

Notice de valorisation des données issues d'Agrosyst

Dernière mise à jour : 19/02/2025

Auteurs

Bérenger Vuittenez, Lucie Baude, Thibault Peyrard, Thomas Badie

Introduction

Agrosyst est un système d'information (SI) géré par l'INRAE et alimenté par des saisies d'ingénieurs du réseau DEPHY dans le cadre du suivi des plus de 2000 fermes et 200 dispositifs expérimentaux du réseau. C'est un SI permettant de saisir les données agronomiques des systèmes de cultures comme les rotations, les itinéraires techniques, le contexte environnemental, les interventions et les utilisations d'intrants ; mais aussi économiques et stratégiques. Le réseau DEPHY est un réseau de fermes ayant pour objectif principal de réduire l'utilisation de pesticides tout en gardant un système viable économiquement.

Les données sont accessibles via le site internet [Datagrosyst](#) dans l'onglet "Exports personnalisés". Cette notice résume les idées clés pour valoriser ces données dans de bonnes conditions.

Table des matières

1) Définitions.....	3
2) Présentation de l'interface de datagrosyst.....	4
3) Ressources supplémentaires.....	6
A. Agrosyst bac à sable anonymisé pour les valorisateurs.....	6
B. Github de partage de code.....	6
D. Schéma des relations entre les tables.....	6
E. Guide des indicateurs Agrosyst.....	6
F. Guide utilisateur Agrosyst.....	7
G. Notice des magasins.....	7
4) Nomenclature et conventions de nommage.....	8
A. Nomenclature utilisée dans les tables.....	8
B. Nomenclature utilisée dans les colonnes.....	8
5) Les outils.....	9
Nettoyage.....	9
Agregation.....	9
Restructuration.....	9
Indicateur.....	10
Interopérabilité (À venir).....	10
6) Les magasins de données.....	10
Magasin CAN :.....	10
7) Divers.....	11
A. Concepts métiers d'Agrosyst.....	11
B. DEPHY FERME, DEPHY EXPE et hors DEPHY.....	12
C. Notion de local à intrants, à cultures et à matériels.....	13
D. Utilisations des tables "composants".....	14
E. Notions de données de performances.....	14
F. Utilisation des données en synthétisé.....	15
G. Identification des états initiaux des systèmes (pz0).....	16
8) FAQ.....	17

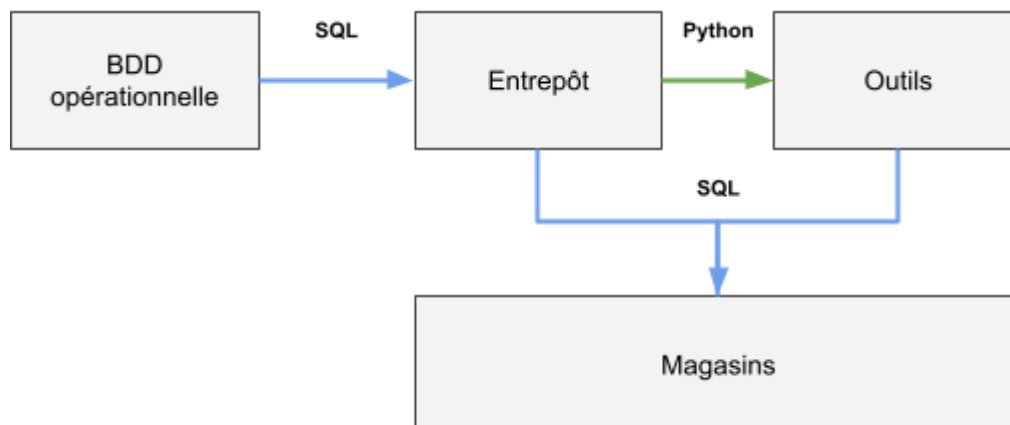
1) Définitions

Entrepôt : l'entrepôt de données contient toutes les données issues d'Agrosyst. Sa structure est très similaire à la base de données opérationnelle d'Agrosyst (données brutes saisies, référentiels et calculs d'indicateurs de performances). Chaque table représente ainsi une entité métier aux responsabilités bien définies : il n'y a pas de duplication d'information à travers les tables. L'entrepôt de données est le socle commun de tout ce qui suit.

Outils : les outils consistent en une série de tables élaborées à partir des données de l'entrepôt. Elles ne sont pas valorisables immédiatement et doivent s'utiliser conjointement avec l'entrepôt.

Magasin : un magasin de données est une restructuration et ou un nettoyage des données d'Agrosyst, conçue pour un usage spécifique. Les magasins de données sont élaborés à partir de l'entrepôt et des outils.

À titre indicatif, la figure 1 représente le processus utilisé pour générer les données disponibles sur **Datagrosyst** :



2) Présentation de l'interface de datagrosyst

L'interface d'exports personnalisés comporte 6 sections :

- **Source**

Dans cette section, l'utilisateur peut choisir quelle source de données utiliser pour son export. Une définition de chacune des sources disponible est consultable dans la partie [définitions](#). La source proposée par défaut est l'entrepôt.

- **Catégories**

Dans cette section, mise à jour lors de la sélection d'une source, l'utilisateur a accès à une classification des différentes tables de la source considérée.

- **Explications**

Dans cette section, mise à jour lors de la sélection d'une source, l'utilisateur a accès à une description succincte de la source et des conditions d'obtention des données

- **Sélection des tables**

Dans cette section, mise à jour lors de la sélection d'une catégorie, l'utilisateur a accès à la liste des tables classées dans la catégorie sélectionnée. L'utilisateur peut ajouter une table dans son export en cliquant sur la case contenant le nom de la table. Il peut cliquer à nouveau pour retirer la table de son export.

- **Documentation**

Dans cette section, mise à jour lors de la sélection d'une table, l'utilisateur peut accéder à une brève description de la table ainsi qu'au détail de toutes les colonnes qui la composent. Les clés primaires et les clés étrangères apparaissent en gras. **Attention : n'importe quel utilisateur a accès à la documentation complète de la table, y compris à la description des champs sensibles qui ne seront pas forcément disponibles à l'export.**

- **Synthèse**

Dans cette section, mise à jour lors de la sélection d'une table, l'utilisateur peut consulter toutes les tables qui seront ajoutées à son export.

Cette section a également quatre boutons :

- "utiliser la dernière configuration" : permet d'utiliser la sélection des tables effectuée précédemment sur le compte actif.

- "Ajouter la documentation" : permet d'ajouter dans l'export la documentation affichée dans la section "Documentation", pour toutes les tables et colonnes existantes (toutes sources confondues)
- "Ajouter la notice" : permet d'ajouter dans l'export la notice de valorisation.
- Soumettre permet de déclencher l'export des tables visible dans la synthèse. L'utilisateur recevra un mail lui indiquant que l'export est prêt à être télécharger en cliquant sur un lien.

Remarque : l'utilisateur peut exporter en une fois des tables issues de catégories et de sources différentes.

3) Ressources supplémentaires

A. Agrosyst bac à sable anonymisé pour les valorisateurs

Afin de comprendre les conditions d'obtention des données d'Agrosyst, une version de l'outil de saisie avec les données anonymisées est disponible depuis votre requête (disponible dans l'onglet profil). Pour vous connecter, utilisez les identifiants transmis dans cette même requête dans la section "valorisation".

Certaines entités sont accessibles directement via l'url. Par exemple, pour visualiser dans l'interface une zone particulière de zone_id = *fr.inra.agrosyst.api.entities.Zone_xxx*

Il suffit de copier-coller dans le navigateur l'url suivant :

https://agrosyst.fr/ags-valorisation/effective/effective-crop-cycles-edit-input.action?zoneTopiaId=fr.inra.agrosyst.api.entities.Zone_xxx

B. Github de partage de code

L'ensemble des scripts utilisés pour l'élaboration des magasins de données ainsi que plusieurs scripts d'exemples de prise en main des données sont disponibles à travers le lien github : https://github.com/beren2/catalogue_script_agrosyst

Toute participation à ce git est la bienvenue si vous pensez que vos travaux peuvent être utiles au reste de la communauté scientifique travaillant sur les données Agrosyst

C. DEPHYGraph

Il est possible d'accéder à certaines informations générales sur les données à partir de *DEPHYGraph*: (par exemple, pour connaître pour chaque campagne la répartition du Nombre d'années dans le réseau : https://dephygraph.fr/#/dashboard?y=c104_campaign_dis&x=c103_networkYears)

D. Schéma des relations entre les tables

Dans la section "documentation" (cf [Présentation de l'interface](#)), vous pouvez accéder au schéma relationnel des différentes tables de l'entrepôt de données.

E. Guide des indicateurs Agrosyst

Le guide des indicateurs Agrosyst rassemble les informations concernant la plupart des indicateurs disponibles sur Agrosyst nativement (dans la catégorie performance de l'entrepôt). Celui-ci est disponible depuis l'interface Agrosyst (en haut à droite sur le "?").

F. Guide utilisateur Agrosyst

Le guide utilisateur Agrosyst peut être un bon moyen de comprendre comment saisir sur Agrosyst. Celui-ci est disponible depuis l'interface Agrosyst (en haut à droite sur le "?").

G. Notice des magasins

Chaque magasin de données dispose d'une notice de valorisation qui lui est propre, explicitant son objectif et les hypothèses sous-jacentes à sa création.

4) Nomenclature et conventions de nommage

A. Nomenclature utilisée dans les tables

xxx_realise: Table qui concerne les données saisies dans le mode “réalisé”

xxx_synthetise : Table qui concerne les données saisies dans le mode “synthétisé”

B. Nomenclature utilisée dans les colonnes

Il est important de comprendre que toutes les entités dans Agrosyst sont millésimées (i.e associé à une année) et identifiées par un id. Pour avoir un suivi d’une entité dans le temps (ex : domaine, culture, sdc ...), il faut utiliser les codes (cf colonne **code**)

id : identifiant d’une entité, clé primaire de la table considérée.

xxx_id : clé étrangère, identifiant de l’entité de la table **xxx**.

code : attribut non millésimé permettant de faire le lien entre les différentes campagnes. (présent pour certaines entités seulement)

Par exemple : si une exploitation saisit sur Agrosyst pendant 3 campagnes, on pourrait trouver 3 lignes dans la table *domaine*, avec trois id différents mais la colonne “code” serait la même.

xxx_code: clé étrangère, code de l’entité de la table **xxx**.

Le recours à ce genre de clé étrangère code est commun lorsqu’on travaille sur le mode de saisie synthétisé.

Il est cependant fréquent que l’usage de cette clé nous conduise à plusieurs entités dans la table cible. Pour choisir la bonne ligne parmi les codes disponibles, on peut alors utiliser l’année du domaine associé au synthétisé en question. Ce travail d’association a été fait au sein du magasin de données nettoyage > catégorie “restructuration”.

5) Les outils

Nettoyage

Les outils de **nettoyage** proposent des tables à différentes échelles issues de scripts qualifiant les données. En utilisant ces tables, l'utilisateur peut, avec un join et un filtre sur les valeurs de ces nouvelles colonnes, nettoyer ses données.

Afin de garantir une liberté maximale à la configuration des filtres utilisés, chaque test donne lieu à une colonne propre.

Par exemple : un utilisateur travaillant en réalisé sur l'échelle intervention peut utiliser la table `intervention_realise_nettoyage` pour tagger toutes les interventions réalisées qui ont une dose de référence aberrante ou un débit de chantier aberrant. La documentation de la colonne permet de comprendre les conditions permettant de conclure que l'entité considérée est aberrante. En général, il s'agit de comparaison avec des valeurs à dire d'expert. Des informations complémentaires sont données sur le github (cf [Ressource supplémentaires](#)).

Agregation

Les outils d'**Agrégation** permettent de sauter rapidement d'une entité à une autre. Ces tables ne contiennent que des identifiants. Elle sont particulièrement utiles lorsqu'on souhaite travailler à une échelle fine (l'utilisation d'intrant par exemple) et néanmoins obtenir des informations provenant d'échelle haute (la filière dans le système de culture par exemple)

Par exemple : permet de savoir immédiatement dans quel domaine a lieu une utilisation d'intrant grâce à la table `utilisation_intrant_realise_agrege`.

Restructuration

Les tables de la catégorie **Restructuration** viennent faciliter la manipulation de certaines entités. Ces tables consistent pour la plupart en une qualification des entités en synthétisé et permettent de ne pas avoir à travailler avec les variable `_code` (cf [Nomenclature utilisée dans les colonnes](#)).

À noter que cette catégorie contient d'autres outils plus spécifiques comme `noeuds_realise_restructure` qui fournit, pour chaque nœud en réalisé, le nœud précédent.

Par exemple : la table `noeuds_synthetise_restructure` permet de connaître, pour chaque `noeuds_synthetisé`, la `culture_id` mobilisée dans le `noeuds_synthetise`.

Indicateur

Les tables de la catégorie **Indicateurs** proposent des informations complémentaires sur certaines tables de l'entrepôt, notamment des opérations complexes à réaliser, comme par exemple la récupération de la dose de référence à la cible non millésimé à l'échelle de l'utilisation d'intrants.

Interopérabilité (*À venir*)

Les tables de la catégorie **Interopérabilité** proposent des clés vers d'autres bases de données.

Par exemple : une table `domaine_localisation` permettra de connaître la maille safran d'un `domaine_id` considéré.

6) Les magasins de données

Un magasin de données est une restructuration et/ou un nettoyage des données Agrosyst, conçu pour un usage spécifique. Un magasin de données peut contenir des données brutes, mais aussi des données calculées. Tous les magasins de données sur Datagrosyst sont conçus à partir des données de l'**entrepôt** et sont accessibles via la section **Source** (cf [Présentation de l'interface](#)). Les scripts d'obtention de ces magasins sont accessibles sur le git. L'url exact est indiqué dans la section **Explications** (cf [Présentation de l'interface](#)).

La présente partie s'attarde à décrire les différents magasins de données actuellement disponibles au sein de Datagrosyst. Comme évoqué précédemment, chaque magasin fait l'objet d'une documentation dédiée (cf [Ressources supplémentaires](#))

Magasin CAN :

Le **Magasin CAN** correspond au magasin de données anciennement appelé "**Exports en masse**" (modulo les renommages de tables et quelques corrections étayées ici). Ce dernier a été conçu conjointement avec la Cellule d'Animation Nationale (CAN) du réseau DEPHY et sert à la mise à jour des données internes du réseau.

7) Divers

A. Concepts métiers d'Agrosyst

Les paragraphes suivants permettent de définir les concepts fondamentaux utilisés dans Agrosyst. Les définitions sont issues du guide utilisateur, disponible sur l'interface bac à sable disponible depuis votre requête Datagrosyst (écrou en haut à droite).

Un **système de culture** est défini, pour une portion d'espace traitée de manière homogène (par ex. : un groupe de parcelles), par les cultures qui y sont pratiquées, leur ordre de succession et les itinéraires techniques (i.e. les interventions culturales) mis en œuvre pour chacune d'elles.

Dans Agrosyst, un **domaine** peut contenir un ou plusieurs systèmes de culture regroupés dans un ou plusieurs dispositifs. La description du système technique du système de culture s'effectue avec les **actes réalisés** et/ou des **actes synthétisés**. La description du système décisionnel s'effectue avec le renseignement du **modèle décisionnel** et des règles de décision.

Les **actes réalisés** permettent de décrire les événements qui se sont déroulés sur les parcelles d'un système de culture pour une campagne donnée : culture(s) présente(s), interventions culturales, mesures et observations.

Les **actes synthétisés** permettent de décrire une synthèse des événements se déroulant au sein du système de culture pour une campagne (il s'agit alors d'un « synthétisé annuel ») ou pour plusieurs campagnes (il s'agit alors d'un « synthétisé pluriannuel »).

Cette séparation en deux modes de saisie (réalisé et synthétisé) explique que dans l'entrepôt, à partir d'une certaine échelle, on ai systématiquement le suffixe correspondant au mode de saisie dans le nom des tables.

Un valorisateur peut choisir de ne valoriser que le réalisé ou que le synthétisé, mais ceci aura un impact sur le nombre d'informations disponibles. La plupart des filières (viticulture, arboriculture, maraîchage, horticulture, cultures tropicales) privilégient le mode synthétisé pour sa simplicité à la saisie. La filière grandes cultures / polyculture-élevage autorise les deux modes de saisie, le mode réalisé permettant en particulier l'interopérabilité avec d'autres outils à travers eDaplos.

Pour se donner une idée de la répartition au sein des modes de saisies, on peut regarder la répartition des filières ([Figure 1](#)) .

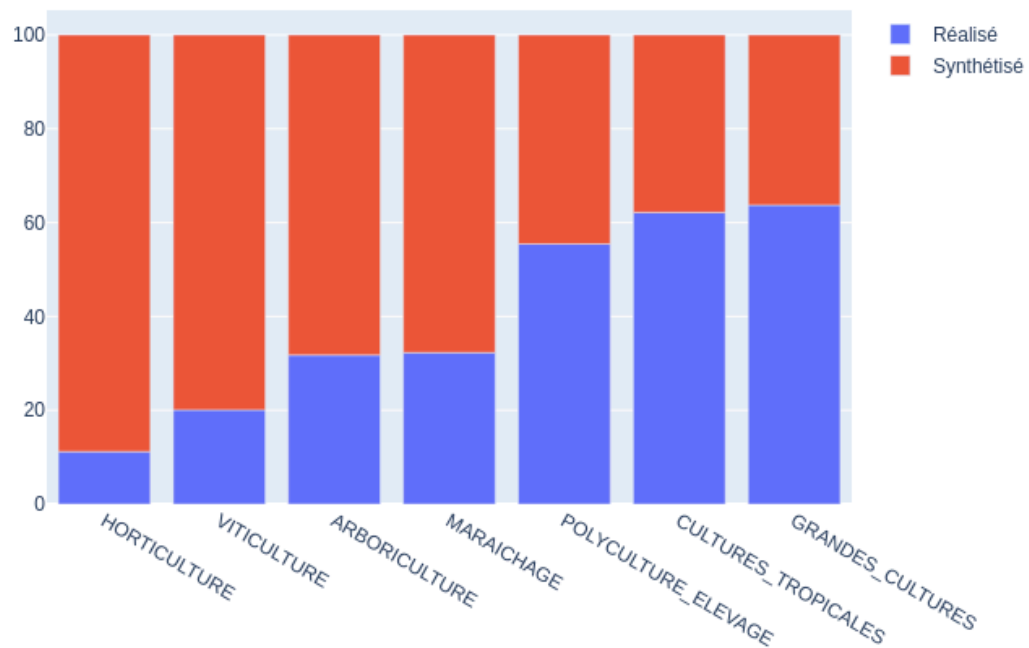


Figure 1 : Pourcentage de système de culture dans chaque mode de saisie pour chaque filière

B. DEPHY FERME, DEPHY EXPE et hors DEPHY

Attention, Agrosyst ne contient pas uniquement les données du dispositif DEPHY FERME ! L'entrepôt contient également les données du dispositif DEPHY EXPE ainsi que des systèmes de cultures issus d'autres réseaux. Cette information est disponible dans la table dispositif. Il faut être vigilant en valorisant conjointement ces données d'origine différentes car les consignes de saisie peuvent beaucoup varier.

C. Notion de local à intrants, à cultures et à matériels

Tous les intrants et toutes les cultures mobilisés dans les itinéraires techniques (ITK) sont déclarés au niveau du domaine. De même, les combinaisons d'outils mobilisées dans les interventions sont déclarées au niveau du domaine.

Du point de vue des données, cela signifie que la description des informations est complétée d'une clé étrangère pointant vers le domaine. On mobilise par la suite ses entités en faisant référence à leur clé primaire dans les ITK. Dit autrement, les intrants et les cultures constituent une boîte à outils dans laquelle l'utilisateur va piocher. Il est donc important de distinguer les tables où on déclare les informations (ex : intrant) des tables où on les mobilise (ex : utilisation_intrant_realise)

Par exemple : Un utilisateur peut créer une nouvelle culture qu'il nomme "Blé" dans son domaine. Cette culture sera alors disponible à l'utilisation dans toutes ces zones (en réalisé) ou dans son système synthétisé. De la même façon, il peut ajouter un intrant de type "produit phytosanitaire" et l'utiliser ensuite lors d'une intervention. Enfin, il peut créer une combinaison d'outils puis la sélectionner sur l'intervention.

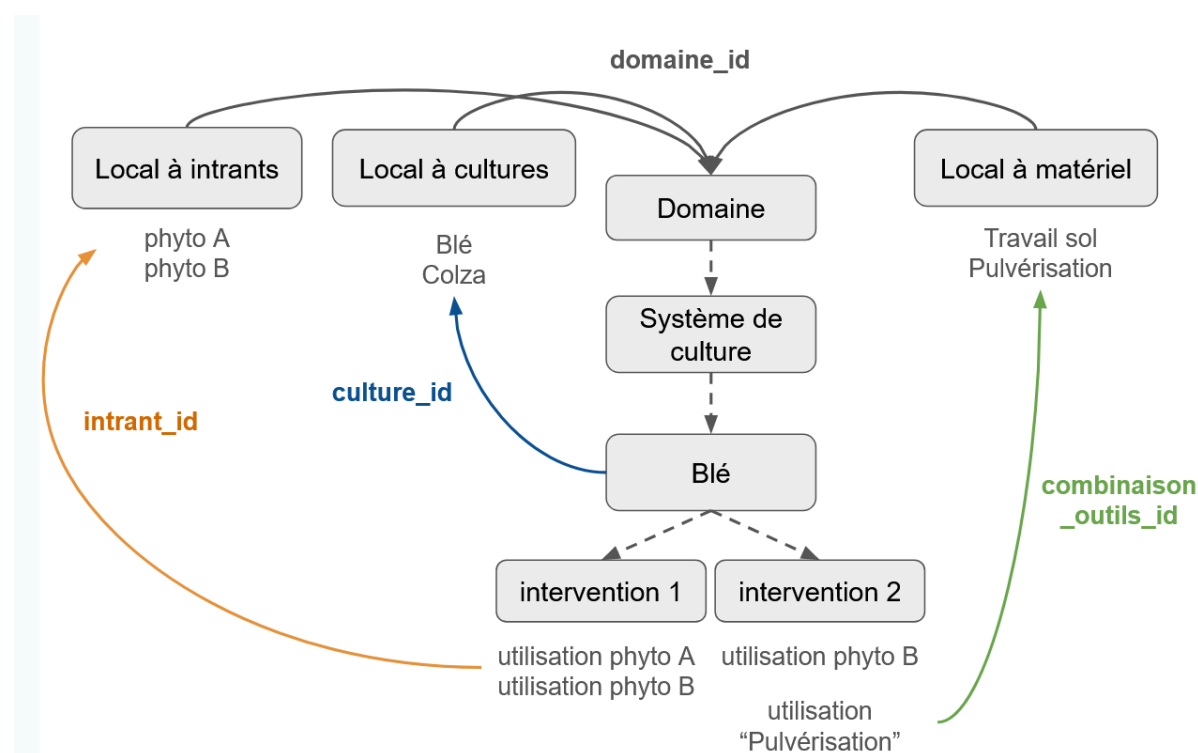


Figure 2 : Schéma de l'architecture des tables du local à intrant, à culture et à matériel. Attention, dans cette figure, le lien entre le système de culture et la sole de blé est un raccourci. En pratique, ce lien n'est pas directement accessible (dépendant du mode de saisie, soit dans le système synthétisé, soit dans la zone)

D. Utilisations des tables “composants”

La table composant_culture et la table combinaison_outil_materiel ont un comportement similaire. Elles informent de ce qui constitue l'entité mère.

Par exemple, la table composant_culture dispose d'une ligne par composant constituant une culture (cas des mélanges d'espèces, des mélanges de variété...). On peut reconstituer tous les composants d'une culture en ne sélectionnant que les composants avec le bon culture_id.

E. Notions de données de performances

Dans la section performance de la source entrepôt, les données disponibles sont calculées à partir des données brutes saisies. Les différentes modalités de calcul de ces indicateurs sont disponibles dans le guide des indicateurs (cf [Guide des indicateurs Agrosyst](#)).

Pour résumer :

- Les indicateurs sont calculés pour plusieurs échelles. Selon l'indicateur, la plus basse échelle est différente ([tableau 1](#))

Tableau 1 : Echelles de calculs des indicateurs

	Intrants	Intervention	Culture	Zone	Parcelle	Sdc	Synthétisé
Indice de Fréquence de Traitement	x	x	x	x	x	x	x
Charges opérationnelles	x	x	x	x	x	x	x
Charge de mécanisation		x	x	x	x	x	x
Coût du carburant		x	x	x	x	x	x
Temps d'utilisation du matériel		x	x	x	x	x	x
Temps de travail manuel		x	x	x	x	x	x
Charges de main d'oeuvre		x	x	x	x	x	x
Charges de main d'oeuvre manuelle		x	x	x	x	x	x
Produit brute			x	x	x	x	x
Marge brute			x	x	x	x	x
Marge semi-nette			x	x	x	x	x
Marge directe			x	x	x	x	x

- Les indicateurs économiques sont calculés en deux versions : **réel et standard**. La version standard utilise des prix issus de référentiels uniquement. La version réelle

calcule l'indicateur avec le prix réel, c'est-à-dire le prix qui a été saisi par l'ingénieur. Si le prix de l'entité n'est pas saisi, le prix de référence est utilisé à la place.

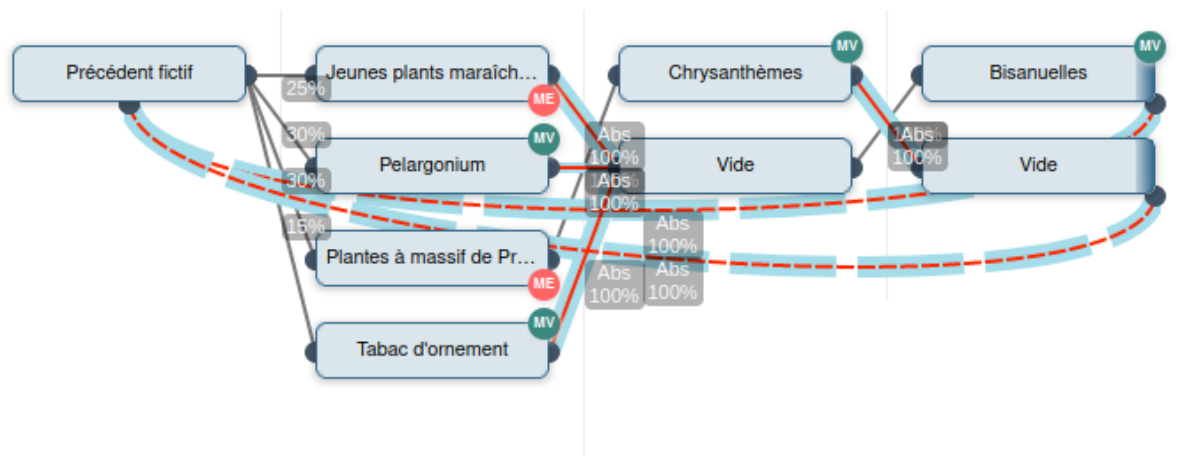
- Les IFT à la cible et à la culture sont disponibles en versions millésimée et non-millésimée. La version millésimée utilisera les doses homologuées en vigueur l'année du traitement, tandis que la version non-millésimée n'utilisera que les doses homologuées les plus récentes. L'IFT le plus communément utilisé pour valoriser les données est **l'IFT à la cible non-millésimé**. L'IFT dit "à l'ancienne" est l'IFT historiquement calculé par Agrosyst avant 2020 et avant que la méthode ministérielle ne soit adoptée. Les référentiels liés ne sont plus maintenus et son utilisation est déconseillée pour de nouveaux projets de valorisation.
- De nombreux indicateurs sont disponibles en version "**hors traitement de semence (Hts)**". Ceci est dû au fait que les saisies des traitements de semence ne sont pas homogènes. Il y a plus de chances que le fait qu'il n'y ait pas de traitement de semence soit un oubli au lieu d'une réelle absence de traitement : il est donc déconseillé de les prendre en compte dans des analyses portant sur l'ensemble du jeu de données.

F. Utilisation des données en synthétisé

L'appropriation des données en synthétisé pour les cultures assolées suppose de bien comprendre les rotations. Lorsqu'un utilisateur mobilise une nouvelle culture depuis sa boîte outils culture, celle-ci doit être reliée aux autres par des connexions. Le poids de la connexion informe sur le % de l'assolement associé à cette culture.

C'est ensuite sur ces connexions, représentant les couples precedent-culture, que les interventions ont lieu.

Par exemple : pour connaître les interventions qui ont eu lieu sur la culture “Chrysanthèmes” avec pour précédent cultural “Plantes à massif de Pr...” dans la figure suivante, il faut connaître l’identifiant de la connexion qui a pour destination “Chrysanthèmes” et pour source “Plantes à massif de Pr...”. On peut alors trouver toutes les interventions qui ont cet identifiant pour clé étrangère dans “connexion_synthese_id”.



L'exemple ci-dessus est particulièrement intéressant puisqu'on y voit une astuce de saisie fréquemment utilisée dans les schémas de rotations des systèmes synthétisés : le précédent fictif. Cette astuce permet de faciliter la saisie en permettant de commencer la rotation par un assolement complexe en rang 2. Normalement, toutes les transitions pointant vers le noeuds "précédent_fictif" sont déclarées absentes dans la rotation (c'est à dire que les couples de culture x précédent_fictif sont absentes dans la rotation) et grâce aux transitions sortantes, on peut facilement représenter les % des différentes soles. En n'ayant pas recours à cette petite astuce, il serait compliqué de trouver les bons pourcentages car il faudrait tenir compte des poids des liens de retour des derniers rangs vers les premiers. Cette méthode de saisie est particulièrement mobilisée pour des systèmes complexes et/ou très diversifiés dans la filière Maraîchage.

G. Bilans de campagnes

Il existe deux types de bilans de campagnes :

- bilans de campagne régionaux : pression en bioagresseur (maladies et ravageurs) de la région et bilan de la pluviométrie de l'année
- bilans de campagne à l'échelle du système de culture : pour un système de culture donnée, les pressions et les maîtrises en bioagresseurs (maladies, ravageurs et adventices) ainsi que la maîtrise de la verse, l'alimentation hydrique et minérale et le rendement et qualité des cultures

Précisions sur les bilans de campagne du système de culture.

Attention : il existe des spécificités selon :

- les filières : certaines filières ont des variables spécifiques ou bien prennent des valeurs différentes. C'est pour cela que les tables concernant la maîtrise des bioagresseurs sont catégorisées selon "cultures assolées", "arboriculture" et "viticulture". Les tables concernant le rendement et l'alimentation hydrique et minérale concernent l'ensemble des filières
- le type de dispositif : le dispositif DEPHY-EXPE a certaines variables différentes qui sont identifiables par un suffixe _EXPE, mais ne font pas l'objet de tables séparées

Pour comprendre les bilans de campagnes à l'échelle du système de culture, il est conseillé de consulter l'outil de saisie (cf [Agrosyst bac à sable anonymisé pour les valorisateurs](#)).

Les bilans de campagnes sdc sont composés de 6 catégories de maîtrise : adventices, maladies, ravageurs, verse, alimentation hydrique et minérale et bilan du rendement.

Lorsque les variables demandées sont les mêmes entre la maîtrise des adventices, maladies et ravageur, les données ont été regroupées dans la même table sous le nom de "maîtrise des bioagresseurs"

Figure 3 displays two screenshots of the Agrosyst tool interface for disease management data entry.

A) Maîtrise des maladies et batérioroses

1. * Quelles cultures concernées ? (1)
- [X] CLEMENTINE COMMUNE
- [] Oranges
- [] Pomelos

2. Affiner la déclaration par espèces/variétés ? (Cliquer pour cacher)
- [X] Clémentinier (2)

3. * Maladie(s) pour ces cultures
- / Gommose parasitaire (3)
- / Acariose (4)

4. Quel niveau d'utilisation de pesticides ? (3)
- Nombre de traitements : 3
- IFT-Fongicide (chimique) : 3
- IFT-fongicide biocontrôle : 0

B) Maladie ciblée (4)

Pression avant intervention
* au moins un des deux

Groupe cible : [dropdown]
Maladie : Gommose parasitaire [dropdown]
Inoculum année(s) précédente(s) : Moyen [dropdown]
Evolution par rapport à l'année précédente : Pression identique à l'année précédente [dropdown]
Echelle de pression : [dropdown]

Résultats obtenus, niveau de maîtrise finale
* Échelle de maîtrise : Symptômes sans effet sur le rendement ou la [dropdown]
% parcelles touchées : > à 50% [dropdown]
% arbres touchés : 2 à 10% [dropdown]

Figure 3 : Exemple des écrans de saisie de la maîtrise des maladies. A) correspond à l'écran de maîtrise globale des maladies pour un ensemble de culture. B) correspond à l'écran de maîtrise d'une maladie spécifique

Les données renseignées dans chaque catégorie ont la même architecture. Un exemple d'écran de saisie est montré [figure 3](#), et un exemple d'architecture entre les tables est montré [figure 4](#). Pour résumer il y a :

- une table de lien entre la maîtrise globale et les cultures [figure 3](#) (1)
- une table de lien entre la maîtrise globale et les composants de cultures espèces et variétés [figure 3](#) (2)

- une table de maîtrise globale [figure 3](#) (3)
- une table de maîtrise spécifique à un bioagresseur [figure 3](#) (4)

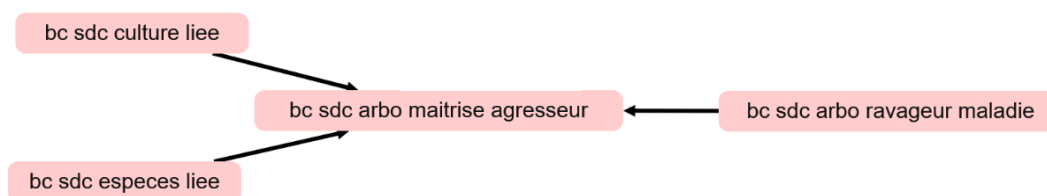


Figure 4 : Liens entre les tables de la maîtrise des ravageurs et maladies pour l'arboriculture

H. Identification des états initiaux des systèmes de culture (pz0)

Définition

Les états initiaux, aussi appelés points zéro (pz0), qualifient les pratiques culturelles du système de culture AVANT son entrée dans le réseau DEPHY FERME. Un système de culture peut être suivi dans le temps grâce au code_dephy. Il existe donc 1 état initial pour 1 code_dephy.

Les consignes de saisie des états initiaux sont de renseigner les pratiques culturelles moyennes avant l'entrée dans le réseau dans un système synthétisé tri-annuel de campagne du domaine année entrée réseau -1. Une spécificité pour la filière viticulture où les pratiques sont souvent très variables d'une année à une autre, la consigne de saisie devient de saisir trois synthétisés monoannuels.

Note : La plupart des pz0 sont donc des synthétisés pluriannuels ayant pour campagne 2009,2010,2011 ou 2014,2015,2016 ou 2019,2020,2021. Ces trois périodes correspondent aux années précédant les trois grandes vagues d'arrivée de nouveaux agriculteurs dans le réseau en 2012, 2017 et 2022.

Cette consigne de saisie n'a pas été respectée à la lettre par tous les utilisateurs et il existe donc un grand nombre d'exceptions dans la base. Par exemple, la saisie des états initiaux de manière mono annuelle a été utilisée par d'autres filières.

Outil identification pz0

Le système d'information n'a pas de variable saisie par l'utilisateur décrivant si le système est un état initial ou bien un système post-pz0. C'est pourquoi un outil dans la catégorie "indicateur" permet d'identifier ces états initiaux parmi les données saisies.

La cellule référence DEPHY et l'équipe Agrosyst ont mis au point un système d'identification des états initiaux dans la base à partir de documents administratifs du réseau DEPHY (dates entrée dans le réseau et sortie) ainsi qu'avec les consignes de saisie des états initiaux.

Attention : l'outil permet d'identifier seulement les états initiaux des dispositifs DEPHY-FERME et un système de culture de modalité de suivi dephy détaillée.

En plus d'identifier les pz0, l'outil qualifie l'ensemble des données de la frise chronologique d'un système de culture avec 3 états : pz0 ou post, ou non attendue (avant ou après la période de l'exploitation) cf [Figure 5](#). Attention, l'outil ne permet pas de rendre compte des données de campagne attendues non saisies (pz0 ou post).



Figure 5 : Schéma d'une frise chronologique d'un système de culture selon le type de donnée attendue. NA : campagne non attendue, NS : non saisi, non détectable par l'outil.

L'outil "identification pz0" se structure de deux colonnes : *entité_id* et *pz0*. La colonne *entité_id* est soit l'identifiant unique d'un système synthétisé soit l'identifiant unique d'une zone. En plus des trois états décrits ci-dessus, la colonne *pz0* a d'autres modalités afin de qualifier les données / frises chronologiques non fiables. Toutes les modalités sont décrites ci dessous :

- pz0 : système avant l'entrée dans le réseaux dephy
- post : système au sein du réseaux dephy
- incorrect : campagne non attendue : zone ou synthétisé de campagne non attendue
- incorrect : chevauchement pz0 : zone ou synthétisé de campagne qui chevauche le pz0 identifié
- incorrect : aucun pz0 saisi : Lorsqu'aucun pz0 n'est détecté, l'ensemble des saisies (systèmes synthétisés et zones) liées au code_dephy, quelque soit la campagne, ont cette modalité attribuée.
- incorrect : saisie pz0 non acceptable : Lorsqu'une ou plusieurs donnée(s) de campagne pz0 ont été saisies mais s'éloignent trop des consignes de saisies pour être acceptable. L'ensemble des saisies (systèmes synthétisés et zones) liées au code_dephy, quelque soit la campagne, ont cette modalité attribuée.

- **incorrect : saisie de plusieurs pz0** : Lorsque plusieurs systèmes synthétisés peuvent correspondre au pz0 mais la qualité de la donnée ne nous permet pas de choisir un pz0. Ce cas de figure apparaît souvent lorsque un même code dephy a été utilisé pour plusieurs systèmes de cultures. On obtient donc plusieurs pz0. L'ensemble des saisies (systèmes synthétisés et zones) liées au code_dephy, quelque soit la campagne, ont cette modalité attribuée.
- **incorrect : code dephy inconnu** : Lorsque le code dephy saisi est inconnu du référentiel administratif, la date d'entrée et de sortie est donc inconnue. L'ensemble des saisies (systèmes synthétisés et zones) liées au code_dephy, quelque soit la campagne, ont cette modalité attribuée.

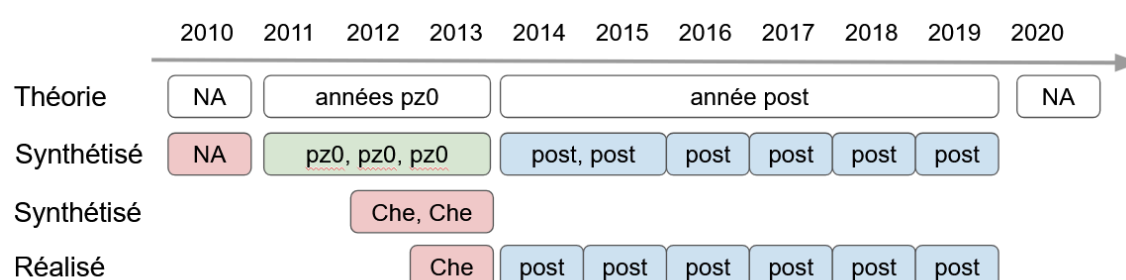


Figure 6 : Frise chronologique des données saisies possibles pour un même code dephy. pz0,pz0,pz0 : synthétisé triannuel pz0; post, post : synthétisé bi annuel post; NA : non attendu; Che : chevauchement avec pz0

Acceptabilité d'un pz0

Préalablement, seuls les **synthétisés et zones avec au minimum une intervention** ont été gardés.

Les seuls pz0 acceptés sont : (par ordre de qualité)

1. synthétisé tri annuel ayant les trois campagnes pz0
2. synthétisé tri annuel ayant un décalage de un an avant ou après cf code dephy n° 1 et 2 [figure 7](#)
3. 1 à 3 synthétisés mono annuel si et seulement si les campagnes des synthétisés ne se chevauchent pas

Il peut arriver que pour un même code dephy plusieurs synthétisés mentionnent une campagne pz0. Comme il ne peut avoir qu'un seul synthétisé par code dephy pz0, un choix est fait selon l'ordre de qualité ci-dessus. Si les deux saisies sont de qualité similaire, aucun choix ne peut être fait et l'ensemble de la frise est qualifié comme "plusieurs pz0 saisis" cf code dephy n°6 [figure 7](#).

Cas particulier des pz0 mono annuels

La saisie des états initiaux avec un système synthétisé mono annuel n'est pas la consigne de saisie sauf exception de la viticulture. Cependant, comme mentionné précédemment, d'autres filières ont saisi leurs pz0 de façon mono annuelle. Nous avons donc choisi de d'identifier pz0 les saisies mono annuelle toutes filières confondues si et seulement si :

- le nombre de synthétisés mono annuels est entre 1 et 3 (2 sont acceptés même si elles ne sont pas continues, ex : 2011 + 2013 cf [figure 7](#))
- si il n'y a pas plusieurs fois la même campagne

Attention : pour utiliser ces pz0, il faut reconstruire le pz0, c'est à dire moyenner les indicateurs pour obtenir les pratiques culturales moyennes d'avant l'entrée dans le réseau. Cela permet également de ne considérer qu'un seul point comme les autres saisies de pz0.

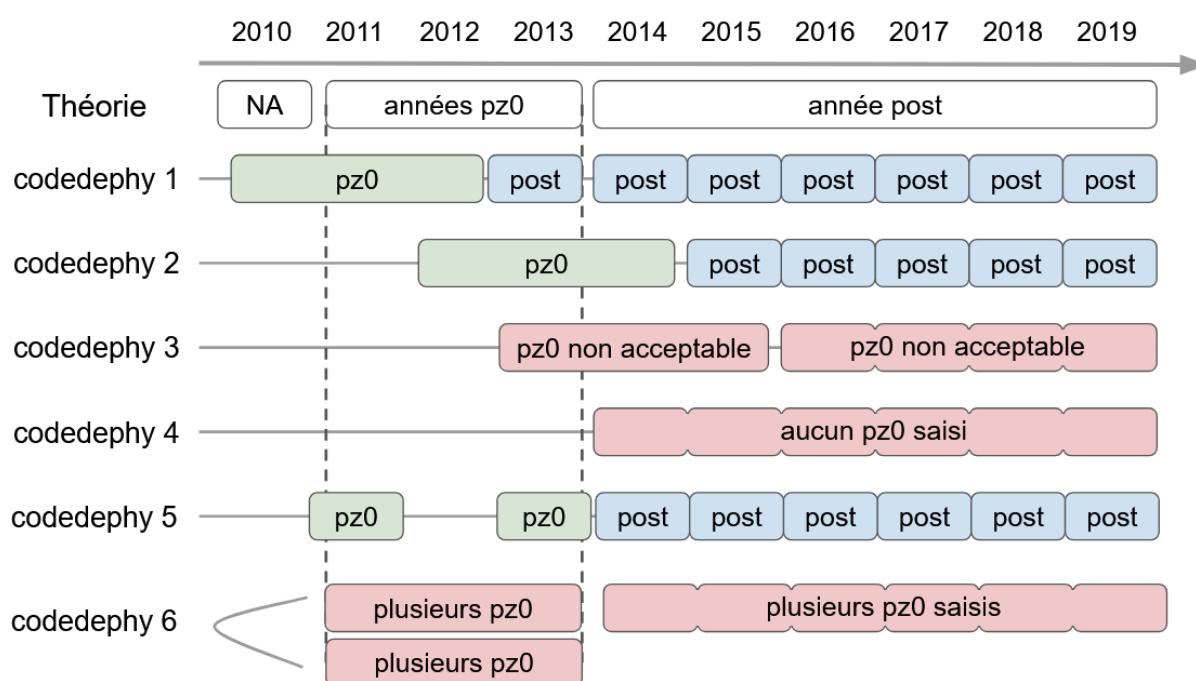


Figure 7 : Frises chronologiques de 6 codes dephy avec les mêmes années d'entrée et de sortie illustrant les pz0 acceptables et non acceptables.

8) FAQ

	Catégorie		
1	culture	Pour récupérer les informations sur la culture précédente , faut-il aller chercher les connexions dont le nœud de la culture présente est la cible ?	Tout à fait. Se souvenir qu'il peut y avoir plusieurs précédents en synthétisé.
2		La majorité des culture_code dans la table de noeuds_synthetise n'est pas présente dans la table composant_culture . Est-ce qu'il y a plusieurs types de codes ? Comment récupérer l'espèce, dans ce cas ?	Confusion probable entre le composant_culture_code (colonne code dans la table composant_culture) et le noeuds_synthetise_culture_code (colonne culture_code dans la table noeuds_synthetise). composant_culture_code est bien un code du composant de culture et non un code de culture --> impossible de les comparer. Le code de culture n'est pas présent dans la table composant_culture . En revanche on peut l'obtenir très facilement car pour un culture_id , on a un et un seul culture_code dans la table culture.
3		Souvent la culture précédente est un précédent fictif , qu'est ce que cela signifie ? A quoi correspond le rendement dans ce cas ?	<p>Le recours à un précédent fictif est une astuce de saisie spécifique au synthétisé (cf Utilisation des données en synthétisé). Cette astuce permet de faciliter la saisie du schéma de rotation en permettant de commencer la rotation par un assolement complexe.</p> <p>Normalement, toutes les connexions pointant vers le noeuds "précédent fictif" sont déclarées absentes dans la rotation (c'est à dire que les couples de culture x précédent fictif sont dans la rotation) et grâce aux transitions restantes, on peut facilement représenter les % des différentes soles.</p> <p>En n'effectuant pas cette astuce, c'est compliqué de trouver les bons pourcentages car il faudrait tenir compte des poids des liens de retour des derniers rangs vers les premiers.</p>
4		Est-ce que cela change quelque-chose que les cultures soient dans la "Même campagne agricole", ou non, car la définition "L'itinéraire technique décrit pour une culture commence dès l'instant où la culture précédente est récoltée (ou détruite) et s'arrête à la récolte de cette	Il peut en effet y avoir plusieurs cultures dans la même campagne agricole. En synthétisé, l'information se trouve dans la table noeuds_synthetise , colonne " memecampagne_noeudprecedent ". On aura dans tous les cas un ITK par culture.

		culture (ou destruction)" est toujours valable ?	ATTENTION : pour le maraîchage, la campagne agricole est définie comme une année civile (de janvier à décembre). Il s'agit là seulement d'une consigne de saisie et il n'y a donc pas de vérification particulière sur l'interface que les champs "memecampagne_noeudprecedent" sont cohérents.
5	Interventions	Pour sélectionner les interventions sur un nœud (= sur la culture que l'on veut étudier), faut-il utiliser source_noeuds_synthetise_id ou cible_noeuds_synthetise_id ?	Les interventions portent sur cible_noeuds_synthetise_id Pour s'en souvenir : disposer du précédent cultural est plus utile que de disposer du suivant cultural.
6		Je ne trouve pas l'information sur le nombre de passages pour les différentes interventions , est ce que c'est le freq_temporelle ?	S'en référer au guide utilisateur (notamment page 87 et page 104 pour les synthétisés) Fréquence temporelle : Nombre de fois où l'intervention a été réalisée divisée par le nombre de campagnes sur lesquelles porte la synthèse. En ce sens, la fréquence temporelle est l'équivalent du nombre de passage des réalisés en synthétisé.
7		Où trouver l'information sur le carburant (type et quantité) utilisé lors de chaque action ?	Un indicateur de performance existe à ce sujet dans la table intervention_realise_performance et intervention_synthetise_performance (conso_carburant)
8		J'ai observé qu'il existe des connexions pour lesquelles il n'y a pas d' intervention de récolte et semis mais uniquement des interventions qui peuvent correspondre à des interventions avant plantation. Est-ce que c'est normal / cela correspond à quelque chose dans les analyses, ou est-ce qu'il faut les enlever ?	Pour certains types de soles, cela peut s'expliquer. Par exemple, si il y a une prairie dans la rotation, utilisée pour du pâturage : il n'y a pas à proprement parler de semis ni de récoltes . Pour la majorité des cultures en revanche, si l'action de semis ou de récolte est absente, ce doit être une mauvaise saisie.
9	matériel	Pour utiliser la table context['combinaison_outil'] en merge sur les interventions, faut-il aussi utiliser le domaine_id sinon on ne sait pas exactement de quel outil il s'agit, ou peut-on approximer et prendre la première occurrence du combinaison_outil_code dans la table combinaison_outil ? Ou autre solution ? Avoir l'info domaine sur l'intervention est	Oui, c'est un problème récurrent avec l'utilisation du "_code" , on a plusieurs choix possibles lors des merges. Une bonne idée est de choisir, parmi les ids possibles, l'id associé à la campagne du domaine dans lequel le synthétisé est déclaré. Une table combinaison_outil_restructure sera bientôt disponible (dans la source Nettoyage) et permettra directement de disposer d'un combinaison_outil_id mobilisé au niveau d'une intervention.

		compliqué car elle n'est pas dans toutes les tables nœuds, interventions etc.	
10		Puissance en cc du matériel : uniquement contenu en champ texte dans les colonnes materiel_caracteristique ? (Semble être dans materiel_caracteristique4 pour les tracteurs, conversion facile en numérique)	Il était complexe de trouver des informations standardisées pour les matériels lors de la création des référentiels : Agrosyst a donc fait le choix d'utiliser des colonnes "caractéristiques". Les informations présentes dans ces colonnes dépendent du type de matériel considéré. Pour les tracteurs en effet, la puissance est disponible dans les colonnes " materiel_caracteristique3 " et " materiel_caracteristique4 ". Nous n'avons pas à notre disposition d'informations plus standardisées.
12	Divers	Est-ce que cela change quelque-chose que les cultures soient dans la "Même campagne agricole", ou non, car la définition "L'itinéraire technique décrit pour une culture commence dès l'instant où la culture précédente est récoltée (ou détruite) et s'arrête à la récolte de cette culture (ou destruction)" est toujours valable ?	Il peut en effet y avoir plusieurs cultures dans la même campagne agricole. En synthétisé, l'information se trouve dans la table noeuds_synthetise , colonne " memecampagne_noeudprecedent ". On aura dans tous les cas un ITK par culture. ATTENTION : pour le maraîchage, la campagne agricole est définie comme une année civile (de janvier à décembre). Il s'agit là seulement d'une consigne de saisie et il n'y a donc pas de vérification particulière sur l'interface que les champs " memecampagne_noeudprecedent " sont cohérents.
13		Est-on sûrs que les personnes en charge de la saisie ont bien rempli les quantités d'intrants en fonction de l'unité (exemple, à l'hectare) et non la valeur totale pour la surface cultivée ?	On ne peut pas avoir de certitude sur la validité de la saisie. Cependant la quantité d'intrant utilisée sert directement au calcul de l'IFT, principal indicateur du réseau DEPHY. De plus, l'interface propose depuis quelque temps déjà, lors de la saisie d'une dose, la dose de référence (à l'hectare). Cela ne garantit pas une donnée valide mais doit éviter beaucoup d'erreurs. La CAN effectue néanmoins des pré-traitements pour certaines valeurs aberrantes.
14		Comment pourrait-on faire pour garder seulement les 3 premières campagnes d'un SDC en synthétisé : l'identifiant du sdc est unique mais le terrain cultivé est le même. Comme un sdc n'est pas exactement égal à un domaine (même si dans la majeure partie des cas il y a une correspondance unique entre domaine et sdc), je ne vois pas comment faire	Tu peux utiliser la colonne " code " dans le synthétisé. Comme tous les autres codes, celui-ci te permet de lier temporellement plusieurs sdc_id
15		Pour le synthétisé, où trouver l'information sur la taille de la parcelle / la taille de la zone cultivée sur laquelle s'appliquent les informations ?	Le concept de parcelle n'existe pas en synthétisé. Il est possible d'obtenir la SAU du domaine dans lequel le synthétisé est déclaré. Par la suite, on peut calculer la superficie des cultures assolées en retranchant la surface imputable aux cultures pérennes (en %). Pour recalculer la surface de chacune des soles d'une rotation en synthétisé, un

		<p>script existe et est disponible sur le git catalogue_script_agrosyst. (https://github.com/beren2/catalogue_script_agrosyst/scripts/utils/get_surfaces_connections_synthetise.py)</p> <p>Si vous obtenez une surface nulle pour une connexion en synthétisé, ça peut vouloir dire :</p> <ul style="list-style-type: none"> - que certaines connexion ont un poids de 0 dans l'assolement. - que le graph est mal-formé, c'est à dire qu'il existe des poids absurdes - qu'aucune intervention n'a été renseignée pour l'ensemble de l'assolement --> il est considéré caduc et on attribue à chaque connexion une surface nulle.
--	--	--