INDICATEURS ET TENDANCES D'ÉVOLUTION : APPROCHES STATISTIQUES ET PERSPECTIVES EN INTELLIGENCE ARTIFICIELLE

Séminaire APRONA : modèles et outils statistiques appliqués à l'hydrogéologie

Laurent Vaute, Hydrogéologue modélisateur – Data Scientist, BRGM Grand Est
30 novembre 2021

RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

AGENCE DE L'EAU RHIN•MEUSE

Géosciences pour une Terre durable
brgm

Sommaire

- **Contexte** : besoin de méthodes plus puissantes pour l'interprétation des données et la prévision des tendances d'évolution de la qualité des eaux souterraines.
- **Nouvelle étude R&D** BRGM – AERM « Typologie des points d'eau ».
- **Principes de l'apprentissage automatique** (machine learning).
- **Méthodologie, algorithmes et outils** retenus pour l'étude.
- **Exemple de regroupement** sur un jeu de données simplifié.

CONTEXTE

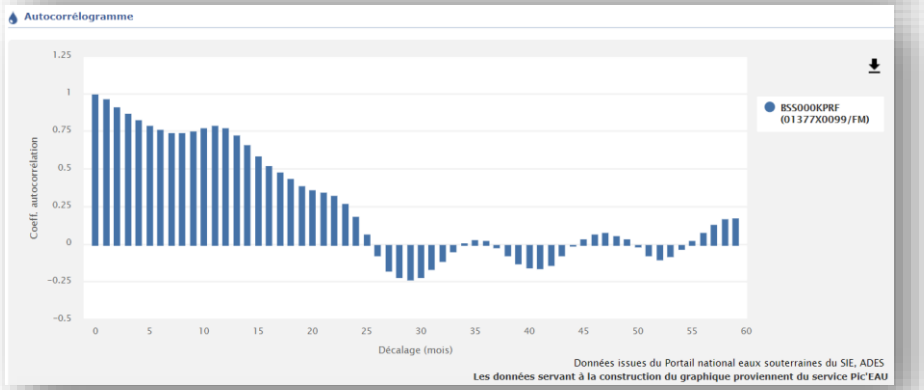
Une grande richesse de données

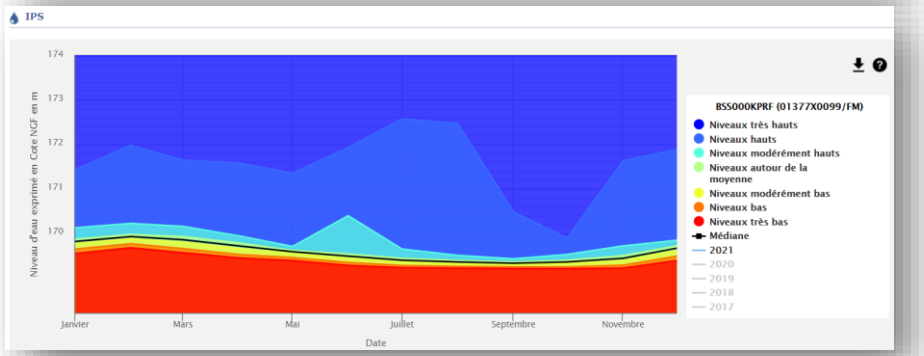
Depuis 20 ans, une politique de l'eau marquée par l'acquisition de très nombreuses données sur la quantité et la qualité de l'eau, et le souhait de reconquérir la qualité de l'eau souterraine

- Mise en place des **réseaux de surveillance** : réseaux patrimoniaux (1999) puis « DCE » en 2007 : réseaux de surveillance (RCS) et opérationnels (RCO).
- Réalisation de **campagnes exceptionnelles** de mesure des polluants émergeants.
- Définition de **captages prioritaires** vis-à-vis des pollutions diffuses agricoles (nitrates, phytosanitaires).

Des banques pour gérer cette grande quantité de données

- La **collecte** et la **bancarisation** des données ont été structurées au niveau des bassins et au niveau national (**ADES**).
- **Pour la quantité**, **des indicateurs statistiques sont disponibles en ligne** sur ADES pour les niveaux piézométriques (autocorrélogramme, IPS, statistiques simples) ; et il existe **de nombreuses méthodes de prévision efficaces** des niveaux piézométriques et des débits.

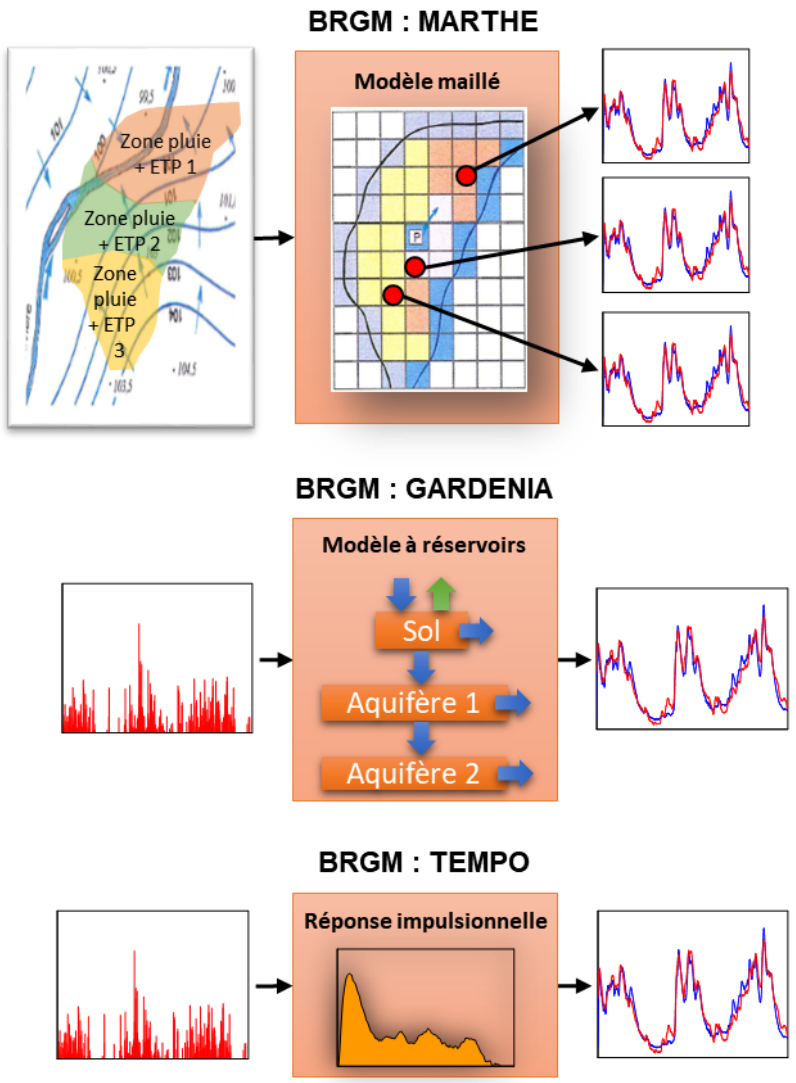




CONTEXTE

Quantité : modèles « classiques » de prévision

(NB : recherche en IA…)



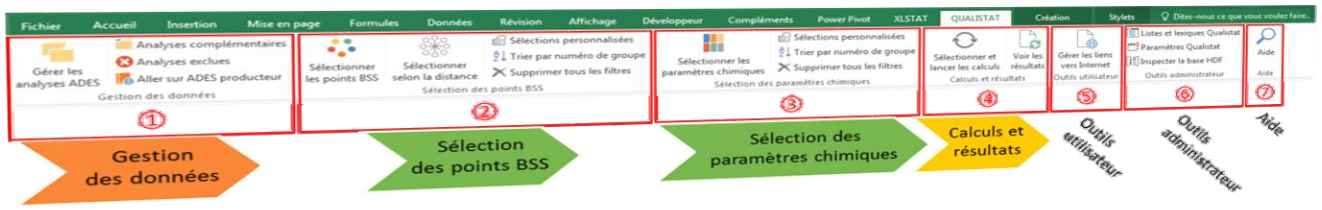
Possibilités et complexité croissantes

Entrées Zones de pluie + ETP, pompages, données géologiques et hydrogéologiques	**Modèle *spatialisé*** Lois physiques des écoulements + géométrie et propriétés de l'aquifère	**Sorties** Niveau et/ou débit en tout point du domaine modélisé
Entrées pluie, ETP, pompages	**Modèle *global*** Fonction(s) mathématique(s) à « sens » physique	**Sortie** Niveau et/ou débit en 1 ou 2 points
Entrées pluie, ETP, pompages	**Modèle *boite noire*** Fonction(s) mathématique(s)	**Sortie** niveau ou débit en 1 seul point

CONTEXTE

Qualité : des outils pour l'exploitation « en masse » des données

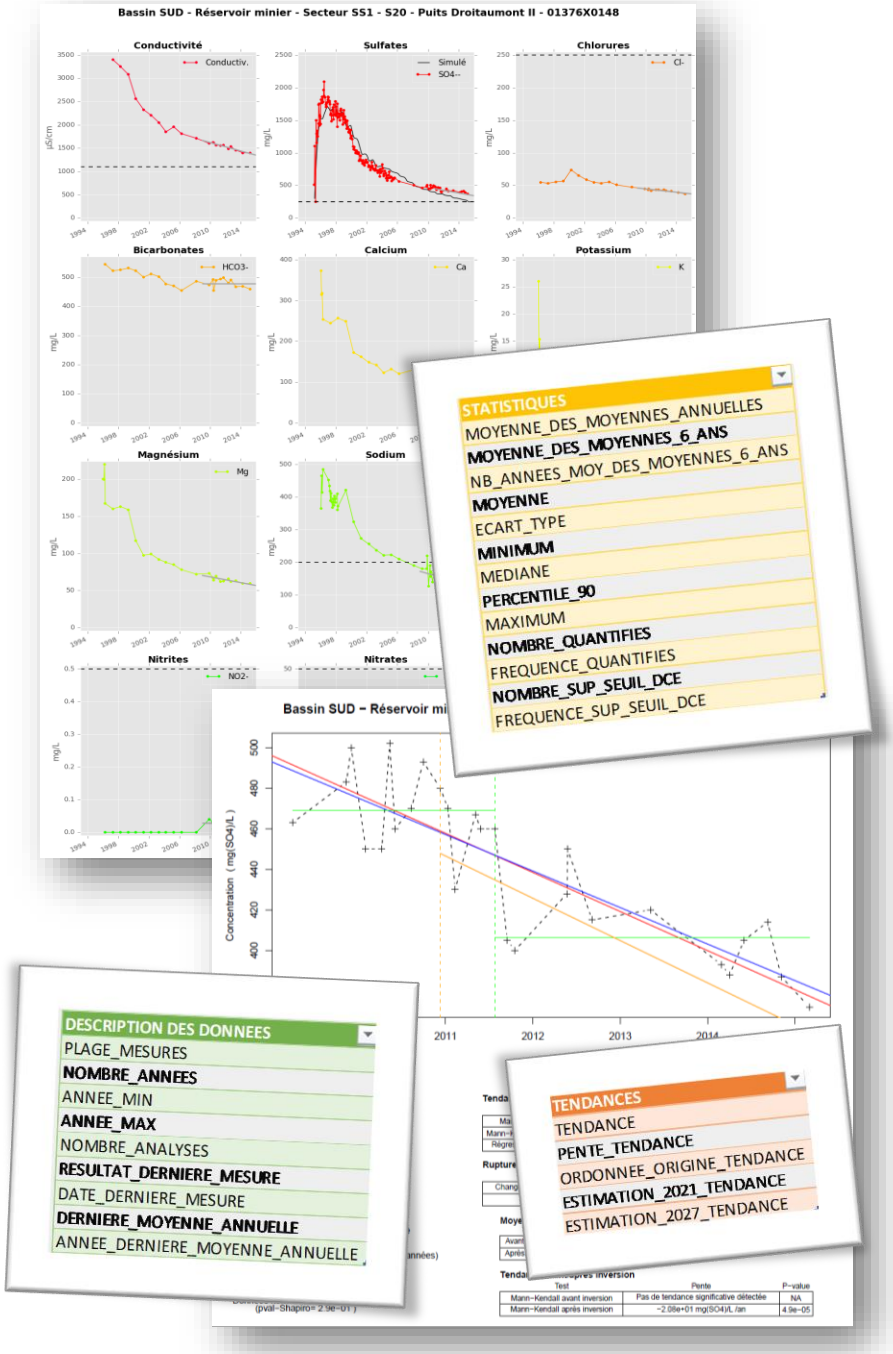
Des outils nécessaires…



Pour les eaux souterraines, un outil de **traitement statistique des données en masse** a été développé par le BRGM pour l'Agence de l'eau Rhin-Meuse : **Qualistat**, utilisé en 2019 pour le calcul de **l'état des masses d'eau souterraines**, l'identification des **paramètres chimiques déclassants**, etc.

…mais pas suffisants !

Mais comment bien interpréter l'énorme quantité de cartes, graphiques et indicateurs statistiques qu'il est possible de produire (jusqu'à 3800 paramètres x 12500 points) ? → besoin de méthodes et d'outils plus puissants pour **synthétiser l'information, identifier les phénomènes importants et construire des modèles de prévisions**.



ETUDE « TYPOLOGIE DES POINTS D'EAU »

Une nouvelle étude de R&D BRGM - AERM

Typologie des points d'eau pour l'interprétation des tendances d'évolution de la qualité de l'eau souterraine

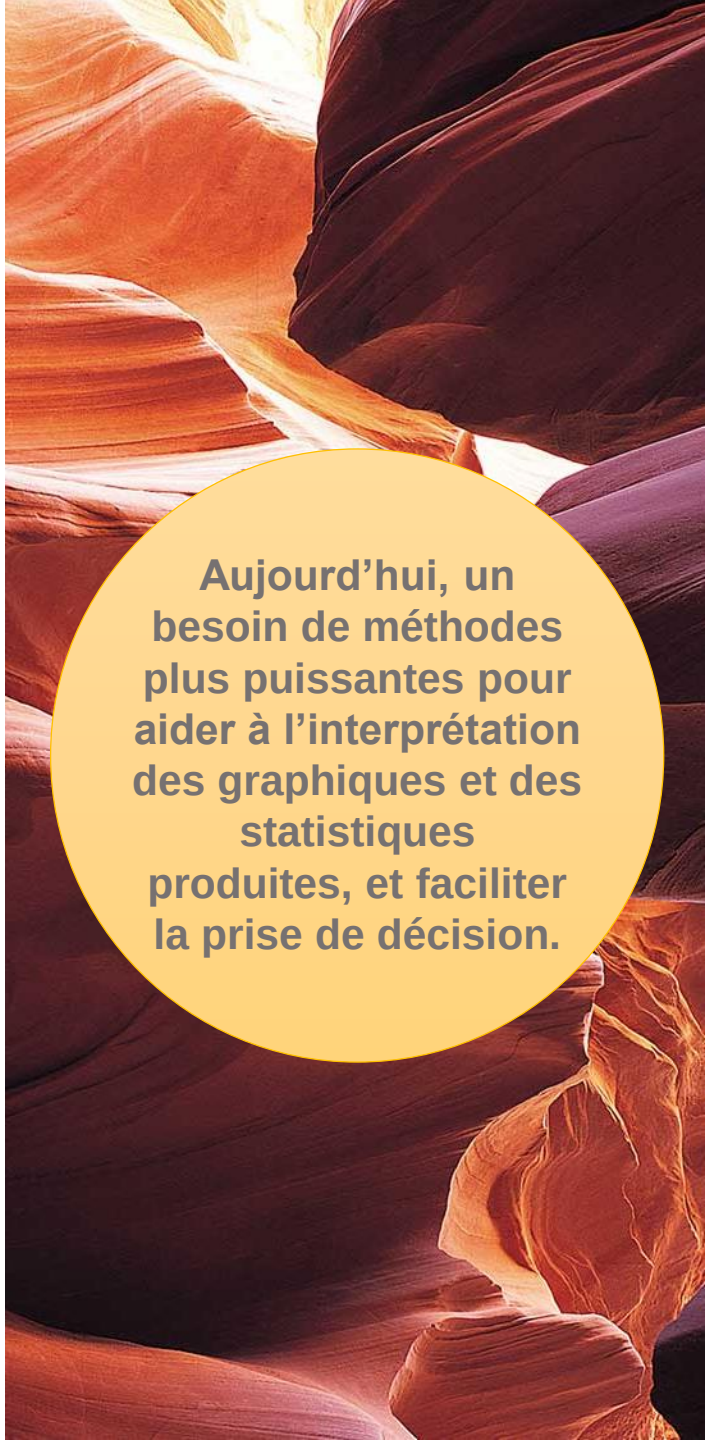
- Nouvelle étude de R&D partagés BRGM / Agence de l'eau Rhin-Meuse (2021-2022) : **regrouper 600 points d'eau selon leur environnement et leur fonctionnement hydrogéologique et leur fonctionnement hydrogéochimique → utiliser les meilleures informations disponibles, quel que soit leur nature et leur format**.

Objectifs opérationnels

- **Aide à l'optimisation de la surveillance des eaux souterraines** : représentativité des points, paramètres importants à surveiller, validation et interprétation des données.
- **Aide à la définition des actions pertinentes de reconquête de la qualité des eaux souterraines** : priorisation de la localisation et des leviers des actions.

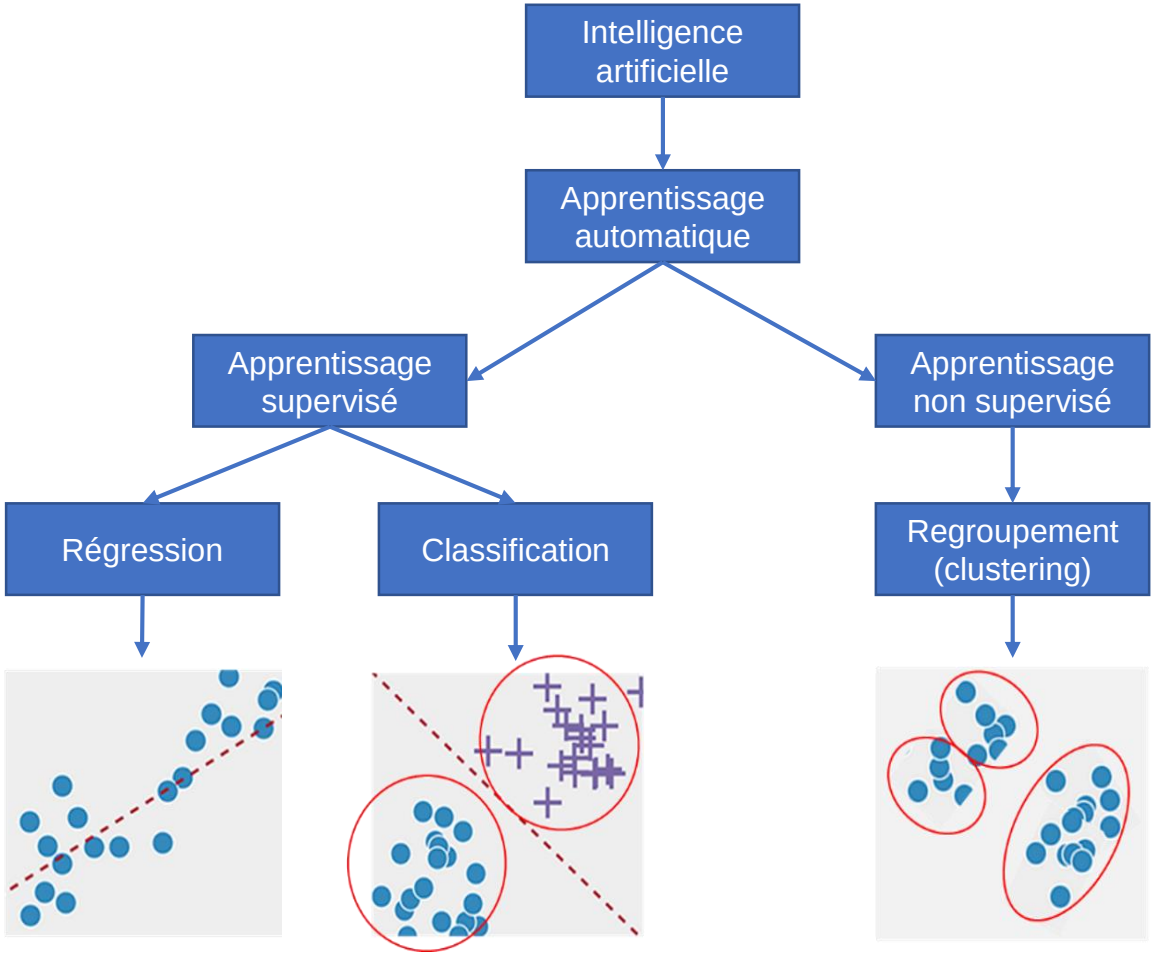
Pour atteindre ces objectifs : la R&D en apprentissage automatique

- Rechercher et développer de **nouvelles méthodes pour « faire parler les données »**, c'est-à-dire découvrir la structure des données : il s'agit d'une **démarche exploratoire**.
- **Méthodologie de l'apprentissage automatique (une branche de l'IA)** : mise en œuvre cohérente d'algorithmes et de méthodes statistiques pour explorer les données et construire des modèles de prévision.



Définitions

- **Apprentissage automatique** (machine learning) : champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour **donner aux ordinateurs la capacité "d'apprendre" à partir de données** (approche data-driven), en améliorant automatiquement leurs performances de prédiction (Wikipedia).
- **Une machine apprend quand elle cherche une formule mathématique** qui, lorsqu'elle est appliquée à des entrées (appelées "données d'entrainement"), produit les sorties attendues.
- **Des algorithmes** d'apprentissage automatique sont constamment améliorés ou créés. Ils permettent par exemple de construire des modèles pour :
  - **prédire** une valeur ou une classe,
  - **identifier des groupes** dans les données.

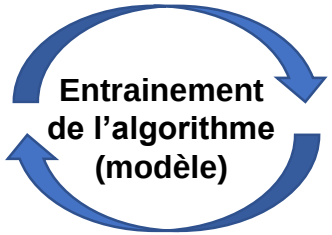


La régression : pour prédire une valeur numérique

- On veut **prédire la valeur d'une variable numérique connue** : la variable **cible**.

EXEMPLES ↓ / VARIABLES EXPLICATIVES →	Indicateur n°1	Indicateur n°2	Indicateur n°3	Indicateur n°4	…	Indicateur n°N
Point d'eau n°1	8000	1,2	Oui	Classe 2	…	Calcaire
Point d'eau n°2	10500	1	Non	Classe 3	…	Alluvions
Point d'eau n°3	900	2	?	Classe 5	…	Calcaire
…	…	…	…	…	…	…
Point d'eau n°600	3000,2	6	Oui	Classe 1	…	Grès
Nouveau point d'eau	**5000**	**< 2**	**Oui**	**Classe 4**	**…**	**Alluvions**

L'algorithme de régression doit s'entraîner à prédire la valeur numérique de la variable cible pour chaque exemple qui lui est fourni :



VARIABLE CIBLE : valeurs connues
12
15,3
<1
…
26
4,2

Prédiction

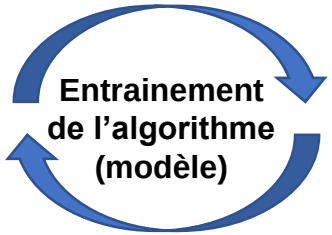
- **C'est un apprentissage supervisé** : on fournit à **l'algorithme de régression** les valeurs de la variable cible pour les exemples qui nous sont connus, et on lui demande d'**apprendre les paramètres du modèle de prévision à l'aide des variables explicatives** : une fois qu'il est entraîné, le modèle peut être utilisé pour faire des prévisions.

La classification : pour prédire une classe (= catégorie)

- On veut **prédire le numéro de classe d'une variable catégorielle connue** : la variable **cible**.

EXEMPLES ↓ / VARIABLES EXPLICATIVES →	Indicateur n°1	Indicateur n°2	Indicateur n°3	Indicateur n°4	…	Indicateur n°N
Point d'eau n°1	8000	1,2	Oui	Classe 2	…	Calcaire
Point d'eau n°2	10500	1	Non	Classe 3	…	Alluvions
Point d'eau n°3	900	2	?	Classe 5	…	Calcaire
…	…	…	…	…	…	…
Point d'eau n°600	3000,2	6	Oui	Classe 1	…	Grès
Nouveau point d'eau	**5000**	**< 2**	**Oui**	**Classe 4**	**…**	**Alluvions**

L'algorithme de classification doit s'entraîner à prédire le numéro de classe de la variable cible pour chaque exemple qui lui est fourni :



VARIABLE CIBLE : classes connues
Classe 1
Classe 3
Classe 3
…
Classe 6

Prédiction → **Classe 3**

- **C'est un apprentissage supervisé** : on fournit à **l'algorithme de classification** les numéros de classe de la variable cible pour les exemples qui nous sont connus, et on lui demande d'**apprendre les paramètres du modèle de prévision à l'aide des variables explicatives** : une fois qu'il est entraîné, le modèle peut être utilisé pour faire des prévisions.

Le regroupement (clustering) : pour construire une typologie

- On veut **regrouper les points d'eau** qui se « ressemblent ».

EXEMPLES ↓ / VARIABLES →	Indicateur n°1	Indicateur n°2	Indicateur n°3	Indicateur n°4	…	Indicateur n°N
Point d'eau n°1	8000	1,2	Oui	Classe 2	…	Calcaire
Point d'eau n°2	10500	1	Non	Classe 3	…	Alluvions
Point d'eau n°3	900	2	?	Classe 5	…	Calcaire
…	…	…	…	…	…	…
Point d'eau n°600	3000,2	6	Oui	Classe 1	…	Grès
Nouveau point d'eau	**5000**	**< 2**	**Oui**	**Classe 4**	**…**	**Alluvions**

L'algorithme doit identifier les groupes de points qui se « ressemblent » et leur donner un numéro de type (il faudra ensuite interpréter chaque « type ») :

Regroupement

NUMERO DE GROUPE à identifier
Groupe 1
Groupe 1
Groupe 8
…
Groupe 3

Prédiction

1

- **Apprentissage non supervisé** : **l'algorithme de regroupement (clustering)** doit se débrouiller « seul » (il n'y a **pas de variable cible**) pour trouver les groupes (les clusters) par des calculs de distance, de densité de points, etc.

Méthodologie (1/2)

3 jeux de données thématiques (tabulaires, spatiales + attributaires (SIG), chronologiques)

Compilation

Points d'eau	Variables
	Environnement du point d'eau • hydrologie dont bassins versants amont du point (SIG) • occupation du sol (SIG) • pressions agricoles (SIG) • pressions industrielles (SIG)

Points d'eau	Variables
	Fonctionnement hydrogéologique • aquifère capté (TAB) • type d'écoulement (TAB) • vulnérabilité (SIG) • infiltration (SIG / CHRO) • piézométrie / débit (CHRO)

Points d'eau	Variables
	Fonctionnement hydrogéochimique • *concentrations en divers composés : majeurs, mineurs, contaminants naturels ou anthropiques (CHRO)*

Extraction de variables

Prétraitement

Indicateurs simples

Indicateurs élaborés

…

Données SIG	Séries chronologiques
- Détermination de la **zone d'alimentation naturelle à l'amont** de chaque point (= AAC ou bassin versant hydrogéologique ou bassin versant hydrologique). *[Verrou : calcul du bassin versant hydrogéologique à l'amont du point d'eau]*	- Suppression des **analyses incohérentes**. - Suppression des **valeurs aberrantes dans chaque série**. *[Verrou : élimination automatique des analyses incohérentes et des valeurs aberrantes]*
- **% de surface de BV pour chaque catégorie** de : - **occupation du sol,** - **pression agricole,** - **nature du sol,** - **vulnérabilité,** - etc.	- **Moyennes**, **écart-types**, **fréquences**... tenant compte correctement des **valeurs censurées**. *[Verrou : calcul de statistiques tenant compte correctement des valeurs censurées (< LQ)]*
	- **Tendance d'évolution**. - **Autocorrélation**. - **Cyclicité**. - **Faciès hydrochimique**. - etc… *[Verrou : calcul automatique d'indicateurs élaborés relatifs aux séries chronologiques]*

Méthodologie (2/2)

Agrégation + Analyse statistique + Choix des variables

	Variables = caractéristiques du point, indicateurs simples, indicateurs élaborés		
Points d'eau	**Environnement du point d'eau** • hydrologie dont bassins versants amont du point • occupation du sol, • pressions agricoles, • pressions industrielles	**Fonctionnement hydrogéologique** • aquifère capté • type d'écoulement • vulnérabilité • *infiltration* • *piézométrie / débit*	**Fonctionnement hydrogéochimique** • *concentrations en divers composés : majeurs, mineurs, contaminants naturels ou anthropiques*

1 tableau de données numériques et catégorielles

nb variables >>>

Normalisation & Encodage

Variable numérique	Variable catégorielle
- **Normalisation** : valeurs ramenées à une échelle commune [0-1] ou [-1,1], - ou **Standardisation** : donner aux valeurs les propriétés d'une distribution normale standard.	- **Encodage** : création d'autant de nouvelles variables numériques binaires (0, 1) ou (0, x) que de catégories présentes dans la variable catégorielle de départ (→ **augmentation potentiellement importante du nombre de variables**).

1 tableau de données numériques

nb variables >

Réduction de dimensions & Visualisation

- **Calcul d'un nombre réduit de nouvelles variables** numériques portant la majeure partie de l'information initiale (AFM, UMAP) et/ou visualisations multivariées (RADVIZ).

Modélisations

- **Ajustement de 5 modèles de classification et/ou de régression** permettant de prédire la valeur d'indicateurs simples ou élaborés (variables catégorielles ou continues), dont les sorties pourraient être eux-mêmes utilisées comme indicateurs si leur fiabilité est suffisante.

[Point d'attention : beaucoup de variables explicatives potentielles, mais peu d'exemples (600)]

- **Regroupement des points d'eau** (HDBSCAN) en recherchant un nombre restreint de **« types » interprétables** (choisir un algorithme permettant de déterminer le type d'un nouveau point).

[Point d'attention : explicabilité du regroupement / visualisation des groupes de points]

Itérations

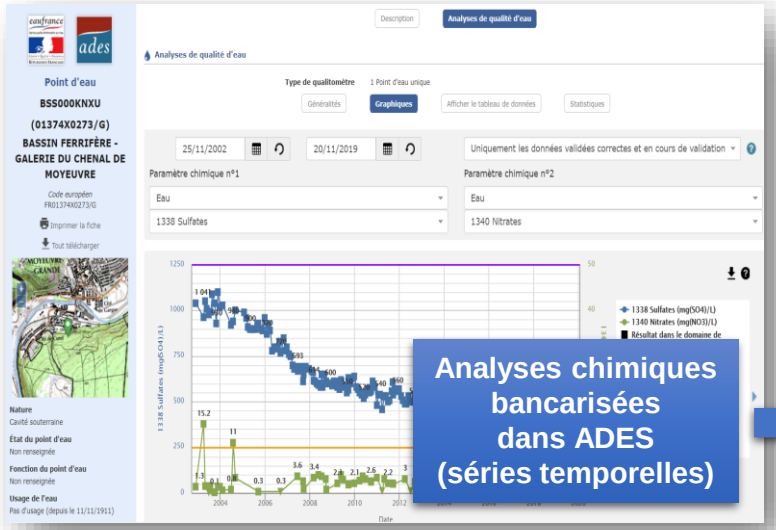
Algorithmes et outils retenus pour l'étude

- **Recherche bibliographique** relative aux méthodes statistiques et algorithmes open-source de :
  - ✓ **détection des anomalies** dans les séries temporelles,
  - ✓ calculs d'indicateurs tenant compte des **données censurées**,
  - ✓ **prévision** de valeurs ou de classes,
  - ✓ **regroupement** (clustering).
- **Tests** d'algorithmes et d'outils**.**
- **Tests de méthodes de regroupement de points** sur un jeu de données simplifié, et choix d'une méthode **adaptée aux données de l'étude**.

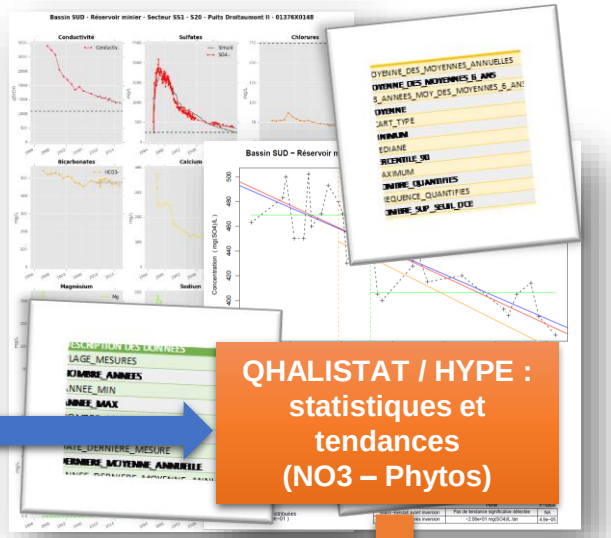
Usage	Algorithme / librairie (Langage)	Jeu de test éventuel	Commentaire
Détection d'anomalies temporelles	**Silverkite** / Greykite (Python)	Données AERM invalidées et anomalies BFL	**Envisagé** : à comparer à 2 approches plus classiques : statistique et modèle auto-arima
Calcul d'indicateurs statistiques et de tendances linéaires	**Qualistat 4.0** (Python et R)		**Retenu** : mis à jour en 2021 car nouveaux formats d'export ADES
Calcul d'indicateurs avec données censurées	**NADA** (R)	Données de la librairie NADA	**Retenu** : testé en 2021
Calcul d'indicateurs spatialisés	**Qgis** (Python)		**Retenu** : standard SIG open-source, compatible Arcgis
Prévision de valeurs ou de classes	**Catboost** (Python)		**Retenu** : 5 modèles à construire permettant éventuellement de calculer de nouveaux indicateurs
Regroupement (clustering)	**AFDM** + **HDBSCAN** + **UMAP** + **RADVIZ** (Python)	Jeu de données simplifié à 9 variables	**Retenu** : approche testée et validée en 2021
Interface de développement	**JupyterLab** (Python et R)		**Retenu** : standard qui respecte les principes de la science ouverte et reproductible
Tableaux de bord	**Tableau Software**		**Retenu** : standard utilisé à l'AERM

EXEMPLE DE REGROUPEMENT

Source des données et variables extraites



Analyses chimiques bancarisées dans ADES (séries temporelles)



QHALISTAT / HYPE : statistiques et tendances (NO3 – Phytos)

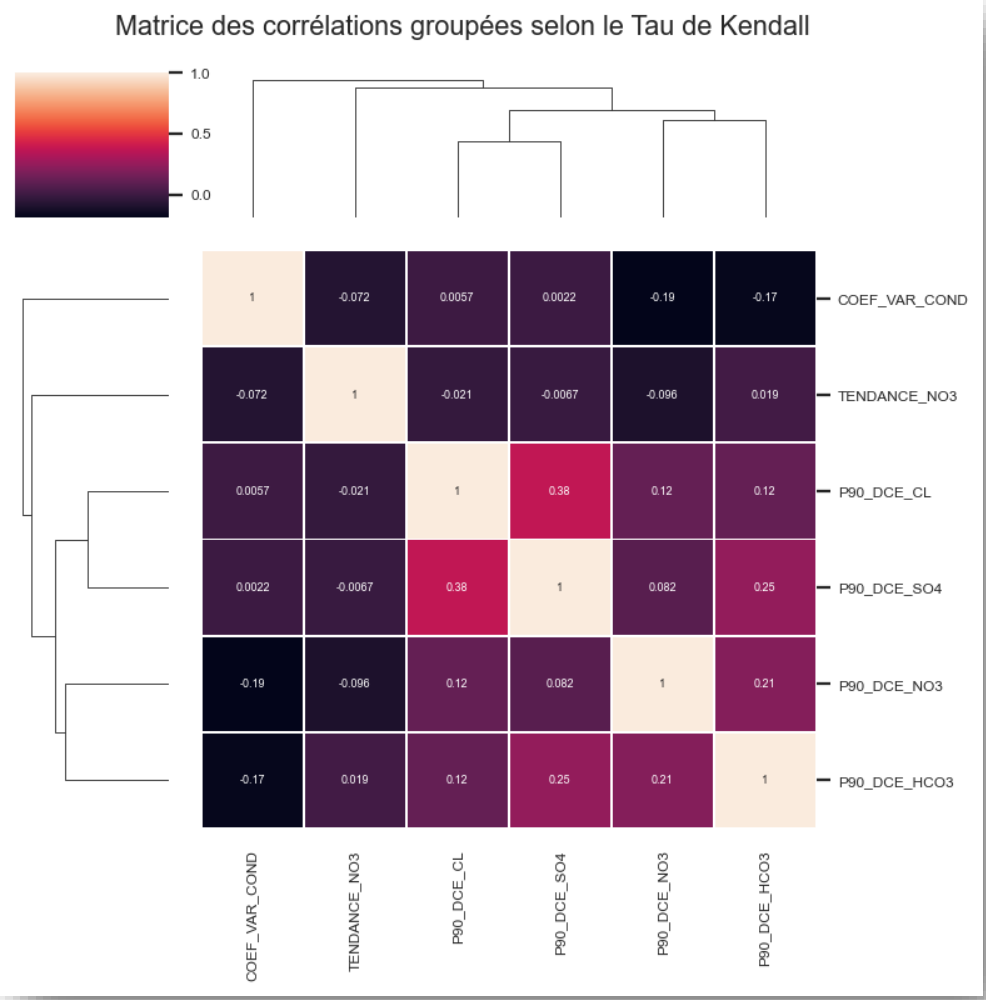
oui	oui	oui	oui	oui
BSS	BSS	ETUDE	ETUDE	ETUDE
Etiquette	Etiquette	Hydrogéologie	Hydrogéologie	Hydrogéologie
INDICE_BSS	LIBELLE_POINT	PROTECTION	PERMEABILITE	MILIEU
00406X0006	Alluvions de la Meuse à AUBRIVES/STATION DE POMPAGE	Libre	Aquifère	Poreux
00535X0003	Alluvions de la Meuse à LAIFOUR/PUITS ADDUCTION EAU	Libre	Aquifère	Poreux
00682X0010	/BELZY - L'ECHELLE	Libre	Aquifère	Mixte
00682X0027	CALCAIRES DU DOGGER DU B.P. À AUBIGNY-LES-POTHEES/FON	Libre	Aquifère	Poreux
00682X0028	CALC DOGGER COTES DE MEUSE ARDENNAISE A AUBIGNY-LES-I	Libre	Aquifère	Poreux
00683X0023	Calcaires du Dogger du B.P. à REMILLY-LES-POTHEES/PREMIER	Libre	Aquifère	Mixte
00684X0011	Grès du Lias inférieur d'Hettange à RENWEZ/LA GOULOTTE LON	Libre	Aquifère	Poreux
00684X0031	Calcaires du Dogger des côtes de Meuse ardennaise à SAINT-MA	Libre	Aquifère	Mixte
00687X0001	Calcaires du Dogger des côtes de Meuse ardennaise à VAUX-VIL	Libre	Aquifère	Poreux
00687X0002	Calc. Dogg. des côtes de Meuse ardennaise à SAINT-MARCEL/F	Libre	Aquifère	Mixte
00687X0003	CALCAIRES DU DOGGER DU B.P. À SAINT-MARCEL/ROUTE DE SU	Libre	Aquifère	Mixte
00687X0011	dennaise à CLAVY-WARBY/BC	Libre	Aquifère	Poreux
00687X0012		Libre	Aquifère	Poreux
00687X0013	Meuse ardennaise à THIS /FON	Libre	Aquifère	Mixte
00687X0014	dennaise à GUIGNICOURT/VE	Libre	Aquifère	Poreux
00687X0015	ENNAISE A GUIGNICOURT-SU	Libre	Aquifère	Poreux
00687X0016	Meuse ardennaise à GUIGNIC	Libre	Aquifère	Poreux
00688X0037	CALC DU GER DES COTES DE MEUSE À GUIGNICOURT SUR V	Libre	Aquifère	Poreux

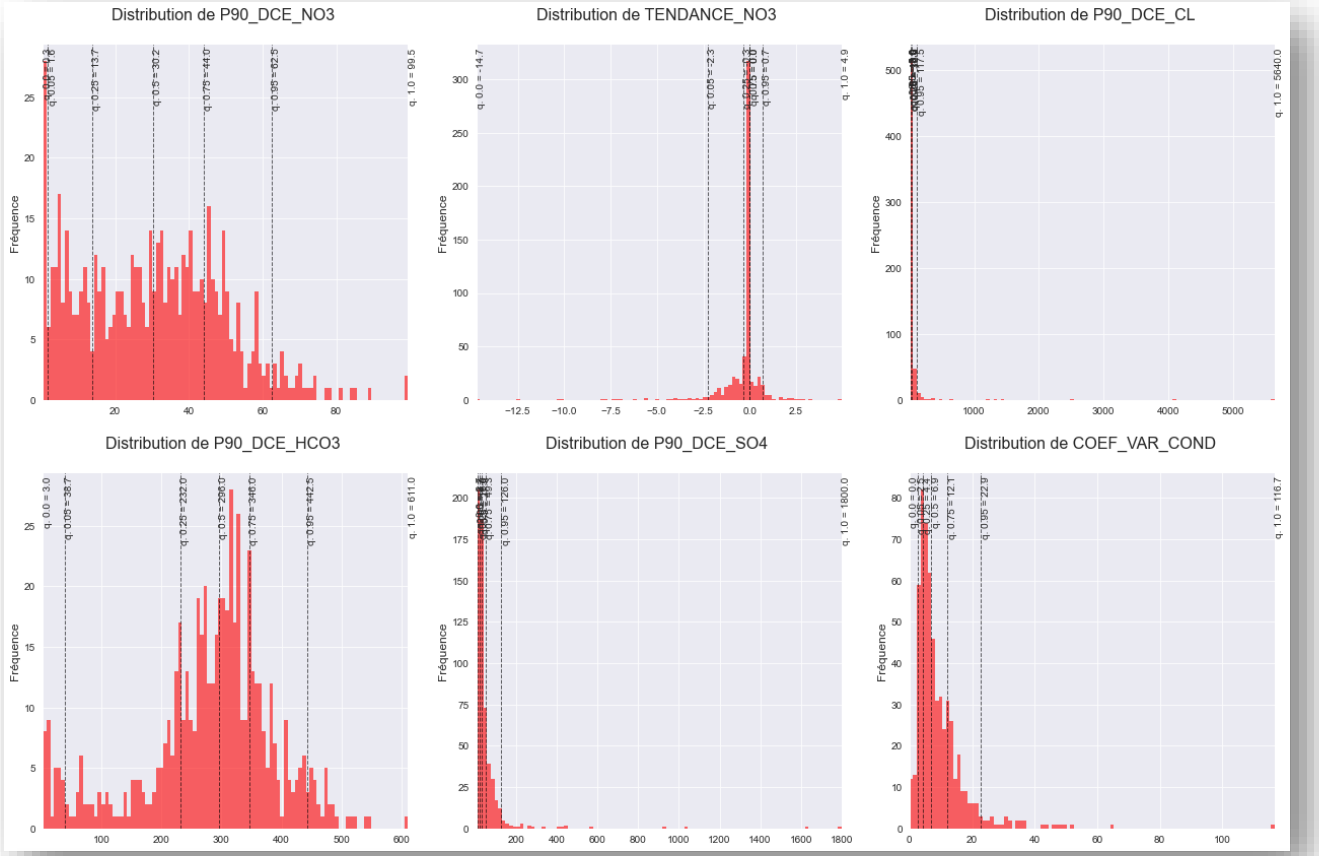
Variables descriptives des points d'eau

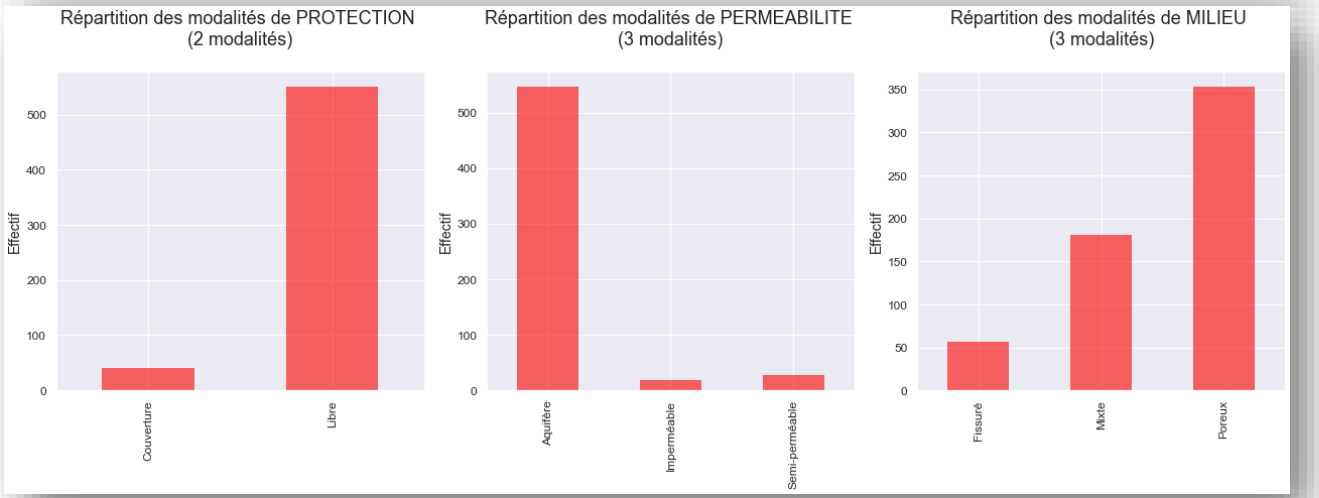
Thème	Indicateur	Signification	Origine	Méthode de calcul	Type
Protection de l'aquifère	PROTECTION	Libre / Couverture	BSS	Imputation manuelle valeurs manquantes	Catégoriel
Type d'écoulement souterrain	MILIEU	Poreux / Mixte / Fissuré	BDLISA	Imputation manuelle valeurs manquantes	Catégoriel
	PERMEABILITE	Aquifère / Semi-perméable / Imperméable	BDLISA	Imputation manuelle valeurs manquantes	Catégoriel
	COEF_VAR_COND	Coefficient de variation de la conductivité	ADES	Statistiques classiques QUALISTAT	Quantitatif
Faciès chimique naturel	P90_DCE_HCO3	Quantile 0,9 concentration en hydrogénocarbonates	ADES	Statistiques classiques QUALISTAT	Quantitatif
Faciès chimique naturel et pression industrielle minière	P90_DCE_SO4	Quantile 0,9 concentration en sulfates	ADES	Statistiques classiques QUALISTAT	Quantitatif
	P90_DCE_CL	Quantile 0,9 concentration en chlorures	ADES	Statistiques classiques QUALISTAT	Quantitatif
Pression agricole	P90_DCE_NO3	Quantile 0,9 concentration en nitrates	ADES	Statistiques classiques QUALISTAT	Quantitatif
	TENDANCE_NO3	Pente de la tendance linéaire nitrates	ADES	Détection et calculs de pente HYPE	Quantitatif

EXEMPLE DE REGROUPEMENT

Propriétés statistiques des variables extraites







EXEMPLE DE REGROUPEMENT

Résultats de l'ensemble des tests

Test de 8 combinaisons d'approches de gestion des données mixtes et d'algorithmes de regroupement

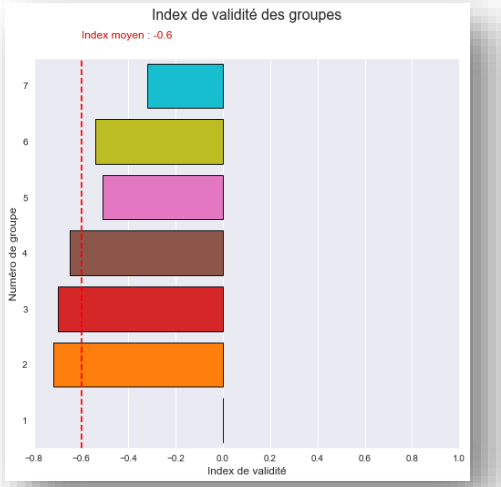
- Le nombre de groupes identifiés varie entre **7 et 9 groupes.**
- Le meilleur regroupement a été obtenu par la combinaison **AFDM - HDBSCAN avec distance euclidienne.**
- Tous les autres tests ont **au moins un critère d'évaluation déclassant, ou un nombre de points atypiques trop élevé** (en rouge dans le tableau).
- Cette analyse est confirmée par les visualisations subjectives.

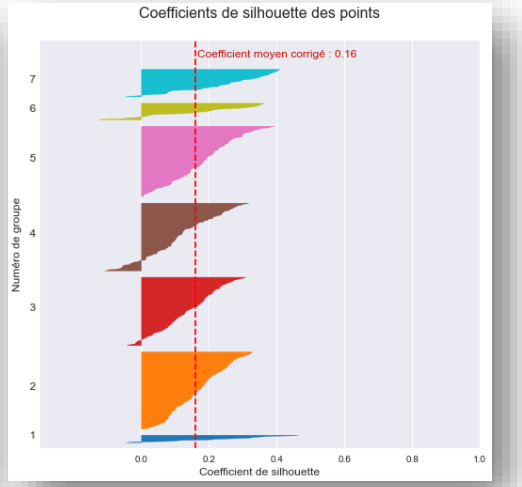
Approches de gestion des jeux de données mixtes		Algorithmes de regroupement	
Type d'approche	**Distance associée**	**K-means**	**HDBSCAN**
Approche n°1 : encodage 0/1 des variables catégorielles	**Distance euclidienne** : adaptée aux variables quantitatives	**K1** **7** groupes Val = **-0,6** / Sil = **0,16**	**H1** **8** groupes **111** points atypiques Val = **0,06** / Sil = **0,13**
	Distance Manhattan : adaptée aux variables quantitatives	-	**H2** **9** groupes **81** points atypiques Val = **0,02** / Sil = **0,12**
	Distance Sorensen-Dice : adaptée aux variables encodées 0/1	-	**H3** **9** groupes **4** points atypiques Val = **-0,59** / Sil = **0,05**
Approche n°2 : discrétisation des variables quantitatives	**Distance Sorensen-Dice** : adaptée aux variables encodées 0/1	-	**H4** **9** groupes **4** points atypiques Val = **-0,53** / Sil = **0,07**
Approche n°3 : métrique de distance hybride Gower	**Distance Gower** : moyenne pondérée Manhattan / Sorensen-Dice	-	**H5** **9** groupes **4** points atypiques Val = **-0,36** / Sil = **0,06**
Approche n°4 : AFDM sans réduction de dimensions	**Distance euclidienne** : adaptée aux variables quantitatives	**K2** **8** groupes Val = **-0,29** / Sil = **0,22**	**H6** **9** groupes **23** points atypiques Val = **-0,02** / Sil = **0,35**

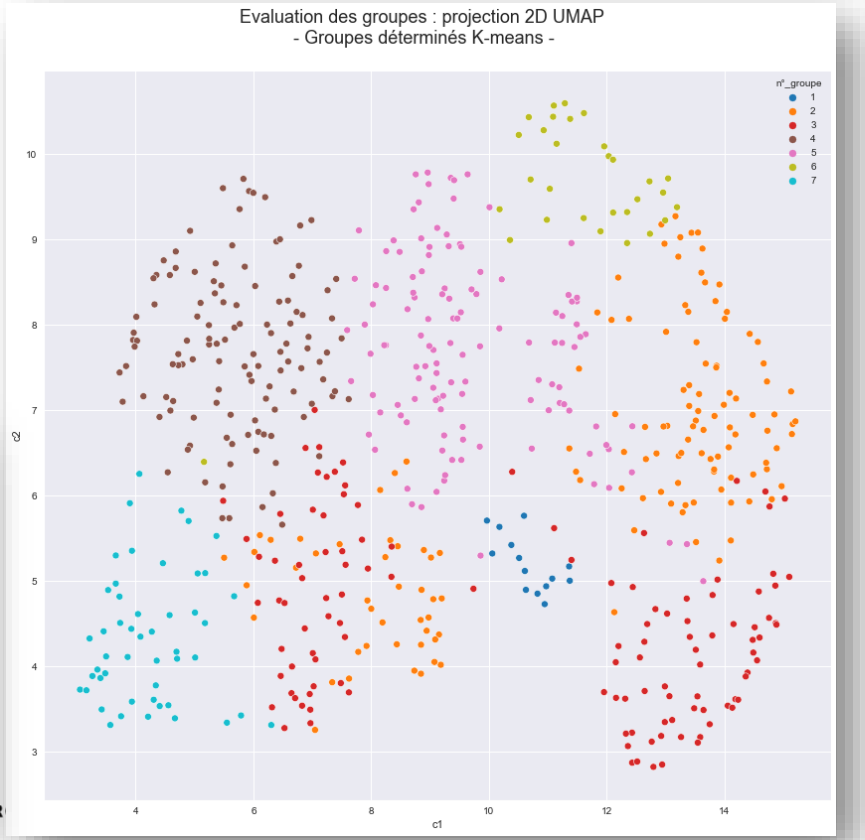
EXEMPLE DE REGROUPEMENT

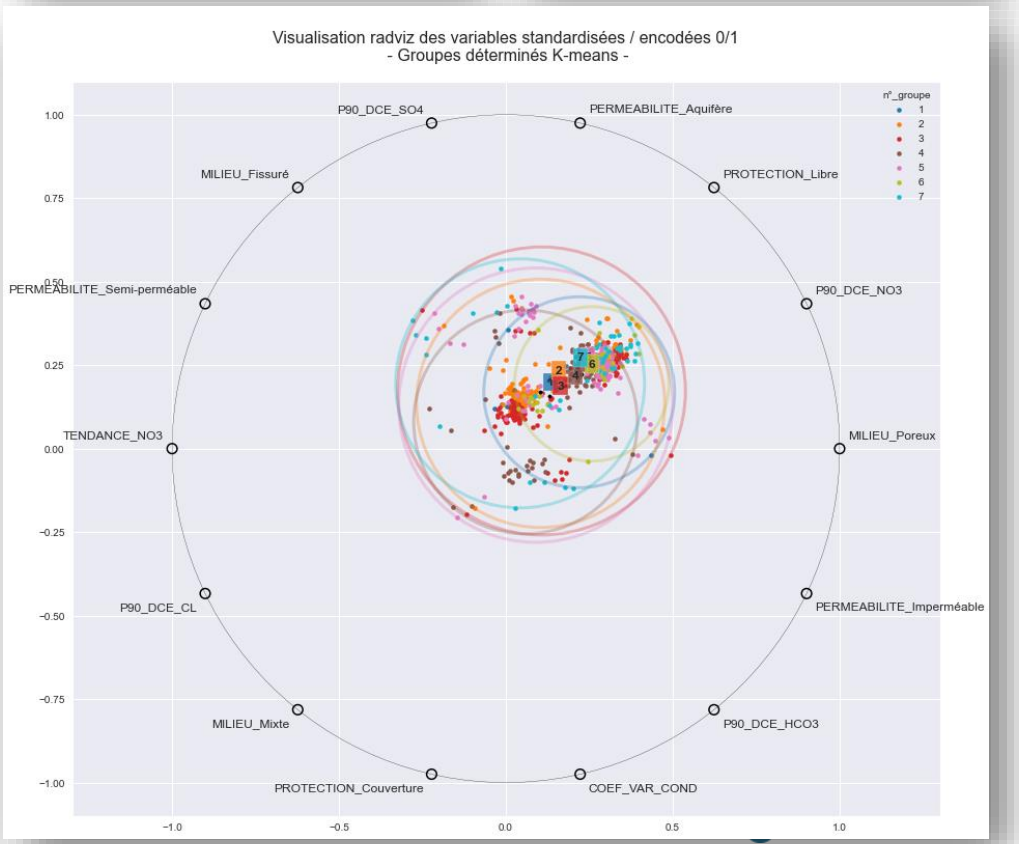
Résultats du test de référence (K1) : encodage 0/1 – K-means

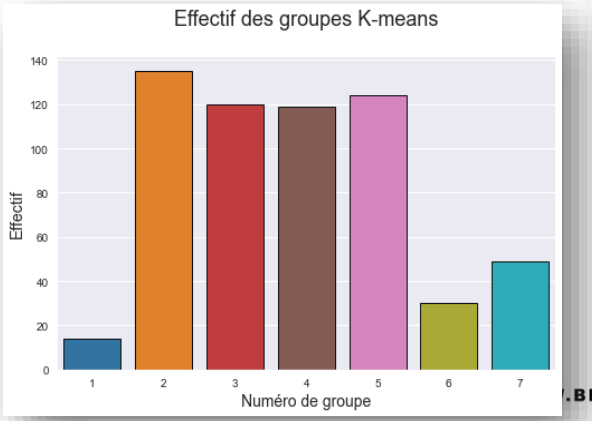
- La combinaison K1 est **incapable d'identifier une structure dans les données**.
- Même combiné avec une AFDM (test K2), **K-means n'est pas assez performant**.









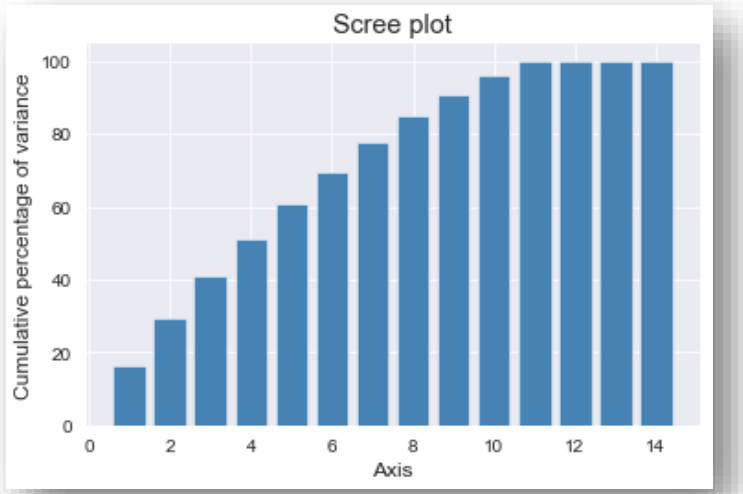


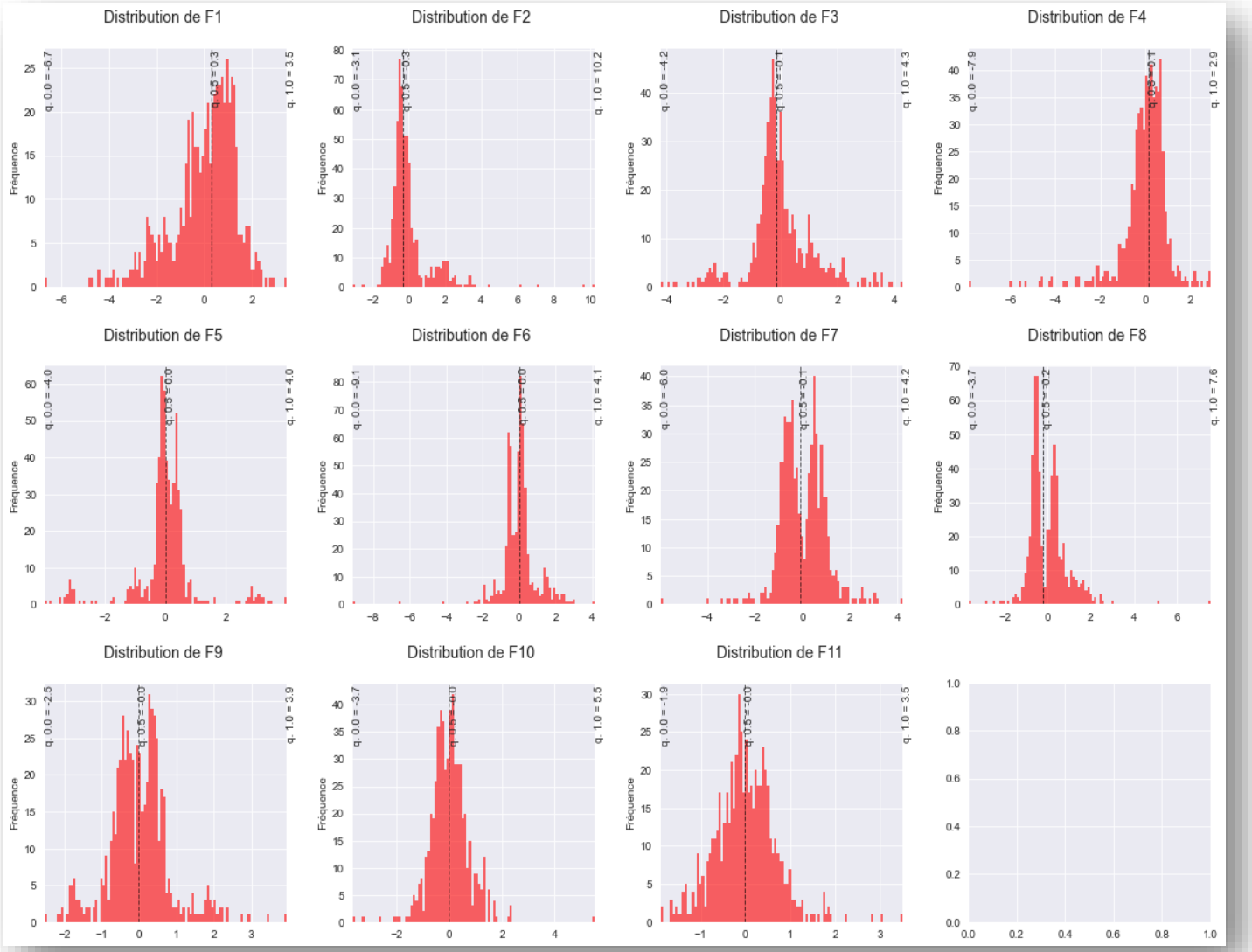
EXEMPLE DE REGROUPEMENT

Résultats du meilleur test (H6) : AFDM - HDBSCAN

Transformation des données par l'AFDM

- **On met à profit la capacité de l'AFDM d'équilibrer l'influence des variables**, sans réduire la dimension du jeu de données.
- On retient donc les **11 premiers facteurs qui expliquent 100% de la variabilité** (les 3 derniers correspondent aux variables indicatrices redondantes).



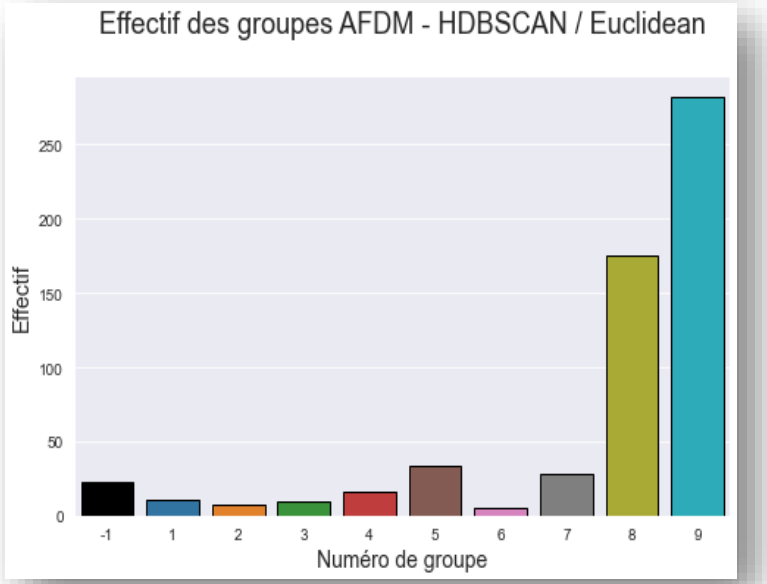


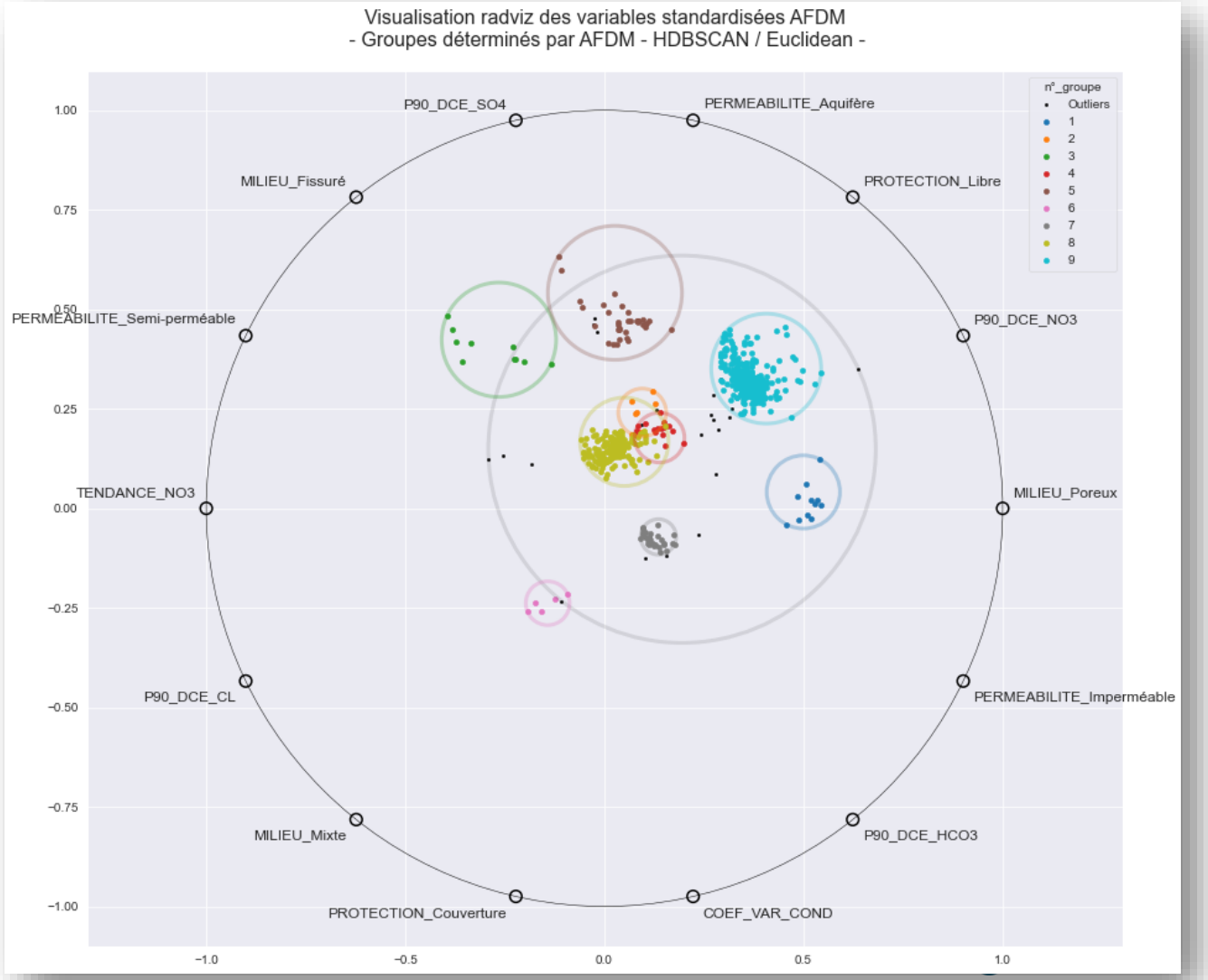
EXEMPLE DE REGROUPEMENT

Résultats du meilleur test (H6) : AFDM - HDBSCAN

Evaluation des groupes

- **9 groupes** dont **7 petits** (5 à 34 points) et **2 très grands** (175 et 282 points).
- **23 points atypiques** (« outliers », groupe n°-1).



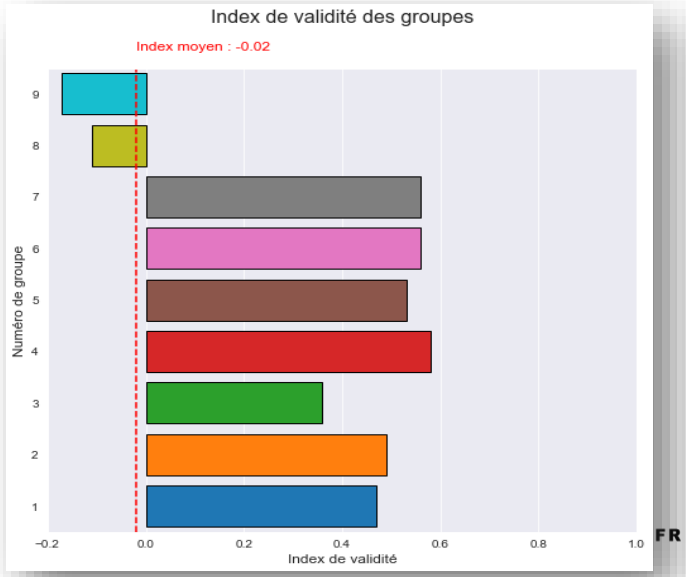


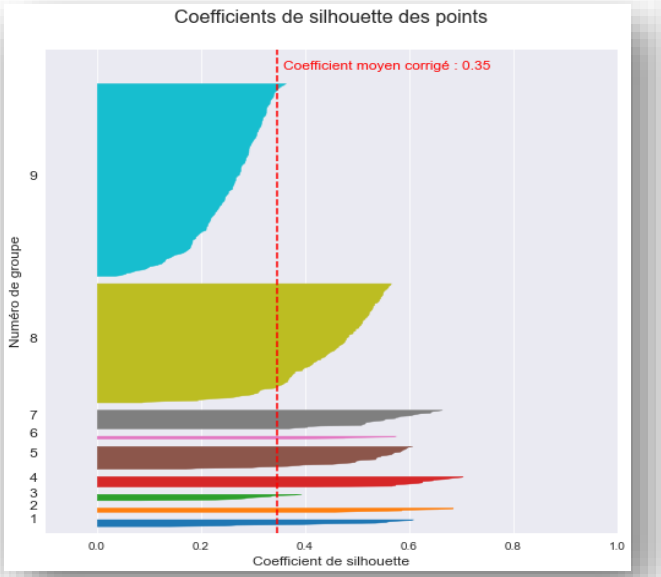
EXEMPLE DE REGROUPEMENT

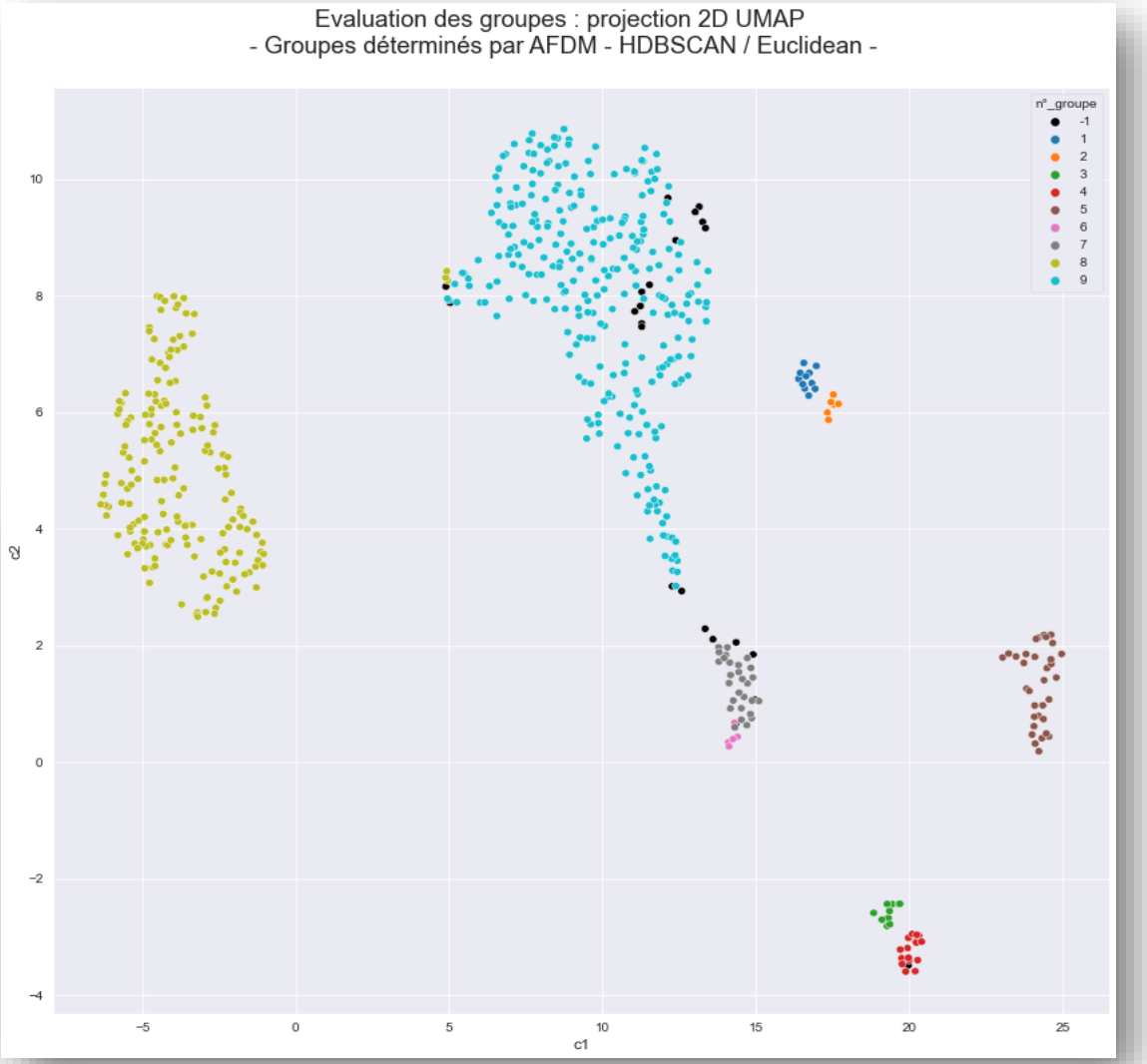
Résultats du meilleur test (H6) : AFDM - HDBSCAN

Evaluation des groupes (suite)

- **Bonne séparation des groupes sur les visualisations** RADVIZ et UMAP…
- **…mais quelques points du groupe n°8 sont mélangés au groupe n°9** sur la projection 2D UMAP…
- **…et l'index de validité**, bon pour les 7 petits groupes, est mauvais pour les 2 grands groupes.



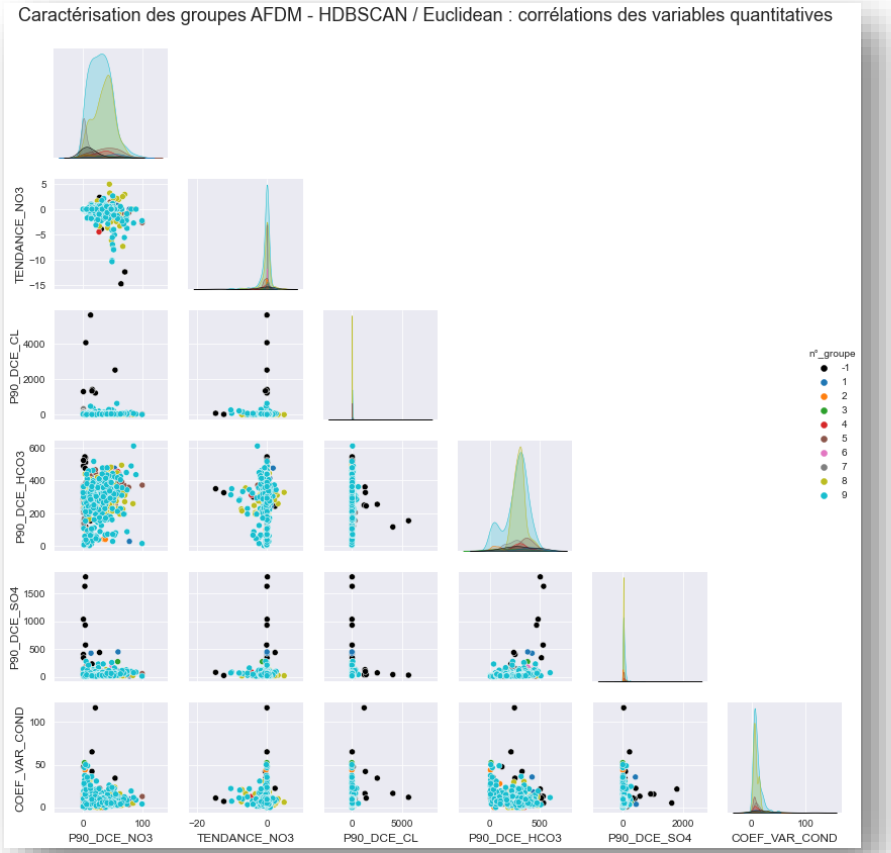


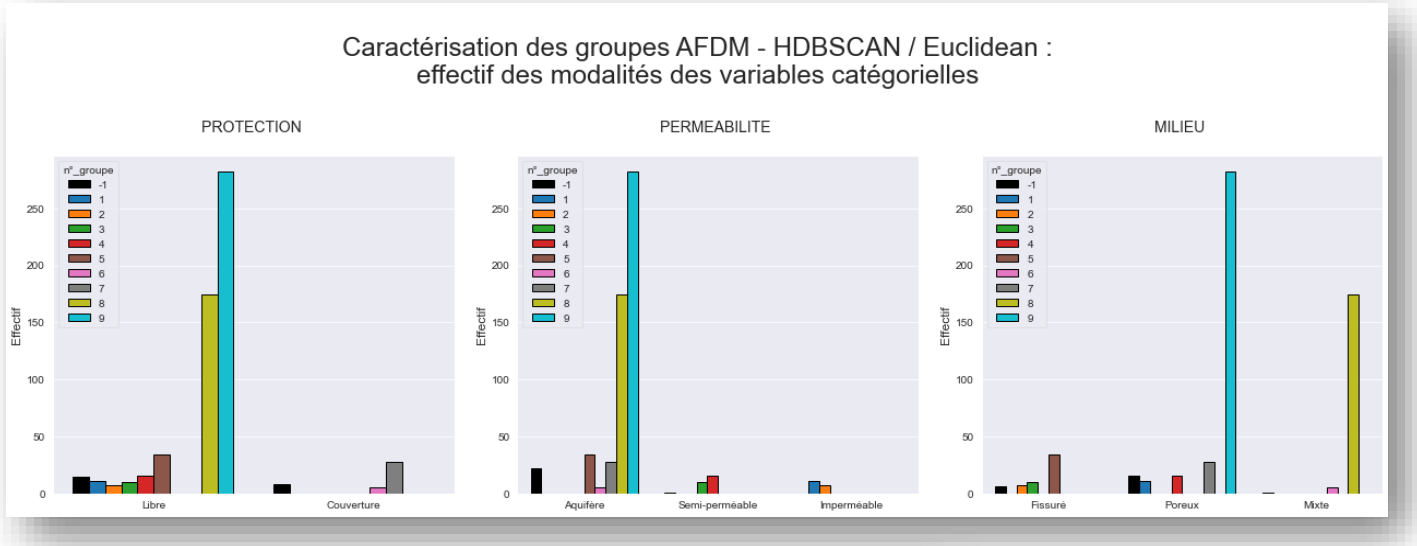


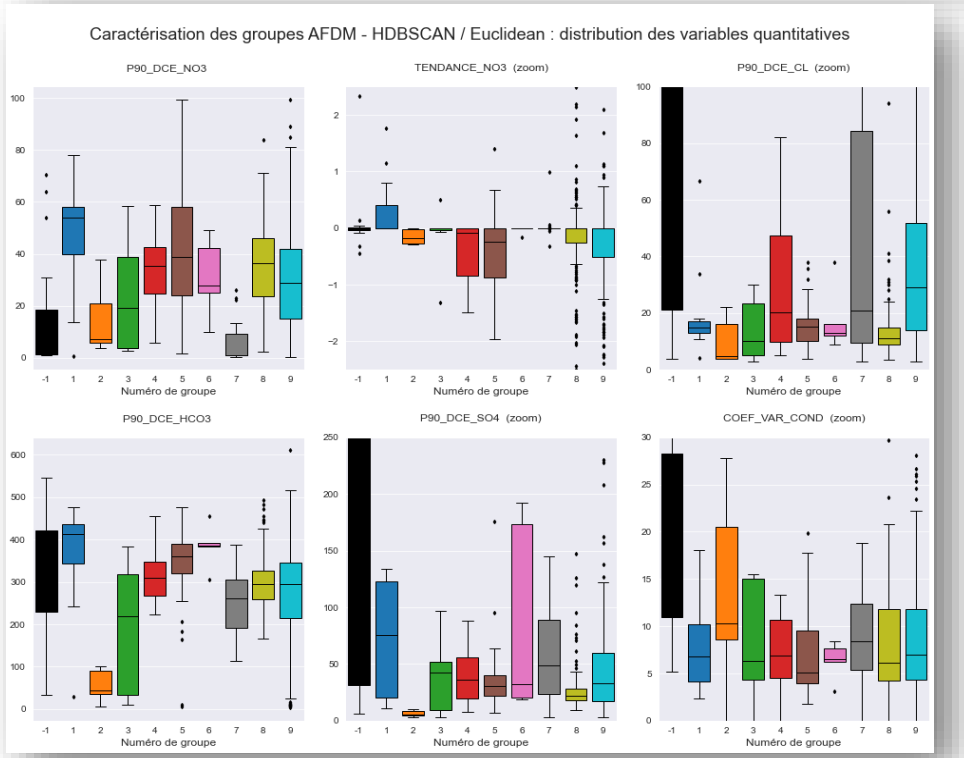
EXEMPLE DE REGROUPEMENT

Résultats du meilleur test (H6) : AFDM - HDBSCAN

Graphiques de caractérisation des groupes







Géosciences pour une Terre durable
brgm

EXEMPLE DE REGROUPEMENT

Résultats du meilleur test (H6) : AFDM - HDBSCAN

Interprétation des groupes

- **Les 3 variables catégorielles dominent** la formation des groupes.
- **Une liaison forte existe** entre certaines combinaisons de modalités et 2 à 5 variables quantitatives.
- **L'interprétation des groupes amène à proposer des actions d'amélioration du regroupement** : corrections d'erreurs, ajout ou remplacement de variables, regroupement distinct sur les grands groupes.

Groupe	Nombre de points d'eau	Index de validité	Modalités discriminantes	Variables quantitatives discriminantes	Commentaire	Action d'amélioration du regroupement
2	7	0,49	Libre Fissuré Imperméable	NO3 très faible TENDANCE NO3 à la baisse HCO3 très faible SO4 très faible COEF_VAR_COND très fort	Milieu vulnérable fissuré peu perméable, sans pression agricole	
1	11	0,47	Libre Poreux Imperméable	NO3 très fort TENDANCE NO3 à la hausse HCO3 très fort	Milieu vulnérable poreux peu perméable, pression agricole forte à la hausse	
3	10	0,36	Libre Fissuré Semi-perméable	NO3 moyen à faible TENDANCE NO3 stable HCO3 très dispersé	Milieu vulnérable fissuré semi-perméable, pression agricole moyenne stable	
4	16	0,58	Libre Poreux Semi-perméable	NO3 moyen TENDANCE NO3 à la baisse	Milieu vulnérable poreux semi-perméable, pression agricole moyenne en diminution	
5	34	0,53	Libre Fissuré Aquifère	NO3 moyen à fort TENDANCE NO3 à la baisse HCO3 fort	Milieu vulnérable fissuré aquifère, pression agricole forte en diminution	
8	175	-0,11	Libre Mixte Aquifère		Milieu vulnérable mixte aquifère, grand groupe aux caractéristiques moyennes	Introduire de nouvelles variables discriminantes et/ou faire un regroupement distinct pour ce groupe
9	282	-0,17	Libre Poreux Aquifère		Milieu vulnérable poreux aquifère, très grand groupe aux caractéristiques moyennes	Introduire de nouvelles variables discriminantes et/ou faire un regroupement distinct pour ce groupe
6	5	0,56	Couverture Aquifère Mixte	NO3 fort HCO3 très fort	Milieu non vulnérable mixte aquifère : incohérence entre la protection par une couverture et NO3 fort	Variable PROTECTION peu pertinente à améliorer ou remplacer
7	28	0,56	Couverture Aquifère Poreux	NO3 très faible HCO3 faible CL très dispersé	Milieu non vulnérable poreux aquifère avec contamination naturelle possible par du sel	
-1	23	Non calculable		NO3 très faible SO4 moyen à très fort CL moyen à très fort COEF_VAR_COND très fort TENDANCE NO3 très négative	- Pression minière avec écoulements rapides (galeries, fractures) - Erreurs de calcul automatique de la tendance	Corriger les erreurs de calcul automatique de la tendance


MERCI POUR VOTRE ATTENTION !
Géosciences pour une Terre durable
brgm