

# De la Donnée à la Décision : Enrichissement par l'Open Data et Prédictions du risque sécheresse en 2023

21/11 à 11h45



**Niels LUBBERS**  
BPCE Assurances IARD



**Anthony LAM**  
BPCE Assurances IARD



**Rémi BELLINA**  
Milliman



**Antoine RAINAUD**  
Milliman

# Notre vision de l'Open Data

Utilisation de l'Open Data pour l'enrichissement de données internes

## Partager ses données publiques pour créer de la valeur

### Création d'une plateforme Open Data

En 2017, création de la plateforme Open Data du Groupe BPCE avec les objectifs suivants :

- Partager les données publiques des entreprises du Groupe BPCE pour accélérer la création de valeur et les usages
- Permettre à nos collaborateurs de créer de la valeur en utilisant l'Open Data
- Faire simple pour nos clients, nos collaborateurs et nos partenaires
- 180 jeux de données à destination du grand public

The screenshot displays the BPCE Open Data platform interface. At the top, there's a navigation bar with 'ACCUEIL', 'CHIFFRES CLEFS', 'DONNÉES', 'EDITEUR CARTES', 'EDITEUR GRAPHES', 'NOS OFFRES D'EMPLOI', and 'CONTACT'. The main content area is divided into several sections:

- 180 jeux de données**: A section for browsing datasets with filters for 'Thème' (e.g., Etudes, Economie du Sport, RSE) and 'Producteur' (e.g., L'Observatoire BPCE, BPCE Digital & Payments).
- Catalogue et statistiques d'utilisation des jeux de données Open Data**: A section providing statistics on the usage of Open Data datasets.
- Groupes BPCE - Offres emploi publiques**: A section for public job offers.
- Agences Caisse d'Epargne**: A section for agency information.
- Agences Banque Populaire**: A section for agency information.
- Evolution abonnés réseaux sociaux**: A section showing the evolution of social media subscribers.
- Abonnés sur Facebook**: A section for Facebook subscribers.
- Abonnés sur Instagram**: A section for Instagram subscribers.
- Abonnés sur LinkedIn**: A section for LinkedIn subscribers.

Each dataset card includes details like 'Producteur', 'Licence', and 'Référence', along with interactive options like 'Tableau', 'Carte', 'Analyse', 'Export', and 'API'.

<http://bpce.opendatasoft.com>

# Notre vision de l'Open Data

## Les sources Open Data utilisées pour l'enrichissement des données à l'adresse

Les rapprochements à l'adresse avec les données en Open Data permettent notamment de :

- Compléter les caractéristiques des logements assurés
- Evaluer le risque climatique physique du logement
- Apporter d'autres informations utiles pour certaines garanties (ex. Vol)



Sur la version 2018 :

- 24 millions de bâtiments
- Géométries au format shapefile
- Sources des bâtiments : cadastre et vue aérienne
- Caractéristiques : année de construction, nombre d'étages, nombre de logements...



Historique de 2013 à 2021:

- 9,9 millions de lignes
- Caractéristiques des logements : Superficie, année de construction



Historique de 2016 à 2021:

- 16 millions de lignes
- Caractéristiques des logements : Superficie, année de construction, nombre de pièces



Liste des logements sociaux en 2019:

- 4,6 millions d'adresses
- Caractéristiques des logements : Superficie, année de construction, nombre de pièces



**Zones inondables** **RGA** **Inondations par remontée de nappe**



**Indices INSEE** **Crimes et délits**



**Listes des commissariats** **Listes des gendarmeries** **Route 500**

*Panorama des données Open Data utilisées pour faire des rapprochements à l'adresse*

# Géocodage d'un portefeuille

La réalisation technique du géocodage et sa valeur



## Réalisation technique

Mise en place d'une surcouche du package de géolocalisation adock :

- Exploitation du package adock mis à disposition sur data.gouv au travers de l'API Adresse de data.gouv
- Différents niveaux de précision pour le géocodage (Home, Street, Locality, City, Post Code)
- Option de pseudo-anonymisation consistant à mélanger les adresses à géocoder avec des adresses prises aléatoirement

## Bénéfices

- Traitement par batch
- Temps de traitement

RUE	CODE POSTAL	VILLE	INSEE	LATITUDE	LONGITUDE	GEOCODING PRECISION
88 avenue de France	75013	Paris	75113	48.8566	2.3522	Home
8 rue de la gare	94000	Créteil	94460	48.7904	2.4556	Home
7 avenue Spinoza	94200	Ivry-Sur-Seine	94041	48.8131	2.3882	Street

Fichier .csv en entrée et formaté

Complétion du fichier par géolocalisation

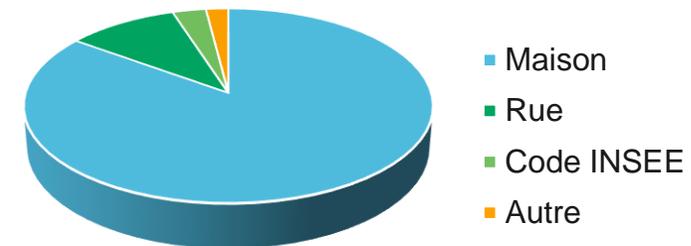


Illustration des précisions de géocodage sur un extrait d'adresses

# Enrichissement à l'adresse par l'Open Data

Démonstration sur la ville de Saint-Jean-de-Luz

Control Panel

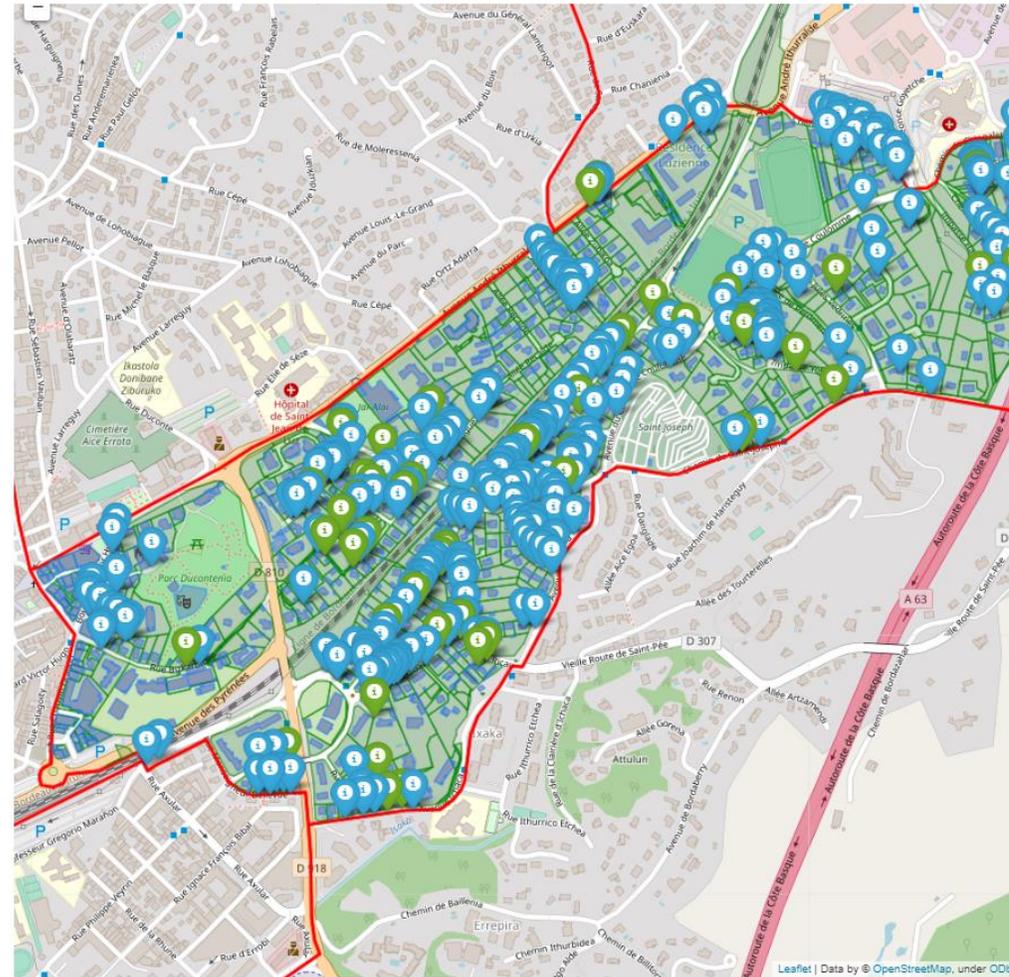
Affichage de zonage

- Afficher les agences CE/BP
- Afficher la base RGA
- Afficher les zones inondables

Affichage de bâtiments à l'IRIS

644830106

- Afficher les adresses BAN
- Afficher les adresses BAN appartements
- Afficher les parcelles cadastrales
- Afficher les batiments BD TOPO

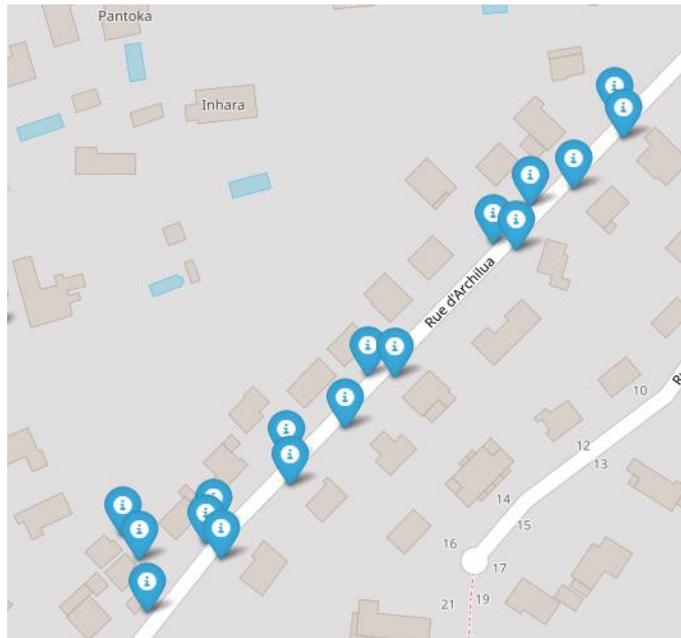


# Exploitation des données géospatiales

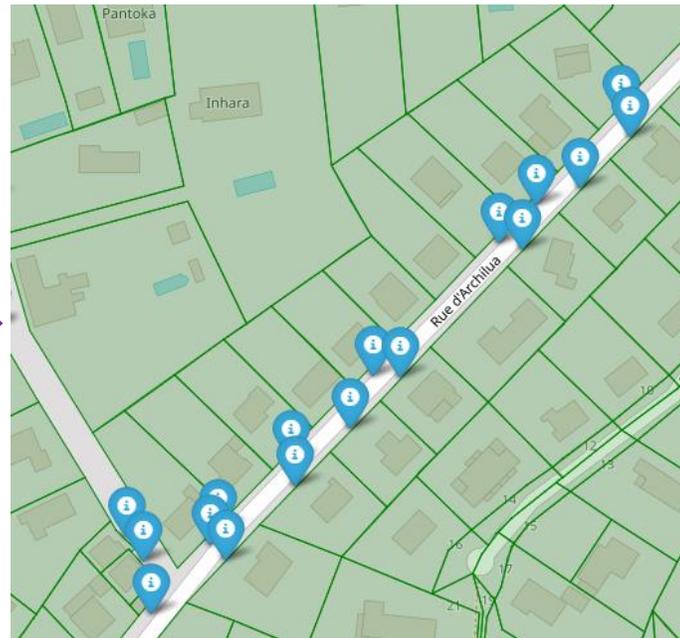
## Comment effectuer des rapprochements pour enrichir les adresses d'un portefeuille

Le rapprochement à l'adresse des données Open Data se fait au travers du géocodage du bien en trois étapes.

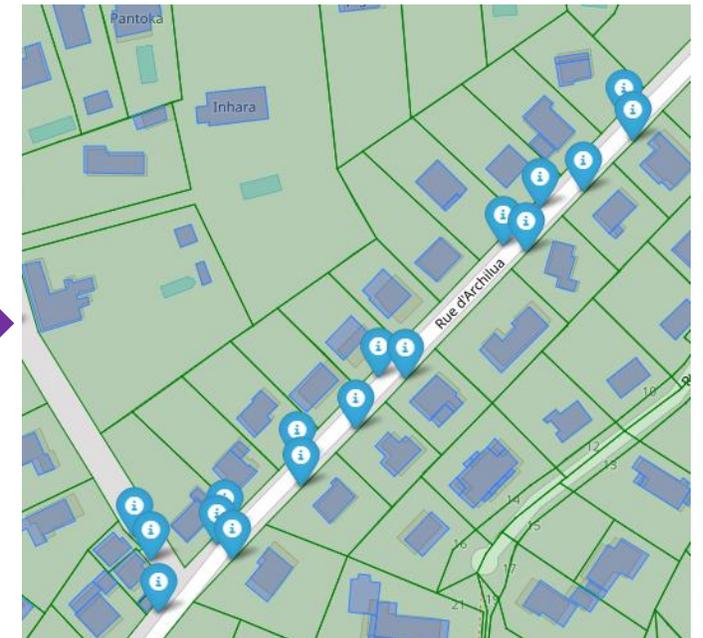
Pour l'exemple du rapprochement à la BD TOPO (pour avoir les caractéristiques du bâti à l'adresse) :



1- On géocode chaque adresse dans le portefeuille



2- On rapproche chaque point de géolocalisation de l'adresse à la parcelle cadastrale correspondante



3- On rapproche le bâti (issu de BD TOPO) de la parcelle cadastrale



Ces étapes sont très chronophages. Pour le cas particulier des rapprochements BD TOPO/Parcelles cadastrales, ces rapprochements sont rendus disponibles dans la base de données BAN PLUS

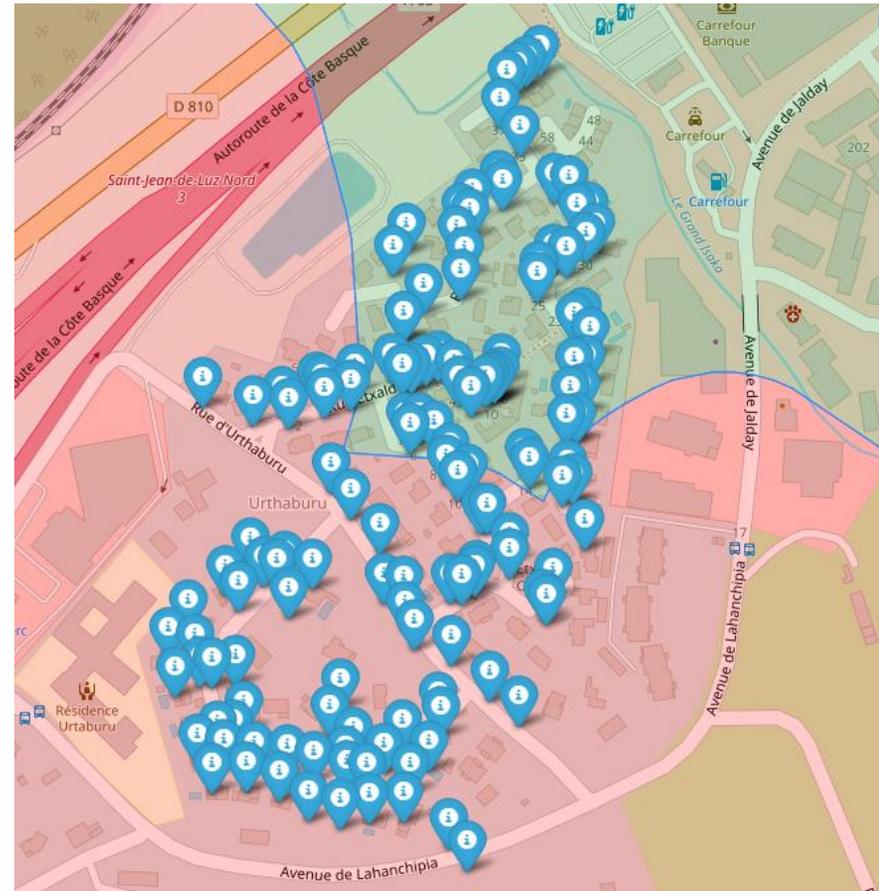
# Exploitation des données géospatiales

Comment effectuer des rapprochements pour évaluer l'exposition à certains aléas

## Rapprochement à l'adresse



De la même façon, les adresses (ensemble de points) peuvent être jointes à des objets de type polygones pour évaluer l'exposition de chaque adresse à l'aléa



**Rapprochement des adresses aux bases RGA**

# Exploitation des données géospatiales

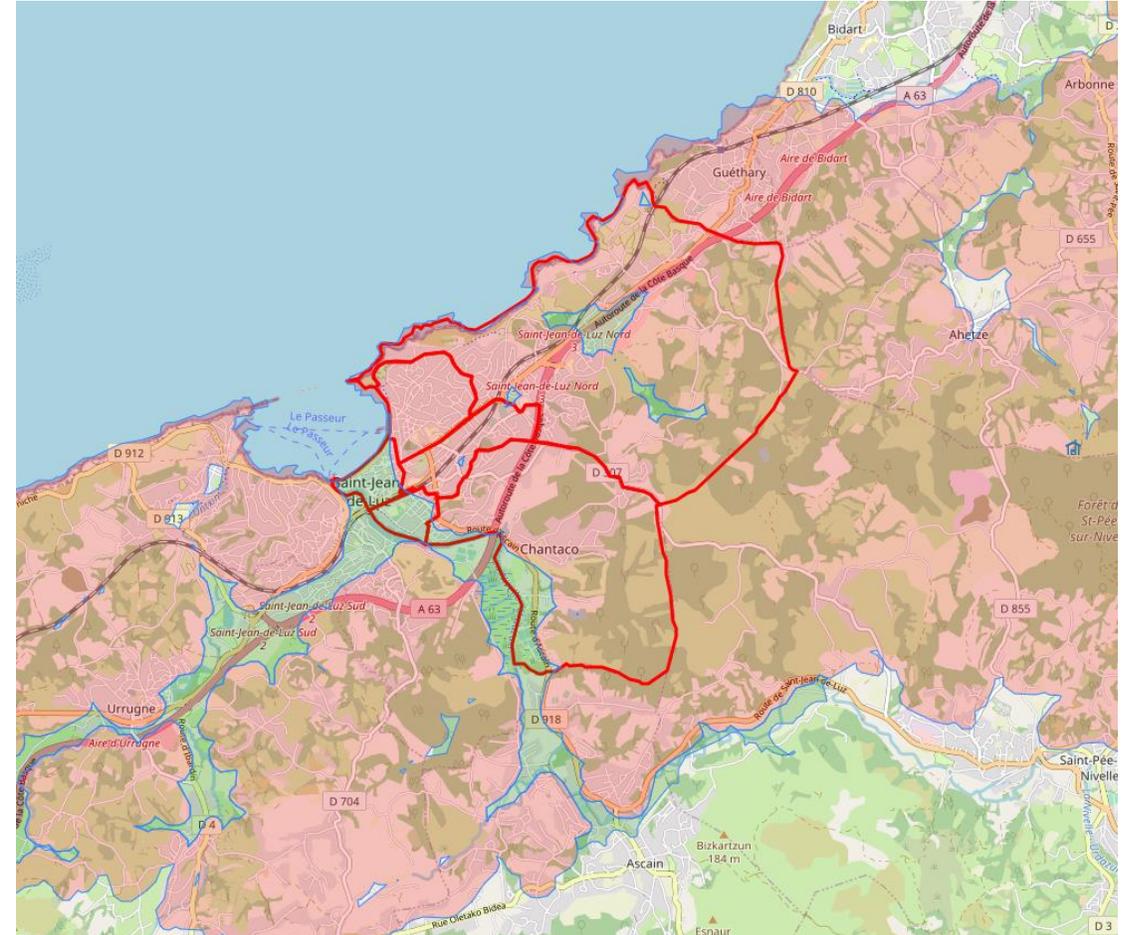
Comment effectuer des rapprochements pour évaluer l'exposition à certains aléas

## Rapprochement à la commune

Les données Open Data disponibles en format shapefile ou geoJSON permettent des jointures facilement réalisables sous Python avec geopandas.

Nous pouvons notamment dans le cadre de la sécheresse :

- Calculer les jointures entre les contours des communes et la base RGA
- Définir ainsi pour chaque commune son taux d'exposition aux différents niveaux d'argile (nécessaire pour le calcul de l'éligibilité des communes)



*Rapprochement de la commune Saint-Jean-de-Luz à la base RGA*



# Le contexte climatique français

# Le contexte climatique français

Comment les risques climatiques sont-ils assurés en France ?

- Ci-dessous, une **vision simplifiée** de l'assurance des risques climatiques :



Généralement **couvert** par les contrats d'assurance

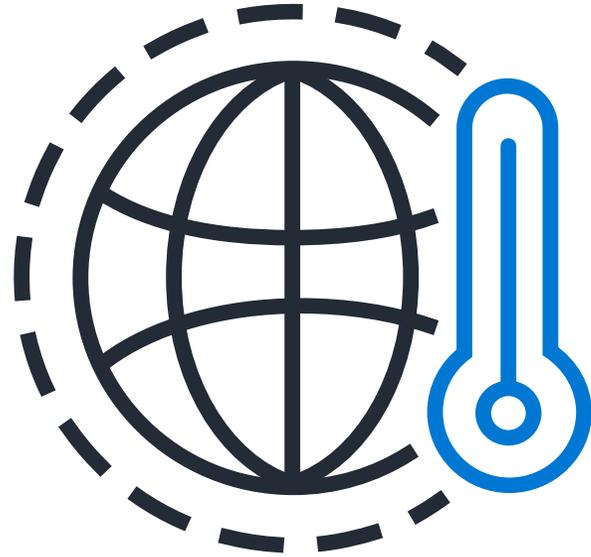
Généralement **pas couvert** par les contrats d'assurance à moins d'une reconnaissance CAT NAT

**Couvert** par les contrats d'assurance **si un arrêté CAT NAT est publié.**

La CCR intervient alors via un traité de réassurance proportionnel à hauteur de 50% sans limite d'intervention avec le support financier de l'Etat.

(\*) Submersion marine

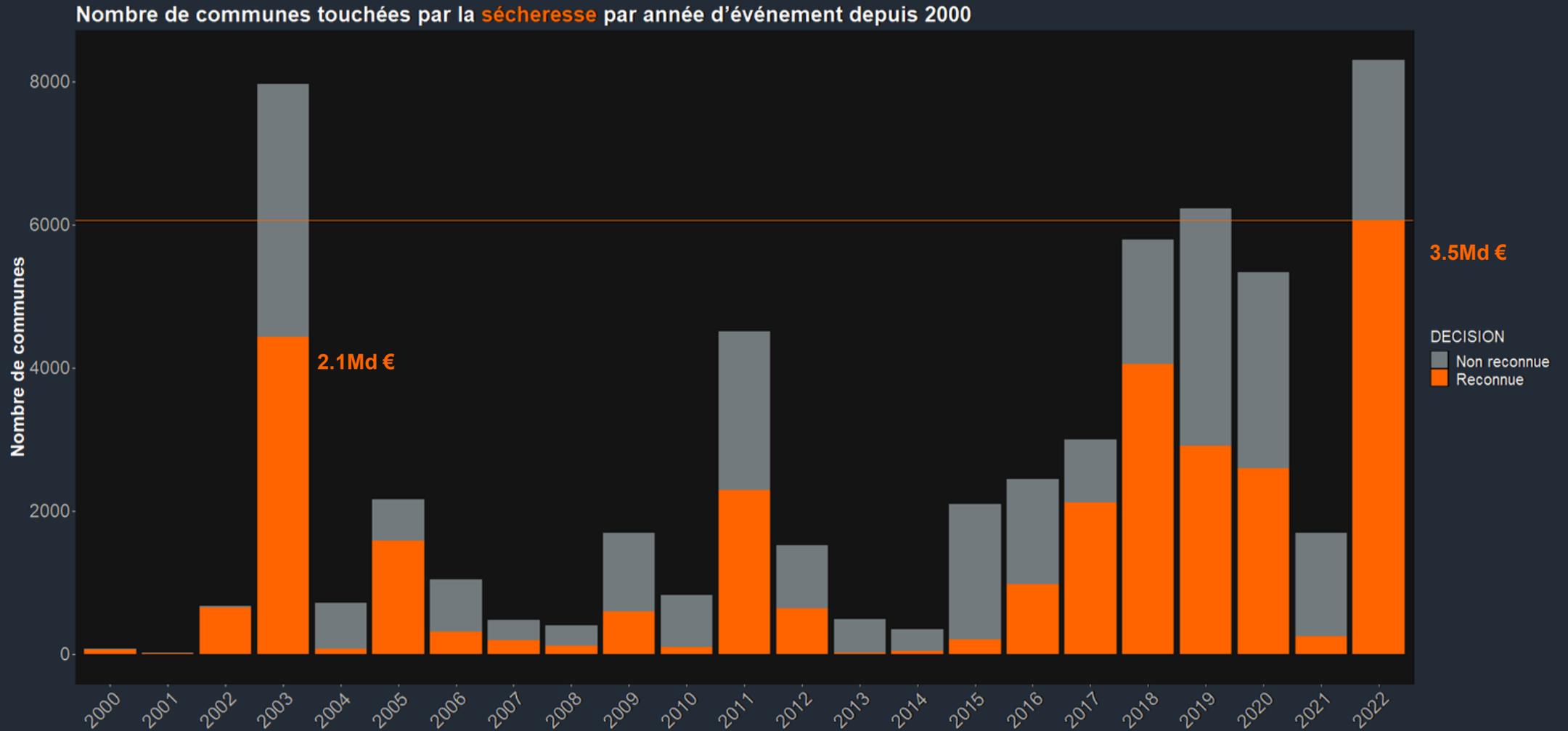
(\*\*) Hors de la France métropolitaine.



# Les spécificité de la sécheresse géotechnique

# Les spécificités de la sécheresse géotechnique

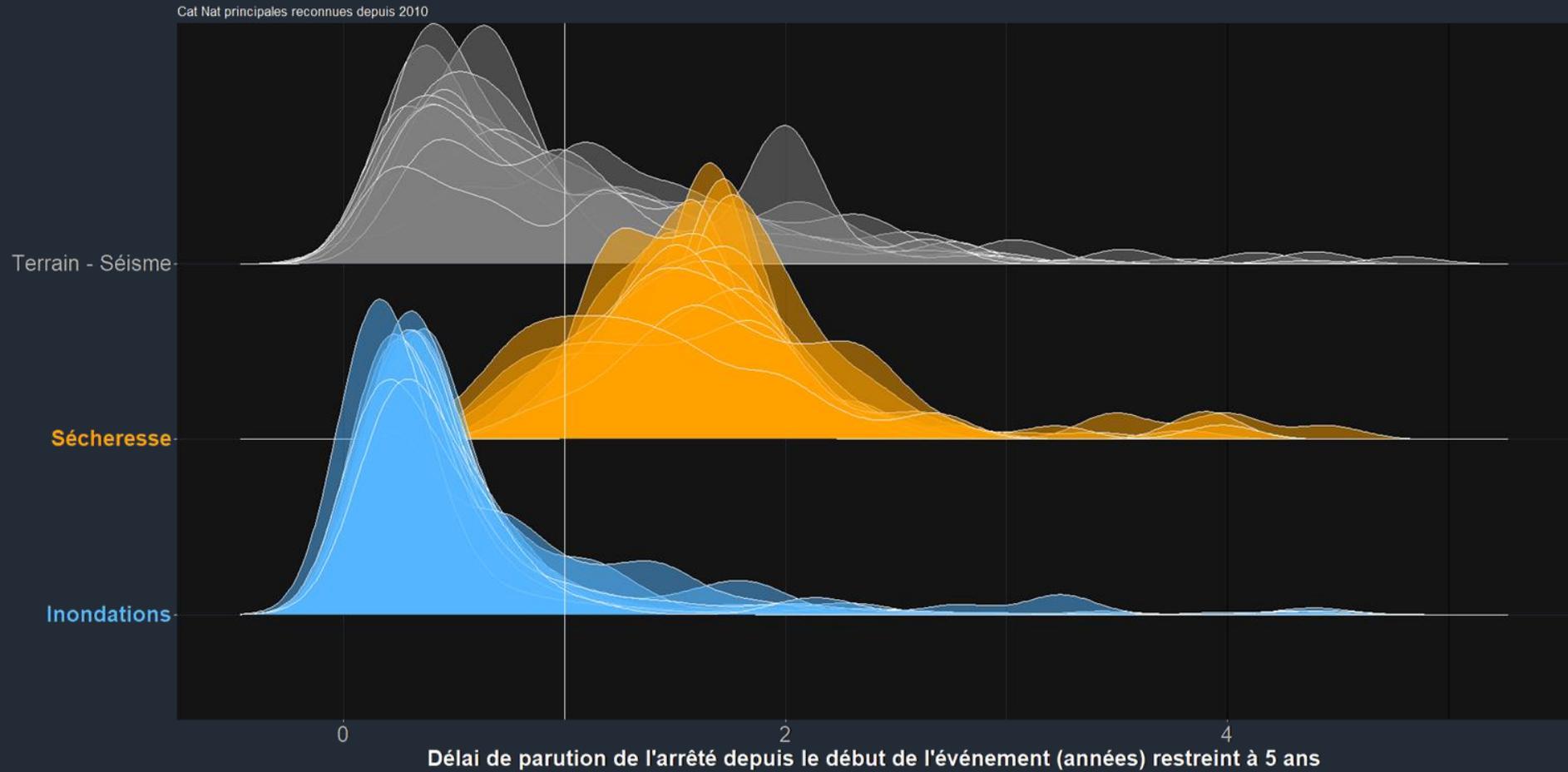
Un volume et coût croissant



Source des données : CCR, France métropolitaine, en date de septembre 2023

# Les spécificités de la sécheresse géotechnique

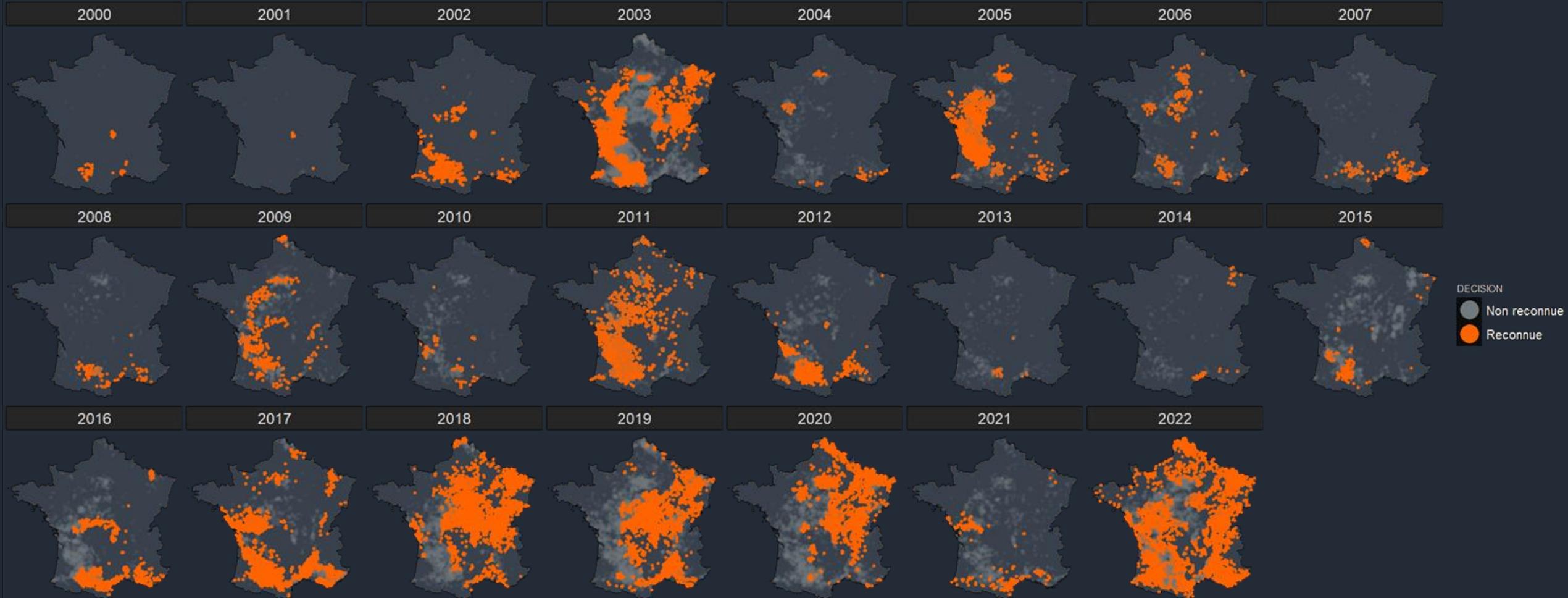
Une temporalité particulière



# Les spécificités de la sécheresse géotechnique

Le pays entier de plus en plus touché (1)

Communes ayant fait une demande de reconnaissance en Cat Nat **sécheresse** depuis 2000



Source des données : CCR, France métropolitaine

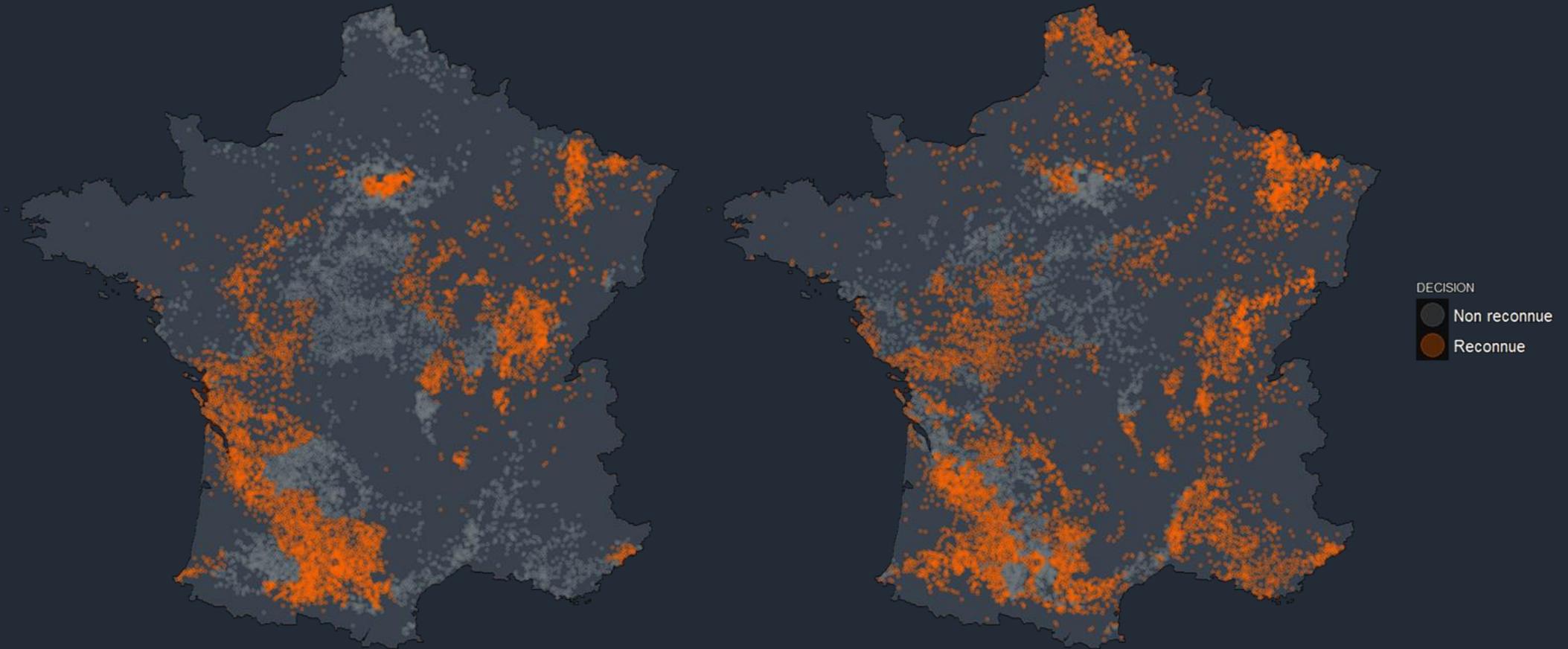
# Les spécificités de la sécheresse géotechnique

Le pays entier de plus en plus touché (2)

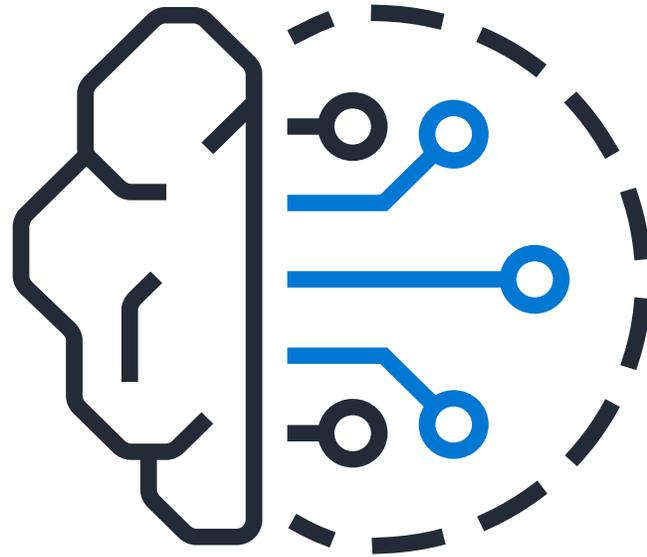
Communes ayant fait une demande de reconnaissance en Cat Nat **sécheresse** en 2003 et 2022

2003

2022



Source des données : CCR, France métropolitaine, en date de septembre 2023



# La modélisation du risque sécheresse

# La modélisation du risque sécheresse

Le processus de reconnaissance CAT NAT pour la sécheresse

## Demande de reconnaissance

Dans le cas de la survenance d'un événement naturel majeur sur une commune, la mairie peut déposer un dossier de reconnaissance CAT NAT au préfet.



## Enquête administrative

Enquête de la préfecture pour étudier la reconnaissance ou non de l'état de catastrophe naturelle.



## Décision de reconnaissance

Décision de reconnaissance CAT NAT par un conseil interministériel. La liste des demandes et reconnaissances est publiée au Journal Officiel.



## Compensation financière

Une compensation financière est versée par la CCR selon la décision de reconnaissance.



# La modélisation du risque sécheresse

Le processus de reconnaissance CAT NAT pour la sécheresse – Le critère géotechnique

- Critère géotechnique

“ La surface exposée au risque RGA doit être supérieure à 3% de la surface totale de la commune ”

- La cartographie correspondante est réalisée par des experts reconnus (BRGM)
- Ce critère permet d'identifier les communes étant **susceptible d'être touchées par la sécheresse RGA**
- Si la surface exposée au risque RGA est < 3%, la commune doit alors fournir une **étude complémentaire du sol**

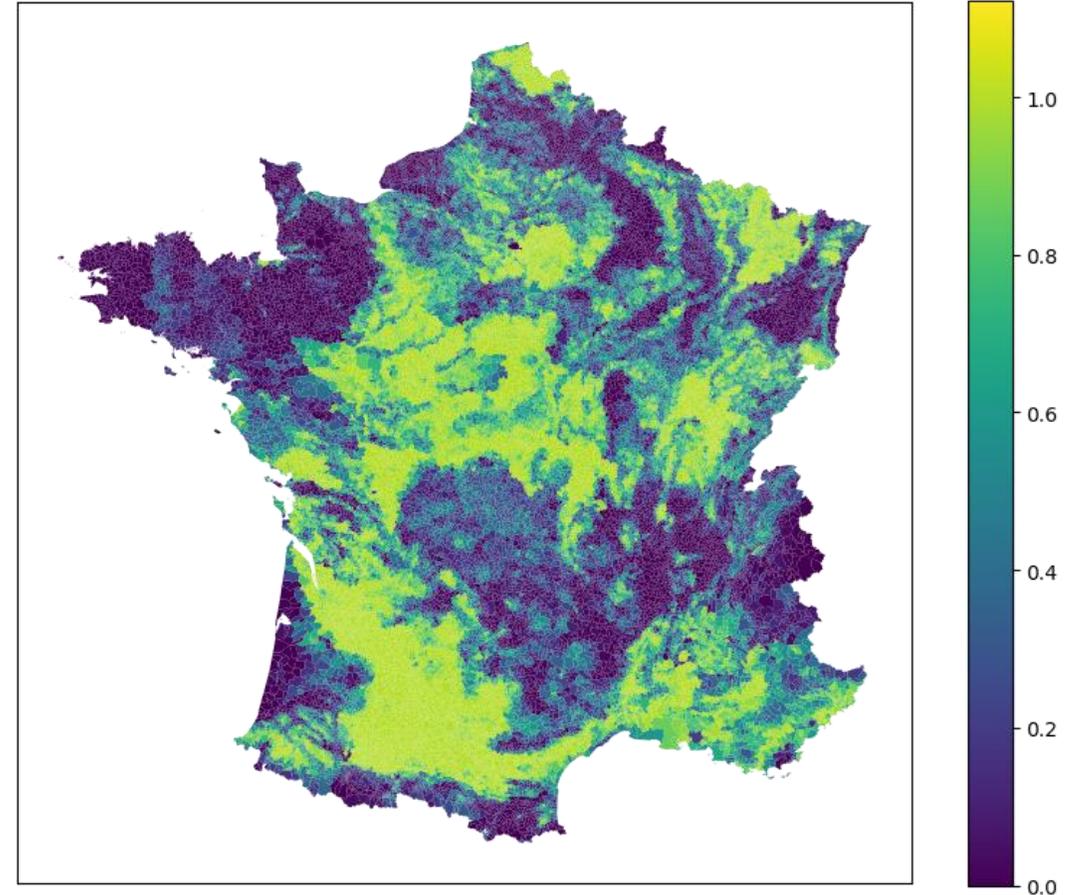
**48,5%**

Part du territoire exposée de manière moyenne ou forte au risque RGA

**29 100**/<sub>35 000</sub>

Nombre de communes éligibles pour le critère géotechnique

Map of RGA Exposure Rate (medium to high)



Visualization of the geotechnical criterion – France  
BRGM (2019) & internal modeling results - Milliman oct. 2023

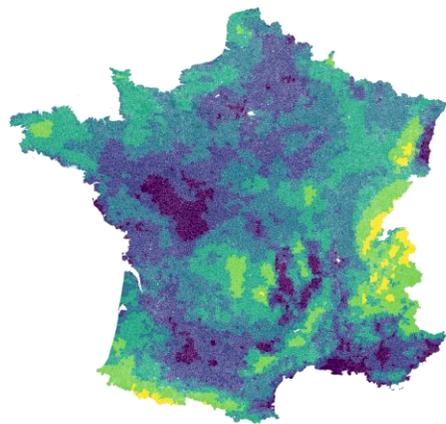
# La modélisation du risque sécheresse

Le processus de reconnaissance CAT NAT pour la sécheresse – Le critère météorologique

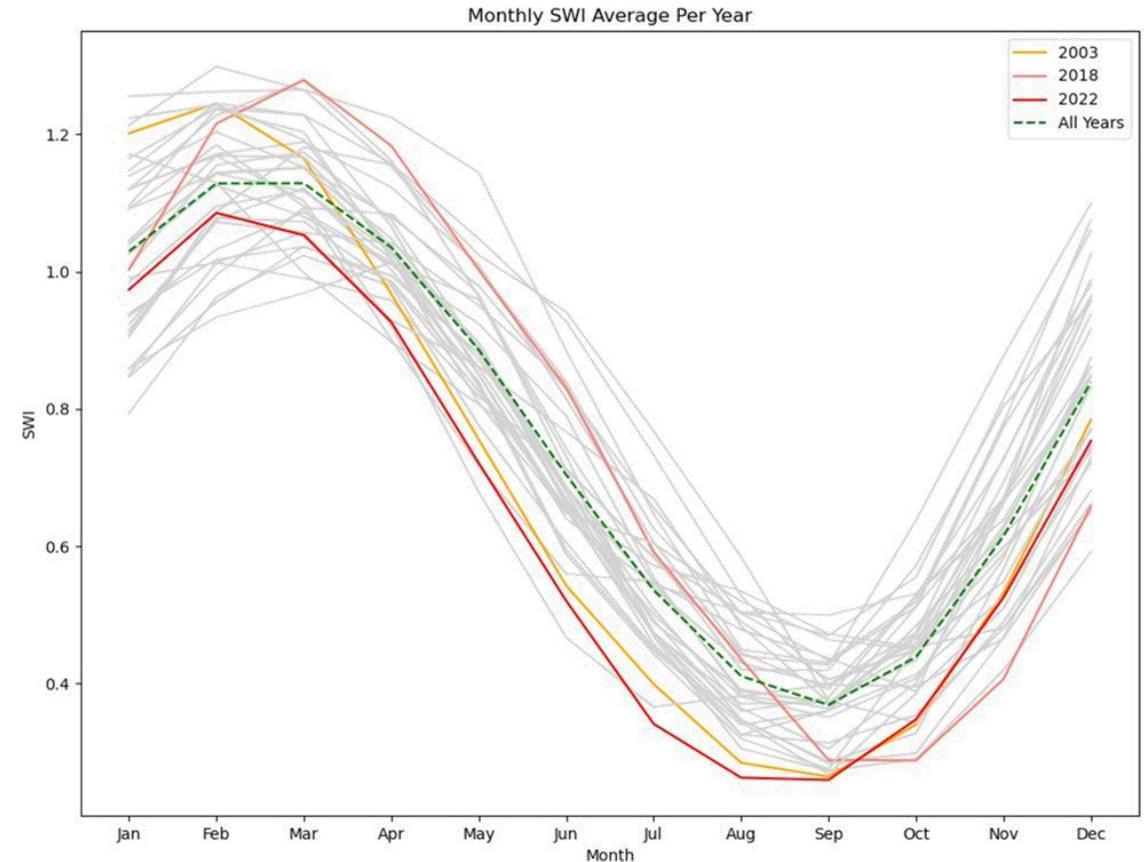
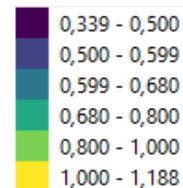
- Critère météorologique :

“ Le SWI (Soil Wetness Index) observé sur l'année doit faire partie des 2 valeurs les plus basses observées sur les 50 dernières années ”

- Le SWI est un **indicateur localisé** établi par **Météo France**
- Cette règle représente une **période de retour de 25 ans**
- Un SWI élevé ~ sol humide VS Un SWI bas ~ sol sec

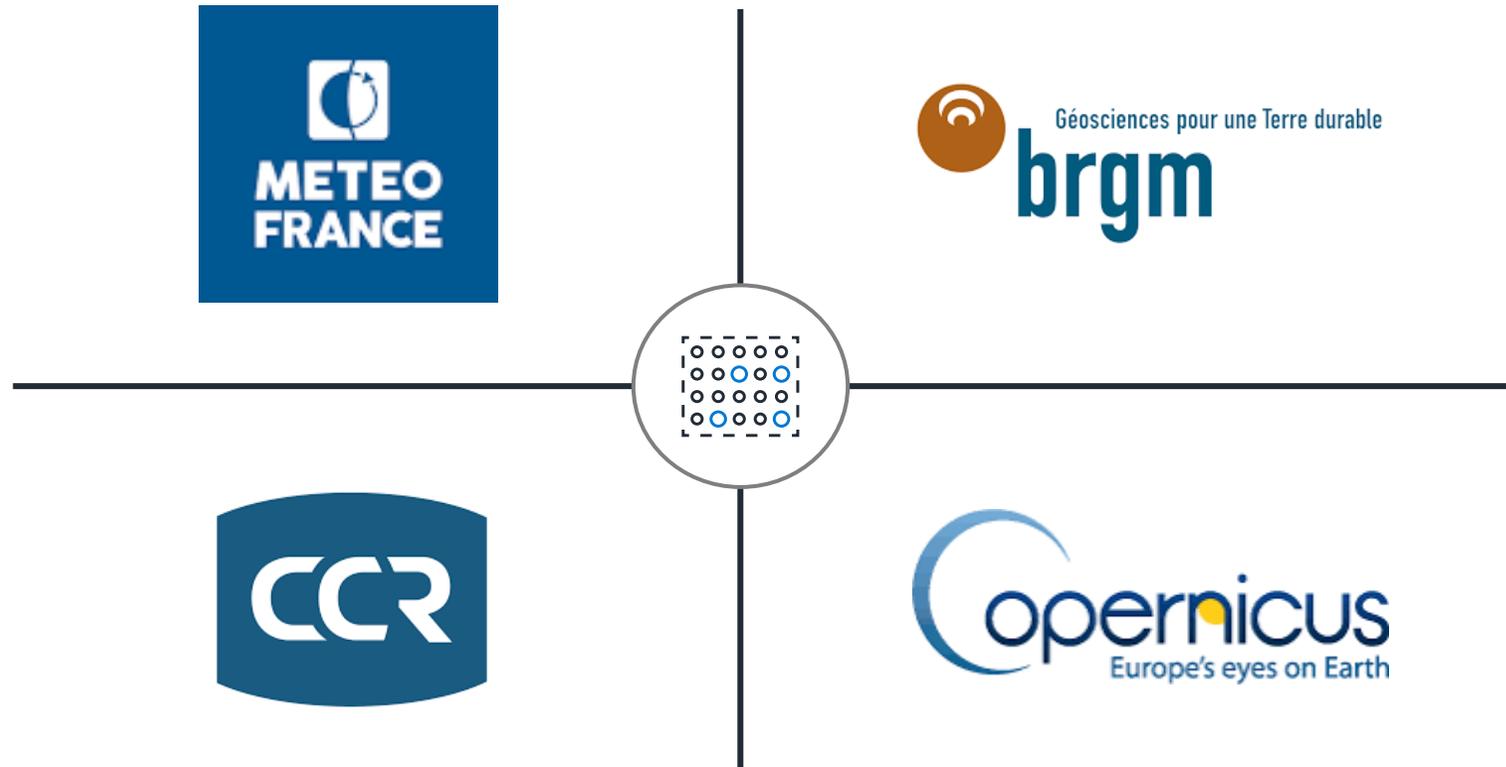


SWI - 2022  
IGN & Météo France - Milliman oct. 2023



# La modélisation du risque sécheresse

Les sources de données

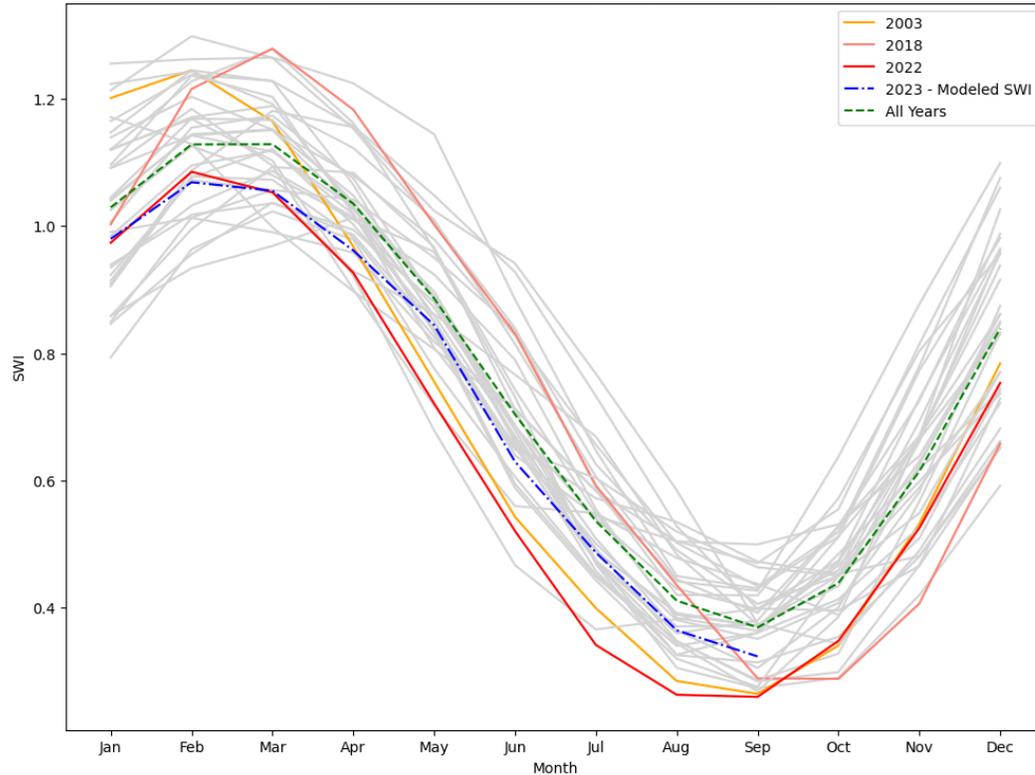


# La modélisation du risque sécheresse

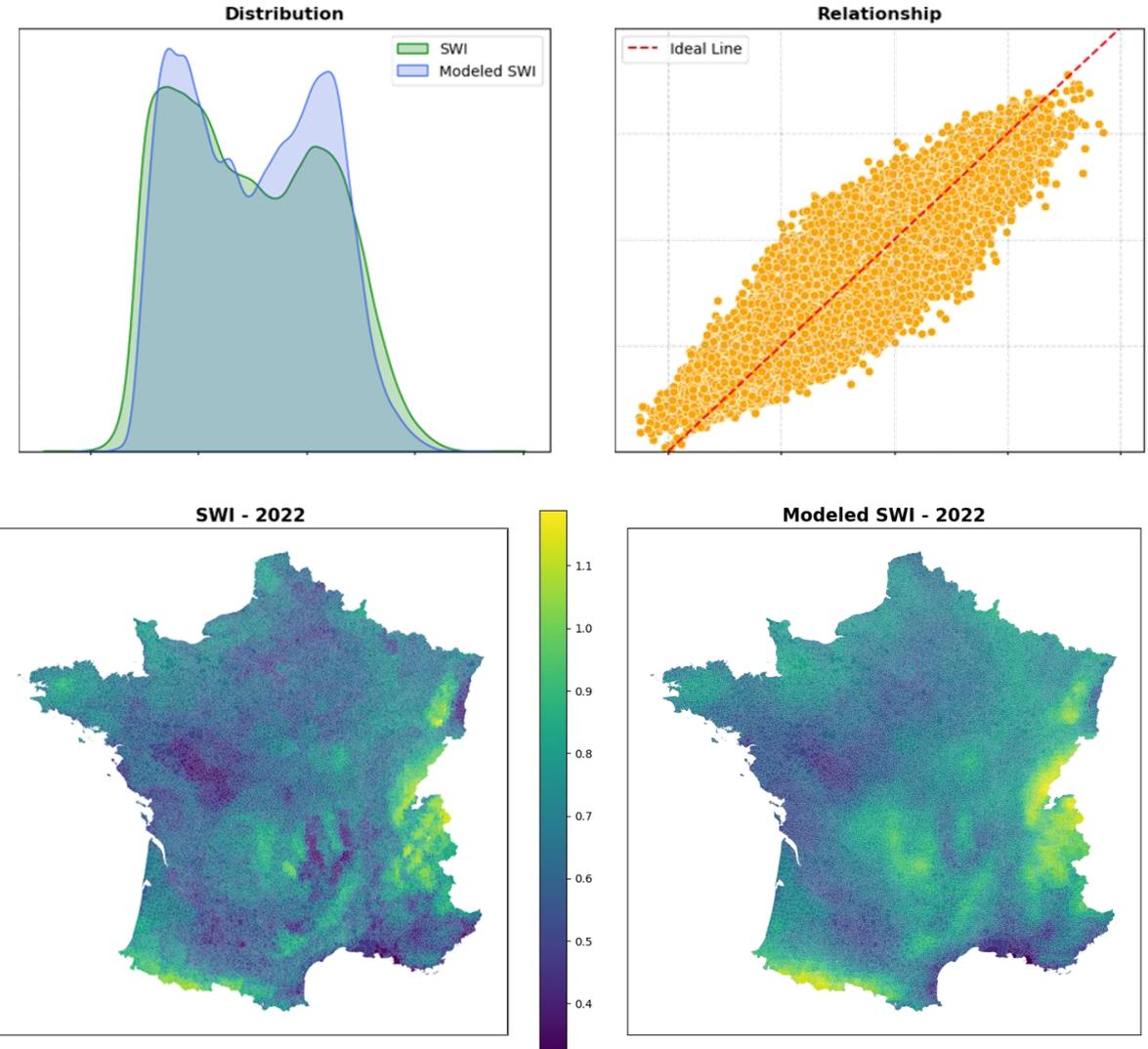
## Modélisation du SWI

- **Prédire les valeurs du SWI** est essentiel pour la modélisation du risque de sécheresse. Le SWI est calculé par **le modèle SIM de Météo France**, basé sur des **'données météorologiques** (température de l'air, précipitations cumulées, rayonnement, vent, etc.)

SWI Mensuel moyen par an



Modèle SWI - 2022



# La modélisation du risque sécheresse

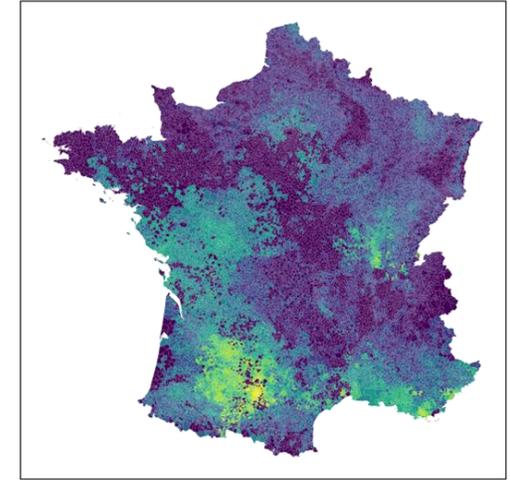
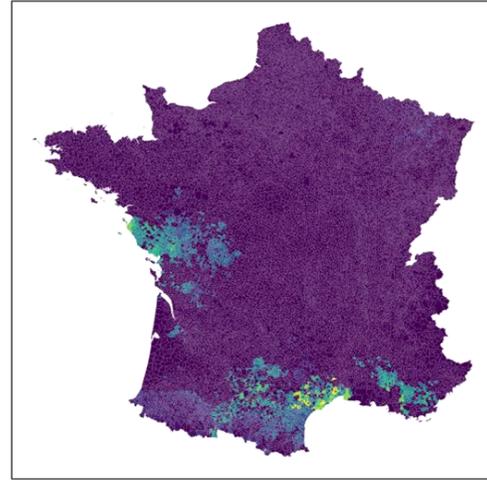
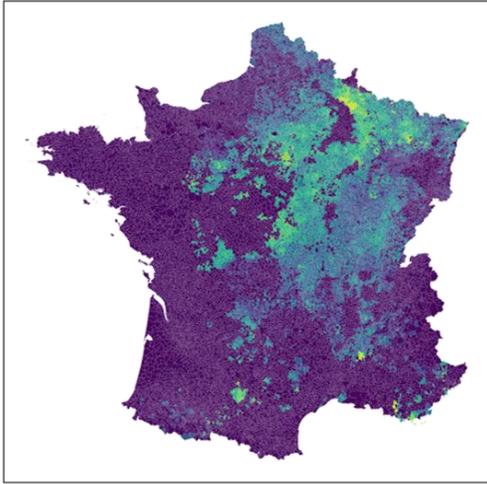
Carte des prédictions – 2020, 2021 et 2022

2020

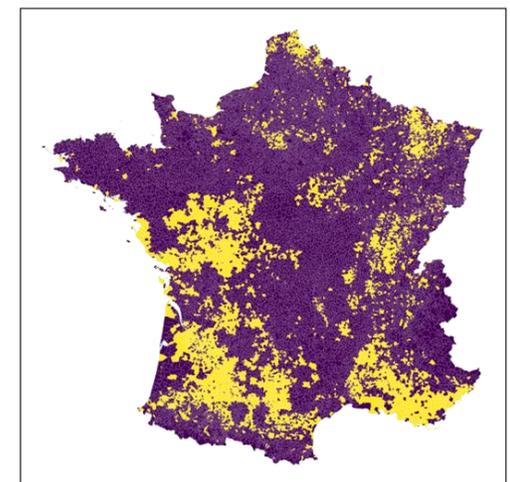
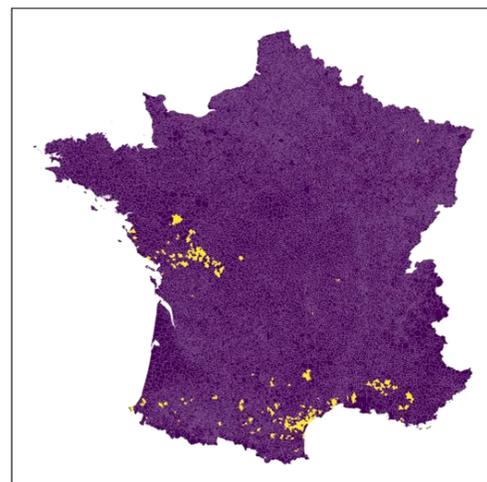
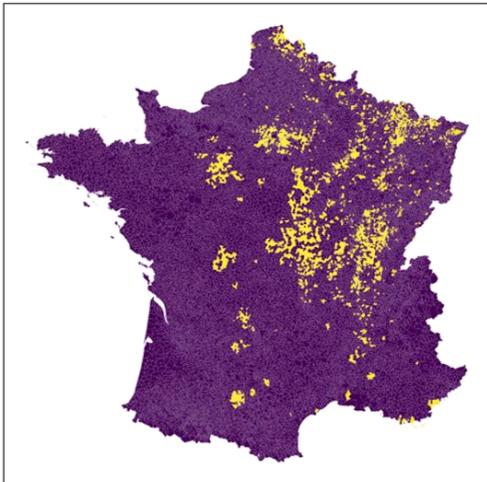
2021

2022

Prédictions



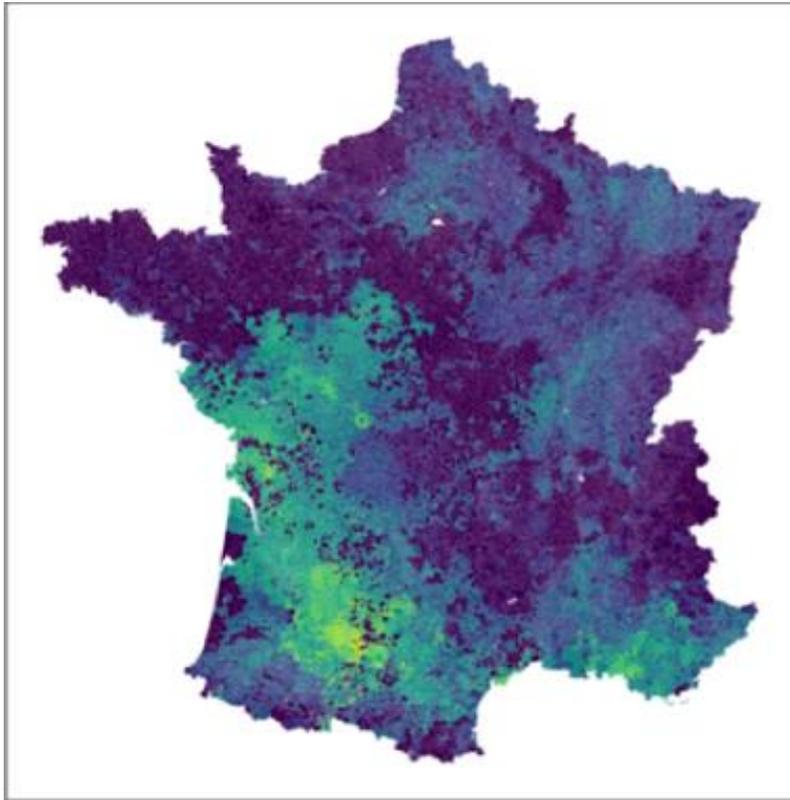
Target



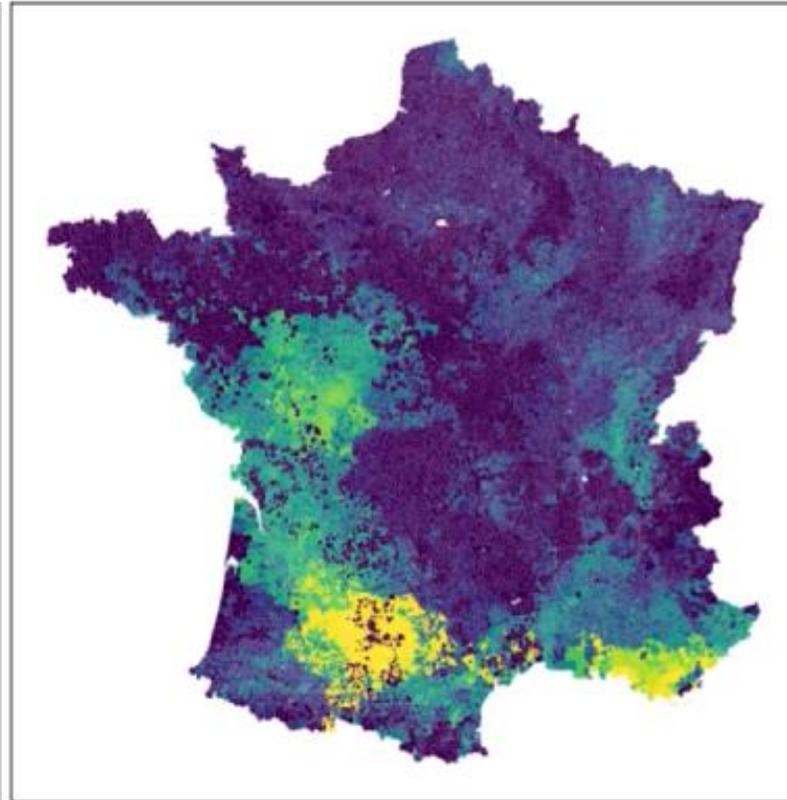
# La modélisation du risque sécheresse

Carte des prédictions au fil de l'année

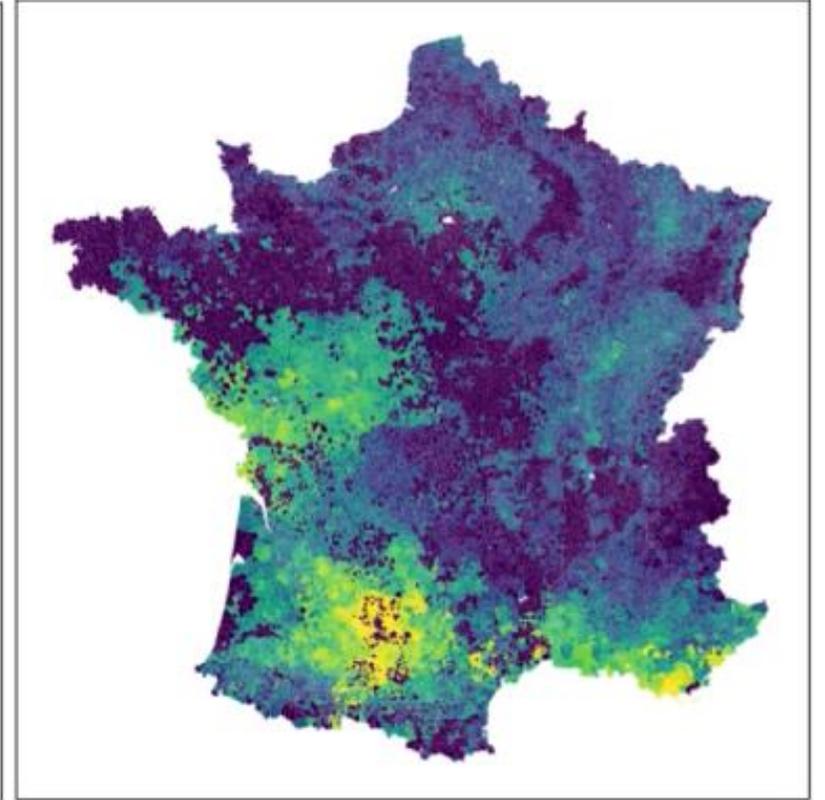
Janvier 2022



Juin 2022



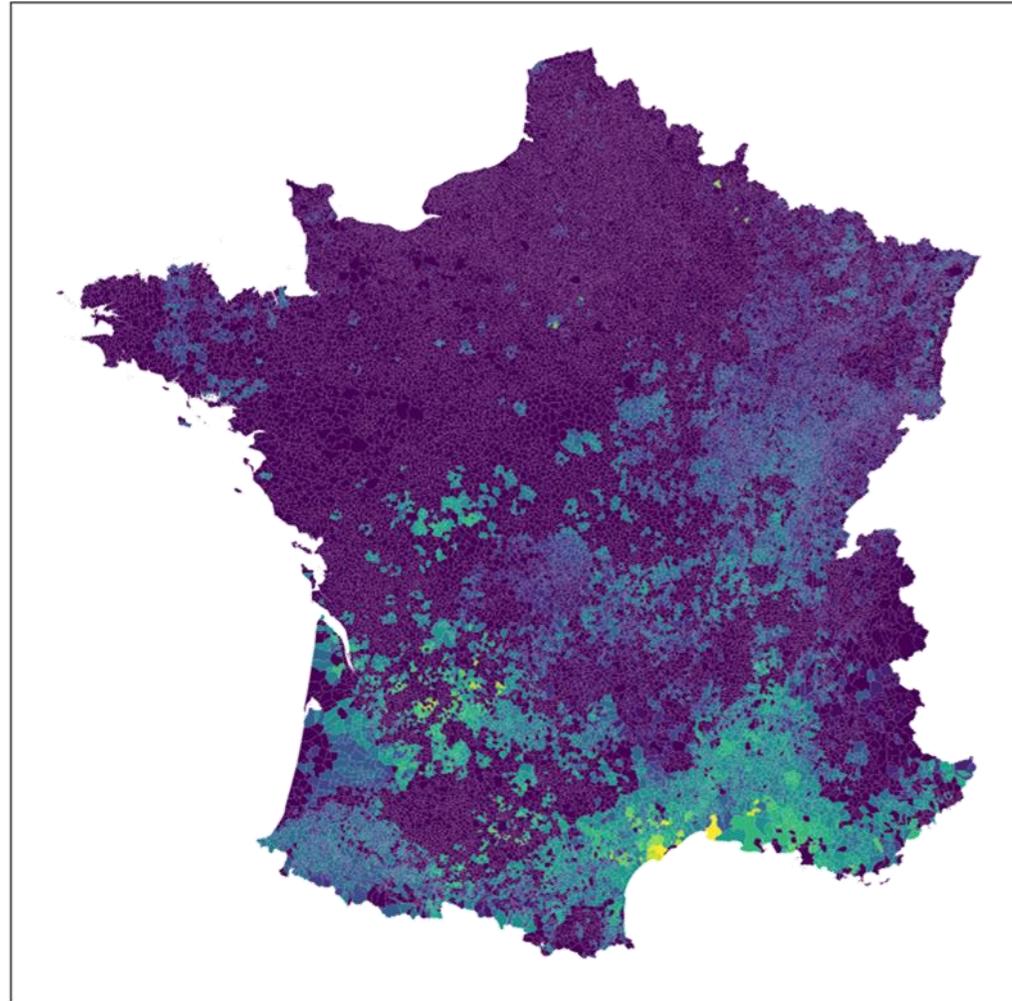
Septembre 2022



# La modélisation du risque sécheresse

Prédictions 2023

Septembre 2023





**Niels LUBBERS**

Data scientist  
**BPCE Assurances IARD**

[niels.lubbers@bpce.fr](mailto:niels.lubbers@bpce.fr)



**Anthony LAM**

Data scientist  
**BPCE Assurances IARD**

[anthony.lam@bpce.fr](mailto:anthony.lam@bpce.fr)



**Rémi BELLINA**

Chief Data Scientist & Senior Manager  
**Milliman France, Paris**

[remi.bellina@milliman.com](mailto:remi.bellina@milliman.com)



**Antoine RAINAUD**

Analyste risques climatiques & Consultant  
**Milliman France, Paris**

[antoine.rainaud@milliman.com](mailto:antoine.rainaud@milliman.com)



**GROUPE  
BPCE**