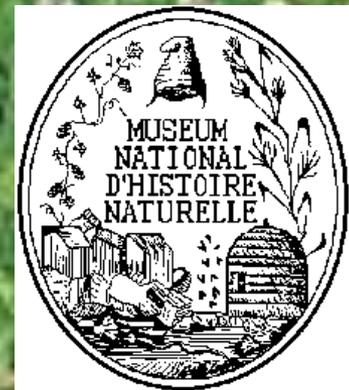


VIGIENATURE

Un réseau de citoyens qui fait avancer la science

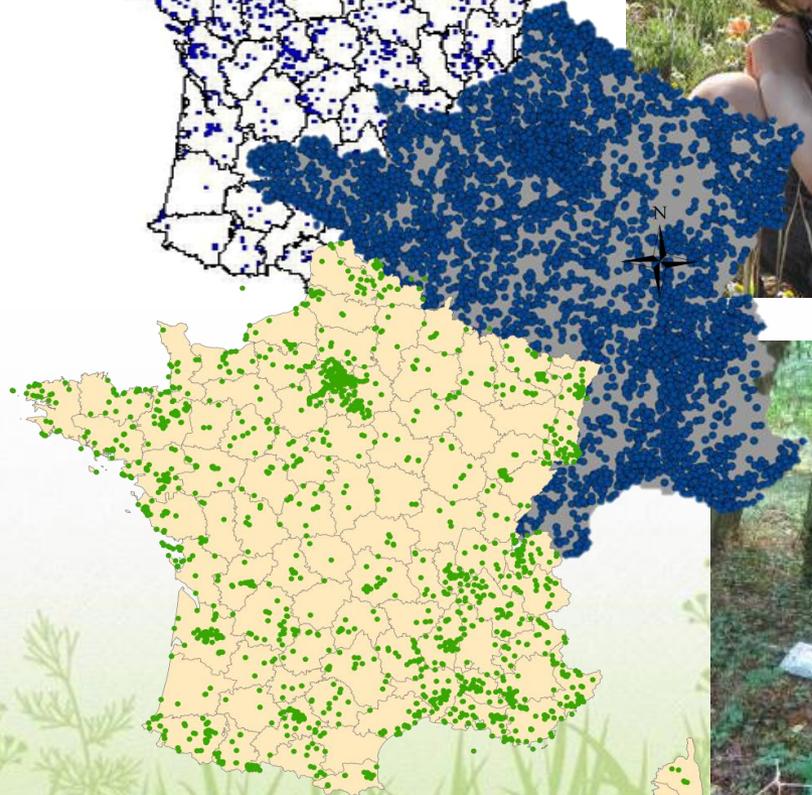
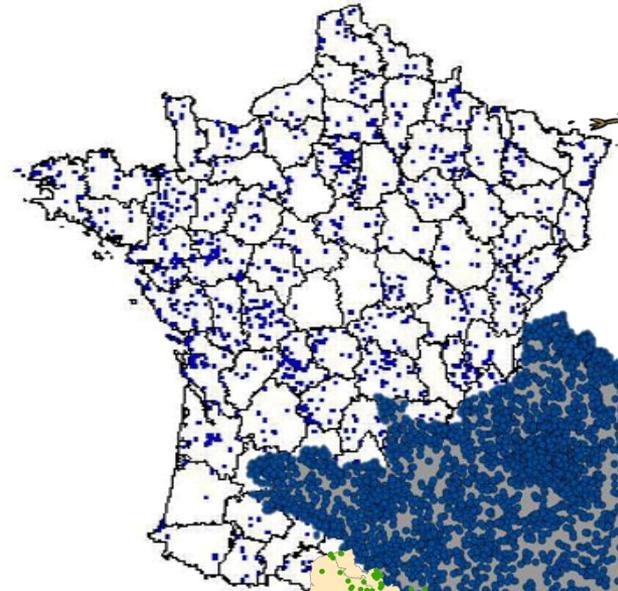
Romain Julliard



Objectifs de Vigie-Nature

Un projet scientifique du Muséum national d'Histoire naturelle

- **surveiller** l'évolution de la biodiversité,
 - » **Mais les espèces sont très diverses !**
- **diagnostiquer** les causes des variations
 - » **Mais causes multiples !**
 - » **Les échelles concernées ne permettent pas l'expérimentation**
- proposer différents **scénarios** d'évolution



0 50 100 200 km

Comment mettre en place un observatoire ?

(1) Réseau d'observateurs :

Compétences pointues
ou *grand nombre de participants ?*

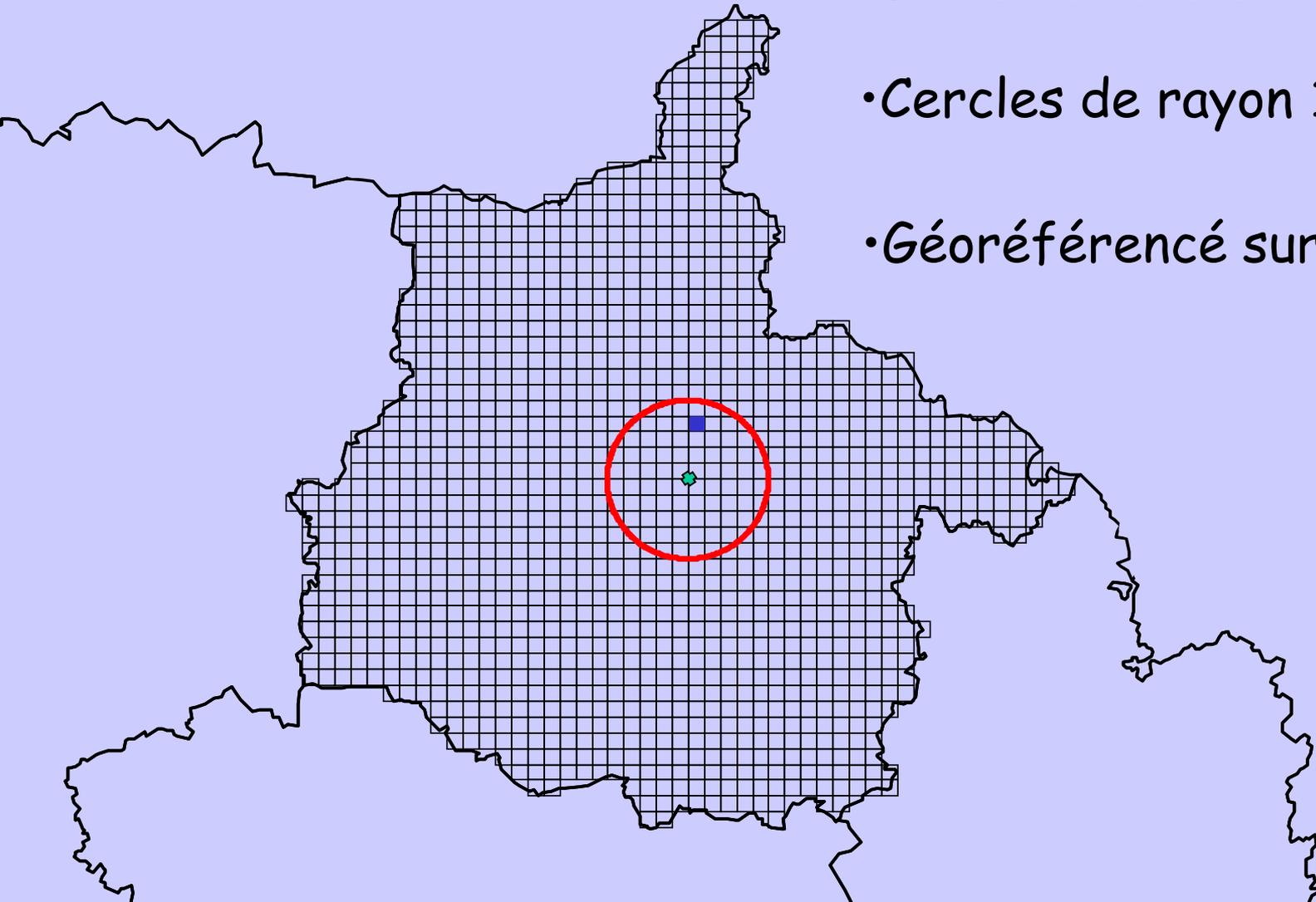
(2) Plan d'échantillonnage :

Meilleure représentativité **mais**
Envoyer les observateurs là où ils n'ont pas envie d'aller ?

=> (3) *Protocole simple*

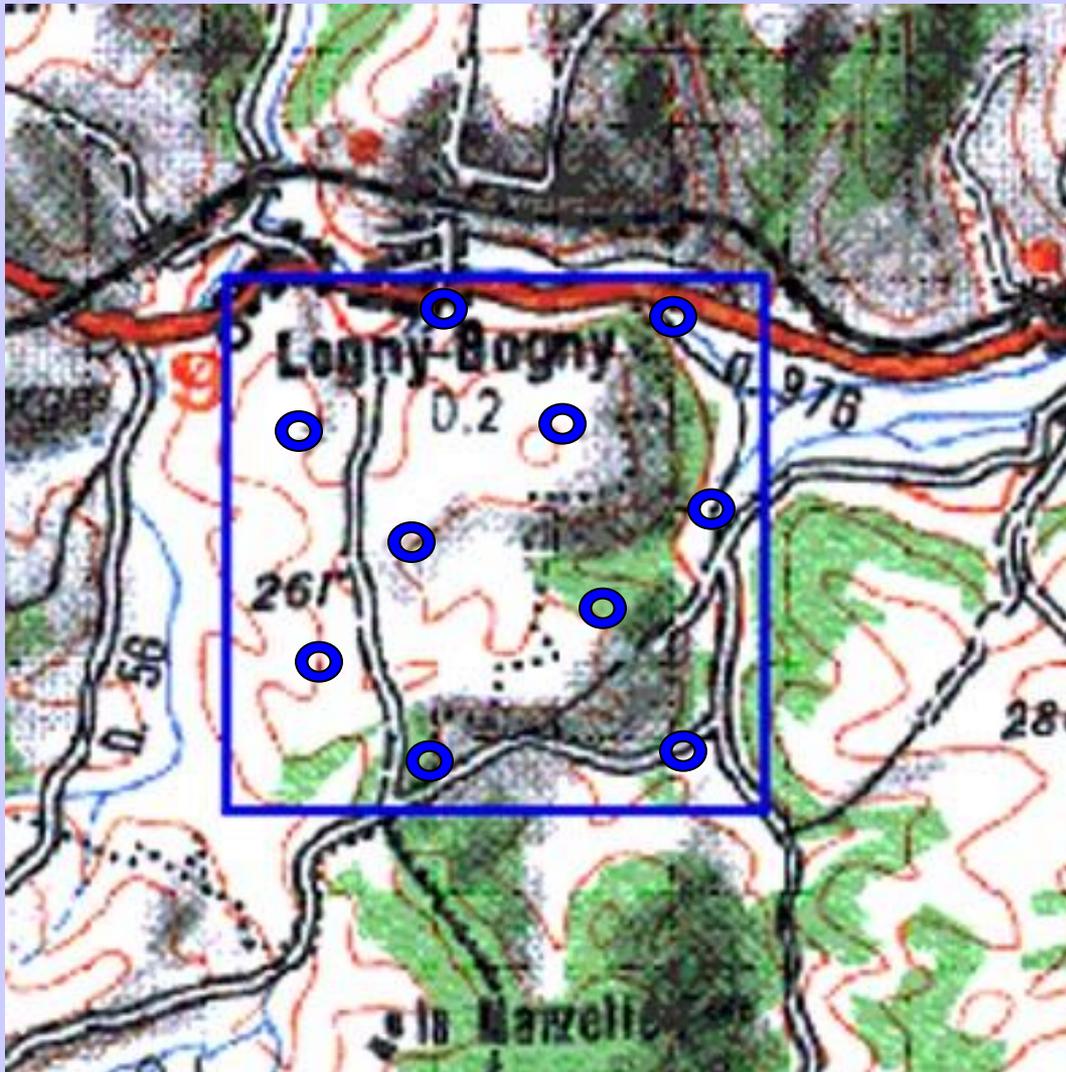
Comment procède-t-on au tirage au sort des sites à prospector ?

- Grille nationale 2 x 2 km
- Cercles de rayon 10 km
- Géoréférencé sur SIG



Distribution homogène des EPS

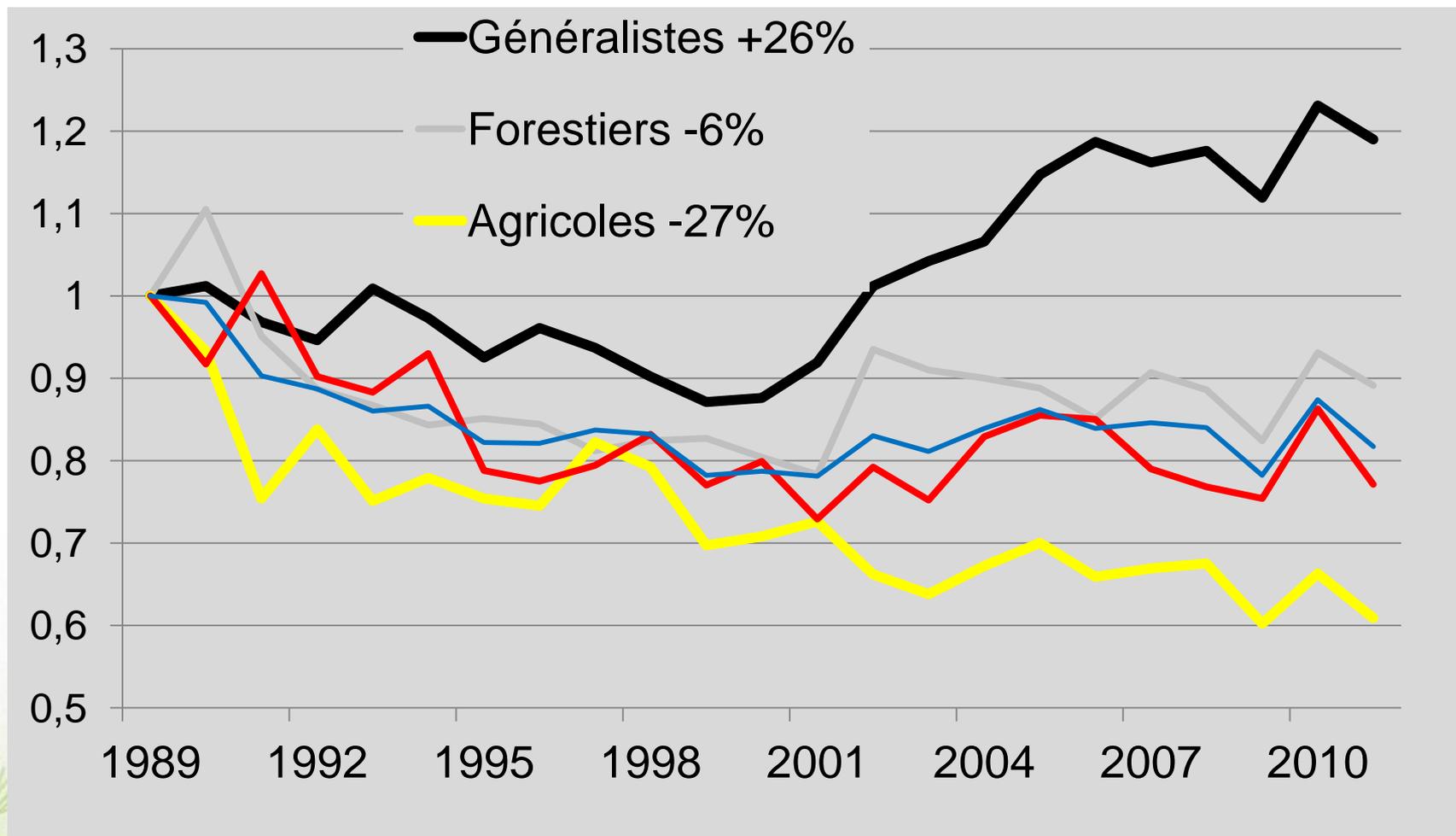
Proportionnelle aux habitats



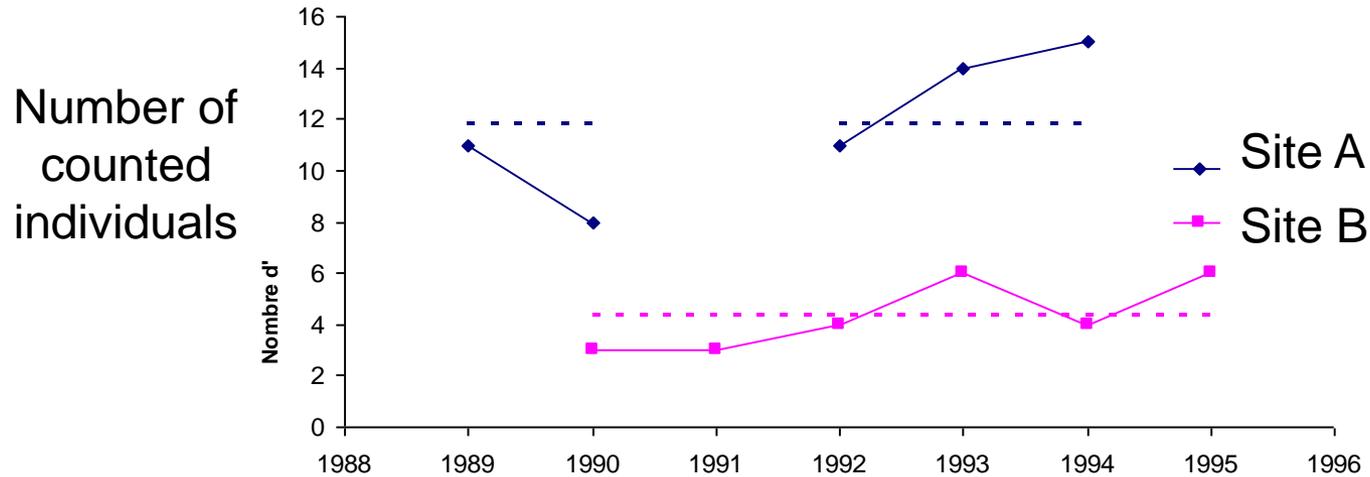
10 EPS dans un carré

Faits deux fois par an
avant et après le 8 mai

**Description
standardisée de
l'habitat**



A statistical parenthesis: estimation



A major hypothesis:

Site A : local variation specific to site A + global variation

Site B : local variation specific to site B + global variation

Etc...

Average over sites => global variation

Unexplored
scale-
dependent
variation



Suivi Photographique des Insectes POLLinisateurs

Une initiative labellisée





AFFICHER

2010-09-17

Nom de la Fleur: Taxon inconnu de la clé (aster)

Il s'agit d'une fleur : plantée

Il s'agit d'une habitat : jardin privé /

BIVILLE-SUR-MER (76098), SEINE-MARITIME (76),

HAUTE-NORMANDIE (23)

par : étamines

[TOUTES SES COLLECTIONS DANS LES GALERIES](#)



Dernière(s) identification(s):
Les Mouches à damier

AFFICHER



Dernière(s) identification(s):
La Syrutte pialante

AFFICHER



Dernière(s) identification(s):
Les Mégachiles rayés

AFFICHER



Dernière(s) identification(s):
Les Mouches aux reflets métalliques

AFFICHER

COMMENTAIRES DES INTERNAUTES

par : cybelle 2011-04-23

Superbe votre collection Etamines, les insectes se bouscuaient pour vous faire plaisir, vous avez dû vous régaler.

par : Fernand 2011-04-23

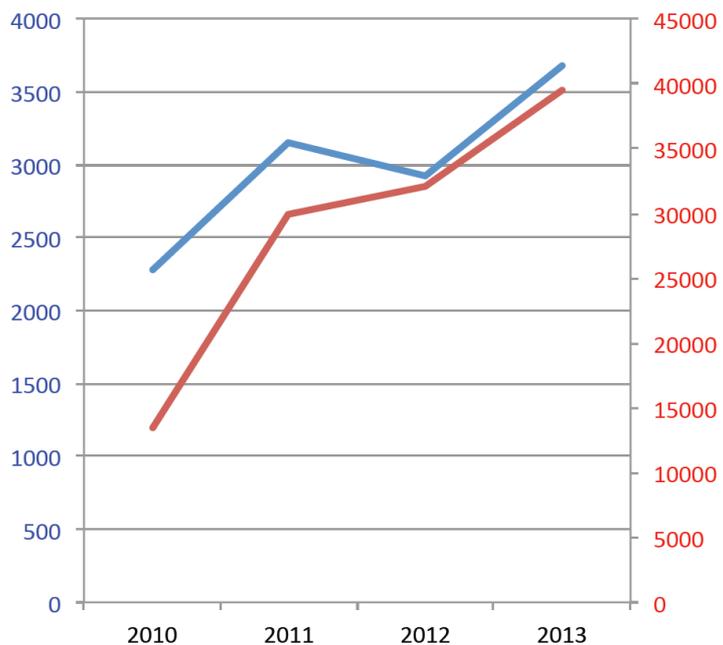
Très belles photos. Les fleurs sont magnifiques et les insectes également. Beau travail étamines .

Des données en quantité



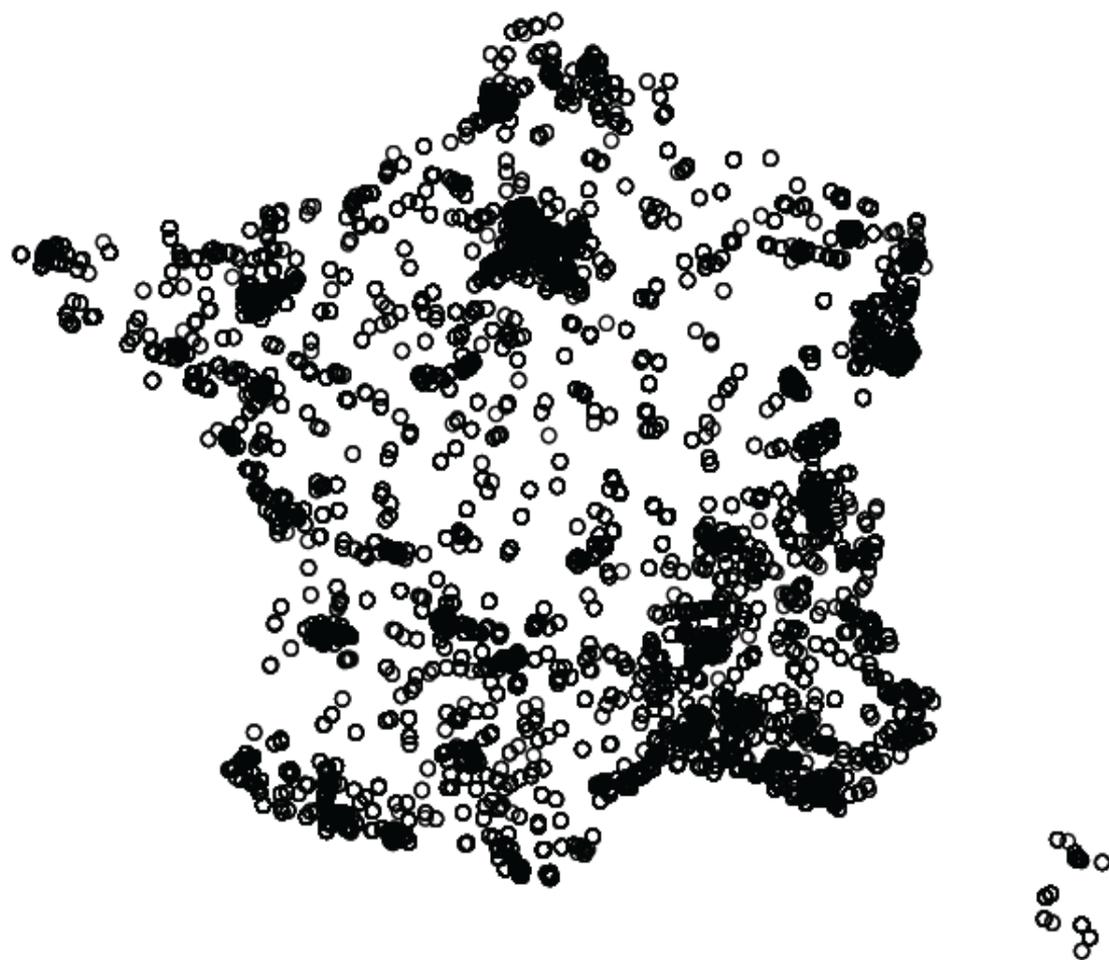
Depuis 2010

- > 1000 observateurs
- > 115000 observations
- > 12000 localités



— nb collections par an

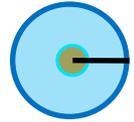
— nb photos par an



Habitat characterization at sampling sites

Three land-use types
(CorineLandCover):

- urbans
- agriculturals
- naturals

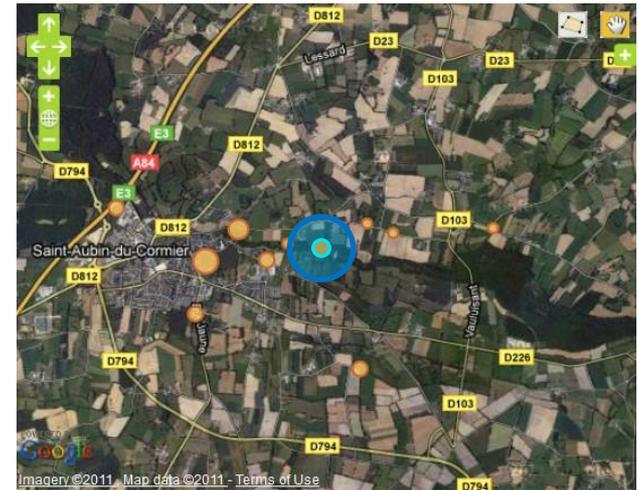
1 km
 **H_c** = proportion (land-use h areas)

... but heterogeneity in:

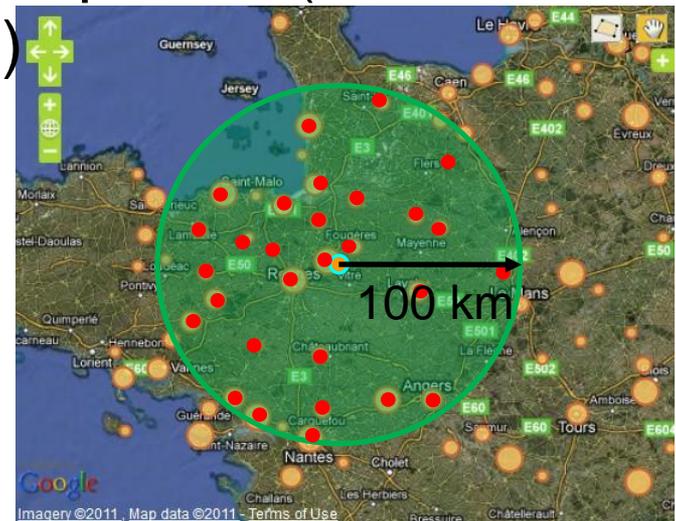
- land-use sampling
- land-use distribution

➔ **M_{c,h}** = **H_c** - **mean(H_r)**
= proportion of land-use locally relatively to regionally

collection c



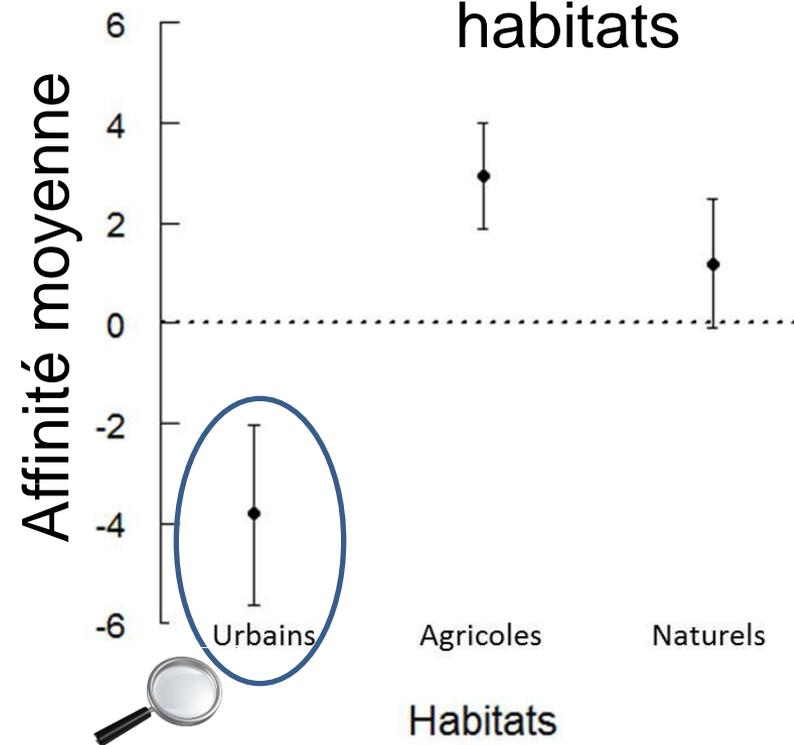
H_r = proportion (land-use h areas)



collection r

En moyenne les pollinisateurs communs fuient la ville, mais apprécient les paysages agricoles

Affinité des insectes pollinisateurs envers les habitats



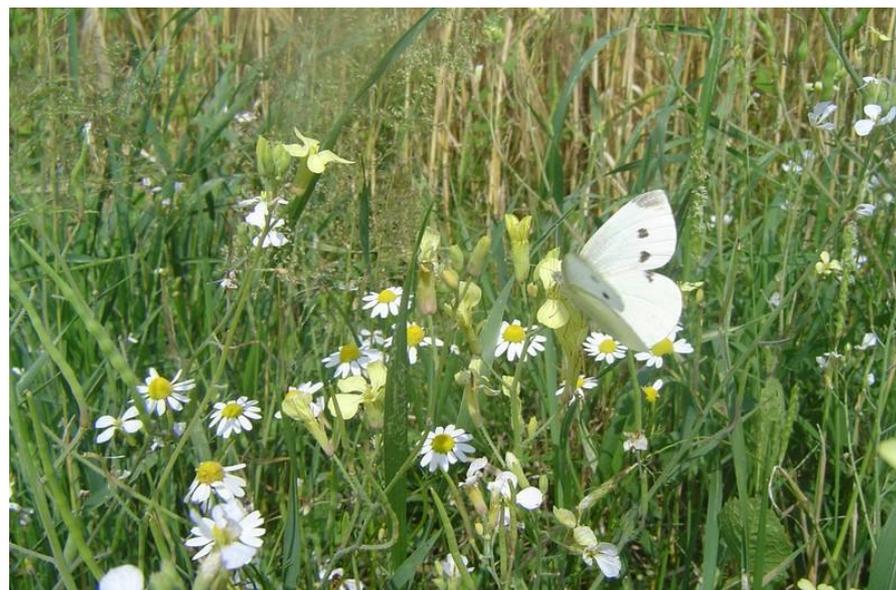
Ce qui fait la qualité d'une donnée naturaliste

⇒ **Précision** (compétence de l'observateur)

⇒ **Répétabilité** (respect du protocole)

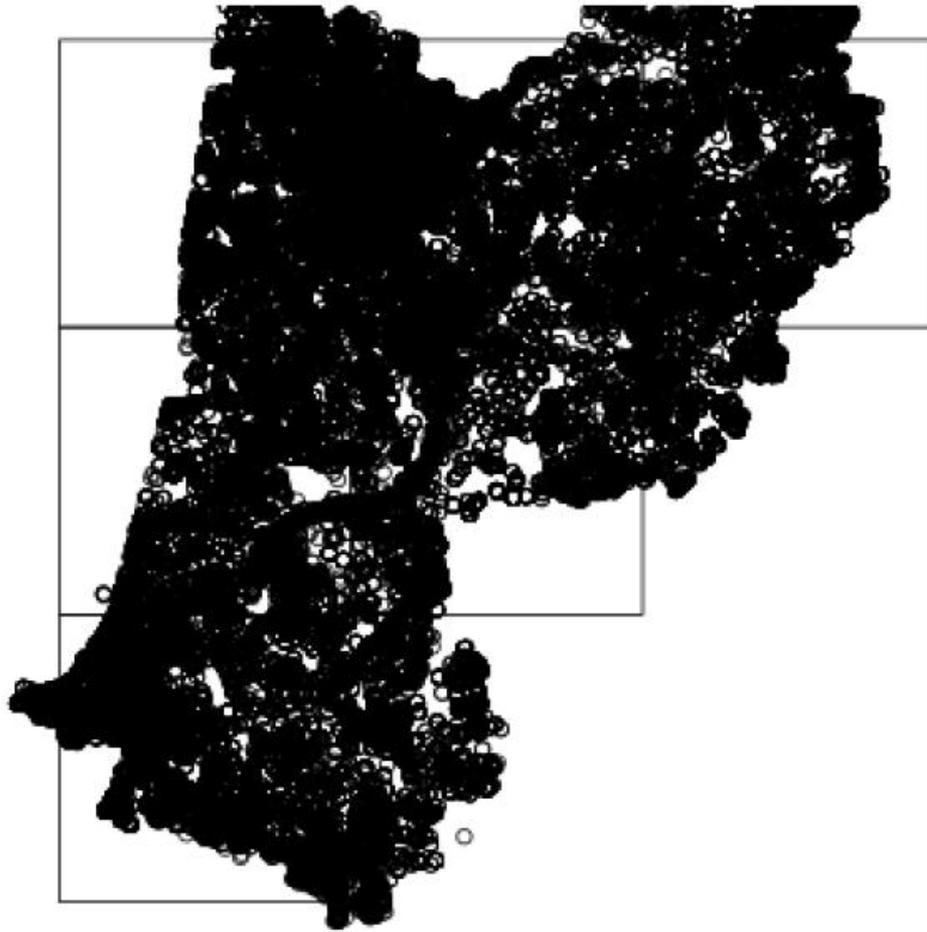
⇒ **Représentativité** (plan d'échantillonnage, post-stratification)

Ex: « j'ai vu 3 papillons blancs dans mon jardin »

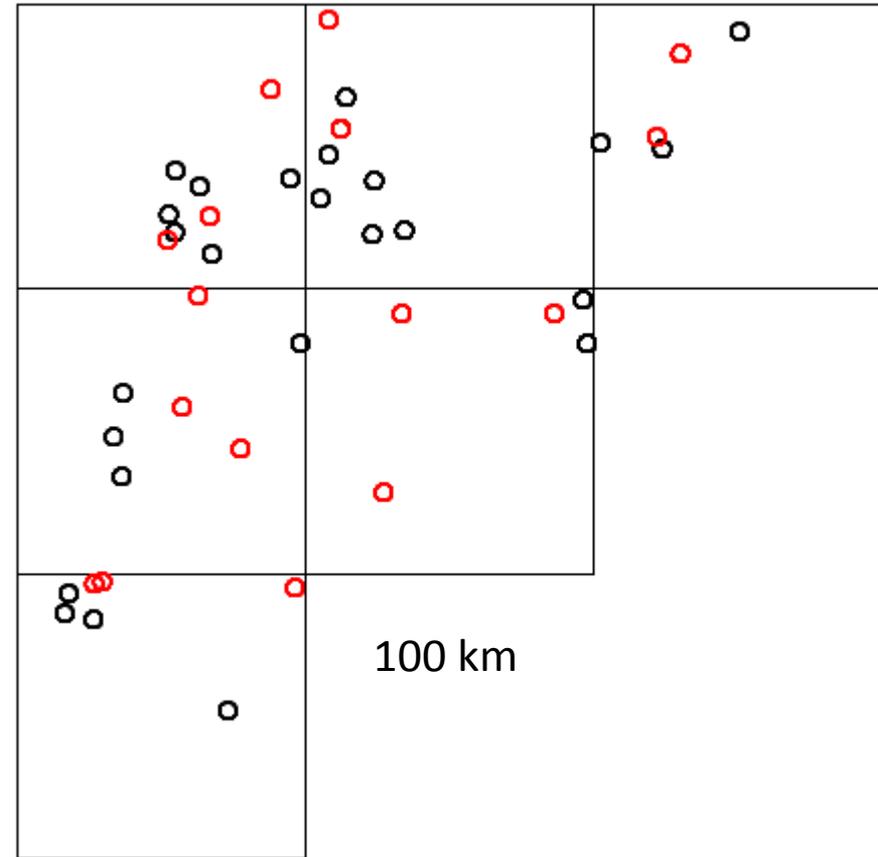


La « data-valanche » touche (enfin) la biodiversité !

- XX^{ème} : pour $n > 30$, loi des grands nombres 😊
- XXI^{ème} : Réseaux d'observateurs + Internet = base de données collaborative (des millions de données dites opportunistes ; crowdsourcing)
- + virage technologique : la saisie en ligne in situ



LPO



STOC1 + **STOC2**

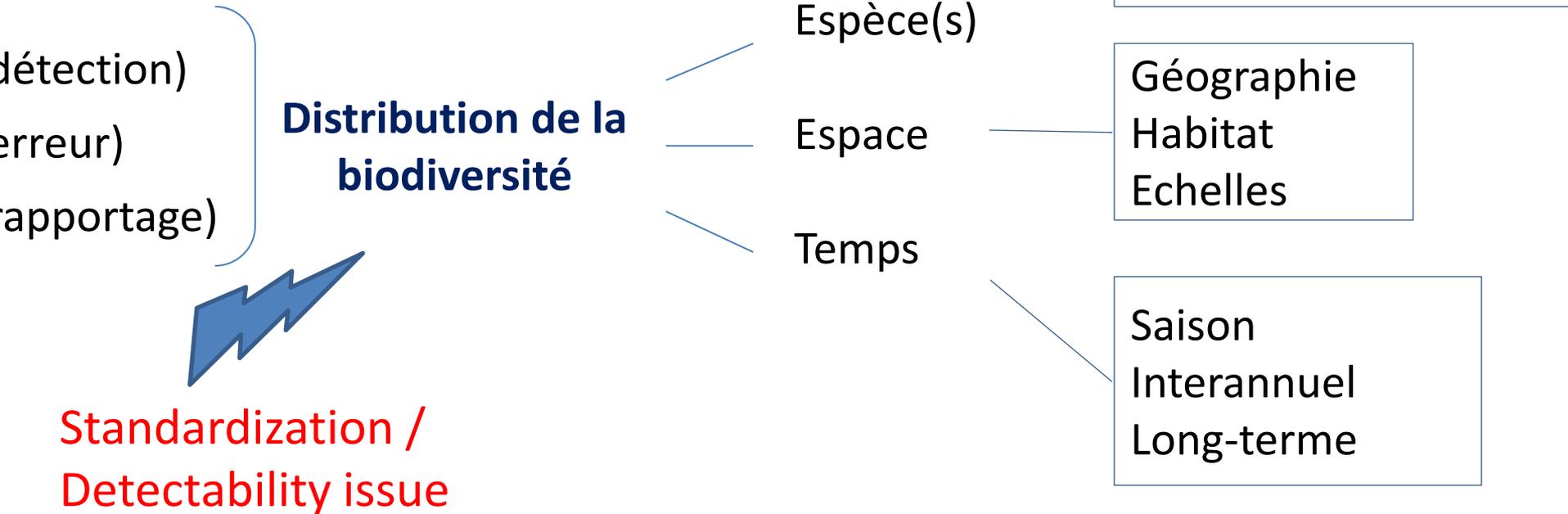
Pour les 22 espèces représentées dans tous les lots

These datasets have the potential to cover gaps:

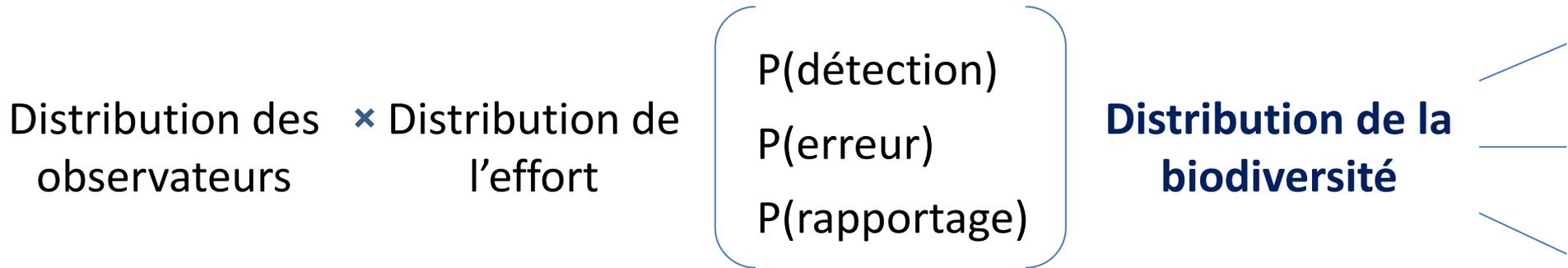
- Rare species, species difficult to monitor (eg reptiles)
- Space (eg Tropical region)
- Finer spatial grain

=> Nouvelles méthodes pour l'analyse

Analyse des données naturalistes : approche classique



Une observation dans une base de données multi-observateurs c'est le résultat :



Echantillonnage préférentiel

Non-indépendance des données

Distribution de l'effort d'observation

Le problème multi-espèces multi-sites multi-protocoles :

$$X_{ijk} \sim \text{Poisson}(A_{ij} \cdot E_{jk} \cdot P_{ik})$$

Probabilité de détection

Effort d'observation

Abondance réelle

=> IJK observations ; IJ + JK + IK paramètres à estimer 😊

Mais ! Pas identifiable...

We show that if effort is known for scheme $k=0$ (standardized data set), then then E_{jk} pour $k \neq 0$ can be estimated, as well as P_{ik} / P_{i0}

So that the relative abundance \tilde{A}_{ij} can be estimated

But \tilde{A}_{ij} can be estimated with the standardized data set alone! What is the gain?

- Indeed we show that the precision of A_{ij} estimates is always improved compared to the use of the standardized data set alone
- The gain in precision may be substantial if opportunistic data are abundant

$$\text{var}(\widehat{A}_{ij}) = \text{var}(\widehat{A}_{ij}^0) \times \frac{P_{i0}A_{ij}}{\sum_l P_{l0}A_{lj}} \times \left(1 + \frac{\gamma_{ij}}{E_{j1}}\right) + O\left(E_{j1}^{-3/2}\right)$$

with

$$\gamma_{ij} = \frac{\sum_{l:l \neq i} P_{l0}^2 A_{lj} + \left(\sum_{l:l \neq i} P_{l0} A_{lj}\right)^2 / A_{ij}}{\sum_l P_{l0} A_{lj}} \times \frac{1 + E_{j0} \sum_l P_{l0} A_{lj}}{\sum_l P_{l0} A_{lj}}.$$

In particular, we have the asymptotic variance reduction factor

$$\frac{\text{var}(\widehat{A}_{ij})}{\text{var}(\widehat{A}_{ij}^0)} \xrightarrow{E_{j1} \gg 1} \frac{P_{i0}A_{ij}}{\sum_l P_{l0}A_{lj}}.$$

What are the key hypotheses?

$$X_{ijk} \sim \text{Poisson}(A_{ij} \cdot E_{jk} \cdot P_{ik})$$

HYP Effort only depend on site and scheme
 Detectability only depend on species and scheme

=> no special effort toward some species in some sites!!

=> Homogénéité du comportement des observateurs

What are the key hypotheses?

$$X_{ijk} \sim \text{Poisson}(A_{ij} \cdot E_{jk} \cdot P_{ik})$$

HYP A_{ij} are the same for standardized and opportunistic schemes

BBS: random sampling design, so that estimated A_{ij} coincides with true A_{ij}

Opportunistic: observers attention likely biased toward some habitat (eg avoid intensive farmland)

Effect will partly be taken into account by P_{ik} (farmland species will be underdetected in opportunistic data)

But this only approximate if habitat heterogeneity does not vary too much among site j and attention is not too biased toward some habitats

Perspectives

Introduire une stratification: l'habitat

$$A_{ij} \frac{S_{ih} V_{hj}}{\sum_{h'} S_{ih'} V_{h'j}} \times E_{j2} \frac{q_{h2}}{\sum_{h' \in l} q_{h'2}} \times P_{ihk}$$

Répartition des espèces en fonction de leur préférence S_{ih} et de l'offre en habitat V_{hj}

Biais d'attention q_h des observateurs vers habitat h parmi ceux présent dans lieu-dit l

Merci de votre attention

www.vigienature.fr

