

RAPPORT DE STAGE

MEKKI Ryan



Département MIASHS, UFR 6 Informatique, Mathématique et Statistique
Université Paul Valéry, Montpellier 3

Juin 2025

RAPPORT DE STAGE CIRAD UMR SENS

Remerciements

Mes plus sincères remerciements vont à mes tutrices de stage Madame COUDEL Emilie et Madame LEBRE Sophie pour les conseils avisés sur mon travail.

15/06/2025.

Résumé

Ce rapport présente le travail réalisé au cours d'un stage de recherche au CIRAD, au sein de l'UMR SENS, dans le cadre du projet *Odyssea* portant sur les effets socio-environnementaux de l'expansion du soja en Amazonie brésilienne. L'objectif principal était d'analyser les profils types d'agriculteurs en fonction de leur recours à l'agroécologie et à l'usage de pesticides, en mobilisant une base de données issue d'une enquête par questionnaire menée auprès de producteurs locaux. Après une phase rigoureuse de nettoyage, de recodage et de création de variables synthétiques (groupe agro/pesticides, participation sociale, etc.), une typologie en quatre groupes a été définie. Ces groupes ont été comparés à l'aide d'analyses descriptives, de tests statistiques (Kruskal-Wallis et χ^2), ainsi que de visualisations graphiques (barres empilées, histogrammes, ACP mixte, clustering). Les résultats mettent en évidence des différences marquées entre les groupes, tant sur les pratiques agricoles que sur les dimensions sociales et territoriales. Les groupes pratiquant l'agroécologie se distinguent par une plus forte participation communautaire, une plus grande distance aux cultures de soja et des perceptions plus critiques des effets environnementaux du modèle productiviste. Les tests statistiques confirment que certaines variables, comme l'usage de produits chimiques ou la participation sociale, sont fortement discriminantes. Des perspectives de prolongement ont été proposées, incluant des modèles statistiques avancés, une approche temporelle, une valorisation cartographique et des applications pour les politiques publiques. Ce travail constitue ainsi un socle analytique robuste pour éclairer les dynamiques agricoles locales et contribuer à la réflexion sur les transitions agroécologiques.

Table des matières

Chapitre 1	Introduction	1
Chapitre 2	Problématique et enjeux pour la structure d'accueil	3
Chapitre 3	Présentation du travail réalisé	4
3.1	Nettoyage et préparation des données	4
3.1.1	Harmonisation des noms de colonnes	4
3.1.2	Recodage de certaines variables (e.g. pratiques agroécologiques, produits chimiques, revenus)	4
3.2	Construction de variables synthétiques	6
3.2.1	Variable groupe : classification des individus en 4 groupes selon l'usage de pesticides et la pratique de l'agroécologie . .	6
3.2.2	Variable niveau_participation : participation sociale (leadership, participation simple, aucune)	7
3.2.3	Variables sur l'activité dans la propriété, embauche de main-d'œuvre, cohabitation/travail	7
3.3	Analyse descriptive	7
3.3.1	Tableaux croisés par groupe (groupe) pour les variables d'intérêt	7
3.3.2	Calculs de pourcentages et moyennes (avec écart-type) selon les groupes	8
3.3.3	Résultats pour la variable : Le soja a porté atteinte à la santé familiale	8
3.3.4	Tests de significativité (e.g. Kruskal-Wallis pour les variables numériques)	9
3.3.5	Tests d'indépendance (χ^2) pour les variables qualitatives . .	10
3.4	Visualisation et exploration initiale	11
3.4.1	Graphiques	11
Chapitre 4	Visualisation, exploration initiale et interprétation des résultats	23
4.1	Participation sociale selon les groupes	23
4.2	Distribution de la distance au soja	23
4.3	Dispersion des groupes selon l'ACM	24
4.4	Clustering des individus (K-means sur ACPM)	24
4.5	Interprétation initiale des résultats	25
Chapitre 5	Perspectives ou travail encore à faire	26
5.0.1	Perspectives à court terme	26
5.0.2	Approfondissement des analyses statistiques	26
5.0.3	Renforcement de la dimension temporelle	26

5.0.4	Valorisation cartographique	27
5.0.5	Intégration de variables supplémentaires	27
5.0.6	Implications pour les politiques publiques	27
Chapitre 6 Conclusion		28
6.1	Récapitulatif des apports du travail	28
6.2	Réponse à la problématique	28
6.3	Apports pour la structure d'accueil (UMR SENS / CIRAD)	29
6.4	Apprentissage personnel	30
6.5	Conclusion	30

CHAPITRE 1

Introduction

Le Centre de coopération internationale en recherche agronomique pour le développement (CIRAD) est un organisme public français dédié à la recherche agronomique au service des pays du Sud. Il dispose de nombreuses implantations à travers le monde, en lien avec ses partenaires de recherche, et regroupe plusieurs unités mixtes impliquées dans l'analyse des dynamiques agricoles, environnementales et sociales.

Mon stage s'est déroulé au sein de l'unité mixte de recherche UMR SENS (Savoirs, Environnements et Sociétés), une équipe pluridisciplinaire réunissant des chercheurs.e.s en sciences sociales et en sciences de l'environnement. Cette unité travaille notamment sur les transformations des systèmes agricoles, les innovations socio-techniques, et les liens entre pratiques agricoles, politiques publiques, environnement et société. Elle développe une approche intégrée fondée à la fois sur des méthodes qualitatives (entretiens, enquêtes de terrain, observation) et quantitatives (traitement de données, analyses statistiques multivariées, spatialisation, etc.), dans une perspective critique de recherche-action.

Ce stage s'inscrit dans le cadre du projet Odysee, une enquête territoriale menée entre 2022 et 2023 dans plusieurs communes de l'État du Pará (Brésil), dont Santarém et Mojuí dos Campos. Le projet vise à documenter les effets sociaux, environnementaux et économiques de l'expansion du soja dans les zones agricoles périurbaines, et à analyser les formes de résistance, d'adaptation ou d'innovation portées par les producteurs. Dans ce cadre, un important travail de collecte de données a été réalisé via des questionnaires détaillés, permettant de croiser des variables relatives aux pratiques agricoles, aux formes de participation sociale, à l'exposition aux intrants chimiques ou encore aux représentations des effets du soja.

L'objectif principal de ce stage est d'explorer et d'analyser les profils types d'acteurs dans cette région, à partir d'une base de données issue de cette enquête. Il s'agit notamment d'identifier des groupes distincts selon leur rapport à l'agroécologie et aux pesticides, et d'examiner leurs caractéristiques sociales, géographiques et perceptives. L'analyse mettra l'accent sur la manière dont les variables individuelles, sociales et territoriales s'articulent, et sur ce que cela révèle des dynamiques agro-environnementales en cours dans la région étudiée.

Structure du rapport. Ce rapport est structuré comme suit :

- Le **chapitre 2** présente la problématique du stage et les enjeux associés pour la structure d'accueil ;
- Le **chapitre 3** détaille le travail effectué, depuis le nettoyage des données jusqu'à la construction de variables synthétiques et les premières analyses descriptives ;

- Le **chapitre 4** est consacré à l'exploration des résultats à l'aide de représentations graphiques, d'analyses factorielles et de clustering, permettant de mieux interpréter la structuration des groupes ;
- Le **chapitre 5** propose des perspectives de prolongement du travail, en lien avec les attentes du projet Odyssea ;
- Le **chapitre 6** conclut le rapport en revenant sur les principaux apports du stage, la réponse à la problématique, les bénéfices pour l'UMR SENS et les apprentissages personnels.

CHAPITRE 2

Problématique et enjeux pour la structure d'accueil

Dans un contexte de transition agroécologique, les pratiques agricoles durables sont de plus en plus encouragées par les politiques publiques et les ONG. Toutefois, la mise en œuvre de l'agroécologie sur le terrain dépend fortement des conditions locales, des ressources disponibles, et de l'implication des acteurs ruraux. Comprendre les facteurs sociaux, économiques et organisationnels qui favorisent ou freinent l'adoption de ces pratiques constitue un enjeu central pour les chercheurs, les décideurs publics, et les acteurs du développement.

L'équipe UMR SENS, en partenariat avec plusieurs acteurs locaux, mène une étude approfondie sur les trajectoires de production dans une région d'Amazonie brésilienne, où coexistent différentes formes d'agriculture, dont certaines relevant de l'agroécologie. Cette étude repose sur l'exploitation d'un jeu de données conséquent issu d'enquêtes de terrain menées auprès de producteurs. Les données couvrent des thématiques variées : pratiques agricoles, accès au foncier, perception des impacts environnementaux, revenus, organisation sociale, etc.

Le travail sur ces données vise à identifier des typologies d'agriculteurs, à comprendre leur rapport à l'agroécologie, et à analyser les effets perçus de ces pratiques sur la santé, l'environnement, ou encore les dynamiques sociales (emploi, participation à des collectifs, leadership). Ces analyses ont pour objectif de fournir à l'équipe des éléments de compréhension solides pour alimenter les réflexions scientifiques, mais aussi pour éclairer les politiques de soutien à l'agroécologie.

CHAPITRE 3

Présentation du travail réalisé

3.1 Nettoyage et préparation des données

La phase de préparation des données a été une étape essentielle avant toute analyse statistique ou exploration factorielle. Elle visait à rendre le jeu de données exploitable, cohérent et pertinent au regard de la problématique.

3.1.1 Harmonisation des noms de colonnes

Dans un premier temps, un travail de sélection des variables a été mené. Il s’agissait d’isoler les variables d’intérêt en lien avec les thématiques agricoles, environnementales et sociales étudiées. On a notamment retenu des variables qualitatives portant sur les pratiques culturelles (comme l’usage d’herbicides ou d’insecticides), l’emploi de main-d’œuvre, la pratique de l’agroécologie, l’exposition aux pesticides, la participation à des groupes sociaux, ou encore les perceptions liées aux impacts du soja sur l’environnement, la santé ou la qualité de l’eau.

Ensuite, un nettoyage des données a été effectué. Cela a inclus l’uniformisation des modalités pour les variables catégorielles, notamment en harmonisant les réponses similaires (ex : “non”, “Non”, ou cases vides ont été traitées de manière cohérente). Les variables contenant trop de valeurs manquantes ont été écartées, mais cette décision a été prise au cas par cas. Aucune exclusion automatique n’a été appliquée sur la base d’un seuil arbitraire.

3.1.2 Recodage de certaines variables (e.g. pratiques agroécologiques, produits chimiques, revenus)

Un point particulier a été porté sur les variables décrivant la participation sociale des individus. Sept variables décrivaient la participation à différents types de groupes : groupe de femmes, syndicat, groupe de jeunes, coopération, communauté associative, église, club de football. À partir de ces informations, une variable synthétique a été créée pour classer chaque individu selon trois niveaux : ceux qui exercent un rôle de leadership, ceux qui participent sans occuper de position de coordination, et ceux qui ne participent à aucun groupe. Cette variable “niveau de participation” s’est révélée importante dans la suite de l’analyse.

Une autre étape importante a été l’élaboration de la variable de regroupement des individus selon leur profil agroécologique. Quatre groupes ont été définis selon deux critères croisés : la pratique ou non de l’agroécologie, et l’usage ou non de produits phytosanitaires. Ces groupes permettent de distinguer, par exemple, les individus en agroécologie sans pesticides, ceux qui ne pratiquent pas l’agroécologie mais utilisent des intrants chimiques, etc. Cette classification a permis de structurer les analyses comparatives et d’orienter les interprétations ultérieures.

La variable Q_6.3.1.3_Pratique.de.l.agroecologie visait à identifier les individus déclarant pratiquer l'agroécologie. Cependant, elle contenait des réponses manquantes ou ambiguës, en particulier chez les personnes n'ayant jamais entendu parler de l'agroécologie (variable Q_6.3.1_entendu_parler_agroecologie).

Pour clarifier cette information, une nouvelle variable binaire, Q_6.3.1.3_Pratique.de.l.agroecologie.corrige, a été construite selon la logique suivante :

- Si un individu n'a jamais entendu parler d'agroécologie, alors il est considéré comme ne pratiquant pas l'agroécologie.
- Dans les autres cas, on reprend la réponse initiale s'il y a une mention explicite de pratique.
- Si la réponse est vide mais que la personne dit avoir entendu parler de l'agroécologie, la valeur est laissée manquante (NA), afin d'éviter toute interprétation abusive.

Ce recodage permet d'éviter des biais d'interprétation dans les analyses comparatives, en garantissant que l'absence de connaissance exclut logiquement la pratique.

```
##          Pratique
## Entendu Non Oui <NA>
##           0   0   2
##      Non 336   0   0
##      Oui  83 113  10

##          Pratique
## Pesticides Non Oui <NA>
##         non 206 73   7
##         oui 213 40   5

##                insecticide
## herbicide           Jamais Tous les 2-3 mois
##                1         2             0
##      Jamais           0         3             5
##      Tous les 2-3 mois 0         1             0

##          Pratique
## Entendu Non Oui <NA>
##           0   0   2
##      Non 336   0   0
##      Oui  83 113  10

##          agroécologie
## insecticides Non Oui
##         non 218 73
##         oui 213 40
```

Enfin, pour les besoins de certaines analyses comme les représentations graphiques et les analyses factorielles, les variables ont été reformulées ou réétiquetées afin d'améliorer leur lisibilité. Certaines modalités ont été regroupées pour simplifier l'analyse, tout en conservant la richesse informative nécessaire.

En somme, cette étape de préparation a permis de partir d'un fichier brut complexe pour aboutir à une base de données propre, cohérente, et adaptée aux objectifs

de l'étude. Elle a constitué un prérequis indispensable à l'interprétation pertinente des résultats qui seront présentés dans les sections suivantes.

3.2 Construction de variables synthétiques

L'une des étapes clés de notre préparation des données a consisté à construire des variables synthétiques. Contrairement aux variables directement issues du questionnaire, ces variables sont obtenues à partir de plusieurs questions et permettent de résumer des informations complexes de manière plus lisible et exploitable dans les analyses statistiques. Elles facilitent notamment la création de profils types, la segmentation des répondants et la comparaison entre groupes.

Trois variables principales ont été construites, détaillées ci-dessous.

3.2.1 Variable groupe : classification des individus en 4 groupes selon l'usage de pesticides et la pratique de l'agroécologie

Cette variable a été construite dans le but de regrouper les individus selon leur profil agroécologique. Elle repose sur deux critères essentiels : d'une part, l'usage ou non de produits phytosanitaires (comme les herbicides et les insecticides), et d'autre part, la pratique ou non de l'agroécologie.

Pour cela, nous avons d'abord recodé la variable relative à l'utilisation de produits chimiques agricoles. Les modalités initiales ont été simplifiées en deux catégories : "oui" si l'individu a déclaré utiliser des produits chimiques, et "non" dans les autres cas. Ensuite, la variable concernant la pratique agroécologique a été corrigée pour tenir compte uniquement des personnes ayant déclaré une pratique effective (et non une simple connaissance du terme). Cela a permis d'exclure les réponses vides ou ambiguës, en distinguant clairement ceux qui pratiquent réellement une forme d'agroécologie.

À partir du croisement de ces deux dimensions – l'usage de produits chimiques et la pratique agroécologique – nous avons défini une variable à quatre modalités. Chaque individu a ainsi été classé dans l'un des groupes suivants :

- Agro / Sans Pesticides : pratique de l'agroécologie sans usage d'intrants chimiques ;
- Sans Agro / Sans Pesticides : ni pratique agroécologique ni usage d'intrants ;
- Sans Agro / Avec Pesticides : recours aux produits chimiques sans agroécologie ;
- Agro / Avec Pesticides : combinaison des deux pratiques (agroécologie + intrants).

```
## [1] Sans Agro / Sans Pesticides Sans Agro / Sans Pesticides
## [3] Sans Agro / Sans Pesticides Sans Agro / Sans Pesticides
## [5] Sans Agro / Sans Pesticides Sans Agro / Avec Pesticides
## 4 Levels: Agro / Sans Pesticides ... Agro / Avec Pesticides
```

Cette variable synthétique, que nous avons appelée "groupe", a joué un rôle central dans les comparaisons réalisées dans la suite de l'analyse. Elle nous a permis de mieux caractériser les profils agricoles et d'identifier les différences d'attitudes ou de conditions sociales associées à chaque configuration.

3.2.2 Variable *niveau_participation* : participation sociale (leadership, participation simple, aucune)

Cette variable a été construite afin de résumer le degré d’implication sociale des individus dans différents types de collectifs ou d’organisations. Le questionnaire initial comportait plusieurs questions distinctes portant sur la participation à divers groupes : groupe de femmes, syndicat, groupe de jeunes, coopérative, communauté associative, église, club de football, etc. Chaque individu pouvait ainsi indiquer, pour chacun de ces groupes, s’il y participait, s’il y exerçait un rôle de coordination ou s’il n’y participait pas du tout.

Afin de synthétiser ces informations dispersées, nous avons créé une variable catégorielle regroupant les individus selon trois niveaux d’implication :

- Leadership : l’individu exerce un rôle de direction, de coordination ou de représentation dans au moins un des groupes ;
- Participation : l’individu participe à un ou plusieurs groupes, mais sans fonction de coordination ;
- Aucune : l’individu ne participe à aucun groupe, ou a répondu “Non” à toutes les questions relatives à la participation.

```
## [1] Leadership      Participation Participation Participation Participation
## [6] Participation
## Levels: Leadership Participation Aucune
```

Cette variable “niveau_participation” permet ainsi de saisir en un seul indicateur la dimension de l’engagement collectif ou associatif des individus. Elle est particulièrement pertinente pour évaluer le lien entre pratiques agricoles, engagement communautaire et dynamiques sociales dans les territoires étudiés.

3.2.3 Variables sur l’activité dans la propriété, embauche de main-d’œuvre, cohabitation/travail

L’ensemble de ce travail de préparation a permis d’enrichir la base de données avec des variables clés, mieux adaptées aux objectifs de notre analyse. Grâce aux recodages, aux regroupements de modalités et à la construction de variables synthétiques, nous disposons désormais d’indicateurs cohérents et directement mobilisables pour explorer les profils des individus.

Dans la section suivante, nous allons tirer parti de cette base structurée pour effectuer une première série d’analyses exploratoires. Nous chercherons à caractériser les différents groupes d’individus selon leurs réponses aux variables qualitatives, en identifiant les fréquences, les associations et les profils types. Cela permettra de mieux comprendre les spécificités des groupes agroécologiques identifiés précédemment, et de dégager les premiers axes d’interprétation.

3.3 Analyse descriptive

3.3.1 Tableaux croisés par groupe (groupe) pour les variables d’intérêt

Cette section présente un tableau de synthèse comparant les quatre groupes d’individus définis selon la combinaison de deux critères : la pratique ou non de l’agroécologie, et l’utilisation ou non de pesticides. Cette classification a donné lieu à quatre groupes distincts :

- Groupe 1 : agroécologie sans pesticides
- Groupe 2 : sans agroécologie, sans pesticides
- Groupe 3 : sans agroécologie, avec pesticides
- Groupe 4 : agroécologie avec pesticides

Pour chacun de ces groupes, plusieurs indicateurs ont été calculés afin de fournir une première caractérisation descriptive :

- **n** : nombre d'individus dans le groupe. Cela permet de vérifier la taille relative de chaque sous-population.
- **pct_agro** et **pct_pesticide** : rappels des modalités structurantes du groupe, avec 1 indiquant la présence de la caractéristique pratique agroécologique ou usage de pesticides, et 0 son absence.
- **munic_mode** : commune municipalité la plus fréquente au sein du groupe. On remarque par exemple que les groupes 1, 3 et 4 sont principalement localisés à Mojuí dos Campos, tandis que le groupe 2 est majoritairement basé à Santarém.
- **zone_mode** : zone géographique la plus représentée dans chaque groupe. Exemples : Centre de légumes, Cipóal, PA Jabuti, etc.

L'objectif de ce tableau est de mettre en évidence des tendances globales et de vérifier la cohérence des profils. On observe par exemple que les groupes en agroécologie (groupes 1 et 4) ont en moyenne des revenus légèrement inférieurs, que le groupe utilisant des pesticides sans agroécologie (groupe 3) est également économiquement plus fragile, ou encore que certaines zones géographiques sont associées de manière privilégiée à certains profils.

Cette étape permet ainsi de poser les bases d'une comparaison plus fine à travers les indicateurs statistiques présentés dans les sections suivantes.

Table 3.1: Tableau de synthèse des groupes

groupe	n	pct_agro	pct_pesticide	munic_mode	zone_mode
Agro / Sans Pesticides	73	1	0	Mojuí dos Campos	M1 Centre de legumes
Sans Agro / Sans Pesticides	218	0	0	Santarem	S1 Cipoal
Sans Agro / Avec Pesticides	213	0	1	Mojuí dos Campos	M4 PA Jabuti
Agro / Avec Pesticides	40	1	1	Mojuí dos Campos	M5 Bonne chance (farine)

3.3.2 *Calculs de pourcentages et moyennes (avec écart-type) selon les groupes*

Pour approfondir la caractérisation des groupes, des indicateurs quantitatifs ont été calculés pour plusieurs variables d'intérêt. Il s'agit notamment de moyennes accompagnées de leur écart-type, calculées pour chaque groupe défini précédemment.

Par exemple, la variable Q_9.4.1.5_soja_a_porte_atteinte_a_la_sante_familiale, qui mesure la perception d'un impact négatif du soja sur la santé familiale, a été analysée ainsi :

3.3.3 *Résultats pour la variable : Le soja a porté atteinte à la santé familiale*

Table 3.2: Moyenne et écart-type pour : Le soja a porté atteinte à la santé familiale

groupe	moyenne	ecart_type
Agro / Sans Pesticides	0.25	0.43
Sans Agro / Sans Pesticides	0.17	0.38
Sans Agro / Avec Pesticides	0.11	0.32
Agro / Avec Pesticides	0.15	0.36

Test de Kruskal-Wallis p-value : 0.0502

- Le groupe 1 (agro sans pesticides) présente une moyenne de 0.25, avec un écart-type de 0.43, ce qui indique une proportion relativement importante de personnes percevant un effet négatif, bien que la dispersion des réponses soit notable.
- Le groupe 2 (sans agro, sans pesticides) affiche une moyenne plus faible, 0.17, avec un écart-type de 0.38.
- Le groupe 3 (sans agro, avec pesticides) présente la moyenne la plus faible (0.11), suggérant que peu de personnes dans ce groupe déclarent un impact du soja sur la santé familiale.
- **munic_mode** : commune municipalité la plus fréquentLe groupe 4 (agro avec pesticides) se situe entre les groupes 2 et 3, avec une moyenne de 0.15.e au sein du groupe. On remarque par exemple que les groupes 1, 3 et 4 sont principalement localisés à Mojuí dos Campos, tandis que le groupe 2 est majoritairement basé à Santarém.

Ces résultats suggèrent que les individus pratiquant l'agroécologie, même sans usage de pesticides, perçoivent davantage d'impacts sanitaires liés au soja, comparativement aux autres groupes. Cela peut refléter une sensibilité accrue à ces enjeux chez les personnes engagées dans des pratiques alternatives, ou un effet de contexte local.

Ce type d'analyse, croisant les valeurs moyennes et les écarts-types, permet de détecter des tendances intéressantes et oriente les tests statistiques qui seront présentés dans la section suivante.

3.3.4 Tests de significativité (e.g. Kruskal-Wallis pour les variables numériques)

Afin de déterminer si les différences observées entre les groupes définis par la variable groupe (agroécologie et usage de pesticides) étaient statistiquement significatives, nous avons appliqué le test non paramétrique de Kruskal-Wallis sur plusieurs variables numériques d'intérêt.

Ce test permet de comparer la distribution d'une variable continue ou ordinale entre plus de deux groupes sans faire l'hypothèse de normalité. Il est particulièrement adapté à notre contexte, où les distributions sont souvent asymétriques ou présentent des valeurs extrêmes.

Résultats des tests de Kruskal-Wallis par variable

- ** Q_9.4.1.5_soja_apporte_une_mauvaise_qualite_de_l_eau ** : p-value = 0.447

- ** Q_9.4.1.5_soja_a_endommagement_l'environnement ** : p-value = 0.00145
- ** Q_9.4.1.5_soja_a_apporte_des_dommages_au_changement_climatique ** : p-value = 0.07
- ** Q_9.5.2_future5ans_veut_continuer_a_renforcer ** : p-value = 0.624
- ** Q_9.4.1.5_soja.a.entraîne.une.perte.de.production ** : p-value = 0.757
- ** Q_11_Dist_soja ** : p-value = 0.0834

Voici quelques résultats clés :

- Q_9.4.1.5_soja_apporte_une_mauvaise_qualite_de_l'eau : la p-value est de 0,0502, frôlant le seuil de significativité conventionnel de 5 %. Cela suggère une différence potentielle entre les groupes, mais elle reste marginale.
- Q_9.4.1.5_soja_a_endommagement_l'environnement : avec une p-value de 0,447, aucune différence significative n'est détectée entre les groupes sur cette perception environnementale.
- Q_9.4.1.5_soja_a_apporte_des_dommages_au_changement_climatique : la p-value est de 0,00145, ce qui indique une différence hautement significative entre les groupes. Il existe donc un lien fort entre le profil agro/pesticides des individus et leur perception des effets climatiques du soja.
- Q_9.5.2_future5ans_veut_continuer_a_renforcer : la p-value de 0,07 ne permet pas de conclure à une différence significative, bien qu'un écart modéré soit observable.
- Q_9.4.1.5_soja.a.entraîne.une.perte.de.production et Q_11_Dist_soja : les p-values de 0,624 et 0,757 respectivement montrent qu'il n'existe pas de différence significative entre les groupes pour ces variables.

En résumé, seule une minorité de variables numériques présentent des écarts statistiquement significatifs selon les groupes. Cela confirme que certains effets perçus du soja sont liés à des profils agroécologiques différenciés, tandis que d'autres perceptions sont partagées de manière plus homogène dans la population étudiée.

3.3.5 Tests d'indépendance (χ^2) pour les variables qualitatives

Afin d'examiner si certaines variables qualitatives diffèrent significativement entre les groupes agro/pesticides, nous avons appliqué des tests du χ^2 d'indépendance sur une sélection de variables catégorielles pertinentes. Ce test permet de vérifier s'il existe une dépendance statistique entre chaque variable et l'appartenance à l'un des quatre groupes définis.

Le tableau ci-dessous présente les résultats, classés par valeur décroissante de la statistique de test :

Table 3.3: Résultats des tests du χ^2 d'indépendance

	variable	statistic	p.value	df
X-squared5	Q_6.2.1_Produits_chimiques_binaire	544.000	1.39e-117	3
X-squared4	Q_6.3.1.3_Pratique.de.l.agroecologie	544.000	1.39e-117	3
X-squared9	Q_9.2.1.2_quelles_nouvelles_activites	373.989	1.24e-02	315
X-squared6	Q_6.1.4.3_herbicide	143.189	1.36e-24	12
X-squared2	Q_5.1.2_type_culture	140.366	2.93e-01	132
X-squared1	Q_2.2.1_zone	88.727	5.32e-07	33
X-squared7	Q_6.1.4.2_Insecticides_chimiques	81.857	1.82e-12	12
X-squared8	Q_6.1.2_emploi_main_d.oeuvre	37.995	1.54e-04	12
X-squared	Q_2.2_municipe	23.123	7.56e-04	6
X-squared3	type_culture_simplifie	7.966	2.41e-01	6

Les résultats indiquent que les deux variables les plus discriminantes sont :

- Q_6.2.1_Produits_chimiques_binaire
- Q_6.3.1.3_Pratique.de.l.agroecologie

Ces deux variables présentent une statistique de test très élevée et une p-value extrêmement faible (10^{-16}), confirmant leur rôle central dans la structure des groupes.

D'autres variables comme l'usage d'herbicides, la participation à des activités collectives (Q_9.2.1.2_quelles_nouvelles_activites), ou encore la zone géographique (Q_2.2.1_zone) contribuent également significativement à différencier les profils. À l'inverse, certaines variables (comme type_culture_simplifie) montrent une dépendance plus faible.

Ces résultats confirment les conclusions visuelles obtenues précédemment à travers les heatmaps et les profils types par groupe, en leur ajoutant une validation statistique formelle.

3.4 Visualisation et exploration initiale

Après les traitements statistiques préliminaires, une phase de visualisation a été engagée afin de mieux appréhender les profils types issus des quatre groupes définis selon la pratique de l'agroécologie et l'usage de pesticides. Ces explorations graphiques et tabulaires ont servi à révéler des tendances générales et à formuler les premières hypothèses d'interprétation.

3.4.1 Graphiques

Cette phase vise à illustrer graphiquement les différences entre les groupes définis à partir des pratiques agroécologiques et de l'usage des pesticides, et à en tirer une première interprétation.

Graphique 1 — Niveau de participation sociale par groupe

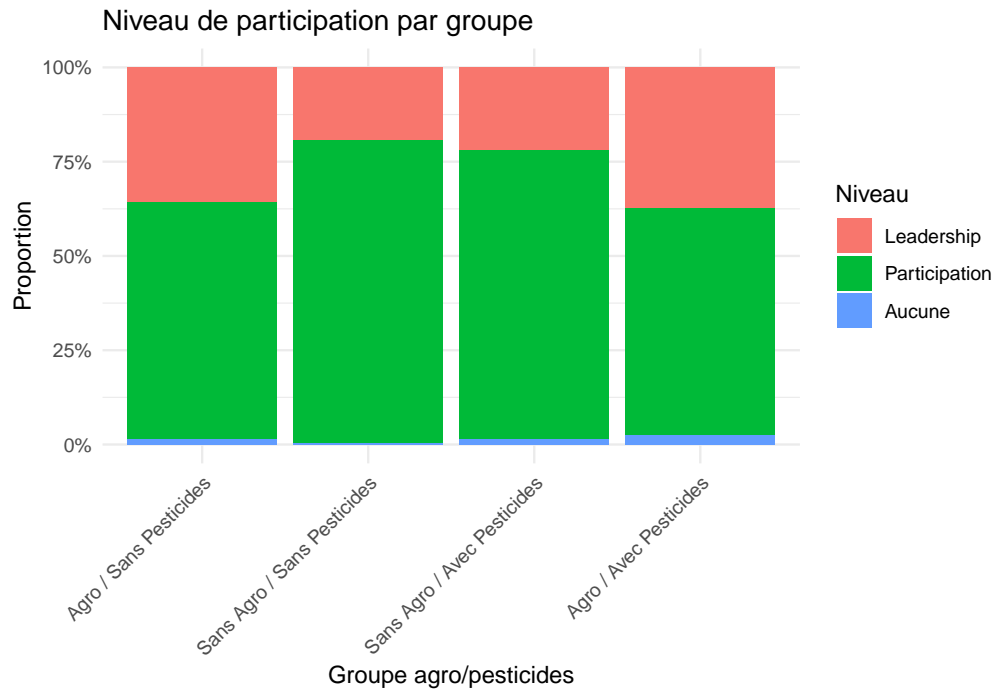


Figure 3.1: Niveau de participation par groupe

Le graphique en barres empilées met en évidence les différences de participation sociale selon les groupes. On observe que les groupes « Agro sans pesticides » et « Agro avec pesticides » comptent davantage d'individus en position de leadership (35,6 % et 37,5 % respectivement), tandis que les groupes « Sans agroécologie » présentent des taux plus élevés de participation simple (80,3 % et 76,5 %). Le niveau « aucune participation » reste marginal, ne dépassant pas 2,5 % dans aucun des groupes.

Ce graphique suggère un lien entre pratiques agroécologiques et implication dans les structures collectives, les groupes en agroécologie étant plus représentés dans les fonctions de coordination ou de responsabilité.

Graphique 2 — Histogramme de la distance au soja (Q_11_Dist_soja)

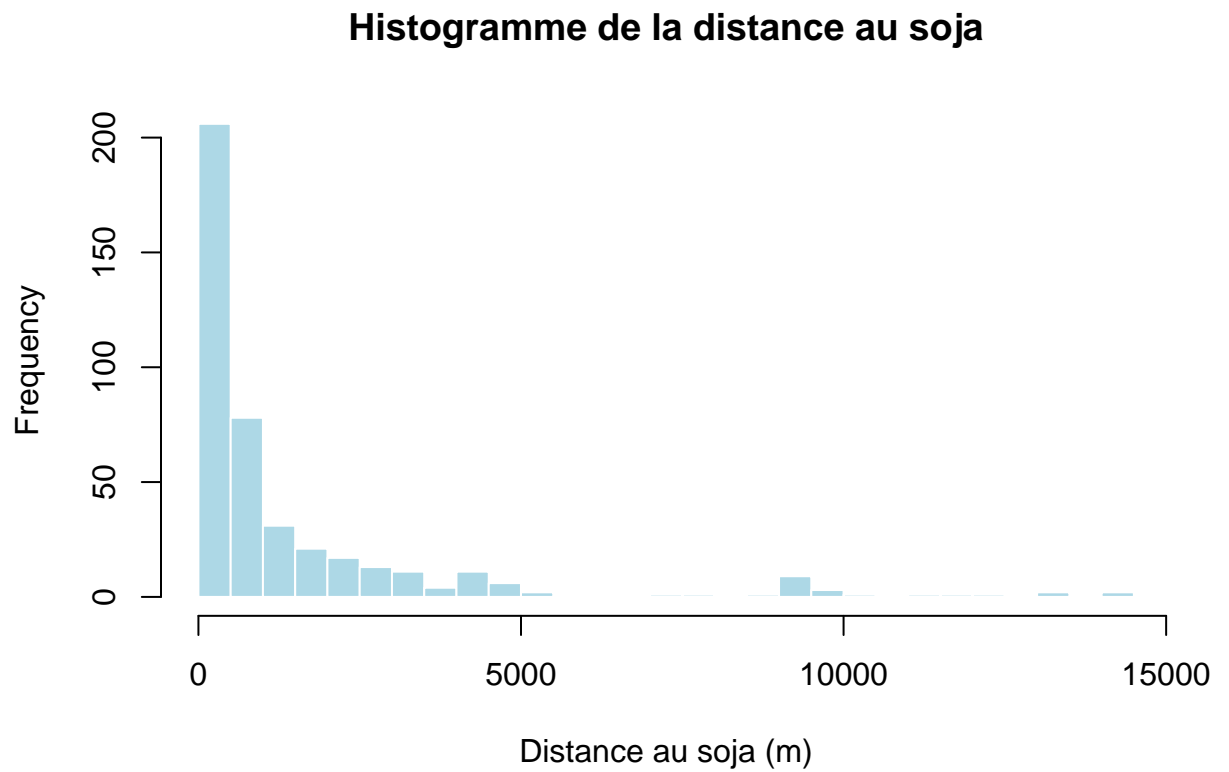


Figure 3.2: Histogramme de la distance au soja

L’histogramme général de la variable de distance au soja indique une distribution fortement asymétrique à droite, avec une concentration marquée d’observations proches des distances nulles. Cela reflète la présence de nombreuses habitations situées à proximité immédiate des cultures de soja.

Graphiques 3, 4, 5 et 6 — Histogrammes par groupe (distance au soja)

Histogramme – Sans Agro / Sans Pesticides

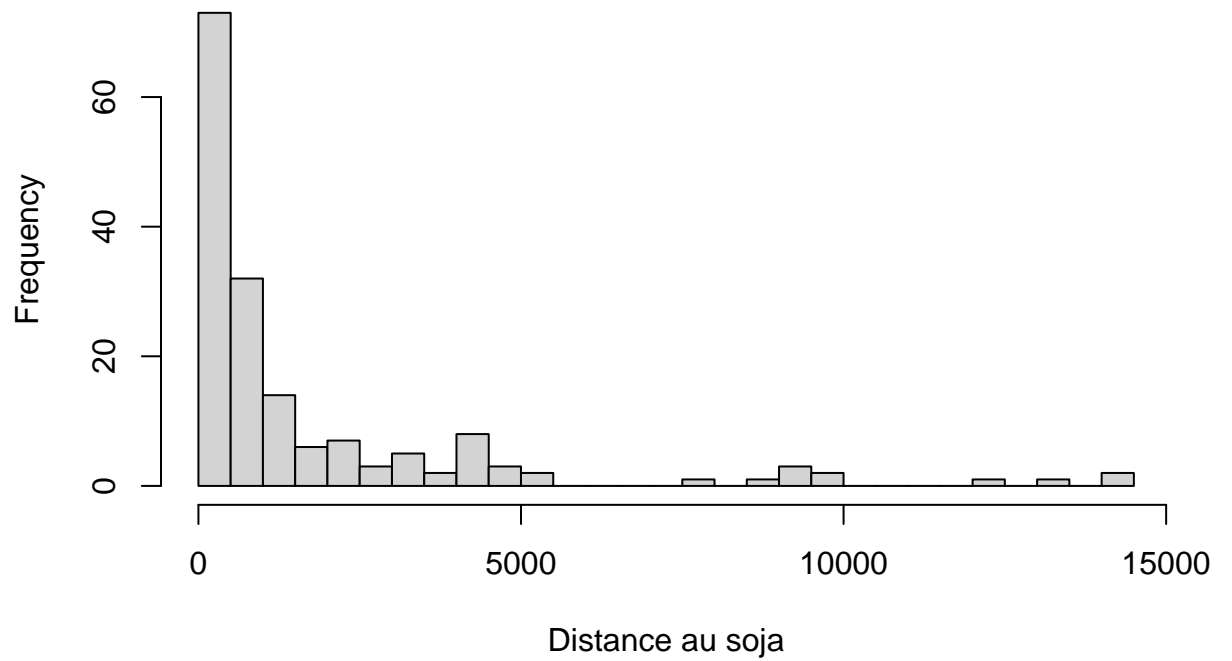


Figure 3.3: Histogramme de la distance au soja pour le groupe 'Sans Agro / Sans Pesticides'

Histogramme – Agro / Sans Pesticides

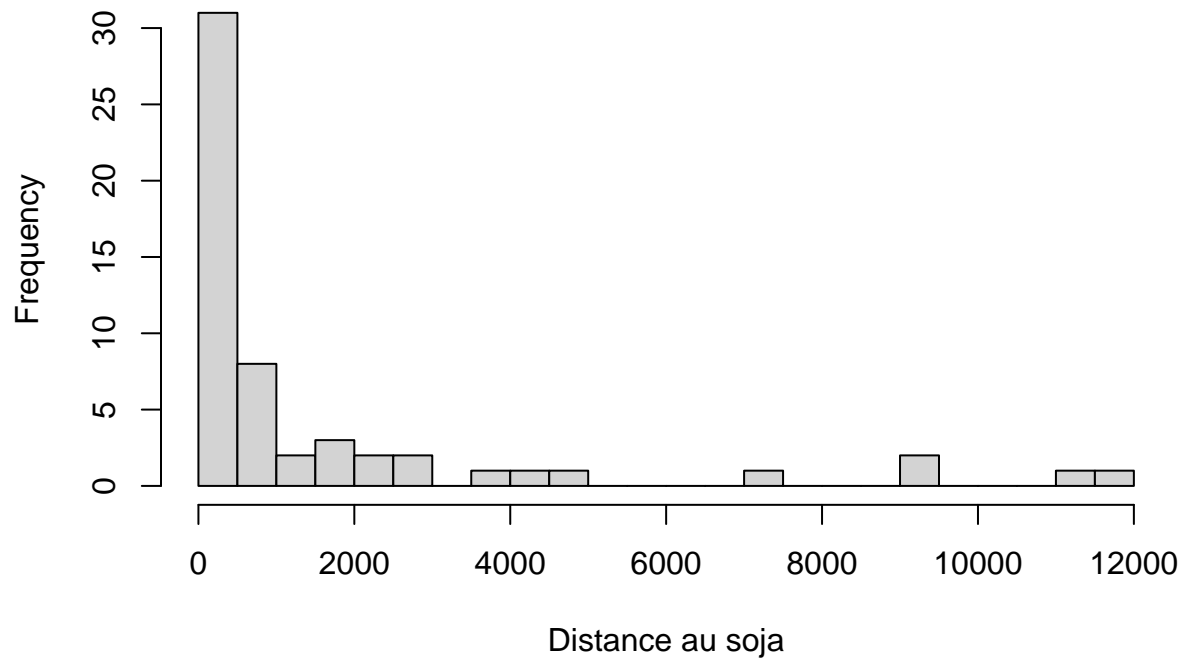


Figure 3.4: Histogramme de la distance au soja pour le groupe 'Agro / Sans Pesticides'

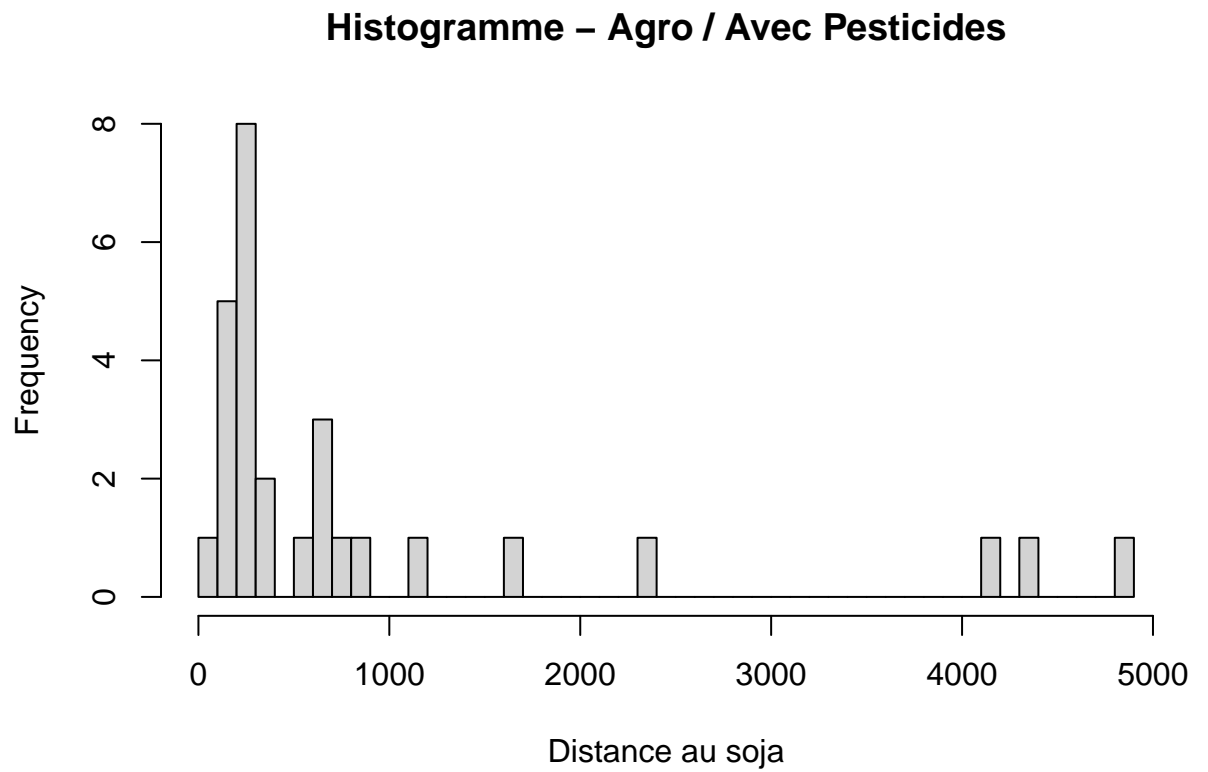


Figure 3.5: Histogramme de la distance au soja pour le groupe 'Agro / Pesticides'

Histogramme – Sans Agro / Avec Pesticides

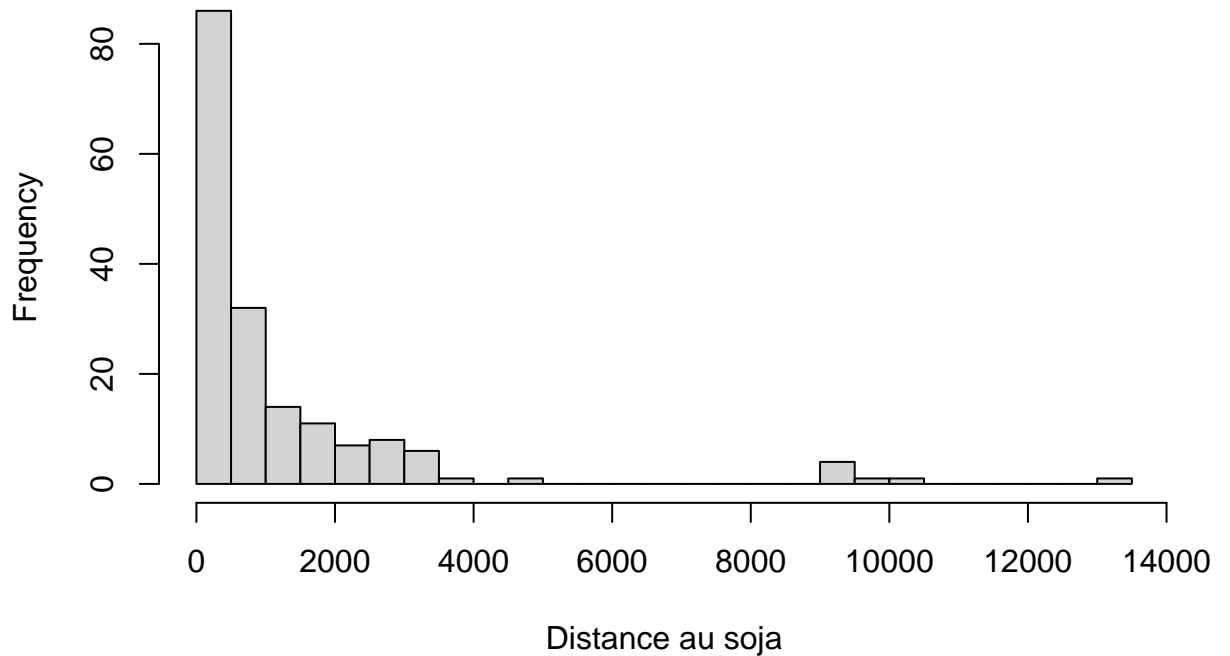


Figure 3.6: Histogramme de la distance au soja pour le groupe 'Sans Agro / Pesticides'

En comparant les histogrammes par groupe, on note que les groupes « Sans agroécologie » (avec pesticides (Figure 3.6) ou sans pesticides (Figure 3.3)) ont tendance à être plus fréquemment exposés à des distances faibles. Cela pourrait refléter leur implantation dans des zones agricoles plus intensives, tandis que les groupes pratiquant l'agroécologie (Figure 3.4 et Figure 3.5) semblent un peu plus dispersés.

Graphique 7 – Fréquence des modalités les plus courantes par groupe

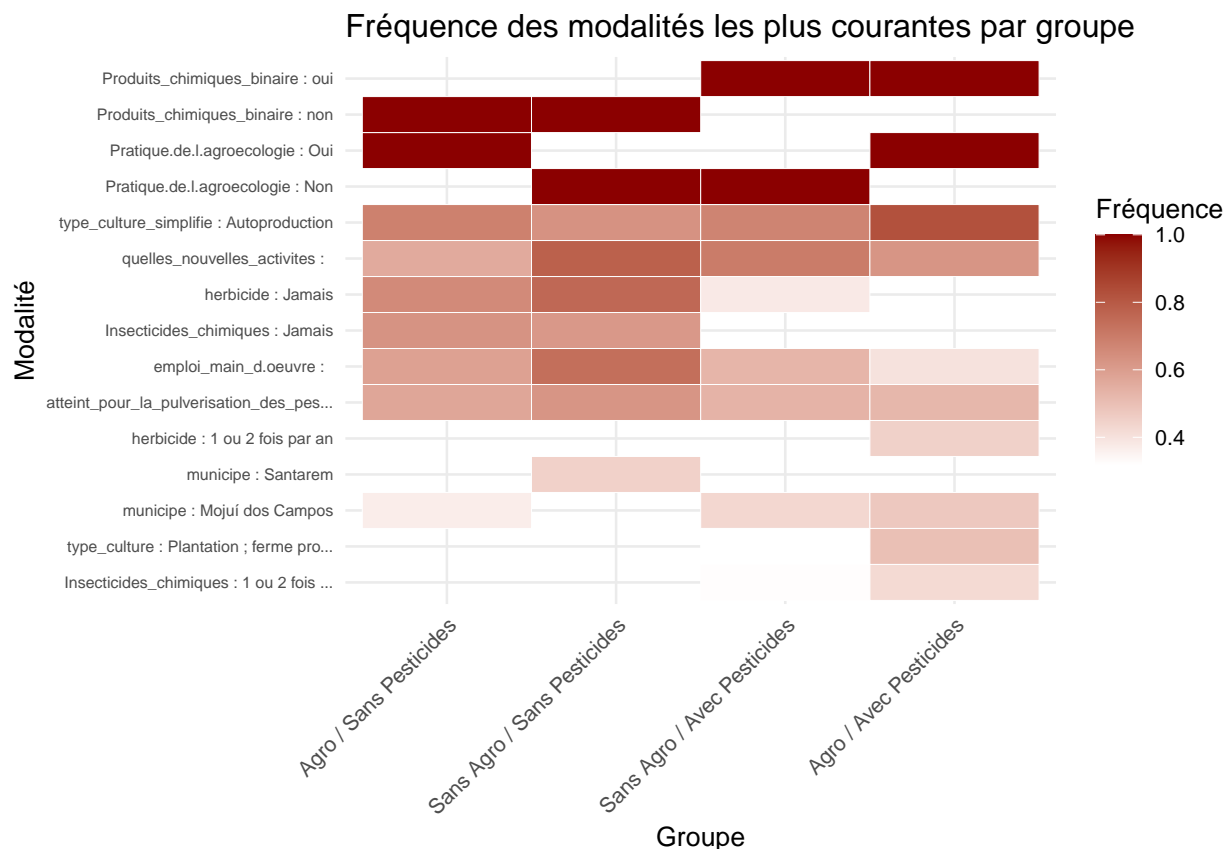


Figure 3.7: Fréquence des modalités les plus courantes par groupe

Ce graphique en tuiles met en évidence la fréquence relative des modalités les plus représentées dans chaque groupe agro/pesticides. Il permet de repérer rapidement les caractéristiques dominantes selon les profils d'agroécologie et d'usage de produits chimiques. Par exemple, on observe que le groupe Agro / Sans Pesticides est fortement associé à l'autoproduction, l'absence de produits chimiques, et une plus grande participation sociale. À l'inverse, le groupe Sans Agro / Avec Pesticides montre une forte exposition aux intrants chimiques et des perceptions environnementales plus critiques.

Graphique 8 – Projection des groupes initiaux dans l'espace ACP mixte

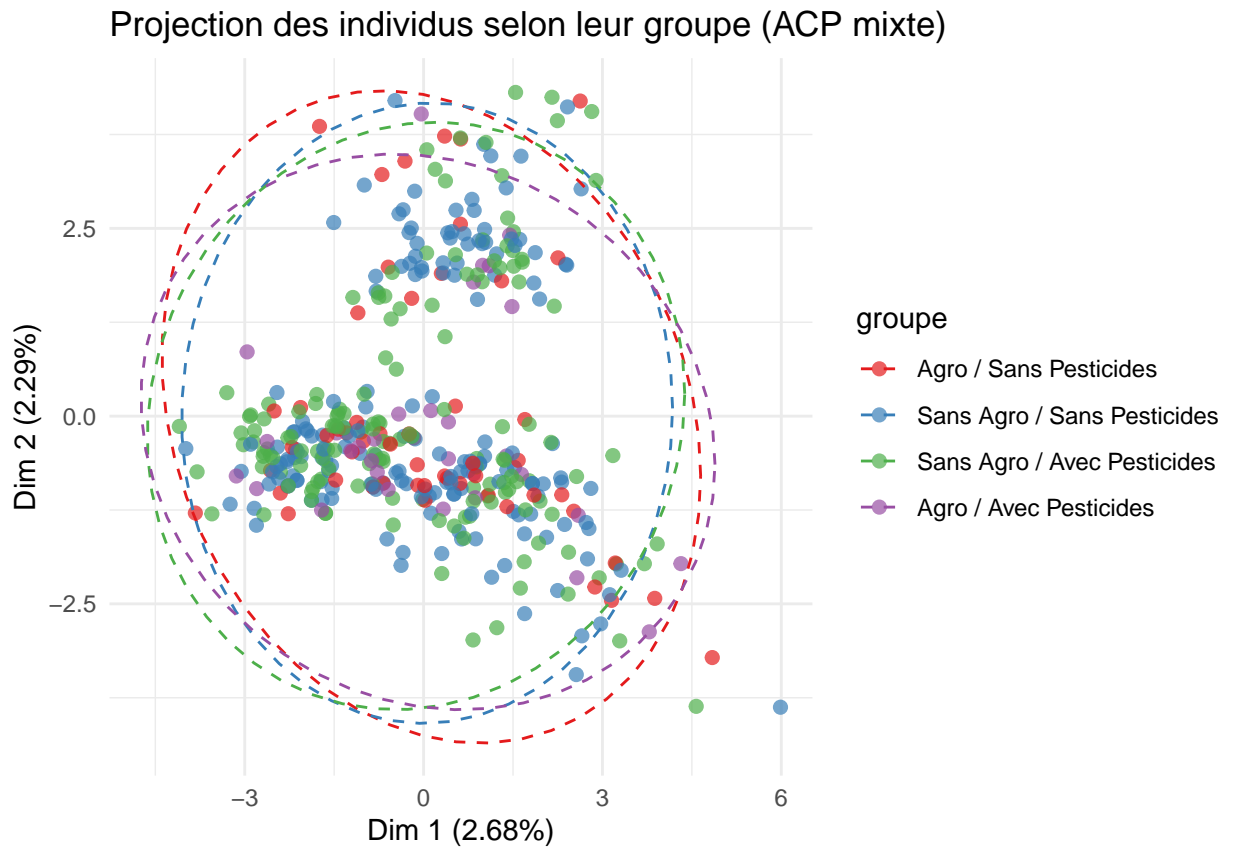


Figure 3.8: Projection des individus selon leur groupe (ACP mixte)

Ce graphique présente la projection des individus selon les deux premières dimensions de l'analyse factorielle mixte (ACPM), en colorant les points selon les groupes définis par la combinaison des pratiques agroécologiques et de l'usage de pesticides.

Les ellipses représentent les zones de concentration des individus pour chaque groupe. Cette visualisation permet d'évaluer la structure spatiale des groupes dans l'espace factoriel. On observe que les groupes agro/pesticides montrent des zones partiellement distinctes, mais avec un fort recouvrement, notamment entre les groupes sans agroécologie. Cela suggère que la typologie initiale ne sépare pas parfaitement les individus selon les deux premières dimensions, bien qu'elle capte certaines tendances générales.

Cela permet de visualiser la structure spatiale des groupes avant de comparer cette organisation avec une classification automatique comme le k-means.

Graphique 9 — Clustering des individus par k-means

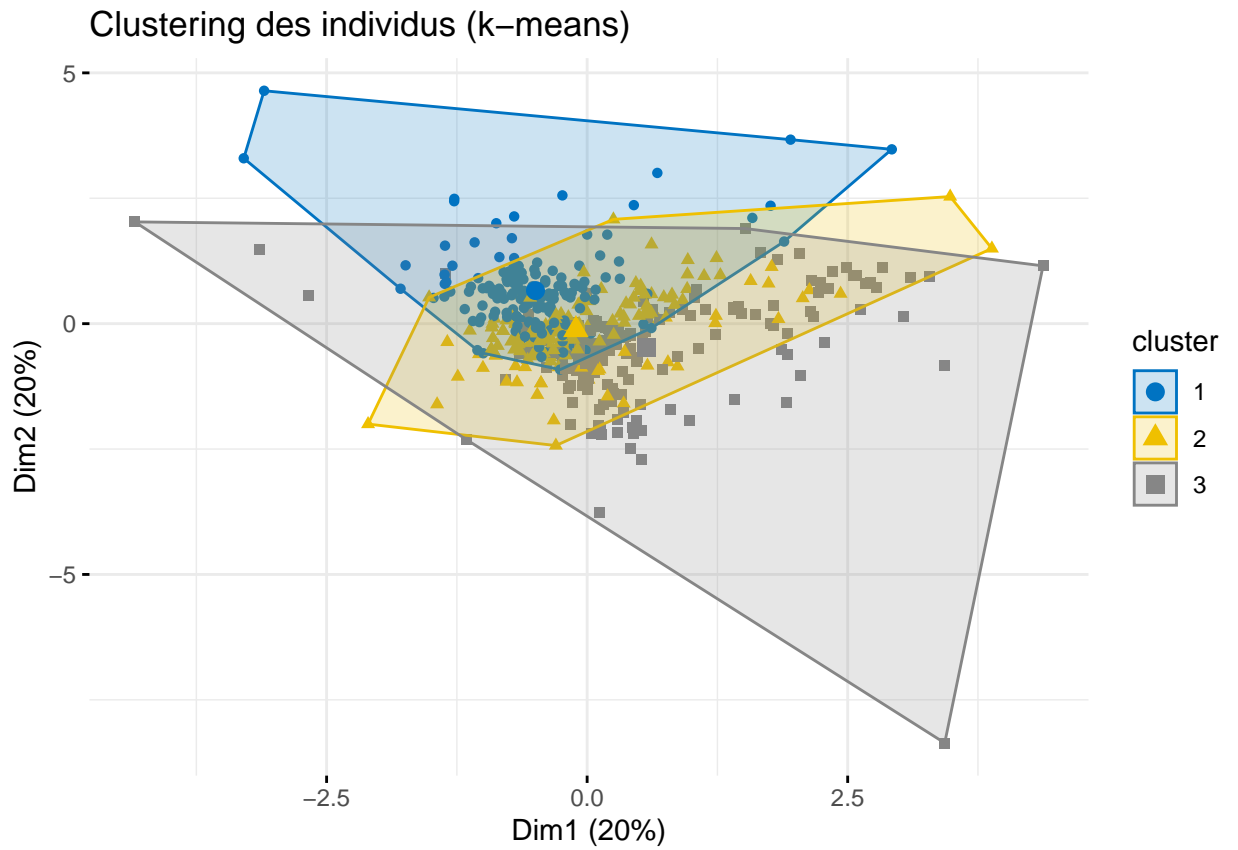


Figure 3.9: Clustering des individus (k-means)

Le graphique du clustering (k-means) projeté sur le plan factoriel montre que les individus peuvent être séparés en trois grandes classes relativement cohérentes dans l'espace réduit. Ces clusters semblent croiser partiellement les groupes agro/pesticides, mais révèlent aussi d'autres dimensions sous-jacentes non directement captées par les variables initiales. Cela ouvre des pistes pour affiner la typologie.

Graphique 10 — Résultats de l'Analyse des Correspondances Multiples

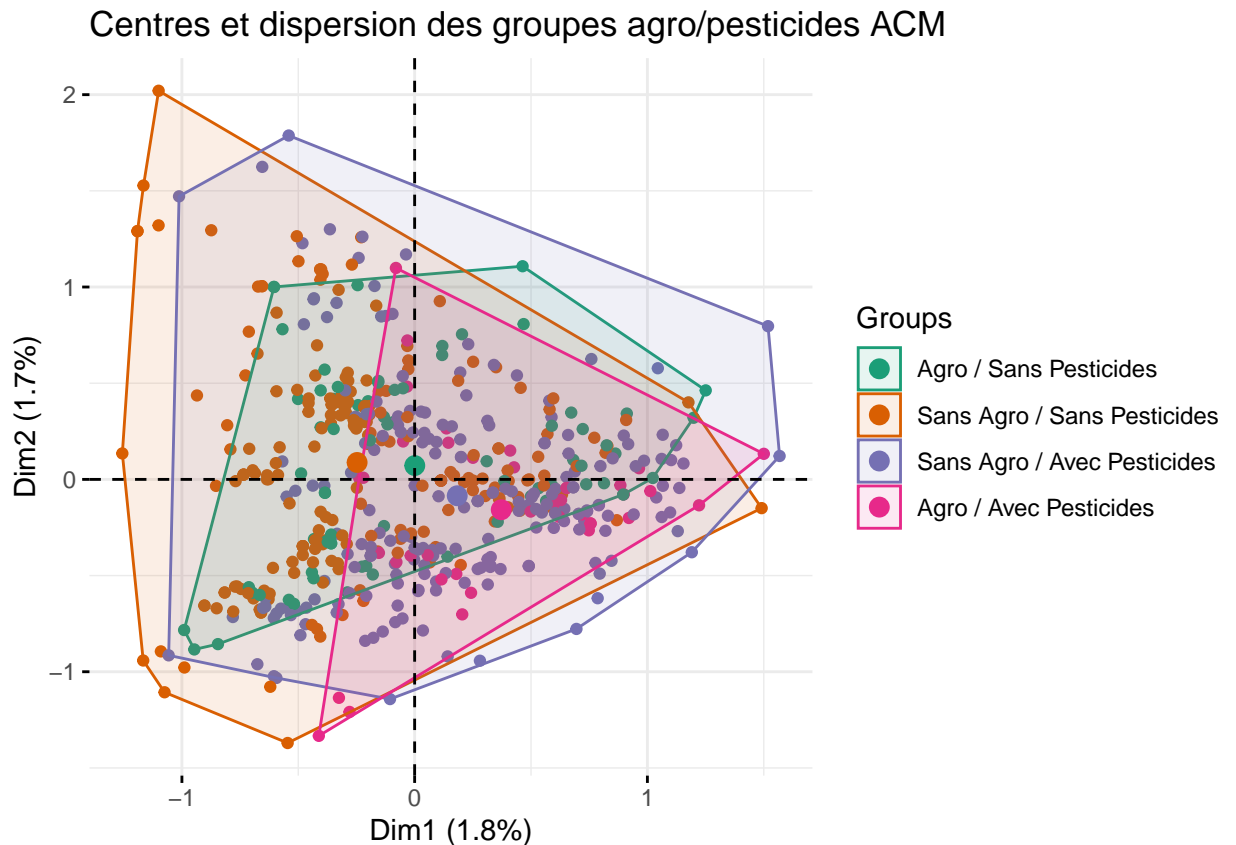


Figure 3.10: Centres et dispersion des groupes agro/pesticides ACM

Ce graphique permet d'explorer la structure spatiale des individus selon leur groupe. Toutefois, il convient de rester prudent dans l'interprétation, car les deux premières dimensions n'expliquent qu'une très faible part de l'inertie totale (moins de 4 %). Les regroupements visibles peuvent donc être partiels ou peu représentatifs de la structure globale. Ce résultat souligne la complexité du jeu de données et suggère que d'autres dimensions factorielles mériteraient d'être explorées pour une compréhension plus complète des profils d'agriculteurs.

On observe certaines tendances, bien que celles-ci doivent être interprétées avec prudence en raison de la faible inertie représentée :

- Le groupe **Agro / Sans Pesticides** semble légèrement plus concentré que les autres, suggérant un profil plus homogène sur les dimensions projetées, mais cela reste à confirmer avec d'autres axes.
- Le groupe **Sans Agro / Sans Pesticides** apparaît plus dispersé sur ce plan, ce qui pourrait refléter une plus grande diversité interne, bien que cette dispersion ne soit significative que si elle se retrouve sur d'autres dimensions.
- Le groupe **Sans Agro / Avec Pesticides** montre une relative compacité, indiquant possiblement un certain degré de cohérence entre les individus, mais cette impression doit être relativisée.
- Le groupe **Agro / Avec Pesticides** occupe une position intermédiaire sur le plan, mais les chevauchements avec d'autres groupes rendent difficile une interprétation claire.

En résumé, ces éléments suggèrent l'existence de différences entre groupes, mais la très faible part de variance expliquée par ce plan (moins de 4 %) limite fortement la robustesse des conclusions. Une exploration de dimensions supplémentaires ou des méthodes complémentaires (e.g. clustering, projection t-SNE) pourrait être nécessaire pour mieux caractériser la structure réelle des profils d'agriculteurs.

CHAPITRE 4

Visualisation, exploration initiale et interprétation des résultats

Cette section vise à compléter l'analyse statistique par des représentations graphiques, afin de mieux comprendre la structure des données, les profils des groupes agroécologiques, et les variables qui les distinguent. Les visualisations permettent également de confirmer ou d'affiner certaines hypothèses émergeant des tableaux croisés ou des tests statistiques.

4.1 Participation sociale selon les groupes

Le premier graphique présente la distribution du niveau de participation sociale selon les quatre groupes définis (croisement agroécologie / usage de pesticides). Il s'agit d'un diagramme en barres empilées, où les proportions de répondants dans chaque niveau de participation (Leadership, Participation simple, Aucune participation) sont visualisées par groupe.

On observe que le leadership est le plus élevé dans le groupe Agro / Avec Pesticides (37,5 %) et dans le groupe Agro / Sans Pesticides (35,6 %). À l'inverse, il est bien plus faible dans les groupes n'adoptant pas l'agroécologie, notamment Sans Agro / Sans Pesticides (19,3 %) et Sans Agro / Avec Pesticides (22,1 %). La participation simple est majoritaire dans tous les groupes, mais elle est particulièrement dominante chez les non-agroécologiques. Enfin, la non-participation reste très marginale ($< 3\%$), avec un maximum de 2,5 % dans le groupe Agro / Avec Pesticides. Ces différences confirment que la dynamique sociale et l'engagement communautaire varient fortement selon les profils agroécologiques.

4.2 Distribution de la distance au soja

Les histogrammes suivants explorent la distribution de la variable Q_11_Dist_soja, c'est-à-dire la distance (en mètres) séparant chaque individu de la culture de soja la plus proche. Le graphique général montre une distribution asymétrique très marquée : la majorité des répondants vivent à moins de 1000 mètres des cultures, avec une concentration autour de 500 mètres. Toutefois, quelques cas extrêmes dépassent 10 000 mètres.

En affinant l'analyse par groupe, on constate que :

- Le groupe *Sans Agro / Sans Pesticides* est celui avec la distribution la plus large. Il comprend de nombreux individus très éloignés des cultures de soja.
- Les groupes *Agro / Sans Pesticides* et *Sans Agro / Avec Pesticides* présentent également une forte concentration en dessous de 2000 mètres, mais avec moins de cas extrêmes.

- Enfin, le groupe *Agro / Avec Pesticides* est aussi représenté dans les zones proches du soja, bien que la distribution semble légèrement plus uniforme.

Ces éléments confirment que les groupes ne se répartissent pas de manière aléatoire dans l'espace, et que la proximité au soja est potentiellement liée au profil agroécologique.

4.3 Dispersion des groupes selon l'ACM

Une analyse en composantes multiples (ACM) a été menée sur plusieurs variables qualitatives (pratiques agricoles, exposition, participation, etc.), afin de mieux comprendre la structure des groupes et leur dispersion.

Le graphique obtenu projette les individus dans un espace factoriel (deux premières dimensions), en les colorant selon leur appartenance au groupe agro/pesticide. Des ellipses convexes ont été tracées autour de chaque groupe pour en visualiser l'étendue et la concentration.

On observe que certains groupes, comme *Sans Agro / Avec Pesticides*, sont plus dispersés, tandis que *Agro / Avec Pesticides* ou *Sans Agro / Sans Pesticides* apparaissent plus resserrés. Cela suggère des profils plus homogènes ou plus hétérogènes selon les cas, et confirme que certaines variables catégorielles permettent de bien distinguer les groupes définis.

4.4 Clustering des individus (K-means sur ACPM)

Un graphique issu d'un clustering k-means réalisé sur les deux premières dimensions de l'ACP mixte permet de visualiser des regroupements spontanés d'individus. Trois clusters principaux émergent. On remarque une certaine correspondance entre ces clusters et les groupes agro/pesticides initiaux, sans être parfaitement superposables. Cela montre que les groupes définis a priori (sur deux variables binaires) capturent une partie de la structure réelle des données, mais pas l'intégralité. Cela peut orienter des analyses futures vers des classifications plus flexibles.

Par ailleurs, avant d'appliquer une méthode de classification automatique, nous avons visualisé la distribution des groupes agroécologiques initiaux dans l'espace factoriel défini par l'ACP mixte. Le **Graphique 6** présente la projection des individus selon les deux premières dimensions de l'ACPM, avec une couleur indiquant leur appartenance au groupe agro/pesticides. Des ellipses de concentration ont été ajoutées pour mieux visualiser les zones de densité par groupe.

Cette représentation permet d'évaluer dans quelle mesure les groupes définis a priori occupent des zones distinctes dans l'espace factoriel. On constate que certains groupes (comme *Sans Agro / Sans Pesticides* et *Agro / Sans Pesticides*) forment des clusters assez bien séparés, tandis que d'autres (comme *Agro / Avec Pesticides*) se trouvent dans une zone de recouvrement. Cette observation souligne la cohérence partielle entre la classification basée sur deux variables binaires (agroécologie, pesticides) et la structure multidimensionnelle révélée par l'analyse factorielle.

Ce graphique sert donc de **pont interprétatif** entre les regroupements initiaux et les résultats du clustering automatique, qui sera présenté dans le graphique suivant.

4.5 Interprétation initiale des résultats

À partir de l'ensemble de ces visualisations, plusieurs interprétations transversales peuvent être formulées :

- Les groupes avec une pratique agroécologique présentent un niveau de participation sociale plus élevé. Cela peut refléter des dynamiques collectives plus structurées, une sensibilisation plus forte, ou encore des réseaux communautaires plus actifs.
- Les groupes sans agroécologie, surtout ceux sans pesticides, semblent plus isolés socialement. Cela peut traduire une position marginale dans le paysage agricole local, ou un repli sur des pratiques autonomes.
- Les perceptions des effets du soja sur l'environnement ou la santé sont hétérogènes. On note que certains groupes expriment plus fortement des inquiétudes, en particulier ceux exposés au soja sans adopter l'agroécologie.
- La distance au soja est un facteur différenciant, mais pas exclusif. Elle permet de mieux comprendre l'exposition potentielle aux externalités du modèle productiviste (pesticides, pollution...).
- Enfin, la structure spatiale et sociale (zones, municipale, emploi, etc.) joue un rôle clair dans l'appartenance aux groupes, ce qui suggère que les pratiques agricoles ne sont pas seulement des choix individuels mais aussi contextuels.

En résumé, cette phase de visualisation et d'exploration initiale a permis de consolider les premières hypothèses issues des analyses statistiques descriptives. Les représentations graphiques ont mis en évidence des contrastes nets entre les groupes agro/pesticides, tant sur le plan des pratiques agricoles que de la participation sociale, des perceptions environnementales ou de la répartition géographique. Ces résultats suggèrent que les profils d'acteurs ne sont pas interchangeables : ils renvoient à des logiques socio-agricoles distinctes, potentiellement influencées par des dynamiques territoriales, institutionnelles ou historiques.

Cette première interprétation servira de socle pour orienter les analyses complémentaires, notamment les analyses factorielles, de classification ou de régression, qui permettront d'explorer plus finement les structures latentes dans les données et d'identifier les variables les plus discriminantes.

CHAPITRE 5

Perspectives ou travail encore à faire

Ce travail a permis d'identifier des profils distincts d'acteurs agricoles selon leur rapport à l'agroécologie et aux pesticides. Toutefois, plusieurs prolongements pourraient enrichir et approfondir l'analyse. Ces perspectives touchent tant à l'outillage statistique qu'aux enjeux opérationnels et politiques.

5.0.1 Perspectives à court terme

Avant d'envisager des prolongements analytiques plus lourds, plusieurs perspectives à court terme peuvent être engagées dès la fin du stage afin d'approfondir les premiers résultats :

- Améliorer l'Analyse des Correspondances Multiples (ACM) en affinant le choix des variables actives et illustratives, ainsi que les paramètres de projection ;
- Étudier la pertinence de la sélection de variables qualitatives utilisées dans les analyses factorielles et les tests statistiques ;
- Réexaminer certaines étapes du traitement des données (recodage, regroupement de modalités) à la lumière des premiers résultats ;
- Mener des échanges avec l'équipe d'accueil pour confronter les résultats à leur expertise de terrain, notamment sur les dynamiques locales, les trajectoires agricoles et les usages du soja ;
- Explorer des visualisations complémentaires afin de rendre les résultats plus intelligibles pour des non-spécialistes.

5.0.2 Approfondissement des analyses statistiques

Des méthodes plus avancées pourraient être mobilisées :

- **Régression logistique multinomiale** pour prédire l'appartenance à un groupe selon diverses variables explicatives ;
- **Analyse discriminante** pour identifier les variables les plus différenciantes ;
- **Modèles mixtes ou hiérarchiques** intégrant les effets des zones ou municipalités ;
- **Robustesse des clusters** via le bootstrap ou la validation croisée.

Ces extensions permettraient d'évaluer la solidité des profils identifiés et d'améliorer la compréhension des déterminants.

5.0.3 Renforcement de la dimension temporelle

L'analyse actuelle est fondée sur des données transversales. Or, les trajectoires agroécologiques évoluent dans le temps :

- Une approche longitudinale (si des données ultérieures sont disponibles) pourrait capter les évolutions ;
- Des *récits de vie* ou entretiens qualitatifs offriraient un éclairage sur les bifurcations vers ou hors de l'agroécologie.

5.0.4 *Valorisation cartographique*

Les données spatiales mériteraient une meilleure valorisation :

- **Cartes thématiques** par groupe et par variable structurante ;
- Intégration dans un **Système d'Information Géographique (SIG)** ;
- Croisements avec des données environnementales (occupation du sol, pollution, proximité des routes...).

Cela permettrait d'identifier les configurations spatiales propices ou défavorables à l'agroécologie.

5.0.5 *Intégration de variables supplémentaires*

Certaines dimensions non exploitées pourraient être ajoutées :

- **Données économiques détaillées** : structure des revenus, endettement, équipement ;
- **Relations sociales et réseaux** : participation à d'autres formes d'organisation, liens avec les ONG ;
- **Système de valeurs** : représentations de la nature, rapport au progrès, croyances en matière de santé.

5.0.6 *Implications pour les politiques publiques*

Les résultats peuvent orienter les actions publiques :

- Ciblage fin des **politiques d'accompagnement** selon les profils ;
- Soutien à la **participation sociale**, levier d'autonomisation et de diffusion ;
- Conception de **dispositifs différenciés** selon les groupes (ex : en transition, isolés, jeunes) ;
- Création d'**indicateurs de suivi** pour évaluer les effets de la transition agroécologique.

Ces perspectives dessinent des prolongements scientifiques et opérationnels concrets, en vue d'un travail partenarial approfondi avec les acteurs locaux et les institutions impliquées.

CHAPITRE 6

Conclusion

6.1 Récapitulatif des apports du travail

Ce travail a permis de structurer et d'analyser un ensemble complexe de données issues d'enquêtes menées auprès de producteurs agricoles au Brésil, dans le contexte de la culture du soja. À travers une démarche rigoureuse mêlant traitement des données, construction de variables synthétiques, analyses statistiques et visualisations graphiques, plusieurs résultats ont été obtenus.

Les principaux apports peuvent être résumés ainsi :

- Une base de données propre, restructurée et cohérente a été constituée à partir du fichier brut, facilitant les traitements ultérieurs.
- Quatre groupes types d'agriculteurs ont été définis à partir du croisement entre la pratique déclarée de l'agroécologie et l'usage de pesticides. Ces groupes ont constitué une grille de lecture pertinente tout au long de l'analyse.
- Des variables synthétiques ont été construites pour décrire les dimensions sociales et spatiales des répondants (ex. niveau de participation, cohabitation, emploi dans la propriété).
- Des analyses statistiques (tests de Kruskal-Wallis et χ^2) ont mis en évidence des différences significatives entre les groupes, notamment en matière de participation sociale, de distance au soja et d'usage d'intrants chimiques.
- Des visualisations ont permis de confirmer et d'illustrer les résultats, notamment en représentant les profils types de chaque groupe, les distances au soja ou encore les dispersions spatiales.
- Des analyses multivariées complémentaires (ACP mixte, clustering, ACM) ont été mobilisées pour explorer les structures latentes dans les données et valider la pertinence des groupes définis.

Ce socle analytique constitue une base solide pour approfondir la compréhension des liens entre pratiques agricoles, conditions sociales, et impacts environnementaux dans les zones concernées par l'extension du soja.

6.2 Réponse à la problématique

L'objectif initial de ce travail était de mieux comprendre comment les pratiques agroécologiques et l'usage des pesticides structurent les profils d'agriculteurs dans les zones affectées par la culture du soja, notamment en termes de conditions sociales, de perception des impacts environnementaux et de positionnement spatial.

Les analyses menées ont permis d'apporter des éléments de réponse clairs :

- Les groupes pratiquant l'agroécologie, en particulier ceux qui n'utilisent pas de pesticides, se distinguent par un niveau de participation sociale plus élevé, une

moindre exposition directe au soja (en termes de distance), et une perception plus critique de ses effets environnementaux.

- À l'inverse, les groupes sans agroécologie — notamment ceux utilisant des produits chimiques — semblent plus isolés socialement, et présentent une plus grande proximité spatiale avec les cultures de soja.
- Ces différences ne sont pas anecdotiques : elles ont été confirmées par des tests statistiques significatifs et par des visualisations qui montrent des profils bien différenciés entre les groupes.
- Enfin, l'ACP mixte et le clustering ont mis en évidence que ces groupes reposent sur des configurations cohérentes de pratiques, de situations sociales et de territoires, ce qui renforce la validité de l'approche typologique adoptée.

En résumé, la problématique posée a trouvé des éléments de réponse empiriques solides : les profils agroécologiques ne sont pas interchangeables avec les profils plus conventionnels, et ils s'inscrivent dans des dynamiques sociales, économiques et géographiques spécifiques. Ces résultats éclairent les tensions entre modèles agricoles dans les régions étudiées et offrent des pistes pour mieux cibler les politiques d'accompagnement de la transition agroécologique.

6.3 Apports pour la structure d'accueil (UMR SENS / CIRAD)

Ce travail s'inscrit pleinement dans les objectifs de recherche portés par l'UMR SENS, en lien avec l'analyse des trajectoires agricoles, des pratiques durables, et des dynamiques sociales en contexte de changement global. Plusieurs apports peuvent être soulignés pour la structure d'accueil :

- **Consolidation méthodologique** : l'approche mobilisée dans ce stage (combinaison d'analyse descriptive, tests statistiques, ACP mixte et classification) constitue une méthode reproductible, claire et rigoureuse pour analyser des profils d'acteurs complexes. Elle pourra être réutilisée dans d'autres terrains suivis par l'UMR SENS, notamment au Brésil, à Madagascar ou au Burkina Faso.
- **Valorisation des données Odyssea** : ce travail a permis de structurer, nettoyer et exploiter une base de données riche mais initialement peu harmonisée. Les scripts et la documentation associés fournissent une base précieuse pour les futures analyses sur les données du projet Odyssea, notamment en vue de publications scientifiques ou de rapports de synthèse.
- **Lecture stratégique des tensions territoriales** : les résultats mettent en lumière des oppositions claires entre différents modèles agricoles (agroécologique vs conventionnel), non seulement en termes de pratiques, mais aussi d'ancrage social et spatial. Ces éléments peuvent nourrir la réflexion de l'équipe sur les leviers de résilience ou de transformation des systèmes agricoles.
- **Support à l'accompagnement des transitions** : en distinguant les profils selon leur exposition, leur mobilisation collective et leurs représentations, cette analyse permet de mieux cibler les dispositifs d'appui, de plaider ou d'animation territoriale promus par l'UMR SENS en lien avec ses partenaires.

Ainsi, ce stage n'a pas seulement contribué à répondre à une problématique scientifique, mais a aussi enrichi les outils, les réflexions et les perspectives opérationnelles de l'unité d'accueil.

6.4 Apprentissage personnel

Ce stage représente une expérience particulièrement formatrice, tant sur le plan méthodologique que personnel. Il me permet de mettre en pratique, dans un cadre concret de recherche appliquée, un ensemble de compétences acquises au cours de ma formation universitaire, tout en découvrant de nouveaux outils et en développant une posture d'analyse autonome.

- **Renforcement des compétences techniques** : j'ai approfondi ma maîtrise de R, en particulier sur des thématiques d'analyse multivariée (ACP mixte, clustering), de traitement des données qualitatives, de production de graphiques reproductibles, ainsi que de rédaction de rapports automatisés en `RMarkdown`/`LATEX`. L'enchaînement des phases de nettoyage, d'exploration, d'analyse et de visualisation m'a permis de mieux intégrer la logique globale d'une démarche d'analyse de données.
- **Apprentissage d'une rigueur scientifique** : travailler au sein d'une unité de recherche comme l'UMR SENS m'a confronté à l'exigence de clarté, de traçabilité et de justification dans chaque étape du traitement des données. J'ai appris à documenter mes choix, à croiser les approches descriptives et exploratoires, et à produire des résultats interprétables en dialogue avec les hypothèses du terrain.
- **Développement d'un regard critique** : au-delà de l'exécution technique, ce stage m'a permis de mieux comprendre les enjeux liés à la structuration des données, à la formulation des variables, à la gestion des valeurs manquantes ou à l'interprétation prudente des résultats. J'ai pris conscience que toute analyse de données nécessite des arbitrages et que le sens des résultats dépend fortement du contexte d'enquête.
- **Ouverture interdisciplinaire** : en m'immergeant dans une problématique à l'interface entre agriculture, environnement et sciences sociales, j'ai pu mesurer l'intérêt et la complexité de croiser des approches quantitatives avec des enjeux territoriaux et politiques concrets. Cela m'a conforté dans mon souhait de m'orienter vers des projets appliqués, mobilisant les outils statistiques au service de problématiques sociales.

En somme, ce stage a été une étape décisive dans mon parcours. Il m'a permis non seulement de valider mes compétences en analyse de données, mais aussi de prendre confiance dans ma capacité à mener un travail autonome, rigoureux et pertinent sur des données réelles. Il constitue un socle solide pour la poursuite de mon parcours en Master et mes futurs projets professionnels.

6.5 Conclusion

Ce stage constitue une opportunité précieuse pour explorer en profondeur un ensemble de données complexes issues du projet Odyssea, en lien avec les dynamiques agricoles et sociales à l'œuvre dans la région de Santarém, au Brésil. À travers une démarche progressive de nettoyage, de structuration, d'analyse et d'interprétation,

nous avons pu mettre en évidence des profils distincts d'agriculteurs, différenciés selon leur recours à l'agroécologie, leur usage des pesticides, leur degré d'intégration sociale et leur exposition au soja.

Les analyses croisées ont révélé des contrastes significatifs entre les groupes, non seulement sur les pratiques agricoles, mais aussi sur les perceptions environnementales, la participation communautaire et la distribution spatiale. L'approche mobilisée, combinant statistiques descriptives, tests de significativité, analyse factorielle mixte et classification, a permis de mettre en évidence certaines différences entre les profils. Toutefois, ces résultats doivent encore être approfondis et validés par une analyse plus poussée, ainsi qu'évalués à la lumière d'une expertise métier, afin de mieux en apprécier la portée et la pertinence dans le contexte local.

Au-delà des résultats empiriques, ce travail a également mis en évidence l'intérêt des méthodes d'analyse de données pour éclairer des enjeux concrets de terrain. Il a permis de renforcer les capacités d'analyse de l'équipe d'accueil, tout en offrant une base exploitable pour de futurs travaux scientifiques ou opérationnels. Enfin, sur le plan personnel, il a représenté une étape structurante dans mon parcours, me permettant de consolider mes compétences techniques, d'affirmer mes choix d'orientation, et de m'approprier une posture analytique rigoureuse et autonome.

Ce travail pourra se prolonger par des analyses plus ciblées, notamment en intégrant les données longitudinales ou en croisant davantage les approches quantitatives et qualitatives. Il ouvre aussi la voie à une réflexion plus large sur les politiques d'appui aux agricultures durables et sur les mécanismes sociaux qui favorisent (ou freinent) leur diffusion dans les territoires ruraux.

References

- [1] Altieri, Miguel A. (1995). *Agroecology: The Science of Sustainable Agriculture*. Westview Press.
- [2] Coudel, Emilie, Hubert, Bernard, & Tonneau, Jean-Philippe (2013). Agroecology as a political practice. *Ecology and Society*, 18(4), 23–35. doi:10.5751/ES-05989-180423
- [3] UMR SENS – CIRAD (2024). *Données du projet Odyssea – enquête territoriale 2022–2023*. Fichier interne.