



UNIVERSITÉ DE
SHERBROOKE

Faculté des sciences

Projet Herbivorie – Modèle conceptuel et logique de la base de données

Auteur	Email	CIP
Hannachi Mallek	hanm1601@usherbrooke.ca	hanm1601
Mamady III Diakite	diam2248@usherbrooke.ca	diam2248
Abdul Magide Djalo	maga5296@usherbrooke.ca	maga5296
Adem Bayounes	Adem.Bayounes@USherbrooke.ca	baya0318

Présenté à :

Pr. Luc Lavoie

Octobre 2025

Table des matières

1	Introduction	4
2	Modèle conceptuel biologiste	4
2.1	Vue générale	4
2.2	Entités principales	5
2.3	Associations et cardinalités	5
3	Modèle conceptuel technique	6
3.1	Vue générale	6
3.2	Caractéristiques générales du MCD technique	7
4	Entités détaillées	7
4.1	SITE	7
4.2	ZONE	7
4.3	PLACETTE	7
4.4	PARCELLE	8
4.5	PLANT	8
4.6	OBSERVATION_BOTANIQUE	8
4.7	OBSERVATION_METEO	8
5	Associations et cardinalités	8
5.1	APPARTENIR_SITE (SITE – ZONE)	8
5.2	APPARTENIR_ZONE (ZONE – PLACETTE)	9
5.3	DECOUPER_PARCELLE (PLACETTE – PARCELLE)	9
5.4	LOCALISER (PARCELLE – PLANT)	9
5.5	OBSERVATION_PLANT (PLANT – OBSERVATION_BOTANIQUE)	9
5.6	OBSERVATION_METEO_SITE (SITE – OBSERVATION_METEO)	9
6	Règles métier synthétiques	9
7	Schéma principal « Herbivorie » : modèle logique normalisé	10
8	Schéma « HerbivorieStaging » : zone tampon pour données brutes	11
9	DVCT : fonctions de validation et conversion	12

10 Scénario simple pour les groupes **15**

11 Conclusion **16**

Table des figures

1	Modèle conceptuel biologiste	4
2	Modèle conceptuel technique normalisé	6

1 Introduction

Le présent rapport documente l'ensemble du travail réalisé dans le cadre du projet **Herbivorie**. Il présente successivement :

- un **modèle conceptuel biologiste** (MCD BIO), orienté terrain et destiné aux biologistes ;
- un **modèle conceptuel technique** (MCD TECH), normalisé et conforme aux exigences du cours ;
- le modèle logique SQL complet normalisé en 5FN ;
- la zone de staging, les fonctions DVCT, les procédures ELT et les opérations ÉMIR ;
- un scénario d'utilisation synthétisant le fonctionnement global du système.

L'ensemble du rapport suit un style technique explicite, cohérent et exempt d'ambiguités.

2 Modèle conceptuel biologiste

2.1 Vue générale

Le MCD biologiste représente le fonctionnement terrain tel qu'observé par les biologistes. Il vise la simplicité et la lisibilité.

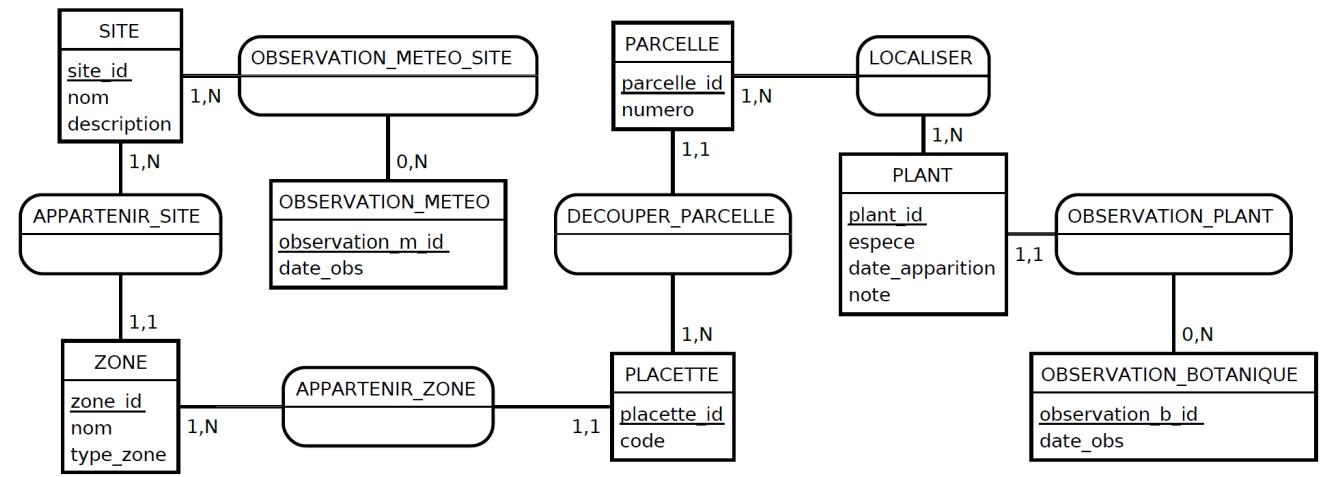


FIGURE 1 – Modèle conceptuel biologiste

2.2 Entités principales

SITE

Représente l'unité géographique générale utilisée pour structurer les observations.

ZONE

Subdivision du site ; groupe plusieurs placettes.

PLACETTE

Unité d'échantillonnage principale utilisée pour les relevés écologiques.

PARCELLE

Subdivision d'une placette améliorant la précision de localisation.

PLANT

Individu végétal suivi dans le temps.

OBSERVATION_BOTANIQUE

Regroupe les mesures botaniques effectuées sur un plant.

OBSERVATION_METEO

Représente les conditions météorologiques journalières du site.

2.3 Associations et cardinalités

- Un site possède **1..N zones**.
- Une zone possède **1..N placettes**.
- Une placette possède **1..N parcelles**.
- Une parcelle contient **1..N plants**.
- Un plant possède **0..N observations botaniques**.
- Un site possède **0..N observations météo**.

3 Modèle conceptuel technique

3.1 Vue générale

Le modèle conceptuel technique reprend les entités et relations du modèle biologiste, mais les structure selon les principes suivants :

- normalisation stricte jusqu'en cinquième forme normale ;
- utilisation de domaines SQL spécialisés pour représenter les règles métier ;
- décomposition systématique des observations botaniques et météorologiques ;
- alignement avec les exigences du cours et les standards CoFELI.

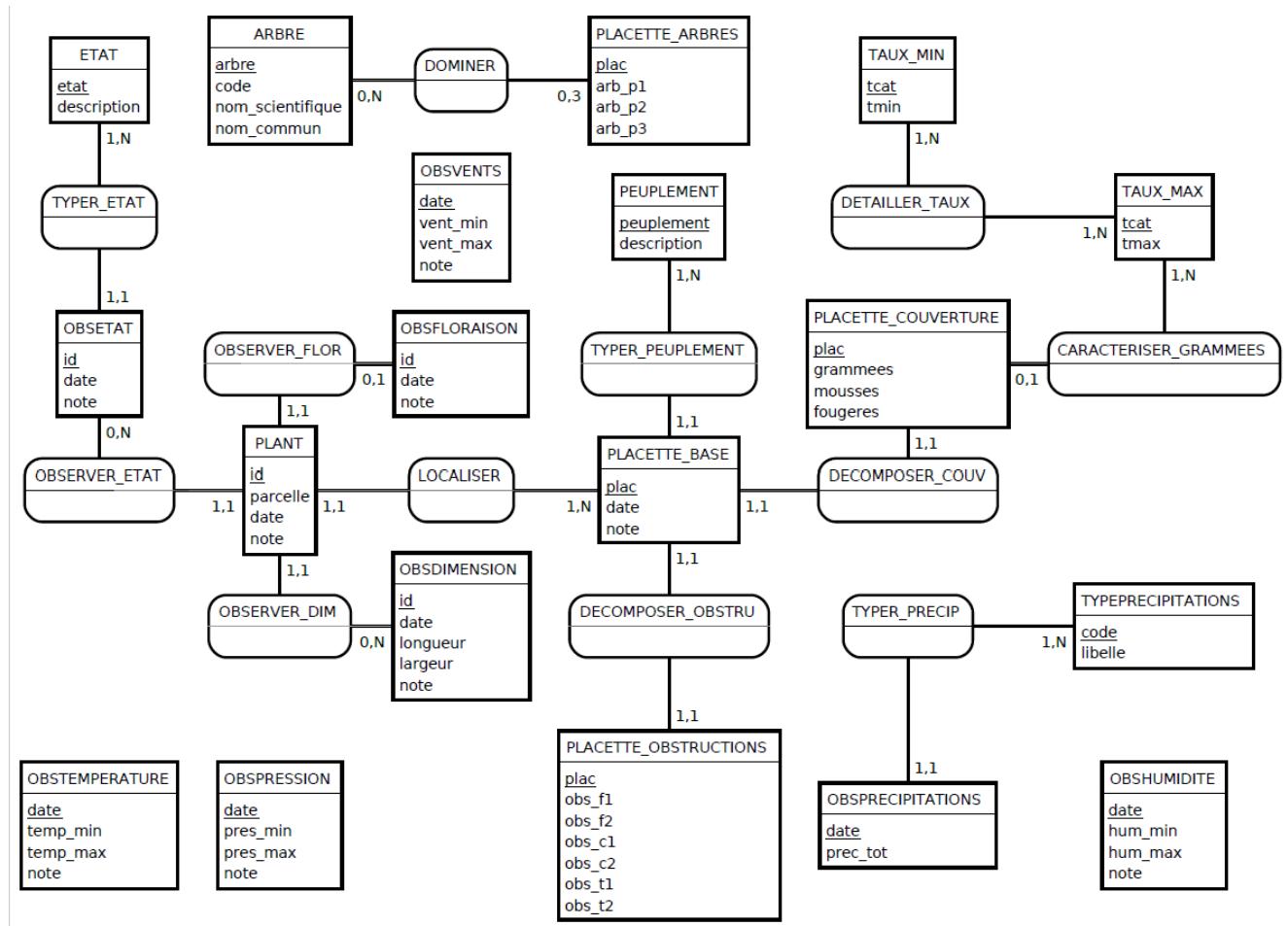


FIGURE 2 – Modèle conceptuel technique normalisé

3.2 Caractéristiques générales du MCD technique

Le modèle est conçu pour permettre une implémentation robuste dans un système relationnel. Il sépare clairement :

- les entités de référence (répertoires);
- les entités opérationnelles (plants, parcelles, placettes);
- les observations botaniques spécialisées;
- les observations météorologiques partitionnées;
- la gestion des taux en 5FN;
- la structuration des placettes en plusieurs tables cohérentes.

4 Entités détaillées

4.1 SITE

- **Identifiant** : site_id
- **Attributs** : nom, description
- **Rôle** : représente un lieu d'étude dans lequel sont organisées les zones et leurs placettes.

4.2 ZONE

- **Identifiant** : zone_id
- **Attributs** : nom, type_zone, site_id
- **Rôle** : partitionne un site en secteurs d'analyse.
- **Dépendance** : chaque zone appartient à un seul site.

4.3 PLACETTE

- **Identifiant** : placette_id
- **Attributs** : code, zone_id
- **Rôle** : unité d'échantillonnage principale du projet.
- **Dépendance** : une placette appartient à une zone.

4.4 PARCELLE

- **Identifiant** : parcelle_id
- **Attributs** : numero, placette_id
- **Rôle** : subdivision de la placette.
- **Dépendance** : une parcelle appartient à une seule placette.

4.5 PLANT

- **Identifiant** : plant_id
- **Attributs** : espece, date_apparition, note, parcelle_id
- **Rôle** : représente un individu végétal observé.

4.6 OBSERVATION_BOTANIQUE

- **Identifiant** : observation_b_id
- **Attributs** : plant_id, date_obs, longueur_feuille, largeur_feuille, etat, floraison, taux_graminees, taux_mousses, commentaire
- **Rôle** : décrit l'état botanique d'un plant à une date donnée.

4.7 OBSERVATION_METEO

- **Identifiant** : observation_m_id
- **Attributs** : site_id, date_obs, temp_min, temp_max, hum_min, hum_max, vent_min, vent_max, pres_min, pres_max, precip_totale, type_precip
- **Rôle** : décrit la météo quotidienne d'un site.

5 Associations et cardinalités

5.1 APPARTENIR_SITE (SITE – ZONE)

Lecture : un site contient plusieurs zones. **Cardinalités** :

- SITE : (1, N)
- ZONE : (1, 1)

5.2 APPARTENIR_ZONE (ZONE – PLACETTE)

Lecture : une zone contient plusieurs placettes. **Cardinalités :**

- ZONE : (1, N)
- PLACETTE : (1, 1)

5.3 DECOUPER_PARCELLE (PLACETTE – PARCELLE)

Lecture : une placette comporte plusieurs parcelles. **Cardinalités :**

- PLACETTE : (1, N)
- PARCELLE : (1, 1)

5.4 LOCALISER (PARCELLE – PLANT)

Lecture : une parcelle accueille plusieurs plants. **Cardinalités :**

- PARCELLE : (1, N)
- PLANT : (1, 1)

5.5 OBSERVATION_PLANT (PLANT – OBSERVATION_BOTANIQUE)

Lecture : un plant possède plusieurs observations botaniques. **Cardinalités :**

- PLANT : (1, N)
- OBSERVATION_BOTANIQUE : (1, 1)

5.6 OBSERVATION_METEO_SITE (SITE – OBSERVATION_METEO)

Lecture : un site est associé à plusieurs observations météo. **Cardinalités :**

- SITE : (1, N)
- OBSERVATION_METEO : (1, 1)

6 Règles métier synthétiques

- Un site comprend toujours au moins une zone.
- Une zone contient au moins une placette.
- Une placette peut être subdivisée en plusieurs parcelles.
- Un plant est toujours localisé dans une parcelle.

- Un plant peut être observé à plusieurs dates.
- Les observations météo sont mutualisées au niveau du site.
- Les mécanismes techniques (staging, rejets, DVCT, ELT, ÉMIR) ne font pas partie du MCD biologiste mais sont définis au niveau du modèle logique.

7 Schéma principal « Herbivorie » : modèle logique normalisé

Le schéma logique reprend l'ensemble des entités du MCD technique et applique :

- une normalisation stricte en 5FN ;
- l'utilisation de domaines spécialisés ;
- la décomposition des observations en tables spécialisées ;
- la documentation métier systématique via COMMENT ON.

Répertoires simples

- Arbre ;
- Peuplement ;
- Etat ;
- TypePrecipitations.

Taux en 5FN

Les intervalles de taux sont répartis en deux tables :

- TauxMin ;
- TauxMax.

Une vue recompose les données et une fonction garantit la compacité des intervalles.

Placette normalisée

- PlacetteBase ;
- PlacetteObstructions ;
- PlacetteCouverture ;
- PlacetteArbres.

Une vue Placette recompose l'entité complète.

Plants et observations

- Plant;
- ObsDimension;
- ObsFloraison;
- ObsEtat.

Météo partitionnée

- ObsTemperature;
- ObsHumidite;
- ObsVents;
- ObsPression;
- ObsPrecipitations.

Une vue MeteoHP regroupe les mesures hors précipitations.

8 Schéma « HerbivorieStaging » : zone tampon pour données brutes

Le schéma HerbivorieStaging est destiné à recevoir les données brutes provenant du terrain ou d'autres groupes du projet. Ce schéma joue un rôle central dans le processus d'intégration des données, car il permet de séparer les opérations de chargement massif des opérations de validation métier.

Tables de staging

Toutes les tables de staging utilisent le type `text` pour l'ensemble des colonnes. Cette décision facilite le chargement sans contrainte, même en présence de données manquantes ou incorrectes.

- **CarnetPlacette** : structure brute des placettes.
- **CarnetPlant** : structure brute des plants.
- **CarnetObservations** : structure brute des observations botaniques.
- **CarnetMeteo** : structure brute des observations météorologiques.

Table de rejets

La table Rejets conserve toutes les lignes non conformes rencontrées lors du traitement ELT. Elle contient :

- un identifiant de rejet;
- le flux concerné (PLACETTEBASE, PLANT, OBSDIMENSION, METEO_TEMPERATURE, etc.);
- un horodatage;
- un motif explicatif;
- les détails de l'erreur;
- la ligne brute en JSONB.

Deux vues simplifient la consultation :

- **Rejets_Detail** : affiche chaque rejet avec le JSON lisible;
- **Rejets_Synthese** : effectue un regroupement par flux avec une première et dernière occurrence.

Ce schéma constitue une zone de transit essentielle pour isoler les données avant validation.

9 DVCT : fonctions de validation et conversion

Le système DVCT (Détection, Vérification, Conversion, Transformation) assure la conformité des données avant leur insertion dans le schéma principal. Toutes les fonctions DVCT sont définies dans le schéma Herbivorie.

Fonctions pour les identifiants structurés

- Placette_id_verif() et Placette_id_conv()
- Peuplement_id_verif()
- Arbre_id_verif()
- Taux_id_verif()
- Plant_id_verif() et Plant_id_conv()
- Parcельe_verif() et Parcельe_conv()

Fonctions pour les dates écologiques

- Date_eco_verif()
- Date_eco_conv()

Fonctions pour les dimensions et états

- Dim_mm_verif() et Dim_mm_conv()
- Etat_id_verif()

Fonctions pour les paramètres météorologiques

- Temperature_verif() / Temperature_conv()
- Humidite_verif() / Humidite_conv()
- Vitesse_verif() / Vitesse_conv()
- Pression_verif() / Pression_conv()
- HNP_verif() / HNP_conv()
- Code_P_verif() / Code_P_conv()

Procédures ELT : chargement contrôlé des données

Les procédures ELT sont construites à partir des fonctions DVCT. Elles assurent un chargement contrôlé des données depuis les tables de staging, effectuent les vérifications et conversions nécessaires, alimentent les tables de production et consignent les rejets. Les étapes suivantes décrivent le fonctionnement de chaque procédure.

- **Placette_elt()** : cette procédure lit les données de CarnetPlacette, vérifie l'identifiant de placette, le peuplement, la date, les taux et les arbres.
 - les lignes valides sont insérées dans Placette_Base, Placette_0bstructions, Placette_Couverture et Placette_Arbres (avec ON CONFLICT DO NOTHING);
 - les lignes invalides sont consignées dans Rejets, avec une indication précise du champ fautif (placette invalide, taux invalide, arbre inconnu, etc.).
- **Plant_elt()** : cette procédure lit CarnetPlant, valide Plant_id, la placette, la parcellle et la date. Elle vérifie que la placette cible existe dans Placette_Base. Les plants valides sont insérés dans Plant et les lignes non conformes sont consignées dans Rejets.

- **Observations_ELT()** : cette procédure lit CarnetObservations et utilise le champ type_obs (DIMENSION, FLORAISON, ETAT). Elle normalise la valeur du type, vérifie l'identifiant, la date, les dimensions, l'état et l'existence du plant.
 - les données valides sont insérées dans ObsDimension, ObsFloraison ou ObsEtat selon le type ;
 - les lignes invalides sont consignées dans Rejets avec un motif explicite ;
 - les types d'observations inconnus sont rejetés avec le motif OBS_TYPE_INCONNU.
- **Meteo_ELT()** : cette procédure lit CarnetMeteo, valide chaque paramètre météo et vérifie les contraintes min max. Les données valides alimentent ObsTemperature, ObsHumidite, ObsVents, ObsPression et ObsPrecipitations. Les incohérences entraînent la création de rejets spécifiques pour chaque sous-flux (METEO_TEMPERATURE, METEO_HUMIDITE, etc.).

La procédure **ELT_Global()** assure l'enchaînement des quatre procédures (Placette_ELT(), Plant_ELT(), Observations_ELT(), Meteo_ELT()). Elle génère des messages NOTICE permettant de suivre les étapes du traitement. Cette procédure constitue l'entrée principale pour réaliser un chargement complet à partir des tables de staging, avec une traçabilité centralisée dans la table Rejets.

Système ÉMIR : consultation, modification, insertion et retrait

Le système ÉMIR regroupe l'EVA, MOD, INS, RET.

a) EVA — Fonctions de consultation

- Rechercher_Plants
- Historique_Plant
- Statistiques_Placette
- Consulter_Meteo_Periode
- Liste_Plants_Floraison
- Plants_Par_Etat

b) MOD — Procédures de mise à jour

- Modifier_Note_Plant
- Modifier_Localisation_Plant
- Modifier_Observation_Dimension
- Modifier_Observation_Etat
- Modifier_Observation_Floraison
- Modifier_Temperatures_Jour

c) INS — Insertion

- Inserer_Plant
- Inserer_Observation_Dimension
- Inserer_Observation_Floraison
- Inserer_Observation_Etat
- Inserer_Observation_Meteo

d) RET — Suppression

- Supprimer_Observation_Dimension
- Supprimer_Observation_Etat
- Supprimer_Plant
- Supprimer_Observations_Meteo_Jour

10 Scénario simple pour les groupes

1. **Réception des données d'un autre groupe** : un groupe transmet un fichier CSV contenant des placettes, des plants, des observations et des données météo. Le fichier est chargé tel quel dans les tables de staging (*CarnetPlacette*, *CarnetPlant*, *CarnetObservations*, *CarnetMeteo*) sans contrainte.

Le staging accepte toutes les données, même incorrectes, afin de permettre une analyse complète avant validation.

2. **Exécution du chargement global** : la procédure `CALL "Herbivorie".ELT_Global();` est exécutée.
 - `Placette_elt()` traite les données de placettes ;

- Plant_ELT() traite les plants ;
- Observations_ELT() traite les dimensions, états et floraisons ;
- Meteo_ELT() traite les données météo.

Les lignes invalides sont consignées dans "Herbivorie_Staging". Rejets avec le flux concerné, le motif et la ligne brute en JSON.

Le script `Herbivorie_cre.sql` installe ainsi un ensemble cohérent comprenant le modèle logique, la zone de staging, les fonctions DVCT, les procédures ELT et le système ÉMIR. Il fournit un environnement complet pour la gestion, la validation et la consultation des données.

11 Conclusion

Le système Herbivorie repose sur une architecture cohérente comprenant un MCD biologiste orienté terrain, un MCD technique normalisé, un schéma logique robuste, un ensemble complet de fonctions DVCT, un pipeline ELT sécurisé et un système ÉMIR structuré. Cette architecture permet de gérer efficacement les données provenant du terrain, de garantir leur qualité et de faciliter leur exploitation scientifique.