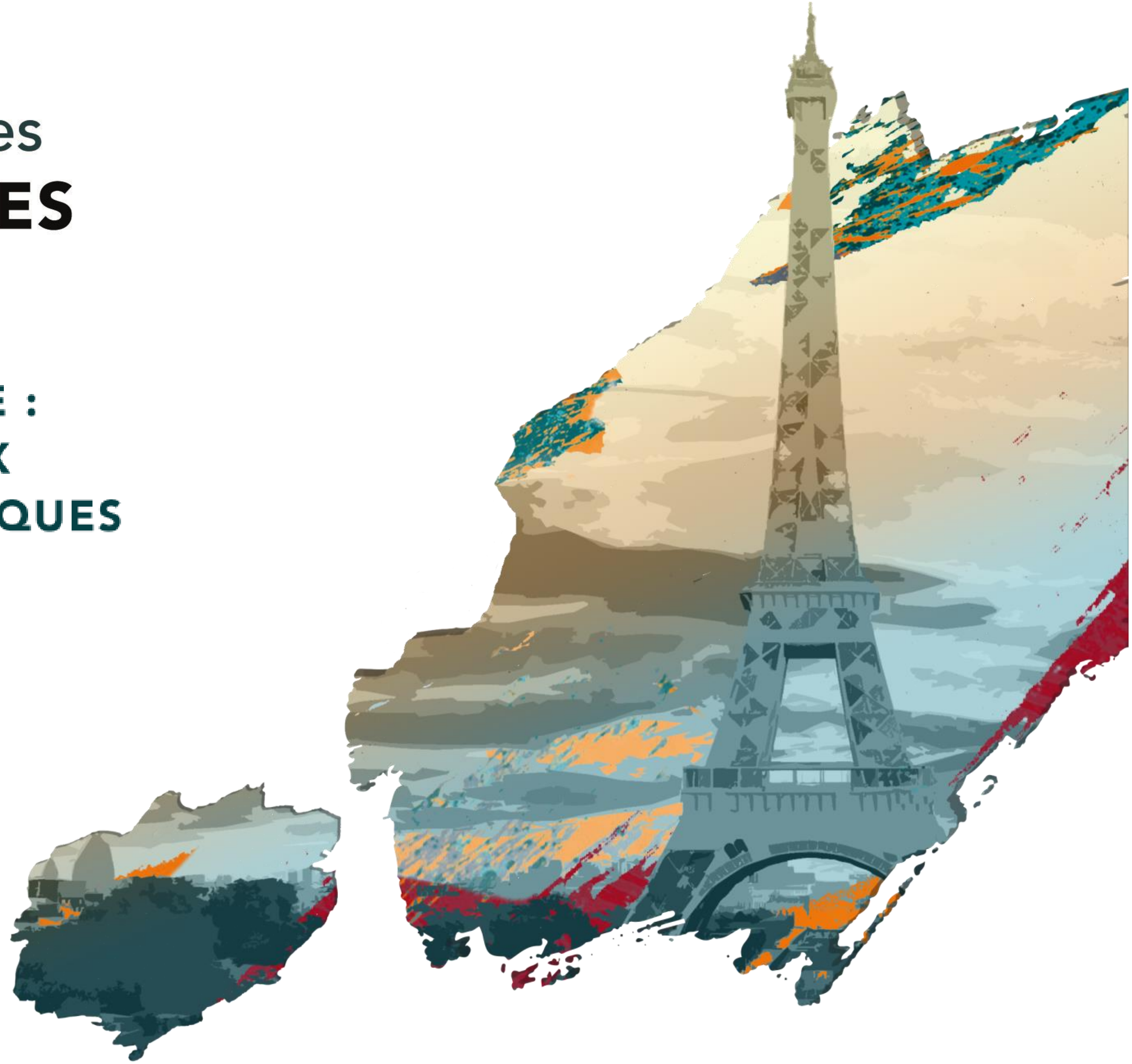
18e Congrès des ACTUAIRES

ASSURANCE ET FINANCE : VENT DEBOUT FACE AUX CHANGEMENTS CLIMATIQUES



INSTITUT DES ACTUAIRES

17 JUIN 2019

addactis

Les données innovantes comme solution aux transformations climatiques

Allianz

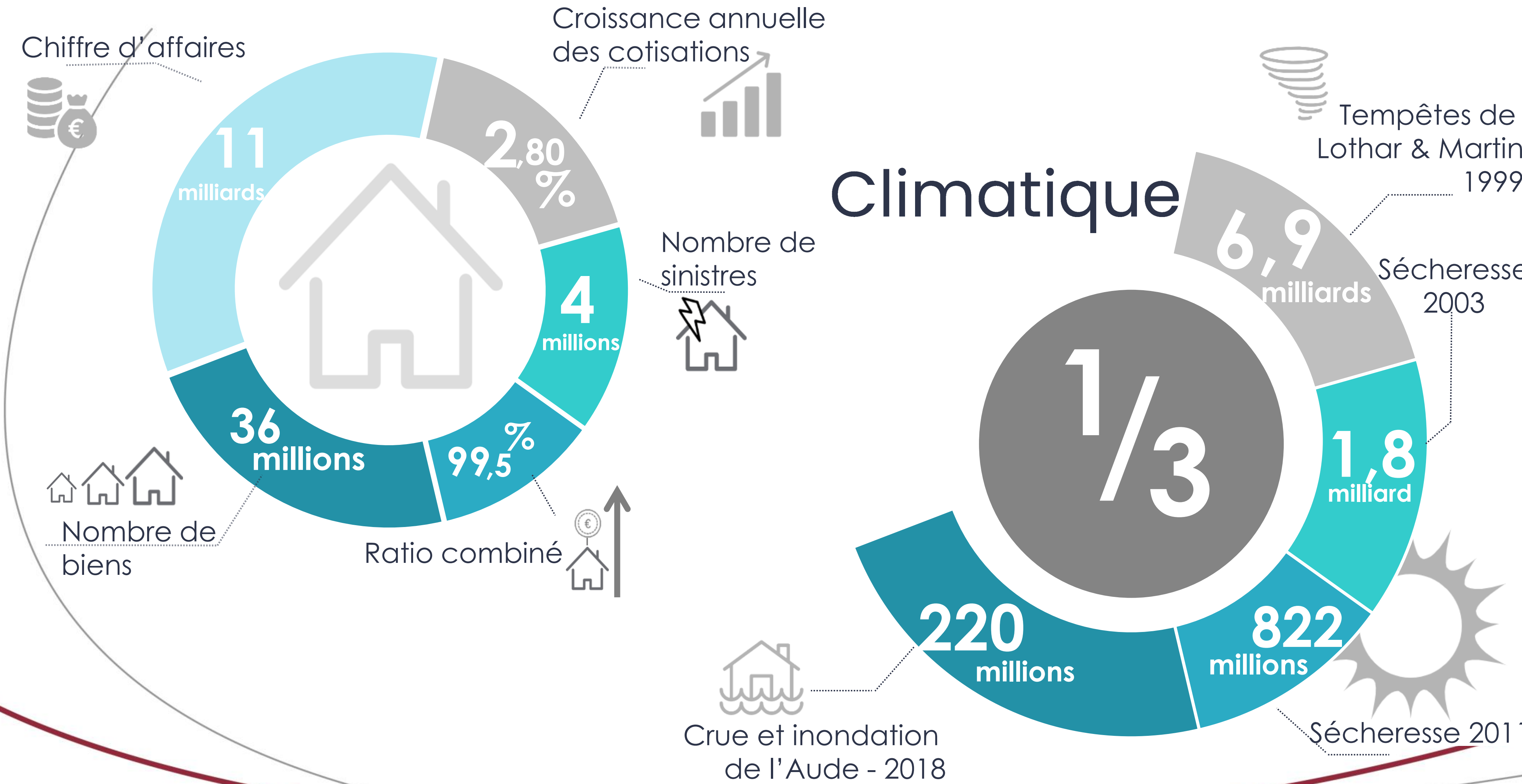
nam.R

MAITRISE DES RISQUES À L’ADRESSE

UNE UTOPIE ?


TARIFICATION MULTIRISQUE HABITATION
AUJOURD'HUI

LES ENJEUX



Éclairer les risques, tracer l'avenir

* Source: CCR

LES ENJEUX

Enjeu économique et financier pour les assureurs

Recherche de moyens pour **connaître**, **maîtriser** et mieux **anticiper** ses risques

Les moyens : Données innovantes sur les habitations

addactis Allianz nam.R

COMMENT ?

INSTITUT DES ACTUAIRES

01 COLLECTE
Smart Data
new

02 TRAITEMENT
Data Cleaning
Feature engineering
Kriging & machine learning

03 SELECTION
Malédiction de la dimension

04 MODELE
GLM
Régression quantile

05 VISUALISATION

ANALYSE DU RISQUE **CLIMATIQUE**

addactis Allianz namR

Big data, big frustrations

Lack of data

- Quality data
- Organized data
- Original data

Research institutes, administrations and business do not have the time and the ability to process, organize and produce original data.

namR

address these issues

Lack of tools

- Data collection
- Data processing & analysis
- Predictive

Research institutes, administrations and business do not have the time, the financials and the technical capacities to develop serious data analytics tools

Overwhelmed administrations

- Collect massive amounts of data but ...
- Don't know how to process and organize their data
- Don't understand their data and their usefulness

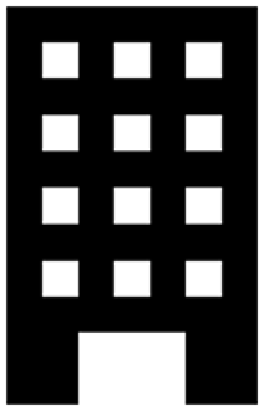
We produce actionable data using all accessible data

- nam.R creates its own **specific** and **proprietary tools**, to harvest massive amount of data, their integration to a **unified referential**. The harvested data is then geolocated, linked with other relevant data and constantly enriched with machine learning algorithms.

- We use **non-personal data**, from **imagery** (satellite and aerial), **text** (web, ads, address..) as well as **geolocated structured informations** (cadastre, urban plans…)

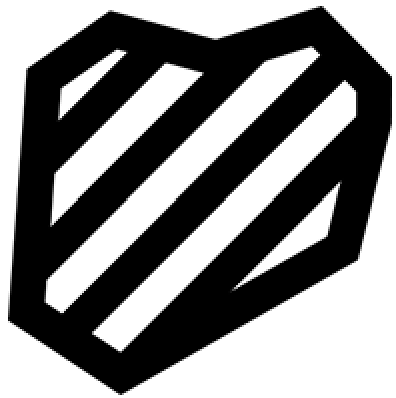
- We **produce new original** data on geolocated entities (buildings, parcels, cities…) to qualify a territory, an asset or an activity.

3 main entities



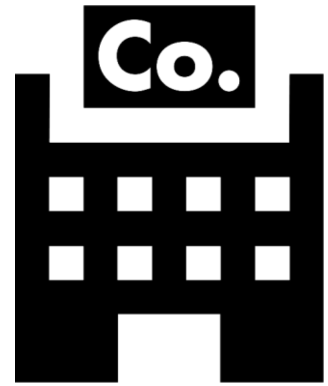
Buildings

Referential of the 34 millions of France's buildings



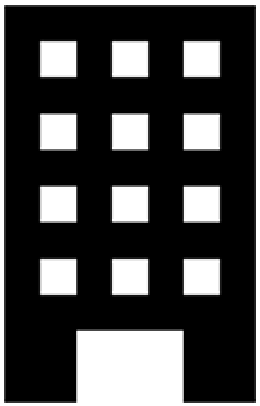
Parcels

Referential of the 88 millions of France's buildings



Companies

Referential of the 10 millions of France's companies



Buildings

Referential of the 34 millions of France's buildings

Morphology

Building morphology at the address

- Shape and footprint
- Roof shape & material
- Construction period

Equipment & Energy

Description of building's equipment and corresponding parcel(s) & information about energetic category

- Heating fuel information
- Glass surface
- Elevator presence

Meteorology

Information about weather at the address

- Wind
- Temperature
- Rain

Surroundings

Informations about neighborhood and close services

- Public services distance
- Number of tree
- Closest waterway

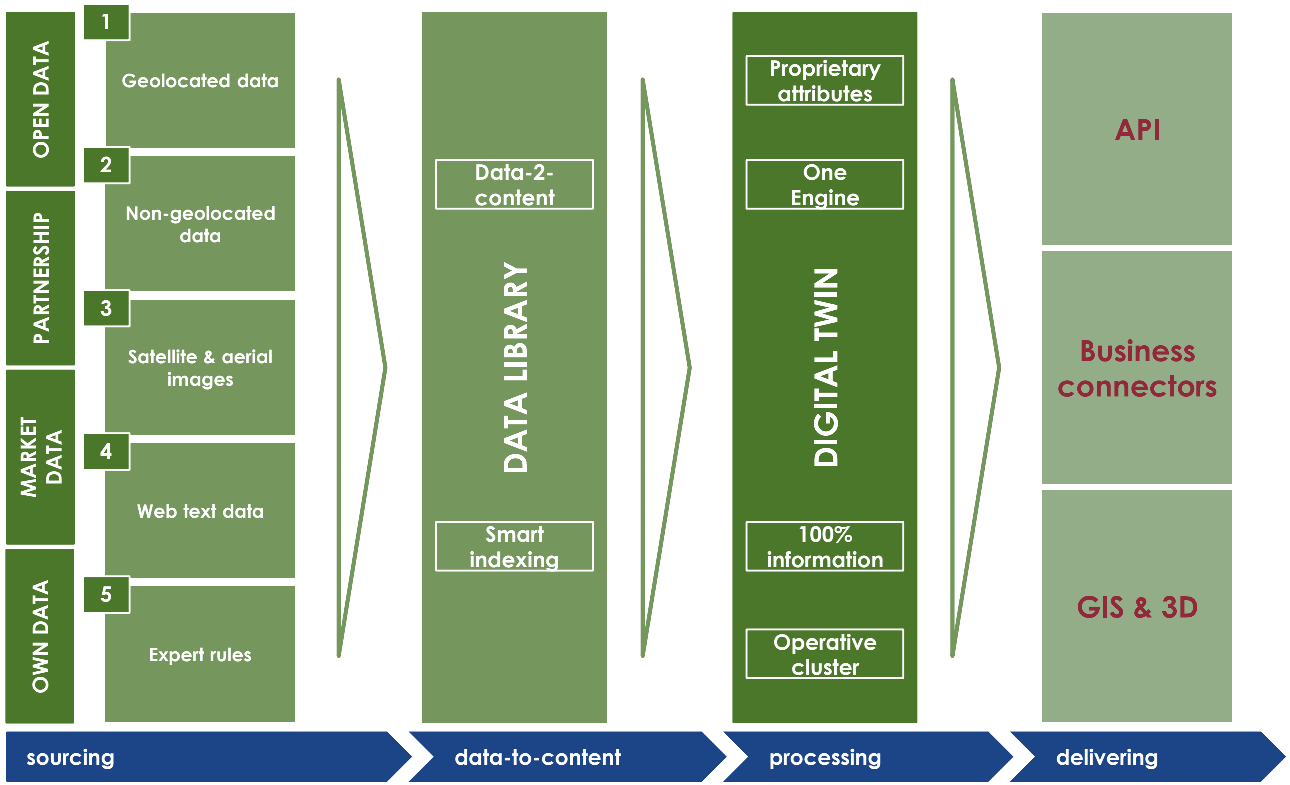
Technical presentation

I. The nam.R core process

I. Geocoding

I. Computer vision

The nam.R core process



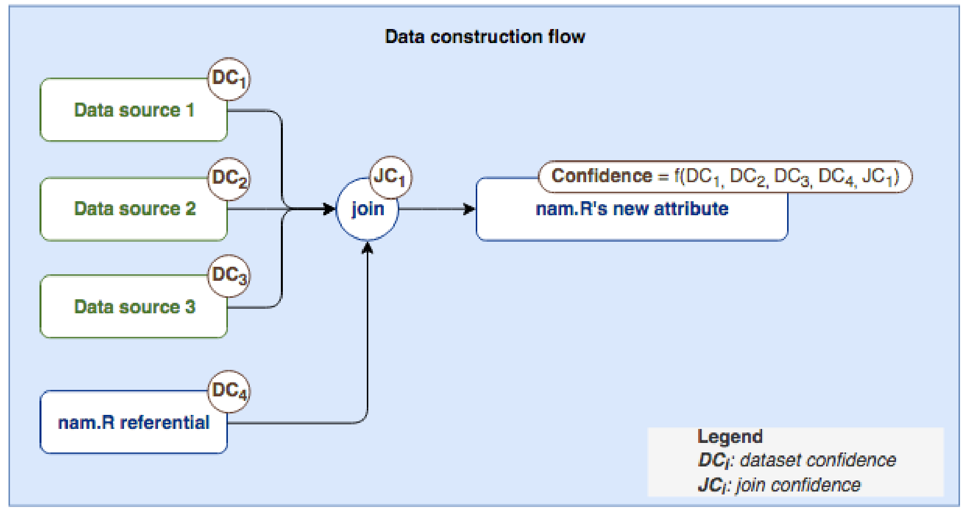
The nam.R core process - confidence

One of most central process in nam.R technical pipe is the evaluation of confidence level. None of nam.R data is going out of our system without being associated with a confidence level which is permanently evaluated and monitored.

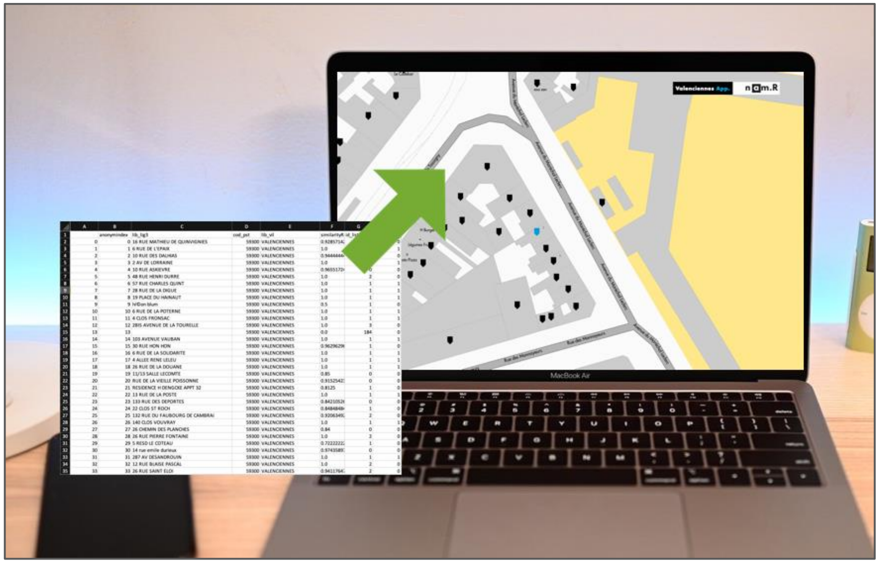
We monitor all the potential uncertainty, for example :

- confidence in data sources
- **confidence in process and techniques**
- algorithmic scores

Therefore, every information in nam.R database is attached with the most conservative combination of confidence level since we keep the minimum of confidence level.



Geocoding : link an address to a building

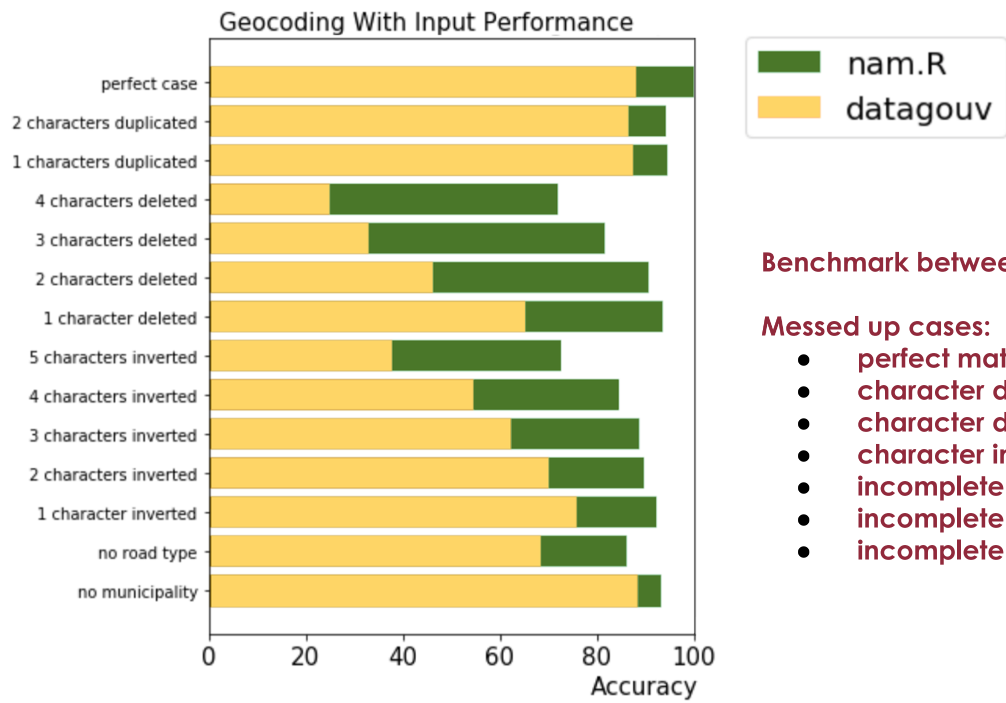


modus operandi

- Create our address referential
- String match : convert “dirty address” to a clean address
- Address - Building link

Geocoding : link an address to a building

geocod.R benchmark



Benchmark between geocod.R and BAN's API adresse.data.gouv.fr

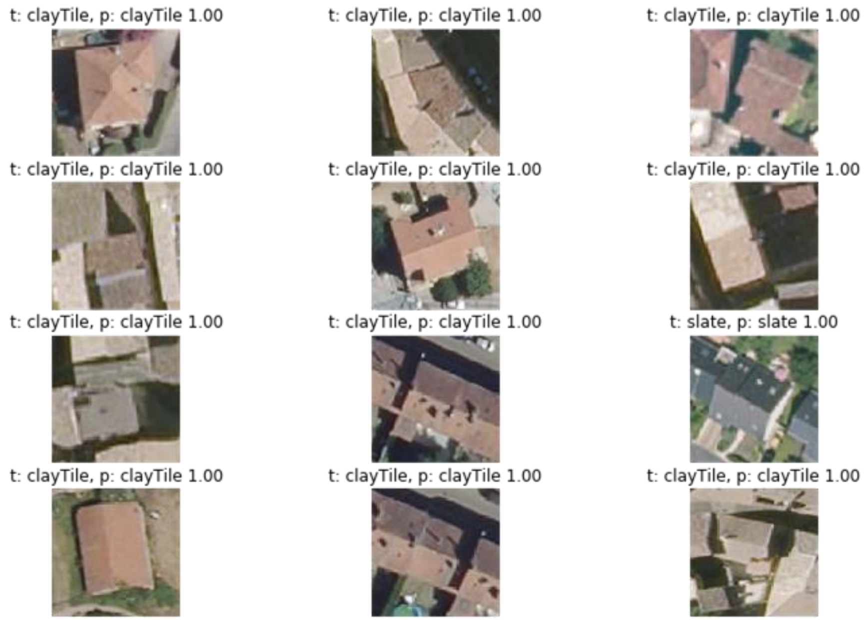
Messed up cases:

- **perfect match** "4 rue Foucault 75016 Paris"
- **character deletion** "4 rue Fouault 75016 Paris"
- **character duplication** "4 rue Fouucault 75016 Paris"
- **character inversion** "4 rue Foucault 75016 Prais"
- **incomplete address** "4 rue Foucault Paris"
- **incomplete address** "4 Foucault 75016 Paris"
- **incomplete address** "4 rue Foucault 75016"

Some example of computer vision

Automatic Detection of Roof Material

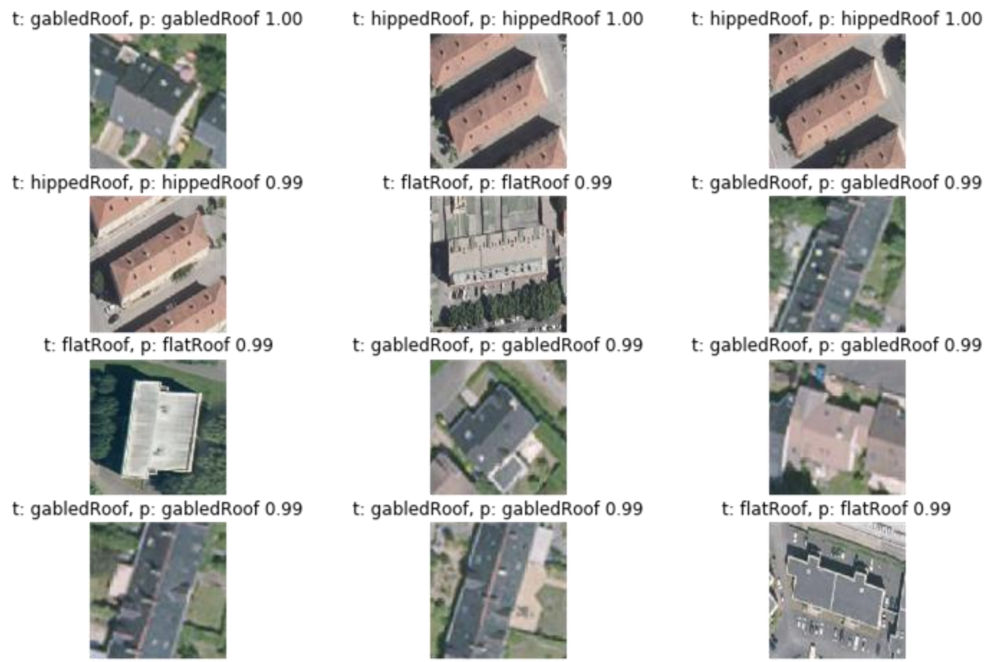
- Extract building footprint from aerial image
- Automatically detect building roof material
- Use of internal advanced Computer Vision algorithms



Some example of computer vision

Automatic Detection of Roof Type

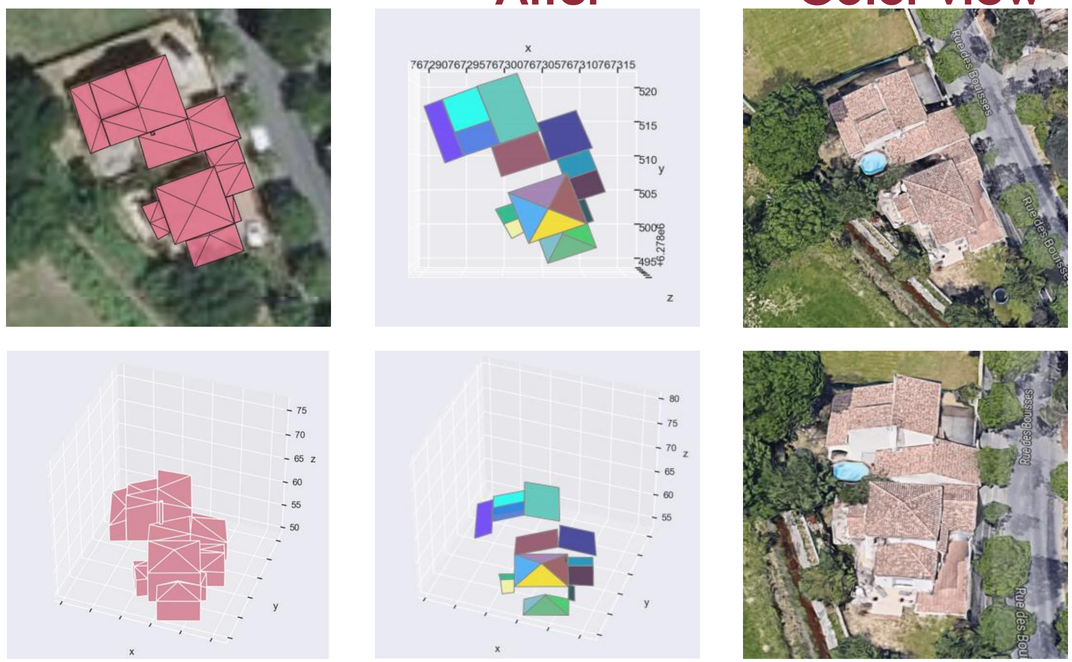
- Automatically extract building ***roof type*** from aerial image
- Follows INSPIRE roof type’s standard labels



Some example of computer vision

With our **3D reconstruction technology**, we can recreate building shapes with some elevation databases to create label to train algorithms to detect roof shapes.

Before **After** **Color view**



Orthogonal view

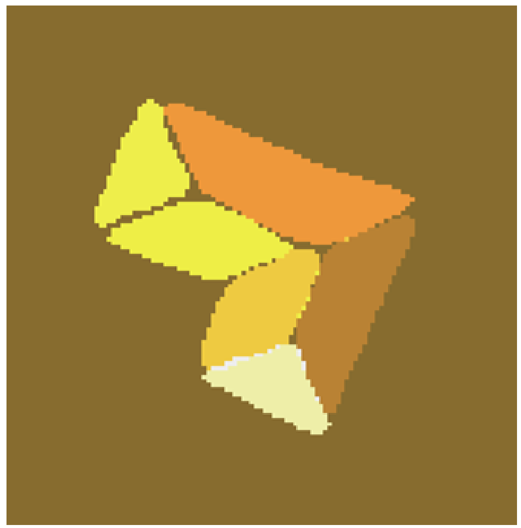
45° view

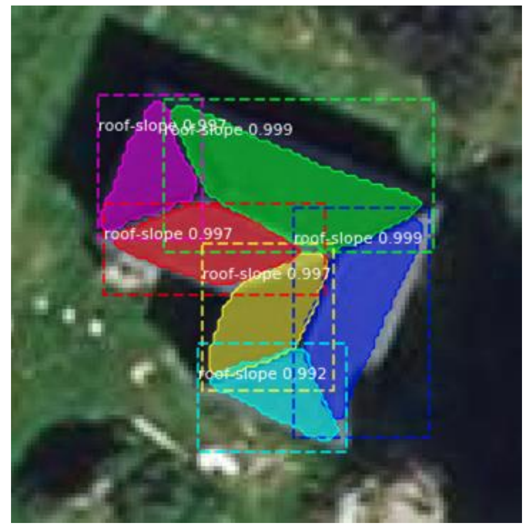
Some example of computer vision

Roof Structure Detection

Advanced aerial image processing to find roof slopes and describe:

- roof surface
- roof orientation
- estimate solar energy potential





Some example of computer vision

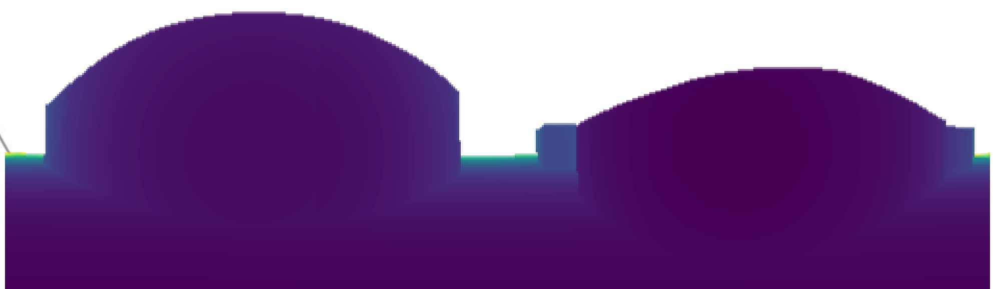
Solar panels detection

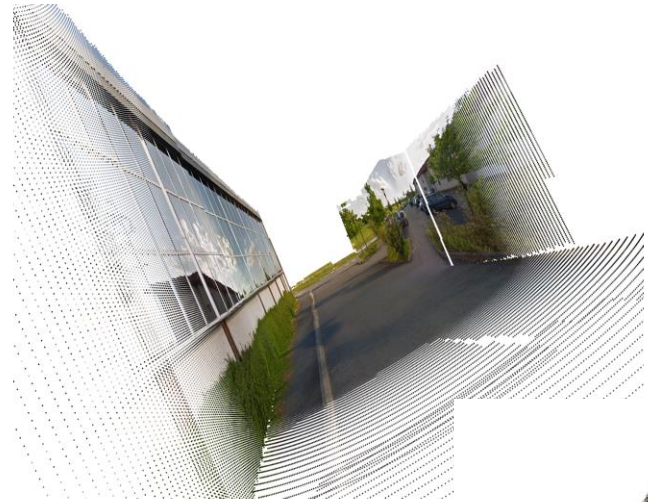
Automatic detection and segmentation of **solar panels** on the roofs of french buildings

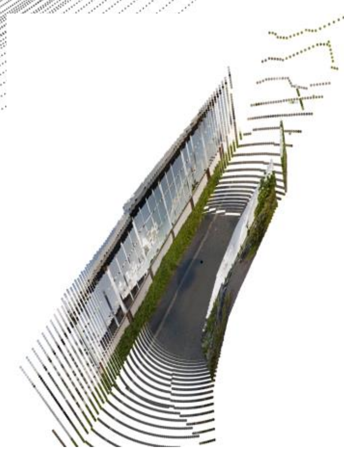
Some example of computer vision

Street View Image Processing

Object-of-interest extraction from Street View panoramas & 3D projection





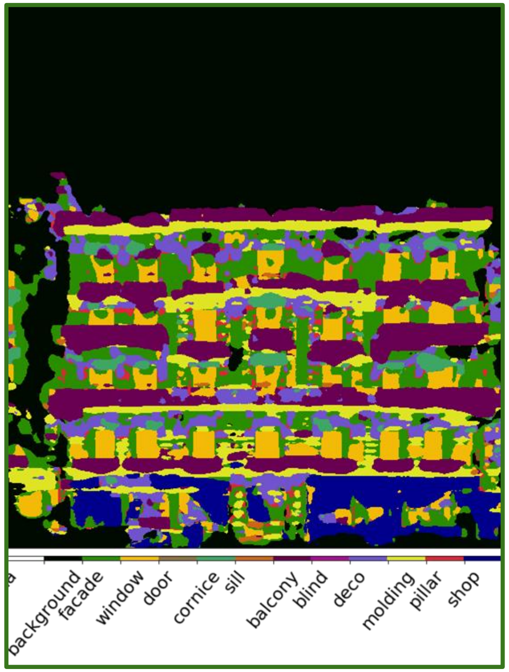


Some example of computer vision

Street View Image Processing

Flat plane projection

DEEP LEARNING



Façade segmentation model

01 COLLECTE
Smart Data

new

02 TRAITEMENT
Data Cleaning
Feature engineering
Kriging & machine learning

03 SELECTION
Malédiction de la dimension

04 MODELE
GLM
Régression quantile

05 VISUALISATION

ANALYSE DU RISQUE **CLIMATIQUE**

INSTITUT DES ACTUAIRES

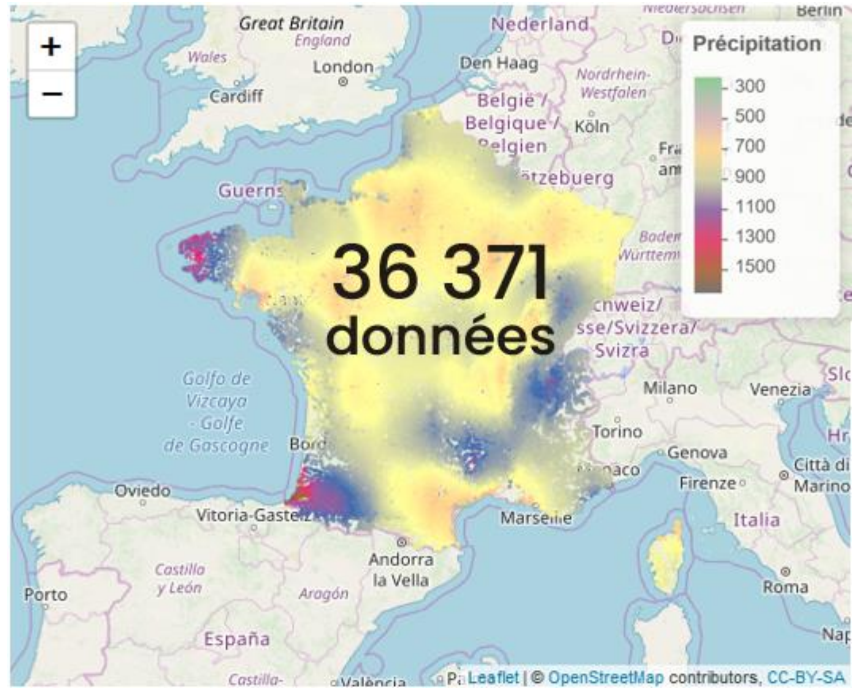
Intérêt du Machine Learning

Traitement, complétion & exhaustivité des données

- Meilleure gestion des données
- Lissage précis pour capter toute l'information
  - Liens entre les données, valeurs extrêmes
- Kriging

KRIGING

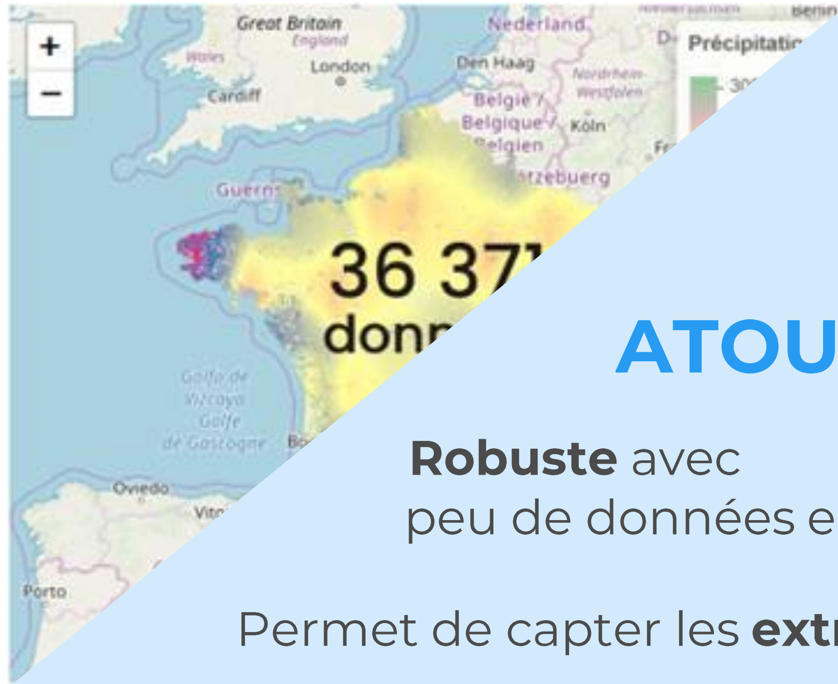
Lissage prédictif par interpolation spatiale



KRIGING

INSTITUT DES ACTUAIRES

Lissage prédictif par interpolation spatiale



ATOUTS :

Robuste avec
peu de données en entrée

Permet de capter les **extrêmes**

Utilise toutes les informations

Analyse et **prédit** un score

Mais aussi

INSTITUT DES ACTUAIRES

Température

Vent

Neige

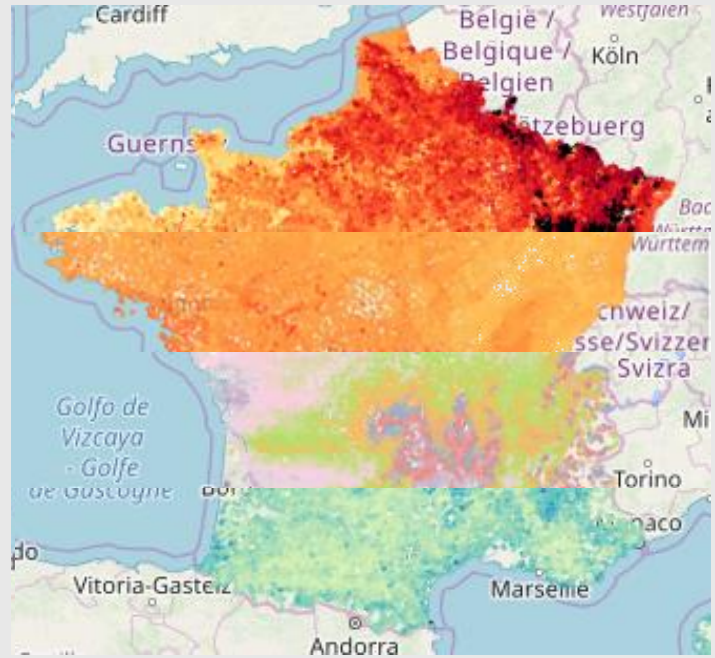
HDD18

HDD0

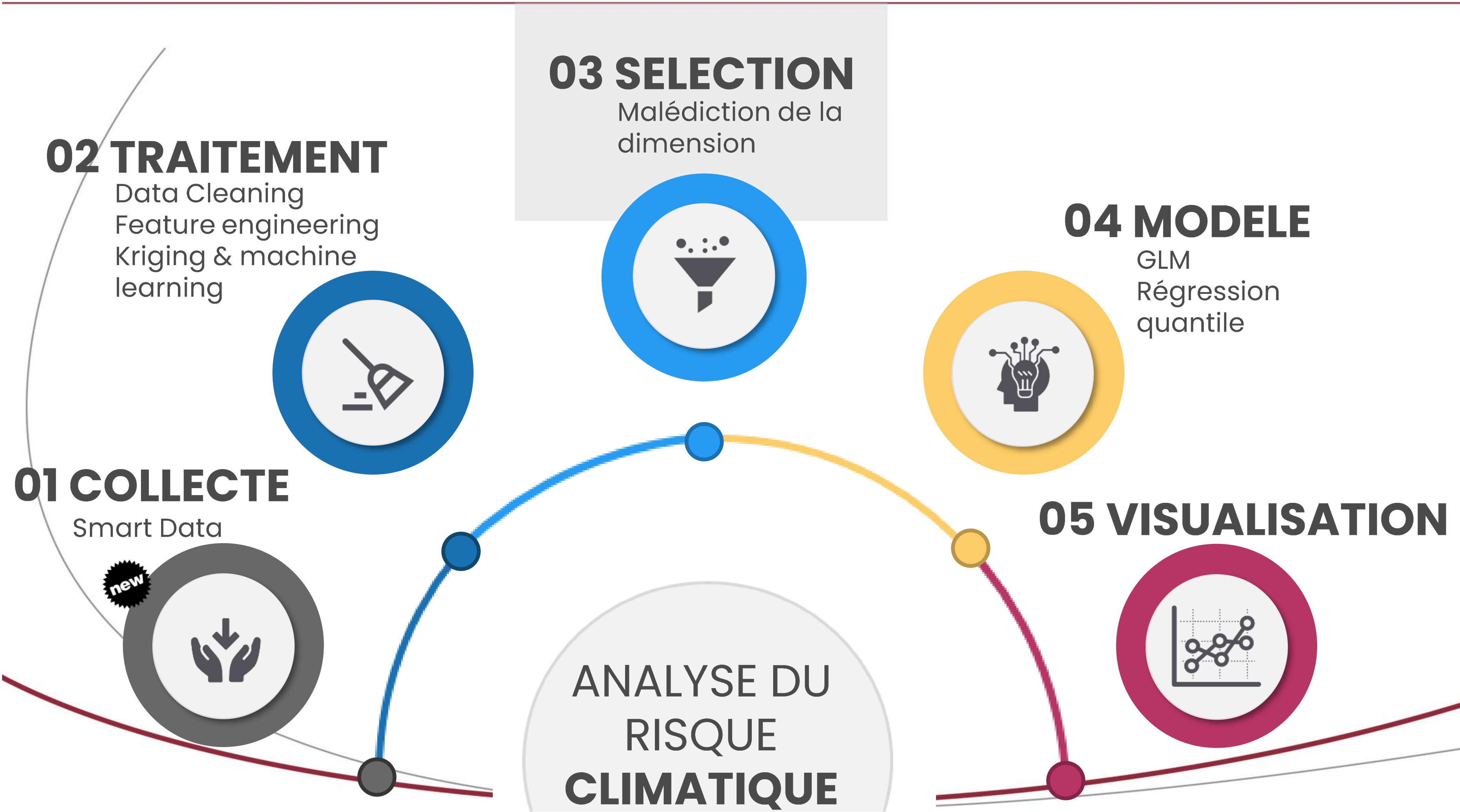
Gel meurtrier

Ensoleillement

. . .



INSTITUT DES ACTUAIRES

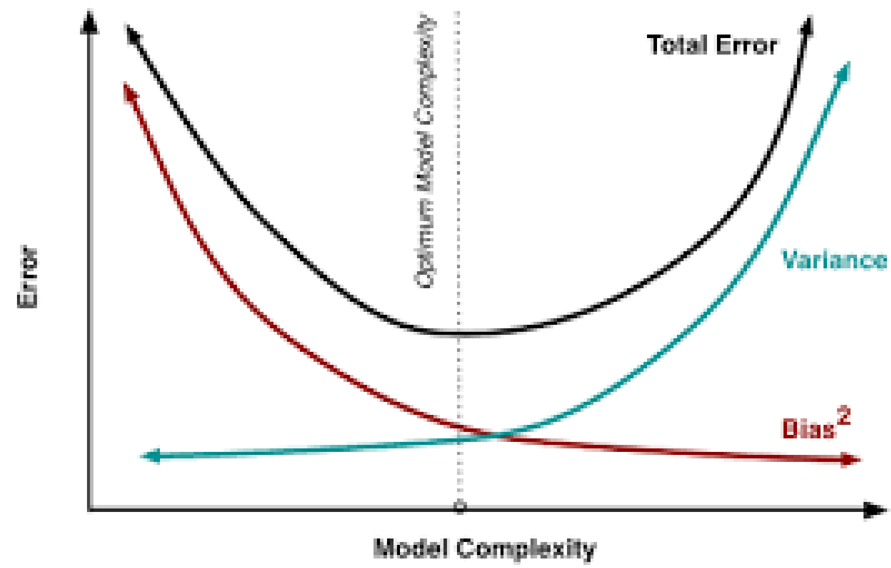


addactis Allianz nam.R

Éclairer les risques, tracer l'avenir

Malédiction de la dimension

Pas de **loi des grands nombres** avec la dimension !



- $Y = f(\mathbf{x}) + \varepsilon$, avec $\mathbf{x} = (x_1, \cdots, x_d)$ à d dimensions
- $\mathrm{Var}(\widehat{f}(\mathbf{x})) \sim \sigma_\varepsilon^2 \, \frac{d}{n}$ (f linéaire)

Nombre de variables

Nombre d'observations

INSTITUT DES ACTUAIRES

01 COLLECTE
Smart Data

new

02 TRAITEMENT
Data Cleaning
Feature engineering
Kriging & machine learning

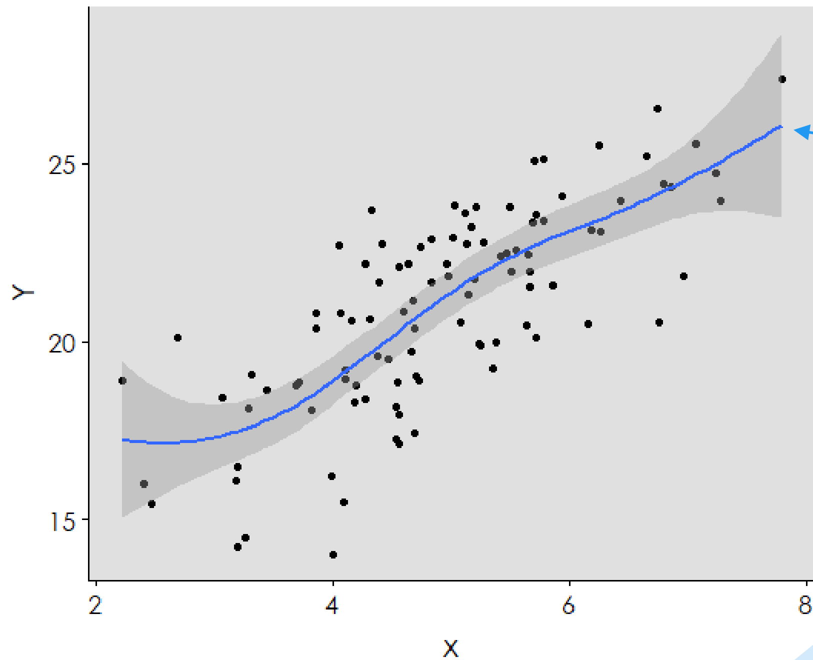
03 SELECTION
Malédiction de la dimension

04 MODELE
GLM
Régression quantile

05 VISUALISATION

ANALYSE DU RISQUE **CLIMATIQUE**

GLM Modélisation de la charge



LIMITES :

Capte l'**effet moyen**
Manque de précision

Effet pur **biaisé**
Par l'interaction entre variables

Itérations **coûteuses**
dans les tests backwards & forwards

Capacités **explicatives** et non prédictives
Interactions approximatives
Info erronée intégrée à la prédiction

Nécessité d'un **grand** jeu de données


SMART DATA
&
MODELE A EFFET MOYEN
?
PIRELLI
Shell
HELIX
06
DALMORE
2CV

INSTITUT DES ACTUAIRES

Intérêts du Machine Learning

Décuplés par la smart data

- Compréhension de l'interaction/lien dans les phénomènes climatiques
  - XGBoost
  - Random Forest

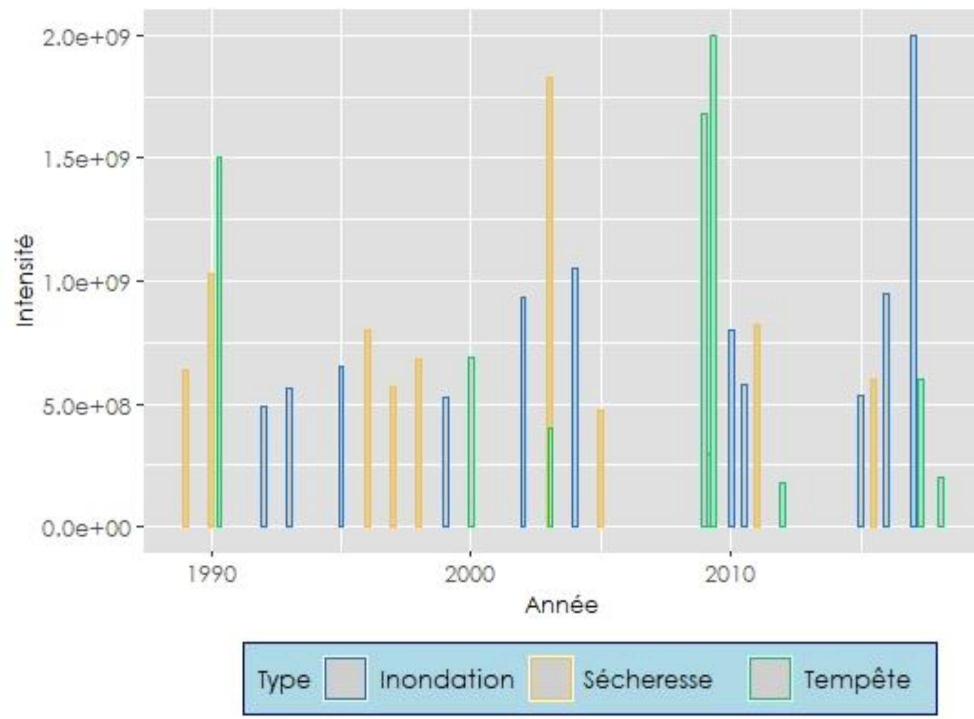
- Capte les phénomènes rares et cas extrêmes
  - Régression quantile

Régression quantile

Evènement climatique atypique

Variable climatique **complexe** à modéliser

Top 20 des évènements CAT NAT en termes de dommages assurés



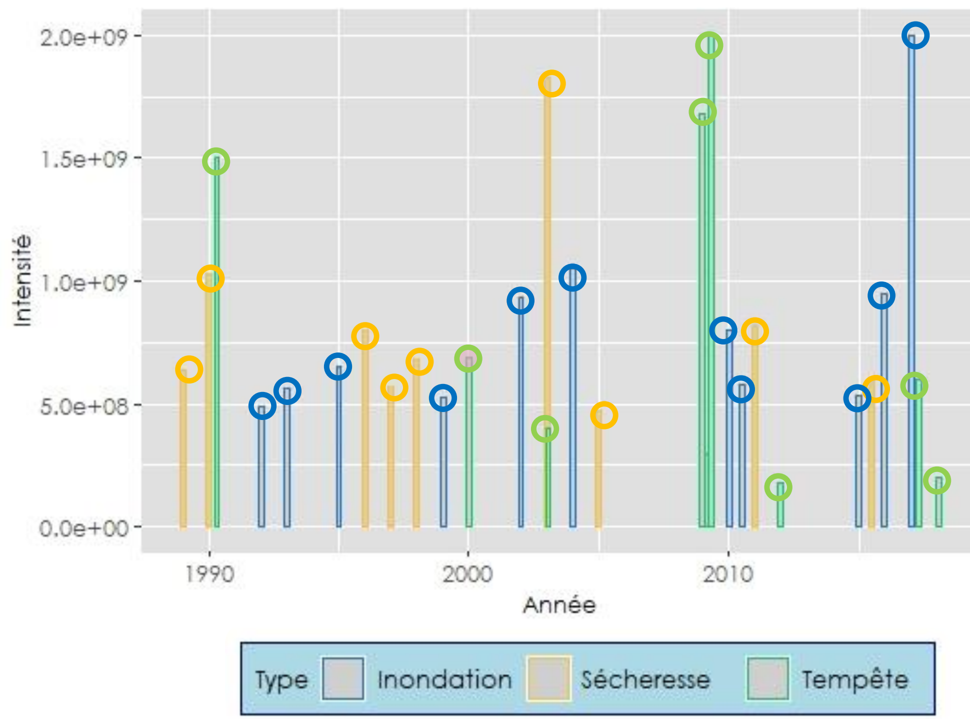
Régression quantile

Evènement climatique atypique

Accélération

Changement de structure

Top 20 des évènements CAT NAT en termes de dommages assurés



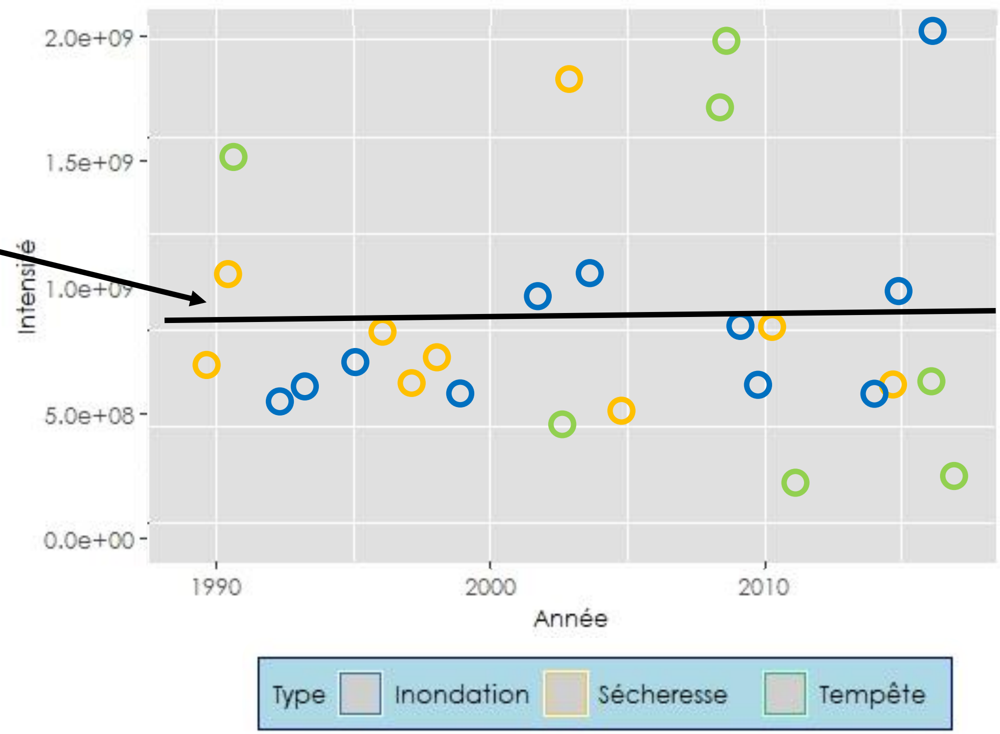
Régression quantile

INSTITUT DES ACTUAIRES

Evènement climatique atypique

Effet moyen GLM

Top 20 des évènements CAT NAT en termes de dommages assurés

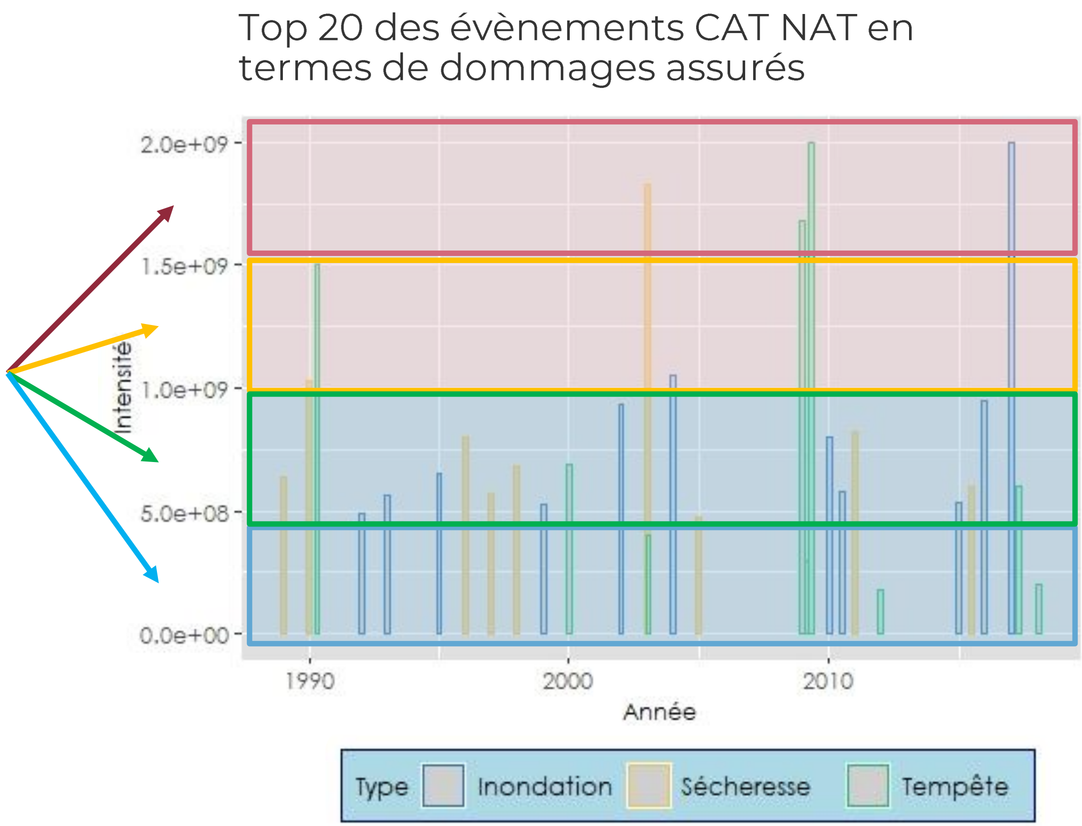


Régression quantile

INSTITUT DES ACTUAIRES

Evènement climatique atypique

Analyse par quantile



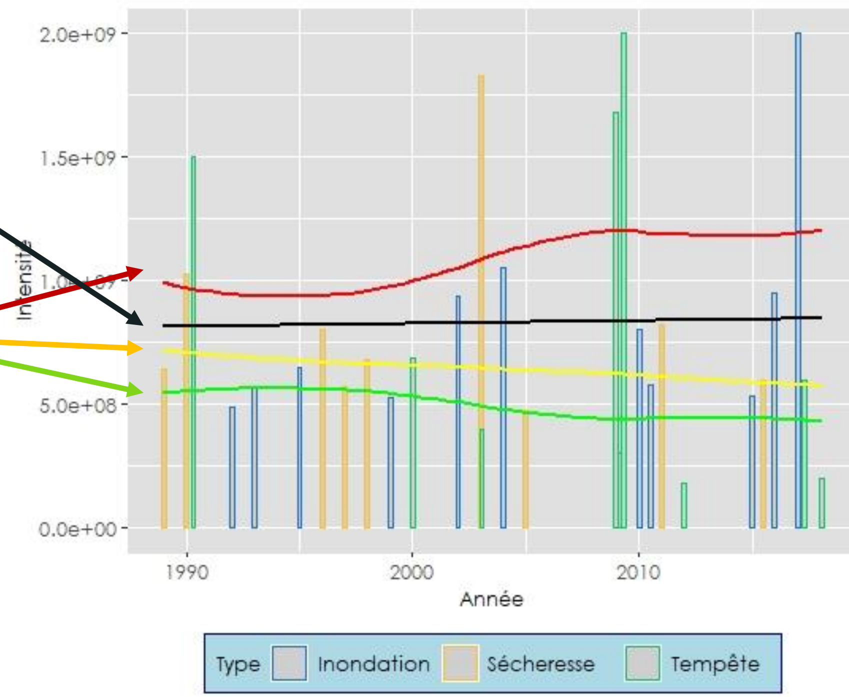
Régression quantile

Evènement climatique atypique

Top 20 des évènements CAT NAT en termes de dommages assurés

GLM Effet moyen

Régression quantile



INSTITUT DES ACTUAIRES

03 SELECTION
Malédiction de la dimension

02 TRAITEMENT
Data Cleaning
Feature engineering
Kriging & machine learning

04 MODELE
GLM
Régression quantile

01 COLLECTE
new
Smart Data

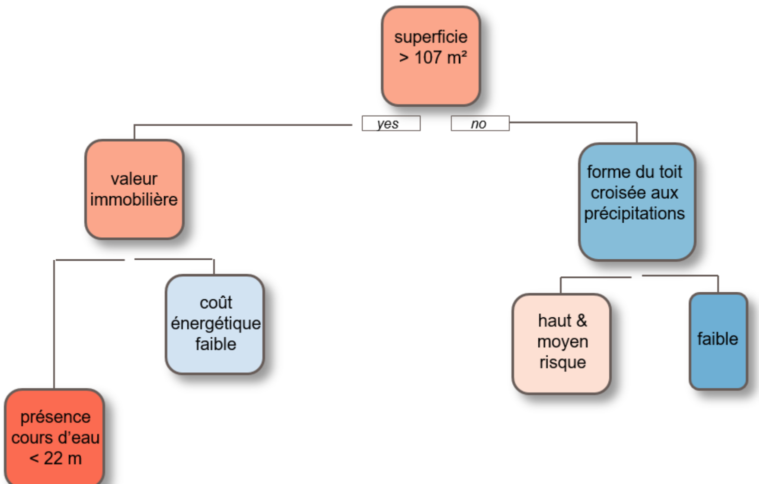
05 VISUALISATION

ANALYSE DU RISQUE **CLIMATIQUE**

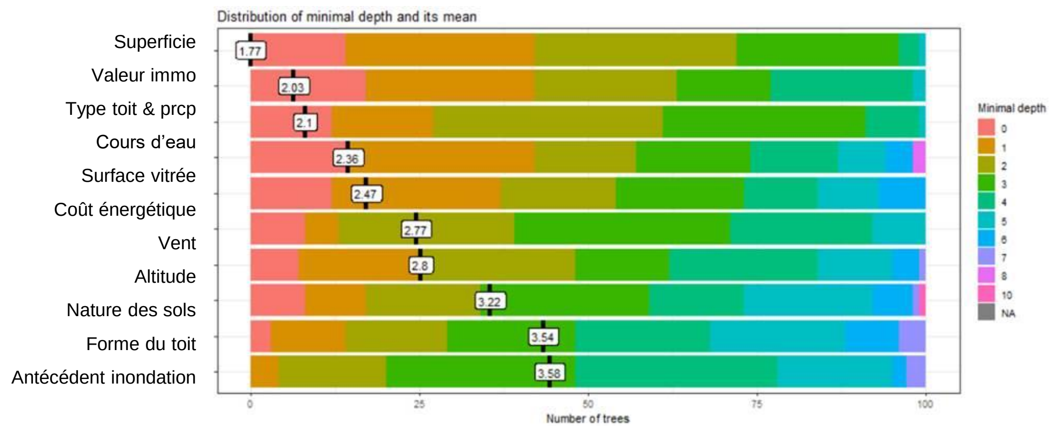
Visualisation

Explainable Machine Learning

Explainable CART



Explainable RF



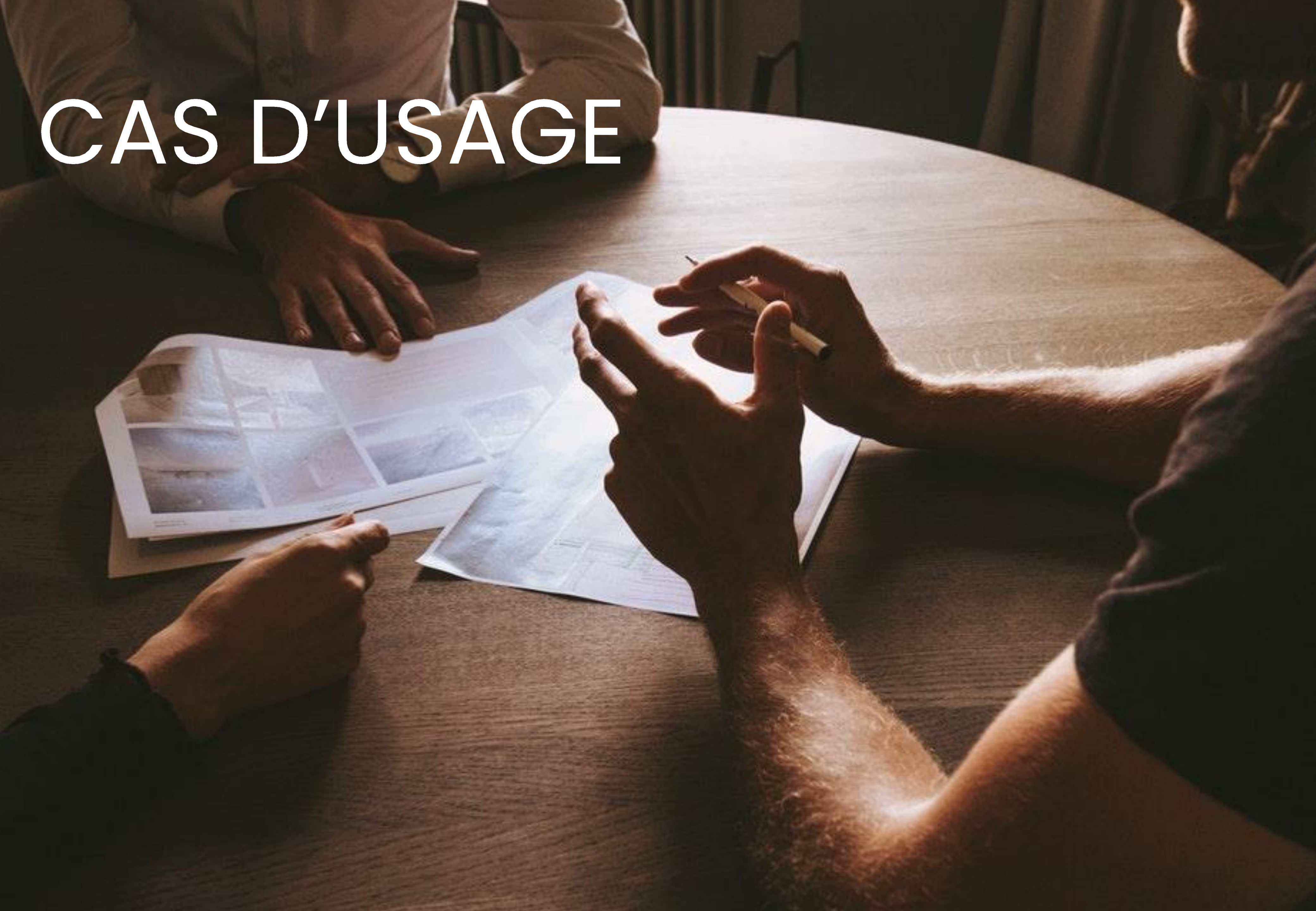
Explainable Boosting Machine
Explainable Deep Learning

ATOUTS :

Interprétabilité

Minimiser **l'effet boîte noire**

CAS D’USAGE



Cas d’usage

1. Souscription

Qualité de la donnée

- Alerte sur la qualité du risque pour l’intermédiaire/visa
- Interdiction de souscription de certains risques
- Ajout automatique de clauses d’exclusion/de limites de garanties

Cas d'usage

INSTITUT DES ACTUAIRES

2. Tarification

Qualité de la donnée

- Pré-remplissage des questions tarifaires
- Complétion du tarif technique par des variables additionnelles
- Suppression de questions
- Tarif « 0 question »

Cas d'usage

INSTITUT DES ACTUAIRES

3. Gestion de portefeuille

Qualité de la donnée

- Ciblage de contrôles de souscription (déclarations erronées)
- Majorations ciblées via des variables additionnelles
- Surveillance spécifique du portefeuille, si lift suffisant sur les graves ou les climatiques

Cas d’usage

INSTITUT DES ACTUAIRES

4. Prévention et gestion de sinistres

Qualité de la donnée

- Application de réduction proportionnelles
- Prévention ciblée sur certains types de risque (alertes climatiques)
- Choisir les risques qui méritent des stratégies d’indemnisation ciblées (réparation en nature, passage d’experts, orientation des prestataires, etc…)

Cas d’usage

5. Ciblage et multi-équipement

Qualité de la donnée

- Campagnes ciblées sur certaines cibles (préférence technique)
- Rebond commercial plus précis (Next Best Product), données prédictives de comportements ou de besoins annexes
- Marketing sortant en envoyant directement un prix (tarif zero question ou soumis à peu de conditions)

Cas d'usage

INSTITUT DES ACTUAIRES

6. Risques Climatiques & gestion des expositions

Qualité de la donnée

- Gestion des accumulations par zones
- Chiffrage de scenario par zone d'accumulation avec estimation fine des sinistres par contrat
- Définition d'une politique d'acceptation des risques climatiques basée sur la géolocalisation et la sensibilité précise de chaque risque aux évènements