



HAL
open science

Plan de gestion de données de l'Observatoire de Recherche sur les Produits Résiduaire Organiques (SOERE PRO)

Christophe Montagnier, Florent Levavasseur, Pierre Benoit, Charles Detaille,
Frédéric Feder, Sabine Houot, Manon Gilles, Thierry Morvan, Camille
Resseguier, Valérie Sappin-Didier, et al.

► **To cite this version:**

Christophe Montagnier, Florent Levavasseur, Pierre Benoit, Charles Detaille, Frédéric Feder, et al..
Plan de gestion de données de l'Observatoire de Recherche sur les Produits Résiduaire Organiques
(SOERE PRO). INRAE. 2025. hal-05230261

HAL Id: hal-05230261

<https://hal.inrae.fr/hal-05230261v1>

Submitted on 1 Sep 2025

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.



Open licence - etalab

PGD du projet "Observatoire de Recherche sur les Produits Résiduaire Organiques"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "INRAE - Trame Structure" fourni par INRAE - Institut national de recherche pour l'agriculture l'alimentation et l'environnement.

Renseignements sur le plan

Titre du plan	PGD du projet "Observatoire de Recherche sur les Produits Résiduaire Organiques"
Livrable	3 sites/8
Version	Version initiale
Objet/périmètre du plan	<p>Le périmètre du PGD se limite aux données produites sur les sites expérimentaux du SOERE PRO : QualiAgro (Feucherolles - INRAE, France), EFELE (Rennes - INRAE, France), PROspective (Colmar - INRAE, France), La Bouzule (INRAE, France), Couhins (INRAE, France), MetaMetha (Nouzilly - INRAE, France), La Mare (La Réunion - CIRAD, France), Sangalkam (CIRAD, Sénégal).</p> <p>Le PGD dans sa version actuelle a été rédigé pour les 3 sites de l'Infrastructure Scientifique Collective (ISC) INRAE (QualiAgro, PROspective et EFELE).</p> <p>Le SOERE PRO compte actuellement 8 sites. C'est pourquoi nous avons structuré le PGD en identifiant chaque site à un produit de recherche (monitoring des données du site). Il y a donc 8 produits de recherche dans le document incluant les 3 sites de l'ISC. Les produits de recherche présentent des parties communes et spécifiques à chaque site.</p> <p>Nous avons donc renseigné en priorité les produits de recherche pour les 3 sites composant l'ISC INRAE pour la campagne de re-labellisation 2025. L'objectif est que le PGD serve de support de manière plus globale pour l'ensemble des sites du SOERE PRO et ce point est actuellement en cours (5 sites restants).</p>

Domaines de recherche (selon classification de l'OCDE)

Earth and related environmental sciences, Agriculture, forestry, and fisheries

Langue

fra

Date de création

2021-07-13

Date de dernière modification

2025-08-29

Identifiant

hal-05230261

Type d'identifiant

HAL Id

Licence

Nom	Etalab Open License 2.0
URL	http://spdx.org/licenses/etalab-2.0.json

Documents (publications, rapports, brevets, plan expérimental....), sites web associés

- Site web : <https://valor-pro.hub.inrae.fr/presentation-de-l-observatoire-soere-pro>
- Espace Nextcloud (accès aux personnes autorisées) : https://nextcloud.inrae.fr/apps/user_saml/saml/selectUserBackEnd?redirectUrl=

Renseignements sur le projet

Titre du projet	Observatoire de Recherche sur les Produits Résiduaire Organiques
Acronyme	PGD_SOERE_PRO
Résumé	L'observatoire « SOERE PRO - réseau d'expérimentations au champ de longue durée pour la recherche sur le recyclage des Produits Résiduaire Organiques en agriculture - a été mis en place en 2008 en réponse aux enjeux sociétaux de retour au sol de PRO. Il a été labellisé Infrastructure Scientifique Collective (ISC) INRAE en 2018 (3 sites). Cet observatoire est un réseau de 8 sites expérimentaux au champ de longue durée permettant d'évaluer les effets d'apports répétés de PRO en agriculture et d'étudier leurs interactions, dans différents contextes agro-pédo-climatiques. Les sites viennent en appui aux politiques publiques (ex. étude GES, étude 4 pour 1000, Esco Mafor) et permettent d'alimenter des outils portés par INRAE, tels que Syst'N, AMG, CERES ou MEANS. Quatre sites fortement instrumentés sont intégrés à l'Infrastructure de Recherche (IR) AnaEE-France : QualiAgro (INRAE EcoSys Palaiseau 91), Prospective (INRAE UEAV Colmar 68), EFELE (INRAE SAS Rennes 35) et SOERE PRO - Réunion (CIRAD Recyclage & Risque Montpellier 34).

Sources de financement

- ANAEE France :
- INRAE :
- Partenaires privés et publics des sites : VEOLIA, SITEUCE, SM4, AERM, SUEZ Environnement, COVED, SEDE, AgriValor, RUNEO, CINOR :
- ADEME :
- Agence Nationale de la Recherche :
- European Commission :

Date de début 2008-01-01

Date de fin

Partenaires

- Veolia (France) <https://ror.org/03mwyp50>
- Syndicat Mixte Recyclage Agricole du Haut Rhin <https://www.smra68.net/>
- Université de Haute-Alsace <https://ror.org/04k8k6n84>
- Écologie Fonctionnelle et Écotoxicologie des Agroécosystèmes <https://ror.org/01z6yh944>
- Unité d'Expérimentation Agronomique et Viticole Colmar <https://ueav.colmar.hub.inrae.fr/>
- GEOSCIENCES UNITE CNRS RENNES <https://geosciences.univ-rennes.fr/>
- Agroécologie <https://ror.org/00mkad321>
- Optimisation des procédés en agriculture, agroalimentaire et environnement <https://opaale.rennes.hub.inrae.fr/>
- Soil Agro and Hydrosystems Spatialization <https://ror.org/03k4s1p46>
- Interactions Sol Plante Atmosphère <https://ror.org/02aswf736>
- Recyclage et risque <https://ror.org/03ze73t75>
- Unité Expérimentale Physiologie Animale de l'Orfrasière <https://uepao.val-de-loire.hub.inrae.fr/>
- Laboratoire Sols et Environnement <https://ror.org/01e8etd40>
- ECOBIO ECOSYSTEMES, BIODIVERSITE, EVOLUTION UNITE CNRS RENNES <https://ecobio.univ-rennes.fr/>

Produits de recherche :

1. Données d'observation et d'analyses dans les compartiments sol, plante, eau, air, PRO - site QualiAgro Feucherolles (78) (Jeu de données)
2. Données d'observation et d'analyses dans les compartiments sol, plante, eau, PRO - site Prospective Colmar (68) (Jeu de données)

3. Données d'observation et d'analyses dans les compartiments sol, plante, eau, air, PRO - site EFELE Le Rheu (35) (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
BENOIT Pierre - https://orcid.org/0000-0001-8681-7078	Écologie Fonctionnelle et Écotoxicologie des Agroécosystèmes - https://ror.org/01z6yh944	
DETAILLE Charles	Recyclage et risque - https://ror.org/03ze73t75	
FEDER Frédéric - https://orcid.org/0000-0001-8434-5193	Recyclage et risque - https://ror.org/03ze73t75	
GILLES Manon	Écologie Fonctionnelle et Écotoxicologie des Agroécosystèmes - https://ror.org/01z6yh944	<ul style="list-style-type: none"> • Personne contact pour les données (Monitoring_COLMAR)
HOUOT Sabine - https://orcid.org/0000-0001-7297-7038	Écologie Fonctionnelle et Écotoxicologie des Agroécosystèmes - https://ror.org/01z6yh944	
LEVAVASSEUR Florent - https://orcid.org/0000-0002-2164-3334	Écologie Fonctionnelle et Écotoxicologie des Agroécosystèmes - https://ror.org/01z6yh944	<ul style="list-style-type: none"> • Responsable du plan
MONTAGNIER Christophe - https://orcid.org/0009-0000-2306-6764	Écologie Fonctionnelle et Écotoxicologie des Agroécosystèmes - https://ror.org/01z6yh944	<ul style="list-style-type: none"> • Coordinateur du projet
MORVAN Thierry - https://orcid.org/0000-0002-6678-2661	Soil Agro and Hydrosystems Spatialization - https://ror.org/03k4s1p46	<ul style="list-style-type: none"> • Personne contact pour les données (Monitoring_EFELE)
RESSEGUIER Camille - https://orcid.org/0009-0004-5325-3164	Écologie Fonctionnelle et Écotoxicologie des Agroécosystèmes - https://ror.org/01z6yh944	<ul style="list-style-type: none"> • Personne contact pour les données (Monitoring_QUALIAGRO)
SAPPIN-DIDIER Valérie - https://orcid.org/0000-0003-3841-9037	Interactions Sol Plante Atmosphère - https://ror.org/02aswf736	
SAVOIE Antoine - https://orcid.org/0000-0002-1606-4020	Unité Expérimentale Physiologie Animale de l'Orfrasière - https://uepao.val-de-loire.hub.inrae.fr/	
VERMEIRE Marie-Liesse - https://orcid.org/0000-0003-2637-0723	Recyclage et risque - https://ror.org/03ze73t75	
WATTEAU Françoise - https://orcid.org/0000-0002-2448-3145	Laboratoire Sols et Environnement - https://ror.org/01e8etd40	

PGD du projet "Observatoire de Recherche sur les Produits Résiduaire Organiques"

Données d'observation et d'analyses dans les compartiments sol, plante, eau, air, PRO - site QualiAgro Feucherolles (78)

Informations sur la structure

Nom de la structure

SOERE PRO : Observatoire de Recherche en Environnement sur les Produits Résiduaire Organiques

Type de structure

- SOERE
- ISC (Infrastructure Scientifique Collective)

SOERE : Systèmes d'Observation et d'Expérimentation au long terme pour la Recherche en Environnement

Site web : <https://valor-pro.hub.inrae.fr/>

Données générales pour la coordination technique dans espace Nexcloud INRAE (accès personnes autorisées)

: https://nextcloud.inrae.fr/apps/user_saml/saml/selectUserBackEnd?redirectUrl=

Identifiant de la structure

Préciser le fournisseur de l'identifiant (ISNI, VIAF, FundRef, DataCite...).

Question sans réponse.

Responsabilités dans la structure

Nom, Prénom	Courriel	Rôle
HOUOT Sabine	sabine.houot@inrae.fr	Responsable coordination scientifique
LEVAVASSEUR Florent	florent.levasseur@inrae.fr	Responsable coordination scientifique
BENOIT Pierre	pierre.benoit@inrae.fr	Responsable coordination scientifique
MONTAGNIER Christophe	christophe.montagnier@inrae.fr	Responsable coordination et animation technique

Etablissement(s) tutelle(s)

- INRAE
- CIRAD

Département de rattachement INRAE (ou anciens départements Inra)

- AGROECOSYSTEM : Agroécosystèmes
- Département de rattachement du CIRAD : PERFORMANCE des SYSTèmes de production et de transformation tropicaux (PERSYST)

Financier(s) (permettant l'acquisition des jeux de données - hors projet)

N/A

Informations sur le plan de gestion

DOI (version publiée du plan de gestion)

Historique des versions

Date	n° de version	Status	Auteur	Affiliation de l'auteur (se reporter à l' annuaire INRAE)	Validé par	Validé le
24/07/2025	V0	Public	MONTAGNIER Christophe	INRAE UMR 1402 ECOSYS Ecologie Fonctionnelle et Ecotoxicologie des agroécosystèmes	Florent LEVAVASSEUR	24/07/2025

Présentation générale des données

Mode d'obtention des données

- Données générées par la structure

L'intégralité des données collectées sont retranscrites dans des fichiers informatisés sous Excel (tableur) dans le format .xlsx ou .csv et sous Word (traitement de texte) dans le format .docx.

Les données générées par la structure :

- données générales (localisation du dispositif, contexte agro-pédo-climatique),
- données du protocole expérimental (plan du dispositif expérimental, modalités de PRO étudiées, équipements du site, suivis réalisés, banques d'échantillons),
- données d'itinéraires techniques culturaux,
- données issues d'analyses dans la matrice SOL (profondeurs 0-28/28-35/35-50/50-90 cm),
- données issues d'analyses dans la matrice PRO,
- données issues d'analyses dans la matrice PLANTES (grain, résidus),
- données issues d'analyses dans la matrice EAUX (lysimètres et pluie),
- données issues d'analyses des reliquats azotés dans les sols (profondeurs 0-30/30-60/60-90 cm),
- données de mesure de la densité apparente du SOL (profondeurs 0-28/28-35/35-50/50-90 cm)
- données de mesure de rendement (grain, résidus),
- données de mesure du statut N des cultures (acquisition avec DUALEX Optical Leafclip Meter), de la concurrence des adventices et du développement des maladies,
- données lysimétriques (mesure volume d'eau drainée à 45 cm et 100 cm),

- données météorologiques (acquisition avec logiciel Hydras / station météo du dispositif) : température, hygrométrie, pluviométrie, rayonnement solaire, direction du vent, vitesse du vent
- données issues de mesures instrumentées (capteurs SOL) : teneur en eau (humidité), tension de l'eau, température entre 0 et 160cm
- données issues de mesures instrumentées (capteurs AIR / GES) : N2O, CO2, NH3

Origine

- Analyse
- Expérimentation
- Observation

Observation = mesure au champ

Voir les exemples de protocoles d'acquisition des données dans la rubrique : **Organisation et documentation des données.**

Type de données

- Dataset
- Text
- Other (à préciser dans la zone "Informations supplémentaires")

Tabulaire (dans tableur type Excel)

Dataset : format tableur type Excel

Text : format traitement de texte type WORD

Nature des données

2 types de données sont générées/gérées pour la conduite du dispositif QualiAgro : données textuelles (texte sous word, présentation powerpoint), données tabulaires (dans un tableur excel).

- Données textuelles de description du dispositif :
 1. Localisation du dispositif, contexte agro-pédo-climatique
 2. Plan du dispositif expérimental, modalités de PRO étudiées, équipements du site, suivis réalisés, banques d'échantillons
 3. Protocoles de réalisation des mesures de terrain (planning)
 4. Protocoles de prélèvement et de préparation des échantillons pour analyses
 5. Protocoles de mesure et d'analyse des échantillons
- Données expérimentales sous forme tabulaire :

1. Compartiment SOL

Type de données : date, profondeur, modalité (traitement)

Type de données : Analyse physico-chimique et des contaminants

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, soufre, CEC, éléments échangeables, conductivité, propriétés physiques, humidité, matière sèche, éléments traces, autres éléments minéraux, granulométrie, pH, calcaire, phosphore, potassium, composés pharmaceutiques, HAP, PCB, dioxines, furanes

Type de données : analyses biologiques

Variable analysée : activité enzymatiques (C, N, S, P), biomasse microbienne

Type de données : analyses des reliquats azotés

Type de données : mesure de la densité apparente

Type de données : Issues de mesures instrumentées (capteurs SOL) : teneur en eau (humidité), tension de l'eau, température entre 0 et 160 cm

2. Compartiment PRO

Type de données : date, type de PRO

Type de données : Analyse physico-chimique et des contaminants

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, soufre, CEC, éléments échangeables, éléments traces, autres éléments minéraux, fractionnement biochimique (ISMO), humidité, matière sèche, pH, calcaire, phosphore, potassium, composés pharmaceutiques, HAP, PCB, dioxines, furanes

Type de données : Minéralisation carbone et azote

3. Compartiment PLANTES

Type de données : date, culture, modalité (traitement)

Type de données : Analyse physico-chimique et des contaminants (grains et résidus)

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, soufre, éléments traces, autres éléments minéraux, humidité, matière sèche, phosphore, potassium, dioxines, furanes

Type de données : Mesure du rendement (grains et résidus)

Type de données : Mesure du statut N des cultures (DUALEX Optical Leafclip Meter), concurrence des adventices et développement des maladies

4. Compartiment EAUX

Type de données : date, profondeur, volume, modalité (traitement)

Type de données : Analyse physico-chimique et des contaminants (lysimétrie : solution du sol)

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, soufre, éléments traces, autres éléments minéraux, pH, calcaire, phosphore, potassium, composés pharmaceutiques

Type de données : Mesure du volume d'eau drainée à 45 cm et 100 cm (lysimétrie : solution du sol)

Type de données : Analyse physico-chimique et des contaminants (eau de pluie)

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, soufre, éléments traces, autres éléments minéraux, pH, calcaire, phosphore, potassium

5. Compartiment AIR

Type de données : date, modalité (traitement)

Type de données : mesures météorologiques

Variables mesurées : température, pluviométrie, hygrométrie, direction du vent, vitesse du vent, rayonnement solaire

Type de données : Issues de mesures instrumentées (capteurs AIR / GES)

Variables mesurées : émission N2O, émission CO2, émission NH3

6. Itinéraires techniques cultureux

Modalité (traitement),

Dates (début, fin, enfouissement PRO),

Type d'intervention (type de culture, travail du sol, semis, récolte, PRO étudié (épandage), Intervention apport (fertilisation, traitement phytosanitaire, nom du produit),

Matériel,

Dose (unité),

Profondeur de travail,

Largeur de travail,

Commentaire

Format des données

- Format texte : .doc; .docx; .txt => données générales et du protocole expérimental
- Format tabulaire : .xls; .xlsx; .csv => tous les autres types de données

Périmètre thématique des données

- Climate
- Farming Systems and Practices
- Plant Breeding and Plant Products
- Soils and soil sciences
- Water resources

Droits de propriété intellectuelle

Les données produites dans les organismes de recherche publics doivent être communicables à toute personne qui en fait la demande si elles n'entrent pas dans le cadre des exceptions légales (sécurité défense, sécurité des populations, patrimoine scientifique et technique, données personnelles, données liées au secret, statistique, etc.), et si elles constituent des documents dits achevés. Selon la nature des données, leur communication peut être obligatoire, interdite, ou soumise à conditions. Pour plus d'information, consulter le guide ["Ouverture des données de recherche. Guide d'analyse du cadre juridique en France"](#) publié par un collectif de juristes issus de l'Enseignement Supérieur et de la Recherche.

Qui détiendra les droits sur les données et les autres informations créées ?

- INRAE 1402 ECOSYS Palaiseau (91)
- VEOLIA (les données sont acquises avec ce partenaire privé et sont donc en copropriété avec lui)

Confidentialité

Identification des jeux de données contenant des données confidentielles

INRAE facilite la diffusion systématique en libre accès des informations issues de ses travaux de recherche financés sur fonds publics, **sous réserve des droits et intérêts des parties prenantes et des tiers, y compris les droits de propriété intellectuelle, la confidentialité, ou le caractère sensible des informations** (sécurité défense, sécurité des populations, patrimoine scientifique et technique, données personnelles, données liées au secret, statistique, etc.) et si elles constituent des **documents dits achevés**.

Dans le cas du **dispositif expérimental QualiAgro**, les données sont acquises avec un partenaire privé et sont donc en copropriété avec ce partenaire. Les données sont confidentielles tant qu'elles n'ont pas été publiées après validation des co-proprétaires. Leur publication permet leur dépôt dans des entrepôts de données ou des bases de données accessibles.

Les métadonnées associées à QualiAgro ne sont pas confidentielles.

Au niveau du foncier non bâti, la parcelle où se situe l'essai est en location chez un agriculteur. **Concernant le site QualiAgro, le RGPD ne s'applique pas** même si les données de pratiques et les caractéristiques de sol d'une parcelle d'un agriculteur relèvent bien du RGPD car cela reflète ses choix personnels. En effet, dans le cas précis de QualiAgro, l'agriculteur applique un protocole défini par INRAE, il n'est pas libre de ses choix et de ses pratiques. Ce ne sont donc plus des données personnelles. Il n'y a par conséquent pas besoin de mentionner le RGPD dans les projets de recherche impliquant QualiAgro. **Cet aspect a été validé par la cellule Informatique et Libertés d'INRAE.**

Quelles sont les mesures prises et les normes auxquelles il est nécessaire de se conformer pour garantir cette confidentialité ?

Les données confidentielles sont collectées et stockées sur serveur de l'unité INRAE 1402 ECOSYS avec accès sécurisé (par identifiant/mot de passe, VPN activé et seulement aux personnes déclarées par l'administrateur) et leur diffusion est réglementée et soumise à validation par les copropriétaires.

Le cas échéant, comment la confidentialité de données fournies par des personnes sera garantie lorsque les données seront partagées ou rendues disponibles pour une analyse de second niveau ?

Néant

Partage des données

Y a t'il une obligation de partage (ou à l'inverse une interdiction ou une restriction) ?

L'Unité ECOSYS dépend d'un organisme de recherche public (INRAE), les données qu'elle produit sont des données FAIR et entrent dans le cadre de la Science Ouverte. Elles doivent être Faciles à trouver, Accessibles, Interopérables et Réutilisables si elles n'entrent pas dans le cadre des exceptions légales (données confidentielles). Les données acquises en copropriété avec un partenaires privé :

- sont librement diffusables si elles ont été publiées,
- pour les données non publiées, leur diffusion se fait en accord avec les copropriétaires.

Les conditions de partage et d'utilisation des données issues du site QualiAgro sont décrites dans la charte interne SOERE PRO : AnaEE_Charte_PRO_Version_Octobre_2024.pdf consultable à l'adresse : <https://valor-pro.hub.inrae.fr/les-sites-experimentaux/charte-pour-l-acces-au-service-pro-anaee-france>.

Quelles sont les réutilisations potentielles de ces données ?

- Publications scientifiques
- Vulgarisation scientifique
- Travaux de Recherche et Développement
- Enseignement

La lecture des données nécessite-t-elle le recours à un logiciel ou un outil spécifique ? Si oui, lequel ?

NON. Les logiciels de la suite Office notamment Excel (tableur), Word (traitement texte), Power Point (présentation) suffisent. Des formats ouverts, plus pérennes et interopérables que ceux de la suite Office pourraient également être envisagés.

Comment les données seront-elles partagées ?

- Publication sur le portail Recherche.Data.gouv.fr (<https://entrepot.recherche.data.gouv.fr/dataverse/observatoire-soere-pro>)
- Envoi par E-mail des données et/ou des fichiers qui ne sont pas sur Recherche.Data.Gouv.fr aux partenaires qui le souhaitent et après accord des copropriétaires

Avec qui ?

- Autre

L'ensemble des données produites seront partagées :

- avec tous pour les données en Open Access (Recherche.Data.Gouv.fr),
- aux agents INRAE authentifiés (LDAP),
- aux partenaires identifiés dans les projets de recherche auxquels ils participent.

Sous quelle licence ?

- Licence ouverte <https://www.etalab.gouv.fr/licence-ouverte-open-licence> (compatible CC-BY)
- Pour les données non confidentielles : sous licence ouverte
- Pour les données confidentielles, pas de licence.

Organisation et documentation des données

Quels méthodes et outils sont utilisés pour acquérir et traiter les données, depuis leur acquisition jusqu'à leur mise à disposition, leur archivage ou leur destruction ?

Utiliser éventuellement un lien vers un schéma illustrant les processus

Les méthodes et outils pour acquérir et traiter les données du dispositif QualiAgro sont gérés à l'échelle de l'UMR 1402 ECOSYS et du SOERE PRO.

- **La collecte des données**

Les textes de référence utilisés pour acquérir les données sur le terrain sont définis dans le cadre de la démarche d'assurance qualité

de l'UMR ECOSYS.

Type de données	Protocoles d'acquisition utilisé / documents de référence correspondant
1. Générales :	/
- localisation du dispositif	QA_Presentation_Dispositif.pdf
- contexte agro-pédo-climatique	QA_Presentation_Dispositif.pdf
2. Protocole expérimental :	/
- plan du dispositif expérimental	QA_Presentation_Dispositif.pdf
- modalités de PRO étudiées	QA_Presentation_Dispositif.pdf
- équipements du site	QA_Presentation_Dispositif.pdf
- suivis réalisés	QA_Presentation_Dispositif.pdf
- banques d'échantillons	En projet de rédaction
3. Protocoles de réalisation des mesures de terrain (planning)	En projet de rédaction
4. Protocoles de prélèvement et de préparation des échantillons pour analyses	/
4a. Compartiment SOL	QA_MO_Sol_Plvmt_RelN_v1.docx; F_ProtocoleExp_Sol_av_epandage_2007_001.doc
4b. Compartiment PRO	QA_MO_EngraisOrg_Ferti_V1.docx; QA_MO_PRO_Epandage_V1_Encours.docx
4c. Compartiment PLANTES	F_MO_Plantes_PlvmtBle_V1_Encours.docx; F_ProtocoleExp_CIPAN_2013_001.doc; QA_ModeOp_Prel_Mais_V0.docx; F_MOE_Luzerne_Recolte_2016_002.doc; F_MO_Plantes_PlvmtFeverole_V1_Encours.docx;
4d. Compartiment EAUX	QA_MO_EauxLysi_PrelevAliquot_V2.docx
5. Protocoles de mesure et d'analyse des échantillons	/
5a. Compartiment SOL	/
- analyse physico-chimique et des contaminants	Catalogue_LAS_2024_V0.pdf; catalogue_analytique_TERANA
- analyse de la biologie	ISO 20130:2018 - Qualité du sol — Mesure de l'activité enzymatique dans des échantillons de sol en utilisant des substrats colorimétriques; SOL_BIO_METbiomassemicrobienne_002
- analyse des reliquats azotés	SOL_CHI_METskalarAzote_minB_MO_022_corrMC05_JNR.doc; SOL_CHI_MET_extractionN_MO_022_MC-JN.docx
- mesure de la densité apparente	F_Protocole_Densité_2004.doc
- mesure instrumentées avec capteurs SOL	QA_MO_Instru_SondesSurface_V1_Encours.docx
5b. Compartiment PRO	/
- analyse physico-chimique et des contaminants	SOL_PRO_MO_2015_V2.docx; Catalogue_LAS_2024_V0.pdf; catalogue_analytique_TERANA; catalogue_analytique_LDAR
- minéralisation carbone et azote	Norme AFNOR XP U 44-163 Décembre 2009 : Caractérisation de la matière organique par la minéralisation potentielle du carbone et de l'azote
5c. Compartiment PLANTES	/
- analyse physico-chimique et des contaminants	catalogue_analytique_TERANA; catalogue_analytique_USRAVE
- mesure de rendement	F_MO_Plantes_PlvmtBle_V1_Encours.docx; F_ProtocoleExp_CIPAN_2013_001.doc; QA_ModeOp_Prel_Mais_V0.docx; F_MOE_Luzerne_Recolte_2016_002.doc; F_MO_Plantes_PlvmtFeverole_V1_Encours.docx
- mesure du statut N des cultures, concurrence des adventices, développement des maladies	Protocole_SPAD_2023_ThM_EFELE_adaptéVF; Protocole_notation_adventices_SOERE_PRO_VF; Protocole_notation_maladies_SOERE_PRO_VF
5d. Compartiment EAUX	/
- analyse physico-chimique et des contaminants (lysémie)	Catalogue_LAS_2024_V0.pdf; SOL_CHI_METcarbliq_MO_01.doc; SOL_CHI_METskalarAzote_minB_MO_022_corrMC05_JNR.doc; Dosage_anions_par_ECZ_MO_V2
- mesure volume d'eau drainée (lysémie)	QA_MO_EauxLysi_PrelevAliquot_V2.docx
- analyse physico-chimique et des contaminants (pluie)	SOL_CHI_ModOp_preparation-eaux-pluie_MO_002.doc
5e. Compartiment AIR	/
	Station météo locale sur site QualiAgro (Feucherolles-78); Note technique Meteo France

- mesures météorologiques N°35 (classification d'un site); Note technique Meteo France N°37 (classification de performance maintenue)
- mesure instrumentées avec capteurs AIR / GES QA_MO_Instru_ChambresGES_Instal_V1_Encours.docx; Calcul du flux de CO2 et N2O.doc

6. Itinéraires techniques culturaux :

type d'intervention / nature intervention N/A

- Le tableau suivant récapitule les étapes, méthodes et outils utilisés pour la gestion des données

Etapes	Méthode(s)	Outil(s)
Préparation	<ul style="list-style-type: none"> • Rédaction de protocoles et modes opératoires et mise à disposition des agents impliqués dans le suivi du dispositif QualiAgro • Formation interne aux gestes techniques 	Traitement de texte
Acquisition	<ul style="list-style-type: none"> • procédures, protocoles et modes opératoires validés par le responsable qualité. • Pour les échantillons : collecte, conditionnement, identification et stockage. 	<ul style="list-style-type: none"> • Saisie papier / ordinateur de terrain • Outil de mesure (balance, centrale d'acquisition Campbell Scientific, etc...)
Traitement des données	<ul style="list-style-type: none"> • recueil, mise en forme des données. • analyse de premier niveau à des fins de présentation 	Tableur (Excel, Expé-R)
Mise à disposition	<ul style="list-style-type: none"> • publication sur Recherche data gouv, • agents en charge du suivi de l'essai, • après validation pour tout scientifique qui en fait la demande autre que les équipes en charge du suivi. 	Entrepôt de données / serveur de stockage en réseau (NAS ECOSYS) / SI-PRO / Messagerie électronique (mails)
Archivage	Espace QualiAgro dans le NAS de l'UMR ECOSYS (VGG-NAS418-1) + Espace QualiAgro dans le SI-PRO	Serveur de stockage en réseau (NAS) + Système d'Information (SI)
Réutilisation	Libre quand publié sur recherche data gouv Sur demande après validation pour les données non publiées	N/A
Destruction	N/A	N/A

Quelles métadonnées seront utilisées pour accompagner le jeu de données ? Quels seront les standards, vocabulaires, taxonomies... utilisés pour décrire et représenter les données et éléments de métadonnées ? Comment les métadonnées seront-elles produites et mises à jour ?

Métadonnées	Origine, mode de production des métadonnées (ex : saisie manuelle, annotation automatique...)	Standard, Vocabulaires associés	Conditions ou fréquence de la mise à jour (si applicable) (ex : changement de l'accessibilité)
Données générales	Géoréférencement, Description de fosses pédologiques	Système de coordonnées géographiques (géodésie) Référentiel Pédologique de l'AFES 2008	N/A
Protocoles expérimentaux	Saisie manuelle	Trame démarche qualité INRAE	En cas de changement dans le protocole général
Méthode d'analyses physico-chimiques	Laboratoires d'analyses	Normes en vigueur (méthodes utilisées pour analyser les différents éléments)	N/A
Itinéraire technique	Saisie manuelle, mesure de doses (engrais organique, PRO)	Thesaurus SOERE PRO (SI-PRO)	Tous les ans
Données climatiques	Station météorologique sur site	Standards Meteo France	En continu

Une documentation complémentaire aux métadonnées est-elle nécessaire pour décrire les données et assurer leur réutilisabilité sur le long terme ?

Chaque Dataset (jeux de données) devra être accompagné d'un fichier nommé "Read-me ou Lisez-moi" lors de leur publication sur le site de l'Entrepôt de Recherche : <https://recherche.data.gouv.fr/fr>. Une trame pour générer plus facilement un Read-me est proposée par l'équipe Recherche Data Gouv : <https://recherche.data.gouv.fr/fr/categorie/33/guide/modele-de-readme>. Ce fichier a pour but d'aider à la compréhension et à l'interprétation des données.

Comment les fichiers de données sont-ils gérés et organisés : contrôle des versions, conventions de nommage des fichiers, organisation des fichiers

• Les données sont gérées et organisées dans des dossiers de fichiers au sein du VGG-NAS418-1 (NAS QualiAgro)

L'autorisation d'accès au NAS est donnée par la gestionnaire de l'essai QualiAgro (création d'un mot de passe aléatoire). Ensuite seulement, il est accessible à l'adresse <https://192.168.12.29:5001/> au moyen d'une connexion avec Identifiant et Mdp LDAP INRAE recommandé (VPN AgroParisTech activé).

Ce dossier partagé est géré par la gestionnaire de l'essai QualiAgro. Il est structuré de la manière suivante :

Dossier racine : PRO => SOERE PRO => QualiAgro

Dans le dossier QualiAgro, on trouve 2 sous dossiers qui regroupent l'essentiel des données intéressant ce Plan de gestion : Documentations et Données

Le dossier Documentations comprend notamment les procédures avec les protocoles expérimentaux servant à l'acquisition des données.

Le dossier Données comprend les fichiers de données classés par compartiment (ex : Données_SOL) et par année. Certains compartiments comme le compartiment Données_Plantes peuvent contenir un niveau supplémentaire (avec l'espèce : le blé par ex.).

• Les données sont gérées et organisées également dans le Système d'Information (SI) PRO. Celui-ci est en production et disponible à l'adresse : <https://si-pro.fr/pro-portage-1709-web-SNAPSHOT/index.jsf>

Il contient des données de thématiques Physico-chimie et d'Itinéraires Techniques (ITK).

Pour la thématique physico-chimie, nous avons des types de données Sol (données élémentaires et moyennées), PRO (élémentaires et moyennées), Plantes (élémentaires et moyennées dont le rendement des cultures) et pour les ITK, nous avons plusieurs types d'interventions à savoir interventions apport, pro-étudié, récolte-coupe, semis plantation, travail du sol.

On trouve également dans le SI des données de référence (métadonnées) qui sont soit partagées par tous les types de données (dispositif, parcelles, traitement, etc.), soit spécifiques à chaque type de données (SOL, ITK...).

Enfin on trouve des données de prélèvements et d'échantillons auxquelles sont liées des données produites au laboratoire pour chaque type de données.

Le SI permet de réaliser le dépôt et l'insertion de données en base de données par l'intermédiaire de fichiers csv au format prédéfinis (référentiels et les données).

L'insertion de données de références en base de données se fait sans versionnement préalable à l'inverse des données produites au laboratoire.

Pour le dépôt des fichiers de données, le SI propose une gestion des versions puis la publication qui se traduit par l'insertion en base de la version souhaitée.

Le SI permet également la consultation et l'extraction de données depuis l'interface web de données.

Les données de l'essai QualiAgro sont consultables en téléchargement au format .csv en suivant le chemin d'accès suivant : Accueil / Gestion des données expérimentales / QualiAgro.

Dans QualiAgro, les données s'organisent ainsi :

Propriétés des sols : sols physico-chimie données élémentaires + sols physico-chimie moyennes / période du jeu de données => téléchargement

Propriétés des plantes : plante physico-chimie données élémentaires + plante physico-chimie moyennes / période du jeu de données => téléchargement

Propriétés des PRO : PRO physico-chimie données élémentaires + PRO physico-chimie moyennes / période du jeu de données => téléchargement

Itinéraires techniques : intervention apport + intervention pro étudié + intervention récolte coupe + intervention semis plantation + intervention travail du sol / période du jeu de données => téléchargement

Quel est le processus de contrôle qualité des données ?

La validation des données consiste à s'assurer que :

- elles ne comportent pas d'erreur de saisie, que les unités de mesure soient correctement transcrites, que la date et le lieu de la collecte sont facilement identifiables et si possible que l'appareil utilisé lors de la collecte apparaisse sur le fichier transmis.
- les données transmises par les laboratoires ne présentent pas de valeurs aberrantes. Si c'est le cas, une contre-analyse peut-être demandée au laboratoire.

La méthode utilisée par la gestionnaire du dispositif pour la recherche des valeurs aberrantes est la suivante : la valeur d'une mesure obtenue dans une des répétitions de l'essai est considérée comme aberrante si sa valeur est 2 fois supérieure à la plus grande valeur des répétitions ou deux fois inférieure à la plus petite valeur des répétitions. Si c'est le cas, la valeur est remplacée par la moyenne des valeurs des autres répétitions.

Elle est réalisée par la gestionnaire du dispositif QualiAgro.

Stockage et sécurité des données

Quels sont les types de flux empruntés par les données et les supports utilisés pour les stocker ?
(Faire éventuellement un lien vers un schéma)

Les flux peuvent variés selon le type de donnée acquis (générales, expérimentales, ...)

En règle générale pour les données expérimentales :

1. Acquisition de la donnée sur matériel spécifique de terrain ou par passage d'échantillon sur appareillage ou par analyse de technicien.
2. Valeur brute stockée sur ordinateur liée au matériel d'acquisition ou sur ordinateur de l'agent ayant réalisé l'analyse, ou directement sur un espace de stockage partagé réseau (type NAS).
3. Traitement du fichier pour sa mise au format BDD (SI-PRO) et dans un format spécifique type tableur EXCEL pour le dataverse (une ligne par échantillon avec éléments analysés en colonnes) sur poste individuel relié au réseau INRAE.
4. Dépôt du fichier traité mis en forme BDD (SI-PRO) et EXCEL (dataverse) déposé sur l'espace VGG-NAS418-1 (NAS QualiAgro).
5. Dépôt et publication des fichiers en BDD (SI-PRO) et EXCEL (dataverse) après validation complète des données selon une fréquence non définie. Les données du SI sont stockés sur les datacenters INRAE.

Le SI-PRO et le dataverse permettent de faire la révision des fichiers.

Quelle est la volumétrie actuelle et prévisionnelle ?

La volumétrie des données est actuellement de 38.64 Go et se répartit de la manière suivante dans les deux espaces de stockage :

- **VGG-NAS418-1 (NAS QualiAgro) : 38.62 Go (46158 fichiers dans 2128 dossiers)**
- **SI-PRO (espace QualiAgro) : Propriété des sols : 8.91 Mo / Propriété des plantes : 8.95 Mo / Propriété des PRO : 1.82 Mo / Itinéraires techniques : 2.07 Mo**

L'augmentation de la production de données à l'avenir n'est pas définie en terme de volumétrie. Les capacités de stockage au niveau du SI-PRO et du NAS ECOSYS devraient être suffisantes.

L'entité hébergeant physiquement les données a-t-elle une politique de sécurité pour son système d'information ?
politique locale, charte des infrastructures de recherche...

[Charte des infrastructures de recherche à l'Inra](#)

Les espaces partagés sur l'espace de stockage commun (NAS QualiAgro) sont en cours de réplique (fréquence à définir) sur un ordinateur de la gestionnaire de l'essai. La sauvegarde se fait en local.

L'espace de stockage du SI-PRO est situé sur le Datacenter INRAE de Toulouse, et répliqué sur celui d'Ile de France selon la politique des datacenter INRAE.

Sécurité - Confidentialité : les données font-elles l'objet d'échange ou de partage avec de tiers acteurs et selon quelles modalités ? comment sont déterminés les droits d'accès aux données avant leur publication ?

Dans le cadre du site QualiAgro, les données sont acquises avec un partenaire privé (VEOLIA) et sont donc en copropriété avec ce partenaire. Les données sont confidentielles tant qu'elles n'ont pas été publiées après validation des co-proprétaires. Leur publication permet leur dépôt dans des entrepôts de données ou des bases de données accessibles. Les données sont transmises au partenaire privé par campagne d'acquisition et sur demande après validation par les copropriétaires pour les autres demandeurs. Dans tous les cas, il n'y a aucun accès direct aux serveurs de stockage d'ECOSYS.

Sécurité - Intégrité - Tracabilité : Quelles sont les mesures de protection mises en œuvre pour suivre la production et l'analyse des données ?

Toutes les interventions sur le terrain sont répertoriées dans les cahiers de suivi d'expérimentation par les gestionnaires de site. Les échantillons sont gérés selon la procédure propre à chaque prélèvement (voir Organisation et documentation des données dans ce PGD).

Chaque fournisseur de données valide sa production et assure la fiabilité des données qu'il fournit ; les fournisseurs de données sont validés et autorisés par les responsables scientifiques du site.

Archivage et conservation des données

Quelles sont les données à conserver sur le moyen ou le long terme et quelles sont les données à détruire ?

L'intégralité des données (brutes ou travaillées) produites par le dispositif QualiAgro (SOERE PRO) sont conservées dans les espaces de stockage dédiés à cet effet (serveur) sur le long terme pour permettre une nouvelle exploitation si nécessaire.

C'est le cas depuis 1998 année du démarrage de l'expérimentation QualiAgro.

Toutes ces données collectées sont archivées, mais ne sont à aucun moment détruites.

Sur quelle plateforme d'archivage pérenne seront archivées les données à conserver sur le long terme ? Sinon, quelles procédures seront mises en place pour la conservation à long terme ?

Les données et métadonnées du SI-PRO sont archivées sur le long terme et répliquées sur les 2 datacenter d'INRAE (Toulouse et Ile de France). Les données stockées sur le VGG-NAS418-1 (NAS QualiAgro) sont conservées également sur le long terme localement.

La politique de l'INRAE permet le partage de données à long terme sur des entrepôts sécurisés. Dans le cas du SOERE PRO, la publication sur des entrepôts de données comme Recherche Data Gouv est prévue dans un futur relativement proche (1er semestre 2025) pour permettre l'accessibilité et le réemploi des jeux de données.

Quelle est la durée de conservation des données ?

Le SOERE PRO a décidé de conserver toutes les données sans limite de durée car nous travaillons en science du sol sur des processus longs (exemple : évolution des stocks de carbone dans le sol, etc...). Elles peuvent-être réexploitées ultérieurement. Cela fait partie du rôle des observatoires à long terme.

Quelles garanties de financements couvriront les coûts associés à la conservation à long terme ?

Les coûts associés à la conservation à long terme des données sont financés à la fois par les fonds institutionnels (INRAE) et plus précisément par la Direction Systèmes d'Information (DSI), mais aussi par les fonds propres du SOERE PRO, de l'IR AnaEE France.

Dans le cadre d'un dépôt long terme sur un entrepôt de données type Recherche data gouv, celui ci est financé par INRAE. L'hébergement sécurisé et la mise à disposition des données sont garantis pendant un minimum de 5 ans renouvelables après la publication. Cette durée constitue un plancher, il ne s'agit pas de la durée maximum de conservation des données qui peut varier selon les jeux de données.

Données d'observation et d'analyses dans les compartiments sol, plante, eau, PRO - site Prospective Colmar (68)

Informations sur la structure

Nom de la structure

SOERE PRO : Observatoire de Recherche en Environnement sur les Produits Résiduaire Organiques

Type de structure

- SOERE
- ISC (Infrastructure Scientifique Collective)

SOERE : Systèmes d'Observation et d'Expérimentation au long terme pour la Recherche en Environnement

Site web : <https://valor-pro.hub.inrae.fr/>

Données générales pour la coordination technique dans espace Nexcloud INRAE (accès personnes autorisées) : https://nextcloud.inrae.fr/apps/user_saml/saml/selectUserBackEnd?redirectUrl=

Identifiant de la structure

Préciser le fournisseur de l'identifiant (ISNI, VIAF, FundRef, DataCite...).

Question sans réponse.

Responsabilités dans la structure

Nom, Prénom	Courriel	Rôle
HOUOT Sabine	sabine.houot@inrae.fr	Responsable coordination scientifique
LEVAVASSEUR Florent	florent.levasseur@inrae.fr	Responsable coordination scientifique
BENOIT Pierre	pierre.benoit@inrae.fr	Responsable coordination scientifique
MONTAGNIER Christophe	christophe.montagnier@inrae.fr	Responsable coordination et animation technique

Etablissement(s) tutelle(s)

- INRAE
- CIRAD

Département de rattachement INRAE (ou anciens départements Inra)

- AGROECOSYSTEM : Agroécosystèmes
- Département de rattachement du CIRAD : PERFORMANCE des SYSTèmes de production et de transformation tropicaux (PERSYST)

Financeur(s) (permettant l'acquisition des jeux de données - hors projet)

N/A

Informations sur le plan de gestion

[DOI \(version publiée du plan de gestion\)](#)

Historique des versions

Date	n° de version	Status	Auteur	Affiliation de l'auteur (se reporter à l' annuaire INRAE)	Validé par	Validé le
24/07/2025	V0	Public	MONTAGNIER Christophe	INRAE UMR 1402 ECOSYS Ecologie Fonctionnelle et Ecotoxicologie des agroécosystèmes	Florent LEVAVASSEUR	24/07/2025

Présentation générale des données

Mode d'obtention des données

- Données générées par la structure

L'intégralité des données collectées sont retranscrites dans des fichiers informatisés sous Excel (tableur) dans le format .xlsx ou .csv et sous Word (traitement de texte) dans le format .docx.

Les données générées par la structure :

- données générales (localisation du dispositif, contexte agro-pédo-climatique),
- données du protocole expérimental (plan du dispositif expérimental, modalités de PRO étudiées, équipements du site, suivis réalisés, banques d'échantillons),
- données d'itinéraires techniques culturaux,
- données issues d'analyses dans la matrice SOL (caractérisation du profil initial (0-100 cm) et données d'analyses détaillées dans chacune des parcelles sur 0-30/30-60 cm),
- données issues d'analyses dans la matrice SOL (profondeurs 0-28 cm, en routine tous les 2 ou 4 ans suivant les éléments analysés),
- données issues d'analyses dans la matrice PRO,
- données issues d'analyses dans la matrice PLANTES (grain, résidus),
- données issues d'analyses dans la matrice EAUX (cases lysimétriques, lysimètres et pluie),
- données issues d'analyses des reliquats azotés dans les sols (profondeurs 0-30/30-60/60-90/90-120 cm),
- données de mesure de la densité apparente du SOL (au départ du site, sur tout le profil puis suivi dans l'horizon labouré 0-30 cm)
- données de mesure de rendement (grain, résidus),
- données de mesure du statut N des cultures (acquisition avec SPAD 502+ de chez Minolta), de la concurrence des adventices et du développement des maladies,
- données lysimétriques (mesure volume d'eau drainée à 45 cm = lysimètres au champ et 90 cm = cases lysimétriques),
- données météorologiques (station météo du centre de Colmar faisant partie du réseau Agroclim) : température, hygrométrie, pluviométrie, rayonnement solaire, direction du vent, vitesse du vent
- données issues de mesures instrumentées (capteurs SOL) : teneur en eau (humidité), tension de l'eau, température entre 0 et 160 cm (thèse Arnaud ISCH)

Origine

- Analyse
- Expérimentation
- Observation

Observation = mesure au champ

Voir les exemples de protocoles d'acquisition des données dans la rubrique : **Organisation et documentation des données.**

Type de données

- Dataset
- Text
- Other (à préciser dans la zone "Informations supplémentaires")

Tabulaire (dans tableur Excel)

Dataset : format tableur type Excel

Text : format traitement de texte type WORD

Nature des données

2 types de données sont générées/gérées pour la conduite du dispositif PROspective : données textuelles (texte sous word, présentation powerpoint), données tabulaires (dans un tableur excel).

- Données textuelles de description du dispositif :

1. Localisation du dispositif, contexte agro-pédo-climatique
2. Plan du dispositif expérimental, modalités de PRO étudiées, équipements du site, suivis réalisés, banques d'échantillons
3. Protocoles de réalisation des mesures de terrain (planning)
4. Protocoles de prélèvement et de préparation des échantillons pour analyses
5. Protocoles de mesure et d'analyse des échantillons

- Données expérimentales sous forme tabulaire :

1. Compartiment SOL

Type de données : date, profondeur, modalité (traitement)

Type de données : Analyse physico-chimique et des contaminants

Variables analysées : carbone organique, azote total, matière organique, soufre, CEC, éléments échangeables (potassium, sodium, calcium et magnésium), propriétés physiques, humidité à des potentiels caractéristiques, éléments traces, autres éléments minéraux, granulométrie avec et sans décarbonatation, pH, calcaire, phosphore Olsen, composés pharmaceutiques, HAP, PCB

Type de données : analyses biologiques

Variable analysée : activité enzymatiques (C, N, S, P), biomasse microbienne

Type de données : analyses des reliquats azotés

Type de données : mesure de la densité apparente

Type de données : Issues de mesures instrumentées (capteurs SOL) : teneur en eau (humidité), tension de l'eau, température entre 0 et 130 cm

2. Compartiment PRO

Type de données : date, type de PRO

Type de données : Analyse physico-chimique et des contaminants

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, phosphore, potassium, magnésium, soufre, CEC, éléments échangeables, éléments traces, autres éléments minéraux, fractionnement biochimique (ISMO), humidité, matière sèche, pH, calcaire, composés pharmaceutiques, HAP, PCB

Type de données : Minéralisation carbone et azote

3. Compartiment PLANTES

Type de données : date, culture, modalité (traitement)

Type de données : Mesure du rendement (grains et résidus)

Type de données : Analyse physico-chimique et des contaminants (grains et résidus)

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, phosphore, potassium, soufre, éléments traces, autres éléments minéraux, humidité, matière sèche

Type de données : Mesure du statut N des cultures (SPAD 502+ de chez Minolta), concurrence des adventices et développement des maladies

4. Compartiment EAUX

Type de données : Date, profondeur, volume, modalité (traitement)

Type de données : Analyse physico-chimique et des contaminants (lysimétrie : solution du sol)

Variables analysées : carbone, azote, matière organique, calcium, sodium, éléments traces, autres éléments minéraux, pH, phosphore, potassium, composés pharmaceutiques

Type de données : Mesure du volume d'eau drainée à 45 cm (lysimétrie : solution du sol), à 90 cm (cases lysimétrique : solution du sol),

Type de données : Analyse physico-chimique et des contaminants (eau de pluie)

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, soufre, éléments traces, autres éléments minéraux, pH, phosphore, potassium

5. Compartiment AIR

Type de données : Date

Type de données : mesures météorologiques

Variables mesurées : température, pluviométrie, hygrométrie, direction du vent, vitesse du vent, rayonnement solaire

6. Itinéraires techniques cultureux

Modalité (traitement),

Dates (début, fin, enfouissement PRO),

Type d'intervention (type de culture, travail du sol, semis, récolte, PRO étudié (épandage), Intervention apport (fertilisation, traitement phytosanitaire, nom du produit),

Matériel,

Dose (unite),

Profondeur de travail,

Largeur de travail,

Commentaire

Format des données

- Format texte : .doc; .docx; .txt => données générales et du protocole expérimental
- Format tabulaire : .xls; .xlsx; .csv => tous les autres types de données

Périmètre thématique des données

- Climate
- Farming Systems and Practices
- Plant Breeding and Plant Products
- Soils and soil sciences
- Water resources

Droits de propriété intellectuelle

Les données produites dans les organismes de recherche publics doivent être communicables à toute personne qui en fait la demande si elles n'entrent pas dans le cadre des exceptions légales (sécurité défense, sécurité des populations, patrimoine scientifique et technique, données personnelles, données liées au secret, statistique, etc.), et si elles constituent des documents dits achevés. Selon la nature des données, leur communication peut être obligatoire, interdite, ou soumise à conditions. Pour plus d'information, consulter le guide "[Ouverture des données de recherche. Guide d'analyse du cadre juridique en France](#)" publié par un collectif de juristes issus de l'Enseignement Supérieur et de la Recherche.

Qui détiendra les droits sur les données et les autres informations créées ?

- INRAE 1402 ECOSYS Palaiseau (91)
- INRAE 0871 UEAV Colmar (68)

Confidentialité

Identification des jeux de données contenant des données confidentielles

INRAE facilite la diffusion systématique en libre accès des informations issues de ses travaux de recherche financés sur fonds publics, **sous réserve des droits et intérêts des parties prenantes et des tiers, y compris les droits de propriété intellectuelle, la confidentialité, ou le caractère sensible des informations** (sécurité défense, sécurité des populations, patrimoine scientifique et technique, données personnelles, données liées au secret, statistique, etc.) et si elles constituent des **documents dits achevés**. Les droits de propriété industrielle portant sur les travaux et résultats issus du site PROspective sont la copropriété du SMRA68 et d'INRAE au prorata de leurs apports respectifs.

Quelles sont les mesures prises et les normes auxquelles il est nécessaire de se conformer pour garantir cette confidentialité ?

Les données confidentielles sont collectées et stockées sur serveur avec accès sécurisé et spécifique (de l'unité INRAE 1402 ECOSYS et de l'unité 0871 UEAV).

Le cas échéant, comment la confidentialité de données fournies par des personnes sera garantie lorsque les données seront partagées ou rendues disponibles pour une analyse de second niveau ?

Néant

Partage des données

Y a t'il une obligation de partage (ou à l'inverse une interdiction ou une restriction) ?

Les unités ECOSYS et UEAV dépendent d'un organisme de recherche public (INRAE). Les données qu'elles produisent sont des données FAIR et entrent dans le cadre de la Science Ouverte. Elles doivent être Faciles à trouver, Accessibles, Interopérables et Réutilisables si elles n'entrent pas dans le cadre des exceptions légales (données confidentielles). Le dépôt de jeu de données dans l'entrepôt Recherche.Data.Gouv.fr (collection SOERE PRO) dans des formats adaptés contribue à rendre ces données FAIR.

Les producteurs des PRO étudiés sur l'essai PROspective, sont avertis des résultats obtenus sur leurs produits et leurs effets sur les compartiments des agrosystèmes avant leur diffusion plus large.

Les conditions de partage et d'utilisation des données issues du site PROspective sont décrites dans la charte interne SOERE PRO : AnaEE_Charte_PRO_Version_Octobre_2024.pdf consultable à l'adresse : <https://valor-pro.hub.inrae.fr/les-sites-experimentaux/charte-pour-l-acces-au-service-pro-anaee-france>.

Quelles sont les réutilisations potentielles de ces données ?

- Publications scientifiques
- Vulgarisation scientifique
- Travaux de Recherche et Développement
- Enseignement

La lecture des données nécessite-t-elle le recours à un logiciel ou un outil spécifique ? Si oui, lequel ?

NON. Les logiciels de la suite Office notamment Excel (tableur), Word (traitement texte), Power Point (présentation) suffisent. Des formats ouverts, plus pérennes et interopérables que ceux de la suite Office pourraient également être envisagés.

Comment les données seront-elles partagées ?

- Publication sur le portail Recherche.Data.gouv.fr (<https://entrepot.recherche.data.gouv.fr/dataverse/observatoire-soere-pro>)
- Envoi par E-mail des données et/ou des fichiers qui ne sont pas sur Recherche.Data.Gouv.fr aux utilisateurs qui le souhaitent et après accord des copropriétaires

Avec qui ?

- Autre

L'ensemble des données produites seront partagées :

- avec tous pour les données en Open Access (Recherche.Data.Gouv.fr),
- aux agents INRAE authentifiés UEAV avec droits d'accès au dossier PRO => Réseau Colmar S:/ SEAV / SEAV / PRO
- aux partenaires identifiés dans les projets de recherche auxquels ils participent.

Certaines données, des années 2000 à 2022, sont accessibles uniquement aux personnes INRAE disposant d'un accès à la copie de: G:\dmonenach sur leur espace personnel. Il y a un projet de réunir ces données sur le réseau Colmar S:/ SEAV / SEAV / PRO.

Sous quelle licence ?

- Licence ouverte <https://www.etalab.gouv.fr/licence-ouverte-open-licence> (compatible CC-BY)
- Pour les données non confidentielles : sous licence ouverte
- Pour les données confidentielles, pas de licence.

Organisation et documentation des données

Quels méthodes et outils sont utilisés pour acquérir et traiter les données, depuis leur acquisition jusqu'à leur mise à disposition, leur archivage ou leur destruction ?

Utiliser éventuellement un lien vers un schéma illustrant les processus

Les méthodes et outils pour acquérir et traiter les données du dispositif PROspective sont gérés à l'échelle de l'UEAV 0871, l'UMR 1402 ECOSYS et du SOERE PRO.

- **La collecte des données**

Les textes de référence utilisés pour acquérir les données sur le terrain sont définis dans le cadre de la démarche d'assurance qualité de l'UEAV.

Type de données	Protocoles d'acquisition utilisé / documents de référence correspondant
1. Générales :	/
- localisation du dispositif	2025_FICHE_Plateforme-PRO'spective_vf.pdf; SOEREPRO_Plaquette_2015_Colmar_ANAEEFrance.pptx
- contexte agro-pédo-climatique	2025_FICHE_Plateforme-PRO'spective_vf.pdf; SOEREPRO_Plaquette_2015_Colmar_ANAEEFrance.pptx
2. Protocole expérimental :	/
- plan du dispositif expérimental	2025_FICHE_Plateforme-PRO'spective_vf.pdf; SOEREPRO_Plaquette_2015_Colmar_ANAEEFrance.pptx
- modalités de PRO étudiées	2025_FICHE_Plateforme-PRO'spective_vf.pdf; SOEREPRO_Plaquette_2015_Colmar_ANAEEFrance.pptx
- équipements du site	2025_FICHE_Plateforme-PRO'spective_vf.pdf; SOEREPRO_Plaquette_2015_Colmar_ANAEEFrance.pptx

- suivis réalisés	2025_FICHE_Plateforme-PRO'spective_vf.pdf; SOEREPRO_Plaquette_2015_Colmar_ANAEEFrance.pptx
- banques d'échantillons	En projet de rédaction
3. Protocoles de réalisation des mesures de terrain (planning)	En projet de rédaction
4. Protocoles de prélèvement et de préparation des échantillons pour analyses	/
4a. Compartiment SOL	Protocole 9 Prélèv terre Horizon surface-V2010.docx; Protocole 8 - Sols Inférieurs.docx; Protocole-7_Prélèv_reliquats_N_version-2011.docx
4b. Compartiment PRO	Protocole 14 Epandage déchets.docx
4c. Compartiment PLANTES	Protocole_1_Conduite_Maïs_2013.docx; Protocole 2 Conduite Blé d'hiver_Version-2010.docx; Protocole-3-ConduiteBetterave-2015.docx; Protocole 4 - Conduite Orge_Version_2012.docx
4d. Compartiment EAUX	Proto 20 Suividel_eauà100cm-versionAencours[2].docx; Protocole 5-Suividel_eauà140cm.docx; Protocole 18 Prélèvement pluviomètre - Proposition modif juin 06_DM.docx; Protocole cases lysimétrique2012.docx; Pas de protocole pour les lysimètres au champ (projet de redaction)
5. Protocoles de mesure et d'analyse des échantillons	/
5a. Compartiment SOL	/
- analyse physico-chimique et des contaminants	Catalogue_LAS_2024_V0.pdf;
- analyse de la biologie	ISO 20130:2018 - Qualité du sol — Mesure de l'activité enzymatique dans des échantillons de sol en utilisant des substrats colorimétriques; SOL_BIO_METbiomassemicrobienne_002
- analyse des reliquats azotés	Protocole-7_Prélèv_reliquats_N_version-2011.docx
- mesure de la densité apparente	Densités-apparentes-cylindre-2008.docx
- mesure instrumentées avec capteurs SOL	Protocole 21 Tensiomètres.docx; 22 - Mise en place bougies, TDR et tensiomètresr.docx
5b. Compartiment PRO	/
- analyse physico-chimique et des contaminants	Protocole 14 Epandage déchets.docx; Catalogue_LAS_2024_V0.pdf; catalogue_analytique_LDAR
- minéralisation carbone et azote	Norme AFNOR XP U 44-163 Décembre 2009 : Caractérisation de la matière organique par la minéralisation potentielle du carbone et de l'azote
5c. Compartiment PLANTES	/
- analyse physico-chimique et des contaminants	Catalogue_analytique_USRAVE
- mesure de rendement	Protocole_1_Conduite_Maïs_2013.docx; Protocole 2 Conduite Blé d'hiver_Version-2010.docx; Protocole-3-ConduiteBetterave-2015.docx; Protocole 4 - Conduite Orge_Version_2012.docx
- mesure du statut N des cultures, concurrence des adventices, développement des maladies	Protocole_SPAD_2023_ThM_EFELE_adaptéVF; Protocole_notation_adventices_SOERE_PRO_VF; Protocole_notation_maladies_SOERE_PRO_VF
5d. Compartiment EAUX	/
- analyse physico-chimique et des contaminants (lysimétrie)	Catalogue_LAS_2024_V0.pdf; Protocole cases lysimétrique2012.docx; Pas de protocole pour les lysimètres au champ (projet de rédaction)
- mesure volume d'eau drainée (lysimétrie)	Protocole cases lysimétrique2012.docx; Pas de protocole pour les lysimètres au champ (projet de rédaction)
- analyse physico-chimique et des contaminants (pluie)	Protocole 18 Prélèvement pluviomètre - Proposition modif juin 06_DM.docx;
5e. Compartiment AIR	/
- mesures météorologiques	Station météo locale du centre de Colmar faisant partie du reseau Agroclim (documentation interne de l'US Agroclim Avignon)

6. Itinéraires techniques

cultureaux : type
d'intervention / nature
intervention N/A

- Le tableau suivant récapitule les étapes, méthodes et outils utilisés pour la gestion des données

Etapes	Méthode(s)	Outil(s)
Préparation	<ul style="list-style-type: none"> • Rédaction de protocoles et modes opératoires et mise à disposition des agents impliqués dans le suivi du dispositif PROspective • Formation interne aux gestes techniques 	Traitement de texte
Acquisition	<ul style="list-style-type: none"> • Procédures, protocoles et modes opératoires validés par le responsable qualité. • Pour les échantillons : collecte, conditionnement, identification et stockage. 	<ul style="list-style-type: none"> • Saisie papier / ordinateur de terrain • Outil de mesure (balance, centrale d'acquisition Campbell Scientific, etc...)
Traitement des données	<ul style="list-style-type: none"> • recueil, mise en forme des données. • analyse de premier niveau à des fins de présentation 	Tableur (Excel, Expé-R)
Mise à disposition	<ul style="list-style-type: none"> • publication sur Recherche data gouv, • agents en charge du suivi de l'essai, • après validation et autorisation pour tout scientifique qui en fait la demande autre que les équipes en charge du suivi. 	Entrepôt de données / Réseau Colmar S:/ SEAV / SEAV / PRO / SI-PRO / G:\dmontenach / Messagerie électronique (mails)
Archivage	Espace PRO dans le NAS du reseau de Colmar + Espace PROspective dans le SI-PRO	Serveur de stockage en réseau (NAS) + Système d'Information (SI)
Réutilisation	Libre quand publié sur recherche data gouv Sur demande après validation pour les données non publiées	N/A
Destruction	N/A	N/A

Quelles métadonnées seront utilisées pour accompagner le jeu de données ? Quels seront les standards, vocabulaires, taxonomies... utilisés pour décrire et représenter les données et éléments de métadonnées ? Comment les métadonnées seront-elles produites et mises à jour ?

Métadonnées	Origine, mode de production des métadonnées (ex : saisie manuelle, annotation automatique...)	Standard, Vocabulaires associés	Conditions ou fréquence de la mise à jour (si applicable) (ex : changement de l'accessibilité)
Données générales	Géoréférencement, Description de fosses pédologiques	Système de coordonnées géographiques (géodésie) Référentiel Pédologique de l'AFES 2008	N/A
Protocoles expérimentaux	Saisie manuelle	Trame démarche qualité INRAE	En cas de changement dans le protocole général
Méthode d'analyses physico-chimiques	Laboratoires d'analyses	Normes en vigueur (méthodes utilisées pour analyser les différents éléments)	N/A
Itinéraire technique	Saisie manuelle, mesure de doses (engrais minéral, digestat, PRO)	Thesaurus SOERE PRO (SI-PRO)	Tous les ans
Données climatiques	Station météorologique sur le centre de Colmar	Standards AGROCLIM / Meteo France	En continu

Une documentation complémentaire aux métadonnées est-elle nécessaire pour décrire les données et assurer leur réutilisabilité sur le long terme ?

Chaque Dataset (jeux de données) devra être accompagné d'un fichier nommé "Read-me ou Lisez-moi" lors de leur publication sur le site de l'Entrepôt de Recherche : <https://recherche.data.gouv.fr/>. Une trame pour générer plus facilement un Read-me est proposée par l'équipe Recherche Data Gouv : <https://recherche.data.gouv.fr/fr/categorie/33/guide/modele-de-readme>. Ce fichier a pour but d'aider à la compréhension et à l'interprétation des données.

Comment les fichiers de données sont-ils gérés et organisés : contrôle des versions, conventions de nommage des fichiers, organisation des fichiers

• Pour les données gérées et organisées dans des dossiers de fichiers au sein du serveur de Colmar

L'autorisation d'accès est donnée par l'informaticien du centre suite à une demande du DU que ce soit pour le réseau Colmar S:/ SEAV / SEAV / PRO ou la copie de G:\dmontenach. Il est prévu de rassembler l'ensemble des données dans le dossier PRO afin d'en faciliter l'accès.

Les fichiers de données sont organisés par année et/ou par thématique

• Pour les données déposées dans l'entrepôt de données Recherche Data Gouv.

Les dépôts des fichiers de données entraînent l'attribution d'un DOI avec une version de publication (actuellement V1). Les fichiers sont accessibles en suivant le chemin suivant dans l'entrepôt :

Recherche Data Gouv/Data INRAE/Experimental - Observation - Simulation/SOERE PRO - Long term field experiment network for research on the recycling of organic waste products in agriculture/PROspective.

Jeux de données :

PROspective's long-term field experiment : <https://doi.org/10.57745/4BYXTK>

PROspective - organic waste products physicochemical dataset : <https://doi.org/10.57745/6GCPX6>

PROspective - plant physicochemical dataset : <https://doi.org/10.57745/BPB34G>

PROspective - soil physicochemical dataset : <https://doi.org/10.57745/IOSBW6>

PROspective - agronomic management dataset : <https://doi.org/10.57745/77AFBH>

• Les données sont gérées et organisées également dans le Système d'Information (SI) PRO. Celui-ci est en production et disponible à l'adresse : <https://si-pro.fr/pro-portage-1709-web-SNAPSHOT/index.jsf>

Il contient des données de thématiques Physico-chimie et d'Itinéraires Techniques (ITK).

Pour la thématique physico-chimie, nous avons des types de données Sol (données élémentaires et moyennées), PRO (élémentaires et moyennées), Plantes (élémentaires et moyennées dont le rendement des cultures) et pour les ITK, nous avons plusieurs types d'interventions à savoir interventions apport, pro-étudié, récolte-coupe, semis plantation, travail du sol.

On trouve également dans le SI des données de référence (métadonnées) qui sont soit partagées par tous les types de données (dispositif, parcelles, traitement, etc.), soit spécifiques à chaque type de données (SOL, ITK...).

Enfin on trouve des données de prélèvements et d'échantillons auxquelles sont liées des données produites au laboratoire pour chaque type de données.

Le SI permet de réaliser le dépôt et l'insertion de données en base de données par l'intermédiaire de fichiers csv au format prédéfinis (référentiels et les données).

L'insertion de données de références en base de données se fait sans versionnement préalable à l'inverse des données produites au laboratoire.

Pour le dépôt des fichiers de données, le SI propose une gestion des versions puis la publication qui se traduit par l'insertion en base de la version souhaitée.

Le SI permet également la consultation et l'extraction de données depuis l'interface web de données.

Les données de l'essai PROspective sont consultables en téléchargement au format .csv en suivant le chemin d'accès suivant : Accueil / Gestion des données expérimentales / PROspective.

Dans PROspective, les données s'organisent ainsi :

Propriétés des sols : sols physico-chimie données élémentaires + sols physico-chimie moyennes / période du jeu de données => téléchargement

Propriétés des plantes : plante physico-chimie données élémentaires + plante physico-chimie moyennes / période du jeu de données => téléchargement

Propriétés des PRO : PRO physico-chimie données élémentaires + PRO physico-chimie moyennes / période du jeu de données => téléchargement

Itinéraires techniques : intervention apport + intervention pro étudié + intervention récolte coupe + intervention semis plantation + intervention travail du sol / période du jeu de données => téléchargement

Quel est le processus de contrôle qualité des données ?

La validation des données consiste à s'assurer que :

- elles ne comportent pas d'erreur de saisie, que les unités de mesure soient correctement transcrites, que la date et le lieu de la collecte sont facilement identifiables et si possible que l'appareil utilisé lors de la collecte apparaisse sur le fichier transmis.
- les données transmises par les laboratoires ne présentent pas de valeurs aberrantes. Si c'est le cas, une contre-analyse peut-être

demandée au laboratoire.

La méthode utilisée par la gestionnaire du dispositif pour la recherche des valeurs aberrantes est la suivante : la valeur d'une mesure obtenue dans une des répétitions de l'essai est considérée comme aberrante si sa valeur est 2 fois supérieure à la plus grande valeur des répétitions ou deux fois inférieure à la plus petite valeur des répétitions. Si c'est le cas, la valeur est remplacée par la moyenne des valeurs des autres répétitions.

Elle est réalisée par le/la gestionnaire du dispositif PROspective.

Stockage et sécurité des données

Quels sont les types de flux empruntés par les données et les supports utilisés pour les stocker ?
(Faire éventuellement un lien vers un schéma)

Les flux peuvent varier selon le type de donnée acquis (générales, expérimentales, ...)

En règle générale pour les données expérimentales :

1. Acquisition de la donnée sur matériel spécifique de terrain ou par passage d'échantillon sur appareillage ou par analyse de technicien.
2. Valeur brute stockée sur ordinateur liée au matériel d'acquisition ou sur ordinateur de l'agent ayant réalisé l'analyse, ou directement sur un le Réseau Colmar S:/ SEAV / SEAV / PRO.
3. Valeur validée stockée sur un Réseau Colmar S:/ SEAV / SEAV / PRO.
4. Traitement du fichier pour sa mise au format BDD (SI-PRO et dataverse) sur poste individuel relié au réseau INRAE.
5. Dépôt et publication des fichiers en BDD (SI-PRO et dataverse) après validation complète des données selon une fréquence non définie. Les données du SI sont stockées sur les datacenters INRAE.

Le SI-PRO et le dataverse permettent de faire la révision des fichiers.

Quelle est la volumétrie actuelle et prévisionnelle ?

La volumétrie des données est actuellement de 191 Go et se répartit de la manière suivante dans les deux espaces de stockage :

- **Espace serveur de Colmar S:/ SEAV / SEAV / PRO : 49 Go (77910 fichiers dans 14683 dossiers)**
- **Espace serveur G:\dmontenach : 141 Go (86813 fichiers dans 8141 dossiers)**
- **Espace PROspective** dans l'entrepôt de données Recherche Data Gouv : 23.23 Mo.
- **SI-PRO (espace PROspective) : Propriété des sols : 6.06 Mo / Propriété des plantes : 11.75 Mo / Propriété des PRO : 2.65 Mo / Itinéraires techniques : 57 Ko**

L'augmentation de la production de données à l'avenir n'est pas définie en terme de volumétrie. Les capacités de stockage au niveau du SI-PRO et du serveur de Colmar devraient être suffisantes. Afin d'optimiser la volumétrie, il y a un projet de rassembler les données de G:\dmontenach sur S:/ SEAV / SEAV / PRO.

L'entité hébergeant physiquement les données a-t-elle une politique de sécurité pour son système d'information ?
politique locale, charte des infrastructures de recherche...

[Charte des infrastructures de recherche à l'Inra](#)

La sécurité du système d'information est gérée au niveau des infrastructures nationales d'INRAE (datacenters de Toulouse et Ile de France).

Sécurité - Confidentialité : les données font-elles l'objet d'échange ou de partage avec de tiers acteurs et selon quelles modalités ? comment sont déterminés les droits d'accès aux données avant leur publication ?

Les données sont confidentielles tant qu'elles n'ont pas été validées par les responsables scientifiques et techniques du site.

Les unités ECOSYS et UEAV dépendent d'un organisme de recherche public (INRAE). Les données qu'elles produisent sont des données FAIR et entrent dans le cadre de la Science Ouverte. Elles doivent être Faciles à trouver, Accessibles, Interopérables et Réutilisables pour toute personne qui en fait la demande si elles n'entrent pas dans le cadre des exceptions légales (données confidentielles).

Toute personne ou scientifique autre que les équipes en charge du suivi du site a accès aux données après leur validation et autorisation par les propriétaires des données.

Les producteurs des PRO étudiés sur l'essai PROspective, sont avertis des résultats obtenus sur leurs produits et leurs effets sur les

compartiments des agrosystèmes avant leur diffusion plus large.

Sécurité - Intégrité - Tracabilité : Quelles sont les mesures de protection mises en œuvre pour suivre la production et l'analyse des données ?

Toutes les interventions sur le terrain sont répertoriées dans les cahiers de suivi d'expérimentation par les gestionnaires de site. Les échantillons sont gérés selon la procédure propre à chaque prélèvement (voir Organisation et documentation des données dans ce PGD).

Chaque fournisseur de données valide sa production et assure la fiabilité des données qu'il fournit ; les fournisseurs de données sont validés et autorisés par les responsables scientifiques du site.

Archivage et conservation des données

Quelles sont les données à conserver sur le moyen ou le long terme et quelles sont les données à détruire ?

L'intégralité des données (brutes ou travaillées) produites par le dispositif PROspective (SOERE PRO) sont conservées dans les espaces de stockage dédiés à cet effet (serveur) sur le long terme pour permettre une nouvelle exploitation si nécessaire.

C'est le cas depuis 2000, année du démarrage de l'expérimentation PROspective.

Toutes ces données collectées sont archivées, mais ne sont à aucun moment détruites.

Sur quelle plateforme d'archivage pérenne seront archivées les données à conserver sur le long terme ? Sinon, quelles procédures seront mises en place pour la conservation à long terme ?

Les données et métadonnées du SI-PRO sont archivées sur le long terme et répliquées sur les 2 datacenter d'INRAE (Toulouse et Ile de France).

La politique de l'INRAE permet le partage de données à long terme sur des entrepôts sécurisés. Dans le cas du SOERE PRO, la publication sur des entrepôts de données comme Recherche Data Gouv est prévue dans un futur relativement proche (1er semestre 2025) pour permettre l'accessibilité et le réemploi des jeux de données.

Quelle est la durée de conservation des données ?

Le SOERE PRO a décidé de conserver toutes les données sans limite de durée car nous travaillons en science du sol sur des processus longs (exemple : évolution des stocks de carbone dans le sol, etc...). Elles peuvent être réexploitées ultérieurement. Cela fait partie du rôle des observatoires à long terme.

Quelles garanties de financements couvriront les coûts associés à la conservation à long terme ?

Les coûts associés à la conservation à long terme des données sont financés à la fois par les fonds institutionnels (INRAE) et plus précisément par la Direction Systèmes d'Information (DSI), mais aussi par les fonds propres du SOERE PRO, de l'IR AnaEE France. Dans le cadre d'un dépôt long terme sur un entrepôt de données type Recherche data gouv, celui ci est financé par INRAE. L'hébergement sécurisé et la mise à disposition des données sont garantis pendant un minimum de 5 ans renouvelables après la publication. Cette durée constitue un plancher, il ne s'agit pas de la durée maximum de conservation des données qui peut varier selon les jeux de données.

Données d'observation et d'analyses dans les compartiments sol, plante, eau, air, PRO - site EFELE Le Rheu (35)

Informations sur la structure

Nom de la structure

SOERE PRO : Observatoire de Recherche en Environnement sur les Produits Résiduaire Organiques

Type de structure

- SOERE
- ISC (Infrastructure Scientifique Collective)

SOERE : Systèmes d'Observation et d'Expérimentation au long terme pour la Recherche en Environnement

Site web : <https://valor-pro.hub.inrae.fr/>

Données générales pour la coordination technique dans espace Nexcloud INRAE (accès personnes autorisées)
: https://nextcloud.inrae.fr/apps/user_saml/saml/selectUserBackEnd?redirectUrl=

Identifiant de la structure

Préciser le fournisseur de l'identifiant (ISNI, VIAF, FundRef, DataCite...).

Question sans réponse.

Responsabilités dans la structure

Nom, Prénom	Courriel	Rôle
HOUOT Sabine	sabine.houot@inrae.fr	Responsable coordination scientifique
LEVAVASSEUR Florent	florent.levasseur@inrae.fr	Responsable coordination scientifique
BENOIT Pierre	pierre.benoit@inrae.fr	Responsable coordination scientifique
MONTAGNIER Christophe	christophe.montagnier@inrae.fr	Responsable coordination et animation technique

Etablissement(s) tutelle(s)

- INRAE
- CIRAD

Département de rattachement INRAE (ou anciens départements Inra)

- AGROECOSYSTEM : Agroécosystèmes
- Département de rattachement du CIRAD : PERFORMANCE des SYSTèmes de production et de transformation tropicaux (PERSYST)

N/A

Informations sur le plan de gestion

DOI ([version publiée du plan de gestion](#))

Historique des versions

Date	n° de version	Status	Auteur	Affiliation de l'auteur (se reporter à l' annuaire INRAE)	Validé par	Validé le
24/07/2025	V0	Public	MONTAGNIER Christophe	INRAE UMR 1402 ECOSYS Ecologie Fonctionnelle et Ecotoxicologie des agroécosystèmes	Florent LEVAVASSEUR	24/07/2025

Présentation générale des données

Mode d'obtention des données

- Données générées par la structure

L'intégralité des données collectées sont retranscrites dans des fichiers informatisés sous Excel (tableur) dans le format .xlsx ou .csv et sous Word (traitement de texte) dans le format .docx.

Les données générées par la structure :

- données générales (localisation du dispositif, contexte agro-pédo-climatique)
- données du protocole expérimental (plan du dispositif expérimental, modalités de PRO étudiées, équipements du site, suivis réalisés, banques d'échantillons)
- données d'itinéraires techniques culturaux,
- données issues d'analyses dans la matrice SOL (caractérisation du profil initial (0-100 cm), par horizons),
- données issues d'analyses dans la matrice SOL (profondeurs 0-25 / 25-30 / 30-40 cm, en routine tous les 4 ans suivant les éléments analysés),
- données issues d'analyses dans la matrice PRO,
- données issues d'analyses dans la matrice PLANTES (grain, paille, plante entière),
- données issues d'analyses dans la matrice EAUX (lysimètres à mèche),
- données issues d'analyses des reliquats azotés dans les sols (profondeurs 0-30/30-60/60-90 cm),
- données de mesure de la densité apparente du SOL (au départ du site, sur le profil 0-100 cm puis suivi tous les 4 ans sur le profil 0-40 cm)
- données de mesure de rendement (grain, résidus),
- données de mesure du statut N des cultures (acquisition avec SPAD 502+ de chez Minolta), de la concurrence des adventices et du développement des maladies,
- données lysimétriques (mesure volume d'eau drainée à 40 et 90 cm de profondeur),
- données météorologiques (station météo Campbell installée sur site : température, hygrométrie, pluviométrie, rayonnement global, direction du vent, vitesse du vent)
- données issues de mesures instrumentées (capteurs SOL) : teneur en eau (humidité), tension de l'eau, température à différentes profondeurs : 5, 13, 25, 40, 60, 80 et 110 cm

Origine

- Analyse
- Expérimentation
- Observation

Observation = mesure au champ

Voir les exemples de protocoles d'acquisition des données dans la rubrique : **Organisation et documentation des données.**

Type de données

- Dataset
- Text

Tabulaire (dans tableur Excel)

Dataset : format tableur type Excel

Text : format traitement de texte type WORD

Nature des données

2 types de données sont générées pour la conduite du dispositif EFELE : données textuelles (texte sous word, présentation powerpoint), données tabulaires (dans un tableur excel ou fichiers texte (txt)).

- Données textuelles de description du dispositif :

1. Localisation du dispositif, contexte agro-pédo-climatique
2. Plan du dispositif expérimental, modalités de PRO étudiées, équipements du site, suivis réalisés, banques d'échantillons
3. Plannings annuels de réalisation des mesures de terrain
4. Protocoles de prélèvement et de préparation des échantillons pour analyses
5. Protocoles de mesure et d'analyse des échantillons

- Données expérimentales sous forme tabulaire :

1. Compartiment SOL

Type de données : date, profondeur, modalité (traitement)

Type de données : Analyse physico-chimique

Variables analysées : carbone organique, azote total, matière organique, soufre, CEC, éléments échangeables (potassium, sodium, calcium et magnesium), propriétés physiques (stabilité structural), éléments traces, autres éléments minéraux, granulométrie, pH, phosphore Olsen,

Type de données : analyses biologiques : activités enzymatiques, communautés microbiennes, communautés lombriciennes

Type de données : analyses des reliquats d'azote

Type de données : mesure de la densité apparente

Type de données : Issues de mesures instrumentées (capteurs SOL) : teneur en eau (humidité), tension de l'eau, température entre 5 et 110 cm de profondeur

2. Compartiment PRO

Type de données : date, type de PRO

Type de données : Analyse physico-chimique et des contaminants

Variables analysées : carbone, azote, matière organique, calcium, manganèse, sodium, phosphore, potassium, magnesium, soufre, CEC, éléments échangeables, éléments traces, autres éléments minéraux, fractionnement biochimique (ISMO), humidité, matière sèche, pH

Type de données : Minéralisation carbone et azote

3. Compartiment PLANTES

Type de données : date, culture, modalité (traitement)

Type de données : Mesure du rendement (grains et résidus de culture)

Type de données : Analyse physico-chimique (grains et résidus de culture)

Type de données : Mesure du statut N des cultures (SPAD 502+ de chez Minolta), concurrence des adventices et développement des maladies

4. Compartiment EAUX

Type de données : Date, profondeur, volume, modalité (traitement)

Type de données : Analyse chimique : anions et cations (lysimétrie : solution du sol)

Type de données : Mesure du volume d'eau drainée par les lysimètres à mèche à 40 et 90 cm de profondeur

5. Compartiment AIR

Type de données : Date

Type de données : mesures météorologiques

Variables mesurées : température, pluviométrie, hygrométrie, direction du vent, vitesse du vent, rayonnement solaire

Type de données : émission des gaz à effet de serre (GES) : CO2 et N2O

6. Itinéraires techniques cultureux

Modalité (traitement),

Dates (début, fin, enfouissement PRO),

Type d'intervention (type de culture, travail du sol, semis, récolte, PRO étudié (épandage), Intervention apport (fertilisation, traitement phytosanitaire, nom du produit),

Matériel,

Dose (unite),

Profondeur de travail,

Largeur de travail,

Commentaire

Format des données

- Format texte : .doc; .docx; .txt => données générales et du protocole expérimental
- Format tabulaire : .xls; .xlsx; .csv => tous les autres types de données

Périmètre thématique des données

- Climate
- Farming Systems and Practices
- Plant Breeding and Plant Products
- Soils and soil sciences
- Water resources

- Greenhouse gazes (GHG)

Droits de propriété intellectuelle

Les données produites dans les organismes de recherche publics doivent être communicables à toute personne qui en fait la demande si elles n'entrent pas dans le cadre des exceptions légales (sécurité défense, sécurité des populations, patrimoine scientifique et technique, données personnelles, données liées au secret, statistique, etc.), et si elles constituent des documents dits achevés. Selon la nature des données, leur communication peut être obligatoire, interdite, ou soumise à conditions. Pour plus d'information, consulter le guide "[Ouverture des données de recherche. Guide d'analyse du cadre juridique en France](#)" publié par un collectif de juristes issus de l'Enseignement Supérieur et de la Recherche.

Qui détiendra les droits sur les données et les autres informations créées ?

INRAE UMR1069 SAS (Sol Agro et hydrosystème Spatialisation) Rennes (35)

Confidentialité

Identification des jeux de données contenant des données confidentielles

INRAE facilite la diffusion systématique en libre accès des informations issues de ses travaux de recherche financés sur fonds publics, **sous réserve des droits et intérêts des parties prenantes et des tiers, y compris les droits de propriété intellectuelle, la confidentialité, ou le caractère sensible des informations** (sécurité défense, sécurité des populations, patrimoine scientifique et technique, données personnelles, données liées au secret, statistique, etc.) et si elles constituent des **documents dits achevés**.

Les droits de propriété industrielle portant sur les travaux et résultats issus du site EFELE sont la propriété d'INRAE.

Quelles sont les mesures prises et les normes auxquelles il est nécessaire de se conformer pour garantir cette confidentialité ?

Les données confidentielles sont collectées et stockées sur serveur avec accès sécurisé et spécifique de l'UMR INRAE 1069 SAS.

Le cas échéant, comment la confidentialité de données fournies par des personnes sera garantie lorsque les données seront partagées ou rendues disponibles pour une analyse de second niveau ?

Néant

Partage des données

Y a-t-il une obligation de partage (ou à l'inverse une interdiction ou une restriction) ?

L'UMR SAS dépend d'un organisme de recherche public (INRAE). Les données qu'elles produisent sont des données FAIR et entrent dans le cadre de la Science Ouverte. Elles doivent être Faciles à trouver, Accessibles, Interopérables et Réutilisables si elles n'entrent pas dans le cadre des exceptions légales (données confidentielles). Le dépôt de jeu de données dans l'entrepôt Recherche.Data.Gouv.fr (collection SOERE PRO) dans des formats adaptés contribue à rendre ces données FAIR.

Les conditions de partage et d'utilisation des données issues du site EFELE sont décrites dans la charte interne SOERE PRO : AnaEE_Charte_PRO_Version_Octobre_2024.pdf consultable à l'adresse : <https://valor-pro.hub.inrae.fr/les-sites-experimentaux/charte-pour-l-acces-au-service-pro-anaee-france>.

Quelles sont les réutilisations potentielles de ces données ?

- Publications scientifiques
 - Vulgarisation scientifique
 - Travaux de Recherche et Développement
 - Enseignement
-

La lecture des données nécessite-t-elle le recours à un logiciel ou un outil spécifique ? Si oui, lequel ?

NON. Les logiciels de la suite Office notamment Excel (tableur), Word (traitement texte), Power Point (présentation) suffisent. Des formats ouverts, plus pérennes et interopérables que ceux de la suite Office pourraient également être envisagés.

Comment les données seront-elles partagées ?

- Publication sur le portail Recherche.Data.gouv.fr (<https://entrepot.recherche.data.gouv.fr/dataverse/efele>)
 - Envoi par E-mail des données et/ou des fichiers qui ne sont pas sur Recherche.Data.Gouv.fr aux utilisateurs qui le souhaitent et après accord des propriétaires des données
-

Avec qui ?

- Autre

L'ensemble des données produites seront partagées :

- avec tous pour les données en Open Access (Recherche.Data.Gouv.fr),
- aux partenaires identifiés dans les projets de recherche auxquels ils participent.
- aux scientifiques intéressés par les données, après validation de la demande par le responsable du site

Sous quelle licence ?

- Autre (à préciser dans la zone d'Informations supplémentaires)
- Licence ouverte <https://www.etalab.gouv.fr/licence-ouverte-open-licence> (compatible CC-BY)
- Pour les données non confidentielles : sous licence ouverte
- Pour les données confidentielles, pas de licence

Organisation et documentation des données

Quels méthodes et outils sont utilisés pour acquérir et traiter les données, depuis leur acquisition jusqu'à leur mise à disposition, leur archivage ou leur destruction ?

Utiliser éventuellement un lien vers un schéma illustrant les processus

Les méthodes et outils pour acquérir et traiter les données du dispositif EFELE sont gérés par l'UMR SAS, en accord avec les règles fixées au sein du SOERE PRO.

- **La collecte des données**

Les textes de référence utilisés pour acquérir les données sur le terrain sont définis dans le cadre de la démarche d'assurance qualité de l'UEAV.

Type de données	Protocoles d'acquisition utilisés / documents de référence correspondant
1. Générales :	/
- localisation du dispositif	https://valor-pro.hub.inrae.fr/les-sites-experimentaux/efele SOEREPRO_Plaquette_2025_EFELE_ANAEFFrance.pptx
- contexte agro-pédo-climatique	https://valor-pro.hub.inrae.fr/les-sites-experimentaux/efele SOEREPRO_Plaquette_2025_EFELE_ANAEFFrance.pptx
2. Protocole expérimental :	/
- plan du dispositif expérimental	https://valor-pro.hub.inrae.fr/les-sites-experimentaux/efele SOEREPRO_Plaquette_2025_EFELE_ANAEFFrance.pptx
- modalités de PRO étudiées	https://valor-pro.hub.inrae.fr/les-sites-experimentaux/efele SOEREPRO_Plaquette_2025_EFELE_ANAEFFrance.pptx
- équipements du site	https://valor-pro.hub.inrae.fr/les-sites-experimentaux/efele SOEREPRO_Plaquette_2025_EFELE_ANAEFFrance.pptx
- suivis réalisés	https://valor-pro.hub.inrae.fr/les-sites-experimentaux/efele SOEREPRO_Plaquette_2025_EFELE_ANAEFFrance.pptx
- banques d'échantillons	En projet de rédaction
3. Protocoles de réalisation des mesures de terrain (planning)	Prévisionnel_Travaux_EFELE_année.docx
4. Protocoles de prélèvement et de préparation des échantillons pour analyses	/
	Proto_Monitoring_sol_année.docx ; Protocole_reliquatsN_année.docx ;

4a. Compartiment SOL	Localisation_sondages_Reliquats.pptx ; Protocole_Bilan Sol_année.docx ; Zones_plt_EFELE_2012_2029.pptx
4b. Compartiment PRO	Proto_Apports_PROs_solides_année.docx ; Proto_Apports-engrais_PROs_liquides_année.docx
4c. Compartiment PLANTES	Protocole_blé_année.docx ; Protocole_maïs_année.docx ; Protocole_sorgho_année.docx ; Protocole_méteil_année.docx ; Protocole_CI_année.docx
4d. Compartiment EAUX	Protocole_lysim_année.docx
5. Protocoles de mesure et d'analyse des échantillons	/
5a. Compartiment SOL	/
- analyse physico-chimique et des contaminants	Catalogue_LAS_2024_V0.pdf;
- analyse de la biologie	Mesures d'activité enzymatiques : DI-RV-17-0021 Cheviron N et al ; DI-RV-23-0007 : Cheviron N et al
- analyse des reliquats azotés	Catalogue Labocea
- mesure de la densité apparente	Méthode interne
- mesure instrumentées avec capteurs SOL	Conversion des données et archivage sur le logiciel Hydras
5b. Compartiment PRO	/
- analyse physico-chimique et des contaminants	Catalogue_LAS_2024_V0.pdf; catalogue_analytique_LDAR
- minéralisation carbone et azote	Norme AFNOR XP U 44-163 Décembre 2009 : Caractérisation de la matière organique par la minéralisation potentielle du carbone et de l'azote - modifiée selon protocole interne
5c. Compartiment PLANTES	/
- analyse physico-chimique et des contaminants	Catalogue_analytique_USRAVE
- mesure de rendement	Protocole_blé_année.docx ; Protocole_maïs_année.docx ; Protocole_sorgho_année.docx ; Protocole_méteil_année.docx ; Protocole_CI_année.docx
- mesure du statut N des cultures, concurrence des adventices, développement des maladies	Protocole_SPAD_2023_ThM_EFELE_adaptéVF; Protocole_notation_adventices_SOERE_PRO_VF; Protocole_notation_maladies_SOERE_PRO_VF
5d. Compartiment EAUX	/
- analyse physico-chimique et des contaminants (lysimétrie)	Catalogue_LAS_2024_V0.pdf;
- mesure volume d'eau drainée (lysimétrie)	Relevés_lysim_année.xls
5e. Compartiment AIR	/
- mesures météorologiques	Station météo Campbell installée sur site ; EFELE_METEO_année.xls
6. Itinéraires techniques cultureux : type d'intervention / nature intervention	Historique_Suivi_Intervention.xls

- Le tableau suivant récapitule les étapes, méthodes et outils utilisés pour la gestion des données

Etapes	Méthode(s)	Outil(s)
Préparation	<ul style="list-style-type: none"> Rédaction de protocoles et modes opératoires et mise à disposition des agents impliqués dans le suivi du dispositif EFELE Formation interne aux gestes techniques 	Traitement de texte
Acquisition	<ul style="list-style-type: none"> Procédures, protocoles et modes opératoires validés par le responsable qualité. Pour les échantillons : collecte, conditionnement, identification et stockage. 	<ul style="list-style-type: none"> Saisie papier / ordinateur de terrain Outil de mesure (balance, centrale d'acquisition Campbell Scientific, etc...)
Traitement des données	<ul style="list-style-type: none"> recueil, mise en forme des données. analyse de premier niveau à des fins de présentation 	Tableur (Excel), Logiciel R
Mise à disposition	<ul style="list-style-type: none"> publication sur Recherche data gouv, agents en charge du suivi de l'essai, après validation et autorisation pour tout scientifique qui en fait la demande autre que les équipes en charge du suivi. 	https://entrepot.recherche.data.gouv.fr/dataverse/efele
Archivage	Espace EFELE sur le serveur de stockage renk de l'UMR SAS	Serveur de stockage renk de l'UMR SAS
Réutilisation	Libre quand publié sur recherche data gouv Sur demande après validation pour les données non publiées	N/A
Destruction	N/A	N/A

Quelles métadonnées seront utilisées pour accompagner le jeu de données ? Quels seront les standards, vocabulaires, taxonomies... utilisés pour décrire et représenter les données et éléments de métadonnées ? Comment les métadonnées seront-elles produites et mises à jour ?

Métadonnées	Origine, mode de production des métadonnées (ex : saisie manuelle, annotation automatique...)	Standard, Vocabulaires associés	Conditions ou fréquence de la mise à jour (si applicable) (ex : changement de l'accessibilité)
Données générales	Géoréférencement, Description de fosses pédologiques	Système de coordonnées géographiques (géodésie) Référentiel Pédologique de l'AFES 2008	N/A
Protocoles expérimentaux	Saisie manuelle	Trame démarche qualité INRAE	En cas de changement dans le protocole général
Méthode d'analyses physico-chimiques	Laboratoires d'analyses	Normes en vigueur (méthodes utilisées pour analyser les différents éléments)	N/A
Itinéraire technique	Saisie manuelle, mesure de doses (engrais minéral, digestat, PRO)	Thesaurus SOERE PRO (SI-PRO)	Tous les ans
Données climatiques	Station météorologique Campbell installée sur site	Standards SI	En continu

Une documentation complémentaire aux métadonnées est-elle nécessaire pour décrire les données et assurer leur réutilisabilité sur le long terme ?

Chaque Dataset (jeux de données) devra être accompagné d'un fichier nommé "Read-me ou Lisez-moi" lors de leur publication sur le site de l'Entrepôt de Recherche : <https://recherche.data.gouv.fr/fr>. Une trame pour générer plus facilement un Read-me est proposée par l'équipe Recherche Data Gouv : <https://recherche.data.gouv.fr/fr/categorie/33/guide/modele-de-readme>. Ce fichier a pour but d'aider à la compréhension et à l'interprétation des données.

Des datapaper sont associés à certains dataset, et permettent une compréhension plus approfondie des données et de leur valorisation

potentielle.

Comment les fichiers de données sont-ils gérés et organisés : contrôle des versions, conventions de nommage des fichiers, organisation des fichiers

- **Pour les données gérées et organisées dans des dossiers de fichiers au sein du serveur renk de l'UMR SAS**

L'autorisation d'accès est donnée par l'informaticien de l'UMR SAS aux personnels impliqués dans le suivi agronomique et météorologique du dispositif.

Les fichiers de données sont organisés par thématique.

- **Pour les données déposées dans l'entrepôt de données Recherche Data Gov.**

Les dépôts des fichiers de données entraînent l'attribution d'un DOI avec une version de publication. Les fichiers sont accessibles en suivant le chemin suivant dans l'entrepôt : <https://entrepot.recherche.data.gouv.fr/dataverse/efele>

Jeux de données :

Soil enzymatic activity data over eight years at the **EFELE** site, a long-term field experiment on effects of organic waste products and tillage practices : <https://doi.org/10.15454/KYYOPH>

Time evolution of nitrate, nitrite and dissolved organic carbon concentrations of the water collected by **EFELE**'s lysimeters :

<https://doi.org/10.15454/RKYCLF>

A dataset of crop yields and exports of C, N, P, K, and Mg over 9 years at the **EFELE** long-term field-experiment site :

<https://doi.org/10.57745/MZ79D8>

- **Les données sont gérées et organisées également dans le Système d'Information (SI) PRO.** Celui-ci est en production et disponible à l'adresse : <https://si-pro.fr/pro-portage-1709-web-SNAPSHOT/index.jsf>

Il contient des données de thématiques Physico-chimie et d'Itinéraires Techniques (ITK).

Pour la thématique physico-chimie, nous avons des types de données Sol (données élémentaires et moyennées), PRO (élémentaires et moyennées), Plantes (élémentaires et moyennées dont le rendement des cultures) et pour les ITK, nous avons plusieurs types d'interventions à savoir interventions apport, pro-étudié, récolte-coupe, semis plantation, travail du sol.

On trouve également dans le SI des données de référence (métadonnées) qui sont soit partagées par tous les types de données (dispositif, parcelles, traitement, etc.), soit spécifiques à chaque type de données (SOL, ITK...).

Enfin on trouve des données de prélèvements et d'échantillons auxquelles sont liées des données produites au laboratoire pour chaque type de données.

Le SI permet de réaliser le dépôt et l'insertion de données en base de données par l'intermédiaire de fichiers csv au format prédéfinis (référentiels et les données).

L'insertion de données de références en base de données se fait sans versionnement préalable à l'inverse des données produites au laboratoire.

Pour le dépôt des fichiers de données, le SI propose une gestion des versions puis la publication qui se traduit par l'insertion en base de la version souhaitée.

Le SI permet également la consultation et l'extraction de données depuis l'interface web de données.

Les données de l'essai EFELE sont consultables en téléchargement au format .csv en suivant le chemin d'accès suivant : Accueil / Gestion des données expérimentales / EFELE.

Dans EFELE, les données s'organisent ainsi :

Propriétés des sols : sols physico-chimie données élémentaires + sols physico-chimie moyennes / période du jeu de données => téléchargement

Propriétés des plantes : plante physico-chimie données élémentaires + plante physico-chimie moyennes / période du jeu de données => téléchargement

Propriétés des PRO : PRO physico-chimie données élémentaires + PRO physico-chimie moyennes / période du jeu de données => téléchargement

Itinéraires techniques : intervention apport + intervention pro étudié + intervention récolte coupe + intervention semis plantation + intervention travail du sol / période du jeu de données => téléchargement

Quel est le processus de contrôle qualité des données ?

La validation des données consiste à s'assurer que :

- elles ne comportent pas d'erreur de saisie, que les unités de mesure soient correctement transcrites, que la date et le lieu de la collecte sont facilement identifiables
- les données transmises par les laboratoires ne présentent pas de valeurs aberrantes. Si c'est le cas, une contre-analyse est demandée au laboratoire.

La méthode utilisée par le gestionnaire du dispositif pour la recherche des valeurs aberrantes est la suivante : la valeur d'une mesure obtenue dans une des répétitions de l'essai est considérée comme aberrante si sa valeur est 2 fois supérieure à la plus grande valeur des répétitions ou deux fois inférieure à la plus petite valeur des répétitions. Si c'est le cas, la valeur est remplacée par la moyenne des valeurs des autres répétitions.

Stockage et sécurité des données

Quels sont les types de flux empruntés par les données et les supports utilisés pour les stocker ?
(Faire éventuellement un lien vers un schéma)

Les flux varient selon qu'il s'agit de données ponctuelles (rendement des cultures, reliquats N...) ou issues de monitoring (données GES, capteurs TDR et température, données météo) :

1. Acquisition de la donnée sur matériel spécifique de terrain ou par passage d'échantillon sur appareillage ou par analyse de technicien.
2. Valeur brute stockée sur ordinateur liée au matériel d'acquisition ou sur ordinateur de l'agent ayant réalisé l'analyse, ou directement sur le serveur renk de l'UMR SAS.
3. Valeur validée stockée sur le serveur renk de l'UMR SAS.
4. Traitement du fichier pour sa mise au format BDD (SI-PRO et dataverse) sur poste individuel relié au réseau INRAE.
5. Dépôt et publication des fichiers en BDD (SI-PRO et dataverse) après validation complète des données selon une fréquence non définie. Les données du SI sont stockées sur les datacenters INRAE.

Quelle est la volumétrie actuelle et prévisionnelle ?

La volumétrie des données se répartit de la manière suivante dans les deux espaces de stockage :

- **Espace serveur renk de l'UMR SAS** : 55.5 Go déclinés en 39 399 fichiers répartis dans 2061 dossiers
- **SI-PRO (espace EFELE) :**

- Essai_TS_MO : Propriété des sols : 0.487 Mo / Propriété des plantes : 0.109 Mo / Propriété des PRO : 0 Mo / Itinéraires techniques : 0 Mo

- Essai PROs : Propriété des sols : 0.794 Mo / Propriété des plantes : 0.231 Mo / Propriété des PRO : 0.264 Mo / Itinéraires techniques : 0 Mo

- **Volumétrie des données brutes GES** : 4.7 Go

L'augmentation de la production de données à l'avenir n'est pas définie en terme de volumétrie. Les capacités de stockage au niveau du SI-PRO et du serveur de l'UMR SAS devraient être suffisantes.

L'entité hébergeant physiquement les données a-t-elle une politique de sécurité pour son système d'information ?
politique locale, charte des infrastructures de recherche...

[Charte des infrastructures de recherche à l'Inra](#)

La sécurité du système d'information est gérée au niveau des infrastructures nationales d'INRAE (datacenters de Toulouse et Ile de France).

Sécurité - Confidentialité : les données font-elles l'objet d'échange ou de partage avec de tiers acteurs et selon quelles modalités ? comment sont déterminés les droits d'accès aux données avant leur publication ?

Les données sont confidentielles tant qu'elles n'ont pas été validées par les responsables scientifiques et techniques du site.

L'UMR SAS appartenant à un organisme de recherche public (INRAE), les données qu'elle produit sont des données FAIR et entrent dans le cadre de la Science Ouverte.

Toute personne ou scientifique autre que les équipes en charge du suivi du site a accès aux données après leur validation et autorisation par les propriétaires des données.

Sécurité - Intégrité - Tracabilité : Quelles sont les mesures de protection mises en œuvre pour suivre la production et l'analyse des données ?

Toutes les interventions sur le terrain sont répertoriées sur un fichier de suivi au format excel par les gestionnaires de site.

Les échantillons sont gérées selon la procédure propre à chaque prélèvement (voir Organisation et documentation des données dans

ce PGD).

Chaque fournisseur de données valide sa production et assure la fiabilité des données qu'il fournit ; les fournisseurs de données sont validés et autorisés par les responsables scientifiques du site.

Archivage et conservation des données

Quelles sont les données à conserver sur le moyen ou le long terme et quelles sont les données à détruire ?

L'intégralité des données (brutes ou travaillées) produites par le dispositif EFELE sont conservées indéfiniment dans les espaces de stockage dédiés, depuis la mise en service de la plateforme.

Sur quelle plateforme d'archivage pérenne seront archivées les données à conserver sur le long terme ? Sinon, quelles procédures seront mises en place pour la conservation à long terme ?

Les données et métadonnées du SI-PRO sont archivées sur le long terme et répliquées sur les 2 datacenter d'INRAE (Toulouse et Ile de France).

La politique de l'INRAE permet le partage de données à long terme sur des entrepôts sécurisés. Dans le cas du SOERE PRO, la publication sur des entrepôts de données comme Recherche Data Gouv est prévue dans un futur relativement proche (1er semestre 2025) pour permettre l'accessibilité et le réemploi des jeux de données.

Quelle est la durée de conservation des données ?

EFELE est une composante du SOERE PRO, au niveau duquel il a été décidé de conserver toutes les données sans limite de durée, du fait qu'il s'agit d'un observatoire ayant pour objectif d'acquérir des données sur le long terme.

Quelles garanties de financements couvriront les coûts associés à la conservation à long terme ?

Les coûts associés à la conservation à long terme des données sont financés à la fois par les fonds institutionnels (INRAE) et plus précisément par la Direction Systèmes d'Information (DSI), mais aussi par les fonds propres du SOERE PRO, de l'infrastructure AnaEE France.

Dans le cadre d'un dépôt long terme sur un entrepôt de données type Recherche data gouv, celui-ci est financé par INRAE. L'hébergement sécurisé et la mise à disposition des données sont garantis pendant un minimum de 5 ans renouvelables après la publication. Cette durée constitue un plancher, il ne s'agit pas de la durée maximum de conservation des données qui peut varier selon les jeux de données.