

## Estimation des paramètres d'un modèle de Gibbs de tessellation en T : application à la modélisation des paysages agricoles.

### Proposition de stage M2

Des tessellations polygonales en T sont caractérisées par deux propriétés suivantes : 1) chaque sommet de tessellation possède exactement trois arêtes incidentes 2) deux parmi ces trois arêtes sont alignées. Les tessellations en T peuvent représenter différents motifs spatiaux, illustrés sur la Figure 1. Des modèles de tessellations aléatoires permettent de reproduire la variabilité de formes et de tailles, observée dans ce type de données. Le modèle de Gibbs de tessellation aléatoire en T, proposé dans [1], est défini par la fonction de potentiel qui dépend d'un ensemble de statistiques de tessellation qu'on cherche à contrôler. En jouant sur les valeurs de paramètres du modèle, on peut générer des tessellations avec des valeurs de statistiques-résumées plus ou moins grandes. Cette propriété confère au modèle la capacité de représenter un large éventail des configurations observées.

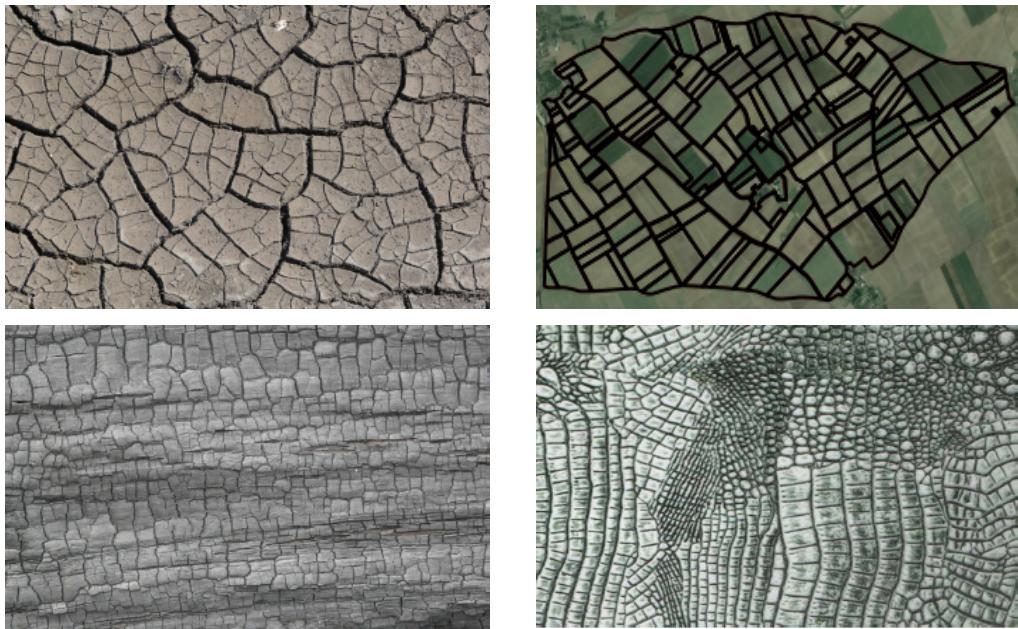


FIGURE 1 – Motifs qui peuvent être approximés par une tessellation en T : un sol craquelé, un paysage agricole, une texture de bois brûlé, un fragment de peau d'un reptile.

**Objectif de stage** L'objectif de stage sera de comparer différentes méthodes d'estimation des paramètres du modèle de Gibbs, défini par une densité non-normalisée. L'algorithme de simulation du modèle, proposé dans [1], rend possible l'application des méthodes basées sur l'échantillonnage dans la loi de tessellation. Parmi les méthodes testées on considérera tout d'abord la méthode de maximum

de vraisemblance Monte Carlo (MCML), proposée dans [2] pour des modèles de processus ponctuels et appliquée dans [6] au modèle de tessellation en T. La deuxième approche c'est l'estimateur de minimum de contraste, proposé dans [5]. La troisième groupe de méthodes sont des méthodes bayesiennes approchées ([3], [7]). L'utilisation de ces méthodes dans le contexte de tessellations gibbsiennes n'a pas été explorée.

La mission de stagiaire consistera à implémenter des algorithmes de calcul des estimateurs, en employant le simulateur du modèle, implanté dans une bibliothèque C++ RLiTe [6]. La comparaison des estimateurs s'effectuera tout d'abord sur des données synthétiques, en commençant par le modèle simple, contrôlant uniquement l'échelle de tessellation et en allant jusqu'à des modèles faisant intervenir plusieurs statistiques de tessellation, potentiellement corrélées. Des méthodes dont les performances s'avèrent les meilleures sur des données synthétiques seront ensuite utilisées dans des modèles à la base de générateur des parcellaires agricoles [4].

**Profil recherché** Une formation en mathématiques appliquées/statistiques ou une formation en agronomie avec des connaissances solides en statistiques. Les compétences en programmation ((C++, R) sont nécessaires. La connaissance des algorithmes de type MCMC sera un atout. Le candidat qui envisage de faire une thèse sera le bienvenu.

**Laboratoire d'accueil** Le stage sera réalisé à l'Institut Ellie Cartan de Lorraine, à Nancy. Il sera encadré par Katarzyna Adamczyk (INRAE, Inria-Pasta) et Radu Stoica (IECL, Inria-Pasta), en collaboration avec d'autres membres de l'équipe Pasta.

**Gratification de stage** Environ 550 euros en fonction de la législation actuelle.

**Dossier de candidature/échéances** Lettre de motivation, CV et le dernier bulletin de notes sont à adresser à Katarzyna.Adamczyk@inrae.fr avant le 20 décembre. L'arbitrage aura lieu avant le 15 janvier. Le stage commencera en février/mars et durera 5 à 6 mois.

## Références

- [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 : Proceedings 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.
- [4] K. Adamczyk-Chauvat, M. Kassa, K. Kiêu, J. Papaïx, and R.S. Stoica. Gibbsian T-tessellation model for agricultural landscape characterization. Preprint HAL, hal-02905984, 2020.
- [5] K. Kiêu and K. Adamczyk-Chauvat. Pseudolikelihood inference for Gibbsian T-tessellations... and point processes. Preprint HAL, hal-01234119, 2015.
- [6] K. Adamczyk-Chauvat and K. Kiêu. LiTe. <http://kien-kieu.github.io/lite>, 2015.
- [7] N. Vihrs, J. Møller and A.E. Gelfand Approximate Bayesian inference for a spatial point process model exhibiting regularity and random aggregation *Scandinavian Journal of Statistics*, 1–26, 2021.