

Asymptotic properties of estimators for partially observed dependent spatial processes in a random environment.

Adélie Erard, Romain Lorrillière and Raphaël Lachièze-Rey

Séminaire de Probabilités du MAP5 – March 20, 2026



Breeding Bird Surveys



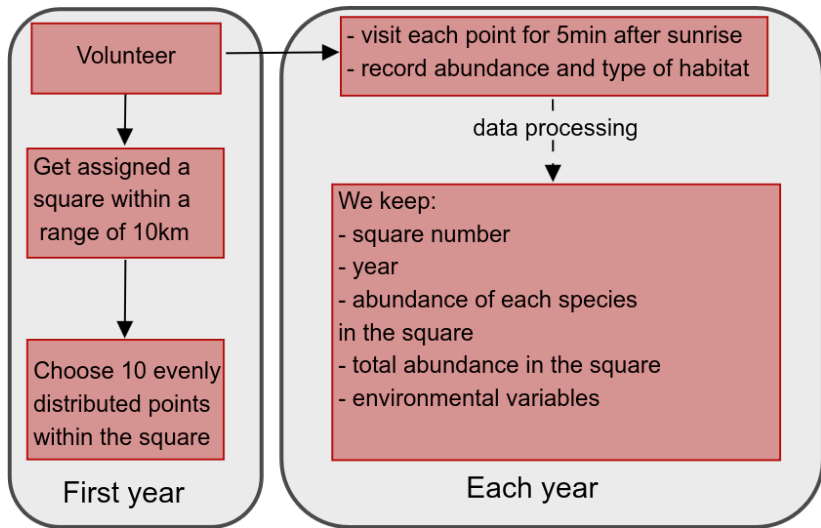
Meadow pipit (*Anthus pratensis*)

From Charles J. Sharp

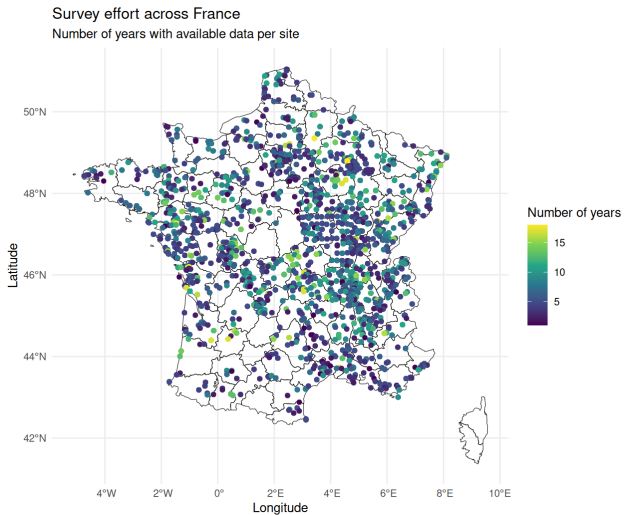
Breeding Bird Surveys (BBS) are long-term, large-scale, international avian monitoring programs designed to track the status and trends of bird populations.

Key features: standardized protocol, geographical and temporal coverage.

French BBS program (STOC)



What's in the data?



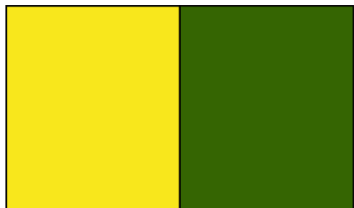
Environmental variables

For each observed point, we retrieve:

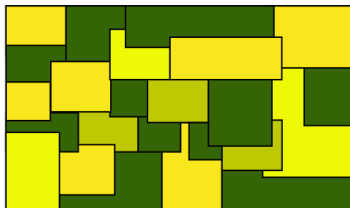
- Climate variables during spring (minimum and maximum temperature, total rain);
- Land uses in the square (% of agricultural, forest and urban land);
- Indices on how the agricultural land is fragmented 10km around the observation.

What did we already do? (But is more in the scope of an Ecology/ Applied Statistic seminar)

Question: Land sharing or land sparing?



Land Sparing



Land Sharing

What did we already do?

- We model bird counts using a **negative binomial regression model** to account for overdispersion.

What did we already do?

- We model bird counts using a **negative binomial regression model** to account for overdispersion.
- The expected count $\lambda(s, t)$ is modeled as:

$$\log \lambda(s, t) = \beta_0 + \sum_k \beta_k X_k(s, t) + \sum_{i, j} \beta_{i, j} X_i(s, t) X_j(s, t) + w_{s, t}$$

- ▶ $X_k(s, t)$: environmental covariate
- ▶ $w_{s, t}$: spatio-temporal randomness

What did we already do?

- We model bird counts using a **negative binomial regression model** to account for overdispersion.
- The expected count $\lambda(s, t)$ is modeled as:

$$\log \lambda(s, t) = \beta_0 + \sum_k \beta_k X_k(s, t) + \sum_{i,j} \beta_{i,j} X_i(s, t) X_j(s, t) + w_{s,t}$$

- ▶ $X_k(s, t)$: environmental covariate
- ▶ $w_{s,t}$: spatio-temporal randomness
- Inference with INLA [Rue et al. 2009]

What did we already do?

- We model bird counts using a **negative binomial regression model** to account for overdispersion.
- The expected count $\lambda(s, t)$ is modeled as:

$$\log \lambda(s, t) = \beta_0 + \sum_k \beta_k X_k(s, t) + \sum_{i, j} \beta_{i, j} X_i(s, t) X_j(s, t) + w_{s, t}$$

- ▶ $X_k(s, t)$: environmental covariate
- ▶ $w_{s, t}$: spatio-temporal randomness
- Inference with INLA [Rue et al. 2009]

Results: Land sparing seems to be better for agricultural birds with conservation concern.

Objective of the work

Give a method to **estimate the variation of abundance** (or future abundance) at a **local scale**.

Birth and death model

Individuals at a time are represented as a point process \mathcal{P} :

$$\mathcal{P} := \sum_{x \in \mathcal{P}} \delta_x$$

Birth and death model

Individuals at a time are represented as a point process \mathcal{P} :

$$\mathcal{P} := \sum_{x \in \mathcal{P}} \delta_x$$

Field of covariates: $\theta := (\theta(x), x \in \mathbb{R}^2)$

Birth and death model

Individuals at a time are represented as a point process \mathcal{P} :

$$\mathcal{P} := \sum_{x \in \mathcal{P}} \delta_x$$

Field of covariates: $\theta := (\theta(x), x \in \mathbb{R}^2)$

The transition to a new state \mathcal{P}' is governed, by:

- birth probability: $b(\tau_x \theta, \tau_x \mathcal{P})$
- death probability: $d(\tau_x \theta, \tau_x \mathcal{P})$

with τ_x a shift operator.

Effort rate

The observers form a random set in \mathbb{R}^2 , $E = (E(x), x \in \mathbb{R}^2)$, that is independent of \mathcal{P} and θ .

Effort rate

The observers form a random set in \mathbb{R}^2 , $E = (E(x), x \in \mathbb{R}^2)$, that is independent of \mathcal{P} and θ .

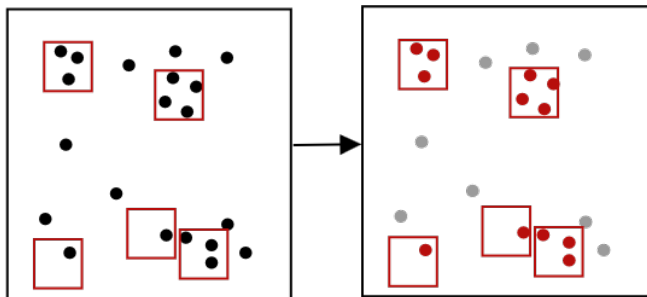
Example

$E_0 = \bigcup_e B(e, \rho)$ initial observed zone with e the location of the observers and ρ the (random) range of observation

To go to the next state of observed zone:

- each point is removed with constant rate;
- new points arrive according to a homogeneous Poisson process

Representation of the model



- birth and death process

□ observed zones

- observed individuals
- non observed individuals

(Non) stationarity assumptions

- No temporal stationarity nor equilibrium
- Not in an high density limit
- We assume our process to be spatially homogeneous

What do we want to predict?

Given a population \mathcal{P}_0 and covariates $\theta = (\theta(x), x \in \mathbb{R}^2)$.

What will be the new state \mathcal{P}_1 of the population around a point x at the next observation time?

What do we want to predict?

Given a population \mathcal{P}_0 and covariates $\theta = (\theta(x), x \in \mathbb{R}^2)$.

What will be the new state \mathcal{P}_1 of the population around a point x at the next observation time?

Two possible observables:

- Next year abundance around x : $N_1(x) = \#\{\mathcal{P}_1 \cap B(x, \rho)\}$
- Variation of abundance between the two states at x :
$$\Delta(x) = N_1(x) - N_0(x)$$

Construction of \hat{N}

Let (θ', \mathcal{C}) be a deterministic configuration of interest - somewhere where we have data at the initial state and want the abundance at the next state.

How do we construct our estimator?

Construction of \hat{N}

Let (θ', \mathcal{C}) be a deterministic configuration of interest - somewhere where we have data at the initial state and want the abundance at the next state.

How do we construct our estimator?

1. Construct a similarity function to compare two contexts.
2. Calculate similarity between (θ', \mathcal{C}) and known contexts (from the data).
3. Do a weighted mean to have an estimation of next year abundance (or variation of abundance).

Construction of \hat{N}

Let k be a similarity function. We define two statistics:

Construction of \hat{N}

Let k be a similarity function. We define two statistics:

$$F_{(\theta', \mathcal{C})}(\theta, \mathcal{P}_0) = \sum_{x_i \in E} N_1(x_i) k(\tau_{x_i}(\theta, \mathcal{P}_0), (\theta', \mathcal{C}))$$

Construction of \hat{N}

Let k be a similarity function. We define two statistics:

$$F_{(\theta', \mathcal{C})}(\theta, \mathcal{P}_0) = \sum_{x_i \in E} N_1(x_i) k(\tau_{x_i}(\theta, \mathcal{P}_0), (\theta', \mathcal{C}))$$

$$Z(\theta, \mathcal{P}_0) = \sum_{x_j \in E} k(\tau_{x_j}(\theta, \mathcal{P}_0), (\theta', \mathcal{C})).$$

Estimator

Let (θ', \mathcal{C}) be a deterministic configuration of interest.

$$\begin{aligned}\hat{N}_n^{(\theta', \mathcal{C})} &:= \frac{F_{(\theta', \mathcal{C})}(\theta, \mathcal{P}_n)}{Z(\theta, \mathcal{P}_n)} \\ &= \frac{1}{\sum_{x_j \in E} k(\tau_{x_j}(\theta, \mathcal{P}_n), (\theta', \mathcal{C}))} \sum_{x_i \in E} N_1(x_i) k(\tau_{x_i}(\theta, \mathcal{P}_n), (\theta', \mathcal{C}))\end{aligned}$$

where $\mathcal{P}_n := \mathcal{P} \cap [-n/2, n/2]^2$

What we want to show

Theorem

Under some conditions:

- 1. on the process \mathcal{P} ;*
- 2. on the covariate field θ ;*
- 3. on the similarity function k .*

We have:

What we want to show

Theorem

Under some conditions:

1. *on the process \mathcal{P} ;*
2. *on the covariate field θ ;*
3. *on the similarity function k .*

We have:

$$\left(\text{Var} \left(\hat{N}_n^{(\theta', \mathcal{C})} \right) \right)^{-1/2} \left(\hat{N}_n^{(\theta', \mathcal{C})} - \mathbb{E}[\hat{N}_n^{(\theta', \mathcal{C})}] \right) \xrightarrow{d} \mathcal{N}(0, 1).$$

Hypothesis on the process

Definition

The p-th order correlation function of \mathcal{P} are defined as follows:

$$\mathbb{E}\left[\prod_{i=1}^p \mathcal{P}(B_i)\right] = \int_{B_1 \times \dots \times B_p} \rho^{(p)}(x_1, \dots, x_p) dx_1 \dots dx_p$$

Hypothesis on the process

Definition (BYY25)

\mathcal{P} has exponentially fast mixing correlation if for all $p, q \in \mathbb{N}$ and all $(x_1, \dots, x_{p+q}) \in (\mathbb{R}^2)^{p+q}$,

$$|\rho^{(p+q)}(x_1, \dots, x_{p+q}) - \rho^{(p)}(x_1, \dots, x_p)\rho^{(q)}(x_{p+1}, \dots, x_{p+q})| \leq C_{p+q}\phi(s)$$

where ϕ is a fast correlation decay function verifying:

$$-\infty \leq \liminf_{r \rightarrow \infty} \frac{\log \phi(r)}{r^b} < 0, \quad 0 < b$$

A class of admissible process

Definition (Cox Process)

Let Λ be a locally finite random measure on \mathbb{R}^d . Let $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}_+$, and conditionally on $\Lambda = \lambda$, let Ψ be a Poisson point process with intensity measure λ . Then Ψ is called a Cox process with intensity measure Λ .

A class of admissible process

Proposition

Let Ψ be a Cox process with intensity Λ . We suppose that Λ has a (random) density with respect to the Lebesgue measure:

$$\Lambda(B) = \int_B \lambda(x) dx.$$

The p -th correlation function of Ψ is given by:

$$\rho^{(p)}(x_1, \dots, x_p) = \mathbb{E}\left[\prod_{i=1}^p \lambda(x_i)\right]$$

If $\{\lambda(x)\}_{x \in \mathbb{R}^d}$ has fast decay of correlation then Ψ has fast decay of correlation.

Log Gaussian Cox processes

Definition

Let Λ be the measure with density $\lambda(u) = \exp(G(u))$, where G is a Gaussian random field. Then the Cox process generated by Λ is a log-Gaussian Cox process.

Log Gaussian Cox processes

Definition

Let Λ be the measure with density $\lambda(u) = \exp(G(u))$, where G is a Gaussian random field. Then the Cox process generated by Λ is a log-Gaussian Cox process.

Proposition

If G is a stationary Gaussian random field and has fast decay of correlation, then the LGCP defined above verifies mixing property.

Example of Gaussian processes that work

Let G be a Gaussian field such that:

- $\mathbb{E}[G_x] = \mu$ and $\text{Cov}(G_x, G_y) = r(x - y)$ (G is stationary)
- $r(x - y) = \exp(-\beta|x - y|)$

Then it verifies the mixing hypothesis.

Hypothesis on the covariate field

Let θ_x be the value of the covariate field at a point $x \in \mathcal{P}$.

Definition (BYY25)

$(\theta_x)_{x \in \mathcal{P}}$ has exponentially fast mixing correlation if for all $p, q \in \mathbb{N}$ and all $(\theta_{x,1}, \dots, \theta_{x,p+q}) \in (\mathbb{R}^2)^{p+q}$,

$$\begin{aligned} & |\mathbb{E}[f(\theta_{x,1}, \dots, \theta_{x,p})g(\theta_{x,p+1}, \dots, \theta_{x,p+q})] \\ & \quad - \mathbb{E}[f(\theta_{x,1}, \dots, \theta_{x,p})]\mathbb{E}[g(\theta_{x,p+1}, \dots, \theta_{x,p+q})]| \leq C_{p+q}\phi(s) \end{aligned}$$

where ϕ is a fast correlation decay function verifying:

$$-\infty \leq \liminf_{r \rightarrow \infty} \frac{\log \phi(r)}{r^b} < 0, \quad 0 < b$$

Stabilization hypothesis

Consider the marked process $\bar{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{(x, \theta_x)}$ and the score function $\xi : (x, \theta_x, \bar{\mathcal{P}}) \mapsto \mathbb{R}$.

Stabilization hypothesis

Consider the marked process $\bar{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{(x, \theta_x)}$ and the score function $\xi : (x, \theta_x, \bar{\mathcal{P}}) \mapsto \mathbb{R}$.

Definition (BYY19)

The radius of stabilization, $R_x^{\bar{\mathcal{P}}}$, of ξ at x and on the marked process $\bar{\mathcal{P}}$ is defined as the smallest r such that:

$$\xi(x, \theta_x, \bar{\mathcal{P}} \cap B(x, r)) = \xi(x, \theta_x, (\bar{\mathcal{P}} \cap B(x, r)) \cup (\mathcal{A} \cap B^c(x, r))),$$

with $\mathcal{A} \subset \mathbb{R}^2$.

Stabilization hypothesis

Consider the marked process $\bar{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{(x, \theta_x)}$ and the score function $\xi : (x, \theta_x, \bar{\mathcal{P}}) \mapsto \mathbb{R}$.

Definition (BYY19)

The radius of stabilization, $R_x^{\bar{\mathcal{P}}}$, of ξ at x and on the marked process $\bar{\mathcal{P}}$ is defined as the smallest r such that:

$$\xi(x, \theta_x, \bar{\mathcal{P}} \cap B(x, r)) = \xi(x, \theta_x, (\bar{\mathcal{P}} \cap B(x, r)) \cup (\mathcal{A} \cap B^c(x, r))),$$

with $\mathcal{A} \subset \mathbb{R}^2$. ξ is said fast to be stabilizing on $\bar{\mathcal{P}}$ if for all $l \in \mathbb{N}$ there exists a fast decaying function ϕ such that:

$$\sup_{1 \leq n \leq \infty} \sup_{x_1, \dots, x_l \in [-n/2, n/2]^2} \mathbb{P}_{x_1, \dots, x_l}(R_x^{\bar{\mathcal{P}}} > t) \leq C\phi(t)$$

Admissible similarity function

Recall that we want to calculate similarity between a given context (θ', \mathcal{C}) and a context around an observed point x with:

$$\xi(x, \theta_x, \overline{\mathcal{P}}) = k(\tau_x(\theta, \mathcal{P}), \theta', \mathcal{C})$$

Admissible similarity function

Recall that we want to calculate similarity between a given context (θ', \mathcal{C}) and a context around an observed point x with:

$$\xi(x, \theta_x, \overline{\mathcal{P}}) = k(\tau_x(\theta, \mathcal{P}), \theta', \mathcal{C})$$

Easy case: we compare (θ', \mathcal{C}) with $\overline{\mathcal{P}} \cap B(x, \rho)$ with ρ a finite range of observation.

Admissible similarity function

Recall that we want to calculate similarity between a given context (θ', \mathcal{C}) and a context around an observed point x with:

$$\xi(x, \theta_x, \overline{\mathcal{P}}) = k(\tau_x(\theta, \mathcal{P}), \theta', \mathcal{C})$$

Easy case: we compare (θ', \mathcal{C}) with $\overline{\mathcal{P}} \cap B(x, \rho)$ with ρ a finite range of observation.

A more challenging case: we compare (θ', \mathcal{C}) with the n closest neighbor of x .

Example of admissible similarity function

In the easy case, we summarize (θ', \mathcal{C}) and $\bar{\mathcal{P}} \cap B(x, \rho)$ as vectors of \mathbb{R}^{p+1} : p covariates and abundance around the center of the observed area.

Example of admissible similarity function

In the easy case, we summarize (θ', \mathcal{C}) and $\bar{\mathcal{P}} \cap B(x, \rho)$ as vectors of \mathbb{R}^{p+1} : p covariates and abundance around the center of the observed area.

For all $x, y \in \mathbb{R}^{p+1}$ we can use:

$$k(x, y) = \frac{\sum_i \min(x_i, y_i)}{\sum_i \max(x_i, y_i)}$$

A useful CLT

Theorem (5.2 of BYY25)

Let $\bar{\mathcal{P}}$ be a marked point process of $\bar{\mathcal{N}}$ having exponential mixing correlations. Let $\xi : \mathbb{R}^d \times \mathcal{M} \times \bar{\mathcal{N}} \rightarrow \mathbb{R}$ be a score function that is:

- fast BL-localizing on finite windows of $\bar{\mathcal{P}}$;
- verifying the p moment condition on finite windows of $\bar{\mathcal{P}}$ for all $p \geq 1$.

If $\text{Var}(\mu_n^\xi) = \Omega(n^\nu)$ for $\nu > 0$. Then, as $n \rightarrow \infty$:

$$\left(\text{Var}(\mu_n^\xi)\right)^{-1/2} \left(\mu_n^\xi - \mathbb{E}[\mu_n^\xi]\right) \xrightarrow{d} \mathcal{N}(0, 1)$$

with $\mu_n^\xi = \sum_{x \in \mathcal{P}_n} \xi((x, U_x), \bar{\mathcal{P}})$

In our case

If \mathcal{P} has exponential mixing correlations and k verifies the stabilization hypothesis we have:

Lemma

$$(\text{Var}(F_{(\theta', \mathcal{C})}(\theta, \mathcal{P}_n)))^{-1/2} (F_{(\theta', \mathcal{C})}(\theta, \mathcal{P}_n) - \mathbb{E}[F_{(\theta', \mathcal{C})}(\theta, \mathcal{P}_n)]) \xrightarrow{d} \mathcal{N}(0, 1)$$

$$(\text{Var}(Z(\theta, \mathcal{P}_n)))^{-1/2} (Z(\theta, \mathcal{P}_n) - \mathbb{E}[Z(\theta, \mathcal{P}_n)]) \xrightarrow{d} \mathcal{N}(0, 1)$$

What about the convergence of the ratio?

In [BYY25] they also have a LLN under the same kind of conditions.

Using Slutsky's lemma and the continuous mapping theorem we should have the convergence of the ratio and thus the following result:

What about the convergence of the ratio?

In [BYY25] they also have a LLN under the same kind of conditions.

Using Slutsky's lemma and the continuous mapping theorem we should have the convergence of the ratio and thus the following result:

Theorem

$$\left(\text{Var}\left(\hat{N}_n^{(\theta', \mathcal{C})}\right)\right)^{-1/2} \left(\hat{N}_n^{(\theta', \mathcal{C})} - \mathbb{E}[\hat{N}_n^{(\theta', \mathcal{C})}]\right) \xrightarrow{d} \mathcal{N}(0, 1)$$

How does it work with a toy data set ?

Square	Year	Abundance	Environmental variables	Abundance next year
10295	2003	5		2
11158	2006	1		4
20204	2015	6		7
30363	2019	8		9
950294	2024	3		?

In our example we say that the similarity only depend on the number of geese and the habitat:

$$k(L_1, L_2) = \frac{1}{2} (\mathbb{1}_{\text{number of geese of } L_1 = \text{number of geese of } L_2} + \mathbb{1}_{\text{same habitat}})$$

Similarity matrix

Square	Year	Abundance	Environmental variables	Abundance next year	Similarity with square of interest
10295	2003	5		2	1
11158	2006	1		4	0.5
20204	2015	6		7	0
30363	2019	8		9	0.5
950294	2024	3		?	1

$$\hat{N} = \frac{2 \times 1 + 4 \times 0.5 + 7 \times 0 + 9 \times 0.5}{1 + 0.5 + 0 + 0.5 + 1}$$
$$= 2.125$$

With a more realistic similarity function

Recall the similarity function:

$$k(x, y) = \frac{\sum_i \min(x_i, y_i)}{\sum_i \max(x_i, y_i)}.$$

We will use this similarity on the lines of our dataset (L_i).

What with real data?

1. Divide into train and test datasets.
2. Similarity calculations:

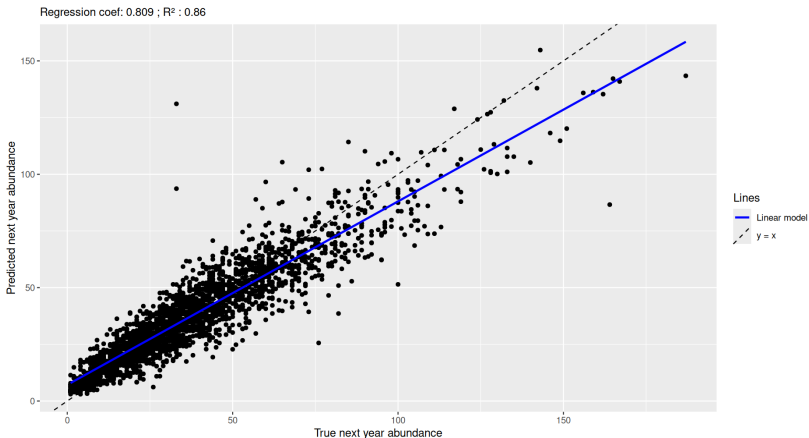
$$K(L_i, L_j) = \mathbb{1}_{k(L_1, L_2) \geq \alpha}$$

with L_i in the train set and L_j in the test set and α a threshold.

3. Apply estimator for each line in the test set :

$$\hat{N}^{L_j} = \frac{1}{\sum_{L_i} k(L_i, L_j)} \sum_{L_i} N(L_i) k(L_i, L_j)$$

Results with real data



Abundance of forest species with $\alpha = 0.8$

What's next?

- Show the mixing property for other kind of processes (Gibbs, Hawkes, ...)
- Simulations
- Try other similarity functions
- Add new environmental variables



Meadow pipit (*Anthus pratensis*)

From Charles J. Sharp.

Thank you !



Błaszczyszyn, B., D. Yogeshwaran, and J. E. Yukich (2019). “Limit Theory for Geometric Statistics of Point Processes Having Fast Decay of Correlations”. In: *The Annals of Probability* 47.2. Publisher: Institute of Mathematical Statistics, pp. 835–895. ISSN: 0091-1798. URL: <http://www.jstor.org/stable/26613430> (visited on 10/31/2023).



— (2025). “Limit theory for statistics of Lipschitz-localized stochastic processes in spatial random models”. In: *preprint*.