

NOV.
2020

ABC'TERRE-2A : APPLICATION PARTICIPATIVE ET APPROPRIATION PAR LES ACTEURS LOCAUX DE LA DEMARCHE ABC'TERRE* (2017-2020)

*ABC'Terre : Atténuation du Bilan gaz à effet de serre incluant le stockage Carbone dans les sols agricoles à l'échelle du Territoire.

Annexes

ADEMEAgence de l'Environnement
et de la Maîtrise de l'Energie

Un projet coordonné par



avec comme partenaires :



Annexe 2 : Guide méthodologique pour l'affectation d'une teneur en carbone organique aux types de sol d'un territoire à partir de la BDAT

Document réalisé par Xavier BOUSSELIN et Olivier SCHEURER (UniLaSalle), 2020

Table des matières

Annexe 2 : Guide méthodologique pour l'affectation d'une teneur en carbone organique aux types de sol d'un territoire à partir de la BDAT	2
Table des matières	3
Avant-propos	4
Introduction	5
1- Présentation de la démarche.....	7
a. Attribution des analyses aux UCS.....	7
b. Tri des analyses et regroupement des UTS en classes de teneurs en argile et en calcaire total : élaboration de méta-UTS	7
2- La méthode étape par étape	10
i. Réalisation des groupes d'UCS selon leur occupation du sol	12
ii. Réalisation des groupes d'UCS selon leur composition en UTS.....	12
c. Demande d'extraction de la BDAT à InfoSol	15
e. Regroupement en classes des analyses et des UTS	16
i. Classification des UTS en méta-UTS	16
ii. Classement des analyses.....	16
iii. Attribution des analyses aux méta-UTS.....	16
f. Calcul des médianes.....	16
i. Calcul de la médiane pour chaque couple [méta-UTS, UCS]	16
ii. Calcul de la médiane pour chaque couple [méta-UTS, meta-UCS]	16
iii. Calcul de la médiane pour chaque méta-UTS à l'échelle du territoire.....	17
iv. Validation agronomique	17
3- Validité et limites de la méthode	18
Bibliographie	20
Annexe 1 : Données « Sol » à extraire du RRP pour la mise en œuvre de la méthode ABC'Terre.	21
Annexe 2 : Exemple de construction de méta-UCS dans le Tardenois	24
Annexe 3 : Evaluation de la teneur en argile à partir de la CEC et de la teneur en carbone organique	32

Avant-propos

Le présent guide a pour objectif d'aider le lecteur à comprendre et mettre en œuvre la méthode d'affectation de teneurs en carbone organique aux types de sol d'un territoire à partir de la Base de Données d'Analyses des Terres (BDAT¹), telle que proposée dans la démarche ABC'Terre.

L'ensemble des éléments présentés et discutés dans ce guide sont établis à partir des données fournies par l'INRA InfoSol (N. Saby) en 2018. Ces données sont issues des analyses de 2005 à 2014 stockées dans la BDAT, les données postérieures à 2014 n'étant pas disponibles.

L'ensemble des éléments relatifs au logiciel RPG Explorer présentés et discutés dans ce guide sont valables pour la version 1.9.5.

¹ BDAT (Base de Données d'Analyses des Terres) : cette base contient des résultats d'analyses de terre réalisées par les agriculteurs, référencées par commune. Ces analyses sont fournies par les laboratoires à Inrae (Unité InfoSol) et contiennent entre 1 et 31 variables qui décrivent les propriétés physico-chimiques des terres dans l'horizon de surface, dont la teneur en carbone organique.

Introduction

Le stockage de carbone dans les sols agricoles pourrait contribuer à l'atténuation du changement climatique. Cette piste d'action a notamment été mise en avant par l'initiative 4 pour 1000, lancée par la France en 2015, lors de la Cop 21. Néanmoins, pour cibler et lancer des actions en faveur du stockage de carbone à l'échelle locale, il est nécessaire d'établir un diagnostic de la situation actuelle.

La démarche ABC'Terre réalise ce diagnostic à l'échelle d'un territoire. Pour ce faire, l'évaluation du stockage ou du déstockage du carbone dans les sols agricoles est réalisée grâce à l'outil SIMEOS AMG. Celui-ci simule l'évolution à long terme du stock de carbone organique de l'horizon de surface sous l'effet d'un système de culture dans un contexte pédoclimatique donné. Chaque couple [rotation, UTS²] inventorié sur le territoire fait l'objet d'une simulation et l'UTS concernée doit être caractérisée par une teneur en carbone organique représentative de la situation actuelle. Cette teneur en Carbone organique initiale est une donnée d'entrée à laquelle le modèle est particulièrement sensible.

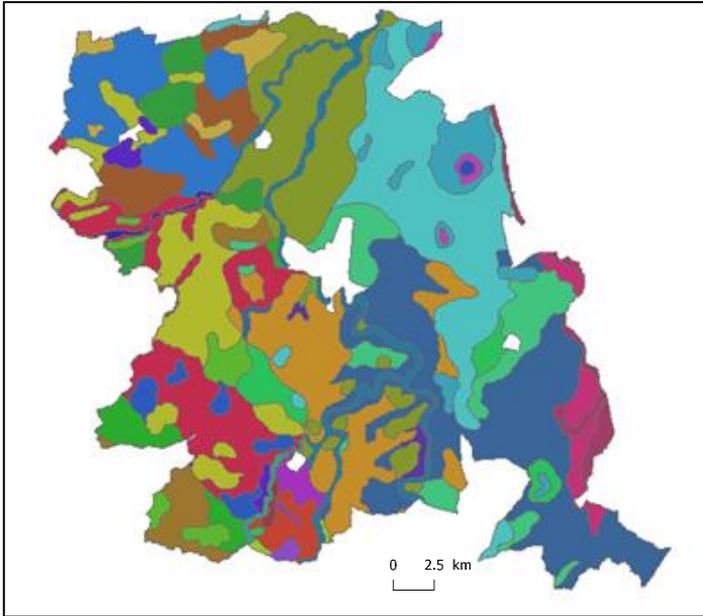
Les bases de données des RRP³ attribuent à chaque UTS une teneur en Carbone organique modale dans l'horizon de surface, mais cette donnée provient souvent d'analyses anciennes et/ou peu nombreuses pour chaque UTS ; sa représentativité a donc été considérée comme insuffisante. A l'opposé, la BDAT fournit des données actualisées et en général nombreuses, mais celles-ci ne sont pas associées explicitement à des types de sol.

La méthode présentée dans ce guide a été élaborée afin de résoudre ce problème : les couples [rotation, UTS] étant reconstitués au niveau spatial de l'Unité Cartographique de Sol (UCS⁴), elle cherche à attribuer une teneur en C organique à chaque couple [UTS, UCS] en exploitant les analyses de la BDAT.

² UTS (Unité Typologique de Sol) ou type de sol

³ RRP (Référentiel régional pédologique) : base de données sur les sols associée à une représentation cartographique au 1/250 000^e. Elle est gérée au niveau régional ou départemental, sous un format national unique (DONESOL) et couvre la quasi-totalité du territoire métropolitain.

⁴ UCS (Unité Cartographique de Sol) : zone géographique délimitée, caractérisée par une composition spécifique en UTS. Voir **Erreur ! Source du renvoi introuvable.**



*Le territoire contient 33 UCS.
Chaque UCS est représentée
par une couleur.*

*Au sein de chaque UCS la
composition en UTS (nature et
proportions) est connue, mais
les UTS ne sont pas
localisées.*

Figure 1 : Représentation du territoire d'étude du Thouarsais (79)

1- Présentation de la démarche

L'affectation d'une analyse de la BDAT à une UTS au sein d'une UCS du territoire d'étude est réalisée sur la base de deux critères : i) un **critère analytique** de similitude des teneurs en argile et en calcaire entre l'analyse et l'UTS, ii) un **critère spatial** d'intersection entre les territoires de la commune de l'analyse et de l'UCS concernée. La démarche globale est schématisée dans la figure 2. Elle se décompose en 3 étapes :

- (a) Attribution des analyses aux UCS
- (b) Tri des analyses et regroupement des UTS en classes de teneurs en argile et en calcaire total, dans chaque UCS
- (c) Affectation des teneurs en carbone organique des analyses aux UTS dans chaque classe et attribution de la médiane aux couples [UTS, UCS]

a. Attribution des analyses aux UCS

La première étape consiste à attribuer les analyses de la BDAT aux UCS. Le croisement spatial entre les contours des communes et ceux des UCS est réalisé à l'aide d'un SIG. La règle utilisée ensuite est que toutes les analyses d'une commune sont attribuées à chacune des UCS qu'elle contient. De ce fait, une même analyse peut être attribuée à plusieurs UCS.

Préalablement, on a sélectionné dans la BDAT les analyses qui renseignent à la fois la teneur en C organique et les teneurs en argile et en calcaire total (CaCO_3) ; celles-ci ne représentent qu'une faible part de l'ensemble des analyses disponibles (de l'ordre de 25%). Dans certains cas, s'il n'y a pas assez d'analyses remplissant ces conditions (teneur en argile manquante), il est possible d'utiliser les analyses – beaucoup plus nombreuses – renseignant la CEC ; les teneurs en argile sont alors estimées à partir de la CEC et de la teneur en C organique. Pour plus d'informations sur ce point, se référer à l'*Annexe 1 : Données « Sol » à extraire du RRP pour la mise en œuvre de la méthode ABC'Terre*.

Rappel des opérations mobilisant les données SOL dans la méthode ABC'Terre :

- Pour l'affectation des rotations aux types de sols via RPG Explorer :
 - o Croisement spatial UCS x ilots de cultures du RPG
 - o Attribution de notes de valeur agronomique aux combinaisons « UTS x rotation de cultures » en fonction des propriétés (contraintes) agronomiques des UTS
- Pour l'affectation d'une teneur initiale en Corg aux UTS, à partir de la BDAT, éventuellement différenciée selon les UCS:
 - o Croisement spatial UCS x communes puis distribution des analyses communales de la BDAT dans les UCS.
 - o Croisement spatial UCS x ilots de culture du RPG pour classification des UCS selon leur type d'occupation du sol (assolement moyen ou autre indicateur)
 - o Classification des UCS selon leur composition en UTS
 - o Regroupement des UTS en classes de teneur en argile et Calcaire dans la strate de surface.
 - o Affectation d'une teneur en Corg par classe d'UTS
- Pour reconstituer les pratiques culturales à l'échelle du territoire :

- Estimation du rendement des cultures en fonction des classes de RUM
- Pour Simeos-AMG :
 - Le C/N et le pH sont deux variables prises en compte dans la fonction de minéralisation k d'AMG (Clivot et al, 2017 et 2019)
- Pour le calcul des émissions GES :
 - Les facteurs d'émissions par volatilisation des apports d'engrais minéraux sont fonction du pH du sol (EMEP EEA, 2016, Tier 2)

Données géographiques nécessaires :

- Shapefile des UCS contenues dans le territoire d'étude, y compris le prolongement des polygones hors des communes du territoire.

Données sémantiques nécessaires :

- Description des UCS :
 - Nom de l'UCS
 - UTS présentes et leur pourcentage surfacique
- Description des UTS :
 - Nom de l'UTS
 - Occupation dominante et secondaire
 - Classe de drainage naturel
 - Profondeur modale
 - RUM estimée si possible
- Caractéristiques de la strate de surface :
 - Texture dominante
 - Taux d'argile modal
 - Taux de calcaire total modal
 - Abondance des éléments grossiers modale
 - CEC et teneur en C organique modales
 - pH eau et rapport C/N

Données sémantiques supplémentaires si RUM non estimée :

Caractéristiques des strates de chaque UTS, soit pour chaque strate :

- N° de strate
- Texture Aisne (ou à défaut Taux d'Argile, Limon, Sable modaux)
- Epaisseur moyenne
- Abondance des éléments grossiers
- Nom des éléments grossiers

Tableau récapitulatif des variables et de leur usage

	Nom de variable dans DONESOL 3	Utilisations				
		Table des valeurs Agronomiques	Classification des UCS	Classement des UTS	Estimation RUM	Détermination rendements, fonction minéralisation (AMG) et volatilisation Nmin
UCS						
Nom de l'UCS	NOM_UCS		x			
UTS						
UTS et pourcentage	POURCENT		x			
Occupation dominante et secondaire	OCCUP1 OCCUP2 (ds L_UCS_UTS)	x	x			
Nom de l'UTS	NOM_UTS		x			
Classe de drainage naturel	DRAI_NAT	x	x			
Profondeur modale	PROF_SOL_MOD		x		x	
RUM	A estimer	x	x			x
STRATES						
Epaisseur moyenne	EPAIS_MOY				xxx	
Texture dominante	TEXTURE (AISNE)		x		xxx	
Taux d'argile	TAUX ARGILE (VAL_MOD)	x	x	x		
Taux de calcaire total	CALC_TOT (VAL_MOD)	x	x	x		
Abondance des éléments grossiers	ABONDANCE_EG (VAL_MOD)	x	x		xxx	
Nom des éléments grossiers	NOM_EG				xxx	
CEC	CEC			x		
Taux de carbone organique	CARBONE			x		
pH eau	PH_EAU					x
rapport C/N	C_N					x

Pour les variables des strates : x = pour la strate de surface xxx = pour chaque strate

)

Cette étape est intégralement prise en charge par Inrae InfoSol, car la base de données localisée au niveau communal est confidentielle. Les analyses fournies comprennent finalement quatre variables : le code de l'UCS, les teneurs en C organique, en argile (analysée ou estimée) et en calcaire total.

b. Tri des analyses et regroupement des UTS en classes de teneurs en argile et en calcaire total : élaboration de méta-UTS

Les UTS sont regroupées en classes de teneur en calcaire total et en argile afin de permettre leur rattachement aux analyses de terre sélectionnées. Ces deux variables comptent parmi les déterminants majeurs du taux annuel de minéralisation du carbone organique des sols ; on admet donc que les groupes d'UTS ainsi constitués sont relativement homogènes de ce point de vue. Ces **groupes d'UTS** définis au sein de chaque UCS seront par la suite dénommés « **meta-UTS** ».

En parallèle, les analyses de terre sont triées en fonction des mêmes classes de teneur en argile et en calcaire total. Pour une UCS donnée, on obtient donc une distribution de teneurs en carbone organique pour chacune des meta-UTS qu'elle contient.

c. Affectation des teneurs en carbone organique des analyses aux UTS et attribution de la médiane aux couples [UTS, UCS]

On admet l'hypothèse qu'au sein d'une UCS, la distribution des teneurs en carbone organique varie peu entre les UTS d'une même meta-UTS ; la variabilité observée dans une meta-UTS résulterait majoritairement de la diversité des histoires culturelles entre parcelles cultivées.

A l'opposé, pour une même meta-UTS (même classe de teneurs en argile et calcaire) présente dans plusieurs UCS, les distributions de teneurs en carbone organique pourraient différer entre UCS. Ces différences pourraient s'expliquer par des spécificités de ces UCS susceptibles d'avoir joué sur les teneurs en carbone organique: histoire de leur occupation du sol (retournement de prairies, évolution des systèmes de culture) ou propriétés des sols (hydromorphie, charge en éléments grossiers, ...).

Conformément à ces hypothèses, l'affectation des teneurs en carbone organique des analyses aux UTS se fait selon la procédure suivante :

- (i) toutes les analyses associées à une meta-UTS sont affectées à chacune des UTS qui la composent, au niveau de chaque UCS
- (ii) chaque couple [UTS, UCS] se voit attribuer la médiane de la distribution des teneurs en carbone organique ainsi obtenue ; on admet que cette médiane est représentative si l'effectif des analyses affectées à une meta-UTS est supérieur ou égal à 10
- (iii) dans les cas -minoritaires- où l'effectif d'une meta-UTS est insuffisant, on cherche à augmenter le nombre d'analyses exploitables, en regroupant les analyses d'UCS présentant des caractéristiques similaires ; on établit pour cela une typologie des UCS sur la base de deux critères : occupation du sol de l'UCS, propriétés des sols. Les **groupes d'UCS** ainsi constitués seront par la suite **dénommés** « **meta-UCS** ». Si l'effectif reste insuffisant malgré ce

regroupement, on utilise par défaut l'ensemble des analyses du territoire, toutes UCS confondues pour calculer la médiane de la meta-UTS.

Les analyses utilisables **pour chaque meta-UCS** doivent être obtenues **à partir d'un croisement géographique spécifique** « communes x meta-UCS » qui doit être réalisé par InfoSol. Dans une meta-UCS donnée, la simple fusion des lots d'analyses attribués à chaque UCS n'est pas possible car elle comptabiliserait plusieurs fois une partie des analyses.

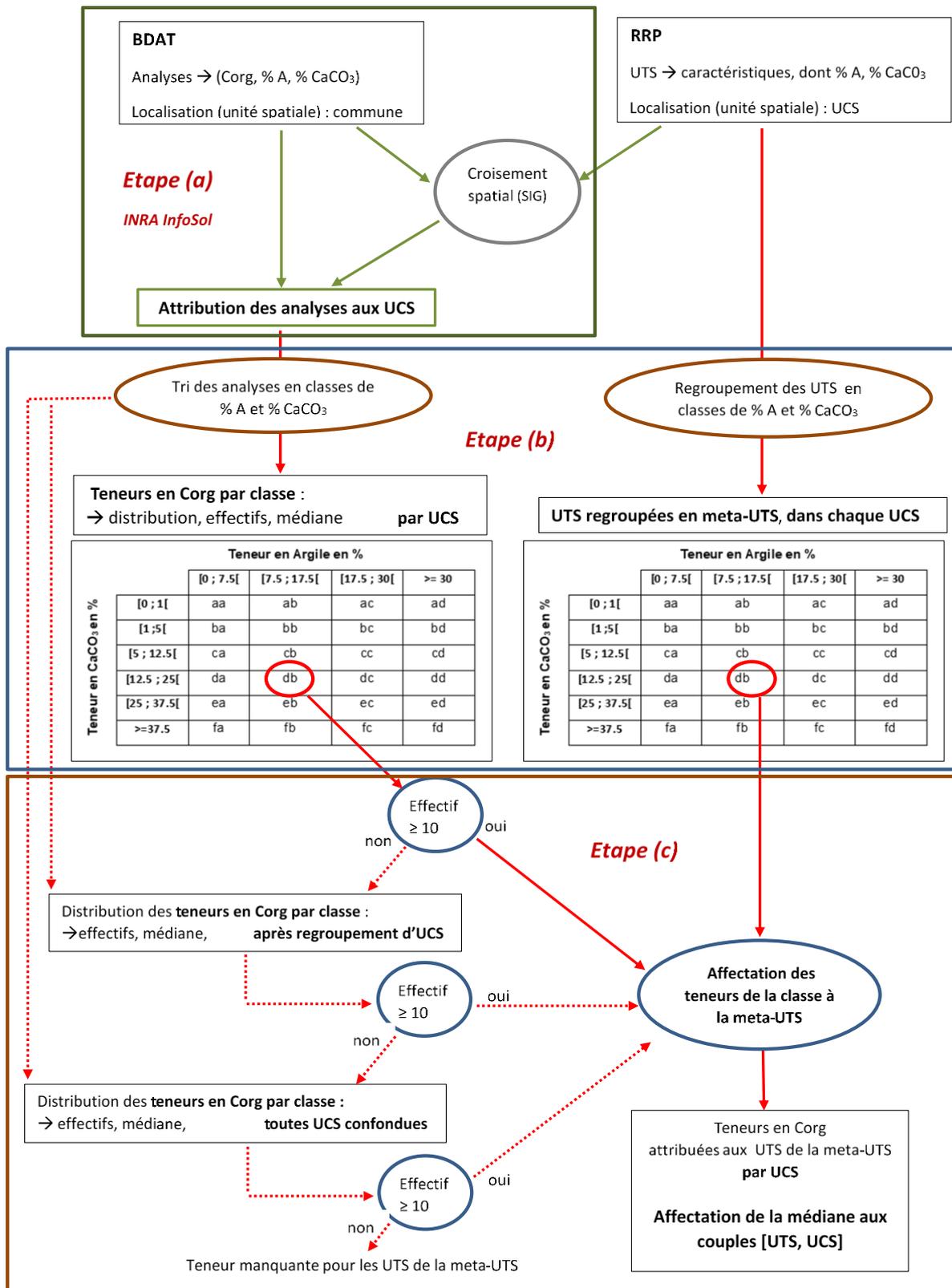


Figure 2 : Schéma de la démarche en 3 étapes

2- La méthode étape par étape

Ce chapitre comporte des redites par rapport au précédent ; l'idée est ici de présenter les étapes dans un ordre et sous un angle pratique et non conceptuel comme dans la partie précédente.

a. Collecte des données

Données SOL :

Les Référentiels Régionaux pédologiques (RRP) constituent la source d'information géographique sur les sols disponible sur la quasi-totalité du territoire métropolitain. Les données doivent être demandées au gestionnaire régional (ou départemental) du RRP concerné par le territoire. La mise à disposition de ces données est conditionnée par la signature d'une licence d'utilisation accordée par le gestionnaire. Des frais minimes de mise à disposition sont en général facturés par le gestionnaire. La carte du découpage national des RRP ainsi que les coordonnées des gestionnaires à contacter sont consultables sur le Geoportail (<https://www.geoportail.gouv.fr/donnees/carte-des-sols>).

Deux types de données sont nécessaires :

- *des données géographiques* constituées par les contours des UCS au format SIG (shapefile)
- *des données sémantiques* comprenant d'une part la description de chaque UCS par son nom et par la liste et les parts surfaciques des UTS qu'elle contient, d'autre part les caractéristiques de chaque UTS.

La demande de mise à disposition des **contours des UCS** ne serait pas strictement nécessaire pour l'affectation des teneurs en carbone aux types de sol car cette donnée est directement accessible pour InfoSol sur l'ensemble du territoire national ; InfoSol peut donc réaliser l'étape (a) dès lors qu'il connaît les limites du périmètre d'étude (numéro INSEE des communes). Cependant, cette donnée est indispensable à l'utilisation de RPG Explorer dans l'étape de reconstitution des assolements de rotations sur le territoire de mise en œuvre de la démarche ABC'Terre. Elle doit donc faire partie des données à demander.

L'annexe 1 donne la liste détaillée des données nécessaires, à communiquer au gestionnaire RRP pour l'établissement de la licence de mise à disposition.

Données du Registre Parcellaire Graphique (RPG) :

Pour pallier les situations où le nombre d'analyses est trop faible, les UCS doivent être regroupées en meta-UCS, en partie sur la base d'indicateurs de l'occupation du sol (et en partie d'après les propriétés des sols). Les données du RPG sont donc nécessaires pour calculer des indicateurs de l'occupation du sol actuelle caractérisant chaque UCS. Il est nécessaire de pouvoir utiliser les données de 3 années culturales successives.

Le calcul des indicateurs est réalisé à l'aide de l'outil RPG Explorer. Pour son installation se référer à la notice de cet outil disponible en ligne sur :

https://tice.agroparistech.fr/coursenligne/courses/RPGEXPLORER/?id_session=0.

b. Création des meta-UCS

L'élaboration des méta-UCS est réalisée en premier lieu, car cela permet de n'interagir qu'une fois avec InfoSol pour extraire les données de la BDAT⁵.

L'objectif de cette classification est de regrouper les UCS en un nombre limité de classes, relativement homogènes (et nettement différenciées entre elles) selon plusieurs facteurs de variation majeurs des teneurs en carbone organique dans les sols (tableau 1) :

- teneurs en argile, CaCO₃, et éléments grossiers dans l'horizon de surface, excès d'eau, RUM, dans les sols
- systèmes de culture selon leur niveau de restitutions humiques dans l'histoire culturale.

Pour ce faire, deux classifications préliminaires des UCS sont réalisées, respectivement sur la base de leur composition en UTS et de leur occupation du sol par les cultures. La classification finale résulte du croisement de ces deux classifications.

Tableau 1 : Facteurs de variation de la teneur en Corg des sols à prendre en compte dans la classification des UCS

Facteur de variation	Effet théorique sur le carbone organique dans les sols	Sens de variation attendu
Teneur en argile des sols (strate de surface)	<i>réduction de la vitesse de minéralisation</i>	La teneur en carbone organique du sol est croissante avec les facteurs observés
Teneur en calcaire des sols (strate de surface)	<i>réduction de la vitesse de minéralisation</i>	
Charge en éléments grossiers (strate de surface)	<i>masse de terre fine plus faible induisant une teneur en Corg plus élevée à bilan humique équivalent</i>	
Excès d'eau (hydromorphie)	<i>réduction de la vitesse de minéralisation</i>	
RUM des sols	<i>production de biomasse et niveau de restitution humique plus élevés</i>	
Part de cultures à restitutions humiques élevées dans l'assolement (colza, céréales à paille, maïs grain, prairies temporaires)	<i>niveau moyen plus élevé des restitutions humiques des systèmes de culture</i>	

⁵ L'extraction des analyses utilisables pour chaque meta-UCS doit être réalisée par InfoSol à partir d'un croisement géographique spécifique (communes x meta-UCS). La simple fusion des lots d'analyses attribués à chaque UCS n'est pas possible car elle comptabiliserait plusieurs fois les mêmes analyses dans une meta-UCS donnée. Si le n° des UCS qui composent chaque meta-UCS créée est communiqué à InfoSol en même temps que les limites du territoire d'étude, il lui est possible de réaliser d'emblée les deux croisements géographiques nécessaires pour affecter les analyses : communes x UCS, communes meta-UCS.

Il est possible de réaliser ces classifications à **direx d'experts ou par une méthode statistique classique proposée par un logiciel tel que R.**

Les indicateurs d'occupation du sol (calculés avec RPG Explorer) sont indispensables si l'on souhaite classer les UCS à l'aide d'une méthode statistique. Ils sont aussi utiles aux experts comme complément d'information si la méthode « à direx d'experts » est choisie.

i. Réalisation des groupes d'UCS selon leur occupation du sol

Comme évoqué précédemment, un indicateur de l'occupation du sol est utilisé pour réaliser des groupes d'UCS les plus homogènes possibles et nettement différenciés entre eux. On fait l'hypothèse que des UCS semblables par leur occupation du sol actuelle ont vraisemblablement subi des histoires culturelles globalement semblables et que la différenciation des groupes est héritée du passé et suffisamment ancienne pour avoir marqué l'évolution des teneurs en carbone dans les sols.

- *Choix et extraction de l'indicateur*

Les indicateurs sont générés avec RPG Explorer. Le choix de l'indicateur dépend du contexte local et de la manière dont RPG Explorer a été utilisé dans les étapes de reconstitution des assolements de rotation.

Il est par exemple possible d'utiliser les séquences de cultures reconnues sur 3 ans au niveau parcellaire (triplets) ou la part surfacique des différents types d'exploitations (Résultat de l'étape 7, voir notice de l'outil), ou l'assolement moyen calculé sur plusieurs années (voir notice de l'outil : le module « indicateur » permet de connaître l'assolement par année). L'indicateur « triplets » semble plus robuste. Il est utilisé dans la suite de ce guide.

Pour générer les triplets, il est préférable de ne pas utiliser les correspondances groupe de culture - culture, pour éviter d'introduire un biais inutile dans les résultats.

- *Classer les triplets*

Dans un premier temps, pour orienter la classification qui suivra, on fait des groupes de triplets de cultures en fonction de leur composition considérée en termes d'effet théorique sur la teneur en C organique des sols.

Ensuite, on définit la part surfacique de chacun de ces groupes dans les UCS.

- *Regrouper les UCS*

Sur la base de ces parts surfaciques, on peut faire une classification à direx d'expert ou utiliser un outil statistique. Il est par exemple possible de faire une classification ascendante hiérarchique (CAH).

Il est important de noter que, dans certains territoires, l'histoire de l'occupation des sols est très homogène (ou l'hétérogénéité existante n'est pas bien segmentée par les UCS). Dans ce cas il est possible de ne faire qu'une classe d'occupation du sol, ce qui revient à ne pas utiliser de critère dans la construction des méta-UCS.

ii. Réalisation des groupes d'UCS selon leur composition en UTS

Cette classification peut se faire soit à direx d'expert à partir du nom des UCS, soit par une méthode statistique en suivant les indications ci-dessous.

- *Classer les UTS en types*

On fait dans un premier temps des regroupements d'UTS en types agronomiques présentant des propriétés communes vis-à-vis de l'évolution du carbone organique. On

utilise pour cela les variables définies précédemment (teneur en argile, éléments grossiers, CaCO₃, réservoir utilisable maximal et excès d'eau).

S'il existe une typologie agronomique régionale TYPTERRES, il est plus simple d'utiliser directement les groupes d'UTS définis par les UTtyperres ; ces derniers sont en effet basés sur les mêmes propriétés des sols.

Les groupes peuvent à défaut être construits grâce à une CAH. Le fichier d'entrée de RPG Explorer qui décrit la composition des UTS peut être utilisé directement. Néanmoins, on veillera à réduire préalablement les variables car elles n'ont pas la même unité.

On cherche ici à ne définir qu'un petit nombre de types.

- *Classer les UCS*

Le regroupement des UCS se fait en fonction de la part surfacique des types (groupes d'UTS) définis ci-dessus. Si le RRP à une structure simple (un type agronomique très dominant par UCS) alors on regroupe les UCS par grand type majoritaire. Si le RRP à une structure complexe (nombreux types agronomiques par UCS en proportions variables) alors on réalise une CAH des UCS à partir de la part surfacique des types constitués.

iii. Croisement des deux classifications

Les meta-UCS sont créées par le croisement des deux classifications.

iv. Exemples d'application

Le tableau 2 donne à titre d'exemple les 2 classifications réalisées dans 3 territoires-tests. Dans chaque territoire, les 3 indicateurs d'occupation du sol évoqués plus haut (assolement, triplets, typologie) ont été comparés; ils montrent des performances similaires.

L'**annexe 2** donne un exemple de mise en œuvre de la méthode statistique (CAH) pour les 2 classifications dans une petite région naturelle de l'Aisne (Tardenois).

La classification en meta-UCS n'est cependant pas toujours nécessaire : la caractérisation des meta-UTS effectuée globalement au niveau du territoire (toutes UCS confondues) peut suffire si les classes de teneurs en argile et calcaire définissant les meta-UTS sont fortement corrélées aux autres critères d'intérêt pour la classification des UCS (charge en cailloux, hydromorphie, successions de cultures dominantes).

Tableau 2 : Description des classes constituées pour créer les méta-UCS dans 3 territoires (avec 3 exemples d'indicateur pour l'occupation du sol dans chaque territoire).

Territoire	Type de classe	Indicateur utilisé	Classe d'UCS	Caractéristiques
Tardenois	Sols	Composition des UCS en UTS et propriétés des UTS	1	Sol très dominant: +RUM, +Hydro, -Cx, -CaCO3
			2	Sol très dominant: -Argile, -Cx,-CaCO3
			3	Sols dominants avec: ++Argile ou ++Cx, ++CaCO3 et +Argile (mais 40% de sols : +RUM, +Hydro, -Cx, -CaCO3)
	Occupation du sol	Assolement	1	+Betteraves, - Prairies
			2	-Betteraves
			3	+Prairies, -blé
		Triplets	1	+ de triplets colza, + protéagineux, + monoculture de céréales, - betteraves
			2	+ de triplets betteraves, - colza, - protéagineux, - monoculture de céréales
			Typologie	1
	2	+de betteraviers spécialisés		
Pays RVGB	Sols	Type d'UTS dominant dans l'UCS	1	Sol très dominant: -RUM, -Hydro, +Cx, +CaCO3
			(2)	(non représenté sur le territoire)
			3	Sol très dominant: ++Hydro, +CaCO3
			4	Sol très dominant: -RU, -Hydro,+Cx, -CaCO3
			5	Sol très dominant: +RUM,-Cx, -CaCO3
	Occupation du sol	Assolement	1	+maïs
			2	-maïs
			Triplets	1
		2		+ de triplets avec maïs et céréales, - de monoculture de maïs
		Typologie	1	+ de céréaliers, -viticulteurs
2	+viticulteurs, - de céréaliers			

Thouarsais	Sols	Substrat	1	UCS du socle ancien
			géologique	2
Occupation du sol	Assolement		1	+blé, +tournesol, +colza, +orge, -maïs, -prairies
			2	+prairies, -blé, -tournesol, -colza
			3	+vignes
			4	++maïs
	Triplets		1	+ de triplets maïs, (+vigne)
			2	++ de triplets colza, +légumineuses, +tournesol, -maïs, -prairies,
			3	++ prairies, - de triplets tournesol, -colza
Typologie		1	+céréaliers et + céréaliers diversifiés	
		2	++ de polyculteurs et + d'éleveurs	

Les signes ++, +, -, précédant les caractéristiques des sols ou de l'occupation du sol indiquent la spécificité de la classe d'UCS par rapport à la moyenne du territoire. Exemples : + RUM = RUM supérieure à la moyenne ; ++ hydro = hydromorphie très supérieure à la moyenne ; - Cx = pierrosité inférieure à la moyenne ; + maïs = fréquence du maïs supérieure à la moyenne.

c. Demande d'extraction de la BDAT à InfoSol

Pour faire une demande d'extraction des analyses de la BDAT à InfoSol, il est nécessaire d'envoyer deux fichiers Excel : un qui liste les codes INSEE des communes du territoire d'étude et un autre qui donne les correspondances entre les numéros des UCS du RRP et les méta-UCS auxquels elles appartiennent.

Un point de vigilance important : il faut s'assurer que le territoire d'étude n'a pas été sujet à des fusions de communes. Si c'est le cas, intégrer à la liste des codes INSEE le code de la « commune nouvelle » (ce code est celui de l'une des anciennes communes) et les anciens codes INSEE des anciennes communes.

InfoSol a directement accès aux contours des UCS des RRP sur l'ensemble de la France. A l'issue du croisement géographique, trois fichiers sont livrés :

- analyses par UCS
- analyses par meta-UCS
- analyses sur l'ensemble du territoire (analyses sans croisement géographique avec les UCS)

Chaque fichier contient les analyses « complètes » (avec 4 variables renseignées : argile, CaCO₃, Corg, CEC) et les analyses « incomplètes » (le taux d'argile est manquant).

d. Estimation du taux d'argile par une fonction de pédotransfert

Si le nombre d'analyses utilisables est insuffisant, le lot d'analyses « complètes » peut être utilisé pour construire statistiquement une fonction de pédotransfert locale qui estime le taux d'argile dans le lot d'analyses incomplètes, à partir de la CEC et du Corg.

La méthode est détaillée dans l'**Annexe 3**. Il serait possible que cette étape soit directement réalisée en amont par InfoSol, mais elle nécessite de s'appuyer sur une expertise pédologique locale.

e. Regroupement en classes des analyses et des UTS

i. Classification des UTS en méta-UTS

Chaque UTS est classée dans une méta-UTS selon ses classes modales de teneur en argile et en calcaire. Les bornes définissant ces classes sont présentées dans le Tableau 3. Le Tableau 3 présente les bornes qui définissent les classes de teneur en CaCO₃ et en argile et les codes attribués aux méta-UTS correspondantes. Les bornes utilisées pour définir ces classes sont définies par les seuils texturaux pour l'argile et aux seuils classiques dans la bibliographie (classes GEPPA in Baize, 2018) pour le calcaire. Il n'est pas utile de les modifier. Il est cependant possible de regrouper des classes voisines de teneurs en calcaire si le nombre d'analyses par classe est trop faible.

Tableau 3: Code des méta-UTS selon leur teneur en calcaire et en argile

		Teneur en Argile en %			
		[0 ; 7.5[[7.5 ; 17.5[[17.5 ; 30[>= 30
Teneur en CaCO ₃ en %	[0 ; 1[aa	ab	ac	ad
	[1 ; 5[ba	bb	bc	bd
	[5 ; 12.5[ca	cb	cc	cd
	[12.5 ; 25[da	db	dc	dd
	[25 ; 37.5[ea	eb	ec	ed
	>=37.5	fa	fb	fc	fd

ii. Classement des analyses

Les analyses de la BDAT sont classées de la même manière que les UTS, selon leurs teneurs en argile et en CaCO₃ (tableau 1).

iii. Attribution des analyses aux méta-UTS

Chaque analyse est attribuée à une méta-UTS selon sa teneur en argile et en calcaire. On obtient finalement une distribution des analyses de la BDAT pour chaque couple [méta-UTS, UCS] ou le cas échéant pour chaque couple [méta-UTS, meta-UCS].

f. Calcul des médianes

i. Calcul de la médiane pour chaque couple [méta-UTS, UCS]

Pour chaque couple on calcule la médiane de la teneur en Corg. Cette opération est réalisée avec un tableur, grâce au fichier d'InfoSol qui liste les valeurs de Corg des analyses attribuées à chaque UCS.

On vérifie pour chaque couple que le nombre d'analyses utilisées pour faire ces calculs est supérieur à 10. Si ce n'est pas le cas, pour un couple donné, on utilise la valeur du couple [méta-UTS, meta-UCS].

ii. Calcul de la médiane pour chaque couple [méta-UTS, meta-UCS]

Cette étape est identique à la précédente mais les médianes sont calculées à partir du fichier d'InfoSol qui liste les valeurs de Corg des analyses attribuées à chaque méta-UCS.

On vérifie pour chaque croisement que le nombre d'analyses utilisées pour faire ces calculs est supérieur à 10. Si ce n'est pas le cas, pour un couple donné, alors on utilise la valeur médiane de la méta-UTS sur l'ensemble du territoire.

iii. Calcul de la médiane pour chaque méta-UTS à l'échelle du territoire

Cette étape est identique à la précédente mais les médianes sont calculées à partir du fichier d'InfoSol qui liste les valeurs de Corg des analyses du territoire sans les attribuer à une sous-unité spatiale.

iv. Validation agronomique

Pour s'assurer de la validité des résultats, il est recommandé de vérifier que les médianes attribuées aux couples [méta-uts, méta-ucs] se classent de manière cohérente avec les caractéristiques des classes des meta-UTS : on s'attend théoriquement à des valeurs croissantes avec la teneur en argile et en calcaire (effet protecteur de l'argile et du calcaire vis-à-vis de la minéralisation des matières organiques).

3- Validité et limites de la méthode

La validation au sens strict de la méthode n'est pas possible faute de jeux de données indépendantes qui permettraient de comparer les teneurs en carbone organique affectées aux UTS à des teneurs observées sur la même période.

A défaut d'une véritable validation, les performances de la méthode ont été évaluées sur des critères de cohérence agronomique des résultats et de généralité. Pour ce faire, la méthode a été appliquée sur 3 territoires-tests choisis pour couvrir une diversité de paysages pédologiques et de systèmes de production agricoles : la petite région naturelle du Tardenois dans l'Aisne, le Pays Rhin-Vignoble-Grand Ballon (RVGB) dans le Haut-Rhin et la Communauté de Communes du Thouarsais dans les Deux-Sèvres. Leur surface est comprise entre 30000 et 45000 hectares. Le tableau 4 en donne les principales caractéristiques pédologiques et agricoles.

Tableau 4 : Caractéristiques des trois territoires-tests

Territoire	Caractéristiques pédologiques	Nombre d'UCS	Nombre d'UTS	Assolement
Tardenois (02)	Plateaux ondulés à dominante de sols limoneux profonds ; sols argileux calcaires caillouteux et sols sablo-limoneux à limono-sableux.	28	77	Betterave, blé, colza, protéagineux
Pays RVGB (68)	Plaines du Rhin et de l'Ill et Piedmont vosgien; sols alluviaux très variés, calcaires ou non, sableux ou limoneux, plus ou moins caillouteux, parfois hydromorphes, limons calcaires loessiques	19	58	Maïs grain prédominant, Blé
CdC Thouarsais (79)	Diversité des sols sur 2 ensembles de matériaux parentaux (socle ancien et formations sédimentaires): limono-sableux, argilo-calcaires caillouteux, argileux et sableux, en partie hydromorphes,	33	26	Blé, colza, protéagineux, tournesol, maïs grain

La méthode et les résultats de cette évaluation sont présentés dans l'article de Scheurer *et al.* (2020) paru dans la revue Etude et Gestion des Sols (<https://www.afes.fr/publications/revue-etude-et-gestion-des-sols/volume-27/>). Les principaux éléments de la conclusion de cet article sont repris ci-dessous.

La méthode d'affectation des analyses de la BDAT aux types de sol sur un territoire local s'appuie sur 2 critères de classement : un double critère analytique (teneur en argile et calcaire) et un critère géographique d'appartenance à une UCS.

Malgré le choix d'un critère analytique simplifié, le nombre d'analyses directement exploitables est parfois insuffisant. La mise au point de fonctions de pédotransfert pour

estimer le taux d'argile en fonction de la CEC et de la teneur en carbone organique permet de s'affranchir de cette limite à condition que ces fonctions soient paramétrées localement. En revanche, le cas des sols très calcaires et comportant une forte proportion de calcaire dans la fraction argileuse pose un problème particulier : suivant le mode de détermination analytique de la teneur en argile (avec ou sans décarbonatation) le taux d'argile d'un même échantillon peut prendre des valeurs très différentes, alors que ce mode de détermination n'est pas renseigné dans la BDAT. Celui-ci étant le plus souvent fixé *a priori* par chaque laboratoire, il faudrait trier les analyses en fonction de leur laboratoire d'origine et du mode de détermination couramment appliqué afin de pouvoir traiter le cas des territoires concernés.

Il subsiste enfin une incertitude sur la localisation réelle de certains groupes d'analyses réalisées à l'initiative d'organismes agricoles (coopératives), référencées sur la commune du commanditaire. L'application de la méthode nécessiterait de pouvoir les repérer puis de les éliminer, au moins dans le cas des territoires très hétérogènes.

L'affectation des analyses au niveau de l'UCS présente plusieurs avantages par rapport aux unités-support couramment utilisées pour exploiter la BDAT que sont le canton ou la petite région agricole. Elle permet : (i) de caractériser assez précisément les types de sol à associer aux analyses de terre, ce qui facilite l'interprétation des résultats, (ii) de prendre en compte –dans une certaine mesure- l'effet des variables autres que A et CaCO_3 sur la teneur en C org estimée, en fonction des caractéristiques dominantes des sols dans l'UCS (hydromorphie, charge en éléments grossiers de la strate de surface), (iii) de prendre en compte –partiellement et indirectement- l'effet de l'histoire culturelle de chaque UCS. L'UCS est utilisée ici comme une unité spatiale au sein de laquelle la dynamique des systèmes de culture est relativement homogène, et a pu générer une évolution spécifique de la distribution des teneurs en carbone dans les sols, notamment à travers le niveau des restitutions humiques des successions culturelles ou les pratiques d'amendements organiques passés.

L'affectation des teneurs en Corg par classes déterminées sur la base des critères analytique et géographique, rend compte d'une part faible de la variabilité observée (13 à 34 % dans les 3 cas étudiés), mais non négligeable au regard de la multiplicité des déterminants de celle-ci. Le rapport de corrélation le plus faible peut s'expliquer par plusieurs facteurs liés respectivement à la méthode, à la source de données, et aux spécificités du territoire : imprécision de la fonction de pédotransfert utilisée pour estimer le taux d'argile, incertitude sur la localisation de certaines analyses de ce territoire (voir *supra*), forte variabilité des histoires culturelles entre parcelles pour un même type de sol au niveau intra-UCS.

Dans les cas minoritaires où le nombre d'analyses est insuffisant pour caractériser une meta-UTS au niveau de l'UCS, le regroupement d'UCS en meta-UCS est opéré selon 2 catégories de critères : propriétés des sols et occupation du sol. Celui-ci repose en partie sur l'hypothèse que des descripteurs de l'occupation du sol actuelle, basés sur les données du RPG et traduisant la diversité des systèmes de culture présents dans chaque UCS, sont des indicateurs pertinents pour agréger ou différencier ces UCS en fonction de leurs histoires culturelles dominantes. On peut s'interroger sur sa validité, notamment dans les territoires marqués par le retournement des prairies permanentes. Cette hypothèse se vérifie néanmoins dans deux territoires tests (Tardenois et Pays RVGB). Par ailleurs, le risque de regrouper par cette méthode des UCS ayant des histoires culturelles très différentes est

minimisé par le fait que l'agrégation repose en même temps sur un critère pédologique qui traduit les contraintes du milieu, comme déterminant fort de l'histoire culturelle d'un territoire.

Un effectif minimal de 10 analyses a été fixé pour pouvoir caractériser une meta-UTS par sa médiane. Le choix de ce seuil vise à minimiser le nombre d'UTS non-renseignées sur le territoire. Il peut paraître trop faible pour limiter le biais d'échantillonnage. Une étude de sensibilité des résultats à des seuils plus élevés serait utile pour compléter l'évaluation de la méthode.

La cohérence agronomique des résultats paraît globalement satisfaisante. Les variations de teneurs observées entre classes (meta-UTS dans les UCS et les meta-UCS) s'expliquent assez bien en fonction des facteurs attendus. Elles mettent en évidence (i) l'effet des propriétés des sols (teneurs plus élevées dans les sols argileux et/ou calcaires, ou hydromorphes ou dans les sols caillouteux), (ii) l'effet du niveau des restitutions humiques des cultures (teneurs plus élevées dans les UCS où les céréales à paille ou le maïs grain sont les plus fréquentes). Ces résultats semblent ainsi conforter la pertinence de la méthode et son potentiel de généralisation dans les territoires d'application de la démarche ABC'Terre.

Bibliographie

Arrouays D., Hardy R., Schnebelen N., Le Bas C., Eimberck M., Roque J., Grolleau E., Pelletier A., Doux J., Lehmann S., Saby N., King D., Jamagne M., Rat D. et Stengel P., 2004. Le programme Inventaire Gestion et Conservation des Sols de France. Etude et Gestion des Sols 11 (3), 187-197.

Baize D., 2018. Guide des analyses en pédologie. Ed. Quae

Laroche B., Richer de Forges A.C., Leménager S., Arrouays D., Schnebelen N., Eimberck M., Toutain B., Lehmann S., Tientcheu E., Héliès F., Chenu J.P., Parot S., Desbourdes S., Girot G., Voltz M. et Bardy M., 2014. Le programme Inventaire Gestion Conservation des Sols de France : volet Référentiel Régional Pédologique. Etude et Gestion des Sols 21, 125-140.

Scheurer O., Bousselin X. et Saby N. - 2020 - Une méthode pour caractériser les teneurs en carbone organique des types de sol d'un Référentiel Régional Pédologique sur un territoire agricole à partir de la Base de Données des Analyses de Terre, Etude et Gestion des Sols, 27, 189-207. <https://www.afes.fr/publications/revue-etude-et-gestion-des-sols/volume-27/>

Annexe 1 : Données « Sol » à extraire du RRP pour la mise en œuvre de la méthode ABC'Terre.

Rappel des opérations mobilisant les données SOL dans la méthode ABC'Terre :

- Pour l'affectation des rotations aux types de sols via RPG Explorer :
 - o Croisement spatial UCS x ilots de cultures du RPG
 - o Attribution de notes de valeur agronomique aux combinaisons « UTS x rotation de cultures » en fonction des propriétés (contraintes) agronomiques des UTS

- Pour l'affectation d'une teneur initiale en Corg aux UTS, à partir de la BDAT, éventuellement différenciée selon les UCS:
 - o Croisement spatial UCS x communes puis distribution des analyses communales de la BDAT dans les UCS.
 - o Croisement spatial UCS x ilots de culture du RPG pour classification des UCS selon leur type d'occupation du sol (assolement moyen ou autre indicateur)
 - o Classification des UCS selon leur composition en UTS
 - o Regroupement des UTS en classes de teneur en argile et Calcaire dans la strate de surface.
 - o Affectation d'une teneur en Corg par classe d'UTS

- Pour reconstituer les pratiques culturales à l'échelle du territoire :
 - o Estimation du rendement des cultures en fonction des classes de RUM

- Pour Simeos-AMG :
 - o Le C/N et le pH sont deux variables prises en compte dans la fonction de minéralisation k d'AMG (Clivot et al, 2017 et 2019)

- Pour le calcul des émissions GES :
 - o Les facteurs d'émissions par volatilisation des apports d'engrais minéraux sont fonction du pH du sol (EMEP EEA, 2016, Tier 2)

Données géographiques nécessaires :

- Shapefile des UCS contenues dans le territoire d'étude, y compris le prolongement des polygones hors des communes du territoire.

Données sémantiques nécessaires :

- Description des UCS :
 - o Nom de l'UCS
 - o UTS présentes et leur pourcentage surfacique

- Description des UTS :
 - o Nom de l'UTS
 - o Occupation dominante et secondaire
 - o Classe de drainage naturel
 - o Profondeur modale
 - o RUM estimée si possible

- Caractéristiques de la strate de surface :
 - Texture dominante
 - Taux d'argile modal
 - Taux de calcaire total modal
 - Abondance des éléments grossiers modale
 - CEC et teneur en C organique modales
 - pH eau et rapport C/N

Données sémantiques supplémentaires si RUM non estimée :

Caractéristiques des strates de chaque UTS, soit pour chaque strate :

- N° de strate
- Texture Aisne (ou à défaut Taux d'Argile, Limon, Sable modaux)
- Epaisseur moyenne
- Abondance des éléments grossiers
- Nom des éléments grossiers

Tableau récapitulatif des variables et de leur usage

	Nom de variable dans DONESOL 3	Utilisations				
		Table des valeurs Agronomiques	Classification des UCS	Classement des UTS	Estimation RUM	Détermination rendements, fonction minéralisation (AMG) et volatilisation Nmin
UCS						
Nom de l'UCS	NOM_UCS		x			
UTS						
UTS et pourcentage	POURCENT		x			
Occupation dominante et secondaire	OCCUP1 OCCUP2 (ds L_UCS_UTS)	x	x			
Nom de l'UTS	NOM_UTS		x			
Classe de drainage naturel	DRAI_NAT	x	x			
Profondeur modale	PROF_SOL_MOD		x		x	
RUM	A estimer	x	x			x
STRATES						
Epaisseur moyenne	EPAIS_MOY				xxx	
Texture dominante	TEXTURE (AISNE)		x		xxx	
Taux d'argile	TAUX ARGILE (VAL_MOD)	x	x	x		
Taux de calcaire total	CALC_TOT (VAL_MOD)	x	x	x		
Abondance des éléments grossiers	ABONDANCE_EG (VAL_MOD)	x	x		xxx	
Nom des éléments grossiers	NOM_EG				xxx	
CEC	CEC			x		
Taux de carbone organique	CARBONE			x		
pH eau	PH_EAU					x
rapport C/N	C_N					x

Pour les variables des strates : **x = pour la strate de surface** **xxx = pour chaque strate**

Annexe 2 : Exemple de construction de méta-UCS dans le Tardenois

Classification des UCS en méta-UCS

La teneur en carbone organique des sols d'une parcelle est fonction de son contexte pédoclimatique et de son histoire culturale. La classification des UCS du territoire d'étude vise la constitution de groupes d'UCS relativement semblables du point de vue de ces 2 facteurs. Ce regroupement sera utilisé dans les cas où le nombre d'analyses exploitables est insuffisant au niveau d'une UCS pour estimer la teneur en Corg d'un type de sol (meta-UTS).

La méthode est présentée au travers de l'exemple du Tardenois. Elle consiste à réaliser 2 classifications des UCS, basées respectivement sur un indicateur de leur occupation du sol (triplets de cultures) et sur leur composition en types de sol. La classification finale sera issue du croisement des 2 précédentes.

I- Groupes d'UCS basés sur les triplets de cultures

Il s'agit de regrouper les UCS qui ont des histoires similaires au travers d'un indicateur (réducteur mais facile à obtenir UCS par UCS) qui est la part surfacique des triplets de cultures reconstitués à l'aide de RPG Explorer sur la période 2011-2013. Un triplet de cultures est défini par une succession de cultures observée sur 3 ans sur une portion de territoire. L'assolement de triplets d'une UCS est défini par la part surfacique de l'ensemble des triplets observés sur l'UCS pour une période donnée.

1) Obtention de l'assolement de triplets et classification des triplets

On obtient un assolement des triplets UCS par UCS grâce aux données RPG et au logiciel de RPG Explorer. Pour ce travail une culture unique est attribuée à chaque groupe de culture ASP (afin de ne pas induire de biais).

Les triplets sont classés suivant leur composition dans différents types adaptés à la région. Ici ils sont classés suivant l'arbre présenté ci-dessous :

Sur les 3 ans :

S'il y a deux ans betteraves ou plus. -> Alors le triplet est classé « mon bett »

Sinon, s'il y a un an de betterave -> Alors le triplet est classé « bett »

Sinon, s'il y a deux ans prairies ou plus. -> Alors le triplet est classé « pp »

Sinon, s'il y a un an de protéagineux ou plus. -> Alors le triplet est classé « prot »

Sinon, s'il y a un an de maïs ou plus. -> Alors le triplet est classé « maïs »

Sinon, s'il y a un an de colza ou plus. -> Alors le triplet est classé « col »

Sinon, s'il y a 3 ans de céréales à paille. -> Alors le triplet est classé « mon_cer »

Sinon -> Alors le triplet est classé « autre »

On calcule ensuite la part surfacique (%) de chaque type.

2) Classification des UCS

Afin de pouvoir classer les UCS en fonction de leur composition en triplets une Classification Ascendante Hiérarchique (CAH) basé sur le critère de Ward et renforcée par la méthode des kmeans est réalisée. Pour cela le logiciel R (R Core Team, 2017) et le package cluster sont utilisés (Maechler et al., 2017). Pour cette classification les données (composition des UCS en types de triplets) ne sont pas réduites, car toutes les variables ont la même unité (% de l'UCS).

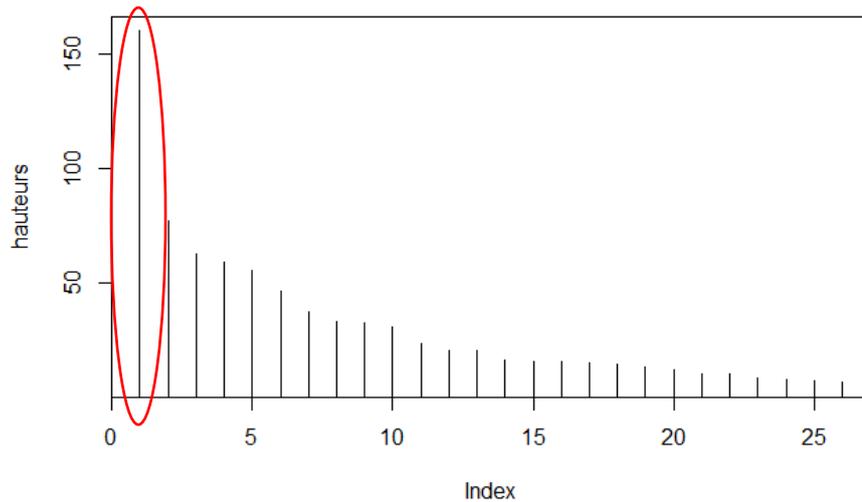


Figure 1 : Diagramme en barres de la distance entre deux regroupements en fonction du nombre de classes créées

Etant donné le fait que nous cherchons à faire un nombre de classes assez faible et que le « saut » entre le découpage en un et deux classes est très important (entouré en rouge sur la figure 1), on découpe le territoire en deux classes d'UCS.

->Description des classes

Afin de caractériser ces classes on utilise la fonction catdes du package FactoMineR (Le et al., 2008). D'autre part, les classes sont projetées sur le plan factoriel d'une ACP créé à partir mêmes données que l'ACP. Les ACP sont réalisées avec le logiciel R (R Core Team, 2017) et au package FactoMineR (Le et al., 2008).

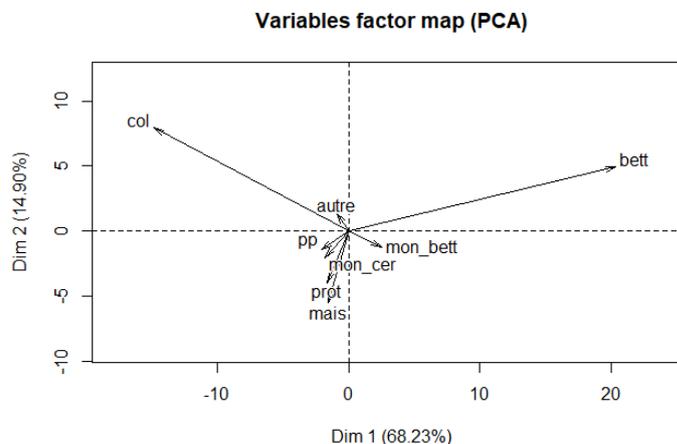


Figure 2 : corrélation des variables entre elles et avec les axes du plan factoriel

On observe que la présence de triplets « betteraviers » est négativement corrélée à la présence de triplet contenant des prairies, du colza ou d'autres « têtes de rotations ». Le premier axe de l'ACP permet de séparer les ucs dominées par des triplets betteraviers (valeurs positives) et les triplets avec du colza (valeurs négatives) (figure 2).

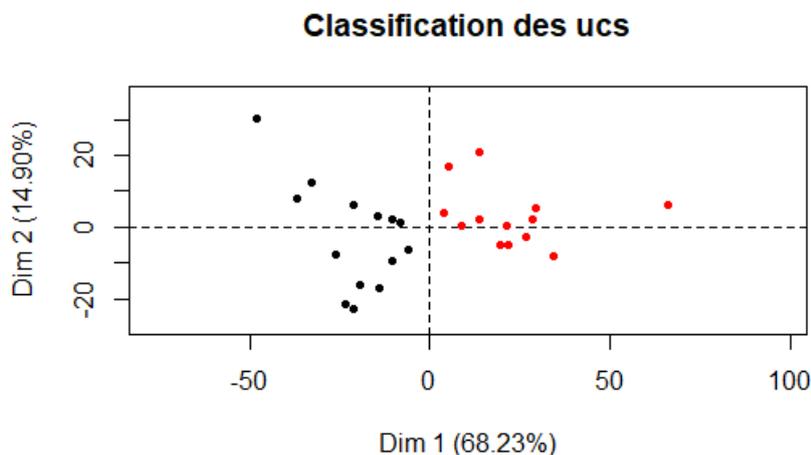


Figure 3 : Projection des groupes sur le plan factoriel

Un point est une ucs. En noire la classe 1 et en rouge la classe 2

Tableau 1 : Caractérisation des classes en fonction de leur composition en triplets

classe	bett		col		prot		mon_cer	
	moy	et	moy	et	moy	et	moy	et
1	8.9	7.3	43.1	14.4	8.3	5.9	10.6	6.1
2	45.8	12.4	21.7	11.6	4.1	3.7	4.9	3.3
tous	26.7	11.6	32.8	17.0	6.3	5.4	7.8	5.7

La CAH réalisée regroupe les UCS en 2 classes. Ces classes se différencient suivant la prévalence des triplets betteraviers qui représentent 9% de la classe 1 contre 46% de la classe 2. La classe 1 est donc plus riche en triplets avec du colza, des protéagineux ou de la monoculture de céréales (Voir tableau 1 et figure 3)

II- Groupes d'UCS basés leur composition en types de sol

L'objectif est ici de regrouper les unités cartographiques de sol (UCS) en fonction de leur composition en types de sols (UTS). Ces UTS sont caractérisées par des variables qui ont une influence sur la dynamique du carbone organique dans le sol : teneur en calcaire, en argile, et en cailloux, note de drainage naturel (excès d'eau) et réservoir utilisable maximal (RUM).

1) Classification des UTS

Avant de classer les UCS en fonction de leur composition en types de sols, il est nécessaire de regrouper les UTS en grands types. Pour cela une Classification Ascendante Hiérarchique (CAH) basée sur le critère de Ward et renforcée par la méthode des kmeans est réalisée. Pour faire cette classification les variables d'entrée sont réduites pour donner la même importance à chacune des variables.

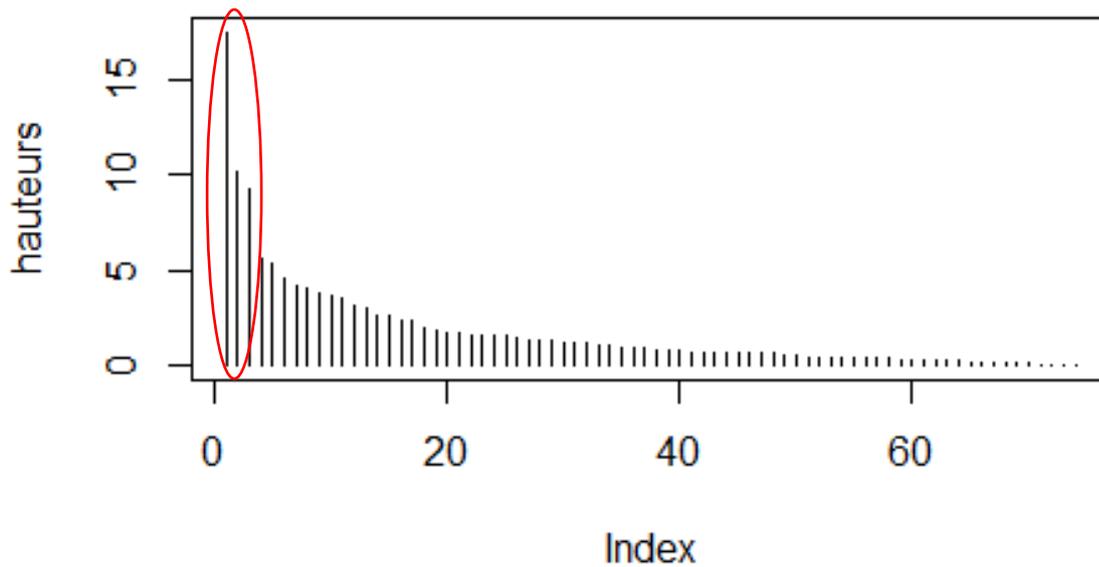


Figure 4 : Diagramme en bars de la distance entre deux regroupements en fonction du nombre de classes

Au regard de la figure 4 il semble cohérent de faire 4 classes.

->Description des classes

Afin de caractériser ces classes on utilise la fonction `catdes` du package `FactoMineR` (Le *et al.*, 2008) et les classes sont projetés sur le plan factoriel d'une ACP créé à partir des 5 variables quantitatives qui ont servi à la réalisation des classes (teneur en argile, RUm, teneur en calcaire, note d'hydromophie et teneur en cailloux). Les données sont centrées et réduites afin de donner la même importance à chacun de ces facteurs dans la construction du plan factoriel. Ces ACP sont réalisées avec le logiciel R (R Core Team, 2017) et au package `FactoMineR` (Le *et al.*, 2008).

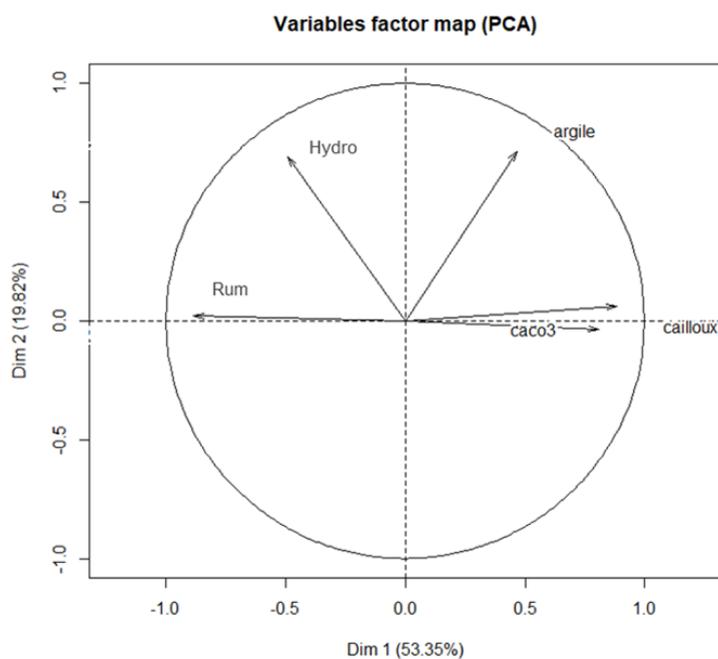


Figure 5 : Corrélation des variables entre elles et avec les axes du plan factoriel

On observe que l'axe 1 de l'ACP permet de projeter plus de la moitié de l'information des cinq variables évoquées précédemment. Cet axe peut être interprété comme étant l'axe du RUm, des cailloux et du calcaire. Plus les coordonnées des UTS sont importantes sur cet axe plus les UCS ont une forte charge en cailloux et sont calcaires et plus ils ont un faible RUm (et inversement) (voir figure 5).

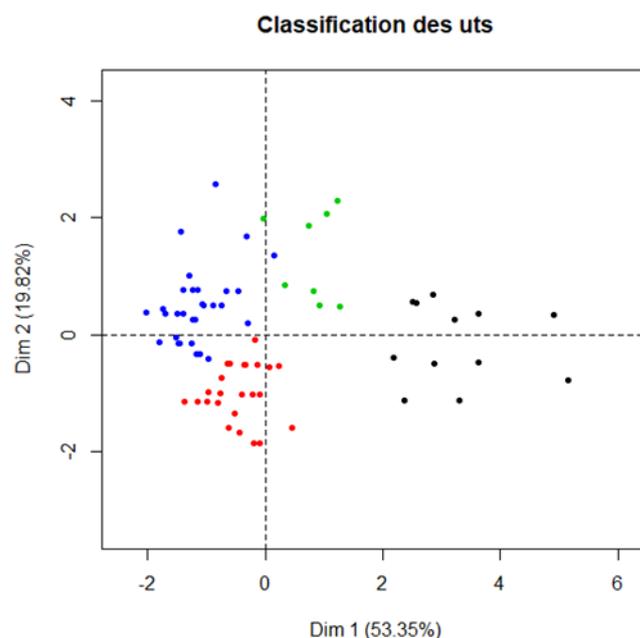


Figure 6 : Projection des classes sur le plan factoriel

Un point est une UTS. En noire la classe 1 en rouge la classe 2 en vert la classe 3 et en bleu la classe 4

Tableau 2 : Description des classes d'UTS

classes	CaCO ₃ (g/kg)		Argile (%)		Cailloux (%)		Hydromorphie		RUm (mm)	
	moy	et	moy	et	moy	et	moy	et	moy	et
1	157	89	26.1	7.9	17.6	5.6	0	0	49	18
2	9	26	16.3	5.7	0.4	2.0	0	0	NS	NS
3	NS	NS	36.3	4.7	NS	NS	NS	NS	NS	NS
4	6	33	17.9	5.5	0.8	3.0	5.3	1.5	161	24
Tous	34	72	20.6	8.7	3.9	7.2	2.4	2.7	134	48

La figure 6 et le tableau 2 décrivent les classes de sol générées par la CAH. Ainsi, la classe 1 regroupe les UTS ayant une forte charge en cailloux, une teneur en CaCO₃ élevée, des sols en moyenne assez argileux, qui ont une faible tendance à l'hydromorphie et un faible RUm. La classe de sol 2 est composée d'UTS qui sont en général pauvres en calcaire et en cailloux, peu hydromorphes. La classe de sol 3 est uniquement caractérisée par une forte teneur en argiles qui est de plus assez peu variable. La classe de sol 4 regroupe les UTS qui sont globalement hydromorphes, avec un fort RUm, avec peu de cailloux, de calcaire et d'argile.

2) Classification des UCS

Grâce à cette classification chaque UCS peut être caractérisée par sa composition en ces quatre groupes de sols. C'est sur la base de ces données que les UCS sont classées grâce à une CAH sans réduire les données (qui sont toutes des proportions). Le logiciel et les extensions utilisées sont les même que dans la partie précédente.

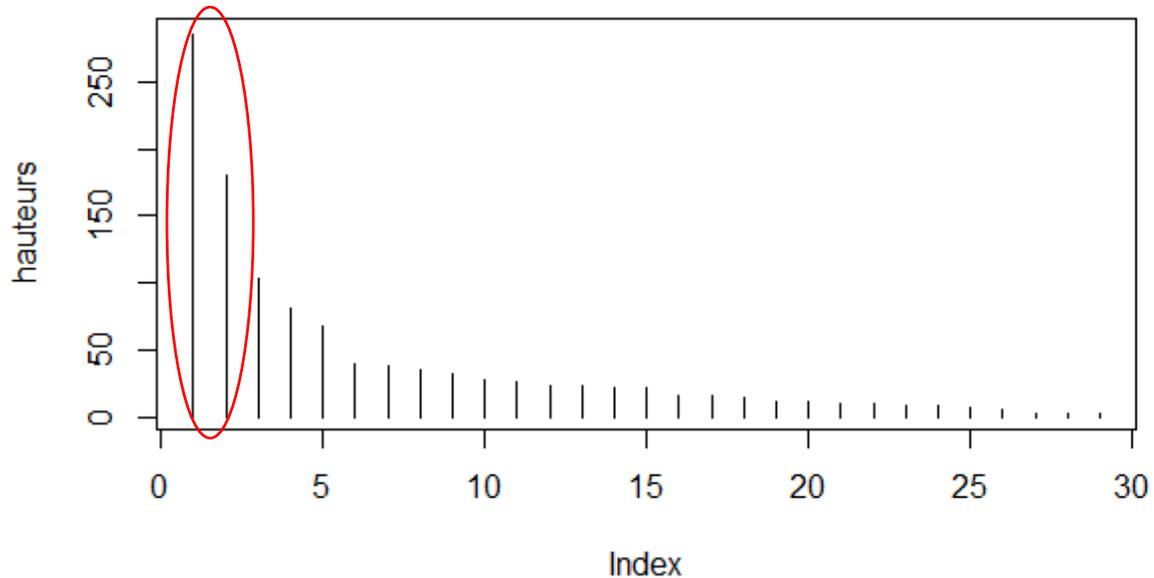


Figure 7 : Diagramme en barres de la distance entre deux classes utilisé pour la CAH

Au regard de la figure 7 il semble cohérent de faire 3 groupes d'UCS.

->Description des classes

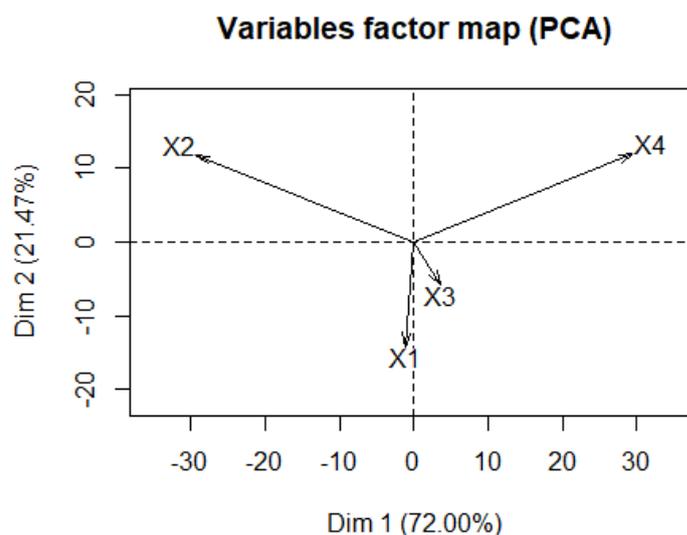


Figure 8 : corrélation des variables entre elles et avec les axes du plan factoriel

X1 : type de sol 1, X2 : type de sol2, X3 : type de sol 3 et X4 : type de sol 4. Voir tableau 2.

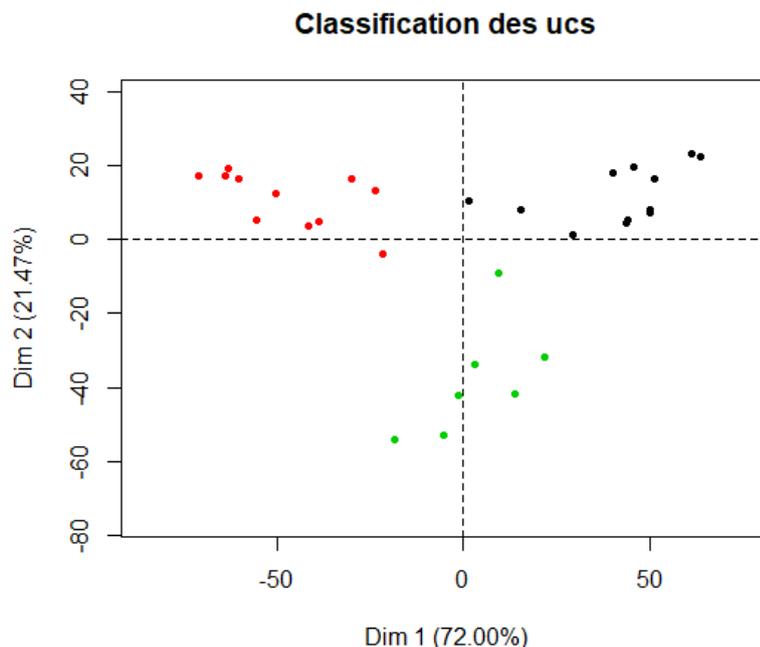


Figure 9 : Projection des groupes sur le plan factoriel

Un point est une UCS. En noir la classe 1, en rouge la classe 2 et en vert la classe 3

Tableau 3 : Description des classes d'UCS

classes	classe de sol 1		classe de sol 2		classe de sol 3		classe de sol 4	
	moy	et	moy	et	moy	et	moy	et
1	1	3	8	10	NS	NS	74	16
2	NS	NS	70	15	0	0	11	10
3	33	16	10	6	18	20	NS	NS
Tous	10	16	31	32	6	13	39	32

On observe que les deux axes de l'ACP projettent plus de 93% de l'information relative à la répartition des différents types de sols. Les sols des UCS de la classe 1 sont largement dominés par des sols de type 4 qui sont des sols avec un RUm globalement élevé et souvent hydromorphes. Les UCS de la classe 2 sont globalement dominées par des sols de type 2 qui sont des sols sableux ou limoneux. La classe d'UCS 3 est plus hétérogène ; elle présente plus de sols calcaires et caillouteux, à faible Rum (classe de sol 1) ou très argileux (classe de sol 3) que les autres UCS. Le reste des sols de cette classe d'UCS ne représentent en moyenne que la moitié de la surface de l'UCS, et se composent de sols à tendance hydromorphe, à fort RUm sur environ 40 % (sol 4) de la surface et par environ 10% de sols limoneux ou sableux (sol 2). (Voir figure 8 et 9 et le tableau 3).

III- Croisement des groupes d'UCS pour obtenir les méta-UCS

Les deux classifications sont croisées pour obtenir les méta-UCS.

Références bibliographiques :

Cornillon, P. A., Husson, F., Jégou, N., Matzner-Lober, E., Josse, J., Guyader, A., Kloareg, M., Rouvière, L. (2010). Statistiques avec R. 3^e édition, col. Pratique de la statistique, Presses universitaires de Rennes. 296 p.

Husson, F., Lê, S., Pagès, J. (2016). Analyse de données avec R. 2nd édition, col. Pratique de la statistique, Presses universitaires de Rennes. 239 p.

Lê, S., Josse, J., Husson, F. (2008). FactoMineR : An R Package for Multivariate Analysis. Journal of Statistical Software, 25(1), 1-18.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K.(2018). Cluster: Cluster Analysis Basics and Extensions. R package version 2.0.7-1.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienne, Autriche. URL <https://www.R-project.org/>.

Annexe 3 : Evaluation de la teneur en argile à partir de la CEC et de la teneur en carbone organique

Pour mettre en œuvre la méthode d'affectation des teneurs en carbone organique aux types de sol, telle que proposée dans la démarche ABC'Terre, il est nécessaire de disposer dans la BDAT d'un nombre maximum d'analyses exploitables par UCS. Ces analyses doivent renseigner à la fois la teneur en Corg et les teneurs en argile et en CaCO₃.

En général, le nombre d'analyses renseignant Corg et la CEC est plus élevé que celui des analyses renseignant Corg et la teneur en argile. Afin de rendre ces analyses utilisables, leur taux d'argile a été estimé à partir de la CEC à l'aide d'une fonction de pédotransfert. La méthode mise en œuvre présentée dans cette annexe.

La fonction de pédotransfert : principes généraux

Formule et méthode

La CEC d'un sol est connue pour être fonction de la teneur en carbone organique et en argile (vraie) de ce sol (Baize, 2018). Cette fonction peut prendre la forme additive :

$$\text{CEC} = a * \text{argile} + b * \text{Corg} + c.$$

L'établissement de la fonction de pédotransfert consiste donc en premier lieu à estimer les coefficients a, b et c à partir d'un jeu de données d'apprentissage. Celui-ci est constitué par un ensemble d'analyses du territoire pour lesquelles on connaît les 3 variables : CEC, argile et Corg. La méthode utilisée est la régression linéaire multiple.

La fonction de pédotransfert obtenue pour estimer le taux d'argile est donc de la forme :

$$\text{Argile} = 1/a * (\text{CEC} - b * \text{Corg} - c)$$

Domaine de validité

Par cette méthode, le taux d'analyses mal classées selon leur teneur argile semble ne pas dépasser 25 à 30 %. A priori, la qualité de la prédiction devrait être d'autant meilleure que le contexte pédologique est homogène. A l'opposé, elle devrait être moins bonne dans les territoires présentant une forte variabilité de la minéralogie des argiles, en général associée à l'hétérogénéité des matériaux parentaux.

Avant d'utiliser la fonction de pédotransfert, il faut par ailleurs vérifier que la distribution des résidus du modèle de régression linéaire n'est pas influencée par la teneur en Corg, faute de quoi les résultats finaux seraient fortement biaisés.

Dans les sols très calcaires et comportant une forte proportion de calcaire dans la fraction argileuse le taux d'argile d'un même échantillon peut prendre des valeurs très différentes: suivant le mode de détermination analytique de la teneur en argile (avec ou sans décarbonatation préalable), alors que ce mode de détermination n'est pas renseigné dans la BDAT.

Celui-ci étant le plus souvent fixé *a priori* par chaque laboratoire, il faudrait identifier le(s) laboratoire(s) qui alimente(nt) principalement la BDAT sur le territoire et le mode de détermination qu'il(s) applique(nt) couramment en routine.

Plusieurs situations peuvent alors se présenter :

- si la méthode « avec décarbonatation » est nettement dominante, les teneurs en argile fournies correspondent à de l'argile « vraie » non calcaire ; une seule fonction de pédotransfert peut être construite sans distinguer les terres calcaires des autres.
- si la méthode « sans décarbonatation » est nettement dominante, la signification d'une teneur en argile peut être très différente selon que la terre est calcaire ou non. Il est dans ce cas nécessaire de construire 2 fonctions de pédotransfert respectivement spécifiques des terres calcaires et non calcaires.
- s'il n'y a pas de méthode dominante, il est préférable de n'utiliser une fonction de pédotransfert que pour les terres non ou peu calcaires (teneur en CaCO_3 inférieure à 250 g/kg).

Comme tout modèle, cette fonction ne peut être utilisée que sur les mêmes plages de valeurs qui ont permis d'estimer ses paramètres.

Validation et taille minimum de l'échantillon d'apprentissage

Afin d'évaluer empiriquement la taille de l'échantillon d'apprentissage nécessaire pour paramétrer la fonction de pédotransfert, des tests ont été menés sur les différents territoires d'étude du projet ABC Terre 2A.

Au regard des résultats présentés dans la figure 1 et du tableau 1, il semble qu'un échantillon d'apprentissage de plus de 50 analyses soit suffisant pour paramétrer une fonction sur un territoire. S'il n'est pas possible d'accéder à 50 analyses, alors 35 analyses semblent encore donner un résultat cohérent bien que légèrement dégradé. Descendre en dessous de 35 analyses induit des problèmes de qualité trop importants.

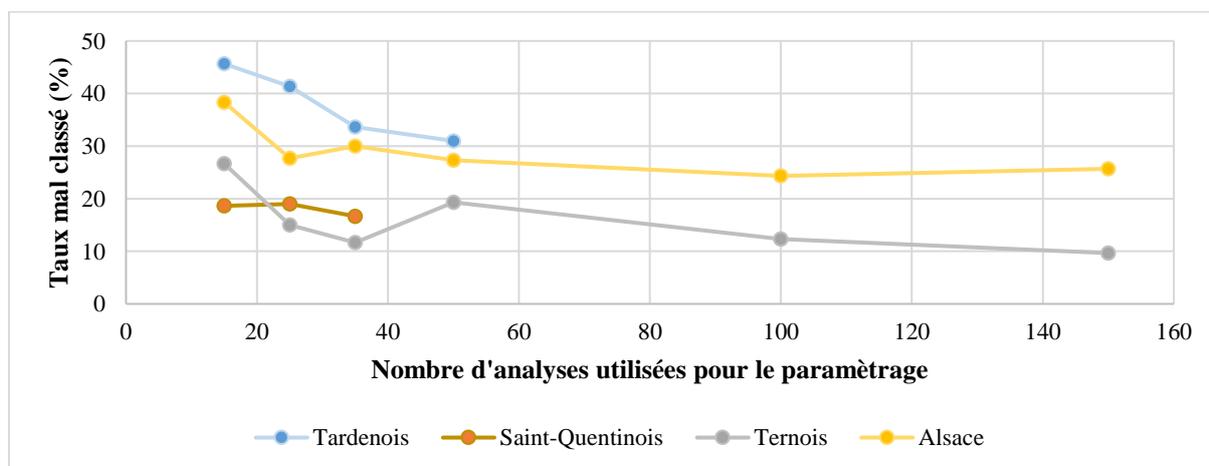


Figure 1 : Pourcentage d'analyses dont la classe de teneur en argile est mal prédite par la fonction de pédotransfert en fonction du nombre d'analyses utilisées dans l'échantillon d'apprentissage

Chaque point est la moyenne de trois tirages aléatoires du nombre d'analyses présenté en abscisse parmi les analyses d'un territoire. Ces analyses ont servi à paramétrer le modèle linéaire : $Y = \theta_1 + \theta_2 X_1 + \theta_3 X_2 + \varepsilon$ où Y: CEC, X_1 : teneur en argile, X_2 : teneur en C organique, θ_1 , θ_2 et θ_3 : les paramètres du modèle et ε : le résidu.

Le taux de « mal classé » est déterminé -pour chaque tirage aléatoire- pour des analyses n'ayant pas servi à déterminer les paramètres du modèle (au minimum 20) : on calcule le pourcentage de cas où ces analyses sont attribuées à la mauvaise classe. Toutes les analyses ayant une teneur en CaCO₃ supérieure à 250 g/kg ont été exclues de l'étude.

Tableau 1 : Taux de mal classé en fonction de la taille de l'échantillon d'apprentissage

	Taille de l'échantillon d'apprentissage					
	15	25	35	50	100	150
Tardenois						
n_val	89	79	69	54	NA	NA
%MC	46	41	34	31	NA	NA
Saint-Quentinois						
n_val	40	30	20	NA	NA	NA
%MC	19	19	17	NA	NA	NA
Ternois						
n_val	196	186	176	161	111	61
%MC	27	15	12	19	12	10
Alsace						
n_val	283	273	263	248	198	148
%MC	38	28	30	27	24	26

n_val : taille de l'échantillon de validation ; **% MC** : pourcentages des analyses de l'échantillon de validation attribuées à la mauvaise classe par la fonction de pédotransfert ;

Chaque point est la moyenne de trois tirages aléatoires du nombre d'analyses (présenté en en tête des colonnes) parmi les analyses « complètes » d'un territoire. Ces analyses ont servi à paramétrer le modèle linéaire : $Y = \theta_1 + \theta_2 X_1 + \theta_3 X_2 + \varepsilon$ où Y : CEC, X1 : teneur en argile, X2 : teneur en C organique, θ_1 , θ_2 et θ_3 : les paramètres du modèle et ε : résidu. Pour les analyses de l'échantillon de validation (analyses n'ayant pas servi à déterminer les paramètres du modèle), on calcule le pourcentage de cas où les analyses sont attribuées à la mauvaise classe. Toutes les analyses ayant une teneur en CaCO₃ supérieure à 250 g/kg ont été exclues.