



**BORDEAUX
SCIENCES
AGRO**



Soutien à la diffusion
des Référentiels Régionaux Pédologiques

Soutenance de mémoire de
pré-spécialisation

UTILISATION D'OUTILS NUMÉRIQUES POUR LA VALORISATION DE DONNÉES PÉDOLOGIQUES

Paul Vilvandré

NumAg : Numérique pour l'Agriculture

Tutrice de stage : **Stéphanie Jalabert**, équipe Sols, Bordeaux Sciences Agro

Le Gis Sol

Pilote des programmes de cartographie en France

Objectifs :

- Constitution et gestion d'un système d'informations sur les sols de France
- Mieux comprendre les sols pour mieux les gérer



Le Gis Sol

Pilote des programmes de cartographie en France

Objectifs :

- Constitution et gestion d'un système d'informations sur les sols de France
- Mieux comprendre les sols pour mieux les gérer

Pourquoi les sols ?

- Principal support des activités humaines
- Soumis à des menaces (érosion, pollutions...)
- Ressource non renouvelable à l'échelle humaine



Le Gis Sol

Pilote des programmes de cartographie en France

Objectifs :

- Constitution et gestion d'un système d'informations sur les sols de France
- Mieux comprendre les sols pour mieux les gérer



Pourquoi les sols ?

- Principal support des activités humaines
- Soumis à des menaces (érosion, pollutions...)
- Ressource non renouvelable à l'échelle humaine

IGCS	RMQS	BDAT	BDETM
Inventaire, Gestion et Conservation des Sols	Réseau de Mesures de la Qualité des Sols	Banque de Données des Analyses de Terre	Base de Données Eléments Traces Métalliques

IGCS

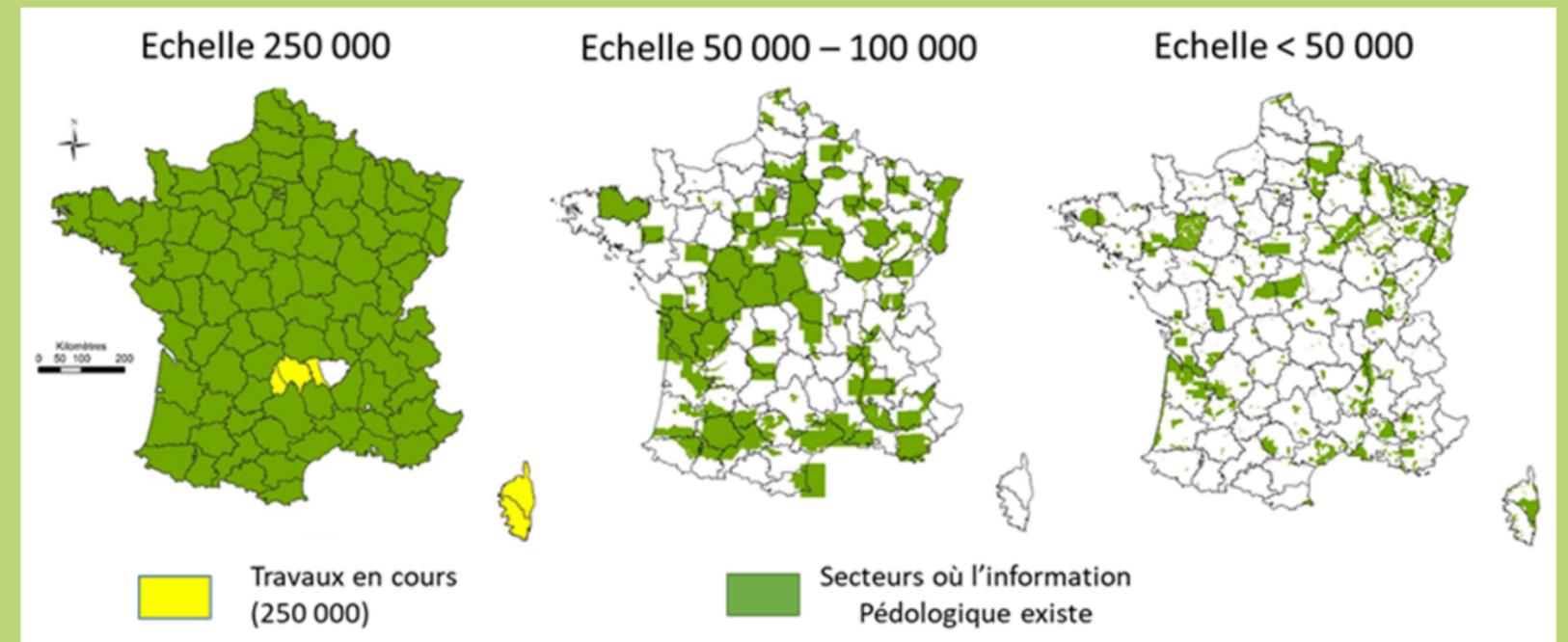
Un programme multi-échelles

Inventaire, Gestion et Conservation des Sols
Cartographie d'inventaire de la **nature des sols** et de leur **distribution spatiale** sur un territoire

Référentiel Régional
Pédologique (RRP)

Volet Connaissance
Pédologiques de la
France (CPF)

Secteurs de
Référence (SR)



IGCS

Un programme multi-échelles

Inventaire, Gestion et Conservation des Sols
Cartographie d'inventaire de la **nature des sols** et de leur **distribution spatiale** sur un territoire

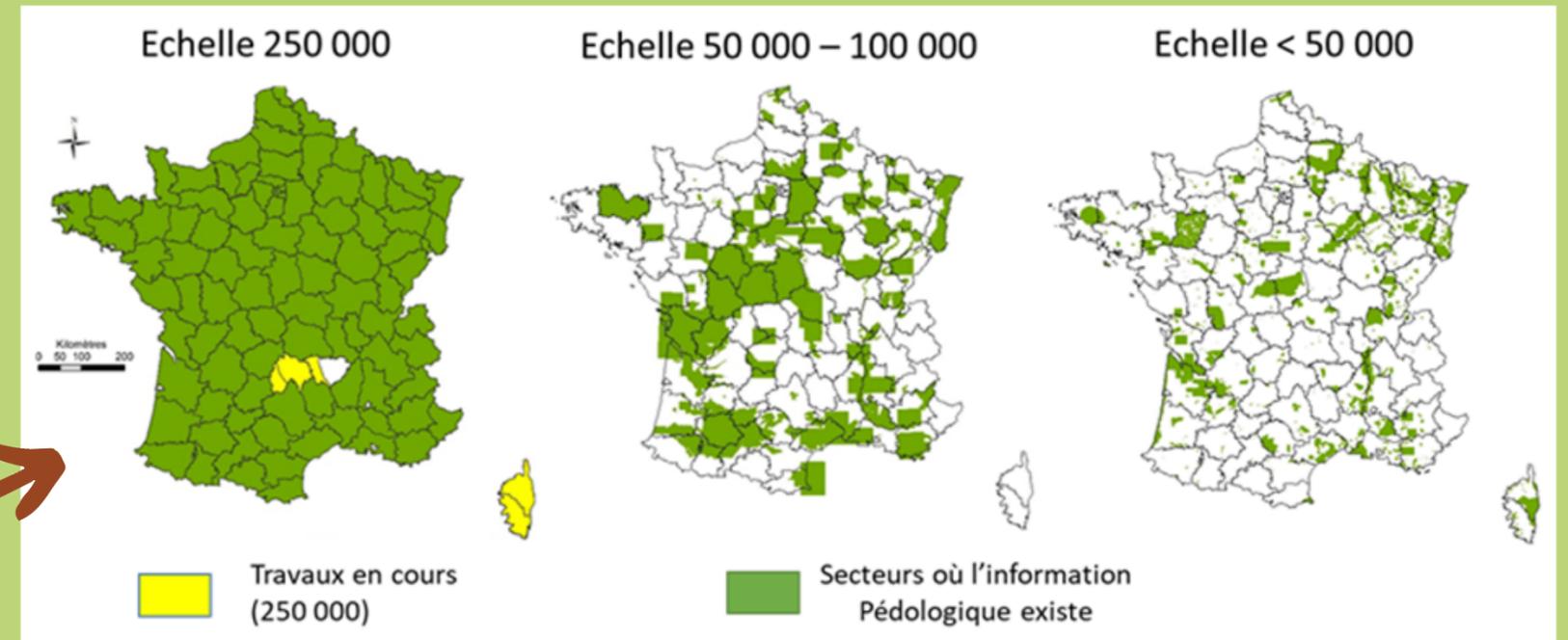
Le volet RRP couvre presque tout le territoire aujourd'hui :

71 RRP finalisés en France métropolitaine

Référentiel Régional Pédologique (RRP)

Volet Connaissance Pédologiques de la France (CPF)

Secteurs de Référence (SR)



Qu'en est-il aujourd'hui de la diffusion de ces RRP sur Internet ?

La diffusion des RRP

A l'échelle nationale : Géoportail

Première valorisation des RRP :
Carte des Sols dominants

En cliquant sur la carte :

Accès aux infos :

- Auteur(s)
- Modalités d'accès aux RRP (notamment **fiches de métadonnées**)

The screenshot displays the Géoportail interface. At the top, there is a search bar with the text "Chercher un lieu, une adresse, une donnée" and a plus sign. The main map area shows a colorful soil map of France. A popup window titled "UNITÉ CARTOGRAPHIQUE DE SOL (UCS) N°61" is open, providing the following information:

- Nom de l'UCS : LANDE MESOPHILE
- Accès à la fiche de l'UCS : [cliquez ici](#)
- Type de sol dominant : Podzolsols (78 %)
- Pour plus d'informations sur ce type de sol : [cliquez ici](#)
- Référence de l'étude : D. ARROUAYS, J. WILBERT, A. RICHER-DE-FORGES, 2015. Référentiel Régional Pédologique d'Aquitaine. Carte des pédopaysages des départements des Landes et de la Gironde (Etude n°30111)
- État de l'étude : Finalisée et labellisée
- Responsable technique pour l'étude : Anne RICHER-DE-FORGES
- Structure(s) propriétaire(s) de l'étude : Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE), US 1106
- Infosol, Orléans
- Pour plus d'informations : [cliquez ici](#)

The popup is connected to the map by a brown arrow. The interface also includes a sidebar on the right with "CARTE EN COURS", "Carte des sols", and buttons for "+ DE DONNÉES" and "ENREGISTRER LA CARTE". At the bottom, there is a scale bar (0-100 km) and a small inset map of Europe.

La diffusion des RRP

A l'échelle nationale : Géoportail

Première valorisation des RRP :
Carte des Sols dominants

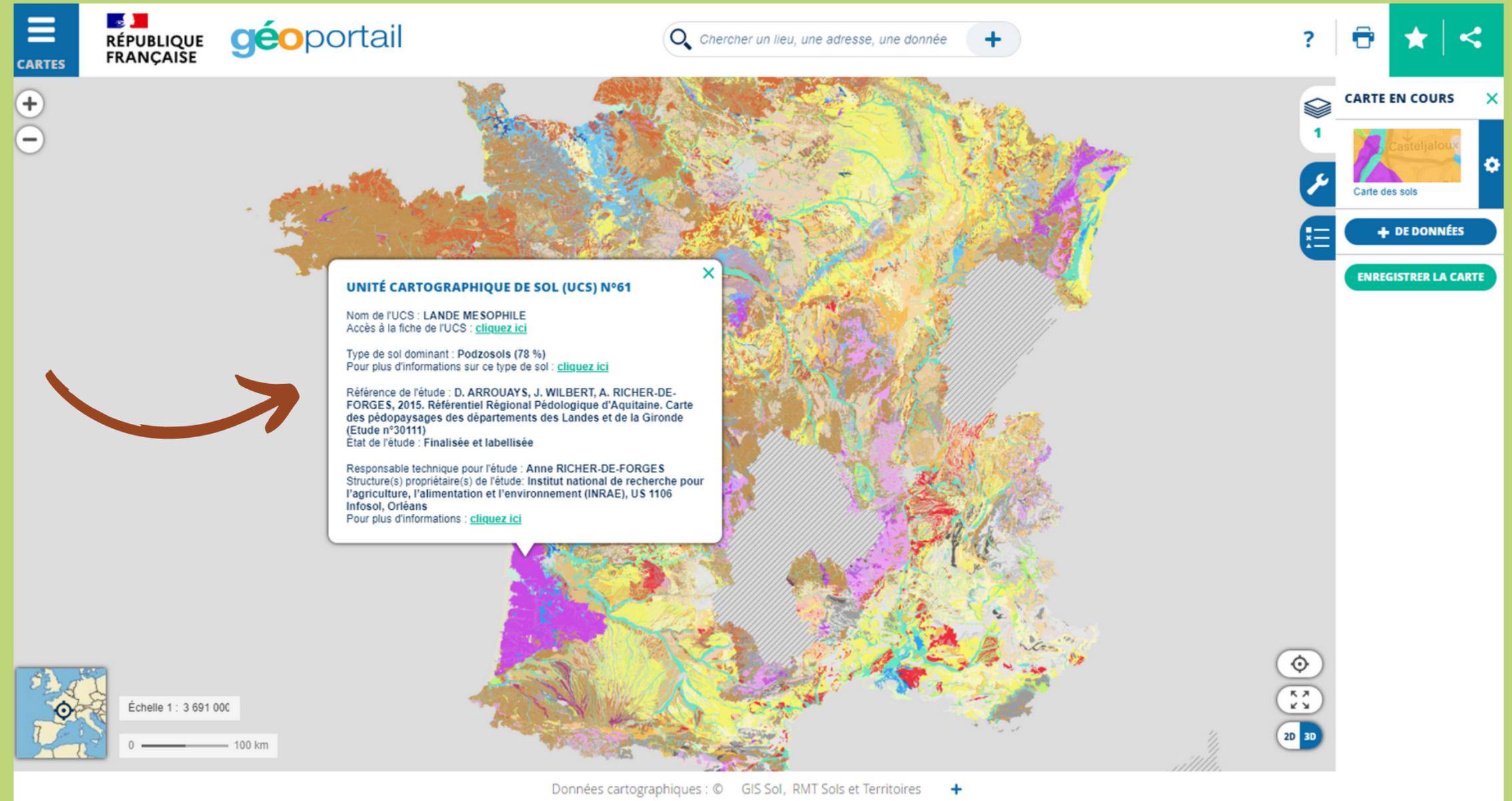
En cliquant sur la carte :

Accès aux infos :

- Auteur(s)
- Modalités d'accès aux RRP (notamment **fiches de métadonnées**)

Principal problème :

RRP **pas tous recensés** en métadonnées



La diffusion des RRP

A l'échelle nationale : Géoportail

Première valorisation des RRP :
Carte des Sols dominants

En cliquant sur la carte :

Accès aux infos :

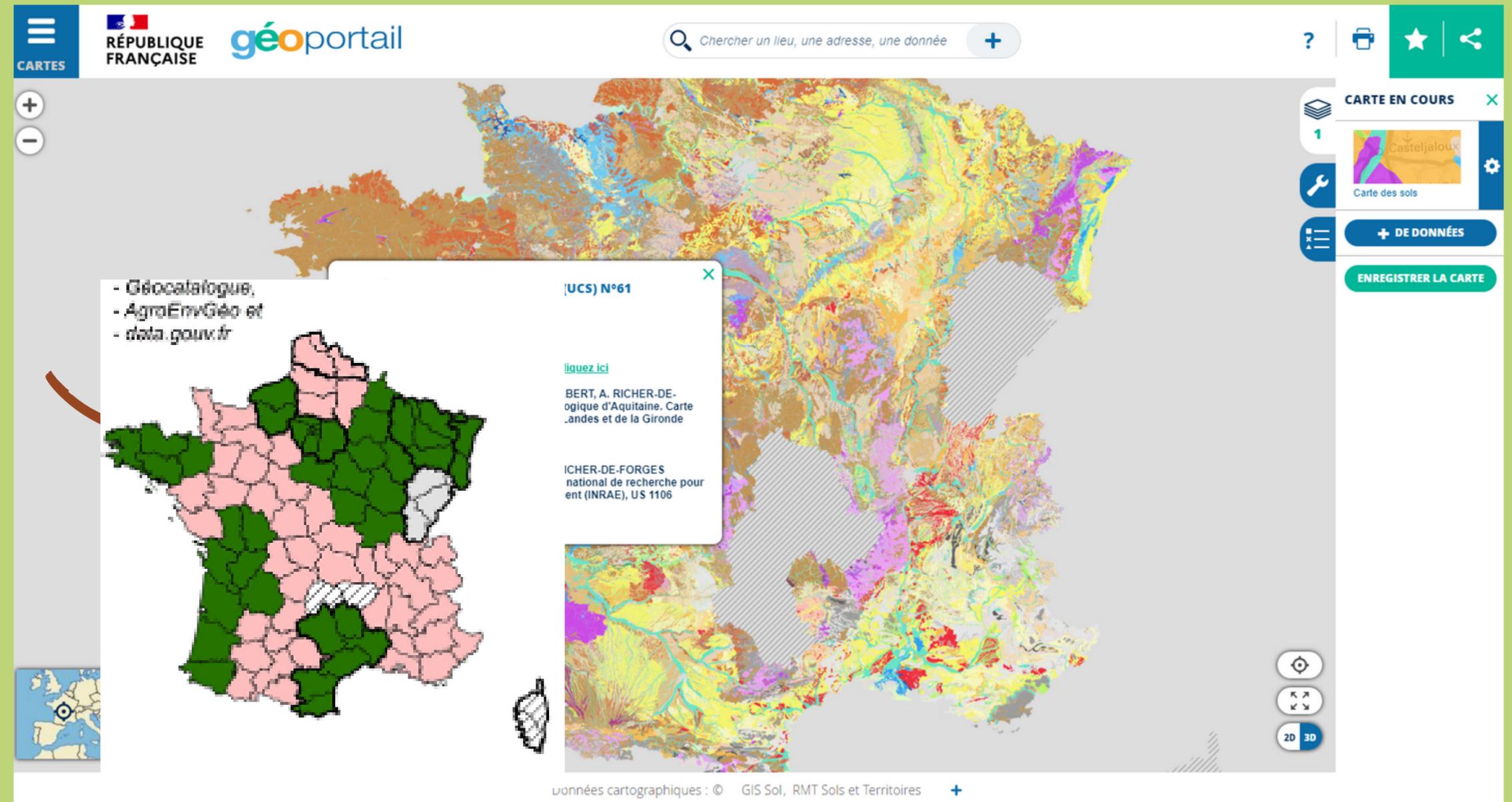
- Auteur(s)
- Modalités d'accès aux RRP (notamment **fiches de métadonnées**)

Principal problème :

RRP **pas tous recensés** en métadonnées

Premier constat :

Seulement 27 fiches de métadonnées pour 71 RRP finalisés



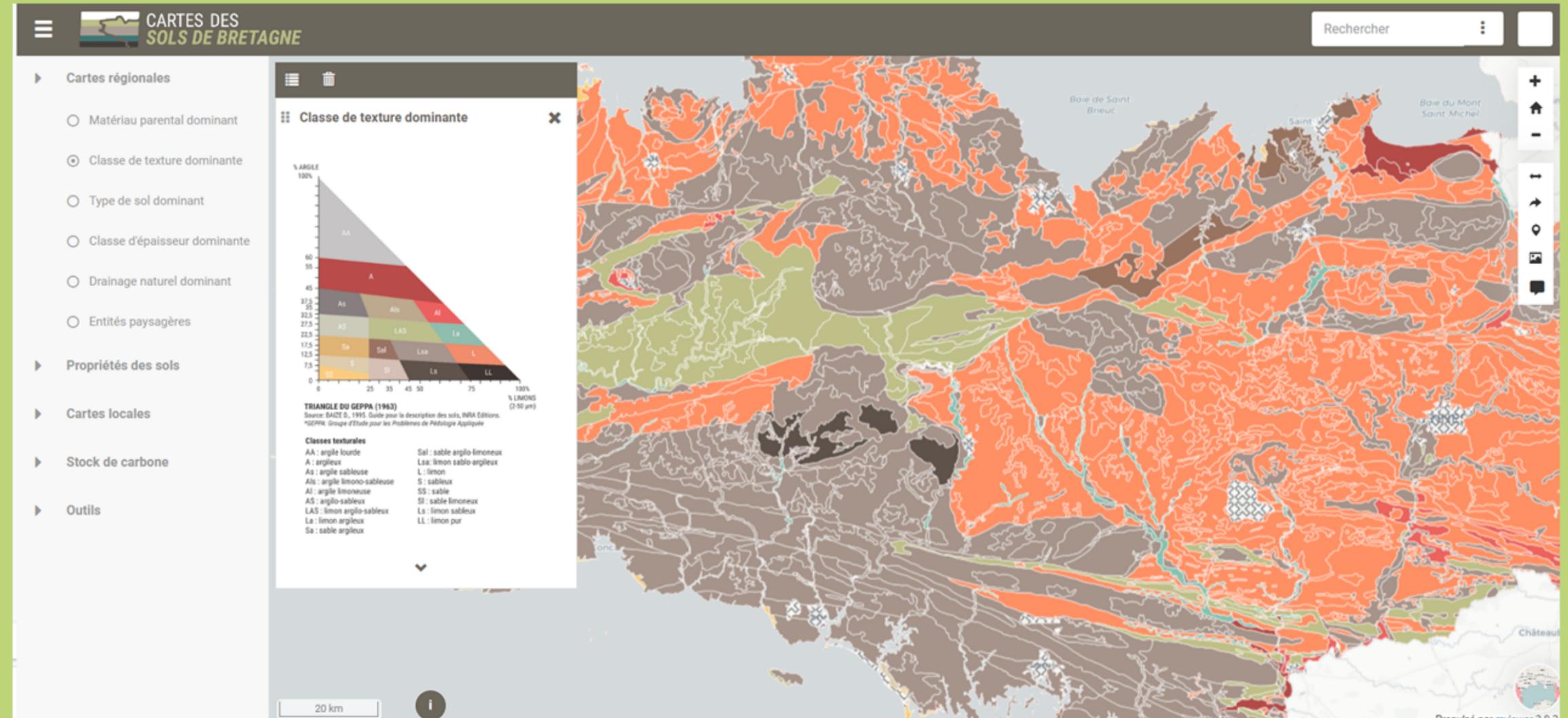
La diffusion des RRP

A l'échelle régionale

Portail thématique Sols de Bretagne : Visionneuse Mviewer

Région la plus avancée sur les supports de valorisation et de transfert de connaissances sur les sols : Bretagne

6 cartes thématiques, 9 propriétés des sols (pH, carbone, érosion...), ...



La diffusion des RRP

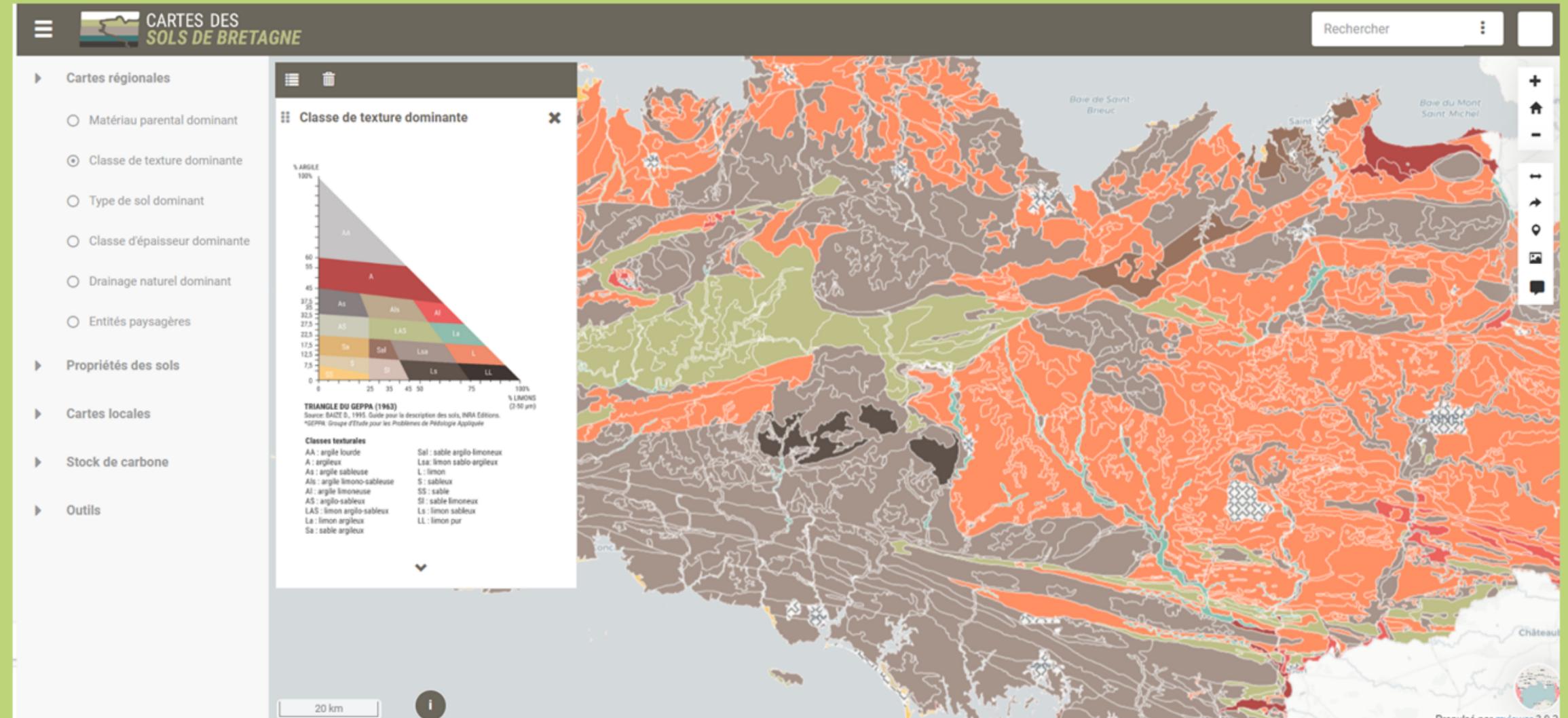
A l'échelle régionale

Portail thématique Sols de Bretagne : Visionneuse Mviewer

Région la plus avancée sur les supports de valorisation et de transfert de connaissances sur les sols : Bretagne

6 cartes thématiques, 9 propriétés des sols (pH, carbone, érosion...), ...

Second constat :
Sols de Bretagne = exemple à suivre mais comment le développer aux autres régions ?



Le projet Solid'R

Pour Soutenir et améliorer la diffusion
des RRP

Appel à projet IGCS, porté par BSA
Avril 2022 – Décembre 2023



BORDEAUX
SCIENCES
AGRO

- Objectif 1 : Accompagnement des partenaires régionaux IGCS dans la mise en conformité et la diffusion des métadonnées des RRP
- Objectif 2 : Aide à la valorisation des données issues des RRP

Le projet Solid'R

Pour Soutenir et améliorer la diffusion des RRP

Appel à projet IGCS, porté par BSA
Avril 2022 – Décembre 2023



BORDEAUX
SCIENCES
AGRO

- Objectif 1 : Accompagnement des partenaires régionaux IGCS dans la mise en conformité et la diffusion des métadonnées des RRP
- Objectif 2 : Aide à la valorisation des données issues des RRP

Phase 1 :
Analyse des opportunités existantes
pour **structurer** et harmoniser la
diffusion des données RRP
Etudiants Master SIGAT



Relais central de la diffusion :
GIS SOL



Pour convaincre, il faut montrer :
exemples de valorisation des
données brutes complexes

Le projet Solid'R

Pour Soutenir et améliorer la diffusion des RRP

Appel à projet IGCS, porté par BSA
Avril 2022 – Décembre 2023



BORDEAUX
SCIENCES
AGRO

- Objectif 1 : Accompagnement des partenaires régionaux IGCS dans la mise en conformité et la diffusion des métadonnées des RRP

Phase 1 :

Analyse des opportunités existantes pour **structurer** et harmoniser la **diffusion des données RRP**
Etudiants Master SIGAT

Relais central de la diffusion :
GIS SOL

Pour convaincre, il faut montrer :
exemples de valorisation des données brutes complexes

- Objectif 2 : Aide à la valorisation des données issues des RRP

Phase 2 :

Guider les gestionnaires RRP en utilisant des outils numériques

 **Mon stage**

Les objectifs du stage

- Construction d'un script permettant de calculer des indicateurs de propriétés des sols à partir des données disponibles dans les bases de données DoneSol
- Développer une première version d'un visualiseur cartographique sous Mviewer "Sols de Nouvelle-Aquitaine"

Enjeux :

- Partager la prise en main des outils en accompagnant les utilisateurs
- Gagner du temps dans la valorisation des données

Plan de la présentation

- Construction d'un script permettant de calculer des indicateurs de propriétés des sols à partir des données disponibles dans les bases de données DoneSol
- Développer une première version d'un visualiseur cartographique sous Mviewer "Sols de Nouvelle-Aquitaine"

Plan de la présentation

- Construction d'un script permettant de calculer des indicateurs de propriétés des sols à partir des données disponibles dans les **bases de données DoneSol**
- Développer une première version d'un visualiseur cartographique sous Mviewer "Sols de Nouvelle-Aquitaine"

I/ La base de données DoneSol

Plan de la présentation

- **Construction d'un script** permettant de calculer des indicateurs de propriétés des sols à partir des données disponibles dans les bases de données DoneSol
- Développer une première version d'un visualiseur cartographique sous Mviewer "Sols de Nouvelle-Aquitaine"

I/ La base de données DoneSol

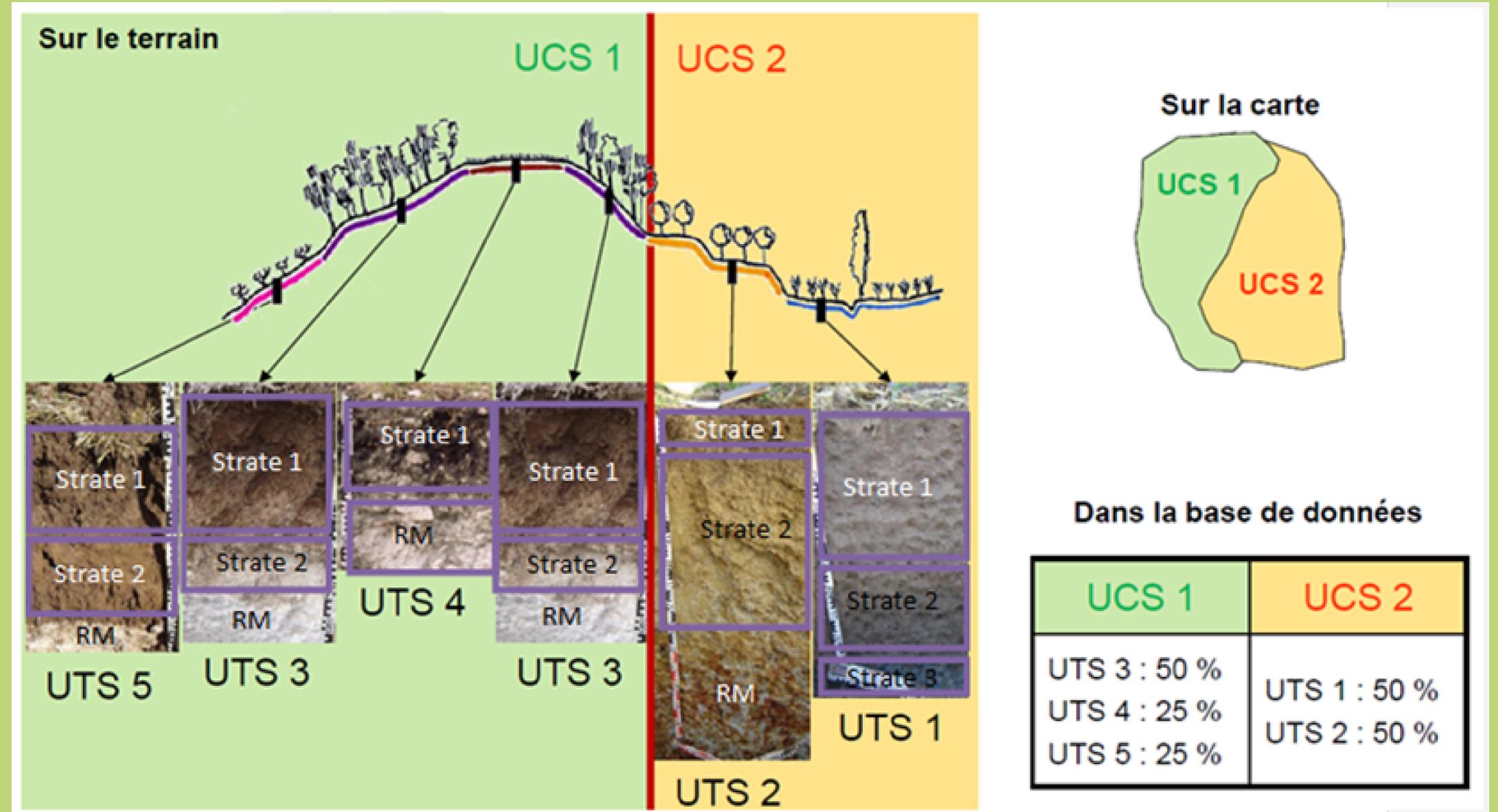
II/ L'organisation du script de valorisation des données

La base de données DoneSol

Base de données nationale regroupant les **données des études pédologiques**

On utilise les **données surfaciques** pour représenter la **variabilité spatiale** des sols

Structure en **unités hiérarchiques**

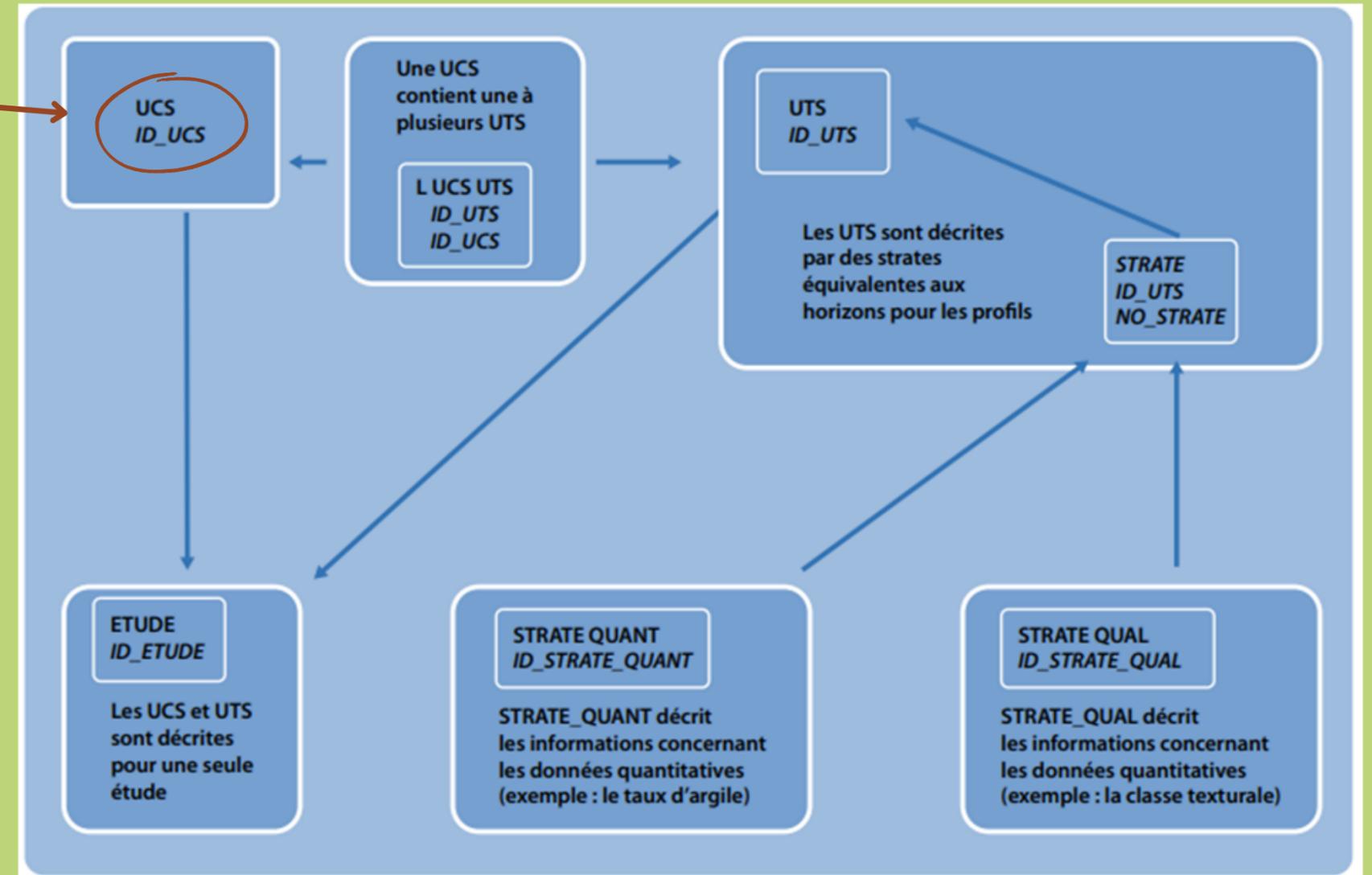
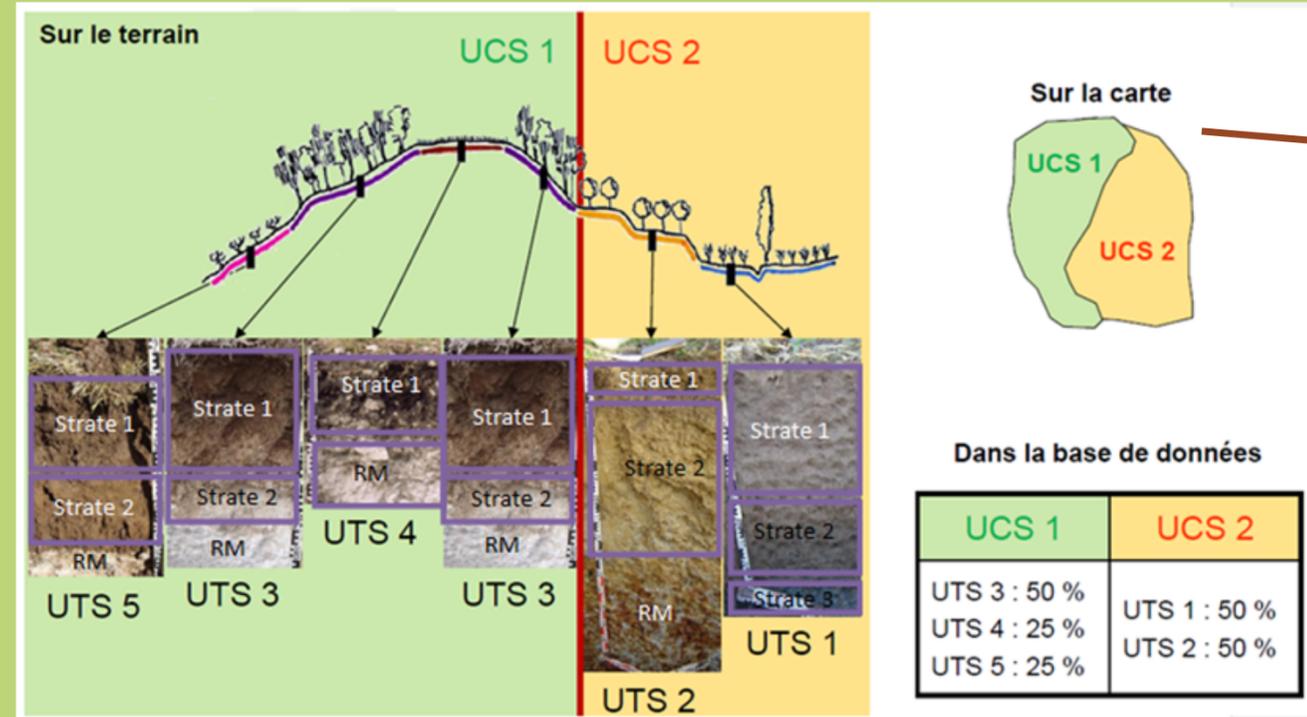


UCS renseignées en fonction des UTS

La base de données DoneSol

Unités hiérarchiques de sols...

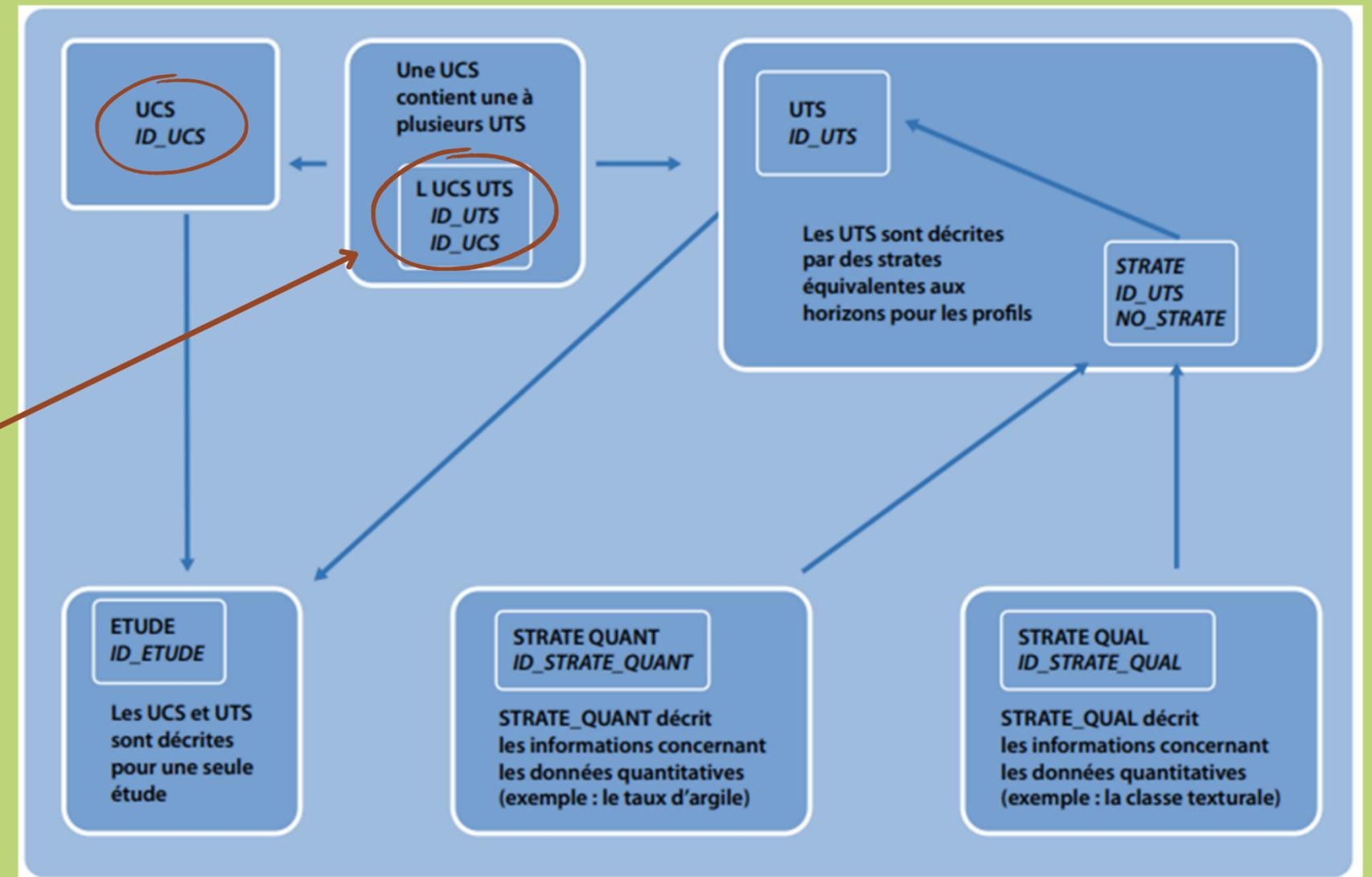
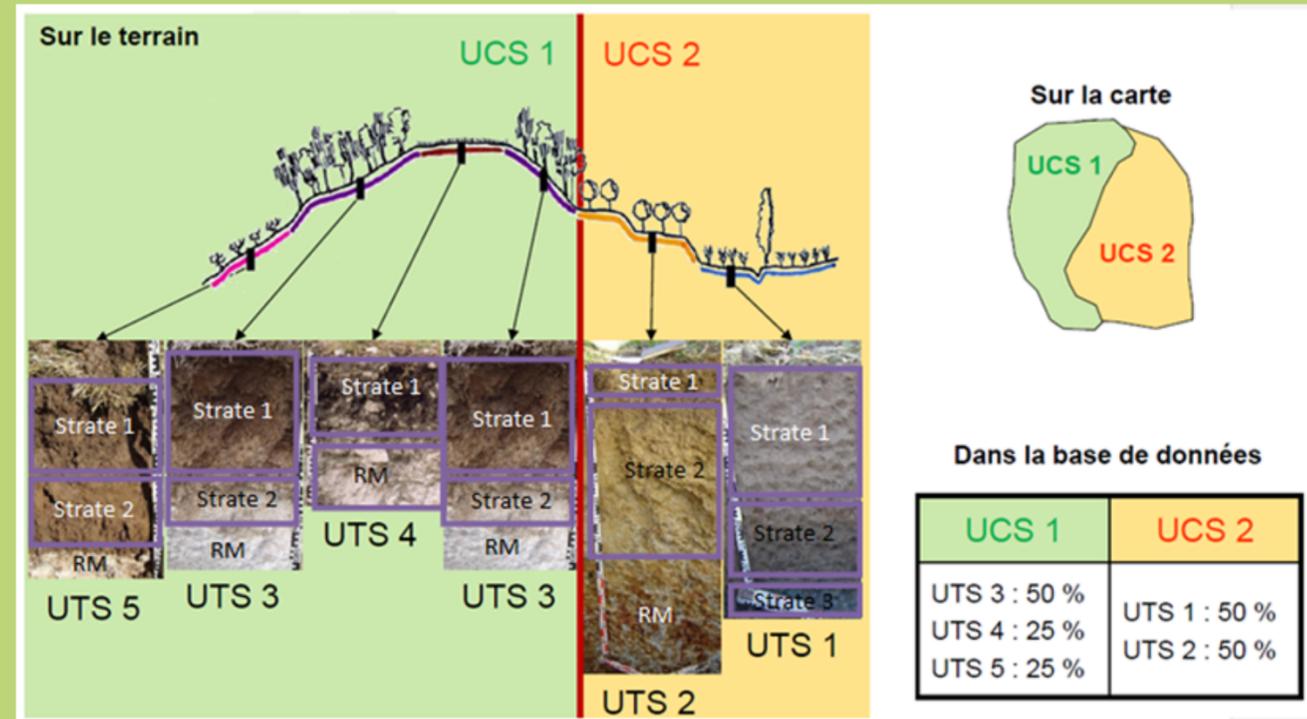
... traduites en tables dans le modèle physique



La base de données DoneSol

Unités hiérarchiques de sols...

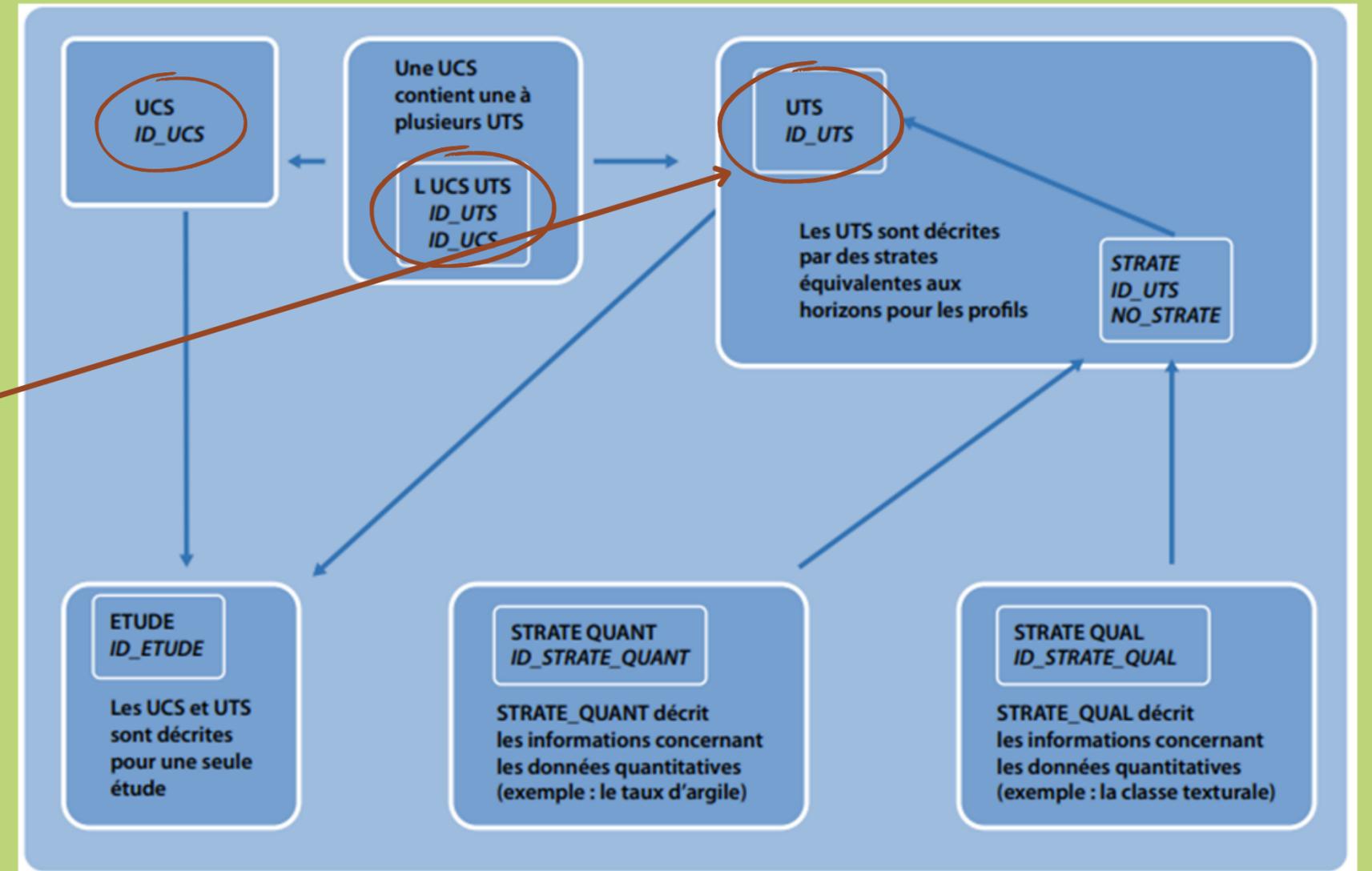
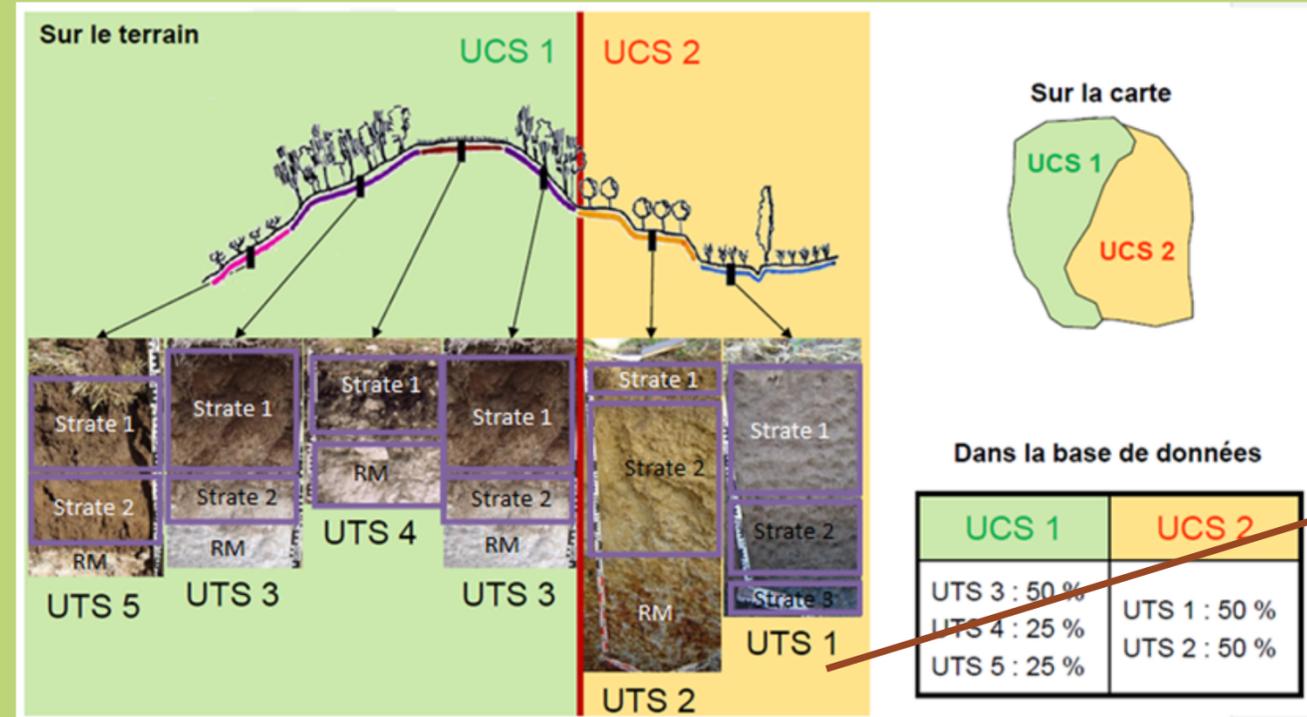
... traduites en tables dans le modèle physique



La base de données DoneSol

Unités hiérarchiques de sols...

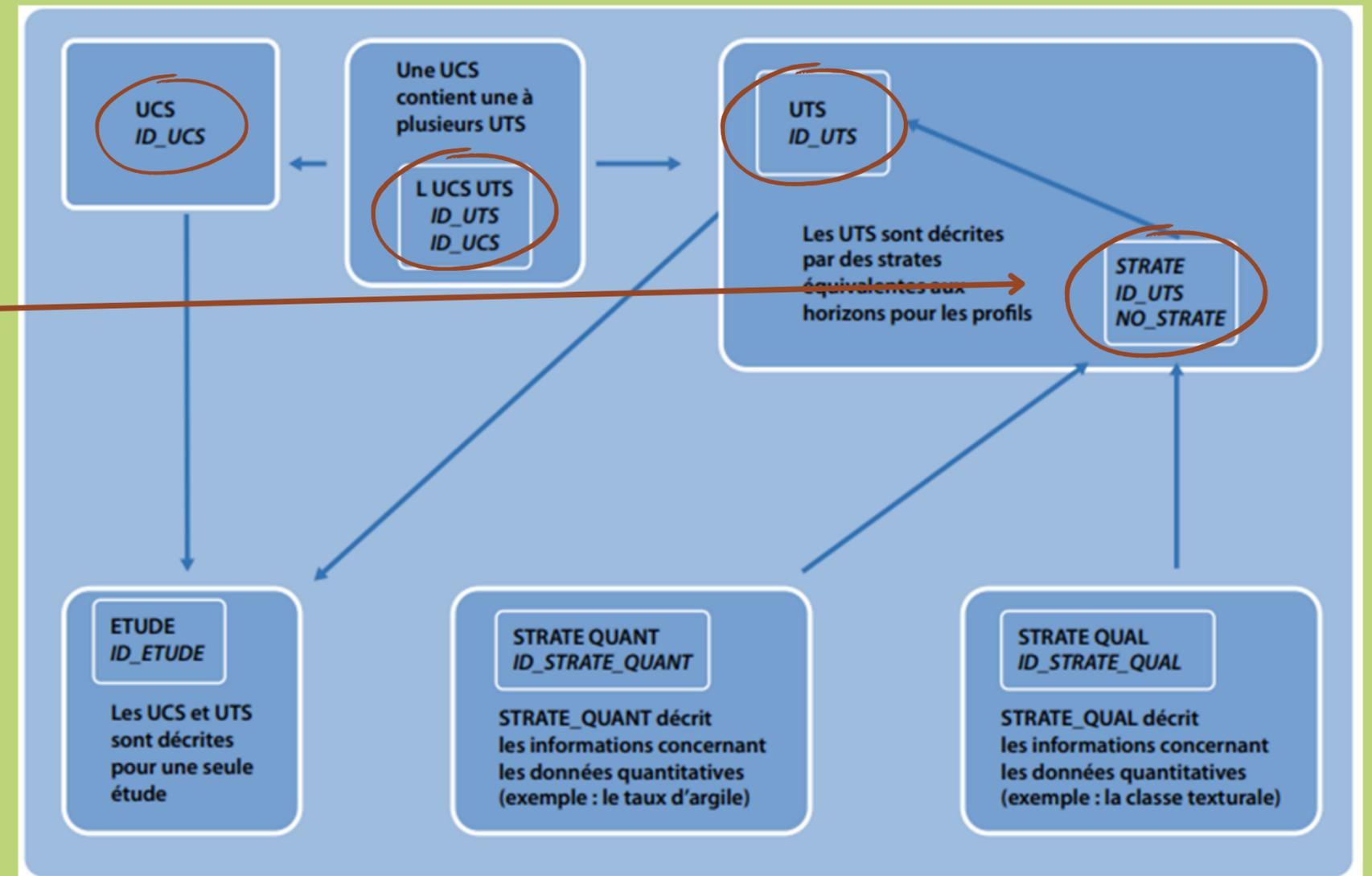
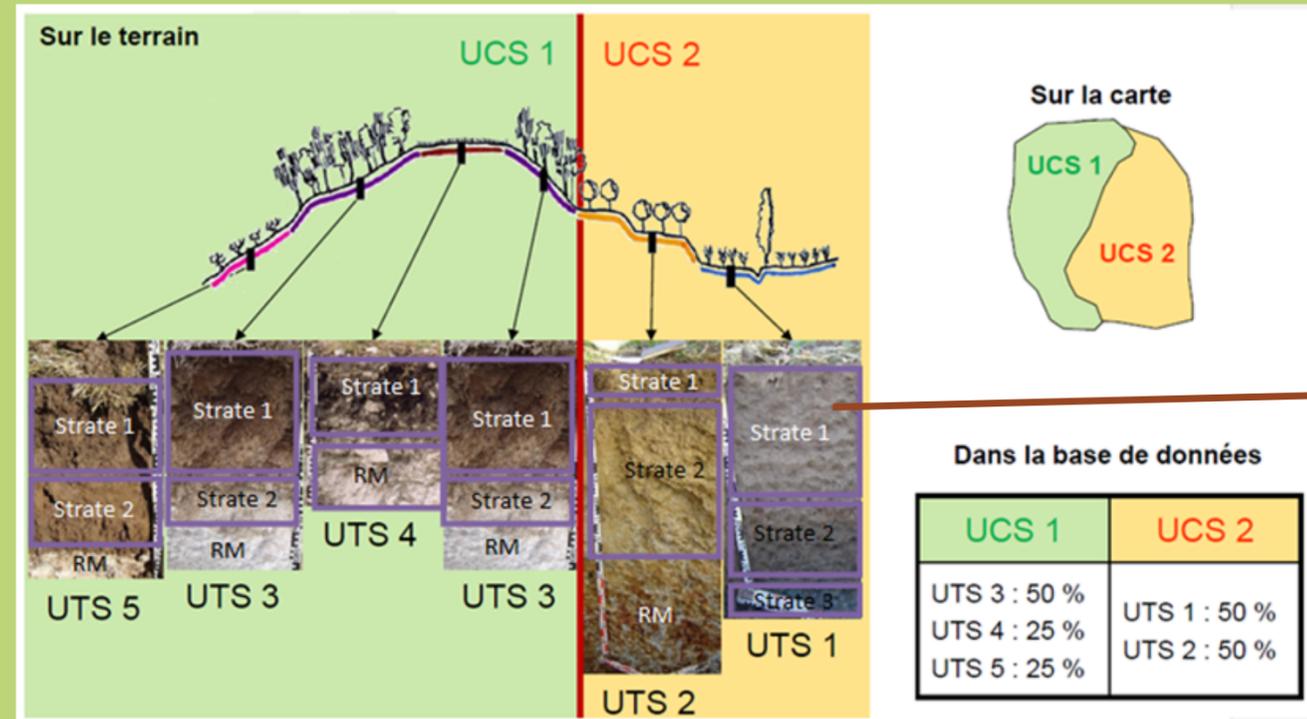
... traduites en tables dans le modèle physique



La base de données DoneSol

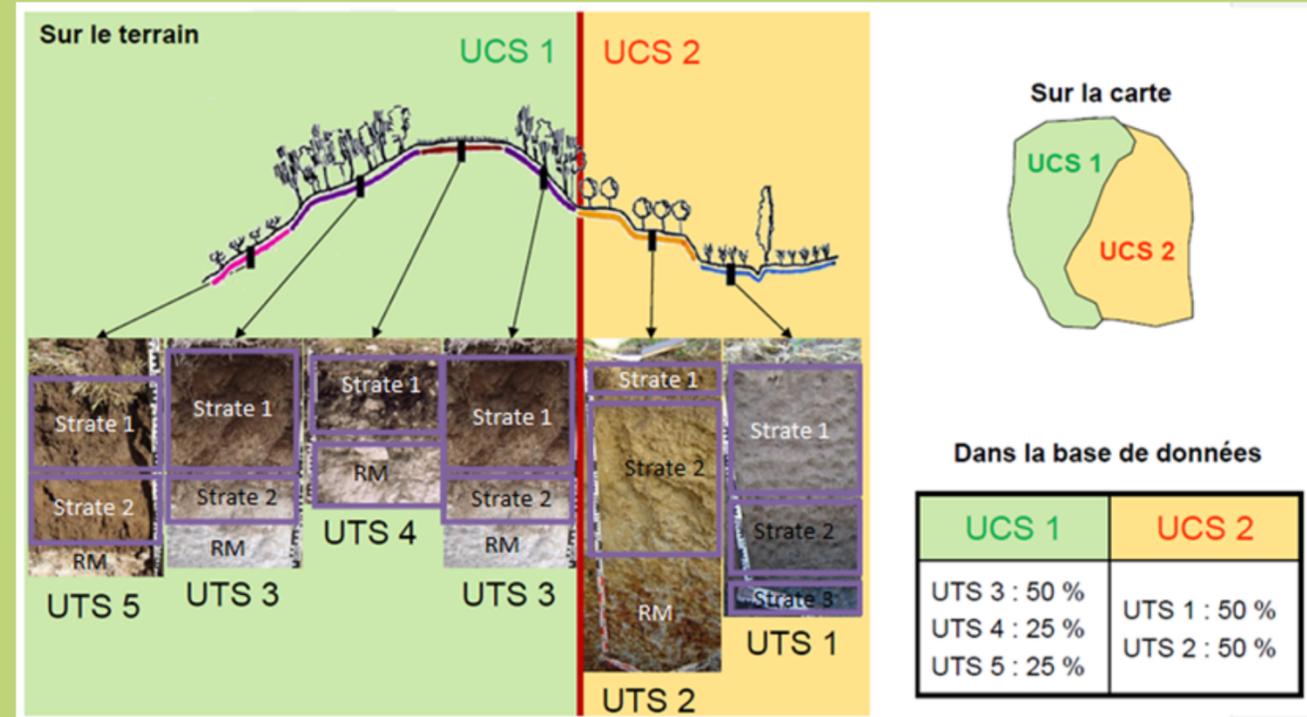
Unités hiérarchiques de sols...

... traduites en tables dans le modèle physique

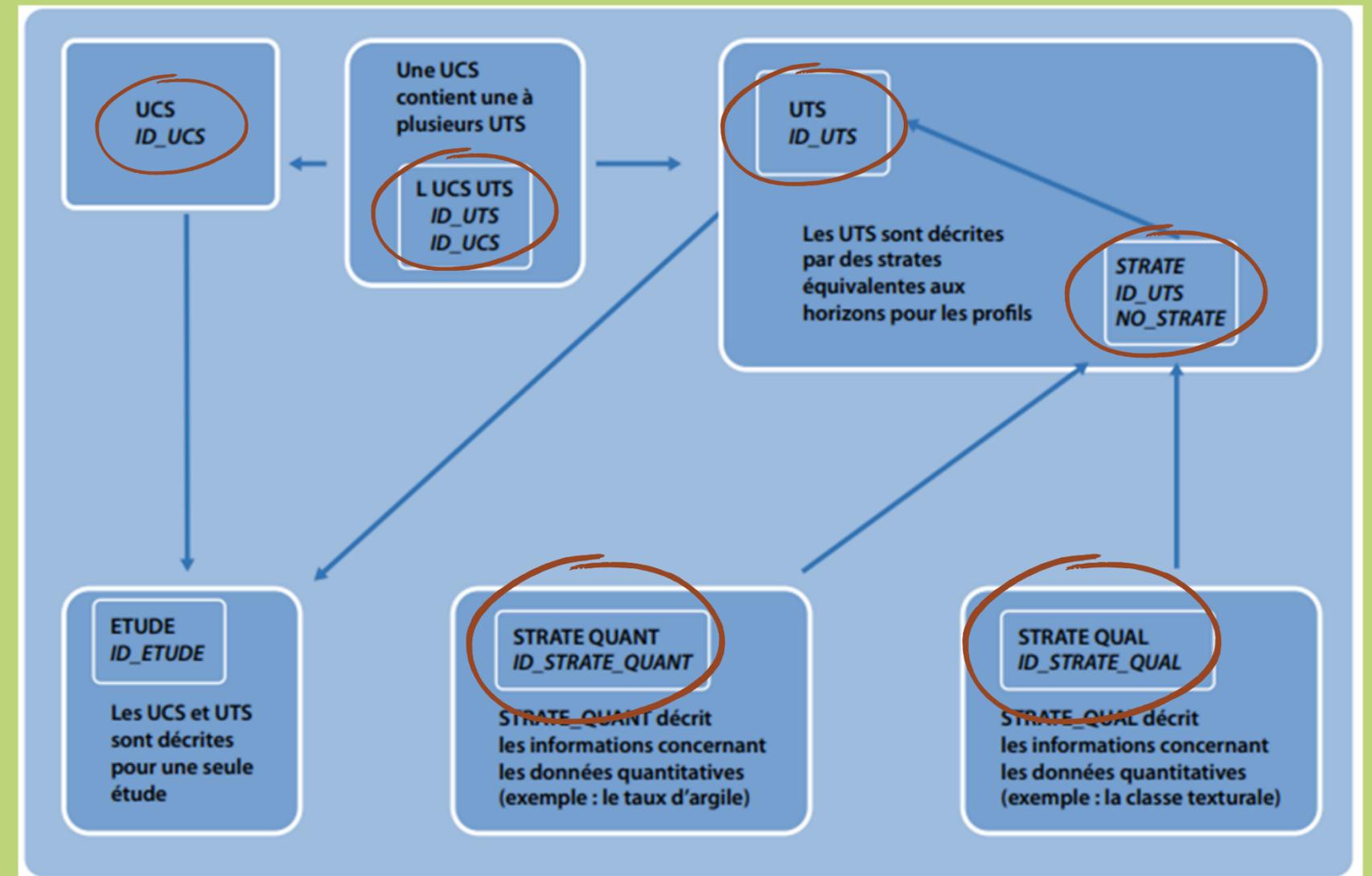


La base de données DoneSol

Unités hiérarchiques de sols...

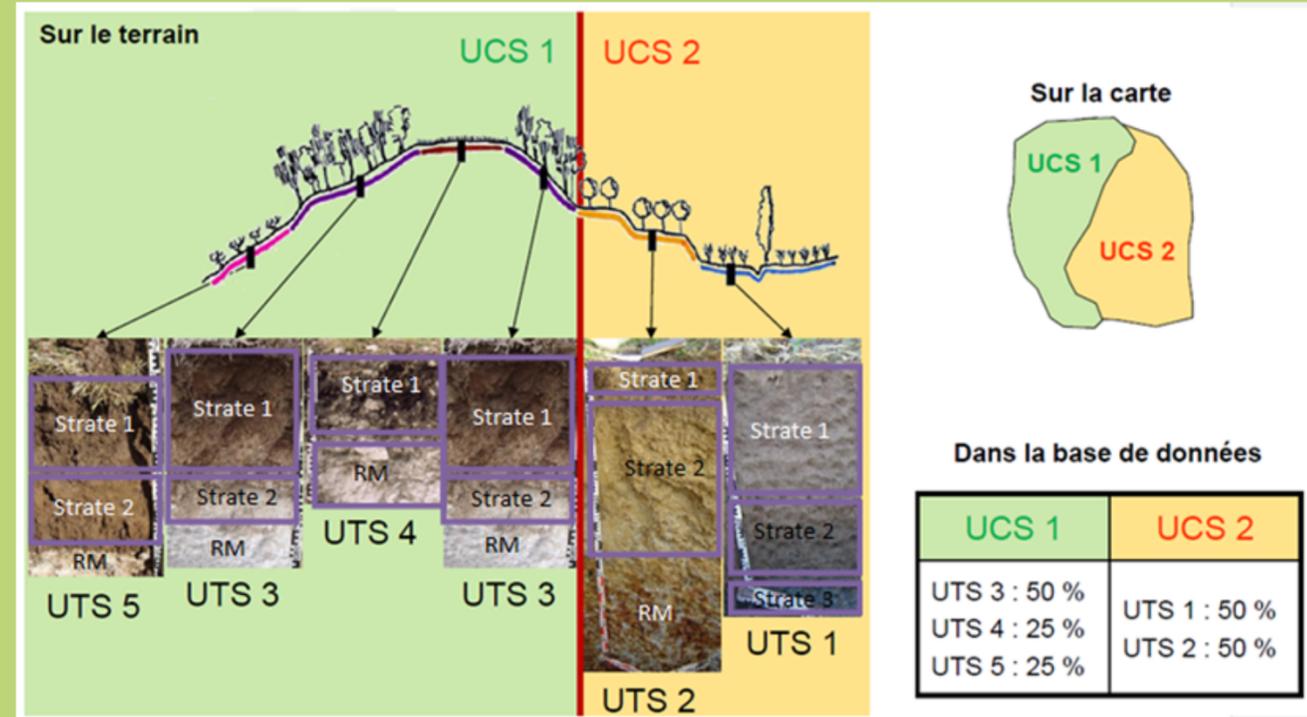


... traduites en tables dans le modèle physique

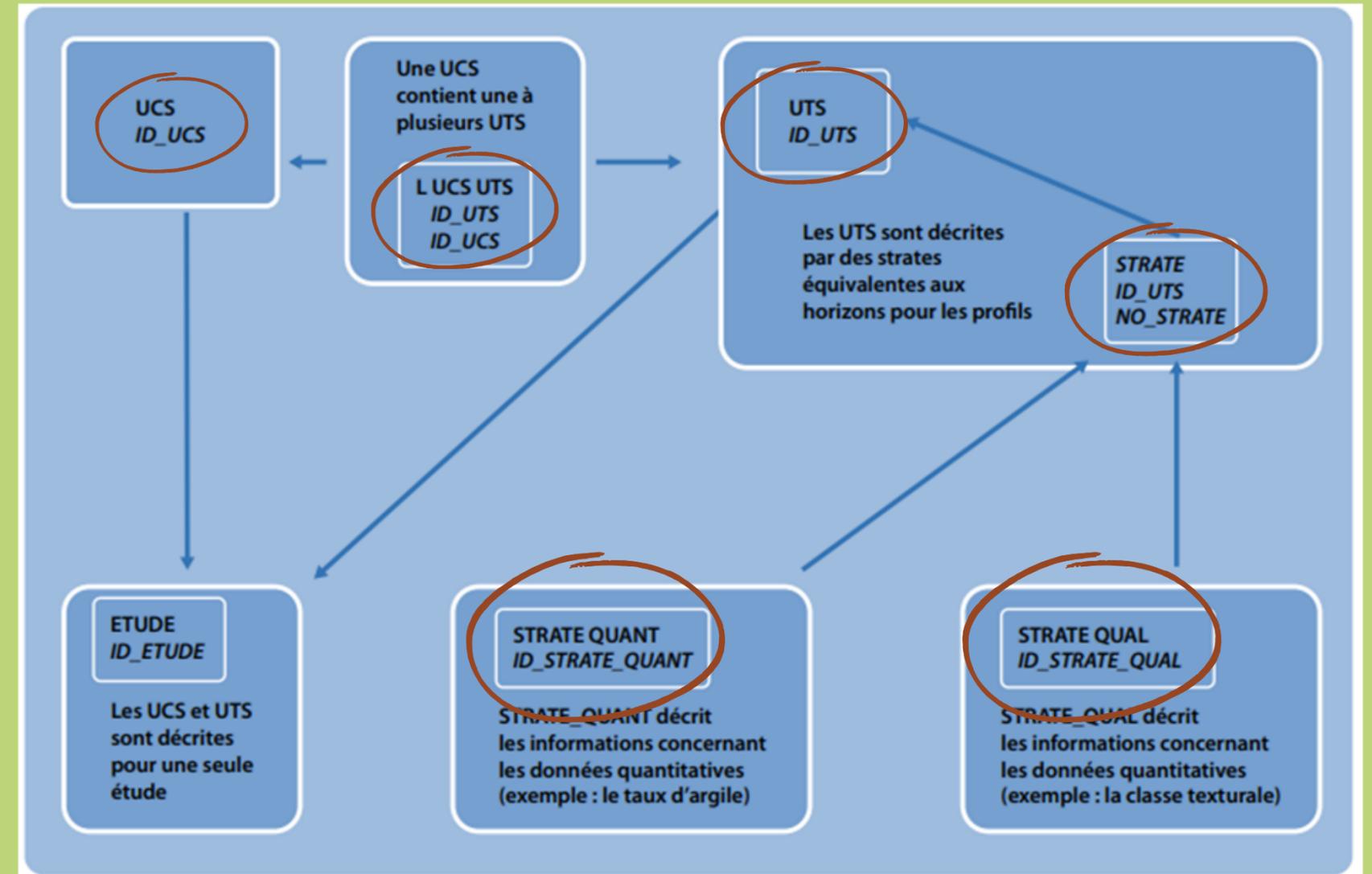


La base de données DoneSol

Unités hiérarchiques de sols...



... traduites en tables dans le modèle physique



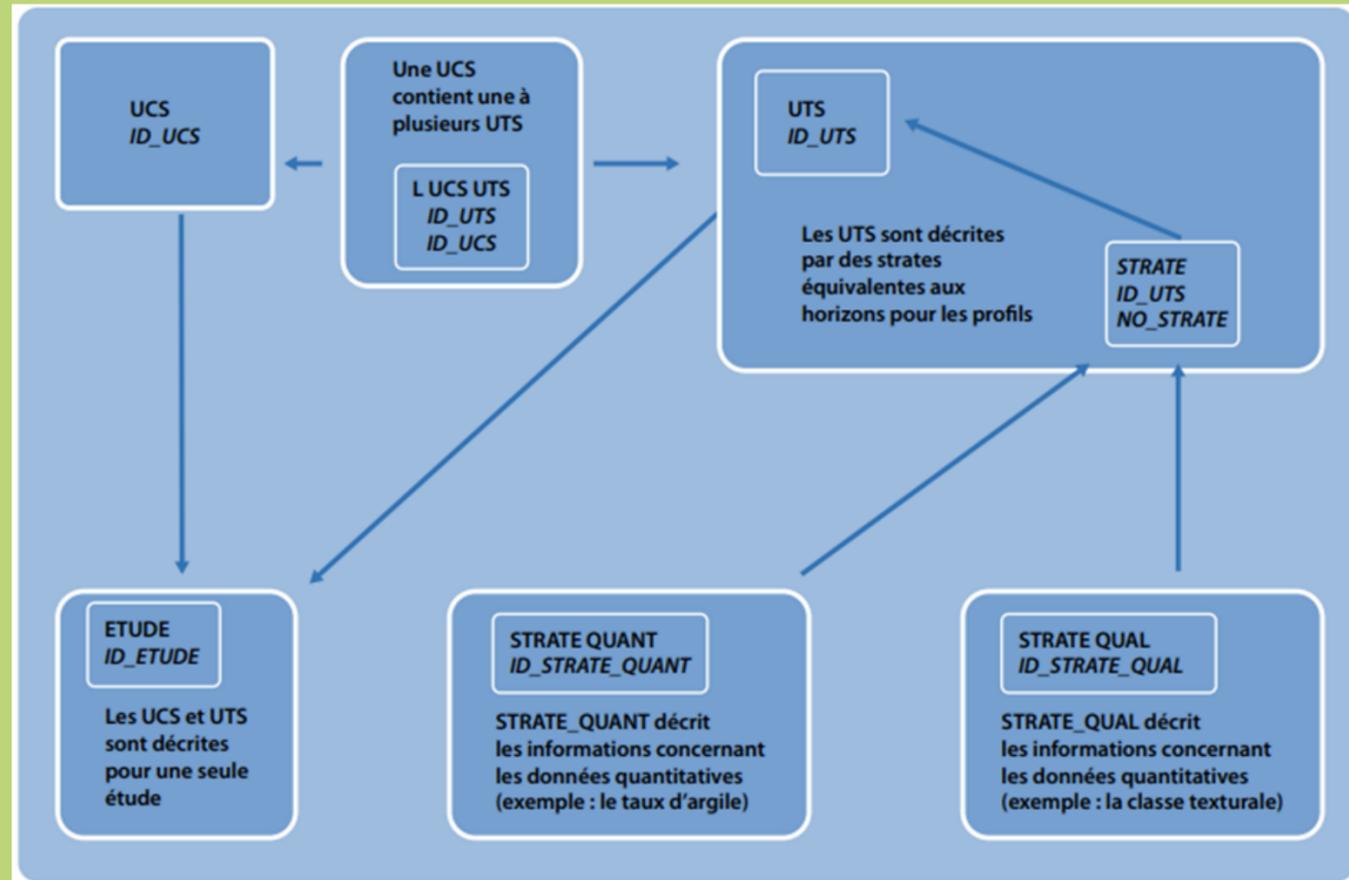
Travail de valorisation des strates jusqu'à l'UCS

Le script prend en compte ces différentes étapes

Le script de valorisation des données

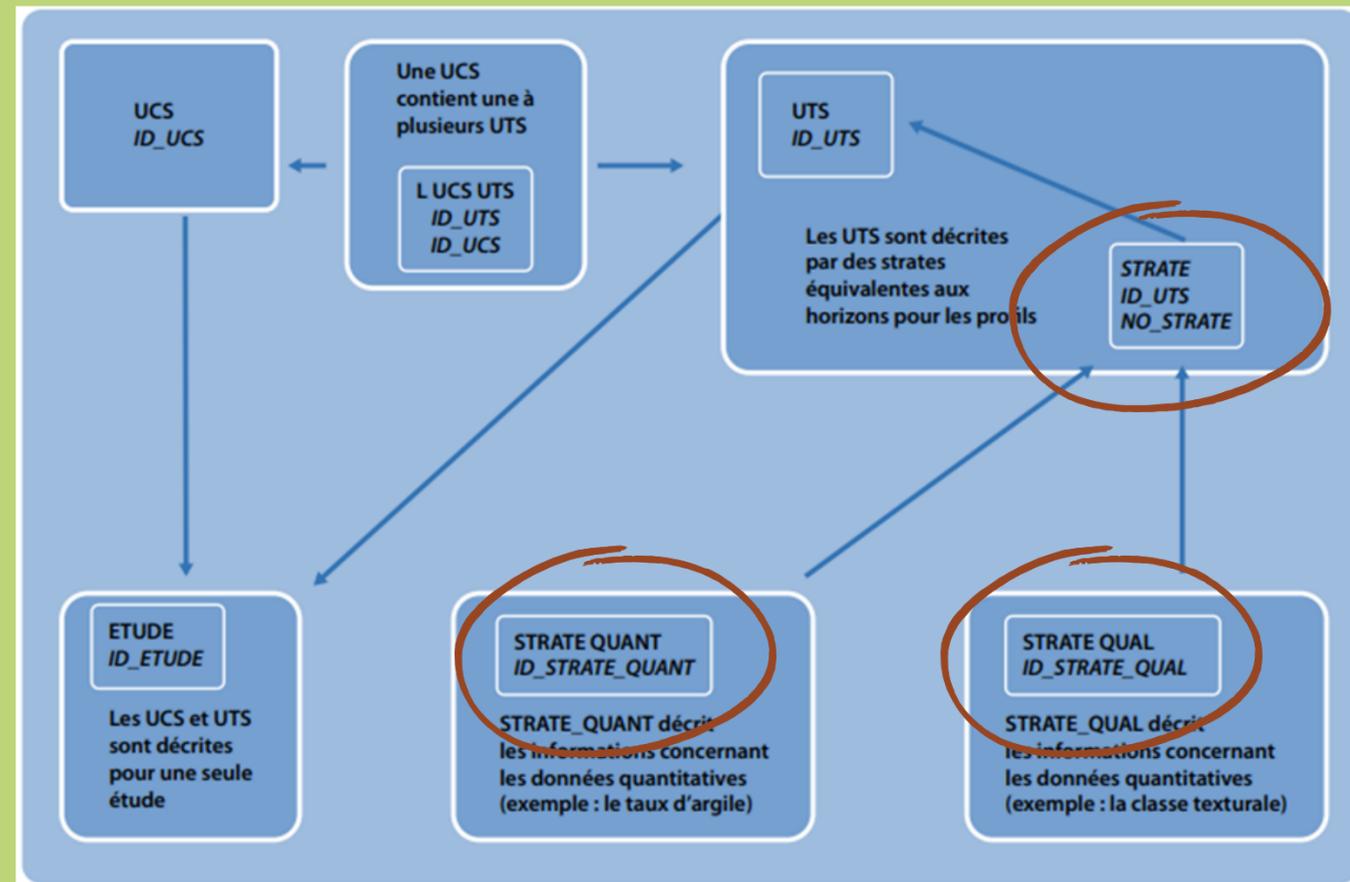
5 étapes

Chargement des données



Le script de valorisation des données

5 étapes



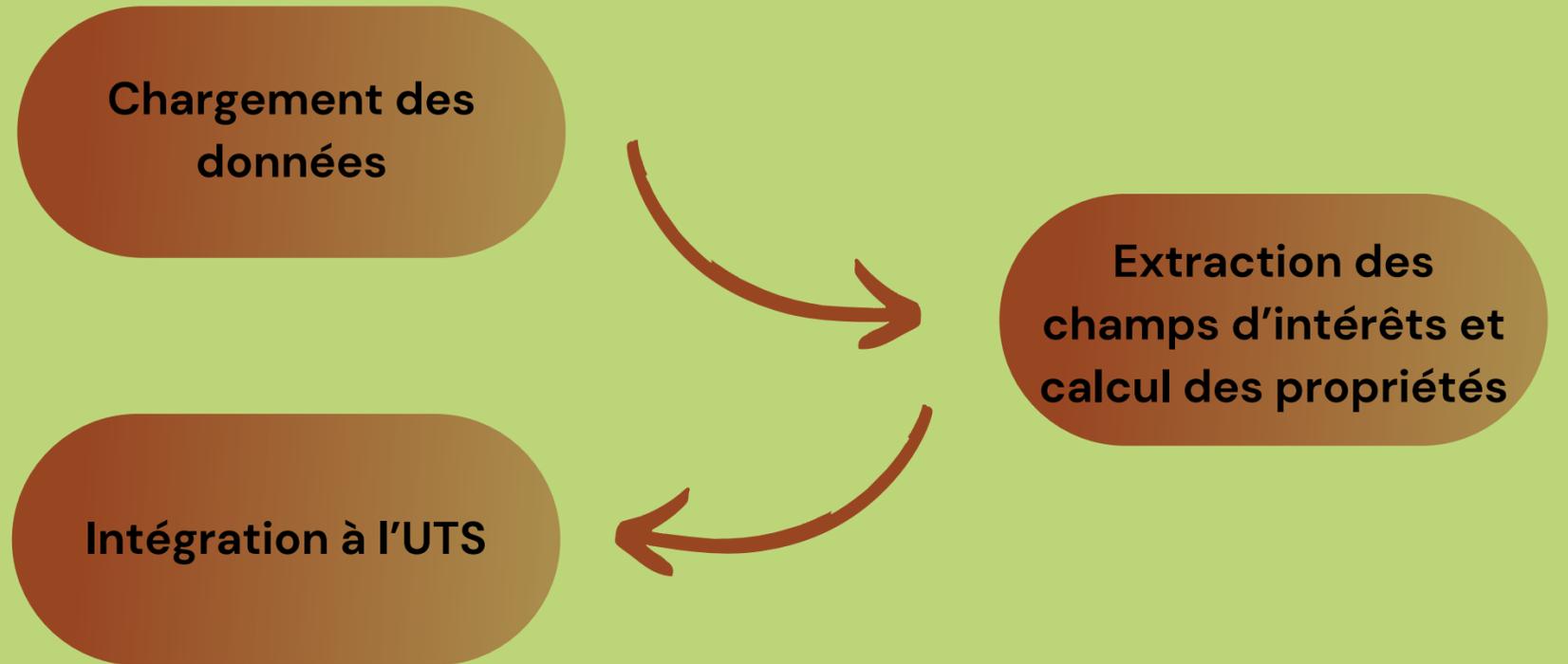
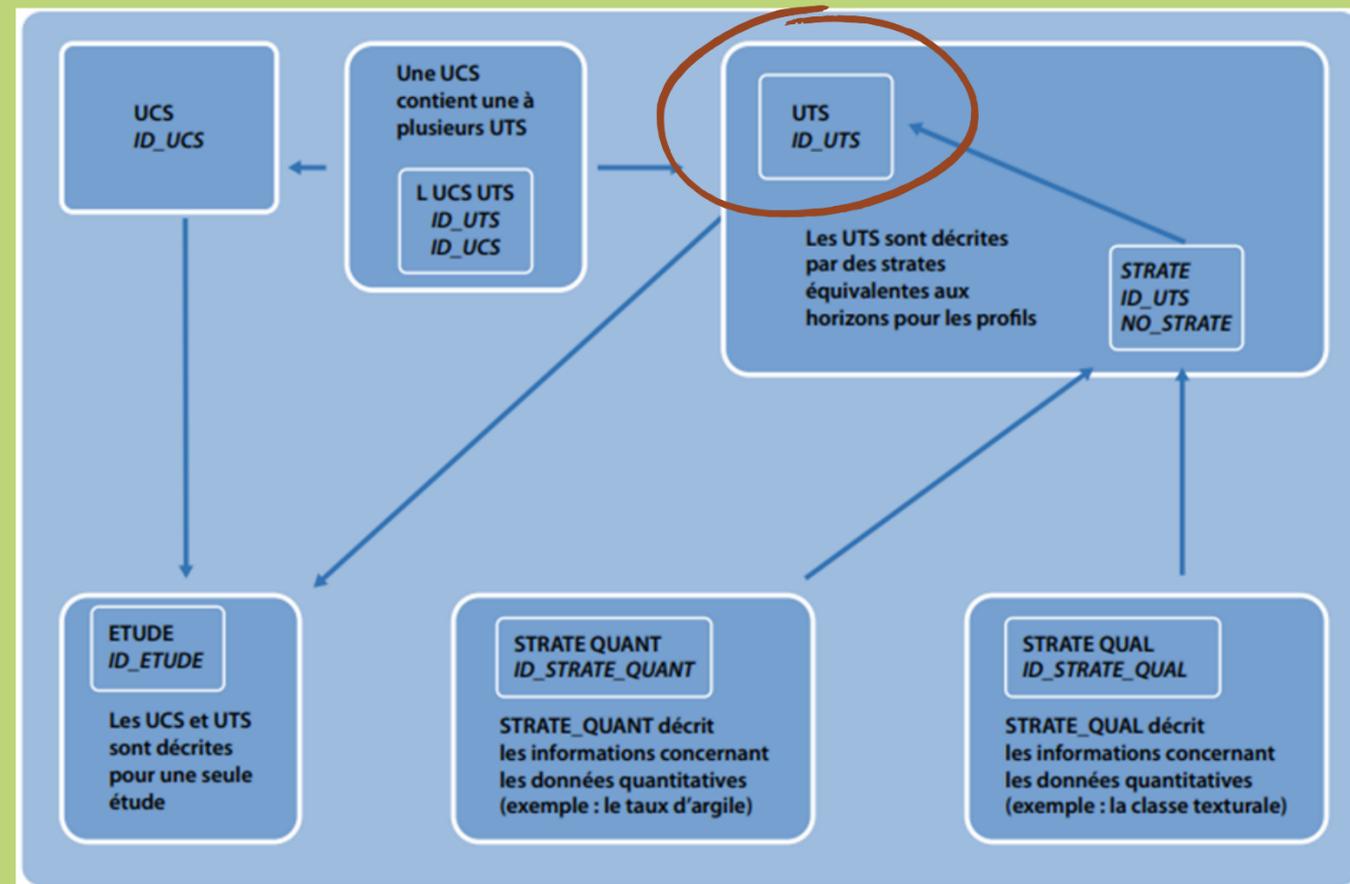
Chargement des données



Extraction des champs d'intérêts et calcul des propriétés

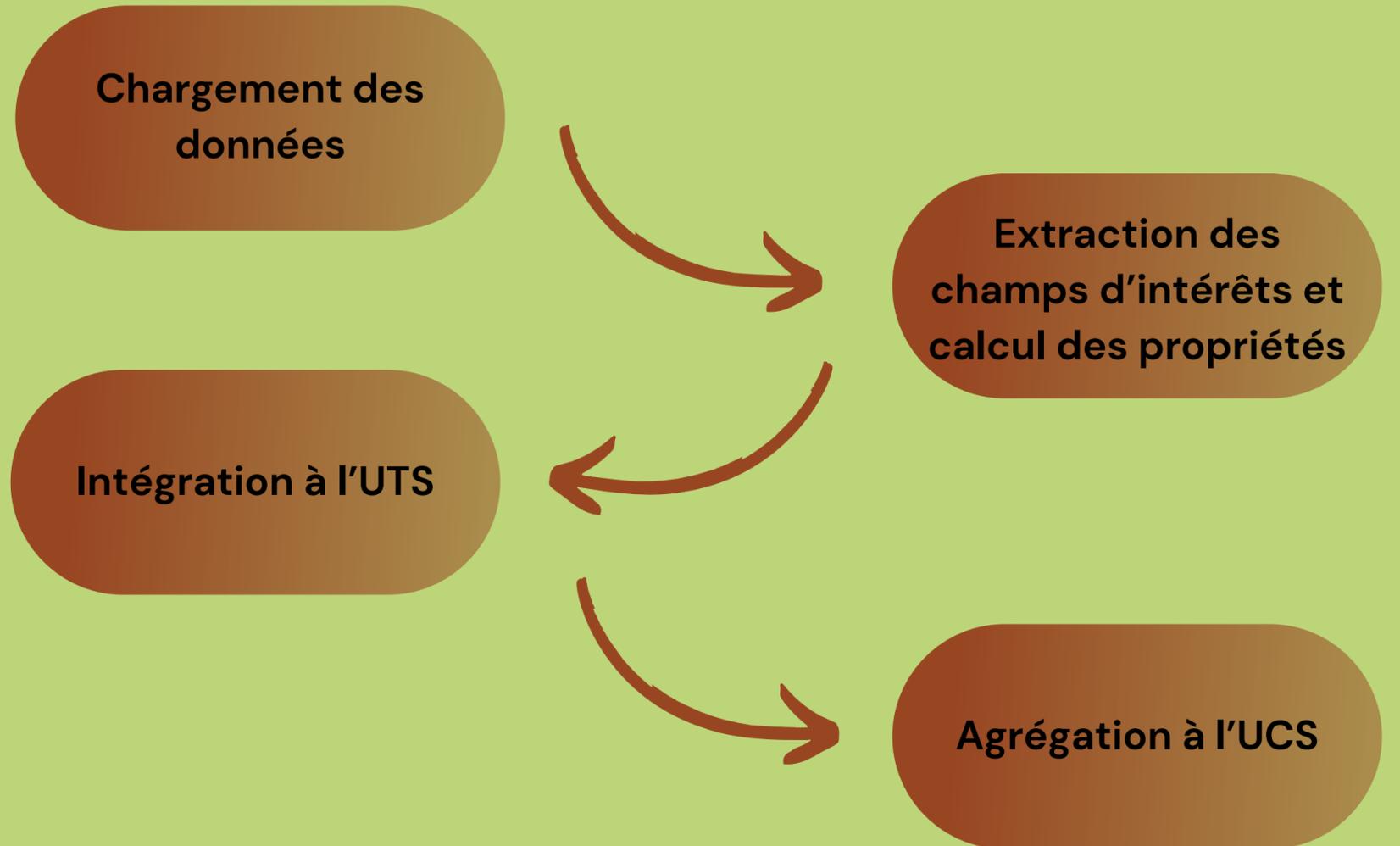
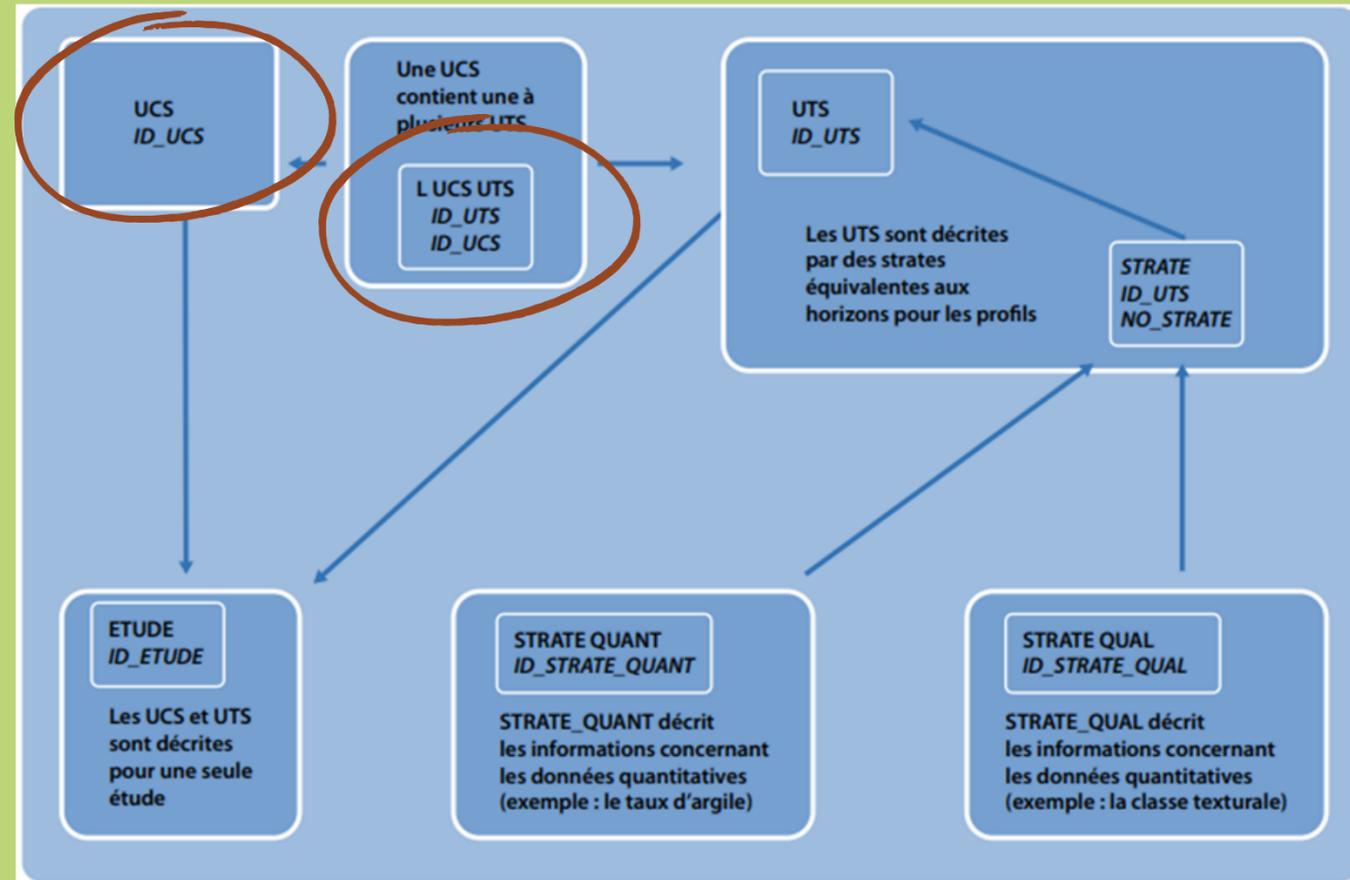
Le script de valorisation des données

5 étapes



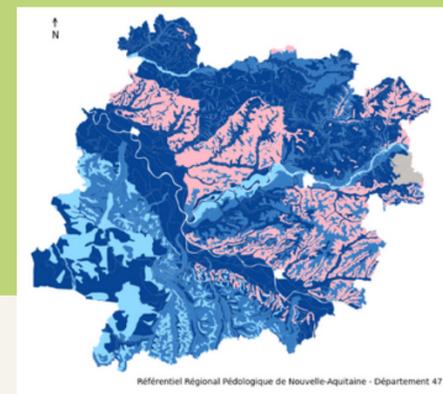
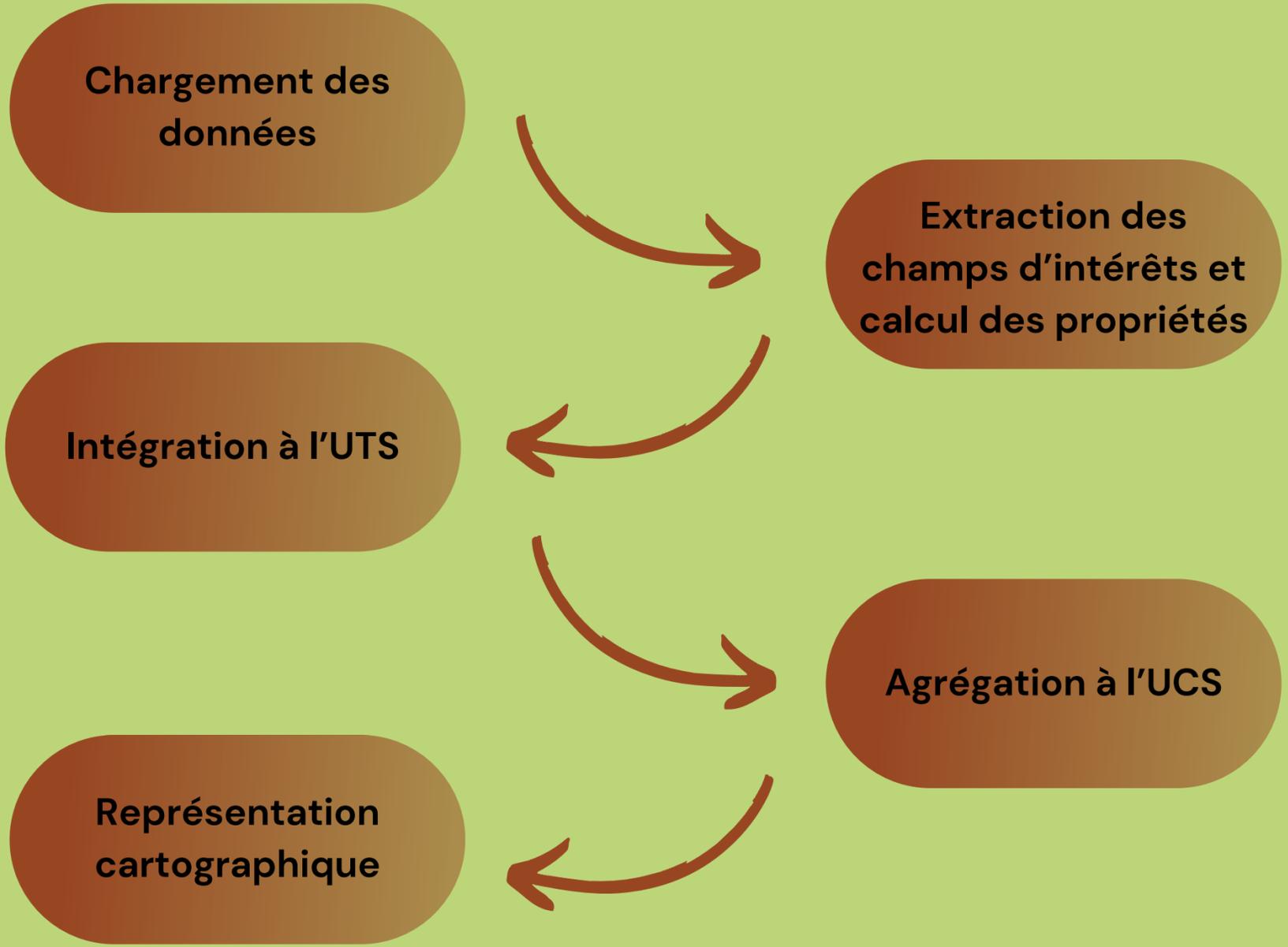
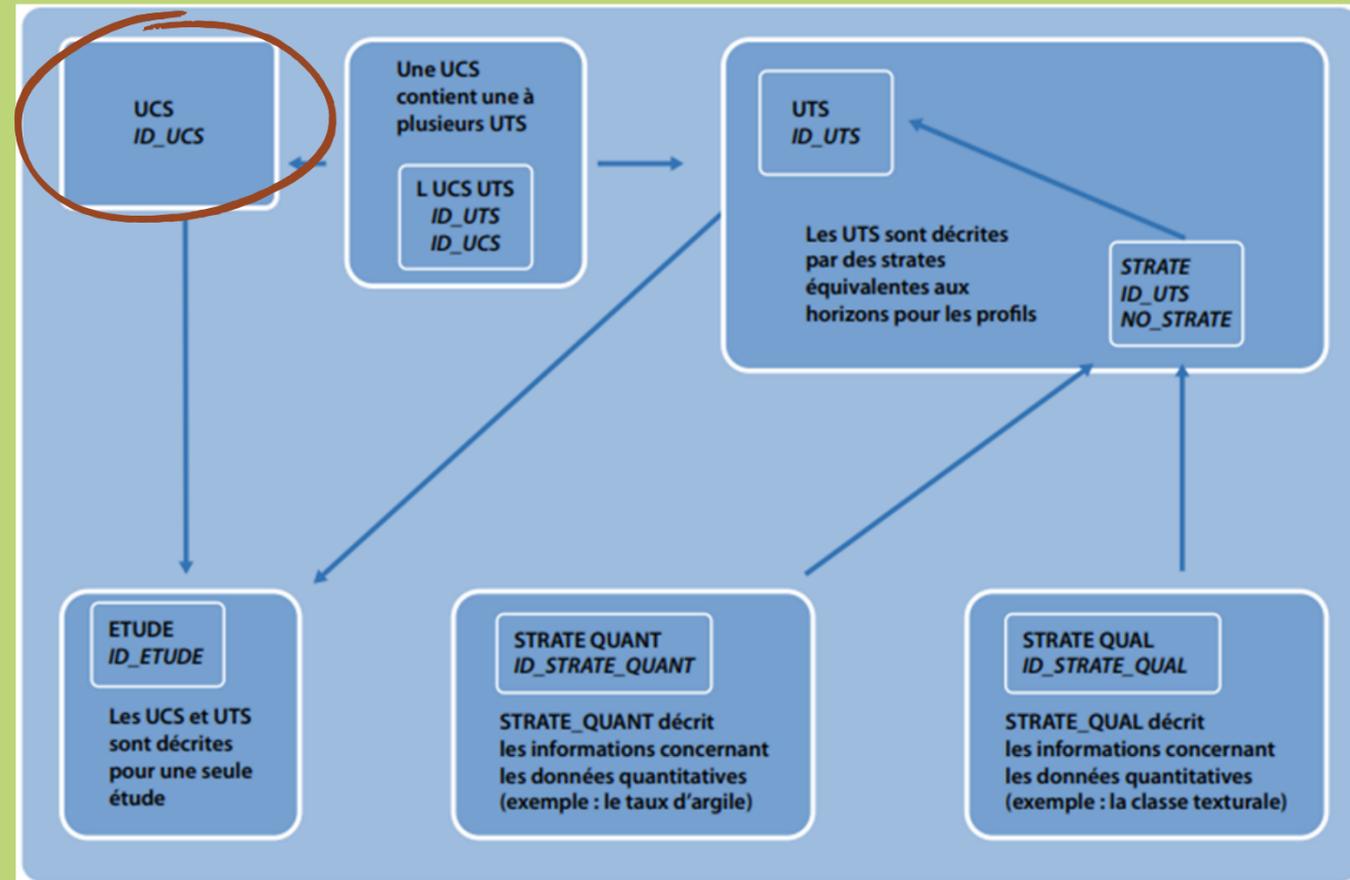
Le script de valorisation des données

5 étapes



Le script de valorisation des données

5 étapes



Le script de valorisation des données

A quoi ressemble le script concrètement ?

Script en langage Python
Développé sous Google Colaboratory

Script Cartographie RRP.ipynb

Fichier Modifier Affichage Insérer Exécution Outils Aide Toutes les modifications ont été enregistrées

Fichiers

- ..
- sample_data
 - 10047_mai_2015_Niveau_2.mdb
 - Strates_avec_RUM.csv
 - Strates_avec_propriétés.csv
 - UCS_avec_RUM_agrege.csv
 - UCS_avec_classes_agregees.csv
 - UTS_avec_RUM.csv
 - UTS_avec_RUM_mise_en_classe.c...
 - sols_nouvelle_aquitaine_rrp47-1.js...

+ Code + Texte

↳ Importation des bibliothèques nécessaires au script Organisé en sections

↳ Importation des données du RPP et préparation du Dataframe

↳ Import des tables à partir du fichier Access

```
[3] # Recherche automatique du fichier et import avec vérification automatique

# Recherche de fichiers contenant ".mdb" dans leur nom
matching_files = glob.glob("*.mdb*")

# Vérifier le nombre de fichiers correspondants
if len(matching_files) == 1:
    # Créer une instance AccessParser pour le fichier correspondant
    db = AccessParser(matching_files[0])

    # Extraire les deux derniers chiffres du nom de fichier
    num_dep_match = re.search(r'\d{2}(?=\D|$)', matching_files[0])
    if num_dep_match:
        num_dep = int(num_dep_match.group())
    else:
        raise ValueError("Impossible de trouver les deux derniers chiffres dans le nom de fichier")

    # Effectuer les opérations souhaitées sur la base de données en utilisant AccessParser
    # Par exemple, afficher les informations sur le catalogue et le num_dep
    print("Informations sur le catalogue :", db.catalog)
    print("Fichier correspondant :", matching_files[0])
    print("num_dep :", num_dep)

elif len(matching_files) > 1:
```

Le script de valorisation des données

A quoi ressemble le script concrètement ?

Script en langage Python
Développé sous Google Colaboratory

Permet d'importer et d'exporter les fichiers facilement

Chargement des données

Script Cartographie RRP.ipynb

Fichier Modifier Affichage Insérer Exécution Outils Aide Toutes les modifications ont été enregistrées

Fichiers

- ..
- sample_data
- 10047_mai_2015_Niveau_2.mdb
- Strates_avec_RUM.csv
- Strates_avec_propriétés.csv
- UCS_avec_RUM_agrege.csv
- UCS_avec_classes_agregees.csv
- UTS_avec_RUM.csv
- UTS_avec_RUM_mise_en_classe.c...
- sols_nouvelle_aquitaine_rrp47-1.js...

+ Code + Texte

↳ Importation des bibliothèques nécessaires au script Organisé en sections

↳ Importation des données du RPP et préparation du Dataframe

↳ Import des tables à partir du fichier Access

```
[3] # Recherche automatique du fichier et import avec vérification automatique

# Recherche de fichiers contenant ".mdb" dans leur nom
matching_files = glob.glob("*.mdb*")

# Vérifier le nombre de fichiers correspondants
if len(matching_files) == 1:
    # Créer une instance AccessParser pour le fichier correspondant
    db = AccessParser(matching_files[0])

    # Extraire les deux derniers chiffres du nom de fichier
    num_dep_match = re.search(r'\d{2}(?=\D|$)', matching_files[0])
    if num_dep_match:
        num_dep = int(num_dep_match.group())
    else:
        raise ValueError("Impossible de trouver les deux derniers chiffres dans le nom de fichier")

    # Effectuer les opérations souhaitées sur la base de données en utilisant AccessParser
    # Par exemple, afficher les informations sur le catalogue et le num_dep
    print("Informations sur le catalogue :", db.catalog)
    print("Fichier correspondant :", matching_files[0])
    print("num_dep :", num_dep)

elif len(matching_files) > 1:
```

Le script de valorisation des données

Extraction et calcul des propriétés

Création d'un **Dataframe**

Extraction des champs d'intérêts et ajout de colonnes pour les propriétés : ajout d'informations sans en supprimer

	no_strate	prof_appar_min	prof_appar_moy	prof_appar_max	epais_min	epais_max	epais_moy	id_uts	forme_strate
0	1	0.0	0.0	0.0	NaN	NaN	30.0	5062	None
1	1	0.0	0.0	0.0	20.0	30.0	28.0	5385	0
2	1	0.0	0.0	0.0	25.0	40.0	30.0	5503	0
3	1	0.0	0.0	0.0	20.0	60.0	30.0	5616	None
4	1	0.0	0.0	0.0	NaN	NaN	20.0	6163	1

Le script de valorisation des données

Extraction et calcul des propriétés

Création d'un **Dataframe**

Extraction des champs d'intérêts et ajout de colonnes pour les propriétés : **ajout d'informations sans en supprimer**

Comment combler les manques ?

	no_strate	prof_appar_min	prof_appar_moy	prof_appar_max	epais_min	epais_max	epais_moy	id_uts	forme_strate
0	1	0.0	0.0	0.0	NaN	NaN	30.0	5062	None
1	1	0.0	0.0	0.0	20.0	30.0	28.0	5385	0
2	1	0.0	0.0	0.0	25.0	40.0	30.0	5503	0
3	1	0.0	0.0	0.0	20.0	60.0	30.0	5616	None
4	1	0.0	0.0	0.0	NaN	NaN	20.0	6163	1

Le script de valorisation des données

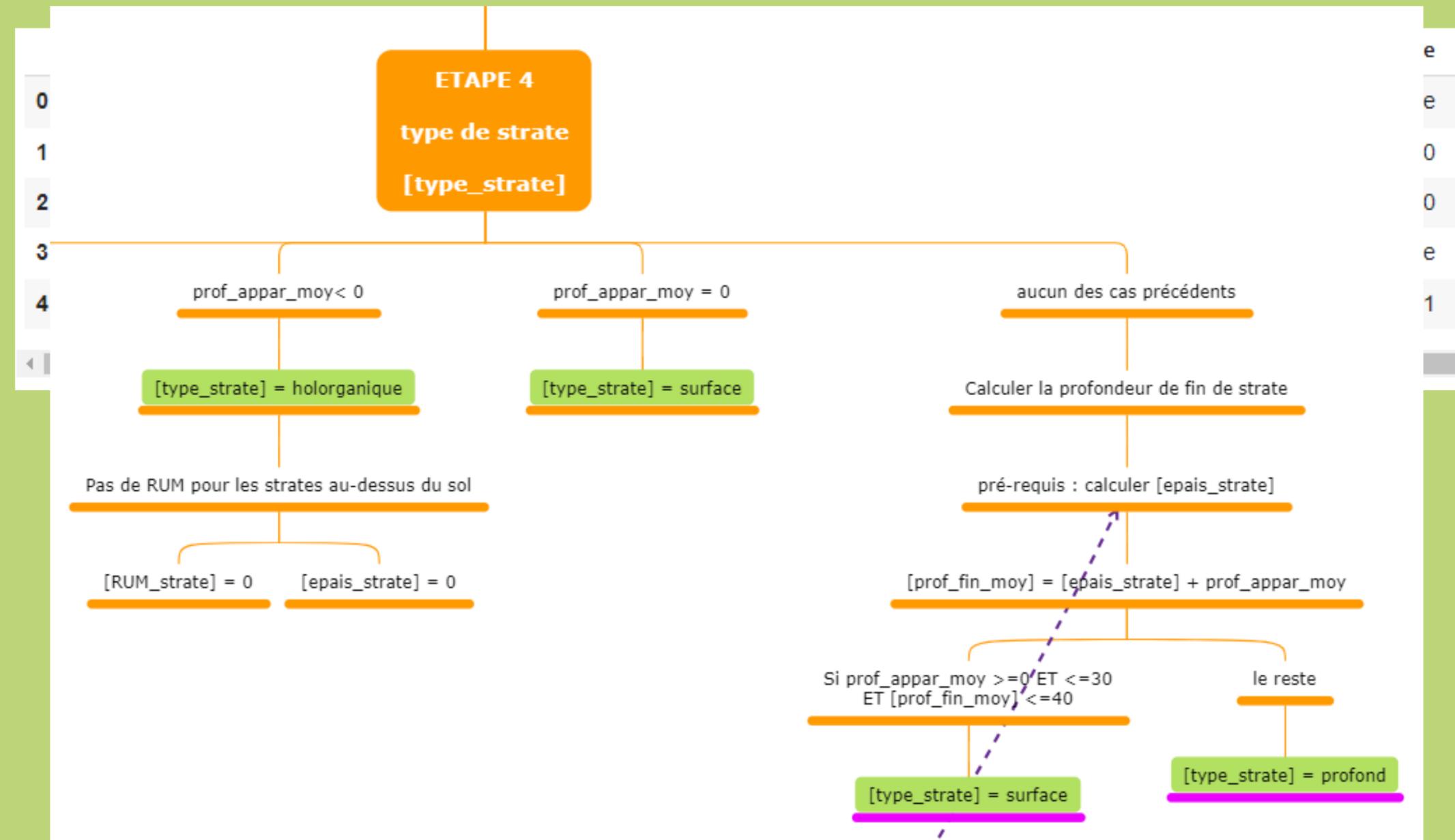
Extraction et calcul des propriétés

Création d'un Dataframe

Extraction des champs d'intérêts et ajout de colonnes pour les propriétés : ajout d'informations sans en supprimer

Comment combler les manques ?

Arbre de décision

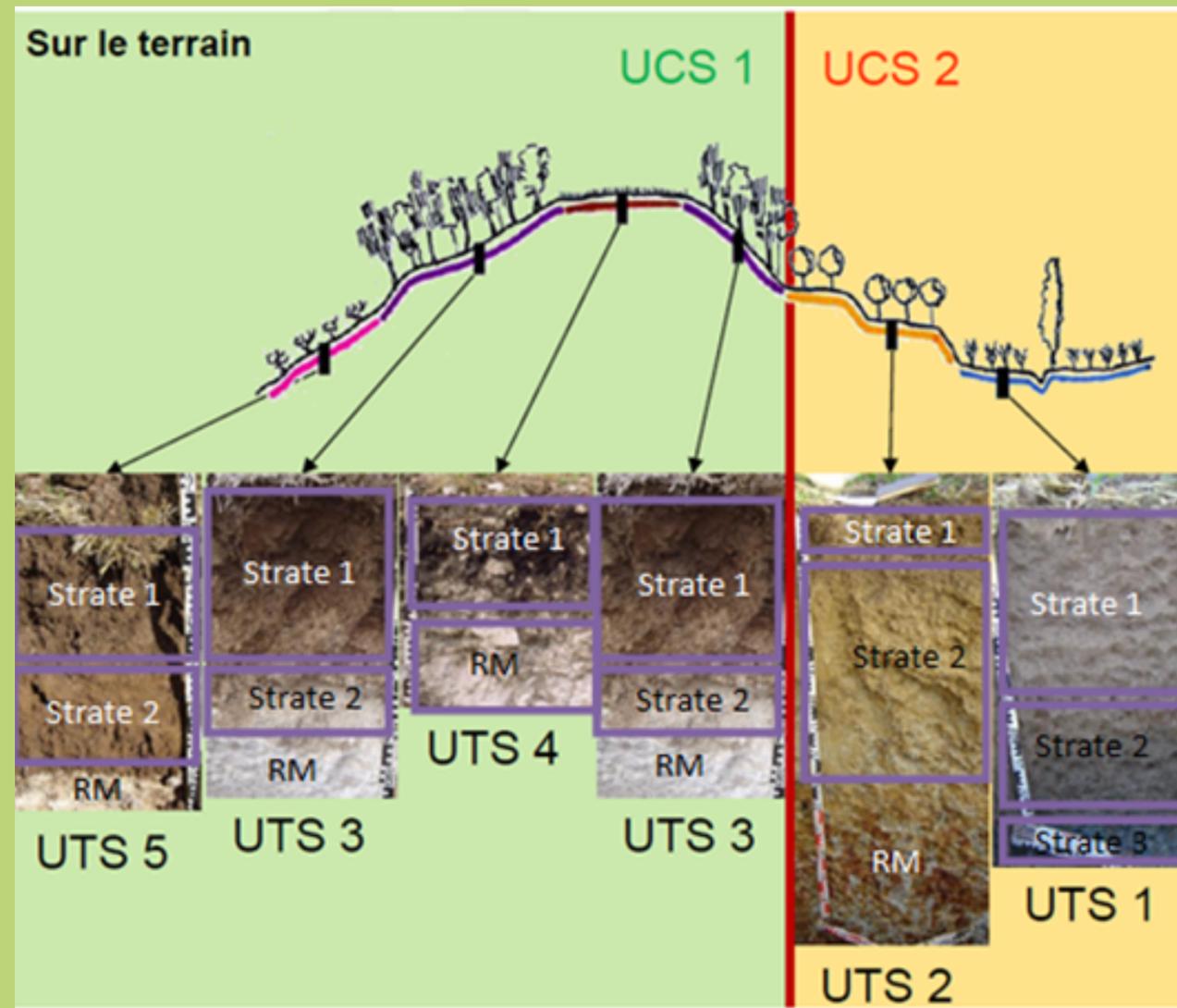


Le script de valorisation des données

Intégration à l'UTS

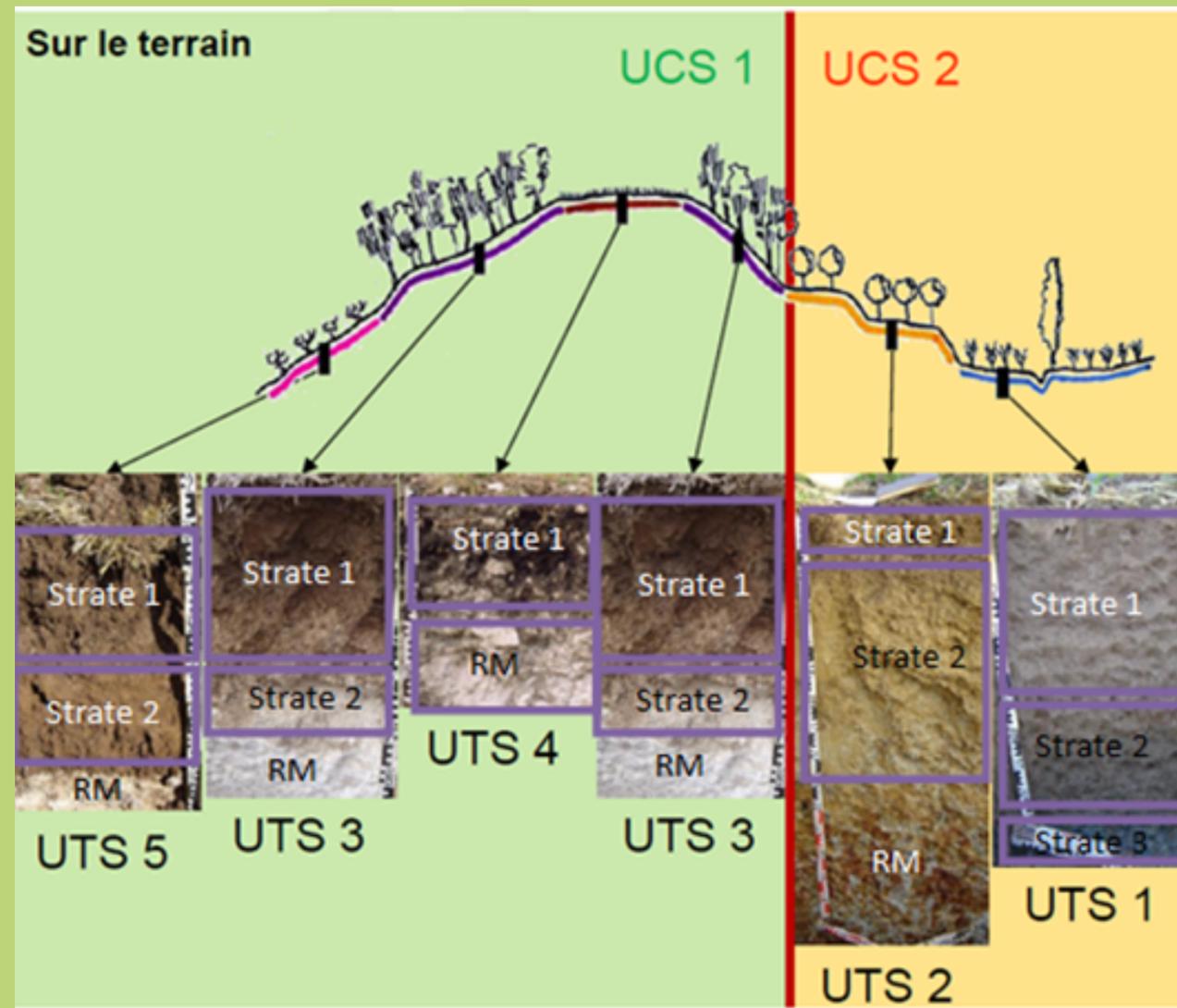
Dépendante de la nature de la propriété considérée

Dans le mémoire, exemple du RUM : stock d'eau par strate
RUM UTS = Somme RUM de chaque strate



Le script de valorisation des données

Intégration à l'UTS



Dépendante de la nature de la propriété considérée

Dans le mémoire, exemple du RUM : stock d'eau par strate
RUM UTS = Somme RUM de chaque strate

Exemples des teneurs : teneur en carbone du sol
Carbone à l'UTS = Moyenne des teneurs en carbone des strates pondérée par leur épaisseur

Le script de valorisation des données

Agrégation à l'UCS

Méthode retenue dans le script :
Agrégation par la classe majoritaire
dans l'UCS

1) Mise en classe de la variable d'intérêt
pour chaque UTS

Classes de RUM utilisées, en mm d'eau :

- 1 : 0 - 25 mm d'eau
- 2 : 25 - 50 mm d'eau
- 3 : 50 - 100 mm d'eau
- 4 : 100 - 150 mm d'eau
- 5 : \geq 150 mm d'eau
- -99 : -99, données insuffisantes pour estimer le RUM

On inclut la borne inférieure dans les classes

UCS	UTS	%	RUM (mm)	Classe de RUM
1	11	35	50	3
1	12	40	170	5
1	13	25	80	3

Le script de valorisation des données

Agrégation à l'UCS

Méthode retenue dans le script :
Agrégation par la classe majoritaire dans l'UCS

1) Mise en classe de la variable d'intérêt pour chaque UTS

2) Choix de la classe majoritaire dans l'UCS

UCS 1 : **60%** de la classe de RUM 3

Classes de RUM utilisées, en mm d'eau :

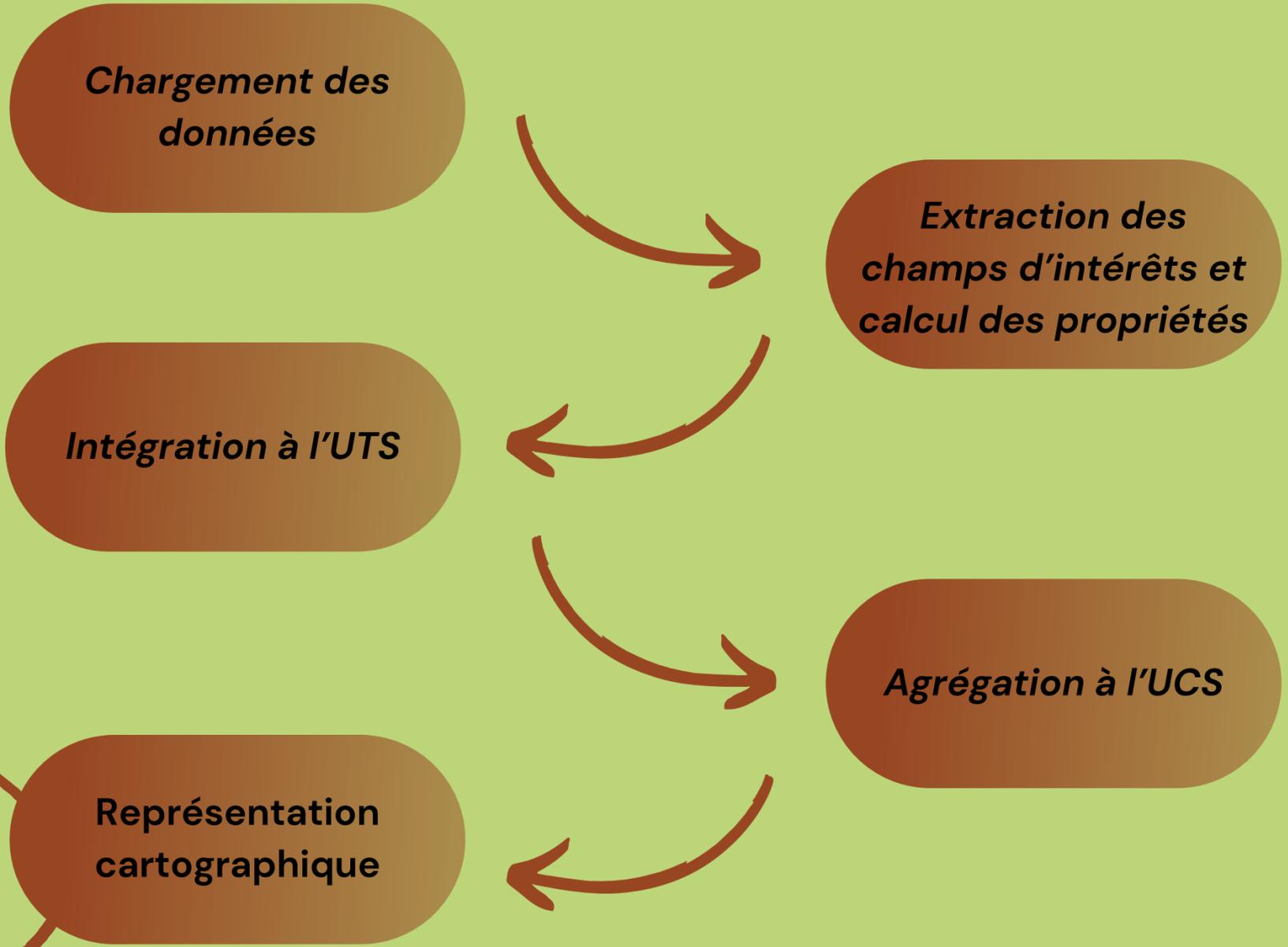
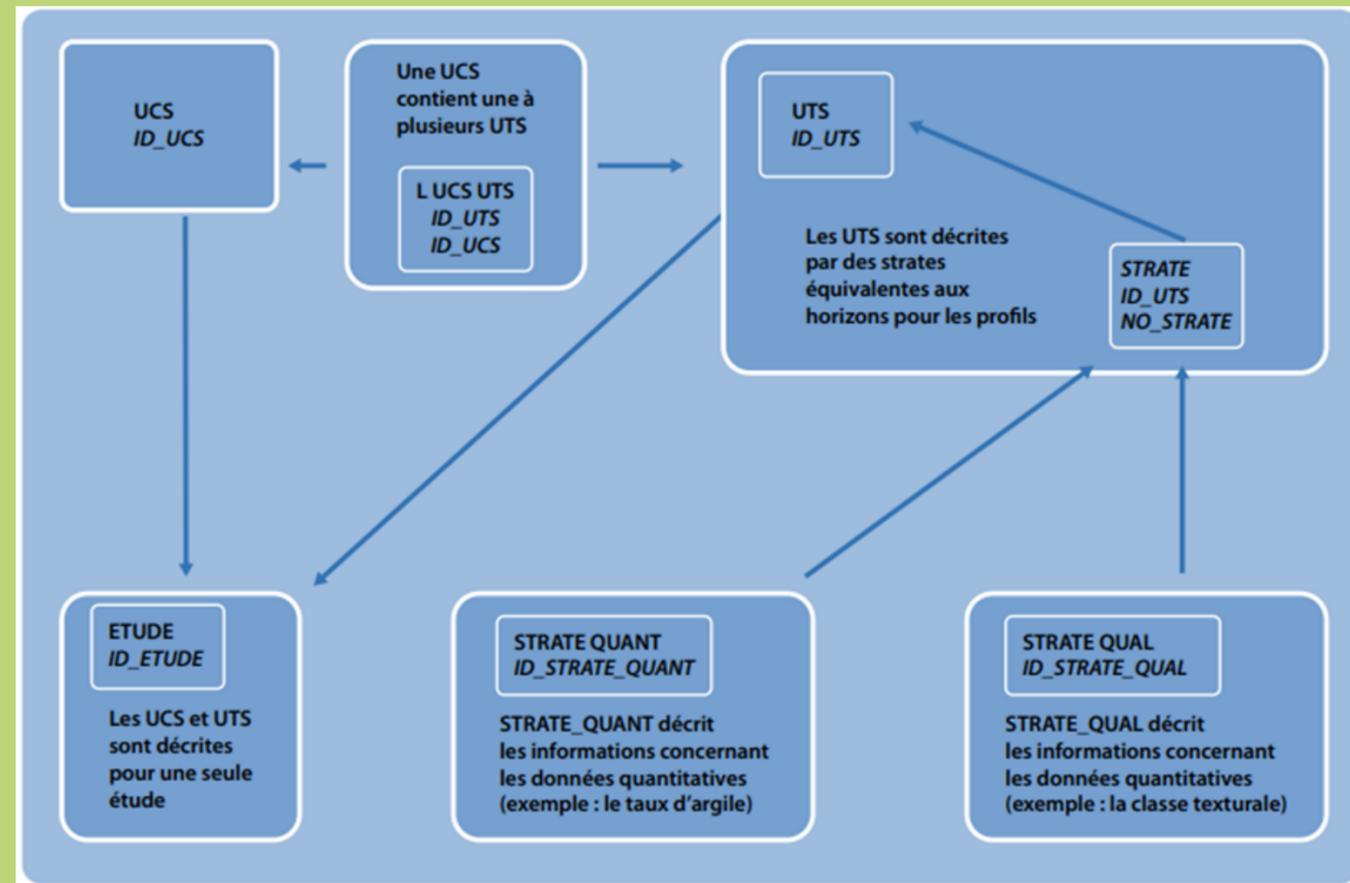
- 1 : 0 - 25 mm d'eau
- 2 : 25 - 50 mm d'eau
- 3 : 50 - 100 mm d'eau
- 4 : 100 - 150 mm d'eau
- 5 : \geq 150 mm d'eau
- -99 : -99, données insuffisantes pour estimer le RUM

On inclut la borne inférieure dans les classes

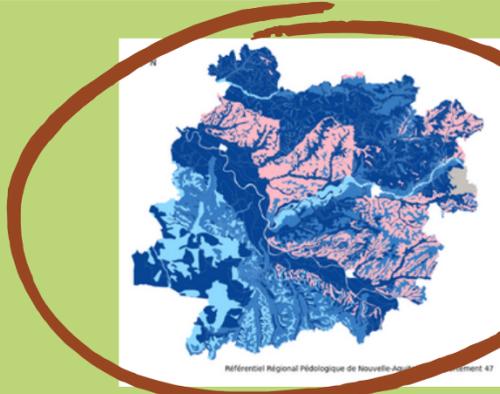
UCS	UTS	%	RUM (mm)	Classe de RUM
1	11	35	50	3
1	12	40	170	5
1	13	25	80	3

Le script de valorisation des données

Dernière étape



Application du script



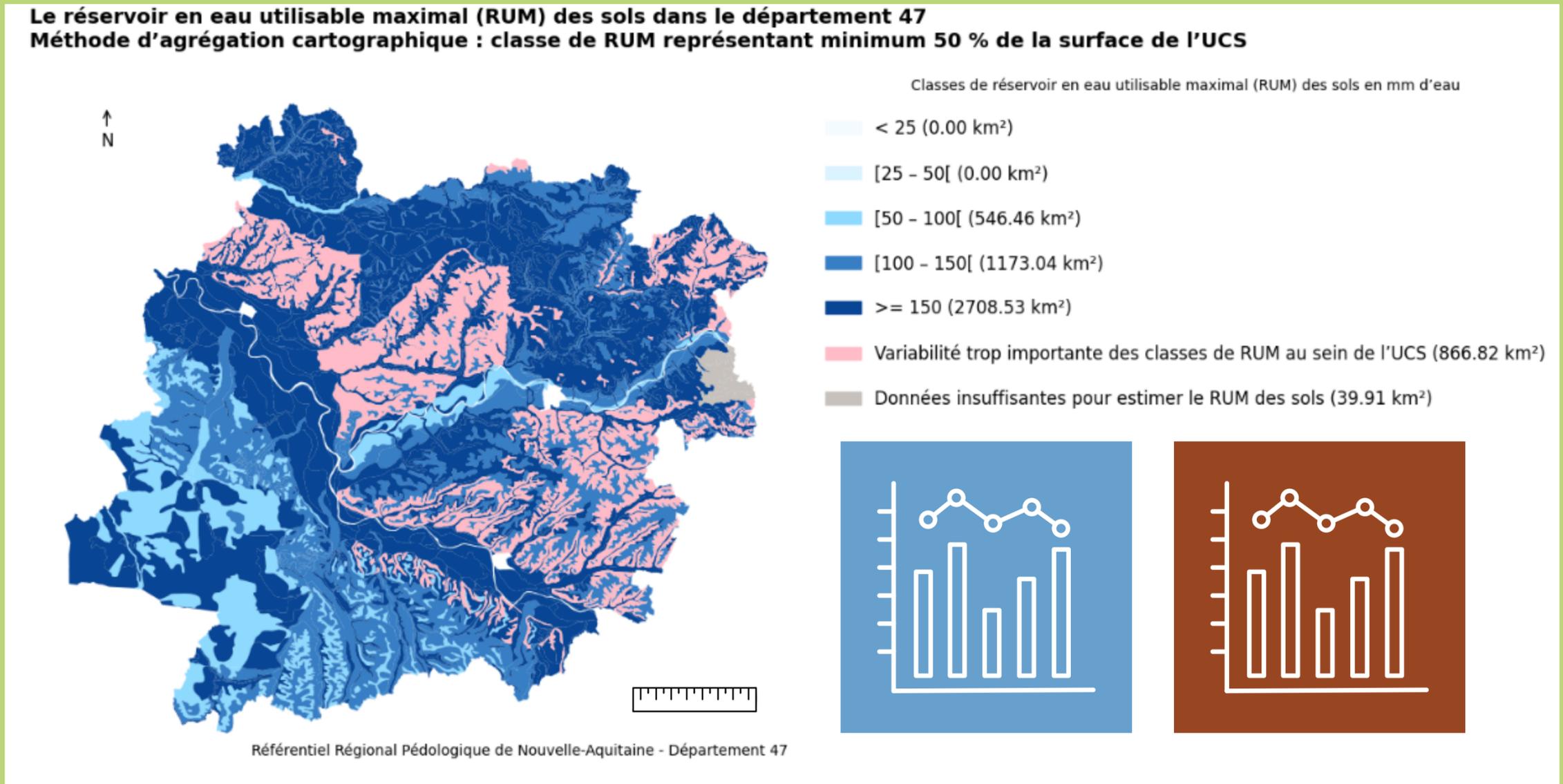
Le script de valorisation des données

Améliorations de la représentation cartographique

Objectif d'ajouter de la Datavisualisation

Quelques points à modifier encore (échelle, fond de carte)

Modules adaptables en fonction des choix des utilisateurs



Conclusions

Objectif 2 du projet Solid'R : Aide à la valorisation des données RRP

- Script python en 5 étapes, valorisant les données Donesol "en un clic"
- Actuellement RUM mais autres propriétés étudiables
- Jeux de données et carte valorisés exportables

Ces données demandent à être partagées



Mviewer Sols de Nouvelle-Aquitaine en cours de développement

Perspectives

Améliorer le script notamment pour produire les livrables prérequis au Mviewer :

- Fichier SLD
- Fiche de métadonnées

Publications dans la revue Etude et Gestion des Sols (revue française de la science des sols):

- Présentation du script
- Comparaison des méthodes de calcul du RUM et des différentes représentations cartographiques

Merci de votre attention !

Bibliographie

Gis Sol. (2023). Gis Sol - Présentation. Consulté le 08 04, 2023, sur Gis Sol - Partageons la connaissance des sols: <https://www.gissol.fr/le-gis>

Laroche, B., Richer de Forges, A., & Jalabert, S. (2023). À la découverte des sols : panorama des données, méthodes et outils existants. GeoDataDays 2023. Reims. Récupéré sur <https://hal.inrae.fr/hal-04208253>

Site du Géoportail, consulté le 15/10/2023, Gis Sol, RMT Sols et Territoires

Travail d'inventaire de Jalabert & Le Bas, 2022

Sols de Bretagne. (2015). Récupéré sur Sols de Bretagne: <https://sols-de-bretagne.fr/>

Laroche, B., Thiry, E. S., Chenu, J., Moulin, J., & Bardy, M. (2014). Les bases de données sur les sols au service des politiques publiques : exemple des zones humides. Etude et Gestion des Sols, 21, 37-49. Récupéré sur <https://hal.inrae.fr/hal-02637382/document>

Lemercier, B., Laroche, B., Armand, R., Chafchafi, A., Détriché, S., Ducommun, C., . . . Lehmann, S. (2017). Guide d'utilisation des bases de données sol pour la production de cartes thématiques. INRA.

Equipe Sols de Bordeaux Sciences Agro. (s.d.). Script de valorisation cartographique - Version du 17/10/2023.

Equipe Sols de Bordeaux Sciences Agro. (s.d.). Arbre Décisionnel - Version du 15/10/2023.

Equipe Sols de Bordeaux Sciences Agro. (s.d.). Mviewer Sols de Nouvelle-Aquitaine - Version du 17/10/2023.

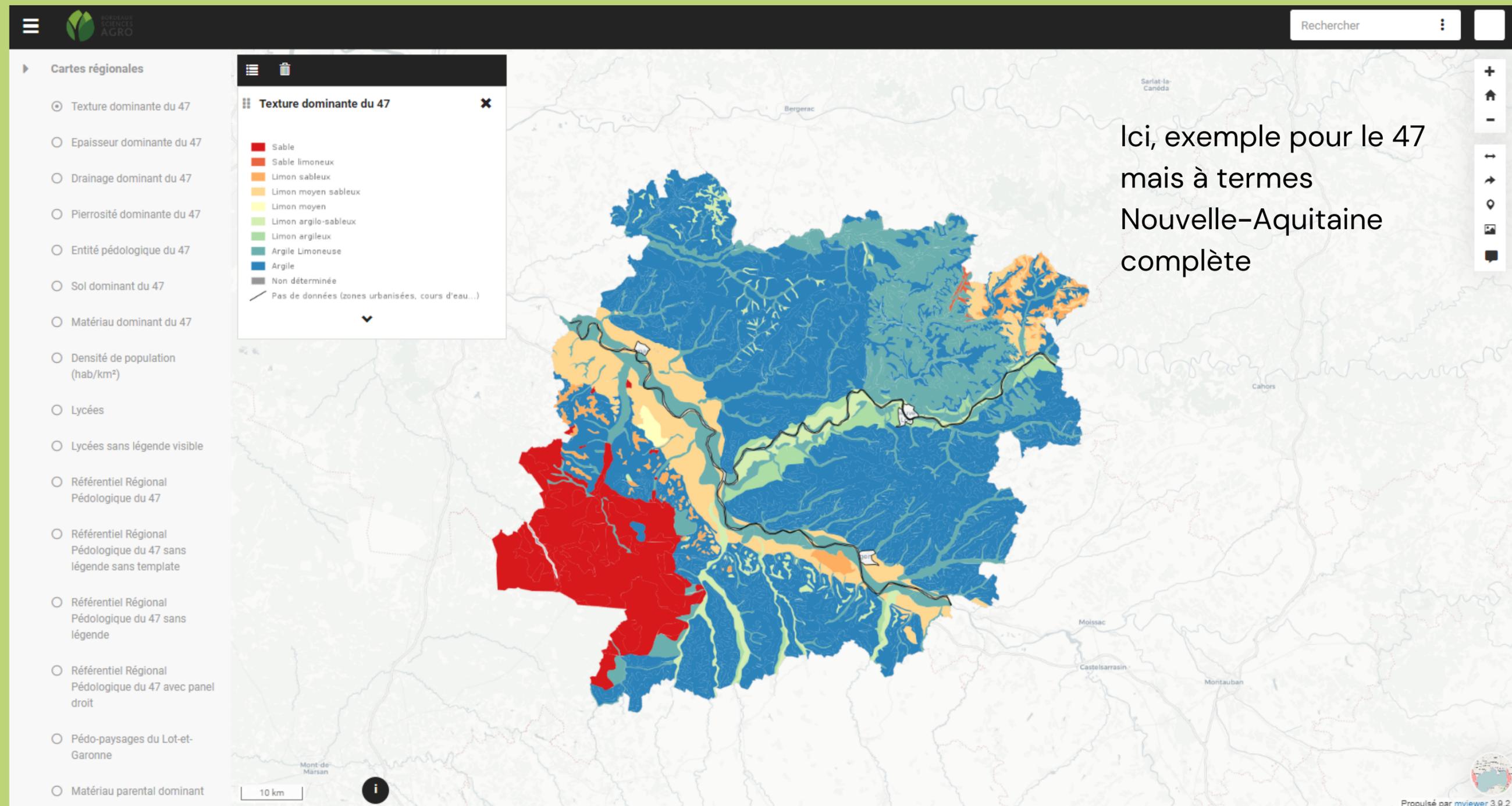
Mviewer Sols de Nouvelle-Aquitaine : première version

Mviewer Sols de Bretagne : code libre sur GitHub

Prérequis : la fiche de métadonnées de la donnée doit être publiée sur une IDG

Flux WMS avec fichier SLD pour la symbologie

Objectif d'ajouter des cartes issues du script sur le Mviewer : besoin d'exporter le fichier SLD du script



Ici, exemple pour le 47
mais à termes
Nouvelle-Aquitaine
complète