



HAL
open science

Statistical inference for random T-tessellations models: application to agricultural landscape modeling.

Katarzyna Adamczyk-Chauvat, Mouna Kassa, Julien Papaix, Kiên Kiêu,
Radu S. Stoica

► **To cite this version:**

Katarzyna Adamczyk-Chauvat, Mouna Kassa, Julien Papaix, Kiên Kiêu, Radu S. Stoica. Statistical inference for random T-tessellations models: application to agricultural landscape modeling.. 11th International Conference on Spatio-Temporal Modelling METMA XI, Lancaster University, Jul 2024, Lancaster, United Kingdom. hal-04669756

HAL Id: hal-04669756

<https://hal.inrae.fr/hal-04669756v1>

Submitted on 9 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Statistical inference for random T-tessellations models: application to agricultural landscape modeling

Katarzyna Adamczyk-Chauvat¹

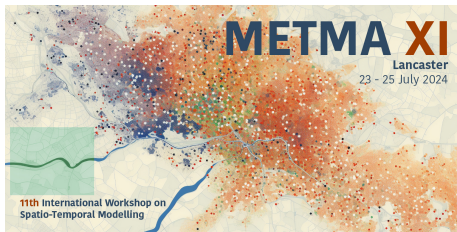
Kiên Kiêu¹

Mouna Kassa²

Radu S. Stoica³

Julien Papaix¹

¹INRAE, Dep. MathNum, Jouy-en-Josas, ²INSA, Rennes, ³Université de Lorraine-CNRS-Inria, Nancy



Outline

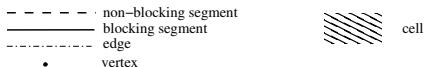
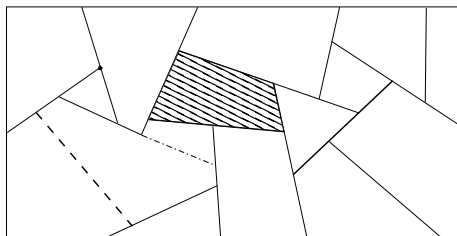
- 1 Random T-tessellation models
- 2 Statistical inference
- 3 Agricultural landscape model

T-tessellation

Definition

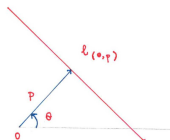
A planar tessellation of a bounded window W is called a *T-tessellation of W* if it satisfies two conditions:

- (i) all the vertices in $\text{int}(W)$ are of degree 3,
- (ii) two of the three edges connected by the same vertex are collinear.



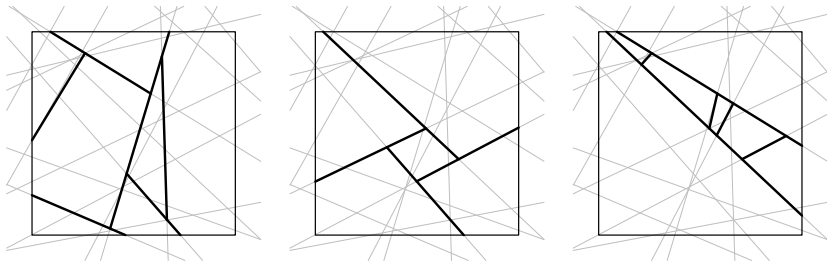
Introducing randomness

$$l_{(\theta, p)} \longleftrightarrow (\theta, p)$$



$$\mathcal{L} = \{(\theta, p) : l_{(\theta, p)} \cap W \neq \emptyset\}$$

T-tessellations supported by the same line set $L \subset \mathcal{L}$:



Poisson lines as a skeleton to define a random T-tessellation

Random T-tessellation models

Definition (Completely Random T-tessellation model)

Let \mathbf{L} be a unit-rate Poisson line process on \mathcal{L} . Let $\mathcal{T}(\mathbf{L})$ be the set of $T \in \mathcal{T}$ with segments lying on the lines of \mathbf{L} and such that for each such line $l_i, l_i \cap T$ consists of exactly one segment. The CRTT model is defined by:

$$\mu(A) = P(T \in A) = \frac{1}{Z} \mathbb{E} \left(\sum_{T \in \mathcal{T}(\mathbf{L})} \mathbb{1}_A(T) \right)$$

for $A \subseteq \mathcal{T}$ and Z - normalising constant.

Random T-tessellation models

Definition (Gibbs model)

Let $h : \mathcal{T} \rightarrow [0, \infty[$ be the non-negative function, integrable with respect to the measure μ . The Gibbs random T-tessellation model is defined by:

$$P(dT) \propto h(T)\mu(dT).$$

The function $U(T) = -\log h(T)$ is called the energy of the model.

Exponential family:

$$h_{\theta}(T) = \exp(t(T)^T \theta)$$

$t(T)$: vector of tessellation statistics, $\theta \in \Theta$: vector of model parameters.

Sampling from the Gibbs model

Metropolis-Hastings algorithm based on splits, merges and flips

Require: initial tessellation T_0 , density $h(T)$, (p_s, p_m, p_f) , uniform measures $(q_s^T(S), q_m^T(M), q_f^T(F))$.

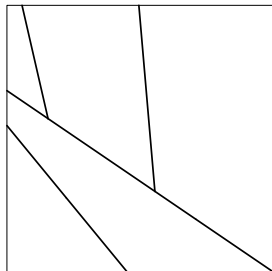
- 1: The current T-tessellation is T_n .
- 2: Draw the update type t from $\{s, m, f\}$ with probabilities (p_s, p_m, p_f) .
- 3: **if** there is no update of type t applicable to T_n **then**
- 4: $T_{n+1} = T_n$
- 5: **else**
- 6: Draw the update U of type t from the distribution $q_t^{T_n}(U)$.
- 7: Compute the Hastings ratio:

$$r_t(T, U) = \frac{h(U(T))}{h(T)} \frac{p_{t-1}}{p_t} \frac{q_{t-1}^{U(T)}(U^{-1})}{q_t^T(U)}$$

- 8: Accept the tessellation $T_{n+1} = U(T_n)$ with probability

$$\min\{1, r_t(T_n, U)\}.$$

- 9: **end if**
-



$\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$: sets of splits, merges, flips applicable to T

q_S^T, q_M^T, q_F^T : uniform measures on $\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$

Sampling from the Gibbs model

Metropolis-Hastings algorithm based on splits, merges and flips

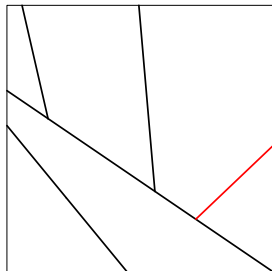
Require: initial tessellation T_0 , density $h(T)$, (p_s, p_m, p_f) , uniform measures $(q_s^T(S), q_m^T(M), q_f^T(F))$.

- 1: The current T-tessellation is T_n .
- 2: Draw the update type t from $\{s, m, f\}$ with probabilities (p_s, p_m, p_f) .
- 3: **if** there is no update of type t applicable to T_n **then**
- 4: $T_{n+1} = T_n$
- 5: **else**
- 6: Draw the update U of type t from the distribution $q_t^{T_n}(U)$.
- 7: Compute the Hastings ratio:

$$r_t(T, U) = \frac{h(U(T))}{h(T)} \frac{p_{t-1}}{p_t} \frac{q_{t-1}^{U(T)}(U^{-1})}{q_t^T(U)}$$

- 8: Accept the tessellation $T_{n+1} = U(T_n)$ with probability $\min\{1, r_t(T_n, U)\}$.

- 9: **end if**



$\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$: sets of splits, merges, flips applicable to T

q_S^T, q_M^T, q_F^T : uniform measures on $\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$

Sampling from the Gibbs model

Metropolis-Hastings algorithm based on splits, merges and flips

Require: initial tessellation T_0 , density $h(T)$, (p_s, p_m, p_f) , uniform measures $(q_s^T(S), q_m^T(M), q_f^T(F))$.

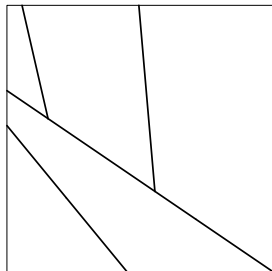
- 1: The current T-tessellation is T_n .
- 2: Draw the update type t from $\{s, m, f\}$ with probabilities (p_s, p_m, p_f) .
- 3: **if** there is no update of type t applicable to T_n **then**
- 4: $T_{n+1} = T_n$
- 5: **else**
- 6: Draw the update U of type t from the distribution $q_t^{T_n}(U)$.
- 7: Compute the Hastings ratio:

$$r_t(T, U) = \frac{h(U(T))}{h(T)} \frac{p_{t-1}}{p_t} \frac{q_{t-1}^{U(T)}(U^{-1})}{q_t^T(U)}$$

- 8: Accept the tessellation $T_{n+1} = U(T_n)$ with probability

$$\min\{1, r_t(T_n, U)\}.$$

- 9: **end if**
-



$\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$: sets of splits, merges, flips applicable to T

q_S^T, q_M^T, q_F^T : uniform measures on $\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$

Sampling from the Gibbs model

Metropolis-Hastings algorithm based on splits, merges and flips

Require: initial tessellation T_0 , density $h(T)$, (p_s, p_m, p_f) , uniform measures $(q_s^T(S), q_m^T(M), q_f^T(F))$.

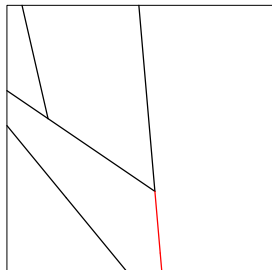
- 1: The current T-tessellation is T_n .
- 2: Draw the update type t from $\{s, m, f\}$ with probabilities (p_s, p_m, p_f) .
- 3: **if** there is no update of type t applicable to T_n **then**
- 4: $T_{n+1} = T_n$
- 5: **else**
- 6: Draw the update U of type t from the distribution $q_t^{T_n}(U)$.
- 7: Compute the Hastings ratio:

$$r_t(T, U) = \frac{h(U(T))}{h(T)} \frac{p_{t-1}}{p_t} \frac{q_{t-1}^{U(T)}(U^{-1})}{q_t^T(U)}$$

- 8: Accept the tessellation $T_{n+1} = U(T_n)$ with probability

$$\min\{1, r_t(T_n, U)\}.$$

- 9: **end if**
-



$\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$: sets of splits, merges, flips applicable to T

q_S^T, q_M^T, q_F^T : uniform measures on $\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$

Sampling from the Gibbs model

Metropolis-Hastings algorithm based on splits, merges and flips

Require: initial tessellation T_0 , density $h(T)$, (p_s, p_m, p_f) , uniform measures $(q_s^T(S), q_m^T(M), q_f^T(F))$.

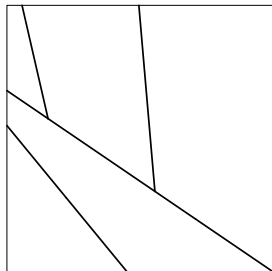
- 1: The current T-tessellation is T_n .
- 2: Draw the update type t from $\{s, m, f\}$ with probabilities (p_s, p_m, p_f) .
- 3: **if** there is no update of type t applicable to T_n **then**
- 4: $T_{n+1} = T_n$
- 5: **else**
- 6: Draw the update U of type t from the distribution $q_t^{T_n}(U)$.
- 7: Compute the Hastings ratio:

$$r_t(T, U) = \frac{h(U(T))}{h(T)} \frac{p_{t-1}}{p_t} \frac{q_{t-1}^{U(T)}(U^{-1})}{q_t^T(U)}$$

- 8: Accept the tessellation $T_{n+1} = U(T_n)$ with probability

$$\min\{1, r_t(T_n, U)\}.$$

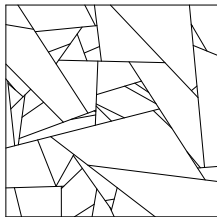
- 9: **end if**
-



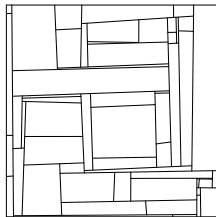
$\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$: sets of splits, merges, flips applicable to T

q_S^T, q_M^T, q_F^T : uniform measures on $\mathbb{S}_T, \mathbb{M}_T, \mathbb{F}_T$

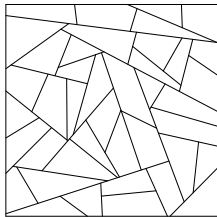
Examples of model simulations



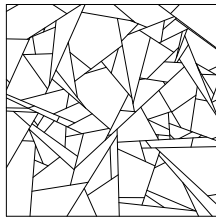
CRTT



Angles



Areas



ACS

Additional tools

Definition

Let $\mathcal{C}_s = \{(s, T) : T \in \mathcal{T}, s \in \mathbb{S}_T\}$. The split Campbell measure of $\mathbf{T} \sim P$ is defined by:

$$C_s^!(B) = \mathbb{E} \sum_{s \in \mathbf{M}_T} \mathbb{1}_B(s, \mathbf{T} \setminus \{s\}) \quad \forall B \subset \mathcal{C}_s$$

$$C_s^!(ds, dT) = \underbrace{\frac{h(T \cup \{s\})}{h(T)}}_{\text{split Papangelou intensity } \lambda_s(s, T)} dsP(dT)$$

Analogous definitions hold for flips.

Outline

- 1 Random T-tessellation models
- 2 **Statistical inference**
- 3 Agricultural landscape model

Minimum contrast estimation

$$\begin{aligned} \rho(\theta; T) &= - \sum_{m \in \mathbb{M}_T} \log \lambda_{s,\theta}(m^{-1}, m(T)) + \int_{\mathcal{S}_T} \lambda_{s,\theta}(s, T) ds \\ &\quad - \sum_{f \in \mathbb{F}_T} \log \lambda_{f,\theta}(f^{-1}, f(T)) + \sum_{f \in \mathbb{F}_T} \lambda_{f,\theta}(f, T) \end{aligned}$$

Proposition

If $\mathbf{T} \sim P_{\theta^*}$ then for every $\theta \in \Theta$:

$$\begin{aligned} \mathbb{E}_{\theta^*} \rho(\theta; \mathbf{T}) &= - \int_{\mathcal{C}_s} \log \lambda_{s,\theta}(s, T) C_{s,\theta^*}^!(ds, dT) + \int_{\mathcal{C}_s} \lambda_{s,\theta}(s, T) ds P_{\theta^*}(dT) \\ &\quad - \int_{\mathcal{C}_f} \log \lambda_{f,\theta}(s, T) C_{f,\theta^*}^!(df, dT) + \int_{\mathcal{C}_f} \lambda_{f,\theta}(f, T) df P_{\theta^*}(dT) \end{aligned}$$

$\mathbb{E}_{\theta^*} \rho(\theta; \mathbf{T})$ has a global minimum at $\theta = \theta^*$.

$$\hat{\theta}_{MPL} = \arg \max_{\theta} (-\rho(\theta; T))$$

Statistical tools extended to T-tessellations

- Monte Carlo Maximum Likelihood to refine the pseudolikelihood estimation

$$\underbrace{\hat{L}(\theta; T)}_{\text{MC log likelihood}} = - \langle t(T), \theta \rangle - \log \frac{1}{n} \sum_{i=1}^n \exp(- \langle t(T_i), \theta - \theta_0 \rangle) + \text{const}$$

- Model assessment based on global envelope tests:

H_0 : observed T-tessellation $\sim P_\theta$

$F = (F(r_1), \dots, F(r_d))$: functional statistic of tessellation

$$\underbrace{F_1}_{\text{data}}, \underbrace{F_2, \dots, F_s}_{\text{model simulations}}$$

Outline

- 1 Random T-tessellation models
- 2 Statistical inference
- 3 Agricultural landscape model**

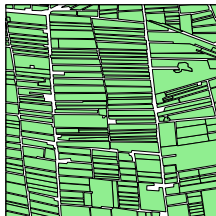
Landscapes approximated by T-tessellations



L_1 : Selommes



L_2 : Kervidy



L_3 : BVD

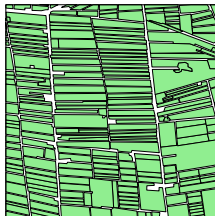
Landscapes approximated by T-tessellations



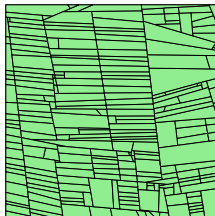
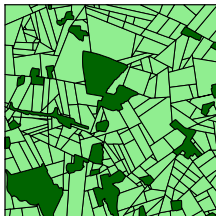
L_1 : Selommes



L_2 : Kervidy



L_3 : BVD



Model construction

Guidelines:

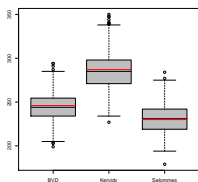
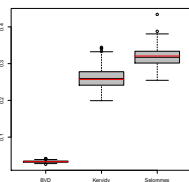
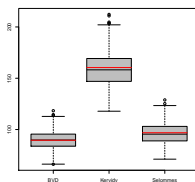
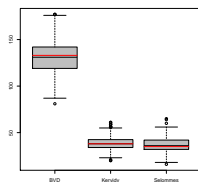
- accounts for typical landscape features
- highlights the main differences between patterns

$$\begin{aligned}
 U_{\theta}(T) &= \theta_1 \text{ number_of_segments}(T) && \leftarrow \text{scale statistic } t_1(T) \\
 &+ \theta_2 \sum_{\text{faces } f} \text{area}(f)^2 && \leftarrow \text{area statistic } t_2(T) \\
 &+ \theta_3 \sum_{\text{vertices } v} \left(\frac{\pi}{2} - \alpha(v) \right) && \leftarrow \text{angle statistic } t_3(T) \\
 &+ \theta_4 \sum_{\text{faces } f} g \left(\frac{\text{length}}{\text{width}}(f) \right) && \leftarrow \text{elongation statistic } t_4(T)
 \end{aligned}$$

Estimation results

Model parameter	Landscape	$\hat{\theta}_{MPL}$	$\hat{\theta}_{MCML}$
θ_1 (scale)	Selommes	1.62	1.99 (0.12)
	Kervidy	1.43	1.70 (0.11)
	BVD	2.32	2.11 (0.13)
θ_2 (areas)	Selommes	68.36	181.19 (66)
	Kervidy	23.73	79.19 (56)
	BVD	1361.46	2561.46 (679)
θ_3 (angles)	Selommes	0.95	2.23 (0.15)
	Kervidy	0.77	1.26 (0.12)
	BVD	1.72	2.21 (0.16)
θ_4 (elongation)	Selommes	0.48	0.28 (0.17)
	Kervidy	0.31	0.07 (0.15)
	BVD	-0.64	-0.63 (0.08)

Model checking


 $t_1(T)$

 $t_2(T)$

 $t_3(T)$

 $t_4(T)$

Distributions of energy statistics calculated over 500 model runs.

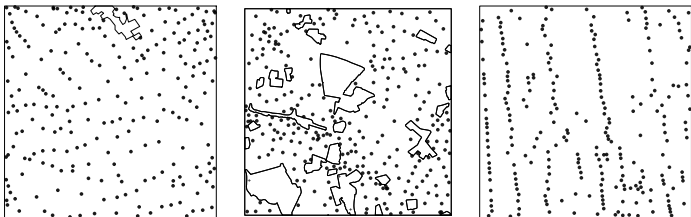


Observed and simulated landscape pattern.

Goodness-of-fit

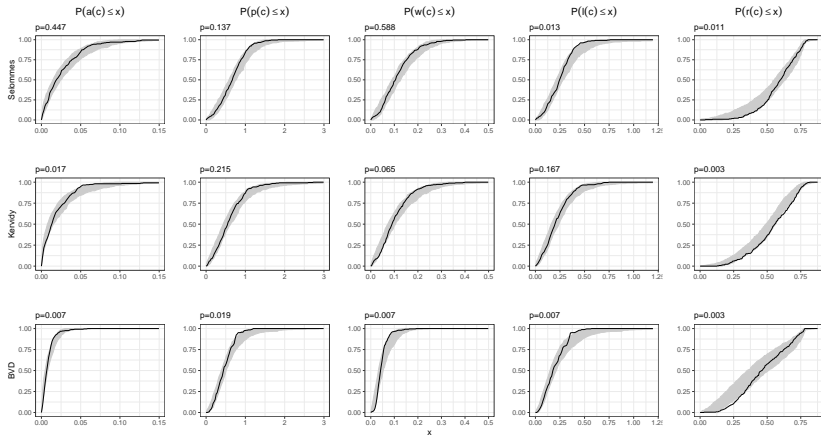
GET for tessellation and/or landscape metrics c.d.f.

- features of a typical cell: the area $a(c)$, the perimeter $p(c)$, the width $w(c)$, the length $l(c)$ of the smallest rectangle encompassing c and the isoperimetric quotient $r(c)$;
- spatial pattern of cell centroids : $G(r)$, $F(r)$, $L(r)$, $D(o)$ (density of the nearest neighbor vector orientation)

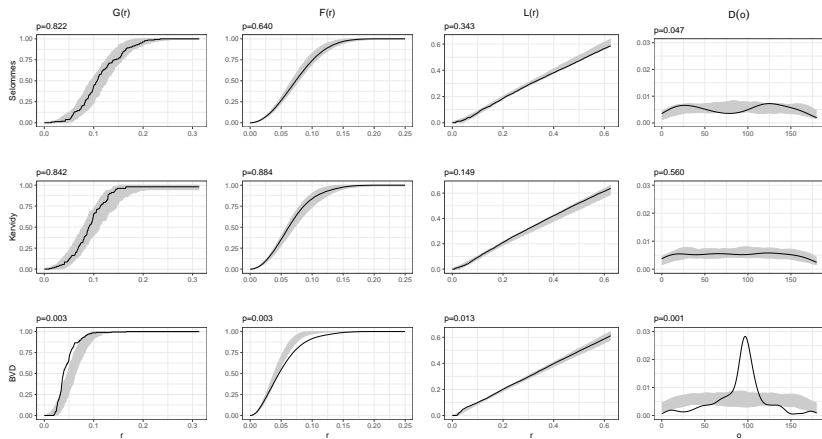


Cell centroid pattern

GET for the features of a typical cell







GET for the summary statistics of cell centroids



Concluding remarks

- Gibbs T-tessellation model: feature-driven model built on Poisson line skeleton;
- statistical methods: MPL estimation, adaptation of point process methods;
- application : characterization of agricultural landscapes... and other patterns that can be approximated by T-tessellations.

References

-  [1] K. Kiêu, K. Adamczyk-Chauvat, H. Monod, R.S. Stoica.
A completely random T-tessellation model and Gibbsian extensions.
Spatial Statistics, 6, 118-138 (2013).
-  [2] K. Kiêu and K. Adamczyk-Chauvat.
Pseudolikelihood inference for Gibbsian T-tessellations... and point processes.
Preprint HAL, hal-01234119, 2015.
-  [3] K. Adamczyk-Chauvat and K. Kiêu
<http://kien-kieu.github.io/lite>.
-  [4] K. Adamczyk-Chauvat, M. Kassa, J. Papaïx, K. Kiêu, R.S. Stoica.
Statistical inference for random T-tessellations models. Application to
agricultural landscape modeling.
Ann Inst Stat Math, 76, 447–479 (2024).