|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nom des dossier** | **Nom des sous-dossiers** | **Objectif du script** | **Fichiers d’entrée** | **Fichiers de sortie** |
| **1.Flore** | 1.1. Key | **2024-01-22\_key\_species\_name.Rmd**  Créé le fichier « key » avec les noms des espèces synonymes, les noms complets des espèces et la clé utilisée par MC pour le nom des espèces : 4 lettres du genre et 7 lettres de l’espèce. But : faire des jointures facilement entre les fichiers avec des noms d’espèces, notamment ceux de BDD de traits | **Flore\_500ENI\_without\_dupplicates\_with\_synonyms.csv** issu de l’archive du post-doc de MC Bopp, téléchargée du nextcloud le 8 janvier 2024. On y trouve le nom de toutes les espèces du jeu de données 500ENI et de leurs synonymes. | **key\_500\_ENI\_name\_species.csv** |
| 1.2. Description flore | **2024-02-12\_calcul\_var\_bota.Rmd**  Calcule à l’échelle de chaque relevé :   * Richesse spécifique * Indice de Shannon, Simpson et inverse simpson * % dicot * % pérennes * % néophytes * % nature-values species * % messicoles * % problematic weeds * Indices fonctionnels sur les traits foliaires | **Flore\_500ENI\_without\_dupplicates\_with\_synonyms.csv**  **key\_500\_ENI\_name\_species.csv**  **Trait\_ENI\_baseflor\_try\_V1.2.csv**  Données de traits issues des différentes bdd explorées durant le post-doc de MC Bopp, remis par G Fried en février 2024. Utilisé pour extraire le % forbs, le % néophytes et le % de pérennes  **Appendix\_species\_list\_v1.csv**  Données issues du supplementary material de l’article Fried et al., 2023 de Botany Letters, utiluisé pour extraire le % nature-value et le % messicoles  *Traçabilité changement doc Trait.docx pour tracer des changements minimes sur les doc de traits afin de permettre des jointures* | **indicateurs\_biodiversite.csv** |
| **2. Traits** | 2.1. Voj\_Lan\_floral\_diameter | **2024-01-17-\_floral\_diameter\_Voj\_Lan.Rmd**  Extraie les données de diamètre floral (à l’échelle de l’unité florale) issues des articles de Vojtko et al., 2022 (New Phytol., https://doi.org/10.1111/nph.18386) et de Lanuza et al., 2023 (Funct. Ecol., https://doi.org/10.1111/1365-2435.14340) pour les espèces représentant 1% des espèces les plus présentes dans les relevés de 500ENI, hors Poaceaes. | **data\_Lanuza\_et\_al\_2023\_all\_sp.csv**  **data\_Vojtko\_et\_al\_2022\_alpinedataset**  **List\_1pc\_most\_frequent\_species.csv**  Les deux jeux de données issues des supplementary des articles:  La liste des espèces 1% les plus fréquentes de 500ENI. Extrait de l’archive de postdoc de MC Bopp  **key\_500\_ENI\_name\_species.csv**  La clé pour joindre les fichiers | **floral\_diameter.csv**: diamètre floral pour 42 espèces du jeu de données  **missing\_sp\_for\_floral\_diameter.csv** :  Liste des 78 espèces du jeu de données pour lesquelles la donnée du diamètre floral est absente des jdd de Vojtko et Lanuza |
| 2.2. TRY\_floral\_diameter | **2024-01-19\_TRY\_floral\_diameter.Rmd**  Extraire les données de diamètre floral issues de TRY, mais après jointure il n’y a qu’une espèce commune à l’extraction TRY et aux espèces manquantes 500ENI | **missing\_sp\_for\_floral\_diameter.csv**  **key\_500\_ENI\_name\_species.csv**  Extraction TRY :  **31339.txt** | **floral\_diameter\_try.csv :**  **diamètre pour la seule espèce issue du jdd TRY (Ambrosia artemisiifolia)**  **missing\_sp\_for\_floral\_diameter\_after\_try.csv** :  Liste des 77 espèces du jeu de données pour lesquelles la donnée du diamètre floral est absente des jdd de Vojtko, Lanuza et TRY |
| 2.3. Genty\_floral\_diameter |  |  |  |
| **3. Pedoclim data** | 3.1. Cleaning soil data | **2024-01-22\_soil\_cleaning.Rmd** nettoie le nouveau fichier de données sol 2024 du réseau 500ENI et sélectionne les variables à utiliser pour les modèles | **ResultatsAnalyseSol\_500ENI\_2022\_Final\_newnames.csv**  Données sol 2024. Fichier extrait du nextcloud 500ENI le 19/01/2024. Il a été transformé en csv. Les noms des variables contenant de accents ou des caractères particuliers empêchaient l’ouverture du fichier sur R, j’ai changé manuellement les noms problématiques. Ces changements sont tracés dans le doc **Traçabilité changement doc sol.docx** | **clean\_soil\_var.csv** |
| 3.2. Clean pedoclim variables | **2024-03-21\_Clean\_pedoclimatic\_variables**  Sélectionne et convertit dans la même unité les variables communes de sol issues de Soil Grids et les mesures sur le réseau nettoyées en 4.1. Seule bulk\_density, présente dans les données issues de soil grid n’est pas présente et est donc retirée. Ajoute à ce jeu de donnée les données climatiques utilisées et extraites par Laura Henckel et MC Bopp. | **clean\_soil\_var.csv**  nouvelles données sol 2024 sélectionnées et nettoyées  **2023-07-12\_Sol\_climat.csv**  Données sol et climat utilisées par Laura et MC. Données climat venues de wordclim et sol de SoilGrids. J’en ai extraits les données climat car j’ai utilisé les données sol du fichier Agroenvironnement, uniquement pour les parcelles où l’on a pas les données sol 2024. Attention les données de température n’avaient pas de virgule et j’ai donc redivisé par 10 pour que ce soit bien degré celsius et pas en degrés celsuis\*10.  **Agroenvironnement.csv**  Données sol, climat et descriptives de la parcelle extraites de l’archive de postdoc de MC Bopp. | **Agroenvironnement\_new.csv**  Fichier agroenvironnement avec les données de sol directement mesurées (Soil\_Nitrogen,Soil\_CEC,Soil\_Clay\_content,Soil\_Ph\_h2o,Soil\_organic\_carbon\_content,Soil\_Sand\_content,Soil\_Silt\_content) pour les 427 parcelles où cela a été mesuré et les données climatiques. Les données de sol pour les autres parcelles sont celles issues de SoilGrids qui ont été conservées. J’ai retiré les axes des ACPs issus du post-doc de Marie-Charlotte. |
| 3.3. Field clustering | **2024-25-01\_field\_clustering.Rmd**  Regroupe les parcelles en clusters pédoclimatiques et paysagers pour pouvoir mener les analyses par sous-groupes pédoclim-paysagers.  Créé également un fichier clean et imputé (pour les quelques données manquantes) de toutes les variables agricoles, climatiques et édaphiques. | **Agroenvironnement\_new.csv**  **dea\_type\_prod\_traite\_Lea.csv**  Bio ou conventionnel. Extrait du nextcloud (« export juin 2023 ») le 24/01/2024 traité selon le doc Traçabilite modifications dea\_type\_prod.docx | **pedoclim\_clusters.csv**  Id de chaque parcelle associée à son numéro de cluster, sa région biogéo d’après VégétalLocal et les variables sol, climat et paysage qui ont été utilisées pour créer les clusters  **cluster\_number.csv**  Id\_parcelle associé à son numéro de cluster  **clusters\_type\_culture.csv**  Nombre de parcelles dans chaque type de culture par clusters  **var\_agroenvi\_input.csv**  Fichier clean et imputé des variables agricoles, climatiques et édaphiques |
| **4. Spatial autocorrelation** |  | **2024-03-26\_spatial\_autocorr.Rmd** | **dataset\_500\_ENI\_Lea.csv**  (voir 5.1.) |  |
| **5. Models\_par\_clusters\_pedoclimatiques** | 5.1. Fusion | **2024-02-13\_fusion\_all\_dataset.Rmd**  Fusionne toutes les données et sélectionne les variables afin de créer le fichier final sur lequel l’analyse par clusters sera menée. Les 11 parcelles sans données de fertilisation sont retirées. | **data\_500ENI\_traits\_floraux.csv**  **dea\_type\_prod\_traite\_Lea.csv**  **cluster\_number.csv**  issu de 4.4.  **indicateurs\_biodiversite.csv**  Issu de 1.2.  **TableFinalPlante2.csv**  Doc issu du travail de Laura Henckel, pour extraire les données de fertilisation uniquement  **parcelle\_adonis.csv**  Données Solagro des IFTs moyennes à l’échelle des communes liées à nos parcelles  **var\_agroenvi\_input.csv** | **dataset\_500\_ENI\_Lea.csv** |
| 5.2. Models | **Glmms\_pedoclim\_clusters.Rmd**  Performe les glmm sur les indicateurs de biodiversité végétale par cluster | **dataset\_500\_ENI\_Lea.csv** |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Tableau « poubelle »

|  |  |  |  |
| --- | --- | --- | --- |
| 4.2. Pedoclimatic PCA with soil variables | **2024-01-22\_pca\_pedoclimatic.Rmd**  Reprend les données climatiques de 500ENI utilisées par MC Bopp + les nouvelles données sol mesurées sur le terrain. Sélectionne et convertit dans la même unité les variables communes. Seule bulk\_density, présente dans les données issues de soil grid n’est pas présente et est donc retirée. Runne l’ACP pédoclimatique et en extraie les axes pour décrire les conditions pédoclimatiques sur le réseau. Pour les parcelles où les données sol n’ont pas été mesurées in situ, les anciennes données sol sont conservées. | **clean\_soil\_var.csv**  nouvelles données sol 2024 sélectionnées et nettoyées  **2023-07-12\_Sol\_climat.csv**  Données sol et climat utilisées par Laura et MC. Données climat venues de wordclim et sol de SoilGrids. Je n’en ai extraits que les données climat car j’ai utilisé les données sol du fichier Agroenvironnement, uniquement pour les parcelles où l’on a pas les données sol 2024.  **Agroenvironnement.csv**  Données sol, climat et descriptives de la parcelle extraites de l’archive de postdoc de MC Bopp. | **Agroenvironnement\_new.csv**  Fichier agroenvironnement avec les données de sol directement mesurées (Soil\_Nitrogen,Soil\_CEC,Soil\_Clay\_content,Soil\_Ph\_h2o,Soil\_organic\_carbon\_content,Soil\_Sand\_content,Soil\_Silt\_content) pour les 427 parcelles où cela a été mesuré et les données climatiques. Les données de sol pour les autres parcelles sont celles issues de SoilGrids qui ont été conservées. J’ai retiré les axes des ACPs issus du post-doc de Marie-Charlotte.  **dim\_pedoclim\_pca\_2024.csv**  Coordonnées des parcelles sur les 3 premiers axes de l’ACP pédoclimatique |
| 4.3. Saving new pedoclim PCA axes | **2024-01-24\_new\_pca\_axes.Rmd**  Remplace les anciens axes de l’ACP dans le fichier fusionné des traits floraux | **dim\_pedoclim\_pca\_2024.csv**  **data\_500ENI\_traits\_floraux.csv** | **data\_500ENI\_traits\_floraux\_2024.csv**  Jeu de données data\_500ENI\_traits\_floraux.csv avec les nouveaux axes d’ACP |