



**HAL**  
open science

## **Rapport technique sur la constitution d'un système d'information géographique à l'échelle de l'Europe, dans le cadre du projet SLAM-B : Volet données sol**

Hugo Senges, Helene Raynal, Hugues Clivot, Julie Constantin

### ► **To cite this version:**

Hugo Senges, Helene Raynal, Hugues Clivot, Julie Constantin. Rapport technique sur la constitution d'un système d'information géographique à l'échelle de l'Europe, dans le cadre du projet SLAM-B : Volet données sol. Inrae. 2025, 46 p. <hal-04920044>

**HAL Id: hal-04920044**

**<https://hal.inrae.fr/hal-04920044v1>**

Submitted on 25 Feb 2026

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ETALAB - Open licence



PROGRAMME  
DE RECHERCHE  
CARBONE ET  
ÉCOSYSTÈMES  
CONTINENTAUX



RÉPUBLIQUE  
FRANÇAISE  
Liberté  
Égalité  
Fraternité

INRAE



# Rapport technique sur la constitution d'un système d'information géographique à l'échelle de l'Europe, dans le cadre du projet SLAM-B : Volet données sol

*Cadrage et framework de simulation STICS - version V1- 29 janvier 2025*

Hugo Senges <sup>(1)</sup>, Hélène Raynal <sup>(1)</sup>, Julie Constantin <sup>(1)</sup>, Hugues Clivot <sup>(2)</sup>

<sup>(1)</sup> Université de Toulouse, INRAE-ENSAT, UMR AGIR, F-31320 Castanet-Tolosan, France

<sup>(2)</sup> Université de Reims Champagne-Ardenne, INRAE, FARE, UMR A 614, Reims, France

Avec les contributions de: C. LeBas, N. Saby, Julien Wengler de l'UMR Info&Sols de INRAE

## Résumé

Ce travail s'inscrit dans le projet SLAM-B qui vise à concevoir et évaluer les trajectoires de bioéconomie des territoires vers la neutralité carbone. Ce rapport synthétise les travaux menés pour constituer un système d'information géographique permettant la création des jeux de données nécessaires pour la simulation du modèle de culture STICS à l'échelle européenne avec une résolution fine. Le volet « données sol » se concentre sur les données sol qui seront intégrées dans le système. Plusieurs bases de données sont mobilisées et comparées, dont la BDAT, RMQS, BDGSF, ESDB, LUCAS et SoilGrid. Chacune offre des informations variées sur les caractéristiques physico-chimiques et structurales des sols. La France sert de cas d'étude avant une extension européenne. Le territoire étudié couvre les zones agricoles, en excluant forêts et cultures pérennes. Le rapport conclut sur une proposition de construction des jeux de données pour la simulation du modèle STICS à partir des données issues de SoilGrid et de LUCAS.

### Pour citer ce rapport :

Senges H., Raynal H., Constantin J., Clivot H., 2024. Rapport technique sur la constitution d'un système d'information géographique à l'échelle de l'Europe, dans le cadre du projet ANR-22-PEXF-0003 – projet SLAM-B : Volet Données Sol

*"Ce travail s'inscrit dans le projet du programme de recherche exploratoire FairCarboN et a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre de France 2030 portant la référence ANR-22-PEXF-0003 – projet SLAM-B »*

*"This work is part of project SLAM-B of the exploratory research program FairCarboN and received government funding managed by the Agence Nationale de la Recherche under the France 2030 program, reference ANR-22-PEXF-0003 – projet SLAM-B]"*

# Table des matières

Rapport technique sur la constitution d'un système d'information géographique à l'échelle de l'Europe, dans le cadre du projet SLAM-B : Volet données sol.....	1
Résumé .....	1
Table des matières .....	2
Tables des illustrations .....	3
Table des annexes.....	3
1. Contexte .....	4
2. Ressources.....	5
2.6. SoilGrids :.....	13
3. Des SMUs aux USMs .....	15
4. Représentation du profil de sol dans les différentes BD .....	18
5 Construction des jeux de données d'entrée pour le modèle STICS .....	19
5.1. Les caractéristiques physico-chimiques des horizons de sol.....	20
5.2. Fonctions de pédotransfert (FPT).....	20
5.2.1. La question de la rétention en eau du sol :.....	21
5.2.2. De la qualification des FPTs.....	22
5.2.3. Remarques méthodologiques.....	28
6. Proposition retenue pour la fondation d'une base de données pédologiques orientée simulation .....	29
6.1. Solution SoilGrid.....	29
6.2. Solutions hybride LUCAS/SoilGrid .....	30
Bibliographie .....	31
Glossaire : .....	33
Chaîne algorithmique : .....	33
Annexes .....	34
1Annexe 1 : Attributs de la table SGDBE .....	34
3 Annexe 3 : Attributs qualifiés et clés d'identification primaires de la base SAPDE/M .....	38
4Annexe 4 : Attributs de l'European Soil Database Derived Datas .....	39
5Annexe 5 : Attributs des données Lucas ponctuelles .....	39
6Annexe 6 : Attributs téléchargeables de l'ISRIC - SoilGrids.....	40
7Annexe 7 : Formules raster utilisées pour le travail de qualification des PFTs.....	41
9Annexe 9 : Table détaillée des variables pédologiques de STICS et des données correspondantes parmi celles investiguées .....	43

## Tables des illustrations

Figure 1: Carte de la CEC, Gis Sol .....	1
Figure 2 : Carte des SMUs du SGDBE, croisés avec la maille Safran .....	1
Figure 3 : Carte de la profondeur de sol disponible pour les racines de l'ESDD.....	2
Figure 4 : Carte de la densité des points d'échantillonnage LUCAS .....	3
Figure 5 : Carte de l'erreur standard de la donnée Topsoil Organic Carbon Lucas.....	3
Figure 6 : Capture d'écran d'une unité pédologique de la BDAT renseignée avec une épaisseur de sol de 30cm ou moins, et pourtant indiquée comme dépassant régulièrement les 100dg/kg de carbone organique au-delà du mètre de profondeur, selon SoilGrids .....	4
Figure 7:Diagramme en boîte des écarts de surfaces entre cellules d'une même grille carrée SAFRAN, étendue à l'échelle européenne .....	5
Figure 8a (ci-dessus) : Carte des USM (Safran X SMU) en fonction de la surface agricole qu'elles contiennent Figure 8b (ci-dessous) : séquence logique de création des USM.....	13
Figure 9 : Courbe de rétention d'eau selon la pression, van Genuchten, 1980.....	15
Figure 10 : Diagrammes en boîtes des séries de valeurs en HCCF et HPF selon les formules et les données mobilisées.....	18
Figure 11 : Carte des écarts de valeurs entre le HMINF calculé grâce aux données Sol Hydro et les données PTRDB.....	19
Tableau 1: Sources de données sur les différentes profondeurs des sols.....	1
Tableau 2: Table des profondeurs disponibles dans les données recensées.....	1
Tableau 3 : Paramètres "sol" du modèles STICS .....	2
Tableau 4 : Comparaison des FPTs investiguées pour la valeur "au champ" .....	3
Tableau 5 : Comparaison des FPTs investiguées pour la valeur "au point de flétrissement" ....	3
Tableau 6 : Résumé des médianes des séries de HCCF .....	4
Tableau 7 : Résumé des médianes des séries de HPF .....	5
Tableau 8 : Proposition de construction du système d'information pédologique STICS sur la base SoilGrids .....	13
Tableau 9 : Proposition de construction du système d'information pédologique STICS sur la base Lucas x SoilGrids.....	15
Tableau 10 : Calculs des épaisseurs de sols par couches dans le modèle STICS, selon les données ESDB.....	18

## Table des annexes

<a href="#">Annexe 1 : Attributs de la table SGDBE</a> .....	<b>Erreur ! Signet non défini.</b> 3
<a href="#">Annexe 2 : Attributs de la table PTRDB</a> .....	34
Annexe 3 : Attributs qualifiés et clés d'identification primaires de la base SAPDE/M.....	36
<a href="#">Annexe 5 : Attributs des données Lucas ponctuelles</a> .....	38
<a href="#">Annexe 6 : Attributs téléchargeables de l'ISRIC - SoilGrids</a> .....	39
<a href="#">Annexe 7 : Formules raster utilisées pour le travail de qualification des FPTs</a> .....	40
<a href="#">Annexe 8 : Tableau de qualification des FPTs en fonctions des paramètres, Dobarco, 2019</a> .	41
<a href="#">Annexe 9 : Table détaillée des variables pédologiques de STICS et des données correspondantes parmi celles investiguées</a> .....	42

# 1. Contexte

Les objectifs de neutralité carbone à 2050 et de durabilité nécessitent de développer une bioéconomie répondant simultanément aux enjeux de production, de transformation alimentaires et non alimentaires et de préservation des ressources naturelles. Pour cela, des approches territoriales combinant agroécologie, valorisation fine des ressources, action collective et développement d'interdépendances vertueuses (ex. : économie circulaire, renforcement du lien-urbain rural) sont à favoriser.

Les approches de modélisation et évaluation intégrées (*Integrated Assessment and Modelling - IAM*) sont reconnues pour leur capacité à accompagner les acteurs dans la conception d'organisations territoriales des activités humaines telles que celles de la bioéconomie. Elles permettent d'intégrer dans des modèles les connaissances pluridisciplinaires génériques et celles empiriques des acteurs et d'utiliser ces modèles pour simuler et évaluer, ex ante, des scénarios d'organisation des territoires sous hypothèses de changements endogènes ou exogènes. Le projet SLAM-B (*ANR-22-PEXF-0003 – projet SLAM-B*) vise à structurer la communauté de recherche française qui développe et applique des approches IAM pour soutenir le développement d'une bioéconomie durable basée sur l'agroécologie. SLAM-B prendra en compte les différentes composantes de la bioéconomie notamment, la diversité des systèmes de production, des matières premières et produits finis, des bioraffineries, des boucles de recyclage, et des organisations de filières. SLAM-B vise à relever trois grands défis scientifiques : (i) développer des approches IAM génériques pour simuler les transitions bioéconomiques et anticiper leurs conséquences, (ii) démontrer la pertinence et la légitimité de ces approches IAM via leur application au sein de Living Labs à portée prospective, appelés Scenario Labs et, (iii) produire des connaissances sur mesure pour les décideurs publics à l'échelle nationale et européenne.

Pour relever ces défis, SLAM-B est organisé en trois groupes de tâches (GT) interconnectés. Le travail présenté s'inscrit dans le GT3 vise à développer et appliquer des approches IAM, à l'échelle de la France et de l'Europe, pour générer des connaissances sur la quantité de résidus de culture et de couverts intermédiaires mobilisables par différentes filières de la bioéconomie, le potentiel de l'intensification écologique des forêts et de l'agriculture et le potentiel de production de biogaz.

Il vise essentiellement le support à la décision publique, tant via le développement d'indicateurs sur mesure pour l'implantation et l'investissement dans des filières bioéconomiques sur les territoires (T.3.1), que de solutions applicables à grande échelle pour mobiliser et utiliser durablement de la biomasse agricole et forestière additionnelle et séquestrer du carbone dans les sols (T3.2-3.7). Le GT3 bâtira sur des données dynamiques et spatialement explicites.

Suite aux premières avancées de ce GT3, les tâches T3.2 (Couverture des sols nus), T3.3 (Interactions résidus de culture et bioéconomie) et T3.7 (Niveau potentiel et conséquences de la production de biogaz), ont été réunies dans une tâche unique de modélisation des systèmes de cultures à l'échelle de l'UE, sur la base du modèle STICS (entre autres), et du cadre conceptuel développé dans le cadre de l'étude 4p1000<sup>1</sup>.

---

<sup>1</sup> Pellerin *et al.*, 2020.

L'objet du présent rapport est la constitution d'une base de données préalable à ce travail de modélisation. Plus précisément, il présente les données satisfaisant aux entrées du modèle de culture STICS sur son compartiment sol. Celles-ci devront être à une résolution fine et à l'échelle de l'Europe

Ci-dessous quelques points de précision sur l'emprise étudiée :

- La France sert de cas d'études dans le présent rapport afin de tester la disponibilité et la qualité des données. La démarche sera ensuite étendue à l'ensemble de l'échelle l'Europe.
- Après arbitrage il est décidé de conserver le Royaume-Uni dans l'emprise étudiée, car les données qui le couvrent pourraient permettre une meilleure calibration du modèle (densité de profils de sols réalisés supérieure à la moyenne).
- Le territoire considéré est celui couvert par des activités agricoles, au sens large, bien qu'excluant *a priori* les forêts et les cultures pérennes, dont les logiques ne sont pas modélisables grâce au modèle STICS.
- Les prairies permanentes, initialement non-retenues au sein de l'espace « cultivé » visé par l'étude, sont considérées à part du reste des emprises cultivées (hors-forêts). Leur dynamique spatiale (extension ou réduction) pourra faire l'objet de scénarios qui à ce titre viendrait impacter les sorties du modèle ou les résultats des scénarios.

## 2. Ressources

Ci-dessous sont présenté les différents jeux de données étudiés dans le cadre de ce travail :

### 2.1. La Base de Données d'Analyses des Terre – BDAT :

Lien : <https://www.gissol.fr/le-gis/programmes/base-de-donnees-danalyses-des-terres-bdat-62>

Description : Quelques 2 millions d'échantillons d'horizons de surface de sols cultivés, prélevés en France entre 1990 et 2014, renseignant 31 paramètres permettant d'évaluer les propriétés physico-chimiques des sols (pH, carbone, capacité d'échange cationique, taux de saturation), leur fertilité (azote, carbone organique, phosphore, potassium, magnésium, sodium), leurs teneurs en métaux et métalloïdes (bore, cuivre, fer, manganèse, zinc) et enfin, leur texture (argile, limon, sable). Les données statistiques agrégées par canton, sont disponibles pour 5 périodes : 1990-1994, 1995-1999, 2000-2004, 2005-2009, 2010-2014.

Commentaire : Base de référence en France, la BDAT est cependant peu aisée à prendre en main d'un point de vue spatial. L'agrégation des statistiques par canton et les formats de livraisons de données (.csv sans géométrie) oblige à reprojeter les informations dans l'espace. De plus les informations sur les différents horizons (au niveau agrégé du canton) ne sont pas renseignées. Elle peut servir à contrôler les autres les données, et à estimer un intervalle de confiance.

## 2.2. Le Réseau de mesures de la qualité des sols – RMQS :

Lien : <https://traitementinfosol.pages.mia.inra.fr/statistiquesrmqs/>

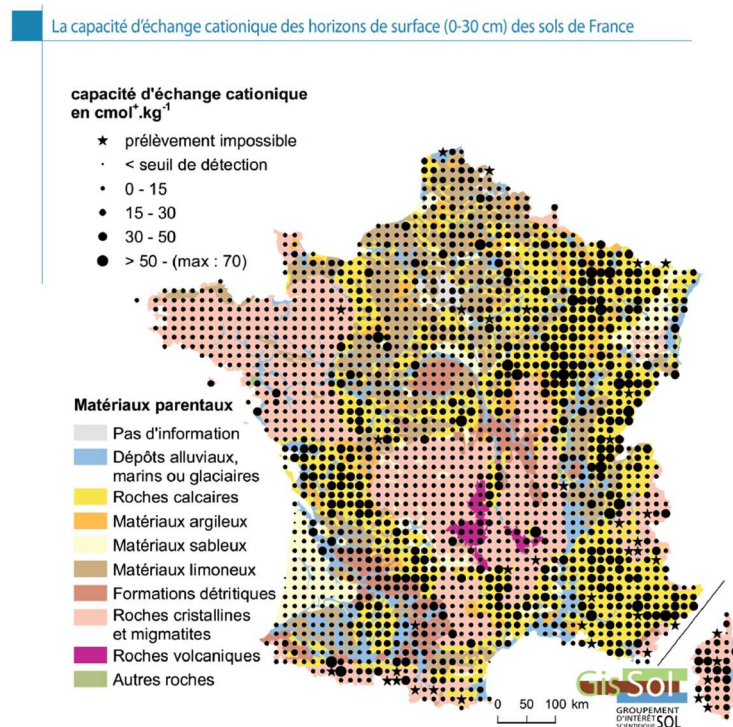
**Description :** Un réseau de 2240 sites échantillonnés et analysés tous les 15 ans, selon une grille systématique de 16 km de côté. La deuxième campagne a démarré en 2016 et devrait s'étendre sur 12 ans. Cela représente environ 180 sites prélevés et analysés chaque année. Les analyses ont été réalisées sur des échantillons composites prélevés à la tarière selon 2 couches de prélèvements (0-30 cm) ou couche travaillée en sol cultivé, appelée composite de surface ou « composite 1 » et la couche sous-jacente jusqu'à 50 cm, appelée composite de sub-surface ou « composite 2 ». Chaque échantillon composite a été constitué à partir du mélange de 25 prélèvements individuels prélevés sur une surface d'échantillonnage de 400 m<sup>2</sup> selon un plan d'échantillonnage aléatoire stratifié

**Paramètres :** Granulométrie / Carbone et azote total / CEC / Cations échangeables / P2O5 / pH / Bore / Eléments majeurs totaux (Fe, K, Mg, Na, Mg), Calcaire total (si pH > 6,5) / Aluminium échangeable (si pH < 6,5)

Ces paramètres sont donnés pour un « Nom de Couche » qui renvoie à la position de la couche étudiée dans la stratigraphie.

**Commentaire :** La base de statistiques agrégées est librement accessible. Les données ponctuelles géolocalisées sont disponibles à l'adresse suivante :

<https://entrepot.recherche.data.gouv.fr/file.xhtml?persistentId=doi:10.15454/AIQ9WS/FXPD DG&version=1.1> . Les paramètres évalués sont très riches et permettent un renseignement complet du fichier sol de STICS. En revanche les études cartographiques produites par Info & Sol qui interpolent ces résultats ne sont disponibles qu'au format PNG (exemple : Figure 1, la carte de la CEC en France par exemple)



Source : Gis Sol, RMQS, 2011 ; Inra, BDGSF, 1998.

Figure 1: Carte de la CEC, Gis Sol

### 2.3. La Base de Données Géographique des Sols de France – BDGSF :

Lien : <https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.15454/BPN57S>

Description : La Base de Données Géographique des sols de France à 1/1 000 000 (BDGSF) correspond à la partie française de la Base de Données Géographique des sols d'Europe<sup>2</sup>. Elle est le fruit d'un projet de collaboration impliquant l'ensemble de l'Union Européenne et des pays voisins. La BDGSF est une représentation simplifiée de la diversité et de la variabilité spatiale des sols en France. La méthodologie utilisée pour différencier et nommer les principaux types de sol est basée sur la terminologie de la légende de F.A.O. pour la Carte des sols du monde à l'échelle du 1/5 000 000. Ces données correspondent à la traduction à dires d'expert, dans la nomenclature du World Reference Base of Soil Resources (version 2008)<sup>3</sup>, du nom du sol dominant de chaque Unité Cartographique de Sol de la Base de Données Géographique des Sols de France à 1/1 000 000.

Commentaire : La BDGSF est une couche SIG composée de polygones représentant des unités cartographiques de sol (SMU). Nous disposons d'une table qui recense pour chaque SMU la liste des unités typologique *STU* et le pourcentage associé dans la maille *POUR*. C'est un découpage du territoire Français en grande zones pédologiques homogènes, ce qui permet de constituer les unités spatiales du plan de simulation (cf Partie « Des SMU<sup>4</sup>s aux USM<sup>5</sup> »). Tous les paramètres STICS ne peuvent être renseignés à cette échelle, du moins pas depuis les informations contenues dans la seule BDGSF (par exemple le taux de carbone organique n'y figure pas). Toutefois des études plus récentes reprennent la base de ce découpage et sont disponibles pour certains paramètres importants. La carte de la profondeur des sols en France en est le meilleur exemple : <https://doi.org/10.15454/7ZDND6>

### 2.4. European Soil DataBase – ESDB V2 :

Lien : <https://esdac.jrc.ec.europa.eu/content/european-soil-database-v20-vector-and-attribute-data>

Description : La Base de Données des Sols d'Europe (ESDB - 2004) à l'échelle du 1 / 1 000 000 et un des produits fournis par l'ESDAC (European Soil Data Center) basé au centre européen Commun de Recherche (JRC). C'est une représentation simplifiée de la diversité spatiale de la couverture de sols. C'est un modèle, sous forme numérique, de la compréhension et de la connaissance du paysage pédologique. Les données sont fournies au niveau d'unités cartographiques des sols appelées « Soil Mapping Unit » SMU qui sont des objets géographiques dont la forme géométrique et la localisation sont claires et bien définies. Certaines SMU peuvent être pures, auquel cas elles sont composées d'une seule unité typologique de sol (STU) couvrant 100 % de la SMU. Dans la majorité des cas, une SMU est composée de plusieurs STU. Une SMU représente donc une association de types de sol. (Pour en savoir plus, consulter [https://esdac.jrc.ec.europa.eu/ESDB\\_Archive/ESDBv2/esdb/sgdbe/metadata/purity\\_maps/purity.htm](https://esdac.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/esdb/sgdbe/metadata/purity_maps/purity.htm)). Le profil de sol est adressé selon le découpage TopSoil (0-50cm) / Subsoil (>50cm)<sup>6</sup>. Pour chacune de ces couches de sol, on a les variables pédologiques associées.

---

<sup>2</sup> Panagos Panos. The European soil database (2006) GEO.

<sup>3</sup> <https://openknowledge.fao.org/server/api/core/bitstreams/1bd0747c-e9d8-4b28-99bf-55684d121e38/content>

<sup>4</sup> SMU: Soil Model Unit

<sup>5</sup> USM: Unit for Simulaiton

<sup>6</sup> « There is no information about soil horizons in the geographical database version 2.0. The soil is arbitrarily divided into two layers : topsoil and subsoil, roughly corresponding to A and B horizons. The depth of differentiation is 50 cm when it is reached ». In King D., Jones R.J.A, Tomasson A.J., *et al.*, 2003, p.135.

Il existe deux versions : sous forme vectorielle (les SMUs y sont les structures élémentaires) et sous forme de raster à la résolution spatiale : 1 pixel = 1km<sup>2</sup>.

L'ESDB embarque quatre jeux de données associés, dont les attributs sont donnés en **Annexes 1 et 2**. Les variables renseignées sont exhaustives mais catégorisées et non quantifiées

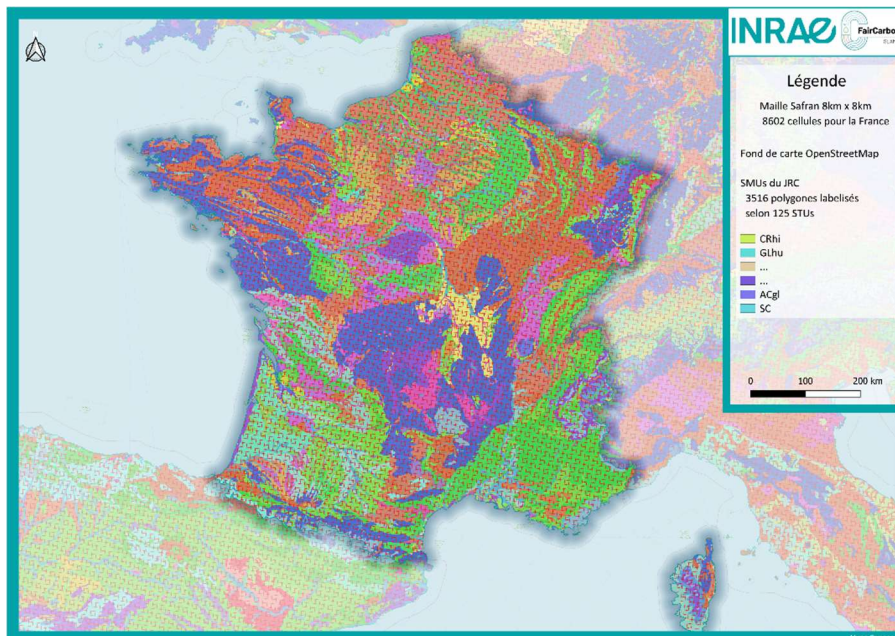
#### 2.4.1. SGDBE : Soil Geographic Database of Eurasia au :1:1 000 000

Description : On trouve dans ce jeu de données les même SMUs que dans l'ESDB, leur composition en STU avec le pourcentage de surface que chaque STU représente au sein de la SMU et les variables décrivant leurs natures et leurs propriétés (informations pédologiques simples) : texture du sol, charge en cailloux, présence de couches imperméables, profondeur du sol, matériau parental, etc.

Commentaire : Ce jeu de données fournit un découpage territorial à l'échelle communautaire, qui permet le cas échéant de baser les Unités de Simulations sur des espaces homogènes du point de vue pédologique (cf. partie « SMU » ci-après). Les informations de profondeur sont également d'intérêt car sans alternative connue à ce jour. Les attributs *Distances to Textural Change* et *Distance to Rock* permettraient ainsi de paramétrer les horizons des unités de simulations dans STICS (cf. partie « Questions de profondeurs », ci-après).

En revanche les attributs sur la texture sont trop imprécis dans leur encodage<sup>7</sup> et dans la résolution spatiale qui les accompagne. Le fonctionnement de la donnée qui dispose parfois sur un même polygone une valeur pour le « sol dominant » et une autre pour le « sol secondaire » rend encore plus difficile l'appréhension de valeurs quantitatives. La figure 2 représente les SMU issues de ce jeu de données

Figure 2 : Carte des SMUs du SGDBE, croisés avec la maille Safran



<sup>7</sup> Exemple d'encodage de l'attribut *TEXT-SRF-DOM*, source :

[https://esdac.jrc.ec.europa.eu/ESDB\\_Archive/ESDBv2/popup/sg\\_attr.htm](https://esdac.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/popup/sg_attr.htm)

0 No information, 9 No mineral texture (Peat soils), 1 Coarse (18% < clay and > 65% sand), 2 Medium (18% < clay < 35% and >= 15% sand, or 18% < clay and 15% < sand < 65%), 3 Medium fine (< 35% clay and < 15% sand), 4 Fine (35% < clay < 60%), 5 Very fine (clay > 60 %)

#### 2.4.2. Pedotransfer Rules Database - PTRDB :

Lien : <https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes>

Description : Les attributs de cette base sont des valeurs dérivées de fonctions de pédotransfert, qui sont des outils d'estimation des propriétés hydrauliques des sols (rétention en eau principalement). Elle se base sur les données de texture de la SGDBE.

Commentaire : L'utilité de ce jeu de données réside dans la comparaison que l'on peut en faire avec d'autres données jugées de « meilleure » qualité *a priori* (spatialement, temporellement et sémantiquement). En effet en reprenant le découpage des SMUs précédent, la distribution spatiale des valeurs reste ici sommaire. De plus on retrouve le même problème d'encodage ambigu des valeurs quantitatives. Il faut donc au préalable convertir les classes de valeurs (de réserve utile « *AWC* » par exemple) en nombres réels. C'est le plus souvent la médiane de la classe qui est retenue, mais la question des classes extrêmes ( « < min » ; « > max ») se pose. De plus les attributs donnés en termes de « réserve utile » ne correspondent pas aux variables hydriques demandées par STICS. On peut inférer les *HMINF* et *HCCF* demandée par STICS depuis les valeurs d'*AWC* données dans PTRDB, mais cela rajoute une étape de calcul, et réduit la reproductibilité du résultat en fonction des règles de pédo-transfert sélectionnées.

#### 2.4.3. Soil Profile Analytical Database SPADE :

Lien : <https://esdac.jrc.ec.europa.eu/content/spadem>

Description : (Soil Profile Analytical Database, that contains measured and estimated soil profiles for Europe). La base de données, initiée en 2004, est une structure commune pour stocker des informations standardisées sur les propriétés des profils pédologiques des principaux sols européens. Elle compte à ce jour 560 profils pour 2656 horizons (cf. la clé de détermination primaire de la base en Annexe). Un processus de validation des données a été mis en place et les profils sont ainsi qualifiés en fonction de leur représentativité. Les attributs investigués permettraient de renseigner intégralement un fichier de paramétrage du sol pour une simulation STICS sur l'aire des profils considérés.

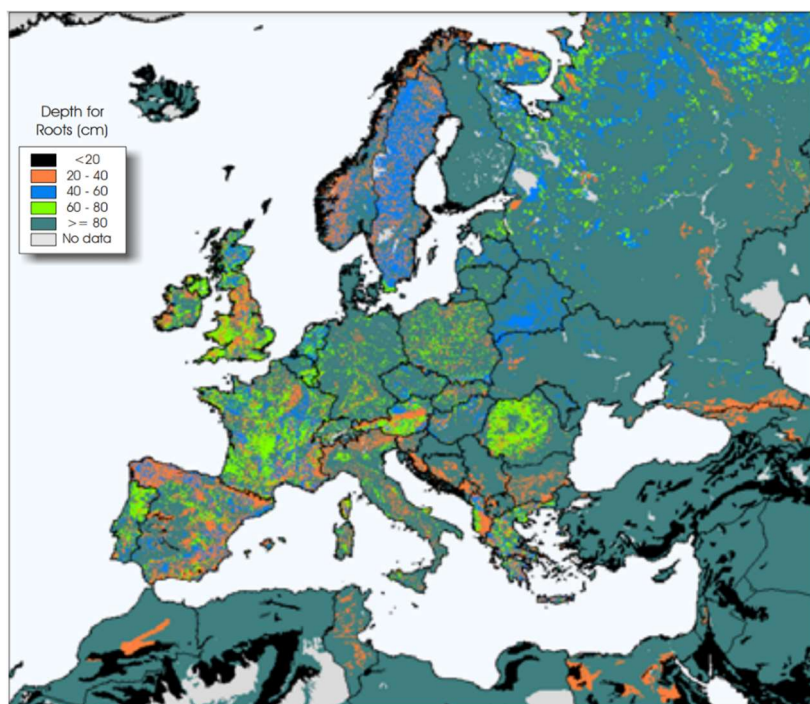
Commentaire : Bien que les informations récoltées soient de qualité, l'inégale répartition spatiale des profils ainsi que l'incertitude quant au positionnement géographique de ces profils rend difficile la mobilisation de ces données ponctuelles. Elles peuvent cependant servir de donnée de contrôle et l'on peut imaginer comparer les sorties d'une simulation STICS basée sur les données plus généralistes et d'une autre simulation basée sur les profils SPADE. On peut également comparer la donnée brute des profondeurs de sol par rapport aux autres sources disponibles

#### 2.4.4. European Soil Database Derived data - ESDDD:

Lien : <https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data>

Description : Enfin, on trouve également, associé à l'ESDB, un jeu de données issus d'un travail de 2013. Un certain nombre de couches pour les propriétés du sol ont été créées à partir des données de la base de données européenne sur les sols en combinaison avec les données de la base de données harmonisée mondiale sur les sols (HWSD) et de la base de données sol-terrain (SOTER). Des règles de pédo-transfert et des interpolations spatiales (et verticale) ont été pratiquées selon les attributs

Figure 3 : Carte de la profondeur de sol disponible pour les racines de l'ESDD



Commentaire : La résolution à 1km de cette donnée est dérivée du SGBDE dont la résolution est moins bonne, et les relevés de terrains plus anciens. Cette augmentation de la résolution est faite par interpolation spatiale. Pour alimenter ce « sur-échantillonnage »<sup>8</sup>, plusieurs sources de données sont compilées (ce qui pose le problème de la colinéarité des variables dérivées). Les données de textures et de rétention en eau peuvent être comparées avec les données plus récentes et nativement mieux résolues (LUCAS). D'autre part la variable « *Depth available to roots* » (voir ci-contre) pourra constituer l'un des indicateurs de

profondeur du sol qui fait référence sur l'aire d'étude.

#### 2.5. Land Use and Coverage Area frame Survey – LUCAS :

Lien : <https://esdac.jrc.ec.europa.eu/projects/lucas>

Description : Enquête périodique par sondage sur l'utilisation des terres/la couverture terrestre pour échantillonner et analyser les principales propriétés pédologiques dans 23 États membres de l'Union européenne. Base de données spatiales cohérente sur la couverture des sols dans l'UE, basée sur des procédures d'échantillonnage et d'analyse standards, elle offre 4 jeux de données principaux.

<sup>8</sup> Opération de définition d'une valeur des sous-pixels, lorsque l'on fait accroître la résolution d'une image. Cela crée systématiquement des artefacts dans le produit sur-échantillonné. Voir : Getreuer P., 2011

### 2.5.1. LUCAS points :

Lien : <https://ec.europa.eu/eurostat/web/lucas/database/2009>

Description : Environ 20.000 points ont été sélectionnés dans la grille LUCAS principale pour le prélèvement d'échantillons de sol. Initié en 2009 la démarche a connu des revisites en 2012, 2015, 2018 et 2022. Les attributs renseignés sont donnés dans un tableau en Annexe n°5.

Commentaire : La principale limite des données LUCAS (brutes ou dérivées) c'est la profondeur investiguée qui ne dépasse pas les 20 premiers centimètres. En revanche la densité spatiale des points échantillonnés (cf figure 4 ci-dessous) et la richesse des attributs renseignés en font la ressource la plus pertinente actuellement en Europe pour les horizons superficiels.

Bien que l'on préférera travailler directement sur les données interpolées présentées ci-après, on peut discriminer les points en fonction d'une couverture des sols tripartite : Forêt / Prairies / Cultures. Cela permettrait de qualifier l'écart de confiance des données sélectionnés dans le framework de simulation retenu.

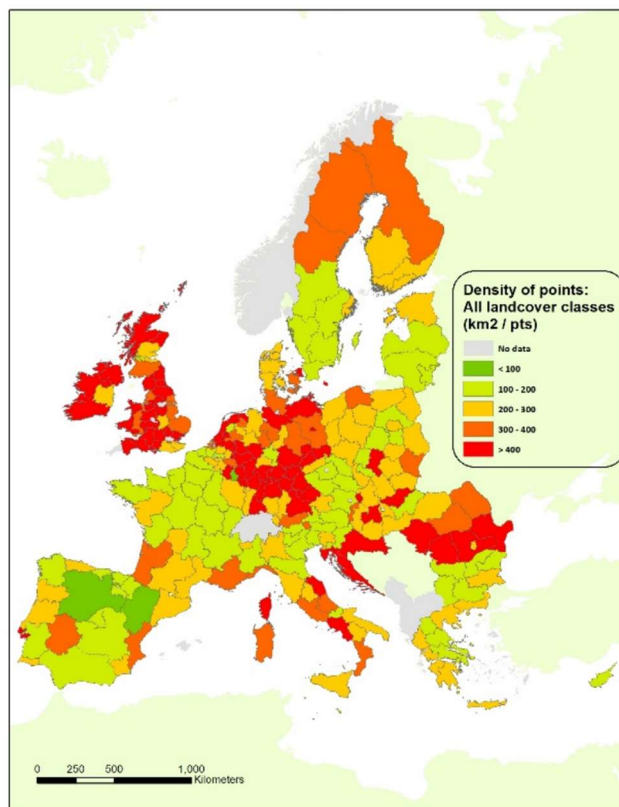


Figure 4 : Carte de la densité des points d'échantillonnage LUCAS

### 2.5.2. LUCAS Soil Chemical properties at European scale:

Lien : <https://esdac.jrc.ec.europa.eu/content/chemical-properties-european-scale-based-lucas-topsoil-data>

Description : Ce jeu de données contient 8 propriétés chimiques : pH, pH (CaCl), capacité d'échange cationique (CEC), carbonates de calcium (CaCO<sub>3</sub><sup>9</sup>), rapport C:N, azote (N), phosphore (P) et potassium (K). Il a été produit en 2019 à l'aide des données ponctuelles sur le sol de l'enquête LUCAS 2009/2012 (environ 22000 points) pour l'UE-26 (non compris Chypre et la Croatie). Les cartes des propriétés chimiques pour l'Union européenne ont été produites à l'aide de modèles de régression par processus gaussien (GPR). Résolution : 500m.

Commentaire : Outre la limitation au-delà des 20 premiers centimètres, ces données font office de référence et affichent une qualification parmi les meilleures du domaine (R<sup>2</sup>=0.65 en moyenne). La résolution spatiale est en outre inégalée à ce stade. A noter que la valeur en carbonate de calcium qui permet de retrouver la bonne valeur d'argile à rentrer dans STICS

<sup>9</sup> Le % de CaCO<sub>3</sub> en-dessous de 30cm n'est pas nécessaire pour STICS ☒ on peut extrapoler le 0-20cm jusqu'à 30cm.

(décarbonatée, grâce à une fonction de pédo-transfert<sup>10</sup>) est la seule disponible pour l'aire d'étude entière<sup>11</sup>. Une opération d'extrapolation verticale devra donc être menée sur cette variable spécifique afin d'approcher la valeur d'argile décarbonatée en-deçà des 20 premiers centimètres.

### 2.5.3. LUCAS physical :

Lien : <https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data>

Description : Produit en 2015 sur la base des 20.000 points LUCAS acquis en 2009, cette autre étude dérivée conserve une résolution de 500m. Elle a mobilisé la méthode des régressions multivariée par splines adaptatives. Elle renseigne les variables suivantes :

- Teneur en argile (%) dans la couche arable (0-20 cm)
- Teneur en limon (%) dans la couche arable
- Teneur en sable (%) dans la couche arable
- Teneur en cailloux (%) dans la couche arable
- Densité apparente dérivée des ensembles de données sur la texture du sol (obtenue à partir de la densité d'emballage et de la teneur en argile du sol suivant l'équation de Jones et al., 2003)
- Classes de texture du sol de l'USDA dérivées de cartes d'argile, de limon et de sable
- Capacité en eau disponible (AWC) pour la fraction de terre fine de la couche arable

Commentaire : On conservera ici les valeurs de textures, de cailloux mais ni la capacité en eau disponible ni la densité apparente ne seront mobilisées. La seconde a bénéficié en 2024 d'un recalcul plus précis (cf. partie 6.3.1 suivante). La première peut être substituée par des indices calculés sur la base d'équation de pédo-transfert mieux adaptées à la demande en entrée de STICS.

### 2.5.4. Bulk density assessment - LUCAS :

Lien : <https://esdac.jrc.ec.europa.eu/content/soil-bulk-density-europe>

Description : Produite en 2024, sur la base de quelques 6.000 points acquis jusqu'en 2018, la carte est de haute résolution (100 m). Les résultats de la modélisation ont montré une très bonne prédiction ( $R^2=0,66$ ) de la densité apparente pour la profondeur de 0 à 20 cm qui surpasse les évaluations précédentes. La résolution spatiale est de 100m.

Commentaire : Cette nouvelle estimation de la densité apparente présente deux innovations principales. La première c'est de discriminer au sein du « Topsoil » les horizons 0-10cm et 10-20cm. Cela permettra de renseigner un horizon 0-10cm au sein de STICS qui prendra au moins comme information cette variation de la densité apparente. La seconde c'est l'usage d'un « Cubist rule-based regression model » qui implique que beaucoup de données ont été mobilisées pour inférer cette carte de densité apparente. Il y a donc un certain risque de colinéarité à utiliser cette valeur comme base de futures estimations d'autres variables.

---

<sup>10</sup> # clay after decarbonatation in % (Beaudoin et al.)

$\text{argi} <- ((\text{soil\_L1\$TOT\_CLAY} * 10) - (-0.00054 * (\text{CaCO3} * 10)^2 + 0.54 * (\text{CaCO3} * 10))) / 10$

<sup>11</sup> Sauf découverte contraire. En date du 19/06/24.

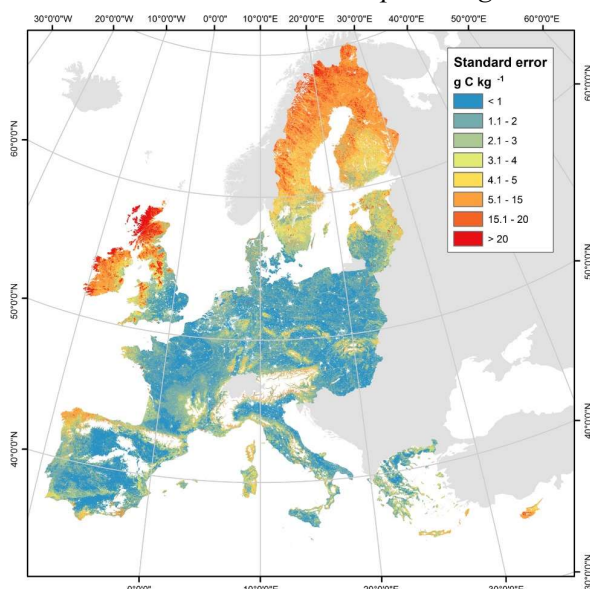
### 2.5.5. Topsoil Soil Organic Carbon - LUCAS :

Lien : <https://esdac.jrc.ec.europa.eu/content/topsoil-soil-organic-carbon-lucas-eu25>

Description : Produite en 2014, cette carte est basée sur une interpolation spatiale sur la base des points acquis en 2009. Elle renseigne en particulier la couche 0-20cm. Elle reprend la méthode des régressions multivariées par splines adaptatives, et un ensemble de covariables environnementales sélectionnées, à savoir la pente, la couverture terrestre, la température annuelle cumulée, la productivité primaire nette, latitude et longitude. Y est associée une carte de l'erreur type des prédictions du modèle OC (cf. ci-dessous). La résolution est de 500m

Commentaire : Carte qui ne se distingue en qualité du jeu de donnée « Chemicals » que par son année de production antérieure et par les covariables en entrée. Le modèle est en revanche bien moins performant ( $R^2=0.27$ ).

Figure 5 : Carte de l'erreur standard de la donnée Topsoil Organic Carbon Lucas



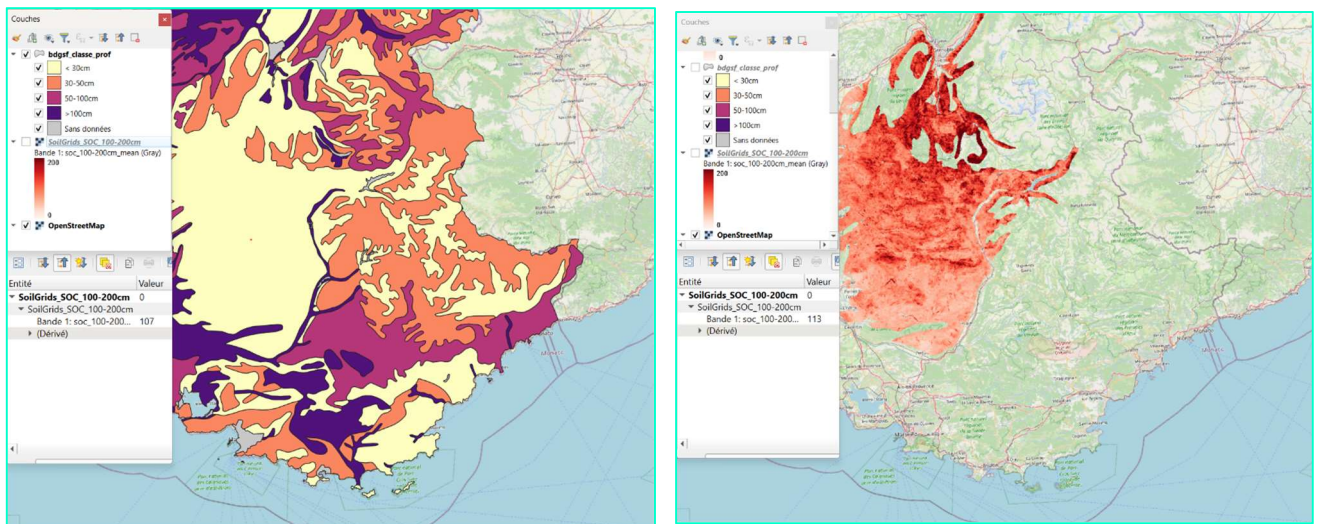
### 2.6. SoilGrids :

Lien : <https://soilgrids.org>

Description : Produite en 2017 et mise à jour en 2020, la base de donnée de l'ISRIC compile un très grand nombre de base de données pédologiques à l'échelle du globe. Elle en produit ensuite une harmonisation sémantique et spatiale qui permet de mettre à disposition un ensemble 11 attributs téléchargeables (liste dans le tableau en Annexe 6). C'est un système de cartographie numérique globale des sols qui utilise des modèles de prédiction ajustés à l'aide de plus de 230 000 observations de profil de sol de la base de données WoSIS et d'une série de covariables environnementales (méthode Digital Soil Mapping). Les covariables ont été sélectionnées à partir d'un pool de plus de 400 couches environnementales à partir de produits dérivés de l'observation de la Terre et d'autres informations environnementales, y compris le climat, la couverture terrestre et la morphologie du terrain. Les variables sont dérivées sur 5 horizons : 0-5 / 5-15 / 15-30 / 30-60 / 60-100 / 100-200cm. Chacune de ces couches spécifiques est de plus doublée de sa couche d'incertitude. La résolution spatiale est de plus ou moins 200m.

**Commentaire :** On commencera par noter que SoilGrids intègre Spade mais pas Lucas. En effet l'une des principales limites « par design » de SoilGrids c'est l'affichage systématique de valeurs sur 0m à 2m de profondeur, sans donner de carte d'estimation de la profondeur des sols, et d'information sur les sources utilisées pour ces interpolations aux couches en-deçà de 60cm. Cela peut donc générer des incohérences. Par exemple, dans des régions connues comme ayant des sols peu profonds, comme en Provence, SoilGrids affiche des taux de carbone organique supérieur à 10%, au-delà du mètre de profondeur (Figure 6). Il est donc nécessaire de relativiser la précision des estimations donnée par SoilGrids. En revanche la richesse sémantique permettrait de renseigner presque intégralement le modèle STICS sur cette base, à l'exception notable de la variable : CaCo3. Enfin, on notera que le téléchargement de la ressource se fait via une API qu'il faut encapsuler dans un script R, livré dans l'archive attachée à ce rapport. Il s'attache notamment à la question de la conversion des systèmes de projection qui induit des déformations relativement importantes de la maille carrée native.

*Figure 6 : Capture d'écran d'une unité pédologique de la BDAT renseignée avec une épaisseur de sol de 30cm ou moins, et pourtant indiquée comme dépassant régulièrement les 100dg/kg de carbone organique au-delà du mètre de profondeur, selon SoilGrids*



### 3. Des SMUs aux USMs

Le choix des données à mobiliser pour alimenter un modèle intégré comme STICS est toujours déterminé par la construction des unités de simulation (USM). La géographie de ces dernières est la base sur laquelle sont agrégées les données. L'étude 4p1000<sup>12</sup> s'appuyait sur une grille avec des mailles de 8km sur 8km sur la France métropolitaine. L'origine de ce découpage vient des données climatiques (maille SAFRAN) dont il est le format de livraison natif. Ces mailles ont été ensuite sub-divisées en fonction des SMU présentes au sein de chaque maille. On prendra cette méthodologie pour le travail SLAM-B.

A titre de prototypage de la solution pour SLAM-B, j'ai réalisé un travail d'extension de cette grille à l'échelle de l'Europe sous PostgreSql. Depuis la projection originale du Jeu de données (EPSG:27572 - NTF (Paris) / Lambert zone II) vers la projection européenne (EPSG:3035 - ETRS89-extended / LAEA Europe), ce qui provoque une légère déformation aux marges orientales et septentrionales du continent. De l'ordre du kilomètre carré, cette déformation est imperceptible sur une carte régionale, la surface médiane des cellules étant de 63km<sup>2</sup>. La répartition de ces déformations est donnée par le diagramme en boîte suivant :

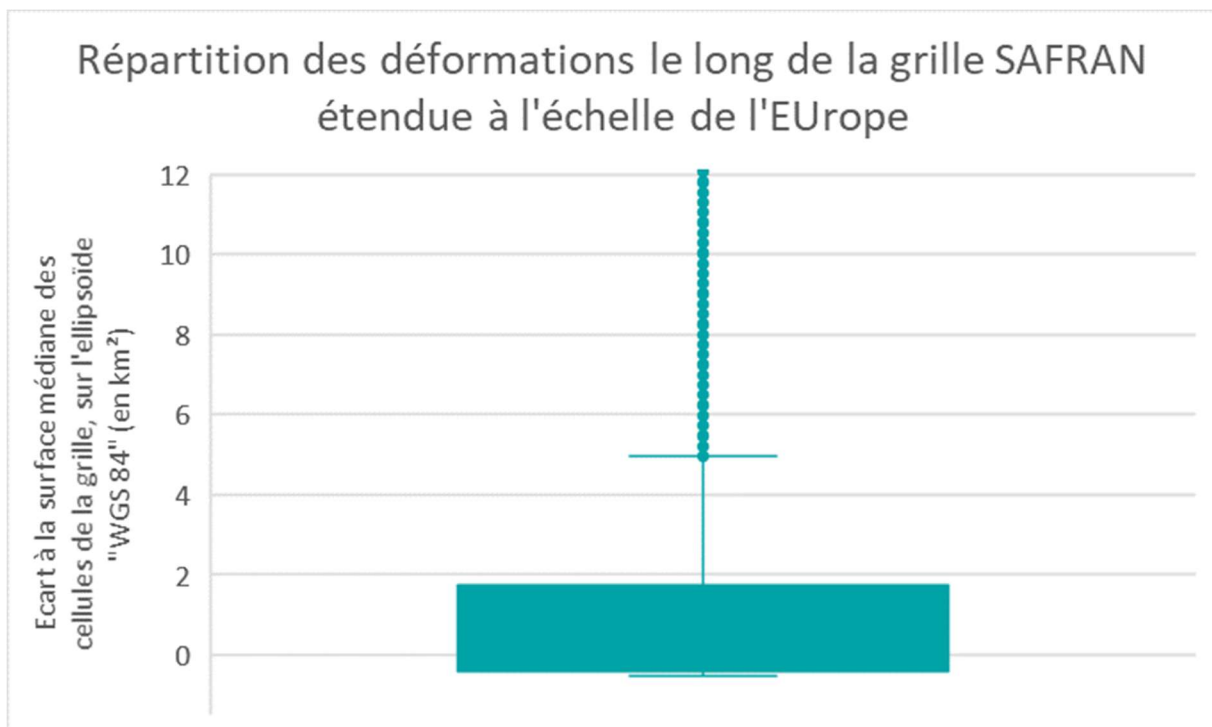


Figure 7: Diagramme en boîte des écarts de surfaces entre cellules d'une même grille carrée SAFRAN, étendue à l'échelle européenne

De la même manière que dans l'étude 4p1000, on sélectionne les polygones d'étude (c'est à dire les pixels sur lesquels se feront les simulations STICS) sur la base d'un critère quantitatif. On retient ici la part de surface de terres agricoles dans la maille. La préparation du filtrage des cellules obtenues selon la surface agricole relative est effectuée sous QGis grâce à la fonction de statistiques de zones. Pour ce faire on procède au téléchargement du Corine Land Cover (CLC) Europe de 2018 (<https://land.copernicus.eu/en/products/corine-land-cover>) , et l'on effectue un traitement sous R pour produire un masque sur les seuls postes agricoles de

<sup>12</sup> Cf. Pellerin *et al.* 2020

l'occupation du sol. La finalisation de cette sélection n'étant pas une priorité (elle peut se faire après le renseignement global du système d'information pédologique à l'échelle de l'Europe).

Les postes sélectionnés sont les suivants :

*Non-irrigated arable land (12)*

*Permanently irrigated land (13)*

*Rice fields (14)*

*Vineyards (15)*

*Fruit trees and berry plantations (16)*

*Olive groves (17)*

*Pastures (18)*

*Annual crops associated with permanent crops (19)*

*Complex cultivation patterns (20)*

*Land principally occupied by agriculture, with significant areas of natural vegetation (21)*

*Agro-forestry areast (22)*

*Natural grasslands (26)*

Il est à noter que cette procédure devra être comparée avec la procédure MAELIA de sélection des sols pertinents (cf. SLAM-B GT.1 et GT.3.1)

Remarque : Ce processus de renseignement de la surface agricole d'après le CLC est associé au script « Filtrage\_CLC.R », tel qu'indiqué dans la chaîne algorithmique.

Une fois obtenu le masque culture / prairie / reste, et les mailles d'intérêts sélectionnées, un travail exploratoire a été mené qui permet de recouper les mailles avec la donnée des « *Soil Mapping Units* » (SMU) de l'ESDB. Ce croisement produit donc un sous-découpage de la maille carrée. Cela permet d'obtenir quelques 14.000 entités (à l'échelle de la France seule) considérées comme homogènes du point de vue pédologique. Ainsi les opérations de moyennes spatiales sur les données quantitatives semblent plus pertinentes qu'elles ne l'auraient été sans cette opération.

Enfin un dernier filtrage est effectué sur la base de la surface occupée par ces entités. Ne sont conservées que celles qui occupent plus de 100 ha de terres agricoles. On conserve plus de 98% de la SAU totale de cette manière.

Remarque : l'utilisation d'une autre donnée météorologique que les données françaises, délivrées sous la maille SAFRAN, pourraient entraîner la reconstruction de ces unités pédoclimatiques sur la base d'une grille différente. Le processus logique reste toutefois strictement le même.

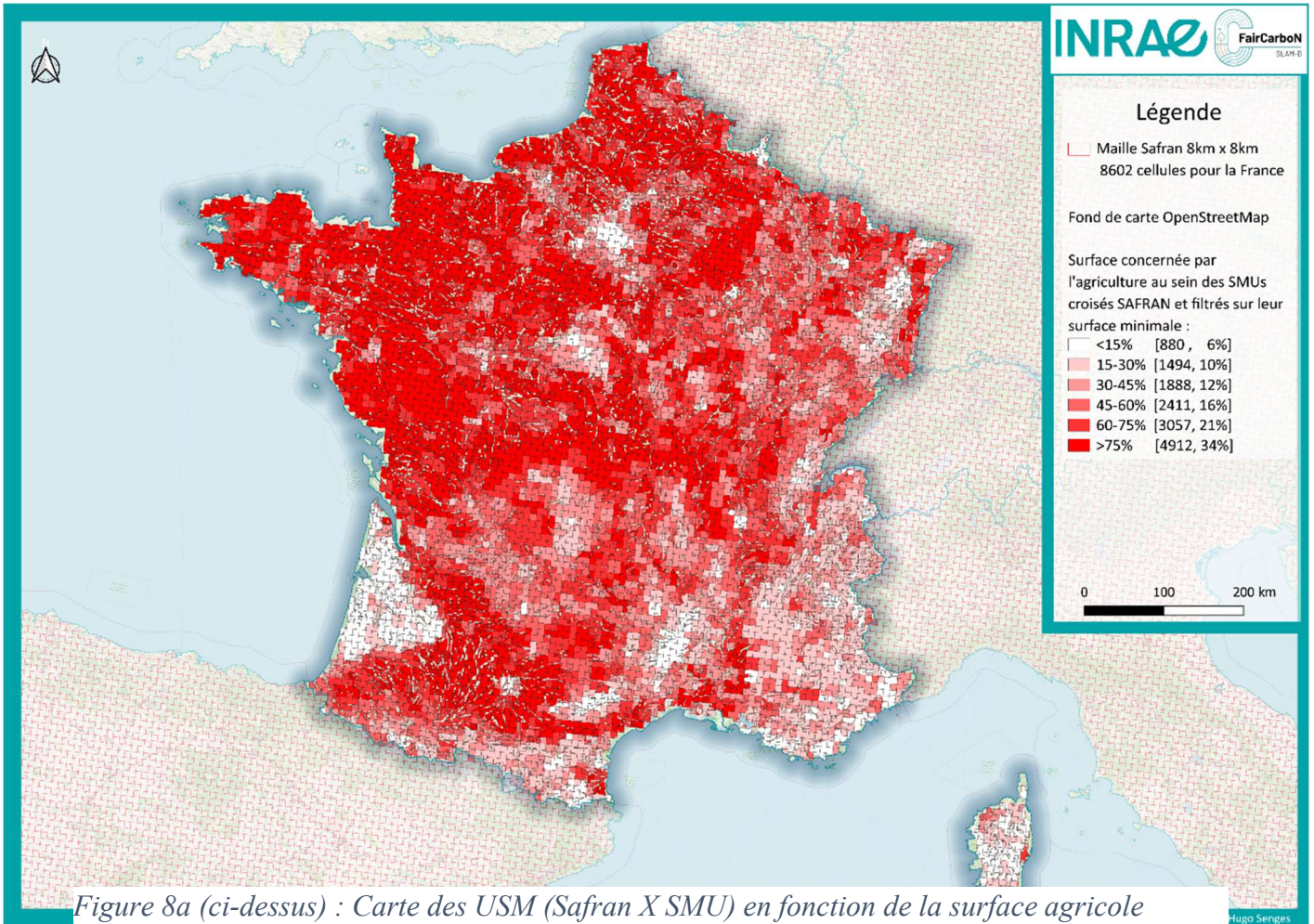
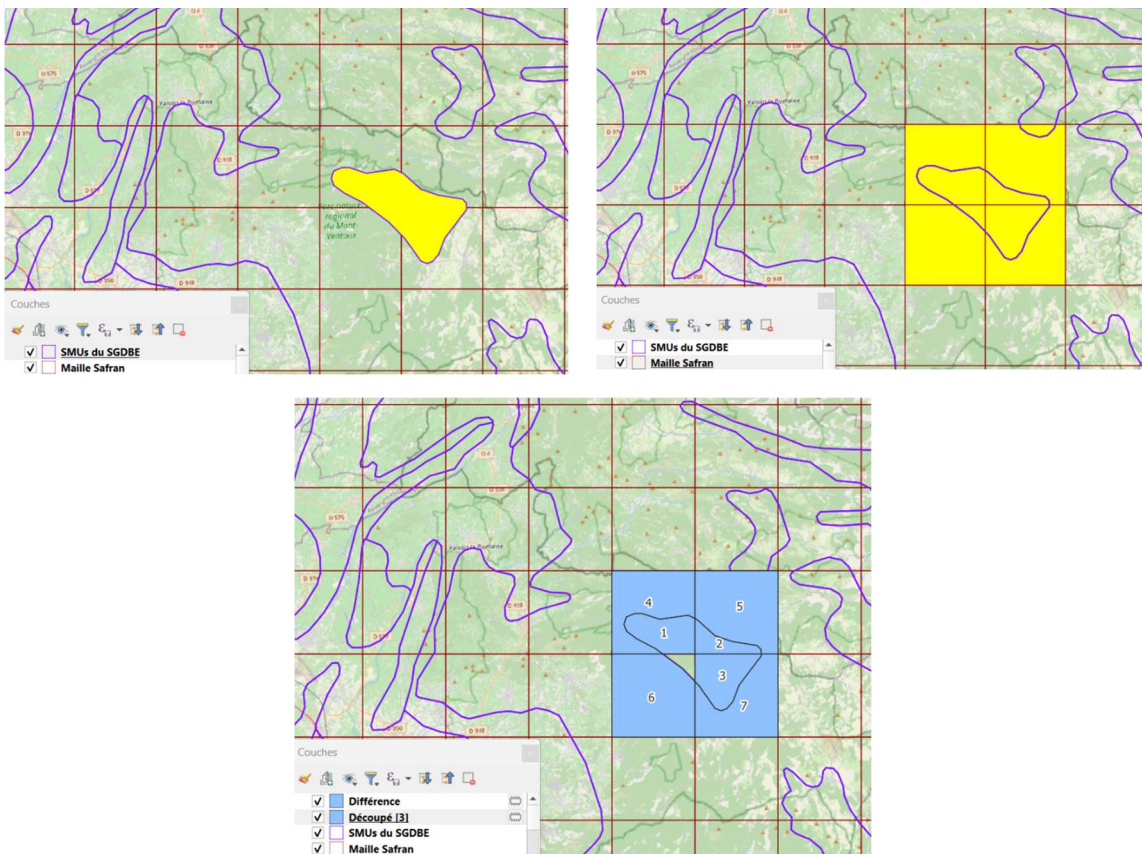


Figure 8a (ci-dessus) : Carte des USM (Safran X SMU) en fonction de la surface agricole qu'elles contiennent

Figure 8b (ci-dessous) : séquence logique de création des USM



## 4. Représentation du profil de sol dans les différentes BD

Une fois établie la dimension planimétrique (l'extension sur le plan de la carte) des unités de simulation (USM), il faut caractériser une dimension supplémentaire : le profil du sol. Les bases de données sol que nous avons étudiées pour caractériser le profil, ne renvoient pas toujours à la même approche en termes de découpage des horizons de sol le long du profil. Ainsi comme vu précédemment dans la description des ressources sol, les données de l'ESDB se limitent à renseigner deux couches d'information par variable d'intérêt : le « Topsoil » (la couche 0-50cm de profondeur par rapport à la surface de sol) et le « Subsoil » en-dessous. On y trouve également différentes informations de profondeurs (« avant changement de texture » / « colonisable par les racines » / « avant la roche ») qui permettent de distinguer trois horizons supplémentaires (dont la profondeur est variable en fonctions des SMU) en plus de ce découpage arbitraire et homogène entre Top/Sub-soil. Les données Lucas se limitent aux 20 premiers centimètres de sol, ce qui constitue une limite. Enfin les données SoilGrids sont renseignées sur un découpage assez précis, mais tous les horizons ainsi dessinés ne sont pas également pertinents, comme cela a été montré précédemment.

Finalement, si l'on veut distinguer pour le modèle STICS différents horizons pédologiques, il nous faut composer avec les informations disponibles à ce sujet. Ci-dessous les variables mobilisables sur la question de la profondeur du sol :

Tableau 1: Sources de données sur les différentes profondeurs des sols

Profondeur renseignée	Intitulé	Source	Lien
Distance au premier horizon argileux	DGH (distance to a gleyed horizon)	PTRDB	<a href="https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes">https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes</a>
Distance au premier horizon imperméable	DIMP (distance to an impermeable layer)		
Distance à la roche	Distance to rock	SGDBE	<a href="https://esdac.jrc.ec.europa.eu/ESDB_Archive/ESDB_Data_Distribution/ESDB_Data_full_distribution/ESDB_data_vx.cfm">https://esdac.jrc.ec.europa.eu/ESDB_Archive/ESDB_Data_Distribution/ESDB_Data_full_distribution/ESDB_data_vx.cfm</a>
Distance au premier changement de texture	Distance to textural change		
Distance d'exploration racinaire	Distance available to roots	Derived data 2013	<a href="https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data">https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data</a>
Distance à la roche	Profondeur des sols	GIS-Sol : Profondeur des sols de France	<a href="https://www.gissol.fr/donnees/cartes/la-profondeur-des-sols-en-france-metropolitaine-1493">https://www.gissol.fr/donnees/cartes/la-profondeur-des-sols-en-france-metropolitaine-1493</a>

Le tableau 2 donne un comparatif des concepts de « profil de sol » dans les différentes ressources (les bordures rouges figurent des profondeurs variables) :

Tableau 2: Table des profondeurs disponibles dans les données recensées

Données	SGDBE	ESDDD	ESDAC	SoilGrid	Lucas	Lucas_BD_Assesment
<b>Profondeur (en cm)</b>						
0-5	Distance to Textural Change	Distance available for Roots	TopSoil ESDAC	0-5	Yes	Yes
0-10				5-15		
10-15				15-30		
15-20				30-60	None	
20-30						
30-50	Distance to Rocks		SubSoil ESDAC	30-60	None	None
50-60				60-100		
60-100				>100		
100-200						

## 5 Construction des jeux de données d'entrée pour le modèle STICS

Les différentes données disponibles ayant été décrites dans la partie précédente, il s'agit dans ce travail d'en sélectionner une composition qui permette de produire les paramètres d'entrée du modèle STICS, sur le compartiment du sol. Le tableau 3 ci-dessous donne ces paramètres d'entrée:

Tableau 3 : Paramètres "sol" du modèles STICS

Mode	alias	Description	Unité
Quantitatif	argi	clay content after decarbonation on profHum	%
	norg	soil organic N content in the first soil layer (constant down to the depth profhum, dry soil)	%
	profhum	maximum soil depth with an active biological activity	cm
	calc	total carbonate content	%
	pH	Initial soil pH (water solution)	pH
	conconseuil	minimum concentration of NO <sub>3</sub> -N in soil (unavailable for leaching and for uptake)	kg/ha/mm
	albedo	albedo of the bare dry soil	SD
	q0	cumulative soil evaporation above which evaporation decreased	mm.d-1
	ruisolnu	fraction of runoff (relative to total rainfall) in a bare soil	SD
	obstarac	soil depth at which root growth is stopped due to physical constraints	cm
	pluibat	minimal amount of rain required to create a soil crust (a value must be given but if the plt.xml vigueurbat parameter is equal to 1 then the parameter is inactive)	mm.d-1
	malchbat	mulch depth at which a crust occurs (a value must be given but if in the plt.xml the vigueurbat parameter is equal to 1 then the parameter is inactive)	cm
	zesx	maximal soil depth affected by soil evaporation	cm
	cfes	parameter defining the soil contribution to evaporation versus depth	SD
	z0solnu	roughness length of bare soil	m
	CsurNsol	initial C to N ratio of soil humus	g.g-1
	finert	initial fraction of inert pool in the soil organic pool (stable SON/ total SON)	SD
penterui	runoff coefficient taking into account the plant mulch	SD	
Booléen	pebbles	option to take into account pebbles in the water and N balances	Y/N
	macroporosity	option to activate calculation of water flux in soil macroporosity	Y/N
	cracks	option to activate an additional water compartment for swelling soils	Y/N
	artificial drainage	option to simulate artificial drainage	Y/N
	capillary rise	option to activate capillary rise	Y/N
	nitrification	option to activate the nitrification model	Y/N
	denitrification	option to activate the denitrification model	Y/N
Quantitatif (+1 classifié) x 5 layers max	epc	thickness of each soil layer	cm
	HCCF	gravimetric water content at field capacity of each soil layer (10 kPa ou 33 kPa)	%
	HMINF	gravimetric water content at wilting point of each soil layer (1500kPa)	%
	DAF	bulk density of the fine earth fraction in each soil layer (without pebbles)	g.cm-3
	cailloux	volumetric content of pebbles per soil layer	%
	typecailloux	Pebbles type : 1= Beauce limestone, 2 = Beauce limestone, 3= Lutetian limestone, 4= Lutetian Brackish marl and limestone, 5=morainic gravels, 6=unweathered flint, sandstone or granite, 7=weathered granite, 8=Jurassic limestone, 9= Pebbles from Magneraud	Code : [1;9]
	infil	infiltrability rate at the base of each soil layer (if codemacropor = 1)	mm.d-1
	epd	thickness of mixing cells in each soil layer (= 2 * dispersion length)	cm

La plupart de ces paramètres peuvent être renseignés directement depuis les données évoquées plus haut. Toutefois on peut distinguer deux catégories au sein des paramètres restants.

### 5.1. Les caractéristiques physico-chimiques des horizons de sol

- Cas des paramètres que l'on peut renseigner avec une constante. La proposition depuis le script Soil.R d'Avril 2024, est de fixer par défaut :
  - Concseuil (« *minimum concentration of NO3-N in soil* ») → 0.01
  - Codenitrif & codedenitrif (« *option to simulate artificial* ») → 1
  - Profhum (« *maximum soil depth with an active biological activity* ») → 30cm<sup>13</sup>
  - Obstrac (« *soil depth at which root growth is stopped due to physical constraints* ») → 200cm

Initialement fixe, le rapport C : N est finalement renseigné grâce à la donnée Lucas et/ou SoilGrid. Il fait l'objet d'un filtre : SI C/N > 25 alors le sol est écarté du plan de simulation (cas par exemple en Sicile de sols avec des valeurs à 48 depuis SoilGrid – visiblement anormales).

Ce filtre concerne 0.5% des sols agricoles totaux (toujours selon la donnée SoilGrid).

- La seconde catégorie de paramètres non immédiatement renseignables avec les données évoquées ci-avant, correspond aux indicateurs dérivés, qui peuvent être calculés (ou estimés) sur la base de fonctions de pédo-transfert (ou de règles de pédo-transfert<sup>14</sup>).

A noter : le compartiment sol de STICS n'est pas le seul concerné par cette question des valeurs « à calculer »<sup>15</sup>.

### 5.2. Fonctions de pédotransfert (FPT)

Les paramètres du compartiment sol de STICS qui demandent à (ou peuvent être) retrouvés par calcul sont les suivants :

- Humidité au point de flétrissement permanent - HMINF,
- Humidité à la capacité au champ - HCCF,
- Taux d'argile décarbonatée,
- Ordre et épaisseurs des couche de sols simulées - epc (X5 max).
- Albédo,
- Évaporation cumulée au sol - q0,

Fonctions de pédotransfert étudiées pour l'albédo, l'évaporation cumulée q0 et la décarbonatation :

- $\text{Albédo} = \text{round}(- (0.0226 * \ln(\text{soc})) + 0.1502, 3)$  <sup>16</sup>

- $q0 = \text{round}(8 / (1 + (\text{argi}/20)^4), 3)$  <sup>17</sup>

---

<sup>13</sup> Ce paramètre a été mis à jour ultérieurement, et est redevenu variable. Cf Senges H., Raynal H., Constantin J., Clivot H., 2025. Rapport technique sur la définition du plan de simulation et l'identification des unités pédoclimatiques pertinentes, dans le cadre du projet ANR-22-PEXF-0003 – projet SLAM-B

<sup>14</sup> La distinction est opérée dans la publication sur PTRDB (King D. et al. 1994) :

Une fonction est identifiée au sens mathématique (continue, de  $\mathbb{R}^x$  dans  $\mathbb{R}$ ) et elle peut être qualifiée statistiquement. Une règle est issue d'un « dire d'expert » est associée souvent une catégorie (typiquement une STU) à une valeur (ou un ensemble de valeur).

<sup>15</sup> L'exemple de l'évapotranspiration (ETP) recalculée par STICS dans le compartiment climat. Il est proposé le 17/06 de faire de cette sortie intermédiaire de STICS un livrable du projet. A comparer avec L'ETP-Penman (ou Penman-Montheith), calculable (de façon déterministe) à partir de données climatiques. L'AMG est également contenue dans STICS.

<sup>16</sup> D'après Gaiser et al. 2010. Les résultats sont à comparer avec les cartes présentées dans « D. Carrer et al 2018 Environ. Res. Lett »

<sup>17</sup> D'après Constantin J. & Raynal H., 2022.

- clay after decarbonation (%) =  $((\text{Clay}\% \cdot 10) - (0.00054 \cdot (\text{CaCO}_3 \cdot 10)^2 + 0.54 \cdot (\text{CaCO}_3 \cdot 10))) / 10$  <sup>18</sup>

- Variables demandées :

- soc, teneur en C organique du sol dans la première couche de sol (en %),
- argi, teneur en argile après décarbonation (en %),
- Clay, teneur en argile avant décarbonation (en %),
- CaCO<sub>3</sub>, teneur totale en carbonate (en %) sur les 30 premiers cm de sol.

L'un des problèmes soulevés par ces FPT, c'est que les seules valeurs existantes de CaCO<sub>3</sub> à l'échelle européenne proviennent de l'étude LUCAS (présentée en partie 2.5.2). La recherche sur les bases de données existantes ne permet pas de renseigner des valeurs de CaCO<sub>3</sub> pour les horizons en-dessous de 20cm (et donc hors du scope LUCAS).

Toutefois la décarbonatation des argiles n'étant prise en compte par STICS que pour le calcul de l'évaporation cumulée  $q\theta$ , uniquement sur l'horizon de surface *profhum* de 0-30cm<sup>19</sup>, et les indices d'humidité (*HCCF* et *HMINF*) ne prenant pas en compte le CaCO<sub>3</sub>, on décide de négliger cette lacune dans les données et de se contenter de l'information fournie par LUCAS<sup>20</sup>.

Le cas, plus spécifique, des FPT sur les capacités de rétention en eau du sol est détaillé dans la sous-partie suivante.

### 5.2.1. La question de la rétention en eau du sol :

La revue bibliographique a mis à jour de nombreuses FPT concurrentes en matière de capacité de rétention en eau. Les plus adaptées à notre problématique sont présentées dans les tableaux 4 et 5 suivants. Ces indices d'humidité des sols sont des variables déterminantes pour le modèle STICS, et il convient donc, pour chacune d'entre elles, d'identifier la FPT la plus adaptée au cadre de notre étude. Une analyse comparative a été menée sur 5 jeux de données et 3 familles de FPT (fonctions de pédo-transfert) afin de déterminer le choix le plus pertinent en matière de détermination des indices de rétention en eau, au champ (*HCCF*) et au point de flétrissement (*HMINF*).

Ces deux indicateurs renvoient à des points sur la courbe de rétention d'eau du sol en fonction de la pression. On identifie dans l'illustration ci-contre le *HCCF* au  $\Theta_r$  (à -100 hPa)<sup>21</sup> et le *HMINF* au  $\Theta_s$  (à -1500 hPa). Ils sont calculés sur la fraction de « terres fine ». On doit donc, avant de procéder au calcul, déduire du

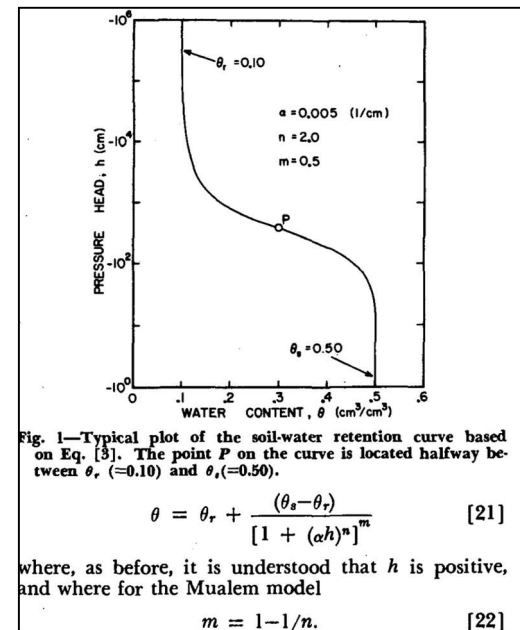


Figure 9 :  
Courbe de rétention d'eau selon la pression, van Genuchten, 1980

<sup>18</sup> D'après Beaudouin *et al.* in Guérif & King, 2007

<sup>19</sup> La valeur de CaCO<sub>3</sub> y sera alors extrapolée depuis celle fournie par LUCAS pour l'horizon 0-20cm.

<sup>20</sup> Une piste à considérer quant au procédé à mobiliser renseigner la valeur « argi » post-décarbonatation pour les horizons inférieurs a été étudiée. Croiser les données SPADE (profil de sols ponctuels contenant l'information CaCO<sub>3</sub>) avec la cartographie de surface CaCO<sub>3</sub> de Lucas pour en induire une sous-couche CaCO<sub>3</sub> extrapolée, permettrait potentiellement de décarbonater les données d'argile.

<sup>21</sup> Il est à noter qu'un arbitrage avec la valeur du  $\Theta$  à -330 hPa a été effectué, étant donnée la littérature qui présente alternativement l'une ou l'autre des valeurs comme « capacité de rétention au champ ». Cet arbitrage s'est fait après le choix des FPT tirées de Dobarco *et al.* 2019, qui privilégie la définition du *HCCF* à -100hPa (ou pF=2.0).

volume total de terre le pourcentage de fragments grossiers (ou « *coarse fragments* » ou « *pebbles* » dans STICS).

On disposait donc *a priori* des données de contrôle suivantes :

- la carte BDGSF de rétention en eau des sols français (mobilisée pour l'étude 4p1000),
- Les valeurs d'AWC<sup>22</sup> de l'ESDB, et de l'ESDDD,
- La valeur pour l'horizon 0-20cm donnée dans Lucas\_Physical,
- Les valeurs affichée dans SoilGrids (mais non publiées en téléchargement).

Un travail de revue de la littérature a permis de mettre à jour les FPT suivantes :

Tableau 4 : Comparaison des FPTs investiguées pour la valeur "au champ"

Dénomination	Fonction HCCF	Fonction « $\theta_{cc}$ »	Fonction « $\theta_{100}$ »
Source	<a href="#">Trouche G. &amp; Morlon P, 1997</a>	<a href="#">Dobarco R., et al., 2021</a>	<a href="#">Cueff, et al., 2021</a>
Formule	$HCCF = 0,5 \text{ Clay}(\%) + 0,218 \text{ Silt}(\%) + 7,35$	$\theta_{cc} = 0,278 + 2,45 \cdot 10^{-3} \text{ argile} - 1,35 \cdot 10^{-3} \text{ sable}$	$\theta_{100\_Lucas} = -9.809 + (0.104 * "Clay@1") - (0.124 * "Bulk\_density@1") + (0.103 * "Silt1@1") + (0.0237 * "N@1") + (0.102 * "Sand1@1")$
Paramètres	Argile et sable	Argile et sable	Argile, densité apparente, Limon, Azote, Sable
Dépendance	Plusieurs jeux de données fournissent ces informations		Lucas ou SoilGrid sont nécessaires pour saisir une valeur en « N »

Tableau 5 : Comparaison des FPTs investiguées pour la valeur "au point de flétrissement"

Dénomination	Fonction HPF	Fonction « $\theta_{fp}$ »	Fonction « $\theta_{1500}$ »
Source	<a href="#">Projet MAELIA</a>	<a href="#">Dobarco R., et al., 2021</a>	<a href="#">Cueff, et al., 2021</a>
Formule	$HPF = \max(0.40 \text{ Clay}\% * BD, 0.005)$	$\theta_{FPF} = 0,08 + 4,01 \cdot 10^{-3} \text{ argile} - 2,93 \cdot 10^{-4} \text{ sable}$	$\theta_{1500\_Lucas} = 0.145 + (0.00256 * "Clay@1") - (0.0856 * "Bulk\_density@1") + (0.007 * "pH\_H2O@1") + (0.000624 * "Silt1@1")$
Paramètres	Argile et Densité apparente	Argile et sable	Argile, densité apparente, pH, Limon
Dépendance	Plusieurs jeux de données fournissent ces informations		Lucas ou SoilGrid sont nécessaires pour saisir une valeur en « N »

### 5.2.2. De la qualification des FPTs

Afin de sélectionner les FPT les plus appropriées pour notre travail, on compare les sorties de chacune d'entre elles lorsqu'on y fait passer l'ensemble des données disponibles en entrée.

Le test mené porte principalement sur les FPT du tableau n°4, permettant d'estimer la capacité de rétention au champ.

Les données d'entrées étant manipulables sous SIG, on a produit sous QGis la chaîne de traitement suivante :

<sup>22</sup> La réserve utile ou « Available water capacity – AWC ») est la différence entre ces indicateur HCCF et HMINF. Les données de « Réserves utiles » identifiées dans la partie 2 ne sont pas suffisamment précises pour être mobilisées.

- i. Pré-traitement des 5 jeux de données téléchargée :  
{SoilGrids ; Lucas ; SGDBE ; SolHydro ; ESDDD}.  
L'opération consiste à :
  - convertir les données dans le même système de projection (EPSG : 3035) ;
  - masquer la donnée sur la même zone (France métropolitaine) ;
  - à supprimer les valeurs nulles ;
  - à éventuellement agréger des données entre elles lorsqu'elles sont livrées en plusieurs fichiers selon l'épaisseur (comme c'est le cas de SoilGrids).
- ii. Production grâce à la calculatrice raster des images « HPF » et « HCCF », selon les 6 formules indiquées en Annexe 7.
- iii. Renseignement par résumé statistique de la couche vectorielle représentant les USMs, en fonction des pixels (HPF, HCCF) contenus dans chaque USM.

A noter : afin de comparer des indicateurs renvoyant à la même réalité, et d'après la limitation de profondeur des données LUCAS, on se limite à l'analyse du HCCF sur les 20 premiers cm.

On se retrouve ainsi avec une couche vectorielle de 14642 entités, dont la table attributaires contient 60 colonnes réparties comme suit :

- [ 15 colonnes donnant pour chaque entités la médiane de la « valeur au champ » selon les appariements (donnée ; FPT) ;
- 15 colonnes donnant les écarts-type de valeurs de pixel contenu par chaque entité pour les même appariements ;
- 15 colonnes de médiane de la « valeur au point de flétrissement », sur le même modèle ;
- 15 colonnes d'écarts-type au « point de flétrissement » sur le même modèle ]

La base de données ainsi constituée, une analyse statistique est menée sur ces résultats intermédiaires. On proposera ensuite une lecture plus cartographique de ces résultats, qui permet de s'assurer de la cohérence du rendu.

La composition de 3 FPT par 5 jeux de données donne 15 séries de HCCF (et symétriquement 15 séries de HPF) dont les tableaux ci-dessous montrent la comparaison des médianes :

*Tableau 6 : Résumé des médianes des séries de HCCF*

Médianes des séries selon les FPT utilisées	Trouche	Saby 2021	Cueff 2021	Moyenne
SoilGrids :	0,303	0,303	0,405	0,337
Lucas - Physical proprieties	0,287	0,288	0,389	0,321
SGBDE	0,288	0,279	0,275	0,280
SolHydro	0,278	0,291	0,295	0,288
« Derived datas », 2013	0,264	0,279	0,292	0,278
Moyenne	0,284	0,288	0,331	0,301

Tableau 7 : Résumé des médianes des séries de HPF

Médianes des séries selon les FPT utilisées	Maelia	Dobarco 2021	Cueff 2021	Moyenne
SoilGrid :	0,112	0,178	0,175	0,155
Lucas - Physical proprieties	0,098	0,150	0,151	0,133
SGBDE	0,108	0,172	0,179	0,153
SolHydro	0,094	0,156	0,169	0,140
« Derived datas », 2013	0,094	0,155	0,157	0,135
Moyenne	0,101	0,162	0,166	0,143

La série moyenne pour la « valeur au champ » devrait donc avoir le résumé suivant :

Moyenne : 0.301 / St-dev : 0.025
----------------------------------

On peut visuellement donner à voir la décomposition des séries, formule par formule :

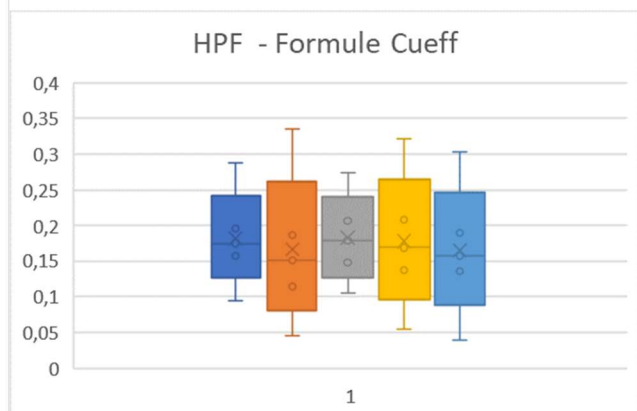
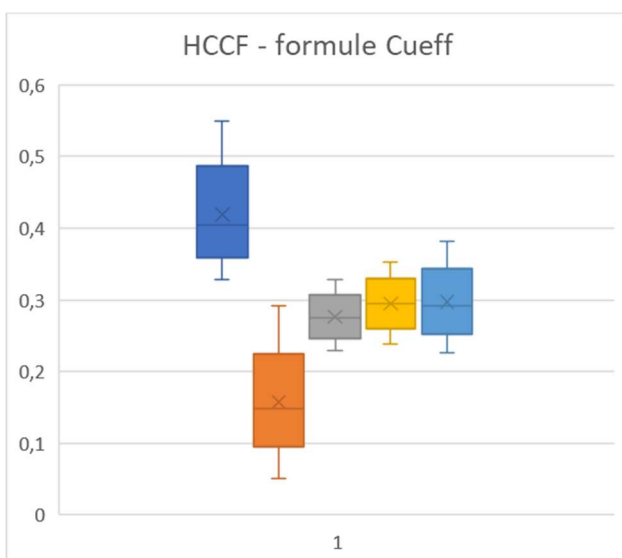
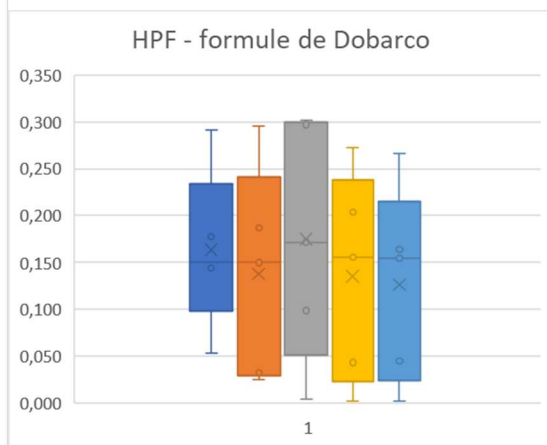
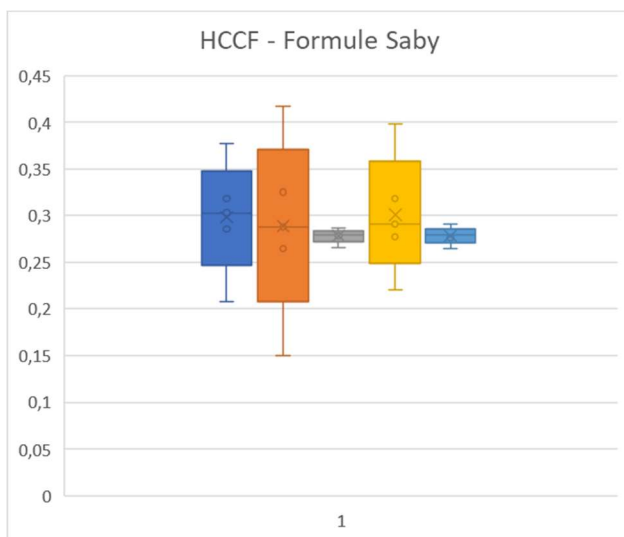
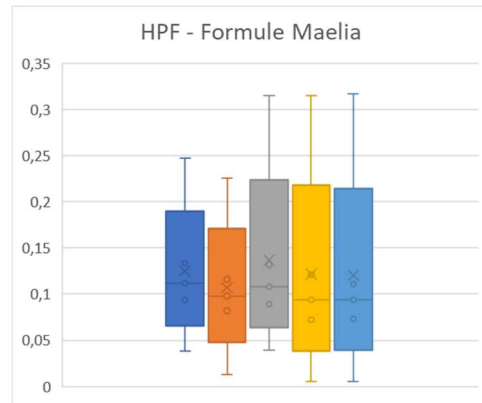
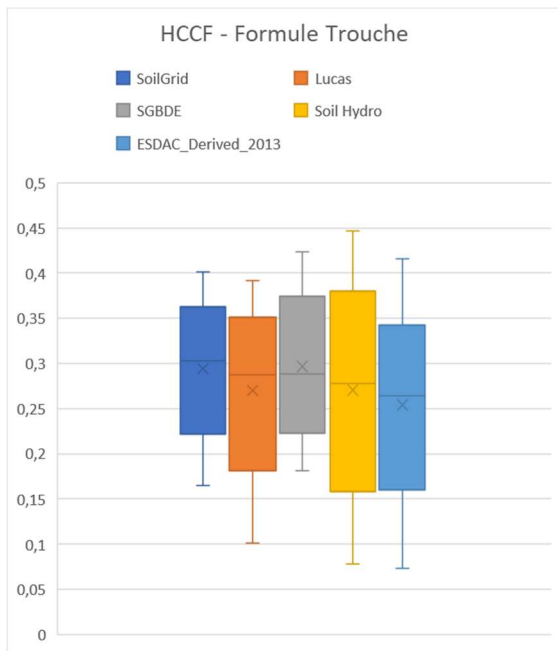


Figure 10 : Diagrammes en boîtes des séries de valeurs en HCCF et HPF selon les formules et les données mobilisées

A ce stade les écarts constatés peuvent s'expliquer par :

- Des erreurs d'appariements entre les unités des données (massiques ou volumiques pour les données de textures) et les unités des résultats des formules (idem).  
C'est visiblement le cas pour les indices HPF dont la version utilisée dans Maelia donne des résultats très en-deçà des deux autres formules mobilisées.
- Une plus faible définition spatiale des séries SGBDE et ESDAC\_Derived qui réduit les valeurs extrêmes
- Des erreurs systématiques dans les sources de données (SoilGrid semble tous le temps légèrement supérieur)

A titre d'illustration des effets induits par les jeux de données en entrée sur le calcul des valeurs en rétention d'eau des sols, on a produit la cartographie suivante. Celle-ci donne à voir les écarts relatifs entre la réserve utile (AWC) obtenue d'après les données de la PTRDB et la même grandeur obtenue d'après les données de texture de SolHydro. On y remarque que pour l'essentiel de la France, l'estimation « PTRDB » est plus de 15% supérieure à l'estimation « SolHydro » :

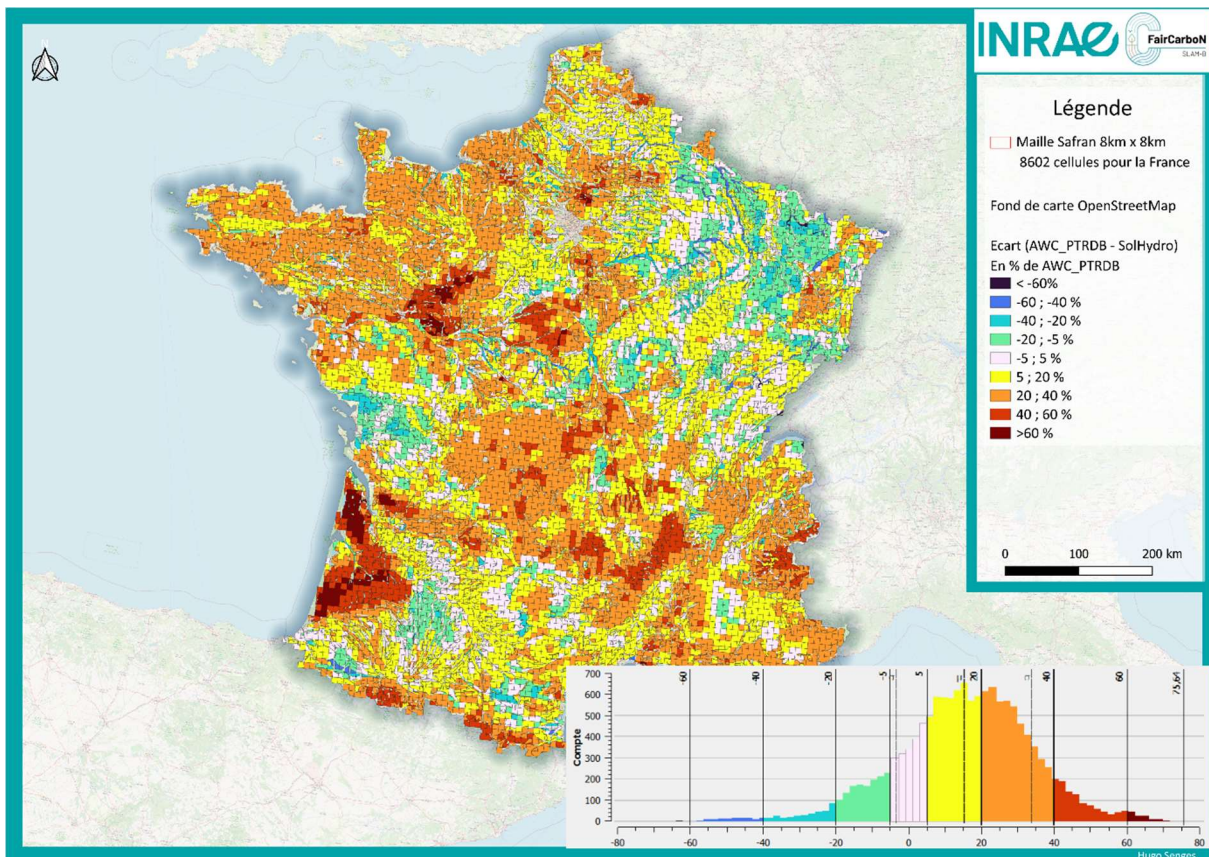


Figure 11 : Carte des écarts de valeurs entre le HMINF calculé grâce aux données Sol Hydro et les données PTRDB

Tout cela invite à considérer Lucas et SoilGrid comme les données disponibles les plus robustes.

Les couples de FPT de Trouche/Maelia et Saby/Dobarco semblent également donner de bons résultats.

On pourrait également tester la FPT proposée par Dobarco<sup>23</sup> en 2019 et incorporant la Bulk Density comme paramètre<sup>24</sup>.

On dispose des données LUCAS pour les horizons 0-10cm et 10-20cm :

$\theta_{FC} =$				$\theta_{pfp} =$			
a	x clay	x sand	x BD	a	x clay	x sand	x BD
0.394	$2.29 \times 10^{-3}$	$-1.24 \times 10^{-3}$	$-7.58 \times 10^{-2}$	0.056	$4.06 \times 10^{-3}$	$-3.14 \times 10^{-4}$	$1.54 \times 10^{-2}$

<sup>23</sup> Voir Annexes pour la comparaison des indices de régression des différentes formules de cette étude

<sup>24</sup> Les fortes incertitudes pesant sur les valeurs de Bulk Density à l'échelle continentale invitent à la précaution dans l'usage de ces données pour le calcul de dérivés (communication personnelle avec Guillaume Coulouma)  
<https://esdac.jrc.ec.europa.eu/themes/soil-bulk-density-assessments>

### 5.2.3. Remarques méthodologiques

- L'un des problèmes soulevés par les FPT est celui de l'interdépendance des données d'entrée qu'elles mobilisent. L'exemple de Bulk Density, qui est elle-même calculée dans le programme LUCAS sur la base de données de texture (% d'argile, de limon et de sable dans la terre fine), montre cette redondance dans les données que l'on pourrait induire en renseignant dans STICS des variables dérivées d'après des grandeurs qui sont elle-même issues de calculs.  
C'est le problème de la colinéarité des jeux de données, qui s'articule aux effets d'échelles différenciés. Alors que certaines variables n'évoluent que progressivement à mesure que l'on se déplace dans le plan, d'autres sont bien plus erratiques. Retrouver les secondes d'après les premières induit donc une perte de variance spatiale.
- La composition d'un jeu de donnée en entrée qui intègre des valeurs provenant à la fois de SoilGrids et de Lucas a été évoqué. Mais les communications avec les pairs ont convergé vers la retenue d'un produit le plus proche possible des jeux déjà publiés. Les discontinuités artificielles que cela pourrait engendrer seraient difficiles à prendre en compte.
- On notera cependant qu'au niveau du modèle, STICS ne fait dialoguer entre elles que des couches du sol adjacentes.
- Afin de limiter les incertitudes supportées par le jeu de données « Sol » que l'on soumettra à STICS, la comparaison entre deux solutions est prévue sur certaines USM. On attachera à cet endroit une importance à leur localisation, afin que celles-ci soient les plus proches possibles des relevés ponctuels qui sous-tendent les interpolations de SoilGrids comme de Lucas<sup>25</sup>.
- Un travail exploratoire mené par Hélène Raynal à MAGELLAN, en Juin 2024, conclue à une part relativement faible dans les variations des sorties de simulations expliquées par les incertitudes liées au variables pédologiques. Cette sensibilité limitée des prévisions agronomiques vis-à-vis de la précision du jeu de données « Sol » invite à concentrer les efforts sur des compartiments plus problématiques, à plus forte incertitude<sup>26</sup>.

---

<sup>25</sup> On pourra par exemple tester ici le [Lucas Harmonized Dataset](#) transmis par Nicolas Gilpart.

<sup>26</sup> On pense ici, par exemple, au problème de la profondeur de sol, dont l'incertitude est d'un ordre de grandeur supérieur à celui des variables de textures et de composition du sol.

## 6. Proposition retenue pour la fondation d'une base de données pédologiques orientée simulation

### 6.1. Solution SoilGrid

Préférée pour son homogénéité et son adéquation avec les contraintes techniques liées à l'usage de STICS, la base SoilGrids servira à construire les fichiers « .soil » du modèle sur l'ensemble de l'aire d'étude.

#### **A NOTER :**

A ce stade cette transcription des données SoilGrids natives vers un système d'information géographique STICS-centré n'est pas investiguée. Le présent rapport n'a pour objet que de détailler et de comparer les différentes sources de données entre elles. L'algorithmie à mobiliser pour effectuer cette transcription sera détaillée dans un rapport dédié<sup>27</sup>.

Il s'agira pour l'essentiel de définir le périmètre des UPC, en lien avec la donnée climatique, et de tirer du tableau des profondeurs donné en partie précédente, les horizons suivants :

*Tableau 8 : Proposition de construction du système d'information pédologique STICS sur la base SoilGrids*

Couche 1	Donnée 0-5cm et 5-15cm de SoilGrid
Couche 2	Donnée 15-30cm de Soilgrid (les sols ayant une profondeur inférieure à 30 cm sont écartés)
Couche 3	Donnée 30-60cm de Soilgrid (éventuellement tronquée en cas de profondeur inférieure à 60cm)
Couche 4	Donnée 60-100cm de Soilgrid (tronquée en cas de profondeur inférieure à 100cm)
Couche 5	Donnée 100-200 de Soilgrid (tronquée en cas de profondeur inférieure à 200cm)

A noter que l'on devra quoi qu'il arrive mobiliser la donnée CaCo3 de Lucas, et qu'il semble judicieux de mobiliser également la donnée en Bulk Density produite sur les relevés Lucas en 2024 sur l'horizon de surface.

---

<sup>27</sup> Cf. Senges H., Raynal H., Constantin J., Clivot H., 2025. Rapport technique sur la définition du plan de simulation et l'identification des unités pédoclimatiques.

## 6.2. Solutions hybride LUCAS/SoilGrid

A priori plus précise sur les horizons de surfaces, cette composition pêche par la discontinuité artificielle qu'elle induit. Elle sera testée dans un second temps sur quelques USM pour évaluer la sensibilité du modèle aux jeux de données pédologiques fournis en entrée :

Tableau 9 : Proposition de construction du système d'information pédologique STICS sur la base Lucas x SoilGrids

Profil Composite	Datas		Layers_STICS
0-10	Lucas BD Assesment	Lucas, Physical & Chemicals	1
0-20	Lucas BD Assesment		2
Depth to Textural Change	Soil_Grid & ESDDD		3
Depth available to Roots			4
Depth to Rocks			5

Le tableau 10 présente la proposition abandonnée de découpage des 5 horizons de sol prévus par STICS et la source d'information **tirée de l'ESDB**.

Tableau 10 : Calculs des épaisseurs de sols par couches dans le modèle STICS, selon les données ESDB

Layers_STICS	Calcul de l'epc (les variables « Distances to » sont en cm depuis la surface)	Epaisseur cumulée	Commentaire
1	10	10	Depuis Lucas_BD_Assessment
2	Min (10 ; « distance to textural change » - 10 )	Min (20 ; D_Text_Chg)	Depuis LUCAS, ou depuis l'ESDB
3	10 - Min (10 ; « distance to textural change » - 10)	Max (20 ; D_Text_Chg)	idem
4	« Depth available to roots »	Variable	Depuis ESDDD
5	« Depth to rocks »	Idem	Depuis ESDB

Il était proposé qu'en cas de « Depth available to roots » supérieur au « Depth to rocks », le cinquième horizon de STICS ne soit pas renseigné. A priori la 5<sup>ème</sup> couche « Distance to rock » ne semble pas nécessaire.

## Bibliographie

- van Genuchten, M. Th. - A closed-form equation for predicting the hydraulic conductivity of unsaturated soils- 1980. Soil Sci. Soc. Am. J.
- King D., Darrousin J., et al. - A geographical Knowledge database on soil properties for environmental studies, based on a 1:1,000,000-scale EC soil map data. - INRA, University of Ghent, Silsoe campus & Land Resource partnership. 1994. 50pp.
- Trouche G., Morlon P. - Comparaison de différentes méthodes d'estimation de la réserve en eau utile des sols (R.U.) dans le périmètre de l'O.G.A.F. Environnement de la zone de Migennes (Yonne) - 1999. Étude et Gestion des Sols, N°6-1
- Wösten J.H.M., Pachepsky Ya.A., Rawls W.J. - Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics - 2001. Journal of Hydrology
- King D., Jones R.J.A, Tomasson A.J., *et al.* - European Land Information System for agri-environmental monitoring - 2003
- Lambert J.J., Daroussin J., Eimberck M., Le Bas C., Jamagne M., King D. & Montanarella D. - Soil Geographical Database for Eurasia & The Mediterranean: Instructions Guide for Elaboration at scale 1:1,000,000 - 2003. INRA & JRC
- Morvan X., Bruand A., Cousin I., Roque J., Baran N. et al. - Prédiction des propriétés de rétention en eau des sols d'un bassin versant à l'aide de fonctions de pédotransfert : influence de la densité apparente et de la teneur en éléments grossiers - 2004. Etude et Gestion des Sols
- IUSS Working Group WRB. 2006. World reference base for soil resources 2006. World Soil Resources Reports No. 103. FAO
- Gaiser T., de Barros I., Sereke F. & Lange F-M. -Validation and reliability of the EPIC model to simulate maize production in small-holder farming systems in tropical sub-humid West Africa and semi-arid Brazil – 2010. Agriculture, Ecosystems & Environment, Volume 135, Issue 4
- Getreuer P. - Image Interpolation with Contour Stencils – CMLA Cachan, IPOL, 2011.
- Hiederer R. Extending Geographic and Thematic Range of SPADE/M with HYPRES Soil Profile Data. JRC & IES. 2011, 49 pp.
- Hiederer R. - Mapping Soil Properties for Europe --- Spatial Representation of Soil Database Attributes - 2013. JRC
- Lugato E., Panagos P., Bampa F., Jones A., Montanarella L.- A new baseline of organic carbon stock in European agricultural soils using a modelling approach - 2013. Global Change Biology
- Miháliková M., Matula S. & Doležalhyprescz F. Database of Soil Hydrophysical Properties in the Czech Republic, 2013, Soil & Water Res., 8: 34–41
- Tóth G., Jones A., Montanarella L., *et al.* - LUCAS TopsoilSurvey methodology, data and results - 2013. JRC
- Breuning-Madsen T., Aagaard Kristensen J. & Balstrøm T. - Final report on the establishment of a SPADE 14 soil profile analytical database connected to the eu soil map at scale 1:1.000.000 - 2014. JRC

- de Brogniez D., Ballabio C. , Stevens A. , Jones R. J. A., Montanarella L., van Wesemael B. - A map of the topsoil organic carbon content of Europe generated by a generalized additive model - 2014. *European Journal of Soil Science*
- Saby, N. P. A., Lemercier, B., Arrouays, D., Leménager, S., Louis, B. P., Millet, F., Paroissien, J. B., Schellenberger, E., Squidadant, H., Swiderski, C., Toutain, B., Walter, C. & Bardy, M. - Le programme Base de Données des Analyses de Terre (BDAT) : Bilan de 20 ans de collecte de résultats d'analyses - 2014. *Etude et Gestion des Sols*
- Ballabio C., Panagos P. & Montanarella L. - Mapping topsoil physical properties at European scale using the LUCAS database - 2015. *Geoderma*
- Ballabio C., Lugato E., Fernández-Ugalde O., Orgiazzi A., Jones A., Borrelli P., Montanarella L. & Panagos P. - Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression - 2019. *Geoderma*
- Dobarco R., Bourennane H., Arrouays D., Saby N., Cousin I. & Martin M. - Uncertainty assessment of GlobalSoilMap soil available water capacity products: A French case study - 2019. *Geoderma*
- Dobarco R., Cousin I., Le Bas C. & Martin M. - Pedotransfer functions for predicting available water capacity in French soils, their applicability domain and associated uncertainty - 2019. *Geoderma*
- Pellerin S. & Bamière L. (pilotes scientifiques), Launay C., Martin R., Schiavo M., Angers D., Augusto L., Balesdent J., Basile-Doelsch I., Bellassen V., Cardinael V., Cécillon L., Ceschia E., Chenu C., Constantin J., Darroussin J., Delacote P., Delame N., Gastal F., Gilbert D., Graux A-I., Guenet B, Houot S., Klumpp K., Letort E., Litrico I., Martin M., Menasseri S., Mézière D., Morvan T., Mosnier C., Roger-Estrade J., Saint-André L., Sierra J., Théron O., Viaud V., Gâteau R., Le Perchec S., Réchauchère O. - Stocker du carbone dans les sols français, Quel potentiel au regard de l'objectif 4 pour 1000 et à quel coût ? Rapport scientifique de l'étude, INRA (France) - 2020.
- Cueff S., Coquet Y., Aubertot J-N., Bel L., Pot V., et al. - Estimation of soil water retention in conservation agriculture using published and new pedotransfer functions - 2021. *Soil and Tillage Research*
- Dobarco R., Mercedes, Bourennane H., Arrouays D., Saby N., Cousin I. & Martin M., - "Réservoir utile des sols de la France métropolitaine". 2021
- Tian Z., Chen J., Cai C., Gao W., Ren T., Heitman J. & Horton R. - New pedotransfer functions for soil water retention curves that better account for bulk density effects - 2021. *Soil and Tillage Research*
- Fernandez-Ugalde, O; Scarpa, S; Orgiazzi, A.; Panagos, P.; Van Liedekerke, M; Marechal A. & Jones. - A. LUCAS 2018 Soil Module. Presentation of dataset and results, Publications Office of the European Union, 2022.
- Liu L., Lu Y., Fu Y., Horton R. & Ren T. - Estimating soil water suction from texture, bulk density and electrical resistivity - 2022. *Geoderma*
- Panagos P., De Rosa D., Liakos L., Labouyrie M., Borrelli P. & Ballabio C. -Soil bulk density assessment in Europe - 2024. *Agriculture, Ecosystems & Environment*

## Glossaire :

JRC : Joint Research Center

FPT : Fonctions de pédo-tranfert

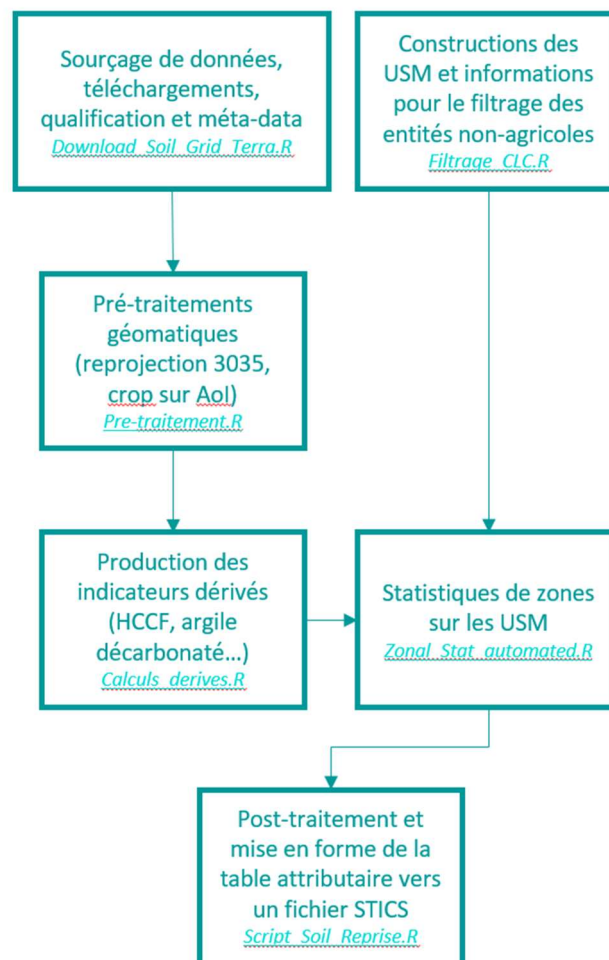
PTR : Pedo-transfer Rule

SMU : Soil Mapping Unit. Unité cartographique cohérente sur le plan pédologique

STU : Soil Taxonomical Unit. Unité typologique de pédologie. Attribué comme libellé à chacune des SMU.

USM : Unité de simulation (entendue ici comme l'extension géographique minimale concernée par une simulation de STICS).

## Chaîne algorithmique :



# Annexes

## Annexe 1 : Attributs de la table SGDBE

NAME	DESCRIPTION	TYPE	SIZE
<b>STU</b>	Soil Typological Unit (STU) identifying number.	Integer number	7
<b>WRB-GRP</b>	Soil Group code of the STU from the World Reference Base (WRB) for Soil Resources.	Character string	2
<b>WRB-ADJ</b>	Soil Adjective code of the STU from the World Reference Base (WRB) for Soil Resources.	Character string	2
<b>WRB-SPE</b>	Complementary code of the STU from the World Reference Base (WRB) for Soil Resources.	Character string	3
<b>FAO90-MG</b>	Soil Major Group code of the STU from the 1990 FAO-UNESCO Soil Revised Legend.	Character string	2
<b>FAO90-UNI</b>	Soil Unit code of the STU from the 1990 FAO-UNESCO Soil Revised Legend.	Character string	1
<b>FAO90-SUB</b>	Soil Sub-unit code of the STU from the 1990 FAO-UNESCO Soil Revised Legend.	Character string	1
<b>SLOPE-DOM</b>	Dominant slope class of the STU.	Integer number	1
<b>SLOPE-SEC</b>	Secondary slope class of the STU.	Integer number	1
<b>ZMIN</b>	Minimum elevation above sea level of the STU (in metres).	Integer number	4
<b>ZMAX</b>	Maximum elevation above sea level of the STU (in metres).	Integer number	4
<b>PAR-MAT-DOM</b>	Code for dominant parent material of the STU.	Integer number	4
<b>PAR-MAT-SEC</b>	Code for secondary parent material of the STU.	Integer number	4
<b>USE-DOM</b>	Code for dominant land use of the STU.	Integer number	2
<b>USE-SEC</b>	Code for secondary land use of the STU.	Integer number	2
<b>AGLIM1</b>	Code for dominant limitation to agricultural use of the STU.	Integer number	2
<b>AGLIM2</b>	Code for secondary limitation to agricultural use of the STU.	Integer number	2
<b>TEXT-SRF-DOM</b>	Dominant surface textural class of the STU.	Integer number	1
<b>TEXT-SRF-SEC</b>	Secondary surface textural class of the STU.	Integer number	1
<b>TEXT-SUB-DOM</b>	Dominant sub-surface textural class of the STU.	Integer number	1
<b>TEXT-SUB-SEC</b>	Secondary sub-surface textural class of the STU.	Integer number	1
<b>TEXT-DEP-CHG</b>	Depth class to a textural change of the dominant and/or secondary surface texture of the STU.	Integer number	1
<b>ROO</b>	Depth class of an obstacle to roots within the soil profile of the STU.	Integer number	1
<b>IL</b>	Depth class of a presence of an impermeable layer within the soil profile of the STU.	Integer number	1
<b>WR</b>	Dominant annual average soil water regime class of the soil profile of the STU.	Integer number	1
<b>WM1</b>	Code for normal presence and purpose of an existing water management system in agricultural land on more than 50 % of the STU.	Integer number	1
<b>WM2</b>	Code for the type of an existing water management system.	Integer number	2
<b>CFL</b>	Code for a global confidence level of the STU description.	Character string	1

Lien vers la nomenclature, attributs par attributs

[https://esdac.jrc.ec.europa.eu/ESDB\\_Archive/ESDB\\_Data\\_Distribution/ESDB\\_Data\\_full\\_distribution/ESDBv2\\_data/popup/sg\\_attr.htm](https://esdac.jrc.ec.europa.eu/ESDB_Archive/ESDB_Data_Distribution/ESDB_Data_full_distribution/ESDBv2_data/popup/sg_attr.htm)

## Annexe 2 : Attributs de la table PTRDB

AGLIM1NNI = Dominant limitation to agricultural use (without no information).

- 0 = No information
- 1 = No limitation to agricultural use
- 2 = Gravelly (over 35% gravels diameter < 7.5 cm)
- 3 = Stony (presence of stones diameter > 7.5 cm, impracticable mechanization)
- 4 = Lithic (coherent and hard rock within 50 cm)
- 5 = Concretionary (over 35% concretions diameter < 7.5 cm near the surface)
- 6 = Petrocalcic (cemented or indurated calcic horizon within 100 cm)
- 7 = Saline (electric conductivity > 4 mS.cm-1 within 100 cm)
- 8 = Sodic (Na/T > 6% within 100 cm)
- 9 = Glaciers and snow-caps
- 10 = Soils disturbed by man
- 20 = Fragic
- 21 = Drained
- 22 = Quasi permanently flooded
- 30 = Eroded phase, erosion
- 31 = Phreatic phase

-----  
AGLIM2NNI = Secondary limitation to agricultural use (without no information).

- 0 = No information
- 1 = No limitation to agricultural use
- 2 = Gravelly (over 35% gravels diameter < 7.5 cm)
- 3 = Stony (presence of stones diameter > 7.5 cm, impracticable mechanization)
- 4 = Lithic (coherent and hard rock within 50 cm)
- 5 = Concretionary (over 35% concretions diameter < 7.5 cm near the surface)
- 6 = Petrocalcic (cemented or indurated calcic horizon within 100 cm)
- 7 = Saline (electric conductivity > 4 mS.cm-1 within 100 cm)
- 8 = Sodic (Na/T > 6% within 100 cm)
- 9 = Glaciers and snow-caps
- 10 = Soils disturbed by man
- 20 = Fragic
- 21 = Drained
- 22 = Quasi permanently flooded
- 30 = Eroded phase, erosion
- 31 = Phreatic phase

-----  
ALT = ELEVATION

- # = No information
- L = Lowlands & intermediate
- U = Uplands & mountains

-----  
ATC

Regrouped accumulated mean annual temperature class (ATC) (source: JRC-MARS)

- H: High (> 3000°C)
- M: Medium (1800-3000°C)
- L: Low (< 1800°C)

-----  
AWC\_SUB = Subsoil available water capacity.

- VL = Very low (~ 0 mm/m)
- L = Low (< 100 mm/m)
- M = Medium (100 - 140 mm/m)
- H = High (140 - 190 mm/m)
- VH = Very high (> 190 mm/m)
- # = No data or not applicable

-----  
AWC\_TOP = Topsoil available water capacity.

- L = Low (< 100 mm/m)
- M = Medium (100 - 140 mm/m)
- H = High (140 - 190 mm/m)
- VH = Very high (> 190 mm/m)
- # = No data or not applicable

-----  
BS\_SUB = Base saturation of the subsoil.

- H = High (> 50 %)
- L = Low (< 50 %)

-----  
BS\_TOP = Base saturation of the topsoil.

- H = High (> 75 %)
- M = Medium (50 - 75 %)
- L = Low (< 50 %)

-----  
CEC SUB = Subsoil cation exchange capacity.

H = High (> 40 cmol(+)/kg)  
M = Medium (15-40 cmol(+)/kg)  
L = Low (< 15 cmol(+)/kg)

-----  
CEC TOP = Topsoil cation exchange capacity.

H = High (> 40 cmol(+)/kg)  
M = Medium (15-40 cmol(+)/kg)  
L = Low (< 15 cmol(+)/kg)

-----  
CRUSTING = Soil crusting class.

1 = Very weak  
2 = Weak  
3 = Moderate  
4 = Strong  
5 = Very strong

-----  
DGH = Depth to a gleyed horizon.

S = Shallow (< 40 cm)  
M = Moderate (40 - 80 cm)  
D = Deep (80 - 120 cm)  
V = Very deep (> 120 cm)

-----  
DIFF = Soil profile differentiation.

H = High differentiation  
L = Low differentiation  
O = No differentiation

-----  
DIMP = Depth to an impermeable layer.

S = Shallow (< 80 cm)  
D = Deep (> 80 cm)

-----  
DR = Depth to rock.

S = Shallow (< 40 cm)  
M = Moderate (40 - 80 cm)  
D = Deep (80 - 120 cm)  
V = Very deep (> 120 cm)

-----  
EAWC SUB = Subsoil easily available water capacity.

VL = Very low (~ 0 mm/m)  
L = Low (< 100 mm/m)  
M = Medium (100 - 140 mm/m)  
H = High (140 - 190 mm/m)  
VH = Very high (> 190 mm/m)  
# = No data or not applicable

-----  
EAWC TOP = Topsoil easily available water capacity.

L = Low (< 100 mm/m)  
M = Medium (100 - 140 mm/m)  
H = High (140 - 190 mm/m)  
VH = Very high (> 190 mm/m)  
# = No data or not applicable

-----  
ERODIBILITY = Soil erodibility class.

1 = Very weak  
2 = Weak  
3 = Moderate  
4 = Strong  
5 = Very strong

-----  
HG = Hydrogeological class.

1R = 1R  
1C = 1C  
1S = 1S  
1L = 1L  
1H = 1H  
1M = 1M  
2 = 2  
3 = 3  
4W = 4W  
4D = 4D

-----  
MIN = Profile mineralogy.

C = Chemical or Geochemical  
M = Mechanical or Physical

MC = Chemical and Mechanical  
ND = No differentiation

---

MIN SUB = Subsoil mineralogy.  
KQ = 1/1 Minerals + Quartz  
KX = 1/1 Min. + Oxy. & Hydroxy.  
MK = 2/1 & 1/1 Minerals  
M = 2/1 & 2/1/1 non swel. Min.  
MS = Swel. & non swel. 2/1 Min.  
S = Swelling 2/1 Minerals  
TV = Vitric Minerals  
TO = Andic Minerals  
NA = Not applicable

---

MIN TOP = Topsoil mineralogy.  
KQ = 1/1 Minerals + Quartz  
KX = 1/1 Min. + Oxy. & Hydroxy.  
MK = 2/1 & 1/1 Minerals  
M = 2/1 & 2/1/1 non swel. Min.  
MS = Swel. & non swel. 2/1 Min.  
S = Swelling 2/1 Minerals  
TV = Vitric Minerals  
TO = Andic Minerals  
NA = Not applicable

---

OC TOP = Topsoil organic carbon content.  
H = High (> 6 %)  
M = Medium (2 - 6 %)  
L = Low (1 - 2 %)  
V = Very low (< 1 %)

---

PD SUB = Subsoil packing density.  
L = Low  
M = Medium  
H = High

---

PD TOP = Topsoil packing density.  
L = Low  
M = Medium  
H = High

---

PEAT = Peat.  
N = No  
Y = Yes

---


PHYS-CHIM = Physico-chemical factor of soil crusting & erodibility.  
1 = Very favourable  
2 = Favourable  
3 = Medium  
4 = Unfavourable  
5 = Very unfavourable

---

PMH = Parent material hydrogeological type.  
R = Porous - Stor. ~ Perm. +  
C = Porous 2 Stor. ~ Perm. +  
S = Porous 1 Stor. + Perm. +  
L = Stor. - Perm. -  
H = Hard. Stor. -- Perm. --  
M = Soft. Stor. -- Perm. --  
# = No information

Annexe 3 : Attributs qualifiés et clés d'identification primaires de la base SAPDE/M

SOURCE	PLOT	METHOD	LAYER
A	X1	M1	L1
			L2
			L3
A	X2	M2	L1
			L2
A	X3	M9	L1
			L2
			L3
B	X3	M2	L1
			L2
			L3



SOURCE	PLOT	METHOD	SEGMENT
A	X1	M1	H1
A	X1	M1	H2
A	X1	M1	H3
A	X2	M2	H1
A	X2	M2	H2
A	X3	M9	H1
A	X3	M9	H2
A	X3	M9	H3
B	X3	M3	H1
B	X3	M3	H2
B	X3	M3	H3

Figure 10: SPADE/M2 Key Attributes

Table 10: Results on Single Parameter Tests of Value Conformity

Field	Range						Status
	Error		Warning		Data		
	MIN E	MAX E	MIN W	MAX W	MIN D	MAX D	
ALT	-500.0	8880.0	-150.0	4500.0	-2.00	2600.00	OK
GWL_N_MEAS	0.0	10000.0	0.1	300.0	25.00	750.00	OK
D_ROO	0.0	10000.0	0.1	300.0	5.00	200.00	OK
D_ROC	0.0	10000.0	0.1	300.0	0.00	235.00	Warning
D_OTHOB	0.0	10000.0	0.1	300.0	0.00	235.00	Warning
HOR_START	0.0	10000.0	0.0	300.0	0.00	235.00	Warning
HOR_END	0.0	10000.0	0.1	1000.0	0.00	999.00	Warning
CLAY	0.0	100.0	0.0	99.0	0.00	81.00	OK
SILT_1	0.0	100.0	0.0	99.0	0.00	77.00	OK
SILT_2	0.0	100.0	0.0	99.0	0.00	92.00	OK
SAND_1	0.0	100.0	0.0	100.0	0.00	96.00	OK
SAND_2	0.0	100.0	0.0	100.0	0.00	92.50	OK
SAND_3	0.0	100.0	0.0	100.0	0.00	100.0	OK
ORG_C	0.0	80.0	0.0	70.0	0.00	56.80	OK
ORG_M	0.0	100.0	0.0	100.0	0.00	109.0	Error
N_TOT	0.0	100.0	0.0	25.0	0.00	21.50	OK
CACO3	0.0	100.0	0.0	100.0	0.00	99.00	OK
CASO4	0.0	100.0	0.0	50.0	0.00	36.00	OK
PH	0.0	14.0	2.5	12.5	2.60	10.0	OK
EC	0.0	100.0	0.0	60.0	0.00	53.60	OK
AR_NA	0.0	100.0	0.0	50.0	0.00	49.00	OK
EXCH_NA_P	0.0	100.0	0.0	75.0	0.00	73.00	OK
EXCH_CA	0.0	1000.0	0.0	115.0	0.00	110.80	OK
EXCH_MG	0.0	500.0	0.0	50.0	0.00	47.60	OK
EXCH_CAMG	0.0	1500.0	0.0	150.0	3.00	91.00	OK
EXCH_K	0.0	100.0	0.0	50.0	0.00	5.00	OK
EXCH_NA	0.0	500.0	0.0	100.0	0.00	65.80	OK
CEC	0.0	1000.0	0.0	200.0	0.20	179.60	OK
BS	0.0	100.0	0.0	100.0	0.20	100.00	OK
WC_1	0.0	100.0	0.0	95.0	0.00	183.00	Error
WC_2	0.0	100.0	0.0	95.0	0.50	140.00	Error
WC_3	0.0	100.0	0.0	95.0	0.50	112.00	Error
WC_4	0.0	100.0	0.0	95.0	0.00	94.30	OK
WC_FC	0.0	100.0	0.0	95.0	0.00	92.90	OK
POR_TOT	0.0	100.0	10.0	95.0	19.00	94.00	OK
BD	0.0	3.0	0.05	2.3	0.08	2.15	OK

Annexe 4 : Attributs de l'European Soil Database Derived Datas

Soil Property	Topsoil (Filename)	Subsoil (Filename)	Unit
Area of STU allocation	STU_EU_ALLOCATE		unitless
Depth available to roots	STU_EU_DEPTH_ROOTS		cm
Clay content	STU_EU_T_CLAY	STU_EU_S_CLAY	%
Sand content	STU_EU_T_SAND	STU_EU_S_SAND	%
Silt content	STU_EU_T_SILT	STU_EU_S_SILT	%
Organic carbon content	STU_EU_T_OC	STU_EU_S_OC	%
Bulk density	STU_EU_T_BD	STU_EU_S_BD	g cm <sup>-3</sup>
Coarse Fragments	STU_EU_T_GRAVEL	STU_EU_S_GRAVEL	%
Total available water content from PTR	SMU_EU_T_TAWC	SMU_EU_S_TAWC	mm
Total available water content from PTF	STU_EU_T_TAWC	STU_EU_S_TAWC	mm

Annexe 5 : Attributs des données Lucas ponctuelles

Field	Description
Depth	Based on sample collected (e.g. 0-20cm, 0-10cm, 20-30cm)
POINTID	LUCAS Point Identifier – link to Eurostat LUCAS Microdata
pH_CaCl2	pH – measured in calcium chloride
pH_H2O	pH – measured in water
EC	Electrical conductivity (milliSiemens per meter – mS m <sup>-1</sup> )
OC	Organic carbon content (g kg <sup>-1</sup> )
CaCO3	Calcium carbonate content (g kg <sup>-1</sup> )
P	Total phosphorus (g kg <sup>-1</sup> )
N	Total nitrogen (g kg <sup>-1</sup> )
K	Extractable potassium (g kg <sup>-1</sup> )
OC (20-30 cm)	Organic carbon content (g kg <sup>-1</sup> ) from different depth
CaCO3 (20-30 cm)	Calcium carbonate content (g kg <sup>-1</sup> ) from different depth
Ox_Al	Oxalate extractable Al (mg kg <sup>-1</sup> )
Ox_Fe	Oxalate extractable Fe (mg kg <sup>-1</sup> )

Field	Description
POINTID	LUCAS Point Identifier – link to Eurostat LUCAS Microdata
BD 0-10	Measured Bulk Density for the depth 0-10 cm (g cm <sup>-3</sup> )
BD 10-20	Measured Bulk Density for the depth 10-20 cm (g cm <sup>-3</sup> )
BD 20-30	Measured Bulk Density for the depth 20-30 cm (g cm <sup>-3</sup> ) – only Portugal
BD 0-20	Bulk Density for the depth 0-20 cm (g cm <sup>-3</sup> ) – arithmetic mean based on the values for 0-10 cm and 10-20 cm

## Annexe 6 : Attributs téléchargeables de l'ISRIC - SoilGrids

Name	Description	Mapped units	Conversion factor	Conventional units
bdod	Bulk density of the fine earth fraction	cg/cm <sup>3</sup>	100	kg/dm <sup>3</sup>
cec	Cation Exchange Capacity of the soil	mmol(c)/kg	10	cmol(c)/kg
cfvo	Volumetric fraction of coarse fragments (> 2 mm)	cm <sup>3</sup> /dm <sup>3</sup> (vol%)	10	cm <sup>3</sup> /100cm <sup>3</sup> (vol%)
clay	Proportion of clay particles (< 0.002 mm) in the fine earth fraction	g/kg	10	g/100g (%)
nitrogen	Total nitrogen (N)	cg/kg	100	g/kg
phh2o	Soil pH	pHx10	10	pH
sand	Proportion of sand particles (> 0.05 mm) in the fine earth fraction	g/kg	10	g/100g (%)
silt	Proportion of silt particles (≥ 0.002 mm and ≤ 0.05 mm) in the fine earth fraction	g/kg	10	g/100g (%)
soc	Soil organic carbon content in the fine earth fraction	dg/kg	10	g/kg
ocd	Organic carbon density	hg/m <sup>3</sup>	10	kg/m <sup>3</sup>
ocs	Organic carbon stocks	t/ha	10	kg/m <sup>2</sup>

Annexe 7 : Formules raster utilisées pour le travail de qualification des PFTs

$\theta_{1500\_Lucas} : 0.145 + (0.00256 * "Clay@1") - (0.0856 * "Bulk\_density@1") + (0.007 * "pH\_H_2O@1") + (0.000624 * "Silt1@1")$

$\theta_{330\_Lucas} : 0.386 + (0.00254 * "Clay@1") - (0.0927 * "Bulk\_density@1") - (0.00271 * "CEC@1") + (0.000172 * "Phosph@1")$

$\theta_{100\_Lucas} : -9.809 + (0.104 * "Clay@1") - (0.124 * "Bulk\_density@1") + (0.103 * "Silt1@1") + (0.0237 * "N@1") + (0.102 * "Sand1@1")$

$\Theta_{CC} =$

$0.00245 * (("clay\_0-5cm\_mean@1"/10) + (3 * "clay\_5-15cm\_mean@1"/10)) /4 - 0.00135 * (("sand\_0-5cm\_mean@1"/10) + (3 * "sand\_5-15cm\_mean@1"/10)) /4 + 0.278$

$\Theta_{pfp} =$

$0.00401 * (("clay\_0-5cm\_mean@1"/10) + (3 * "clay\_5-15cm\_mean@1"/10)) /4 - 0.000293 * (("sand\_0-5cm\_mean@1"/10) + (3 * "sand\_5-15cm\_mean@1"/10)) /4 + 0,08$

HCCF\_SH :

$(0.5 * "argile\_moy\_0-15@1") + (0.218 * "limon\_moy\_0-15@1") + 7.35$

HPF :

$\text{Max}(0.40 * "argile\_moy\_0-15@1" * Bulk\_density@1 ; 0.005)$

AWC\_033-1500\_SH

$"\theta_{330\_Lucas\_SH@1}" - "\theta_{1500\_Lucas\_SH@1}"$

AWC\_033-1500\_SH\_EG :

$"AWC\_033-1500\_SH@1" * (100 - "Cailloux\_0\_15\_mean\_mean@1") / 100 * 20 / 10$

Dans une [étude de Dobarco](#), l'auteur trouvait les données de validations croisées pour des FPT à différents nombre de paramètres suivants :

*Annexe 8 : Tableau de qualification des FPTs en fonctions des paramètres, Dobarco, 2019*

**Table 7. Cross-validation of continuous-PTFs**

Horizon type	Model	pF = 2.0		pF = 4.2	
		RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>
<b>Non-stratified</b>	Clay + Sand	0.043	0.69	0.034	0.82
	Clay + Sand + soil organic C	0.042	0.69	<b>0.034</b>	<b>0.83</b>
	Clay + Sand + Bulk density	<b>0.041</b>	<b>0.70</b>	0.034	0.82
	Clay + Sand + Bulk density + soil organic C	0.042	0.69	0.034	0.83
<b>Topsoil</b>	Clay + Sand	0.039	0.73	0.027	0.84
	Clay + Sand + soil organic C	<b>0.037</b>	<b>0.77</b>	<b>0.027</b>	<b>0.85</b>
	Clay + Sand + Bulk density	0.039	0.73	0.028	0.83
	Clay + Sand + Bulk density + soil organic C	0.037	0.77	0.027	0.85
<b>Subsoil</b>	Clay + Sand	0.043	0.68	<b>0.036</b>	<b>0.80</b>
	Clay + Sand + soil organic C	0.043	0.66	0.036	0.80
	Clay + Sand + Bulk density	<b>0.041</b>	<b>0.70</b>	0.036	0.80
	Clay + Sand + Bulk density + soil organic C	0.042	0.69	0.036	0.80



					2019	JRC>ESDAC >Chemicals Propreties_L UCAS	<a href="https://esdac.jrc.ec.europa.eu/content/chemical-properties-european-scale-based-lucas-topsoil-data">https://esdac.jrc.ec.europa.eu/content/chemical-properties-european-scale-based-lucas-topsoil-data</a>	pH in H2O 20,000 soil samples were collected in 24 EU countries The chemical properties maps for the European Union were produced using Gaussian process regression (GPR) models.	500m	<a href="https://www.science-direct.com/science/article/pii/S0016706119304768">https://www.science-direct.com/science/article/pii/S0016706119304768</a>	R2=0.65 RMSE = 0,78	
coneseuil	minimum concentration of NO3-N in soil (unavailable for leaching and for uptake)	real	kg.ha-1	[0, 0.5]								
albedo	albedo of the bare dry soil	rea	SD	[0.05, 0.6]								
q0	cumulative soil evaporation above which evaporation decreased	real	mm.d-1	[0, 50]								
ruisolnu	fraction of runoff (relative to total rainfall) in a bare soil	real	SD	[0, 1]								
obstarac	soil depth at which root growth is stopped due ti physical constraints	real	cm	[10, 1000]								
pluibat	minimal amount of rain required to create a soil crust (a value must be given but if the plt.xml vigueurbat parameter is equal to 1 then the parameter is inactive)	real	mm.d-1	[5, 100]								
malchbat	muchl depth at which a crust occurs (a value must be given but if in the plt.xml the vigueurbat parameter is equal to 1 then the parameter is inactive)	real	cm	[0, 2]								
zesx	maximal soil depth affected by soil evaporation	real	cm	[10, 150]								
cfes	parameter defining the soil contribution to evaporation versus depth	real	SD	[0.5, 4]								
z0solnu	roughness lenght of bare soil	real	m	[0.01, 0.2]								
CsurNso	initial C to N ratio of soil humus	real	g.g-1	[8, 20]	2019	JRC>ESDAC >Chemicals Propreties_ LUCAS	<a href="https://esdac.jrc.ec.europa.eu/content/chemical-properties-european-scale-based-lucas-topsoil-data">https://esdac.jrc.ec.europa.eu/content/chemical-properties-european-scale-based-lucas-topsoil-data</a>	C : N 20,000 soil samples were collected in 24 EU countries The chemical properties maps for the European Union were produced using Gaussian process regression (GPR) models.	500m	<a href="https://www.science-direct.com/science/article/pii/S0016706119304768">https://www.science-direct.com/science/article/pii/S0016706119304768</a>	R2=0.91 RMSE = 1.97	
finert	initial fraction of inert pool in the soil organic pool (stable SON/ total SON)	real	SD	[0,1]								
penterui	runoff coefficient taking into account the plant mulch	real	SD	[0, 5]								

Quanti (+1 class) x 5 layers	epc	thickness of each soil layer	int	cm	[1, 1000]	JRC>ESDAC >ESDB>PTR DB	<a href="https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes">https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes</a>	DIFF Soil profile differentiation.	1km	<a href="https://www.researchgate.net/profile/Juan-Ibanez-5/publication/254559741_The_pedodiversity_concept_and_its_measurement_Application_to_soil_information_systems/links/00b7d51feac40bde1300000/The-pedodiversity-concept-and-its-measurement-Application-to-soil-information-systems.pdf#page=73">https://www.researchgate.net/profile/Juan-Ibanez-5/publication/254559741_The_pedodiversity_concept_and_its_measurement_Application_to_soil_information_systems/links/00b7d51feac40bde1300000/The-pedodiversity-concept-and-its-measurement-Application-to-soil-information-systems.pdf#page=73</a>	None		
						JRC>ESDAC >ESDB>SG DBE	<a href="https://esdac.jrc.ec.europa.eu/content/sgdbe-attributes#TEXT_DEP_CHG">https://esdac.jrc.ec.europa.eu/content/sgdbe-attributes#TEXT_DEP_CHG</a>	TEXT-DEP-CHG Depth class to a textural change of the dominant and/or secondary surface texture of the STU.	1km				
						2013	JRC>ESDAC>ESDDD	<a href="https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data">https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data</a>		1000m	Compilation de données (ESDB - SOTWIS - HWSD), Règles depédo-transfert et interpolation spatiale (et verticale) selon les attributs	<a href="https://esdac.jrc.ec.europa.eu/ESDB_Archive/eusoils_docs/ot/her/EUR26082EN.pdf">https://esdac.jrc.ec.europa.eu/ESDB_Archive/eusoils_docs/ot/her/EUR26082EN.pdf</a>	None
	HCCF	gravimetric water content at field capacity of each soil layer (10 kPa ou 33 kPa)	real	%	[10, 110]								
						2009	JRC>ESDAC >LUCAS	<a href="https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data">https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data</a>	Available Water Capacity (AWC) for the topsoil fine earth fraction. Soil point data from the LUCAS 2009 soil survey (around 20,000 points) for EU-25. Topsoil (0-20cm) modelled by Multivariate Additive Regression Splines	500m	<a href="https://www.sciencedirect.com/science/article/pii/S0016706115300173">https://www.sciencedirect.com/science/article/pii/S0016706115300173</a>	R <sup>2</sup> =0,5 RMSE= 10%	
						1995	JRC>ESDAC >ESDB>PTR DB	<a href="https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes">https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes</a>	AWC_SUB Subsoil available water capacity. BS_SUB Base saturation of the subsoil. EAWC_SUB Subsoil easily available water capacity	1km	<a href="https://www.researchgate.net/profile/Juan-Ibanez-5/publication/254559741_The_pedodiversity_concept_and_its_measurement_Application_to_soil_information_systems/links/00b7d51feac40bde1300000/The-pedodiversity-concept-and-its-measurement-Application-to-soil-information-systems.pdf#page=73">https://www.researchgate.net/profile/Juan-Ibanez-5/publication/254559741_The_pedodiversity_concept_and_its_measurement_Application_to_soil_information_systems/links/00b7d51feac40bde1300000/The-pedodiversity-concept-and-its-measurement-Application-to-soil-information-systems.pdf#page=73</a>	None	There is no information about soil horizons in the geographical database version 2.0. The soil is arbitrarily divided into two layers : topsoil and subsoil, roughly corresponding to A and B horizons. The depth of differentiation is 50 cm when it is reached.  These are based on expert judgement, mainly qualitative, and assume that a due weight is given to the confidence level of individual inferred attributes. The results provided by the application of these rules are therefore only qualitative estimates. At the 1:1,000,000 scale it is difficult to provide accurate information from the few data contained in the Soil Database. The application of the pedotransfer rules from the PTRDB to the SGDBE results in a table that contains for each STU a set of values for each STU, attaching a confidence level to each output value, which can highlight those areas for which the results are not so reliable
						2013	JRC>ESDAC >ESDDD	<a href="https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data">https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data</a>	Compilation de données (ESDB - SOTWIS - HWSD), Règles depédo-transfert et interpolation spatiale (et verticale) selon les attributs	1000m	<a href="https://esdac.jrc.ec.europa.eu/ESDB_Archive/eusoils_docs/ot/her/EUR26082EN.pdf">https://esdac.jrc.ec.europa.eu/ESDB_Archive/eusoils_docs/ot/her/EUR26082EN.pdf</a>	None	Travail comparable à celui que l'on mène en ce qu'ils rapportent et comparent plusieurs sources de données entre elles;

HMNF	gravimetric water content at wilting point of each soil layer (1500kPa)	real	%	[2, 25]					1km			
DAF	bulk density of the fine earth fraction in each soil layer (without pebbles)	real	g.cm-3	[0.8, 2]	2016	JRC>ESDAC >LUCAS	<a href="https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data">https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data</a>	Soil point data from the LUCAS 2009 soil survey ( 6,000 measured point). Derived from soil texture datasets (obtained from the packing density and themapped clay content following the equation of Jones et al. 2003).	500m	<a href="https://www.sciencedirect.com/science/article/pii/S0016706115300173">https://www.sciencedirect.com/science/article/pii/S0016706115300173</a>	R <sup>2</sup> =0,5 RMSE= 10%	In <a href="https://www.sciencedirect.com/science/article/pii/S0167880924000252">https://www.sciencedirect.com/science/article/pii/S0167880924000252</a> "At each chosen LUCAS location, before collecting the soil cores, stones larger than 6 cm, vegetation residues, grass, and litter were removed from the soil surface. The surveyors were advised to select the locations without stones and sample the fine soil in the ring" "As a major drawback for this method, we refer to the small volumes collected (which may not be spatially representative of the location) and to the importance of coarse fragments"
					2024	JRC>ESDAC >LUCAS_B D	<a href="https://esdac.jrc.ec.europa.eu/content/soil-bulk-density-europe">https://esdac.jrc.ec.europa.eu/content/soil-bulk-density-europe</a>	bulk density for Europe in 0-20cm, 0-10cm, 10-20cm layers generated with advanced Cubist rule-based regression model.	500m	<a href="https://publications.jrc.ec.europa.eu/repository/handle/JRC135156">https://publications.jrc.ec.europa.eu/repository/handle/JRC135156</a>	R2 = 0.66	
cailloux	volumetric content of pebbles per soil layer	real	%	[0, 100]								
typecailloux	Pebbles type : 1= Beauce limestone, 2 = Beauce limestone, 3= Lutetian limestone, 4= Lutetian Brackish marl and limestone, 5=morainic gravels, 6=unweathered flint, sandstone or granite, 7=weathered granite, 8=Jurassic limestone, 9= Pebbles from Magneraud	int	SD	[1,10]								
infil	infiltrability rate at the base of each soil layer (if codemacropor = 1)	real	mm.d-1	[0.1, 100]								
epd	thickness of mixing cells in each soil layer (= 2 * dispersion lenght)	int	cm	[1, 50]								

Bool	pebbles	option to take into account pebbles in the water and N balances	1995	JRC>ESDAC >ESDB>PTR DB	<a href="https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes">https://esdac.jrc.ec.europa.eu/content/ptrdb-attributes</a>	VS = volume of stones (%) [0/10/15/20%]	1km
			2016	JRC>ESDAC >Physical_ propreties_ LUCAS	<a href="https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data">https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data</a>	Coarse fragments %, for the 0-20cm topsoil Soil point data from the LUCAS 2009 soil survey (around 20,000 points) for EU-25. Topsoil (0-20cm) modelled by Multivariate Additive Regression Splines	500m
			2013	JRC>ESDAC >ESDDD	<a href="https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data">https://esdac.jrc.ec.europa.eu/content/european-soil-database-derived-data</a>	Compilation de données (ESDB - SOTWIS - HWSD), Règles depédo-transfert et interpolation spatiale (et verticale) selon les attributs	1000
	macroporosity	option to activate calculation of water flux in soil macroporosity					
	cracks	option to activate an additional water compartment for swelling soils					
	artificial drainage	option to simulate artificial drainage					
	capillary rise	option to activate capillary rise					
	nitrification	option to activate the nitrification model					
denitrification	option to activate the denitrification model						