angevin F., Klein E.K., Choimet C., Gauffreteau A., Lavigne C., Messéan A. and Meynard J.M., 2008. Modelling impacts of cropping systems and climate on maize cross pollination in agricultural landscapes: The MAPOD model, *European Journal of Agronomy*, 28(3), p. 471-484.

Baddeley A. and R. Turner, 2005. Spatstat: an R package for analyzing spatial point patterns, *Journal of Statistical Software*, 12(6), p. 1-42.

Gardner R.H., Milne B.T., Turner M.G. and O'Neill R.V., 1987. Neutral models for the analysis of broad-scale landscape pattern, *Landscape Ecology*, 1(1), p. 19–28.

Gaucherel C., Fleury D., Auclair D. and Dreyfus P., 2006a. Neutral models for patchy landscapes, *Ecological Modelling*, 197(1-2), p. 159-170.

Gaucherel C., Giboire N., Viaud V., Houet T., Baudry J. and Burel F., 2006b. A domain-specific language for patchy landscape modelling: The Brittany agricultural mosaic as a case study, *Ecological Modelling* 194 (1–3), p. 233–243.

Gaucherel C., 2008. Neutral models for polygonal landscapes with linear networks. *Ecological Modelling*, 219, p. 39–48.

Lazrak G., Mari J.F. and Benoît M., 2010. Landscape regularity modelling for environmental challenges in agriculture, *Landscape Ecology,* in press.

Le Ber F., Benoît M., Schott C., Mari J.F. and Mignolet C., 2006. Studying crop sequences with CarrotAge, a HMM-based data mining software, *Ecological Modelling*, 191(1), p. 170-185.

Le Ber F., Lavigne C., Adamczyk K., Angevin F., Colbach N., Mari J.F. and Monod H., 2009. Neutral modelling of agricultural landscapes by tessellation methods - Application for gene flow simulation, *Ecological Modelling*, 220, p. 3536-3545.

Li H. and Reynolds J.F., 1994. A simulation experiment to quantify spatial heterogeneity in categorical maps, *Ecology*, 75(8), p. 2446--2455.

Mackisack M.S. and Miles R.E.. Homogeneous rectangular tessellations, *Advances in Applied Probability*, 28, p. 993-1013.

Okabe A., Boots B., Sugihara K. and Chiu S.N., 2000. *Spatial tessellations: concepts and applications of Voronoi diagrams*. John Wiley & Sons, Inc., 2nd edition.

O'Rourke J., 1998. *Computational Geometry in C.*, Cambridge University Press, 2nd edition.