

---

**ANNEE UNIVERSITAIRE 2022/2023**

---

**Département Génie Biologique**

**RAPPORT DE STAGE**  
**Option : SEE**

**Tanguy Leon**

**Lieu du stage :**      **INRAE Haut-de-france**  
**2 Chaussée Brunehaut**  
**80200 Estrees-Mons**

**IUT Rue de Kergoat - CS 93937 29238 BREST Cedex 3 / 02 98 01 60 73**

Stage de 2<sup>ème</sup> année de BUT génie biologique 2022 - 2023



Collection de miscanthus de l'INRAE haut de France

source : © Aline Waquet - INRAE

**Réalisation d'une ontologie sur le miscanthus à des fins  
de partage de données provenant de l'INRAE**

## **Remerciements**

Je souhaite remercier mon maître de stage, Mme Hulmel Maryse de m'avoir accueilli au sein de l'INRAE d'Estrées-Mons. Elle m'a beaucoup appris sur le partage de données et sur le miscanthus. De plus, elle a fait preuve de confiance et m'a guidée tout le long de mon stage.

Je tiens également à remercier Mme Zapater Marion et M. Iqbal Shehyar pour leur bienveillance, mais aussi pour m'avoir aidé et accompagné sur les différentes missions qui m'étaient confiées.

Je remercie aussi l'unité expérimentale en particulier, Mme Heumez Marie et M. Lelong Bastien, pour leur accueil et leur gentillesse lors des manipulations.

Je souhaite également remercier tout le personnel de l'INRAE qui m'a accueilli et qui m'a permis de passer un très bon stage.

## Sommaire

Table des figures.....	
I – Introduction.....	1
II – Présentation de l’INRAE et de l’unité de recherche.....	1
III – La science ouverte à l’INRAE .....	2
IV – Réalisation d’une ontologie pour miscanthus durant mon stage.....	4
V – Conclusions et perspectives.....	5
VI – Références bibliographiques .....	6
VII – Annexe .....	7
Résumé.....	9

## Table des figures

*Figure 1: Exemple de caractères pour l'ontologie créée par R. Raverdy sur la plateforme GnpIS*

*Figure 2 : Exemple de construction de l'ontologie*

## I – Introduction

J'ai rejoint une équipe menant des recherches au sein de l'INRAE, un institut de recherche spécialisé en agriculture, alimentation et environnement. Cette équipe, spécialisée en écophysiologie et génétique, étudie une culture récente en France, la culture du miscanthus (Brancourt-Hulmel, 2022). Cette graminée pérenne est reconnue pour sa forte production en biomasse tout en ne nécessitant pas de produits chimiques, à l'exception de désherbage lors de l'installation de la culture. Ses débouchés sont nombreux et visent entre autres la bioénergie, la litière animale ou horticole, les produits biosourcés (composites, béton allégé, isolants...) et divers services écosystémiques. Ces services peuvent concerner la limitation de l'érosion, la protection de zones sensibles telles les aires d'alimentation de captage en eau, les zones non traitées autour des habitations. Les recherches sont menées selon un tripode de caractères relatifs à la production de biomasse, sa valorisation et l'apport en services écosystémiques. Au niveau écophysiologique, une attention particulière porte sur l'étude du recyclage de l'azote par la plante (Leroy, 2021).

Mon stage s'est déroulé en deux étapes. J'ai d'abord commencé par participer au phénotypage de caractères écophysiologiques, le phénotypage consistant en l'évaluation des caractères de la plante au champ et au laboratoire. Les caractères en question étaient relatifs au recyclage de l'azote par la plante et le matériel végétal concernait une descendance de miscanthus (espèce *Miscanthus sinensis*), qu'on appelle communément "population". A l'aide du phénotypage de cette population, l'équipe sera ensuite en mesure d'identifier des QTL, c'est-à-dire des régions chromosomiques associées à ces caractères quantitatifs. Dans les grandes lignes, le phénotypage des caractères consistait à mesurer la production de biomasse aérienne et souterraine et les composants de cette production à différentes dates. Des échantillons des organes de la plante ont ensuite été broyés pour des analyses chimiques ultérieures afin d'en déterminer la concentration en azote. Cela permet de calculer ensuite les différents flux d'azote entre organes et entre dates afin d'appréhender l'étude du recyclage de l'azote par la plante. Pour ces mêmes variables, j'ai ensuite été amené à compléter une base de données associée aux données pour lesquelles j'ai participé au phénotypage des plantes, ce qui a constitué l'objectif principal de mon stage.

Pour la rédaction du rapport de stage, nous avons choisi avec ma maître de stage, une présentation de rapport selon la deuxième possibilité intitulée "stage moins scientifique". En effet, même si le stage s'est déroulé dans le domaine de la recherche, il ne s'agit pas d'un stage qui répond strictement à une présentation de résultats scientifiques comme c'est généralement le cas. Celui-ci se place dans un contexte plus général de diffusion des résultats de la recherche avec un objectif de développement d'un outil afin de permettre l'accès aux données issues de la recherche. Aussi, nous proposons un plan de rapport de stage plus libre où il nous paraît important d'introduire, entre autres, la notion de science ouverte au sein de l'institut où l'objectif du stage prend toute sa place. Après une présentation générale de l'INRAE, j'introduirai la notion de science ouverte, en m'appuyant sur quelques exemples issus de l'institut, puis j'aborderai la façon dont on peut participer à la science ouverte par le partage de données. Pour cela, je décrirai la plateforme utilisée par l'équipe avec laquelle j'ai travaillé. Enfin, je décrirai les différentes étapes qui m'ont permis de compléter l'ontologie destinée à partager les données liées à l'étude du recyclage de l'azote par le miscanthus.

## II – Présentation de l'INRAE et de l'unité de recherche

L'INRAE est un institut de recherche dont les principales activités sont de mener des expérimentations concernant l'agriculture, l'alimentation et l'environnement. Créée en 2020 par l'association de l'INRA (lui-même créé en 1946) et de l'IRSTEA, elle est implantée dans la France entière grâce à ces 18 centres de recherche ([inrae.fr/centres](http://inrae.fr/centres)). Son objectif principal est d'améliorer les techniques agricoles et d'élevage, mais aussi de pallier les problèmes environnementaux grâce à l'agronomie. Les enjeux mondiaux tels que le réchauffement climatique, le manque de ressources ou encore l'augmentation de la population sont au centre des recherches. L'INRAE est réparti sur plusieurs centres en France dont celui des Hauts-de-France qui comprend une unité INRAE où mon stage s'est déroulé. Cette unité fait elle-même partie de l'unité mixte de recherche transfrontalière (UMRt) BioEcoAgro. Elle se compose de plus de 400 chercheurs, techniciens et doctorants ([bioecoagro.eu](http://bioecoagro.eu)). Cette UMRt provient d'une collaboration

étroite entre l'INRAE (basé sur les sites de Laon et d'Estrées-Mons), l'université de Lille, celle de Picardie Jules Verne et celle de Liège. Une unité expérimentale de l'INRAE se joint aussi aux projets afin de permettre l'acquisition des données sur le terrain ou dans les laboratoires. Cette collaboration permet une meilleure efficacité dans l'obtention de nouvelles connaissances mais aussi de former de nouveaux chercheurs, ingénieurs et techniciens lors de stage ou de projet doctoral. L'ensemble des données produites sont diffusées publiquement, c'est ce qu'on appelle la science ouverte.

### III – La science ouverte à l'INRAE

#### a) Que signifie la science ouverte ?

La science ouverte est une nouvelle approche du partage de connaissances et répond à une problématique majeure. En effet, plusieurs mêmes études peuvent être répétées, en France ou à l'étranger. Cela peut engendrer des coûts importants qui ne sont pas nécessaires et donc entraîner une diminution des connaissances produites. La science ouverte permet de diffuser les connaissances acquises à un large public, que ce soit des scientifiques ou tout simplement à un public curieux d'apprendre ([INRAE engagé pour une science ouverte | INRAE INSTIT](#)). Cela a l'avantage d'éviter les redondances et de consolider les connaissances en comparant, voire en agrégeant, les données issues de plusieurs études pour réaliser des analyses approfondies appelées méta-analyses.

De ce fait, des bases de données scientifiques ont vu le jour et sont alimentées en connaissances au fur et à mesure que les travaux progressent. De plus, les résultats bruts sont eux aussi partagés afin de laisser la liberté de réutiliser les données.

Le but final est d'accroître les connaissances pour répondre aux problématiques, le mieux et le plus rapidement possible. A l'INRAE, la science ouverte est omniprésente dans tous les domaines de recherche, toutes les données d'un domaine scientifique sont stockées au même endroit sur une plateforme dédiée et sont gérées par des pôles numériques dédiés ([Recherche Data Gouv, les données en commun | INRAE INSTIT](#)). A l'horizon 2030, l'INRAE ambitionne d'atteindre 100 % de publications librement accessibles en privilégiant leur dépôt dans les archives ouvertes telles que HAL INRAE (Inrae DipSO, 2023).

#### b) Quelques exemples du partage de la science

Le partage de la science peut se faire de diverses manières (<https://www.static.inrae.fr/actualites/prix-liber-science-ouverte-transparente-gratuite>). Bien évidemment, la majorité se fait sur internet grâce notamment à des publications scientifiques, lesquelles peuvent s'accompagner de fichiers de données (exemple des données indiquées dans l'article de Raverdy et al, 2022 <https://doi.org/10.15454/1NVRNJ>). Il est très simple de les retrouver. Des sites web comme des dictionnaires en ligne sont aussi publiés, permettant à tout type de personne (scientifique ou non) de s'informer sur un sujet précis. L'existence de base de données en ligne apparaît peu à peu et permet une diffusion de résultats bruts et traités. De plus, des publications scientifiques sur l'ensemble des travaux réalisés sont publiées dans des revues scientifiques et sont disponibles sur google scholar, HAL INRAE... Cette large accessibilité à la connaissance est un point fort pour tous les bureaux d'études et instituts de recherche. Le public visé est donc plus important.

#### c) Comment partager les données ?

Dans le domaine des plantes et des champignons, le partage de données relatives aux recherches menées à l'INRAE se réalise à l'aide de la plateforme GnpIS (<https://urgi.versailles.inrae.fr/gnpis/>), un système d'information intégratif multi-espèces qui a été créé et géré par une unité de recherche dédiée, l'Unité de Recherche en Génomique-Info (URGI) localisée à l'INRAE de Versailles.

La plateforme GnpIS est utilisée pour des données générées dans le cadre de grands projets internationaux et nationaux conduits par les chercheurs de l'INRAE spécialisés dans le domaine de la biologie, génétique et amélioration des plantes et plus largement les sciences végétales. Elle est régulièrement améliorée et

mise à jour plusieurs fois par an. GnpIS est accessible par un portail web et permet de parcourir différents types de données, soit indépendamment à travers des interfaces dédiées ou simultanément en utilisant une recherche rapide ('google like search') ou des outils de recherche avancée (<https://scholar.google.com/>).

La plateforme rassemble des données de différente nature qui peuvent concerner des ressources génétiques, des données de phénotypage (évaluation de caractéristiques des plantes et des champignons), des données de polymorphisme, de génétique d'association, de carte génétique, marqueurs et détection de régions chromosomiques associées à des caractères quantitatifs (QTL ou Quantitative Trait Loci), des données de séquençage de génome, de synténie, de séquences, d'éléments transposables, de transcriptomique et enfin de banques de séquences. Les données relatives à mon objectif de stage concernent le phénotypage (<https://urgi.versailles.inrae.fr/>).

Pour faciliter l'interrogation des données de phénotypage sur ce site, une ontologie a été créée et permet de simplifier la compréhension des différents caractères mesurés. Pour les données de phénotypage dans le domaine de la génétique des plantes, une ontologie a été développée spécifiquement pour chaque plante cultivée et permet de classer différents types d'objets, mais aussi ses différentes caractéristiques et ses propriétés. Ainsi, l'ontologie va pouvoir permettre à une machine de produire des raisonnements, dits « inférences » à partir de données. De ce fait, beaucoup d'informations essentielles peuvent être récupérées à partir d'une ontologie, ce qui permet d'instruire le lecteur sur les recherches avant même d'avoir lu la publication ou parcouru les données reliées à cette ontologie.

L'ontologie présente sur la plateforme GnpIS fait apparaître toutes les variables mesurées *in situ* (c'est-à-dire dans le champ, la chambre climatique, la serre...) ou dans les laboratoires d'analyses (<https://urgi.versailles.inra.fr/ephep/ephep/ontologyportal.do>). La méthode de mesure utilisée est indiquée, de même que l'acronyme (sigle) lié au nom de la variable. L'ensemble des variables est alors classé selon la nature de la variable et selon la plante étudiée (figure 1).

The screenshot displays the GnpIS 'Ontologies' page. The main content area shows a hierarchical tree for the 'Miscanthus' ontology. Under the 'Morphological' trait class, the 'Circonférence de plante' trait is highlighted, with its specific variable 'C\_cm: Circonférence de plante' selected. A sidebar on the right provides detailed information for this variable, including its identifier (BFF:1000011), name, entity (Plant), attribute (Circumference), and class (Morphological). It also lists the measurement method 'Mesure directe' and the unit 'cm'.

Figure 1: Exemple de caractères pour l'ontologie créée par R. Raverdy sur la plateforme GnpIS

## IV – Réalisation d’une ontologie pour miscanthus durant mon stage

La création d’une ontologie suit plusieurs étapes afin d’être aboutie. En effet, celle-ci doit être effectuée dans un ordre logique pour ne pas créer de confusion entre les différentes variables associées aux caractères étudiés. Pour développer une ontologie spécifique aux variables relatives au recyclage de l’azote, je me suis appuyé sur l’ontologie “miscanthus” (<https://urgi.versailles.inra.fr/ephep/ephep/ontologyportal.do>) qui a été initiée par Raphaël Raverdy dans le cadre d’une thèse sur “l’étude du déterminisme génétique de la production et de la composition de biomasse chez *Miscanthus sinensis*” (2021). J’ai d’abord commencé par répertorier toutes les variables en lien avec le recyclage de l’azote et qui avaient été en grande partie étudiées au sein de la thèse réalisée par Julie Leroy (2021) sur “l’étude écophysiological temporelle de la gestion des réserves carbonées et azotées par trois géotypes de miscanthus”. J’en ai exclu les variables qui étaient communes au travail de thèse de Raphaël Raverdy pour éviter les redondances. Pour chaque variable de l’ontologie “miscanthus”, j’ai ensuite renseigné les éléments de la variable (nom et synonyme), ceux du caractère correspondant (nom, entité, classe), de la méthode (nom, description, classe) et enfin l’échelle. Un caractère donné peut être apprécié par plusieurs variables. Par exemple, on pourrait avoir différentes façons de mesurer la hauteur de canopée, ce qui correspondrait à plusieurs variables.

En observant l’ontologie existante, j’ai pu constater qu’il existe quatre classes de caractères : phénotypique, morphologique, agronomique ou de qualité. Pour ma liste de caractères, j’ai dû ajouter une nouvelle classe “physiologique” de caractères pour inclure des caractères tels que “Nitrogen use efficiency” (efficacité d’utilisation de l’azote). Chaque classe comprend alors plusieurs caractères/variables. L’entité de l’organe mesuré va ensuite être déterminée : par exemple, cela peut concerner les tiges, les feuilles ou le rhizome (tige souterraine). Il en est de même pour l’attribut du caractère qui peut correspondre à une date, une hauteur, une concentration, ... Une fois classée, chaque variable nécessite un acronyme afin de permettre une utilisation simple lors de la rédaction. L’acronyme est très important car il se doit d’être facile à comprendre dès la première lecture.

Si j’ai pu être autonome pour renseigner les différents éléments de l’ontologie (grâce à la thèse de Julie Leroy), le travail le plus délicat a cependant concerné le choix du sigle correspondant au nom de la variable. J’ai eu plusieurs réunions de travail à ce sujet avec l’équipe à savoir, Maryse Hulmel, ma maître de stage qui est spécialisée en génétique, Marion Zapater, une ingénieure de recherche spécialisée en écophysiological et Shehyar Iqbal, un doctorant de l’équipe. Lors de ces réunions, ces trois chercheurs ont confronté leurs différentes propositions afin d’être unanimes sur un choix unique et non équivoque.

Pour proposer les nouveaux noms, nous avons prêté une attention particulière à la structure des noms déjà existants dans l’ontologie qui a été initiée par Raphaël Raverdy. Celui-ci est souvent composé des premières lettres de chaque mot de la variable suivie par l’unité de cette variable (figure 2). Par exemple, “ABM\_tDMha” désigne le nom pour “Aboveground biomass production” (production de biomasse aérienne) et s’exprime en tonnes de matière sèche (DM pour Dry Matter) par hectare. Par analogie, nous avons donc proposé BBM\_tDMha pour “Belowground biomass production” (production de biomasse souterraine). De la même façon, PEm\_DOY désigne “Plant emergence date” (date d’émergence de la plante lorsqu’elle sort du sol) et s’exprime en quantités (DOY pour days of year). Pour toutes les variables associées à des stades, nous avons donc procédé de même. Par exemple, nous avons proposé “FLEm\_DOY” pour “Flag leaf emergence” (date de sortie de la feuille étendard correspondant à la dernière feuille émise par la plante) exprimée en quantités également.

Variable name	Trait name	Scale name	Trait class	Entity	Attribute	Method class
ABMMax_tDMha	Maximum aboveground biomass	tDM/ha	Agro-nomical	Aboveground part	Biomass Yield	Computation
BBMMax_tDMha	Maximum belowground biomass	tDM/ha	Agro-nomical	Belowground part	Biomass Yield	Computation
BBM_tDMha	Belowground (RhIM+Rhi2LL+RhiBuds) biomass production	tDM/ha	Agro-nomical	Belowground part	Biomass Yield	Computation
ABM_tDMha	Aboveground biomass production harvested in February	tDM/ha	Agro-nomical	Aboveground part	Biomass Yield	Computation
TBM_tDMha	Total (A+BS+FL+B) biomass production	tDM/ha	Agro-nomical	Plant	Biomass Yield	Computation
BSBM_tDMha	Basal stem biomass production	tDM/ha	Agro-nomical	Basal shoot	Biomass Yield	Computation

Figure 2 : Exemple de construction de l'ontologie

La dernière étape a été de tester si les acronymes des variables ne posent aucun problème lors de l'importation des variables avec le logiciel R, un logiciel d'analyse de données très couramment utilisé en recherche. Tous les noms ayant été testés et validés par les chercheurs de l'équipe, l'ontologie va donc pouvoir être soumise d'ici peu à la plateforme GnpIS. Au total, j'ai répertorié 75 variables durant mon stage (cf annexe).

## V – Conclusions et perspectives

La science ouverte est de nos jours une nécessité dans le domaine de la recherche et permet la diffusion des connaissances acquises. L'INRAE étant un organisme public de recherche, les résultats produits par l'institut sont donc pour la plupart disponibles sur internet *via* des bases de données, des publications scientifiques ou encore des sites web. Suite aux données acquises dans le cadre de la thèse de Julie Leroy sur l'étude du recyclage des réserves azotées et carbonées chez miscanthus, l'équipe de recherche encadrante a cherché à les rendre accessibles sur la plateforme GnpIS, une plateforme dédiée dans le domaine des sciences végétales. Pour faciliter l'interrogation sur la plateforme, une ontologie est nécessaire et permet de simplifier la compréhension des différents caractères mesurés. Dans ce contexte, la réalisation d'une ontologie sur le recyclage de l'azote par le miscanthus m'a été confiée. Elle va permettre d'informer le lecteur sur les différentes variables ou caractères en lien avec le recyclage de l'azote et qui ont été mesurés au champ ou en laboratoire. Pour la développer, je me suis appuyé sur celle initiée par un ancien doctorant de l'équipe. À l'aide de la thèse de Julie Leroy, j'ai d'abord répertorié toutes les variables en lien avec le recyclage de l'azote. Puis j'ai complété les différents éléments de la variable (nom et synonyme), ceux du caractère correspondant (nom, entité, classe), de la méthode (nom, description, classe) et enfin l'échelle. Pour chaque nouvelle variable, j'ai proposé un nouvel acronyme en concertation avec les chercheurs de l'équipe qui sont concernés par les différentes variables. Cet outil sera utile à plusieurs niveaux. Pour l'équipe, il va d'abord constituer une aide dans l'analyse des données, chaque variable étant parfaitement définie, puis lors de la rédaction de la partie "matériel et méthodes" des futures publications scientifiques. De plus, l'ontologie permettra l'unification des abréviations utilisées entre les différentes publications à venir. Les noms de variables pourront en outre être importées sans difficulté avec le logiciel R pour analyser les données, ce qui va éviter de les recoder inutilement. Cette unification permettra aussi de connecter des plateformes entre elles. Enfin, pour les autres chercheurs, l'ensemble de ce travail sera par la suite intégré à la plateforme GnpIS, ce qui permettra l'utilisation des données par d'autres communautés de chercheurs, que ce soit au niveau national ou international.

## VI – Références bibliographiques

Brancourt-Hulmel M., 2022. Miscanthus : une culture nouvelle en France. Les mots de l'agronomie - Histoire et critique, sous la direction de P. Morlon, INRAE-ACT. [en ligne]. Consulté le 11.04.2023 [Miscanthus : une culture nouvelle en France — Les Mots de l'agronomie \(inrae.fr\)](#)

Leroy, J., 2021. Etude écophysiological temporelle de la gestion des réserves carbonées et azotées par trois génotypes de miscanthus. Thèse AgroParisTech, Paris.

<https://www.inrae.fr/centres/>

<https://www.bioecoagro.eu/>

[INRAE engagé pour une science ouverte | INRAE INSTIT](#)

[Recherche Data Gouv, les données en commun | INRAE INSTIT](#)

Inrae DipSO Direction Pour La Science Ouverte, 2023 Bonnes pratiques pour atteindre 100 % de publications librement accessibles en 2030. [en ligne]. Consulté le 17.04.2023 ff10.17180/zgmf-be70ff.ffhal-04028495

<https://www.static.inrae.fr/actualites/prix-liber-science-ouverte-transparente-gratuite>

<https://urgi.versailles.inrae.fr/gnpis/>

<https://scholar.google.com/>

<https://urgi.versailles.inrae.fr/>

<https://urgi.versailles.inra.fr/epheis/epheis/ontologyportal.do>

Raverdy, R., 2021. Study of the genetic determinism related to biomass production and composition traits in *Miscanthus sinensis*. Thèse AgroParisTech, Paris

Raverdy, R., Mignot, E., Arnoult, S., Fingar, L., Bodineau, G., Griveau, Y., Volant, S., Brancourt-Hulmel, M., 2022. Estimation of Genetic Parameters of Biomass Production and Composition Traits in *Miscanthus sinensis* Using a Staggered-Start Design. BIOENERGY Res. 15, 735–754. <https://doi.org/10.1007/s12155-022-10459-5>

## VII – Annexe

### Annexe 1: Liste des nouvelles variables

Variable name	Trait name	Scale name	Trait class	Entity	Attribute	Method class
<u>PEm_DOY</u>	<u>Plant emergence date</u>	<u>DOY</u>	<u>Phenological</u>	<u>Plant</u>	<u>Time</u>	<u>Measurement</u>
<u>SLA_cm2g</u>	<u>Specific leaf area</u>	<u>cm<sup>2</sup>/g</u>	<u>Morphological</u>	<u>Leaf</u>	<u>Area</u>	<u>Measurement</u>
<u>GLBM_tDMha</u>	<u>Green leave biomass</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Leaf</u>	<u>Biomass yield</u>	<u>Computation</u>
<u>BLBM_tDMha</u>	<u>Brown leave biomass</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Leaf</u>	<u>Biomass yield</u>	<u>Computation</u>
<u>LAI</u>	<u>Leaf area index</u>	<u>m<sup>2</sup>/m<sup>2</sup></u>	<u>Morphological</u>	<u>Leaf</u>	<u>Area</u>	<u>Measurement</u>
<u>RDen_m2</u>	<u>Root density</u>	<u>roots/m<sup>2</sup></u>	<u>Morphological</u>	<u>Root</u>	<u>Density</u>	<u>Counting</u>
<u>RDep_m</u>	<u>Root depth</u>	<u>m</u>	<u>Morphological</u>	<u>Root</u>	<u>Depth</u>	<u>Measurement</u>
<u>PLNb</u>	<u>Plant leaf number</u>	<u>leaves/plant</u>	<u>Morphological</u>	<u>Leaf</u>	<u>Number</u>	<u>Counting</u>
<u>PSNb</u>	<u>Plant stem number</u>	<u>stem/plant</u>	<u>Morphological</u>	<u>Stem</u>	<u>Number</u>	<u>Counting</u>
<u>PSNbJune</u>	<u>Plant stem number in June</u>	<u>stem/plant</u>	<u>Morphological</u>	<u>Stem</u>	<u>Number</u>	<u>Counting</u>
<u>PSNbJune_MaxPSNb%</u>	<u>Plant stem number in June in proportion to maximal plant stem number</u>	<u>%</u>	<u>Morphological</u>	<u>Stem</u>	<u>Number</u>	<u>Computation</u>
<u>HGR_cmDD</u>	<u>Height growth rate</u>	<u>cm/DD</u>	<u>Morphological</u>	<u>Plant</u>	<u>Height</u>	<u>Computation</u>
<u>FlagEm_DD</u>	<u>Flag leaf emergence in degree days</u>	<u>DD</u>	<u>Phenological</u>	<u>Plant</u>	<u>Time</u>	<u>Measurement</u>
<u>FlagEm_DOY</u>	<u>Flag leaf emergence</u>	<u>DOY</u>	<u>Phenological</u>	<u>Plant</u>	<u>Time</u>	<u>Measurement</u>
<u>HD_DOY</u>	<u>Heading date = panicle emergence</u>	<u>DOY</u>	<u>Phenological</u>	<u>Plant</u>	<u>Time</u>	<u>Measurement</u>
<u>Senes50_DD</u>	<u>Cumulative degree days to reach 50% relative stem height of the first senescent cohort</u>	<u>DD</u>	<u>Phenological</u>	<u>Plant</u>	<u>Time</u>	<u>Measurement</u>
<u>LLH_cm</u>	<u>Last ligulated leaf height</u>	<u>cm</u>	<u>Phenological</u>	<u>Plant</u>	<u>Height</u>	<u>Measurement</u>
<u>SenLLH_cm</u>	<u>Senescent last ligulated leaf height</u>	<u>cm</u>	<u>Phenological</u>	<u>Plant</u>	<u>Height</u>	<u>Measurement</u>
<u>SenLLH_LLH%</u>	<u>Senescent last ligulated leaf height in proportion to last ligulated leaf height</u>	<u>%</u>	<u>Phenological</u>	<u>Plant</u>	<u>Height</u>	<u>Measurement</u>
<u>SenSc</u>	<u>Senescence score (1 to 4, 4=max)</u>		<u>Phenological</u>	<u>Plant</u>	<u>Senescence</u>	<u>Measurement</u>
<u>ABMMax_tDMha</u>	<u>Maximum aboveground biomass</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Aboveground part</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>BBMMax_tDMha</u>	<u>Maximum belowground biomass</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Belowground part</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>BBM_tDMha</u>	<u>Belowground (RhiM+Rhi2LL+RhiBuds) biomass production</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Belowground part</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>ABM_tDMha</u>	<u>Aboveground biomass production harvested in February</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Aboveground part</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>TBM_tDMha</u>	<u>Total (A+BS+FL+B) biomass production</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Plant</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>BSBM_tDMha</u>	<u>Basal stem biomass production</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Basal shoot</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>ABSBM_tDMha</u>	<u>Aboveground (A) + basal stems (BS) biomass production</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Plant</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>FLBM_tDMha</u>	<u>Fallen leaves biomass production</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Leaf</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>AFLBM_tDMha</u>	<u>Aboveground (A) + fallen leaves (FL) biomass production</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Plant</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>ABSFLBM_tDMha</u>	<u>Aboveground (A) + basal stems (BS) + fallen leaves (FL) biomass production = total aboveground biomass production</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Plant</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>YSBM_tDMha</u>	<u>Young shoot biomass production</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Young shoot</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>Rhi2LLBM_tDMha</u>	<u>Biomass production of 4 cm-rhizome associated to the stems with at least two ligulated leaves</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Rhizome</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>RhiBudsBM_tDMha</u>	<u>Biomass production of 4 cm-rhizome associated to buds</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Rhizome</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>RhiMBM_tDMha</u>	<u>Biomass production corresponding to the main part of rhizome</u>	<u>tDM/ha</u>	<u>Agronomical</u>	<u>Rhizome</u>	<u>Biomass Yield</u>	<u>Computation</u>
<u>RUE_gMJ</u>	<u>Radiation use efficiency</u>	<u>g/MJ<sup>2</sup></u>	<u>Physiological</u>	<u>Plant</u>	<u>Radiation use efficiency</u>	<u>Computation</u>
<u>PAR_nm</u>	<u>Photosynthetic active radiation</u>	<u>nm</u>	<u>Physiological</u>	<u>Plant</u>	<u>Photosynthetic active radiation</u>	<u>Measurement</u>
<u>tPAR_nm</u>	<u>Transmitted photosynthetic active radiation</u>	<u>nm</u>	<u>Physiological</u>	<u>Plant</u>	<u>Photosynthetic active radiation</u>	<u>Computation</u>
<u>NUE_kgDMkgN</u>	<u>Nitrogen use efficiency</u>	<u>kgDM/kgN</u>	<u>Physiological</u>	<u>Plant</u>	<u>Nitrogen use efficiency</u>	<u>Computation</u>
<u>ANC_gNkgDM</u>	<u>Aboveground nitrogen concentration (harvested height &gt;7cm)</u>	<u>gN/kgDM</u>	<u>Quality</u>	<u>Aboveground part</u>	<u>Nitrogen concentration</u>	<u>Computation</u>
<u>BNC_gNkgDM</u>	<u>Belowground (RhiBuds+Rhi2LL+RhiM) nitrogen concentration</u>	<u>gN/kgDM</u>	<u>Quality</u>	<u>Belowground part</u>	<u>Nitrogen concentration</u>	<u>Computation</u>

RhiBudsNC_gNkgDM	Nitrogen concentration of 4 cm-rhizome associated to buds	gN/kgDM	Quality	Rhizome	Nitrogen concentration	Computation
Rhi2LLNC_gNkgDM	Nitrogen concentration of 4 cm-rhizome associated to the stems with at least two ligulated leaves	gN/kgDM	Quality	Rhizome	Nitrogen concentration	Computation
RhiMNC_gNkgDM	Nitrogen concentration associated to the main part of rhizome	gN/kgDM	Quality	Rhizome	Nitrogen concentration	Computation
BSNC_gNkgDM	Basal stem nitrogen concentration	gN/kgDM	Quality	Basal shoot	Nitrogen concentration	Computation
TNC_gNkgDM	Total (A+BS+FL+B) weighted average nitrogen concentration of plant	gN/kgDM	Quality	Plant	Nitrogen concentration	Computation
ABSNC_gNkgDM	Weighted average nitrogen concentration of aboveground (A) + basal stems (BS)	gN/kgDM	Quality	Plant	Nitrogen concentration	Computation
FLNC_gNkgDM	Fallen leaves nitrogen concentration	gN/kgDM	Quality	Leaf	Nitrogen concentration	Computation
AFLNC_gNkgDM	Weighted average nitrogen concentration of aboveground (A) + fallen leaves (FL)	gN/kgDM	Quality	Plant	Nitrogen concentration	Computation
ABSFLNC_gNkgDM	Weighted average nitrogen concentration of aboveground (A) + basal stems (BS) + fallen leaves (FL) nitrogen concentration = total aboveground nitrogen concentration	gN/kgDM	Quality	Plant	Nitrogen concentration	Computation
YSNC_gNkgDM	Young shoot nitrogen concentration	gN/kgDM	Quality	Young shoot	Nitrogen concentration	Computation
ANQ_kgNha	Aboveground nitrogen quantity (harvested height > 7cm)	kgN/ha	Quality	Aboveground part	Nitrogen quantity	Computation
BNQ_kgNha	Belowground (RhiBuds+Rhi2LL+RhiM) nitrogen quantity	kgN/ha	Quality	Belowground part	Nitrogen quantity	Computation
RhiBudsNQ_kgNha	Nitrogen quantity of 4 cm-rhizome associated to buds	kgN/ha	Quality	Rhizome	Nitrogen quantity	Computation
Rhi2LLNQ_kgNha	Nitrogen quantity of 4 cm-rhizome associated to the stems with at least two ligulated leaves	kgN/ha	Quality	Rhizome	Nitrogen quantity	Computation
RhiMNQ_kgNha	Nitrogen quantity associated to the main part of rhizome	kgN/ha	Quality	Rhizome	Nitrogen quantity	Computation
BSNQ_kgNha	Basal stem nitrogen quantity	kgN/ha	Quality	Basal shoot	Nitrogen quantity	Computation
TNQ_kgNha	Total (A+BS+FL+B) nitrogen quantity	kgN/ha	Quality	Plant	Nitrogen quantity	Computation
ABSNO_TDMha	Aboveground (A) + basal stems (BS) nitrogen quantity	kgN/ha	Quality	Plant	Nitrogen quantity	Computation
FLNQ_kgNha	Fallen leaves nitrogen quantity	kgN/ha	Quality	Leaf	Nitrogen quantity	Computation
AFLNQ_kgNha	Aboveground (A) + fallen leaves (FL) nitrogen quantity	kgN/ha	Quality	Plant	Nitrogen quantity	Computation
ABSFLNQ_kgNha	Aboveground (A) + basal stems (BS) + fallen leaves (FL) nitrogen quantity = total aboveground nitrogen quantity	kgN/ha	Quality	Plant	Nitrogen quantity	Computation
YSNQ_kgNha	Young shoot nitrogen quantity	kgN/ha	Quality	Young shoot	Nitrogen quantity	Computation
Ndff_kgNha	Quantity of nitrogen derived from fertilizer in the plant	kgN/ha	Quality	Plant	Nitrogen quantity	Computation
Nrec	Quantity of nitrogen derived from fertilizer in the plant in proportion of quantity of applied nitrogen	%	Physiological	Plant	Nitrogen uptake	Computation
Ndff_MaxTNQ%	Quantity of nitrogen derived from fertilizer in the plant in proportion of maximum nitrogen quantity in the plant	%	Physiological	Plant	Nitrogen uptake	Computation
ARbNQ_kgNha	Autumn nitrogen remobilisation flux calculated from aboveground part	kgN/ha	Physiological	plant	Nitrogen remobilization	Computation
ARbNQ_MaxANQ%	Autumn nitrogen remobilisation efficiency calculated from aboveground part	%	Physiological	Plant	Nitrogen remobilization	Computation
ARbNQ_kgNha	Autumn nitrogen remobilisation flux calculated from belowground part	kgN/ha	Physiological	Plant	Nitrogen remobilization	Computation
ARbNQ_MaxANQ%	Autumn nitrogen remobilisation efficiency calculated from belowground part	%	Physiological	Plant	Nitrogen remobilization	Computation
LossNQ_kgNha	Loss of nitrogen quantity by the whole plant	kgN/ha	Physiological	Plant	Nitrogen loss	Computation
LossNQ_MaxTNQ%	Loss of nitrogen quantity by the whole plant in proportion of maximum total nitrogen quantity of the plant	%	Physiological	Plant	Nitrogen loss	Computation
NUMax_kgNha	Maximum nitrogen uptake from the soil	kgN/ha				
Nother_kgNha	Nitrogen uptake not originating from fertilizer	kgN/ha	Physiological	Plant	Nitrogen uptake	Computation
Nother_MaxTNQ%	Nitrogen originating from other nitrogen sources than fertilizer and remobilisation in proportion of maximum total nitrogen quantity of the plant	%	Physiological	Plant	Nitrogen uptake	Computation
SRNQ_kgNha	Spring nitrogen remobilization flux	kgN/ha	Physiological	Plant	Nitrogen remobilization	Computation

## Résumé

L'unité mixte de recherche BioEcoAgro de l'INRAE Hauts-de-France partage l'intégralité des connaissances acquises à toutes les personnes désirant s'informer sur la production de biomasse par le miscanthus aux niveaux génétique et écophysiologiques. L'ensemble des données est alors disponible sur une plateforme numérique nommée « GnpIS ». Cette ouverture de la science permet d'éviter la répétition des travaux de recherche que d'autres organismes pourraient mener.

Les travaux effectués durant ce stage portent sur le miscanthus, une plante pérenne ayant de nombreux débouchés dus à sa production de biomasse importante : chauffage, alimentation, paillage, ... De nombreux caractères/variables sont alors étudiés, de ce fait beaucoup de données sont produites.

Mon principal objectif fut de compléter l'ontologie déjà existante, en y intégrant de nouvelles variables écophysiologiques. Une ontologie facilite la compréhension des caractères mesurés et simplifie donc le partage de données. Elle permet aussi de s'accorder sur les abréviations de chaque variable. Ainsi, 75 nouvelles variables provenant de la thèse de Julie Leroy (une ancienne doctorante de l'équipe) ont été ajoutées et seront prochainement intégrées à la plateforme GnpIS.