

Introduction

Il est souvent dit que Madagascar regorge de ressources naturelles exceptionnelles. Pourtant, il est classé parmi les pays les moins développés de la planète et la majorité de sa population, entre autres la population rurale, vit dans une extrême pauvreté (PAM, 2014). La Grande Ile est devenue un pays de toutes les vulnérabilités : économique, climatique, sanitaire, institutionnelle et par-dessus tout alimentaire. Cette situation de vulnérabilité s'est généralisée dans le Pays et s'empire de plus en plus dans le monde rural et le secteur agricole notamment en matière de sécurité alimentaire. Dans ce contexte, une problématique de fond émerge : les petites exploitations agricoles ont du mal à assurer leur sécurité alimentaire en tant que ménage malgré les actions de développement apporté pour redresser la situation. Cette problématique alimentait la recherche consacrée à cette partie suivant deux questions de recherche majeures sont :

- comment et vers quoi évoluait la sécurité alimentaire des PPD depuis la fin des projets d'appuis ?
- quelle est la vulnérabilité des PPD en matière de sécurité alimentaire ?

L'objectif de cette partie est de mettre en relief la différenciation des PPD en matière de sécurité alimentaire. Les objectifs spécifiques consistent à :

- définir et de comprendre les modes de sécurité alimentaire des PPD,
- identifier par la suite leurs points de vulnérabilité.

Les hypothèses de recherche sont que :

- les PPD ont différents modes de sécurité alimentaire,
- les PPD sont encore vulnérables après les projets.

Les résultats attendus seront :

- les modes de sécurité alimentaire adoptés par les PPD au lendemain des projets d'appuis,
- la mise en évidence de la différenciation de leur vulnérabilité en matière de sécurité alimentaire.

Seront ainsi dissertés succinctement les matériels et les méthodes complémentaires utilisés, les résultats relatifs aux analyses sur la différenciation de la sécurité alimentaire des PPD et les différentes discussions sur les modes de sécurité alimentaire des PPD et leur vulnérabilité respective.

2.1 Matériels et méthodes

Les variables de la sécurité alimentaire à Madagascar

La FAO a confirmé dans son rapport d'évaluation que le riz reste la denrée principale de l'alimentation des Malagasy (FAO & PAM, 2010). Par conséquent, étudier la sécurité alimentaire de la population de Madagascar revient en grande partie à étudier sa sécurité en riz. Moyennant de trois dimensions du concept de la sécurité alimentaire (PAM, 2005), on peut retenir comme étant pertinentes les variables qui suivent pour étudier ainsi la sécurité alimentaire de la population de la Grande Ile :

- **La Disponibilité** de la nourriture par la capacité de production de riz qui est traduite par la surface rizicole de l'exploitation.

- **L'Accessibilité** à la nourriture par l'autosuffisance en riz, la consommation de ce que l'on a produit, elle est traduite par la durée de la Suffisance en riz ou bien de l'autonomie en riz.

- **L'Utilisation** de la nourriture, avec deux variables :

➤ La qualité de l'alimentation *i.e.* la diversification de production facilite la diversification alimentaire donc une meilleure qualité nutritionnelle qui est donnée par le taux de diversification ou bien par la diversité alimentaire.

➤ L'Insécurité Alimentaire Saisonnière ou IAS où le ménage doit entreprendre un changement temporaire de son régime alimentaire et nutritionnel. Pendant cette période de soudure, la quantité et la qualité changent et se dégradent par rapport à la période normale ; le niveau de satisfaction psychologique et culturelle n'est pas la même en cette période (EPP PADR, 2005). IAS est la durée de la période de soudure de chaque ménage.

En outre, l'étude de la différenciation de la sécurité alimentaire des PPD comprend aussi l'analyse de risques pour chacune de ces variables autrement dit une analyse de vulnérabilité de la sécurité alimentaire des PPD. La surface des cultures vivrières est étudiée en complément pour pouvoir pousser les analyses en profondeur.

2.1.2 Méthode de traitement des données

A noter que cette étude a été menée sur deux populations : la population PSA et la population ERI mais avec les mêmes variables de la sécurité alimentaire qui sont exposées précédemment. Chaque base de données des deux échantillons est ainsi traitée séparément mais avec les mêmes méthodes statistiques décrites ci-après :

- l'Analyse en Composantes principales ou ACP,
- l'Analyse Factorielle des Correspondances ou AFC,
- la Classification Ascendante Hiérarchique ou CAH,
- l'Analyse Factorielle Discriminante ou AFD,
- le test d'égalité d'échantillons indépendants : test de normalité, test de Levene et test t-Student,
- la régression linéaire multiple,
- l'Analyse de variances ou ANOVA

L'Analyse de Composantes Principales

L'ACP regroupe les variables en variables latentes. Ce traitement comprend l'extraction des variables, la variance totale expliquée, la matrice des composantes et le diagramme des composantes.

La qualité de l'extraction des variables est donnée par le tableau de la qualité de représentation. Plus le coefficient d'extraction est élevé, plus la qualité de représentation des variables est meilleure. La variance totale expliquée (Eigen value) permet de décider sur le nombre de composantes à retenir pour l'étude. Le pourcentage de variance totale expliquée ne doit pas être inférieur à 70% pour réduire le plus possible la perte d'inertie et avoir une bonne qualité de représentation des données.

La matrice des composantes donne la contribution de chaque variable sur chacune des composantes principales. Elle permet ainsi d'identifier les variables qui composent chacun des facteurs. Le diagramme des composantes donne la projection des variables dans l'espace et permet de visualiser le poids de chaque variable sur les composantes. Le recours à la rotation varimax permet d'avoir une optimisation de la projection. Ainsi, la matrice des composantes après rotation varimax et la représentation graphique permettent de configurer mieux la composition des deux facteurs et relier les variables à la composante correspondante.

A partir de ces équations, les scores de chaque groupe sont calculés et attribués à chacune des observations. Ceci permet de les affecter manuellement au groupe Subsistance ou au groupe Productivité suivant leurs plus grands scores respectifs, de procéder par la suite à l'Analyse Factorielle Discriminante.

2.1.2.2 L'Analyse Factorielle Discriminante

L'AFD a permis de regrouper les observations selon les composantes principales et obtenir une typologie confirmée. En premier lieu, AFD donne les fonctions déterminantes qui

sont au nombre de $n-1$ des variables indépendantes qui est dans ce cas-ci le nombre des composantes principales retenues. Le tableau des valeurs propres donne le pourcentage de la variation expliquée par la fonction déterminante. Plus ce pourcentage est élevé, plus la fiabilité de l'analyse est meilleure. A la fin, AFD donne les résultats d'affectation des observations avec leurs scores respectifs et confirme ainsi la typologie des observations.

2.1.2.3 Le test d'égalité de proportions

Pour savoir si deux proportions de deux modalités sont statistiquement égales, il faut procéder au test binomial. Le test a comme hypothèse nulle H_0 l'égalité de deux proportions dans un échantillon. H_0 est à retenir si p-value est supérieur à α qui est le degré de signification retenue. Dans l'étude, la valeur retenue pour α est 0,05. Dans le cas contraire *i.e.* p-value inférieur à α , H_0 est à rejeter et les deux proportions sont probablement différentes.

Le test d'égalité d'échantillons indépendants

Les tests d'égalité d'échantillons indépendants commencent par le test de normalité puis le test de Levene et le test t-Student. Ceci pour réaliser des analyses comparatives des deux échantillons indépendants. Les échantillons indépendants étudiés sont respectivement PSA et ERI. Le test d'égalité d'échantillons indépendants a pour but de savoir si les moyennes des scores d'une variable dans deux échantillons indépendants sont égales ou non. Le test d'égalité d'échantillons pose comme postulat que les distributions suivent une loi normale à 95%. Une distribution suit une loi normale à 95 si ses coefficients d'aplatissement et de symétrie sont compris entre -2 et +2. Les configurations des distributions dans les histogrammes et les diagrammes de normalité reflètent aussi cette normalité. Si ce postulat est confirmé, le test paramétrique peut se faire pour tester l'égalité des échantillons, il s'agit dans ce cas de test t-Student et de Levene. Au cas où le postulat n'est pas vérifié, il faut procéder au test non paramétrique pour vérifier l'égalité des deux échantillons. Dans cette partie, le postulat de la normalité est vérifié et les observations sont indépendantes. Par conséquent, les conditions du recours au test t-Student et au test de Levene sont remplies pour tester l'égalité des deux échantillons indépendants PSA et ERI.

Pour comparer deux moyennes d'une variable dans deux échantillons différents, la variable dépendante doit être une variable d'échelle, la variable indépendante quant à elle est dichotomique. L'hypothèse nulle H_0 est l'égalité des deux moyennes ($H_0 : \mu_1 = \mu_2$). Si le niveau de signification du test t est supérieur au seuil de signification retenu, il faut retenir et conclure que les moyennes de la variable sont égales dans les deux échantillons. Dans le cas

contraire, il faut rejeter H_0 et conclure que les deux moyennes de la variable sont probablement différentes dans les deux échantillons.

Le premier tableau expose les moyennes calculées, c'est la statistique descriptive des échantillons par rapport à la variable. Ce tableau permet de lire entre autres les moyennes et les écart-types respectifs. Pour conclure sur la signification de la différence des moyennes, il faut lire le deuxième tableau et le test proprement dit. Il comprend les résultats du test d'égalité de variances de Levene et du test t-Student.

Le test de Levene qui passe en premier lieu permet de vérifier l'égalité de variances de la variable étudiée dans les deux échantillons. L'hypothèse nulle H_0 est l'égalité des variances dans les deux échantillons ($H_0: \sigma_1^2 = \sigma_2^2$), l'hypothèse alternative H_1 est la différence des variances ($\sigma_1^2 \neq \sigma_2^2$). De cette signification dépend la lecture des résultats du test t-Student. Si la signification p-value du test de Levene est supérieur à $\alpha=0,05$, il faut retenir l'hypothèse nulle H_0 et conclure une égalité des variances. Par la suite il faut lire la première ligne du tableau pour le test t-Student et conclure sur la signification du test et l'égalité des moyennes de la variable sur les deux échantillons. Si le niveau de signification p-value du test de Levene est inférieur à $\alpha=0,05$, il faut rejeter l'hypothèse nulle, les variances sont donc différentes et il faut dans ce cas pour continuer lire la deuxième ligne et conclure sur la signification du test et d'égalité des moyennes.

2.1.2.5 La régression linéaire multiple

La régression multiple est nécessaire à la modélisation en produisant l'équation du modèle de Sécurité alimentaire de chaque échantillon. Faisant suite à la typologie produite par AFD, la modélisation est faisable par le biais de la régression linéaire multiple. En premier lieu, on a le tableau récapitulatif des modèles qui comprend le nom du modèle, le coefficient de corrélation R de Pearson, le R-deux et le R-deux ajusté. R exprime l'intensité et le sens de la corrélation tandis que R-deux donne la proportion de la variable dépendante expliquée par la variable indépendante. En second lieu, il y a le tableau ANOVA qui permet surtout de statuer sur le niveau de significativité de l'analyse de régression. En dernier lieu se trouve le tableau des coefficients de régression permettant à la fois d'obtenir les coefficients non standardisés de l'équation du modèle et les coefficients de régression standardisés β .

A noter que le coefficient de régression non standardisé donne la valeur de la pente partielle décrivant le changement de Y quand une variable indépendante X change d'une unité en même temps que l'effet de l'autre variable indépendante sur la variable dépendante Y est contrôlé. Les signes des coefficients de régression non standardisés indiquent le sens

d'augmentation ou de réduction de Y par rapport à X. La grandeur des coefficients de régression indiquent l'importance de l'effet de chaque variable sur Y tout en contrôlant l'effet de l'autre VI sur VD.

Ainsi, l'équation du modèle est $Y = a + A_1 X_1 + A_2 X_2$

- Y : score prédit de la variable dépendante Y
- a : constante *i.e.* intersection du plan de régression et l'axe de la variable dépendante
- A_1 et A_2 sont des coefficients de régression non standardisés :
- A_1 : pente du plan par rapport à la variable indépendante X_1
- A_2 : pente du plan par rapport à la variable indépendante X_2
- X_1 : score de la variable indépendante X_1
- X_2 : score de la variable indépendante X_2

Le coefficient de régression standardisé β indique l'importance de la variable indépendante sur la variable dépendante, plus β est grande, plus la variable indépendante a une influence sur la variable dépendante. En effet, il indique le changement en écart-type de Y si X_1 change 01 écart-type tout en contrôlant les effets des autres VI

L'étude de l'effet combiné des deux composantes est le fait d'assembler toutes les variables indépendantes dans l'analyse de régression. L'étude de l'effet combiné des variables de la sécurité alimentaire des exploitations agricole permet de pousser encore plus loin les analyses. En effet, on peut considérer, pour cette partie, que la sécurité alimentaire totale SA est égale à la somme des deux composantes retenues. Ce qui attribue un nouveau score à chacune des observations. Les scores ainsi obtenus sont traités en régression linéaire multiples et donnent les résultats.

2.1.2.6 La Classification Ascendante Hiérarchique

La CAH est une méthode de classification qui permet de regrouper les données dans un certain nombre de classes adéquates. Il fait ressortir les modalités nécessaires et permet de constituer le tableau de contingence pour l'Analyse Factorielle de Correspondance. Le mode de classification utilisé est le regroupement des lignes par dissimilarité de la distance euclidienne conjuguée avec la méthode d'agrégation en lien complet. Ceci permet d'obtenir des classes plus compactes.

Les observations sont classées par le biais de CAH selon leur score en sécurité alimentaire. Les résultats de la classification sont donnés par le tableau des points centraux et

le tableau synoptique des classes. En premier lieu, le tableau des points centraux donne les points centraux de chaque classe et permet ainsi de catégoriser et hiérarchiser ces classes selon la valeur de leur point central respectif. En second lieu, le tableau synoptique des classes est constitué par les moyennes, les minimums et les maximums de chacune des classes ainsi que leur proportion respective.

En outre, les scores moyens de chaque classe ainsi que les minimums et les maximums servent à établir l'échelle de la variable dépendante étudiée qui est la sécurité alimentaire pour cette partie.

2.1.2.7 L'Analyse Factorielle des Correspondances

L'Analyse Factorielle des correspondances ou AFC permet d'étudier l'association entre les variables et représenter graphiquement les proximités entre les modalités. Dans cette partie, les modalités étudiées sont celles du score de sécurité alimentaire et les modes de sécurité alimentaire des échantillons. Trois résultats d'analyse sont particulièrement retenus pour l'étude ; ce sont les profils lignes, les distances khi-deux et le graphique asymétrique.

L'AFC est basée sur l'analyse des profils. Un profil est une fréquence relative égale à l'ensemble des fréquences divisées par leur total. Il reflète la façon dont la catégorie d'une variable varie selon les catégories de l'autre variable. Le tableau des profils lignes donne ainsi les profils lignes et permet d'étudier la proximité entre les groupes. Le tableau des distances khi-deux affiche les distances khi-deux entre les points lignes et les points colonnes. Il permet ainsi de détecter les cas similaires dont la distance est petite.

Le graphique asymétrique des lignes est une représentation pseudo-barycentrique qui représente les colonnes dans l'espace des lignes. Les distances entre lignes et colonnes peuvent être interprétées en projetant les points-lignes sur les vecteurs-colonnes. La modalité correspondante au point-ligne qui est le plus éloigné de l'origine est celle qui est la plus liée à la modalité correspondant au vecteur.

2.1.2.8 L'analyse de variance

L'ANOVA sert à analyser la relation entre l'autonomie en riz et la superficie moyenne des cultures vivrières. L'ANOVA ou Analyse de Variance permet de modéliser une variable quantitative avec une variable explicative qualitative qui est aussi appelée facteur. Ceci permettra de mettre en exergue la relation de variance entre les deux variables.

2.2 Résultats

Sont exposés succinctement dans ce paragraphe le regroupement des variables de l'hypothèse en composantes principales, la typologie, la comparaison des deux échantillons de l'univers d'étude, la modélisation, la classification et l'analyse des correspondances entre les variables de la sécurité alimentaire ainsi que l'analyse de la vulnérabilité de la sécurité alimentaire des PPD.

2.2.1 Evolution de la sécurité alimentaire des PPD

2.2.1.1 Sécurité alimentaire en deux composantes

L'ACP a permis d'identifier deux composantes principales selon la qualité de représentation donnée par le tableau 5.

Tableau 5 : Qualité de représentation

| | PSA | | ERI | |
|-----------|---------|------------|---------|------------|
| | Initial | Extraction | Initial | Extraction |
| Capacité | 1,00 | 0,53 | 1,00 | 0,56 |
| Autonomie | 1,00 | 0,99 | 1,00 | 1,00 |
| IAS | 1,00 | 0,99 | 1,00 | 1,00 |
| Diversité | 1,00 | 0,80 | 1,00 | 0,54 |

Tous les coefficients d'extraction sont largement supérieurs à 0,30. Ce qui permet de conclure une bonne qualité d'extraction et de retenir ainsi toutes les variables initiales.

Le nombre de composantes est obtenu à travers le cumul du pourcentage de la variance (tableau 6).

Tableau 6 : Variance totale expliquée

| Composante | Valeurs propres initiales PSA | | | Valeurs propres initiales ERI | | |
|------------|-------------------------------|------------------|--------------|-------------------------------|------------------------|--------------|
| | Total | % de la variance | % cumulés | Total | % de la variance | % cumulés |
| 1 | 2,33 | 58,26 | 58,26 | 2,13 | 53,34 | 53,34 |
| 2 | 0,97 | 24,35 | 82,60 | 0,96 | 24,07 | 77,41 |
| 3 | 0,70 | 17,40 | 100,00 | 0,90 | 22,59 | 100,00 |
| 4 | 0,00 | 0,00 | 100,00 | $-4,64 \cdot 10^{-17}$ | $-1,16 \cdot 10^{-15}$ | 100,00 |

D'après ces résultats, les deux premières composantes permettent d'expliquer 82,6% de la variation pour PSA et 77,41% pour ERI ; la projection de chacune des variables sur les composantes est donnée par le tableau 7.

Tableau 7 : Matrice des composantes

| | PSA | | | | ERI | | | |
|-----------|----------------|------|----------------|------------|----------------|-------|----------------|------------|
| | Avant rotation | | Après rotation | | Avant rotation | | Après rotation | |
| | Composante | | Composante | | Composante | | Composante | |
| | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| Capacité | 0,6 | 0,4 | 0,3 | 0,6 | 0,3 | 0,66 | 0,1 | 0,7 |
| Autonomie | 0,9 | -0,3 | 0,9 | 0,2 | 0,9 | -0,23 | 0,9 | 0,13 |
| IAS | -0,9 | 0,3 | -0,9 | -0,2 | -0,9 | 0,23 | -0,9 | -0,13 |
| Diversité | 0,4 | 0,8 | 0,0 | 0,9 | 0,3 | 0,64 | 0,1 | 0,7 |

La composante 1 est constituée par Autonomie et IAS tandis que la composante 2 par Capacité et Diversité.

La matrice des composantes après rotation varimax et la représentation graphique permettent de configurer mieux la composition des deux facteurs ; ils sont composés respectivement des variables ayant les plus grandes valeurs absolues :

- Composante F1 : Autonomie et IAS qui sous entendent SUBSISTANCE
- Composante F2 : Capacité et Diversité qui sous entendent PRODUCTIVITE

De tout ce qui précède découlent les équations primaires pour chacun des modèles cités ci-dessus :

- Modèle PSA
 - ⇒ Subsistance = 0,9 * Autonomie – 0,9 * IAS
 - ⇒ Productivité = 0,6 * Capacité + 0,9 * Diversité
- Modèle ERI :
 - Subsistance= 0,9 (Autonomie - IAS)
 - Productivité = 0,7 Capacité + 0,7 Diversité

Ces équations ont permis la scorification de chaque observation pour chaque composante et de procéder à l'affectation manuelle puis à l'AFD afin de confirmer le groupe d'appartenance de chaque observation.