A method for estimating population growth at a local scale.

An application using French breeding bird surveys data.

Adélie Erard [1,2], Raphaël Lachièze-Rey[1] and Romain Lorrillière [2]

[1]MAP5, Université Paris Cité

[2]CESCO, Muséum national d'Histoire naturelle

March 3, 2025

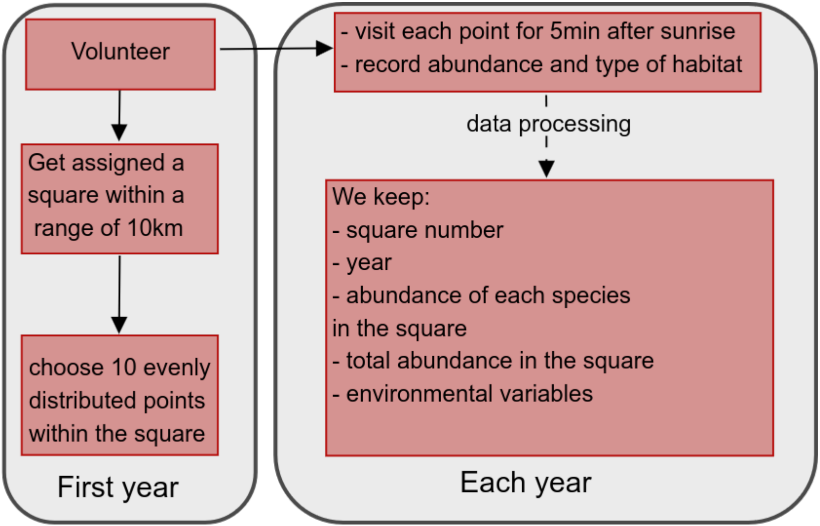
Breeding Bird Surveys

Our goal: giving a method to estimate the growth of a population at a local scale.

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)



Base layers of the model

Birds at time t are represented as a point process

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

Observers at time t are represented as a point process

$$O_t := \sum_{y \in O_t} \delta_y$$

Environmental variables are represented by a random field

$$\Theta(\cdot)$$

Observed zone

Let (O_t) be a spatial birth and death process with birth intensity $p_t(y)$ and death intensity γ

Let C_y be the square centered in y with 2km side for all $y \in X$

Then the observer process $(\mathcal{O}_t)$ is defined by the following closed random set:

$$\mathcal{O}_t = \bigcup_{y \in O_t \cap O_{t+1}} C_y$$

Final process

For all $x \in \mathcal{P}$ pose:

$$U_x = (\mathbb{1}_{x \in \mathcal{O}_t}, \Theta(\cdot - x))$$

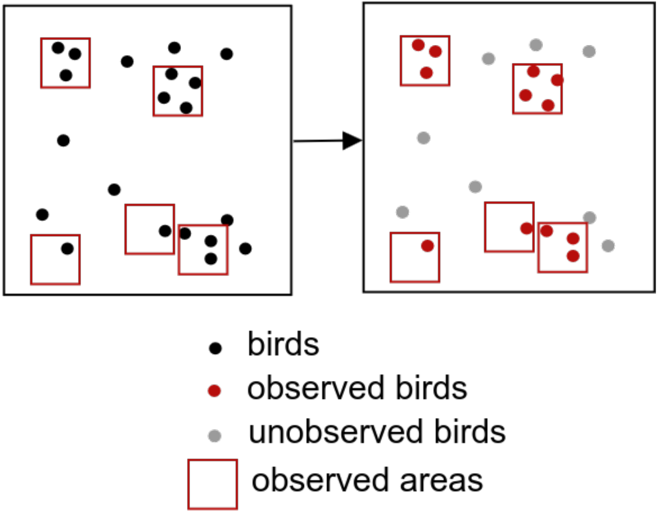
Local variation:

$$\beta_x := \beta(U_x, B(x, R) \cap \mathcal{P})$$

Final marked process:

$$\overline{\mathcal{P}} := \sum_{x \in \mathcal{P}} \delta_{(x, U_x, \beta_x)}$$

Representation of the model



Estimator

Let $\mathcal{C}$ be a deterministic configuration of marked points (x, U_x), set:

$$\hat{s}_n^{\mathcal{C}}(\overline{\mathcal{P}}) := \frac{1}{\sum_{x \in \mathcal{P}_n} k(\overline{C_x}, \mathcal{C})} \sum_{x \in \mathcal{P}_n} k(\overline{C_x}, \mathcal{C}) \beta_x$$

where k is a similarity function, and $\mathcal{P}_n := \mathcal{P} \cap [-n/2, n/2]^2$

Similarity function

Let x, y be two lines of the database and p the number of variables. We set:

$$k(x,y) := \frac{1}{p}\sum_{i=1}^{p} s_i(x,y),$$

$$s_i(x,y) := 1 - \frac{|x_i - y_i|}{R_i} \text{ for quantitative variables,}$$

$$s_i(x,y) := \mathbb{1}_{x_i = y_i} \text{ for categorical variables.}$$

Stabilization theory & convergence results

Theorem (2.1 of [Baryshnikov and Yukich, 2005])

If ξ is stabilizing and satisfies the p-th moment condition for some $p > 1$, then for all $f \in B(\mathbb{R}^d)$:

$$\lim_{\lambda \to \infty} \frac{\mathbb{E}[\langle f, \mu_\lambda^\xi \rangle]}{\lambda} = \int f(x)\mathbb{E}[\xi(0; \mathcal{P}_1)]dx$$

with $\mu_\lambda^\xi = \sum_{x \in \mathcal{P}} \xi(x, U_x, \mathcal{P})$

In our case:

$$\lim_{n \to \infty} \hat{s}_n^{\mathcal{C}}(\overline{\mathcal{P}}) = \mathbb{E}[\beta_{\mathcal{C}}]$$

Method for data analysis

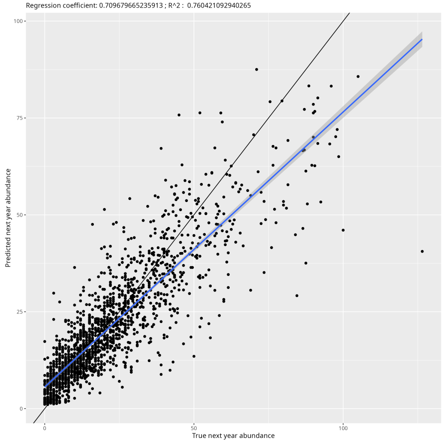
1. Calculate the similarity between lines based on all variables except the target;
2. Split the database into train and test;
3. For each line in test set, calculate $\hat{s}_n^{\mathcal{C}}(\overline{\mathcal{P}})$ with $\overline{\mathcal{P}}$ as the train set.

What's in the database?

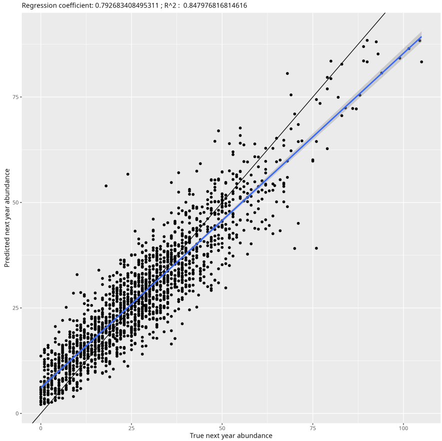
Square	Year	Abundance	Environmental variables	**Abundance next year**
10295	2003	5	🦜🦜 🦢 🦢 🌡️☀️ 🌲	2
11158	2006	1	🦢 🦢 🌧️🌡️ 🌽	4
20204	2015	6	🐦 🐦 🐦 ☀️ ❄️ 🌽	7
30363	2019	8	🐦 🐦 🦢🦜 🌧️ ❄️ 🌽	9
950294	2024	3	🦢 🦢 🐦 🌧️🌡️🌲	?

The abundance of the square 950294 in 2025 is estimated by taking the mean over the most similar rows, e.g first and second row, and thus should be 3.

Results with real data



Agricole species



Forest species

Perspectives

- Understand why it better works with certain species/group of species.
- Do simulations do validate theory.
- Push the theory further with more dependencies between marks with [Błaszczyszyn et al., 2019]
- Develop a method to identify which parameters influence the most population growth.

Thank you for your attention!

Baryshnikov, Y. and Yukich, J. E. (2005).
Gaussian limits for random measures in geometric probability.
The Annals of Applied Probability, 15(1A):213–253.
Publisher: Institute of Mathematical Statistics.

Błaszczyszyn, B., Yogeshwaran, D., and Yukich, J. E. (2019).
Limit Theory for Geometric Statistics of Point Processes Having Fast Decay of Correlations.
The Annals of Probability, 47(2):835–895.
Publisher: Institute of Mathematical Statistics.