

Explorer l'espace avec les données naturalistes : « Sortez des sentiers battus ! »



Marc Dufrêne
Natagora - Namur - 03.12.2023



Journal of Biogeography (1989) 16, 301–303

A
biogeographical
short story

A dot on a map

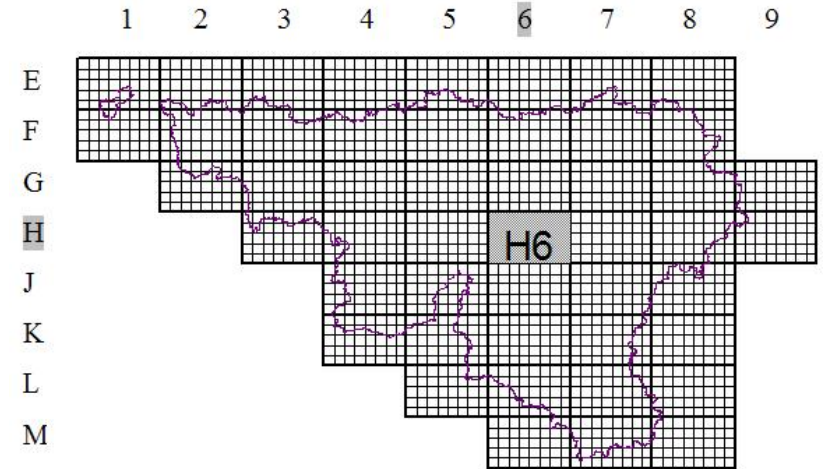
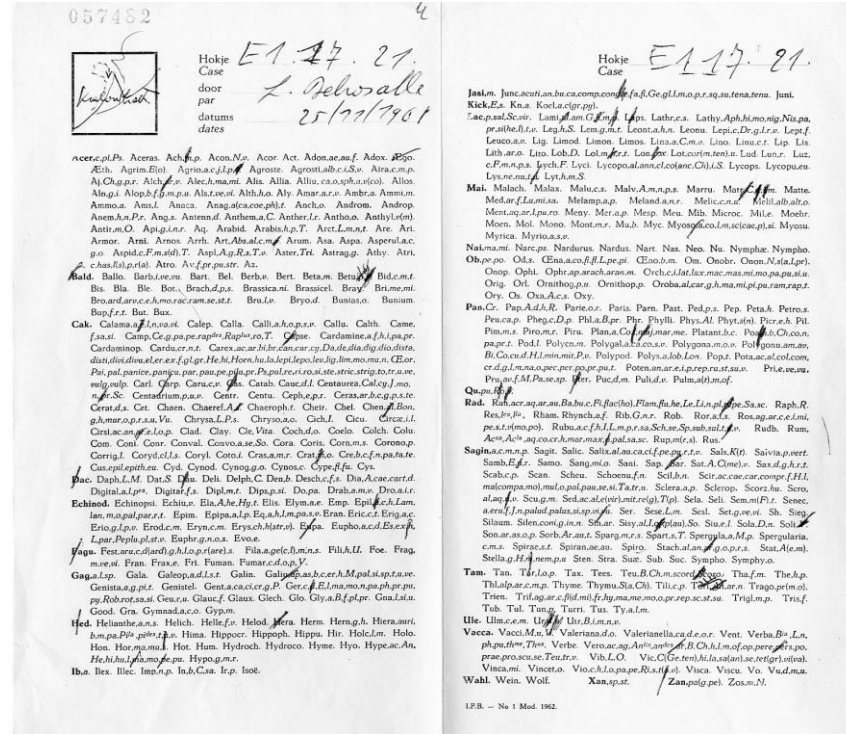
Dedicated to Célestin-Joseph Dovern of Hesdin and to all local naturalists

(1)

M. Goupil was dying. He was now confined to his bed in a musty room above the pharmacy on the Rue Charles Quint. Dried herbs hung dustily from the beams, roots of orris were heaped in corners, bottled fungi lined the shelves, and manuscript papers littered the

L'histoire des données naturalistes

- Atlas de la flore belge et luxembourgeoise de van Rompaey & Delvosalle (1972, 1979)

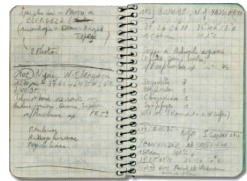


Relevé de terrain avec des fiches à cocher
 Cartographie à la main sur les cartes pour réaliser un atlas

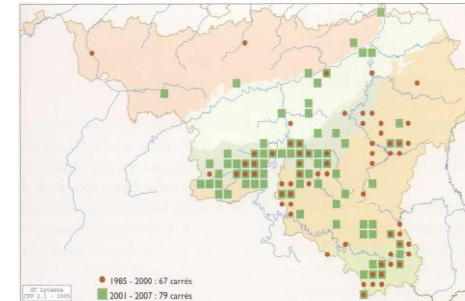
=> 1.640 taxons
 => 680.000 données

L'histoire des données naturalistes

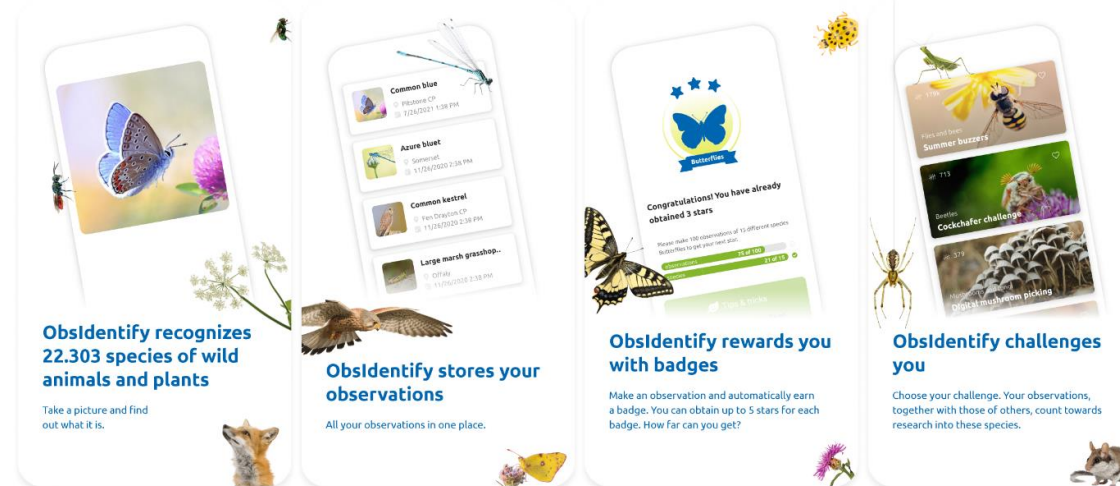
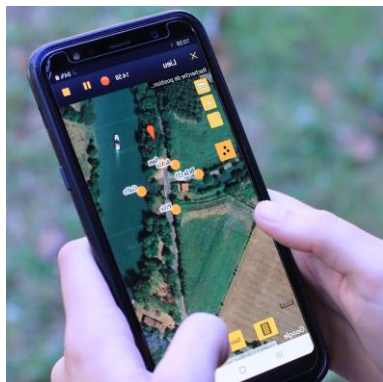
- Depuis les années 90 => informatisation



88414	Projet :	003 Encodage Yvan B	Code :	Toponyme:
	Coord. originales :	Lambert belge 1972	THS1	Terril d'He
	X	100752	Y	127345
	Précision:	10	Datum:	BEL72
	<input type="checkbox"/> Coordonnée conventionnelle	Coordonnées calculées		
	Grille utilisateur			
	Deg. déc./Greenwich	UTM alphanumérique		
	50.45573	67652	51UER46	
	DMS/Greenwich			
	50°27'20.7"N	003°40'35.5"E		
	Utm: 100752 127345			
	S.G.I.B.:	Cadastr: Hensies		



- Depuis les années 2000 => informatisation - support à l'identification



ObsIdentify recognizes 22.303 species of wild animals and plants
Take a picture and find out what it is.

ObsIdentify stores your observations
All your observations in one place.



ObsIdentify rewards you with badges
Make an observation and automatically earn a badge. You can obtain up to 5 stars for each badge. How far can you get?

ObsIdentify challenges you
Choose your challenge. Your observations, together with those of others, count towards research into these species.



Quels sont les usages des bases de données ?

Les hotspots de biodiversité floristique à protéger

 **La biodiversité en Wallonie** Plan du site - Contact - Extranet Recherche 

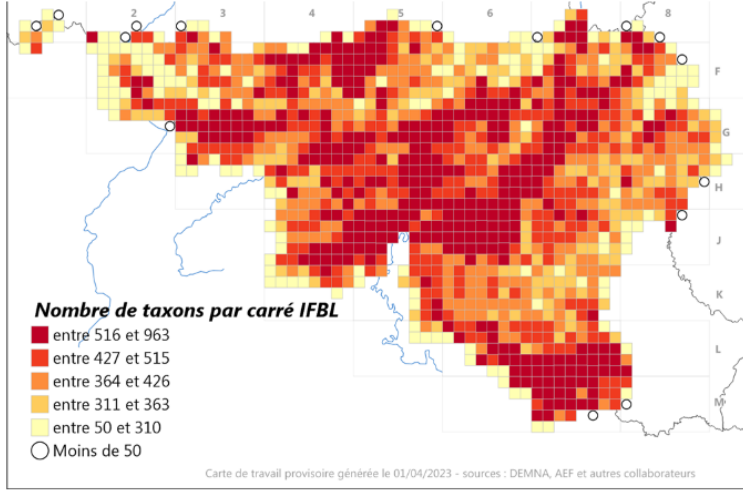
[Accueil](#) [Espèces](#) [Biotopes](#) [Sites](#) [Invasives](#) [Législation](#) [Agir](#) [Natura 2000](#) [Actualités](#) [A propos](#)

Vous êtes sur : [Accueil](#) > [Espèces](#) > [Flore](#) > [Atlas permanent](#) > [Atlas en ligne](#)

Atlas en ligne

Derniers chiffres

Nombre total de données	5 042 176
Sources de données	31
Nombre de taxons	2 737
Donnée la plus récente	03/12/2022
Nombre de taxons répertoriés dans les carrés 4x4	2 729
Nombre de taxons répertoriés dans les carrés 1x1	2 716
Nombre de carrés 4x4 avec au moins une donnée	1 199
Nombre de carrés 1x1 avec au moins une donnée	16 891



Nombre de taxons par carré IFBL

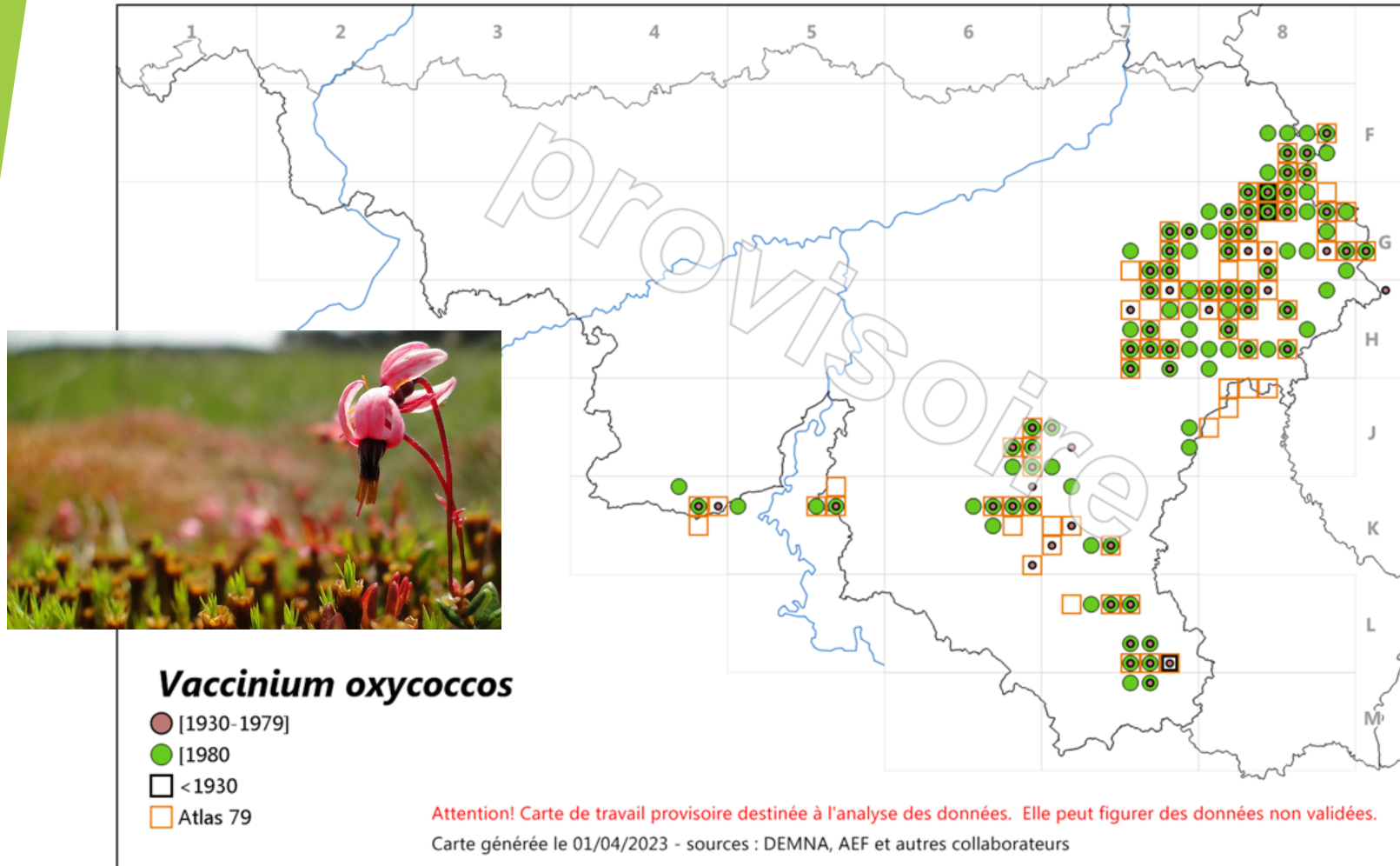
- entre 516 et 963
- entre 427 et 515
- entre 364 et 426
- entre 311 et 363
- entre 50 et 310
- Moins de 50

Carte de travail provisoire générée le 01/04/2023 - sources : DEMNA, AEF et autres collaborateurs

[Carrés prioritaires & version interactive](#)

Quels sont les usages des bases de données ?

La réalisation de cartes de répartition dynamique



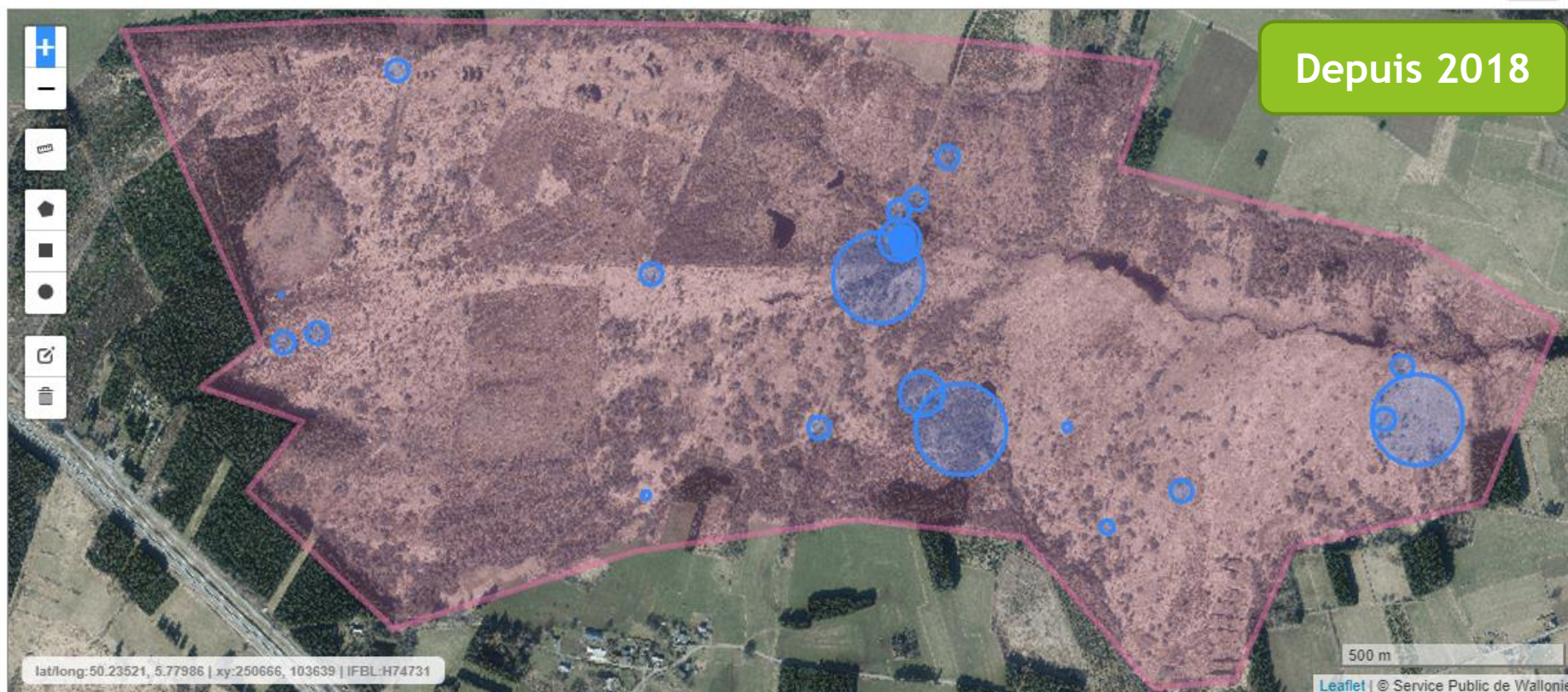
Pour identifier :

- les zones où l'espèce est présente ou a été présente,
- la fragmentation de l'aire,
- les affinités géographiques avec des facteurs écologiques,
- ...

Quels sont les usages des bases de données ?

L'identification des enjeux locaux

Espèces protégées ou patrimoniales sur le territoire wallon (prototype)



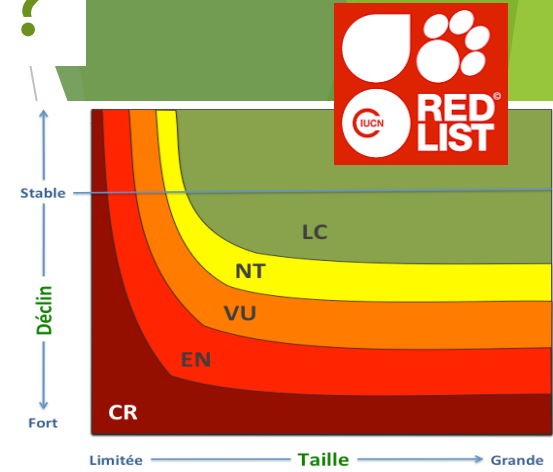
=> Pour la protection, la gestion, la restauration

Quels sont les usages des bases de données ?

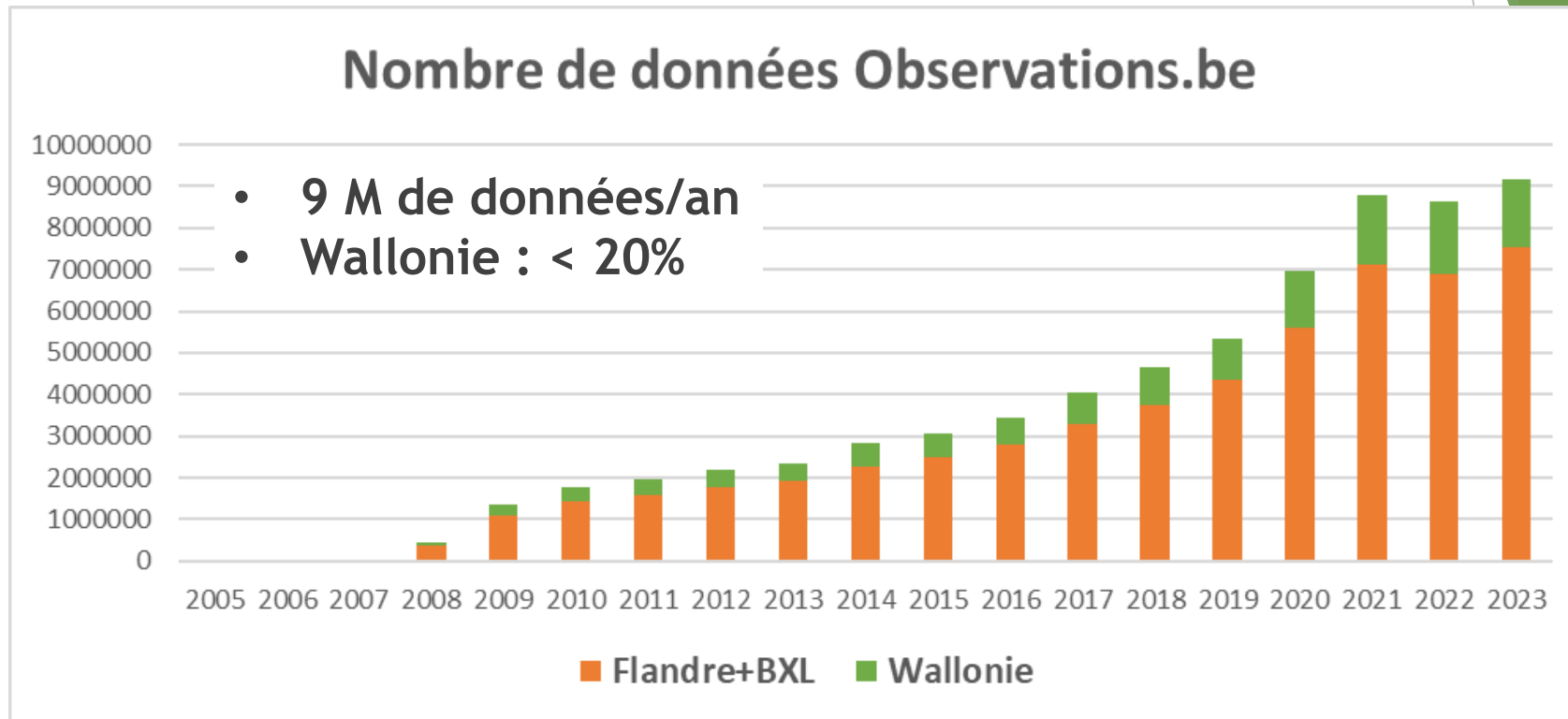
L'analyse de l'évolution de la répartition

- Identification de la **vulnérabilité des taxons** qui peut conduire à les protéger
- **Protection légale** des taxons menacés avec imposition de contraintes aux tiers, notamment en cas de développement de projets (EIE)
- Critères essentiel de la **zonation du réseau écologique**
- Identification de **priorités d'actions** pour la protection des sites et des mesures de restauration

Mieux vaut dès lors ne pas se tromper !



Explosion des bases de données

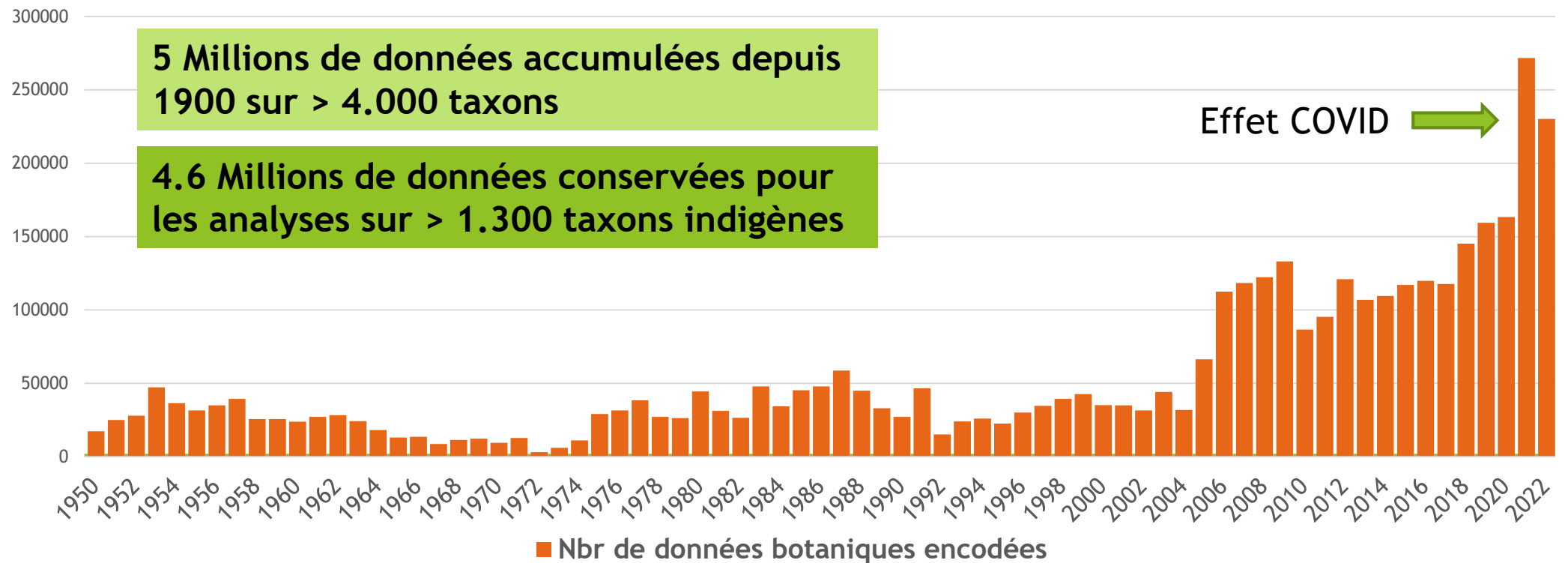


=> Développement des supports d'encodage virtuel (multiplication)

=> Développement des outils d'aide à la détermination (démocratisation)

Explosion des bases de données

Nbre d'observations botaniques en Wallonie



=> Des systèmes d'informations qui sont censés devenir de plus en plus performants pour des diagnostics de qualité

=> Logique du gestionnaire de BDs : plus il y en a, mieux c'est ...

Plus de données = plus d'informations ?

1. Exactitude de l'information : est-ce le bon taxon ?

- Problème des **variations des sources taxonomiques**
=> maintenir un référentiel de synonymies le plus complet
=> pour la flore, problème des variétés horticoles à gérer

- Problème de la **validation de la détermination**
 - Photographie
 - Cohérence géographique et temporelle
 - Expertise de l'observateur



Orvet fragile
Anguis fragilis (51%)

■ relativement commun

📍 Beausaint

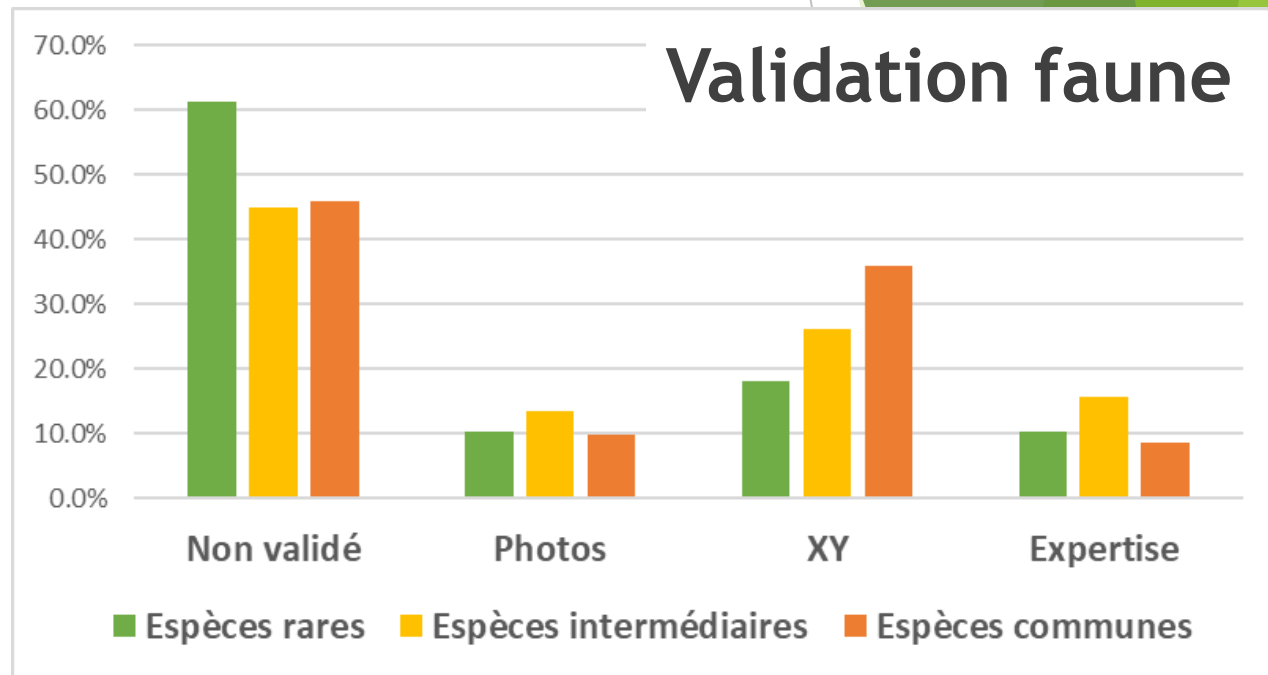
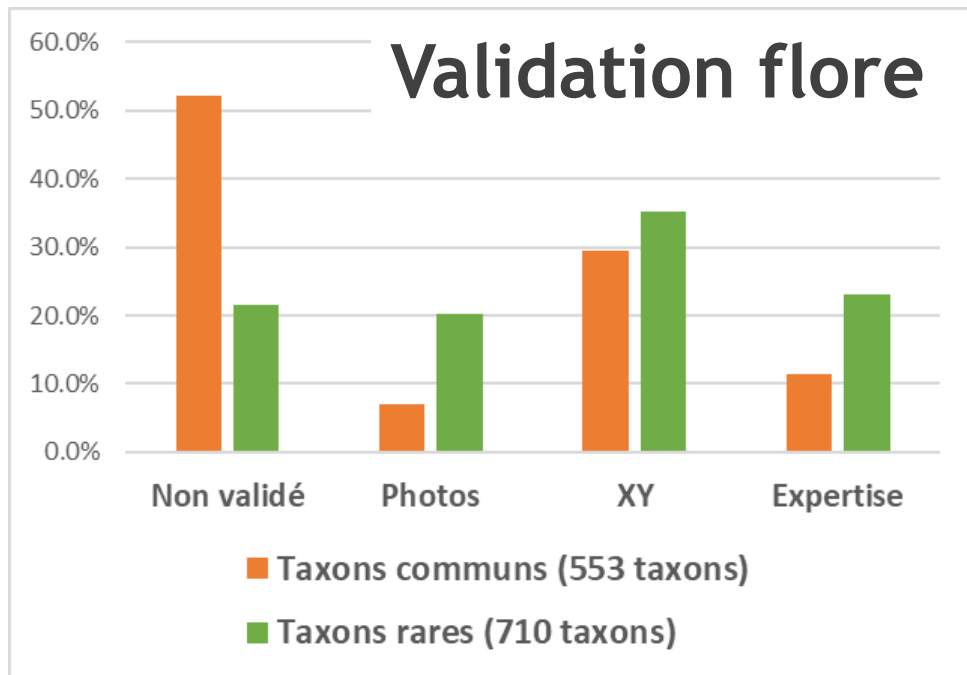
📅 23/4/2023 11:38



=> Prudence lors des encodages en renseignant l'incertitude de la détermination

Plus de données = plus d'informations ?

1. Exactitude de l'information : est-ce le bon taxon ?



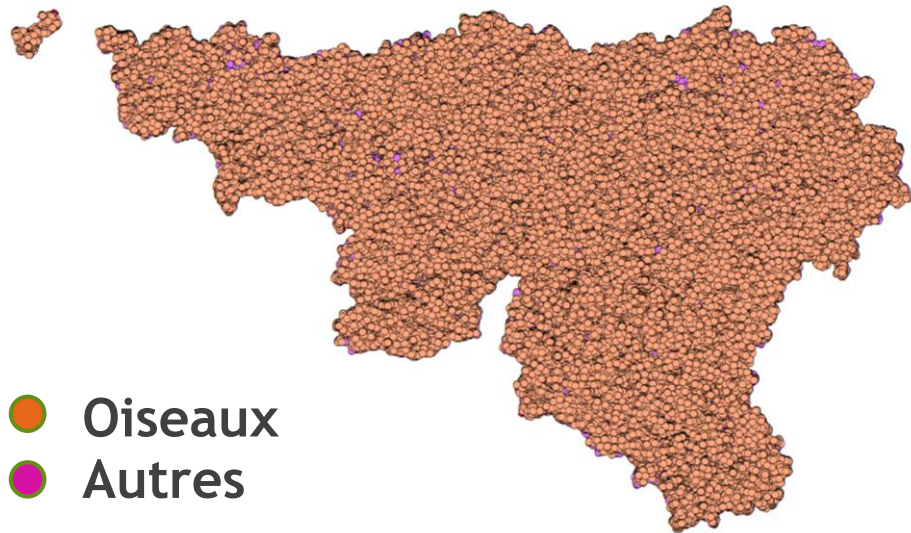
Encore 20-60% d'observations d'espèces non-validées, y compris des espèces rares avec des enjeux biologiques (200.000 données inutilisables)

=> Saturation du processus de validation qui est pourtant essentiel (confiance / enjeux de conservation)

Plus de données = plus d'informations ?

2. Pertinence de l'information : est-ce une donnée utile ?

Saturation des données « Faune » par les données « Oiseaux »



Comportement	Période de reproduction		
	Autres	Avril-Mai-Juin	Total
1-Nicheur possible	1.8%	5.1%	6.9%
2-Nicheur probable	0.6%	1.4%	1.9%
3-Nicheur certain	0.5%	0.6%	1.1%
Autres codes	12.7%	5.1%	17.8%
Non défini	45.7%	26.6%	72.3%
Total	61.3%	38.7%	100.0%

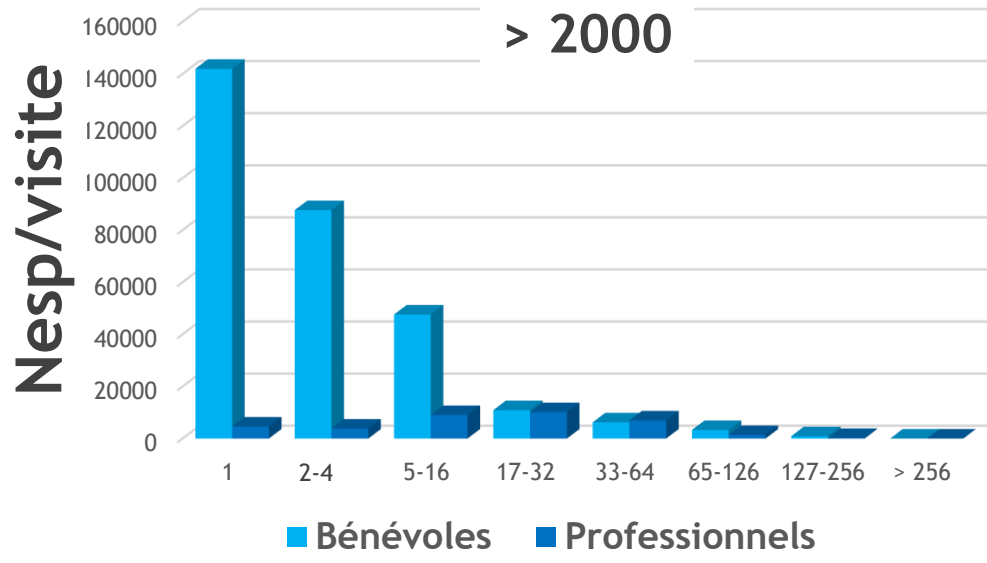
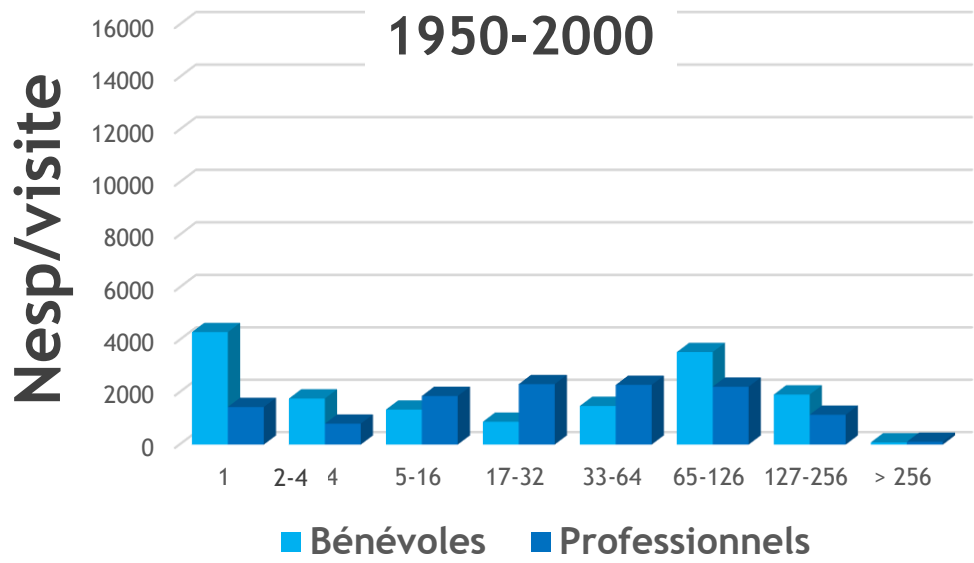
A peine 7% (voire même **2%** !) des données « Oiseaux » avec des codes de comportement permettant de « supposer » une reproduction

=> Il est essentiel pour un groupe si important de bien identifier les comportements pour être pris en compte

Plus de données = plus d'informations ?

3. Changement de comportement des observateurs

Nombre d'espèces encodées par visite

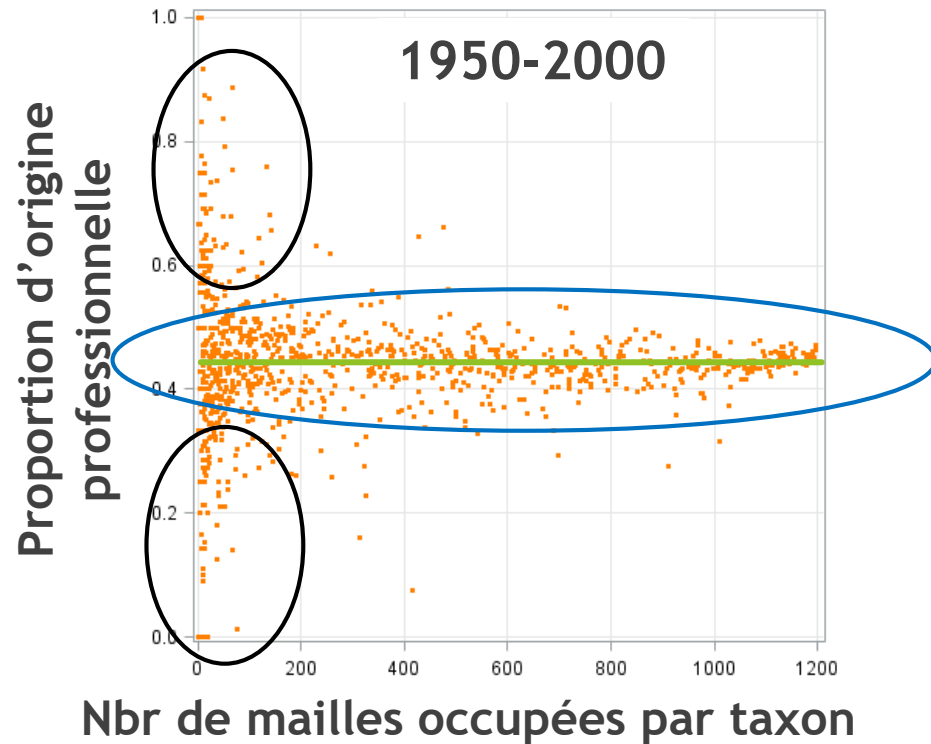


Les données ponctuelles opportunistes dominant dans les données récentes des naturalistes bénévoles => listes très partielles

Ciblent-elles certains groupes d'espèces comme les espèces rares ?

Plus de données = plus d'informations ?

3. Changement de comportement des observateurs

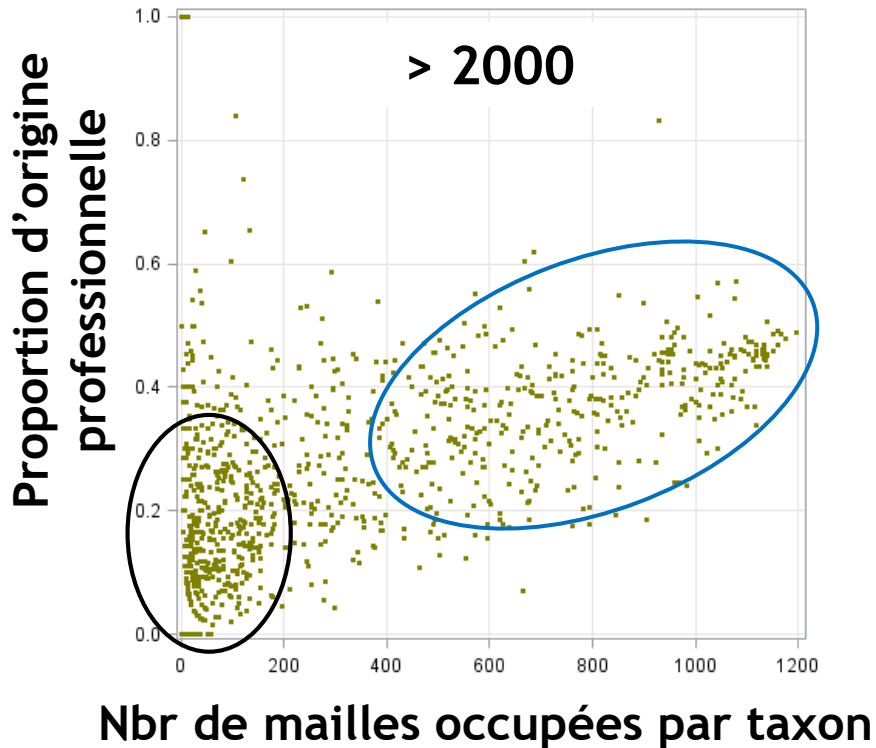


Avant 2000, les observateurs « professionnels » représentaient 44% des données pour 35% des observateurs

Quelques espèces plus rares sont plus souvent observées par un groupe que par un autre mais cela concerne les deux groupes

Plus de données = plus d'informations ?

3. Changement de comportement des observateurs



Après 2000, si les professionnels sont proportionnellement 30 x moins nombreux, ils assurent encore plus de 30% des observations.

Mais de nombreuses rares sont nettement plus souvent observées par les amateurs que les professionnels

=> Soutenir la réalisation de listes d'espèces qui sont bien plus intéressantes que des observations ponctuelles opportunistes

=> Développer des projets d'inventaires locaux

Plus de données = plus d'informations ?

3. Changement de comportement des observateurs

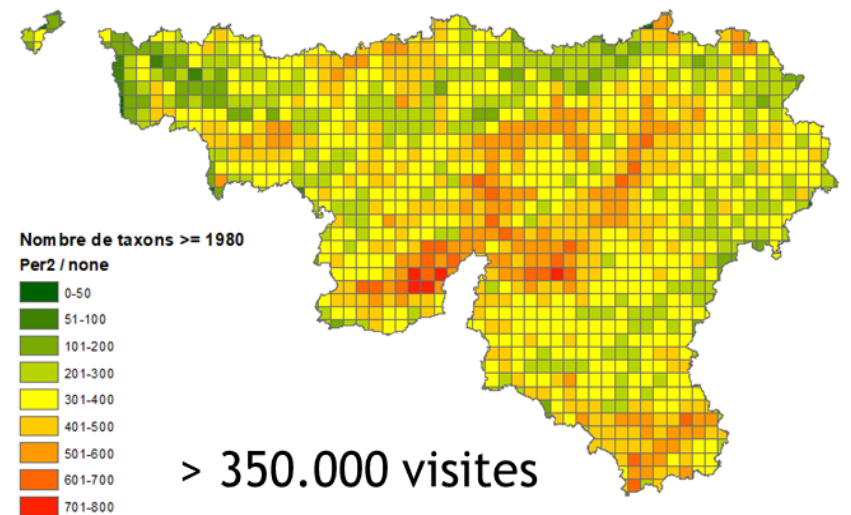
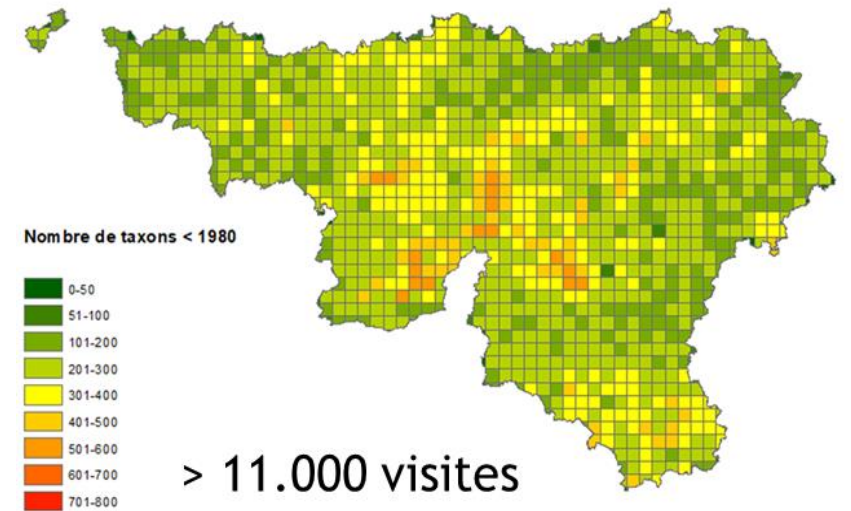
Flore : 4.6 millions de données

- Avant 1980 : 19% des données
- 1980-2000 : 16% des données
- > 2000 : 65% des données

Echantillonnage très différent :

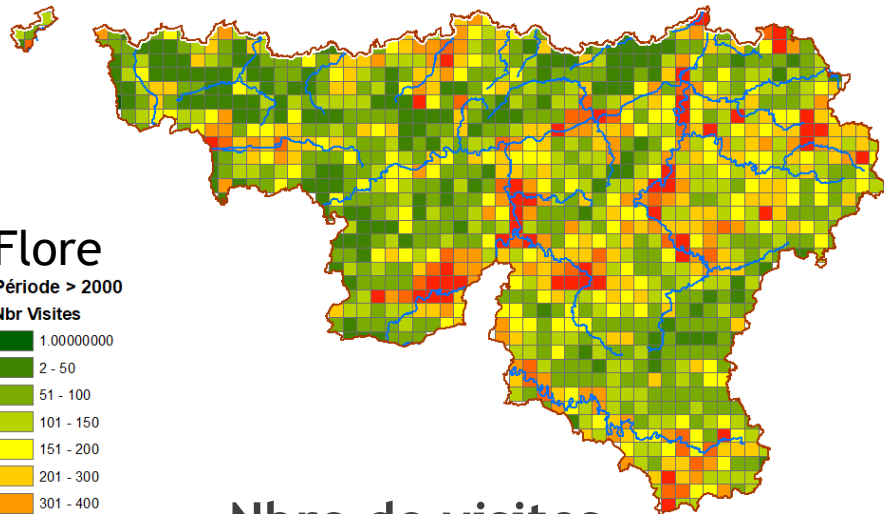
- Avant 1980 : 11.400 visites
- 1980-2000 : 16.000 visites
- > 2000 : 290.000 visites

Plus de **30 fois** plus de visites !

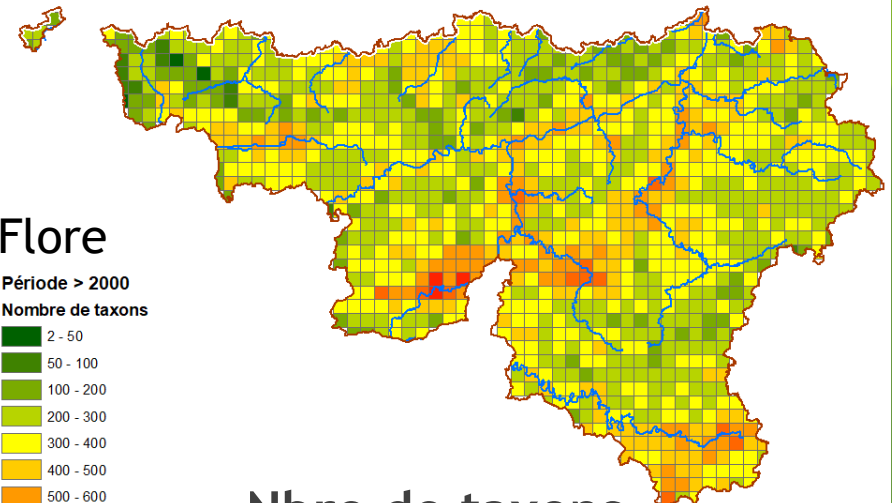


Plus de données = plus d'informations ?

3. Changement de comportement des observateurs

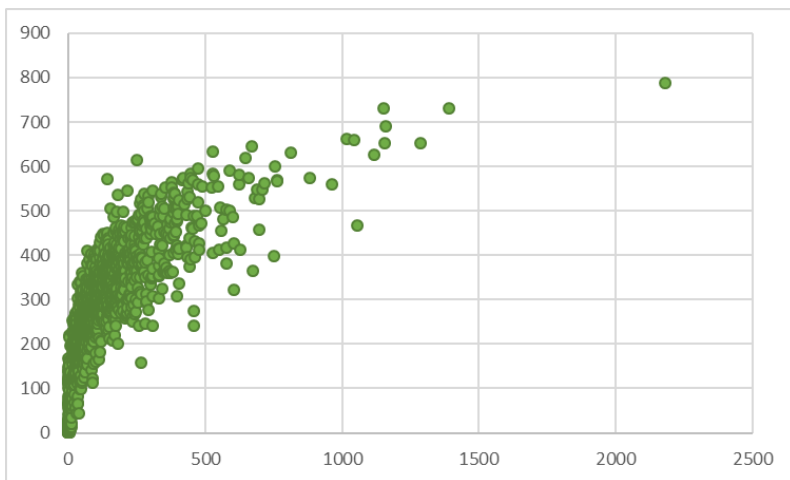


Nbre de visites



Nbre de taxons

Nbre de taxons



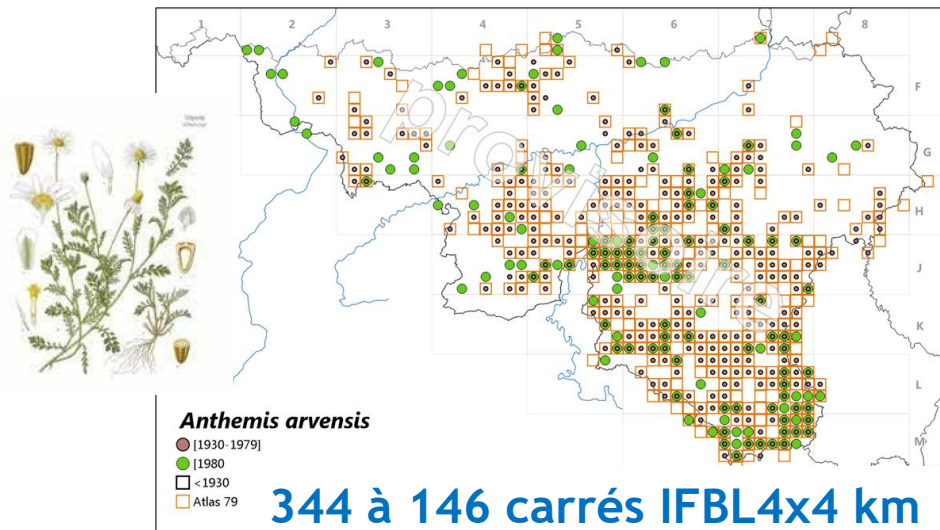
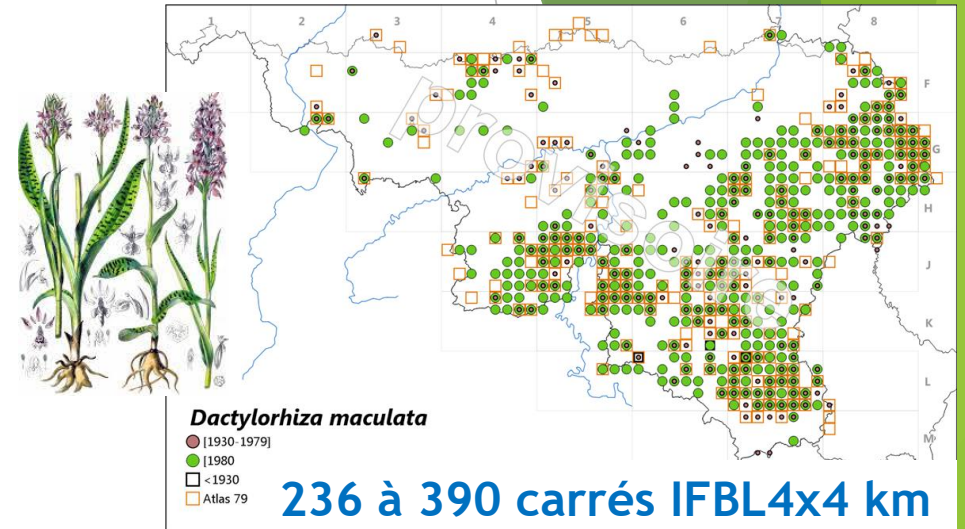
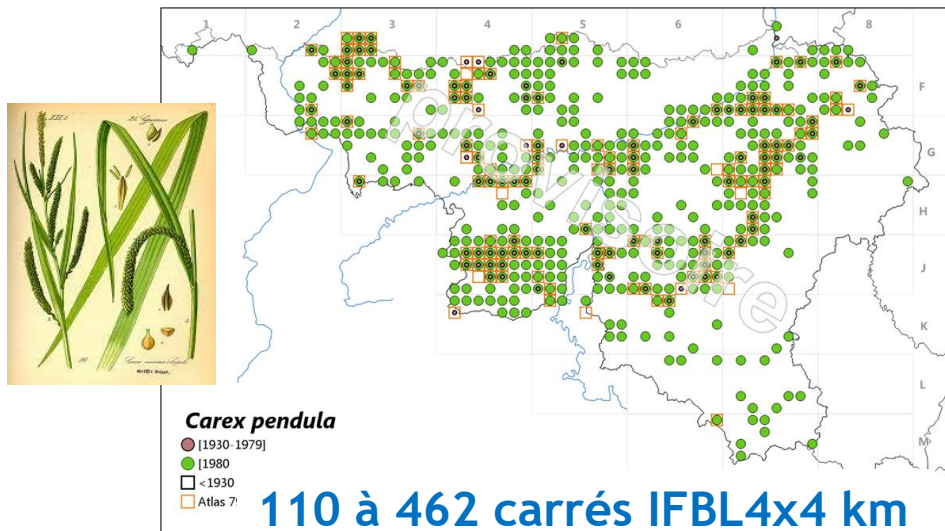
Nbre de visites

Les régions les plus riches
sont les plus visitées

Ce qui n'exclut pas qu'elles
soient les plus riches ...

Plus de données = plus d'informations ?

3. Changement de comportement des observateurs



Comment évaluer les tendances des répartition des espèces avec une telle variation de l'échantillonnage ?



La détectabilité a fondamentalement changé !

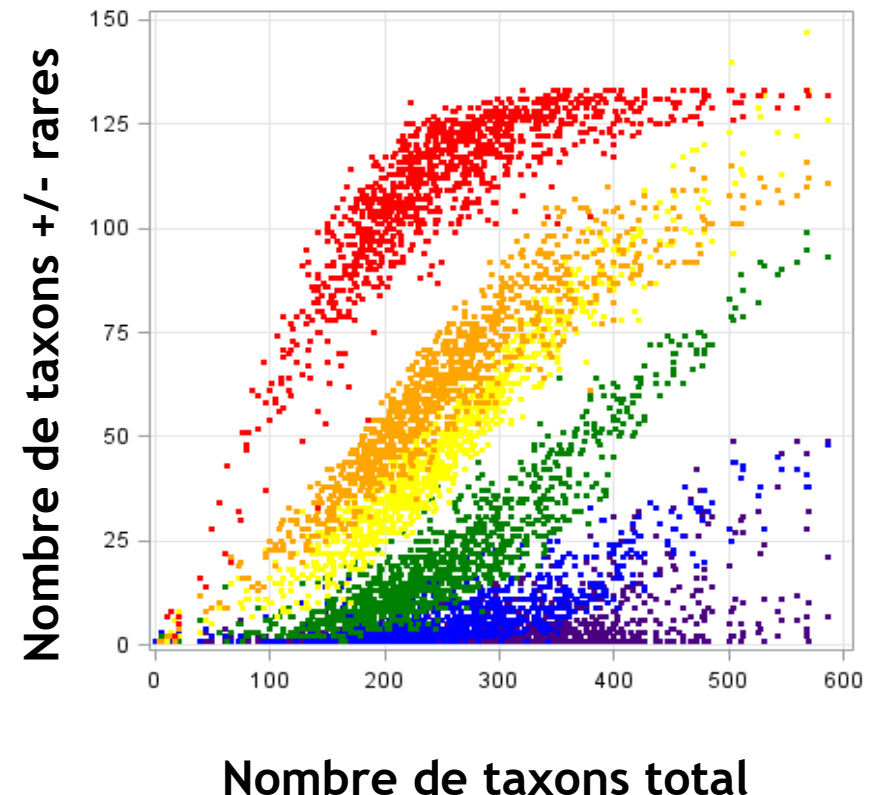
Plus de données = plus d'informations ?

3. Changement de comportement des observateurs

Influence critique des différences de détectabilité en $f(\text{rareté})$

On définit 8 classes de rareté :

- ≤ 5 mailles : exceptionnelle (1-E)
- 6-15 : très rare (2-RR)
- 16-50 : rare (3-R)
- 51-100 : assez rare (4-AR)
- 101-250 : peu commune (5-PC)
- 251-500 : assez commune (6-AC)
- 501-750 : commune (7-C)
- > 750 : très commune (8-CC)

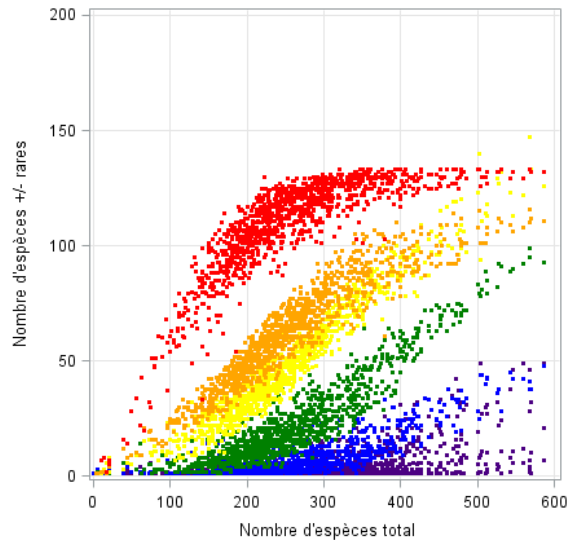


Plus de données = plus d'informations ?

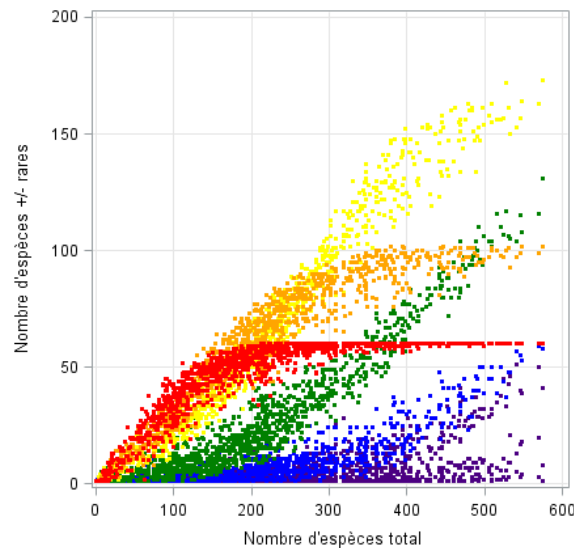
3. Changement de comportement des observateurs

Influence critique des différences de détectabilité en f(rareté)

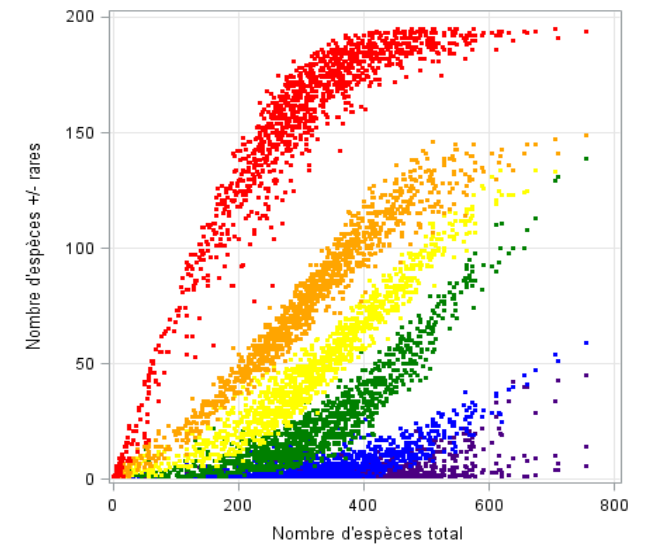
Atlas (< 1980)



1980-2000



> 2000



Inventaire systématique

Inventaire « conservation »

Inventaire « opportuniste »

On n'a pas échantillonné de la même manière !

=> Il faut tenter d'en tenir compte dans les analyses de données ...

Plus de données = plus d'informations ?

4. Précision de la localisation des observations

Importance des localisations précises pour des usages locaux

Faune



Données documentées comme précises mais improbables là ...

Données documentées comme imprécises mais correctes ...

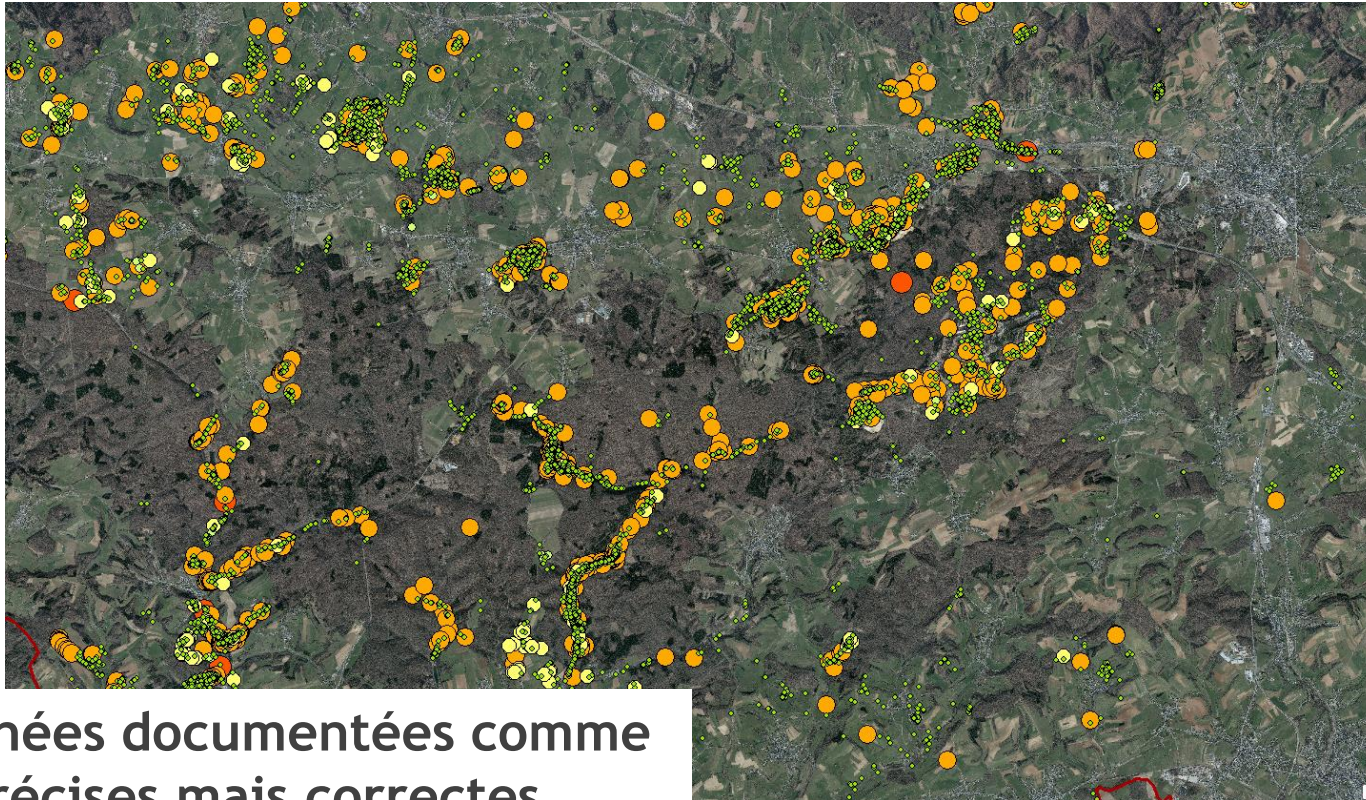


Plus de données = plus d'informations ?

4. Précision de la localisation des observations

Importance des localisations précises pour des usages locaux

Faune



Données documentées comme imprécises mais correctes ...

=> Il est nécessaire de documenter correctement la précision des observations pour pouvoir les associer à des sites précis

Pour toutes les espèces (> 2000)

Plus de données = plus d'informations ?

Importance de la validation et documentation

Groupes biologiques	Non-validé	Validé	Pct validé
Amphibiens	34 256	24 780	42.0%
Coléoptères	41 411	27 323	39.8%
Crustacés	2 327	1 722	42.5%
Dictyoptères	440	476	52.0%
Diptères	22 589	1 685	6.9%
Homoptères	10 568	998	8.6%
Hyménoptères	18 946	15 344	44.7%
Hétérocères	41 910	13 112	23.8%
Libellules	83 727	179 458	68.2%
Mammifères	148 197	105 861	41.7%
Mollusques	15 361	11 468	42.7%
Orthoptères	21 992	11 684	34.7%
Papillons	286 075	733 440	71.9%
Poissons	12 252	2 806	18.6%
Reptiles	22 505	35 098	60.9%
Total	762 556	1 170 040	60.5%

=> 60% des données sont validées

Pour toutes les espèces (> 2000)

Plus de données = plus d'informations ?

Importance de la validation et documentation

Groupes biologiques	Non-validé	Validé	Pct validé	Validé+100m	Pct précis	Pct conservés
Amphibiens	34 256	24 780	42.0%	23 121	93%	39%
Coléoptères	41 411	27 323	39.8%	26 310	96%	38%
Crustacés	2 327	1 722	42.5%	1 701	99%	42%
Dictyoptères	440	476	52.0%	453	95%	49%
Diptères	22 589	1 685	6.9%	1 613	96%	7%
Homoptères	10 568	998	8.6%	988	99%	9%
Hyménoptères	18 946	15 344	44.7%	14 967	98%	44%
Hétérocères	41 910	13 112	23.8%	11 255	86%	20%
Libellules	83 727	179 458	68.2%	153 309	85%	58%
Mammifères	148 197	105 861	41.7%	93 799	89%	37%
Mollusques	15 361	11 468	42.7%	10 319	90%	38%
Orthoptères	21 992	11 684	34.7%	10 919	93%	32%
Papillons	286 075	733 440	71.9%	649 003	88%	64%
Poissons	12 252	2 806	18.6%	2 296	82%	15%
Reptiles	22 505	35 098	60.9%	31 130	89%	54%
Total	762 556	1 170 040	60.5%	1 035 942	89%	54%

=> Plus de 50% des observations sont valorisables localement

Plus de données = plus d'informations ?

Importance de la validation et documentation

Groupes biologiques	Non-validé	Validé	Pct validé
Amphibiens	1319.0	5013.0	79.2%
Coléoptères	2238.0	4831.0	68.3%
Crustacés	401.0	640.0	61.5%
Dictyoptères	323.0	474.0	59.5%
Diptères	61.0	54.0	47.0%
Homoptères	34.0	49.0	59.0%
Hyménoptères	3888.0	4381.0	53.0%
Hétérocères	2148.0	3396.0	61.3%
Libellules	54.0	23869.0	99.8%
Mammifères	4392.0	10928.0	71.3%
Mollusques	152.0	3086.0	95.3%
Orthoptères	4057.0	5972.0	59.5%
Papillons	3706.0	96103.0	96.3%
Poissons	3099.0	953.0	23.5%
Reptiles	763.0	21430.0	96.6%
Total	26635.0	181179.0	87.2%

=> Plus de 85% des observations d'espèces protégées et LR sont validées

Plus de données = plus d'informations ?

Importance de la validation et documentation

Groupes biologiques	Non-validé	Validé	Pct validé	Validé+100m	Pct précis	Pct conservés
Amphibiens	1319.0	5013.0	79.2%	4880.0	97.3%	77%
Coléoptères	2238.0	4831.0	68.3%	4510.0	93.4%	64%
Crustacés	401.0	640.0	61.5%	639.0	99.8%	61%
Dictyoptères	323.0	474.0	59.5%	451.0	95.1%	57%
Diptères	61.0	54.0	47.0%	40.0	74.1%	35%
Homoptères	34.0	49.0	59.0%	47.0	95.9%	57%
Hyménoptères	3888.0	4381.0	53.0%	4178.0	95.4%	51%
Hétérocères	2148.0	3396.0	61.3%	3067.0	90.3%	55%
Libellules	54.0	23869.0	99.8%	21414.0	89.7%	90%
Mammifères	4392.0	10928.0	71.3%	10231.0	93.6%	67%
Mollusques	152.0	3086.0	95.3%	3037.0	98.4%	94%
Orthoptères	4057.0	5972.0	59.5%	5481.0	91.8%	55%
Papillons	3706.0	96103.0	96.3%	86543.0	90.1%	87%
Poissons	3099.0	953.0	23.5%	867.0	91.0%	21%
Reptiles	763.0	21430.0	96.6%	18431.0	86.0%	83%
Total	26635.0	181179.0	87.2%	163816.0	90.4%	79%

=> Près de 90% des données validées sont suffisamment précises

Plus de données = plus d'informations ?

Importance de la validation et documentation

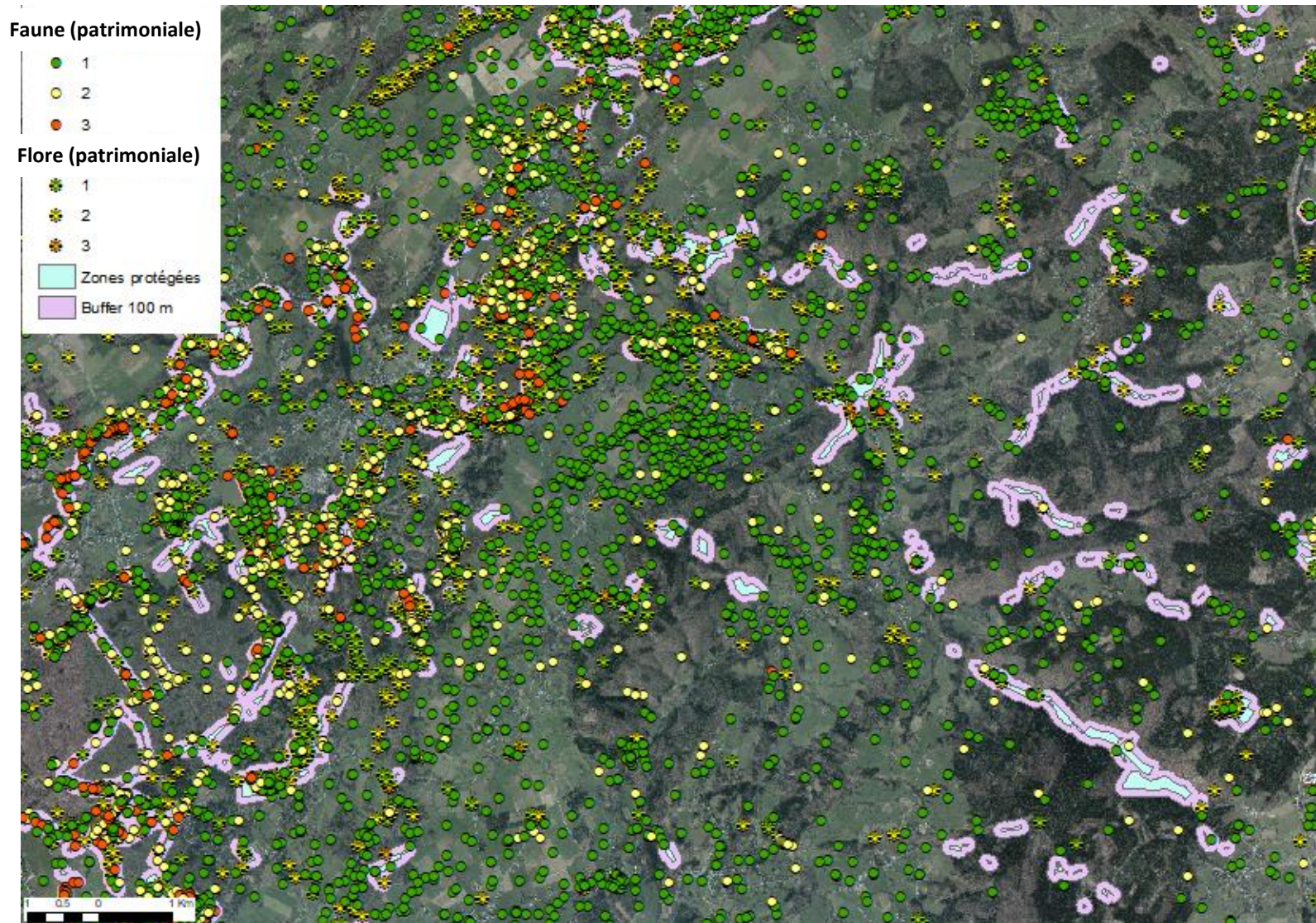
Espèces protégées et LR

Groupes biologiques	Non-validé	Validé	Pct validé	Validé+100m	Pct précis	Pct conservés	Pct initial
Amphibiens	1319.0	5013.0	79.2%	4880.0	97.3%	77%	8.3%
Coléoptères	2238.0	4831.0	68.3%	4510.0	93.4%	64%	6.6%
Crustacés	401.0	640.0	61.5%	639.0	99.8%	61%	15.8%
Dictyoptères	323.0	474.0	59.5%	451.0	95.1%	57%	49.2%
Diptères	61.0	54.0	47.0%	40.0	74.1%	35%	0.2%
Homoptères	34.0	49.0	59.0%	47.0	95.9%	57%	0.4%
Hyménoptères	3888.0	4381.0	53.0%	4178.0	95.4%	51%	12.2%
Hétérocères	2148.0	3396.0	61.3%	3067.0	90.3%	55%	5.6%
Libellules	54.0	23869.0	99.8%	21414.0	89.7%	90%	8.1%
Mammifères	4392.0	10928.0	71.3%	10231.0	93.6%	67%	4.0%
Mollusques	152.0	3086.0	95.3%	3037.0	98.4%	94%	11.3%
Orthoptères	4057.0	5972.0	59.5%	5481.0	91.8%	55%	16.3%
Papillons	3706.0	96103.0	96.3%	86543.0	90.1%	87%	8.5%
Poissons	3099.0	953.0	23.5%	867.0	91.0%	21%	5.8%
Reptiles	763.0	21430.0	96.6%	18431.0	86.0%	83%	32.0%
Total	26635.0	181179.0	87.2%	163816.0	90.4%	79%	8.5%

=> Moins de 10% des observations peuvent avoir un impact local

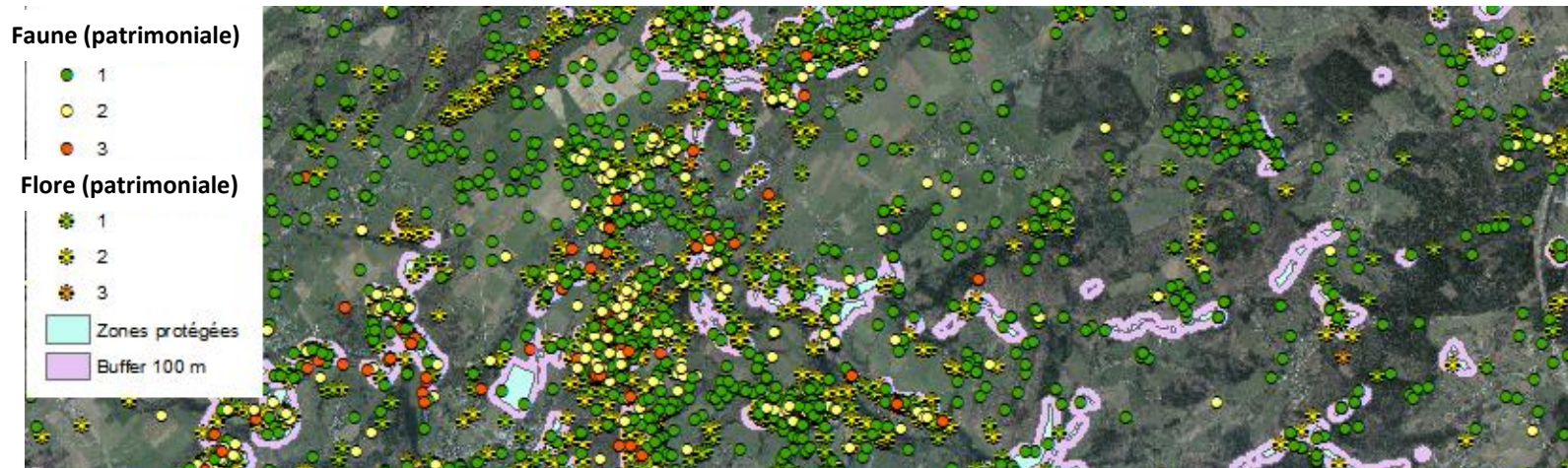
Plus de données = plus d'informations ?

5. Renforcement des zones déjà bien connues

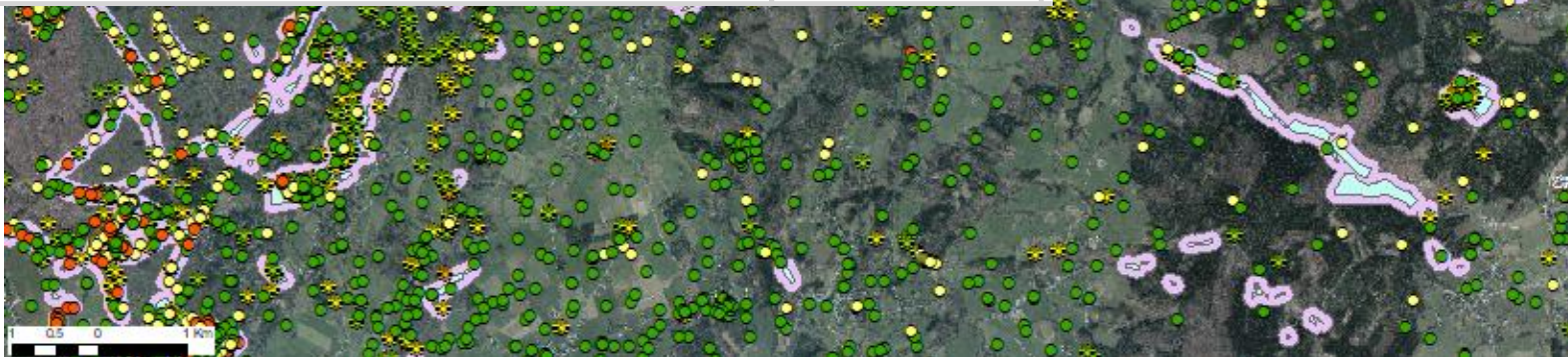


Plus de données = plus d'informations ?

5. Renforcement des zones déjà bien connues

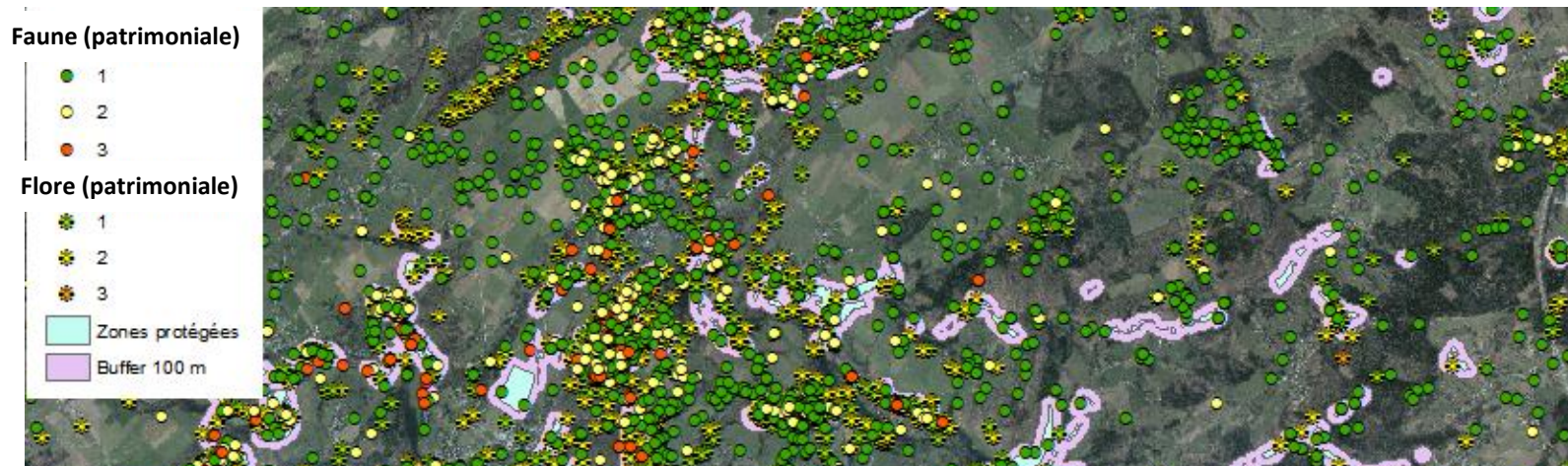


Type de zones	Surface
Zones protégées	2.5%
Zones protégées + 100 m	9.0%



Plus de données = plus d'informations ?

5. Renforcement des zones déjà bien connues



Type de zones	Surface	Flore	Faune	Oiseaux	Autres
Zones protégées	2.5%	21.3%	30.1%	22.4%	33.9%
Zones protégées + 100 m	9.0%	36.6%	45.1%	33.6%	50.8%

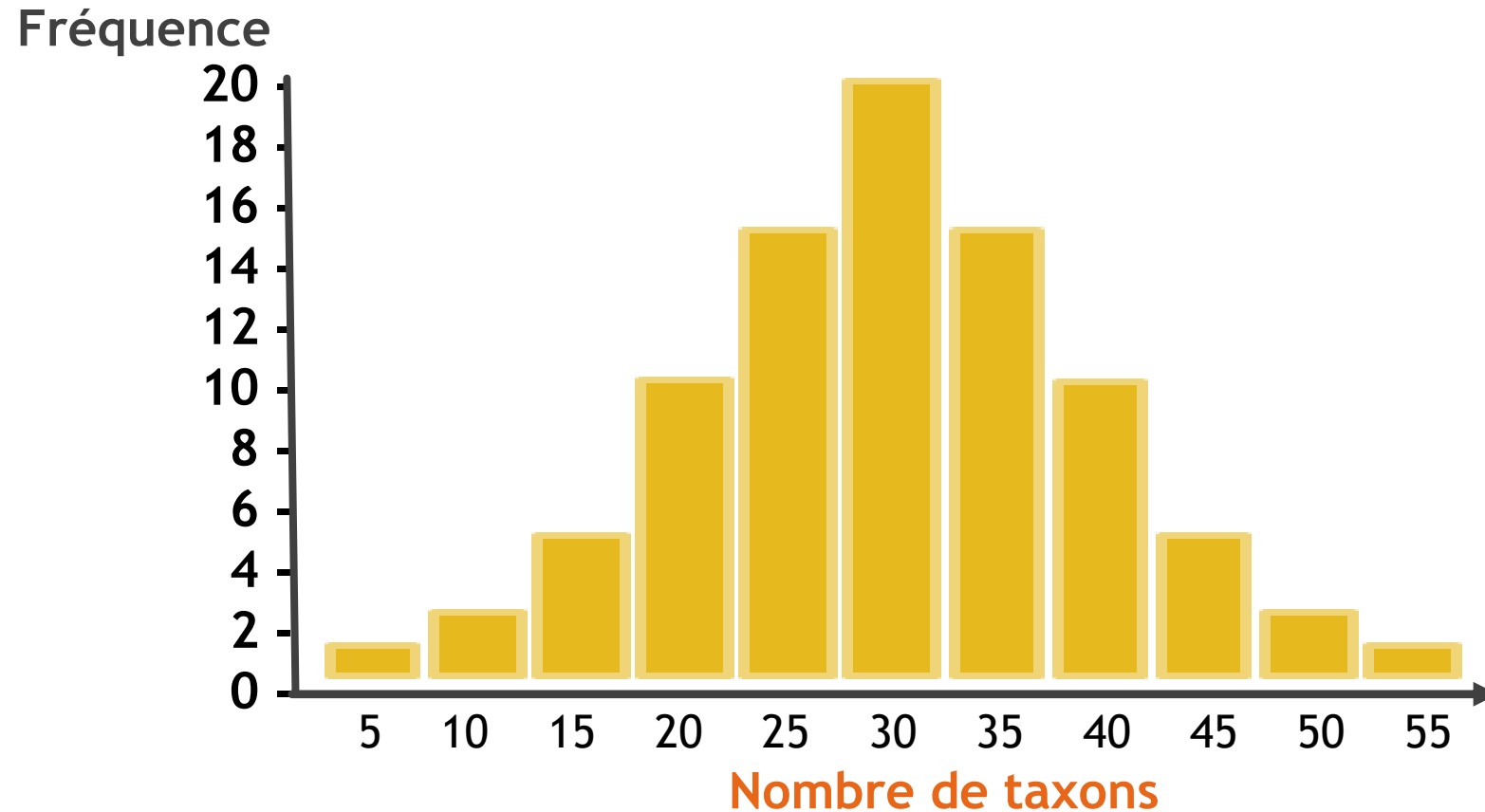


Plus de 36% (Flore) et 45% (Faune) des observations sont réalisées dans moins de 10% du territoire qui est déjà bien connu et protégé !



Plus de données = plus d'informations ?

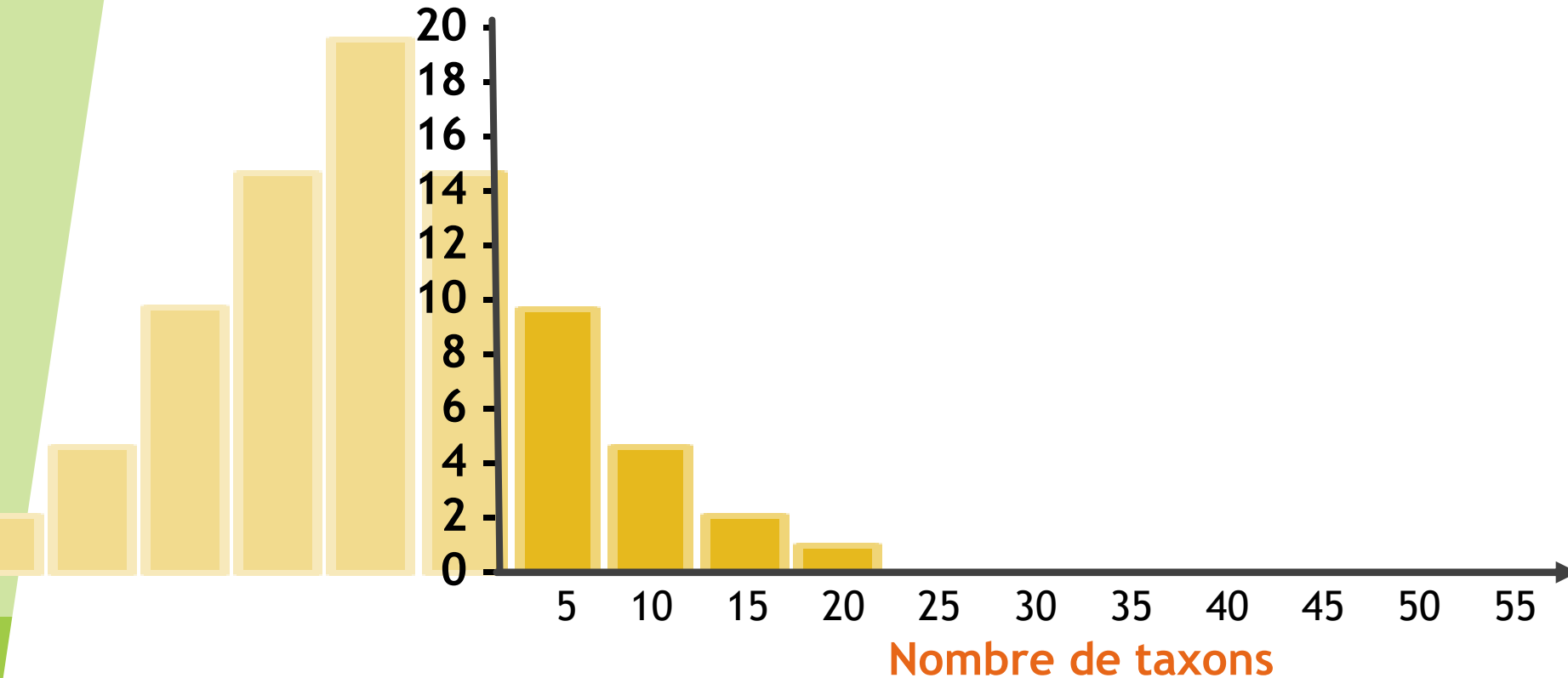
L'importance de la répartition spatiale de l'échantillonnage



Des atlas suffisamment exhaustifs se caractérisent par une distribution des richesses/maille en forme cloche

Plus de données = plus d'informations ?

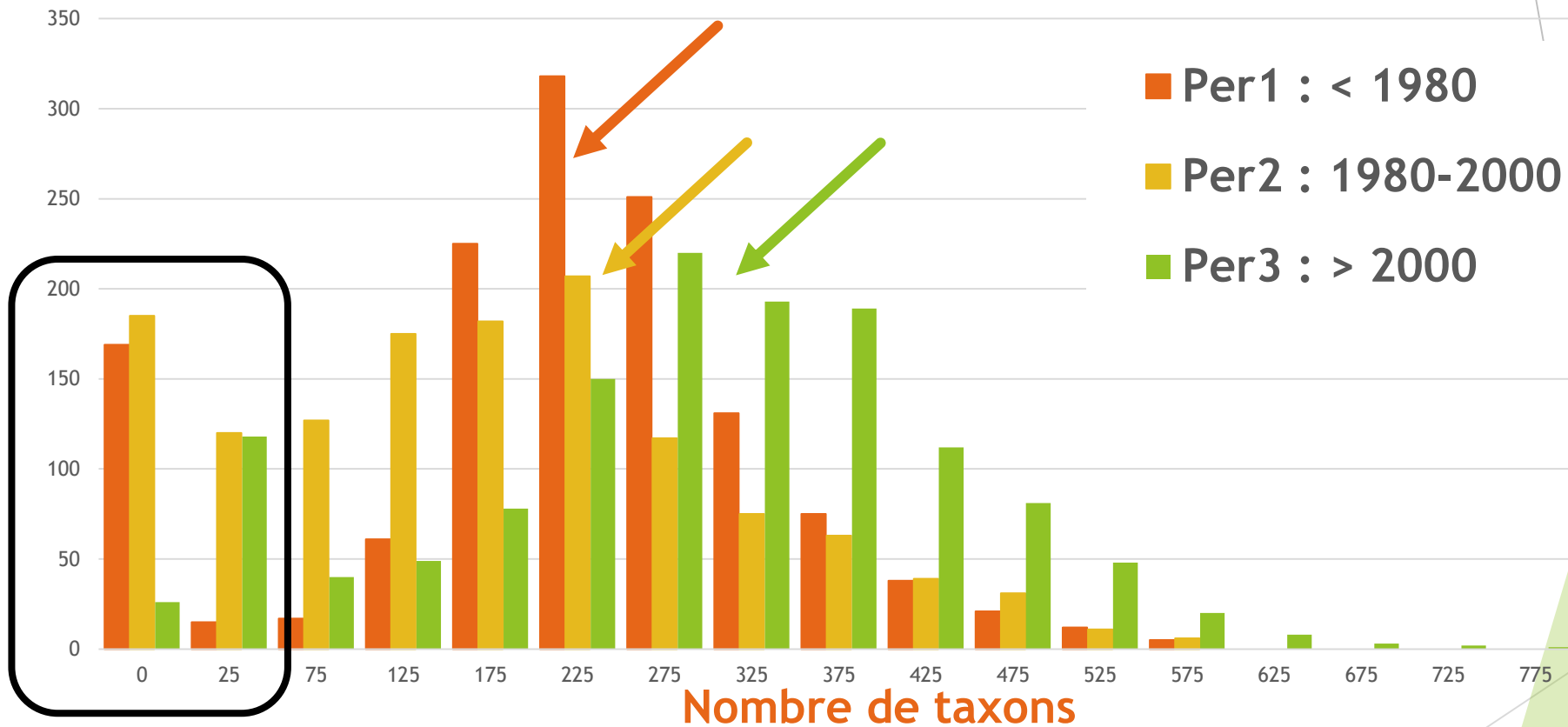
L'importance de la répartition spatiale de l'échantillonnage



Plus on inventorie de mailles et dans les mailles, plus on approche le nombre moyen d'espèces par maille

Plus de données = plus d'informations ?

L'importance de l'échantillonnage (maille 4x4 km)



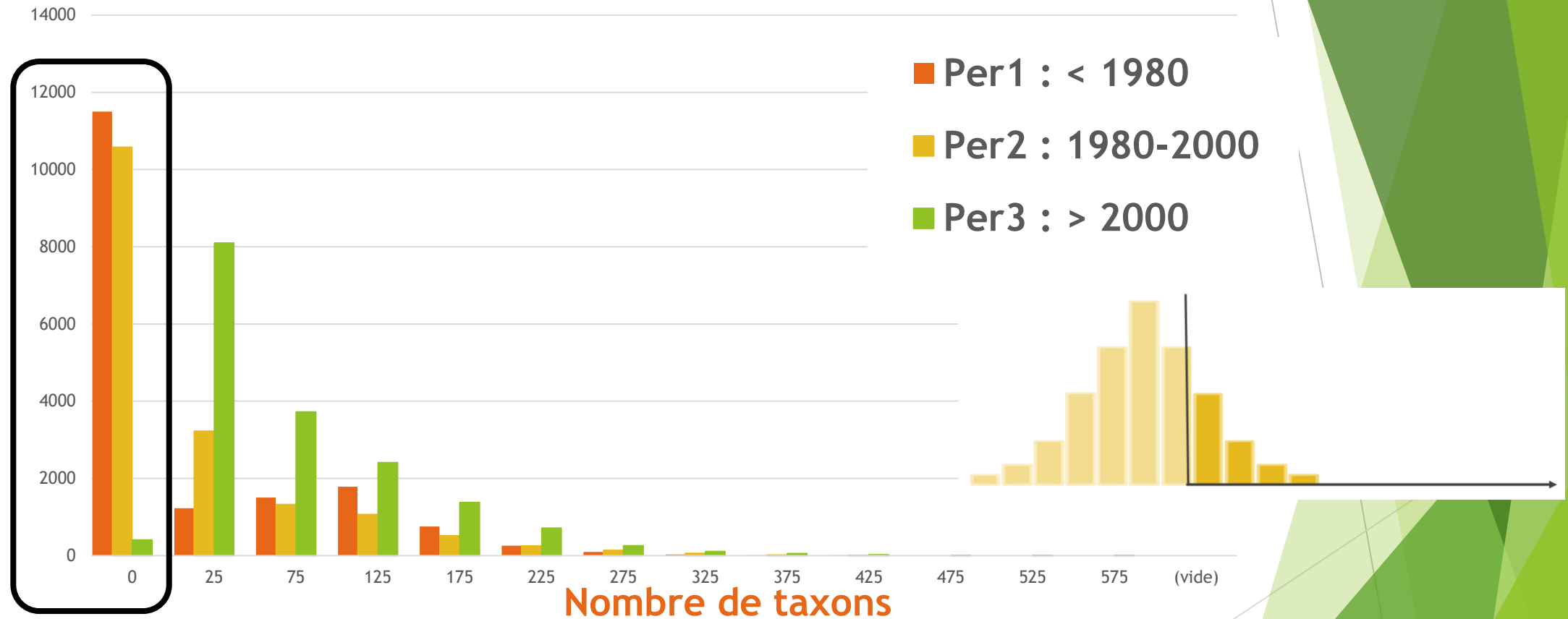
Flore

Plus d'espèces en moyenne depuis 2000 qu'avant 1980 et 2000

Peu de mailles peu inventoriées depuis 2000 (< 50 taxons)

Plus de données = plus d'informations ?

L'importance de l'échantillonnage (maille 1x1 km)

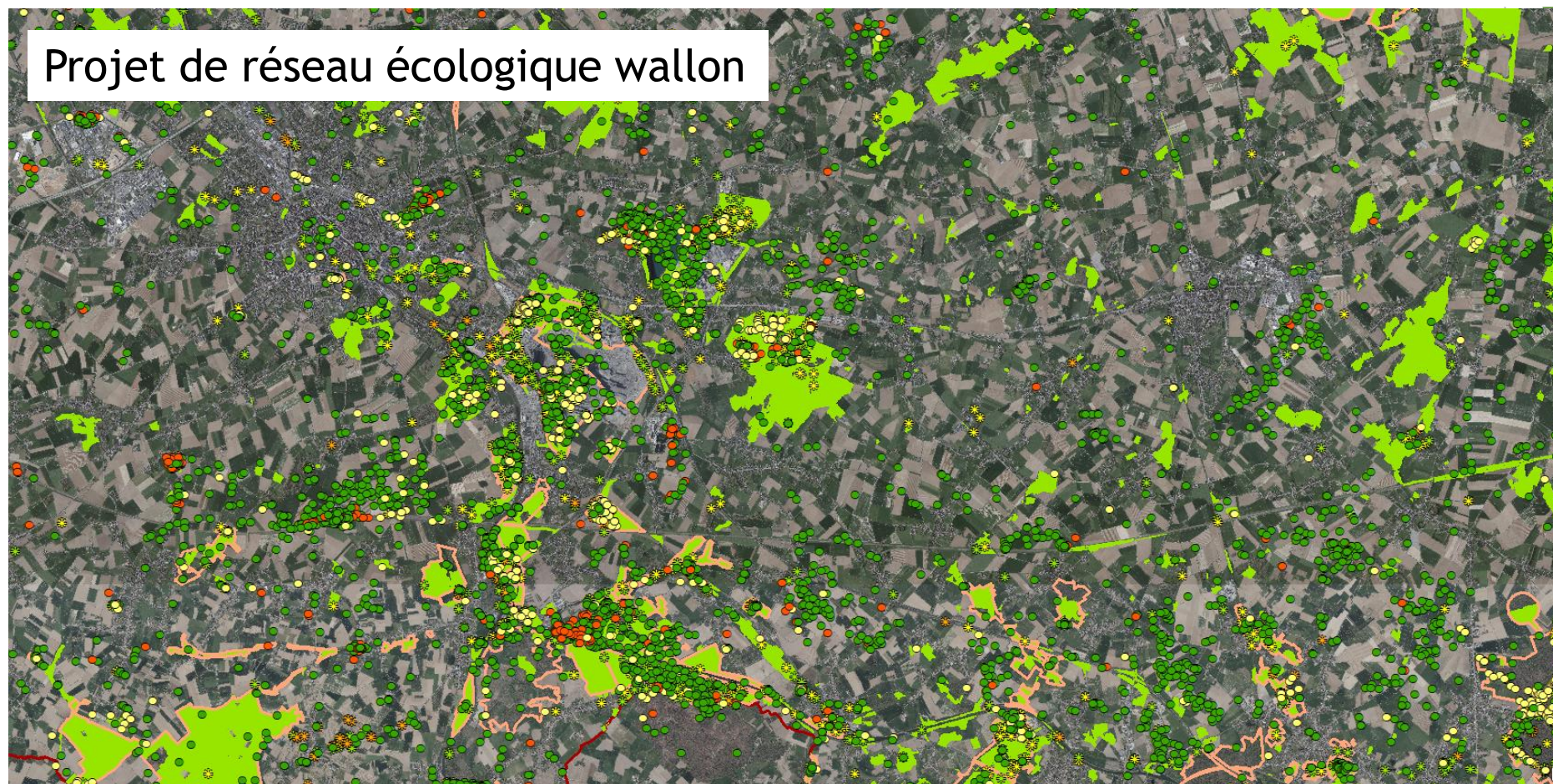


Flore

Dés qu'on zoome un peu, on est bien loin d'avoir une vision correcte de biodiversité présente

Plus de données = plus d'informations ?

L'importance de l'échantillonnage (échelle des sites)



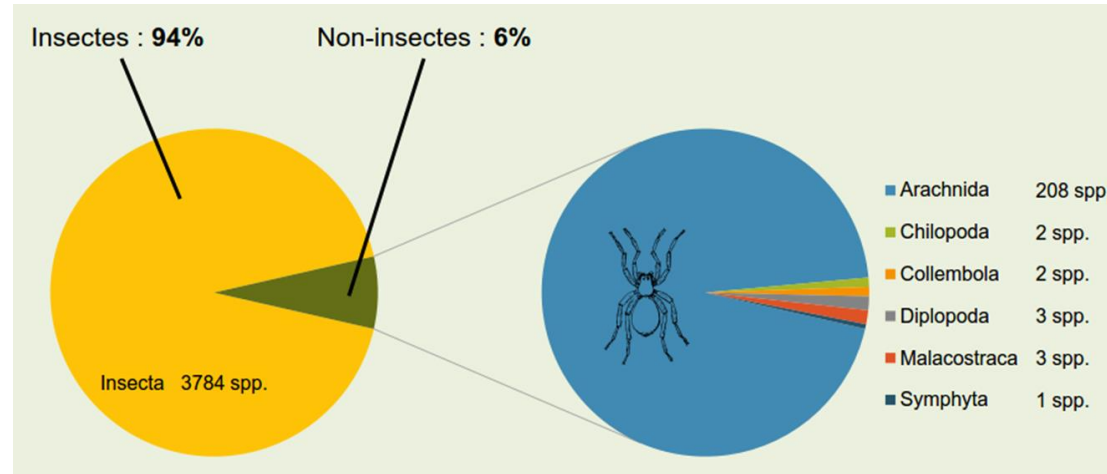
Des lacunes importantes existent, empêchant la mise en évidence d'un intérêt biologique stratégique qui existe très probablement

Plus de données = plus d'informations ?

Quand on cherche, on trouve !



Jardin Jean Massart à Bruxelles



- Grande partie de la faune belge des Coccinelles (62%), Carabides (28%), Cucurlionidés (32%), Scolytinae (40%), ...
 - **160 espèces nouvelles pour la faune belge, 3 espèces nouvelles pour la science, ...**
- => On ne connaît en fait qu'une toute petite partie des enjeux de la biodiversité qui peut être présente partout ...
- => A quand l'obligation de faire un inventaire systématique préventif en cas de destruction de sites potentiellement intéressants ?

Plus il y a de données, mieux c'est ?

Oui ! Mais si elles sont de qualité, c'est encore mieux !

- Encoder le niveau de certitude des déterminations
- Structurer et participer à la validation des données (taxo, XY, comportement, ...)
- Décrire au mieux les observations d'espèces « sensibles »
- Ne pas hésiter à faire des listes d'espèces, de renseigner des biotopes originaux, ...



- Et puis, sortir des sentiers battus pour que la Wallonie montre sa vraie richesse !

