



SINP

Système d'Information
sur la Nature et le Paysage



Analyse des données du Dépôt Légal de Biodiversité

Journées techniques du SINP
21 mars 2019



AGENCE FRANÇAISE
POUR LA BIODIVERSITÉ
ÉTABLISSEMENT PUBLIC DE L'ÉTAT



Mathias Laville
mathias.laville@mnhn.fr

Analyse des données DEPOBIO – Quels objectifs ?

Des jeux de données sont versés sur la plateforme depuis 6 mois ...

Au regard du standard de l'INPN et des exigences en terme de qualité de la donnée, où en est-on ?

Peut-on intégrer l'intégralité des données directement ?

- Sinon, quelles données peuvent être intégrées au SINP ?
- Suivant quels critères ?
- Quelles précautions faut-il adopter ?
- Des points particuliers auxquels il faut être attentif ?

Analyse des données DEPOBIO – Quels objectifs ?

- Identifier les problèmes récurrents dans la construction ou le contenu des jeux de données et des métadonnées.
-> Quoi ?
- Identifier les principaux facteurs déterminants dans la construction des jeux de données et l'émergence des problèmes éventuels.
-> Comment ? Pourquoi ?
- Appréhender les risques potentiels associés aux problèmes identifiés.
-> Quelles conséquences potentielles ?
- Proposer des améliorations pérennes et / ou solutions à cours terme pour faire face aux risques identifiés.

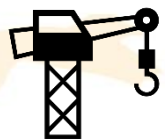
Etat des lieux

718 jeux de données



Etat des lieux

299 cadres d'acquisition
= Projets



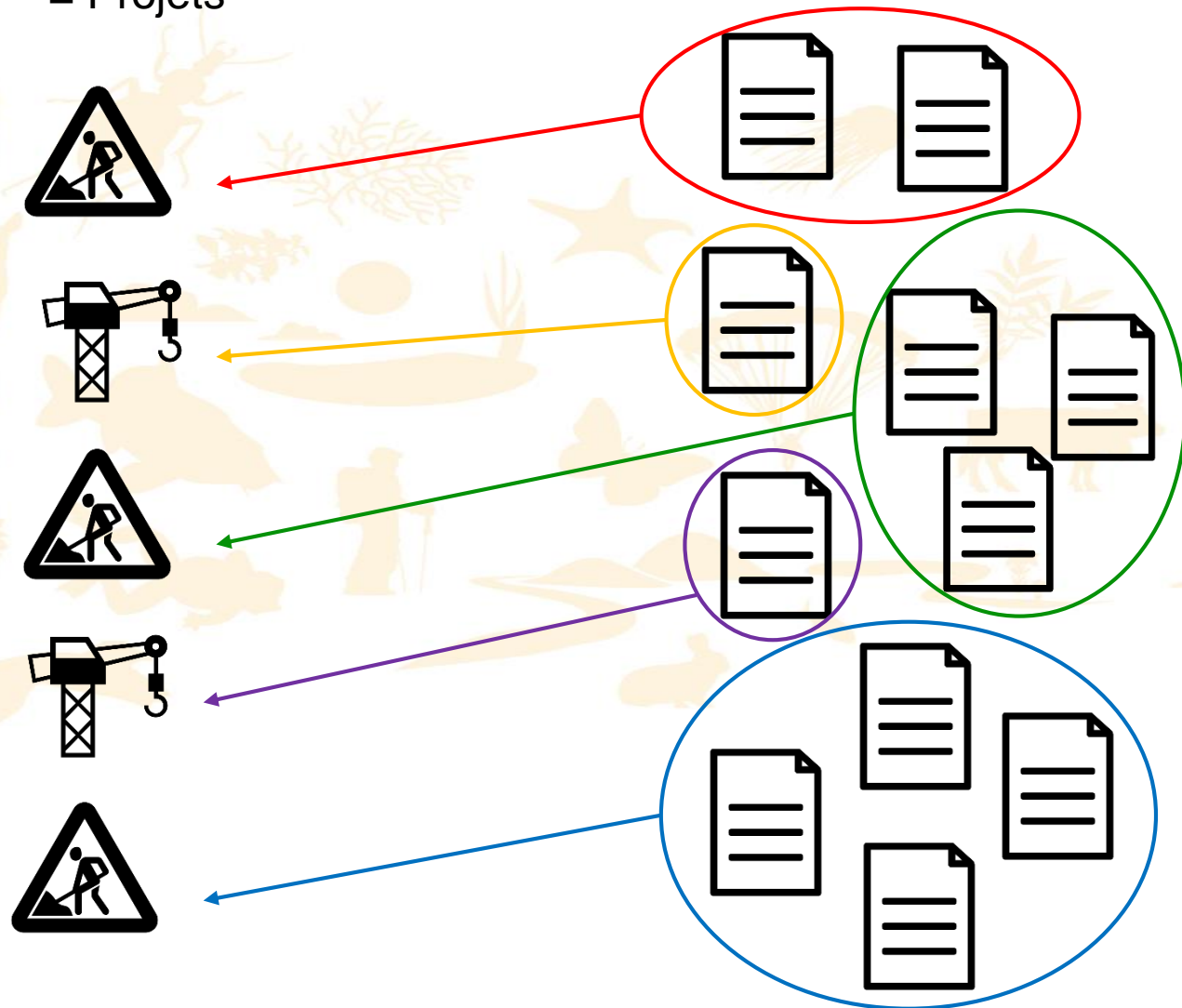
718 jeux de données



Etat des lieux

299 cadres d'acquisition
= Projets

718 jeux de données



Minimum = 1 JDD

Maximum = 35 JDD

Médiane = 1 JDD

Moyenne = 2,4 JDD

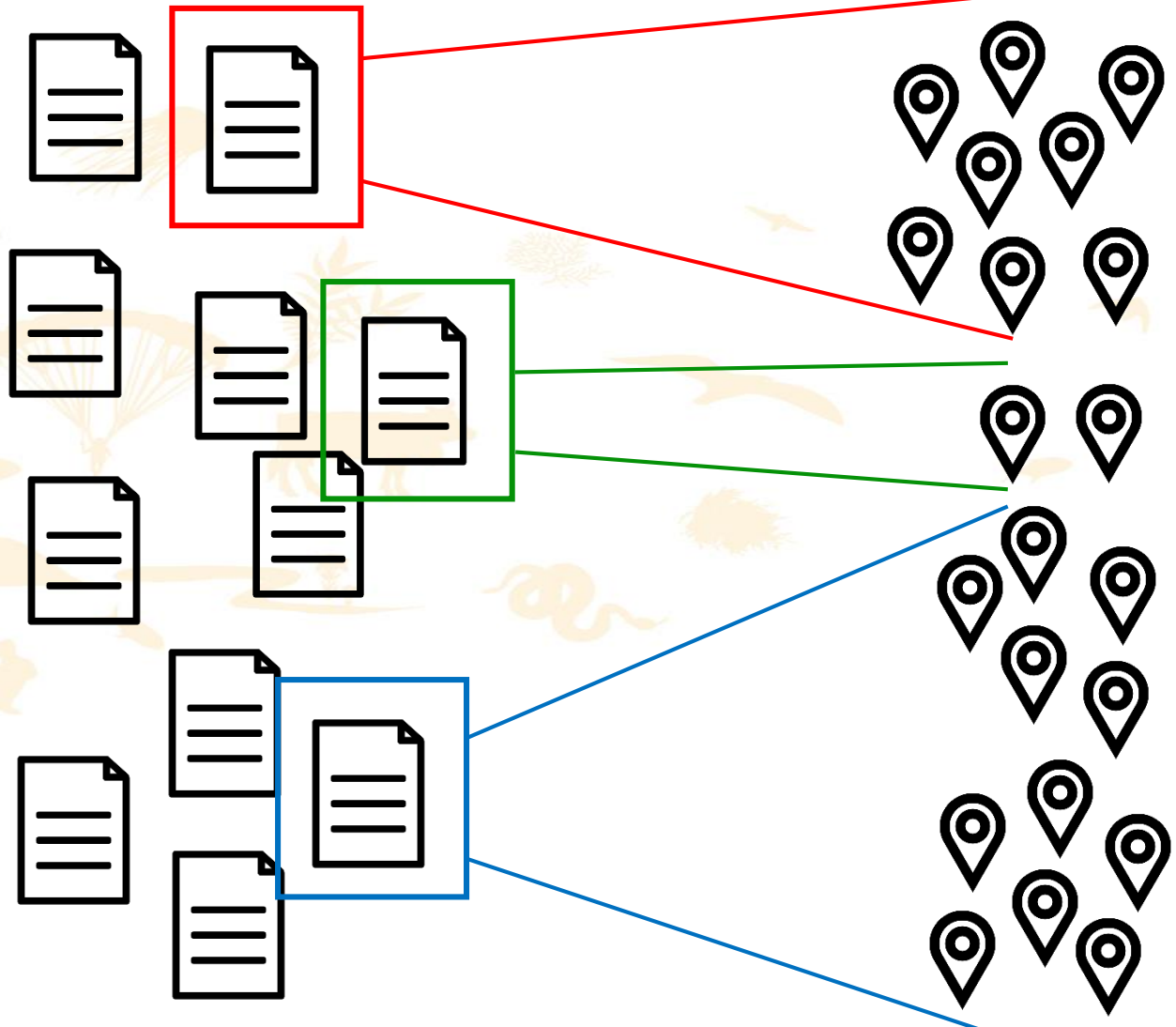
Etat des lieux

718 jeux de données

479 339 données



Minimum = 1 donnée
Maximum = 26 754 données
Médiane = 132 données
Moyenne = 668 données

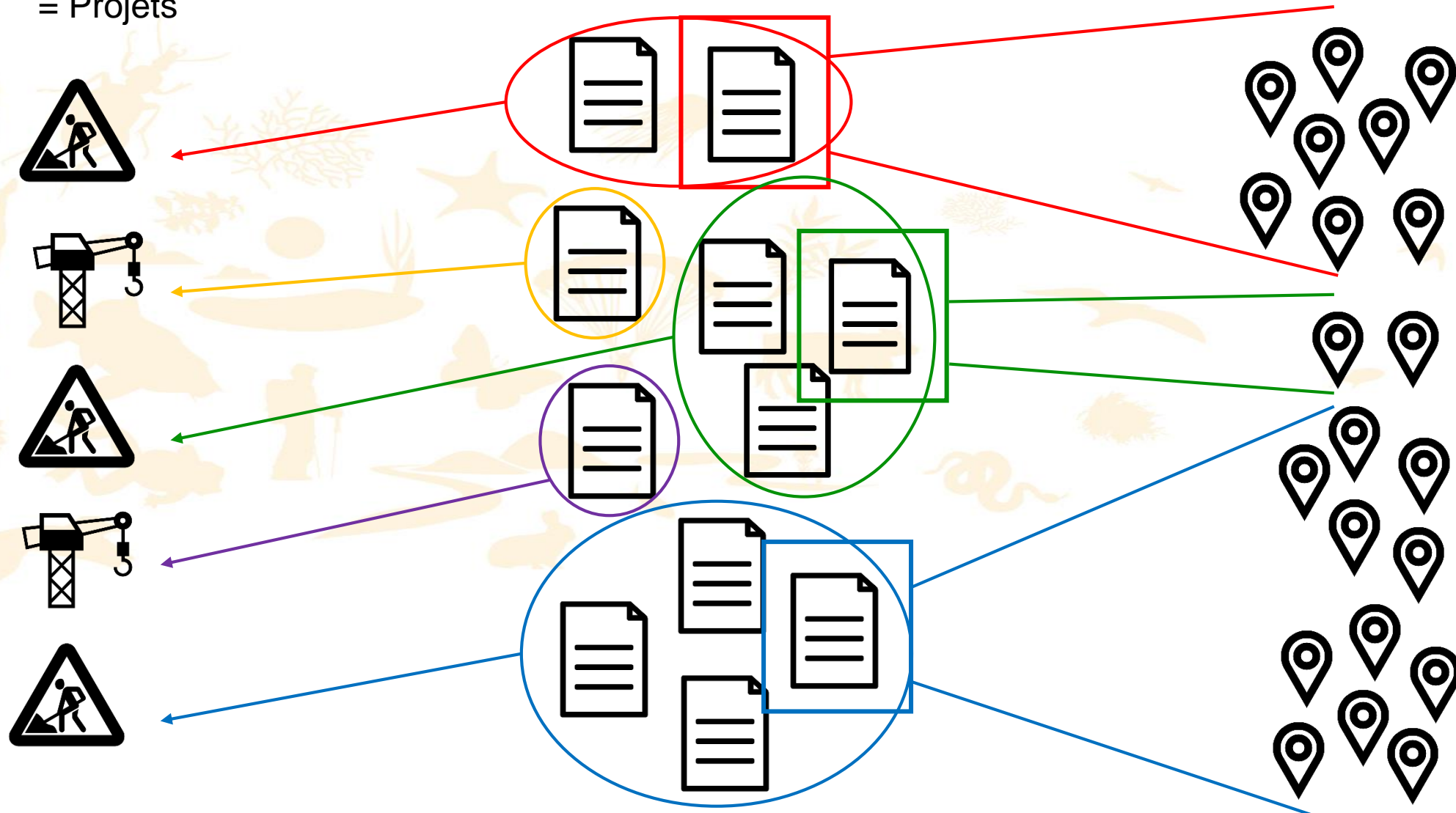


Etat des lieux

299 cadres d'acquisition
= Projets

718 jeux de données

479 339 données

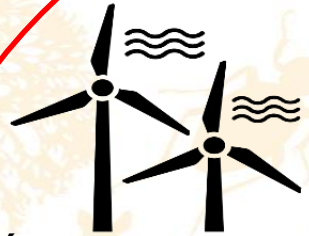


Etat des lieux – Projets & infrastructures

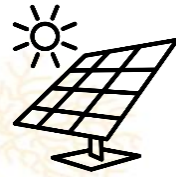


Etat des lieux – Projets & infrastructures

Énergie Ressources



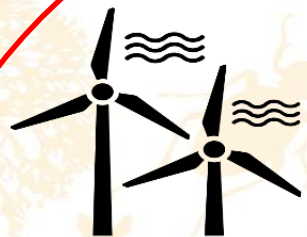
Éolienne : 145



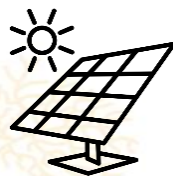
Photovoltaïque : 62

Etat des lieux – Projets & infrastructures

Énergie Ressources



Éolienne : 145



Photovoltaïque : 62



Hydroélectrique : 8



Carrières / mines : 28



Cimenterie : 2



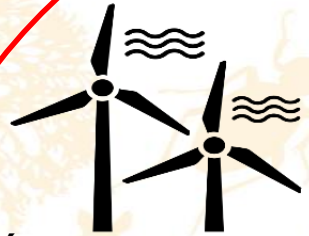
Gaz : 7



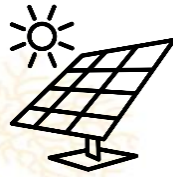
Hydrocarbures : 1

Etat des lieux – Projets & infrastructures

Énergie Ressources



Éolienne : 145



Photovoltaïque : 62



Hydroélectrique : 8



Carrières / mines : 28



Cimenterie : 2



Gaz : 7

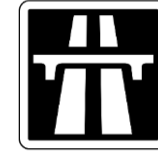


Hydrocarbures : 1

Logistique Réseaux de transports



Logistique : 8



Routes : 6



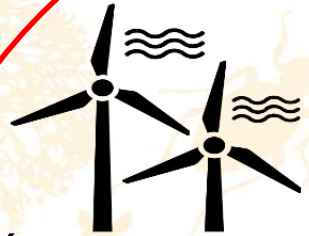
Aires de stationnement : 2



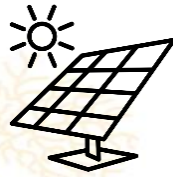
Divers transports (par câble, tramway, ferroviaire, port de plaisance) : 4

Etat des lieux – Projets & infrastructures

Énergie Ressources



Éolienne : 145



Photovoltaïque : 62



Hydroélectrique : 8



Carrières / mines : 28



Cimenterie : 2



Gaz : 7

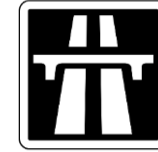


Hydrocarbures : 1

Logistique Réseaux de transports



Logistique : 8



Aires de stationnement : 2

Routes : 6



Divers transports (par câble, tramway, ferroviaire, port de plaisance) : 4

Gestion des déchets Traitement de matériaux



Déchets : 6



Traitement de matériaux : 4

+ ZAC, centre commercial : 3

+ Autres / Non déterminés : 13

Etat des lieux – Informations géographiques et territoires

390 119 données associées à une **géométrie** :

- **Point** -> 316 677 données (66 %)
- **Polygone** -> 65 673 données (14 %)
- **Ligne** -> 7 769 données (2 %)

72 954 données rattachées à une entité :

- **Commune** -> 69 789 données (15 %)
- **Maille 10 x 10 km** -> 1 789 données (- 1 %)
- **Département** -> 1 376 données (- 1 %)

+ 16 266 données non localisées (3 %)

Etat des lieux – Informations géographiques et territoires

390 119 données associées à une **géométrie** :

- **Point** -> 316 677 données (66 %)
- **Polygone** -> 65 673 données (14 %)
- **Ligne** -> 7 769 données (2 %)

72 954 données rattachées à une entité :

- **Commune** -> 69 789 données (15 %)
- **Maille 10 x 10 km** -> 1 789 données (- 1 %)
- **Département** -> 1 376 données (- 1 %)

+ 16 266 données non localisées (3 %)

Les territoires concernés :

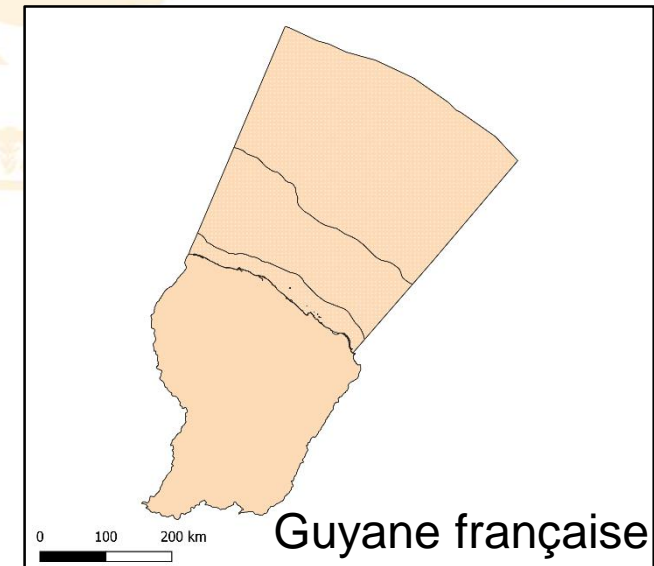
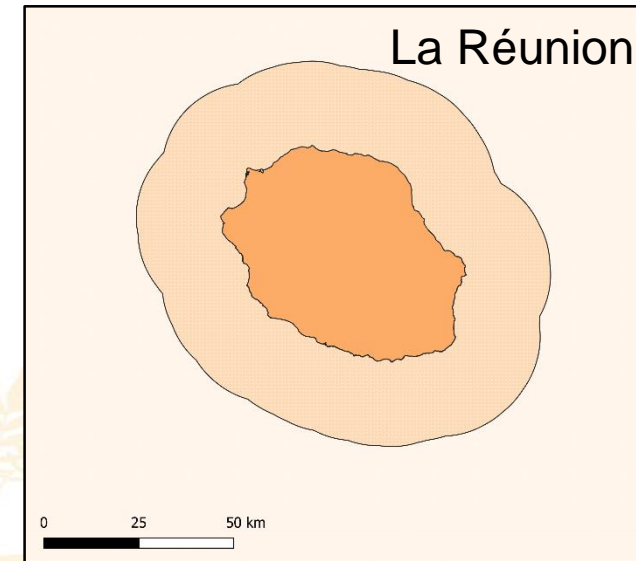
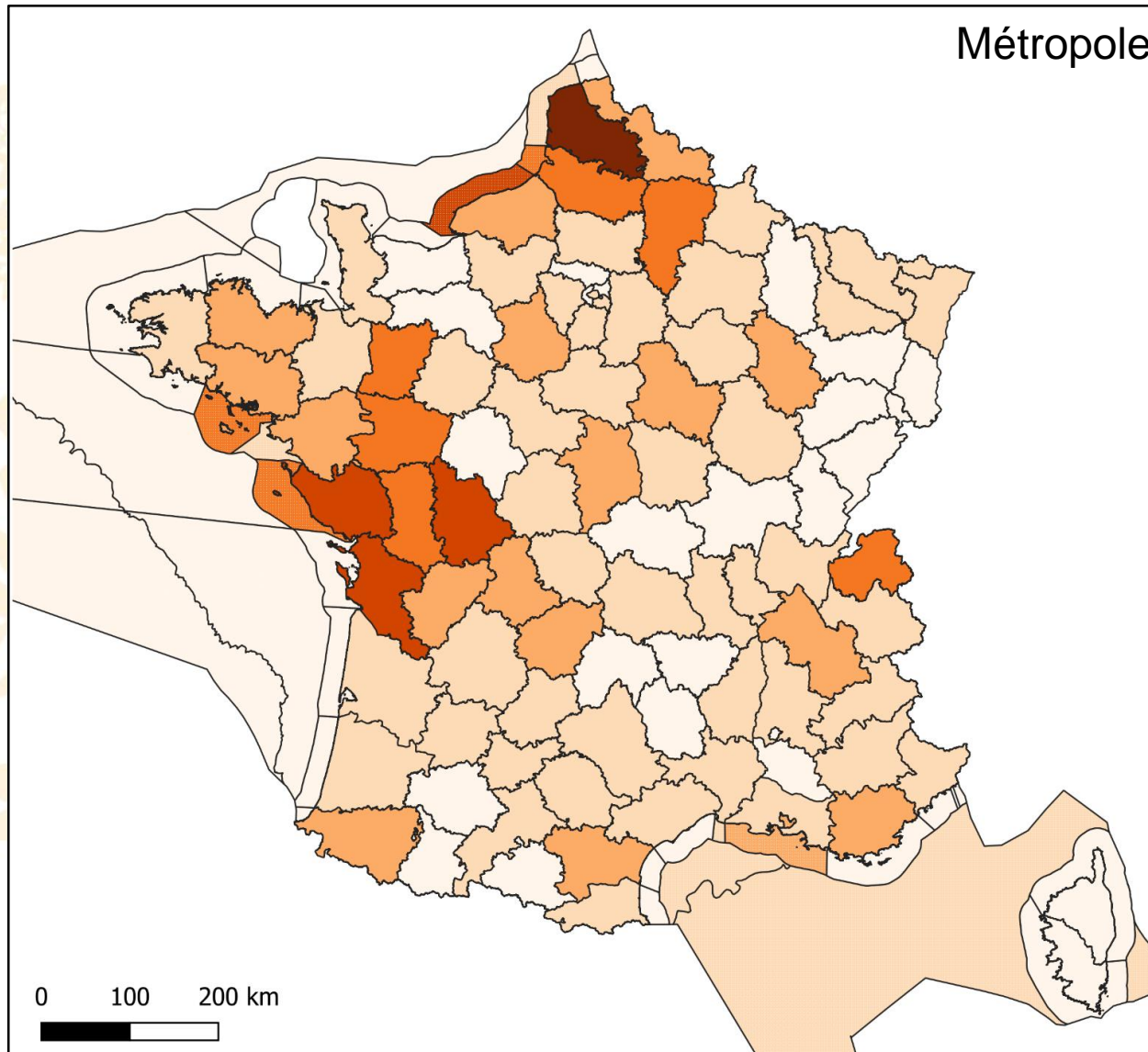
Métropole : 381 747 données

La Réunion : 3 681 données

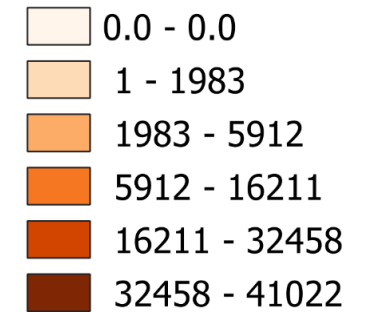
Guyane française : 721 données

Hors France : 76 924 données

Etat des lieux – Emprise spatiale des données

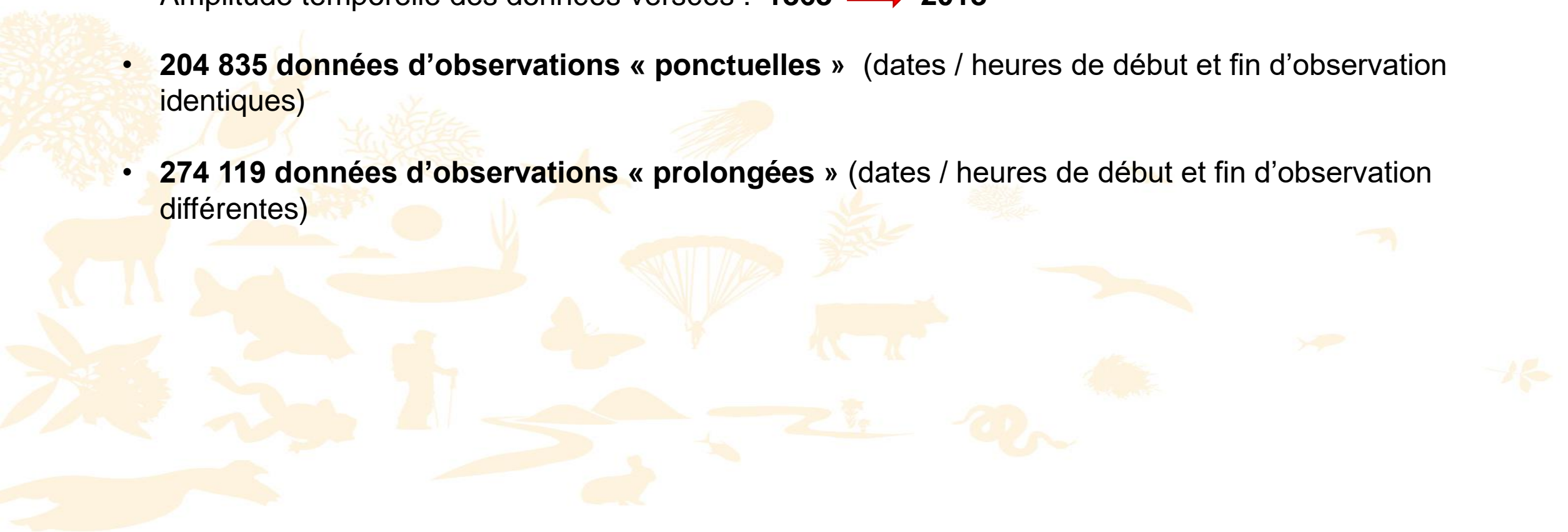


Nombre de données



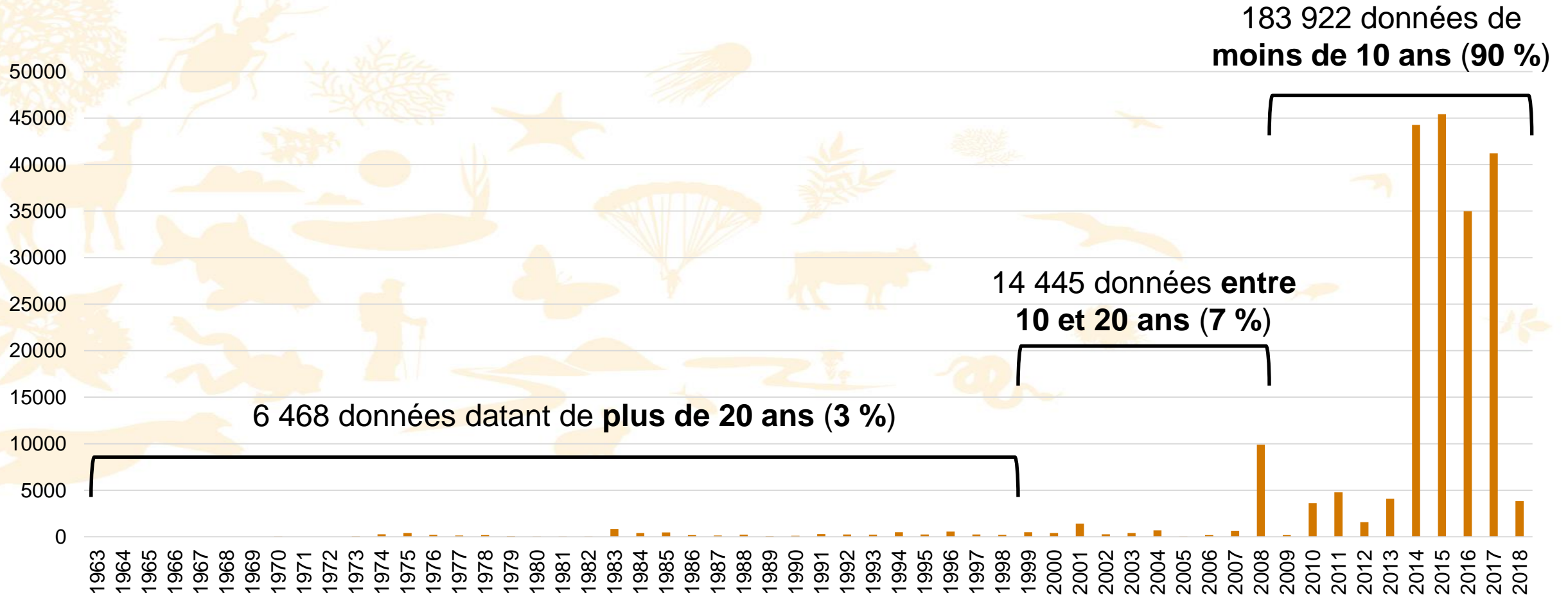
Etat des lieux – Temporalité

- Amplitude temporelle des données versées : **1863** ➔ **2018**
- **204 835 données d'observations « ponctuelles »** (dates / heures de début et fin d'observation identiques)
- **274 119 données d'observations « prolongées »** (dates / heures de début et fin d'observation différentes)



Etat des lieux – Temporalité

- Données d'observations « ponctuelles » :



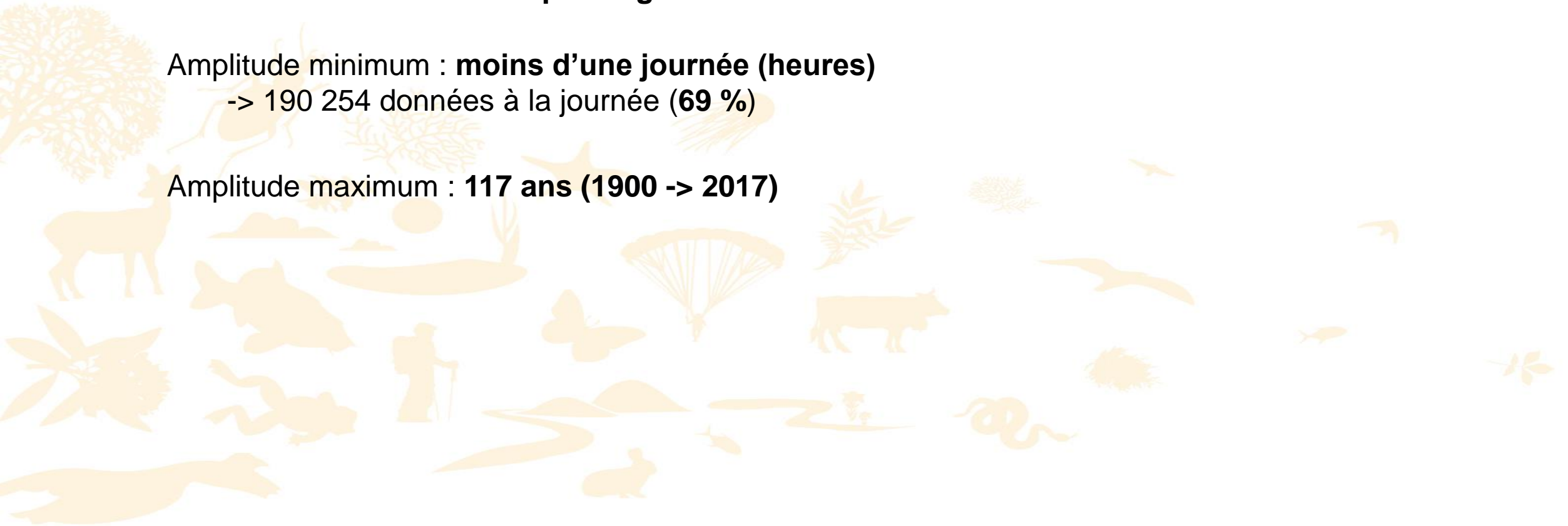
Etat des lieux – Temporalité

- **Données d'observations « prolongées » :**

Amplitude minimum : **moins d'une journée (heures)**

-> 190 254 données à la journée (69 %)

Amplitude maximum : **117 ans (1900 -> 2017)**



Etat des lieux – Source des données

Détail des différents statuts de données :

- **Privée** : **420 674** données (88 %)
- **Acquise** : **29 411** données (6 %)
- **Publique** : **17 997** données (4 %)
- **Indéterminé** : **11 257** données (2 %)

Provenance des données :

- **Terrain** : **477 631** données
- **Littérature** : **25** données
- **Collection** : **2** données
- **Indéterminé** : **1 681** données

Etat des lieux – Source des données

Détail des différents statuts de données :

- **Privée** : **420 674** données (88 %)
- **Acquise** : **29 411** données (6 %)
- **Publique** : **17 997** données (4 %)
- **Indéterminé** : **11 257** données (2 %)

Provenance des données :

- **Terrain** : **477 631** données
- **Littérature** : **25** données
- **Collection** : **2** données
- **Indéterminé** : **1 681** données



Aucune déclaration de réutilisation de données précédemment versées dans le SINP !

Etat des lieux – Noms cités et CD_NOM

Problèmes identifiés sur les noms cités (champ obligatoire) :

- Encodage
- Quelques valeurs aberrantes

Avant la
réconciliation
taxonomique

42,3 % des données sans CD_NOM

Etat des lieux – Noms cités et CD_NOM

Problèmes identifiés sur les noms cités (champ obligatoire) :

- Encodage
- Quelques valeurs aberrantes

Avant la
réconciliation
taxonomique

42,3 % des données sans CD_NOM



Après la
réconciliation
taxonomique

17,6 % des données sans CD_NOM

Etat des lieux – Noms cités et CD_NOM

Problèmes identifiés sur les noms cités (champ obligatoire) :

- Encodage
- Quelques valeurs aberrantes

Avant la
réconciliation
taxonomique

42,3 % des données sans CD_NOM



Réconciliation pour **58,4 %** des données sans CD_NOM

Après la
réconciliation
taxonomique

17,6 % des données sans CD_NOM

Analyses – Démarche d'analyse

Les analyses sont pratiquées à **différentes échelles** : à la fois sur **les métadonnées** renseignées, **les jeux de données** dans leur globalité et sur **les données** en elles mêmes.

Elles s'articulent en deux phases :

- 1) Une phase d'analyses quantitatives permettant de dégager les principaux patrons d'organisation.
- 2) Une phase plus qualitative sur la base d'échantillonnages stratifiés aux vues des patrons identifiés en phase 1.

-> sollicitation probable des différents acteurs mis en jeu dans le processus

Analyses – Démarche d'analyse

Les analyses vont se structurer autour de différents axes dont notamment :

Les acteurs

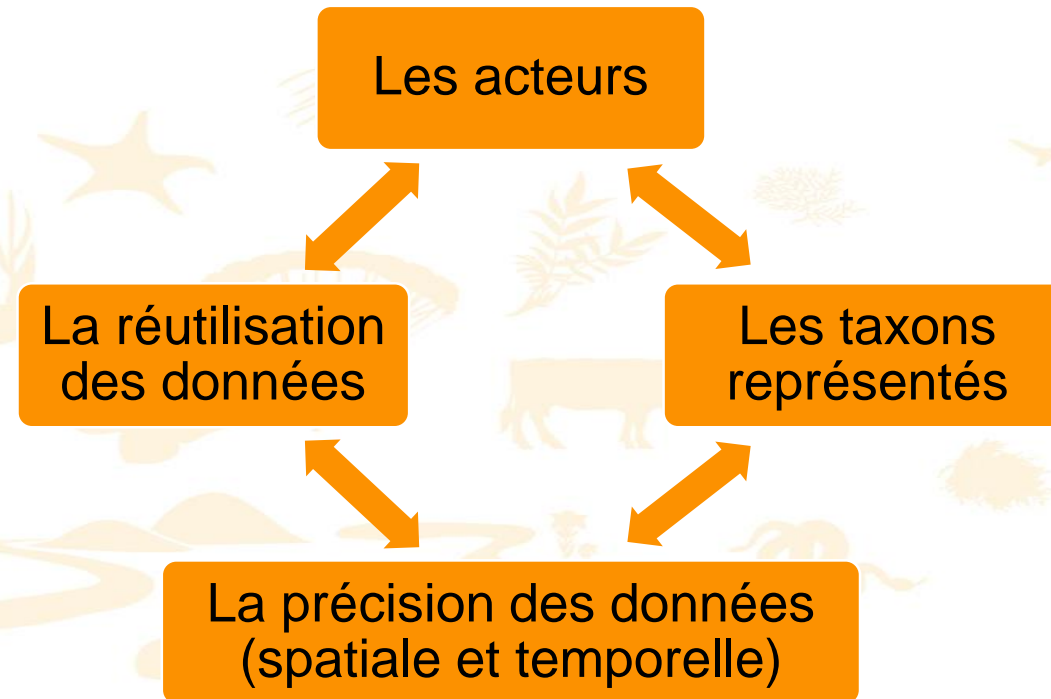
La réutilisation
des données

Les taxons
représentés

La précision des données
(spatiale et temporelle)

Analyses – Démarche d'analyse

Les analyses vont se structurer autour de différents axes dont notamment :



L'objectif étant de pouvoir **décrire les principaux patrons observés** pour chaque axe et **comment ils s'articulent entre eux**.

Merci de votre attention

Crédits icônes :

solar by Atif Arshad from the Noun Project

Windmill by ProSymbols from the Noun Project

marble quarry by veysel from the Noun Project

Mine by arif fajar yulianto from the Noun Project

hydroelectric by Olena Panasovska from the Noun Project

Hydropower by sahua d from the Noun Project

Gas by priyanka from the Noun Project

warehouse by Mohamad Arif Prasetyo from the Noun Project

logistic by Komkrit Noenpoempisut from the Noun Project

Highway by Myly from the Noun Project

Waste by Akash from the Noun Project

Waste by Rose Alice Design from the Noun Project

Mountain by Vadim Solomakhin from the Noun Project

Bricks by Dinosoft Labs from the Noun Project

meteorite by mangin maëlys from the Noun Project

Crane by Iconic from the Noun Project

work in progress by Gleb Khorunzhiy from the Noun Project

Map Marker by Role Play from the Noun Project

Data by projecthayat from the Noun Project

Train by ladhina from the Noun Project

Boat by Orin zuu from the Noun Project

cable car by anbilero adaleru from the Noun Project

Fuel by Payungkead Im-anong from the Noun Project

Concrete Mixer by DPlcons from the Noun Project

Parking by Margot Lnrs from the Noun Project